Immunogenetics (2000) 51 : 905–914
Digital Object Identifier (DOI) 10.1007/s002510000221

ORIGINAL PAPER

Sirje Timmusk · Sylvie Partula · Lars Pilström

Different genomic organization and expression of immunoglobulin light-chain isotypes in the rainbow trout

Received: 23 December 1999 / Revised: 6 June 2000 / Published online: 2 August 2000

Abstract cDNA studies have distinguished two isotypes of the rainbow trout (*Oncorhynchus mykiss*) immunoglobulin (Ig) light chain (designated L1 and L2). This study characterized genomic clones of these isotypes. *L1* genes are arranged in clusters with single copies of variable (*V*), joining (*J*), and constant (*C*) segments. The transcriptional orientation of the *V* genes is opposite to that of the *J* and *C* segments, indicating that the *V* genes must be rearranged by inversion. *L2* is also organized in clusters, consisting of two or three *V*, one *J*, and one *C* exon, all in the same transcriptional orientation. *L1* and *L2* of rainbow trout are similar to the previously identified cod and catfish clusters. Repeat sequences were found upstream of each *J* segment in the *L2* genes, each of which includes a 16-bp sequence similar to the conserved κ sequence motifs of mammalian *J*κ*1* genes. Sequence analyses showed that the regions upstream of *L1* and *L2* genes have several putative *cis*-acting elements also present in the promoter regions of *Ig* genes of other organisms. Octamer motifs, a TATA box, and an E-box were found in the 5′ region of an *IgL1V* gene. A κ-Y element, a CCCT element, a TATA box, an E-box but no classical octamer were found in the 5′ region of the *IgL2* gene. Northern blot analyses showed that *L1* and *L2* are expressed in spleen, head kidney, excretory kidney, thymus, and heart. The expressed ratio of L1 and L2 is estimated to 85:15% in blood and lymphoid tissues.

Key words Evolution · Gene expression · B lymphocytes · Rainbow trout · Immunoglobulins

The nucleotide sequence data reported in this paper have been submitted to the EMBL/GenBank databases and have been assigned the accession numbers AJ251647–AJ251651

S. Timmusk · L. Pilström (✉)
Immunology Programme,
Department of Cell and Molecular Biology, BMC,
Uppsala University, Box 596, 75124 Uppsala, Sweden
E-mail: lars.pilstrom@icm.uu.se
Phone: +46-18-4714218
Fax: +46-18-557579

S. Partula
Institute of Immunology, Johannes-Gutenberg University,
Obere Zahlbacher Strasse 67, 55101 Mainz, Germany

Introduction

Immunoglobulins (Igs) are the effector molecules of specific humoral immunity in all jawed vertebrates. The Ig molecules consist of two heavy (H) and two light (L) chains and the different isotypes of the H chain give the Ig molecule its effector properties. Teleost fish express IgM (of predominantly tetrameric form) and a putative homologue of IgD (Hordvik et al. 1999; Wilson et al. 1997). There are also different isotypes of the L chain, which in mammals are classically referred to as κ and λ. The light chain has no known functional properties except contributing to the antigen-binding structure of the antibodies. Several isotypes of the L chain will thus increase the repertoire of antibodies. All studied vertebrate species, except birds, have more than one L-chain isotype (Pilström et al. 1998).

There are three major types of *Ig* locus organization, denoted multicluster (as in elasmobranchs), minimalistic (as in birds), and translocon (or mammalian). The *IgL* loci of elasmobranch and teleost fish are organized in multiple clusters [variable (*V*)*L*-joining (*J*)*L*-constant (*C*)*L*]. The *VL* segments are found in opposite transcriptional orientation to the *CL* genes in Atlantic cod (*Gadus morhua*) and channel catfish (*Ictalurus punctatus*), and are sometimes duplicated, as in catfish (Warr 1995). The *IgL* locus of Siberian sturgeon (*Acipenser baeri*) is organized in a translocon manner (Lundqvist et al. 1996). Two L-chain variants have been described in the rainbow trout with the aid

of monoclonal antibodies (Sanchez et al. 1989). The gene of the first isotype (L1) was isolated by screening expression libraries with antibodies (Daggfeldt et al. 1993) and the second isotype (L2) using degenerated primers in a polymerase chain reaction (PCR) (Partula et al. 1996).

Studies of the regulatory mechanisms for the expression of the immunoglobulin loci are few. The enhancer of the *IgH* locus in channel catfish has been studied in detail (reviewed in Magor et al. 1999). The presence of a multicluster organization in teleosts and elasmobranchs raises several questions concerning the mechanisms of *IgL* gene regulation, allelic exclusion, and clonal selection. Although much is known about transcriptional regulation of mammalian *IgL* genes (Blomberg et al. 1991; Glozak and Blomberg 1996; Hagman et al. 1990; Meyer and Neuberger 1989; Picard and Schaffner 1984; Rudin and Storb 1992), very little is known about *IgL* transcriptional control in poikilothermic vertebrates. Recently, Bengtén and co-workers (2000) published studies on the enhancer activity of one *IgL* isotype of Atlantic cod, demonstrating that not all clusters contain their own enhancer region. However, no data have been published on the promoter of any isotype of a poikilotherm vertebrate.

In this investigation we studied the genomic organization of the two L-chain isotypes in rainbow trout and their expression in different tissues, and provide a short overview of the first 200 bp 5′ of the *VL* start site, corresponding to the promoter region.

Material and methods

Isolation and mapping of L1 and L2 genomic clones

Duplicate filters of a rainbow trout recombinant genomic library (described by Lee et al. 1993) (amplified from 2×10^6 primary recombinants) were hybridized with trout *IgL1*-specific (Daggfeldt et al. 1993) and *IgL2*-specific (Partula et al. 1996) cDNA probes. The probes were labelled with ^{32}P dCTP using a random priming labelling kit (Amersham, UK). The filters were hybridized in a solution containing 0.5 M NaH_2PO_4/Na_2HPO_4 (pH 7.2), 1 mM ethylenediamine tetraacetic acid 7% (w/v) sodium dodecyl sulfate (SDS) (modified from Church and Gilbert 1984) and the labelled probe. The filters were washed either at low stringency [2×standard sodium citrate buffer (SSC) and 0.1% SDS, at 55 °C] or high stringency (0.2×SSC and 0.1% SDS, at 65 °C).

Isolation of phage DNA and restriction mapping was performed by standard methods (Sambrook et al. 1989). DNA of *L1* and *L2* genomic clones was digested with restriction nucleases either alone or in combination and analyzed by electrophoris on agarose gels. The DNA was transferred to Hybond N^+ nylon membranes (Amersham) and hybridized with rainbow trout *VL*- and *CL*-specific probes. Positive fragments were selected for further analyses and subcloned into the pUC19 plasmid (Pharmacia Biotech, Uppsala, Sweden).

RNA preparation and Northern blot analysis

Total RNA was isolated from brain, eye, dura mater, thymus, heart, head kidney (pronephros), spleen, excretory kidney (the posterior one-third of the organ), liver, muscle, gut, and ovary (pooled organs of three individuals) by the acidic phenol extraction method (Chomczynski and Sacchi 1987). The RNA (10 µg per sample) was fractionated on an agarose gel containing 8.6% formaldehyde and transferred to Hybond N^+ nylon membrane (Amersham) in 20×SSC by capillary blot. The filters were fixed by UV cross-linking (150 mJ), prehybridized and hybridized. Probes including *V*, *J*, and *C* regions were used for Northern blot analyses and the lengths for *L1* were 348 bp and for *L2* 393 bp. Both probes included the last part of the *V*, the whole *J*, and the first half of the *C* region. Probes were radiolabelled by random priming (Megaprime DNA Labelling System, Amersham) and washed at low (2×SSC, 0.1% SDS at 55 °C) or high (0.2×SSC, 0.1% SDS at 65 °C) stringencies and exposed overnight to Kodak XAR films at –70 °C. The same filters were hybridized with a trout *EfTu-1* (ribosomal elongation factor) probe to standardize for the amount of RNA in each lane (Hansen 1997).

