BIOINFORMATIC ANALYSIS OF CHICKEN CHEMOKINES, CHEMOKINE RECEPTORS, AND TOLL-LIKE RECEPTOR 21

A Thesis

by

JIXIN WANG

Submitted to the Office of Graduate Studies of
Texas A&M University
in partial fulfillment of the requirements for the degree of

MASTER OF SCIENCE

August 2006

Major Subject: Poultry Science

BIOINFORMATIC ANALYSIS OF CHICKEN CHEMOKINES, CHEMOKINE RECEPTORS, AND TOLL-LIKE RECEPTOR 21

A Thesis

by

JIXIN WANG

Submitted to the Office of Graduate Studies of
Texas A&M University
in partial fulfillment of the requirements for the degree of

MASTER OF SCIENCE

Approved by:

Co-Chairs of Committee,	James J. Zhu
	Luc R. Berghman
Committee Member,	David L. Adelson
Head of Department,	Alan R. Sams

August 2006

Major Subject: Poultry Science

ABSTRACT

Bioinformatic Analysis of Chicken Chemokines, Chemokine Receptors, and Toll-Like Receptor 21. (August 2006)

Jixin Wang, B.S., Tarim University of Agriculture and Reclamation;

M.S., South China Agricultural University

Co-Chairs of Advisory Committee: Dr. James J. Zhu
Dr. Luc R. Berghman

Chemokines triggered by Toll-like receptors (TLRs) are small chemoattractant proteins, which mainly regulate leukocyte trafficking in inflammatory reactions via interaction with G protein-coupled receptors. Forty-two chemokines and 19 cognate receptors have been found in the human genome. Prior to this study, only 11 chicken chemokines and 7 receptors had been reported. The objectives of this study were to identify systematically chicken chemokines and their cognate receptor genes in the chicken genome and to annotate these genes and ligand-receptor binding by a comparative genomics approach. Twenty-three chemokine and 14 chemokine receptor genes were identified in the chicken genome. The number of coding exons in these genes and the syntenies are highly conserved between human, mouse, and chicken although the amino acid sequence homologies are generally low between mammalian and chicken chemokines. Chicken genes were named with the systematic nomenclature used in humans and mice based on phylogeny, synteny, and sequence homology. The

independent nomenclature of chicken chemokines and chemokine receptors suggests that the chicken may have ligand-receptor pairings similar to mammals.

The TLR family represents evolutionarily conserved components of the pattern-recognizing receptors (PRRs) of the innate immune system that recognize specific pathogen-associated molecular patterns (PAMPs) through their ectodomains (ECDs). TLR's ECDs contain 19 to 25 tandem copies of leucine-rich repeat (LRR) motifs. TLRs play important roles in the activation of pro-inflammatory cytokines, chemokines and modulation of antigen-specific adaptive immune responses. To date, nine TLRs have been reported in chicken, along with a non-functional TLR8. Two non-mammalian TLRs, TLR21 and TLR22, have been identified in pufferfish and zebrafish. The objectives of this study were to determine if there is the existence of chicken genes homologous to fish-specific TLRs, and if possible ligands of these receptors exist. After searching the chicken genome sequence and EST database, a novel chicken TLR homologous to fish TLR21 was identified. Phylogenetic analysis indicated that the identified chicken TLR is the orthologue of TLR21 in fish. Bioinformatic analysis of potential PAMP binding sites within LRR insertions showed that CpG DNA is the putative ligand of this receptor.

DEDICATION

To my parents who have supported me so many years for my education.

ACKNOWLEDGEMENTS

I sincerely appreciate my advisor Dr. James J. Zhu for his great support and guidance during my graduate studies. With keen insight and broad knowledge, he introduced me into this exciting area of Bioinformatics and Genomics and led me through this research project.

I would like to thank Dr. Luc R. Berghman for serving as my committee co-chair and thesis advisor with consideration and affection. I also would like to thank Dr. David L. Adelson for serving as one of my committee members and assisting in the revision of the genomic organization, annotation, and ligand-receptor inferences of chicken chemokines and chemokine receptor genes based on comparative genomics manuscript.

I appreciated the help Dr. Ahmet Yilmaz gave me in DNA sequencing, and thank Mr. Suresh Xavier for his assistance in setting up the chicken EST database for this study. I would also like to thank Dr. Sing-Hoi Sze and Yuan Jin in the Computer Science Department at Texas A&M University for helping with designing computer programs for this project. I also want to extend my gratitude to all persons who contributed to chicken genome project, EST sequencing, and the UCSC Genome Browser.

Thanks to Dr. Alan R. Sams, Head of the Department of Poultry Science, and to Mr. Robert L. Pottberg, Ms. Sarah Robertson, Ms. Pattie F. Horsman, and Ms. Elizabeth M. Hirschler for their support during my graduate studies.

TABLE OF CONTENTS

LIST OF FIGURES

FIGURE Page

FIGURE Page

LIST OF TABLES

CHAPTER I

INTRODUCTION

TOLL-LIKE RECEPTORS

The immune response is broadly categorized into innate and adaptive immunity. The innate immune system is the first line of defense against infections and it can recognize a few highly conserved molecular structures produced by microbial pathogens such as bacterial lipopolysaccharides (LPS), lipoteichoic acid (LTA), and double-stranded RNA (dsRNA) (Yamamoto 2004). These conserved molecular structures are collectively referred to as pathogen-associated molecular patterns (PAMPs). The receptors of the innate immune system that recognize PAMPs and trigger various effector responses are called pattern-recognizing receptors (PRRs). These receptors may be secreted or expressed on the surface of immune cells. The Toll-like receptors (TLRs) are a family of transmembrane proteins that serve as PRRs for many microbe-derived molecules and play important roles in the activation of pro-inflammatory cytokines and modulation of antigen-specific adaptive immune responses (Armant and Fenton 2002).

The Toll receptor was originally discovered in *Drosophila* and was identified as an important mediator of development. It was later found to play an essential role in antifungal responses (Anderson et al. 1985). As the Toll receptor is involved in the

---

This thesis follows the style of Immunogenetics.

invertebrate immune system, these receptors may be conserved in vertebrates as an ancient mechanism of innate immune recognition. Subsequently, a family of toll like receptors was discovered in *Drosophila*, as well as in birds and mammals. Currently 10 TLRs (hTLR 1 - hTLR10) have been found in the human genome and the mouse genome contains twelve genes that encode TLRs (mTLR1-9, 11, 12, 13). TLRs differ from each other in their ligand specificities, the cell type they activate, expression patterns and the signaling pathways they utilize. The TLRs recognize distinct PAMPs (Table 1.1), including lipopolysaccharide (LPS) from gram negative bacteria(TLR4) , lipoprotein and peptidoglycan from gram positive bacteria (TLR1 , 2 and 6), double-stranded RNA(TLR3), CpG- containing DNA(TLR9), flagellin(TLR5) and single-stranded RNA(TLR7) (Armant and Fenton 2002; Takeda and Akira 2003; O'Neill 2004)

.

Table 1.1 TLRs, their ligands, role in host defense and disease, co-receptors and adaptor usage. Abbreviations: IBD, inflammatory bowel disease; LPS, lipopolysaccharide; RA, rheumatoid arthritis; RSV, respiratory syncytial virus; SLE, systemic lupus erythrometosis; TLRs, Toll-like receptors; UTIs, urinary tract infections (O'Neill 2004)

TLR	Ligands	Microbial target	Possible role in disease	Co-receptors	Adapters
TLR1	Triacyl lipopeptides	Mycobacteria		TLR2	MyD88
TLR2	Peptidoglycan	Gram positive bacteria	Sepsis, RA, IBD	TLR1, TLR6, CD14, Dectin-1	MyD88, Mal
	Lipoteichoic acid	Gram positive bacteria			
	GPI-linked proteins	Trypanosomes			
	Atypical LPS	Gram negative bacteria			
	Lipoproteins	Mycobacteria			
	Zymosan	Fungi			
TLR3	DsRNA	Viruses		None known	Trif
TLR4	LPS	Gram negative bacteria	Sepsis, RA, IBD, asthma	CD14, MD2, CR3	MyD88, Mal, Trif, Tram
	F protein	RSV			
	Hsp60?	Host			
	FN fragments?	Host			
TLR5	Flagellin	Many bacteria	IBD, Legionnaire's disease	None known	MyD88
TLR6	Diacyl lipopeptides	Mycobacteria		TLR2	MyD88
	Zymosan	Fungi			
TLR7	ssRNA	Many viruses		None known	MyD88
	Imiquimod, R848	Synthetic			
	Loxiribine	Synthetic			
TLR8	ssRNA	Many viruses		None known	MyD88
	R848	Synthetic			
TLR9	Herpes virus DNA	Virus	SLE	None known	MyD88
	CpG ODNs	Synthetic			
TLR10 (mouse)	?				
TLR11 (mouse)	?	Uropathogenic bacteria	UTIs	None known	MyD88
TLR12 (mouse)	?			None known	?
TLR13 (mouse)	?			None known	?

Upon activation by ligand, TLRs induce the production of reactive oxygen and nitrogen intermediates. They also initiate adaptive immunity by activating antigen-presenting cells (APC) by inducing the production of various pro-inflammatory cytokines and chemokines via nuclear factor- κB (NF-κB) or IFN regulatory factor (IRF) signal transduction pathways (Barton and Medzhitov 2002; Yamamoto et al. 2004) (Fig.

1.1). In addition, TLR signaling can regulate T-helper-1 (Th1) and T-helper-2(Th2) responses and stimulate the proliferation and maturation of dendritic cells to the draining lymph nodes in some species (Dabbagh and Lewis 2003; Akira et al. 2001). So signaling through TLRs leads to the immune induction of innate and adaptive immunity in host defense.

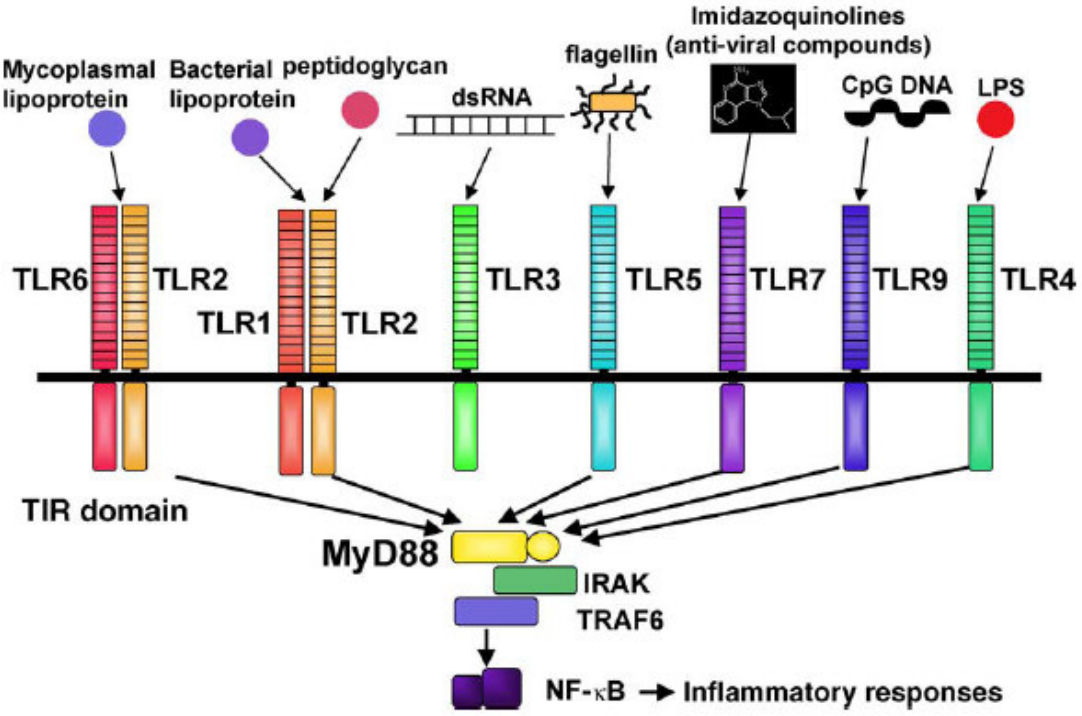


Fig. 1.1 TLRs recognize molecular pattern associated with bacterial pathogens. triacylated lipoprotein for TLR1; peptidoglycan for TLR2; double-stranded RNA for TLR3; lipopolysaccharide (LPS) for TLR4; flagellin for TLR5; diacylated lipoprotein for TLR6; imidazoquinoline and its derivative R-848, for TLR7; and bacterial unmethylated CpG DNA for TLR9. MyD88 associates with the TIR domain of TLRs and transduces signals to induce immune responses (Yamamoto et al. 2004)

CHEMOKINES AND CHEMOKINE RECEPTORS

Following ligand-TLR binding, many genes such as chemokines are induced in activated T cells, mononuclear phagocytic cells and endothelial cells (Lichtman and Abbas 1997). Chemokines are a group of small proteins that regulate leukocyte migration through interactions with G-protein coupled receptors in the immune system (Zlotnik and Yoshie 2000).

Chemokines are divided into two major (CC and CXC) and two minor (XC and CX_3C) subfamilies based on the four conserved cysteines at the N terminus of the polypeptide. The first two cysteines in the two major subfamilies are either adjacent (CC) or separated by one amino acid (CXC). The first two cysteines in the CX_3C chemokines are separated by three amino acids, whereas the XC chemokines contain only two of the cysteines (Murphy 2003) (Fig. 1.2).

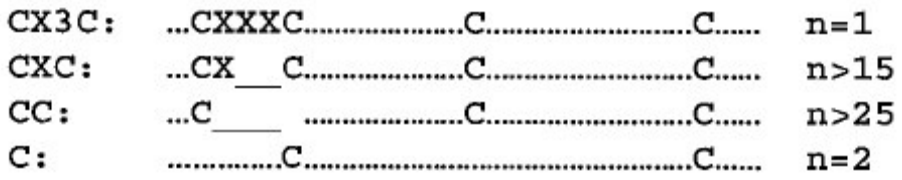


Fig. 1.2 Structural classification of the chemokine family by signature cysteines. The number of members in each subclass is listed at the right of each structure. Underlines indicate gaps in the alignment; X, an amino acid other than cysteine; and dots, other amino acids. Spacing between cysteines is similar in all four groups. The N and C termini can vary in length (Murphy 2002)

The CC chemokines can be further divided into two subcategories, MCPs (monocyte chemoattractant proteins) and MIPs (macrophage inflammatory proteins) based on their structural similarities (Van Coillie et al. 1999). The members of these two CC subcategories specifically attract mononuclear cells but not neutrophils. The CXC chemokines can also be divided into two subfamilies, one with an ELR (a conserved Glu-Leu-Arg preceding the first cysteine) motif, which is angiogenic and attracts neutrophils, and the other without the ELR motif, whose members do not attract neutrophils (Laing and Secombes 2004).

Chemokine genes are characterized by their chromosomal locations and similar gene structure. Most human CC and CXC genes are organized in gene clusters in mammalian genomes, such as human Chromosomes 4 and 17, and mouse Chromosomes 5 and 11 (Nomiyama et al. 2001). The genes encoding the CC subfamily contain three exons, whereas the CXC chemokine genes contain four exons (Forssmann et al. 2001; Kaiser et al. 1999). The XC subfamily of chemokines contains two members in human but only one in mouse. CX_3CL1 is the only known member of the CX_3C subfamily in human, mouse, rat, and monkey. There are extensive conserved syntenies in the chromosomal regions containing chemokine genes between human and mouse.

In humans, many transcription factor binding sites (TFBS), including sites for NF-kB, hepatocyte NF-1, and AP-1, have been located in the promoter region of chemokine genes such as IL-8, MCP-1, and CCL5. These TFBS play important roles in the transcriptional control of expression of chemokine genes (Martin et al. 1997; Rovin et al. 1995).

Unlike chemokines, chemokine receptors share a higher degree of sequence identity within a species and between species. They have characteristic seven alpha-helix transmembrane domains, are between 340-370 amino acids long, and have up to 80% amino acid identity (Olson and Ley, 2002; Horuk, 2001) (Fig. 1.3). They also share an acidic amino terminus, a conserved sequence in the second intracellular loop, and one cysteine in each extracellular domain (Murphy, 1996). Most receptors can bind several chemokines with high affinity but only from a single class (Horuk, 2001). Like chemokines, most chemokine receptors are clustered in a few chromosomal regions, such as human Chromosomes 2 and 3 (Onuffer and Horuk, 2002). Almost all the chemokine receptors have their amino acid sequences encoded in a single exon.

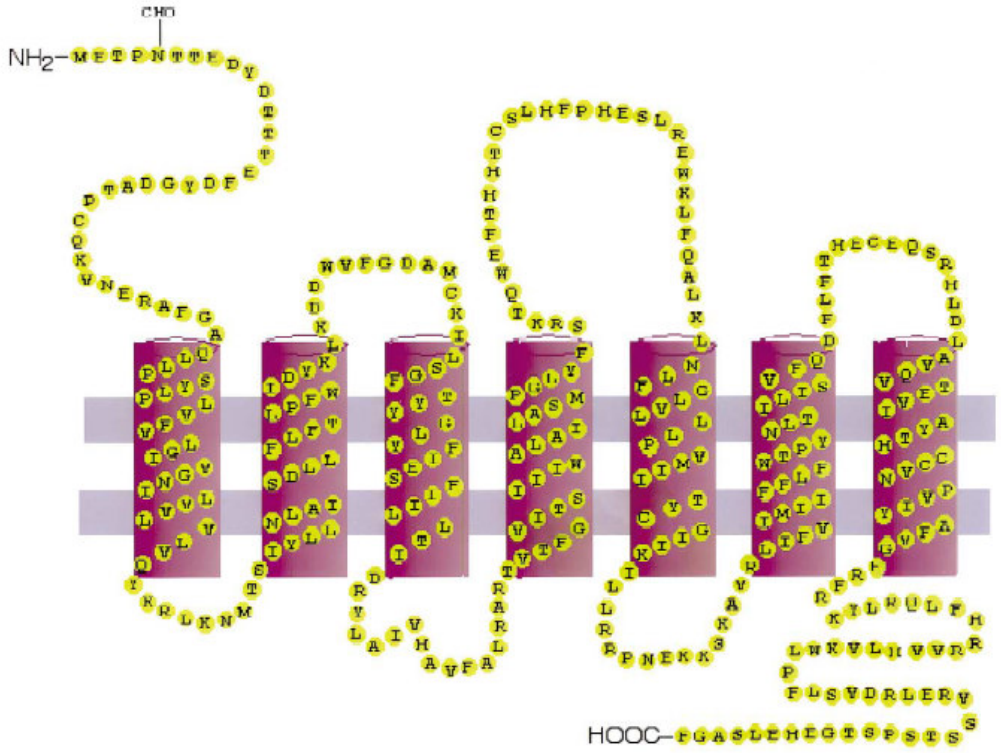


Fig. 1.3 A membrane topography of CCR1. The chemokine receptors had seven conserved transmembrane helices, four extracellular loops with a conserved cysteine, and a conserved DRYLAIVHA sequence in the second intracellular domain. CHO is potential N-linked glycosylation sites (Horuk, 2001)

The chicken represents a useful model organism to study disease resistance in the immune system. However, many chicken homologues of mammalian chemokines have not been been identified.

Comparative genomics, which provides information based on the sequence conservation between organisms, is a powerful approach to genome annotation. It has been used extensively to identify novel genes with bioinformatics tools and computational methods. The chicken draft genome sequence released on March 1, 2004 by National Institutes of Health is available at the University of California at Santa Cruz (UCSC) Genome Browser (http://genome.ucsc.edu), and offers great opportunities for identification and characterization of novel genes. The objectives of this study were to identify novel chemokines, chemokine receptors and Toll like receptor genes in chickens using comparative genomics techniques, and to make functional inferences from the sequence analyses of the novel genes.

CHAPTER II

GENOME-WIDE IDENTIFICATION, LIGAND-RECEPTOR INFERENCES OF CHICKEN CHEMOKINES AND CHEMOKINE RECEPTOR GENES BASED ON COMPARATIVE GENOMICS*

INTRODUCTION

Chemokines are a family of small chemoattrative peptides that were originally recognized to be involved in host defense as regulators of leukocyte trafficking, but more recently have been shown to have roles in organogenesis, hematopoiesis, neuronal communication and HIV infection (Olson and Ley, 2002; Baggiolini et al. 1997). Their cognate receptors belong to a Class A subfamily of the largest superfamily of G-protein coupled receptors (Onuffer and Horuk, 2002). Chemokines are believed to have originated from gene duplications and these genes underwent selection during recent evolutionary time (Zlotnik and Yoshie, 2000). All chemokines have a characteristic cysteine motif. Similarly, chemokine receptors may also be derived from a common ancestor through gene duplications. All chemokine receptors share high homology with the prototypical family member, rhodopsin (Onuffer and Horuk, 2002).

---

*Reprinted from "Genomic organization, annotation, and ligand-receptor inferences of chicken chemokines and chemokine receptor genes based on comparative genomics" by Jixin Wang, David L Adelson, Ahmet Yilmaz, Sing-Hoi Sze, Yuan Jin and James J Zhu, 2005, BMC Genomics, 6(1):45.

