Genomic and transcriptomic analysis of genes involved in disease resistance against enteric septicemia of catfish (ESC) in catfish

by

Suxu Tan

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
December 15, 2018

Keywords: catfish, ESC, QTL, GWAS, alternative splicing

Approved by

Zhanjiang Liu, Chair, Professor Emeritus of School of Fisheries,
Aquaculture, and Aquatic Sciences
Rex Dunham, Co-chair, Professor of School of Fisheries, Aquaculture, and Aquatic Sciences
Charles Chen, Professor of Department of Crop, Soil and Environmental Sciences
Aaron Rashotte, Associate Professor of Department of Biological Sciences

Abstract

Catfish is the leading aquaculture species in the United States, accounting for over 50% of all U.S. aquaculture production. In recent years, catfish industry has been facing numerous challenges, including increased costs of feed and fuel, international competition, and infectious diseases. Among infectious diseases, enteric septicemia of catfish (ESC), caused by the bacterium *Edwardsiella ictaluri*, resulted in the most significant losses. Sustainable and effective prevention/control strategies of ESC have been elusive, and improved understanding of the underlying molecular mechanism could help develop broodstocks with enhanced ESC resistance.

For the first part of my dissertation work, I conducted a genome-wide association study (GWAS) to map quantitative trait loci (QTL) for ESC disease resistance using the first-generation backcrosses. Channel catfish and blue catfish exhibit great contrast in ESC resistance, with channel catfish being highly susceptible and blue catfish being very resistant. The interspecific backcrosses provide an ideal system for the identification of QTL, because both disease resistance gene and trait are segregating in this population. Two significant QTL on LG1 and LG23 were identified as revealed by a mixed linear model and family-based association test. Examination of the resistance alleles indicated their origin from blue catfish. The significantly associated markers should be useful for marker-assisted introgression to develop catfish breeds with ESC resistance. Candidate

genes were also investigated, suggesting the importance of involved pathways of phagocytosis and T-cell activation in ESC resistance.

The second part of my dissertation work focused on post-transcriptional regulation of the transcriptome, specifically the regulation of alternative splicing (AS) after infection with *E. ictaluri.* AS can generate more than one transcript from a single pre-mRNA in eukaryotes, playing crucial roles in gene regulation. My questions were if *E. ictaluri* infection affected alternative splicing, and if so, what genes were affected? What was the mode of alternative splicing catfish adopted to cope with *E. ictaluri* infection? I utilized genomic information and RNA-Seq datasets to characterize AS profiles and their induced changes after bacterial infection with *E. ictaluri* in channel catfish. A total of 27,476 AS events from 9,694 genes were identified in channel catfish. AS was observed to be greatly induced by *E. ictaluri* infection. Intriguingly, genes involved in RNA binding and RNA splicing themselves were significantly enriched in differentially alternatively spliced (DAS) genes after infection. These results will help in understanding the molecular mechanisms underlying ESC resistance and host responses to *E. ictaluri* in catfish.

Acknowledgments

I would like to express heartfelt thanks to my advisor Dr. Zhanjiang Liu for his invaluable guidance and tremendous support. He has been a role model for excellence in research, mentoring, and education of graduate students. I would also like to extend sincere gratitude to my co-advisor Dr. Rex Dunham, my committee members Dr. Charles Chen and Dr. Aaron Rashotte, and my University reader Dr. Mark Liles for their valuable feedback and assistance. Their passion and attitude towards research have a great impact on my own research.

I have been studying and working in a great research group. I would like to thank Wenwen Wang, Dr. Changxu Tian, Dr. Xin Geng, Dr. Jun Yao, Dr. Shikai Liu, Dr. Lisui Bao, Dr. Sen Gao, Dr. Tao Zhou, Dr. Xiaozhu Wang, Dr. Ning Li, Dr. Donghong Niu, Dr. Xiaoxiao Zhong, Dr. Yulin Jin, Dr. Yujia Yang, Dr. Zihao Yuan, and Ms. Dongya Gao and all other colleagues in the laboratory. I am very grateful to them for their help, collaboration, and friendship.

I would like to thank the Chinese Scholarship Council for its financial support.

Finally, I would like to express my deepest appreciation to my beloved girlfriend Wenwen Wang and my parents Gang Tan and Hongtao Li for their endless love and constant encouragement.

Table of Contents

Abstract ii

Acknowledgments iv

List of Tables viii

List of Figures ix

List of Abbreviations x

Chapter 1 Introduction 1

1.1 The catfish industry 1

1.2 Enteric septicemia of catfish 2

1.3 Genome-wide association study 4

1.4 Alternative splicing 6

References 9

Chapter 2 A genome-wide association study in interspecific backcross catfish for QTL controlling enteric septicemia of catfish 16

2.1 Abstract 16

2.2 Background 17

2.3 Methods 20

2.3.1 Ethics statement 20

2.3.2 Experimental fish, bacterial challenge and sample collection 20

2.3.3 DNA isolation and genotyping 22
2.3.4 Statistical analysis 23
2.3.5 Sequence analysis 24
2.4 Results 25
2.4.1 Experimental fish and sample structure 25
2.4.2 QTL for ESC resistance 27
2.4.3 Genes within the associated QTL for ESC resistance 35
2.4.4 Correlation of the SNPs with ESC resistance 38
2.5 Discussion 39
References 50
Chapter 3 Increased alternative splicing as a host response to *Edwardsiella ictaluri* infection in catfish 60
3.1 Abstract 60
3.2 Background 61
3.3 Methods 63
3.3.1 Data generation 63
3.3.2 Identification of alternative splicing in channel catfish 63
3.3.3 Identification of alternative splicing after infection with *E. ictaluri* 64
3.3.4 Identification of differential alternative splicing 65
3.3.5 Functional enrichment analysis 65
3.3.6 Gene analysis 65
3.3.7 Quantitative real-time PCR analysis 66

3.4 Results 68

3.4.1 Alternative splicing profiles in channel catfish 68

3.4.2 Infection induced increase of alternative splicing 71

3.4.3 Enriched DAS genes after *E. ictaluri* infection 74

3.4.4 Gene analysis 76

3.4.5 Quantitative real-time PCR analysis 77

3.5 Discussion 80

References 84

List of Tables

Table 1. The catfish samples used in the GWAS analysis 25

Table 2. SNPs associated with ESC disease resistance in the first-generation backcrosses. 31

Table 3. Immune-related genes within the genome-wide significant QTL on linkage group 1 and 23 37

Table 4. Primers used for qPCR validation of seven differential alternative splicing (DAS) events 67

Table 5. Summary of RNA-Seq paired-end reads from nineteen tissues of a doubled haploid channel catfish (*Ictalurus punctatus*) before and after quality control 69

Table 6. Alternative splicing events in double haploid channel catfish (*Ictalurus punctatus*) 69

Table 7. Number of the alternative splicing event and alternatively spliced genes identified from each chromosome of the channel catfish, *Ictalurus punctatus* 70

Table 8. Summary of RNA-Seq paired-end reads from intestine tissues during *Edwardsiella ictaluri* infection before and after quality control 72

Table 9. Alternative splicing events predicted from intestine tissues after *Edwardsiella ictaluri* infection 72

Table 10. Differential alternative splicing (DAS) events at each time point after infection with *Edwardsiella ictaluri* as compared with the control. 74

Table 11. PANTHER overrepresentation analysis of differentially alternatively spliced (DAS) genes at 24 h after infection with *Edwardsiella ictaluri* 75

Table 12. PANTHER overrepresentation analysis of differentially alternatively spliced (DAS) genes at 72 h after infection with *Edwardsiella ictaluri* 76

List of Figures

Figure 1. Backcross strategy for the development of the experimental population..... 21

Figure 2. Mortality rate of the first-generation backcrosses after *Edwardsiella ictaluri* infection. 26

Figure 3. Sample structure identified by principal component analysis with two major principal components. 27

Figure 4. Manhattan plot of genome-wide association analysis for ESC disease resistance generated from EMMAX (A) and QFAM (B). 30

Figure 5. Genes within the significant region associated with ESC resistance on LG1 36

Figure 6. Genes within the significant region associated with ESC resistance on LG23. 37

Figure 7. Signal transduction pathway involved in Fcγ receptor-mediated phagocytosis. 42

Figure 8. Signal transduction pathway leading to T-cell activation. 43

Figure 9. Alternative splicing profile in channel catfish, *Ictalurus punctatus*. 70

Figure 10. Changes of alternative splicing after *Edwardsiella ictaluri* infection. 73

Figure 11. Enrichment of functional categories in differentially alternatively spliced (DAS) genes after infection with *Edwardsiella ictaluri*. 75

Figure 12. Schematic presentation of the alternative splicing isoforms of the catfish *srsf2* gene. 77

Figure 13. Quantitative real-time PCR of seven differential alternative splicing (DAS) events of six genes. 80

List of Abbreviations

AS	Alternative splicing
DAS	Differential alternative splicing
EMMAX	Efficient Mixed-Model Association eXpedited
ESC	Enteric septicemia of catfish
GWAS	Genome-wide association study
IBS	Identity by state
LD	Linkage disequilibrium
LG	Linkage group
MAF	Minimum allele frequency
PCA	Principal component analysis
QFAM	Family-based association tests for quantitative traits
qPCR	Quantitative real-time PCR
QTL	Quantitative trait loci
SNP	Single nucleotide polymorphism

Chapter 1 Introduction

1.1 The catfish industry

Ictalurid catfish are primary aquaculture species in the United States with respect to the production and economic impact. The catfish production accounted for over 50% of all U.S. aquaculture production (NMFS 2016). In 2017, catfish producers had sales of $380 million with a production quantity of 342 million pounds (www.nass.usda.gov). Most of the U.S. catfish are produced in four states, including Mississippi, Alabama, Arkansas, and Texas. Channel catfish (*Ictalurus punctatus*) and blue catfish (*I. furcatus*) were the most and second most cultured catfish in the United States, respectively. In recent years, the interspecific hybrid (channel catfish female × blue catfish male) has become the best catfish for culture due to desirable traits, including growth (Dunham et al. 1987, 1990, 2008), disease resistance (Wolters et al. 1996; Arias et al. 2012), hypoxia tolerance (Dunham et al. 1983), fillet yield (Argue et al. 2003), and seinability (Smitherman and Dunham 1985). The hybrid now represents over 70% of U.S. catfish production (Dunham and Elaswad 2018).

The catfish industry has been shrinking in recent years, with over 50% decreasing from 2003 (Hanson and Sites 2015). Encountered challenges consist of fierce international competition, increased feed and energy costs, and devastating disease (Plumb and Hanson 2011; Hanson and

Sites 2015). The three most important diseases threatening the catfish industry are enteric septicemia of catfish (ESC), columnaris disease, and motile Aeromonas septicemia (Aeromonas disease), which are caused by the Gram-negative bacterial pathogens, *Edwardsiella ictaluri*, *Flavobacterium columnare*, and *Aeromonas hydrophila*, respectively. Among them, ESC caused by *E. ictaluri* is the leading fish disease affecting farm-raised catfish.

1.2 Enteric septicemia of catfish

Enteric septicemia of catfish (ESC) was first reported in 1976 (Hawke et al. 1981). It is the most prevalent and devastating disease in catfish farming (Hawke 1979; Hawke et al. 1981; Plumb and Hanson 2011), leading to economic loss of $40–60 million in the U.S. catfish industry annually (Shoemaker et al. 2009). In addition to channel catfish, cases have been reported worldwide in white catfish (*Ameiurus catus*), brown bullhead catfish (*Ameiurus nebulosus*), yellow catfish (*Pylodictis olivaris*), striped catfish (*Pangasianodon hypophthalmus*), walking catfish (*Clarias batrachus*), and species of fish other than catfish (Hawke et al. 1981; Kasornchandra et al. 1987; Ye et al. 2009; Dong et al. 2015). ESC occurs most often within a specific temperature range between 20 and 28 °C (Hawke et al. 1998). This disease causes enteritis and septicemia, characterized by ulcerative skin lesions, petechial hemorrhages around the mouth and abdomen, and pimples or holes between eyes (Areechon and Plumb 1983; Hawke et al. 1998). The causative agent *E. ictaluri* is a Gram-negative, rod-shaped bacterium of the family Enterobacteriacae. *E. ictaluri* can enter catfish through the intestinal tract, the nares, the gills, and the skin (Hawke et al.

1998; Menanteau-Ledouble et al. 2011). *E. ictaluri* is internalized in leukocytes of channel catfish, particularly in macrophages (Miyazaki and Plumb 1985; Shotts et al. 1986; Baldwin and Newton 1993). The initial detection of *E. ictaluri* in the internal organs can occur as early as 15 min post infection by gastric intubation (Baldwin and Newton 1993). Bacteria are released in the blood, and colonize other organs within 48 h (Menanteau-Ledouble 2009). The earliest lesions were observed microscopically at 2 days after an immersion challenge of channel catfish (Newton et al. 1989).

ESC disease control in catfish industry can be achieved through different programs, including feed restriction, vaccination, antibiotics, and genetic improvement (Wise et al. 2004; Elaswad and Dunham 2017). Wise and Johnson (1998) reported that feeding frequency significantly affected survival of channel catfish exposed to *E. ictaluri* and supported the observation that restricting feed during ESC epizootic is an effective way to reduce the impact of disease. Vaccination could provide disease control because it activates immune responses by mimicking the natural infection. The commercial ESC vaccine Aquavac-ESC is available for channel catfish, but it has not been widely used due to little economic return (Bebak and Wagner 2012). Recently, another live attenuated *E. ictaluri* vaccines have been developed for protecting catfish fry and fingerlings (Wise et al. 2015; Abdelhamed et al. 2018). Vaccination could only provide short-term disease control, and some vaccines quite effective in the laboratory were not very useful in the commercial pond environment (Dunham and Elaswad 2018). Use of antibiotic-added feed is another means to control ESC (Plumb et al. 1995). Adding feed with the antibiotics, florfenicol, was shown to be effective to decrease the mortality associated with ESC in channel

catfish, blue catfish, and their hybrids (Gaunt et al. 2015). However, extensive use of antibiotics was associated with the development of antibiotic resistance in *E. ictaluri* (De La Pena and Espinosa-Mansilla 2009). Moreover, it may lead to environmental problems and antibiotic residues in fish products. Genetic improvement of disease resistance in catfish will provide long-term protection, which can be inherited from generation to generation. Approaches of family selection (Bilodea-Bourgeois et al. 2007), intraspecific crossbreeding (Dunham and Smitherman 1985), interspecific hybridization (Wolters et al. 1996), and gene transfer (Dunham et al. 2002) were demonstrated to improve the resistance to ESC in catfish. However, the molecular mechanism underlying ESC disease resistance is not well known.

1.3 Genome-wide association study

Genome-wide association study (GWAS) is based upon the principle of linkage disequilibrium (LD) which is the nonrandom association between alleles at different loci. It is created by evolutionary forces such as mutation, drift, and selection and is broken down by recombination (Hartl et al. 1997). The relationship between phenotypic differences in individuals and genetic polymorphism has been of fundamental biological interest. GWAS can link genotype and phenotype which examine genome-wide set of genetic variants in individuals to find variants that associate with a trait. The traits not only include whole organism phenotypes such as length, weight, milk production, reproduction rate and disease resistance, but also include phenotype such as mRNA transcript abundances of a particular gene (Hernandez et al. 2012; Nicolae et al. 2010).

Many types of molecular markers have been used in traditional quantitative trait loci (QTL) mapping, including random amplified polymorphic DNA (RAPD), restriction fragment length polymorphism (RFLP), amplified fragment length polymorphism (AFLP), simple sequence repeats (SSRs, or microsatellites), and single nucleotide polymorphisms (SNPs) (Collard et al. 2005). However, not all of them are suitable for GWAS. SNPs are now the markers of choice for two reasons. First, SNPs are the most abundant genetic variations widely distributed throughout the genomes of most species. For example, the average distance between two nearby SNPs is approximately 116 bp in channel catfish (Sun et al. 2014). Second, they are generally bi-allelic polymorphisms amenable to automated genotyping. GWAS was pioneered more than a decade ago in human genetics. The first successful GWAS was published in 2005, and it identified two SNPs significantly associated with age-related macular degeneration (Klein et al. 2005). With the development of high-density SNP arrays and reduced cost in genotyping, the quantity and quality of QTL analysis were significantly enhanced in human beings, crop, and animals (Manolio 2013).

The ideal samples used in GWAS should be homogeneous without population stratification and highly contrasted in trait of interest. Population stratification is generated from different allele frequencies among subpopulations, resulting in false positive results. To deal with this problem, experimental population structure and statistical analyses have been designed. Samples from natural population and family-based population can be used in GWAS (Ionita-Laza et al. 2013). Population-based design is time-saving and cost-effective. However, it is easy to obtain false positive results in the presence of population stratification, because individuals could be derived

from a limited number of founders to form subpopulations. Recent developments in statistical models have accounted for the problem using natural population. The family-based design is more resistant to population stratification, since individuals from a few families are used rather than unrelated sample. Individuals in family-based population have relatively homogenous genetic background, making correction of population stratification easier compared with the natural population design. Some disadvantages of family-based design should be concerned. For instance, the causal alleles throughout the entire population may not be detected in families because they may be homozygous in relatively small size of family-based samples. Moreover, it is time-consuming to generate families, e.g., a generation of catfish takes three to four years.

1.4 Alternative splicing

Eukaryotic cells mainly rely on gene expression regulation to maintain homeostasis. Regulation of gene expression includes a wide variety of mechanisms, from chromatin structure remodeling, to transcriptional and post-transcriptional regulation, and to translational and post-translational regulation. One of the important post-transcriptional regulation is alternative splicing (AS), which is a pivotal step in the generation of proteomic diversity and fine tuning of gene regulation in eukaryotes (Stamm et al. 2005). The concept of AS was first introduced in the 1970s. Exons and introns were discovered in the Adenovirus *hexon* gene in 1977 (Sambrook 1997). Walter Gilbert proposed that exons could be alternatively spliced together to generate multiple mRNA isoforms (Gilbert 1978).

Previous studies indicated that levels of mRNA and protein expression were not linearly correlated in eukaryotes, suggesting post-transcriptional regulation including AS (Ghazalpour et al. 2011). Moreover, transcriptional regulation and AS were indicated to be largely independent processes in the same tissues or cells, highlighting the roles of AS in regulation of gene expression (Pan et al. 2008).

AS allows for the creation of multiple distinct mRNA transcripts from a single gene (Graveley 2001). Based on the changes in splice sites and the resulted spliced RNA products, AS can be classified into five basic types: exon skipping, mutually exclusive exon, alternative 5' splice site, alternative 3' splice site, and intron retention (Graveley 2001). In humans, over 90% of genes have been reported to be alternatively spliced (Pan et al. 2008; Wang et al. 2008). Similarly, more than 60% of multi-exon genes underwent AS in plants (Syed et al. 2012). One extreme example of AS was the *Dscam* gene of *Drosophila melanogaster* with more than 38,000 alternatively spliced variants (Graveley 2005).

Splice site selection is determined by both *cis* and *trans* elements: splicing factors bind specific sequences in the precursor mRNAs (pre-mRNAs) called exon and intron splicing enhancers (ESEs and ISEs, respectively) or silencers (ESSs and ISSs, respectively), affecting the recruitment of spliceosome components to respective splice sites (Matlin et al. 2005; Nilsen and Graveley 2010). The transcript variants and protein isoforms produced by pre-mRNA splicing may have different stabilities, sub-cellular localizations, and functions (Wang et al. 2008). In many cases, each mRNA transcript has the potential to encode a unique protein, often with alternations

in localization, activity and function (Kalam et al. 2017). In other cases, some mRNA transcripts are degraded by the nonsense-mediated decay (NMD) to eliminate abnormal transcripts and modulate levels of normal gene expression (Lejeune and Maquat 2005; Maquat 2004; Wollerton et al. 2004). Approximately a third of alternatively spliced variants were targets of NMD (Lewis et al. 2003). In the splicing process, the spliceosome leaves behind protein marks called exon junction complexes (EJCs) at exon-exon junctions of the spliced mRNA, facilitating mRNA export (Le Hir et al. 2000, 2001; Green et al. 2003). During the pioneer round of mRNA translation, the ribosome displaces the EJCs (Ishigaki et al. 2001). NMD would be elicited if one or more EJC remains on the mRNA. In other words, NMD is triggered if translation terminates > 50-55 nucleotides upstream of the final exon-exon junction of an mRNA (Ishigaki et al. 2001; Green et al. 2003; Lejeune and Maquat 2005; Ni et al. 2007).

