Developmental and Comparative Immunology 49 (2015) 79–95

Contents lists available at ScienceDirect

Developmental and Comparative Immunology

journal homepage: www.elsevier.com/locate/dci

Chemokine receptors in Atlantic salmon

Unni Grimholt [a,*], Helena Hauge [b], Anna Germundsson Hauge [b], Jong Leong [c], Ben F. Koop [c]

[a] Soeren Jaabaeksgate 10B, 0460 Oslo, Norway
[b] Norwegian Veterinary Institute, P.O. Box 750 Sentrum, 0106 Oslo, Norway
[c] Centre for Biomedical Research, Department of Biology, University of Victoria, PO Box 3020 STN CSC, Victoria, Canada

ARTICLE INFO

Article history:
Received 8 July 2014
Revised 9 November 2014
Accepted 10 November 2014
Available online 15 November 2014

Keywords:
Atlantic salmon
Chemokine receptors
Whole genome duplication
Inflammation

ABSTRACT

Teleost sequence data have revealed that many immune genes have evolved differently when compared to other vertebrates. Thus, each gene family needs functional studies to define the biological role of individual members within major species groups. Chemokine receptors, being excellent markers for various leukocyte subpopulations, are one such example where studies are needed to decipher individual gene function. The unique salmonid whole genome duplication that occurred approximately 95 million years ago has provided salmonids with many additional duplicates further adding to the complexity and diversity. Here we have performed a systematic study of these receptors in Atlantic salmon with particular focus on potential inflammatory receptors.

Using the preliminary salmon genome data we identified 48 chemokine or chemokine-like receptors including orthologues to the ten receptors previously published in trout. We found expressed support for 40 of the bona fide salmon receptors. Eighteen of the chemokine receptors are duplicated, and when tested against a diploid sister group the majority were shown to be remnants of the 4R whole genome duplication with subsequent high sequence identity. The salmon chemokine receptor repertoire of 40 expressed bona fide genes is comparably larger than that found in humans with 23 receptors. Diversification has been a major driving force for these duplicate genes with the main variability residing in ligand binding and signalling domains.

1. Introduction

Chemokine receptors (CRs) and their ligands play an important role in coordination of cell trafficking in many biological processes. They are predominantly expressed on the surface of leukocytes and because they dictate migration of these cells between tissues, they are crucial for an effective immune response. Chemokines and their receptors have traditionally been divided into two main functional categories; homeostatic chemokines and their receptors are involved in basal cell trafficking and homing while inducible chemokines and their receptors are involved in inflammatory responses. There are also a few receptors with dual function in addition to some atypical chemokine receptors (Bonecchi et al., 2010; Cancellieri et al., 2013).

CRs belong to the large family of G-protein-coupled seven transmembrane receptors with four extracellular and four intracellular domains. The extracellular N-terminal part of the receptor is responsible for ligand binding while the intracellular domains including the C-terminus are involved in intracellular signalling (Neel et al., 2005; Szpakowska et al., 2012). CRs are named according to the chemokine class they bind. CCRs bind to CC-chemokines, CXCRs bind to CXC chemokines, XCR binds to XC chemokines and CX3CR binds to CX3C chemokines where X is any amino acid and C is cysteine (Allen et al., 2007; Charo and Ransohoff, 2006).

In general there are fewer receptors than chemokines, with approximately 20 receptors versus 50 ligands identified in mammals, so most receptors bind more than one ligand. With the exception of atypical CRs, ligand binding causes conformational changes in the receptors that in turn trigger intracellular signals causing cellular events such as directional cellular migration.

Homeostatic chemokines are constitutively expressed and are important for normal cell trafficking and for the development and maintenance of the immune system (Moser and Loetscher, 2001; Proudfoot, 2002). The human homeostatic receptors are CCR7, CCR9, CCR10, CXCR4 and CXCR5, where for instance CCR7 is expressed on cells destined for lymph nodes, CCR9 directs leukocytes to the intestine, CCR10 directs T-cells to skin and intestine and CXCR5 directs B-cells to lymph node follicles (Charo and Ransohoff, 2006). CXCR4 is widely expressed with multiple functions including a role in the central nervous system (Bonecchi et al., 2010; Tran and Miller, 2003).

