The chemokine superfamily: II. The 64 CC chemokines in channel catfish and their involvement in disease and hypoxia responses

Qiang Fu^{a, b}, Yujia Yang^b, Chao Li^c, Qifan Zeng^b, Tao Zhou^b, Ning Li^b, Yang Liu^b, Yun Li^b,
Xiaozhu Wang^b, Shikai Liu^b, Daoji Li^a, Zhanjiang Liu^{b, *}

ABSTRACT

Chemokines are a superfamily of structurally related chemotactic cytokines exerting significant roles in regulating cell migration and activation. Based on the arrangement of the first four cysteine residues, they are classified into CC, CXC, C and CX3C subfamilies. In this study, a complete set of 64 CC chemokine ligand (CCL) genes was systematically identified, annotated, and characterized from the channel catfish genome. Extensive phylogenetic and comparative genomic analyses supported their annotations, allowing establishment of their orthologies, revealing fish-specific CC chemokines and the expansion of CC chemokines in the teleost genomes through lineage-specific tandem duplications. With 64 genes, the channel catfish genome harbors the largest numbers of CC chemokines among all the genomes characterized to date, however, they fall into 11 distinct CC chemokine groups. Analysis of gene expression after bacterial infections indicated that the CC chemokines were regulated in a gene-specific and time-dependent manner. While only one member of CCL19 (CCL19a.1) was significantly up-regulated after ESC infection, all CCL19 members (CCL19a.1, CCL19a.2 and CCL19b) were significantly induced after columnaris infection. In addition, CCL19a.1, CCL19a.2 and CCL19b were also drastically up-regulated in ESC-susceptible fish, but not in resistant fish, suggesting potential significant roles of CCL19 in catfish immune responses. High levels of expression were mostly observed to be related to susceptibility to diseases or intolerance to hypoxia, suggesting that high levels of CC chemokine expression are detrimental to the host.

Key words: chemokine, fish, genome, innate immunity, disease resistance, hypoxia

1. Introduction

Chemokines are a superfamily of chemotactic cytokines exerting pivotal roles in regulating cell migration under both inflammatory and physiological conditions (Nomiyama et al., 2008; Peatman and Liu, 2007; Zlotnik and Yoshie, 2000). They are key immune regulators, acting as a bridge between innate and adaptive immunity (Alejo and Tafalla, 2011; Raman et al., 2011). Specifically, chemokines promote leukocyte mobilization and regulate the immune responses and differentiation of the recruited cells (Alejo and Tafalla, 2011; Esche et al., 2005). In addition to their roles in immunity, chemokines also participate in angiogenesis (Arenberg et al., 1997; Keane et al., 1998), neurological development (Belmadani et al., 2006; Gordon et al., 2009), organogenesis and germ cell migration (Doitsidou et al., 2002; Knaut et al., 2003), hypoxia responses (Bosco et al., 2006; Schioppa et al., 2003), and normal development and growth (Baoprasertkul et al., 2005; David et al., 2002; Zeng et al., 2016). Chemokines are structurally related small peptides, typically 8-15 kDa, with majority of them containing four conserved cysteine residues (Peatman and Liu, 2007; Zlotnik and Yoshie, 2000). Based on the arrangement of the first two of the cysteine residues, chemokines are divided into four subfamilies, CC, CXC, C and CX3C (Bacon et al., 2002). The CC and CXC chemokines are the two largest subfamilies of chemokines.

CC chemokines constitute the largest subfamily of chemokines with 28 members being identified from mammalian species, among which human had the largest number with 24 distinct CC chemokines, and each gene has one single copy (Bacon et al., 2002). The mammalian species

do not have the largest numbers of CC chemokine genes. To date before this work, zebrafish had the largest number of CC chemokine genes with 62 CC chemokine genes (Alejo and Tafalla, 2011; Nomiya et al., 2008; <http://zfin.org/>). However, the diversity of fish CC chemokines is comparatively small because the largest numbers of fish CC chemokine genes were derived from extensive lineage-specific, intrachromosomal tandem duplications.

The majority of human, murine, chicken and zebrafish CC chemokine genes are organized in gene clusters within their genomes. The largest clusters are found on human chromosome 17, mouse chromosome 11, chicken chromosome 19, and zebrafish chromosome 25 (Nomiya et al., 2008; Wang et al., 2005; Zlotnik et al., 2006). There are correlations between genomic architecture and the inducibility of their expression, with inflammatory CC chemokines constituting the large clusters, and a few homeostatic CC chemokines distributed among several chromosomes (Peatman et al., 2006). In addition, orthologies across species are relatively high between the non-clustered chemokines, but low when comparing the clustered CC chemokines of several species (Peatman et al., 2006; Wang et al., 2005).

Seven large groups of fish CC chemokines were established through phylogenetic analysis: the CCL19/21/25 group, the CCL20 group, the CCL27/28 group, the CCL17/22 group, the macrophage inflammatory protein (MIP) group, the monocyte chemotactic protein (MCP) group and a fish-specific group (Peatman and Liu, 2007). Since sequence conservation of CC chemokines is low, and duplication and divergence may have occurred at a significantly faster rate than in other genes, establishing orthology between fish and mammalian CC chemokines has been problematic (Peatman et al., 2006). In addition, concrete orthologues cannot be identified for the majority of CC chemokine genes based on either sequence identities or phylogenetics in

the previous studies of teleost CC chemokines (Laing and Secombes, 2004; Peatman et al., 2005). However, genomic location of CC chemokines is quite important to trace the origins of CC chemokines in teleosts and higher vertebrates. Comparisons of syntenic relationships of CC chemokines should help to understand the evolutionary dynamics of CC chemokine genes, their modes of duplication and divergence that led to the present repertoire of vertebrate CC chemokines. Hence, the major objective of the present study was to determine the entire CC chemokinome of catfish, to establish their orthologies through phylogenetic and syntenic analysis. Additional objectives include analysis of their expression after biotic and abiotic stresses such as after bacterial infections and under hypoxia challenges.

Channel catfish (*Ictalurus punctatus*) is the primary aquaculture species in the United States (Naylor et al., 2000), but its sustainable production is threatened by severe disease outbreaks. In particular, the enteric septicemia of catfish (ESC) caused by *Edwardsiella ictaluri*, columnaris disease caused by *Flavobacterium columnare*, and the emerging Aeromonas disease caused by *Aeromonas hydrophila*, all bacterial diseases, cause major economic losses to the catfish industry. The disease incidents were increased with exposures to environmental stresses, especially when exposed to hypoxia (Geng et al., 2014). Understanding of molecular linkage between disease and hypoxia responses should help strategic thinking of solving these problems in genetic enhancing programs. We previously identified 26 distinct CC chemokine transcripts (SCYa101-SCYa126) from catfish and, in subsequent studies, obtained the genomic sequences and structures of 24 CC chemokine genes (Bao et al., 2006; He et al., 2004; Peatman et al., 2006). However, their orthologies were still obscure and the nomenclature of catfish CC chemokines need to be updated (Zlotnik and Yoshie, 2000). In addition, their involvement in

disease and hypoxia responses has not been fully understood. Here we report a complete repertoire of 64 CC chemokines from channel catfish, the largest among all organisms to date, and their evolutionary dynamics and expression in relation to bacterial infection and hypoxia stress.