References

Abdelhamed H, Lawrence ML, Karsi A (2018) Development and characterization of a novel live attenuated vaccine against enteric septicemia of catfish. Front Microbiol 9:1819

Areechon N, Plumb JA (1983) Pathogenesis of *Edwardsiella ictaluri* in channel catfish, *Ictalurus punctatus*. J World Aquac Soc 14:249–260

Argue BJ, Liu Z, Dunham RA (2003) Dress-out and fillet yields of channel catfish, *Ictalurus punctatus*, blue catfish, *Ictalurus furcatus*, and their F1, F2 and backcross hybrids. Aquaculture 228:81–90

Arias CR, Cai W, Peatman E, Bullard SA (2012) Catfish hybrid *Ictalurus punctatus* × *I. furcatus* exhibits higher resistance to columnaris disease than the parental species. Dis Aquat Organ 100:77–81

Baldwin TJ, Newton JC (1993) Pathogenesis of enteric septicemia of channel catfish, caused by *Edwardsiella ictaluri*: bacteriologic and light and electron microscopic findings. J Aquat Anim Health 5:189–198

Bebak J, Wagner B (2012) Use of vaccination against enteric septicemia of catfish and columnaris disease by the US catfish industry. J Aquat Anim Health 24:30–36

Bilodeau-Bourgeois AL, Bosworth BG, Wolters WR (2007) Reductions in susceptibility of channel catfish, *Ictalurus punctatus*, to enteric septicemia of catfish through two generations of selection. J World Aquac Soc 38:450–453

Collard BC, Jahufer MZ, Brouwer JB, Pang EC (2005) An introduction to markers, quantitative trait loci (QTL) mapping and marker-assisted selection for crop improvement: the basic concepts. Euphytica 142:169–196

De La Pena AM, Espinosa-Mansilla A (2009) Analysis of antibiotics in fish samples. Anal Bioanal Chem 395:987–1008

Dong HT, Nguyen VV, Phiwsaiya K, Gangnonngiw W, Withyachumnarnkul B, Rodkhum C, Senapin S (2015) Concurrent infections of *Flavobacterium columnare* and *Edwardsiella ictaluri* in striped catfish, *Pangasianodon hypophthalmus* in Thailand. Aquaculture 448:142–150

Dunham RA, Brummett RE, Ella MO, Smitherman RO (1990) Genotype-environment interactions for growth of blue, channel and hybrid catfish in ponds and cages at varying densities. Aquaculture 85:143–151

Dunham RA, Elaswad A (2018) Catfish Biology and Farming. Annu Rev Anim Biosci 6:305–325

Dunham RA, Smitherman RO (1985) Improved growth rate, reproductive performance, and disease resistance of crossbred and selected catfish from AU-M and AU-K lines. Circular 279, Alabama Agricultural Experiment Station, Auburn University, Auburn University, AL

Dunham RA, Smitherman RO, Goodman RK (1987) Comparison of mass selection, crossbreeding, and hybridization for improving growth of channel catfish. Prog Fish Cult 49:293–296

Dunham RA, Smitherman RO, Webber C (1983) Relative tolerance of channel × blue hybrid and channel catfish to low oxygen concentrations. Prog Fish Cult 45:55–57

Dunham RA, Umali GM, Beam R, Kristanto AH, Trask M (2008) Comparison of production traits of NWAC103 channel catfish, NWAC103 channel catfish × blue catfish hybrids, Kansas Select 21 channel catfish, and blue catfish grown at commercial densities and exposed to natural bacterial epizootics. N Am J Aquac 70:98–106

Dunham RA, Warr GW, Nichols A, Duncan PL, Argue B, Middleton D, Kucuktas H (2002) Enhanced bacterial disease resistance of transgenic channel catfish *Ictalurus punctatus* possessing cecropin genes. Mar Biotechnol 4:338–344

Elaswad A, Dunham RA (2017) Disease reduction in aquaculture with genetic and genomic technology: current and future approaches. Rev Aquac 0:1–23

Gaunt PS, Chatakondi N, Gao D, Endris R (2015) Efficacy of florfenicol for control of mortality associated with *Edwardsiella ictaluri* in three species of catfish. J Aquat Anim Health 27:45–49

Ghazalpour A, Bennett B, Petyuk VA, Orozco L, Hagopian R, Mungrue IN, Farber CR, Sinsheimer J, Kang HM, Furlotte N, Park CC, Wen PZ, Brewer H, Weitz K, Camp DG, Pan C, Yordanova R, Neuhaus I, Tilford C, Siemers N, Gargalovic P, Eskin E, Kirchgessner T,

Smith DJ, Smith RD, Lusis AJ (2011) Comparative analysis of proteome and transcriptome variation in mouse. PLoS Genet 7:e1001393

Gilbert W (1978) Why genes in pieces?. Nature 271:501

Graveley BR (2001) Alternative splicing: increasing diversity in the proteomic world. Trends Genet 17:100–107

Graveley BR (2005) Mutually exclusive splicing of the insect *Dscam* pre-mRNA directed by competing intronic RNA secondary structures. Cell 123:65–73

Green RE, Lewis BP, Hillman RT, Blanchette M, Lareau LF, Garnett AT, Rio DC, Brenner SE (2003) Widespread predicted nonsense-mediated mRNA decay of alternatively-spliced transcripts of human normal and disease genes. Bioinformatics 19:i118–i121

Hanson T, Sites D (2015) 2014 U.S. Catfish Database. USDA National Agricultural Statistics Service (NASS), Mississippi Agricultural Statistics Service (MASS)

Hartl DL, Clark AG (1997) Principles of Population Genetics, 3rd edn. Sinauer Associates, Inc., Sunderland, Massachusetts

Hawke JP (1979) A bacterium associated with disease of pond cultured channel catfish, *Ictalurus punctatus*. J Fish Board Canada 36:1508–1512

Hawke JP, Durborow RM, Thune RL, Camus AC (1998) ESC: enteric septicemia of catfish. Southern Regional Aquaculture Center Stoneville, Mississippi

Hawke JP, McWhorter AC, Steigerwalt AG, Brenner DJ (1981) *Edwardsiella ictaluri* sp. nov., the causative agent of enteric septicemia of catfish. Int J Syst Bacteriol 31:396–400

Hernandez DG, Nalls MA, Moore M, Chong S, Dillman A, Trabzuni D, Gibbs JR, Ryten M, Arepalli S, Weale ME, Zonderman AB, Troncoso J, O'Brien R, Walker R, Smith C, Bandinelli S, Traynor BJ, Hardy J, Singleton AB, Cookson MR (2012) Integration of GWAS SNPs and tissue specific expression profiling reveal discrete eQTLs for human traits in blood and brain. Neurobiol Dis 47:20–28

Ionita-Laza I, Lee S, Makarov V, Buxbaum JD, Lin X (2013) Family-based association tests for sequence data, and comparisons with population-based association tests. Eur J Hum Genet 21:1158

Ishigaki Y, Li X, Serin G, Maquat LE (2001) Evidence for a pioneer round of mRNA translation: mRNAs subject to nonsense-mediated decay in mammalian cells are bound by CBP80 and CBP20. Cell 106:607–617

Kalam H, Fontana MF, Kumar D (2017) Alternate splicing of transcripts shape macrophage response to *Mycobacterium tuberculosis* infection. PLoS Pathog 13:e1006236

Kasornchandra J, Rogers WA, Plumb JA (1987) *Edwardsiella ictaluri* from walking catfish, *Clarias batrachus* L., in Thailand. J Fish Dis 10:137–138

Klein RJ, Zeiss C, Chew EY, Tsai JY, Sackler RS, Haynes C, Henning AK, SanGiovanni JP, Mane SM, Mayne ST, Bracken MB, Ferris FL, Ott J, Barnstable C, Hoh J (2005) Complement factor H polymorphism in age-related macular degeneration. Science 308:385–389

Le Hir H, Gatfield D, Izaurralde E, Moore MJ (2001) The exon–exon junction complex provides a binding platform for factors involved in mRNA export and nonsense-mediated mRNA decay. EMBO J 20:4987–4997

Le Hir H, Moore MJ, Maquat LE (2000) Pre-mRNA splicing alters mRNP composition: evidence for stable association of proteins at exon–exon junctions. Genes Dev 14:1098–1108

Lejeune F, Maquat LE (2005) Mechanistic links between nonsense-mediated mRNA decay and pre-mRNA splicing in mammalian cells. Curr Opin Cell Biol 17:309–315

Lewis BP, Green RE, Brenner SE (2003) Evidence for the widespread coupling of alternative splicing and nonsense-mediated mRNA decay in humans. Proc Natl Acad Sci 100:189–192

Manolio TA (2013) Bringing genome-wide association findings into clinical use. Nat Rev Genet 14:549–558

Maquat LE (2004) Nonsense-mediated mRNA decay: splicing, translation and mRNP dynamics. Nat Rev Mol cell Biol 5:89–99

Matlin AJ, Clark F, Smith CWJ (2005) Understanding alternative splicing: towards a cellular code. Nat Rev Mol cell Biol 6:386–398

Menanteau-Ledouble S (2009) Pathogen entrance and development of disease during infection of the American channel catfish *Ictalurus punctatus* by the enterobacterium *Edwardsiella ictaluri*. Dissertation, Mississippi State University

Menanteau-Ledouble S, Karsi A, Lawrence ML (2011) Importance of skin abrasion as a primary site of adhesion for *Edwardsiella ictaluri* and impact on invasion and systematic infection in channel catfish *Ictalurus punctatus*. Vet Microbiol 148:425–430

Miyazaki T, Plumb JA (1985) Histopathology of *Edwardsiella ictaluri* in channel catfish, *Ictalurus punctatus* (Rafinesque). J Fish Dis 8:389–392

Newton JC, Wolfe LG, Grizzle JM, Plumb JA (1989) Pathology of experimental enteric septicaemia in channel catfish, *Ictalurus punctatus* (rafinesque), following immersion-exposure to *Edwardsiella ictaluri*. J Fish Dis 12:335–347

Ni JZ, Grate L, Donohue JP, Preston C, Nobida N, O'Brien G, Shiue L, Clark TA, Blume JE, Ares M (2007) Ultraconserved elements are associated with homeostatic control of splicing regulators by alternative splicing and nonsense-mediated decay. Genes Dev 21:708–718

Nicolae DL, Gamazon E, Zhang W, Duan S, Dolan ME, Cox NJ (2010) Trait-associated SNPs are more likely to be eQTLs: annotation to enhance discovery from GWAS. PLoS Genet 6:e1000888

Nilsen TW, Graveley BR (2010) Expansion of the eukaryotic proteome by alternative splicing. Nature 463:457–463

NMFS (National Marine Fisheries Service) (2015) Fisheries of the United States 2014. NMFS, Office of Science and Technology, Silver Spring, Maryland

Pan Q, Shai O, Lee LJ, Frey BJ, Blencowe BJ (2008) Deep surveying of alternative splicing complexity in the human transcriptome by high-throughput sequencing. Nat Genet 40:1413–1415

Plumb JA, Hanson LA (2011) Health maintenance and principal microbial diseases of cultured fishes. John Wiley & Sons

Plumb JA, Sheifinger CC, Shryock TR, Goldsby T (1995) Susceptibility of six bacterial pathogens of channel catfish to six antibiotics. J Aquat Anim Health 7:211–217

Sambrook J (1977) Adenovirus amazes at cold spring harbor. Nature 268:101

Shoemaker CA, Klesius PH, Evans JJ, Arias CR (2009) Use of modified live vaccines in aquaculture. J World Aquacult Soc 40:573–585

Shotts EB, Blazer VS, Waltman WD (1986) Pathogenesis of experimental *Edwardsiella ictaluri* infections in channel catfish (*Ictalurus punctatus*). Can J Fish Aquat Sci 43:36–42

Smitherman RO, Dunham RA (1985) Genetics and breeding. In: Channel Catfish Culture, Elsevier Scientific Publishing, Amsterdam, pp 283–316

Stamm S, Ben-Ari S, Rafalska I, Tang YS, Zhang ZY, Toiber D, Thanaraj TA, Soreq H (2005) Function of alternative splicing. Gene 344:1–20

Sun L, Liu S, Wang R, Jiang Y, Zhang Y, Zhang J, Bao L, Kaltenboeck L, Dunham R, Waldbieser G, Liu Z (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

Syed NH, Kalyna M, Marquez Y, Barta A, Brown JWS (2012) Alternative splicing in plants–coming of age. Trends Plant Sci 17:616–623

Wang ET, Sandberg R, Luo S, Khrebtukova I, Zhang L, Mayr C, Kingsmore SF, Schroth GP, Burge CB (2008) Alternative isoform regulation in human tissue transcriptomes. Nature 456:470–476

Wise DJ, Camus AC, Schwedler TE, Terhune JS (2004) Health management. In: Tucker CS, Hargreaves JA (eds) Developments in Aquaculture and Fisheries Science, Elsevier, pp 444–503

Wise DJ, Greenway TE, Byars TS, Griffin MJ, Khoo LH (2015) Oral vaccination of channel catfish against enteric septicemia of catfish using a live attenuated *Edwardsiella ictaluri* isolate. J Aquat Anim Health 27:135–143

Wise DJ, Johnson MR (1998) Effect of feeding frequency and Romet-medicated feed on survival, antibody response, and weight gain of fingerling channel catfish *Ictalurus punctatus* after natural exposure to *Edwardsiella ictaluri*. J World Aquac Soc 29:169–175

Wollerton MC, Gooding C, Wagner EJ, Garcia-Blanco MA, Smith CWJ (2004) Autoregulation of polypyrimidine tract binding protein by alternative splicing leading to nonsense-mediated decay. Mol Cell 13:91–100

Wolters WR, Wise DJ, Klesius PH (1996) Survival and antibody response of channel catfish, blue catfish, and channel catfish female × blue catfish male hybrids after exposure to *Edwardsiella ictaluri*. J Aquat Anim Health 8:249–254

Ye S, Li H, Qiao G, Li Z (2009) First case of *Edwardsiella ictaluri* infection in China farmed yellow catfish *Pelteobagrus fulvidraco*. Aquaculture 292:6–10

Chapter 2 A genome-wide association study in interspecific backcross catfish for QTL controlling enteric septicemia of catfish

2.1 Abstract

Infectious diseases pose significant threats to the catfish industry. Enteric septicemia of catfish (ESC) caused by *Edwardsiella ictaluri* is the most devastating disease for catfish aquaculture, causing huge economic losses annually. Channel catfish and blue catfish exhibit great contrast in resistance against ESC, with channel catfish being highly susceptible and blue catfish being highly resistant. As such, the interspecific backcross progenies provide an ideal system for the identification of quantitative trait locus (QTL). We previously reported one significant QTL on linkage group (LG) 1 using the third-generation backcrosses, but the number of founders used to make the second and third-generation backcross progenies were very small. Although the third-generation backcross progenies provided a greater power for fine mapping than the first-generation backcrosses, some major QTL for disease resistance may have been missing due to the small numbers of founders used to produce the higher-generation backcrosses. In this study, we performed a genome-wide association study (GWAS) using first-generation backcrosses with the catfish 690K SNP arrays to identify additional ESC disease resistance QTL, especially those at the species level. Two genomic regions on LG1 and LG23 were determined to be significantly

associated with ESC resistance as revealed by a mixed linear model and family-based association test. Examination of the resistance alleles indicated their origin from blue catfish, indicating that at least two major disease resistance loci exist among blue catfish populations. Upon further validation, markers linked with major ESC disease resistance QTL should be useful for marker-assisted introgression, allowing development of highly ESC resistant breeds of catfish.

2.2 Background

Disease resistance describes the overall capacity and efficiency of the host to cope with infections, which is composed of the innate and adaptive immunities. The innate immune system provides physical and chemical barriers, senses pathogens through various receptors to activate microbial defense and stimulate the adaptive immune response (Medzhitov 2007). The adaptive immune system comprises of T and B lymphocytes and their mediated processes. T lymphocytes, including T-helper (Th) cells ($CD4^+$ T cells) and cytotoxic T cells ($CD8^+$ T cells), recognize antigenic peptides bound to major histocompatibility complex (MHC) via antigen-specific extracellular receptors. Cytotoxic T cells induce apoptosis of infected cells or kill target cells (Andersen et al. 2006), while Th cells differentiate into effector cells (Th1, Th2, and Th17) upon antigen stimulation, producing distinct sets of cytokines for antigen-specific activation of innate host-defense cells such as macrophage and granulocytes as well as for the activation of B cells to generate antigen-specific antibodies (Medzhitov 2007; Annunziato et al. 2009). Additionally, B cells can recognize antigens by directly binding to them. The various modules of innate immunity

cooperate with cell-mediated and humoral immune responses of adaptive immunity, providing effective protection against infection for vertebrates.

Immunogenetic analysis in fish species has been rare. The practical approaches have focused on the identification of disease resistance-associated DNA markers for marker-assisted selection. The interest for the identification of genetic markers and genomic regions for resistance against diseases in aquaculture species is increasing because rapid progress can be made in this area to reduce disease problems and enhance aquaculture production without deep analysis of the immunological and molecular mechanisms. Among various approaches, quantitative trait locus (QTL) mapping and genome-wide association study (GWAS) have been the most useful approaches for disease resistance studies with aquaculture species.

Disease resistance QTL have been identified in various aquaculture species (For a recent review, see Abdelrahman et al. 2017). For example, QTL have been identified in rainbow trout (*Oncorhynchus mykiss*) for resistance against infectious pancreatic necrosis (Ozaki et al. 2007), infectious hamatopoietic necrosis (Campbell et al. 2014), whirling disease (Baerwald et al. 2011), and bacterial cold water disease (Liu et al. 2015; Palti et al. 2015). In Asian seabass (*Lates calcarifer*), QTL have been detected for viral nervous necrosis disease resistance (Liu et al. 2016a), and for iridovirus resistance (Wang et al. 2017a). For Japanese flounder (*Paralichthys olivaceus*), a single major genetic locus was found to explain 50% of the total variation of disease resistance to lymphocystis disease (Fuji et al. 2006), and the linked DNA marker was successfully used for marker-assisted selection (Fuji et al. 2007). In Atlantic salmon (*Salmo salar*), QTL associated with

resistance to Salmon Rickettsial Syndrome were identified (Correa et al. 2015). Moreover, a major QTL in Atlantic salmon that explained most of the genetic variance for resistance to infectious pancreatic necrosis has been commercially utilized by breeding companies (Houston et al. 2008; Moen et al. 2009).

Catfish is the primary aquaculture species in the United States. Channel catfish (*Ictalurus punctatus*) used to be the major cultured species. In recent years, the interspecific hybrid [channel catfish female × blue catfish (*Ictalurus furcatus*) male] is increasingly used by the catfish industry because of its superiority for disease resistance, growth rate, fillet yield, and harvestability (Dunham et al. 2008; Lucas and Southgate 2012). The interspecific hybrids not only promote the aquaculture production but also serve as an excellent model for the analysis of QTL associated with traits such as disease resistance. With enteric septicemia of catfish (ESC), blue catfish is very resistant while channel catfish is highly susceptible (Wolters and Johnson 1994; Wolters et al. 1996). Therefore, the interspecific hybrid provides an ideal system for genetic analysis of ESC disease resistance.

Genomic resources of catfish have been made available to facilitate GWAS in catfish, including a large number of SNPs (Liu et al. 2011; Sun et al. 2014), the high-density SNP arrays (Liu et al. 2014; Zeng et al. 2017), and the reference genome sequence (Liu et al. 2016b). Using the 250K SNP array, several GWAS analyses have been conducted with performance and production traits of catfish, including disease resistance against columnaris (Geng et al. 2015), hypoxia tolerance (Wang et al. 2017b; Zhong et al. 2017), heat tolerance (Jin et al. 2017), albinism

(Li et al. 2017b), body weight (Li et al. 2017a), body conformation (Geng et al. 2017a) and head size (Geng et al. 2016). With ESC disease resistance, GWAS analysis was conducted using the third-generation backcrosses (Zhou et al. 2017). The use of higher-generation backcrosses could have provided greater powers for the detection of QTL in a smaller genomic region because of accumulated chromosomal recombinations. However, some significant QTL may have been undetected if such QTL were not included in the few individuals of the first-generation backcrosses used as founders of the higher-generation backcrosses.

The QTL by definition must be there in the first generation. However, it would go undetected if the appropriate allelic variation was missing. In this study, we utilized the first-generation backcrosses for determining the genomic regions associated with ESC disease resistance, to ensure detection of species-level QTL, especially those offered from blue catfish.

2.3 Methods

2.3.1 Ethics statement

All the procedures involving the handling and treatment of fish were 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.

2.3.2 Experimental fish, bacterial challenge and sample collection

The bacterial challenge experiment was performed with one-year-old catfish (average 14.8 cm) produced by crossing male F1 hybrid catfish (channel catfish female × blue catfish male) with female channel catfish (Figure 1). The female channel catfish were Marion strain (Dunham and Smitherman 1984), including the maternal parent of the F1 hybrid catfish. Four backcross families were used, and the grandparents among families were different. A total of 1,200 experiential fish were randomly selected (300 fish per family) at the Auburn University Fish Genetics Facility and were distributed in four tanks (300 L water) by family. A control group containing 200 fish was reared in another tank. The average temperature was maintained at 28℃, and oxygen level was controlled at optimal level by aeration.

