 doi:10.20944/preprints202202.0211.v2

Article

Comparative analysis of the pIgR gene from the Antarctic tele-ost Trematomus bernacchii reveals distinctive features of cold adapted Notothenioidei

Alessia Ametrano[1], Simona Picchietti[2], Laura Guerra[2], Stefano Giacomelli[1], Umberto Oreste[1], Maria Rosaria Coscia[1],*

[1] Institute of Biochemistry and Cell Biology, National Research Council of Italy, Via P. Castellino 111, 80131 Naples, Italy; alessia.ametrano@ibbc.cnr.it (A.A.); stefano77g@hotmail.it (S.G.); umberto.oreste@ibbc.cnr.it (U.O.); mariarosaria.coscia@ibbc.cnr.it (M.R.C)
[2] Department for Innovation in Biological, Agro-food and Forest Systems, University of Tuscia, Largo dell'Università, 01100 Viterbo, Italy; picchietti@unitus.it (S.P.); lauraguerra@unitus.it (L.G.)

* Correspondence: mariarosaria.coscia@ibbc.cnr.it; Tel: +0039 081 6132556 (M.R.C.)

Abstract: The IgM and IgT classes were previously identified and characterized in the Antarctic teleost *Trematomus bernacchii*, a species belonging to the Perciform suborder Notothenoidei. Herein we characterized the gene encoding the polymeric immunoglobulin receptor (pIgR) in the same species and compared it to pIgR of multiple teleost species belonging to five perciform sub-orders, including 11 Antarctic and one non-Antarctic (*Cottoperca gobio*) notothenioid species, the latter living in less cold periantarctic sea. Antarctic pIgR genes displayed particularly long in-trons marked by sites of transposable elements and transcription factors. Furthermore, analysis of *T. bernacchii pIgR* cDNA unveiled multiple amino acid substitutions unique to Antarctic species, all introducing adaptive features, including N-glycosylation sequons. Interestingly, *C. gobio* shared most features with the other perciforms rather than with the cold adapted relatives. T. bernacchii pIgR transcripts were predominantly expressed in mucosal tissues, as indicated by q-PCR and in situ hybridization analysis. These results suggest that in cold adapted species pIgR preserved its fundamental role in mucosal immune defense, although remarkable gene structure modifications occurred.

Keywords: pIgR; gene structure; cold environment; gene expression; teleost immunity; adaptive evolution; mucosal tissues; genome alteration; Notothenioidei; IgV domains.

1. Introduction

The polymeric immunoglobulin receptor (pIgR) emerged early during evolution in teleost fishes [1] and coevolved with mucosal Ig isotypes, ensuring a mucosal protection [2]. It has a conserved structure, consisting of an extracellular region composed of varying numbers of IgV domains, increasing across the evolutionary scale [3], a transmembrane region and a cytoplasmic tail [4]. In mammals, pIgR has five IgV domains, D1-D5, except for the bovine and rabbit IgRs having three (D1, D4, and D5) of the five domains, derived from alternative splicing [5, 6]. In birds, reptiles and amphibians, pIgR consists of four domains, corresponding to mammalian D1, D3, D4 and D5, respectively [7-9]. Fish pIgR shows the simplest topology, based on comparative sequence analyses, comprising only two Ig-like domains, which are homologous to mammalian D1 and D5 [10]. Teleost pIgR is known to be expressed in mucosa-associated lymphoid tissues, e.g., intestine, gills, skin, buccal and pharyngeal cavity, and olfactory system [11]. pIgR has been shown to bind both IgM and IgT, although the exact Ig-binding site of pIgR has not yet been clarified, given that

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 27 June 2022 doi:10.20944/preprints202202.0211.v2

polymeric Igs (pIgs) are devoid of a J chain, contrary to cartilaginous fish in which the J chain has been identified [12].
Over recent years, great attention has been paid to the function of fish pIgR, while studies aimed at examining its gene structure are still limited. Full-length transcripts from the pIgR gene were characterized in various teleost species belonging to different orders [11], and the binding sites of cytokine-inducible regulatory elements, well known in mammals, were also predicted and considered as potential regulators of the transcription of pIgR in teleost fish [13-15].
At present, no data are available on pIgR from teleost species living under extreme conditions, such as Notothenioidei (Perciform suborder) that represent the prevalent component of the Antarctic fish fauna. During their evolutionary history, Nototheniodei have undergone extraordinary challenges to adapt to the constantly cold marine environment of Antarctica. The Antarctic notothenioid families have been proposed as a superfamily, named Cryonotethenioidea [16], to be distinguished from the non-Antarctic families that are considered the most phyletically basal branch, having remained in periantarctic seawaters, under temperate conditions [17, 18]. This taxonomic group has long been considered an attractive model to study biochemical, physiological and morphological adaptations, although poorly investigated at a molecular level. In previous studies, we investigated the genes encoding IgM and IgT isotypes in several cold-adapted and temperate notothenioid species [19-21] and highlighted the first evidence for a pos-sible hepato-biliary transport of Ig in the Antarctic species *Trematomus bernacchii* [22].
The most recent advances in collecting omics data from notothenioid fish provided a source of fundamental information about molecular and genetic features, allowing conduct of also evolutionary studies on notothenioids in comparison with other perciform species. Thus, the main goal for this study is to investigate the specificities of the pIgR gene related to evolutionary adaptation, through a comparative analysis, based on the genomes and transcriptomes available for T. bernacchii and for 11 Antarctic species belonging to the same suborder as *T. bernacchii* (Notothenoidei). In particular, we ex-tended the analysis to *Trematomus loenbergii*, *Dissostichus eleginoides*, *Dissostichus mawsoni*, *Notothenia coriiceps* (family Nototheniidae), *Harpagifer antarcticus* (family Harpagiferidae), *Gymnodraco acuticeps* (family Bathydraconidae), *Pseudochaenichthys georgianus*, *Chaenocephalus aceratus*, *Chionodraco myersi* and *Chionodraco hamatus* (family Channichthyidae), all adapted to live in the extreme environment of Antarctica. Moreover, another notothenioid species, *Cottoperca gobio*, belonging to the family Bovichtidae, the ancestral notothenioid family living in more temperate periantarctic seawaters, was added for comparison. Additionally, 26 perciform species, belonging to five different families, were included for comparative analysis. Finally, the expression of the *T. bernachii pIgR* gene was evaluated through q-PCR and *in situ* hybridization (ISH), which allowed transcript localization.
Taken together, these findings underline several peculiar features that may be considered as the hallmark of cold pIgRs and underpin the primary role of pIgR in mucosal immune response and host protection in a cold adapted teleost species.

2. Results

2.1. Analysis of T. bernacchii pIgR gene locus

The whole pIgR genomic sequence (8310 nt) was retrieved as a single-copy gene from the *T. bernacchii* genome (GenBank assembly accession: NW_022987689) by using NCBI Genome Data Viewer. The *pIgR* gene structure consists of eight exons, interrupted by seven introns. The first exon includes the 5’ UTR and encodes the leader peptide (245 nt), the second exon (345 nt) encodes the D1 domain; the third exon (82 nt) and the 5’ 208-nt of the fourth exon encode the D2 domain; the 3’ end of the fourth (19 nt) exon, the fifth (88 nt) exon, and the 5’- most first end of

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 27 June 2022 doi:10.20944/preprints202202.0211.v2

the sixth (10 nt) exon encode the Extracellular Membrane Proximal Domain (EMPD); 5' end of the sixth exon (51 nt) encodes the transmembrane domain (TM); the 3' end of the sixth exon (31 nt) along with the seventh (50 nt) and eighth (81 nt) exons encode the cytoplasmic region. The terminal sequence is 348 nt, including the stop codon (Figure S1). Along with those found in the 3' UTR, an alternative poly-adenylation signal was identified in the fourth intron, and its functionality was assessed using the bioinformatic tool Poly(A) Signal Miner [23]. Although this additional site was predicted to have a low confidence score compared to that of the other two, it might be inferred a possible involvement in the transcription of the messenger RNA encoding the secreted form of pIgR.

Regarding the gene locus organization, the *T. bernacchii pIgR* gene is flanked by the dad1 (5383 nt upstream) and lrrc24 (21790 nt downstream) genes, as shown in zebrafish and other teleost species [15]. In addition to the pIgR gene, a further pIgR-like gene of 6481 nt was identified in another genomic scaffold (NW_022988066.1).

To investigate the evolution of the pIgR gene locus, we performed a comparative analysis by considering two Antarctic species belonging to the same suborder Notothenoidei as *T. bernacchii* (family Nototheniidae), *G. acuticeps* (family Bathydraconidae) and *P. georgianus* (family Channichthyidae), and the non-Antarctic species, *C. gobio* (family Bovichtidae), the latter living in more temperate periantarctic seawater. This analysis was possible due to the considerable amount of sequencing data obtained from the Antarctic notothenioid fish genomes project at Sanger Institute, aimed at deepening the very relevant topic of molecular adaptation to extreme conditions.

To identify specific features of the notothenioid *pIgR* gene, we searched databases for homologous sequences from multiple temperate species belonging to other perciform suborders. Twenty-six perciform species, belonging to five different suborders and used for comparative analysis are indicated in Table S5 (see also Material and Methods section).

As a first step in the characterization of the *pIgR* gene locus, we investigated the intronic regions. The second and third introns were significantly longer in *T. bernacchii* as well as in the other two Antarctic species than those of the temperate species considered, ranging from 2372 to 3080 nt, and from 538 to 552 nt, respectively (Table 1). Interestingly, in the case of the non-Antarctic notothenioid *C. gobio*, we observed that the size of the third intron was in line with that of temperate species. Conversely, the other introns were found to be conserved in length across the species considered (Table 1).

Table 1. The polymeric Ig receptor (*pIgR*) gene size and intron length (nt) in the perciform suborders Notothenioidei, Percoidei, Serranoidei, Scorpaenoidei and Cottoidei.

Suborder	Species	*pIgR* gene size	1st intron	2nd intron	3rd intron	4th intron	5th intron	6th intron	7th intron
Notothenioidei	*Trematomus bernacchii*	8310	628	2372	548	2825	124	146	111
	Gymnodraco acuticeps	8235	583	2866	552	2906	124	85	111
	Pseudochaennichthys georgianus	9013	582	3080	538	2939	133	85	111
	Cottoperca gobio	5930	596	1519	204	1713	156	129	115
Percoidei	*Sander lucioperca*	4818	649	124	211	1080	161	155	115
	Perca fluviatilis	3976	590	126	211	1237	171	137	122
	Perca flavescens	6669	554	125	211	1229	171	156	115
	Etheostoma spectabile	5939	579	124	153	1014	143	137	115
	Etheostoma cragini	3700	595	124	207	985	143	145	117
Serranoidei	*Epinephelus lanceolatus*	5831	588	117	206	2969	163	93	107
	Plectropomus leopardus	7090	545	1304	183	2135	161	150	118
					208	2631	634	133	118

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 27 June 2022 doi:10.20944/preprints202202.0211.v2

Scorpaenoidei	*Sebastes umbrosus*	6887	601	779					
Cottoidei	*Pungitius pungitius*	5126	579	461	196	872	136	167	96
	Anarrhichthys ocellatus	5432	595	593	205	1369	176	153	106
	Gasterosteus aculeatus	4095	577	274	195	1210	89	159	99

Thus, we conclude that the larger introns accounted for the larger size of the *pIgR* gene of *T. bernacchii a*s well as the other two Antarctic species analyzed.
Given that the *T. bernacchii* pIgR gene carried extraordinary long introns, we searched for factors possibly accounting for this modification by using the MEME tool [24], which allows the identification of conserved motifs repeatedly occurring in a sequence dataset. A comparative analysis, conducted on each of the seven introns from notothenioid and temperate species, allowed the identification of distinct conserved sequence motifs in Antarctic species (Figure 1). In particular, we found two regions, named motif 4.2 (50 nt) and motif 5.2 (48 nt) exclusively at the 5' end of the second intron of Antarctic *pIgR* (Figure 1a). Other motifs (motif 1.3, motif 2.3, motif 3.3 and motif 4.3), found in the third intron, were shared by most species, including notothenioids (Figure 1b). Of note, motif 5.3, 50 nt in length, repeated in tandem on both DNA strands of the third intron represented a distinctive feature of Antarctic fish, as both were absent in the non-Antarctic relative *C. gobio* and in all the other perciforms analyzed.