2. Materials and methods

2.1. Gene identification and sequence analysis

The channel catfish CC chemokine genes were identified by searching against the channel catfish genome sequence (Liu et al., 2016) and transcriptome datasets, using all available sequences of CC chemokines from zebrafish (*Danio rerio*), salmon (*Salmo salar*), medaka (*Oryzias latipes*), tilapia (*Oreochromis niloticus*), fugu (*Takifugu rubripes*), frog (*Xenopus laevis*), chicken (*Gallus gallus*), mouse (*Mus musculus*), cow (*Bos Taurus*) and human (*Homo sapiens*) retrieved from the NCBI (<http://www.ncbi.nlm.nih.gov/>), Ensemble (<http://www.ensembl.org>) and ZFIN (<http://zfin.org/>). At the same time, the amino acid sequence of 24 previously reported CC chemokines (Bao et al., 2006) identified from catfish EST datasets were also retrieved from NCBI and used in the following steps. To identify the initial catfish CC chemokine sequences, TBLASTN was performed against channel catfish RNA-Seq database, generated by genome-guided assembly (unpublished), with a cutoff E-value of $1e^{-5}$. Other transcriptome assemblies generated from various RNA-Seq studies were also used for the most comprehensive blast search (Li et al., 2012; Liu et al., 2012; Sun et al., 2012). After that, Clustal Omega (<http://www.ebi.ac.uk/Tools/msa/clustalo/>) was used to eliminate duplicates in the initial sequence pool and each sequence was further confirmed to be unique based on their position

information on the genome. Then, the coding sequences were predicted using ORF (opening reading frames) finder (<http://www.ncbi.nlm.nih.gov/gorf/gorf.html>), which were further validated by BLASTP against NCBI non-redundant (nr) protein database. For the purpose of verifying copy number and sequence accuracy of CC chemokines, BLASTN was performed against the channel catfish genome sequence (Liu et al., 2016), using the initial pool of CC chemokine mRNA sequences as queries, with a cutoff E-value of $1e^{-10}$. Fgenesh program of Molquest software (Softberry Int.) was used to predict the genes from retrieved genomic scaffold sequences (Solovyev et al., 2006). The characteristic functional domains were identified using the simple modular architecture research tool (SMART <http://smart.embl-heidelberg.de>) and further endorsed by conserved domain predicted through BLASTP.

2.2. Phylogenetic analysis

Phylogenetic analysis was conducted using the amino acid sequences of CC chemokine from various organisms along evolutionary spectrum including human, mouse, chicken, frog and several fish species with inclusion of channel catfish, zebrafish, tongue sole, fugu and medaka. MUSCLE (multiple sequence comparison by log-expectation) was used to conduct alignment of multiple amino acid sequences with default parameters (Edgar, 2004). MEGA 6 was used to construct the unrooted phylogenetic tree of CC chemokines using the maximum likelihood method (Tamura et al., 2013). Based on the alignment results, JTT (Jones-Taylor-Thornton) + I (invariant sites) + G (gamma distribution for modeling rate heterogeneity) model was selected (Darriba et al., 2011). The bootstrapping with 1000 replications was performed to test the phylogenetic tree and gaps/missing data treatment was analyzed using all sites. Bootstrap values

less than 50 are not shown.

2.3. Synteny and schematic genomic organization analysis

Syntenic analysis was conducted to provide extra evidence to the orthologies of channel catfish CC chemokine genes, basing on the comparison the neighbouring genes of CC chemokines in channel catfish with zebrafish. BLASTN was conducted using the catfish CC chemokine coding sequences as queries against the catfish genome sequence database (Liu et al., 2016), with a cutoff E-value of $1e^{-10}$. The flanking genes of catfish CC chemokines were predicted from the catfish genomic scaffolds using Fgenesh program Molquest software (Softberry Int.) (Solovyev et al., 2006) and validated by running BLASTP against NCBI non-redundant database. The genomic pattern of CC chemokine genes and their neighboring genes in zebrafish were obtained from NCBI database. For the nomenclatures of channel catfish CC chemokine genes, since no clear orthology to mammalian chemokines could be established for most fish chemokines (Nomiya et al., 2008; Peatman et al., 2006), we named them after zebrafish whenever possible based on orthologues which determined by phylogenetic and syntenic analysis because zebrafish is the model organism and has the most updated nomenclature system (<http://zfin.org/>). When the orthologous relationship was obscure, especially for those tandem duplicated CC chemokine copies originated from one single gene (Peatman et al., 2005; Wang et al., 2005), we named them based on their genomic locations on the chromosome (Nomiya et al., 2008; Peatman et al., 2006).

Schematic genome structure of catfish CC chemokine genes was analyzed in order to visually and systematically clarifying the distribution of CC chemokine genes in catfish genome

and reveal the potential relationship between structure and function. The schematic genomic organization of catfish CC chemokine genes demonstrated the detailed information of catfish CC chemokine genes. Such information was collected and decided from the identifications, annotations of catfish CC chemokine.

2.4.Expression analysis following E.ictaluri and F.columnare infections

The expression profiles of CC chemokines in response to ESC and columnaris infections (*E.ictaluri* and *F.columnare*) were determined with meta-analysis of RNA-Seq datasets. The bacterial challenges were performed as previously described (Li et al., 2012; Sun et al., 2012; Wang et al., 2013) and all experiments were approved by the Institutional Animal Care and Use Committee (IACUC) at Auburn University. On the one hand, the expression of catfish CC chemokines after bacterial infection with time variation were determined. Briefly, marion channel catfish (mean weight 35 g) were maintained in 30 L tanks and acclimatized at a temperature of 28 °C for 1 week before experiment. Bacterial was cultured from a single colony, re-isolated from a symptomatic fish and biochemically confirmed before being inoculated in the shaker incubator overnight. The concentration of bacteria was determined using colony forming unit (CFU) per mL by plating 10 µL of 10-fold serial dilutions onto plates. For ESC infection, fish were challenged in 30 L aquaria with three control and three treatment groups. Aquaria were randomly divided into three sampling time-points: 3 h, 24 h, and 72 h post-infection for both control and treatment groups. During challenge, *E.ictaluri* (MS-S97-773) with a concentration of 4×10^8 CFU/mL was added into the treatment aquaria for 2 h with the water flow turned off. Control fish were exposed to the same routines as treatment fish, but were immersed in sterilized

media alone. At each time-point, 30 fish were collected from the control and treatment aquaria and anesthetized with MS-222 (300 mg/L). The entire intestinal tracts from 10 fish (3 replicates of 10 fish each) were pooled together, flash frozen in liquid nitrogen and store at -80 °C until RNA extraction. Similarly, for columnaris challenge, fish were randomly divided into six 30 L aquaria of which three aquaria were used as control (4 h, 24 h and 48 h) and the other three were used as challenge groups. *F.columnare* (BGFS-27; genomovar II) with a concentration of 3×10^6 CFU/mL were added into the treatment aquaria for 2 h. Gills of 18 fish were collected and randomly divided into three replicate pools (6 fish each) at each sampling time-point. Samples were flash-frozen in liquid nitrogen and stored at -80 °C until RNA extraction.

On the other hand, the expression patterns of catfish CC chemokines in susceptible and resistant fish to ESC were also determined. Catfish (mean weight 35 g) were challenged in 500 L (400 L water) aquaria with control group containing 400 fish and treatment group containing 1200 fish. The MS-S97-773 isolate of *E.ictaluri* bacteria was obtained from a natural outbreak and utilized in the experimental challenge. All the experimental procedures including bacterial isolation, cultivation and infection were same with ESC challenge described above. During 3-5 days after challenge, all dying fish with classical ESC clinic signs were collected as susceptible fish. After two weeks of the challenge, all survival fish were collected as resistant fish and the fish in control group were also collected at that time. 72 fish (24 fish/group each for resistant, susceptible and control) were collected and anesthetized with MS-222 (300 mg/L). Equal amounts of liver tissue from 8 fish (3 replicates of 8 fish each) were pooled together, flash frozen in liquid nitrogen and store at -80 °C until RNA extraction.