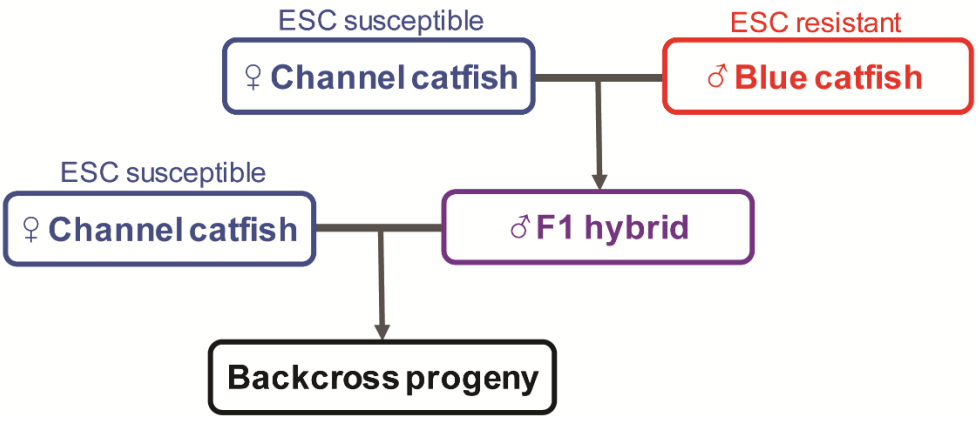

Figure 1. Backcross strategy for the development of the experimental population.

The infestation with *E. ictaluri* was carried out after acclimation. The bacteria challenge procedure was conducted as previously described (Wang et al. 2013). The bacteria were cultured from a single colony, re-isolated from a symptomatic fish, and biochemically confirmed. A single colony was cultured in brain heart infusion (BHI) medium and incubated at 28℃ overnight. The concentration of the bacteria was determined using colony forming unit (CFU) per mL by plating ten-fold serial dilution onto BHI agar plates. The infestation was conducted by immersion exposure for 2 h at a final bacterial concentration of 2×10^7 CFU/mL, whereas control fish were treated with an equal volume of sterile BHI medium. Moribund catfish were continuously sampled by taking blood, and the sampling time was recorded. Fish for further genotyping were selected from the extremes of the disease resistance distribution for each family based on the selective genotyping method (Darvasi and Soller 1992).

2.3.3 DNA isolation and genotyping

DNA was isolated using standard protocols. Briefly, the blood samples in cell lysis solution were incubated at 55°C overnight. Protease K and protein precipitation solution were used to remove proteins. Extracted DNA was precipitated with isopropanol and collected by brief centrifugation, washed twice with 70% ethanol, air-dried, and rehydrated in TE buffer (pH 8.0). DNA was quantified by spectrophotometry (NanoDrop Technologies, Wilmington, DE, USA) and checked by 1% agarose gel electrophoresis stained with ethidium bromide for integrity. DNA was diluted to 50 ng/μL with a final volume of 10 μL. Genotyping with the catfish 690K SNP array

was performed at GeneSeek (Lincoln, Nebraska, USA). Genotype callings were generated using the Axiom Analysis Suite software.

2.3.4 Statistical analysis

To identify association between SNPs and resistance to *E. ictaluri*, the single SNP test was performed using the SVS software package (SNP & Variation Suite, Version 8.3) and PLINK (Version 1.9) (Purcell et al. 2007). Quality control was performed for samples and markers. Samples with genotype missingness larger than 5% were removed. The markers with a minor allele frequency (MAF) lower than 0.05 or a call rate lower than 95% were excluded from the analysis. Linkage disequilibrium (LD) pruning was then conducted to achieve a set of independent SNPs with a window size of 50 SNPs, window increment of 5, and r^2 threshold of 0.5. With the independent SNP markers, identity-by-state (IBS) distances were estimated to reflect relatedness between pairs of samples, and the population structure was analyzed by principal component analysis (PCA).

EMMAX (Efficient Mixed-Model Association eXpedited) analyses using all qualified SNPs were conducted with the first three principal components and body weight as covariates (Kang et al. 2010). The model is listed as follows:

$$Y = X\beta + Zu + e$$

where Y is the vector of phenotype; β is the coefficient vector of fixed effects including first three principal components and fish body weight; u is the vector of the random effect, $Var(u) = G\sigma_g^2$,

where σ_g^2 is the additive genetic variance and G is the genomic kinship matrix using the IBS; e is the vector of random residuals; X is the matrix of fixed effects and Z is the matrix of random additive genetic effects.

The Manhattan plot of the $-\log_{10}$ (*P* value) was generated using the SVS software, with the genetic marker map according to channel catfish genome sequence (Version Coco1.2, Liu et al. 2016). The threshold *P* value for genome-wide significance was calculated based on Bonferroni-correction with estimated number of independent markers and LD blocks.

Family based association test for quantitative traits (QFAM) was conducted using PLINK (Purcell et al. 2007). QFAM partitions the genotypes into between- and within- family components, and the within-family component is free of sample structure (Fulker et al. 1999; Abecasis et al. 2000). The QFAM-Within procedure in PLINK performed a simple linear regression of phenotype on the within-family component of genotype, and then used permutations to correct for family structure. The model is

$$\hat{y}_{ij} = \mu + \beta_b b_i + \beta_w w_{ij}$$

where y_{ij} denotes the phenotype of individual j in family i; μ is the population mean; β_b is the vector of between-family effect and β_w is the vector of within family effect; $b_i = (\Sigma g_{ij})/n_i$ if parental genotypes are unknown (g_{ij}: genotype of the jth offspring in the ith family), and $b_i = (g_{iF} + g_{iM})/2$ if parental genotypes are available (g_{iF}: genotype of father, g_{iM}: genotype of mother); $w_{ij} = g_{ij} - b_i$.

2.3.5 Sequence analysis

Genes within ±1 Mb of associated regions were predicted from catfish genome sequence (Liu et al. 2016b) using FGENESH (Solovyev et al. 2006) and annotated by BLAST against the NCBI-nr database (ftp://ftp.ncbi.nlm.nih.gov/blast/db/). Synteny analyses and catfish genome information from the NCBI were utilized to provide supporting evidence for the proper annotation.

2.4 Results

2.4.1 Experimental fish and sample structure

A total of 1,200 catfish from four families were challenged with *E. ictaluri*, and 288 fish with extreme phenotype were chosen for genotyping. The mortality rate after infection is shown in Figure 2. The information of catfish samples utilized in this study is summarized in Table 1. PCA analysis was conducted using eigenvalues as coordinates to visualize the sample structure. As shown in Figure 3, each family was grouped into a separate cluster and the four families were distantly related.

Table 1. The catfish samples used in the GWAS analysis

Family ID	Dam	Sire	Sample number	Susceptible sample number	Resistant sample number
1	Channel 1	Hybrid 1	71	36	35
2	Channel 2	Hybrid 2	70	34	36
3	Channel 3	Hybrid 3	70	36	34
4	Channel 4	Hybrid 4	77	36	41

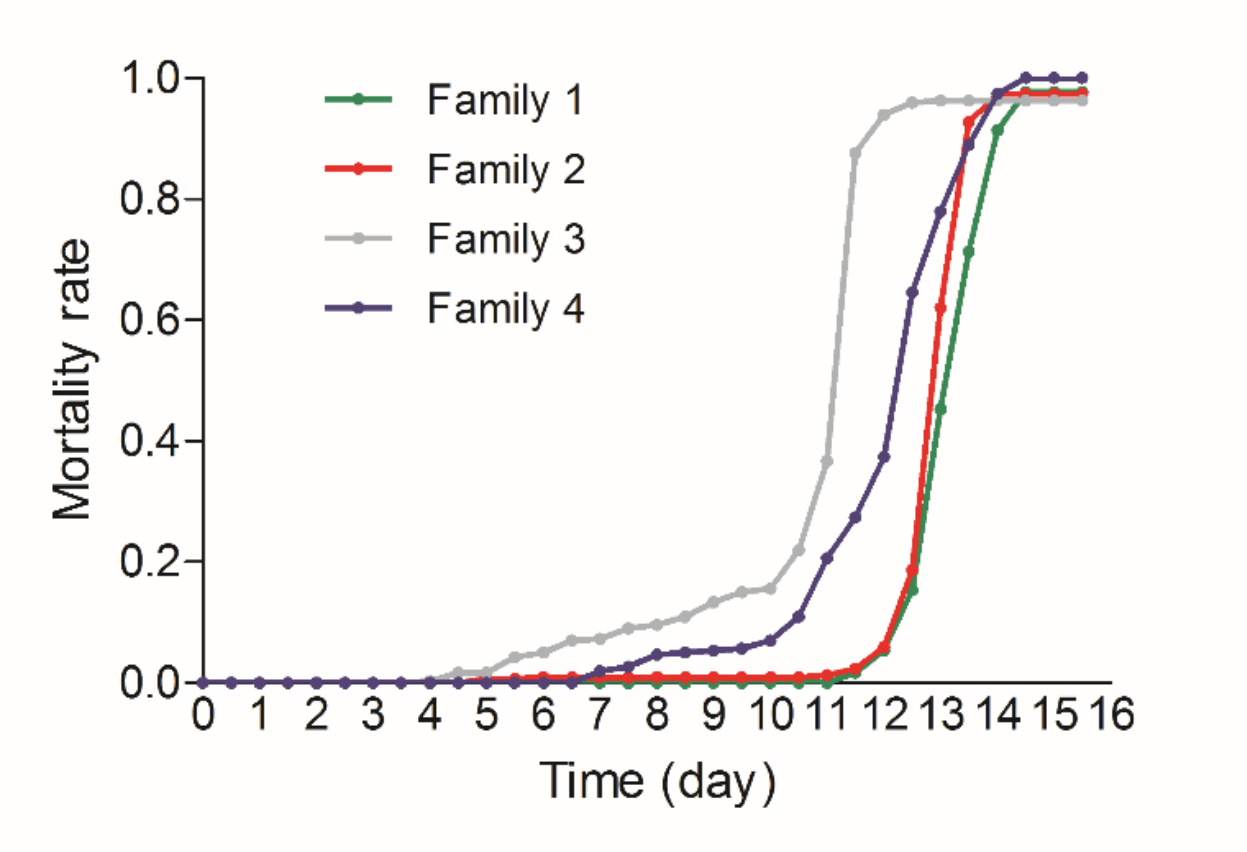

Figure 2. Mortality rate of the first-generation backcrosses after *Edwardsiella ictaluri* infection.

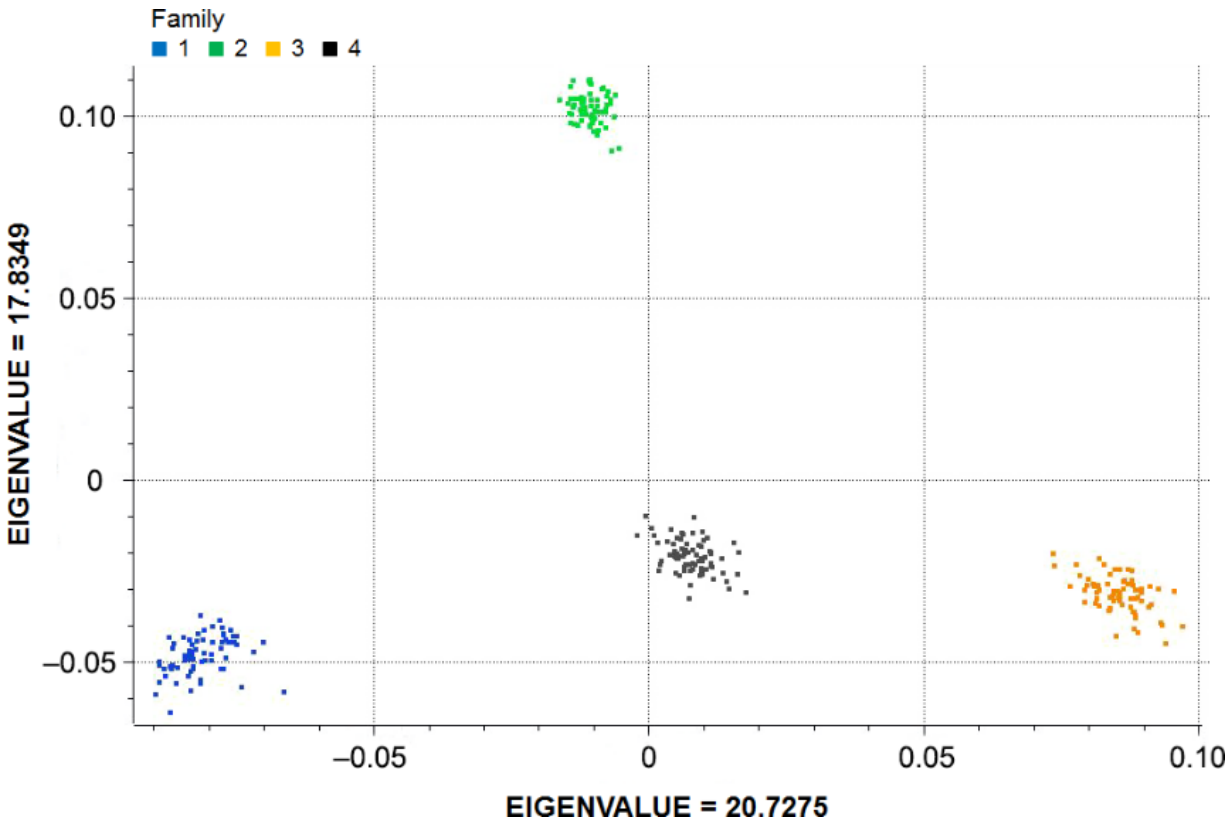

Figure 3. Sample structure identified by principal component analysis with two major principal components.

2.4.2 QTL for ESC resistance

A total of 407,196 SNPs were kept after filtering out SNP markers with genotyping errors, MAF < 0.05, or call rate < 95%. The number of independent SNPs and LD blocks was 15,061 after LD pruning, and the genome-wide threshold for statistical significance was calculated to be 3.32e–

6 (0.05/15,061) with $-\log_{10}$ (*P* value) = 5.48. The threshold of $-\log_{10}$ (*P* value) for suggestive association was arbitrarily set as 5 (Geng et al. 2016).

The Manhattan plots generated from EMMAX and QFAM are shown in Figure 4A and Figure 4B, respectively. In general, the association results of EMMAX and QFAM were positively correlated. EMMAX showed one genome-wide significant region on linkage group (LG) 23 and two suggestively associated regions on LG1 and LG14, while QFAM revealed two genome-wide significant regions on LG23 and LG1 as well as one suggestive region on LG1. Generally, QFAM is more robust in correcting family structure when using family-based samples than EMMAX (Geng et al. 2017b). In the following sections, we mainly describe the identified regions according to the results generated by QFAM.

SNPs associated with ESC resistance are listed in Table 2. Forty-five genome-wide significant SNPs were found on LG1 from 33,233,001 to 33,712,554, spanning approximately 479.6 Kb. The most significant SNP (AX-157732970) on LG1 reached the genome-wide significance level with $-\log_{10}$ (*P* value) = 6.05. Similarly, three significantly associated SNPs were located on LG23 in genomic region from 7,793,018 to 8,136,119, spanning approximately 343.1 Kb. The $-\log_{10}$ (*P* value) of the most significant SNP (AX-157695826) was 6.25. In addition, 20 SNPs were identified within the suggestive QTL on LG1 from 5,052,936 to 9,817,492. Based on the phenotype and the SNPs placed on the catfish 690K SNP array (Zeng et al. 2017), the resistance allele of associated SNPs (Table 2) as well as their origins could be determined. Four significant SNPs and four suggestively associated SNPs are all interspecific, which means they are SNPs from

interspecific origin and not intraspecific variants. For these interspecific SNPs, all resistance alleles originated from blue catfish. EMMAX was utilized to investigate the contribution of significantly associated QTL to the phenotypic variance. The most significant SNP was used to represent the corresponding significant region (Geng et al. 2015), with AX-157732970 on LG1 and AX-157695826 on LG23 accounting for 6.8% and 10.4% of the phenotypic variance, respectively. These two loci alone would indicate a minimum heritability of 0.17 in a first-generation backcross population. According to the phenotypic variance calculated in SVS (3827.5), fish with the "good" allele of AX-157732970 and AX-157695826 will live 16.1 h and 20.0 h longer, respectively.

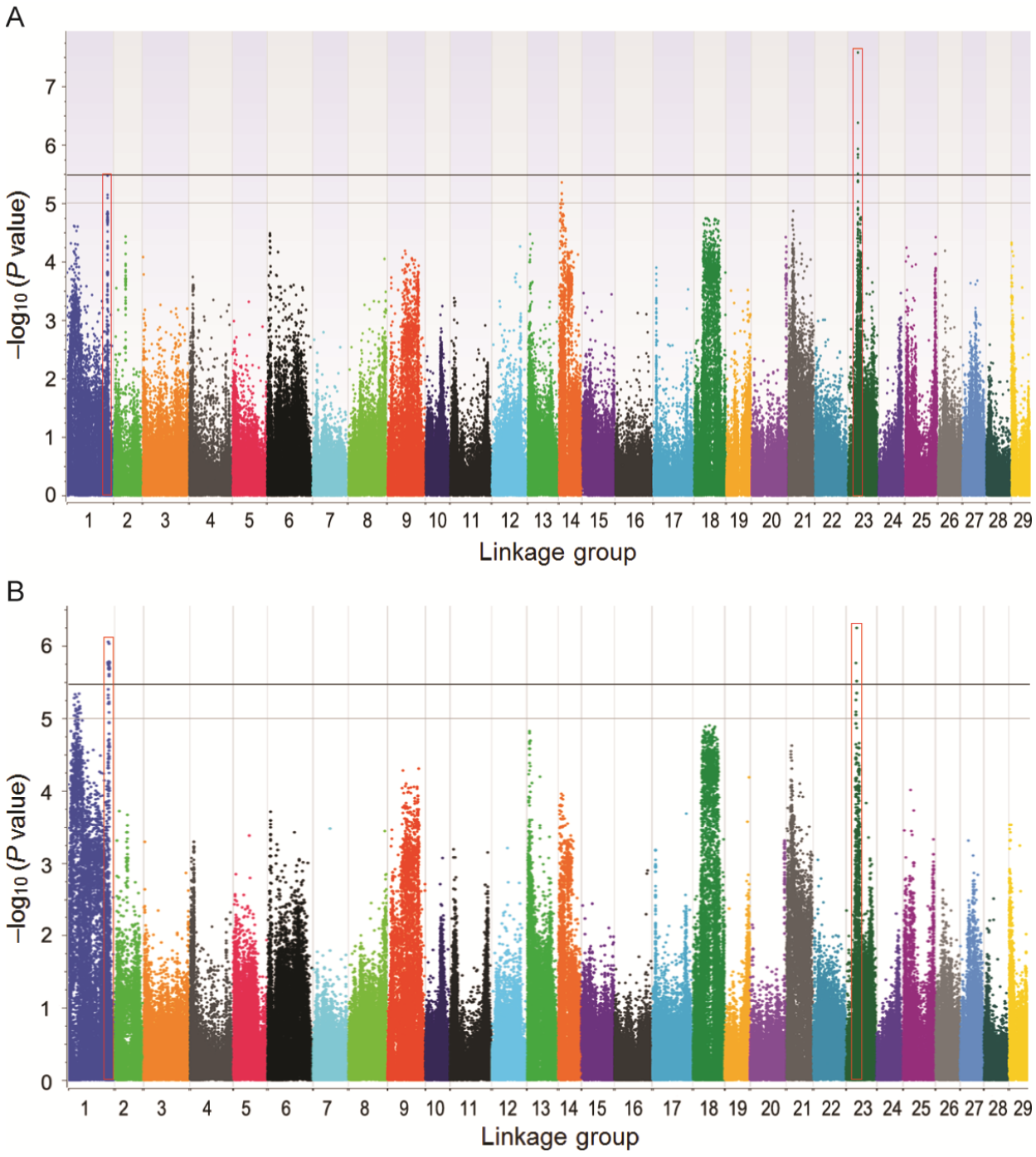

Figure 4. Manhattan plot of genome-wide association analysis for ESC disease resistance generated from EMMAX (A) and QFAM (B). The black solid line indicates the threshold *P*

value for genome-wide significance. The grey solid line indicates the threshold *P* value for significance of "suggestive association". The red box represents the shared QTL using two methods.

Table 2. SNPs associated with ESC disease resistance in the first-generation backcrosses. Significantly associated SNPs are indicated in bold.