For quantitative purposes, total RNA (7 µg per well) from trout white blood cells (three individuals separately), spleen (three individuals separately), and head kidney (two individuals separately) as well as homologous cDNA (*L1* and *L2*) of known concentrations (1 ng, 500 pg, and 100 pg) was separated on gels and hybridized to a ^{32}P-labelled probe corresponding to the full cDNA length of *L1* and *L2*. Quantification of *L1*- and *L2*-specific mRNA to which the probe hybridized was performed in a chemiluminescent imagescreen (GS-250 Molecular Imager; Bio-Rad, Hercules, Calif.) using Image Quant Software and the hybridization signals were used to estimate the relative expression of the respective mRNAs.

PCR analysis of L1 and L2 expression

cDNA from rainbow trout tissue samples and blood was synthesized with reverse transcriptase (RT Superscript II; Life Technologies, Gaithersburg, Md.) using 5 µg of total RNA as template and oligo-$(dT)_{15}$ as primer. PCR reactions were performed in a volume of 25 µl containing 5% of the RT reaction as a template, 20 pmol of each primer, 200 pM dNTP, 2 mM $MgCl_2$, and 0.5 units of *Taq* polymerase (Boeringer Mannheim, Germany). DNA was amplified in a PTC-100 TM thermocycler (MJ Research, Watertown, Mass.) using the following conditions: 94 °C (2 min), 35 cycles of 94 °C (40 s), 55 °C (40 s), and 72 °C (1 min). PCR products were cloned into pBluescript II KS^+vector (Stratagene, California) and sequenced. The following primers were used:

- L1 Vs: CTTCACTCTGACCATCAGTG
- L1Vas: GAATCTAGATGGAGTCCCAGACTGA
- L1Js: GTAGCGGCACCAGACTTGATGTT
- L1 Cas:TCACCTCTCCTGCCTTGGTCCACT
- L2 Vs: TCTAAAGCCACGTCTGATAA
- L2Vas: ACTGAAGGTGTCTGTGTCACTGCT
- L2 Js: AACCAAGCTCTTTGTTACT
- L2 Cas: CACTTTGTCCTGGTTCCATTCTGA

where "s" is the forward primer and "as" is the reverse primer.

Southern blot analysis

Rainbow trout erythrocyte DNA (7 µg per well) was isolated as described by Sambrook and co-workers (1989) and digested with *Hin*dIII, *Eco*RI, and *Bam*HI. The DNA was separated on a 0.8% agarose gel and transferred to a Hybond N^+ nylon membrane (Amersham) by blotting in 0.4 M NaOH. The filters were UV cross-linked (150 mJ), and labelled with ^{32}P using the random primer method (see above). The duplicate filters were hybridized with *V* and *C* probes in a buffer modified from Church and Gilbert (see above) at 65 °C overnight. For genomic Southern blot analyses, the following probes were used: *L1* cDNA probes from nucleotide positions 1 to 254 [containing a leader (*L*) and *V* region] and nucleotide positions 457–587 (a *C* probe) of clone *rtSgE* (Daggfeldt et al. 1993) and *L2* cDNA

probes from nucleotides 1 to 307 containing *L*, *V*, and *J* regions, and nucleotides 422–746 (*C* probe) (Partula et al. 1996). The filters were then washed at 65 °C in 0.1×SSC and 0.1% SDS for 30 min and exposed to X-ray film at –70 °C overnight.

Sequencing

The Sanger dideoxy nucleoside triphosphate chain termination method (Sanger et al. 1977) with T7 DNA polymerase was used for sequencing. All exons were sequenced on both strands and the nucleotide sequences were analyzed using DNAStar software (DNASTAR, Wisconsin). The relative transcriptional orientation of the gene segments in selected genomic clones was determined by sequencing overlapping fragments and by PCR using sense and antisense oligonucleotide primers specific for the *VL* and *JL* genes as well as an antisense primer for the *CL* segment. The presence of two or more *V* exons in the cluster was sometimes determined by PCR using *V* sense and *V* antisense primers.

Results

Structure of IgL1 genes of rainbow trout

Screening the rainbow trout genomic library with a cDNA probe specific for the *L1C* region yielded two genomic clones. PCR analyses, restriction enzyme mapping, and sequencing revealed that one of the isolated clones (*L1E2*) contains one *V*, one *J*, and one *C* exon in the cluster, the *V* exon having an opposite transcriptional orientation to the *J* and *C* exons (Fig. 1) as in Atlantic cod, but this cluster spans 3824 bp compared to approximately 2.1 kb in cod (Bengtén et al. 2000). The other clone (*L1A*) contains one *V*, one *J* and two *C* segments. One of the *C* segments is a pseudogene, since it contains just one half of the exon. The *V* is in opposite transcriptional orientation and is also a pseudogene due to one frameshift. In both *L1* clones, a κ-like recombination signal sequence (RSS) is present at the 3′ end of the *V* segment and at the 5′ end of the *J* segment (the 12-nucleotide spacer is associated with the *V* segment and the 23-nucleotide spacer with the *J* segment). The leader exons encode 15 amino acids (aa) and are separated from the *V* gene by introns of 154 (*L1E2*) and 187 (*L1A*) bp.

Structure of IgL2 genes of rainbow trout

Three different *L2* hybridizing genomic clones, approximately 14–15 kb in length, were isolated. They were characterized using PCR analyses, restriction enzyme mapping, and sequencing (Fig. 1). The *L2* genes have two or three *V*, one *J*, and one *C* segment in each cluster but the direction of transcription is identical for all the segments. The size of a complete cluster with three *V* segments (clone *L2C31*) is 5172 bp. A κ-like RSS is present at the 3′ end of all *V* segments and at the 5′ end of the *J* segments, suggesting that all the *V* segments in the cluster may be functional. The well-conserved (both in sequence and length) leader exons, encoding 13 aa, are separated from the *V* exons by introns 90–100 bp in length.