Illumina-based RNA-Seq reads were retrieved from our three previous transcriptome datasets

described above: intestine samples from catfish challenged with ESC (3 h, 24 h and 72 h post infection) (SRA accession number SRP009069) (Li et al., 2012) and gill samples from catfish challenged with columnare (4 h, 24 h and 48 h post infection) (SRA accession number SRP012586) (Sun et al., 2012) were used to determine the expression signatures of CC chemokine genes with time variation, meanwhile, liver samples from catfish challenged with ESC (the comparison between resistant and susceptible catfish) (SRA accession number SRP028159) (Wang et al., 2013) was used to compare the expression differences of CC chemokine genes between susceptible and resistant catfish. The trimmed high quality RNA-Seq reads were first mapped to reference containing the whole 26,661 unique catfish cDNA transcripts together with all cDNA sequences of channel catfish CC chemokine genes identified in this study. Mapping parameters were set as 95% or greater sequence identity with a maximum of two mismatches. The number of total mapped reads for each chemokine genes was determined and RPKM was calculated. The expression fold change of each chemokine gene was calculated based on normalized RPKM using proportions-based Kal's test. Chemokine genes with absolute fold change ≥ 2 , total reads number > 5 and p-value < 0.05 were regarded as differentially expressed genes (DEGs) (Fu et al., 2016a; Fu et al., 2016b; Sun et al., 2015).

2.5.Expression analysis following hypoxia stress

For hypoxia challenge, 350 catfish (average body mass 21.73 g) were acclimated in an aquarium [$90 \times 60 \times 30$ cm (L $\times$ W $\times$ H)] for one week with flow-through water. After acclimation, 50 fish were moved to another tank as control group, the remaining fish were exposed to hypoxia stress. Dissolved oxygen levels in the challenge group were reduced from 8.5

mg/L to 1.00 mg/L after bubbling nitrogen gas into the aquarium for 1.5 h. The oxygen level was measured every hour to ensure the oxygen concentration was properly maintained. Starting at 1.5 h when first fish start to lose balance, samples were collected continuously until 45 fish were collected. At 5 h after low oxygen challenge, 45 tolerant fish were randomly collected. 15 fish were euthanized with MS-222 (300 mg/L) and collected for hypoxia intolerant, tolerant and control group. Gill samples were collected and pooled for RNA analysis. The similar meta-analysis was conducted to determine the expression signatures of CC chemokine between hypoxia-tolerant and hypoxia-intolerant fish using our previous Illumina-based RNA-Seq datasets (SRA accession number SRP039612) (Feng, 2013). The differential expression analyses were conducted the same as described above in bacterial infectious tissues (see 2.4).

3. Results

3.1. Identification of CC chemokine genes in channel catfish

In the present study, a total of 64 CC chemokine genes were identified in the channel catfish genome. The characteristics of the 64 CC chemokines including their names, open reading frame lengths, the CC domain positions, and chromosomal locations are all included in Table 1. Their sequences were deposited to GenBank, and their accession numbers are also included in Table 1. The complete coding sequences of the catfish CC chemokines vary between 78 to 195 amino acids (Table 1). Among those 64 CC chemokines, 24 were reported in our previous work (Bao et al., 2006) from channel catfish EST datasets while the rest 40 were identified in this study (Table 1).

3.2. Phylogenetic and syntenic analysis of channel catfish CC chemokine genes

Phylogenetic analysis was conducted to annotate the channel catfish CC chemokine genes using the amino acid sequences of human, mouse, chicken, frog, channel catfish, zebrafish, tongue sole, fugu and medaka (Fig. 1). Overall, phylogenetic analysis supported the annotation of catfish CC chemokines. As shown in Figure 1, most of the channel catfish CC chemokines were clustered into clades with their respective counterparts in zebrafish with strong bootstrap support, with exception of CCL19a, CCL25a and CCL35. Of the 64 catfish CC chemokines, a number of them are highly related duplicates, which included 24 genes for CCL36, nine genes for CCL33, eight genes for CCL39, seven genes for CCL34, four genes for CCL20, three genes for CCL19 and CCL27, and two genes for CCL25 and CCL32 (Fig. 1 and Table 2). Many of these CC chemokines were named previously, but the nomenclature was not standardized, and in addition, their relationship was not accurately reflected in those studies because the lack of the full repertoire of the CC chemokines. Here they were re-named according to phylogenetic or syntenic analysis, and their correct names are included in Table 1.

Though phylogenetic relationships provided strong support for the identities of most catfish CC chemokines, syntenic analyses were necessary to provide additional evidence for annotations and identifications of CC chemokines. On the one hand, the syntenic analysis provided strong evidence for the identification of catfish CCL19a, CCL25a and CCL35 whose phylogenetic relationship were somewhat obscure. As shown in Figure 2 (C, E and F), the syntenic block containing CCL19a, CCL25a and CCL35 were well conserved between zebrafish and catfish. On the other hand, the synteny also supported the annotations of remaining CC chemokines and their

duplicates, especially for those clustered CC chemokines (Peatman et al., 2006). For example, as shown in Figure 2 (B, C, G and H), 13 members of CCL36, 6 members of CCL33 and CCL39, three members of CCL34b, two members of CCL19a and CCL20b were apparently derived from tandem duplications. In addition, the synteny block containing above genes were also well conserved between catfish and zebrafish, with some local differences in gene orientations. Overall, the syntenic analysis validated the annotation of CC chemokine genes (Figure 1 and Table 1).

3.3. Comparative analysis of copy numbers of CC chemokine genes

Copy number of CC chemokine genes was investigated in channel catfish, along with various other vertebrate species. Taken together, the numbers of CC chemokine genes varied greatly among different species, ranging from 7 (chicken and frog) to 64 (channel catfish). The results are summarized in Table 2, and several interesting observations were made: 1) The channel catfish has the largest numbers of CC chemokine genes with 64 genes, followed by zebrafish with 62 genes, while the frog and chickens have the smallest numbers of chemokine genes, each with 7 chemokine genes; 2) In spite of the largest numbers of chemokine genes in teleost fish, their diversity is quite low; they were derived mostly from intrachromosomal tandem duplications. In contrast, the mammalian species have the highest levels of CC chemokine diversity; 3) There were apparently group-specific chemokines: With minor exceptions, 17 chemokines were mammal-specific (CCL2, CCL3, CCL6, CCL7, CCL8, CCL9, CCL11, CCL12, CCL13, CCL14, CCL15, CCL16, CCL18, CCL22, CCL23, CCL24 and CCL26), eight groups of chemokines were fish-specific (CCL32, CCL33, CCL34, CCL35, CCL36, CCL38,

CCL39 and CCL44), two chemokines (CCL1 and CCL17) were present in higher vertebrates starting from chickens to mammals; four chemokines (CCL4, CCL5, CCL21 and CCL28) were in tetrapod species ranging from frog to mammals, but not in teleost fish, and four chemokines (CCL19, CCL20, CCL25 and CCL27) were shared by all the vertebrate species under analysis.

3.4. Genomic clusters of channel catfish CC chemokine genes

In order to systematically determine the distribution of CC chemokine genes in catfish genome, genomic locations of the catfish CC chemokine genes were analyzed. Overall, 46 of 64 identified channel catfish CC chemokines are distributed on 8 different chromosomes (chromosome 5, 7, 8, 10, 16, 20, 22 and 23) (Fig. 3A and Fig. 3B), but the chromosome locations of the other 18 genes have not yet been mapped (Fig. 3C). There are two large CC chemokine gene clusters in channel catfish genome, located on chromosome 7 and 8, containing 18 and 15 genes, respectively (Fig. 3A). Other genes are present in mini-clusters or as single genes on 6 different chromosomes (Fig. 3B).