Chemokines are highly basic proteins, 70 to 125 amino acids long. Sequence identity among chemokines is usually low; however, all share a typical overall tertiary structure, which consists of at least four cysteines that form two disulfide bonds.

At present, 42 chemokine genes have been identified in human (24 CXC, 15 CC, 1 CX_3C, and 2 XC) and 36 (21 CXC, 13 CC, 1 CX_3C, and 1 XC) in mouse, whereas there are 11 receptors for CCLs, 6 for CXCLs, 1 for CX3CL, and 1 for XCL in human and mouse. Only 11 chicken chemokines including 4 CXC, 6 CC, and 1 XC, and seven chicken chemokine receptors including 2CXCR and 5CCR have been reported in the literature (Petrenko et al., 1995; Petrenko et al. 1997; Sugano et al., 1987; Li et al., 2000; Hughes et al., 2001; Liang et al., 2001, Smith et al., 2004; Read et al., 2005; Sick et al. 2000; Hartl et al., 2000; Rossi et al., 1999). Chicken chemokines share low sequence identity with mammals (Kaiser et al. 1999). Therefore, it is very difficult to assign chicken chemokines to a specific mammalian counterpart based on sequence data alone. Because of limited sequence homology, most of the reported chicken chemokines were not named in accordance with the systematic nomenclature of mammalian chemokines. The newly available chicken draft genome sequence and a large number ESTs has allowed systematic identification and annotation of chicken chemokine and cognate receptor genes. The objectives of this study were to systematically identify chemokine and chemokine receptor genes in the chicken genome, to name these genes according to existing systematic nomenclature, and to make ligand-receptor binding inferences based on comparative sequence analysis. The systematic nomenclature for these chicken genes was based on the phylogenetic trees and syntenies of chicken, human, and mouse genes,

and ligand-receptor binding inferences were according to the recommendations of the IUIS/WHO Subcommittee on Chemokine Nomenclature (Bacon et al. 2002 and Murphy 2002).

MATERIALS AND METHODS

Gene identification

To identify syntenies, genes closely linked to human and mouse chemokines were identified and localized on the chicken genome using the UCSC genome browser (http://genome.ucsc.edu). Expressed Sequence Tags (ESTs) and chicken mRNA sequences in the corresponding chromosomal regions were then identified and, if necessary, assembled with the CAP3 program (http://pbil.univ-lyon1.fr/cap3.php) (Huang and Madan 1999). These sequences were aligned with the corresponding chicken genomic sequence and any deletions or insertions corrected. Sequences were then submitted to ORF Finder (Open Reading Frame Finder) (http://www.ncbi.nlm.gov/gorf/gorf.html) and the open reading frames were used as queries in BLASTP searches against the non-redundant protein database in Genbank (http://www.ncbi.nih.gov/Genbank). Sequences that produced significant alignments with chemokines were identified as putative chicken chemokine sequences. To identify chicken chemokine receptors, all sequences of putative chicken chemokine receptors including ESTs, mRNAs, and predicted sequences were retrieved from the UCSC Genome Browser. The identified ESTs were used to determine the translation start sites for the receptors. If the translation start sites could not be determined from ESTs,

translation start sites were based on the most likely predicted sequences from non-chicken reference genes in the UCSC Genome Browser.

Sequence analyses

Complete amino acid sequences of currently known human and mouse chemokines were retrieved from Genbank (http://www.ncbi.nih.gov/Genbank/). The amino acid sequences of all putative chicken chemokines were predicted based on the open reading frames of the expressed nucleotide sequences (ESTs or mRNAs). The amino acid sequences were grouped according to CC, CXC, and CX3C motifs and aligned using the ClustalW program (http://www.ebi.ac.uk/clustalw). The percent similarity of the amino acid sequences was determined based on alignments with the most likely human or mouse orthologs. Human CCR1 was included in the multiple alignments of chicken chemokine receptors for comparison. The seven transmembrane domains were predicted using the SMART program (http://194.94.45.211/smart/show_motifs.pl).

For comparison, human chemokines hCCL1 (GenBank accession number: (NM_002981), hCCL2 (BC009716), hCCL3 (BC071834), hCCL4 (NM_002984), hCCL5 (BC008600), hCCL7 (NM_006273), hCCL8 (NM_005623), hCCL11 (BC017850), hCCL13 (BC008621), hCCL14(BC045165), hCCL15 (NM_032964), hCCL16 (NM_004590), hCCL17 (BC069107), hCCL18 (BC069700), hCCL19 (BC027968), hCCL20 (BC020698), hCCL21 (BC027918), hCCL22 (BC027952), hCCL23 (NM_145898), hCCL24 (BC069072), hCCL25 (NM_005624), hCCL26 (BC069394), hCCL27 (AJ243542), hCCL28 (AF220210), hCXCL1 (BC011976),

hCXCL2 (BC015753), hCXCL3 (BC065743), hCXCL4 (NM_002619), hCXCL5 (BC008376), hCXCL6 (BC013744), hCXCL7 (BC028217), hCXCL8 (BC013615), hCXCL9 (BC063122), hCXCL10 (BC010954), hCXCL11 (BC012532), hCXCL12 (BC039893), hCXCL13 (BC012589), hCXCL14 (BC003513), and hCXCL16 (BC017588), and hCX_3CL1(NM_002996) and mouse chemokines CCL1 (NM_011329), mCCL2 (NM_011333), mCCL3 (NM_011337), mCCL4 (NM_013652), mCCL5 (BC033508), mCCL6 (BC002073), mCCL7 (BC061126), mCCL8 (NM_021443), mCCL9 (NM_011338), mCCL10 (U15209), mCCL11 (NM_011330), mCCL12 (BC027520), mCCL17 (BC028505), mCCL19 (BC051472), mCCL20 (BC028504), mCCL25 (NM_009138), mCCL27 (BC028511), mCCL28 (BC055864), and mCX_3CL1 (BC054838) were retrieved from the GenBank. Reported chicken chemokines K60 (Y14971), cCAF (M16199), MIP-1β (AJ243034), k203 (Y18692), AH294 (AY037859), AH221 (AY037860), AH189 (AY037861), JSC (AF285876), SDF-1(BX936268), Clone 391 (L34552) and lymphotactin (AF006742) are included in this study. Rat (BC070938) and monkey (AF449286) CX_3CL1 were also retrieved for CX_3CL sequence analysis. There are several human chemokine-like genes in the human genome, which were not included in this study.

Human and mouse chemokine receptors hCCR1 (NM_001295), hCCR2 (NM_000647), hCCR3 (NM_001837), hCCR4 (NM_005508), hCCR5 (NM_000579), hCCR6 (NM_004367), hCCR7 (NM_001838), hCCR8 (NM_005201), hCCR9 (NM_006641), hCCR10 (AY429103), hCCR11 (AF193507), hCXCR1 (NM_000634), hCXCR2 (BC037961), hCXCR3 (NM_001504), hCXCR4 (AY728138), hCXCR5

(NM_032966), hCXCR6 (NM_006564), hCX3CR1 (NP_001328), and hXCR1 (NM_005283), mCXCR1 (AY749637), mCXCR2 (NM_009909), mCXCR3 (NM_009910), mCXCR4 (NM_009911), mCXCR5 (NM_007551), mCXCR6 (NM_030712), mCCR1 (NM_009912), mCCR2 (NM_009915), mCCR3 (NM_009914), mCCR4 (NM_009916), mCCR5 (NM_009917), mCCR6 (NM_009835), mCCR7 (NM_007719), mCCR8 (NM_007720), mCCR9 (NM_0099130), mCCR10 (AF215982), mCCR11 (AF306532), mCX3CR1 (NM_009987), and mXCR1 (NM_011798), and reported chicken cCCR2 (CAF28776), cCCR5 (CAF28777), cCCR8L1(CAF28778), cCCR9 (CAF28781), cCXCR1 (AAG33964), cCXCR4 (NP_989948), and cXCR1 (CAF28779), were also retrieved from GenBank for comparisons.

Phylogenetic analyses of protein sequences of chicken, human, and mouse chemokines and chemokine receptors were based on the amino acid sequences using neighbor-joining with options selected for bootstrap test, pairwise deletion and Poisson correction, using MEGA3 (Kumar et al., 2004). For ligand-receptor inference, the first 20 amino acids (leading peptide) of all chemokines were removed before the phylogenetic analysis and chicken CCLs were divided into two groups, one group located on chromosomes 4, 19, and the other from other chromosomes. Syntenies, phylogenetic trees, and sequence homologies were the combined information used for naming chicken chemokine and their cognate receptor according to the recommendations of the IUIS/WHO Subcommittee on Chemokine Nomenclature (Bacon et al., 2002). These chicken genes were named according to their closest human

or mouse predicted orthologs if all information supports the nomenclature. If there was more than one chicken gene similar to a human and/or mouse gene, these gene was named as in the human and/or mouse followed by a letter with alphabet order. If a specific human or mouse ortholog could not be reliable determined, the chicken genes were named according to a closest human or mouse ortholog followed by an "L" and a number based on the information available. This nomenclature also used the existing systematic names reported in the literature to avoid confusion.

Transcription factor binding sites (TFBS)

The promoter sequence of the chicken chemokine genes were predicted from a fragment of 3000bp upstream of the transcription start site using a Neural Network Promoter Prediction program (http://www.fruitfly.org/seq_tools/promoter.html) (Reese, 2001). The DNA sequence 3000 nucleotides upstream and downstream of the TATA box in the predicted promoter of the chicken chemokine genes was used to identify the transcription factor binding sites using the MATCH 2.0 software and a TFBS matrix for all vertebrates in the TRANSFAC database (http://unix.cognia.com/cgi-bin/biobase/transfac/8.2/bin/start.cgi).

Polymerase chain reaction (PCR) and DNA sequencing

Chicken EST or mRNA sequences were identified for all chemokine genes. All sequences contained complete putative open reading frames except for CX_3CL1. However, partial chicken CX_3CL1 gene sequences (BM426140, BI066258, and CR389767) were identified, with a gap of 123 nucleotides between the ESTs. Forward (TGTGACATCGGGAGTCGCTAC) and reverse (AAAATCCCCAGCGTTTGCTACT)

PCR primers were used to amplify across the gap using cDNA prepared from white blood cells. PCR was performed as follows: An initial denaturation step at 94^0C for 2 min and 35 cycles of denaturation, annealing, and extension at 94^0C for 30 sec, 59^0C for 45 sec, and 72^0C for 1 min., and a final extension step was carried out at 72^0C for 10 min. Unincorporated nucleotides were removed from amplified PCR products using BioMax spin-50 mini-columns (Millipore, Billerica, MA). BigDye terminator cycle sequencing reaction kits and an ABI Prism 377XL DNA Sequencer (Applied Biosystems) were used for DNA sequencing.

RESULTS

Chicken chemokines and chemokine receptors

In addition to the 11 previously reported, 12 new chicken chemokine were identified. These included 7 new CC chemokines named CCL1L1 (BX935885), CCL3L1 (CF258095), CCL/MCP-L2 (CK610627), CCL/MCP-L3 (CF251629), CCL17 (BI067703), CCL19 (BX929857), and CCL21 (CR522995), 4 new chicken CXC chemokines named CXCL13a (BX262175), CXCL13b (BX264625), and CXCL13c (CR352598), CXCL15 (BX929947), and 1 CX_3CL1 chemokine (assembled from CR389767, BI066258, BM426140, and our sequence: AY730688). Eleven reported chicken genes were also named accordingly as CCL1L2 (L34552), CCL5 (ah294, AY037859), CCL4L1 (MIP-1β, AJ243034), CCL/MCP-L1 (ah221, AY037860/BX933162), CCL16 (k203, Y18692), CCL20 (ah189, AY037861), CXCL8a (cCAF, M16199), CXCL8b (K60, Y14971), CXCL12 (SDF-1, BX936268), CXCL14 (JSC, AF285876), and XCL1 (lymphotectin, AF006742). In summary, there were 13

CCL, 8 CXCL, 1 CX_3CL, and 1 XCL genes identified in the chicken genome. The information used for the nomenclature is shown in the comparative genome maps and phylogenetic trees.

Chicken chemokine amino acid sequence alignment showed that all chicken CC chemokines have four conserved cysteines with two adjacent cysteines at the N-terminus (Fig. 2.1), whereas all chicken CXC chemokines have the conserved four cysteines with the first two cysteines separated by one amino acid (Fig. 2.2). Both chicken CCLs and CXCLs show higher degrees of sequence similarity to each other in the signal peptide sequences and sequence regions containing the last two cysteines. CXCL8a, CXCL8b and the newly identified chemokine CXCL15 contain the ELR (Glu-Leu-Arg) motif. Only one chicken CX_3C chemokine was found (Fig. 2.3). The number of amino acid residues between conserved cysteines in all chemokines was highly conserved between chicken and human (Table 2.1).

Fig. 2.1 Alignment of amino acid sequences of chicken chemokine CC subfamily. Alignment gaps are indicated by dashes. Sequences with identical amino acid in at least 50% of chicken chemokines are highlighted in gray and conserved cysteine residues in dark gray

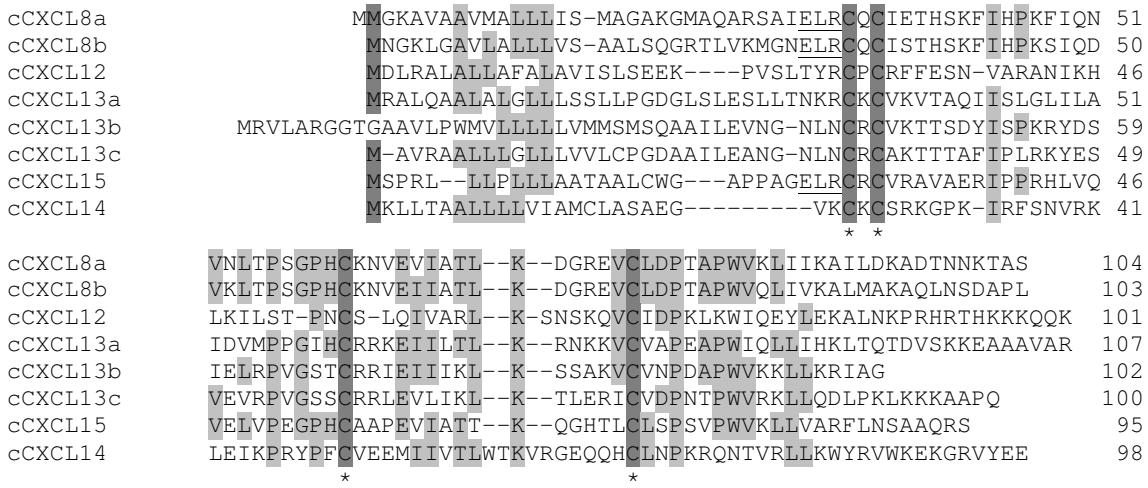


Fig. 2.2 Alignment of amino acid sequences of chicken chemokine CXC subfamily. Alignment gaps are indicated by dashes. Sequences with identical amino acid in at least 50% of chicken chemokines are highlighted in gray and conserved cysteine residues in dark gray. The conserved ELR motifs are underlined

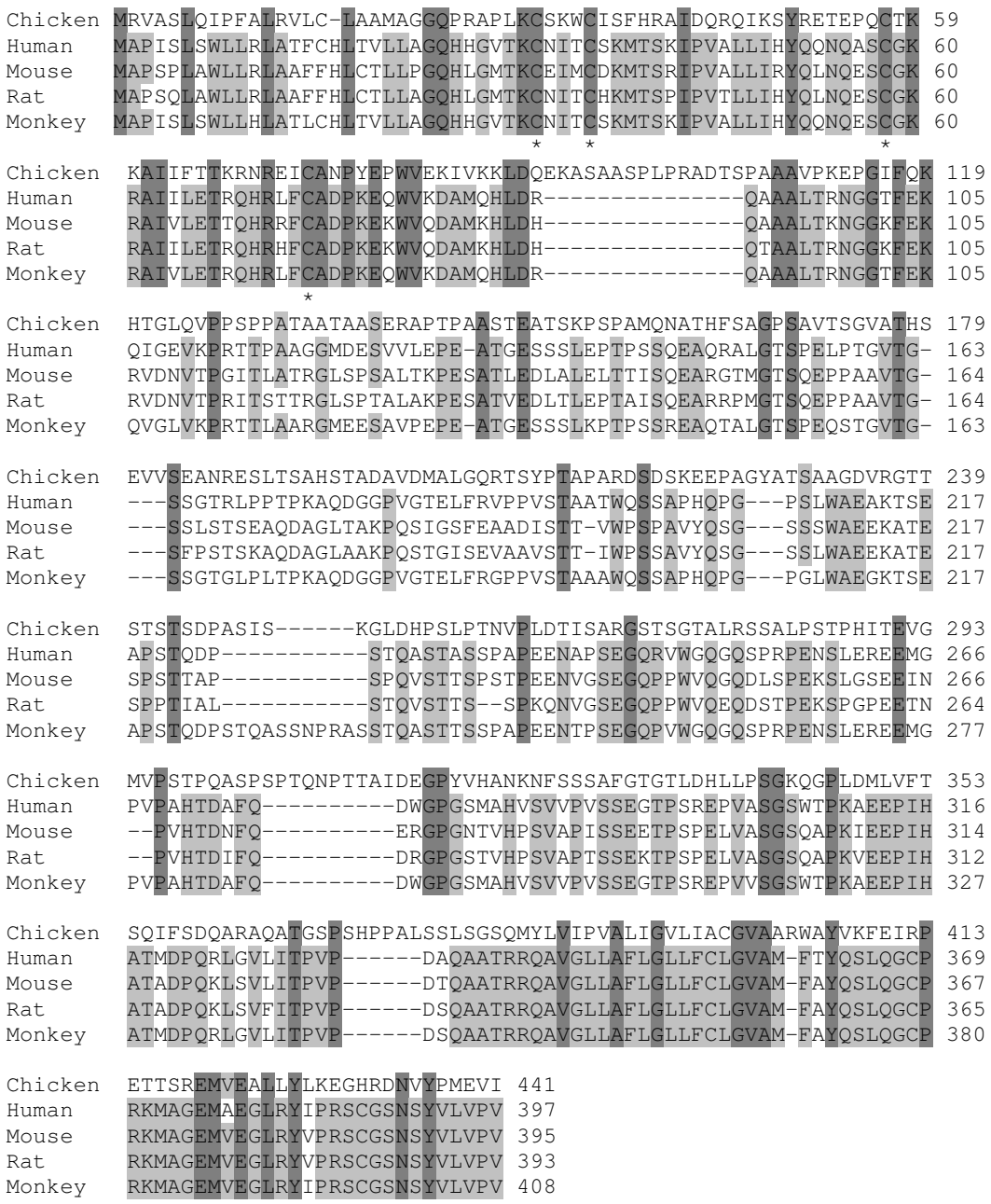


Fig. 2.3 Alignment of amino acid sequences of chicken, human, mouse, rat and monkey chemokine CX_3CL1. Alignment gaps are indicated by dashes. Sequences identical in all species are highlighted in gray. The asterisk represents the conserved cysteine residues

Table 2.1 Chicken and corresponding human chemokine cysteine motifs

Families	Chemokines	Chicken motif	Human motif
CC	CCL1L1	$CCX_{25}CX_{15}C$	$CCX_{22}CX_{15}C$
	CCL1L2	$CCX_{24}CX_{15}C$	$CCX_{22}CX_{15}C$
	CCL3L1, CCL4, CCL5	$CCX_{22}CX_{15}C$	$CCX_{22}CX_{15}C$
	CCL16	$CCX_{22}CX_{15}C$	$CCX_{21}CX_{15}C$
	CCL17	$CCX_{22}CX_{15}C$	$CCX_{22}CX_{15}C$
	CCL19	$CCX_{25}CX_{15}C$	$CCX_{24}CX_{15}C$
	CCL20	$CCX_{24}CX_{15}C$	$CCX_{24}CX_{15}C$
	CCL21	$CCX_{25}CX_{15}C$	$CCX_{24}CX_{17}C$
	CCL13L1, CCL13L2, CCL13L3	$CCX_{22}CX_{15}C$	$CCX_{22}CX_{15}C$
CXC	CXCL8a, CXCL8b	$CXCX_{24}CX_{15}C$	$CXCX_{24}CX_{15}C$
	CXCL13a, CXC13b, CXCL13c	$CXCX_{24}CX_{15}C$	$CXCX_{24}CX_{15}C$
	CXCL12	$CXCX_{22}CX_{15}C$	$CXCX_{22}CX_{15}C$
	CXCL14	$CXCX_{23}CX_{19}C$	$CXCX_{23}CX_{20}C$
	CXCL15	$CXCX_{24}CX_{15}C$	N/A[1]
XC	XCL1	$CX_{36}C$	$CX_{36}C$
CX_3C	CX_3CL1	$CX_3CX_{21}C_{15}C$	$CX_3CX_{21}C_{15}C$

[1] CXCL15 is not found in humans.