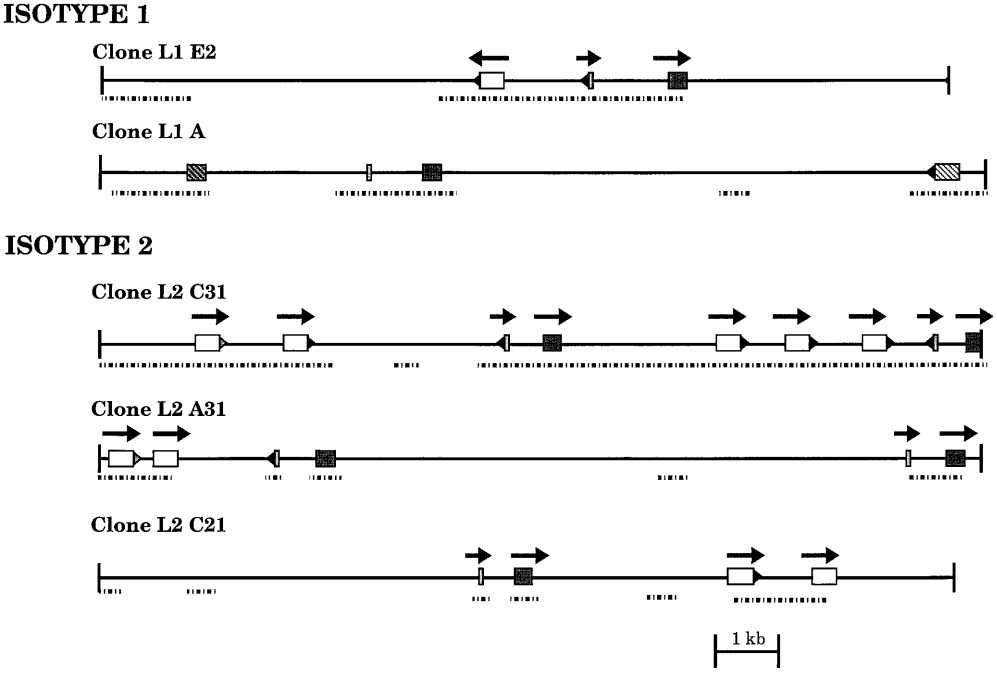


Fig. 1 Physical map of rainbow trout recombinant genomic phage clones of the two light-chain isotypes (*L1* and *L2*). The insert lengths are approximately 14–15 kb. *Open boxes* indicate the localization of *V*, *small filled boxes* of *J*, and *filled boxes* of *C* exons. *Hatched boxes* indicate pseudogenes found in clone *L1A*. *Arrows* show the transcriptional orientation of the gene segments. The heptamer and nonamer signal sequences are marked with *triangles* and sequenced areas are denoted with *hatched lines* below each clone. All segments are drawn to scale. Accession numbers (size of the sequenced part): *L1E2* AJ251651 (5061 bp); *L2C31-1* AJ251647 (2621 bp); *L2C31-2* AJ251648 (8959 bp); *L2A31* AJ251649 (1706 bp); *L2C21* AJ251650 (1602 bp)

Analysis of the IgL1 and IgL2 V regions

Compared to each other, the *V* genes of the *L1* isotype have a variable framework region 1 (FR1), and well-conserved FR2 and FR3. Complementarity-determining region (CDR)1 and 3 exhibit great variability but CDR2 appears quite conserved. The *V* genes also vary in total length (Fig. 2). The inferred *IgL1* sequence derived from clone *L1E2* shares only 65% of its residues at the amino acid level with the previously published cDNA sequence *RtSgE* (Daggfeldt et al. 1993).

In the *L2* isotype, the CDR1 and 3 regions as well as the beginning of the FR3 region show variability; FR1, FR2 and, surprisingly, the CDR2 regions are the most conserved (Fig. 2). Sequences of the complete *IgL2V* regions were compared to the previously published cDNA *VL2a* sequences (Partula et al. 1996). Identities range from 45 to 94% (at the amino acid level). The three clones *L2A31V1*, *L2C31V3*, and *L2C31V4* have high similarity to each other and to *VL2a.* Four other genomic *V* genes (Fig. 2) share up to 62% of residues with each other and 45–51% with *VL2a*. Thus they can be considered as emerging from different families of *VL* genes, here denoted as *VL2-1*, *VL2-2*, *VL2-3*, and *VL2-4* (Fig. 2).

P-nucleotide and putative N-nucleotide additions were examined in *L2* cDNA clones by sequencing from FR3 of the *V* regions to the second half of the *C* regions. All clones encoded unique rearranged *L2* sequences, which exhibited variability in CDR3. No nucleotide additions were detected in the 3′ end of the *V* region of the 32 sequenced cDNA clones. However, deletion of two nucleotides was found in the 3′ end of the *V* genes of one clone (data not shown). The inferred amino acid sequences of the last part of the variable regions of ten *IgL2* cDNA clones from peripheral blood lymphocytes (PBLs) are compared in Fig. 3 with the genomic *L2* sequences. Analysis of the amino acid sequence of the genomic *IgL2V* genes indicates that there are some differences in the FR3 region and a good variability in the CDR3 region.

Genomic *J* regions of *L2* are quite conserved and differ by three nucleotides, resulting in two amino acid changes. Likewise, the genomic *C* regions of *L2* are very similar with only three amino acid substitutions. As the cDNA analysis of ten clones shows, each of the rainbow trout *IgL2J* and *C* regions are very similar with few amino acid exchanges. Four unique *J* and ten unique *C* regions were identified (Fig. 3), indicating that there are likely at least ten clusters in the locus. This is comparable to observations in other "multicluster organization" *CL* regions of e.g., Atlantic cod, channel catfish, common carp (*Cyprinus car-*

Fig. 2A,B Inferred amino acid sequences of the variable regions of *IgL1* and *IgL2* genomic clones. The framework regions (FR) and complementarity-determining regions (CDR) have been delineated according to Kabat and co-workers (1991) (**A**) and Lefranc (1999) (**B**). *Dots* indicate identity to the published cDNA sequences and *dashes* are used to maximize the alignment. Alignment after Lefranc (IMGT) needs an insertion of five amino acids into CDR2 and gaps in FR1 to fit the positions of conserved amino acids (**B**). The inferred *IgL1E2* clone and already published *L1* cDNA clone *RtSgE* (Daggfeldt et al. 1993) are 65% identical at the amino acid level and form two different families. Seven *IgL2* clones analyzed share 45–94% of their residues with the already published *L2* cDNA clone *VL2a* (Partula et al. 1996). New families found are named at the end of each sequence

A

```
L1                   FR1                 CDR1           FR2               CDR2     FR3                                             CDR3

cDNA V     QITVTQTPEMKAFQTGDAVTLRC RFNKPPACSPPCVA WYQQKPGGAPGLLIY-- YATTLQS GTPSRFSGSGSGSD-------------FTLTISGVQAEDAGDYYC QSYHSGDV
E2   V     .YV.....VV..VVPEQT.S.N. K-TSSNVYNNNYL. ..........K....-- L...... ..............-------------.................. ..F.YPG

L2

cDNA, V    QKAVTQTPSVLTVSTKGTATFHC DITKGEGNVVI--- WYKQVPGGAPQYVLR-- YYHTWSS FTSSSPDQYGYGFSSDRFTSKATSDKDYQFIISNVEETDSAVYYC QTWYSSVKVH
A31, V1    .....................D. .M..D.K.D.S--- ...............-- ....... ............................................. ...D......V  VL2-1
C31, V4    ....................... ....D..S..---  ..............K-- ....... S......K..S................................... H..D......I
C31, V3    ....................... ....D..S..---  ..............K-- ...S... S......K..S................................... H..D......I
C31, V2    VTV...K.P.V..RKGE...LD. NLGTVQ-NSAAR-- ...........F..YF- -..D--- ---N.APS..S....-K...DH..ES..RL..NT.......... ........EYV  VL2-2
C31, V1    VTV...K.P.V..RKGE...MD. NLGTVTEKL.TN-- ........V..SCDKVL PWIW.-- -----.ILM.S....PK.......ES..R...N....E....... K..D....EYV  VL2-3
C31, V5    VIVL..K.Y.Q.KT.GEEV.MD. N.GRFD.NYVH--- .......AV.....S-- FIYGD-- ---.LLNIM.H.......Q..S...L......KQ..AG....... H.YDD.AEEFV  VL2-4
C21, V2    VIVL..K...Q.KT.GEEV.MD. N.GRFD.TMSI--- ....-S..V.....S-- FIYGD-- ---..SEY..PP...S..NY.S...L......KQ..AG....... ..YDD.AEEFV
```