3.5. Expression of CC chemokine genes following bacterial infection

The expression of 64 catfish CC chemokines was determined after bacterial infections with two pathogens, *E. ictaluri* and *F. columnare* using meta-analysis of RNA-Seq datasets. After ESC infection, seven CC chemokine genes were significantly regulated, of which five (CCL19a.1, CCL25b, CCL33.7, CCL36.13 and CCL36.22) were significantly up-regulated while two (CCL34b.4 and CCL36.10) were significantly down-regulated (Fig. 4). Interestingly, the time of their induced or suppressed expression was quite different, with three genes (CCL36.10,

CCL36.13 and CCL36.22) significantly differentially expressed at 3 hour, two genes (CCL33.7 and CCL34b.4) significantly differentially expressed at 24 hour, and two genes (CCL19a.1 and CCL25b) significantly up-regulated at 72 hour (Fig. 4).

Expression of CC chemokine genes after columnaris infection was shown in Figure 5. Ten of 64 CC chemokine genes (CCL19a.1, CCL19a.2, CCL19b, CCL32a.2, CCL34a.1, CCL34a.2, CCL34a.3, CCL35, CCL36.13 and CCL36.21) were significantly regulated. Most of these CC chemokine genes were induced after columnaris infection, with only two exceptions: CCL34a.1 at 48 h and CCL36.13 at 24 h after infection. Most induced CC chemokine expression was relatively rapid, with significant induced expression observed at 4 h and 24 h after infection (Fig. 5).

3.6. Expression of CC chemokine in resistant and susceptible fish after ESC infection

Expression of CC chemokines in the liver of resistant fish and susceptible fish were determined after ESC infection and the results were shown in Figure 6. Four CC chemokine genes (CCL19b, CCL25b, CCL34a.2 and CCL34b.4) were mildly induced after ESC infection in resistant fish. For the susceptible fish, 13 CC chemokines (CCL19a.1, CCL19a.2, CCL19b, CCL20a.3, CCL25b, CCL34a.1, CCL34a.2, CCL34b.1, CCL34b.4, CCL36.7, CCL36.8, CCL26.17 and CCL36.22) were significantly induced while one gene (CCL36.10) were mildly suppressed after ESC infection (Fig. 6). The levels of induced expression were much more dramatic in susceptible fish than in resistant fish, especially with CCL19a.1, CCL19a.2 and CCL20.3 (Fig. 6). Not only the level of disease-induced differential expression was dramatic, the absolute expression levels of certain CC chemokine genes were very high. For instance,

CCL19a.2 was induced over 200 fold after ESC infection in susceptible fish, and the RPKM of this gene was as high as 9043.56 (Fig. 6).

3.7. Expression profile of CC chemokine genes following hypoxia stress

To get insights to the potential involvement of CC chemokine under abiotic stress, the expression profiles of catfish CC chemokines under hypoxia conditions were determined for both low-oxygen tolerant and intolerant fish. Interestingly, all significantly differentially expressed CC chemokines (CCL27a, CCL34a.1 and CCL34a.2) after hypoxia challenge in intolerant fish were up-regulated, while the majority of significantly differentially expressed CC chemokines (CCL34a.1, CCL34a.2 and CCL36.7) after hypoxia challenge in tolerant fish were down-regulated, with exception of CCL36.4 that is up-regulated (Fig. 7).

4. Discussion

Chemokines are a family of structurally related chemotactic cytokines, serving as key regulators for cell migration and activation, especially under inflammatory conditions (Nomiyama et al., 2008; Zlotnik and Yoshie, 2000). CC chemokines constitute the largest subfamily of chemokines with 28 CC chemokines identified from mammalian species. Before this study, the largest number of CC chemokines found in a single species is 62 from Zebrafish. In the present study, we identified a complete repertoire of 64 CC chemokine genes in the channel catfish, making the channel catfish as a species with the largest number of CC chemokine genes.

This study provided the systematic annotation of the catfish CC chemokines. Although

previous studies were devoted to the analysis of the catfish CC chemokines (Bao et al., 2006; He et al., 2004; Peatman et al., 2006; Peatman and Liu, 2007), the proper annotation of these genes were not in place. Here in this study, a series of phylogenetic and syntenic analyses allowed us to annotate the 64 channel catfish CC chemokines into 11 CC chemokine groups. A total of 12 CC chemokine groups were identified from fish, of which zebrafish and channel catfish shared 11 groups, with no genes found from the group of CCL38 in channel catfish genome, where six genes were present in the zebrafish genome.

Analysis of the copy numbers of CC chemokines in teleost fish and tetrapods provided insights to the evolution of CC chemokine subfamily. Firstly, the chemokine system was proposed to appear about 650 million years ago, at the emergence of vertebrates, since no putative chemokine or chemokine receptors have been identified in any analyzed invertebrate species (DeVries et al., 2006). In addition, the present study endorses and extends the expansion of CC chemokines in teleost fish (Alejo and Tafalla, 2011; Nomiya et al., 2008). 64 CC chemokines were identified in catfish in the present study (Table 2), compared to seven found in amphibians (Frog) and birds (chicken), 19 found in cow, 23 found in mouse and 24 found in human. Similar to catfish CXC chemokines (Fu et al., 2016b), there are two major factors leading to the expansion of CC chemokines in teleost fish, the duplications and the existence of fish-specific chemokines. On the one hand, duplications for some genes (CCL19, CCL20, CCL25, CCL27, CCL32, CCL33, CCL34, CCL36 and CCL39) were only found in teleosts, through both whole genome duplication and lineage-specific tandem gene duplications. On the other hand, the existence of fish-specific CC chemokines (CCL32, CCL33, CCL34, CCL35, CCL36, CCL38, CCL39 and CCL44) were also observed in the present study, which have been

reported in several fish species such as zebrafish and rainbow trout (Alejo and Tafalla, 2011; Nomiya et al., 2008; Wiens et al., 2006). Moreover, CCL19 and CCL20 were found in all analyzed species, which might suggest that CCL19 and CCL20 are the two most evolutionarily conserved members in CC chemokines, similar to CXCL12 and CXCL14 in CXC chemokines (Fu et al., 2016b; Zlotnik and Yoshie, 2012; Zlotnik et al., 2006). Last but not least, a set of 16 CC chemokines including CCL2, CCL3, CCL6, CCL7, CCL8, CCL9, CCL12, CCL13, CCL14, CCL15, CCL16, CCL18, CCL22, CCL23, CCL24 and CCL26 were only harbored by mammalian species and were lacking in other species (Table 2). Those mammal-specific CC chemokines are believed to have derived by gene duplication after the divergence of the bird and mammalian ancestral lineages (Wang et al., 2005).

Phylogenetic and syntenic analyses indicated that the mammalian species have the highest levels of CC chemokine diversity with a total of 28 distinct CC chemokine gene groups. In contrast, teleost fish have the lowest levels of CC chemokine diversity even though their numbers of the CC chemokine genes were the largest. This was because the largest numbers of fish CC chemokines were derived from intrachromosomal tandem duplications. The paralogues shared a high level of sequence identities, while the mammalian chemokines may have been independently derived in evolution, sharing low sequence identities among them. Of all the CC chemokines, eight chemokine gene clusters were teleost-specific; 17 were mammal-specific; two were shared by higher vertebrates and four were shared in tetrapods but not present in teleost fish (Table 2). These results suggested that gene duplication may have been a predominant evolutionary mechanism for fish to adapt to their environment. In contrast, emergence of new genes may have been more popular in mammals.