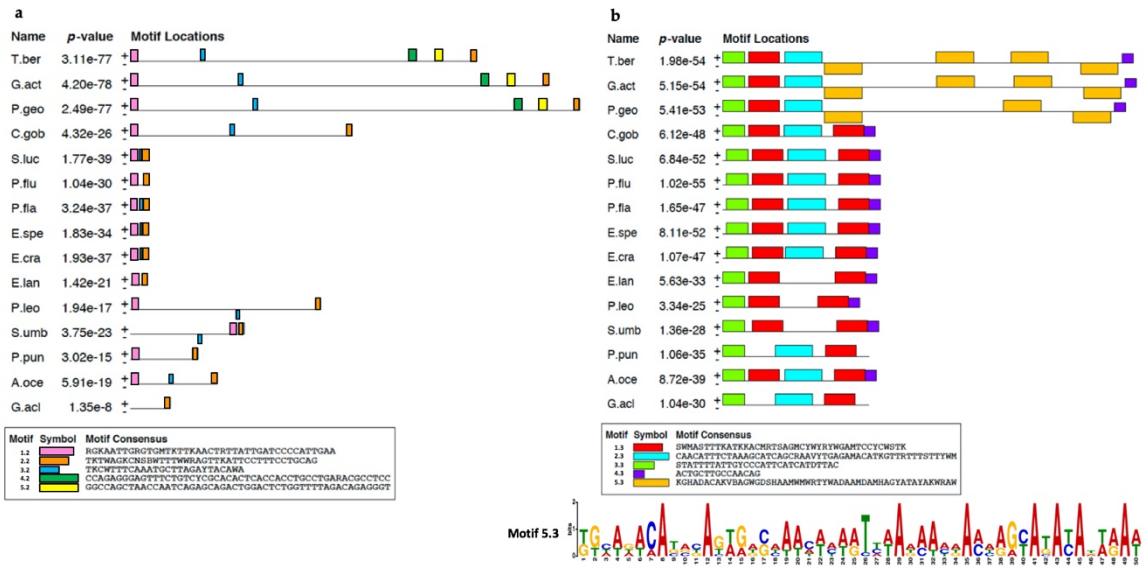


Figure 1. Conserved sequence motifs (colored boxes) identified by MEME in the second (a) and third introns (b) of pIgR from *Trematomus bernacchii* (T.ber), *Gymnodraco acuticeps* (G.act), *Pseudochaennichthys georgianus* (P.geo), *Cottoperca gobio* (C.gob), *Sander lucioperca* (S.luc), *Perca fluviatilis* (P.flu), *Perca flavesvens* (P.fla), *Etheostoma spectabile* (E.spe), *Etheostoma cragini* (E.cra), *Epinephelus lanceolatus* (E.lan), *Plectropomus leopardus* (P.leo), *Sebastes umbrosus* (S.umb), *Pungitius pungitius* (P.pun), *Anarrhichthys ocellatus* (A.oce), *Gasterosteus aculeatus* (G.acl). The WebLogo representation of motif 5.3 (ocher box) is reported at the bottom of panel b.

To elucidate other reasons to account for the presence of the very long intronic regions of pIgR in *T. bernacchii* and in the other Antarctic species, we searched for transposable elements (TEs), which are known to cause genome modifications. We used the RepeatMasker software [25], which is based on the annotations available at Ensembl and compares data against a curated library of repeats using local alignment methods. A systemic analysis of the distribution of TE elements in each intron in all Perciformes indicated that the second intron in *T. bernacchii* was

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 27 June 2022 doi:10.20944/preprints202202.0211.v2

characterized by the presence of one type of TEs (Figure 2a), while the third intron appeared more heterogeneous due to the presence of SINEs and LINEs, the same as found in *P. georgianus*, or similarly to *G. acuticeps*, which showed LINEs and LTR elements (Figure 2b). Moreover, only the second intron of the *C. gobio* pIgR gene was characterized by the presence of DNA transposons, like the *T. bernacchii* second intron (Figure 2a).

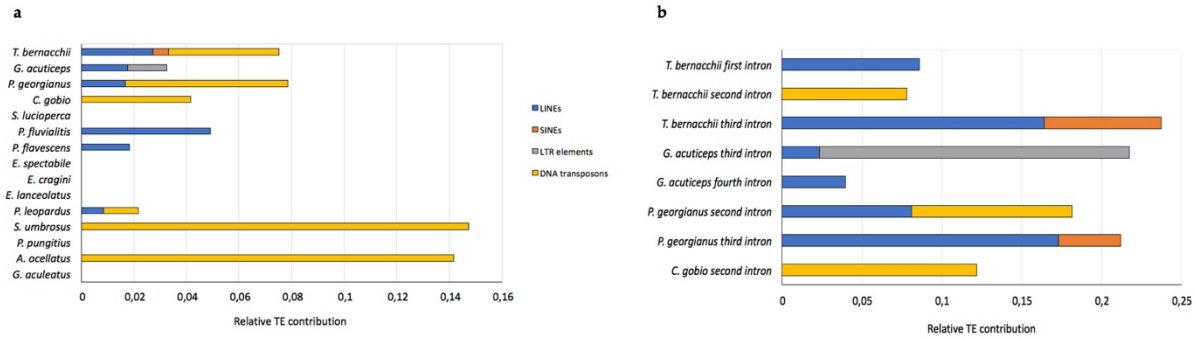
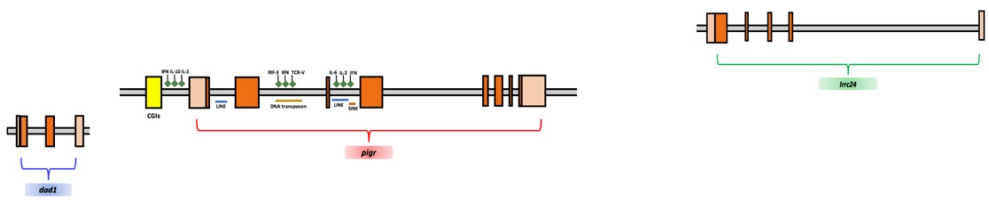


Figure 2. Presence of transposable elements (TEs) in the *pIgR* gene from notothenioids and other perciform species. (a) Distribution of TEs in overall pIgR intronic sequences of representatives of each perciform suborder; (b) Distribution of TEs found in each of the first four pIgR introns of Antarctic and non-Antarctic species.

To assess whether such long intronic sequences found in the *T. bernacchii pIgR* gene could influence its regulation, we investigated the presence of regulatory signals in the 5' flanking region and in all introns by using Tfsitescan [26]. Interestingly, in the second and third introns along with the 5' flanking region, we identified several putative transcription-factor binding sites, sharing a very high score (up to 12), many of them involved in innate immune response, e.g., IL-10 and IFN (Table S1). Similar results were obtained for the other two Antarctic species, *G. acuticeps* and *P. georgianus*. No statistically significant sites were found in the fourth intron either of the notothenioids nor in the intronic regions of the temperate species, including the non-Antarctic species.

Of special interest was the search of CpGs islands (CGIs), known as indicators of transcription promoter sequences [27]. Using the UCSC Genome Browser on *T. bernacchii* and on the two other Antarctic species assemblies, we detected a CGI with lengths ranging from 210 to 348 nt, harboring 21-35 CpGs, located about 500 nt upstream of the 5' flanking region of the pIgR gene (Figure S1). We found that the ObsCpG/ExpCpG ratio ranged from 1.35 to 1.50 in the Antarctic pIgR gene locus. Remarkably, no CGIs were found in the non-Antarctic relative *C. gobio*, nor in the temperate counterparts. The main features of the *T. bernacchii pIgR* gene locus described above are summarized in Figure 3.



Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 27 June 2022 doi:10.20944/preprints202202.0211.v2

Figure 3. Schematic representation of the organization of the pIgR gene locus in *T. bernacchii.* Exons are indicated in orange boxes, promoter and terminal sequences are in bright orange, the CpG islands (CGIs) in yellow. Putative transcription-factor binding sites (green diamonds) and TEs (blue, ocher and brown lines) are indicated above and below the scheme, respectively. dad1 and lrrc24 flanking genes are also reported.

2.2. Analysis of T. bernacchii pIgR cDNA

Initially, a partial cDNA sequence of the pIgR was obtained from total RNA extracted from the spleen of a *T. bernacchii* specimen as described in the Methods section. The primers used for PCR experiments were designed on the nucleotide sequences encom-passing the pIgR D1-D2 domains from *Epinephelus coioides*. The amplicon obtained (529 nt) was cloned and sequenced.
Subsequently, to extend the 5’ end of cDNA, a 5’ RACE was performed by using the gene specific primer pIGR1Rev, designed at the beginning of the D1 domain; the antisense primer pIGRIIr, designed in the middle of the D2 domain, was used for the nested PCR amplification. The amplicon obtained (235 nt) was cloned and sequenced.
To complete the sequence at the 3’ end of the pIgR transcript, a 3’ RACE was carried out by using pIGR1Fwd as a sense specific primer. To verify the correct amplification, a nested PCR reaction was performed, using pIGRII as a sense primer. The amplicon ob-tained (650 nt) was cloned and sequenced.
The full-length cDNA sequence encoding *T. bernacchii* pIgR consisted of 1414 nt, including a 5' UTR (38 nt) and a 3' UTR (359 nt) (Figure 4). Two polyadenylation signals were identified, one canonical, at position 1387, and a second one, non-canonical, at position 1395 (Figure 4).
Thirty positions were polymorphic because differences were found in the transcript variants. Five carried non-synonymous substitutions. Interestingly, a fairly critical mutation was identified in the fourth position of the canonical polyadenylation signal, which blocks its function. In this circumstance, the non-canonical poly(A)-site more likely becomes functional.
The functional protein domains were identified in the deduced amino acid sequence using the bioinformatics tools SignalP, Prosite, and TMpred. Following a 21-aa long signal peptide, two immunoglobulin domains, D1 and D2 were found, consisting of 108 and 95 residues respectively, separated by a short linker sequence (Figure 4). The D1 and D2 domains were both identified as IgV domains, being the two conserved cysteine residues forming the intrachain disulfide bond spaced by 68 aa (D1) or 62 aa (D2), which is a distance greater than that found in IgC-and IgI- type domains. Furthermore, the presence of two additional cysteine residues, spaced by seven amino acid residues in both domains, is a typical feature of the IgV domains of pIgRs. It was noted that, while D1 is encoded by a single exon, the D2 nucleotide sequence comprises also one intron (Figure S1), which is an unusual feature for IgV domains. A 39-aa extracellular proximal domain (EMPD) sequence, with a high theoretical pI (11.0), rich in prolines (17.9%), was preceding, at the carboxy-terminus of D2, the transmembrane domain (TM). The latter consisted of 20 residues and was characterized by a preponderance of leucine residues (30%) and by the presence of a cysteine residue, very infrequent in transmembrane proteins (Figure S2-S3). The sequence ended with a 51-aa long basic cytoplasmic tail (Figure 4). The amino acid composition of *T. bernacchii* pIgR is reported in Table S2.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 27 June 2022 doi:10.20944/preprints202202.0211.v2

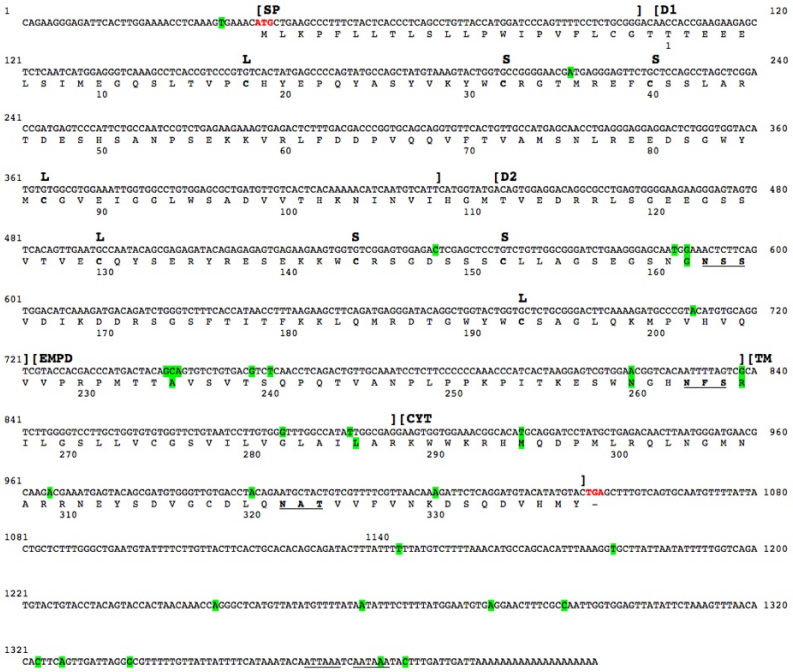


Figure 4. Nucleotide and deduced amino acid sequences of the gene encoding *T. bernacchii* pIgR. The start and the stop codons are given in bold red, the polyadenylation signals, located down-stream of the stop codon, are underlined. The boundaries of the signal peptide (SP), the D1, D2 and EMPD domains, and the transmembrane (TM) and cytoplasmic (CYT) regions are indicated by square brackets above the nucleotide sequence. Polymorphic sites are highlighted in green. Amino acid residues are reported in one-letter code below the nucleotide sequence. The large (L) and the small (S) loops are indicated above codons for the cysteine residues forming them. N-glycosylation sites defined by the NXS/T sequons are bold and underlined. Numbering of nucleotides and of amino acids residues is reported above and below each sequence, respectively.

Once the cDNA sequence was obtained, it was aligned against the two pIgR transcript variants X1 and X2, predicted from the reference genome (Figure 5).

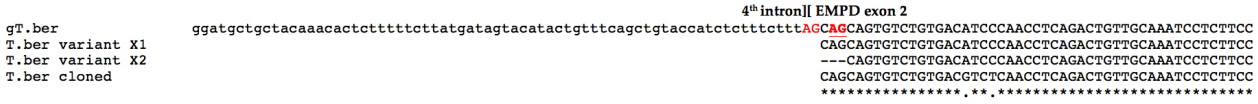


Figure 5. Alignment of the *T. bernacchii* pIgR genomic sequence with that of the three transcript variants. Of the pIgR genomic sequence (gT.ber), retrieved from the *T. bernacchii* genome, only the region spanning the 3' end of the fourth intron and the 5' end of the EMPD exon 2 is shown. Two pIgR variants (T.ber variant X1 and T.ber variant X2) are predicted from the genome; the pIgR cDNA variant (T.ber cloned) has been isolated in the present study. The canonical acceptor splicing site is depicted in red. The cryptic splicing site is in bold red and underlined. Gaps are indicated by dashes. Identical nucleotides are indicated with an asterisk below the alignment; positions differing in one nucleotide are marked in dot.

