Marine Genomics xxx (2015) xxx–xxx

Contents lists available at ScienceDirect

Marine Genomics

New insights into evolution of IgT genes coming from Antarctic teleosts

Stefano Giacomelli [a], Francesco Buonocore [b], Fabio Albanese [a], Giuseppe Scapigliati [b], Marco Gerdol [c], Umberto Oreste [a], Maria Rosaria Coscia [a,*]

[a] Institute of Protein Biochemistry, CNR, Napoli, Italy
[b] Department for Innovation in Biological, Agro-food and Forest Systems, University of Tuscia, Largo dell'Università snc, Viterbo, Italy
[c] Department of Life Sciences, University of Trieste, Trieste, Italy

ARTICLE INFO

Article history:
Received 31 March 2015
Received in revised form 15 June 2015
Accepted 16 June 2015
Available online xxxx

Keywords:
Teleost IgT
Antarctic fish
Immunoglobulin heavy chain constant domains
Exon loss
Genome evolution

ABSTRACT

Cloning and characterization of IgT heavy chain genes were performed in the Antarctic Notothenioid teleost *Trematomus bernacchii* and in a non-Antarctic Notothenioid species, *Bovichtus diacanthus*, belonging to a phylogenetically basal lineage of Notothenioids. Compared to IgT from other non-Antarctic teleost species, including *B. diacanthus*, *T. bernacchii* IgT lacked most of the second constant domain but maintained only a few amino acid residues, which could be aligned to *B. diacanthus* CH2 domain. By analyzing several cDNA clones from a single specimen, three differently sized IgT transcript variants, named Long, Short and Shortest, were identified. Genomic analysis of *T. bernacchii* and *B. diacanthus* IgH loci revealed that, in the case of *T. bernacchii*, within the intron between the exons coding for the entire first and second constant domains a reminiscence of the ancestral second exon was present. The Long and Short variants were found to be encoded by indel alleles, whereas the Shortest variant was generated by alternative splicing that led to the CH2 exonic remnant skipping. Through comparison between genomic and cDNA sequences we hypothesized the presence of three different copies of the IgT heavy chain gene, one of which being considered the functional gene since the corresponding transcripts were identified. Moreover, either Long or Short exonic variants were found to be used in IgT heavy chain membrane form in an unbiased manner, as seen for the secretory form. Phylogenetic analysis was performed on the constant region from all teleost IgT available to date, including IgT from another Antarctic Notothenioid species, *Notothenia coriiceps*, identified by searching the transcriptome. The loss of almost an entire domain together with the conservation of some amino acids such as proline, glycine and cysteine in the CH2 domain remnant, could be interpreted as another distinctive feature of the Antarctic fish genome evolution, providing also new insights into the structural variation of teleost immunoglobulin genes.

1. Introduction

Immunoglobulins (Igs) are humoral mediators playing a key role in the vertebrate immune system to protect the organism from a wide variety of pathogens. Igs are composed of two heavy (H) chains and two light (L) chains, each with a variable region at the N-terminus and a constant region at the C-terminus; H and L chains are encoded by separate loci, the IgH locus and IgL locus, respectively. In teleost fish the genomic structure of the IgH locus, referred to as translocon configuration, comprises multiple gene segments responsible for antibody diversity. The immunoglobulin heavy chain repertoire is generated by random rearrangement of variable (*VH*), diversity (*D*), and joining (*JH*) gene segments. Moreover, different constant region genes (CH) account for different effector functions. Based on their constant region, five IgH isotypes have been identified in mammals: IgM (μ chain), IgD (δ chain), IgG (γ chain), IgA (α chain) and IgE (ε chain).

In teleost fish only three isotypes are known to date: IgM, mainly tetrameric, the predominant immunoglobulin in the serum; IgD, whose function is still unclear; IgT, representing an Ig isotype found only in teleosts and thought to be specialized in mucosal immunity (Zhang et al., 2011), as observed in the gut (Zhang et al., 2010) and skin mucus (Xu et al., 2013) of *Oncorhynchus mykiss*. The structure of IgM is highly conserved across species, in contrast to the structures of IgT and IgD that are more variable. IgD are chimeric molecules and usually contain seven or more CH domains, the first one being an IgM constant domain. The IgT isotype was first identified in few teleost species belonging to different orders, under different names according to the species: IgT in *O. mykiss* (Salmoniformes) (Hansen et al., 2005), novel IgH in *Takifugu rubripes* (Tetraodontiformes) (Savan et al., 2005a), while it was named IgZ in *Danio rerio* (Cypriniformes) (Danilova et al., 2005) and referred to as IgM-IgZ chimera in *Cyprinus carpio* (Savan et al., 2005b). IgT are present in serum as monomers, whereas in the

Abbreviations: Cμ, IgM heavy chain constant domain; Cτ, IgT heavy chain constant domain; IgH, immunoglobulin heavy chain gene locus.