Expressions of inflammatory chemokines are induced by mediators such as tumour necrosis factor, interferon-γ, or microbial products associated with infection or trauma (Charo and Ransohoff,

* Corresponding author. Soeren Jaabaeksgate 10B, 0460 Oslo, Norway. Tel.: +47 92661039.
E-mail address: unni.grimholt@gmail.com (U. Grimholt).

http://dx.doi.org/10.1016/j.dci.2014.11.009
0145-305X/

2006). A classic example would be a pathogen recognised by a toll-like receptor which then induces expression of secreted chemokine (Kaisho, 2012). In humans, the inflammatory chemokine receptors are CCR1-3, CCR5, CXCR1, CXCR2, and CX3CR1. The receptors CCR4, CCR6, CCR8, CXCR3, CXCR6, and XCR1 have dual functions participating in both inflammatory as well as homeostatic processes.

Most of the atypical or silent CRs also bind chemokines, but this binding does not induce a signalling cascade with subsequent cell migration. Instead, several of these receptors have regulatory roles. Human atypical receptors are CCBP2 (D6), CCRL1 (CCX-CKR), CCRL2, DARC, and CXCR7 (RDC1). CCBP2 has been shown to act as a scavenger for CC-chemokines and can drastically reduce the amount of ligand available for other CRs (Graham et al., 2012). The chemokine-like receptor CMKLR1 does not bind to a chemokine, but may have multiple functions as it can regulate CCRL2 activity through competitive binding to the ligand chimerin (Yoshimura and Oppenheim, 2011). Some atypical receptors have recently been renamed to ACKRs where CCBP2 is now ACKR2, CXCR7 is ACKR3 and CCRL1 is ACKR4 (Bachelerie et al., 2014).

CRs have been identified in many teleost species with the primary focus on teleosts with sequenced genomes i.e. fugu, tetraodon, medaka, stickleback and zebrafish (Bajoghli et al., 2009; Chang et al., 2007; DeVries et al., 2006; Diotel et al., 2010; Huising et al., 2003b; Liu et al., 2009; Nomiyama et al., 2011; Oehlers et al., 2010; Sasado et al., 2008; Verburg-van Kemenade et al., 2013; Xu et al., 2010). Results from these studies show clear-cut teleost orthologues to mammalian homeostatic receptors, but orthology to mammalian inflammatory CRs is less obvious. A few publications also exist on CRs in salmonids i.e. CCR6, CCR7, CCR9/9b, CCR13 (CCR3), IL8R (CXCR1), CXCR2, CXCR3a, CXCR3b and CXCR4 (Daniels et al., 1999; Dixon et al., 2013; Ordas et al., 2012; Xu et al., 2014; Zhang et al., 2002). There are also a few accepted teleost chemokine receptor ligand pairs such as CCR6 with CCL20-like ligands and CCR7 with CCL19/21-like ligands (Laing and Secombes, 2004). However, the functional role of most fish CRs remains unresolved. Understanding the specific function of individual receptors and identifying their ligands is essential for understanding teleost homeostasis and inflammation.

To broaden our understanding of CRs in salmonids, we made use of the salmon genome (Davidson et al., 2010) to identify receptors and study their evolution and potential function. From a disease prevention point of view we paid particular attention to receptors potentially involved in inflammation. Of the 48 receptors identified several have potential roles in inflammation being expressed in immunologically important tissues. Most importantly, the salmonid-specific whole genome duplication event approx. 95 million years ago (Macqueen et al., 2014) has had a significant impact on the receptor repertoire.