The cloned sequence that resulted was identical to the variant X1, whereas the variant X2 had a deletion of a triplet that encodes an alanine, in the 5' end of the EMPD exon 2. Therefore, we evaluated putative alternative exon isoforms by applying a computational tool for the identification of potential splice sites. We highlighted that the site at position 73 in the region encompassing the 3' end of the fourth intron and the 5' end of the EMPD exon 2 showed the highest score as a cryptic acceptor splicing site, thus accounting for the variant X2 (Table S3).
To further extend the analysis of the pIgR gene structure in cold adapted species, two more Antarctic species, *G. acuticeps* and *P. georgianus*, were added along with the non-Antarctic species *C. gobio*, since complete reference genomes were available for all of them. Two (in *P. georgianus*)

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 27 June 2022 doi:10.20944/preprints202202.0211.v2

or three (in *G. acuticeps*) pIgR transcripts, referred to as variant X1, variant X2 and variant X3, had been annotated in the respective genome assemblies (Figure 6).

a

```
                          EMPD exon 1                                ][EMPD exon 2
T.ber variant X1   TGCATGTGCAGGTCGTACCACGACCCATGACTACAGCAGTGTCTGTGACATCCCAACCTC
T.ber variant X2   TGCATGTGCAGGTCGTACCACGACCCATGACTACAG---TGTCTGTGACATCCCAACCTC
G.act variant X1   TGCATGTGCAGGTCAAACCACGACCCATGACTACAG---TGTCTGTGACATCCCAACCTC
G.act variant X2   TGCATGTGCAGGTCAAACCACGACCCATGACTA------TGTCTGTGACATCCCAACCTC
G.act variant X3   TGCATGTGCAGGTCAAACCACGACCCATGACTACAG---TGTCTGTGACATCCCAACCTC
P.geo variant X1   TGCATGTGCAGGTCAAACCACGACCCATGACTACAGCAGTGTCTGTGACATCCCAACCTC
P.geo variant X2   TGCATGTGCAGGTCAAACCACGACCCATGACTACAG---TGTCTGTGACATCCCAACCTC
C.gob variant X1   TGCATGTGCAGGTCACACCCCGACCCACGACTACAGCAGTGTCTGTGACATCCCCGCCTC
C.gob variant X2   TGCATGTGCAGGTCACACCCCGACCCACGACTACAG---TGTCTGTGACATCCCCGCCTC
                   **************  *** ******* *****      ***************  ****
```

b

```
T.ber isoform X1   HVQVVPRPMTTAVSVTSQP
T.ber isoform X2   HVQVVPRPMTT-VSVTSQP
G.act isoform X1   HVQVKPRPMTT-VSVTSQP
G.act isoform X2   HVQVKPRPMT--MSVTSQP
G.act isoform X3   HVQVKPRPMTT-VSVTSQP
P.geo isoform X1   HVQVKPRPMTTAVSVTSQP
P.geo isoform X2   HVQVKPRPMTT-VSVTSQP
C.gob isoform X1   HVQVTPRPTTTAVSVTSPP
C.gob isoform X2   HVQVTPRPTTT-VSVTSPP
                   **** ***:*  .****:*
```

Figure 6. Multiple alignments of the EMPD sequences. (a) Multiple alignment of cDNA sequences spanning the two EMPD exons of the pIgR variants identified in the Antarctic species *T. bernacchii* (T.ber variants X1 and X2), *G. acuticeps* (G.act variants X1, X2 and X3), *P. georgianus* (P.geo variants X1 and X2), and in the non-Antarctic species *C. gobio* (C.gob variants X1 and X2). The exonic region in which a cryptic splicing site was predicted is depicted in bold red and underlined; (b) Multiple alignment of deduced amino acid residues of the EMPD domain. Amino acid residues involved in the usage of the cryptic splice site are depicted in bold red. Gaps are indicated by dashes. Identical nucleotides or amino acid residues are marked with an asterisk below alignments; positions differing in one nucleotide or amino acid residue are marked with a dot.

As previously shown by the AT content analysis of notothenioid IgT genes [21], a remarkably high AT content of *pIgR* exons was shared within Antarctic species, reaching a peak (55.8 %) in the family Nototheniidae, which comprises *T. bernacchii* and *Dissostichus eleginoides* (Figure 7). This finding reflects a peculiar feature not shared with either the non-Antarctic species *C. gobio*, their closest relative, nor with all the other Perciformes.

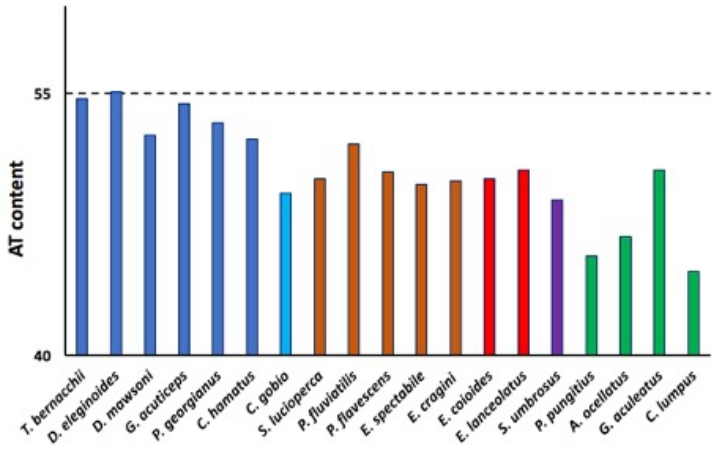


Figure 7. Percentage of AT content of the pIgR exons of Antarctic species (blue bars) and the non-Antarctic notothenioid species *C. gobio* (light blue bar) compared to representative species of the temperate perciform suborders Percoidei (brown bars), Serranoidei (red bars), Scorpaenoidei (purple bar), Cottoidei (green bars).

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 27 June 2022 doi:10.20944/preprints202202.0211.v2

2.3. Analysis of T. bernacchii pIgR deduced amino acid sequence

To identify adaptive characteristics unique to notothenioid pIgR sequences, we extended our analysis to other notothenioid species and to representatives of the five perciform suborders, on the basis of the available data (see Material and Methods Section 4.1). However, since many sequences resulted fragmentary and/or incomplete, for convenience, the overall comparisons were only referred to the overlapping regions.
Some conserved motifs, previously identified in other fish and suggested to help stabilize the secondary structure of pIgR [14], were also found in D1 (CWDC, KYWC and DxGxYxC motifs) and D2 of *T. bernacchii* pIgR (KxWC and DxGWYWC) (Figure 8).

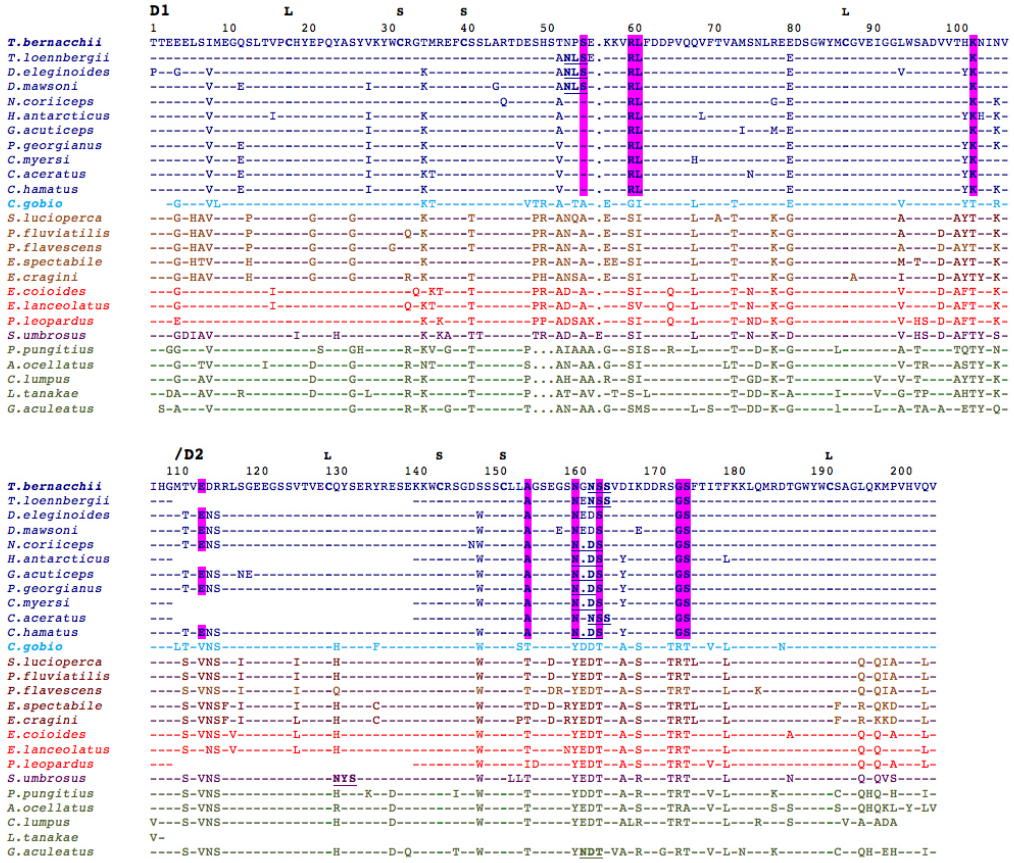


Figure 8. Multiple amino acid sequence alignment of pIgR D1-D2 domains from Notothenioidei (in blue; *T. bernacchii* in bold blue and *C. gobio* in bold light blue) and other species of the perciform suborders Percoidei (in brown), Serranoidei (in red), Scorpaenoidei (in purple), and Cottoidei (in green). Antarctic notothenioid specific residues are reported in bold and highlighted in magenta. The highly conserved cysteine residues characteristic of the immunoglobulin fold are in bold. The large (L) and the small (S) loops are indicated above the cysteine residues forming them. Putative N-glycosylation sites are in bold and underlined. Amino acid residues that are iden-tical to those shown in the sequence of *T. bernacchii* are indicated by dashes. Gaps are indicated by dots. The complete alignment is shown in Figure S3.

Twenty amino acid positions were found to be specific for Antarctic pIgRs, as present in all Antarctic sequences, but absent in all sequences from the other species analyzed, as well as in the non-Antarctic notothenioid *C. gobio* (Figure S3). Ten were present in the secretory component; four positions (E55, R60, L61 and K103) were localized in the D1 domain, six in D2 (E114, A155, N161, S164, G174 and S175) (Figure 8), one in the EMPD, three in the TM domain and six in the cytoplasmic tail (Figure S3). In addition, it is noteworthy that five out of the ten notothenioid-specific residues, which are present in the extracellular portion of the receptor, introduce or abolish an electrostatic charge; N161 and S164 are convergent substitutions since introducing a

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 27 June 2022 doi:10.20944/preprints202202.0211.v2

glycosylation sequon NXS/T (being X different from P). Two additional residues (E80 and T112) were shared by Antarctic and non-Antarctic nothotenioid species, but they were absent in the other perciform suborders (Figure 8).
Glycosylation was found to be another distictive feature of the notothenioid pIgR. Up to four N-glycosylation sequons were found in Antarctic fish compared to the other perciform species, which showed no sequons at all, except for *Sebastes umbrosus* and *G. aculeatus*, harboring just one (Figure 9). Interestingly, the site in D2, which is present in 9 out of 12 sequences, is alternately located at two asparagines spaced by a single residue (Figure 8).

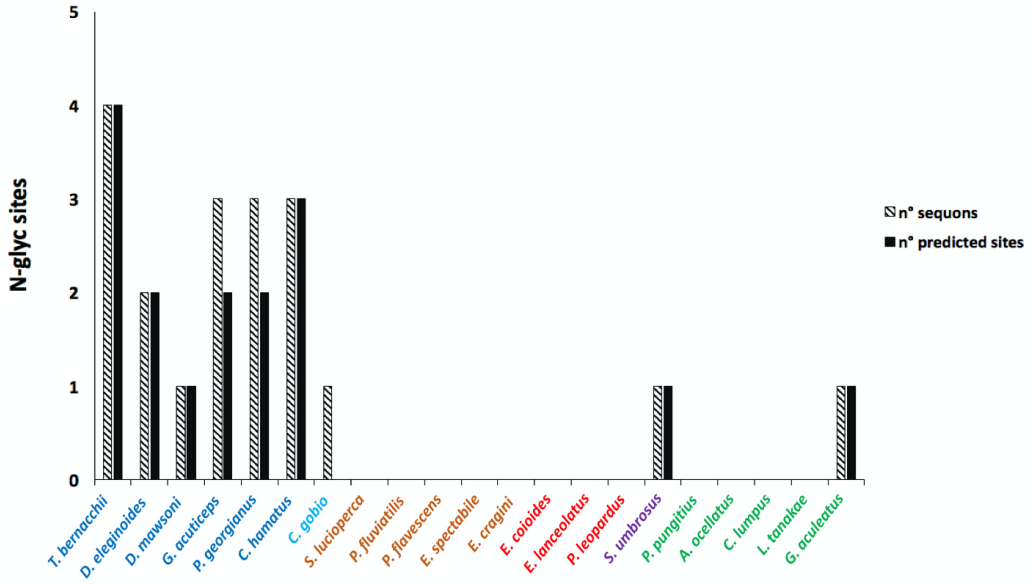


Figure 9. Distribution of potential N-glycosylation sequons (diagonal black bars) and glycosylated sequons (black bars) in the pIgR of Antarctic (blue) and non-Antarctic (light blue) notothenioid species, compared to representative species of the temperate perciform suborder Percoidei (brown), Serranoidei (red), Scorpaenoidei (purple), Cottoidei (green).

Interestingly, all the sequons identified were predicted to be glycosylated in most Antarctic species, except for two out of three in *G. acuticeps* and *P. georgianus*. Of the four notothenioid N-glycosylation sites, one occurs in each of the D1 and D2 domains, one at the boundary between the TM region and the extracellular portion, another one is in the cytoplasmic tail (Figure S3). The distance tree generated from the multiple alignment clearly confirms phylogenetic relationships among the species analyzed (Figure 10).

