Fish & Shellfish Immunology 28 (2010) 40–48

Contents lists available at ScienceDirect

Fish & Shellfish Immunology

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

The immunoglobulin gene loci in the teleost *Gasterosteus aculeatus*

Yonghua Bao [a,*], Tao Wang [a], Yongchen Guo [a], Zhihui Zhao [b], Ning Li [a], Yaofeng Zhao [a,c,*]

[a] State Key Laboratory of Agrobiotechnology, College of Biological Sciences, China Agricultural University, Beijing 100193, PR China
[b] College of Animal Science and Veterinary Medicine, Agricultural Division, Jilin University, Changchun 130062, PR China
[c] College of Animal Science and Technology, Qingdao Agricultural University, Qingdao 266109, PR China

ARTICLE INFO

Article history:
Received 26 June 2009
Received in revised form
16 September 2009
Accepted 17 September 2009
Available online 24 September 2009

Keywords:
Stickleback
Immunoglobulin
Genome
Locus
Cluster pattern
Evolution

ABSTRACT

In the present study, we have annotated both the immunoglobulin heavy (IgH) and light (IgL) chain genes in the stickleback (*Gasterosteus aculeatus*), based on the recently released genome data. The IgH gene locus is arranged in a configuration of $(V_n\text{-}D\text{-}J\text{-}C_\zeta\text{-}D_3\text{-}J_4\text{-}C_\mu\text{-}C_\delta)_3\text{-}V_6\text{-}D\text{-}J\text{-}C_\zeta$, which is structurally different from any of the known teleost IgH loci. The μ genes consistently exhibit a 4-CH encoding structure and all the ζ genes encode only three CH domains (lacking the equivalent exon of the zebrafish ζCH2). As in many other teleosts, the stickleback δ genes contain multiple CH exons, but exist as three copies. The members of four V_H gene families, containing 47 segments, were interspersed in the germline. The stickleback IgL chain genes are also organized in multiple clusters and located in three chromosomes (10, 11, and 15). Sequence and phylogenetic analyses revealed that two isotypes, L1 (κ, including two subgroups, 1A and 1B) and L2 (σ) could be identified. The transcriptional orientations of the V_L segments were found to be either the same (only in L2 isotype) or opposite to (in L1A, 1B and 2 isotypes) those of the J_L and C_L segments, indicating that these segments would undergo rearrangement by deletion or inversion when expressed.

1. Introduction

Structural analysis of the Ig loci is important for a better understanding of the genetic basis of antibody diversity and evolutionary divergence of the locus in vertebrates [1]. Using potent molecular biology tools, the genomic organization of the Ig gene loci in many species has been characterized. Most investigators preferred classification and nomenclature that reflected relationships and evolutionary descent. Although all vertebrate Ig genes are clearly related, some species present distinctive characteristics due to multiple gene duplications and losses that have undoubtedly occurred during evolution.

In mammals, there are five major Ig isotypes (IgM, IgD, IgG, IgE, and IgA) that possess distinct effector functions. The IgH locus has been well described in humans and mice, where a "translocon" arrangement of multiple V gene segments followed by D, J, and C domains, occurs [2]. The C_μ region is immediately 3′ to the J segments and is followed by δ, γ, ε, and α. In contrast, the immunoglobulin genes of all cartilaginous fish studied to date are arranged in a "multiple clusters" configuration. Each cluster contains one V segment, one or more D segments, one J segment, and a single C gene [3]. The somatic rearrangement of these gene segments is restricted to a cluster [4]. Three isotypes, IgM, IgW (also known as IgX or IgNARC (New Antigen Receptor from Cartilaginous fish), dependent upon the species identified) and IgNAR (New Antigen Receptor), have been identified thus far in cartilaginous fish [5].