B

```
L1                   FR1                   CDR1         FR2                CDR2                FR3                            CDR3

cDNA V     QITVTQTPEMKAFQTGDAVTLRCRFN KPPACSPPC VAWYQQKPGGAPGLLIY  -----YATTLQS-- ----GTPSRFSGSGSGSDFTLTISGVQAEDAGDYYC  QSYHSGDV
E2   V     .YV.....VV..VVPEQT.S.N.K-T SSNVYNNNY L...........K....  -----L......-- ----..................................  ..F.YPGS

L2

cDNA, V    QKAVTQTPSVLTVSTKGTATFHCDIT  KGEGNV-- VIWYKQVPGGAPQYVLR YYHTWSSFTSSSPDQ YGYGFSSDRFTSKATSDKDYQFIISNVEETDSAVYYC QTWYSSVKVH
A31, V1    .....................D..M.  .D.K.D-- .S............... ............... ..................................... ....D......V  VL2-1
C31, V4    ..........................  .D...S-- ................K ......S......K ..S.................................. H..D......I
C31, V3    ..........................  .D...S-- ................K ...S...S......K ..S.................................. H..D......I
C31, V2    VTV...K.P.V..RKGE...LD.NLG  TVQ-NSA- AR..........F..Y F..DN.APS..S... .K...DH..ES..RL..NT.................. ..........EYV  VL2-2
C31, V1    VTV...K.P.V..RKGE...MD.NLG  TVTEKLV- TN........V..SCDK VLPWIW.-------- ILM.S....PK.......ES..R...N....E..... ..K..D....EYV  VL2-3
C31, V5    VIVL..K.Y.Q.KT.GEEV.MD.N.G  RFD.NY-- VH.......AV.....S FIYGD.LL------- NIM.H.......Q..S...L......KQ..AG..... ..H.YDD.AEEFV  VL2-4
C21, V2    VIVL..K...Q.KT.GEEV.MD.N.G  RFD.TM-- SI....-S..V.....S FIYGD..S------- EY..PP...S..NY.S...L......KQ..AG..... ....YDD.AEEFV
```

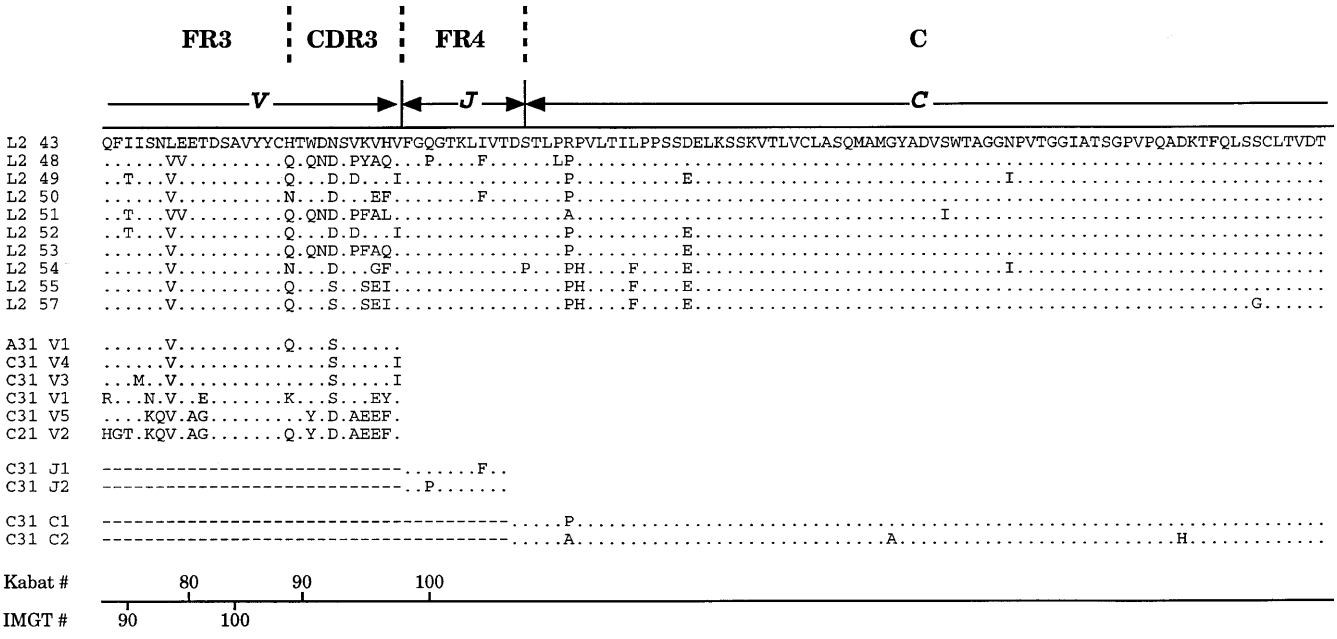


Fig. 3 Inferred amino acid sequences of codons 95–223 of the *C* region of ten *L2* cDNA clones from peripheral blood lymphocytes are compared with those *IgL2* sequences found in genomic clones. FR3 and CDR3 are delineated according to Kabat and co-workers (1991). Analysis of the amino acid sequences of *IgL2* genomic *V* segments and cDNA *V* segments indicates that families may be found at the genomic level which may be less common at the mRNA level. Four unique *J* and ten unique *C* regions were identified. Amino acid positions are denoted below the sequences

pio), horned shark (*Heterodontus francisci*), and sandbar shark (*Carcharhinus plumbeus*) (Daggfeldt et al. 1993; Ghaffari and Lobb 1993, 1997; Rast et al. 1994; Schluter et al. 1989; Shamblott and Litman 1989a, 1989b; Tomana et al. 1999).

Putative promoter regions of L1 and L2

Consensus sequences of the transcription factor-binding sites were studied in both isotypes from an alignment of the putative promoter regions (position –200 to 0) from all potentially functional genes within the *L2* clones, and one pseudogene and one functional gene within the *L1* clones (Fig. 4). Since the *VL* genes in trout *IgL1* are in an opposite transcriptional orientation, introns between *VL* and *JL* of the *L1* isotype contain promoter-like sequences that are also in the opposite orientation. These studies showed that *L1E2* has two conserved octamer motifs (-TTTGCAT) at positions –140 bp and –158 bp, an E-box (CACAGT) at position –128 bp and TATA boxes at positions -102 bp and –63 bp (Fig. 4). In contrast, the *L2* sequences contain no conserved octamer motifs. Two octamer-like sequences are present in the 5′ region of each *L2* segment, but in one of them only 4 bp (-TTTG–-) match the consensus sequence (ATTTGCAT). This octamer-like sequence was found in all *L2* promoter regions at various distances from the ATG (–195 to –344 bp, not shown in the figure). The second octamer-like motif is located at a constant distance from the ATG (always at –94 bp) in each of the seven clones but only 3 bp (-T-T–-T) match the consensus octamer. Both of these octamer-like motifs are unlikely to be functional. Putative TATA boxes in all seven *L2* promoter regions were either –30 or –97 bp from the ATG with consensus sequences AGTAA and TTATA, respectively. Overlapping the proximal octamer-like sequence (and 5′ of it), an 18-bp well-conserved region was found at –99 bp from the ATG where 6 bp (TCTCCTTA) of 8 bp matched the consensus sequence of the κ-Y element (with consensus CTTCCTTA; known as PU.1- and elf-1-binding site). A putative CCCT element (an octamer costimulatory site) was found at –55 bp and an E2A-type E-box (CAGCTG; bHLH-binding site) in all seven sequences –38 bp from the ATG.

Each *L2 J* segment also contains a conserved κ sequence motif (KI and KII) at the 5′ end which is repeated four to six times in the genome (Fig. 5). These sequences are hypothesized to act as binding sites for the protein KLP, the κ locus rearrangement-enhancing element. The repeated sequences in the trout genome include a 16-bp conserved sequence which is similar to KI motifs of human and mouse $J\kappa$ genes. This structure differs from KI of human and mouse $Ig\kappa$ genes at four positions (Fig. 5). No KI- or KII-like sequences were found in the region 5′ of the *J* exon of the trout *L1* isotype.