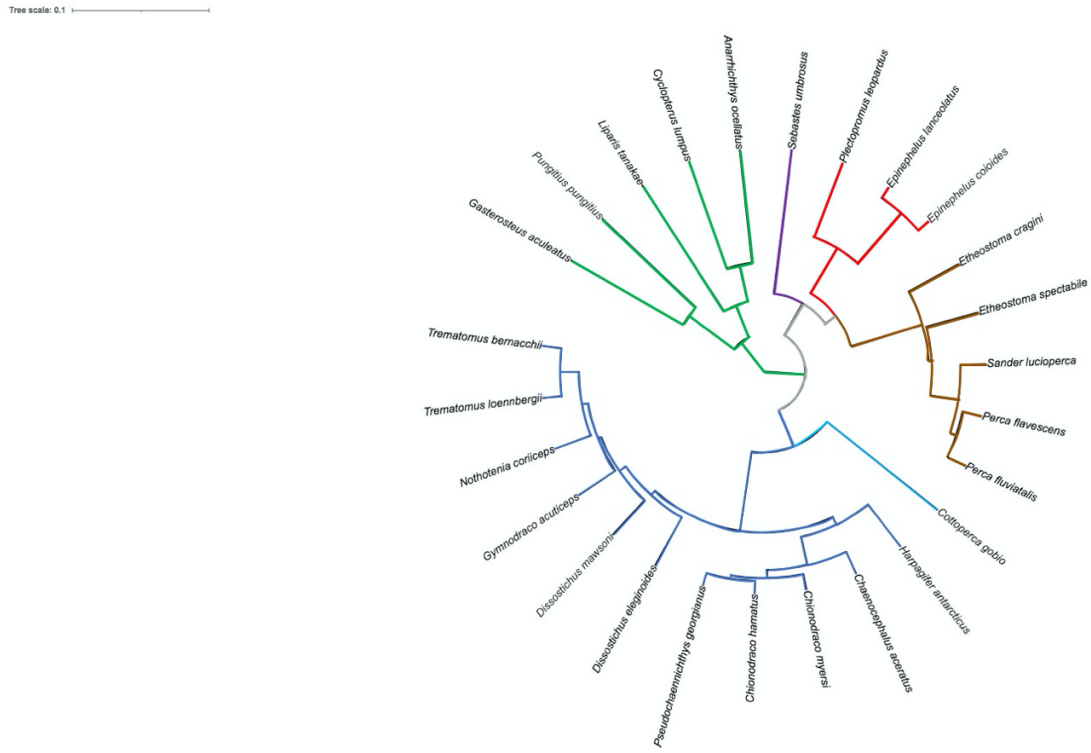
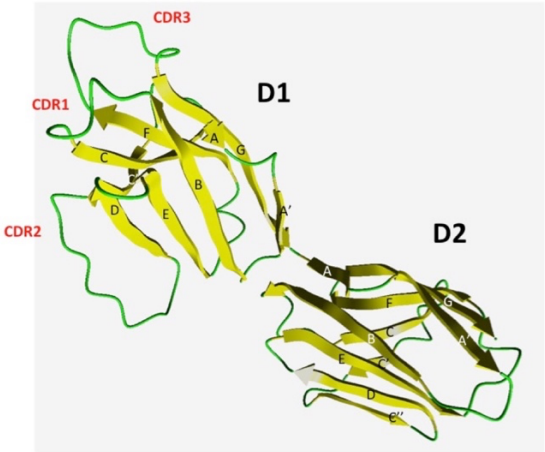


Figure 10. Distance tree of the pIgR D1 domain from teleost species belonging to the perciform suborders Notothenioidei (Antarctic species, blue lines; non-Antarctic species, light blue line), Percoidei (brown lines), Serranoidei (red lines), Scorpaenoidei (purple lines), and Cottoidei (green lines). The tree was generated by the Clustal Omega tool. The sequences used are as in Figure 8.

2.4. Structural analysis of T. bernacchii pIgR

We constructed a molecular model of the ectodomain of *T. bernacchii* pIgR. It consists of two tandem IgV domains, D1 and D2, whose axes diverge at about 120°. Differently from other IgV domains, a disulfide bridge connects C and C' strands in both D1 and D2 domains. The regions that are defined as complementary determining regions (CDRs) in the antibody VH and VL domains were recognized in the D1 domain (Figure 11). D1 CDR loops are more extended, while the structure of D2 appears more compact and contains an additional C'' strand. Minor structural differences between trout, selected as template strcuture, and T. bernacchii are related to the size of D1 CDR2, which is longer in T. bernacchii (Figure 11).



Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 27 June 2022 doi:10.20944/preprints202202.0211.v2

Figure 11. Ribbon representation of the molecular model of *T. bernacchii* pIgR extracellular region. The β-strands are labeled with uppercase letters. Loops that correspond to the three Complementary Determining Regions in the D1 domain are labeled (CDR) and colored (green).

The N-glycosylation sites of D2 were exposed to solvent (Figure S4), suggesting a role of the attached carbohydrate moiety. Notably, all substitutions introducing electrostatic charges in notothenioids were also found to expose the side chain to the solvent as well as the glycosylated N161, suggesting that the solubility of the molecule increases at the very low temperature of Antarctic seawater. The position of the Antarctic species-specific charged residues R60, K103, E114 and N161 is shown in Figure S4.

2.5. Basal expression analysis of pIgR transcripts in T. bernacchii mucosal tissues and lymphoid organs

To gain some insights into the constitutive expression of *pIgR* in a cold adapted teleost, a relative mRNA expression pattern of the gene was determined by q-PCR in mucosal tissues and lymphoid organs of *T. bernacchii*.
As shown in Figure 12, pIgR transcripts were expressed in all the tested tissues, with the highest abundance detected in gills and the lowest in muscle, as expected. Higher levels were also found in the intestinal segments. In particular, the mRNA levels in the middle intestine were 2.7-fold (adjusted $p < 0.01$) and 3.9-fold (adjusted $p < 0.05$) lower than those in the anterior and posterior ones, respectively. No statistically significant transcriptional differences were found between the anterior and posterior segments. A moderate pIgR expression was detected in liver and head kidney, the former being 3.9-fold and the latter 10-fold lower than gills (adjusted $p < 0.001$). These findings are consistent with the predominant role of pIgR in mucosal compartments.

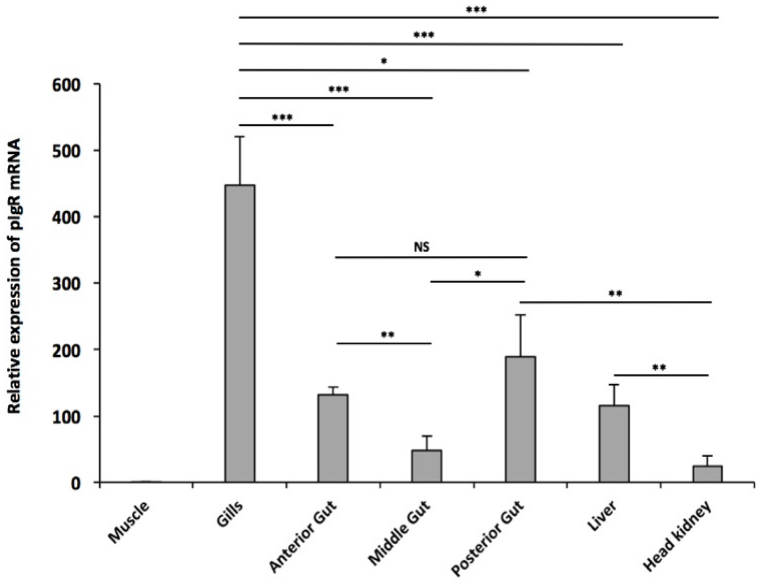


Figure 12. Relative expression levels of pIgR in different tissues from *T. bernacchii*. Data from three independent experiments are presented as mean gene expression relative to the housekeeping β–actin (± SD). The muscle tissue was used as a negative control. Levels of transcription were evaluated by q-PCR in duplicates using three T. bernacchii specimens. *$p < 0.05$; **$p < 0.01$; ***$p < 0.001$; NS, not significant (two-tailed Student's t test with Bonferroni correction).

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 27 June 2022 doi:10.20944/preprints202202.0211.v2

2.6. pIgR expressing cells in T. bernacchii intestinal and hepatic tissues

To identify pIgR-producing cells in *T. bernacchii* tissues we performed ISH analysis with anti-sense and sense RNA DIG labelled probes. Given the results obtained by q-PCR, we paid special attention to the gut-liver communication axis. The posterior in-testine, which displayed higher pIgR expression levels (Figure 12), was lined by a simple columnar epithelium of polarized cells (enterocytes) (Figure 13a-d). All the enterocytes were pIgR-expressing cells (Figure 13a-c). In particular, the highest staining intensity was detected around the nucleus and on the enterocyte basolateral surface, that face the basement membrane (Figure 13b). The apical surface of most epithelial cells did not display any staining. Scattered cells, localized both in the epithelium (Figure 13b) and the underlying lamina propria, were stained (Figure 13c). Moreover, the staining with the anti-sense probe (Figure 13e) revealed a strong signal throughout the liver. Notably, the expression of the pIgR gene was mainly detected around the nucleus of most hepatocytes. ISH with the pIgR sense probe did not result in any staining both in the posterior intestine and liver, as expected (Figure 13 d-f).

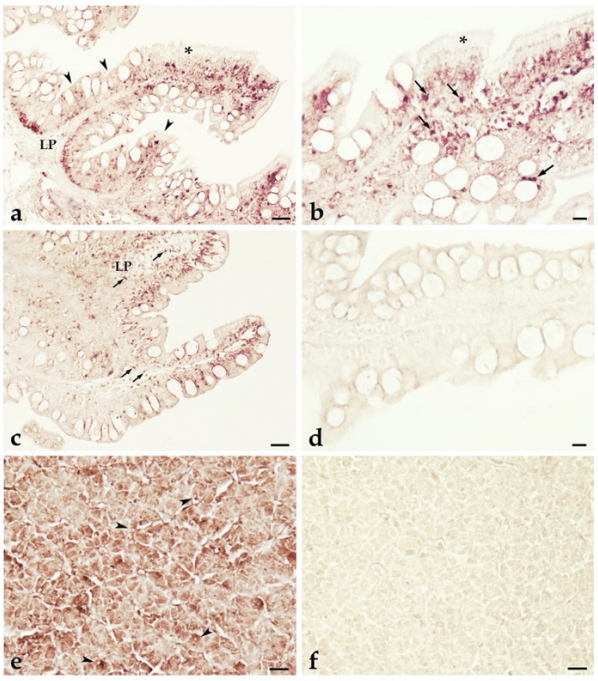


Figure 13. ISH of *T. bernacchii* pIgR. pIgR-expressing cells detected in *T. bernacchii* (posterior) gut (a-c) and liver (e) using an antisense probe. No signals were detected in either intestine (d) or liver with a sense probe (f). Stained enterocytes and liver epithelial cells are shown by arrowheads. Scattered cells containing pIgR transcripts are mainly in the epithelium and in the lamina propria (arrows). LP, lamina propria; * apical surface of enterocytes. Scale bars: 20 μm.

3. Discussion

The pIgR plays a crucial role in mammalian immune responses, since it ensures multifaceted immune functions [14, 28]. The pIgR has been studied in multiple teleost species, highlighting similarities and differences in its structure, as reviewed by Kortum et al. [15]. However, data about teleost pIgR functions are very limited and mainly referred to the transport of pIgs across the mucosal epithelial cells [2, 11, 13-15]. The pIgR gene locus organization still remains poorly investigated. The pIgR gene has been identified as a single copy in zebrafish, along with a wide multigene family that com-prises pIgR-like (pIgRL) genes, differentially expressed in lymphoid and

 doi:10.20944/preprints202202.0211.v2

myeloid cells [15]. A BLAST search of the *Takifugu rubripres* genome database allowed identification of a *pIgR* homologous gene [29].

To expand the current knowledge in this field, we isolated and characterized, for the first time, the pIgR gene in the cold adapted species *T. bernacchii*. Thanks to the increasing number of sequenced genomes from teleosts [30] over the past few years, we searched the annotated genome of *T. bernacchii* [31] to perform a full analysis of the *pIgR* gene structure. Additionally, the genome assemblies of two other cold adapted species *G. acuticeps* and *P. georgianus* [31], and of *C. gobio* [32], a temperate notothenioid species phyletically basal for the Antarctic Clade [18], were searched for a sequence-based comparison. We identified a single *pIgR* gene composed of eight exons and seven introns, as reported for other fish pIgRs. Notably, we found that Antarctic species have the largest pIgR gene (8235-9013 nt), compared to the other Perciformes included in the present study (3700-7090 nt). *C. gobio* showed a 5930-nt long gene, which was shorter than Antarctic species, but closer to the size of the pIgR gene from the temperate species. Further investigation found that the main difference with the temperate orthologs was in the length of introns, e.g., the third and fourth introns were twice or three times longer. The second intron was particularly long in the Antarctic species, while the third intron in *C. gobio* showed a standard length, as found in the temperate perciforms. These findings faithfully mirror the phylogenetic relationships among Notothenoidei [33], as confirmed also by the distance tree of pIgR D1 domain from notothenioid fish and species belonging to the other perciform suborders.