Chicken chemokines have limited amino acid sequence similarity compared to their human counterparts. Generally, chicken CXC chemokines shared 27 to 60% amino acid identity with their human homologs, except for CXCL12, which shared 73% identity with human CXCL12. The length of chicken chemokine CXCL polypeptides ranged from 95 to 107 amino acids. Compared to their human homologs, chicken CXCL chemokine amino acid sequences are shorter except for chicken CXCL8a, CXCL8b, and CXCL12, which were 4, 5, 12 amino acids longer than their respective

human homologs. In contrast, the sequence identities between human and chicken CCL chemokines are generally greater than for CXCL chemokines, ranging from 25 to 56%. Chicken chemokine CCL polypeptides range from 89 to 108 amino acids in length. Chicken CCL1L2, CCL5, and CCL17 have the same amino acid length as their human counterparts. Chicken CCL4L1 is shorter than corresponding human CCL4, whereas chicken CCL1L1, CCL3L1, CCL19 and CCL20 are longer than the corresponding human CCLs. These differences in length between human and chicken chemokines are exclusively in the N- and C- termini.

Chicken CX_3CL1 encodes 441 amino acids, longer than all mammalian CX_3CL1 examined. It shares 20, 22, 20, and 22% amino acid identity with human, mouse, rat, and monkey CX_3CL1, respectively. Sixty-six amino acid residues in chicken CX_3CL1 are identical to residues in mammals, but the sequence identity between mammals is, as expected, much higher than that between chickens and mammals (Fig. 2.3). There are more identical amino acid residues between chicken and the mammals at both ends of the sequences.

Unlike the chemokines, all the chicken chemokine receptor genes were aligned with non-chicken chemokine receptor reference genes in the chicken genome browser. There was at least one chicken EST sequence aligned to each receptor gene except for CCR4. In addition to 7 reported receptors, 7 new chicken chemokine receptors were identified and named as CCR4 (predicted sequences: ENSGALT00000019505.1/ chr2_792.1), CCR6 (CV039916, BU451770, and CK987456), CCR7 (predicted sequence: chr27_random_59.1), CCR8a (AJ720982), CXCR2 (BX258468), CXCR5

(AJ450829), CX3CR1 (CF252942, BU204148, and AJ443633). In contrast to chicken chemokines, chicken chemokine receptors share significant amino acid identity with their human receptor counterparts. The percents of amino acid identity between chicken and human chemokine receptors ranged from 48 to 81%. The lengths of these chicken receptors range from 335 to 382 amino acids. The complete sequence of chicken CXCR2 is unknown due to a sequence gap in the chicken genome sequence. The CXCR2 EST and a partial genome sequence contain the last 170 amino acids of the C-terminus.

Forty-four amino acid residues were highly conservative (>85% homology) among all chicken chemokine receptors (Fig. 2.4). These receptors all have seven transmembrane helices and three extracellular loops. Of the seven transmembrane helices, helix 1 and 7 show higher degrees of sequence similarity than the other helices. The similarity between the extracellular domains of the chicken receptors is lower, but all had a conserved cysteine residue. In contrast, the intracellular domains (except at the C-terminus) generally had higher degrees of sequence similarity than the extracellular domains. The second intracellular domains contain a highly conserved DRYLAIV sequence.

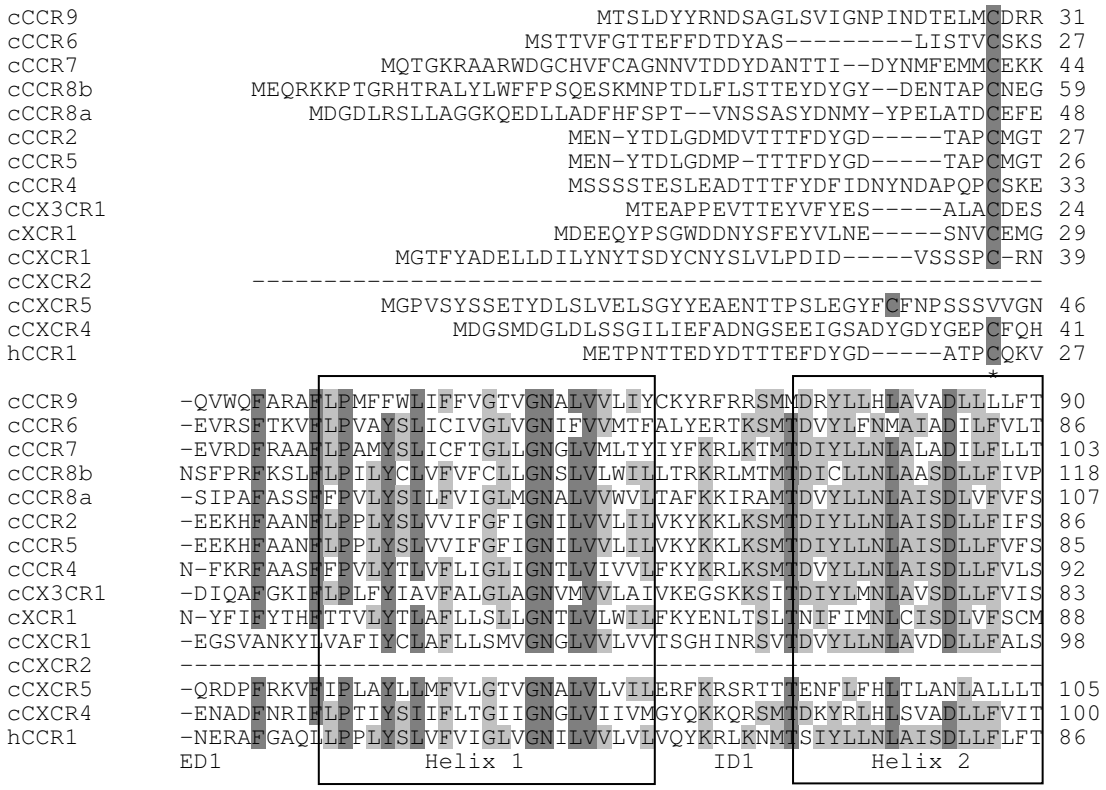


Fig. 2.4 Alignment of amino acid sequences of chicken chemokine receptors with human CCR1. Alignment gaps are indicated by dashes. Sequences with identical amino acid in at least 50% or 85% of the chicken chemokines are highlighted in gray and dark gray, respectively. Asterisks represent the conserved cysteine residues. ED and ID denote extracellular and intracellular domains, respectively. Seven transmembrane spanning domains of chicken chemokine receptors were predicted using the SMART program and these consensus domains are indicated with a box. The N-terminal sequence of chicken CXCR2 is currently unknown

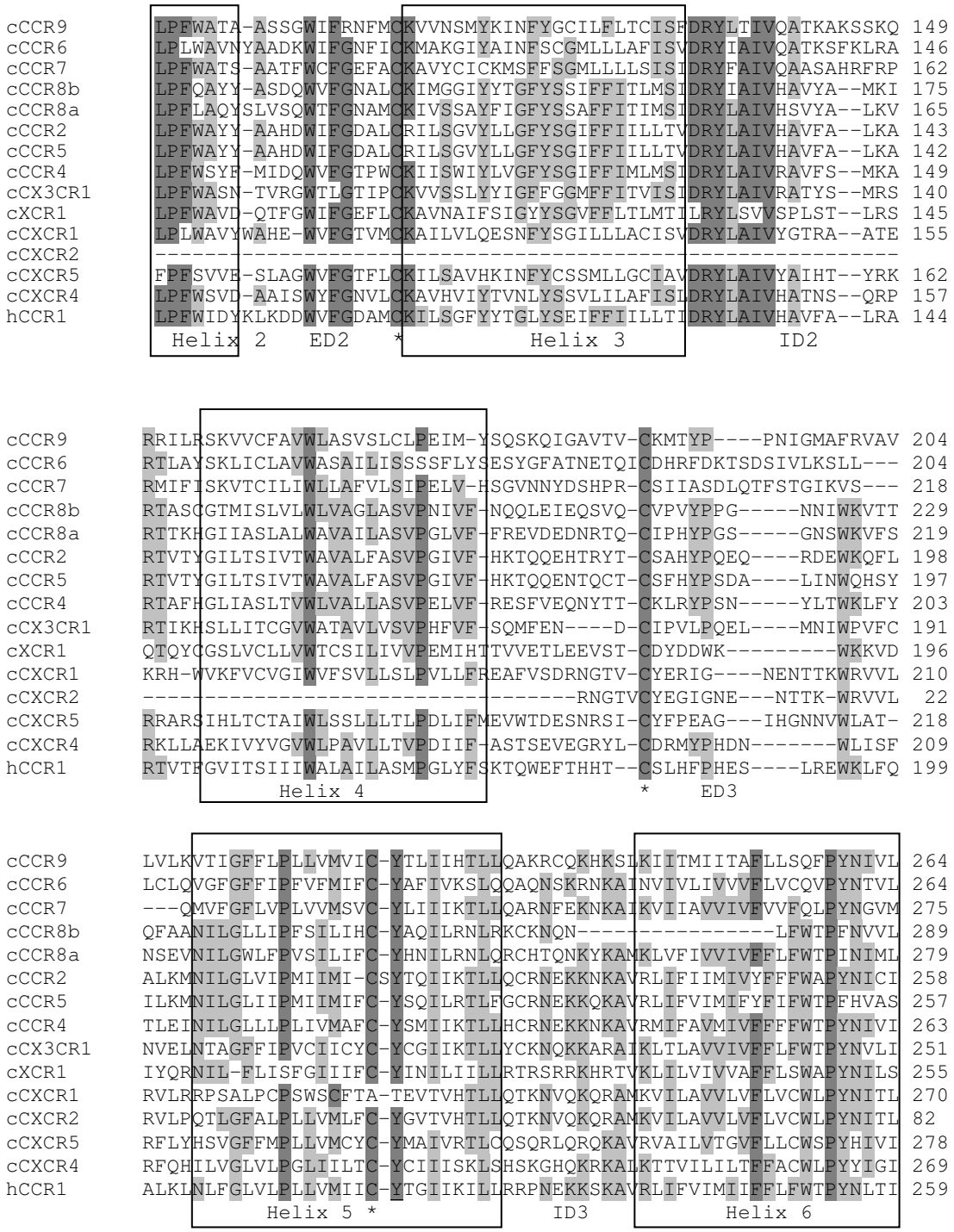
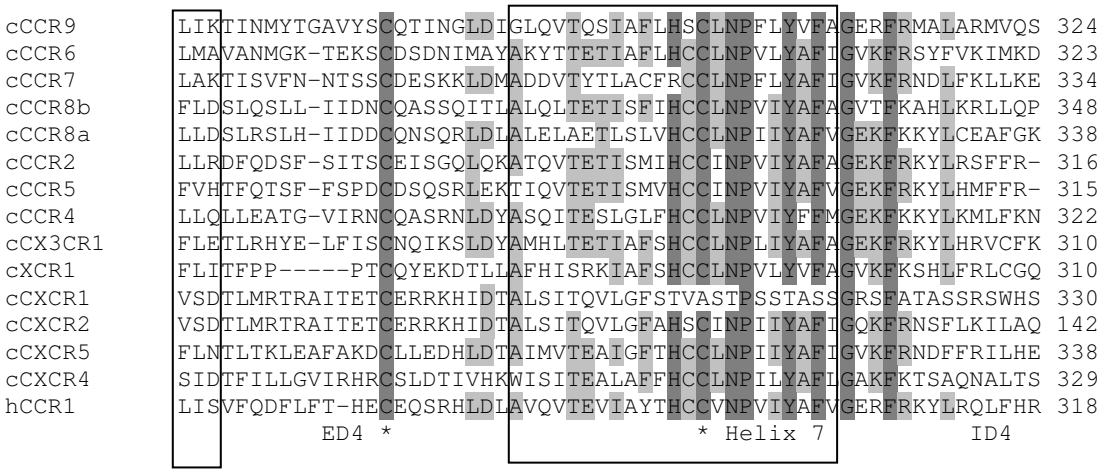


Fig. 2.4 (continued)



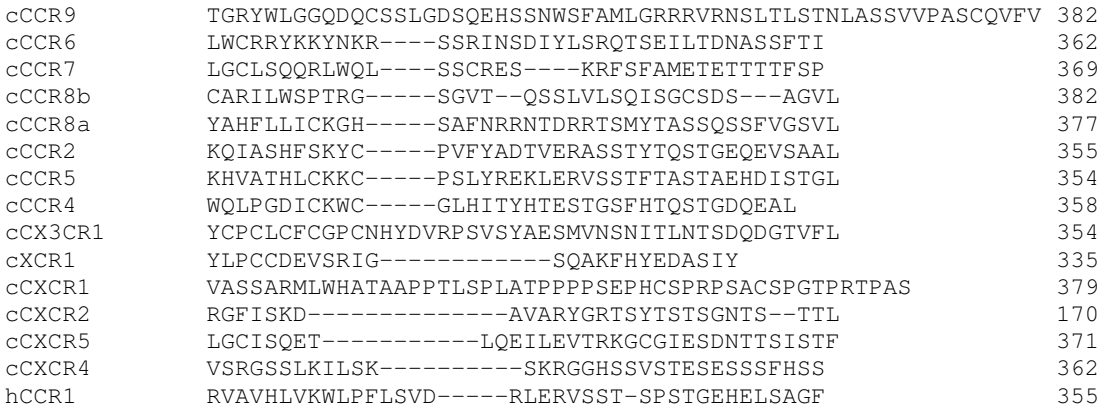
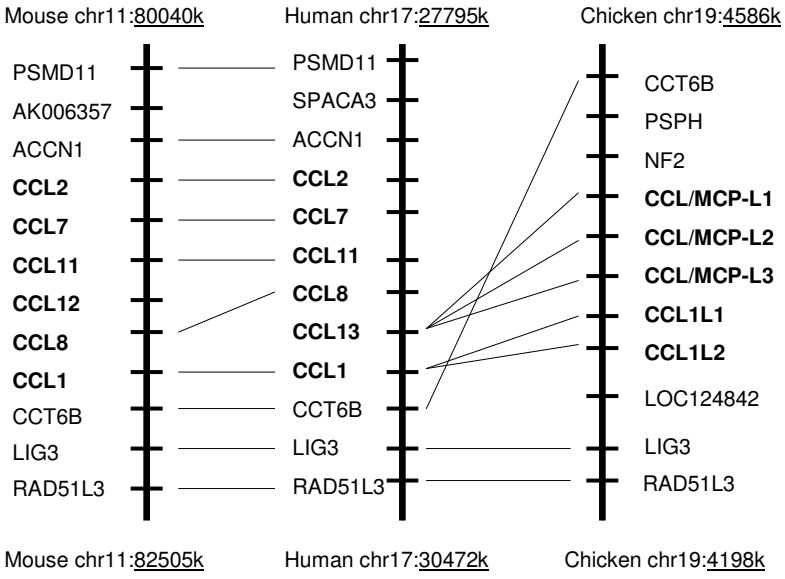


Fig. 2.4 (continued)

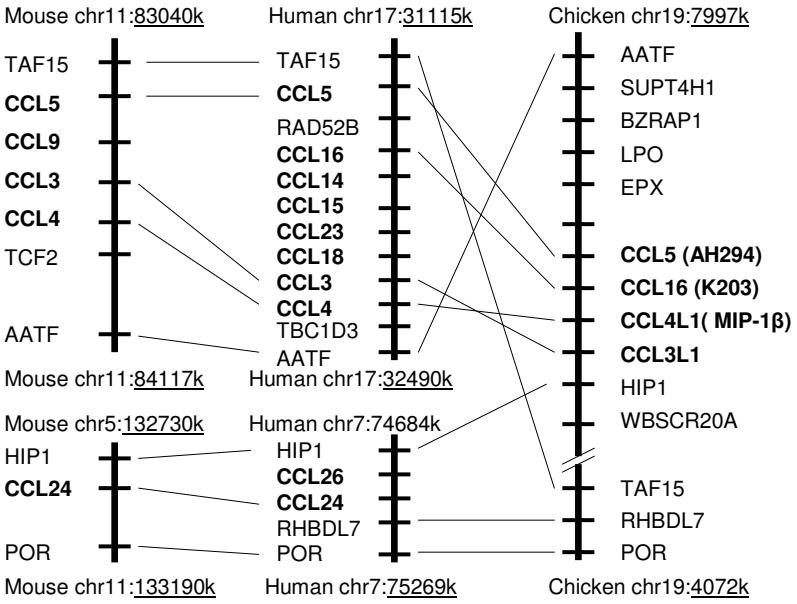
Chromosomal locations and syntenies

Comparisons of the chromosomal segments containing chemokines in human, mouse, and chicken indicated that the organization of the chemokine genes was generally conserved between chicken and mammals (Fig. 2.5 and Fig. 2.6). Chicken CC and CXC chemokines are located on Chromosomes 1, 4, 6, 9, 13, 19, and Z. Like

human and mouse, there were two large clusters in the chicken genome, located on Chromosome 19, containing 9 CCL genes. Two CCL1-like (CCL1L1 and 2) and three chicken MCP-like (CCL/MCP-L1, L2, and L3) genes related to human and mouse MCPs, such as CCL2, 7, 8, 11, and 13, are in one cluster (Fig. 2.5A), and CCL5, CCL16, CCL3L1, and CCL4L1 genes in another cluster (Fig. 2.5B). Another CCL cluster is located on Chromosome Z containing two genes, CCL19 and CCL21 (Fig. 2.5C). Two CXCL gene clusters were located on Chromosome 4 and contained 6 genes, two CXCL8 (CXCL8a and b) and one CCL15 genes in one cluster (Fig. 2.6A) and three CXCL13 (CXCL13a, b, and c) genes in another (Fig. 2.6B). Chicken shares the syntenies with mouse and human in all these regions. There is one composite cluster containing one CX_3CL1 and one CCL17 genes (Fig. 2.6E). Synteny was conserved in chicken on one side of this cluster. Chicken CCL20, CXCL12, CXCL14, and XCL1 are individually located on Chromosomes 9 (Fig. 2.5D), 6 (Fig. 2.6C), 13 (Fig.2.6D), and 1 (Fig. 2.6F), respectively, and the syntenies were highly conserved between chicken, mouse, and human in these four locations. Mammalian CCL25, CCL28, and CXCL16 were not found in the chicken genome, although the syntenies associated with CCL25 and CCL28 were also conserved in chickens. A number of human chemokines, including CCL2, 7, 8, 11, 15, 18, 23, 24, and 26, CXCL1, 2, 3, 4, 5, 6, 7, 9, 10, and 11 in chemokine clusters that share syntenies with chicken clusters on Chromosome 4, and 19 were not found in the chicken genome, indicating gene duplications in mammals.