The schematic genome organization of catfish CC chemokines were analyzed not only to determine the distribution of CC chemokines in catfish genome but also to explore the potential relationship between clustered location and functions of catfish CC chemokines. As shown in Figure 3, there are two large CC chemokine clusters located on chromosome 7 and 8 of catfish genome, containing 18 and 15 genes, respectively. Such large CC chemokines clusters were also observed in some model species and the largest clusters are found on human chromosome 17, mouse chromosome 11, chicken chromosome 19 and zebrafish chromosome 25 (Nomiya et al., 2008; Wang et al., 2005; Zlotnik et al., 2006). However, in cases in mammals, those chemokines were annotated as independent chemokines although they may be evolutionarily related, perhaps because of their relatively high levels of sequence divergence. Of the 64 CC chemokines from channel catfish, the chromosomal locations of 46 chemokines were determined, leaving those for 18 CC chemokines unknown at present. Future studies of their genomic location through genetic linkage mapping should be conducted, because genomic organization, their expression, and functions may be related. Previous studies indicated that there are correlations between genomic architecture and the inducibility of their expression, with inflammatory CC chemokines constituting a large cluster (Peatman et al., 2006), which was supported by the present study. As shown in Figure 4, 5, 6 and 7, more than half of the differentially expressed genes after bacterial infection and hypoxia stress are located on the two largest CC chemokine clusters, suggesting that the CC chemokines with the same evolutionary origin still share a similar regulatory framework, allowing their coordinated expression.

It has been well studied that CC chemokines not only exert vital roles in development and homeostasis but also in inflammatory responses in humans (Raman et al., 2011; Zlotnik and

Yoshie, 2000, 2012). However, in fish species, the involvements of CC chemokines in immune responses were only sporadically reported. Therefore, we examined the expression profiles of 64 CC chemokines following the infection of *E. ictaluri* and *F. columnare* in channel catfish. Interestingly, the regulation trend of catfish CC chemokines were quite similar to CXC chemokines (Fu et al., 2016b). In general, CC chemokines were induced after bacterial infections, with exception of CCL34b.4 and CCL36.10 down-regulated after ESC infection, and CCL34a.1 down-regulated at 48 h after columnaris infection (Fig. 4 and 5). In addition, CCL19a.1 was induced after ESC infection, and CCL19a.1, CCL19a.2 and CCL19b were up-regulated after columnaris infection both bacterial infections, which might indicate the significant role that CCL19 play in fish immune activities. Moreover, fish-specific CC chemokines might play vital roles in fish immune responses, as several copies of CCL33, CCL34 and CCL 36 were observed to be involved in two bacterial immune responses and some of them showed drastic expression patterns (Fig. 4 and 5). Although the functions of fish-specific CC chemokines in fish immunity are currently poorly understood (Chen et al., 2013), the expression patterns should be useful for future studies concerning functional analysis of CC chemokines.

Comparisons of CC chemokine expression between resistant and susceptible fish after ESC infection suggesting that expression of the majority of CC chemokines was correlated with disease susceptibility. As shown in Figure 6, all the differentially expressed genes in both susceptible and resistant group were induced after ESC infection except the mild down-regulation of CCL36.10 in resistant group, and the expression of all induced CC chemokines was much more drastic in susceptible fish than in resistant fish after ESC infection. Among these, CCL19a.2 was expressed at extremely high level, over two hundreds fold in susceptible fish.

Besides, the expression levels of CCL19a.1, CCL20a.3 and CCL34b.1 were also relatively high, more than 20 times in susceptible fish. Thus, the expression of these genes can be used as an indicator of susceptibility after ESC infection. Interestingly, CCL19 not only showed drastic up-regulation in susceptible fish after ESC infection, but also participated in the immune responses after ESC and columnaris infection, as discussed above, suggesting the important role it play in immune activities. As one of the homeostatic chemokines, CCL19 and its receptor (CCR7) are expressed constitutively within lymphoid tissues, regulating lymphoid and dendritic cell (DC) homing (Forster et al., 1999; Muller et al., 2003). Previous studies showed that CCL19 enhanced T cell proliferation by inducing maturation of DCs, leading to the up-regulation of costimulatory molecules and the production of proinflammatory cytokines. Moreover, CCL19 programmed DCs for the induction of T helper type (Th) 1 rather than Th2 responses (Marsland et al., 2005). While we do not know the exact mechanism of how CCL19 participate in the immune activities in teleost fish, a similar mechanism may be conserved in fish, as it does in humans (Marsland et al., 2005).

The expression pattern of CC chemokines was similar to that of CXC chemokines (Fu et al., 2016b), with the majority of CC chemokines being expressed at high levels in hypoxia-intolerant fish than in tolerant fish. The only exception is CCL36.4 that is significantly up-regulated in hypoxia-tolerant fish (Fig. 7). This suggest that high expression of CCL36.4 might be correlated with increased tolerance against low oxygen (Sun et al., 2015). On the contrary, the significantly induced expression of most CC chemokine genes in intolerant fish may indicate that high levels of induced CC chemokines are indicative of sensitivity of fish for stress tolerance. This parallels with the disease susceptibility. In other words, a set of chemokines were expressed significantly

higher in susceptible fish and in intolerant fish than in resistant and tolerant fish (Fu et al., 2016b). This correlation is very interesting because stresses including hypoxia stress tended to induce disease incidents (Geng et al., 2014). There is apparently a common linkage between bacterial infection and low oxygen stresses, and chemokines could be a common set of molecular signals amplifying the adverse effect of infections or stresses. Future studies are required to provide detailed molecular mechanisms of stress responses and their effect on disease susceptibility.

5. Conclusion

In this study, the entire CC chemokinome, with a total of 64 CC chemokine genes, was systematically identified and characterized in the channel catfish genome. Phylogenetic and syntenic analysis allowed annotations of these genes, and provided insights into their evolutionary dynamics. A total of 8 fish-specific CC chemokine clusters were identified, which, along with commonly shared 4 CC chemokines, expanded through tandem duplications, making teleost fish have the largest numbers of CC chemokine genes. Most differentially expressed CC chemokine genes were up-regulated after disease infection or hypoxia stress, with significant differences in resistant and susceptible fish. Therefore, expression of certain CC chemokines, such as CCL19a.1, can be used as indicators of disease and stress susceptibility.

Acknowledgements

This project was supported by USDA National Institute of Food and Agriculture through a

grant from Animal Disease Program (2015-67015-22975), and from Agriculture and Food Research Initiative Animal Genomics, Genetics and Breeding Program (2015-67015-22907), and partially supported by USDA Aquaculture Research Program Competitive Grant (2014-70007-22395). We are grateful to the availability of RNA-Seq datasets produced by our group. Qiang Fu was supported by a scholarship from the China Scholarship Council (CSC).