Linkage group	SNP ID	Position (bp)	-log$_{10}$ (*P* Value)	SNP origin	SNP type	Resistant allele
1	**AX-157732970**	33308205	6.05	Interspecies	A/C	A
	AX-157907862	33712554	6.02	Channel catfish	A/G	G
	AX-157907656	33409686	5.78	Channel catfish	A/C	A
	AX-157907686	33447757	5.78	Channel catfish	A/G	A
	AX-157874495	33233001	5.77	Channel catfish	T/C	C
	AX-157907535	33236185	5.77	Channel catfish	T/C	C
	AX-157907548	33249228	5.77	Channel catfish	A/G	A
	AX-157777440	33264961	5.77	Blue catfish	T/C	T
	AX-157907565	33277191	5.77	Channel catfish	T/C	T
	AX-157907602	33331297	5.77	Channel catfish	A/G	G
	AX-157907622	33355721	5.77	Channel catfish	T/C	T
	AX-157907629	33370700	5.77	Channel catfish	A/G	A

AX-157907660	33412696	5.77	Channel catfish	T/G	G
AX-158389527	33419990	5.77	Channel catfish	T/G	T
AX-157907669	33423263	5.77	Channel catfish	A/G	G
AX-158404804	33463988	5.77	Blue catfish	C/G	G
AX-158404805	33464522	5.77	Blue catfish	C/G	C
AX-157907695	33466925	5.77	Channel catfish	A/C	C
AX-157907706	33482124	5.77	Channel catfish	A/G	G
AX-157907676	33431026	5.75	Channel catfish	T/C	T
AX-157765607	33286616	5.75	Interspecies	T/C	C
AX-157737553	33301636	5.75	Interspecies	T/C	C
AX-158404077	33464303	5.75	Blue catfish	A/G	G
AX-157720493	33266726	5.73	Channel catfish	A/G	A
AX-157777442	33541569	5.72	Blue catfish	A/G	A
AX-157907712	33488813	5.69	Channel catfish	A/G	G
AX-157907716	33495093	5.69	Channel catfish	A/G	G
AX-157907737	33518488	5.69	Channel catfish	A/G	A
AX-157907739	33519681	5.69	Channel catfish	A/G	G
AX-157769590	33554011	5.69	Interspecies	T/C	T
AX-157777443	33564726	5.69	Blue catfish	A/G	A

AX-158379196	33570168	5.69	Channel catfish	A/G	G
AX-157907782	33574139	5.69	Channel catfish	T/C	T
AX-157907785	33579001	5.69	Channel catfish	T/G	G
AX-157772089	33587740	5.69	Channel catfish	A/G	G
AX-157907793	33588637	5.69	Channel catfish	T/C	T
AX-157907795	33590255	5.69	Channel catfish	T/C	C
AX-157907798	33600624	5.69	Channel catfish	A/C	C
AX-157812014	33601261	5.69	Channel-Blue	T/C	C
AX-157907758	33540480	5.67	Channel catfish	A/G	A
AX-157812012	33566514	5.67	Channel-Blue	T/C	T
AX-157907801	33619552	5.67	Channel catfish	T/C	C
AX-157907780	33571670	5.67	Channel catfish	A/G	A
AX-157907803	33623499	5.60	Channel catfish	A/G	A
AX-157907634	33374908	5.58	Channel catfish	T/C	C
AX-157907551	33254088	5.40	Channel catfish	T/C	T
AX-157893863	8870666	5.34	Channel catfish	T/G	T
AX-157892258	5651238	5.33	Channel catfish	A/C	A
AX-157812005	33307993	5.32	Channel-Blue	A/G	A
AX-157892819	6628233	5.30	Channel catfish	A/G	A

AX-157892501	5068593	5.28	Channel catfish	A/G	G
AX-157777441	33265128	5.27	Blue catfish	A/G	A
AX-157776277	7919526	5.23	Blue catfish	A/G	G
AX-157907753	33536979	5.22	Channel catfish	T/C	C
AX-157907549	33250880	5.20	Channel catfish	A/G	A
AX-157731312	33606466	5.20	Interspecies	T/C	T
AX-157892493	5052936	5.18	Channel catfish	T/G	T
AX-157765090	9690072	5.17	Interspecies	T/G	T
AX-157776297	8747394	5.15	Blue catfish	A/G	A
AX-157776352	9223134	5.13	Blue catfish	T/C	C
AX-158404242	6356950	5.13	Channel catfish	T/C	C
AX-157893457	8077965	5.11	Channel catfish	A/G	G
AX-157892640	6225087	5.10	Channel catfish	A/G	G
AX-157892526	6030173	5.10	Channel catfish	A/G	A
AX-157907816	33642422	5.08	Channel catfish	T/C	C
AX-157776382	9817492	5.07	Blue catfish	T/C	C
AX-157766971	8915931	5.07	Interspecies	T/C	T
AX-157810508	7800853	5.06	Channel-Blue	T/C	T
AX-157892504	5076713	5.03	Channel catfish	T/C	C

	AX-157892154	5378742	5.03	Channel catfish	A/C	A
	AX-157737169	7785264	5.02	Interspecies	T/C	C
	AX-157776258	7620722	5.01	Blue catfish	A/G	G
23	**AX-157695826**	8083472	6.25	Channel catfish	A/T	A
	AX-158291343	7793018	5.76	Channel catfish	A/C	C
	AX-158291811	8136119	5.51	Channel catfish	T/C	C
	AX-158291798	8121263	5.35	Channel catfish	T/C	T
	AX-157717445	8157082	5.35	Channel catfish	T/C	T
	AX-158291671	7968747	5.25	Channel catfish	A/G	G
	AX-158291673	7971570	5.08	Channel catfish	A/G	G
	AX-158291325	7771356	5.04	Channel catfish	A/G	A
	AX-158291644	7937157	5.04	Channel catfish	T/C	C

2.4.3 Genes within the associated QTL for ESC resistance

To explore the potential genes involved in ESC resistance, the ±1 Mb regions around associated SNPs were examined. Synteny analyses were conducted to compare the extend regions around the significant SNPs. The conserved synteny was observed between catfish and zebrafish as shown in Figure 5 and Figure 6. A set of 43 genes were identified within the significantly associated region on LG1. Of these genes, seven genes were found to have known functions in

immunity (Table 3), including NCK adaptor protein 1 (*nck1*), angiotensin II receptor type 1 (*agtr1*), transient receptor potential cation channel subfamily C member 1 (*trpc1*), abl interactor 1 (*abi1*), Rap1-GTP-interacting adaptor molecule (*apbb1ip*), ARP3 actin related protein 3 homolog B (*actr3b*), and vav guanine nucleotide exchange factor 3 (*vav3*). Within the significant QTL on LG23, 45 genes were detected and three genes had immune-related functions (Table 3), including mannose receptor C-type 1 like (*mrc1l*), protein kinase C theta (*prkcq*), and GATA binding protein 3 (*gata3*). Moreover, 49 genes were immune-related, of a total of 271 genes identified in the suggestive region on LG1.

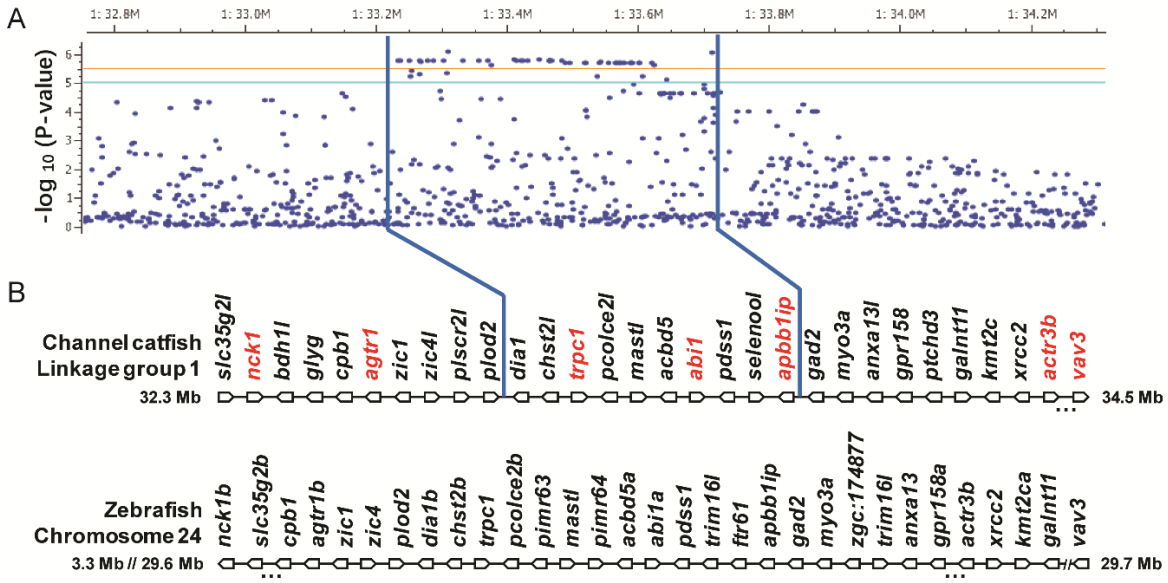

Figure 5. Genes within the significant region associated with ESC resistance on LG1. A) Regional Manhattan plot for the QTL on LG1. The yellow horizontal line indicates the threshold *P* value for genome-wide significance. The blue horizontal line indicates the threshold *P* value for

significance of "suggestive association". B) Synteny analysis between catfish and zebrafish. Immune-related genes in channel catfish are highlighted in red.

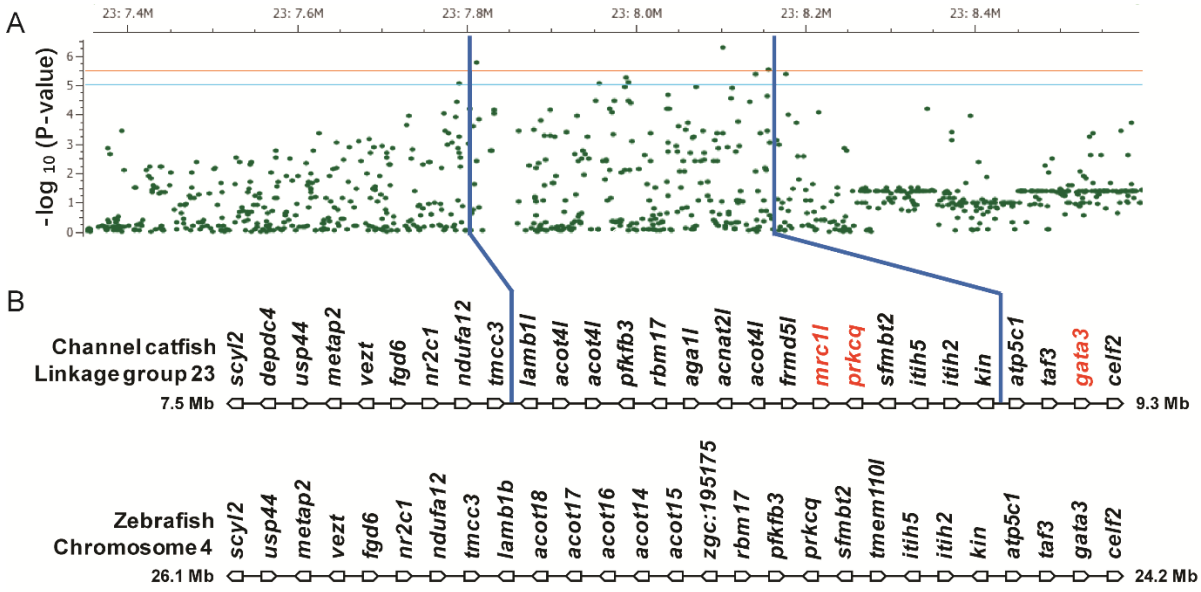

Figure 6. Genes within the significant region associated with ESC resistance on LG23. A) Regional Manhattan plot for the QTL on LG23. The yellow horizontal line indicates the threshold *P* value for genome-wide significance. The blue horizontal line indicates the threshold *P* value for significance of "suggestive association". B) Synteny analysis between catfish and zebrafish. Immune-related genes in channel catfish are highlighted in red.

Table 3. Immune-related genes within the genome-wide significant QTL on linkage group 1 and 23

Linkage group	Gene	Location (bp)	Function
1	*nck1*	32,352,283-32,413,183	actin filament organization
			Phagocytosis
			T cell activation
			B cell receptor signaling
	agtr1	32,480,471-32,494,698	regulation of inflammatory response
	trpc1	33,472,942-33,504,891	calcium ion transport
			B cell receptor signaling
	abi1	33,560,300-33,609,297	actin polymerization or depolymerization
			Phagocytosis
	apbb1ip	33,632,904-33,677,852	T cell activation
	actr3b	34,277,661-34,297,129	actin nucleation
			Phagocytosis
	vav3	34,542,282-34,629,025	Phagocytosis
			B cell receptor signaling
23	*mrc1l*	7,943,812-7,946,175	cellular response to lipopolysaccharide
			endocytosis
			T cell activation
	prkcq	7,954,997-7,964,953	inflammatory response
			T cell activation
	gata3	8,239,408-8,259,340	inflammatory response
			T cell differentiation
			humoral immune response

Gene function was collected form Uniprot (http://www.uniprot.org) and REACTOME (http://www.reactome.org).

2.4.4 Correlation of the SNPs with ESC resistance

Conditioned analyses were performed to determine whether the associations detected on a linkage group were independent (Nishimura et al. 2012). The lead SNP on each associated QTL was included as a covariate in the mix linear model. After conditioning, associations of the surrounding SNP on the same linkage group disappeared, while SNPs on other linkage groups

generally remained the same, indicating there was no independently associated SNP marker within the same linkage group.

2.5 Discussion

Identification of polymorphisms affecting ESC disease resistance could facilitate effective breeding through marker-assisted selection and introgression, and thereby reduce losses for the catfish aquaculture industry. Here, GWAS was conducted using the channel catfish × blue catfish backcross progenies genotyped with the newly designed catfish 690K SNP arrays. We successfully identified QTL associated with ESC resistance on LG1 and LG23. Channel catfish and blue catfish exhibit great contrast in the resistance to ESC, with blue catfish being highly resistant, almost never becoming infected naturally and having low mortality when challenged artificially (Bilodeau-Bourgeois et al. 2008; Hanson 2006; Wolters and Johnson 1994; Wolters et al. 1996). In the backcross progenies, both genes and the disease resistance traits are segregating, allowing mapping of disease resistance to chromosomal locations. Although the F2 intercrosses of hybrids between the two species with phenotypic divergence are appropriate populations for mapping QTL (Cnaani et al. 2003; Poompuang and Hallerman 1997), it is extremely difficult to produce the interspecific F2 hybrids due to the very low hatching rate (Dunham and Argue 2000). Therefore, the backcross progenies provide a best available alternative for the study of interspecific ESC resistance/susceptibility-associated QTL.

Two methods, EMMAX and QFAM, were utilized for family-based samples in the present study. Both are effective for correcting population stratification which can lead to biased or spurious results. The statistical approaches of EMMAX and QFAM to control the population stratification are different. EMMAX uses a pairwise relatedness matrix as random effect to correct for sample structures including hidden relatedness and population stratification (Kang et al. 2010). QFAM partitions the genotypes into between- and within- family components (Fulker et al. 1999; Abecasis et al. 2000). The within-family analysis is robust to population stratification, which assesses transmission of alleles within a family, but without making use of allelic association observed across families. When using family-based samples, QFAM was found to be more robust in correcting family structure than EMMAX (Geng et al. 2017b). Family-based samples were utilized in the present study and the suggestive QTL on LG14 found by EMMAX but not by QFAM was suggested to be false positive due to family structure. Furthermore, the suggestive region on LG1 was not found by EMMAX but identified by QFAM in this study. Geng et al. (2016) reported that QFAM possessed more power compared with EMMAX in a family-based study. Recent work has shown that inclusion of candidate markers in the relatedness matrix could lead to decreased power due to double-fitting of candidate markers in the mix linear model (Yang et al. 2014). Therefore, QFAM results were chosen for further analysis.

Examination of the associated SNPs revealed that the superior disease resistance alleles were from blue catfish. Four suggestively associated SNPs and four significant SNPs including the most significant SNP (AX-157732970) on LG1 were all interspecific. All of the resistance

alleles were found to be derived from blue catfish, explaining the very strong resistance of blue catfish against the ESC disease. For LG23, all associated SNPs were channel catfish-specific, which means the single nucleotide variation on these loci occurred only in channel catfish. Further investigation showed that all the SNPs that could be mapped to blue catfish genome sequence are homologous with the two resistance alleles. The results implied that certain blue catfish alleles were associated with increased ESC resistance.

Within the two significantly associated regions on LG1 and LG23 (indicated by red boxes in Figure 4), an examination of genes and related pathways provided insights into the molecular mechanisms underlying ESC resistance. The most studied, closest relative of catfish, zebrafish, has 696 genes involved in the immune system process in PANTHER database (http://www.pantherdb.org). The channel catfish has 27,143 genes based on NCBI annotation. The approximate proportion of immune genes in the genome is 0.026. The proportion of immune genes within our QTL region for LG1 and LG23 is 0.163 (7/43) and 0.067 (3/45), which in the case of LG1 is more than expected (Chi-square test, $P < 0.001$) compared to the rest of the genome. The immune-related genes observed in the significant QTL were found to be mainly involved in phagocytosis (Figure 7) and T-cell activation (Figure 8). Phagocytosis is a principal component of the innate immunity in which phagocytes, including macrophages, neutrophils and dendritic cells, internalize targets in an actin-dependent manner (Botelho and Grinstein 2011). *E. ictaluri* is internalized in catfish phagocytes after infection, especially in macrophages (Miyazaki and Plumb 1985; Shotts et al. 1986; Baldwin and Newton 1993), suggesting that phagocytic activities play

crucial roles in immunity for combating ESC. T cells have a central role in adaptive immunity by regulating immune responses of various phagocytes and B cells through cytokine secretion or by directly destroying antigen-bearing cells (Medzhitov 2007). Russo et al. (2009) demonstrated that macrophages from vaccinated fish were more efficient in rapid clearance of infection upon re-exposure to virulent *E. ictaluri*, reflecting that memory lymphocytes were involved and lymphocytes in adaptive immunity were important in enhancing macrophage activities during immune responses.

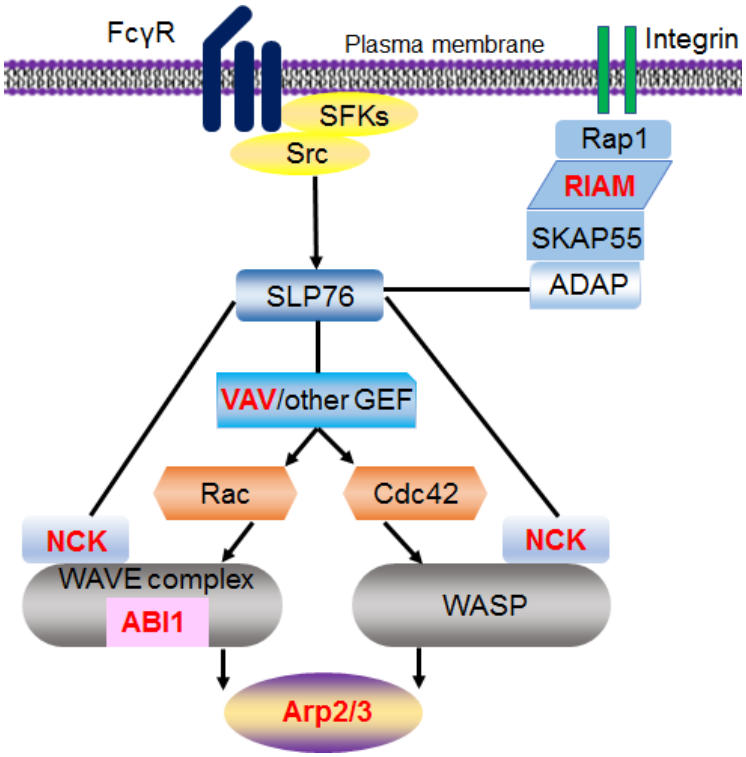

Figure 7. Signal transduction pathway involved in Fcγ receptor-mediated phagocytosis. Corresponding proteins of immune-related genes identified in this study are highlighted in red.

Abbreviations: SFKs, Src family tyrosine kinase; Src, proto-oncogene tyrosine-protein kinase Src; SLP76, Src homology 2 domain-containing leukocyte protein of 76 kilodaltons; VAV, VAV guanine nucleotide exchange factor; Rac, small GTPase Rac; Cdc42, small GTPase Cdc42; NCK, NCK adaptor protein; WASP, Wiskott-Aldrich syndrome protein; WAVE, WASP family Verprolin-homologous protein; ABI1, abl interactor 1; Arp2/3, the actin nucleation complex Arp2/3, ADAP, Adhesion and degranulation-promoting adapter protein; SKAP55, Src kinase-associated phosphoprotein of 55 kDa; RIAM, Rap1-GTP-interacting adaptor molecule; Rap1, small GTPase Rap1.

