Published in final edited form as:
Immunogenetics. 2013 May ; 65(5): 387–396. doi:10.1007/s00251-013-0678-9.

Immunoglobulin light chains in medaka (*Oryzias latipes*)

Susana Magadán-Mompó,
Virologie et Immunologie Moleculaires, Institut National de la Recherche Agronomique (INRA), Domaine de Vilvert, 78352 Jouy-en-Josas, France

Anastasia M. Zimmerman,
Department of Biology, College of Charleston, 66 George Street, Charleston, SC 29424, USA

Christian Sánchez-Espinel, and
Shared Unit of Immunology, University of Vigo–Vigo University Hospital Complex (Hospital Meixoeiro), Edificio de Ciencias Experimentales, Rua das Abeleiras, Campus As LagoasMarcosende, Vigo 36310 Pontevedra, Spain

Francisco Gambón-Deza
Unidad de Inmunología, Hospital do Meixoeiro, Servizo Galego de Saúde (SERGAS), Carretera de Madrid s/n, Vigo 36210 Pontevedra, Spain

Susana Magadán-Mompó: smagadan@jouy.inra.fr

Abstract

The gene segments encoding antibodies have been studied in many capacities and represent some of the best-characterized gene families in traditional animal disease models (mice and humans). To date, multiple immunoglobulin light chain (IgL) isotypes have been found in vertebrates and it is unclear as to which isotypes might be more primordial in nature. Sequence data emerging from an array of fish genome projects is a valuable resource for discerning complex multigene assemblages in this critical branch point of vertebrate phylogeny. Herein, we have analyzed the genomic organization of medaka (*Oryzias latipes*) IgL gene segments based on recently released genome data. The medaka IgL locus located on chromosome 11 contains at least three clusters of IgL gene segments comprised of multiple gene assemblages of the kappa light chain isotype. These data suggest that medaka IgL gene segments may undergo both intra- and inter-cluster rearrangements as a means to generate additional diversity. Alignments of expressed sequence tags to concordant gene segments which revealed each of the three IgL clusters are expressed. Collectively, these data provide a genomic framework for IgL genes in medaka and indicate that Ig diversity in this species is achieved from at least three distinct chromosomal regions.

Keywords

Immunoglobulin genes; Teleost immunity; Antibody diversity; Evolution

Correspondence to: Susana Magadán-Mompó, smagadan@jouy.inra.fr.

Electronic supplementary material The online version of this article (doi:10.1007/s00251-013-0678-9) contains supplementary material, which is available to authorized users.

Introduction

Gene segments encoding antibodies have been identified in a range of animals including sharks, teleosts, amphibians, reptiles, birds, monotremes, marsupials, and mammals (Litman et al. 1999; Belov et al. 2002; Flajnik 2002; Marchalonis et al. 2002). Despite extensive efforts to identify orthologous genes in more primitive vertebrates (hagfish and lamprey) and invertebrates (sea urchins, crustaceans, worms, and insects), none have been found (Klein 1989; Marchalonis and Schluter 1990; Smith and Davidson 1992; Rast et al. 2006). These findings suggest that Ig segments emerged early during the evolutionary history of vertebrates, conceivably during the transition between jawless fishes (Agnatha) and jawed animals (Gnathostomata), as such teleosts are a critical branch point in vertebrate phylogeny to investigate the genomic repertoire of Ig gene segments.

To date, in all studied vertebrate species, with the exception of birds, bats, and snakes, more than one immunoglobulin light chain isotype can be found (Lundqvist et al. 2006; Das et al. 2010; Gambón-Deza et al. 2012). The conventional classification of IgL into kappa and lambda isotypes was initially designated as a means to classify mammalian IgL (Edelman and Gally 1962). Over the years, additional isotypes have been described which deviate from the traditional kappa and lambda model in a variety other vertebrate groups (Pilström 2002). The emerging classification system currently differentiates between four ancestral clans: kappa (κ/elasmobranch type III/NS4/Teleost L1, L3, F, G/Xenopus r), lambda (λ/ elasmobranch type II), sigma (σ/teleost L2/elasmobranch type IV), and sigma cart (σ-cart).

Criscitiello and Flajnik (2007) have proposed an IgL classification system based on criteria of sequence homology, and the spacing of heptamer and nonamers of recombination signal sequences (RSS). In addition, the genomic configuration of IgL gene segments and the length of the complementary determining regions (CDR) of corresponding V_L gene segments support a syntenic approach to IgL classification. Several studies have also used molecular sequence markers to distinguish immunoglobulin light isotypes (Das et al. 2008; Edholm et al. 2009, 2011).