References

- Alejo, A., Tafalla, C., 2011. Chemokines in teleost fish species. *Developmental and comparative immunology* 35, 1215-1222.
- Arenberg, D.A., Polverini, P.J., Kunkel, S.L., Shanafelt, A., Hesselgesser, J., Horuk, R., Strieter, R.M., 1997. The role of CXC chemokines in the regulation of angiogenesis in non-small cell lung cancer. *J Leukoc Biol* 62, 554-562.
- Bacon, K., Baggiolini, M., Broxmeyer, H., Horuk, R., Lindley, I., Mantovani, A., Maysushima, K., Murphy, P., Nomiyama, H., Oppenheim, J., Rot, A., Schall, T., Tsang, M., Thorpe, R., Van Damme, J., Wadhwa, M., Yoshie, O., Zlotnik, A., Zoon, K., Nomenclature, I.W.S.o.C., 2002. Chemokine/chemokine receptor nomenclature. *Journal of interferon & cytokine research : the official journal of the International Society for Interferon and Cytokine Research* 22, 1067-1068.
- Bao, B., Peatman, E., Peng, X., Baoprasertkul, P., Wang, G., Liu, Z., 2006. Characterization of 23 CC chemokine genes and analysis of their expression in channel catfish (*Ictalurus punctatus*). *Developmental and comparative immunology* 30, 783-796.
- Baoprasertkul, P., He, C., Peatman, E., Zhang, S., Li, P., Liu, Z., 2005. Constitutive expression of three novel catfish CXC chemokines: homeostatic chemokines in teleost fish. *Mol Immunol* 42, 1355-1366.
- Belmadani, A., Tran, P.B., Ren, D., Miller, R.J., 2006. Chemokines regulate the migration of neural progenitors to sites of neuroinflammation. *J Neurosci* 26, 3182-3191.
- Bosco, M.C., Puppo, M., Santangelo, C., Anfosso, L., Pfeffer, U., Fardin, P., Battaglia, F., Varesio, L., 2006. Hypoxia modifies the transcriptome of primary human monocytes: modulation of novel immune-related genes and identification of CC-chemokine ligand 20 as a new hypoxia-inducible gene. *J Immunol* 177, 1941-1955.
- Chen, J., Xu, Q., Wang, T., Collet, B., Corripio-Miyar, Y., Bird, S., Xie, P., Nie, P., Secombes, C.J., Zou, J., 2013. Phylogenetic analysis of vertebrate CXC chemokines reveals novel lineage specific groups in teleost fish. *Dev Comp Immunol* 41, 137-152.
- Darriba, D., Taboada, G.L., Doallo, R., Posada, D., 2011. ProtTest 3: fast selection of best-fit models of protein evolution. *Bioinformatics* 27, 1164-1165.
- David, N.B., Sapede, D., Saint-Etienne, L., Thisse, C., Thisse, B., Dambly-Chaudiere, C., Rosa, F.M., Ghysen, A., 2002. Molecular basis of cell migration in the fish lateral line: role of the chemokine receptor CXCR4 and of its ligand, SDF1. *Proc Natl Acad Sci U S A* 99, 16297-16302.
- DeVries, M.E., Kelvin, A.A., Xu, L., Ran, L., Robinson, J., Kelvin, D.J., 2006. Defining the origins and evolution of the chemokine/chemokine receptor system. *J Immunol* 176, 401-415.
- Doitsidou, M., Reichman-Fried, M., Stebler, J., Kopranner, M., Dorries, J., Meyer, D., Esguerra, C.V., Leung, T., Raz, E., 2002. Guidance of primordial germ cell migration by the chemokine SDF-1. *Cell* 111, 647-659.
- Edgar, R.C., 2004. MUSCLE: a multiple sequence alignment method with reduced time and space complexity. *BMC bioinformatics* 5, 113.
- Esche, C., Stellato, C., Beck, L.A., 2005. Chemokines: key players in innate and adaptive immunity. *J Invest Dermatol* 125, 615-628.
- Feng, J., 2013. Transcriptome and expression profiling of response to acute hypoxia in the gill of adult catfish, Plant and Animal Genome XXI Conference. Plant and Animal Genome.
- Forster, R., Schubel, A., Breitfeld, D., Kremmer, E., Renner-Muller, I., Wolf, E., Lipp, M., 1999. CCR7 coordinates the primary immune response by establishing functional microenvironments in secondary lymphoid organs. *Cell* 99, 23-33.
- Fu, Q., Li, Y., Yang, Y., Li, C., Yao, J., Zeng, Q., Qin, Z., Liu, S., Li, D., Liu, Z., 2016a. Septin genes in channel catfish (*Ictalurus punctatus*) and their involvement in disease defense responses. *Fish Shellfish Immunol* 49, 110-121.
- Fu, Q., Zeng, Q., Li, Y., Yang, Y., Li, C., Liu, S., Zhou, T., Li, N., Yao, J., Jiang, C., Li, D., Liu, Z., 2016b. The chemokine superfamily in channel catfish: I. CXC subfamily and their involvement in disease defense and hypoxia responses. *Fish Shellfish Immunol*.
- Geng, X., Feng, J., Liu, S., Wang, Y., Arias, C., Liu, Z., 2014. Transcriptional regulation of hypoxia inducible factors alpha (HIF-alpha) and their inhibiting factor (FIH-1) of channel catfish (*Ictalurus punctatus*) under hypoxia. *Comp Biochem Physiol B Biochem Mol Biol* 169, 38-50.
- Gordon, R.J., McGregor, A.L., Connor, B., 2009. Chemokines direct neural progenitor cell migration following striatal cell loss. *Mol Cell Neurosci* 41, 219-232.
- He, C., Peatman, E., Baoprasertkul, P., Kucuktas, H., Liu, Z., 2004. Multiple CC chemokines in channel catfish and blue catfish as revealed by analysis of expressed sequence tags. *Immunogenetics* 56,

379-387.

Keane, M.P., Arenberg, D.A., Moore, B.B., Addison, C.L., Strieter, R.M., 1998. CXC chemokines and angiogenesis/angiostasis. *Proc Assoc Am Physicians* 110, 288-296.

Knaut, H., Werz, C., Geisler, R., Nusslein-Volhard, C., Tübingen Screen, C., 2003. A zebrafish homologue of the chemokine receptor Cxcr4 is a germ-cell guidance receptor. *Nature* 421, 279-282.

Laing, K.J., Secombes, C.J., 2004. Trout CC chemokines: comparison of their sequences and expression patterns. *Mol Immunol* 41, 793-808.

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.

Liu, S., Zhang, Y., Zhou, Z., Waldbieser, G., Sun, F., Lu, J., Zhang, J., Jiang, Y., Zhang, H., Wang, X., Rajendran, K.V., 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, 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, R., Li, Z., Sonstegard, T., Feng, J., Danzmann, R.G., Schroeder, S., Scheffler, B., Duke, M.V., Ballard, L., Kucuktas, H., Kaltenboeck, L., Liu, H., Armbruster, J., Xie, Y., Kirby, M.L., Tian, Y., Flanagan, M.E., Mu, W., Waldbieser, G.C., 2016. The channel catfish genome sequence provides insights into the evolution of scale formation in teleosts. *Nature communications* 7, 11757.

Marsland, B.J., Battig, P., Bauer, M., Ruedl, C., Lassing, U., Beerli, R.R., Dietmeier, K., Ivanova, L., Pfister, T., Vogt, L., Nakano, H., Nembrini, C., Saudan, P., Kopf, M., Bachmann, M.F., 2005. CCL19 and CCL21 induce a potent proinflammatory differentiation program in licensed dendritic cells. *Immunity* 22, 493-505.

Muller, G., Hopken, U.E., Lipp, M., 2003. The impact of CCR7 and CXCR5 on lymphoid organ development and systemic immunity. *Immunol Rev* 195, 117-135.

Naylor, R.L., Goldburg, R.J., Primavera, J.H., Kautsky, N., Beveridge, M.C., Clay, J., Folke, C., Lubchenco, J., Mooney, H., Troell, M., 2000. Effect of aquaculture on world fish supplies. *Nature* 405, 1017-1024.

Nomiyama, H., Hieshima, K., Osada, N., Kato-Unoki, Y., Otsuka-Ono, K., Takegawa, S., Izawa, T., Yoshizawa, A., Kikuchi, Y., Tanase, S., Miura, R., Kusuda, J., Nakao, M., Yoshie, O., 2008. Extensive expansion and diversification of the chemokine gene family in zebrafish: identification of a novel chemokine subfamily CX. *BMC genomics* 9, 222.