The expression of the *pIgR* gene still needs further analysis, as a single transcript of this gene is usually found in most fishes. However, in *Ctenopharyngodon idella*, seven pIgR splicing transcripts were identified, a full-length and six truncated variants, two gener-ated by exon skipping and the other four having different motif arragements at the 3' end [34]. In the present study, we determined that the introns of the pIgR gene lengthened during evolution of Notothenioidei without impairing the splicing process, as confirmed by the identification of transcript variants relative to the 5' EMPD exon 2. Using a splice prediction tool, we examined several sequence features of constitutive and cryptic sites, and identified alternative isoforms. Alternative splice sites are generally defined as weaker than the constitutive ones [35], however in the case of the T. bernacchii pIgR gene, the score relative to the variant X2, which derived from the usage of a cryptic site, was almost as high as that of the constitutive acceptor. The presence of the variant X2 remains to be assessed, as it only predicted by genome evaluation. Coversely, the genomic variant X1 was readily aligned to the cDNA sequence.

Intron size can vary due to the accumulation over time of non-homologous recom-bination, insertions and deletions, and TE activity. TEs account for a significant portion of vertebrate genomes and are known to play a role in genome modifications [36, 37]. Hotspots of retrotransposons (Rex-like) and DNA transposons (Tc1-like) elements have been reported in multiple teleost fishes, including Notothenioidei [31, 38, 39]. These findings support the idea that TEs could have contributed to the evolutionary process that led to the elongation of introns observed in Antarctic *pIgR* genes. The search for TEs in the introns of nototheniods and of the other five perciform suborders led to the identification of several elements. Notothenioidei contain all types of TEs, whereas some temperate species possess just one type, e.g, DNA transposons in *S. umbrosus* (Scorpanoidei) and *A. ocellatus* (Cottoidei) or LINEs in *P. fluvialitis* and *P. flavescens* (Percoidei). Prevalance of DNA transposons in the temperate species, including the non-Antarctic species *C. gobio*, along with the absence of TEs in multiple species belonging to different suborders, may be related to the evolutionary modifications of the whole teleost genome [40]. It is notable that the retrotransposons became prevalent in Antarctic fish. A detailed analysis of the distribution of TE elements in each intron, interestingly underlined that in *T. bernacchii* the first and second introns were characterized by the presence of one type of TEs,

while the third intron appeared more heterogeneous due to the presence of SINEs and LINEs. More interestingly, SINE elements found in the third intron of *T. bernacchii* represented the highest percentage (7.3 %) of TEs identified in the other cold adapted species. It is well known that SINEs are frequently found in Trematominae genera, accounting for their rapid radiation/genome rearrangement [41, 42]. To our knowledge this finding is not surprising. Previous studies we conducted on the Ig heavy chain gene locus of Notothenioidei revealed peculiar rearrangements of the intronic regions of both IgT and IgM heavy chain genes [21, 43].

In light of these data, we can hypothesize that in cold adapted species the long intronic sequences other than the first ones, may have evolved under selective constraint since they are functionally relevant. The primary role of some introns in regulating gene expression has been well described over the past decades [44]. Prediction of the binding sites for NF-κB, STAT6 and IRF1 in the sequences of fish pIgR genes available in GenBank, allowed their consideration as potential regulators in fish [14], in line with the cytokine-inducible regulatory elements, which are well known to direct the transcription of pIgR in mammals [13, 45]. Our analysis clearly evidenced that the second and third introns in Antarctic species were enriched for transcriptional factors, different from the other introns and from the other species analyzed. Given the structure of the *pIgR* gene, which comprises more exons with multiple variants, the presence of intragenic transcription factors identified in Antarctic teleosts may suggest a fine-tuned regulation of its expression. At present, the regulatory mechanism of *pIgR* expression has not yet been clarified in teleost fish nor in cold adapted species. Some insights into the adaptive strategies employed to ensure efficient gene expression under cold conditions were described by Lau et al. [46]. These authors proposed that Antarctic fishes may have adjusted transcription of their globin genes by duplicating cis-acting regulatory elements.

Another clue for a possible fine-tuning gene regulation of *T. benacchii pIgR* was provided by the presence of CpG islands (CGIs). It is well known that in vertebrate genomes CGIs are involved in transcriptional regulation [27]. The identification of promoter-associated CGIs has interesting implications for a possible key role of epigenetic regulation in cold adapted pIgR genes. In addition, Varriale and Bernardi [47] reported an increase of CGIs and DNA methylation levels in Antarctic fishes, most likely being ac-companied by the progressive cooling of Antarctic seawaters. However, the nature of regulatory elements and the mechanisms underlying transcription at low temperature are still poorly known and deserve further study. It is tempting to verify experimentally the pIgR gene regulation in cold adapted teleosts. Assessing whether CGIs sites are differentially methylated in a tissue-specific manner might help evaluate DNA methylation patterns in response to environmental threats.

In contrast to human pIgRs, which are extensively glycosylated, fish pIgRs show limited or even no glycosylation sites. These data raise the question whether N-glycosylation of pIgR is necessary for its function. This is an interesting issue that became even more relevant when we highlighted the presence of up to four N-glycosylation sites exclusively in pIgRs from Antarctic species. In addition, all sites were predicted to be glycosylated, thus appearing as a distinctive feature of cold water species. In particular, the asparagine residue of the glycosylation sequon present in the extracellular portion of *T. bernacchii* pIgR was found exposed to the solvent in the molecular model we built, indicating the carbohydrate availability for binding. As a general role, carbohydrates are involved in protein folding, stability and protection from proteolytic attacks. Glycosylation can also modulate a correct balance between protein solubility and structural flexibility. This role can be viewed as a special adaptive response to a cold environment, such as in the case of Antarctic fish IgT and IgM, found to be highly glycosylated [21, 48]. A further element that indicates an increase in solubility is related to the greater surface electrostatic charge. In fact, some amino acid residues, specific for the cold-adapted species, introduced electrostatic charges

 doi:10.20944/preprints202202.0211.v2

located on the surface of our mo-lecular model. These findings underline several peculiar features that may be considered as the hallmarks of cold pIgRs.

Furthermore, pIgR expression was analyzed in mucosal tissues and lymphoid organs, including the anterior, middle and posterior intestine, gills and head kidney. *T. bernacchii* pIgR transcripts were predominately expressed in the mucosal tissues, showing similar expression patterns to those of other teleosts species [10, 15, 29, 49, 50, 51]. The highest pIgR level detected in the gills is one of the most interesting aspects. Considering that gills represent one of the first lines of defence in teleosts, it will be challenging to assess whether the pIgR gene expression is upregulated in a tissue-specific manner in *T. bernachii*. Studies on CGIs and /or DNA methylation can provide additional insight into the regulation of the pIgR in response to pathogens, which are naturally occurring in the cold adapted fish investigated, e.g., the well documented nematode parasite infection [52, 53]. In other teleost species, high pIgR expression was found in gills after induced bacterial infection [11, 14]. Additionally, pIgR mRNA had higher expression in the posterior intestine than in the middle segment and did not differ from the anterior segment. These results, in line with the findings collected from other teleost fish [49, 54], further confirmed that the posterior intestine plays an important role in the mucosal immune response and host defence., High staining intensity was found through ISH in the cytoplasm of the enterocytes, suggesting that *T. bernacchii* pIgR mediates transepithelial transcytosis, as described in mammals [28]. Scattered pIgR-expressing cells were also detected in the mucosa, as previously observed in *Cyprinus carpio* [49]. However, currently, there are not yet any markers to provide evidence on nature of these pIgR-expressing cells, although staining may suggest an additional role played by pIgR molecules in teleost fish. In mammals, pIgR ensures multifaceted immune functions such as i) improving pIg stability, ii) pro-tection of secretory Ig from proteolytic degradation, iii) exclusion of pathogens from mucosal surface, iv) intracellular neutralization of invading pathogens, v) removal of pathogen-secretory Ig complexes from infected tissues, and vi) nonspecific microbial scavenger function of free secretory component (SC) [28]. Data about these functions of pIgR in teleost mucosal immunity are very limited, although new findings provided direct evidence for pIgR-mediated immune excretion of IgM–antigen complexes in *Paralichthys olivaceus* [55].

The role of the liver was considered to expand the current knowledge about the pIgR expression and pIgR-mediated transport of SIgs in teleost hepatocytes [22, 56, 57]. The results obtained by q-PCR demonstrated that pIgR was expressed in *T. bernacchii* liver, and ISH demonstrated that the transcripts were localized around the nucleus of the hepatocytes. In this species, IgM-immunoreactivity was detected in the perisinusoidal cells, bile canaliculi, pre-ductules and in the intraluminal mucus of the anterior intestine [22]. The data strongly substantiate that *T. bernacchii* IgM may be transported via pIgR across the hepatocytes to be secreted into the bile and subsequently into the gut. This implies the putative formation of a receptor-Ig complex, as is the case in mammals [56, 57]. Numerous observations indicated that teleost pIgR binds IgM [29, 51, 58] and IgT [59, 60], while it remains to be elucidated whether this can occur for IgD. We consider that to improve the current understanding of the T. bernacchii pIgR/SIg system and reveal the specific mechanisms of intracellular Ig transport, it will be necessary to rely on the availability of specific markers and adopt novel technologies from the field of molecular cell biology. Addressing these and other future studies will aid in further dissecting the complex roles of pIgR in mucosal immune defence of teleost fish.

In conclusion, our findings highlight several peculiar features acquired by cold adapted pIgR and underpin its pivotal role played in mucosal immune defence. The genome information gained in the present study will be useful for comparative and functional genomic analyses, contributing to advance the current knowledge of the *pIgR* gene in teleost fish.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 27 June 2022 doi:10.20944/preprints202202.0211.v2

4. Materials and Methods

4.1. Biological samples

Adult female specimens of *Trematomus bernacchii* (family Nototheniidae) were caught by use of nets in the Ross Sea, in the proximity of the Italian "Mario Zucchelli" Station at 74°42' S, 164°07' E, 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. Fish specimens (average weight 350 g) were kept in aquaria with running, aerated seawater until sacrificed. Tissues were collected and immediately frozen in liquid nitrogen.

4.2. Cloning of pIgR transcript

Total RNA was extracted, using a SV Total RNA Isolation System kit (Promega), from 150 mg of head kidney collected from a *T. bernacchii* specimen, homogenized by Potter-Elvehjem glass-Teflon. RNA quality was assessed on a 2% agarose gel and by measuring A260/A280 ratio, the concentration by reading absorbance at 260 nm with a NanoDrop 1000 Spectrophotometer (Thermo Scientific). RNA was then subject to DNase I treatment (Thermo Scientific, #EN0521), in order to avoid DNA genomic contamination for downstream analysis. cDNA was obtained from 1 μg of total RNA using Maxima H Minus Reverse Transcriptase (Thermo Scientific, #EP0751). The oligonucleotides used as primers to perform the first round-PCR amplification, were designed on the nucleotide sequence coding for the D1-D2 domains of *E. coioides* pIgR (accession number FJ803367), available in the GenBank database. The target sequence was amplified in a final volume of 25 μl using 2 μl cDNA (20 ng), 1,25 μM of specific primers (1.0 μM), 0.5 μl of dNTP Mix (0.2 μM), 2.5 μl 10X DreamTaq Buffer, 0,5 μl (1 U) of DreamTaq DNA polymerase (Thermo Scientific, #EP0705), up to volume with H2O, as follows: 95 °C for 3 min, 35 cycles of 95 °C (30 s), 60 °C (30 s), and 72 °C (1 min) with a final extension at 72 °C for 10 min. In order to improve the yield of the specific target amplification, the PCR product was then subject to a second amplification, following the same conditions as the primary PCR. Primers used in all the PCR experiments are shown in Table S4, which also reports the target domains (D1-D2) of pIgR for each primer. PCR products were analyzed on a 1.5% agarose gel, subsequently purified by NucleoSpin® Gel and PCR Clean-up (Macherey-Nagel) and finally cloned into pGEM®-T Easy Vector (Promega, #A1360). Positive clones were screened by the blue/white method and sequenced on both strands on an ABI PRISM 3100 automated sequencer at Eurofins Genomics Europe Sequencing GmbH (Jakob-Stadler-Platz 7, 78467 Konstanz, Germany).

4.3 3' and 5' Rapid Amplification of cDNA Ends (RACE)

In order to complete the 3' cDNA region of the T. bernacchii pIgR, 3' Rapid Ampli-fication of cDNA Ends (3'RACE) was performed using a commercial kit (Invitrogen™) according to the manufacturer's instructions. First-strand cDNA was synthesized as de-scribed above, using AP as a specific primer. PCR amplification was then carried out with pIGR1Fw as a sense primer and AUAP as an antisense primer. Subsequently a nested PCR was performed with pIGRII as a sense primer and AUAP as antisense primer (Table S4). The amplification was performed as follows: 95 °C for 5 min, 40 cycles of 95 °C (30 s), 55 °C (30 s) and 72 °C (1 min) with a final extension at 72 °C for 15 min. 5' Rapid Am-plification of cDNA Ends (5' RACE) was carried out on *T. bernacchii* pIgR cDNA using 5' RACE System for Rapid Amplification of cDNA Ends version 2.0 (Invitrogen™), fol-lowing the manufacturer's instructions. First-stranded cDNA was synthesized using an antisense specific primer pIGR1Rev. Subsequent PCR amplification was performed with pIGR1Rev and Oligo d(T)-anchor primer (AAP), supplied by the kit as sense primer. A nested PCR was performed with

pIGRIIr and AAP. The amplification was performed as follows: 95 °C for 3 min, 40 cycles of 95 °C (30 s), 60 °C (30 s), 72 °C (1.30 min) with a final extension at 72 °C for 10 min. 3' and 5' RACE products were cloned and sequenced as described above.