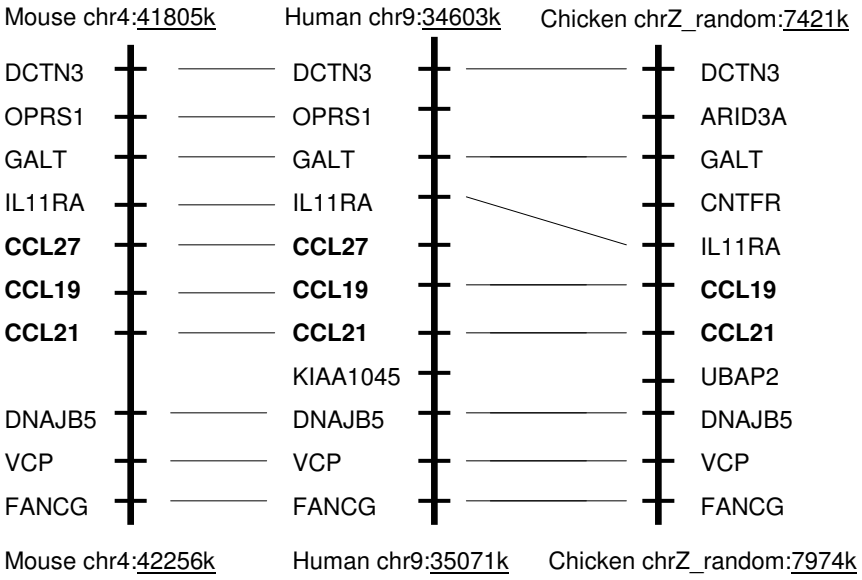


A

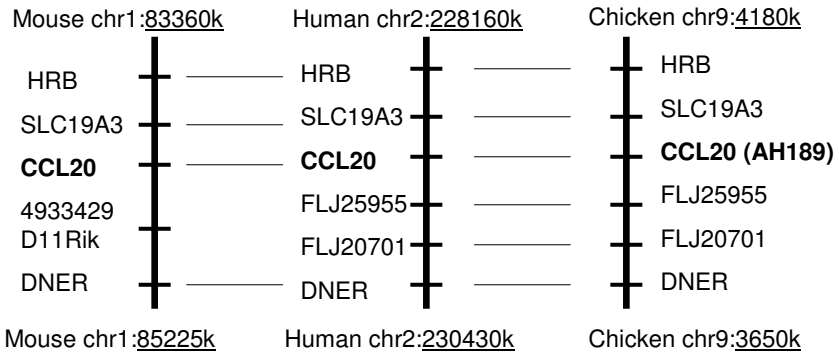


B

Fig. 2.5 Genomic organization (syntenies) of human, mouse, and chicken CCLs, CX3CLs, and XCLs

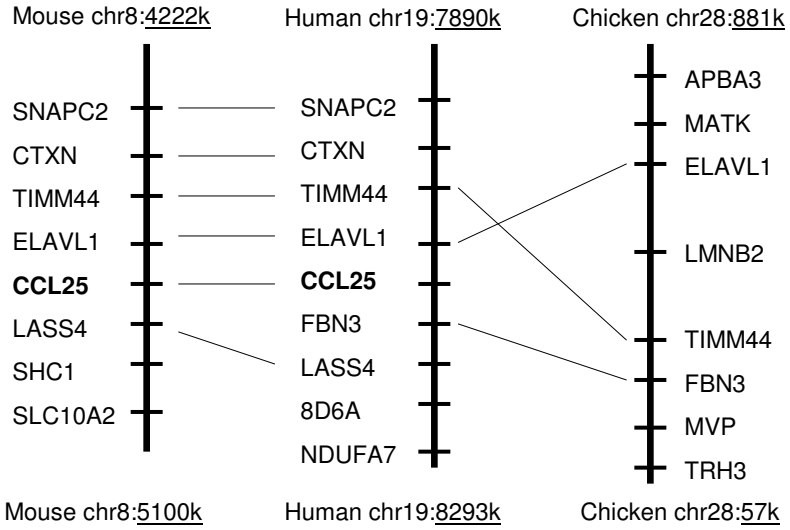


C

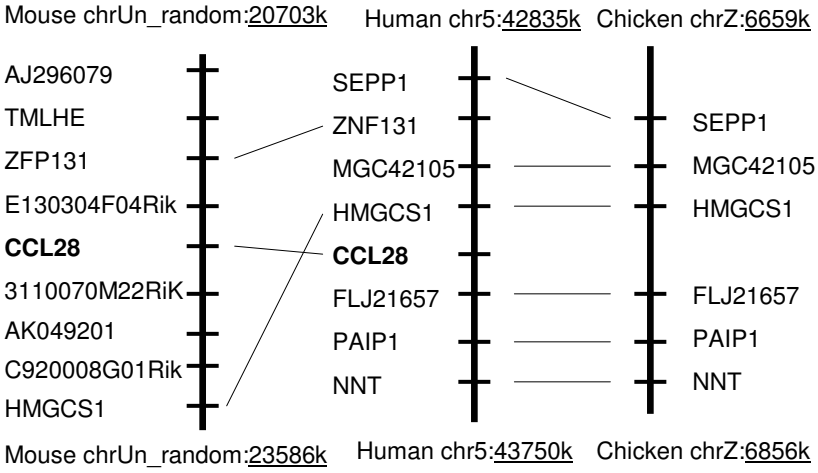


D

Fig. 2.5 (continued)

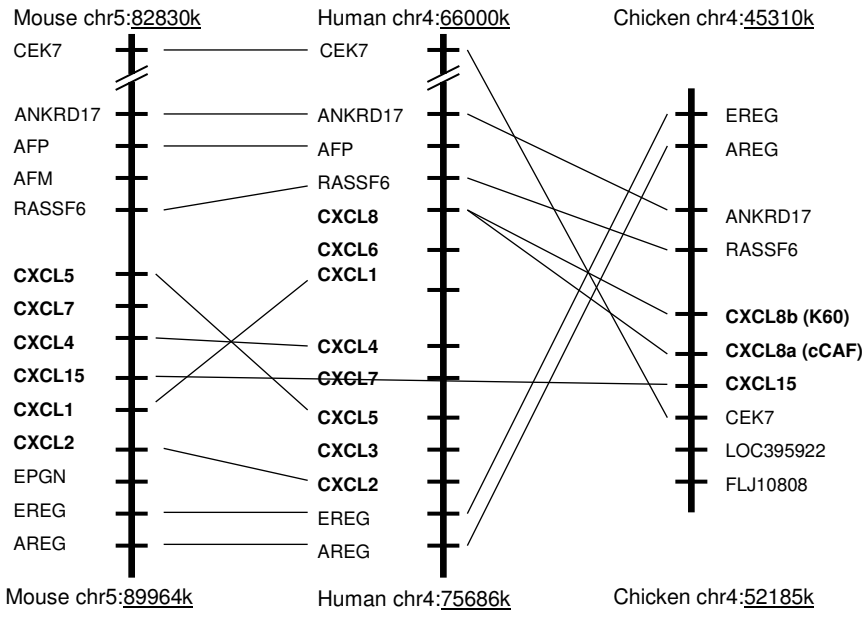


E

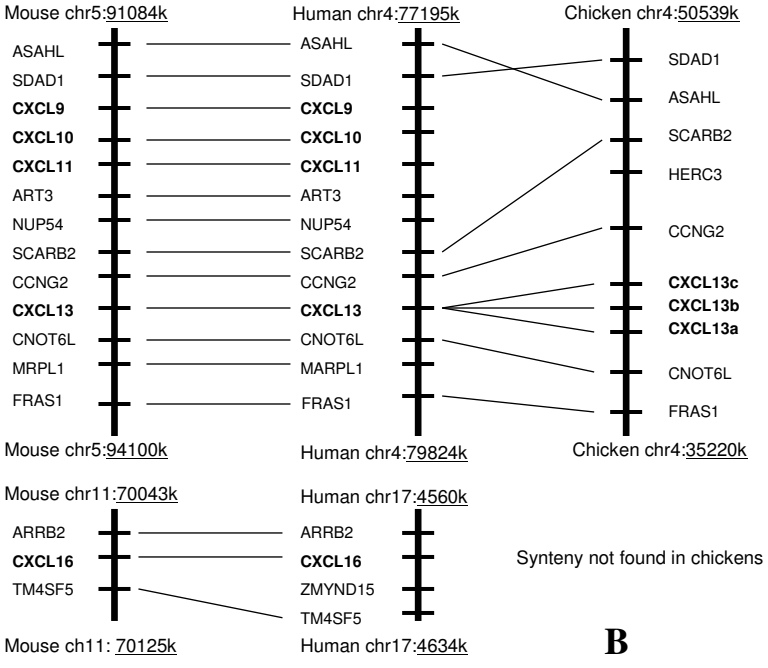


F

Fig. 2.5 (continued)

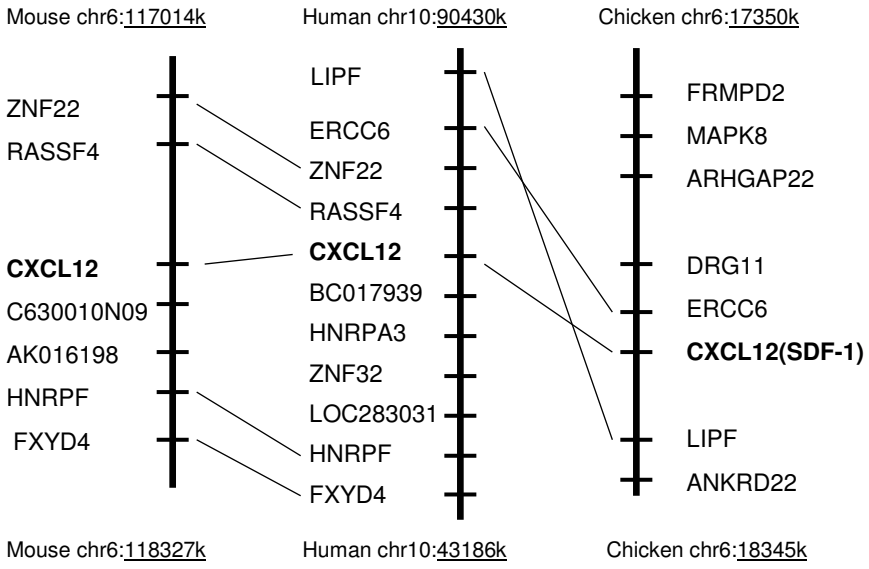


A

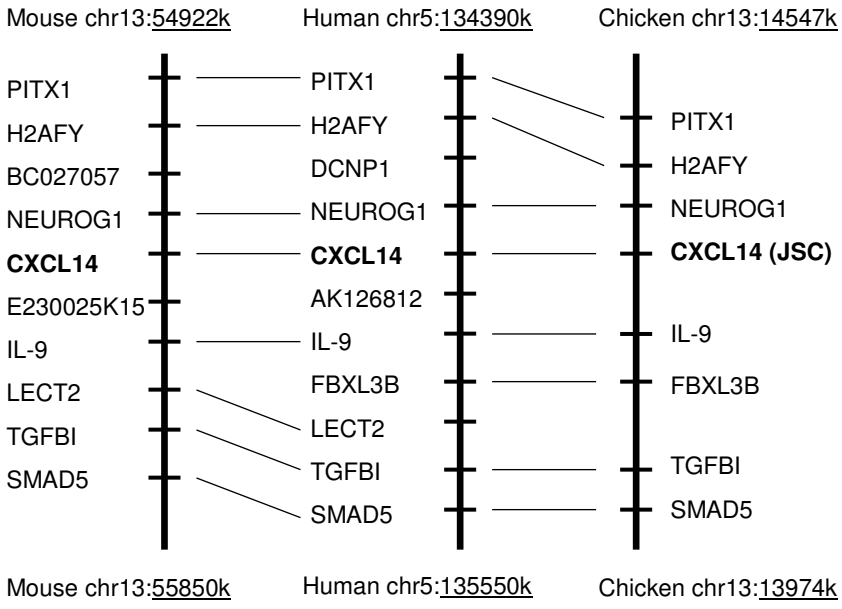


B

Fig. 2.6 Genomic organization (syntenies) of human, mouse, and chicken CXCLs

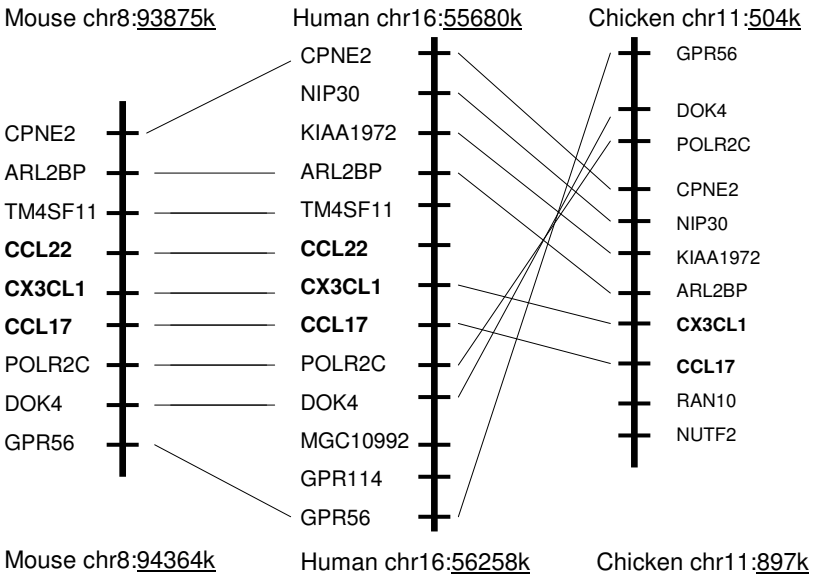


C

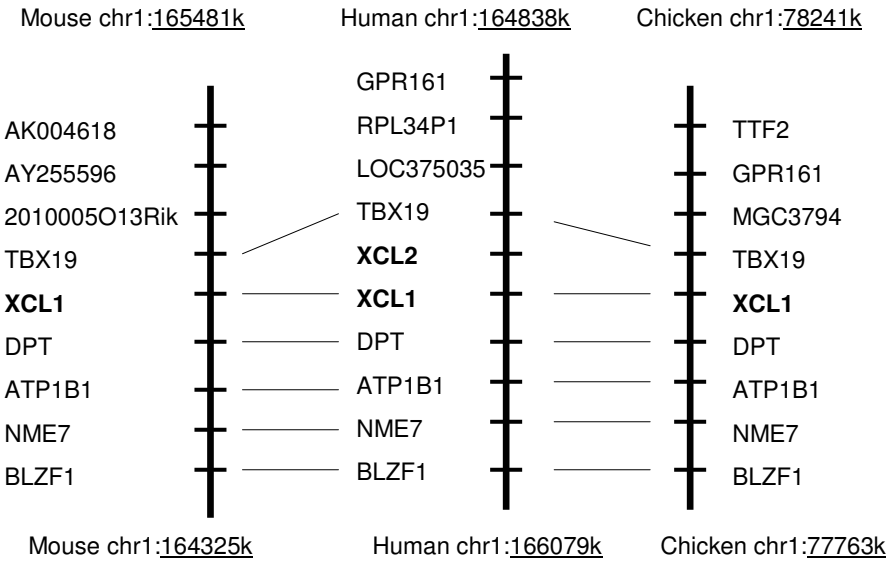


D

Fig. 2.6 (continued)



E



F

Fig. 2.6 (continued)

Chemokine receptor genes were also highly conserved between chicken, human, mouse, and were similarly clustered. The largest cluster of chicken chemokine receptors was found on Chromosome 2, where 5 receptor genes (CCR2, CCR5, CCR8L1, CCR9, and XCR1) were identified. Another cluster on Chromosome 2 contained CCR4, CCR8, and CX_3CR1 genes. Chicken CXCR1 and CXCR2 are also clustered as in mammals, but the chromosome segment is unknown. The remaining CCR and CXCR genes are individually located on Chromosomes 3 (CCR6), 27 (CCR7), 7 (CXCR4), and 24 (CXCR5). Several human chemokine receptors, such as CCR1, CCR3, CCR10, CXCR3, and CXCR6 were not found in the chicken genome, though the syntenies associated with these receptors are present in the chicken genome.

Gene structure

According to the chicken genome sequence, chicken chemokine genes shared the typical three-exon CC and four-exon CXC gene structures with mammals except for CXCL13a and CXCL13b, which have only three exons. Chicken chemokine genes are shorter than the corresponding human genes due to shorter introns in chickens. The gene structure of chemokine receptors was also conserved between chicken and mammals. The EST sequences indicated that chicken chemokine receptor genes could have up to 5 exons though the complete sequences were not available. However, the expressed sequences showed that the amino acid sequences of identified chicken receptors are mostly encoded in a single exon as are most of the mammalian chemokine receptors. Chicken ESTs aligned with the chicken genome sequence indicated that these receptor mRNAs had approximately 2 kb of 5' UTR, as did those found in humans.

Phylogenetic analysis and nomenclatures

The phylogenetic trees (Figures 2.7, 2.8 and 2.9) showed that chicken CCL5, 16, 17, 19, and 20 and all seven CXCLs were closely related to single specific human and/or mouse chemokine orthologs. The phylogenetic trees together with the syntenies associated with these genes (Figs. 2.5 and 2.6) strongly indicate that these genes are the orthologs of those found in mammals; therefore, they are named accordingly. The phylogenetic results show that chickens have two CXCL8 and three CXCL13 genes (only one copy each in mammals), indicating gene duplications in aves.One chicken CXCL related to mouse CXCL15 but not to human CCLs is named as cCXCL15, which is also supported by the synteny of the chemokine cluster (Fig. 2.6A) . Chicken CCL21 is named according to relatedness to the human and mouse and the highly conserved synteny (Fig. 2.5C). According to the phylogenetic tree in Figure 2.8, two directly linked chicken CCLs are remotely related to human and mouse CCL1. The synteny associated with these genes also indicated that they may be CCL-like genes (Fig. 2.5A); therefore, they are named as CCL1L1 and CCL1L2. Three closely related chicken CCLs that are directly linked to the CCL1-like genes are related to a group of clustered human and mouse MCP CCLs (2, 7, 8, 11, and 13) in the phylogenetic tree.

The synteny and phylogenetic tree do not provide information to a specific mammalian ortholog, though these three chicken genes are somewhat more similar to human CCL13 and mouse CCL2. The results indicate that these genes are chicken MCP-like (Fig. 2.5A and 2.8); therefore, they are named ad CCL/MCP like (CCL/MCP-L1, -L2, and -L3). A chicken CCL gene that is directly linked to CCL16 and CCL5 (Fig. 2.5B) is distantly related to chicken CCL5 in the tree. This gene has been reported as MIP-1ß-like chemokine (Petrenko et al. 1995), which is CCL4 in human and mouse. Therefore, it is named as CCL4L1 in order to conform to the report. Another CCL in this cluster that does not display relatedness to other CCLs in the phylogenetic tree (Fig. 2.8) is named as CCL3L1 because this chemokine display highest sequence similarity to a human CCL3-like chemokine, and it shares synteny with human CCL3 (Fig. 2.5B). Overall, CXCLs are more conservative amone chicken, human, and mouse than CCLs.

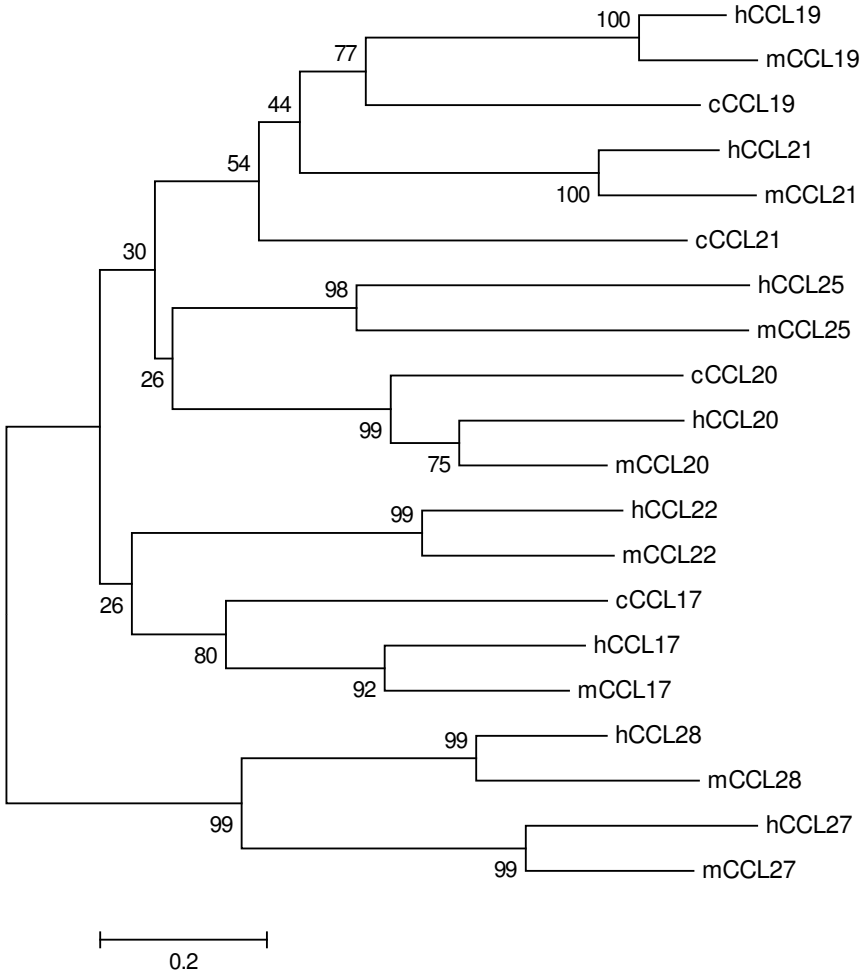


Fig. 2.7 Phylogenetic tree of the chemokine CC subfamily members that are not located on Chromosome 19. The tree was constructed using the amino acid sequences of chicken, human, and mouse chemokines. The numbers on the branches are bootstrap values (percentage that the simulation supports the original interpretation). Human, mouse, and chicken are abbreviated as h, m, and c, respectively. The scale bar reflects the horizontal distance at which amino acid sequences differ by 20% between two sequences

Fig. 2.8 Phylogenetic tree of the chemokine CC subfamily members that are located on Chromosome 19. The tree was constructed using the amino acid sequences of chicken, human, and mouse chemokines. The numbers on the branches are bootstrap values (percentage that the simulation supports the original interpretation). Human, mouse, and chicken are abbreviated as h, m, and c, respectively. The scale bar reflects the horizontal distance at which amino acid sequences differ by 20% between two sequences.