2. Material and methods

2.1. Bioinformatics

Using available CR sequences from published articled and/or retrieved from GenBank, BLASTN (Altschul et al., 1997) and TBLASTN (Schaffer et al., 2001) searches were initially performed against both expressed and genomic Atlantic salmon resources at NCBI GenBank, cGRASP [cGRASP, Internet 2009; (Rondeau et al., 2014)] and/or the SalmonDB in Chile (Di Genova et al., 2011). Identified genomic sequences from the latest Atlantic salmon genome assembly (GenBank: AGKD00000000.3) or from the Northern pike genome assembly (AZJR00000000.1) were subjected to gene prediction analysis using GenScan (Burge and Karlin, 1997), FGENESH (Solovyev et al., 2006) and/or Augustus (Stanke et al., 2006). Predicted ORFs were tested through alignment with similar sequences from other species and sometimes changed using expressed match in Spidey (Wheelan et al., 2001). Spidey was also used to define exon intron boundaries.

To assess evolutionary relationships and orthology, all identified amino acid sequences were aligned using ClustalX2.0.11 (Larkin et al., 2007). ClustalX was also used to calculate percentage identity. Phylogenetic analyses were performed using the neighbor-joining method (Saitou and Nei, 1987). The percentages of replicate trees in which the sequences clustered together in the bootstrap test (1000 replicates) are shown next to the branches (Felsenstein, 1985). The tree is drawn to scale, with branch lengths in the same units as those of the evolutionary distances used to infer the phylogenetic tree. The evolutionary distances were computed using the *p*-distance method (Nei and Kumar, 2000) and are in the units of the number of amino acid differences per site. All ambiguous positions were removed for each sequence pair. Evolutionary analyses were conducted in MEGA5 (Tamura et al., 2011).

Secondary structure including transmembrane domains were predicted and visualised using Rhythm (Rose et al., 2009). Prediction of N-linked glycosylation sites was performed using NetNGlyc 1.0 (Gupta and Brunak, 2002) while tyrosine sulfation sites were predicted using Sulfinator (Monigatti et al., 2002). O-linked glycosylation sites were predicted using NetOGlyc 3.1 (Julenius et al., 2005).

2.2. Northern pike cDNA and genomic DNA

Pike (*Esox lucius*) genome, transcriptome and genetic map data are described fully in Rondeau et al. (2014). In brief, DNA from a single pike individual (Leong et al., 2010) was submitted directly to BGI (http://www.genomics.cn/en/index) for Illumina sequencing; DNA libraries of 180 bp were constructed for paired-end sequencing and libraries of 2 kb and 6 kb fragments were constructed for mate-pair sequencing and assembly. Fragment assembly used ALLPATHS-LG (Gnerre et al., 2011). The resulting contigs ≥200 bp were screened and trimmed for vector and contamination, which produced 94,267 contigs (N50 = 16,909, bioproject PRJNA221548, accession GenBank:AZJR00000000) and 5688 scaffolds ≥1000 bp with a total genome size of 877,777,613 bp.

2.3. Tissue transcriptomes and analysis

For transcriptome data, tissues were extracted from a single, 1 year old, presmolt juvenile male Atlantic salmon. RNA from 11 tissues – brain, eye, gill, hind gut, head kidney, heart, kidney, liver, muscle, stomach, spleen – were extracted and submitted to BGI for Illumina sequencing. Contig assembly used Trinity (Haas et al., 2013). The resulting set of transcripts was reduced by retaining those with a significant BLASTX (Altschul et al., 1997) match to the SwissProt or Gene Ontology protein databases ($\leq 10^{-5}$) or had a predicted open reading frame ≥300 bp. Further, only those transcripts that mapped to our genome assembly using BLAT (Kent, 2002) were retained. To remove possible alleles from our assembly, we retained a single, longest representative of transcripts that were ≥98% similar over a minimum length of 300 bp, as determined by BLASTN (Altschul et al., 1997). This curated set represents our RNA-seq reference transcriptome. FPKM values were then determined for each transcript for each of the 11 different tissues.

Pike chemokine receptors were identified using known salmon genes as queries that were BLASTed against the pike genome and transcriptome (Rondeau et al., 2014). Identified contigs were further examined as earlier.

2.4. RNA extraction

Three healthy Atlantic salmon weighing 70–80 g (AquaGen breed) kept in a freshwater flow system at 12 °C with regular feeding were sacrificed by overexposure to Finquel® (ScanAqua AS) and tissues