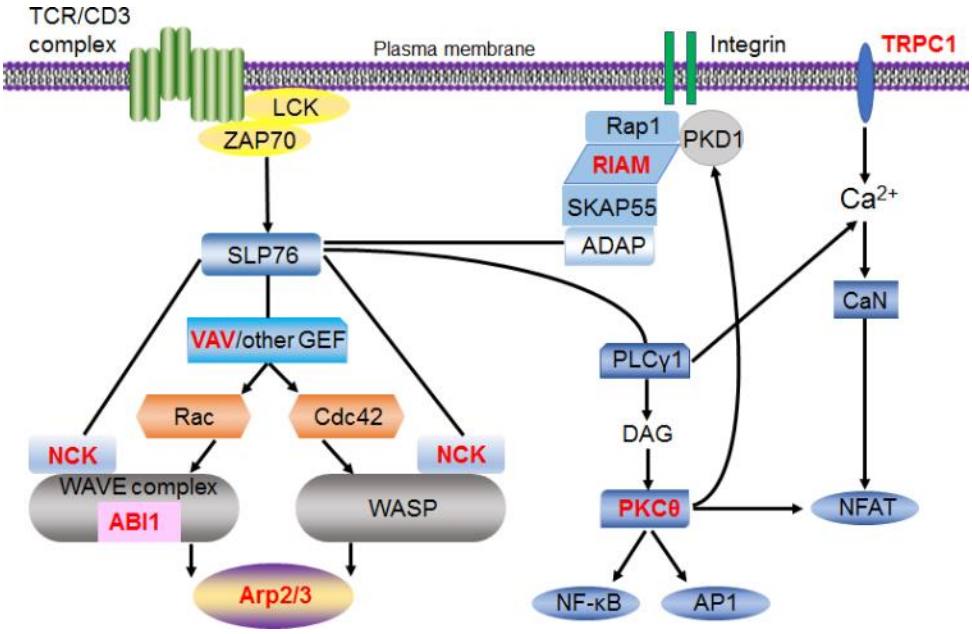

Figure 8. Signal transduction pathway leading to T-cell activation. Corresponding proteins of immune-related genes identified in this study are highlighted in red. Abbreviations: LCK, LCK

proto-oncogene, Src family tyrosine kinase; ZAP70, zeta chain of T-cell receptor associated protein kinase 70; SLP76, Src homology 2 domain-containing leukocyte protein of 76 kilodaltons; VAV, VAV guanine nucleotide exchange factor; Rac, small GTPase Rac; Cdc42, small GTPase Cdc42; NCK, NCK adaptor protein; WASP, Wiskott-Aldrich syndrome protein; WAVE, WASP family Verprolin-homologous protein; ABI1, abl interactor 1; Arp2/3, the actin nucleation complex Arp2/3, ADAP, Adhesion and degranulation-promoting adapter protein; SKAP55, Src kinase-associated phosphoprotein of 55 kDa; RIAM, Rap1-GTP-interacting adaptor molecule; Rap1, small GTPase Rap1; PLCγ1, phospholipase C-γ1; DAG, diacylglycerol; PKCθ, protein kinase C theta; PKD1, polycystin 1, transient receptor potential channel interacting; TRPC1, transient receptor potential cation channel subfamily C member 1; CaN, calcineurin; NF-κB, nuclear factor-kappa B; AP-1, activator protein-1; NFAT, nuclear factor of activated T cells.

One of the similar process of both phagocytosis and T-cell activation is cytoskeleton rearrangement. In phagocytes, local polymerization of actin filaments supports the protrusion of pseudopodia that facilitates the engulfment of pathogens or particles. In T cells, cytoskeleton reorganization is crucial for the formation of immunological synapse, which is crucial for cell adhesion and T-cell activation. Four genes within the significantly associated region on LG1, including *nck1, vav3, abi1,* and *actr3b,* were involved in actin cytoskeleton reorganization. Nck proteins are adapter proteins which play pivotal role in the phagocytic process (Coppolino et al.

2001) and T-cell receptor (TCR) induced actin cytoskeleton reorganization during T-cell activation (Lettau et al. 2009). Upon the phosphorylation of tyrosine kinases, SH2 domain-containing leukocyte phosphoprotein of 76kDa (SLP-76) recruits Nck. Nck then promotes the recruitment of the multidomain adapter protein Wiskott-Aldrich syndrome protein (WASP) and WASP family Verprolin-homologous protein (WAVE), whose activations facilitate the actin-related protein 2/3 (Arp2/3) complex to drive actin filament formation. SLP-76 also recruit the Vav proteins, promoting the GTP-loading of Rho GTPases to facilitate the activation of WASP family proteins. The Vav family proteins (Vav1, Vav2, and Vav3) are guanine nucleotide exchange factors (GEFs) for Rho family GTPases (Hornstein et al. 2004; Tybulewicz 2005), playing crucial roles in the regulation of actin dynamics during phagocytosis (Patel et al. 2002; Hall et al. 2006) and T-cell activation (Villalba et al. 2000, 2002). Abl interactor 1 (Abi1) is crucial in actin cytoskeleton dynamics due to its participation in the WAVE complex, which stimulates Arp2/3 complex-dependent actin assembly (Innocenti et al. 2004; Kheir et al. 2005). Arp2/3 complex plays a major role in the regulation of the actin cytoskeleton (Machesky and Gould 1999), essential for actin polymerization in the phagocytic cup and subsequent particle engulfment in both FcγR- and CR3-mediated phagocytosis (May et al. 2000). Moreover, the Arp2/3 complex was involved in actin cytoskeleton remodeling during T-cell activation (Krause et al. 2000). Actin-Related Proteins ARP3 (*actr3*) is one subunit of the Arp2/3 complex. In the present study, the homology of *actr3*, *actr3b*, was in the significantly associated region on LG1 that associated with ESC disease resistance.

Another shared process during T-cell activation and phagocytosis is the integrin activation, where Rap1-GTP-interacting adaptor molecule (RIAM, *apbb1ip*) is implicated. Integrins are transmembrane heterodimers responsible for mediating cell/cell or cell/matrix adhesions (Hynes 2002), which are crucial for development, tissue maintenance and repair, immune response and hemostasis (Hynes 2002; Harburger and Calderwood 2009). Binding affinity and avidity of integrins for their ligands facilitate T-cell activation by providing stable contact with APCs and extracellular proteins (Burbach et al. 2007). A key integrin regulator is the small GTPase Ras-proximity-1 (Rap1), which enhances T-cell activation by mediating TCR-induced adhesion to intercellular adhesion molecule (ICAM) (Katagiri et al. 2000; Sebzda et al. 2002). In addition, FcγR-mediated responses have been shown to require the presence of β2 integrins, include adhesion to immune complexes and enhanced Fc-dependent phagocytosis (Ortiz-Stern and Rosales 2003).

The calcium ion (Ca^{2+}) is the main common second messenger involved in signaling transduction following immunoreceptor activation. For the T-cell activation, TCR-induced increase in intracellular Ca^{2+} level is crucial for many cellular functions, including the translocation of NFAT (nuclear factor of activated T cells) to the nucleus through calcineurin (CaN) activation (Macian 2005). TRPC1, a membrane-spanning subunit of cation channel, was involved in store-operated Ca^{2+} entry (SOCE) by interacting with calcium release-activated calcium channel protein 1 (Orai1) and stromal interaction molecule 1 (STIM1) (Ong et al. 2007), both of which were up-regulated during T-cell activation (Lioudyno et al. 2008). Furthermore, TRPC1 was shown to be

expressed in a human T cell line (HPB-ALL) and contributed to cannabinoid-induced Ca^{2+} influx in these cells independently of intracellular Ca^{2+} store depletion (Rao and Kaminski 2006).

Pattern recognition receptors (PRRs) are well known for their significant roles in immunity. Mannose receptor C type 1 (MRC1) is a C-type lectin primarily present on macrophages and dendritic cells. MRC1 played significant roles in both innate and adaptive immune responses against various microorganisms (Taylor et al. 2005), and it was implicated in pathogen recognition, phagocytosis, cytokine production, antigen processing and presentation, cell adhesion and migration (Apostolopoulos and McKenzie 2001; East and Isacke 2002; Taylor et al. 2005; Gazi and Martinez-Pomares 2009). Due to the crucial roles in immune responses, much work on MRC1 has been done in human and mice, with a few studies in fish as well. In grass carp (*Ctenopharyngodon idella*), MRC1 was significantly up-regulated after *Aeromonas hydrophila* infection in liver, spleen, head kidney and intestine (Wang et al. 2014). Similarly, the expression of MRC1 was dramatically induced in the spleen of tilapia (*Oreochromis niloticus*) at all three studied timepoints (5 h, 50 h, and 7 d) following *Streptococcus iniae* infection (Zhu et al. 2015). In this study, the gene *mrc1 like* was within the QTL on LG23, and its expression in catfish was significantly up-regulated at 3h and 72 h after *E. ictaluri* infection in a previous study (Li et al. 2012).

The gene *prkcq* was located within the associated region on LG23, coding for protein kinase C theta (PKCθ). Predominantly expressed in T lymphocytes, PKCθ plays essential roles for T-cell activation, proliferation, differentiation, survival, and cytokine production (Anderson et al.

2006; Isakov and Altman 2002; Barouch-Bentov et al. 2005; Hayashi and Altman 2007). PKCθ is the only serine/threonine protein kinase C (PKC) recruited selectively to the center of the immunological synapse following stimulation with APCs, required for T-cell activation and downstream signaling (Monks et al. 1997; Bi et al. 2001). T lymphocytes in PKCθ-deficient mice displayed impaired activation of transcription factors including NF-κB (nuclear factor-kappa B), AP-1 (activator protein-1) and NFAT (Pfeifhofer et al. 2003; Sun et al. 2000). PKCθ was found to be significantly up-regulated in intestine of *Pelodiscus sinensis* after intragastric challenge with lipopolysaccharide (LPS) (Xu et al. 2016).

One interesting finding of our study was that the functionally related genes in immunity were in relatively close proximity on the chromosome. Of the seven immune-related genes within the significantly associated region on LG1, six genes were implicated in phagocytosis (*nck1, vav3, abi1, actr3b, apbb1ip,* and *agtr1*), and six genes were involved in T-cell activation (*nck1, vav3, abi1, actr3b, apbb1ip,* and *trpc1*). In plants, some resistance genes often cluster together in certain chromosome, sometimes so tightly that they can be considered as complex loci (Vale et al. 2001). The structural organization may facilitate coordinated expression and function (Michalak 2008). Further studies need to be conducted to confirm the involvement of these genes and their specific function in ESC disease resistance.

Our long-term goal is to enhance catfish stocks with a broad spectrum of disease resistance as well as other favorable traits, and support a sustainable and profitable aquaculture industry. To achieve this goal, the genetic basis underlying desirable traits should be understood, especially the

accurate location of QTL affecting the traits. In this study, we identified two genomic regions associated with ESC resistance on LG1 and LG23. We previously reported one QTL in LG1 using the third-generation backcrosses (Zhou et al. 2017), which was produced by mating with channel catfish female in each generation. The QTL was validated here in the first-generation backcrosses, suggesting that this QTL was operating in various populations of a broad genetic background. SNPs which were associated with ESC resistance were identified in the present study. Individuals must be generated that are homozygous for these key SNPs, identified and selected as these SNPs are likely associated with the blue catfish resistance alleles that are incompletely dominant since blue catfish are almost totally resistant to ESC, but the F1 channel catfish female × blue catfish male (heterozygous) is not totally resistant though improved. Research should continue to identify the exact location and identity of the ESC resistance alleles using more families or populations. This would allow additional options to interspecific introgression to establish the ESC resistance alleles in a homozygous state in cultured catfish. These include gene transfer or even CRISPR/Cas9 technology to mutate the channel catfish alleles into their blue catfish ESC resistance counterparts (Dunham et al. 2014).

References

Abdelrahman H, ElHady M, Alcivar-Warren A, Allen S, Al-Tobasei R, Bao L, Beck B, Blackburn H, Bosworth B, Buchanan J, Chappell J, Daniels W, Dong S, Dunham RA, Durland E, Elaswad A, Gomez-Chiarri M, Gosh K, Guo X, Hackett P, Hanson T, Hedgecock D, Howard T, Holland L, Jackson M, Jin Y, Khalil K, Kocher T, Leeds T, Li N, Lindsey L, Liu S, Liu Z, Martin K, Novriadi R, Odin R, Palti Y, Peatman E, Proestou D, Qin G, Reading B, Rexroad C, Roberts S, Salem M, Severin A, Shi H, Shoemaker C, Stiles S, Tan S, Tang KFJ, Thongda W, Tiersch T, Tomasso J, Prabowo WT, Vallejo R, Steen H, Vo K, Waldbieser G, Wang H, Wang X, Xiang J, Yang Y, Yant R, Yuan Z, Zeng Q, Zhou T (2017) Aquaculture genomics, genetics and breeding in the United States: current status, challenges, and priorities for future research. BMC Genomics 18:191

Abecasis GR, Cardon LR, Cookson WOC (2000) A general test of association for quantitative traits in nuclear families. Am J Hum Genet 66:279–292

Andersen MH, Schrama D, thor Straten P, Becker JC (2006) Cytotoxic T cells. J Invest Dermatol 126:32–41

Anderson K, Fitzgerald M, DuPont M, Wang T, Paz N, Dorsch M, Healy A, Xu Y, Ocain T, Schopf L, Jaffee B, Picarella D (2006) Mice deficient in PKC theta demonstrate impaired *in vivo* T cell activation and protection from T cell-mediated inflammatory diseases. Autoimmunity 39:469–478

Annunziato F, Cosmi L, Liotta F, Maggi E, Romagnani S (2009) Type 17 T helper cells—origins, features and possible roles in rheumatic disease. Nat Rev Rheumatol 5:325–331

Apostolopoulos V, McKenzie IFC (2001) Role of the mannose receptor in the immune response. Curr Mol Med 1:469–474

Baerwald MR, Petersen JL, Hedrick RP, Schisler GJ, May B (2011) A major effect quantitative trait locus for whirling disease resistance identified in rainbow trout (*Oncorhynchus mykiss*). Heredity (Edinb) 106:920

Baldwin TJ, Newton JC (1993) Pathogenesis of enteric septicemia of channel catfish, caused by *Edwardsiella ictaluri*: bacteriologic and light and electron microscopic findings. J Aquat Anim Health 5:189–198

Barouch-Bentov R, Lemmens EE, Hu J, Janssen EM, Droin NM, Song J, Schoenberger SP, Altman A (2005) Protein kinase C-θ is an early survival factor required for differentiation of effector $CD8^+$ T cells. J Immunol 175:5126–5134

Bi K, Tanaka Y, Coudronniere N, Sugie K, Hong S, van Stipdonk MJB, Altman A (2001) Antigen-induced translocation of PKC-θ to membrane rafts is required for T cell activation. Nat Immunol 2:556

Bilodeau-Bourgeois L, Bosworth BG, Peterson BC (2008) Differences in mortality, growth, lysozyme, and toll-like receptor gene expression among genetic groups of catfish exposed to virulent *Edwardsiella ictaluri*. Fish Shellfish Immunol 24:82–89

Botelho RJ, Grinstein S (2011) Phagocytosis. Curr Biol 21:R533–R538

Burbach BJ, Medeiros RB, Mueller KL, Shimizu Y (2007) T-cell receptor signaling to integrins. Immunol Rev 218:65–81

Campbell NR, LaPatra SE, Overturf K, Towner R, Narum SR (2014) Association mapping of disease resistance traits in rainbow trout using restriction site associated DNA sequencing. G3 (Bethesda) 4:2473–2481

Cnaani A, Hallerman EM, Ron M, Weller JI, Indelman M, Kashi Y, Gall GAE, Hulata G (2003) Detection of a chromosomal region with two quantitative trait loci, affecting cold tolerance and fish size, in an F2 tilapia hybrid. Aquaculture 223:117–128

Coppolino MG, Krause M, Hagendorff P, Monner DA, Trimble W, Grinstein S, Wehland J, Sechi AS (2001) Evidence for a molecular complex consisting of Fyb/SLAP, SLP-76, Nck, VASP and WASP that links the actin cytoskeleton to Fcγ receptor signalling during phagocytosis. J Cell Sci 114:4307–4318

Correa K, Lhorente JP, López ME, Bassini L, Naswa S, Deeb N, Di Genova A, Maass A, Davidson WS, Yáñez JM (2015) Genome-wide association analysis reveals loci associated with resistance against *Piscirickettsia salmonis* in two Atlantic salmon (*Salmo salar* L.) chromosomes. BMC Genomics 16:854

Darvasi A, Soller M (1992) Selective genotyping for determination of linkage between a marker locus and a quantitative trait locus. Theor Appl Genet 85:353–359

Dunham RA, Argue BJ (2000) Reproduction among channel catfish, blue catfish, and their F1 and F2 hybrids. Trans Am Fish Soc 129:222–231

Dunham RA, Smitherman RO (1984) Ancestry and breeding of catfish in the United States. Circular 273, Ala Agric Exp Stat, Auburn University, Alabama

Dunham RA, Taylor JF, Rise ML, Liu Z (2014) Development of strategies for integrated breeding, genetics and applied genomics for genetic improvement of aquatic organisms. Aquaculture 420-421:S121–123

Dunham RA, Umali GM, Beam R, Kristanto AH, Trask M (2008) Comparison of production traits of NWAC103 channel catfish, NWAC103 channel catfish× blue catfish hybrids, Kansas Select 21 channel catfish, and blue catfish grown at commercial densities and exposed to natural bacterial epizootics. N Am J Aquac 70:98–106

East L, Isacke CM (2002) The mannose receptor family. Biochim Biophys Acta (BBA)-General Subj 1572:364–386

Fuji K, Hasegawa O, Honda K, Kumasaka K, Sakamoto T, Okamoto N (2007) Marker-assisted breeding of a lymphocystis disease-resistant Japanese flounder (*Paralichthys olivaceus*). Aquaculture 272:291–295

Fuji K, Kobayashi K, Hasegawa O, Coimbra MRM, 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

Fulker DW, Cherny SS, Sham PC, Hewitt JK (1999) Combined linkage and association sib-pair analysis for quantitative traits. Am J Hum Genet 64:259–267

Gazi U, Martinez-Pomares L (2009) Influence of the mannose receptor in host immune responses. Immunobiology 214:554–561

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

Geng X, Liu S, Yuan Z, Jiang Y, Zhi D, Liu Z (2017a) A genome-wide association study reveals that genes with functions for bone development are associated with body conformation in catfish. Mar Biotechnol 19:570–578

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

Geng X, Zhi D, Liu Z (2017b) Genome-wide association studies of performance traits. In: Liu Z (ed) Bioinformatics in aquaculture: principles and methods, 1st edn. John Wiley & Sons, Hoboken, pp 415–433

Hall AB, Gakidis MAM, Glogauer M, Wilsbacher JL, Gao S, Swat W, Brugge JS (2006) Requirements for Vav guanine nucleotide exchange factors and Rho GTPases in FcγR-and complement-mediated phagocytosis. Immunity 24:305–316

Hanson LA (2006) Enteric septicaemia of catfish. In: Manual of diagnostic tests for aquatic animals, 5th edn. OIE, Paris (France), pp 228–235

Harburger DS, Calderwood DA (2009) Integrin signalling at a glance. J Cell Sci 122:159–163

Hayashi K, Altman A (2007) Protein kinase C theta (PKCθ): a key player in T cell life and death. Pharmacol Res 55:537–544

Hornstein I, Alcover A, Katzav S (2004) Vav proteins, masters of the world of cytoskeleton organization. Cell Signal 16:1–11

Houston RD, Haley CS, Hamilton A, Guy DR, Tinch AE, Taggart JB, McAndrew BJ, Bishop SC (2008) Major quantitative trait loci affect resistance to infectious pancreatic necrosis in Atlantic salmon (*Salmo salar*). Genetics 178:1109–1115

Hynes RO (2002) Integrins: bidirectional, allosteric signaling machines. Cell 110:673–687

Innocenti M, Zucconi A, Disanza A, Frittoli E, Areces LB, Steffen A, Stradal TEB, Di Fiore PP, Carlier M-F, Scita G (2004) Abi1 is essential for the formation and activation of a WAVE2 signalling complex. Nat Cell Biol 6:319

Isakov N, Altman A (2002) Protein kinase Cθ in T cell activation. Annu Rev Immunol 20:761–794

Jin Y, Zhou T, Geng X, Liu S, Chen A, Yao J, Jiang C, Tan S, Su B, Liu Z (2017) A genome-wide association study of heat stress-associated SNPs in catfish. Anim Genet 48:233–236

Kang HM, Sul JH, Zaitlen NA, Kong S, Freimer NB, Sabatti C, Eskin E (2010) Variance component model to account for sample structure in genome-wide association studies. Nat Genet 42:348–354

Katagiri K, Hattori M, Minato N, Irie S, Takatsu K, Kinashi T (2000) Rap1 is a potent activation signal for leukocyte function-associated antigen 1 distinct from protein kinase C and phosphatidylinositol-3-OH kinase. Mol Cell Biol 20:1956–1969

Kheir WA, Gevrey J-C, Yamaguchi H, Isaac B, Cox D (2005) A WAVE2-Abi1 complex mediates CSF-1-induced F-actin-rich membrane protrusions and migration in macrophages. J Cell Sci 118:5369–5379

Krause M, Sechi AS, Konradt M, Monner D, Gertler FB, Wehland J (2000) Fyn-binding protein (Fyb)/SLP-76–associated protein (SLAP), Ena/vasodilator-stimulated phosphoprotein (VASP) proteins and the Arp2/3 complex link T cell receptor (TCR) signaling to the actin cytoskeleton. J Cell Biol 149:181–194

Lettau M, Pieper J, Janssen O (2009) Nck adapter proteins: functional versatility in T cells. Cell Commun Signal 7:1

Li C, Zhang Y, Wang R, Lu J, Nandi S, Mohanty S, Terhune J, Liu Z, 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 Shellfish Immunol 32:816–827

Li N, Zhou T, Geng X, Jin Y, Wang X, Liu S, Xu X, Gao D, Li Q, Liu Z (2017a) Identification of novel genes significantly affecting growth in catfish through GWAS analysis. Mol Genet Genomics https://doi.org/10.1007/s00438-017-1406-1