* Corresponding author at: Institute of Protein Biochemistry, CNR, Via Pietro Castellino 111, 80131 Napoli, Italy. Tel.: +39 0816132556; fax: +39 0816132277.
E-mail address: mr.coscia@ibp.cnr.it (M.R. Coscia).

http://dx.doi.org/10.1016/j.margen.2015.06.009
1874-7787/

gut mucus they are found mainly as multimers, similarly to the tetrameric IgM, but non-covalently linked (Zhang et al., 2010). The domain number of IgT varies among species. In fact, *O. mykiss* and *D. rerio* IgT/Z comprise four constant domains (Danilova et al., 2005; Hansen et al., 2005), whereas IgT of *Gasterosteus aculeatus* (Gasterosteiformes) have three constant domains (Gambón-Deza et al., 2010; Bao et al., 2010). *T. rubripes* IgZ comprise only two domains, corresponding to the first and fourth constant domains of *D. rerio* IgT, respectively (Savan et al., 2005a); in *C. carpio* IgT have been shown to be chimeric molecules consisting of the Cμ1 domain and a second domain similar to Cτ4 of *D. rerio* (Savan et al., 2005b). Few years later, a new IgZ variant, named IgZ2 has been identified in *D. rerio* and in *C. carpio* (Ryo et al., 2010; Hu et al., 2010). It has been shown that these two IgZ variants have different functions: IgZ1 is mainly expressed systemically against blood pathogens, and the IgZ2 chimera is preferentially expressed in the mucosal compartment to respond to parasite infections. Interestingly, in *D. rerio* a membrane-bound IgZ2 co-localizes with the IgM membrane-bound form on B cell surface, suggesting that IgZ2 may represent a novel B cell receptor (Hu et al., 2010).

Later on, more and more sequences from other teleost species have become available, disclosing a wide variability in the organization of the IgH loci among teleosts (Fillatreau et al., 2013). In *D. rerio* and *O. mykiss*, the IgH locus has been shown to present Cτ exons located upstream of Cμ and Cδ exons, together with their own set of *D* and *JH* gene segments (Danilova et al., 2005; Hansen et al., 2005). This organization of IgH loci recalls the mouse T cell receptor α/δ locus (Danilova et al., 2005). In fugu IgT heavy chain constant genes are similarly found upstream of Cμ and Cδ genes and this Ig isotype has its own *D* and *JH* segments, but the gene organization differs significantly from *D. rerio* and *O. mykiss* IgT (Savan et al., 2005a). In the IgH locus of *G. aculeatus* Cτ, Cμ, and Cδ exons have been found tandem duplicated three times and separated by *VH*, *D*, and *JH* segments. Moreover, a fourth Cτ gene exists at the 3′ end of the locus (Gambón-Deza et al., 2010; Bao et al., 2010). In *Thunnus orientalis* IgM and IgT are each composed of four constant domains. Interestingly, IgM and IgT share three *VH* gene families whereas IgM exclusively uses one. Moreover, both IgM and IgT use the same *DH* segment, whereas the *JH* gene segments are isotype specific (Mashoof et al., 2014). Cypriniformes can also have different types of IgH loci. *D. rerio* has only one IgH locus with the canonical structure (Danilova et al., 2005) whereas the two *C. carpio* IgZ subisotypes are encoded by two distinct loci (Henkel et al., 2012). The family Salmonidae has been shown to possess a very complex IgH locus characterized by the presence of two parallel IgH isoloci (*IGHA* and *IGHB*) that can be attributed to its tetraploid past (Yasuike et al., 2010).

At present, in two well-known teleost species, *Ictalurus punctatus* and *Oryzias latipes*, ortholog IgT genes have not yet been found (Bengtén et al., 2006; Magadán-Mompó et al., 2011).