4.4. Data availability

The cDNA sequence coding for pIgR from *T. bernacchii* has been deposited in the GenBank database (https://www.ncbi.nlm.nih.gov/genbank/) under the accession number MZ540772. The nucleotide sequences from the other species used for molecular analysis are reported in Table S5. Genome assemblies and predicted transcripts of pIgR from the Antarctic species *T. bernacchii* (v. fTreBer1.1), *G. acuticeps* (v. fGymAcu1.1) and *P. georgianus* (v. fPseGeo1.1) [31 - for all species above], from the temperate notothenioid species *C. gobio* (v. fCotGob3.1) [32]. The scorpaenoid *S. umbrosus* genome assembly (v. fSebUmb1.pri) is within the framework of the Vertebrate Genome Project (https://www.ncbi.nlm.nih.gov/bioproject/PRJNA562006/) [30]

pIgR genome assemblies and predicted transcripts from *S. lucioperca* (v. SLUC_FBN_1.2) [61]; *P. fluviatilis* (v. GENO_Pfluv_1.0) [62], *P. flavescens* (v. PFLA_1.0) [63], *E. spectabile* (v. UIUC_Espe_1.0) [64], *E. cragini* (v. CSU_Ecrag_1.0) [65], *E. lanceolatus* (v. ASM528154v1) [66], *P. leopardus* (v. YSFRI_Pleo_2.0) [67], *P. pungitius* (v. NSP_V7) [68], *A. ocellatus* (v. GSC_Weel_1.0) [69] and *G. aculeatus* (v. GAculeatus_UGA_version5) [70] were retrieved from research articles that first reported them.

pIgR transcripts from the Antarctic species *D. eleginoides* [71], *N. coriiceps* [72] and *C. hamatus* [73], and from the serranoid species *E. coioides* [10] were retrieved from transcriptome shotgun assemblies, whereas for the Antarctic species *D. mawsoni* [74], *H. antarcticus* [31], *C. myersi* [75] and *C. aceratus* [76], from genome assemblies.

4.5. Gene sequence analyses

The pIgR intronic sequence dataset derived from all the species considered was scanned by using different tools.

The MEME (Multiple Em for Motif Elicitation) tool within the MEME suite v.5.4.1 environment [24] was used to discover conserved motifs by using default parameters.

In order to verify whether introns harbour repeated elements, all sequences were analyzed with RepeatMasker v.4.09 [25] and screened the repeats against the Dfam v.3.5 [77] library of known repeats found in the genomes of Actinopterygii, excluding simple repeats or low complexity DNA.

All introns were queried for the presence of transcription factor binding sites using the online tool Tfsitescan (www.ifti.org/cgi-bin/ifti/Tfsitescan.pl) [26].

Sequencing chromatogram was visualized using the program FinchTV (version 1.3.0). The nucleotide sequence obtained was verified by sequence similarity searches against the GenBank database, using the BLAST program. The prediction of cryptic and constitutive splice sites in pIgR from *T. bernacchii* was performed by using the Alternative Splice Site Predictor (ASSP) Tool [78].

Polyadenylation signals in the pIgR nucleotide sequence were predicted using Poly(A) Signal Miner [23]. The AT content of cDNA sequences was calculated with the GC Content Calculator (Biologics International Corp, Indianapolis USA).

4.6. Deduced amino acid sequence analyses

The amino acid sequences were deduced from nucleotide sequences using the ExPASy Translate Tool (https://web.expasy.org/translate/).

The amino acid composition was analyzed using the ProtParam [79] and Pep-Calc (www.pepcalc.com) tools.

The transmembrane-spanning regions and their orientation were predicted by using the TMpred tool [80].
The presence of the signal peptide was determined by using the SignalP-6.0 Sever tool [81].
Multiple sequence alignments were performed with Clustal Omega (https://www.ebi.ac.uk/Tools/msa/clustalo/) [82] and the out alignments were obtained in ClustalW format [83]. A distance-based tree of the D1 sequences Clustal Omega out alignment was reconstructed and analysed by iTOL program (https://itol.embl.de/) [84].
Sequons and putative N-glycosylation sites were identified using the NetNGlyc 4.0 Server [85].
A 3D molecular model was built for the *T. bernacchii* pIgR ectodomain using the Phyre2 tool (http://www.sbg.bio.ic.ac.uk/phyre/html/) [86]. 97% of the amino acid residues were modeled at 100% confidence using the 5f1s PDBe template (*Oncorhyncus mykiss* pIgR).
A Phyre molecular model was also built for the transmembrane domain. The highest confidence was 69% with the 6rx4 PDBe template. The obtained PDB models were analyzed by the molecular graphics program YASARA [87] (www.yasara.org).

4.7. Expression analysis of pIgR by qReal-time PCR
Total RNA was extracted, using a SV Total RNA Isolation System kit (Promega), from 150 mg each of anterior, middle and posterior gut, liver, gills, head kidney and muscle collected from three T. *bernacchii* specimens. Quantitative PCR-based expression analysis was performed on *T. bernacchii* cDNA using the Light Cycler 480 (Roche). The reaction consisted of 2 μl of cDNA diluted 1:10 and mixed with 5 μl of PowerUp™ SYBR™ Green Master Mix 2X (Applied Byosystem™) in a final volume of 10 μl with a final concentration of 0.3 μM of each primer, according to the manufacturer's instructions. TbrtpIgRFwd and TbrtpIgRRev are primers designed on the D1 and D2 domains respectively for the amplification of products from *T. bernacchii* pIgR (Table S4). qPCR was performed three times and samples, including DEPC water as negative control, were run in duplicate each time. The PCR amplification conditions were: 95 °C for 2 min, followed by 40 cycles of 95 °C (15 s), 60 °C (15 s), 72 °C (1 m). In order to assess the amplification specificity and the absence of primer dimers, a final dissociation step was run to generate a melting curve. In all melting curve analyses, single specific peaks were observed. The relative expression of pIgR was determined with the 2-ΔΔCq method, using β-actin as housekeeper gene (Table S4) and muscle as negative control.
Comparison between mucosal and lymphoid tissues was performed using two-tailed paired Student's t-tests adjusted by Bonferroni post hoc test. Data are presented as means ± standard deviation. p values <0.05 are considered as statistically significant, and shown as * p <0.05, ** p <0.01, *** p <0.001.

4.8. In situ hybridization (ISH)
4.8.1. Synthesis of RNA probes
Cells from gills of *T. bernacchii* were obtained by tissue teasing and suspended in Tripure (Roche). Total RNA was isolated and resuspended in DEPC-treated water. For reverse transcription, the BioScript RNase H minus (Bioline) enzyme was employed, using 1 μg of total RNA and 0.5 μg of random primers [pd(N)6]. Specific PCR primers (Table S6) were designed to amplify a 419 nt product corresponding to the T. bernacchii pIgR sequence.
Reactions were carried out in an Eppendorf Mastercycler personal (Milano, Italy). The cycling conditions were: 1 cycle of 94 °C for 5 min, 35 cycles of 94 °C for 45 s, 52 °C for 45 s, 72 °C for 45 s, followed by 1 cycle of 72 °C for 10 min. The resulting DNA was purified using the QIAquick Gel Extraction Kit (QIAgen), inserted into the pGEM-T Easy vector (Promega) and transfected into competent JM109 E. coli cells. Plasmid DNA from three independent clones was purified using the

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 27 June 2022 doi:10.20944/preprints202202.0211.v2

Wizard Plus SV Minipreps DNA Purification System (Promega) and sequenced using Eurofins Genomic Sequencing Services. Sequence similarity searching was carried out using the BLAST program. Selected plasmid clones were used as target in PCR reactions to synthesize the anti-sense and sense probes (see primers in Table S6). PCR conditions for anti-sense probe were: 1 cycle of 94 °C for 5 min, 35 cycles of 94 °C for 45 s, 54 °C for 45 s, 72 °C for 45 s, followed by 1 cycle of 72 °C for 10 min. The cycling protocol used for sense probe (the negative control of ISH ex-periments) was: 1 cycle of 94 °C for 5 min, 35 cycles of 94 °C for 45 s, 48 °C for 45 s, 72 °C for 45 s, followed by 1 cycle of 72 °C for 10 min. The PCR-products obtained were purified from agarose gel using QIAquick gel extraction kit (QIAgen) and used to synthesize DIG-labelled RNA probes with the DIG-RNA Labeling Kit (Roche).

4.8.2. Staining procedures

All steps were carried out according to Picchietti et al. [88]. In detail, posterior in-testine and liver from adult specimens (N= 3) were fixed overnight at room temperature (RT) in 4% paraformaldehyde in 0.01 M, pH 7.4 phosphate buffered saline (PBS), then dehydrated, embedded in paraffin wax and cut into 7μm-thick sections using a rotary microtome. Serial sections were collected on poly-L-lysine coated slides, air-dried overnight at 37 °C and stored at RT for subsequent investigation. After dewaxing in xylene and rehydration in graded ethanol series, sections were washed with DEPC water before proteinase K (Sigma-Aldrich) digestion. The concentration of proteinase K was titrated for posterior intestine and liver and the best results were obtained with 1 μg/ml. The di-gestion was stopped by immersion in cold DEPC water. Acetylation was performed by incubating sections in 0.25% acetic anhydride in 85 mM Tris-HCl buffer, containing 0.2% acetic acid and 0.02 M ethylenediaminetetraacetic acid (EDTA) for 10 min. Following rinses in DEPC water, sections were gradually dehydrated and incubated overnight at 45 °C with the probes (concentrations varying from 0.3 to 0.6 ng/ml) and the optimal one resulted at 0.45 ng/ml. Subsequently, sections were washed with 2 x saline-sodium citrate (SSC) buffer at RT, then with 0.2 x SSC at 55 °C for 90 min and incubated for 30 min with 20μg/ml RNAase A in 0.01 M Tris-HCl containing 0.5MNaCl and 1mMEDTA. Sections were transferred for 1 h to Buffer 1 (0.1 M Tris containing 0.15 M NaCl and 1% blocking reagent), then for 30 min to Buffer 2 (0.1 M Tris containing 0.15 M NaCl, 0.5% BSA and 0.3% Triton X- 100). Then, sections were incubated for 2 h at RT with alkaline phos-phatase-conjugated anti-digoxigenin antibody (Fab fragment; Roche Diagnostic, Germany) diluted 1:1000 in Buffer 2, then washed with 0.1 M Tris containing 0.15 M NaCl and Buffer 3 (0.1 M Tris containing 0.1 M NaCl and 50 mM MgCl2). Following staining with nitro blue tetrazoliumchloride and 5-bromo-4-chloro-3-indolyl-phosphate (Roche Diagnostic, Germany), sections were mounted with Aquatex, aqueous mounting agent for microscopy (Merck, USA) and examined under bright-field illumination. Images were acquired by a Zeiss Axioskop 2 plus a microscope equipped with the Axiocam MRC camera and the Axiovision software (Carl Zeiss, Oberkochen, Germany).

Supplementary Materials: Figure S1. Complete sequence of *T. bernacchii* polymeric Ig receptor (pIgR) gene. **Figure S2.** 3D molecular model of the transmembrane helix of *T. bernacchii* pIgR built with the Phyre2 tool (http://www.sbg.bio.ic.ac.uk/phyre/html/). **Figure S3.** Multiple alignment of the deduced amino acid sequences of pIgRs. **Figure S4.** Detail of R60, K103, E114, and N161 residues in the 3D molecular model built for the secretory component of *T. bernacchii* pIgR. **Table S1.** Predicted transcription factor-binding sites in the 5'-flanking region, and in the second and third introns of *T. bernacchii* polymeric Ig receptor (*pIgR*) gene. **Table S2.** Amino acid composition of *T. bernacchii* pIgR and the respective regions SC (Secretory Component), EMPD (Extracellular proximal domain), TM (Transmembrane domain), Cyt (Cytoplasmic tail). **Table S3.** Splice-site

prediction for the region encompassing the 3' end of the fourth intron and the 5' end of the EMPD exon of *T. bernacchii* pIgR. **Table S4.** List of primers used in PCR experiments. **Table S5.** List of perciform suborders and respective species investigated for *pIgR* genomic and transcript sequences available. **Table S6.** Specific primers used for RT-PCR and sense and anti-sense probes.

Author Contributions: Conceptualization, A.A., S.P. and M.R.C.; validation, A.A., S.P. and U.O.; formal analysis, A.A., S.P., L.G., U.O and M.R.C.; investigation, A.A., S.P., L.G. and S.G.; resources, S.P. and M.R.C.; writing—original draft preparation, A.A., U.O. and M.R.C.; writing—review and editing, A.A., S.P., S.G., U.O. and M.R.C.; visualization, A.A., S.P., L.G., U.O. and M.R.C.; project administration, S.P. and M.R.C.; funding acquisition, S.P. and M.R.C. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by the Italian National Program for Research in Antarctica, grant number PNRA18_00077.

Institutional Review Board Statement: The study was conducted according to the "Protocol on environmental protection to the Antarctic Treaty" Annex V.

Informed Consent Statement: Not applicable.

Data Availability Statement: The data presented in this study are available in the Material and Methods Section, the Paragraph 4.4.

Acknowledgments: We thank Iliana Bista and the Antarctic notothenioid fish genomes project (https://www.sanger.ac.uk/project/antarctic-notothenioids-fish-genomes-project/) for allowing us use of their data, and the VGP for early access to the genome of *S. umbrosus*, specifically Peter Sudmant, Olivier Fedrigo, Martin Pippel, Erich Jarvis and the Rockefeller University Vertebrate Genomes Lab for the assembly. We are grateful to Ennio Cocca (IBBR, CNR, Naples, Italy) and Salvatore Fioriniello (IGB Buzzati-Traverso, CNR, Naples, Italy) for useful discussion and suggestions.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Akula, S.; Mohammadamin, S.; Hellman, L. Fc receptors for immunoglobulins and their appearance during vertebrate evolu-tion. PLoS One. 2014, 9, e96903, doi: 10.1371/journal.pone.0096903.
2. Kaetzel, C. S. Coevolution of mucosal immunoglobulins and the polymeric immunoglobulin receptor: evidence that the com-mensal microbiota provided the driving force. ISRN Immunology. 2014, e541537, doi:10.1155/2014/541537.
3. Oreste, U.; Ametrano, A.; Coscia, M. R. On origin and evolution of the antibody molecule. Biology. 2021, 10, 140, doi:10.3390/biology10020140.
4. Mostov, K.E.; Friedlander, M.; Blobel, G. The receptor for transepithelial transport of IgA and IgM contains multiple immu-noglobulin-like domains. Nature. 1984, 308, 37–43, doi:10.1038/308037a0.