Li Y, Geng X, Bao L, Elaswad A, Huggins KW, Dunham RA, Liu Z (2017b) A deletion in the Hermansky–Pudlak syndrome 4 (*Hps4*) gene appears to be responsible for albinism in channel catfish. Mol Genet Genomics 292:663–670

Lioudyno MI, Kozak JA, Penna A, Safrina O, Zhang SL, Sen D, Roos J, Stauderman KA, Cahalan MD (2008) Orai1 and STIM1 move to the immunological synapse and are up-regulated during T cell activation. Proc Natl Acad Sci 105:2011–2016

Liu P, Wang L, Wan ZY, Ye BQ, Huang S, Wong S-M, Yue GH (2016a) Mapping QTL for resistance against viral nervous necrosis disease in Asian seabass. Mar Biotechnol 18:107–116

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

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

Liu S, Vallejo RL, Palti Y, Gao G, Marancik DP, Hernandez AG, Wiens GD (2015) Identification of single nucleotide polymorphism markers associated with bacterial cold water disease resistance and spleen size in rainbow trout. Front Genet 6:298

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

Lucas JS, Southgate PC (2012) Aquaculture: Farming aquatic animals and plants. John Wiley & Sons

Machesky LM, Gould KL (1999) The Arp2/3 complex: a multifunctional actin organizer. Curr Opin Cell Biol 11:117–121

Macian F (2005) NFAT proteins: key regulators of T-cell development and function. Nat Rev Immunol 5:472

May RC, Caron E, Hall A, Machesky LM (2000) Involvement of the Arp2/3 complex in phagocytosis mediated by FcγR or CR3. Nat Cell Biol 2:246

Medzhitov R (2007) Recognition of microorganisms and activation of the immune response. Nature 449:819

Michalak P (2008) Coexpression, coregulation, and cofunctionality of neighboring genes in eukaryotic genomes. Genomics 91:243–248

Miyazaki T, Plumb JA (1985) Histopathology of *Edwardsiella ictaluri* in channel catfish, *Ictalurus punctatus* (Rafinesque). J Fish Dis 8:389–392

Moen T, Baranski M, Sonesson AK, 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

Monks CRF, Kupfer H, Tamir I, Barlow A, Kupfer A (1997) Selective modulation of protein kinase C-omega during T-cell activation. Nature 385:83

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 Genet 13:40

Ong HL, Cheng KT, Liu X, Bandyopadhyay BC, Paria BC, Soboloff J, Pani B, Gwack Y, Srikanth S, Singh BB, Gill D, Ambudkar IS (2007) Dynamic assembly of TRPC1-STIM1-Orai1 ternary complex is involved in store-operated calcium influx. Evidence for similarities in store-operated and calcium release-activated calcium channel components. J Biol Chem 282:9105–9116

Ortiz-Stern A, Rosales C (2003) Cross-talk between Fc receptors and integrins. Immunol Lett 90:137–143

Ozaki A, Khoo S-K, Yoshiura Y, Ototake M, Sakamoto T, Dijkstra JM, Okamoto N (2007) Identification of additional quantitative trait loci (QTL) responsible for susceptibility to infectious pancreatic necrosis virus in rainbow trout. Fish Pathol 42:131–140

Palti Y, Vallejo RL, Gao G, Liu S, Hernandez AG, Rexroad III CE, Wiens GD (2015) Detection and validation of QTL affecting bacterial cold water disease resistance in rainbow trout using restriction-site associated DNA sequencing. PLoS One 10:e0138435

Patel JC, Hall A, Caron E (2002) Vav regulates activation of Rac but not Cdc42 during FcγR-mediated phagocytosis. Mol Biol Cell 13:1215–1226

Pfeifhofer C, Kofler K, Gruber T, Tabrizi NG, Lutz C, Maly K, Leitges M, Baier G (2003) Protein kinase C θ affects Ca^{2+} mobilization and NFAT activation in primary mouse T cells. J Exp Med 197:1525–1535

Poompuang S, Hallerman EM (1997) Toward detection of quantitative trait loci and marker-assisted selection in fish. Rev Fish Sci 5:253–277

Purcell S, Neale B, Todd-Brown K, Thomas L, Ferreira MAR, Bender D, Maller J, Sklar P, De Bakker PIW, Daly MJ (2007) PLINK: a tool set for whole-genome association and population-based linkage analyses. Am J Hum Genet 81:559–575

Rao GK, Kaminski NE (2006) Induction of intracellular calcium elevation by Δ9-tetrahydrocannabinol in T cells involves TRPC1 channels. J Leukoc Biol 79:202–213

Russo R, Shoemaker CA, Panangala VS, Klesius PH (2009) *In vitro* and *in vivo* interaction of macrophages from vaccinated and non-vaccinated channel catfish (*Ictalurus punctatus*) to *Edwardsiella ictaluri*. Fish Shellfish Immunol 26:543–552

Sebzda E, Bracke M, Tugal T, Hogg N, Cantrell DA (2002) Rap1A positively regulates T cells via integrin activation rather than inhibiting lymphocyte signaling. Nat Immunol 3:251

Shoemaker CA, Klesius PH, Evans JJ, Arias CR (2009) Use of modified live vaccines in aquaculture. J World Aquacult Soc 40:573–585

Shotts EB, Blazer VS, Waltman WD (1986) Pathogenesis of experimental *Edwardsiella ictaluri* infections in channel catfish (*Ictalurus punctatus*). Can J Fish Aquat Sci 43:36–42

Solovyev V, Kosarev P, Seledsov I, Vorobyev D (2006) Automatic annotation of eukaryotic genes, pseudogenes and promoters. Genome Biol 7:S10

Sun L, Liu S, Wang R, Jiang Y, Zhang Y, Zhang J, Bao L, Kaltenboeck L, Dunham RA, Waldbieser G, Liu Z (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

Sun Z, Arendt CW, Ellmeier W, Schaeffer EM (2000) PKCθ is required for TCR-induced NF-κB activation in mature but not immature T lymphocytes. Nature 404:402

Taylor PR, Gordon S, Martinez-Pomares L (2005) The mannose receptor: linking homeostasis and immunity through sugar recognition. Trends Immunol 26:104–110

Tybulewicz VLJ (2005) Vav-family proteins in T-cell signalling. Curr Opin Immunol 17:267–274

Vale FXR, Parlevliet JE, Zambolim L (2001) Concepts in plant disease resistance. Fitopatol Bras 26:577–589

Villalba M, Bi K, Hu J, Altman Y, Bushway P, Reits E, Neefjes J, Baier G, Abraham RT, Altman A (2002) Translocation of PKCθ in T cells is mediated by a nonconventional, PI3-K–and Vav-dependent pathway, but does not absolutely require phospholipase C. J Cell Biol 157:253–263

Villalba M, Coudronniere N, Deckert M, Teixeiro E, Mas P, Altman A (2000) A novel functional interaction between Vav and PKCθ is required for TCR-induced T cell activation. Immunity 12:151–160

Wang L, Bai B, Huang S, Liu P, Wan ZY, Ye B, Wu J, Yue GH (2017a) QTL mapping for resistance to iridovirus in Asian seabass using genotyping-by-sequencing. Mar Biotechnol 19:517–527

Wang L, Liu L, Zhou Y, Zhao X, Xi M, Wei S, Fang R, Ji W, Chen N, Gu Z, Liu X, Wang W, Asim M, Liu X, Lin L (2014) Molecular cloning and expression analysis of mannose receptor C type 1 in grass carp (*Ctenopharyngodon idella*). Dev Comp Immunol 43:54–58

Wang R, Sun L, Bao L, Zhang J, Jiang Y, Yao J, Song L, Feng J, Liu S, Liu Z (2013) Bulk segregant RNA-seq reveals expression and positional candidate genes and allele-specific expression for disease resistance against enteric septicemia of catfish. BMC Genomics 14:1

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 RA, Liu Z (2017b) Multiple across-strain and within-strain QTLs suggest highly complex genetic architecture for hypoxia tolerance in channel catfish. Mol Genet Genomics 292:63–76

Wolters WR, Johnson MR (1994) Enteric septicemia resistance in blue catfish and three channel catfish strains. J Aquat Anim Health 6:329–334

Wolters WR, Wise DJ, Klesius PH (1996) Survival and antibody response of channel catfish, blue catfish, and channel catfish female × blue catfish male hybrids after exposure to *Edwardsiella ictaluri*. J Aquat Anim Health 8:249–254

Xu J, Zhao J, Li Y, Zou Y, Lu B, Chen Y, Ma Y, Xu H (2016) Evaluation of differentially expressed immune-related genes in intestine of *Pelodiscus sinensis* after intragastric challenge with lipopolysaccharide based on transcriptome analysis. Fish Shellfish Immunol 56:417–426

Yang J, Zaitlen NA, Goddard ME, Visscher PM, Price AL (2014) Advantages and pitfalls in the application of mixed-model association methods. Nat Genet 46:100–106

Zeng Q, Fu Q, Li Y, Waldbieser G, Bosworth B, Liu S, Yang Y, Bao L, Yuan Z, Li N, 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. Sci Rep 7:40347

Zhong X, Wang X, Zhou T, Jin Y, Tan S, Jiang C, Geng X, Li N, Shi H, Zeng Q, Yang Y, Yuan Z, Bao L, Liu S, Tian C, Peatman E, Li Q, Liu Z (2017) Genome-wide association study reveals multiple novel QTL associated with low oxygen tolerance in hybrid catfish. Mar Biotechnol 1–12

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. Mol Genet Genomics 292:231–242

Zhu J, Li C, Ao Q, Tan Y, Luo Y, Guo Y, Lan G, Jiang H, Gan X (2015) Trancriptomic profiling revealed the signatures of acute immune response in tilapia (*Oreochromis niloticus*) following *Streptococcus iniae* challenge. Fish Shellfish Immunol 46:346–353

Chapter 3 Increased alternative splicing as a host response to *Edwardsiella ictaluri* infection in catfish

3.1 Abstract

Alternative splicing (AS) is the process of generating multiple transcripts from a single pre-mRNA used by eukaryotes to regulate gene expression and increase proteomic complexity. Although AS profiles have been well studied in mammalian species, they have not been well studied in aquatic species, especially after biotic stresses. In the present study, genomic information and RNA-Seq datasets were utilized to characterize AS profiles and their induced changes after bacterial infection with *Edwardsiella ictaluri* in channel catfish (*Ictalurus punctatus*). A total of 27,476 AS events, derived from 9,694 genes, were identified in channel catfish. Exon skipping was the most abundant while mutually exclusive exon was the least abundant type of AS. AS was greatly induced by *E. ictaluri* infection with 21.9% increase in AS events. Interestingly, genes involved in RNA binding and RNA splicing themselves were significantly enriched in differentially alternatively spliced (DAS) genes after infection. Sequence analyses of splice variants of a representative alternatively spliced gene, splicing factor *srsf2*, revealed that certain spliced transcripts may undergo nonsense-mediated decay (NMD), suggesting functional significance of induced AS. Although statistical analysis was not possible with such large datasets,

results from quantitative real-time PCR from representative DAS events provided general validation of the bacterial infection-induced AS. This is the first comprehensive study of AS and its changes in response to bacterial infection in fish species, providing insights into the molecular mechanisms of host responses to biotic stresses.

3.2 Background

Alternative splicing (AS) is a crucial mechanism that increases transcriptomic diversity and regulates eukaryotic gene expression. It is prevalent in eukaryotes. For instance, over 90% of genes were shown to be alternatively spliced in humans (Pan et al. 2008; Wang et al. 2008). The presence of alternative splicing transcripts is supported by EST (expressed sequence tag) and cDNA sequence (Modrek et al. 2001; Harrow et al. 2012), microarray data (Pan et al. 2004; Sánchez-Pla et al. 2012), and RNA-seq data (Wang et al. 2008; Ozsolak and Milos 2011).

AS is a tightly regulated process and its patterns constantly change under various physiological and environmental conditions (Kelemen et al. 2013). Studies showed that AS could influence key molecules in the Toll-like receptor (TLR) signaling pathway to modulate innate immune responses (O'Connor et al. 2015; Wells et al. 2006). For example, an alternatively spliced TLR4 mRNA was significantly induced by lipopolysaccharides (LPS) of Gram-negative bacteria in a mouse macrophage cell line (Iwami et al. 2000). The corresponding TLR4 protein significantly inhibited LPS-mediated TNF-α production and NF-κB activation, suggesting that this protein functioned as a feedback mechanism to inhibit the excessive LPS responses in mouse macrophages.

Furthermore, AS could play crucial roles on the differentiation and activation of lymphocytes in adaptive immunity (Lynch 2004; Yabas et al. 2015). For instance, CD45 (lymphocyte common antigen) pre-mRNA was alternatively spliced to produce proteins with distinct extracellular domains in response to T-cell activation (Altin and Sloan 1997; Lynch and Weiss 2000).

Catfish is the major aquaculture species in the United States. However, the catfish industry has been hindered in recent years due to disease problems. Among them, enteric septicemia of catfish (ESC) is the most devastating disease in the catfish industry (Hawke et al. 1981; Plumb and Hanson 2011), causing an economic loss of millions of dollars annually (Shoemaker et al. 2009). ESC is caused by a rod-shaped, Gram-negative bacterium *Edwardsiella ictaluri*, which has spread throughout the catfish industry, infecting both young fish and mature food-size fish.

Significant progress has been made in genome research with catfish. Milestones include generation of a high-quality reference genome sequence of channel catfish (*Ictalurus punctatus*) (Liu et al. 2016), understanding of genomic variations including millions of polymorphic markers (Liu et al. 2011; Sun et al. 2014) and development of high-density SNP arrays (Liu et al. 2014; Zeng et al. 2017) used for mapping the whole genome (Li et al. 2014; Liu et al. 2016; Zeng et al. 2017) and for the analysis of quantitative trait loci (QTL) for important performance traits such as disease resistance (Geng et al. 2015; Zhou et al. 2017; Tan et al. 2018; Shi et al. 2018), low oxygen tolerance (Wang et al. 2017; Zhong et al. 2017), heat tolerance (Jin et al. 2017), growth rate (Li et al. 2017), head size (Geng et al. 2016), and body conformation (Geng et al. 2017). However, the role of AS in regulation of gene expression, especially its dynamic change after bacterial infections

remains unknown. Although AS profiles detected by EST and cDNA data have been reported in catfish (Lu et al. 2010), the scope of the study was limited because of the low levels of transcriptome coverage. High-throughput cDNA sequencing (RNA-Seq) has become a powerful and cost-effective technology for detection and quantification of AS, as it has allowed the identification of novel transcript isoforms and quantitation of their dynamic changes with high sensitivity and accuracy (Park et al. 2013). Taking the advantage of RNA-Seq, here we report global AS profiles, and the impact of *E. ictaluri* infection on AS in catfish.

3.3 Methods

3.3.1 Data generation

To gain insights into the AS profiles of channel catfish, Illumina RNA-Seq datasets were downloaded from NCBI Sequence Read Archive (SRA) (Liu et al. 2012). Channel catfish genome sequence and annotation file were also downloaded (Liu et al. 2016). AS analyses after bacterial infection were conducted using available RNA-Seq datasets from our previous study of channel catfish in response to *E. ictaluri* infection (Li et al. 2012).

3.3.2 Identification of AS in channel catfish

Channel catfish reference genome sequence and annotation (Liu et al. 2016) and RNA-Seq datasets generated from 19 pooled RNA samples (Liu et al. 2012) were utilized for the analysis of AS profiles. First, the quality of paired-end raw data was evaluated by the FastQC tool (Andrews

2010). They were then subjected to quality control with FASTX toolkit (Gordon and Hannon 2010) to trim biases in the 5′ sequences (Hansen et al. 2010), with Trimmomatic (Bolger et al. 2014) to trim adaptor sequences and low-quality reads (quality score less than 20 and read length shorter than 35). The cleaned reads were tested once again to ensure the read quality using FastQC. High-quality reads were aligned against the channel catfish reference genome sequence using TopHat2 (Kim et al. 2013). The transcripts were assembled from the mapped reads with Cufflinks 2.2.1 using the "-g" option, by which the transcript assembly output includes the reference transcripts as well as novel assembled isoforms (Trapnell et al. 2012). Then, ASTALAVISTA was used to identify and classify the different types of AS (Foissac and Sammeth 2007). This tool can classify not only five basic AS types, but also complex types. Alternatively spliced (AS) genes were also determined.

3.3.3 Identification of AS after infection with *E. ictaluri*

AS profiles and their associated responding genes after *E. ictaluri* infection were determined by using the bioinformatic pipeline mentioned above with minor modifications. The transcripts were assembled from the mapped reads with Cufflinks 2.2.1 without using the "-g" option. Previous studies have shown that the number of AS genes and the number of AS events were positively correlated with sequencing depth when sequencing data was limited (Huang et al. 2015). Thus, the same number of high-quality reads (based on the smallest number of high-quality

reads among samples in a study) was randomly chosen for the subsequent identification of AS genes and AS events to compare AS profiles between control and challenged samples.

3.3.4 Identification of differential alternative splicing

rMATS (Shen et al. 2014) was used to detect differential alternative splicing (DAS) in each of the post-infection sample with the control sample. rMATS is a computational tool which defines DAS by computing and comparing the inclusion level [percent spliced in (PSI)] of certain AS event between two sets of RNA-Seq datasets. For example, rMATS counts the ratio of junction and exon reads for the exon to all reads (junction and exon reads for the exon and its two flanking exons) (Shen et al. 2014; Vitulo et al. 2014). An FDR-adjusted *P* value less than 0.05 was used as criteria for DAS. DAS genes were also determined.

3.3.5 Functional enrichment analysis

To identify enriched functional classes in the DAS gene sets, a statistical overrepresentation test (Cho and Campbell 2000) was performed using the PANTHER classification system (http://pantherdb.org). To use PANTHER, zebrafish orthologs were searched through BLAST program and ENSEMBL database (http://www.ensembl.org).

3.3.6 Gene analysis

To provide an insight into the potential mechanisms and functional implication of AS regulation, one representative DAS gene, *srsf2*, was selected. Sequence analyses of isoform switches were conducted based on rMATS results and the assembled transcripts from RNA-Seq data.

3.3.7 Quantitative real-time PCR analysis

To assess the results from bioinformatic analysis, representative DAS genes identified by RNA-Seq were validated by quantitative real-time PCR (qPCR). Samples were obtained from our recent ESC challenge (Jin et al. 2018). Total RNA was extracted from intestine using RNeasy Mini kit (Qiagen, USA) following the manufacturer's instruction. RNA concentration and integrity were measured using NanoDrop 2000 spectrophotometer (NanoDrop Technologies, USA) and agarose gel electrophoresis. Reverse transcription was performed using the iScript cDNA Synthesis Kit (Bio-Rad, USA). All the cDNA products were then diluted to 250 ng/μL. The *srsf2* gene and five other representative genes (*ikzf1*, *mst1r-like*, *thrb*, *nbr1-like*, and *lmo7*) with DAS events of inclusion-level difference larger than 0.4 were selected for qPCR. Two DAS events of *srsf2* occurred at both 24 h and 72 h post-infection, while the other five events showed DAS at 72 h after infection. For each of seven significant AS events, two pairs of primers were designed for inclusion isoforms and exclusion isoforms. The specific primers were designed by Primer3 software (Rozen and Skaletsky 2000) and listed in Table 4. The primer specificity was accessed by aligning with the channel catfish reference genome database (Liu et al. 2016) using BLASTN with E-value of

1e-10. Standard PCR was first conducted to ensure the expected PCR product sizes. The qPCR was performed using PerfeCTa SYBR Green FastMix Reaction Mixes (Quanta Bio, USA) on a CFX96 real-time PCR detection system (Bio-Rad, USA). Expression of the *18S* gene was utilized as an internal reference to standardize the quantification of target cDNA. Relative quantification of gene expression was analyzed with the Bio-Rad CFX Manager Software.