IgH loci from amphibians and reptiles, like those of mammals, are also organized in a "translocon" configuration [6,7]. The IgH locus in the amphibian *Xenopus tropicalis* has been shown to be arranged as $V_H\text{-}D_H\text{-}J_H\text{-}C_\mu\text{-}C_\delta\text{-}C_\chi\text{-}C_\upsilon\text{-}C_\Phi$, where the C_φ encodes the first (genetic) hinge-containing Ig class in non-mammalian vertebrates [8,9]. Recently, four IgH isotypes, IgM, IgD, IgY, and IgA, were identified in reptiles, although the IgA encoding gene shows a punctate distribution [10–12]. Similar to reptiles, birds express the characteristic IgM, IgA, and IgY [13,14]. A number of V_H pseudogenes, the absence of IgD and opposite IgA gene transcription orientation discriminated birds from other vertebrate groups [13,15–17].

Bony fish have been shown to express IgM [18], IgD [19], and IgZ/T/H [20–24]. An unusual pattern of genomic organization was uncovered in this category of fish, where IgZ/T gene (ζ/t) segments are sandwiched between V_H genes and D_H/J_H segments, resembling the mouse locus encoding the T cell receptor α (TCR α) and TCR δ [25]. It has been reported that the C region of the zebrafish ζ is encoded by four CH exons. Trout τ is predicted to lie in two IgH loci, both containing four CH exons. However, the fugu IgH gene is composed of only two CH exons, corresponding to the zebrafish

Abbreviations: C, constant gene; V, variable gene; D, diversity gene; J, joining gene; IgH, immunoglobulin heavy chain; IgL, immunoglobulin light chain.

* Correspondence address: State Key Laboratory of Agrobiotechnology, College of Biological Sciences, China Agricultural University, Beijing 100193, PR China. Tel.: +86 1062734945; fax: +86 1062733904.
E-mail addresses: baoyonghua2009@yahoo.com.cn (Y. Bao), yaofengzhao@cau.edu.cn (Y. Zhao).

1050-4648/$ – see front matter © 2009 Elsevier Ltd. All rights reserved.
doi:10.1016/j.fsi.2009.09.014

ζCH1 and ζCH4 respectively, and common carp possesses at least three copies of IgM/IgZ chimera of carp μCH1 domain and zebrafish ζCH4 domain analog [20–24].

In the L chain locus, the V_L is located upstream of the J_L, and is followed by C_L. To date, three major types of organization of these segments have been uncovered: multiclusters (e.g. in elasmobranches), minimalistic (e.g. in birds), and translocons (e.g. in mammals) [26]. The teleost IgL locus has been confirmed to be a cluster-like organization as has that of cartilaginous fish [27–30]. With the fourth isotype identified in elasmobranches [31] and κ and λ found in reptiles [32], vertebrate IgL isotypes discovered so far could be characterized: bird: λ; mammal: κ and λ; reptiles: κ and λ; amphibian: ρ (κ), type III (λ) and σ; teleost: L1/L3 (κ like), L2 (σ like) and λ [33], elasmobranch: type 1 (σ-cart), type 2 (λ like), type 3 (κ like) and type 4(σ).

In this study, we performed a detailed analysis of the stickleback IgH and IgL gene loci, revealing the presence of multiple copies of μ, δ, and ζ in the IgH locus and two types of light chain genes.

2. Materials and methods

2.1. Genomic analysis of the stickleback Ig gene loci

Available stickleback EST sequences of μ, δ, and ζ were used to identify the corresponding genomic sequences by BLAST searching against the stickleback genome database (www.ensembl.org). A 350 kb genomic contig was found to contain the V_H, D_H, J_H, and CH segments.

Similarly, cDNA sequences (GenBank accession numbers AY278356 and BT026591) of stickleback IgL genes were used as queries to scan the whole genome. The acquired matching regions were assigned to chromosomes 10, 15 (AY278356), and 11 (BT026591).

In addition, we also conducted searches in the stickleback genome using the known teleost IgL lambda V (GenBank accession numbers: Ip-EU925385, Ip-EU925383, If-CK403484, Om-BX861772, Gm-AJ293807, and Gm-AJ293808) and C sequences (GenBank accession numbers: Ip-EU872022, If-CK403484, Om-BX861350, and Gm-AJ293807) from channel catfish (Ip), blue catfish (If), rainbow trout (Om), and Atlantic cod (Gm) as queries. However, no additional Ig light chain genes could be identified.