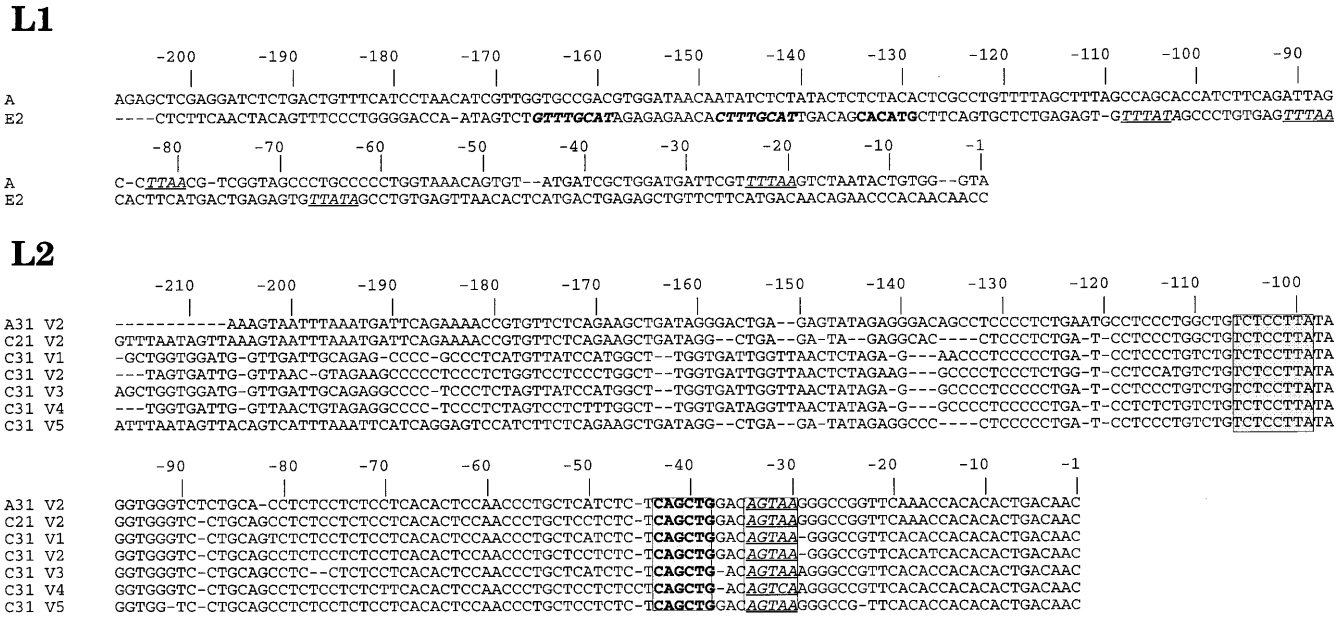
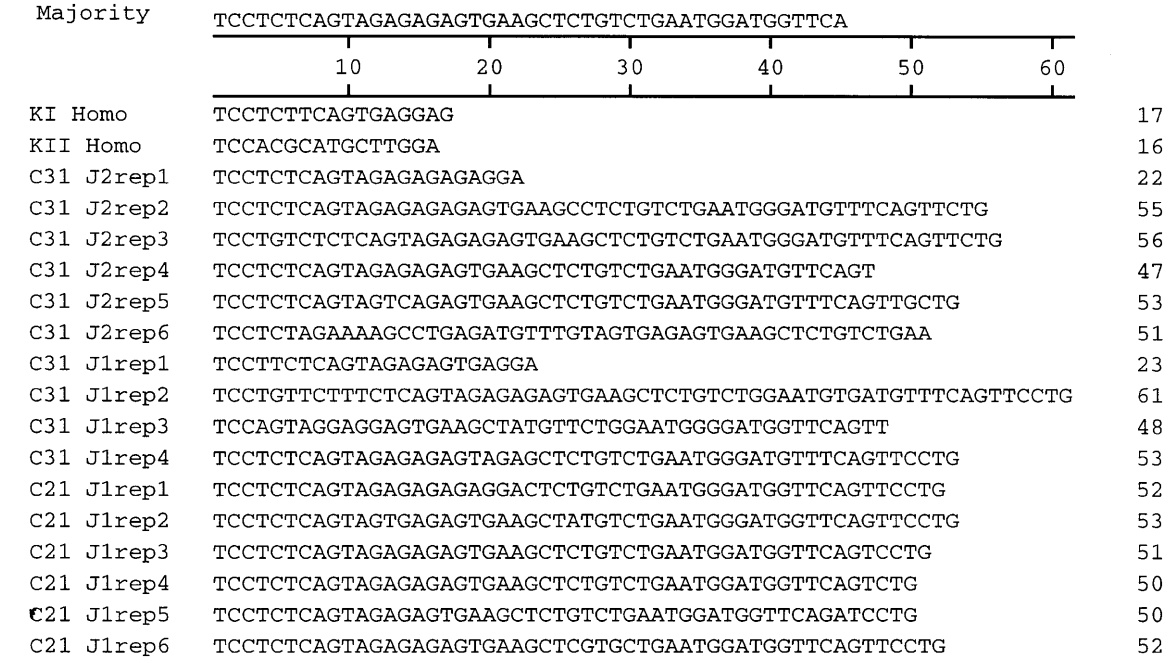


Fig. 4 Alignment of consensus sequences of the putative promoter regions from all potentially functional genes (*L2*), and one pseudogene (*L1A*) and one functional gene (*L1E2*) within the clones. The analyzed sequences are from position –200 to 0 bases (where the ATG codon starts at position +1). Putative promoter elements of *L2* are *boxed*. E-boxes are denoted in *bold*, putative TATA boxes in *underlined italics* and octamer in *bold italics*. The putative κ-Y element is located from –99 to –106 bp, and a CCCT element (not marked) at –55 bp

Analyses of L1 and L2 mRNA expression

Northern blot analyses were performed to compare the expression of *L1* and *L2* isotypes in different tissues. *L1* transcripts were expressed at high levels in the typical lymphoid tissues like spleen and head kidney but at low levels in the excretory kidney, thymus, and heart. Similarly, *L2* transcripts were also expressed at a high level in spleen and head kidney, with lower levels in heart, thymus, and excretory kidney (Fig. 6). *IgL1* had one predominant message with an approximate length of 1 kb, while *L2* had two different mRNA transcripts of approximately 1 and 0.7 kb.

Fig. 5 Alignment of repeated sequences in the trout genome. The 5′ end of each *J* segment of trout *IgL2* contains a conserved κ sequence motif (*KI* and *KII*), repeated in the genome four to six times. Human κ locus KI and KII sequences are shown on rows 1 and 2, respectively. Joining segments of clone *C31* are enumerated from 5′ to 3′ *J1* and *J2*, respectively



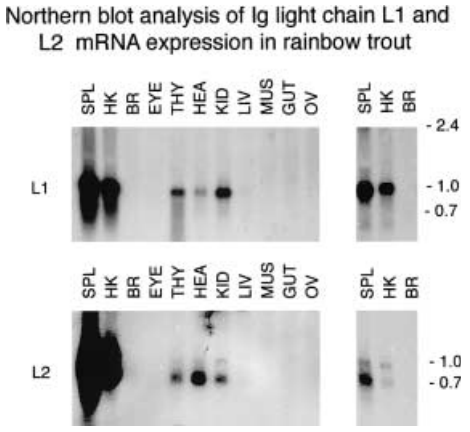


Fig. 6 Northern blot analysis of total RNA (10 μg per well). From the left: trout spleen (*SPL*), head kidney (*HK*), brain (*BR*), eye (*EYE*), thymus (*THY*), heart (*HEA*), kidney (the last one-third of the organ, *KID*), liver (*LIV*), muscle (*MUS*), gut (*GUT*), and ovary (*OVA*). In the three rightmost lanes: total RNA (7 μg per well) from trout spleen (*SPL*), head kidney (*HK*), and brain (*BR*). The filters were washed at low stringency (*upper part*) and at high stringency (lower *part*). Sizes in kilobases are indicated to the *right*