Table 4. Primers used for qPCR validation of seven differential alternative splicing (DAS) events

Gene	Gene symbol	Inclusion/exclusion isoforms	Primers
ikaros family zinc finger 1	*ikzf1*	Inclusion	F: 5' CGGAGTGATGAGGAGAACGG R: 5' AGGAAGCACCACACTGACTG
		Exclusion	F: 5' ATCGCACTGACAAACCCACT R: 5' ACTTGTGGGGCTTTCCAACA
macrophage-stimulating protein receptor-like	*mst1r-like*	Inclusion	F: 5' GGCATGAGAGTAAGGTTTGCC R: 5' AGGGCTGACAGGGATATGGT
		Exclusion	F: 5' AGGGAGACTACAGACGCGAG R: 5' GGGCTGACAGGGATATGGTC
thyroid hormone receptor beta	*thrb*	Inclusion	F: 5' AGAAACGCAAATTCCTGAGTGC R: 5' ACATAGGCAGCTTTTTGGCG
		Exclusion	F: 5' ACGCAAATTCCTGGCTGAAG R: 5' GCTCTCGGAGTCATAACGCA
next to BRCA1 gene 1 protein-like	*nbr1-like*	Inclusion	F: 5' CTACCTCAAATGCACCCGCT R: 5' GTCAGAAGTCCAGCCGATGT
		Exclusion	F: 5' CTCCAGTCCAGAGGGTCCCAA R: 5' GTCAGAAGTCCAGCCGATGTT
LIM domain 7	*lmo7*	Inclusion	F: 5' AAAACCTTAAAAGGAGATCGGAGT R: 5' CGCTCTTGCTCTTTTTGCCA

		Exclusion	F: 5’ TGAGAAACAAAGGTGGATCTGAA R: 5’ TTTGCCATTTCTCCTGTCGC
serine/arginine-rich splicing factor 2	*srsf2*	Inclusion (DAS event 1)	F: 5’ TGGCCAAAATGTAGCTGTCCT R: 5’ TGTCTGGGTTCTTGCCCATT
		Exclusion (DAS event 1)	F: 5’ TGAGCGCGAGTTCAGATGAT R: 5’ GACCGAGTTTGCAACAACCTT
		Inclusion (DAS event 2)	F: 5’ TGCGCTTAAGTGGATTCTTAGGT R: 5’ CGTGCATGTGCTAGGGCTAA
		Exclusion (DAS event 2)	F: 5’ TCAGACCAAAAGGCGCAGT R: 5’ AATCAAGGCGTTACACGGGA
18S rRNA	*18S*	-	F: 5’ GAGAAACGGCTACCACATCC R: 5’ GATACGCTCATTCCGATTACAG

3.4 Results

3.4.1 AS profiles in channel catfish

Using the genome information and transcriptome from pooled RNA samples isolated from nineteen tissues of a doubled haploid channel catfish (Liu et al. 2012; Table 5), 27,476 AS events from 9,694 AS genes were identified in channel catfish. The most abundant event was exon skipping, accounting for 29.5% of all AS events, followed by alternative 3’ splice site (24.0%), complex splicing (21.1%), alternative 5’ splice site (12.8%), and retained intron (10.3%); mutually exclusive exon occurred in only 2.2% of the AS events (Figures. 9a and 9b, Table 6). The distribution of AS genes and AS events on each chromosome is summarized in Table 7. The numbers of AS genes from the 29 chromosomes were largely proportional to the chromosome sizes, ranging from 483 (Chr 1) to 206 (Chr 29). The average number of AS events per gene was

approximately three. The distribution of AS events along the genome sequence is shown in Figure 9C, and apparently there were some AS hot spots along the genome, especially on chromosomes 1, 20 and 25. For example, at location 15.8-16.0 Mb of chromosome 25, the number of AS events was 165, while the average number of AS events was 35 per Mb.

Table 5. Summary of RNA-Seq paired-end reads from nineteen tissues of a doubled haploid channel catfish (*Ictalurus punctatus*) before and after quality control

RUN	Number of raw reads	Avg.length of raw reads (bp)	Number of reads after trimming	Percentage retained
SRR392744	266,626,644	100	221,764,482	83.17%

Table 6. Alternative splicing events in double haploid channel catfish (*Ictalurus punctatus*)

	Alternative splicing event	Percentage
Exon skipping	8,100	29.5%
Mutually exclusive exon	608	2.2%
Alternative 5’ splice site	3,530	12.8%
Alternative 3’ splice site	6,594	24.0%
Intron retention	2,838	10.3%
Complex	5,806	21.1%
Total	27,476	100%

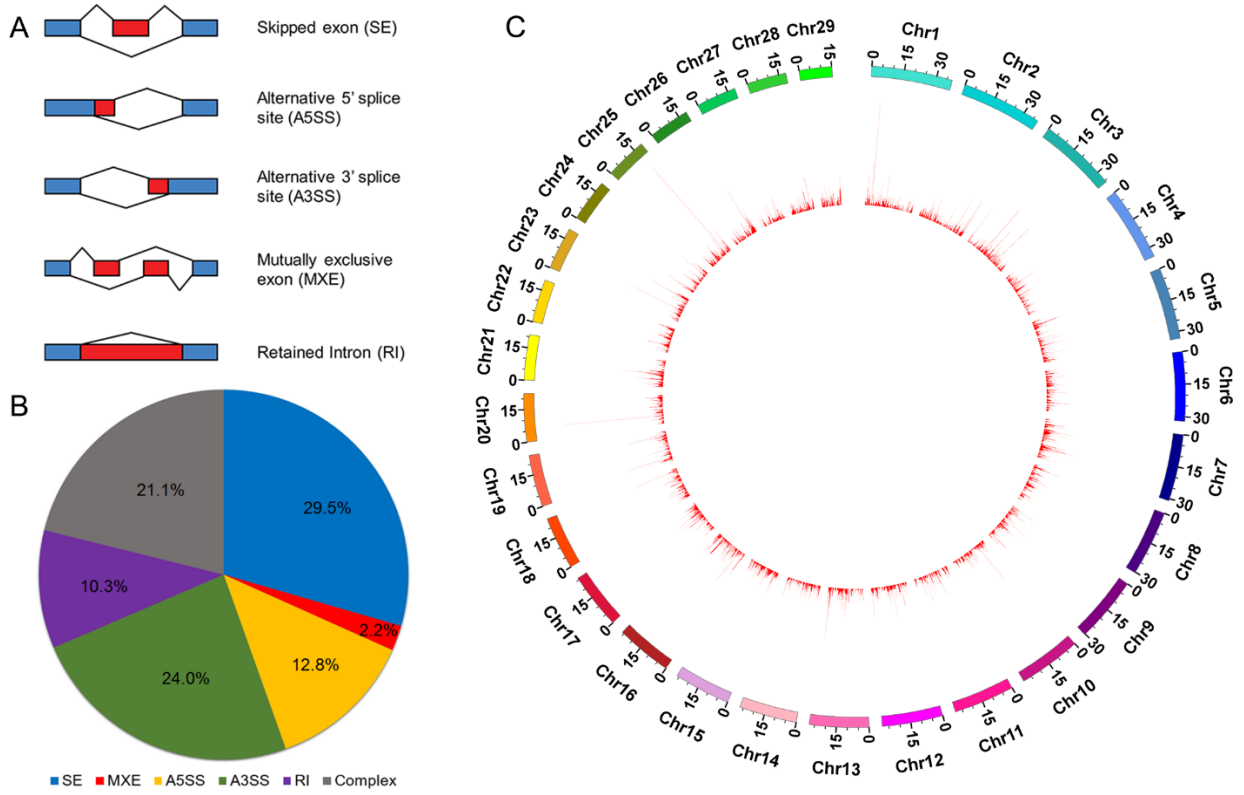

Figure 9. Alternative splicing profile in channel catfish, *Ictalurus punctatus*. A) Schematic representation of five basic types of AS events. Alternative exons are shown as red boxes, and flanking constitutive exons are shown as blue boxes. B) Pie chart showing the percentage distribution of AS events. C) A Circos plot showing the distribution of AS events along the channel catfish genome.

Table 7. Number of the alternative splicing events and alternatively spliced genes identified from each chromosome of the channel catfish, *Ictalurus punctatus*

Chromosome	Alternative splicing events	Alternatively spliced genes	Events per gene
1	1,397	483	2.89

2	1,284	464	2.77
3	1,263	428	2.95
4	898	341	2.63
5	1,137	445	2.56
6	1,033	361	2.86
7	1,263	410	3.08
8	937	369	2.54
9	973	370	2.63
10	882	344	2.56
11	862	292	2.95
12	997	375	2.66
13	1,111	348	3.19
14	837	343	2.44
15	943	359	2.63
16	1,080	336	3.21
17	902	329	2.74
18	928	349	2.66
19	859	252	3.41
20	849	262	3.24
21	800	270	2.96
22	673	274	2.46
23	744	242	3.07
24	674	237	2.84
25	868	269	3.23
26	807	288	2.80
27	616	218	2.83
28	731	254	2.88
29	685	206	3.33
Unassigned	443	176	2.52
Total	27,476	9,694	2.83

3.4.2 Infection induced increase of AS

Summary of RNA-Seq data from the studies of *E. ictaluri* infection is shown in Table 8. The distribution of AS events at each time point after *E. ictaluri* infection is shown in Figure 10A. Detailed information of the distribution of AS events during infection is listed in Table 9. The dynamic changes of the number of AS genes and AS events during the ESC progression are shown in Figure 10B. After the bacterial challenge, the number of AS genes and AS events both increased quickly at early time points. As the disease progressed, the infection-induced AS declined but stayed at higher levels than the control.

Table 8. Summary of RNA-Seq paired-end reads from intestine tissues during *Edwardsiella ictaluri* infection before and after quality control

Sample type	RUN	Number of raw reads	Avg.length of raw reads (bp)	Number of reads after trimming	Percentage retained
Control	SRR357325	52,779,050	101	47,989,506	90.93%
3 h	SRR357322	54,204,696	101	49,289,416	90.93%
24 h	SRR357323	45,803,946	101	41,561,296	90.74%
72 h	SRR357324	44,771,458	101	40,432,304	90.31%
Total		197,559,150		179,272,522	

Table 9. Alternative splicing events predicted from intestine tissues after *Edwardsiella ictaluri* infection

	Control	3 h	24 h	72 h
Exon skipping	997 (30.9%)	1293 (32.8%)	1093 (29.8%)	1053 (28.8%)
Mutually exclusive exon	49 (1.5%)	62 (1.6%)	64 (1.7%)	67 (1.8%)

Alternative 5’ splice site	453 (14.0%)	550 (14.0%)	541 (14.7%)	512 (14.0%)
Alternative 3’ splice site	1135 (35.1%)	1323 (33.6%)	1297 (35.3%)	1333 (36.4%)
Intron retention	334 (10.3%)	399 (10.1%)	422 (11.5%)	409 (11.2%)
Complex	262 (8.1%)	311 (7.9%)	256 (7.0%)	287 (7.8%)
Total	3230	3938	3673	3661

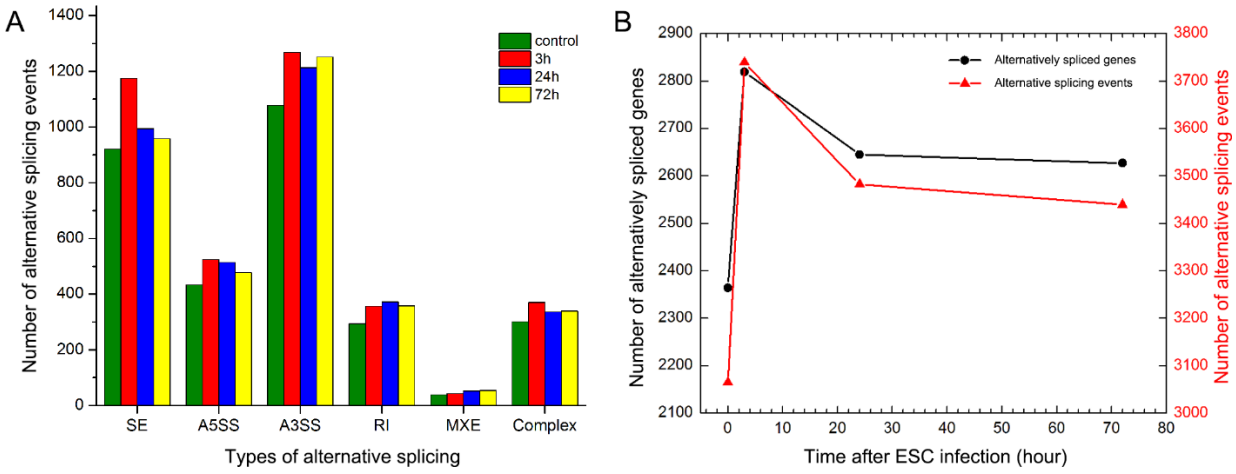

Figure 10. Changes of alternative splicing after *Edwardsiella ictaluri* infection, showing the distribution of different AS types (A), and the number of AS genes and AS events (B) at each time point after *E. ictaluri* infection. For the control samples, a master pool containing equal amounts of each replicate at all three post-infection time points was formed together and indicated as 0 h.

DAS genes and events were determined by comparing those after infection with those of the controls. A total of 5, 24 and 58 DAS events were identified at 3 h, 24 h and 72 h after *E. ictaluri* infection (Table 10), corresponding to 5, 23 and 54 DAS genes, respectively.

Table 10. Differential alternative splicing (DAS) events at each time point after infection with *Edwardsiella ictaluri* as compared with the control. The two numbers in the parentheses separated by colon (:) indicate the number of significant AS events that have higher inclusion levels of the specified events for control sample (first number) or after infection (second number).

	Control_3 h	Control _24 h	Control _72 h
Exon skipping	0 (0:0)	2 (1:1)	34 (15:19)
Mutually exclusive exon	3 (0:3)	0 (0:0)	1 (1:0)
Alternative 5' splice site	0 (0:0)	6 (2:4)	5 (1:4)
Alternative 3' splice site	2 (1:1)	7 (2:5)	3 (0:3)
Intron retention	0 (0:0)	9 (3:6)	15 (9:6)
Total	5	24	58

3.4.3 Enriched DAS genes after *E. ictaluri* infection

Zebrafish orthologs of catfish DAS genes were searched and used for the PANTHER overrepresentation test to reveal the enrichment of functional categories. From the initial list of 5, 23, and 54 DAS genes at 3 h, 24 h, and 72 h after *E. ictaluri* infection, PANTHER was able to map 5, 20, and 47 genes, respectively. Statistically significant ($P < 0.05$, Bonferroni correction for multiple testing) molecular functions and biological processes of the DAS genes were identified (Figure 11 and Tables 11-12). The overrepresentation tests revealed RNA binding and RNA splicing as the enriched molecular function and biological process after *E. ictaluri* infection, respectively.

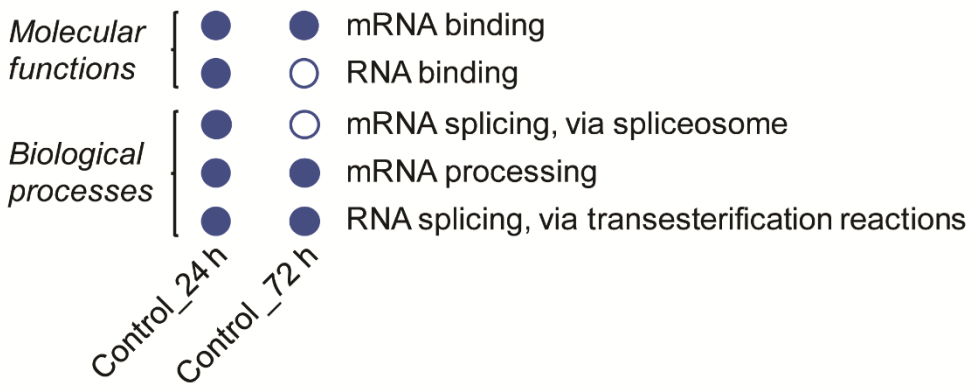

Figure 11. Enrichment of functional categories in differentially alternatively spliced (DAS) genes after infection with *Edwardsiella ictaluri*. The solid circle and open circle represent the presence or absence, respectively, of certain molecular function or biological process that was enriched in the DAS genes.

Table 11. PANTHER overrepresentation analysis of differentially alternatively spliced (DAS) genes at 24 h after infection with *Edwardsiella ictaluri*

PANTHER classification category	Number of genes in reference list	Number of genes in target list	Expected	Fold enrichment	Over-/under-representation (+/-)	P value
Molecular functions						
mRNA binding	146	4	0.12	34.31	+	9.04E-04
RNA binding	391	5	0.31	16.01	+	2.06E-03
Biological processes						
RNA splicing, via transesterification reactions	154	5	0.12	40.65	+	2.89E-05
mRNA splicing, via spliceosome	176	5	0.14	35.57	+	5.57E-05

mRNA processing	269	5	0.21	23.27	+	4.44E-04

Table 12. PANTHER overrepresentation analysis of differentially alternatively spliced (DAS) genes at 72 h after infection with *Edwardsiella ictaluri*

PANTHER classification category	Number of genes in reference list	Number of genes in target list	Expected	Fold enrichment	Over-/under-representation (+/-)	P value
Molecular functions						
mRNA binding	146	4	0.27	14.60	+	2.93E-02
Biological processes						
RNA splicing, via transesterification reactions	154	4	0.29	13.84	+	4.73E-02
mRNA processing	269	6	0.50	11.88	+	2.59E-03

3.4.4 Gene analysis

GO analyses using DAS genes revealed the overrepresentation of genes associated with RNA binding and RNA splicing. Among the splicing factors, the most well studied and important ones were serine/arginine-rich (SR) splicing factors (Long and Caceres 2009). We found that one SR splicing factor, *srsf2*, was alternatively spliced at 24 h and 72 h after infection, with significantly differential intron retention. Three transcript isoforms were identified (isoforms 1-3, Figure 12), isoform 2 and 3 had different introns excluded. The isoform 2 may undergo NMD based on the transcript structure with the translation terminating 1406 (> 50-55) nucleotides

upstream of the final EJC. The inclusion level of the alternative intron significantly decreased at 24 h and 72 h after the bacterial infection. For the isoform 3, the translation terminated 12 (< 50-55) nucleotides upstream of the final EJC, which may be free of NMD.

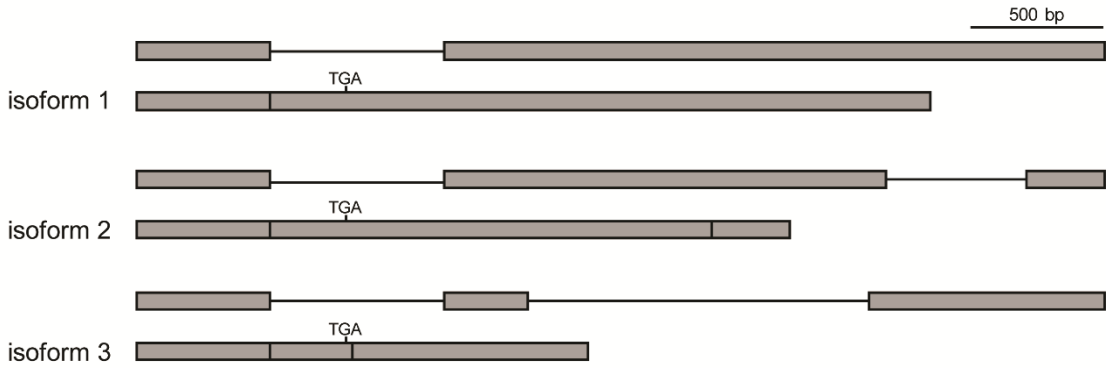

Figure 12. Schematic presentation of the alternative splicing isoforms of the catfish *srsf2* gene. Exons are indicated by boxes, and introns by lines. Transcript isoforms are indicated on the left margin. The stop codon is indicated by TGA.

3.4.5 Quantitative real-time PCR analysis

We selected *srsf2* and five other representative DAS genes, corresponding to seven DAS events, for qPCR validation. Agarose gel electrophoresis and melting curve analysis using CFX Manager Software revealed a single product for each pair of designed primers. For both inclusion and exclusion isoforms of one DAS event, relative expression values comparing infected group

with control group were generally consistent with the results of RNA-Seq analysis (Figure 13). Since the isoform expression change of *srsf2* at 24 h and 72 h post-infection were quite similar, only the fold change at 72 h post-infection was shown in Figure 13. It was noted that the relative expression change of inclusion isoforms was different from the change of exclusion isoforms for each selected DAS event.