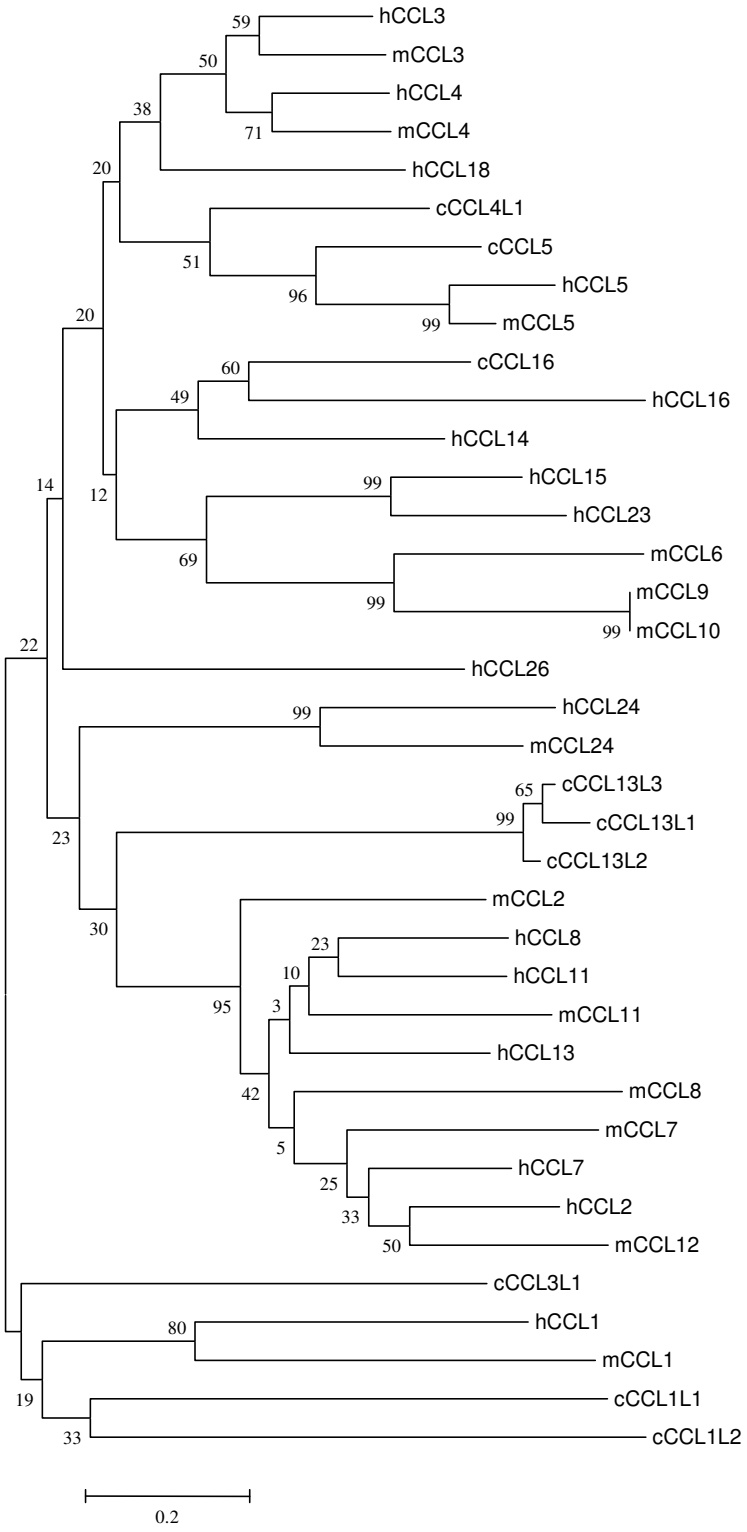
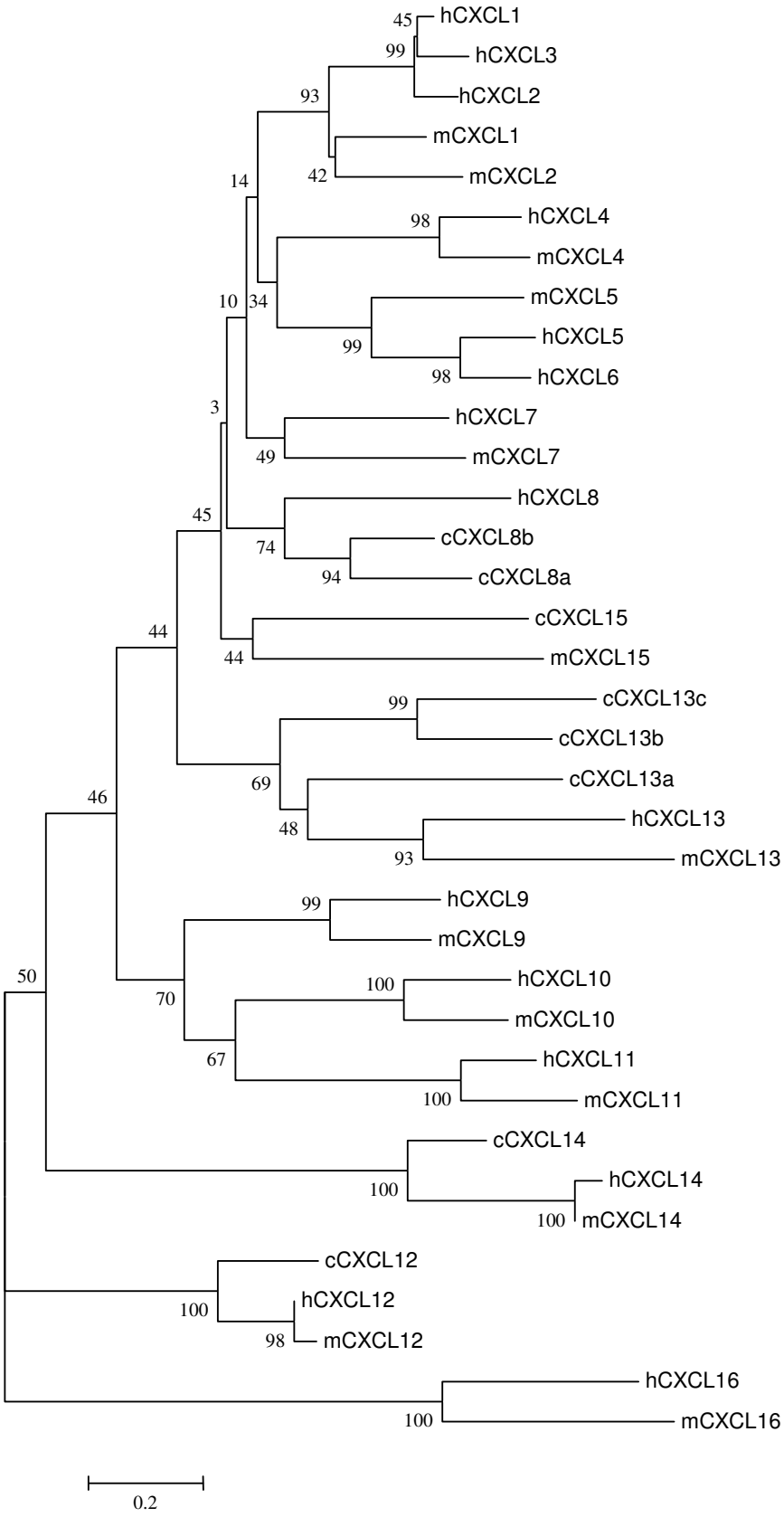

hCCL3
mCCL3
hCCL4
mCCL4
hCCL18
cCCL4L1
cCCL5
hCCL5
mCCL5
cCCL16
hCCL16
hCCL14
hCCL15
hCCL23
mCCL6
mCCL9
mCCL10
hCCL26
hCCL24
mCCL24
cCCL13L3
cCCL13L1
cCCL13L2
mCCL2
hCCL8
hCCL11
mCCL11
hCCL13
mCCL8
mCCL7
hCCL7
hCCL2
mCCL12
cCCL3L1
hCCL1
mCCL1
cCCL1L1
cCCL1L2
59
50
38
71
20
51
96
99
20
60
49
14
12
99
69
99
99
22
99
65
99
23
30
23
10
95
3
42
5
25
33
50
80
19
33
0.2

Fig. 2.9 Phylogenetic trees of the chemokine CXC subfamily constructed using the amino acid sequences of chicken, human, and mouse chemokines. The numbers on the branches are bootstrap values (percentage that the simulation supports the original interpretation). Human, mouse, and chicken are abbreviated as h, m, and c, respectively. The scale bar reflects the horizontal distance at which amino acid sequences differ by 20% between two sequence


hCXCL1
hCXCL3
hCXCL2
mCXCL1
mCXCL2
hCXCL4
mCXCL4
mCXCL5
hCXCL5
hCXCL6
hCXCL7
mCXCL7
hCXCL8
cCXCL8b
cCXCL8a
cCXCL15
mCXCL15
cCXCL13c
cCXCL13b
cCXCL13a
hCXCL13
mCXCL13
hCXCL9
mCXCL9
hCXCL10
mCXCL10
hCXCL11
mCXCL11
cCXCL14
hCXCL14
mCXCL14
cCXCL12
hCXCL12
mCXCL12
hCXCL16
mCXCL16
45
99
93
42
14
98
10
34
99
98
3
49
45
74
94
44
44
99
69
46
48
93
99
50
100
70
67
100
100
100
100
98
100
0.2

Chicken chemokine receptors can also be named according to mammalian nomenclature based on phylogenetic analysis (Fig. 2.10) and syntenies. The genetic distances appear to be shorter between chicken and mammalian chemokine receptors than those between chicken and mammalian chemokines, which is probably due to highly conserved transmembrane domains in these receptors. Chicken CCR4, CCR6, CCR7, CCR9, CXCR2, CXCR4, CXCR5, CX3CR1, and XCR1 were closely related to a mammalian ortholog based on the phylogenetic analysis. There were two distantly related (relative to the distance between mice and human CCR8) CCR8 genes in chickens, the one closer to human CCR8 as CCR8a and the other as CCR8b. There were also two chicken CCRs closely related to human and mouse CCR2 and CCR5, but the phylogenetic analysis could not distinguish them as either CCR2 or CCR5. Because these two receptors were located in a conserved chromosomal region on chicken, human, and mouse chromosomes, these two chicken CCRs were named as CCR2 and CCR5 based on the synteny in which CCR2 is closer to XCR1 than CCR5.

According to the ligand-receptor binding information discovered in humans (all but the receptors of CXCL14 and CXCL15 and the ligand of CCR11 are known) and the systematic nomenclature of chicken genes in the present study, all chicken chemokines and the receptors that recognize the identified chemokines were co-identified in this study except the ligand of CCR9, which is CCL25 in mammals (Table 2.2). These putative ligand-receptor pairings can now be experimentally tested in the lab.

Fig. 2.10 Phylogenetic trees of chemokine receptors constructed using the amino acid sequences of chicken, human, and mouse chemokine receptors. The numbers on the branches are bootstrap values (percentage that the simulation supports the original interpretation). Human, mouse, and chicken chemokines are abbreviated as h, m, and c, respectively, followed by the receptor named. The scale bar reflects the horizontal distance at which amino acid sequences differ by 10% between two sequences.

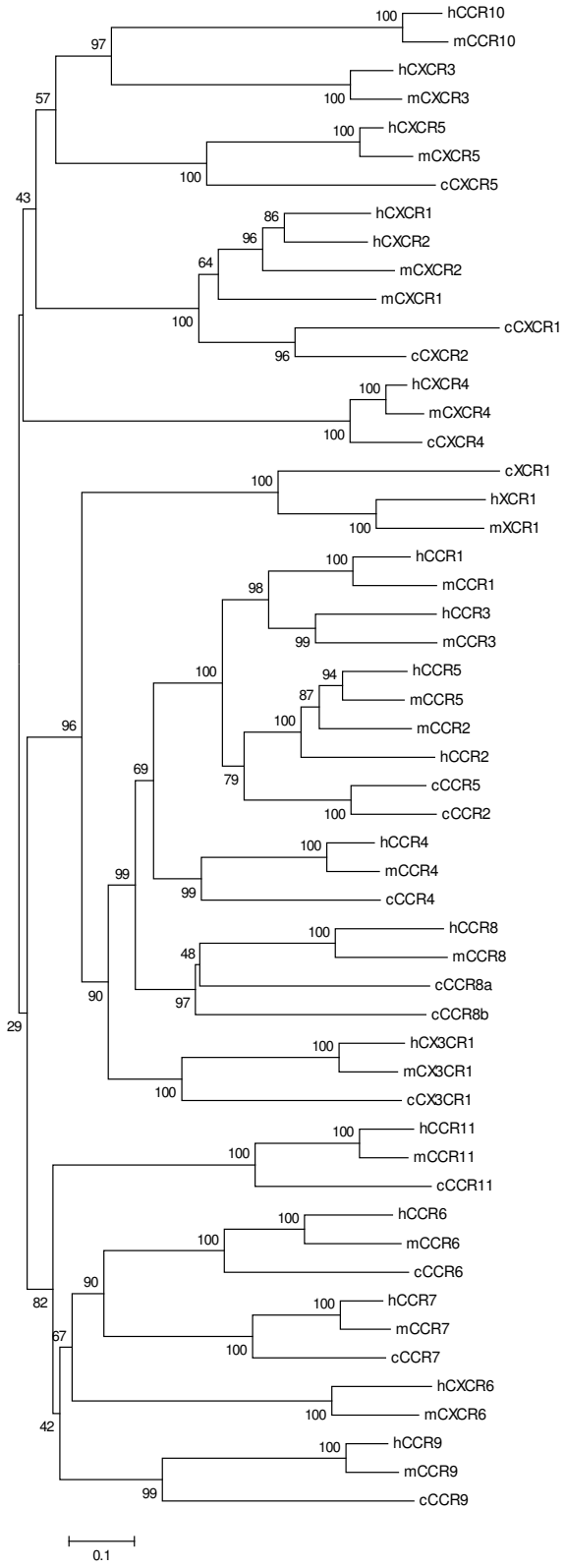

hCCR10
mCCR10
hCXCR3
mCXCR3
hCXCR5
mCXCR5
cCXCR5
hCXCR1
hCXCR2
mCXCR2
mCXCR1
cCXCR1
cCXCR2
hCXCR4
mCXCR4
cCXCR4
cXCR1
hXCR1
mXCR1
hCCR1
mCCR1
hCCR3
mCCR3
hCCR5
mCCR5
mCCR2
hCCR2
cCCR5
cCCR2
hCCR4
mCCR4
cCCR4
hCCR8
mCCR8
cCCR8a
cCCR8b
hCX3CR1
mCX3CR1
cCX3CR1
hCCR11
mCCR11
cCCR11
hCCR6
mCCR6
cCCR6
hCCR7
mCCR7
cCCR7
hCXCR6
mCXCR6
hCCR9
mCCR9
cCCR9
0.1

Table 2.2 Systematic names, chromosomal locations, and putative identified cognate receptors of chicken chemokines[1]

Nomenclature	Chromosomal location (kb)	Ligand name	Putative receptor	Chromosomal location (kb)
CCL1L1	chr19:4,488-4,499	I-309/TCA	CCR8a	chr2:43,465-43,469
CCL1L2	chr19:4,491-4,492		CCR8b	chr2:41,804-41,811
CCL13L1	chr19:4,495-4,495	MCP-?		
CCL13L2	chr19:4,499-4,499	MCP-?	CCR2	chr2:41,765-41,773
CCL13L3	chr19:4,507-4,508	MCP-?		
CCL16	chr19:258-260	HCC		
CCL3L1	chr19:240-241	MIP-1α		
CCL4	chr19:251-252	MIP-1β	CCR5	chr2:41,782-41,790
CCL5	chr19:263-265	RANTES		
CCL17	chr11:768-770	TARC	CCR4	chr2:43,498-43,503
CCL19	chrZ_random:7,805-7,808	MIP-3β	CCR7	chr27_random:665-675
CCL21	chrZ_random:7,810-7,810	SLC		
CCL20	chr9:4,119-4,121	MIP-3α	CCR6	chr3:38,589-38,596
Not found[2]		CCL25	CCR9	chr2:41,880-41,882
Unknown[3]			CCR11	chr2:41,409-41,413
CXCL8b	chr4:51,462-51,465	IL-8	CXCR2	chrUn:136,108-136,109
CXCL8a	chr4:51,475-51,478		CXCR1[4]	chrUn:25,460-25,462
CXCL12	chr6:18,184-18,194	SDF-1	CXCR4	chr7:31,441-31,443
CXCL13a	chr4:35,453-35,454	BCA-?		
CXCL13b	chr4:35,456-35,458	BCA-?	CXCR5	chr24:5,242-5,249
CXCL13c	chr4:35,455-35,456	BCA-?		
CXCL14	chr13:14,231-14,238	BRAK	Unknown[3]	
CXCL15	chr4:51,500,039-51,501,326	Lungkine	Unknown[3]	
CX_3CL1	chr11:758,710-764,541	Fractalkine	CX_3CR1	chr2:43,480,576-43,490,112
XCL1	chr1:780,81,417-78,086,769	Lymphotactin	XCR1	chr2:41,831,759-41,832,766

[1] The systematic naming, ligand naming, and putative receptors are according to Bacon et al. (2002).

[2] CCL25 was not identified in this study.

[3] The information is currently not available in humans or mice.

[4] The ligand and receptor binding has been experimentally tested (Li et al., 2005)

Putative TFBS in novel chicken chemokines

Promoter sequences with score range from 0.80 to 1, with 1 being best, were identified. Most of the promoters in chicken chemokine genes contain the typical TATA box (Table 2.3). In addition, the TATA box in the predicted promoter sequence is located 61 to 156 nucleotides upstream of the first ATG in exon 1 of chicken chemokines. Several transcription factor binding sites including sites for TATA-binding protein, HSF1 (heat shock factor 1), C/EBPβ (CCAAT/enhancer-binding protein β), nuclear factor- κB (NF-κB), GATA motif, STAT2 (signal transducer and activator of transcription 2), IRF1 (interferon regulatory 1), and AP-1 have been identified in the 3000 nucleotides upstream and downstream of the TATA box in the predicted promoter of chicken chemokines.

Analysis of TFBS also indicate that binding sites for NF-κB were mostly located upstream of the TATA box, whereas the binding sites for C/EBPβ and GATA motif were located downstream of the TATA box. The sites for STAT3, STAT4, and IRF1 were present on the both sides of the TATA box.

Table 2.3 The predicted promoter sequence and promoter score of chicken chemokine genes. The chicken chemokine genes are shown in italic and the TATA box in the chicken promoter is shown in bold

Gene	Predicted Promoter	PS
CXCL8-a(cCAF)	CACAAAACTG**TATAA**AAGCACAATGGGTCTCCCTGTTCAGACATCAGCAA	0.97
CXCL8-b(K60)	CAGGACAGTA**TATAA**TCAAGGCAGGTACCAGAGCACAGCCACTGACCACC	0.99
CXCL12	CGCACCAGGAT**TAAA**GGACCGGGGACTGGGGTTAAAGGATCGCGCTGCGA	0.96
CXCL13a	CCGGGTTGGG**TATAAA**GCAGCCTAGCAGTGGCTCGGGGGCAGTGAGGATG	1.00
CXCL13b	GGGCGATGGG**TGAAA**ACCACCCCCAACCCCTGGGCCTCCCATGCGCCGCC	0.88
CXCL13c	ATCTGACTGCAA**TAAAA**AAAACCCCTTGCATGGGGCCTGTGTTCTGATGT	0.99
CXCL14(JSC)	GCGGGAGCGCA**TTTAA**AAGGGGACGAGCGGGGGGGAACGCGGCAGAGCTC	1.00
CXCL15	TACCGAGAGAA**TAAAA**GCAATGGGGACAGCGAGGCGGCACAGCCAGCACG	0.90
CCL/MCP-L1(Ah221)	CAGCTGCGGGA**TAAAA**GTGGCAGTGCAGCAGGGGCTGAGCATTGTGCAGA	0.97
CCL/MCP-L2	CAGCGGCGGGT**TAAAA**GTGGCAGTGCAGCAGGGGCTGAGCATTGTGCAGA	1.00
CCL/MCP-L3	GAGGCTGGGGA**TAAAA**GTGGCAGTGCAGCAGGGGCTGAGCATTGTGCAGA	0.97
CCL1L1	GGCCAGGGGCTA**TATAA**TGCCGAGGTACAGGCAGGTAGCACTCAGCAGCA	0.98
CCL1L2	TGCTGCACCA**TAAAA**GCAGCCCAGCGGTCCCATCCCAAGCAGAAGGGAAA	0.97
CCL3L1	CCTGGTTGC**ATAAA**GAAAGCCCCATCATTCTGGGTAGTGCTGCAAGAAGG	0.96
CCL4L1(MIP-1β)	AGCCACACGCCC**TATAA**ATTGGGCAGTTGCTGAATGTCTCAGCATCACAG	0.99
CCL5(Ah294)	CACCAGCCCGA**TAAAA**GGCGAGGAGCATCCCAGCCACTCCACACCACACC	0.97

Table 2.3 (continued)

CCL16(K203)	GGACTCCCTAT**TAAAA**GGGATCTGCATCCATCTGCTGCAGACCGCCCCGC	0.97
CCL17	CTGGAGAGG**ATAAA**TTGCAGCAGCTCCCTGCCGAGCTAGCAGAGCTCCCT	0.85
CCL19	CTCTATGGTGGA**AATAA**GGGGAGGAGTGTGGGTTATTAAGTGGGATAAAT	0.80
CCL20(Ah189)	CAACTGCTT**TATAA**GCAGGTGGAGAAGCGTTGAGGCTGGCAGTACAGGCA	0.96
CCL21	AAGCGACCGCAG**ATAAA**TTGGAGCAGCTGAAGCTCAGGACAGCAGTGTGT	0.88
CX3CL1	TGGGTCAACC**TAGAA**GAGCCGCTCTGGAGGTGGTGGGAGCTCTTGGATGG	0.80
XCL1(Lymphotactin)	ACTGACAATA**TATAA**ACTCTGTGAAAGGCACCTCACCATAACCACTTTCT	0.99

DISCUSSION

We systematically searched for all chicken chemokine and chemokine receptor genes in the recently available draft chicken genome. Without this information, it may have taken years to find chicken chemokines and their receptors. In order to find additional chicken chemokines after using the comparative genomic approach, I conducted low stringency BLAST against all chicken ESTs. No additional chemokines were identified, which indicated that the comparative genomics approach based on syntenies was very effective in finding chicken chemokines. The independent nomenclature of chicken chemokines and their cognate receptors, and mammalian chemokine-receptor binding information suggest that most of the genes have been identified. The sole exception was CCL25, the only known ligand of CCR9 in mammals, which was not found in this study, although its receptor was identified. Likewise, CXCL14, and CXCL15 were identified in both chickens and mammals, but their receptors are unknown. Therefore, it is very likely that there are additional chicken chemokine and chemokine receptor genes in the chicken genome and that they may lie in genomic regions that lack sufficient sequence coverage.