Peatman, E., Bao, B., Baoprasertkul, P., Liu, Z., 2005. In silico identification and expression analysis of 12 novel CC chemokines in catfish. *Immunogenetics* 57, 409-419.

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

Peatman, E., Liu, Z., 2007. Evolution of CC chemokines in teleost fish: a case study in gene duplication and implications for immune diversity. *Immunogenetics* 59, 613-623.

Raman, D., Sobolik-Delmaire, T., Richmond, A., 2011. Chemokines in health and disease. *Exp Cell Res* 317, 575-589.

Schioppa, T., Uranchimeg, B., Saccani, A., Biswas, S.K., Doni, A., Rapisarda, A., Bernasconi, S., Saccani, S., Nebuloni, M., Vago, L., Mantovani, A., Melillo, G., Sica, A., 2003. Regulation of the chemokine receptor CXCR4 by hypoxia. *J Exp Med* 198, 1391-1402.

Solovyev, V., Kosarev, P., Seledsov, I., Vorobyev, D., 2006. Automatic annotation of eukaryotic genes, pseudogenes and promoters. *Genome biology* 7 Suppl 1, S10 11-12.

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

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

Tamura, K., Stecher, G., Peterson, D., Filipowski, A., Kumar, S., 2013. MEGA6: Molecular Evolutionary Genetics Analysis version 6.0. *Molecular biology and evolution* 30, 2725-2729.

Wang, J., Adelson, D.L., Yilmaz, A., Sze, S.H., Jin, Y., Zhu, J.J., 2005. Genomic organization,

annotation, and ligand-receptor inferences of chicken chemokines and chemokine receptor genes based on comparative genomics. *BMC Genomics* 6, 45.

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, 929.

Wiens, G.D., Glenney, G.W., Lapatra, S.E., Welch, T.J., 2006. Identification of novel rainbow trout (*Onchorynchus mykiss*) chemokines, CXCD1 and CXCD2: mRNA expression after *Yersinia ruckeri* vaccination and challenge. *Immunogenetics* 58, 308-323.

Zeng, Q., Liu, S., Yao, J., Zhang, Y., Yuan, Z., Jiang, C., Chen, A., Fu, Q., Su, B., Dunham, R., Liu, Z., 2016. Transcriptome Display During Testicular Differentiation of Channel Catfish (*Ictalurus punctatus*) as Revealed by RNA-Seq Analysis. *Biology of reproduction* 95, 19.

Zlotnik, A., Yoshie, O., 2000. Chemokines: a new classification system and their role in immunity. *Immunity* 12, 121-127.

Zlotnik, A., Yoshie, O., 2012. The chemokine superfamily revisited. *Immunity* 36, 705-716.

Zlotnik, A., Yoshie, O., Nomiya, H., 2006. The chemokine and chemokine receptor superfamilies and their molecular evolution. *Genome Biol* 7, 243.

Figure legends:

Fig. 1. Phylogenetic analysis of 64 CC chemokine genes using amino acid sequences. The unrooted phylogenetic tree was conducted using maximum likelihood algorithm under the JTT + I + G model using MEGA6 software. The red clusters indicate fish CC chemokine genes and black dots indicate catfish CC chemokine genes. Bootstrap (1,000 replications) support values appear on the branches. Bootstrap values less than 50 are not shown. The Latin name and accession numbers were provided in Appendix: Supplementary File S1 (Table S1).

Fig. 2. Syntenic analysis of channel catfish chemokine genes with zebrafish. (A) CCL27a; (B) CCL36; (C) CCL33, CCL35 and CCL39; (D) CCL25b and CCL44; (E) CCL25a; (F) CCL19a.1 and CCL19a.2; (G) CCL34b.1, CCL34b.2 and CCL34b.4; (H) CCL20b.1 and CCL20b.2; (I) CCL20a.3. These syntenies were generated with the information obtained from the NCBI. Full gene names were provided in Appendix: Supplementary File S1 (Table S2).

Fig. 3. Schematic genomic organization of catfish CC chemokine subfamily genes. Arrowheads indicate CC chemokine genes and their transcriptional orientation. The numbers under the chemokine genes indicated the numbers of coding exons.

Fig. 4. Expression of CC chemokine genes in intestine after ESC infection at 3 h, 24 h and 72h. The y-axis indicated fold change of CC chemokine genes and the x-Axis provided the names of the genes. Only the significantly expressed genes ($|\text{fold change}| \geq 2$, with the p-value < 0.05) were presented.

Fig. 5. Expression of CC chemokine genes in gill after *F. columnare* infection at 4 h, 24 h and 48 h. The y-axis indicated fold change of CC chemokine genes and the x-Axis provided the names of the genes. Only the significantly expressed genes ($|\text{fold change}| \geq 2$, with the p-value < 0.05) were presented.

Fig. 6. Comparison of the expression of CC chemokine between resistant and susceptible catfish in liver

after ESC infection. The y-axis indicated fold change of CC chemokine genes and the x-Axis provided the names of the genes. Only the significantly expressed genes ($|\text{fold change}| \geq 2$, with the $p\text{-value} < 0.05$) were presented.

Fig. 7. Comparison of CC chemokine expression between tolerant and intolerant catfish in gill after hypoxia stress. The y-axis indicated fold change of CC chemokine genes and the x-Axis provided the names of the genes. Only the significantly expressed genes ($|\text{fold change}| \geq 2$, with the $p\text{-value} < 0.05$) were presented.

Table 1. Identification of CC chemokine genes in channel catfish genome.

Chemokine genes	CDS (bp)	CC domain	Chr	Ori	location	Resource	Accession no.	Previous name	Reference
CCL19a.1	110	32-92	chr22	+	4478046-4479346	G, T, E	XP_017307008	SCYA106	(Bao et al., 2006)
CCL19a.2	112	37-101	chr22	-	4473890-4474878	G, T, E	XP_017307006	SCYA109	(Bao et al., 2006)
CCL19b	107	32-94	chr16	-	2920205-2920809	G, T, E	JT218590	SCYA113	(Bao et al., 2006a)
CCL20a.1	107	20-80	chr7	+	7204965-7206596	G, T, E	NP_001187124	SCYA112	(Bao et al., 2006a)
CCL20a.3	104	26-86	US	+	84624-85406	G, T	XP_017317373		This study
CCL20b.1	100	26-86	chr23	+	9220057-9221676	G, T	XP_017308519		This study
CCL20b.2	94	26-86	chr23	+	9224596-9226275	G, T	XP_017308520		This study
CCL25a	106	24-86	chr20	-	15062000-15063332	G, T	XP_017351899		This study
CCL25b	86	24-85	chr10	-	21393065-21393806	G, T, E	XP_017332798	SCYA103	(Bao et al., 2006a)
CCL27a	97	25-87	chr5	+	16717616-16720083	G, T	XP_017322714		This study
CCL27b.1	88	21-78	chr16	+	2577679-25777372	G, T	JT340209		This study
CCL27b.2	92	27-73	NA		NA	T, E	ABA54971	SCYA126	(Bao et al., 2006a)
CCL32a.1	107	29-88	chr7	+	7237486-7238101	G, T, E	XP_017328673	SCYA116	(Bao et al., 2006a)