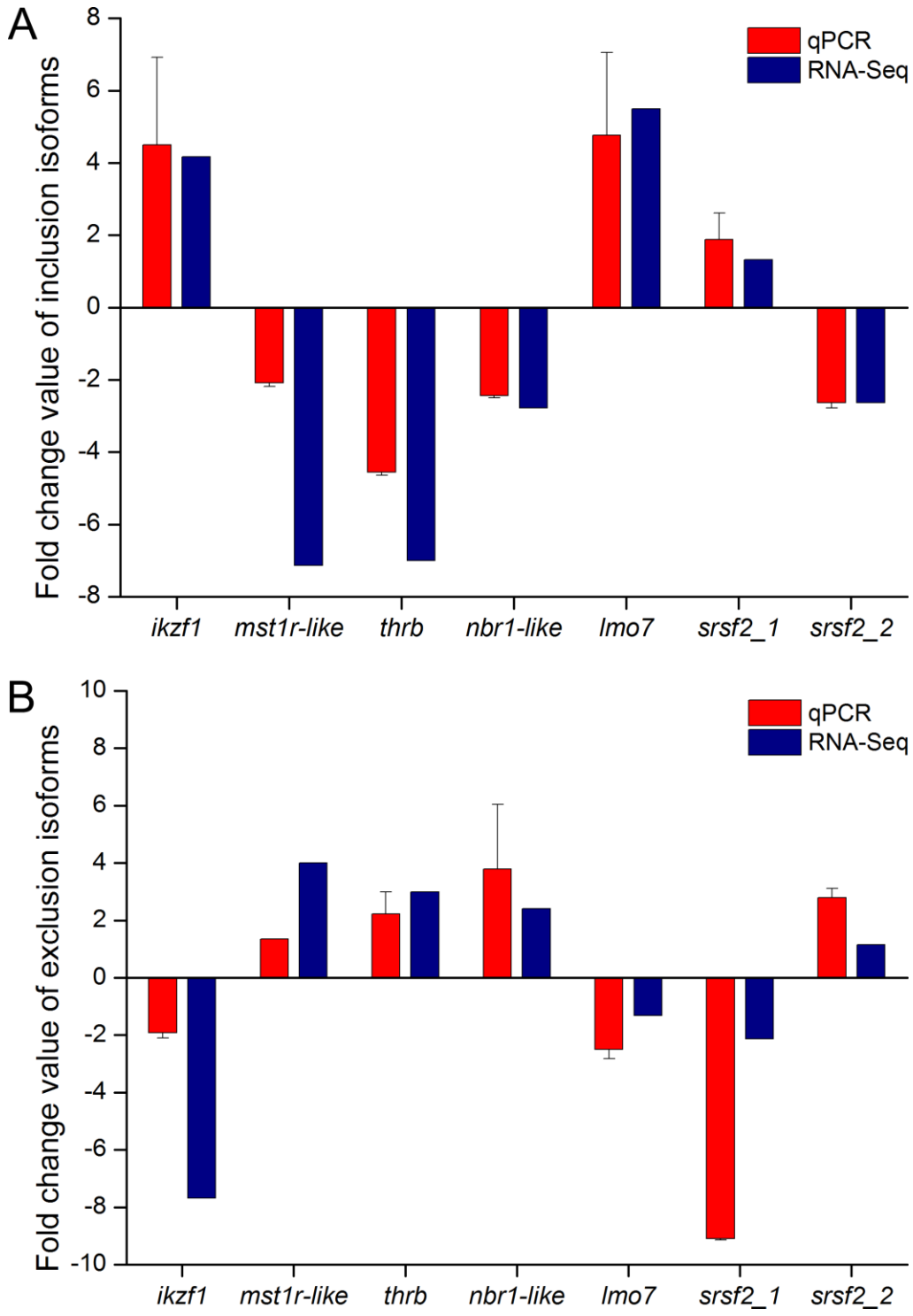
A
Fold change value of inclusion isoforms
8
6
4
2
0
-2
-4
-6
-8
qPCR
RNA-Seq
ikzf1
mst1r-like
thrb
nbr1-like
lmo7
srsf2_1
srsf2_2
B
Fold change value of exclusion isoforms
10
8
6
4
2
0
-2
-4
-6
-8
-10
qPCR
RNA-Seq
ikzf1
mst1r-like
thrb
nbr1-like
lmo7
srsf2_1
srsf2_2

Figure 13. Quantitative real-time PCR of seven differential alternative splicing (DAS) events of six genes. Fold change value indicates the relative expression of inclusion isoforms (A) or exclusion isoforms (B) at 72 h post-infection in comparison with control group. The number following the symbol "_" represents different DAS event of a same gene.

3.5 Discussion

Aquatic organisms, including fish, have a close relationship with their aquatic environments. They are exposed intimately with multiple stressors such as infectious bacteria. To adapt to the frequent and unpredictable environmental variations, they must possess rapid and effective responses, including regulation of gene expression at both transcriptional and post-transcriptional levels. As one important mechanism of post-transcriptional regulations, AS plays critical roles in tuning the orchestra of gene regulation and serves as a main source of transcriptome and proteome diversity in metazoans (Chen and Manley 2009; Maniatis and Tasic 2002). Investigations on AS after disease infections should provide better understanding on AS regulation, facilitating the enhancement of immune regulatory networks.

Our study was designed to investigate the AS profiles in catfish and the changes of AS in response to the *E. ictaluri* infection. Utilizing genomic data and large RNA-Seq datasets of channel catfish (Liu et al. 2012, 2016), we observed that exon skipping was the most abundant AS type,

accounting for 29.5% of all AS events (Figure 9B and Table 6). Mutually exclusive exon represented the least AS type, occurring in only 2.2% of the AS events. These results are similar to those observed in human, mouse and chicken (Chacko and Ranganathan 2009). Interestingly, complex AS events accounted for a relatively large proportion of all AS events, consistent with the situation found in many other species (Sammeth et al. 2008). Moreover, AS events occurred throughout the genome (Figure 9C), indicating its prevalence in channel catfish.

The change in AS was not random but tightly regulated (Palusa et al. 2007). Kelemen et al. (2013) stated that AS patterns in an organism constantly change in response to environmental stimuli. In this study, analyses of AS profiles after *E. ictaluri* infection revealed temporally dynamic and complex patterns of AS (Figure 10). In the ESC challenge, the number of AS genes and AS events increased sharply to high levels at early stages after infection, suggesting the acute changes in AS regulation may be linked with immune responses to *E. ictaluri*.

Many methods have been proposed to analyze and interpret DAS, and the accurate detection and quantification of DAS remain a bioinformatic challenge (Alamancos et al. 2014; Lahat and Grellscheid 2016). DAS can be analyzed at the level of differentially expressed transcript isoform or at the level of individual splicing event. When analyzing splicing mechanism and regulation, focusing on individual splicing events was preferable. In this study, we used rMATS which utilized Bayesian statistics to compare splicing event-based differential splicing between samples (Shen et al. 2014). Compared with control samples, DAS genes were examined at each time point after *E. ictaluri* infection in catfish. Shared DAS genes were observed with

different comparisons, reflecting that some AS genes and events were differentially regulated at various stages of disease progression.

GO analysis provides a classification of genes by molecular function, cellular component and biological process (Ashburner et al. 2000). Functional GO categories of DAS genes between each pair of treatment were identified in this study. One intriguing observation was that genes involved in the biological process of RNA splicing were significantly enriched after infection (Figure 11). This observation was consistent with the finding that splicing factors themselves often undergo regulation by AS in response to environmental cues (Staiger and Brown 2013; Stamm et al. 2005). The most important splicing factors were SR splicing factors, functioning as activators or repressors of both constitutive and AS (Bourgeois et al. 2003). Their activities were tightly linked to cellular and environmental cues (Yabas et al. 2015). In this study, we found that SR splicing factors were differentially spliced after bacterial infection, suggesting their important roles in immune regulation processes.

The transcripts of *srsf2* underwent differential intron retention after *E. ictaluri* infection. Sequence analyses showed that some transcript isoforms may be degraded by NMD due to AS of the intron in the 3' UTR. Interestingly, previous studies showed that *srsf2* could autoregulate the splicing of its own mRNA by AS of one intron in the 3' UTR coupled with NMD (Sureau et al. 2001). Moreover, the AS in 3' UTR of *srsf2* served as a key element in regulating the *srsf2* expression during hematopoietic differentiation in human (Sureau and Perbal 1994). The present

study showed that the proportion of *srsf2* transcript isoform which would be targeted by NMD increased at 24 h and 72 h after *E. ictaluri* infection.

In summary, in this study, we determined the AS profiles in catfish and discovered highly induced levels of AS after *E. ictaluri* infection. One intriguing finding was that RNA processing factors were significantly overrepresented among DAS genes after infection. We also presented one case of *srsf2* gene where isoform switches may be functionally related to immune responses, indicating functional significance of AS regulation in response to biotic stresses. The qPCR validation showed relatively good consistency with RNA-Seq results. Further studies using more tissues and larger sequencing datasets will be warranted to provide a more comprehensive analysis of AS after bacterial infections in catfish.

Note: Chapter 2 was published on Molecular Genetics and Genomics (Tan et al. 2018); Chapter 3 was published on Marine Biotechnology (Tan et al. 2018)

References

Alamancos GP, Agirre E, Eyras E (2014) Methods to study splicing from high-throughput RNA sequencing data. In: Hertel KJ (ed) Spliceosomal Pre-mRNA Splicing, Humana Press, Totowa, NJ, pp 357–397

Altin JG, Sloan EK (1997) The role of CD45 and CD45-associated molecules in T cell activation. Immunol Cell Biol 75:430–445

Andrews S (2010) FastQC: a quality control tool for high throughput sequence data. Cambridge, UK: Babraham Institute

Ashburner M, Ball CA, Blake JA, Botstein D, Butler H, Cherry JM, Davis AP, Dolinski K, Dwight SS, Eppig JT, Harris MA, Hill DP, Issel-Tarver L, Kasarskis A, Lewis S, Matese JC, Richardson JE, Ringwald M, Rubin GM, Sherlock G (2000) Gene Ontology: tool for the unification of biology. Nat Genet 25:25–29

Bolger AM, Lohse M, Usadel B (2014) Trimmomatic: a flexible trimmer for Illumina sequence data. Bioinformatics 30:2114–2120

Bourgeois CF, Lejeune F, Stévenin J (2003) Broad specificity of SR (serine/arginine) proteins in the regulation of alternative splicing of pre-messenger RNA. Prog Nucleic Acid Res Mol Biol 78:37–88

Chacko E, Ranganathan S (2009) Comprehensive splicing graph analysis of alternative splicing patterns in chicken, compared to human and mouse. BMC Genomics 10:S5

Chen M, Manley JL (2009) Mechanisms of alternative splicing regulation: insights from molecular and genomics approaches. Nat Rev Mol cell Biol 10:741–754

Cho RJ, Campbell MJ (2000) Transcription, genomes, function. Trends Genet 16:409–415

Harrow J, Frankish A, Gonzalez JM, Tapanari E, Diekhans M, Kokocinski F, Bronwen LA, Barrell D, Zadissa A, Searle Stephen, Barnes I, Bignell A, Boychenko Veronika, Hunt T, Kay M, Mukherjee G, Rajan J, Despacio-Reyes G, Saunders G, Steward C, Harte R, Lin M, Howald

C, Tanzer A, Derrien T, Chrast J, Walters N, Balasubramanian S, Pei B, Tress M, Rodriguez JM, Ezkurdia I, Baren JV, Brent M, Haussler D, Kellis M, Valencia A, Reymond A, Gerstein M, Guigó R, Hubbard TJ (2012) GENCODE: the reference human genome annotation for The ENCODE Project. Genome Res 22:1760–1774

Foissac S, Sammeth M (2007) ASTALAVISTA: dynamic and flexible analysis of alternative splicing events in custom gene datasets. Nucleic Acids Res 35:W297–W299

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

Geng X, Liu S, Yuan Z, Jiang Y, Zhi D, Liu Z (2017) A genome-wide association study reveals that genes with functions for bone development are associated with body conformation in catfish. Mar Biotechnol 19:570–578

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

Gordon A, Hannon GJ (2010) Fastx-toolkit. FASTQ/A short-reads preprocessing tools (unpublished) http://hannonlab.cshl.edu/fastx_toolkit

Hansen KD, Brenner SE, Dudoit S (2010) Biases in Illumina transcriptome sequencing caused by random hexamer priming. Nucleic Acids Res 38:e131–e131

Hawke JP, McWhorter AC, Steigerwalt AG, Brenner DJ (1981) *Edwardsiella ictaluri* sp. nov., the causative agent of enteric septicemia of catfish. Int J Syst Bacteriol 31:396–400

Huang J, Gao Y, Jia H, Liu L, Zhang D, Zhang Z (2015) Comparative transcriptomics uncovers alternative splicing changes and signatures of selection from maize improvement. BMC Genomics 16:363

Iwami K, Matsuguchi T, Masuda A, Kikuchi T, Musikacharoen T, Yoshikai Y (2000) Cutting edge: naturally occurring soluble form of mouse Toll-like receptor 4 inhibits lipopolysaccharide signaling. J Immunol 165:6682–6686

Jin Y, Zhou T, Geng X, Liu S, Chen A, Yao J, Jiang C, Tan S, Su B, Liu Z (2017) A genome-wide association study of heat stress-associated SNPs in catfish. Anim Genet 48:233–236

Jin Y, Zhou T, Li N, Liu S, Xu X, Pan Y, Tan S, Shi H, Yang Y, Yuan Z, Wang W, Luo J, Gao D, Dunham R, Liu Z (2018) JAK and STAT members in channel catfish: Identification, phylogenetic analysis and expression profiling after *Edwardsiella ictaluri* infection. Dev Comp Immunol 81:334–341

Kelemen O, Convertini P, Zhang Z, Wen Y, Shen M, Falaleeva M, Stamm S (2013) Function of alternative splicing. Gene 514:1–30

Kim D, Pertea G, Trapnell C, Pimentel H, Kelley R, Salzberg SL (2013) TopHat2: accurate alignment of transcriptomes in the presence of insertions, deletions and gene fusions. Genome Biol 14:R36

Lahat A, Grellscheid SN (2016) Differential mRNA Alternative Splicing. In: Aransay AM, Trueba JLL (eds) Field Guidelines for Genetic Experimental Designs in High-Throughput Sequencing, Springer, Cham, pp 105–119

Li C, Zhang Y, Wang R, Lu J, Nandi S, Mohanty S, Terhune J, Liu Z, 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 Shellfish Immunol 32:816–827

Li N, Zhou T, Geng X, Jin Y, Wang X, Liu S, Xu X, Gao D, Li Q, Liu Z (2017) Identification of novel genes significantly affecting growth in catfish through GWAS analysis. Mol Genet Genomics 293:587–599

Li Y, Liu S, Qin Z, Waldbieser GC, Wang R, Sun L, Bao L, Danzmann RG, Dunham RA, Liu Z (2014) Construction of a high-density, high-resolution genetic map and its integration with BAC-based physical map in channel catfish. DNA Res 22:39–52

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

Liu S, Zhang Y, Zhou Z, Waldbieser GC, Sun F, Lu J, Zhang J, Jiang Y, Zhang H, Wang X, Rajendran K, Khoo L, Kucuktas H, Peatman E, Liu Z (2012) Efficient assembly and annotation of the transcriptome of catfish by RNA-Seq analysis of a doubled haploid homozygote. BMC Genomics 13:595

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

Long JC, Caceres JF (2009) The SR protein family of splicing factors: master regulators of gene expression. Biochem J 417:15–27

Lu J, Peatman E, Wang W, Yang Q, Abernathy J, Wang S, Kucuktas H, Liu Z (2010) Alternative splicing in teleost fish genomes: same-species and cross-species analysis and comparisons. Mol Genet Genomics 283:531–539

Lynch KW (2004) Consequences of regulated pre-mRNA splicing in the immune system. Nat Rev Immunol 4:931–940

Lynch KW, Weiss A (2000) A model system for activation-induced alternative splicing of CD45 pre-mRNA in T cells implicates protein kinase C and Ras. Mol Cell Biol 20:70–80

Maniatis T, Tasic B (2002) Alternative pre-mRNA splicing and proteome expansion in metazoans. Nature 418:236–243

Modrek B, Resch A, Grasso C, Lee C (2001) Genome-wide detection of alternative splicing in expressed sequences of human genes. Nucleic Acids Res 29:2850–2859

O'Connor BP, Danhorn T, De Arras L, Flatley BR, Marcus RA, Farias-Hesson E, Leach SM, Alper S (2015) Regulation of toll-like receptor signaling by the SF3a mRNA splicing complex. PLoS Genet 11:e1004932

Ozsolak F, Milos PM (2011) RNA sequencing: advances, challenges and opportunities. Nat Rev Genet 12:87

Palusa SG, Ali GS, Reddy ASN (2007) Alternative splicing of pre-mRNAs of Arabidopsis serine/arginine-rich proteins: regulation by hormones and stresses. Plant J 49:1091–1107

Pan Q, Shai O, Lee LJ, Frey BJ, Blencowe BJ (2008) Deep surveying of alternative splicing complexity in the human transcriptome by high-throughput sequencing. Nat Genet 40:1413–1415

Pan Q, Shai O, Misquitta C, Zhang W, Saltzman AL, Mohammad N, Babak T, Siu H, Hughes TR, Morris QD, Frey BJ, Blencowe BJ (2004) Revealing global regulatory features of mammalian alternative splicing using a quantitative microarray platform. Mol Cell 16:929–941

Park JW, Tokheim C, Shen S, Xing Y (2013) Identifying differential alternative splicing events from RNA sequencing data using RNASeq-MATS. In: Shomron N (ed) Deep Sequencing Data Analysis, Humana Press, Totowa, NJ, pp 171–179

Plumb JA, Hanson LA (2011) Health maintenance and principal microbial diseases of cultured fshes. Wiley, Oxford

Rozen S, Skaletsky H (2000) Primer3 on the WWW for general users and for biologist programmers. In: Misener S, Krawetz SA (eds) Bioinformatics Methods and Protocols, Humana Press, Totowa, NJ, pp 365–386

Sammeth M, Foissac S, Guigó R, Brenner S, Spieth J (2008) A general definition and nomenclature for alternative splicing events. PLoS Comput Biol 4:e1000147

Sánchez-Pla A, Reverter F, de Villa MCR, Comabella M (2012) Transcriptomics: mRNA and alternative splicing. J Neuroimmunol 248:23–31

Shen S, Park JW, Lu Z, Lin L, Henry MD, Wu Y, Zhou Q, Xing Y (2014) rMATS: robust and flexible detection of differential alternative splicing from replicate RNA-Seq data. Proc Natl Acad Sci 111:E5593–E5601

Shi H, Zhou T, Wang X, Yang Y, Wu C, Liu S, Bao L, Li N, Yuan Z, Jin Y, Tan S, Wang W, Zhong X, Qin G, Gao D, Dunham RA, Liu Z (2018) Genome-wide association analysis of intra-specific QTL associated with the resistance for enteric septicemia of catfish. Mol Genet Genomics https://doi.org/10.1007/s00438-018-1463-0

Shoemaker CA, Klesius PH, Evans JJ, Arias CR (2009) Use of modifed live vaccines in aquaculture. J World Aquacult Soc 40:573–585

Staiger D, Brown JWS (2013) Alternative splicing at the intersection of biological timing, development, and stress responses. Plant Cell 25:3640–3656

Stamm S, Ben-Ari S, Rafalska I, Tang Y, Zhang Z, Toiber D, Thanaraj TA, Soreq H (2005) Function of alternative splicing. Gene 344:1–20

Sun L, Liu S, Wang R, Jiang Y, Zhang Y, Zhang J, Bao L, Kaltenboeck L, Dunham RA, Waldbieser GC, Liu Z (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

Sureau A, Gattoni R, Dooghe Y, Stevenin J, Soret J (2001) SC35 autoregulates its expression by promoting splicing events that destabilize its mRNAs. EMBO J 20:1785–1796

Sureau A, Perbal B (1994) Several mRNAs with variable 3'untranslated regions and different stability encode the human PR264/SC35 splicing factor. Proc Natl Acad Sci 91:932–936

Tan S, Zhou T, Wang W, Jin Y, Wang X, Geng X, Luo J, Yuan Z, Yang Y, Shi H, Gao D, Dunham RA, Liu Z (2018) GWAS analysis using F2 interspecific hybrids reveals superior blue catfish alleles responsible for strong resistance against enteric septicemia of catfish. Mol Genet Genomics 293:1107–1120

Trapnell C, Roberts A, Goff L, Pertea G, Kim D, Kelley DR, Pimentel H, Salzberg SL, Rinn JL, Pachter L (2012) Differential gene and transcript expression analysis of RNA-seq experiments with TopHat and Cufflinks. Nat Protoc 7:562–578

Vitulo N, Forcato C, Carpinelli EC, Telatin A, Campagna D, D'Angelo M, Zimbello R, Corso M, Vannozzi A, Bonghi C, Lucchin M, Valle G (2014) A deep survey of alternative splicing in grape reveals changes in the splicing machinery related to tissue, stress condition and genotype. BMC Plant Biol 14:99

Wang ET, Sandberg R, Luo S, Khrebtukova I, Zhang L, Mayr C, Kingsmore SF, Schroth GP, Burge CB (2008) Alternative isoform regulation in human tissue transcriptomes. Nature 456:470–476

Wang X, Liu S, Jiang C, Geng X, Zhou T, Li N, Bao L, Li Y, Yao J, Yang Y, Dunham RA, Liu Z (2017) Multiple across-strain and within-strain QTLs suggest highly complex genetic architecture for hypoxia tolerance in channel catfish. Mol Genet Genomics 292:63–76

Wells CA, Chalk AM, Forrest A, Taylor D, Waddell N, Schroder K, Himes SR, Faulkner G, Lo S, Kasukawa T, Kawaji H, Kai C, Kawai J, Katayama S, Carninci P, Hayashizaki Y, Hume DA, Grimmond SM (2006) Alternate transcription of the Toll-like receptor signaling cascade. Genome Biol 7:R10

Yabas M, Elliott H, Hoyne GF (2015) The role of alternative splicing in the control of immune homeostasis and cellular differentiation. Int J Mol Sci 17:3

Zeng Q, Fu Q, Li Y, Waldbieser GC, Bosworth B, Liu S, Yang Y, Bao L, Yuan Z, Li N, 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. Sci Rep 7:40347

Zhong X, Wang X, Zhou T, Jin Y, Tan S, Jiang C, Geng X, Li N, Shi H, Zeng Q, Yang Y, Yuan Z, Bao L, Liu S, Tian C, Peatman E, Li Q, Liu Z (2017) Genome-wide association study reveals multiple novel QTL associated with low oxygen tolerance in hybrid catfish. Mar Biotechnol 19:379–390

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. Mol Genet Genomics 292:231–242