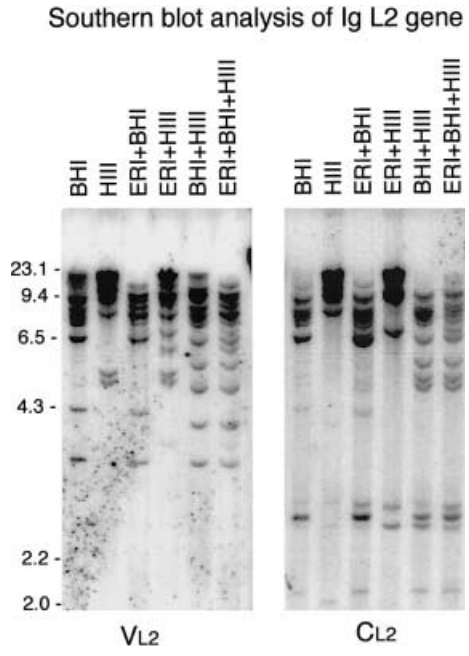


Fig. 7 Southern blot analysis of genomic (erythrocyte) DNA for isotype *L2* of rainbow trout with a probe containing the variable domain (*left*) and the constant domain (*right*). The restriction enzymes used were *Bam*HI (*BHI*), *Hin*dIII (*HIII*), and *Eco*RI (*ERI*), and the position of size markers (in kb) are shown to the *left*. *Bam*HI cuts within the constant exon. The Southern blot of *L1* digested with the same restriction enzymes is illustrated by Daggfeldt and co-workers (1993)

Table 1 Relative expression (%) of mRNA of isotypes *L1* and *L2* in leukocytes, spleen, and head kidney of rainbow trout. The data are mean values from two (head kidney) and three (leukocytes and spleen) different individuals. For details see Material and methods

Isotype	Tissue	Mean value	Range
L1	Leukocytes	84.5	80–87.5
	Spleen	87.7	85–93
	Head kidney	83	81–85
L2	Leukocytes	15.5	13–20
	Spleen	12.3	7–15
	Head kidney	17	15–19

For quantitative purposes, *L1* and *L2* mRNA expression were measured with a phosphoimager, and relative values (%) were calculated relative tousing homologous cDNA (*L1* or *L2*) of known concentration as a standard. *L1* is expressed at about 85% and *L2* at 15% both in blood and in lymphoid tissues (Table 1). The amount of RNA in each lane was approximately the same as determined by expression of EfTu-1 in the same lanes (data not shown).

Discussion

The present study confirms an earlier hypothesis, based on Southern blot analysis, that the two trout Ig light-chain gene loci have a multicluster-type organization (Daggfeldt et al. 1993; Partula et al. 1996). An extensive Southern blot analysis of genomic trout DNA and the mapping of genomic clones with *V* and *C* probes of trout *L2* indicates that there is more than one *V* exon in each cluster (Figs. 1, 7). The results of the genomic sequence studies verify this, and show that trout *IgL2* has two or three *V* genes in each cluster, all in the same transcriptional orientation. The structure of trout *IgL2* clusters is not typical of those described to date in teleost fish. While two inverted *V* genes in the same cluster have previously been found in the catfish Ig light-chain locus of the G isotype (Ghaffari and Lobb 1993), three *V* genes in the same cluster is unique.

Physical mapping of genomic clones of trout *IgL1* indicated a 1:1 ratio of *V* and *C* exons in each cluster. Sequencing confirmed this and revealed an inverted *V* gene, as reported for cod (Bengtén et al. 2000).

We can hypothesize that in fish phylogeny, different evolutionary strategies were used to create antibody variability. The original unit of a light chain, $(VJC)_1$, may either have been multiplied completely in total length, giving rise to the classical multicluster type of organization (cartilaginous fish and teleosts), or partially [(VJ) or (V),(J)], giving rise to the translocon type of organization (chondrostei and tetrapoda). Those fish species which migrate between different environments, i.e., from salt to fresh water, or which live in different temperatures, need to protect themselves against a broader spectrum of pathogens specific for each habitat. Under such selective pressure, such fish (and perhaps different light-chain classes) probably evolved distinct genetic features to

increase variability. Acquiring either an inverted *V* gene in a cluster, which may allow rearrangement between clusters (cod *L1*, trout *L1*) or creating several *V* genes inside one cluster (trout *L2*) or both (catfish) will increase the diversity. This hypothesis will require further studies of the *L* loci to be verified or rejected.

Stretch region

A particular feature of trout *VL2* genes is an unusual stretch of 11 aa in the FR3 region (Partula et al. 1996). This stretch was predicted to form a coiled structure lying between the presumptive D and E β strands. This loop is rich in charged (Arg, Lys, Asp) and hydrophilic (Thr, Ser, Ala) residues and could be near the CDR1–3 loops that form the antigen recognition site of the V domain (Partula et al. 1996). This conserved region was found in all *V* genes of the *L2* isotypes, but not in the *L1* isotype nor in light chains of any other species. However, another interpretation is also possible. A longer CDR2 as depicted in Fig. 2B will allow, according to IMGT URL (http://imgt.cines.fr:8104), a folding of the molecule, and might affect the repertoire estimation.

KI and KII sequences

Conserved κ motifs were located 5′ of each *J* exon in *IgL2* clusters (close to the nonamer). The motifs are 50–56 bp in length, and are repeated four to six times in each *L2* cluster. These motifs always include a 16-bp conserved sequence similar to the K-sequence motifs of human and mouse *Jκ1* genes. Two *cis*-acting elements, KI and KII, separated by 38 bp from each other, are located upstream of the *Jκ* nonamer-heptamer RSS sequences (Weaver and Baltimore 1987). The KI and KII sites are the DNA-binding sites of the κ locus protein (KLP; other names EBB-1 or Pax 5) which is expressed in pre-B- and B-cell lines, but not in plasma cell lines and nonlymphoid cell lines (Weaver and Baltimore 1987). Studies of knockout mice mutated in the KI and KII sites showed that the sites in the region upstream of Jκ are the "enhancer" elements of κ light-chain gene rearrangement; disruption of KLP binding by mutation of the KI and KII sites significantly decreased κ-chain rearrangement, but did not inhibit κ germline transcription (Ferradini et al. 1996). Thus KLP binding to KI and KII sites enhances κ locus rearrangement. KLP was recently found to be the only lineage-specific transcription factor to fulfil this criterion, as pro-B cells from *Pax5–/–* knockouts can differentiate along multiple myeloid and lymphoid lineages except the B-cell pathway. Restoration of *Pax5* activity overcomes the B-cell developmental block (Nutt et al. 1999). Similar repeats were found in three nonrearranged *JC* transcripts of trout *IgL2* (Partula et al. 1996). No κ-like sequences were found 5′ of the *J* exon of the trout *L1* isotype.

Promoter region

Expression of *Ig* genes in mouse and human is regulated at multiple levels. The rearrangements of *Ig* genes are unique for B cells and other cell types cannot express *Ig* genes even if they express the relevant transcription factors. Thus, there appears to be regulation at the level of accessibility for rearrangement of *Ig* genes. One hypothesis is that sterile germline transcripts, initiated from *Ig* promoters and pseudopromoters open the loci for access by the *RAG* gene products that initiate the rearrangement (Blackwell et al. 1986; Kemp et al. 1980; Schlissel et al. 1991). The regulatory regions for *Ig* transcription are similarly organized in the different Ig chains in mammals. All *H* and *L V* regions in mammals have distinct promoters, and depending on the rearrangement events, a different promoter will be used. All the *Vκ* gene promoters have an octamer element. *Igκ* promoter sequences are centered around an octamer (ATTTGCAT) and usually contain an E-box (-CANNTG-), κ-Y element (Atchison et al. 1990), the CCCT element (Högbom et al. 1991) and an initiator-like sequence. The initiator-like sequence is at approximately –2 up to 39 bp (basepair) nonconserved sequence, overlapping with and downstream of ATG site (Pelletier et al. 1997).