5. Kulseth, M.A.; Krajci, P.; Myklebost, O.; Rogne, S. Cloning and characterization of two forms of bovine polymeric immuno-globulin receptor cDNA. DNA Cell Biol. 1995, 14, 251–256, doi:10.1089/dna.1995.14.251.
6. Kuhn, L.C.; Kocher, H.P.; Hanly, W.C.; Cook, L.; Jaton, J.C.; Kraehenbuhl, J.P. Structural and genetic heterogeneity of the recep-tor mediating translocation of immunoglobulin A dimer antibodies across epithelia in the rabbit. J Biol Chem. 1983, 258, 6653–6659, doi:10.1016/S0021-9258(18)32462-1.
7. Wieland, W.H.; Orzaez, D.; Lammers, A.; Parmentier, H.K.; Verstegen, M.W.; Schots, A. A functional polymeric immuno-globulin receptor in chicken (Gallus gallus) indicates ancient role of secretory IgA in mucosal immunity. Biochem J. 2004, 380, 669-676, doi:10.1042/bj20040200.
8. Magadán-Mompo', S.; Sánchez-Espinel, C.; Gambón-Deza, F. IgH loci of American alligator and saltwater crocodile shed light on IgA evolution. Immunogenet. 2013, 65, 531–541, doi:10.1007/s00251-013-0692-y.
9. Braathen, R.; Hohman, V.S.; Brandtzaeg, P.; Johansen, F.-E. Secretory antibody formation: conserved binding interactions between J chain and polymeric Ig receptor from humans and amphibians. J Immunol. 2007, 178, 1589–1597, doi:10.4049/jimmunol.178.3.1589.
10. Feng, L.-N.; Lu, D.-Q.; Bei, J.-X.; Chen, J.-L.; Liu, Y.; Zhang, Y.; Liu, X.-C.; Meng, Z.-N.; Wang, L.; Lin, H.-R. Molecular cloning and functional analysis of polymeric immunoglobulin receptor gene in orange-spotted grouper (Epinephelus coioides). Comp Biochem Physiol B, Biochem Mol Biol. 2009, 154, 282–289, doi:10.1016/j.cbpb.2009.07.003.
11. Salinas, I.; Fernández-Montero, Á.; Ding, Y.; Sunyer, J.O. Mucosal immunoglobulins of teleost fish: A decade of advances. Dev Comp Immunol. 2021, 121, 104079, doi:10.1016/j.dci.2021.104079.
12. Hohman, V.S.; Stewart, S.E.; Rumfelt, L.L.; Greenberg, A.S.; Avila, D.W.; Flajnik, M. F.; Steiner, L.A. J chain in the nurse shark: implications for function in a lower vertebrate. J Immunol. 2003, 170, 6016 – 6023, doi:10.4049/jimmunol.170.12.6016.
13. Kaetzel, C. S. The polymeric immunoglobulin receptor: bridging innate and adaptive immune responses at mucosal surfaces. Immunol Rev. 2005, 206, 83–99, doi:10.1111/j.0105-2896.2005.00278.x.
14. Kong, X.; Wang, L.; Pei, C.; Zhang, J.; Zhao, X.; Li, L. Comparison of polymeric immunoglobulin receptor between fish and mammals. Vet Immunol and Immunopath. 2018, 202, 63–69, doi:10.1016/j.vetimm.2018.06.002.
15. Kortum, A.N.; Rodriguez-Nunez, I.; Yang, J.; Shim, J.; Runft, D.; O'Driscoll, M.L.; Haire, R.N.; Cannon, J.P.; Turner, P.M.; Lit-man, R. T.; et al. Differential expression and ligand binding indicate alternative functions for zebrafish polymeric immuno-globulin receptor (PIgR) and a family of pigr-like (PIGRL) proteins. Immunogenet. 2014, 66, 267–279, doi:10.1007/s00251-014-0759-4.
16. Near, T.J.; Ghezelayagh, A.; Ojeda, F.P.; Dornburg, A. Recent diversification in an ancient lineage of notothenioid fishes (Bo-vichtus: Notothenioidei). Polar Biol. 2019, 42, 943–952, doi:10.1007/s00300-019-02489-1.
17. Eastman, J.T. Antarctic fish biology: evolution in a unique environment vol.30, Academic Press, San Diego, 1993, 59-60.
18. Near, T.J.; Dornburg, A.; Harrington, R.C.; Oliveira, C.; Pietsch, T.W.; Thacker, C.E.; Satoh, T.P.; Katayama, E.; Wainwright, P.C.; Eastman, J.T.; et al. Identification of the notothenioid sister lineage illuminates the biogeographic history of an Antarctic adaptive radiation. BMC Evol Biol. 2015, 15, 109, doi:10.1186/s12862-015-0362-9.
19. Coscia, M. R.; Varriale, S.; Giacomelli, S.; Oreste, U. Antarctic teleost immunoglobulins: more extreme, more interesting. Fish Shellfish Immunol. 2011, 31, 688–696, doi:10.1016/j.fsi.2010.10.018.

20. Giacomelli, S.; Buonocore, F.; Albanese, F.; Scapigliati, G.; Gerdol, M.; Oreste, U.; Coscia, M. R. New insights into evolution of IgT genes coming from Antarctic teleosts. Mar Genomics. 2015, 24, 55–68, doi:10.1016/j.margen.2015.06.009.
21. Ametrano, A.; Gerdol, M.; Vitale, M.; Greco, S.; Oreste, U.; Coscia, M.R. The evolutionary puzzle solution for the origins of the partial loss of the Cτ2 exon in notothenioid fishes. Fish Shellfish Immunol. 2021, 116, 124–139, doi:10.1016/j.fsi.2021.05.015.
22. Abelli, L.; Coscia, M.R.; De Santis, A.; Zeni, C.; Oreste, U. Evidence for hepato-biliary transport of immunoglobulin in the Antarctic teleost fish Trematomus bernacchii. Devel Comp Immunol. 2005, 29, 431–442, doi:10.1016/j.dci.2004.09.004.
23. Liu, H.; Han, H.; Li, J.; Wong, L. DNAFSMiner: a web-based software toolbox to recognize two types of functional sites in DNA sequences. Bioinformatics 2005, 21 (5), 671–673, doi:10.1093/bioinformatics/bth437.
24. Bailey, T.L.; Boden, M.; Buske, F.A.; Frith, M.; Grant, C.E.; Clementi, L.; Ren, J.; Li, W.W.; Noble, W.S. MEME Suite: tools for motif discovery and searching. Nucleic Acids Res. 2009, 37, W202–W208, doi:10.1093/nar/gkp335.
25. Smit, A.F.A.; Hubley, R.; Green P. RepeatMasker Open-3.0, 1996.
26. Ghosh, D. Object-Oriented Transcription Factors Database (OoTFD). Nucleic Acids Res. 2000, 28, 308–310, doi:10.1093/nar/28.1.308.
27. Deaton, A. M.; Bird, A. CpG islands and the regulation of transcription. Genes Dev. 2011, 25 (10), 1010–1022, doi:10.1101/gad.2037511.
28. Wei, H.; Wang, J.-Y. Role of polymeric immunoglobulin receptor in IgA and IgM transcytosis. Int J Mol Sci. 2021, 22, 2284, doi:10.3390/ijms22052284.
29. Hamuro, K.; Suetake, H.; Saha, N.R.; Kikuchi, K.; Suzuki, Y. A teleost polymeric Ig receptor exhibiting two Ig-like domains transports tetrameric IgM into the skin. J Immunol. 2007, 178, 5682–5689, doi:10.4049/jimmunol.178.9.5682.
30. Koepfli, K.-P.; Paten, B.; Genome 10K community of scientists; O'brien, S. J. The genome 10K Project: a way forward. Annu Rev Anim Biosci 3. 2015, 57–111, doi:10.1146/annurev-animal-090414-014900.
31. Bista, I.; Wood, J. M. D.; Desvignes, T.; McCarthy, S.A.; Matschiner, M.; Ning, Z.; Tracey, A.; Torrance, J.; Sims, J.; Chow, W.; et al. Genomics of cold adaptations in the Antarctic notothenioid fish radiation. Correspondence: Iliana Bista (ilianabi-sta@gmail.com); Richard Durbin (rd109@cam.ac.uk), Wellcome Sanger Institute, Tree of Life, Wellcome Genome Campus, Hinxton, CB10 1SA, United Kingdom; Department of Genetics, University of Cambridge, Downing Street, Cambridge, CB2 3EH, United Kingdom 2022, preprint doi: https://doi.org/10.1101/2022.06.08.494096
32. Bista, I.; McCarthy, S.A.; Wood, J. M. D.; Ning, Z.; Detrich III, H.W.; Desvignes, T.; Postlethwait, J.; Chow, W.; Howe, K.; Tor-rance, J.; et al. The genome sequence of the channel bull blenny, Cottoperca gobio (Gunther, 1861). Wellcome Open Res. 2020, 5, 148, doi:10.12688/wellcomeopenres.16012.1.
33. Near, T. J.; MacGuigan, D. J.; Parker, E.; Struthers, C. D.; Jones, C. D.; Dornburg, A. Phylogenetic analysis of Antarctic notothe-nioids illuminates the utility of RADseq for resolving Cenozoic adaptive radiations. Mol Phylogenet Evol 2018, 129, 268–279, doi: 10.1016/j.ympev.2018.09.001.
34. Pei, C.; Sun, X.; Zhang, Y.; Li, L.; Gao, Y.; Wang, L.; Kong, X. Molecular cloning, expression analyses of polymeric immuno-globulin receptor gene and its variants in grass carp (Ctenopharyngodon Idellus) and binding assay of the recombinant im-munoglobulin-like domains. Fish Shellfish Immunol. 2019, 88, 472–479, doi:10.1016/j.fsi.2019.03.024.

35. Zheng, C.L.; Fu, X-D.; Gribskov, M. Characteristics and regulatory elements defining constitutive splicing and different modes of alternative splicing in human and mouse. RNA. 2005, 11, 1777-1787, doi:10.1261/rna.2660805.
36. Mills, R. E.; Bennett, E. A.; Iskow, R. C.; Devine, S. E. Which transposable elements are active in the human genome? Trends Genet. 2007, 23, 183–191, doi:10.1016/j.tig.2007.02.006.
37. Sela, N.; Kim, E.; Ast, G. The role of transposable elements in the evolution of non-mammalian vertebrates and invertebrates. Genome Biol. 2010, 11, R59 doi:10.1186/gb-2010-11-6-r59.
38. Daane, J. M.; Detrich, H. W. III. Adaptation and diversity of Antarctic fishes: a genomic perspective. Annu Rev Anim Biosci. 2022, 10, 39–62, doi: 10.1146/annurev-animal-081221-064325.
39. Cocca, E.; Iorio, S. D.; Capriglione, T. Identification of a novel helitron transposon in the genome of Antarctic fish. Mol Phyl Evol. 2011, 58, 439–446, doi:10.1016/j.ympev.2010.12.020.
40. Shao, F.; Han, M.; Peng, Z. Evolution and diversity of transposable elements in fish genomes. Sci Rep. 2019, 9, 15399, doi:10.1038/s41598-019-51888-1.
41. Auvinet, J.; Graça, P.; Belkadi, L.; Petit, L.; Bonnivard, E.; Dettaï, A.; Detrich, W. H.; Ozouf-Costaz, C.; Higuet, D. Mobilization of retrotransposons as a cause of chromosomal diversification and rapid speciation: the case for the Antarctic teleost genus Trematomus. BMC Genomics. 2018, 19, 339, doi:10.1186/s12864-018-4714-x.
42. Pisano, E.; Coscia, M. R.; Mazzei, F.; Ghigliotti, L.; Coutanceau, J.-P.; Ozouf-Costaz, C.; Oreste, U. Cytogenetic mapping of im-munoglobulin heavy chain genes in Antarctic fish. Genetica. 2007, 130, 9–17, doi:10.1007/s10709-006-0015-4.
43. Coscia, M. R.; Varriale, S.; De Santi, C.; Giacomelli, S.; Oreste, U. Evolution of the Antarctic teleost immunoglobulin heavy chain gene. Mol Phyl Evol. 2010, 55, 226–233, doi:10.1016/j.ympev.2009.09.033.
44. Dwyer, K.; Agarwal, N.; Pile, L.; Ansari, A. Gene architecture facilitates intron-mediated enhancement of transcription. Front Mol Bio. 2021, 8, 276, doi:10.3389/fmolb.2021.669004.
45. Kushiro, A.; Sato, T. Polymeric immunoglobulin receptor gene of mouse: sequence, structure and chromosomal location. Gene. 1997, 204, 277–282, doi:10.1016/S0378-1119(97)00482-4.
46. Lau, D. T.; Saeed-Kothe, A.; Parker, S. K.; Detrich, H. W. III. Adaptive evolution of gene expression in Antarctic fishes: diver-gent transcription of the 5′-to-5′ linked adult α1- and β-globin genes of the Antarctic teleost Notothenia coriiceps is controlled by dual promoters and intergenic enhancers1. Am Zool. 2001, 41, 113–132, doi:10.1093/icb/41.1.113.
47. Varriale, A.; Bernardi, G. DNA methylation and body temperature in fishes. Gene. 2006, 385, 111–121, doi:10.1016/j.gene.2006.05.031.
48. Coscia, M.R.; Morea, V.; Tramontano, A.; Oreste, U. Analysis of a cDNA sequence encoding the immunoglobulin heavy chain of the Antarctic teleost Trematomus bernacchii. Fish Shellfish Immunol. 2000, 10, 343–357, doi:10.1006/fsim.1999.0244.
49. Rombout, J.; Vandertuin, S.; Yang, G.; Schopman, N.; Mroczek, A.; Hermsen, T.; Tavernethiele, J. Expression of the polymeric immunoglobulin receptor (PIgR) in mucosal tissues of common carp (Cyprinus carpio l.). Fish Shellfish Immunol. 2008, 24, 620–628, doi:10.1016/j.fsi.2008.01.016.
50. Tadiso, T. M.; Sharma, A.; Hordvik, I. Analysis of polymeric immunoglobulin receptor- and CD300-like molecules from At-lantic salmon. Mol Immunol. 2011, 49, 462–473, doi:10.1016/j.molimm.2011.09.013.