CCL32a.2	195	27-86	chr7	+	7247805-7250321	G, T, E	XP_017326742	SCYA118	(Bao et al., 2006a)
CCL33.1	81	30-77	chr8	+	1228702-1230367	G, T, E	JT348749	SCYA121	(Bao et al., 2006a)
CCL33.2	78	30-92	chr8	+	1236059-1236980	G, T, E	XP_017330655	SCYA117	(Bao et al., 2006a)
CCL33.3	81	30-76	chr8	-	22113983-2211668 5	G, T, E	XP_017330182	SCYA111	(Bao et al., 2006a)
CCL33.4	85	29-85	chr8	-	22128055-2213350 4	G, T	XP_017330183		This study
CCL33.5	100	45-97	chr8	-	22144181-2214513 7	G, T	XP_017329053		This study
CCL33.6	86	31-83	chr8	-	22152385-2215317 7	G, T	XP_017330184		This study
CCL33.7	78	27-72	chr8	-	22160432-2216147 6	G, T	XP_017328905		This study
CCL33.8	78	29-76	chr8	-	22190679-2219240 0	G, T	JT396455		This study
CCL33.9	80	31-75	NA		NA	T, E	ABA54969	SCYA122	(Bao et al., 2006a)
CCL34a.1	114	25-106	chr7	-	7305183-7306106	G, T, E	ABA54949	SCYA101	(Bao et al., 2006a)
CCL34a.2	92	28-82	chr7	-	7310154-7310890	G, T	x		This study
CCL34a.3	96	25-94	NA		NA	T, E	ABA54954	SCYA107	(Bao et al., 2006a)
CCL34b.1	111	30-87	chr23	+	1925979-1927276	G, T, E	XP_017308298	SCYA114	(Bao et al., 2006a)
CCL34b.2	122	24-81	chr23	-	1893222-1893997	G, T	XP_017308139		This study
CCL34b.3	118	24-81	NA		NA	T, E	ABA54966	SCYA119	(Bao et al., 2006a)
CCL34b.4	100	36-95	chr23	+	1930203-1931344	G, T, E	NP_00118712 3	SCYA102	(Bao et al., 2006a)
CCL35	99	31-89	chr8	+	22008624-2201053 9	G, T, E	XP_017328933	SCYA115	(Bao et al., 2006a)
CCL36.1	95	28-86	chr7	-	29226673-2922784 3	G, T	XP_017326661		This study
CCL36.2	97	28-86	chr7	-	29252932-2925379 0	G, T	XP_017328326		This study
CCL36.3	99	28-88	chr7	-	29269022-2928597 5	G, T	XP_017328322		This study
CCL36.4	90	28-86	chr7	-	29299665-2930215 7	G, T	XP_017326596		This study
CCL36.5	91	28-86	chr7	-	29310357-2931125 5	G, T	XP_017326595		This study

CCL36.6	95	31-89	chr7	+	29333865-2933515 6	G, T	XP_017328678		This study
CCL36.7	101	28-86	chr7	+	29347053-2935726 3	G, T	XP_017328677		This study
CCL36.8	101	28-86	chr7	-	29400829-2940175 7	G, T	XP_017326670		This study
CCL36.9	95	28-86	chr7	-	29584265-2958529 7	G, T	XP_017326447		This study
CCL36.10	95	28-86	chr7	+	29584520-2958529 7	G, T	NP_00118709 7		This study
CCL36.11	101	28-86	chr7	+	29737229-2973864 8	G, T	x		This study
CCL36.12	95	28-86	chr7	-	30002283-3000327 2	G, T	x		This study
CCL36.13	93	26-84	chr7	-	30021467-3002192 7	G, T	x		This study
CCL36.14	96	23-81	NA		NA	T	JT471639		This study
CCL36.15	123	51-109	NA		NA	T	x		This study
CCL36.16	69*	1-49	NA		NA	T	x		This study
CCL36.17	101	28-86	US	-	1502-2198	G, T	XP_017317951		This study
CCL36.18	76*	1-61	NA		NA	T	x		This study
CCL36.19	95	31-89	US	-	1232-2537	G, T	XP_017318019		This study
CCL36.20	95	28-86	NA		NA	T, E	AAT52138	SCYA104	(Bao et al., 2006a)
CCL36.21	79	9-67	NA		NA	T, E	AAT52139	SCYA105	(Bao et al., 2006a)
CCL36.22	101	28-86	NA		NA	T, E	ABA54955	SCYA108	(Bao et al., 2006a)
CCL36.23	91	28-86	NA		NA	T, E	ABA54967	SCYA120	(Bao et al., 2006a)
CCL36.24	95	28-86	NA		NA	T, E	ABA54970	SCYA124	(Bao et al., 2006a)
CCL39.1	132	62-120	chr8	+	22018474-2201927 8	G, T	XP_017330604		This study
CCL39.2	129	62-122	chr8	+	22036030-2203681 1	G, T	XP_017330611		This study
CCL39.3	99	29-87	chr8	+	22044290-2204496 9	G, T	XP_017330612		This study
CCL39.4	99	29-87	chr8	+	22050473-2205112 4	G, T	XP_017330613		This study
CCL39.5	99	29-87	chr8	+	22055069-2205574 4	G, T	XP_017330606		This study

CCL39.6	99	29-87	chr8	+	22063555-2206508 7	G, T	XP_017330614	This study
CCL39.7	155	86-144	US	-	841-1665	G, T	XP_017318153	This study
CCL39.8	99	29-87	NA		NA	T, E	ABA54957	SCYA110 (Bao et al., 2006a)
CCL44	112	28-87	chr10	+	21352284-2135568 0	G, T	XP_017334100	This study

Asterisk (*) in the CDS column indicated partial coding sequences. US in the Chr column indicated those CC chemokines were located on different unplaced scaffolds. G, T, E in the Resource column indicated those CC chemokine genes were identified from channel catfish genome, transcriptome and EST databases, respectively. NA indicated the information were not available.

Table 2. Copy numbers of CC chemokine genes in several representative vertebrate genomes. Yellow shading indicated the mammal-specific CC chemokine genes. Gray shading indicated the CC chemokine genes harbored by mammals and chickens. Green shading indicated the CC chemokine genes harbored by tetrapod species. Blue shading indicated the CC chemokines harbored by all species. Red shading indicated the teleost-specific CC chemokines.

Genes	Catfish	Zebrafish	Frog	Chicken	Cow	Mouse	Human
CCL2	0	0	0	0	1	1	1
CCL3	0	0	0	0	1	1	1
CCL6	0	0	0	0	0	1	0
CCL7	0	0	0	0	0	1	1
CCL8	0	0	0	0	1	1	1
CCL9	0	0	0	0	0	1	0
CCL11	0	1	0	0	1	1	1
CCL12	0	0	0	0	0	1	0
CCL13	0	0	0	0	0	0	1
CCL14	0	0	0	0	1	0	1
CCL15	0	0	0	0	0	0	1
CCL16	0	0	0	0	1	0	1
CCL18	0	0	0	0	0	0	1
CCL22	0	0	0	0	1	1	1
CCL23	0	0	0	0	0	0	1
CCL24	0	0	0	0	1	1	1
CCL26	0	0	0	0	1	1	1
CCL1	0	0	0	1	1	1	1
CCL17	0	0	0	1	1	1	1
CCL4	0	0	1	1	1	1	1
CCL5	0	0	1	1	1	1	1
CCL21	0	0	1	0	1	3	1
CCL28	0	0	1	1	1	1	1
CCL19	3	3	1	1	1	1	1
CCL20	4	5	1	1	1	1	1
CCL25	2	2	0	0	1	1	1
CCL27	3	2	1	0	1	1	1

CCL32	2	8	0	0	0	0	0
CCL33	9	3	0	0	0	0	0
CCL34	7	17	0	0	0	0	0
CCL35	1	2	0	0	0	0	0
CCL36	24	2	0	0	0	0	0
CCL38	0	6	0	0	0	0	0
CCL39	8	10	0	0	0	0	0
CCL44	1	1	0	0	0	0	0
total	64	62	7	7	19	23	24

Fig.1

Fig. 2

Fig. 4

Fig.5

Fig. 6

Fig. 7