All *H* and *L* promoters share a highly conserved octamer motif (ATTTGCAT) that is positioned approximately 70 bp upstream of the transcription start site. Although the complementary sequence is preferred, only seven nucleotides (-TTTGCAT) are strictly conserved in human and mouse. Synthetic promoters containing just an octamer site and a TATA box can direct accurate and B-lineage restricted transcription, thus highlighting the importance of the octamer element in the tissue-specific expression of *Ig* promoters (Dreyfus et al. 1987; Wirth et al. 1987). The promoter octamer motif is well-established as an important component in regulation of mammalian *Ig* gene expression (Bergman et al. 1984; Falkner and Zachau 1984). In the chicken light-chain promoter, mutation of the octamer motif sequence (ATTTGCAT to ATATTCAT) completely inactivated a construct that was active before mutation (Heltemes et al. 1997). At present, relatively little is known concerning the regulation of *IgL* genes in fish. Bengtén and co-workers (2000) recently published studies on the enhancer activity of *IgL1* in Atlantic cod. The lack of functional promoter studies make it difficult to draw conclusions concerning what constitutes a fish *IgL* promoter. Inferences can only be made by comparing the sequences. Promoter regions of five sandbar shark (*C. plumbeus*) genomic clusters show variation in TATA box position, sequences, and orientation of the octamer (Hohman et al. 1995). The rainbow trout *IgL2* promoter region does not have a typical octamer. Two octamer-like sequences can be found although they are probably not functional. Pro-

moter regions in type I L-chain loci from horned shark (*H. francisci*) (Shamblott and Litman 1989b) and *L* chains type II from sandbar shark, spotted ratfish (*Hydrolagus colliei*), horned shark, and little skate (*Raja erinacea*) each contain the consensus octamer ATTTGCAT (Rast et al. 1994). However, the type I L-chain locus in little skate does not contain a consensus octamer. So functional studies are needed to reveal which motifs are functional and important in regulating the expression of the two *IgL* loci in trout.

Interestingly, the octamer element has been shown to enhance human *Vλ* gene rearrangement (Stiernholm and Berinstein 1995). Although trout *IgL2* is neither κ nor λ like, a variant octamer could result in a lower frequency of rearrangements, which might explain why the *L2* isotype has an additional rearrangement-enhancing element, the putative KLP protein-binding site found 5′ of the *J* region. In all seven *L2* promoter regions, putative TATA boxes are at about the same distance (–29 or –30 bp) from the ATG with the consensus sequence AGTAA. A mammalian consensus TATA box (TTATA) can be found at –97 bp from the ATG but this sequence partly overlaps with the κ-Y motif at –99 bp. To determine which is the functional TATA box, additional expression studies will be needed.

From the *L1* isotype, two putative promoter regions were analyzed. One (from clone *L1A*) is a pseudo *V* (with one frameshift) and the sequence 200 bp 5′ of its start codon does not contain any identifiable promoter-like elements indicating possible new DNA insertion(s) to the area or *V* exon translocation elsewhere. The other *V* gene of *L1* clone *L1E2* has TATA boxes (TTATA) –63 and –102 bp from the start ATG site and two typical octamers (-TTTGCAT) at positions –140 and –158 bp, respectively. Fish are the earliest vertebrates in which Igs are found (Warr 1995). The fact that *IgL1* of rainbow trout has similar octamer sequences to those of human and mouse may indicate that *L1* is an evolutionarily younger isotype. *L2*, on the other hand, might be more ancient and may derive from a period when regulation of *Ig* transcription was not yet so well developed.

Expression

Northern blot analysis showed that *IgL1* has one predominant (1-kb) mRNA. The transcription and splicing of nonrearranged or truncated *IgL1* transcripts is uncommon and has so far only been described in Atlantic cod (Daggfeldt et al. 1993). *IgL2* has two different transcripts of 1 and 0.7 kb. Studies were performed on a large set of tissue samples to test whether *L2* is expressed in nonlymphoid organs at levels detectable by Northern blot. *L1* and *L2* transcripts are both expressed at high levels in spleen and head kidney and at lower levels in the excretory kidney, thymus, and heart. All other tested tissues gave negative results. However, *L2* but not *L1* expression was detected at low levels by PCR in several nonlymphoid tissues (data not shown). One explanation might be the missing octamer motif in the *L2* promoters. The TATA box and initiator site are not tissue specific and nonlymphoid transcription factors possibly influence *L2* gene expression. The other explanation could be contamination by blood, but *L1* expression in the nonlymphoid tissues was not detected, although *L1* has a higher level of expression in blood cells than *L2* (Table 1).

In summary, the results presented here demonstrate that the two light-chain gene types of rainbow trout have different cluster organizations. First, *L1* has one inverted *V*, one *J*, and one *C* in the cluster, while *L2* has two or three *V*, one *J*, and one *C* in the cluster, all in the same transcriptional orientation. The amino acid sequence comparison confirmed that some *V* genes of a trout *L2* cluster are probably tandemly duplicated within the cluster (similarity 99%), while others have likely been transposed from elsewhere (similarity 45–51%). Second, the cluster size of *L1* is 3–4 kb and of *L2* around 5 kb. Third, *L1* is expressed at approximately fivefold higher levels than *L2* both in blood and in lymphoid tissues. Fourth, *L1* has putative promoter sequences, which contain both a classical octamer and a TATA box. The putative promoter of *L2* has a well-conserved sequence in the first 100 bp 5′ of the ATG but neither an octamer nor a TATA box were conserved. Fifth, *L1* is expressed with one transcript and *L2* with two transcripts. Thus, *L2* deviates in several respects from the common features of an Ig light-chain locus.

Acknowledgements We acknowledge the skilful technical assistance of Ann-Sofie Lundbäck and Anna-Lena Matson. We thank Dr. Gregory Warr, Dr. Melanie Wilson, Dr. Tönis Timmusk, and anonymous reviewers for critical comments on the manuscript. This work was supported by grants from The Swedish Forestry and Agricultural Council (SJFR 40.0484/96) and from the European Community (FAIR CT 95-0666 and CT98-4026).

References

Atchison M, Delmas V, Perry RP (1990) A novel upstream element compensates for an ineffectual octamer motif in an immunoglobulin Vκ promoter. EMBO J 9:3109–3117

Bengtén E, Strömberg S, Daggfeldt A, Magor G, Pilström L (2000) Transcriptional enhancers of immunoglobulin light chain genes in Atlantic cod (*Gadus morhua*). Immunogenetics 51:647–658

Bergman Y, Rice D, Grosschedl R, Baltimore D (1984) Two regulatory elements for immunoglobulin κ light chain gene expression. Proc Natl Acad Sci USA 81:7041–7045

Blackwell TK, Moore MW, Yancopoulos GD, Suh H, Lutzker S, Selsing E, Alt FW (1986) Recombination between immunoglobulin variable region gene segments is enhanced by transcription. Nature 324:585–589

Blomberg BB, Rudin CM, Storb U (1991) Identification and localization of an enhancer for the human λ L chain Ig gene complex. J Immunol 147:2354–2358

Chomczynski P, Sacchi N (1987) Single-step method of RNA isolation by acid guanidinium thiocyanate-phenol-chloroform extraction. Anal Biochem 162:156–159

Church GM, Gilbert W (1984) Genomic sequencing. Proc Natl Acad Sci USA 81:1991–1995