Although most of the systematic nomenclature of the chicken genes was unambiguous based on both of phylogenetic trees and syntenies, the information that was used to name seven chicken CCLs as CCL1L1, CCL1L2, CCL3L1, CCL4L1, CCL/MCP-L1, CCL/MCP-L2, and CCL/MCP-L3 and to distinguish two chicken chemokine receptors into CCR2 and CCR5 may have been inadequate. CCR2 and CCR5 are closely related and tightly linked in the human, mouse, and chicken genomes. The

phylogenetic analysis indicates these genes were duplicated after the divergence between mammals and aves. Further lab testing of ligand binding will make them more distinguishable in terms of biological functions. Chicken CCL1/MCP-L1, -L2, and -L3 were related to a group of clustered mouse and human MCP CCLs; therefore, specific cognate receptors must be tested to distinguish them. In humans, the chemokines of this MCP group and MIPs, such as CCL3, CCL4, and CCL5 can bind to more than one receptor, such as CCR1, 2, 3, and/or 5, but not both, such as CCR2 for MCPs and CCR5 for MIPs. Therefore, chicken CCR2 and CCR5 may be two receptors that recognize these two groups of CCL chemokines. Interestingly, two CCL1 like (CCL1L1 and CCL1L2) and two CCL1 receptor (CCR8a and CCR8b) genes were found in the chicken genome. The ligand-receptor binding for these four genes can not be determined in this study. Nerveless, the names assigned based on comparative analysis in this study may prove useful in order to apply the functional and physiological knowledge from other species to chickens. Further lab testing must be carried out to confirm the ligand and receptor binding and to understand their biological functions.

Chicken chemokine ESTs were highly represented in the EST database. There were several ESTs aligned to each identified chicken chemokine gene in the UCSC Genome Browser. The sequences assembled from ESTs probably contained most, if not all, of the full-length chemokine mRNA sequences. Promoter sequences with a typical TATAA were detected with promoter prediction software (Table 2.3). However, there were only a few ESTs that partially cover chicken chemokine receptor genes. Some of these ESTs contained translation start sites. These EST sequences and reported

complete coding sequences indicate that the amino acid sequences of chicken chemokine receptors were mostly encoded in one exon. The predicted amino acid sequences were of the expected length and aligned very well with the coding sequences of non-chicken reference genes in the UCSC Genome Browser. The conserved gene structure of this receptor family and high sequence similarity between chicken and mammals suggested that the predicted coding sequences were very accurate, especially for those with ESTs containing translation start sites. CCR4 was the only predicted gene that did not have a matching EST and CXCR2 was the only identified gene with partial sequence. Further study including sequencing expressed sequences is needed to confirm these genes.

CHAPTER III

A NOVEL TOLL-LIKE RECEPTOR HOMOLOGOUS TO TOLL-LIKE RECEPTOR 21 IN FISH

INTRODUCTION

TLRs are a family of type I transmembrane receptors with an amino- terminal extracellular domains and a carboxy-terminal cytoplasmic region called Toll/Interleukin-1 receptor (TIR) domains (Bell et al. 2003) (Fig. 3.1). The extracellular region contains 19-25 tandem copies of leucine-rich repeat (LRR) motifs, which are involved in ligand binding. The typical LRR consensus sequence for the TLRs is arranged in a 24-residue repeat motif, consisting of $XL^{2}XXL^{5}XL^{7}XXN^{10}XL^{12}XXL^{15}XXXXF^{20}XXL^{23}X$, where X can be any amino acid; and the leucine residues, at positions 2,5,7,12,15, and 23 are occasionally replaced by other hydrophobic amino acids (Buchanan and Gay 1996). F at position 20 is a conserved phenylalanine; and N is asparagine and is often replaced by a cysteine or threonine.

The intracellular TIR domains are mostly conserved in TLRs and mediate protein-protein interactions between the TLRs and different adapter proteins such as MyD88, resulting in the activation of nuclear factor κB and mitogen-activated protein kinase signal transduction pathways (Barton and Medzhitov, 2003). TLR family members can be divided into five subfamilies based on the comparison of the amino acid sequences: the TLR2, 3, 4, 5, and 9 subfamilies (Takeda and Akira 2003). The TLR2

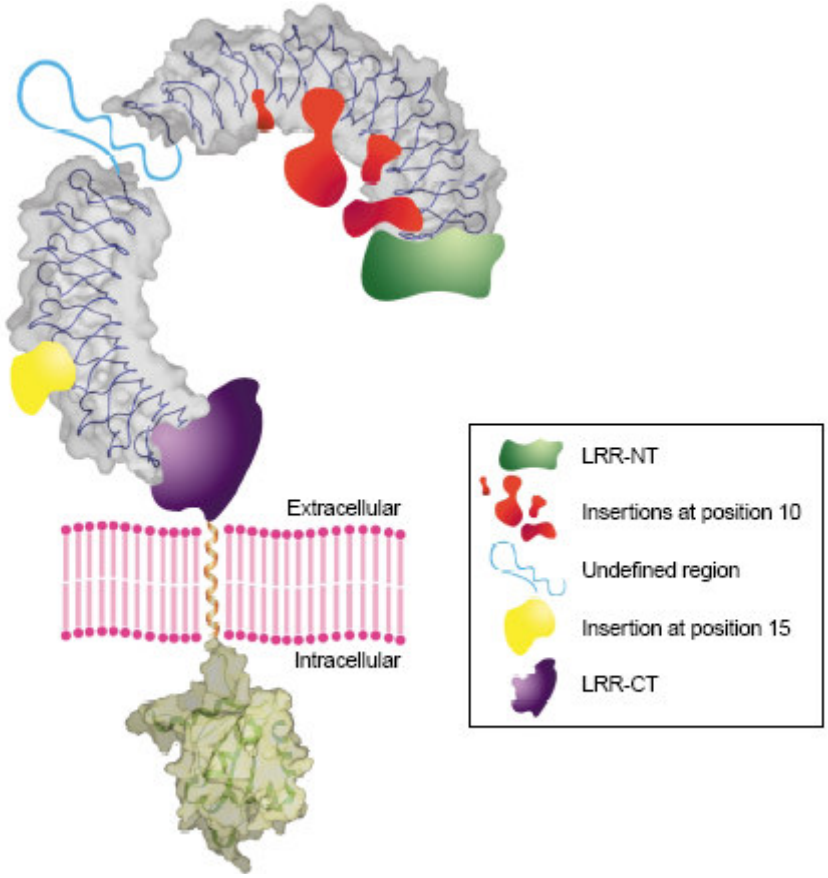


Fig. 3.1 A Toll-like receptor (TLR). All TLRs are integral membrane glycoproteins with an N-terminal ectodomain and a single transmembrane domain. The ectodomain of a TLR7, TLR8 and TLR9 family member is depicted, with the LRR solenoid shown with a gray molecular surface, and the N- and C- terminal flanking regions shown in green and purple, respectively. An undefined region present in TLR7, TLR8 and TLR9 but not in the other TLRs, is shown as a light-blue string. Insertions within LRRs at position 10 are indicated in red and might contribute to the formation of the pathogen-associated molecular pattern (PAMP) binding site. An insert at position 15 is indicated in yellow, and is expected to originate on the convex face of the TLR. The TIR domain structure was taken from pdb 1FYV (Bell et al. 2003)

subfamily is composed of TLR1, 2, 6, and 10; the TLR9 subfamily is composed of TLR7, 8, and 9.

In the chicken, nine different TLRs (chTLR) have been identified and characterized (Fukui et al. 2001; Leveque et al. 2003; Philbin et al. 2005; Yilmaz et al. 2005; Higgs et al. 2006) and a non-functional TLR8 (Iqbal et al. 2005). Two types of chicken TLR2 were cloned and both genes were mapped to the same chromosomal segment on chromosome 4, suggesting that they arose by gene duplication (Fukui et al. 2001). Allelic variation in chTLR4 was associated with resistance to infection with *Salmonella enterica* Serovar Typhimurium in chickens (Leveque et al. 2003). Predictions of the TLR family in the pufferfish (*Fugu* rubripes) and zebrafish (*Danio* rerio) genomes were recently reported (Oshiumi et al. 2003; Jault et al. 2004; Meijer et al. 2004), showing that the counterparts of most of known mammalian TLR genes were identified in the pufferfish genome with the exception of TLR4 and TLR 6, while two novel TLRs, TLR21 and TLR22, were found (Oshiumi et al. 2003). Analysis of the zebrafish genome resulted in the discovery of the orthologues of human TLR genes with a homologue of TLR4 in addition to TLR21 and TLR22 (Jault et al. 2004; Meijer et al. 2004). The human and mouse genomes do not contain TLR21 and TLR22 genes. The newly available chicken draft genome sequence and a large number ESTs has allowed systematic identification and annotation of novel chicken TLR genes. The objective of this study was to test whether TLR21 and TLR22 genes exist in the chicken using bioinformatics, to compare their predicted structures with fish TLR homologues, analyze their expression patterns and ligand inference based on comparative genomics .

MATERIALS AND METHODS

Identification of a novel toll-like receptor

BLAST searches were carried out with TBLASTN program using fugu TLR21 and TLR22 sequences as queries at the National Center for Biotechnology Information (NCBI) database (Altschul et al. 1997). Expressed Sequence Tags (ESTs) and chicken mRNA sequences in the corresponding chromosomal regions were then identified and, if necessary, assembled with the CAP3 program. These sequences were aligned with the corresponding chicken genomic sequence and any deletions or insertions corrected. Sequences were then submitted to ORF Finder (Open Reading Frame Finder). Also, the chicken draft genome sequence released on March 1, 2004 by the National Institutes of Health (http://mgc.ucsc.edu/cgi-bin/hgGateway) was used for BLAT searches using the following queries; Fugu TLR21 (BAC66138), Fugu TLR22 (AAW69372), zebrafish TLR21 (AY389459) and zebrafish TLR22 (AY389460) respectively.

The retrieved ESTs were used to design primers with Primer Express 2.0 (Applied Biosystems, Foster City, CA) to amplify cDNA fragments overlapping full coding sequences of the TLR with three PCR reactions. PCR was performed as follows: An initial denaturation step at 94^0C for 2 min and 35 cycles of denaturation, annealing, and extension at 94^0C for 30 sec, 59^0C for 45 sec, and 72^0C for 1 min, and a final extension step was carried out at 72^0C for 10 min. Unincorporated nucleotides were removed from amplified PCR products using BioMax spin-50 mini-columns (Millipore, Billerica, MA). BigDye terminator cycle sequencing ready reaction kits and an ABI

Prism 377XL DNA Sequencer (Applied Biosystems) were used for DNA sequencing in the Gene Technologies Laboratory at Texas A&M University.

Sequence analysis

The amino acid sequence of the putative chicken TLR21 was predicted based on the open reading frame of the expressed nucleotide sequences (ESTs or mRNAs). The initiation codon was predicted using the ATGpr program (Salamov et al. 1998). The protein sequences were used as queries with BLASTP to search against the non-redundant protein database in Genbank for homologous hits. Complete amino acid sequences of currently known human, mouse and chicken TLRs were retrieved from Genbank for comparison and sequence analysis. The functional domain structures were predicted using the SMART 4.0 program (http://194.94.45.211/) (Ponting et al, 1999) and the hydrophobic profile analysis for predicting transmembrane domain (TM) was performed by the program of the Weizmann Institute of Science (http://bioinformatics.weizmann.ac.il/hydroph/plot_hydroph.html). Signal peptide sequences were predicted using the Signal 3.0 program (http://www.cbs.dtu.dk/services/SignalP/) (Bendtsen et al. 2004) with the settings 'eukaryotes' using both the neural network and hidden Markov model methods. The alignment of TLR21 amino acid sequences was performed using the MultiAlin program (Corpet, 1988).

Phylogenetic trees were constructed by the MEGA program (Kumar et al. 2004) using the neighbor-joining method based on the alignment of the amino acid sequences

of the extracellular LRR domains and intracellular TIR domains of TLRs individually, predicted by SMART software.

The sequence data used for the phylogenetic tree construction with accession numbers of human, mouse, chicken, Fugu and zebrafish are: human toll like receptors hTLR1 (GenBank accession number: NM_003263), hTLR2(NM_003264), hTLR3(NM_003265), hTLR4(U88880), hTLR5(NM_003268), hTLR6(NM_006068), hTLR7(NM_016562), hTLR8(AF245703), hTLR9(AF245704), hTLR10(AF296673), Mouse toll like receptors from mTLR1 through 13 were mTLR1 (NM_030682), mTLR2 (NM_011905), mTLR3 (NM_126166), mTLR4 (NM_021297), mTLR5 (AF186107), mTLR6 (AF314636), mTLR7 (NM_133211), mTLR8 (NM_133212), mTLR9 (NM_031178), mTLR11 (NM_205819), mTLR12 (NM_205823), mTLR13 (NM_205820). Chicken toll like receptors were chTLR1 type 1 (NM_001007488), chTLR1 type 2 (AY633573), chTLR2 type 1(AB050005), chTLR2 type 2 (AB046533), chTLR3 (AY633575), chTLR4 (AY064697), chTLR5 (NM_001024586), chTLR7 (NM_001011688). Fugu and zebrafish toll like receptors with accession numbers were Fugu TLR21 (BAC66138), Fugu TLR22 (AAW69372), zTLR2 (NM_212812), zTLR3 (NM_001013269), zTLR4b (NM_212813), zTLR21 (AY389459) and zTLR22 (AY389460).

The MEGA program constructs the phylogenetic trees with the options of bootstrap test, pairwise deletion, and poisson correction. The topological stability of the neighbor-joining trees was evaluated with 1000 bootstrapping replications.

Tissue expression profiles

Total RNA was isolated from whole blood cells, kidney, liver, lung, oviduct, small intestine, large intestine, and spleen of a laying White Leghorn hen *(Gallus domesticus)* and extracted using TRIZOL reagent (Invitrogen, Carlsbad, CA) according to the manufacturer's instructions. White blood cells were isolated from the blood samples using Histopaq 1.099 (Sigma, St Louis, MO). RNA samples were quantified with a UV spectrophotometer (Brinkmann Instruments, Westbury, NY) and the quality of the RNA assessed by agarose gel electrophoresis. Genomic DNA was removed by incubation with RNase-free DNase I (Invitrogen). Each tissue was reverse transcribed into cDNA by random hexamer priming using the ThermoScript kit (Invitrogen). The primer sequences used for RT-PCR were designed based on the cDNA sequence of the predicted chTLR21 using Primer Express 2.0 program (Applied Biosystems) with forward (5'- GAGCTGCAGACGCTGGATTTA -3') and reverse (5'-CCACTGTGAGATGTACTGCAGGTT -3') primers. Approximately 50 ng of the first-strand cDNA was amplified in a 10 µL PCR reaction that contained 1 µL 10 X Taq reaction buffer (500 mM KCl, 100 mM Tris-HCl, and 15 mM Mg $(OAC)_2$, pH 8.3), 0.2 µM of each of the forward and reverse primers (Table 1), 0.2 mM of dNTP mix, and 0.3 U Taq DNA polymerase (New England Biolabs, Beverly, MA). PCR was carried out using a thermocycler (MJ research, Watertown, MA). For an internal marker, RT-PCR for chicken cytoplasmic β-actin (X00182) was carried out under the same conditions. The primers for β-actin were 5' -CTGATGGTCAGGTCATCACCATT- 3'(sense) and 5'-TACCCAAGAAAGATGGCTGGA-3' (antisense). PCR conditions described to amplify

expressed TLR21 sequence (5'-UTR and 3'-UTR) were also used to detect tissue expression. PCR products were analyzed by 1.5 % agarose gel electrophoresis and stained with ethidium bromide. The PCR fragments were directly sequenced to confirm the identity of the amplified product.

Ligand inference of *in silico* analysis

A bioinformatic approach was used to predict possible ligands of the new chicken Toll-like receptor. Several methods of identifying conserved domains and functional sites from protein sequence were used. Leucine-rich repeat (LRR) motifs were predicted with the SMART program with the default parameters (http://smart.embl-heidelberg.de/). The pathogen-binding ectodomains were used as queries to search similar known motifs using the PROSITE program (http://expasy.nhri.org.tw/prosite/) against the protein pattern database. The BLOCKS program (http://blocks.fhcrc.org/blocks/blocks_search.html) was used against the Blocks database. MotifFinder program searches were performed against Motif Libraries (http://motif.genome.jp/). BLASTP searches were also carried out against the non-redundant protein database in Genbank to find homologs.

ChTLR21 protein subcellular localization prediction

LOCtree (http://cubic.bioc.columbia.edu/cgi/var/nair/loctree/), PSORT II (http://psort.nibb.ac.jp), TMHMM (http://www.cbs.dtu.dk/services/TMHMM/) and SignalP program (http://www.cbs.dtu.dk/services/SignalP/) with different algorithms for detecting sorting signals were exploited to determine chTLR21 protein subcellular localization.

RESULTS

Novel chicken putative TLR21

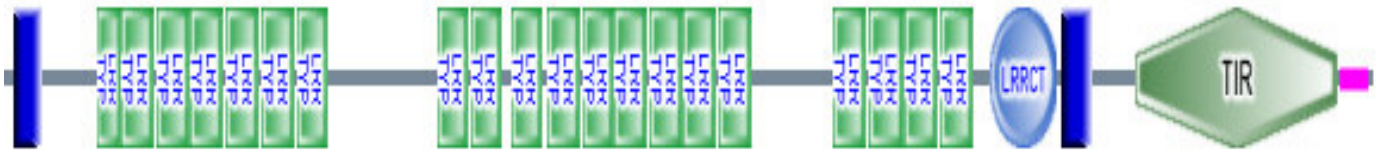
Fish TLR21 and TLR22 were retrieved from Genbank to find orthologs in chicken using fugu and zebrafish TLR21 and TLR22 protein sequences. Blast analysis of the chicken genome database identified one mRNA sequence that encoded protein with homology to both of fugu TLR21 and TLR22 but no clone similar only to TLR 22 was obtained. The newly identified gene was named as chTLR21 (GenBank accession number: AJ720600) based on comparison of the identities of protein sequences and phylogenetic analysis. The amino acid sequence of the putative chicken TLR21 was predicted based on the open reading frame of the expressed nucleotide sequences. The amino acid sequence of the putative chTLR 21 is highly homologous to sequences of fugu TLR21 and fugu TLR22. The partial 5′UTR and 3′UTR of chicken TLR21 was then cloned (GeneBank Accession number: DQ158908 and DQ198090) and sequenced and shown to extend the 5′ end by 393 nt and the 3′ end by 150 nt of the published cDNA (Caldwell et al., 2005). The presence of an upstream in-frame terminator codon in the ORF of the chicken TLR21 gene, one ATG codon downstream from the stop codon in the 5′ -upstream region and the existence of a signal peptide sequence showed that this ATG is a good candidate for the translation start codon. The program for predicting the initiation codon (http://www.hri.co.jp/atgpr/) also indicated that this ATG is the start codon. Based on these results, we conclude that the chicken TLR21 ORF consists of 3346bp that encodes 959 amino acids. Analysis of the 5′-upstream regulatory

region of chicken TLR21 gene demonstrated that a putative TATA box (TATAA) is located at 892bp upstream from the putative translation initiation site (data not shown).

Sequence homology revealed that the putative chicken TLR21 protein was 43% identical to Fugu TLR21, and had 59% similarity with a partial sequence of zebrafish TLR21. In addition, the degree of homology was higher in the trans-membrane and cytoplasmic domains than the extracellular domain. The homology in the extracellular domain was 40% whereas the homology in the cytoplasmic domain was 53% between chicken and Fugu TLR21.