51. Xu, Z.; Parra, D.; Gómez, D.; Salinas, I.; Zhang, Y.-A.; Jørgensen, L. von G.; Heinecke, R. D.; Buchmann, K.; LaPatra, S.; Sunyer, J. O. Teleost skin, an ancient mucosal surface that elicits gut-like immune responses. PNAS. 2013a, 110, 13097–13102, doi:10.1073/pnas.1304319110.
52. Palm, H. W. Ecology of Pseudoterranova decipiens (Krabbe, 1878) (Nematoda: Anisakidae) from Antarctic Waters. Parasitol Res. 1999, 85, 638–646, doi:10.1007/s004360050608.
53. Orecchia, P.; Mattiucci, S.; D'Amelio, S.; Paggi, L.; Plötz, J.; Cianchi, R.; Nascetti, G.; Arduino, P.; Bullini, L. Two new members in the Contracaecum osculatum Complex (Nematoda, Ascaridoidea) from the Antarctic. Int J Parasitol. 1994, 24, 367–377, doi:10.1016/0020-7519(94)90084-1.
54. Rombout, J. H. W. M.; Taverne-Thiele, A. J.; Villena, M. I. The gut-associated lymphoid tissue (galt) of carp (Cyprinus carpio l.): an immunocytochemical analysis. Dev Comp Immunol. 1993, 17, 55–66, doi:10.1016/0145-305X(93)90015-I.
55. Sheng, X.; Qian, X.; Tang, X.; Xing, J.; Zhan, W. Polymeric immunoglobulin receptor mediates immune excretion of mucosal igm–antigen complexes across intestinal epithelium in flounder (Paralichthys olivaceus). Front Immunol. 2018, 9, doi:10.3389/fimmu.2018.01562 .
56. Salinas, I.; Parra, D. 6 - Fish mucosal immunity: intestine. In Mucosal Health in Aquaculture; Beck, B. H., Peatman, E., Eds.; Academic Press: San Diego, 2015; pp 135–170, doi:10.1016/j.dci.2021.104079.
57. Brandl, K.; Kumar, V.; Eckmann, L. Gut-liver axis at the frontier of host-microbial interactions. Am J Physiol-Gastr L. 2017, 312, G413–G419, doi:10.1152/ajpgi.00361.2016.
58. Xu, G.; Zhan, W.; Ding, B.; Sheng, X. Molecular cloning and expression analysis of polymeric immunoglobulin receptor in flounder (Paralichthys olivaceus). Fish Shellfish Immunol. 2013b, 35, 653–660, doi:10.1016/j.fsi.2013.05.024.
59. Zhang, Y.-A.; Salinas, I.; Li, J.; Parra, D.; Bjork, S.; Xu, Z.; LaPatra, S. E.; Bartholomew, J.; Sunyer, J. O. IgT, a primitive immu-noglobulin class specialized in mucosal immunity. Nat Immunol. 2010, 11, 827–835, doi:10.1038/ni.1913.
60. Yu, Y.; Liu, Y.; Li, H.; Dong, S.; Wang, Q.; Huang, Z.; Kong, W.; Zhang, X.; Xu, Y.; Chen, X. et al. Polymeric immunoglobulin re-ceptor in dojo loach (Misgurnus anguillicaudatus): molecular characterization and expression analysis in response to bacte-rial and parasitic challenge. Fish Shellfish Immunol. 2018, 73, 175–184, doi:10.1016/j.fsi.2017.12.019.
61. Nguinkal, J.A.; Brunner, R.M.; Verleih, M.; Rebl, A.; Ríos-Pérez, L. de los; Schäfer, N.; Hadlich, F.; Stüeken, M.; Wittenburg, D.; Goldammer, T. The first highly contiguous genome assembly of pikeperch (Sander lucioperca), an emerging aquaculture spe-cies in Europe. Genes. 2019, 10, 708, doi:10.3390/genes10090708.
62. Roques, C.; Zahm, M.; Cabau, C.; Klopp, C.; Bouchez, O.; Donnadieu, C.; Kuhl, H.; Gislard, M.; Guendouz, S.; Journot, L. et al. A chromosome-scale genome assembly of the European perch, Perca fluviatilis. Submitted (JUN-2019) to the EMBL/GenBank/DDBJ databases.
63. Feron, R.; Morvezen, R.; Bestin, A.; Haffray, P.; Klopp, C.; Zahm, M.; Cabau, C.; Roques, C.; Donnadieu, C.; Bouchez, O. et al. A chromosome-scale genome assembly of the yellow perch, Perca flavescens. Submitted (JAN-2019) to the EMBL/GenBank/DDBJ databases.
64. Moran, R.L.; Catchen, J.M.; Fuller, R.C. A chromosome-level genome assembly, high-density linkage maps, and genome scans reveal the genomic architecture of hybrid incompatibilities underlying speciation via character displacement in darters (Percidae: Etheostominae). Submitted (AUG-2019) to the EMBL/GenBank/DDBJ databases.
65. Reid, B.N.; Moran, R.L.; Kopack, C.J.; Fitzpatrick, S. W. Rapture-ready darters: Choice of reference genome and genotyping method (whole-genome or sequence capture) influence population genomic inference in Etheostoma. Mol Ecol Resour. 2021, 21, 404–420, doi:10.1111/1755-0998.13275.

66. Zhou, Q.; Gao, H.; Zhang, Y.; Fan, G.; Xu, H.; Zhai, J.; Xu, W.; Chen, Z.; Zhang, H.; Liu, S.; et al. A chromosome-level genomea of the giant grouper (Epinephelus lanceolatus) provides insights into its innate immunity and rapid growth. Mol Ecol Resour. 2019, 19, 1322–1332, doi:10.1111/1755-0998.13048.
67. Zhou, Q.; Guo, X.; Huang, Y.; Gao, H.; Xu, H.; Liu, S.; Zheng, W.; Zhang, T.; Tian, C.; Zhu, C.; et al. De novo sequencing and chromosomal-scale genome assembly of leopard coral grouper, Plectropomus leopardus. Mol Ecol Resour. 2020, 20, 1403–1413, doi:10.1111/1755-0998.13207.
68. Varadharajan, S. Pungitius pungitius genome assembly, contig: LG2, whole genome shotgun sequence. Submitted (OCT-2019) to the EMBL/GenBank/DDBJ databases.
69. Culibrk, L.; Leelakumari, S.; Taylor, G.A.; Tse, K.; Cheng, D.; Chuah, E.; Kirk, H.; Pandoh, P.; Troussard, A.; Zhao, Y.; et al. The genome of the wolf eel (Anarrhichthys ocellatus). Submitted (JAN-2019) to the EMBL/GenBank/DDBJ databases.
70. Nath, S.; Shaw, D.E.; White, M. A. Improved contiguity of the threespine stickleback genome using long-read sequencing. G3. 2021, 11, jkab007, doi:10.1093/g3journal/jkab007.
71. Touma, J.; García, K.K.; Bravo, S.; Leiva, F.; Moya, J.; Vargas-Chacoff, L.; Reyes, A.; Vidal, R. De novo assembly and charac-terization of patagonian toothfish transcriptome and develop of EST-SSR markers for population genetics. Front in Mar Sci. 2019, 6, 720, doi:10.3389/fmars.2019.00720.
72. Shin, S.C.; Kim, S. J.; Lee, J.K.; Ahn, D.H.; Kim, M.G.; Lee, H.; Lee, J.; Kim, B.-K.; Park, H. Transcriptomics and comparative analysis of three Antarctic notothenioid fishes. PLoS One. 2012, 7, e43762, doi:10.1371/journal.pone.0043762.
73. Song, W.; Li, L.; Huang, H.; Jiang, K.; Zhang, F.; Wang, L.; Zhao, M.; Ma, L. Tissue-based transcriptomics of Chionodraco hamatus: sequencing, de novo assembly, annotation and marker discovery. J Fish Biol. 2019, 94, 251–260, doi:10.1111/jfb.13882.
74. Lee, S. J.; Kim, J.-H.; Jo, E.; Choi, E.; Kim, J.; Choi, S.-G.; Chung, S.; Kim, H.-W.; Park, H. Chromosomal assembly of the Antarc-tic toothfish (Dissostichus mawsoni) genome using third-generation DNA sequencing and Hi-C technology. Zool Res. 2021, 42, 124–129, doi:10.24272/j.issn.2095-8137.2020.264.
75. Bargelloni, L.; Babbucci, M.; Ferraresso, S.; Papetti, C.; Vitulo, N.; Carraro, R.; Pauletto, M.; Santovito, G.; Lucassen, M.; Mark, F. C.; et al. Draft genome assembly and transcriptome data of the icefish Chionodraco myersi reveal the key role of mitochon-dria for a life without hemoglobin at subzero temperatures. Commun Biol. 2019, 2, 1–11, doi:10.1038/s42003-019-0685-y.
76. Jakobsen, S.K.; Tørresen, K.O., Boessenkool, S.; Malmstrøm, M.; Star, B.; Jakobsen, S.K.; Riiser, S.E. The Chaenocephalus acera-tus whole genome shotgun project. Submitted (MAR-2018) to the EMBL/GenBank/DDBJ database.
77. Hubley, R.; Finn, R.D.; Clements, J.; Eddy, S.R.; Jones, T.A.; Bao, W.; Smit, A.F.A.; Wheeler, T.J. The Dfam database of repetitive DNA families. Nucleic Acids Res. 2016, 44, D81–D89, doi:10.1093/nar/gkv1272.
78. Wang, M.; Marín, A. Characterization and prediction of alternative splice sites. Gene, 2006, 366, 219–227, doi:10.1016/j.gene.2005.07.015.
79. Gasteiger, E.; Hoogland, C.; Gattiker, A.; Duvaud, S.; Wilkins, M. R.; Appel, R. D.; Bairoch, A. Protein identification and anal-ysis tools on the ExPASy Server. In The Proteomics Protocols Handbook; Walker, J. M., Ed.; Humana Press: Totowa, NJ, 2005, 571–607.
80. Cuthbertson, J. M.; Doyle, D. A.; Sansom, M. S. P. Transmembrane helix prediction: A comparative evaluation and analysis. Protein Eng Des Sel. 2005, 18, 295–308, doi:10.1093/protein/gzi032.

81. Teufel, F.; Armenteros, J. J. A.; Johansen, A. R.; Gíslason, M. H.; Pihl, S. I.; Tsirigos, K. D.; Winther, O.; Brunak, S.; Heijne, G. von; Nielsen, H. SignalP 6.0 achieves signal peptide prediction across all types using protein language models. bioRxiv. 2021, 447770.
82. Sievers, F.; Higgins, D. G. Clustal Omega, accurate alignment of very large numbers of sequences. In Multiple Sequence Alignment Methods; Russell, D. J., Ed.; Methods in Molecular Biology; Humana Press: Totowa, NJ. 2014, 105–116.
83. Thompson, J.D.; Higgins, D.G.; Gibson, T.J. CLUSTAL W: improving the sensitivity of progressive multiple sequence align-ment through sequence weighting, position-specific gap penalties and weight matrix choice. Nucleic Acids Res. 1994, 22, 4673–4680, doi:10.1093/nar/22.22.4673.
84. Letunic, I.; Bork, P. Interactive Tree f Life (ITOL): An online tool for phylogenetic tree display and annotation. Bioinformatics. 2007, 23, 127–128, doi:10.1093/bioinformatics/btl529.
85. Gupta, R.; Brunak, S. Prediction of glycosylation across the human proteome and the correlation to protein function. Pac Symp Biocomput. 2002, 310-322.
86. Kelley, L.A.; Mezulis, S.; Yates, C.M.; Wass, M.N.; Sternberg, M.J.E. The Phyre2 web portal for protein modeling, prediction and analysis. Nat Protoc. 2015, 10, 845–858, doi:10.1038/nprot.2015.053.
87. Krieger, E.; Vriend, G. Models@Home: distributed computing in bioinformatics using a screensaver based approach. Bioin-formatics 2002, 18 (2), 315–318, doi:10.1093/bioinformatics/18.2.315.
88. Picchietti, S.; Abelli, L.; Guerra, L.; Randelli, E.; Proietti Serafin, F.; Belardinelli, M.C.; Buonocore, F.; Bernini, C.; Fausto, A.M.; Scapigliati, G. MHC II-b□chain gene expression studies define the regional organization of the thymus in the developing bony fish Dicentrarchus labrax (L.). Fish Shellfish Immunol. 2015, 42, 483-493, doi:10.1016/j.fsi.2014.11.012.