Daggfeldt A, Bengtén E, Pilström L (1993) A cluster type organization of the loci of the immunoglobulin light chain in Atlantic cod (*Gadus morhua* L.) and rainbow trout (*Oncorhynchus mykiss* Walbaum) indicated by nucleotide sequences of cDNA and hybridization analysis. Immunogenetics 38:199–209

Dreyfus M, Doyen N, Rougeon F (1987) The conserved decanucleotide from the immunoglobulin heavy chain promoter induces a very high transcriptional activity in B-cells when introduced into an heterologous promoter. EMBO J 6:1685–1690

Falkner FG, Zachau HG (1984) Correct transcription of an immunoglobulin κ gene requires an upstream fragment containing conserved sequence elements. Nature 310:71–74

Ferradini L, Gu H, De Smet A, Rajewsky K, Reynaud C-A, Weill J-C (1996) Rearrangment-enhancing element upstream of the mouse immunoglobulin kappa chain J cluster. Science 271:1416–1420

Ghaffari SH, Lobb CJ (1993) Structure and genomic organization of immunoglobulin light chain in the channel catfish: an unusual genomic organizational pattern of segmental genes. J Immunol 151:6900–6912

Ghaffari SH, Lobb CJ (1997) Structure and genomic organization of a second class of immunoglobulin light chain genes in the channel catfish. J Immunol 159:250–258

Glozak MA, Blomberg BB (1996) The human λ immunoglobulin enhancer is controlled by both positive elements and developmentally regulated negative elements. Mol Immunol 33:427–438

Hagman J, Rudin CM, Haasch D, Chaplin D, Storb U (1990) A novel enhancer in the immunoglobulin λ locus is dublicated and functionally independent of NFκB. Genes Dev 4:978–992

Hansen JD (1997) Characterization of rainbow trout terminal deoxynucleotidyl transferase structure and expression: TdT and RAG1 co-expression define the trout primary lymphoid tissues. Immunogenetics 46:367–375

Heltemes LM, Tuggle CK, Lamont SJ (1997) Octamer function in the chicken lambda immunoglobulin light chain promoter. Immunogenetics 47:73–76

Högbom E, Magnusson A-C, Leanderson T (1991) Functional modularity in the SP6 κ promoter. Nucleic Acids Res 19:4347–4354

Hohman VS, Schluter SF, Marchalonis JJ (1995) Diversity of Ig light chain clusters in the sandbar shark (*Carcharhinus plumbeus*). J Immunol 155:3922–3928

Hordvik I. Thevarajan J, Samdal I, Bastani N, Krossøy B (1999) Molecular cloning and phylogenetic analysis of the Atlantic salmon immunoglobulin D gene. Scand J Immunol 50:202–210

Kabat EA, Wu TT, Perry HM, Gottesman KS, Foeller C (1991) Sequences of proteins of immunological interest, 5th edn. US Department of Health and Human Services, Public Health Servic, National Institute of Health, Bethesda, Md

Kemp DJ, Harris AW, Adams JM (1980) Transcripts of the immunoglobulin Cμ gene vary in structure and splicing during lymphoid development. Proc Natl Acad Sci USA 77:7400–7404

Lee MA, Bengtén E, Daggfeldt A, Rytting A-S, Pilström L (1993) Characterisation of rainbow trout cDNAs encoding a secreted and membrane-bound Ig heavy chain and the genomic intron upstream of the first constant exon. Mol Immunol 30:641–648

Lefranc MP (1999) The IMGT unique numbering for immunoglobulins, T cell receptors and Ig-like domains. Immunologist 7:132–135

Lundqvist M, Bengtén E, Strömberg S, Pilström L (1996) Ig light chain gene in the Siberian sturgeon (*Acipenser baeri*): implication for the evolution of the immune system. J. Immunol 157:2031–2038

Magor BG, Ross DA, Pilström L, Warr GW (1999) Transcriptional enhancers and the evolution of the IgH locus. Immunol Today 20:13–17

Meyer KB, Neuberger MS (1989) The immunoglobulin κ locus contains a second, stronger B-cell-specific enhancer which is located downstream of the constant region. EMBO J 8:1959–1964

Nutt SL, Heavey B, Rolink AG, Busslinger M (1999) Commitment to the B-lymphoid lineage depends on the transcription factor Pax5. Nature 401:556–562

Partula S, Schwager J, Timmusk S, Pilström L, Charlemagne J (1996) A second immunoglobulin isotype in the rainbow trout. Immunogenetics 45:44–51

Pelletier MR, Hatada EN, Scholz G, Scheidereit C (1997) Efficient transcription of an immunoglobulin κ promoter requires specific sequence elements overlapping with and downstream of the transcriptional start site. Nucleic Acids Res 25:3995–4003

Picard D, Schaffner W (1984) A lymphocyte-specific enhancer in the mouse immunoglobulin κ gene. Nature 307:80–82

Pilström L, Lundqvist ML, Wermenstam NE (1998) The immunoglobulin light chain in poikilothermic vertebrates. Immunol Rev 166:123–132

Rast JP, Anderson MK, Ota T, Litman RT, Margittai M, Shamblott MJ, Litman GW (1994) Immunoglobulin light chain class multiplicity and alternative organizational forms in early vertebrate phylogeny. Immunogenetics 40:83–99

Rudin CM, Storb U (1992) Two conserved essential motifs of the mouse immunoglobulin λ enhancers bind B-cell-specific factors. Mol Cell Biol 12:309–320

Sambrook J, Fritsch EF, Maniatis T (1989) Molecular cloning: a laboratory manual. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY

Sanchez C, Dominguez J, Coll J (1989) Immunoglobulin heterogeneity in the rainbow trout, *Salmo gairdneri* Richardson. J Fish Dis 12:459–465

Sanger F, Nicklen S, Coulson AR (1977) DNA sequencing with chain-terminating inhibitors. Proc Natl Acad Sci USA 74:5463–5467

Schlissel MS, Corcoran LM, Baltimore D (1991) Virus-transformed pre-B cells show ordered activation but not inactivation of immunoglobulin gene rearrangement and transcription. J Exp Med 173:711–720

Schluter SF, Hohman VS, Edmundson AB, Marchalonis JJ (1989) Evolution of immunoglobulin light chains: cDNA clones specifying sandbar shark constant regions. Proc Natl Acad Sci USA 86:9961–9965

Shamblott MJ, Litman GW (1989a) Genomic organization and sequences of immunoglobulin light chain genes in a primitive vertebrate suggest coevolution of immunoglobulin gene organization. EMBO J 8:3733–3739

Shamblott MJ, Litman GW (1989b) Complete nucleotide sequence of primitive vertebrate immunoglobulin light chain genes. Proc Natl Acad Sci USA 86:4684–4688

Stiernholm NBJ, Berinstein NL (1995) A mutated promoter of a human IgVλ gene segment is associated with reduced germline transcription and a low frequency of rearrangement. J Immunol 154:1748–1761

Tomana M, Nakao M, Moritomo T, Fujiki K, Yano T (1999) Isolation of cDNA encoding immunoglobulin light chain from common carp (*Cyprinus carpio* L.). Fish Shellfish Immunol 9:71–80

Warr GW (1995) The immunoglobulin genes of fish. Dev Comp Immunol 19:1–12

Weaver D, Baltimore D (1987) B lymphocyte-specific protein binding near an immunoglobulin κ-chain gene J segment. Proc Natl Acad Sci USA 84:15516–1520

Wilson M, Bengtén E, Miller NW, Clem LW, Du Pasquier L, Warr GW (1997) A novel chimeric Ig heavy chain from a teleost fish shares similarities to IgD. Proc Natl Acad Sci USA 94:4593–4597

Wirth T, Staudt L, Baltimore D (1987) An octamer oligonucleotide upstream of a TATA motif is sufficient for lymphoid-specific promoter activity. Nature 329:174–178