Sequence analyses

The deduced protein sequence of the identified chicken TLR21 gene had typical TLR structure (Fig. 3.2) and similar domain arrangement as Fugu TLR21. Like all reported mammalian TLR molecules, chicken TLR21 had several LRRs in the N-terminal region, a transmembrane domain and the TIR domain at the C-terminal end predicted using SMART program. There were eight more LRRs in Chicken TLR21 than Fugu TLR21. In addition, chicken TLR21 was composed of the most populated typical motifs whereas the LRRs in Fugu 21 belong to outlier motifs.



ChTLR21

Fig. 3.2 Secondary structures of chicken TLR21 polypeptide. The vertical dark blue bars represent trans-membrane domains. Pink bars represent simple sequences. Abbreviations used: TIR: Toll/Interleukin-1 receptor, LRR: outlier type leucine-rich repeats, LRR_TYP: leucine-rich repeats (typical and most populated), LRRCT: leucine-rich repeat at the C-terminal domain

The complete chicken TLR21 amino acid sequence was aligned with currently available TLR21 in other species using the Clustal W algorithm (Fig. 3.3). In this study, we also considered zebrafish TLR21 which is incomplete and contains only TIR domain. The transmembrane domain was 23 amino acids in length in all of TLRs. The TIR domain was similar in length but fugu TLR21 was 7 amino acids longer than chicken TLR21 while zebrafish TLR21 was 11 shorter respectively. Alignment of amino acid residues also suggests that the extracelullar regions are less conserved than the rest of the domains and the residues in TIR domain were highly conserved among all species (Fig. 3.3).

As described previously in zebrafish and higher vertebrates (Jault et al, 2004; O'Neill et al., 2003), the three main conserved amino acid motif were also located in the chicken TLR 21 core TIR structure. The first one seemed to conserve the described (F/Y) DA consensus motif. The second one contained three conserved amino acid residues, namely R, D and G, reported to constitute a loop in the TIR domain of human TLR2.

The last box is characterized by the conserved FW motif, which is found in the last α-helix of TIR of human TLR2. Examples of other such conserved amino acids include 13 residues in chicken TLR21 (AVYNSRKTVCVVS) that differ by 2 amino acids from Fugu TLR21. In addition, VLLRRTYLRWP (11 residues) in chTLR21 differ only by 4 amino acids from fugu or zebrafish TLR21.

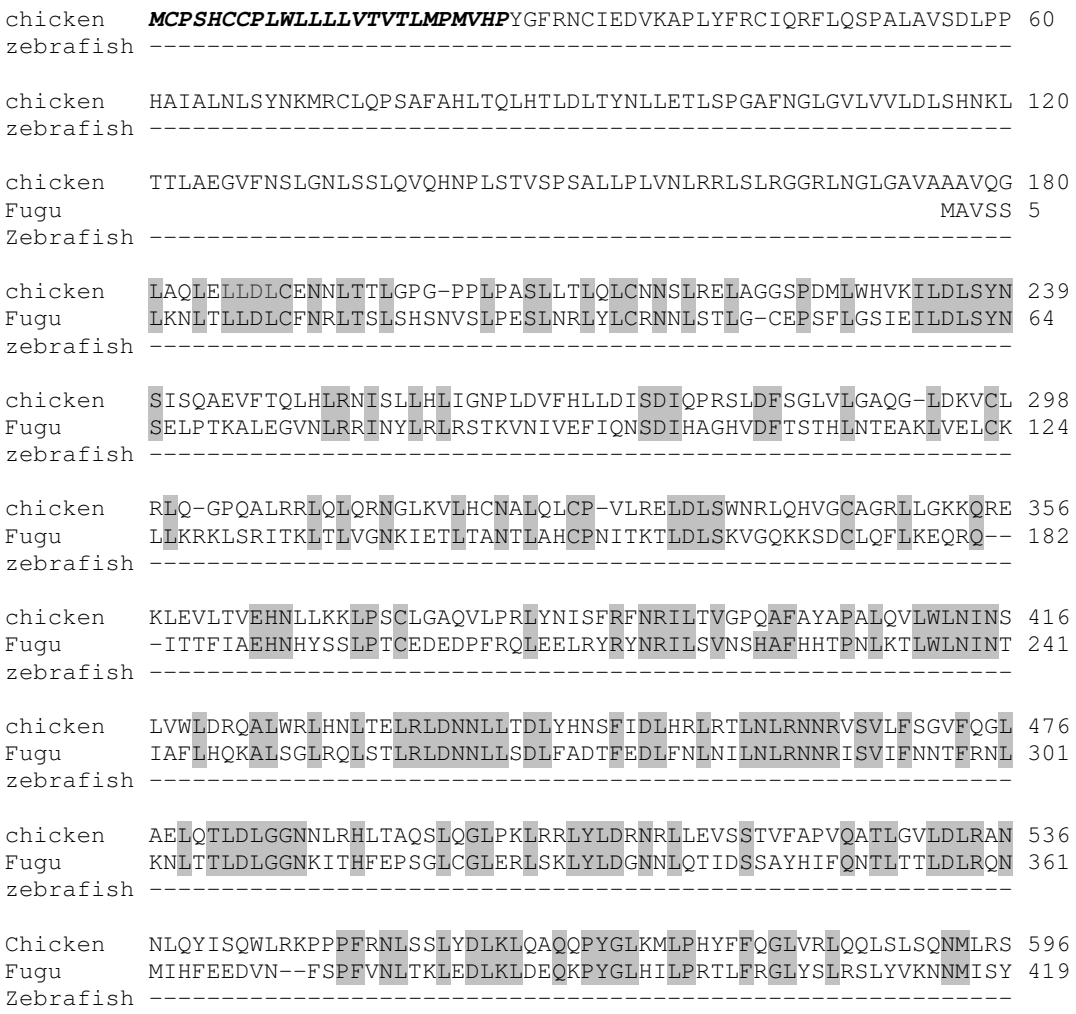


Fig. 3.3 Alignment of amino acid sequences of chicken TLR21, Fugu TLR21, and zebrafish TLR21. Alignment gaps are indicated by dashes. Sequences with identical amino acid in at least 50% of the TLR21 are highlighted in gray. Sequences identical in all species are highlighted in dark gray. The predicted transmembrane and Toll/Interleukin-1 receptor (TIR) domains are underlined and double underlined, respectively. The signal peptide cleavage site was predicted using SignalP 3.0 (http://www.cbs.dtu.dk/services/SignalP/) and the signal peptide is bold and italic

```
chicken   IPPDVFEDLGQLRSLALADSSNGLHDLPDGIFRNLGNLRFLDLENAGLHSLTLEVFGNLS 656
Fugu      LAADVFRDLKHLNFLSLDNCCVGPTHLPAGIFKDLTNLTILTVENMGIQNLSTEVFGNIS 479
zebrafish ------------------------------------------------------------

chicken   RLQVLHLARNELKTFNDSVASRLSSLRYLDLRKCPLSCTCDNMWLQGW-LNNSRVQVVYP 715
Fugu      QLKKIQLNHNVMQTFPVTVLQSLTKLQYLDIRNVPLSCTCENSLLRNWTVNNQKVQMIYL 539
zebrafish ------------------------------------------------------------


chicken   YNYTCGSQHNAYIHSFDTHVCFLDLGLYLFAGTAPAVLLLLVVPVVYHRAYWRLKYHWYL 775
Fugu      YSLPCPHDPKVKFFNFDTSVCNIDLGQYLFFCTAPWIFLFTVWPLLYVKLYWKIKYSYYV 599
zebrafish ------------------------------------------------------------

chicken   LRCWVNQRWRRE---EKCYLYDSFVSYNSADESWVLQKLVPELEHG--AFRLCLHHRDFQ 830
Fugu      FRSWFSEQWRRLREKEENCKYDAFISYNSSDELWVMNELLPNLEGNGSSFKICLHHRDFE 659
Zebrafish -----GEQWRRLRDQEEKYNYDAFVSYNSADEDWVMEQLLPNLEG--SSFRLCLHHRDFE 53

chicken   PGRSIIDNIVDAVYNSRKTVCVVSRSYLRSEWCSLEVQLASYRLLDERRDILVLVLLEDV 890
Fugu      PGRYIIDNIVSAVYSSRKTICVVSRNFLSSEWCSLEIQLASYRLFDEHRDVLLLVLLEPI 719
Zebrafish LGRDIVDNIVAAVYGSRKTICVVSQSFLRSEWCSLEIQLASYRLFQEMQDVLLLVFLEPI 113

chicken   GDAELSAYHRMRRVLLRRTYLRWP----LDP-AAQPLFWARLKRALRWVEGGEEEEEEGL 945
Fugu      SERQLSSYHRMRKVMLKKTYLQWPGSECTNPLQAQGLFWSQLRRAIGTTSRIETEEKGTR 779
zebrafish PERQLSAYHRMRKVMLKKTYLQWPGSNCSDPNSAKE---------------------- 149

chicken   GGGTGRPREGDKQM 959
Fugu      VANKEDADASDNHV 793
zebrafish --------------
```

Fig. 3.3 (continued)

Phylogenetic trees were constructed based on multiple sequence alignments of the amino acid sequences of the conserved LRR and TIR domains individually (Figs. 3.4 and 3.5) among mammal and other species using the neighbor-joining method (Saitou and Nei, 1987) within the Molecular Evolutionary Genetics Analysis [MEGA(3.0)] package (Kumar et al. 2004). Data was analyzed using poisson correction. As shown in figures 3.4 and 3.5, phylogenetic trees drawn based on the amino acid sequences of the N-terminal LRR and C-ternimal TIR are similar. TLRs can be grouped into five subfamilies of TLR2, 3, 4, 5, and 9 in both phylogenetic trees. The naming of the chicken TLR21 was further supported by the phylogenetic trees. The phylogenetic trees also indicate that TIR domains from TLR22 and 21 have a common ancestor in fish and Fugu before separation. In addition, it can be found that the TIR domains of TLR21 genes form a group with TLR22. This branch is only present in fish and chickens. The main difference between these two phylogenetic trees is that hTLR1 and hTLR6 are more closely related to mTLR1 and mTLR6 in the tree constructed based on the extracellular LRR domains.

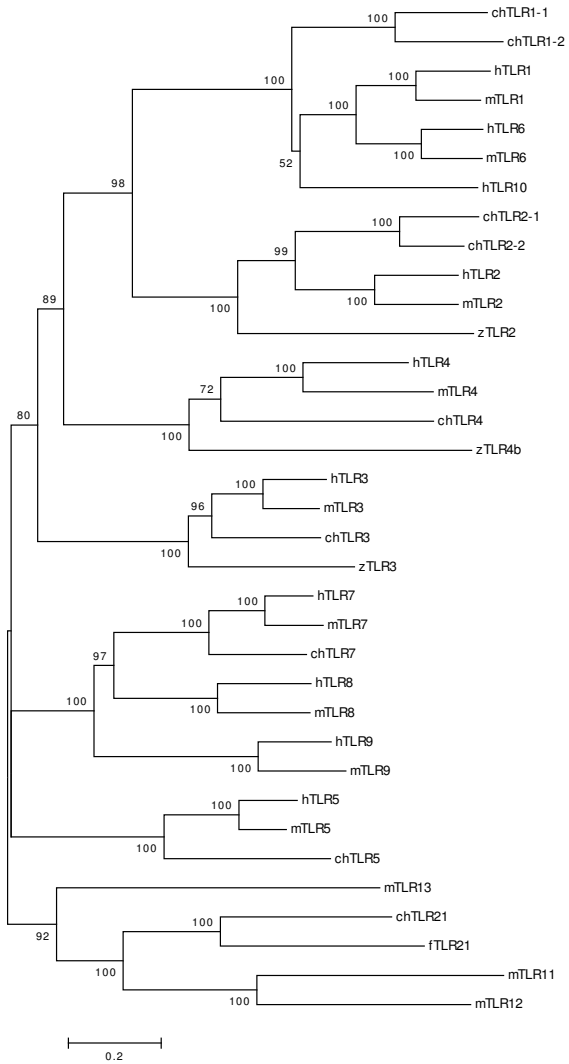


Fig. 3.4 Phylogenetic tree constructed based on the amino acid sequences of amino-terminal leucine-rich repeats (LRR) domains of chicken, human, mouse, pufferfish and zebrafish Toll-like receptors. Data was analyzed using Poisson correction and gaps were removed by complete deletion. The bootstrapping values are indicated by the numbers at the nodes. Human, mouse, and chicken Toll-like receptors are abbreviated as hTLR, mTLR, and chTLR, respectively. ChTLR1-1, chTLR1-2, chTLR2-1, and chTLR2-2 represent chTLR1 type 1, chTLR1 type 2, chTLR2 type 1, and chTLR2 type 2, respectively

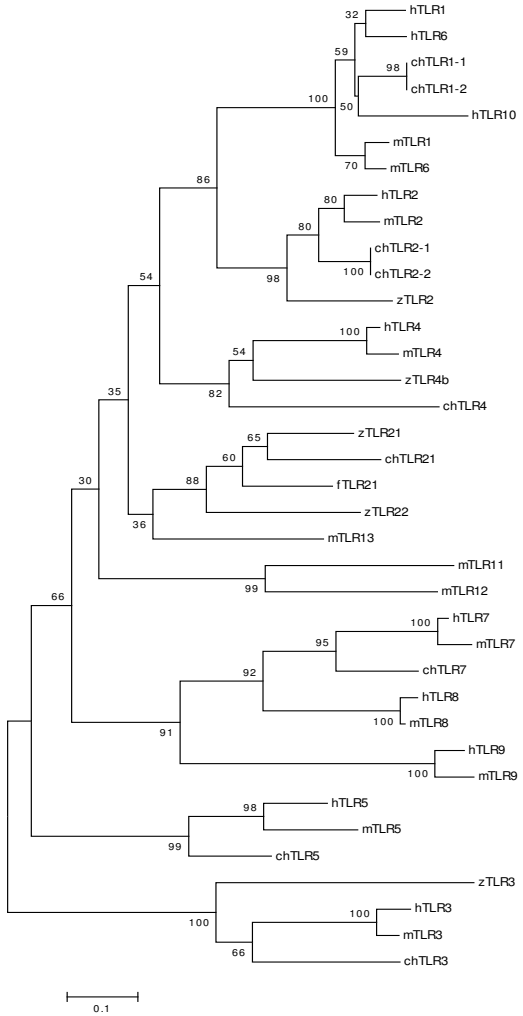


Fig. 3.5 Phylogenetic tree constructed based on the amino acid sequences of carboxy-terminal Toll/Interleukine-1 receptor (TIR) domains of chicken, human, mouse, pufferfish and zebrafish Toll-like receptors. Data was analyzed using Poisson correction and gaps were removed by complete deletion. The bootstrapping values are indicated by the numbers at the nodes. Human, mouse, and chicken Toll-like receptors are abbreviated as hTLR, mTLR, and chTLR, respectively. ChTLR1-1, chTLR1-2, chTLR2-1, and chTLR2-2 represent chTLR1 type 1, chTLR1 type 2, chTLR2 type 1, and chTLR2 type 2, respectively

Conservation of synteny

Analyses of human, mouse and chicken genomes demonstrate the conserved syntenies. This novel chicken toll like receptor lies between SLC7A6 and NDRG4 genes (Fig. 3.6). The order, chromosomal location and orientation of these two genes preserve in both of human and mouse genomes although there is no evidence of a TLR between them. Both of fugu and chicken TLR21 contains one single exon in the coding region although whether the chromosomal locations are conserved or not is unknown due to the unassembled fugu genome. The chicken TLR21 gene consists of two exons. Exon one and 2 in fugu gene seem to be a single exon.

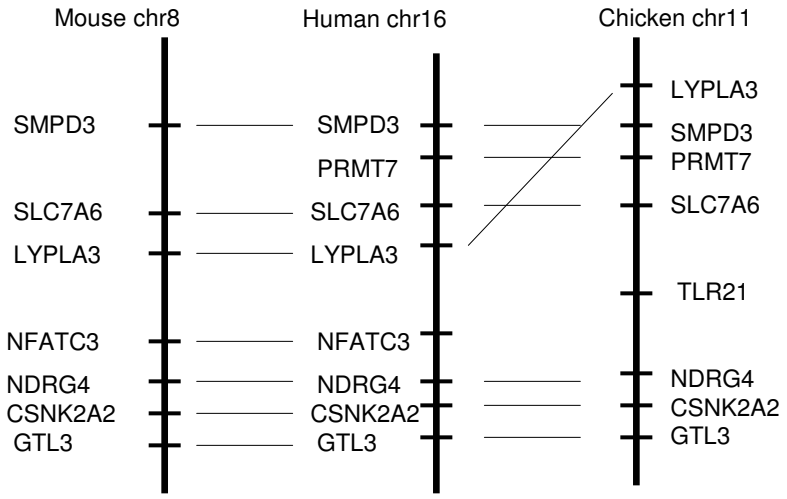


Fig. 3.6 Comparative genomic synteny in human, mouse, and chicken for flanking genes of chicken TLR21

Tissue expression profiles

To assess whether the predicted chicken TLR21 gene was expressed, the expression of chicken TLR21 mRNAs in different tissues such as the blood, kidney, liver, lung, oviduct, small intestine, large intestine and spleen was examined by RT-PCR. Chicken TLR21 was widely expressed in all eight tissues tested (Fig. 3.7), with strong expression in spleen, small intestine and large intestine.

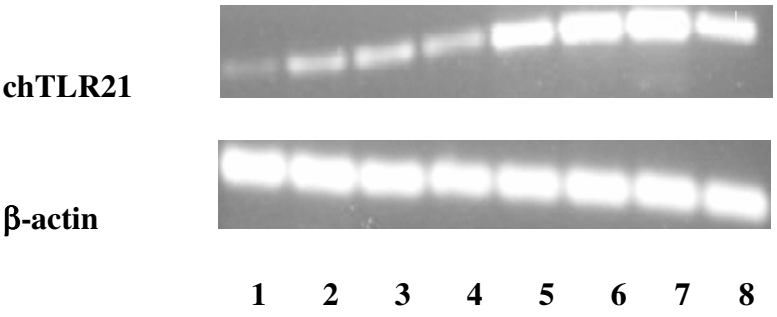


Fig. 3.7 RT-PCR analysis of chicken Toll-like receptor 21 gene expression among different tissues. (1) kidney, (2) liver, (3) lung, (4) oviduct, (5) small intestine, (6) large intestine, (7) spleen and (8) white blood cells. The chicken β-actin gene was used as a control for constitutive expression

Predicted ligand

A leucine zipper motif (PS00029, Accession number in PROSITE database) (periodic repetition of leucine residues at every seventh position) was detected within the LRR8 and LRR9 sequences of the chicken TLR21 by PROSITE program. In addition, this leucine zipper motif is present in human toll like receptor 3, 7, 9 and 10, mouse toll like receptor 3, 7, 8 and 9, chicken toll like receptor 2 types 1, 3, 7 and 21. The sequence homology search of the protein database showed that there is a block (sequence DLSHNKL) having 100% identity with mouse TLR9 by BLOCKS program. Two DNA binding pattern (IPB001275 and IPB006821) were detected in chTLR21 by MotifFinder program.

The LRRs of chTLR21-ECDs follow the typical 24-residue LRR consensus sequence with characteristically spaced hydrophobic amino acid residues (Fig. 3.8). There are three insertions at the position 15 of the LRR7, LRR8 and LRR10 in chTLR21. When chTLR21 insertion sequences were BLASTP searches against the NCBI database, it hits human toll like receptor 9 with high score. The alignment between chicken chTLR21 insertion sequences and hitted human toll like receptor 9 sequence is shown in Figure 3.9. This hitted sequence is contained in the insertion of the human TLR9 LRR.

Fig. 3.8 Alignment of leucine-rich repeat motifs in chTLR21.The chTLR21-ECD sequence contains 21 LRRs and a C-terminal cap (LRR-CT). Alignment gaps are indicated by dashes. The consensus sequence is displayed on the bottom and residues consistent with the consensus motif are highlighted in dark gray and the hydrophobic residues in the LRR consensus are highlighted in gray. The four cysteine residues in the C-terminal cap are underlined

ChTLR21 Protein subcellular localization prediction showed that chicken TLR21 and human TLR9 are extracellular protein respectively by Both of LoCtree and PSORT II programs (data not shown). A signal peptide was detected in both chicken TLR21 and human TLR9 respectively.

```
chTLR21      LTLQLCNNSLRELAGGSPDMLWHVKILDLSYNSISQAEVFTQLH-LRNISLLHLIGNPLD 59
hTLR9        FTLDLSRNNLVTVQPEMFAQLSHLQCLRLSHNCISQAVNGSQFLPLTGLQVLDLSHNKLD 60

chTLR21     VFHLLDISDIQ---PRSLDFSGLVLGAQGLDKVCLRLQGPQALRRLQLQRNGLKVLHCNA 116
hTLR9       LYHEHSFTELPRLEALDLSYNSQPFGMQGVGHNFSFVAHLRTLRHLSLAHN--NIHSQVS 118

chTLR21         LQLCP-VLRELDLSWNRLQHVGCAGRL 142
hTLR9           QQLCSTSLRALDFSGNALGHMWAEGDL 145
```

Fig. 3.9 BLASTP result of alignment of amino acid sequences of chicken chTLR21 insertion with hTLR9 hit. Alignment gaps are indicated by dashes. Sequences with identical amino acid are highlighted in dark gray

DISCUSSION

Sequence analysis of chTLR21 indicated that the predicted start codon in the open reading frame reported in the present study is probably true start codon because there are several stop codons in the frame before the start codon (Kozak 1996). The chicken TLR21 gene consists of two exons. Exon 1 and 2 in fugu gene seem to be a single exon. This presumably reflects intron loss in the fugu or intron gain in chickens.