2.2. Sequence analysis

According to conserved recombination signal sequence (RSS) motif specificity, FUZZNUC, an online software package (http://anabench.bcm.umontreal.ca/anabench/Anabench-Jsp/Applications/fuzznuc.jsp?APPLICATIONID=81&APPLICATIONNA ME=fuzznuc) was used to survey V, D, and J gene segments. Constant exons were identified by a conventional TBLASTN approach.

Editing and comparison of all sequences were performed using the DNAstar program. Multiple sequence alignments were made using the *Clustal*W method and optimized manually.

Phylogenetic trees were constructed with PHYLIP 3.67 software combined with the TreeView package. A neighbor-joining model was used in phylogenetic analysis (Bootstrapping with 1000 repeats).

3. Results

3.1. Stickleback heavy chain

3.1.1. Genomic organization

Our analysis showed that the IgH gene locus was arranged in a configuration of $(V_n\text{-D-J-C}\zeta\text{-D}_3\text{-J}_4\text{-C}_\mu\text{-C}_\delta)_3\text{-V}_6\text{-D-J-C}\zeta$, where a block of V_n-D-J-Cζ-D_3-J_4-$C_μ$-$C_δ$ was duplicated three times (Fig. 1). A similar genomic organization (V_n-D-J-Cζ-D_n-J_n-$C_μ$-$C_δ$) has also been reported in other teleosts, but with only one such copy [21–23]. This core IgH skeleton is specific to teleosts. Downstream of the three clusters, a single ζ gene and corresponding V, D, and J segments were found.

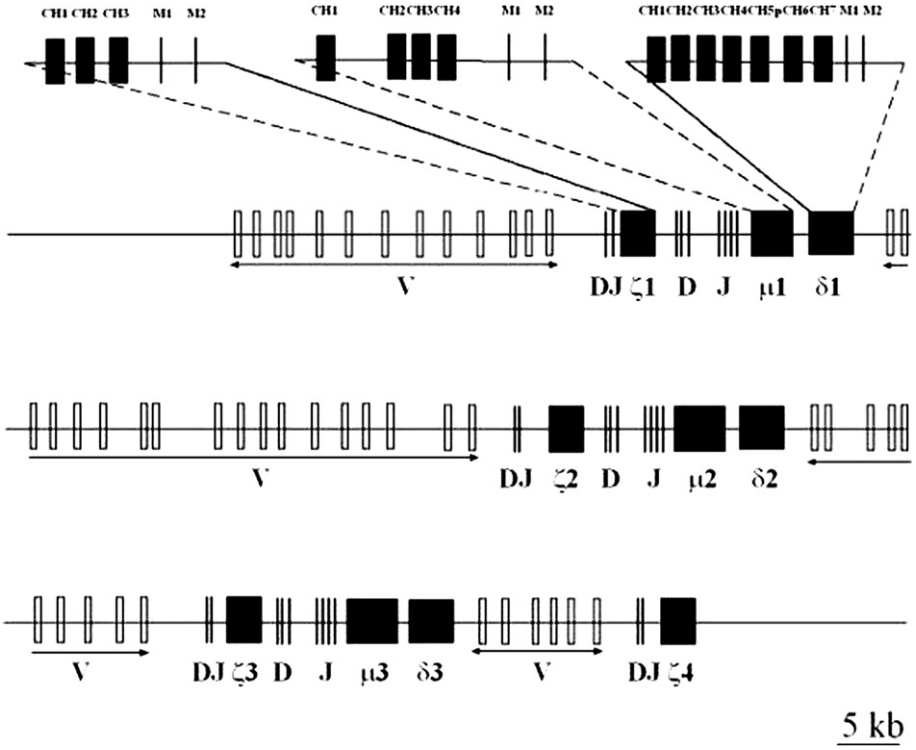

Fig. 1. The stickleback IgH locus. The genomic locations of V (V, D, and J) and C (μ, δ, and ζ) genes are shown to scale. Bidirectional arrowheads indicate the extent of V segments in the genome.