We have previously investigated IgM from several species of the Perciform suborder Notothenioidei (Coscia and Oreste, 2003; Coscia et al., 2000, 2010) that is the main component of the modern Antarctic ichthyofauna (Eastman, 2005). Notothenioids acquired peculiar features to adapt to the extremely cold Antarctic environment, such as development of antifreeze glycoproteins (Chen et al., 1997), and compensation for the lack of the swim bladder by reduction of skeletal mineral density and increase in lipid deposition (Schluter, 2000; Eastman, 2005). Another unique feature is the absence of hemoglobin in one of the five Antarctic Notothenioid families (Channichthyidae), due to the lack of the functional globin genes (di Prisco et al., 2002). Several studies have indicated that evolutionary adaptation of Notothenioids was accompanied by genome expansion (Detrich and Amemiya, 2010; Coppe et al., 2013).

In the present study, IgT heavy chain constant region genes of two Notothenioid species, the Antarctic teleost *Trematomus bernacchii* (Nototheniidae family) and the non-Antarctic species *Bovichtus diacanthus* (Bovichtidae family) were sequenced and characterized. The latter species was chosen for comparison since it represents the phyletically basal lineage of Notothenioid species that inhabited non-Antarctic more temperate waters before Antarctica became isolated from other continents (Eastman, 1993; Clarke and Johnston, 2003). Compared to IgT from other teleost species, including the non-Antarctic Notothenioid one, *T. bernacchii* IgT lack the heavy chain second constant domain, maintaining only a few amino acid residues compared to the *B. diacanthus* CH2 domain. Genomic analysis allowed us to clarify that the few residues left were encoded by a short region of different sizes, according to the genomic sequence analyzed, within the intron between the first and the following constant exons. This sequence can be considered an exonic remnant of the ancestral second constant exon, since it retains the canonical splicing sites and is regularly included in the mature transcript. These data shed further light on the distinctive features characterizing the adaptive evolution Antarctic fishes underwent, providing also new insights into the complexity of teleost immunoglobulin genes.

2. Materials and methods

2.1. Biological samples

Specimens of the species *T. bernacchii* Boulenger, 1902 were collected by use of gill nets or traps in Tethys bay (Ross Sea) at 74° 41′ S, 164° 26′ E, out of the specially protected areas, during the XXV Italian Antarctic Expedition (2009–2010). The activity permit, released by Italian National Program for Antarctic Research (PNRA), was in agreement with the "Protocol on environmental protection to the Antarctic Treaty" Annex V. Intestines were collected and immediately frozen in liquid nitrogen.

In addition, *B. diacanthus* Carmichael, 1819 specimens were collected during the ICEFISH (International Collaborative Expedition to collect and study Fish Indigenous to Sub-Antarctic Habitats) expedition 2004. Spleens were dissected from adult specimens and frozen immediately. Tissues from both species were stored at −80 °C until use.

2.2. RNA extraction and PCR amplification of IgT heavy chain cDNA

Total RNA was extracted from the intestine of *T. bernacchii* using SV Total RNA Isolation System kit (Promega), starting with 150–200 mg of tissue, following the manufacturer's instructions. Tissue was homogenized by Ultra Turrax T-25 IKA-Werke, performing three cycles of 30 s each at 21.500 rpm on ice. The same method was performed for *B. diacanthus* spleen. The RNA obtained, whose quality was checked on agarose gel and measured by NanoDrop 1000 Spectrophotometer (Thermo Scientific), was subjected to reverse transcription using Transcriptor First Strand cDNA Synthesis Kit (Roche). To accomplish the first PCR amplification of cDNA, the oligonucleotides used as primers were designed on conserved regions of IgT heavy chain nucleotide sequences available in GenBank from *Epinephelus coioides* (GU182366) and *Siniperca chuatsi* (DQ016660). The amplification was performed with DreamTaq DNA polymerase (Thermo Scientific) as follows: 95 °C for 5 min, 40 cycles of 95 °C (30 s), 58 °C (30 s), and 72 °C (1 min) with a final extension at 72 °C for 10 min. Where necessary, the PCR product was then subjected to a second amplification, carried out in the same conditions as the primary PCR, in order to increase the amount of the specific amplicon. Oligonucleotides used in all PCR experiments reported in this work are listed in Table 1. The PCR products were then analyzed on 1% agarose gel, purified by DNA Gel Extraction kit (Agilent), and cloned into pGEM®-T Easy Vector (Promega). The sequence of positive clones, identified by blue/white selection, was determined with an ABI PRISM 3100 automated sequencer at Eurofins Genomics S.r.l. (Vimodrone, Milan, Italy).

Please cite this article as: Giacomelli, S., et al., New insights into evolution of IgT genes coming from Antarctic teleosts, Mar. Genomics (2015), http://dx.doi.org/10.1016/j.margen.2015.06.009