The novel TLRs in fugu (TLR21 and TLR22) were presumed to have been lost in the mammalian lineage and are fish-specific TLRs (Oshiumi et al. 2003). In this study, a

homologue of fish TLR21 was found in the chicken. The syntenic relationship showed that this novel chicken toll like receptor might have arisen due to gene duplication after the divergence between aves and mammals. Chicken TLR22 was not detected in the chicken genome and its absence could be explained either because it does not exist, or as an artifact of the incomplete chicken genome sequence.

The leucine zipper pattern has been shown to be a new DNA-binding motif and common to DNA-binding proteins (Landschulz et al. 1988; Busch and Sassone-Corsi 1990). The detection of a leucine zipper pattern by PROSITE program and a DNA binding region by the MotifFinder program suggested that chTLR21 might be a DNA binding protein.

In addition, the insertion at position 15 found in the chicken TLR21 LRR motif is consistent with the report that large insertions occur after positions 10 or 15 in LRR motifs of TLRs (Bell et al., 2003). The TLR ectodomains are responsible for ligand binding and the potential PAMP binding sites are more likely to be formed by insertions/loops in the LRR sequence (Choe et al., 2005; Bell et al., 2003). The human TLR9 LRR insertion was the highest scoring hit by BLASTP, with chTLR21 leucine LRR as a query; showing that chTLR21 may have the same ligand as human TLR9. TLR9 is essential for responses to bacterial DNA and the immune stimulatory effect of bacterial DNA is due to the presence of unmethylated deoxycytidylate-phosphate-deoxyguanylate (CpG) dinucleotides (Hemmi et al., 2000). Synthetic oligodeoxynucleotides (ODN) containing these unmethylated CpG motifs can mimic

bacterial DNA and activate both innate and acquired immune responses that have evolved to protect against intracellular infections (Krieg 2002).

The same predicted protein sorting signal of chicken TLR21 and human TLR9 indicated that they could have similar biological function. TLRs are integral membrane proteins and most of the TLRs are expressed on the cell surface. However, TLR9 appears to function intracellularly and require internalization and endosomal maturation for CpG-DNA to activate TLR9 (Bauer et al., 2002). How TLR9 ligands move from outside to inside the cell is not well known (Ulevitch et al., 2004). Based on these results, we conclude that the novel chicken toll like receptor 21 play its biological function by binding to CpG DNA ligand.

The ortholog of TLR9 has not been identified in the chicken genome (Yilmaz et al, 2005; Roach et al. 2005) and chicken may not have a TLR9. However, the chicken's immune system can be activated by bacterial DNA containing C_pG motif (Gomis et al. 2003). In this study, we predicted a novel chicken toll like receptor whose sequence characteristic motif can recognize nucleic acid DNA. Therefore, the function of mammalian TLR 9 may be carried out by this novel toll like receptor by binding invading bacterial DNA. Many cell types have shown response to CpG ODN stimulation but these responses may activate either direct stimulation through TLR9 or indirect activation through CpG ODN-induced cytokine secretion (Griebel et al., 2005). The ligand inference results of this study may be tested by RNA interference (RNAi) technology in chickens.

CHAPTER IV

SUMMARY AND CONCLUSIONS

In summary, 23 chemokine and 14 chemokine receptor genes were identified from the chicken genome in this study. Many chicken genes displayed high degrees of similarity with their human and mouse orthologs in terms of gene structure and synteny. Chicken has significantly fewer CCLs, CXCLs, CCRs, and CXCRs than mammals, but it has the same number of CX_3C, XC, and cognate receptors as mouse. The results of phylogenetic analysis generally agree with the comparative chromosomal locations and syntenies of the genes. The independent nomenclature of chicken chemokines and chemokine receptors suggests that the chicken may have ligand-receptor pairings similar to mammals. The organization of these genes suggests that there were a substantial number of these genes present before divergence between aves and mammals and more gene duplications and CC, CXC, CCR, and CXCR subfamilies in mammals than in aves after the divergence.

Based on the organization, syntenies, and phylogenetic trees of chicken, mouse, and human chemokine and chemokine receptor genes, it may be concluded that there were a substantial number of chemokine and cognate receptor genes before divergence between aves and mammals and the organization of chicken chemokine genes represents a prototype of ancient genomic structure of chemokine genes before the divergence of mammals and aves. The presence of a few chicken chemokine and chemokine receptor paralogs and orthologs of the mammalian genes indicated that most

chicken chemokine and their receptor genes shared common ancestors with the human and mouse genes. There were significantly more gene duplications in CC, CXC, CCR, and CXCR subfamilies in mammals than in aves after the divergence of mammals and aves. The mammalian and chicken genome sequences and genes identified in this study can be used for further investigation of the molecular evolution of these gene families and as a model for the study of the divergence between aves and mammals. Avian and mammalian species may share very similar chemokine-receptor binding patterns. The results of this study may be used as functional inferences for these chicken genes before they are experimentally tested.

In this study the sequence and expression pattern of chicken TLR21 were identified and characterized based on the draft chicken assembly. Sequence analysis of chTLR21 indicated that the predicted start codon in the open reading frame reported in the present study are probably true start codons because there are several stop codons in the frame before the start codon (Kozak 1996). Recently TLR21 in Fugu and zebrafish has been isolated (Oshiumi et al., 2003; Meijer et al., 2004; Jault et al., 2004). The Fugu genome has orthologues of all human TLRs except TLR4 and TLR6. The TIR domain of chicken TLR21 has high sequence identity with the TIR domains of Fugu and zebrafish TLR21. The phylogenetic analysis of amino acid sequences, resemblance in the primary, secondary and tertiary structures (motifs) of polypeptides indicated that it might be an orthologue of Fugu and zebrafish TLR21 although the number of exons of chicken TLR21 is different from that of Fugu in the gene structure.

We know that CpG-ODN's immunostimulatory activity is mediated by pattern recognition receptor Toll-like receptor 9 in human and mouse (Bauer et al. 2001; Hemmi et al. 2000). The chicken TLR21-ECD LRR insertion analysis and insertion sequence similarity with human TLR9 showed that chTLR21 could have the same ligand as human TLR9. This information can be used for further functional inferences of chicken TLR 21 in the immune response.

REFERENCES

Akira S, Takeda K, Kaisho T (2001) Toll-like receptors: critical proteins linking innate and acquired immunity. Nat Immunol 2:675-680

Altschul SF, Madden TL, Schäffer AA, Zhang J, Zhang Z, Miller W, Lipman DJ (1997) Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. Nucleic Acids Research 25:3389-402

Anderson KV, Jurgens G, Nusslein-Volhard C(1985) Establishment of dorsal-ventral polarity in the Drosophila embryo: genetic studies on the role of the Toll gene product. Cell 42:779-789

Armant MA, Fenton MJ (2002) Toll-like receptors: a family of pattern-recognition receptors in mammals. Genome Biol 3(8)

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 (2002) Chemokine/chemokine receptor nomenclature. J Interferon Cytokine Res 22:1067-1068

Baggiolini M, Dewald B, Moser B (1997) Human chemokines: an update. Annu Rev Immunol 15:675-705

Barton GM, Medzhitov R (2002) Control of adaptive immune responses by Toll-like receptors. Curr Opin Immunol 14:380-383

Barton GM, Medzhitov R (2003) Toll-like receptor signaling pathways. Science 300:1524-1525

Bauer S, Kirschning CJ, Hacker H, Redecke V, Hausmann S, Akira S, Wagner H, Lipford GB (2001) Human TLR9 confers responsiveness to bacterial DNA via species-specific CpG motif recognition. Proc Natl Acad Sci U S A 98:9237-9242

Bauer S, Wagner H. Bacterial CpG-DNA licenses TLR9 (2002) Curr Top Microbiol Immunol 270:145-154

Bell JK, Mullen GE, Leifer CA, Mazzoni A, Davies DR, Segal DM (2003) Leucine-rich repeats and pathogen recognition in Toll-like receptors. Trends Immunol 24:528-533

Bendtsen JD, Nielsen H, von Heijne G, Brunak S (2004) Improved prediction of signal peptides: SignalP 3.0. J Mol Biol 340:783-795

Buchanan SG, Gay NJ (1996) Structural and functional diversity in the leucine-rich repeat family of proteins. Prog Biophys Mol Biol 65:1-44

Busch SJ, Sassone-Corsi P (1990) Dimers, leucine zippers and DNA-binding domains. Trends Genet 6:36-40

Caldwell,R.B., Kierzek,A.M., Arakawa,H., Bezzubov,Y., Zaim,J.,Fiedler,P., Kutter,S., Blagodatski,A., Kostovska,D., Koter,M.,Plachy,J., Carninci,P., Hayashizaki,Y. and Buerstedde,J.M (2005) Full-length cDNAs from chicken bursal lymphocytes to facilitate gene function analysis. Genome Biol 6(1):R6

Choe J, Kelker MS, Wilson IA (2005) Crystal structure of human toll-like receptor 3 (TLR3) ectodomain. Science 309:581-585

Corpet, F (1988) Multiple sequence alignment with hierarchical clustering. Nucleic Acids Res 16:10881-10890

Dabbagh K, Lewis DB (2003) Toll-like receptors and T-helper-1/T-helper-2 responses. Curr Opin Infect Dis 16:199-204

Forssmann U Magert HJ Adermann K, Escher SE, Forssmann WG (2001) Hemofiltrate CC chemokines with unique biochemical properties: HCC-1/CCL14a and HCC-2/CCL15. J Leukoc Biol 70:357-366

Fukui A, Inoue N, Matsumoto M, Nomura M, Yamada K, Matsuda Y, Toyoshima K, Seya T (2001) Molecular cloning and functional characterization of chicken toll-like receptors. A single chicken toll covers multiple molecular patterns. J Biol Chem 276:47143-47149

Gomis S, Babiuk L, Godson DL, Allan B, Thrush T, Townsend H, Willson P, Waters E, Hecker R, Potter A (2003) Protection of chickens against Escherichia coli infections by DNA containing CpG motifs. Infect Immun 71:857-863

Griebel PJ, Brownlie R, Manuja A, Nichani A, Mookherjee N, Popowych Y, Mutwiri G, Hecker R, Babiuk LA (2005) Bovine toll-like receptor 9: a comparative analysis of molecular structure, function and expression. Vet Immunol Immunopathol 108:11-16

Hartl M, Bister,K (2000) Suppression of genes in jun-transformed avian fibroblasts. jun-suppressed chemokine (JSC) mRNA with accession number AF285876, unpublished report, EMBL nucleotide sequence database

Hemmi H, Takeuchi O, Kawai T, Kaisho T, Sato S, Sanjo H, Matsumoto M, Hoshino K, Wagner H, Takeda K, Akira S (2000) A toll-like receptor recognizes bacterial DNA. Nature 408:740-745

Higgs R, Cormican P, Cahalane S, Allan B, Lloyd AT, Meade K, James T, Lynn DJ, Babiuk LA, O'farrelly C (2006) Induction of a novel chicken Toll-like receptor following Salmonella enterica serovar Typhimurium infection. Infect Immun 74:1692-1698

Horuk R (2001) Chemokine receptors. Cytokine Growth Factor Rev 12:313-335 Hughes S Haynes A O'Regan M, Bumstead N (2001) Identification, mapping, and phylogenetic analysis of three novel chicken CC chemokines. Immunogenetics 53:674-683

Huang, X. and Madan, A (1999) CAP3: A DNA sequence assembly program. Genome Res 9:868-877

Hughes S, Haynes A, O'Regan M, Bumstead N (2001) Identification, mapping, and phylogenetic analysis of three novel chicken CC chemokines. Immunogenetics 53:674-683

Iqbal M, Philbin VJ, Withanage GS, Wigley P, Beal RK, Goodchild MJ, Barrow P, McConnell I, Maskell DJ, Young J, Bumstead N, Boyd Y, Smith AL (2005) Identification and functional characterization of chicken toll-like receptor 5 reveals a fundamental role in the biology of infection with Salmonella enterica serovar typhimurium. Infect Immun 73:2344-2350

Jault C, Pichon L, Chluba J (2004) Toll-like receptor gene family and TIR-domain adapters in Danio rerio. Mol Immunol 40:759-771

Kaiser P, Hughes S, Bumstead N (1999) The chicken 9E3/CEF4 CXC chemokine is the avian orthologue of IL8 and maps to chicken chromosome 4 syntenic with genes flanking the mammalian chemokine cluster. Immunogenetics 49:673-684

Kozak M (1996) Interpreting cDNA sequences: some insights from studies on translation. Mamm Genome 7:563-574

Krieg AM (2002) CpG motifs in bacterial DNA and their immune effects. Annu Rev Immunol 20:709-760

Kumar S, Tamura K & Nei M (2004) MEGA3: Integrated software for molecular evolutionary genetics analysis and sequence alignment. Briefings in Bioinformatics 5:150-163

Laing KJ, Secombes CJ (2004) Chemokines. Dev CompImmunol 28:443-460

Landschulz WH, Johnson PF, McKnight SL (1988) The leucine zipper: a hypothetical structure common to a new class of DNA binding proteins. Science 240:1759-1764

Leveque G, Forgetta V, Morroll S, Smith AL, Bumstead N, Barrow P, Loredo-Osti JC, Morgan K, Malo D (2003) Allelic variation in TLR4 is linked to susceptibility to Salmonella enterica serovar Typhimurium infection in chickens. Infect Immun 71:1116-1124

Li QJ, Lu S, Ye RD, Martins-Green M (2000) Isolation and characterization of a new chemokine receptor gene, the putative chicken CXCR1. Gene 257:307-317

Li QJ, Yao M, Dueck M, Parpura V, Martins-Green M (2005) cCXCR1 is a receptor for cIL8 (9E3/cCAF) and its N- and C-terminal peptides and is also activated by hIL8 (CXCL8). J Leukoc Biol 77:421-431

Liang TS, Hartt JK, Lu S, Martins-Green M, Gao JL, Murphy PM (2001) Cloning, mRNA distribution, and functional expression of an avian counterpart of the chemokine receptor/HIV coreceptor CXCR4. J Leukoc Biol 69:297-305

Lichtman AH, Abbas AK (1997) T-cell subsets: recruiting the right kind of help. Curr Biol 7:R242-R244

Martin T, Cardarelli PM, Parry GC, Felts KA, Cobb RR (1997) Cytokine induction of monocyte chemoattractant protein-1 gene expression in human endothelial cells depends on the cooperative action of NF-kappa B and AP-1. Eur J Immunol 27:1091-1097

Meijer AH, Gabby Krens SF, Medina Rodriguez IA, He S, Bitter W, Ewa Snaar-Jagalska B, Spaink HP (2004) Expression analysis of the Toll-like receptor and TIR domain adaptor families of zebrafish. Mol Immunol 40:773-783

Murphy PM (1996) Chemokine Receptors: cloning strategies. Methods 10:104-118

Murphy PM (2002) International union of pharmacology. XXX. Update on chemokine receptor nomenclature. Pharmacol Rev 54:227-229

Murphy P.M (2003) Chemokines. In: William EP (ed), Fundamental Immunology. Lippincott Williams & Wilkins, New York, pp 801-840

Nomiyama H, Mera A, Ohneda O, Miura R, Suda T, Yoshie O (2001) Organization of the chemokine genes in the human and mouse major clusters of CC and CXC chemokines: diversification between the two species. Genes Immun 2:110-113

Olson TS, Ley K (2002) Chemokines and chemokine receptors in leukocyte trafficking. Am J Physiol Regul Integr Comp Physiol 283:R7-R28

O'Neill LA (2004) TLRs: Professor Mechnikov, sit on your hat. Trends Immunol 25:687-693

Onuffer JJ, Horuk R (2002) Chemokines, chemokine receptors and small-molecule antagonists: recent developments. Trends Pharmacol Sci 23:459-467

Oshiumi H, Tsujita T, Shida K, Matsumoto M, Ikeo K, Seya T (2003) Prediction of the prototype of the human Toll-like receptor gene family from the pufferfish, Fugu rubripes, genome. Immunogenetics 54:791-800

Petrenko O, Ischenko I, Enrietto PJ (1995) Isolation of a cDNA encoding a novel chicken chemokine homologous to mammalian macrophage inflammatory protein-1 beta. Gene 160:305-306

Petrenko,O., Ischenko,I. and Enrietto,P.J (1997) Characterization of changes in gene expression associated with malignant transformation by the NF-kappaB family member, v-Rel. Oncogene 15:1671-1680

Philbin VJ, Iqbal M, Boyd Y, Goodchild MJ, Beal RK, Bumstead N, Young J, Smith AL (2005) Identification and characterization of a functional, alternatively spliced Toll-like receptor 7 (TLR7) and genomic disruption of TLR8 in chickens. Immunology 114:507-521

Ponting CP, Schultz J, Milpetz F, Bork P (1999) SMART: identification and annotation of domains from signalling and extracellular protein sequences. Nucleic Acids Res 27:229-232

Read LR, Cumberbatch JA, Buhr MM, Bendall AJ, Sharif S (2005) Cloning and characterization of chicken stromal cell derived factor-1. Dev Comp Immunol 29:143-152

Reese MG (2001) Application of a time-delay neural network to promoter annotation in the Drosophila melanogaster genome. Comput Chem 26,51-56

Roach JC, Glusman G, Rowen L, Kaur A, Purcell MK, Smith KD, Hood LE, Aderem A (2005) The evolution of vertebrate Toll-like receptors. Proc Natl Acad Sci U S A 102:9577-9582

Rossi D, Sanchez-Garcia J, McCormack WT, Bazan JF, Zlotnik A (1999) Identification of a chicken "C" chemokine related to lymphotactin. J Leukoc Biol 65:87-93

Rovin BH, Dickerson JA, Tan LC, Hebert CA (1995) Activation of nuclear factor-kappa B correlates with MCP-1 expression by human mesangial cells. Kidney Int 48:1263-1271

Saitou N, Nei M (1987) The neighbor-joining method: a new method for reconstructing phylogenetic trees. Mol Biol Evol 4:406-425

Salamov AA, Nishikawa T, Swindells MB (1998) Assessing protein coding region integrity in cDNA sequencing projects. Bioinformatics 14:384-390

Sick C, Schneider K, Staeheli P, Weining KC (2000) Novel chicken CXC and CC chemokines. Cytokine 12:181-186

Smith J, Speed D, Law AS, Glass EJ, Burt DW (2004) In-silico identification of chicken immune-related genes. Immunogenetics 56:122-133

Sugano,S., Stoeckle,M.Y. and Hanafusa,H (1987) Transformation by Rous sarcoma virus induces a novel gene with homology to a mitogenic platelet protein. Cell 49:321-328

Takeda K, Akira S (2003) Toll-like receptors. Annu Rev Immunol 21:335-376

Ulevitch RJ, Mathison JC, da Silva Correia J (2004) Innate immune responses during infection. Vaccine 6;22 Suppl 1:S25-30

Van Coillie E, Van Damme J, Opdenakker G (1999) The MCP/eotaxin subfamily of CC chemokines. Cytokine Growth Factor Rev 10:61-86

Yamamoto M, Takeda K, Akira S (2004) TIR domain-containing adaptors define the specificity of TLR signaling. Mol Immunol 40:861-868

Yilmaz A, Shen S, Adelson DL, Xavier S, Zhu JJ (2005) Identification and sequence analysis of chicken Toll-like receptors Immunogenetics 56:743-753

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

VITA

JIXIN WANG
Department of Poultry Science
Texas A&M University
Phone: 979-845-1931
E-mail: jixinwang@gmail.com

EDUCATION

2006 M.S. in Poultry Science, Texas A&M University, College Station, Texas
2002 M.S. in Physiology, South China Agricultural University, Guangzhou, China
1999 B.S. in Animal Science, Tarim University of Agriculture and Reclamation, Alar, China

SELECTED PUBLICATIONS

1. 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(1):45.

2. Wang, J., Jiang, Q. Y., and Fu, W. L. 2002. Establishment of in vitro culture for porcine skeletal muscle satellite cell. Journal of South China Agricultural University, 23(1), 89-91.

CONFERENCE PRESENTATIONS

1. Zhu, J.J., Wang, J., Adelson, D. L., Yilmaz, A. Chicken chemokines and their genomic organization. Plant & Animal Genomes XIII Conference, San Diego, CA, January 2005.

2. Wang, J, Jiang, Q., Fu, W. A new method for isolation and culture of porcine myogenic satellite cells. The 7th National Symposium on Animal Physiology and Biochemistry Forum, Zhengzhou, China, August 2001.