The isotypes of IgL found in teleosts can currently be classified as being either κ, λ, and σ (Edholm et al. 2009, 2011; Ghaffari and Lobb 1993, 1997; Bao et al. 2010). To date, the κ, λ, and σ isotypes found in teleosts have been found to exist on different chromosomes in a cluster assemblage (Daggfeldt et al. 1993; Bao et al. 2010; Edholm et al. 2009, 2011; Zimmerman et al. 2008, 2011). The immunoglobulin heavy chain (IgH) loci of rainbow trout (*Oncorhynchus mykiss*) (Hansen et al. 2005), zebrafish (*Danio rerio*) (Danilova et al. 2005), stickleback (*Gasterosteus aculeatus*) (Bao et al. 2010; Gambón-Deza et al. 2010), and medaka (Magadán-Mompó et al. 2011) are of (V-(D)-J-C) translocon type configuration typified in mice and human. The IgL gene segments of humans are also in a translocon type of arrangement where a number of V segments lie upstream of several J and finally one or more constant (C) region genes.

A departure from a single cluster can be found in the lambda IgL isotype of mice as the lambda IgL are situated in a two-cluster (V_2-$(J\text{-}C)_2$-V-$(J\text{-}C)_2$) configuration (Gerdes and Wabl 2002). Extrapolating from the two λ clusters in mice, it has been conventional to

broadly define a single Ig "cluster" as any number of V regions upstream of one or more (D), J, and C segments. To date, the most extensive number of IgH and IgL clusters has been found in cartilaginous fishes (sharks and rays) where upwards of several hundred (V-(D)-J-C) clusters are predicted to exist. The presence of multiple kappa clusters on one or more chromosomes (Daggfeldt et al. 1993; Bao et al. 2010; Edholm et al. 2009, 2011; Zimmerman et al. 2008; 2011) indicates that cluster duplication and expansions likely played a major role in the generation of antibody diversity in teleost fishes. In this study, we have combined a suite of bioinformatics-based approaches coupled with expressed sequence tag (EST) data to annotate and fit VJ-C transcripts to concordant genomic regions. Collectively, these analyses reveal three distinct IgL cluster assemblages in the medaka genome, all of which can be classified as being of the kappa isotype. This annotation should prove useful for future efforts to understand how relative gene orders and Ig cluster configurations contribute to the functional regulation and diversification of antibody gene expression in an emerging teleost model for comparative immunology.

Material and methods

Identification of the IGL locus

Genome builds of *Oryzias latipes* (assembly HdrR, October 2005; version 56.1i) available from NCBI (www.ncbi.nlm.nih.gov) (Wheeler et al. 2001) and Ensembl databases (http://www.ensembl.org/index.htlm) (Stabenau et al. 2004) were examined to locate antibody light chain genes. Previously published sequences from IgL of other teleost fishes (Haire et al. 2000; Edholm et al. 2009, 2011; Bao et al. 2010) were used as queries in BLAST alignments to identify genomic scaffolds and chromosomes containing immunoglobulin genes. These sequences (Scaffolds 379, 157, 1550, 92, 355, chromosome 11) were downloaded and annotated using the Vector-NTI (Invitrogen, available in www.invitrogen.com) (Lu and Moriyama 2004).

Identification of exons coding for C_L domains were discerned by aligning genomic sequences with previously published immunoglobulin mRNAs (Edholm et al. 2009; Zimmerman et al. 2008; Bao et al. 2010). Boundaries of exons were deduced using the software packages FGENESH (www.softberry.com) (Solovyev et al. 2006) and Augustus (http://augustus.gobics.de/submission) (Stanke et al. 2004).V_L and J_L segments of medaka were identified by several criteria including the presence of canonical (allowing one or two nucleotide mismatches) RSS, by the presence of AG/GT splice sides flanking open reading frames, and pattern searches for RSS with 23 or 12 bp spacers flanking the 3′ or 5′ends of gene segments. Exon boundaries were further deduced by using alignments of *O. latipes* EST sequences (retrieved from NCBI and http://www.shigen.nig.ac.jp/medaka) (Sasado et al. 2010) to the resultant annotated scaffolds.

IGL loci functionality

Identified immunoglobulin constant (C_L) exons from medaka and other fishes were used in iterative alignments for homologous sequences in the medaka ESTs database (http://www.shigen.nig.ac.jp/medaka) (Sasado et al. 2010). A total of 11 cDNA libraries generated from different tissues of the HdrR inbred medaka strain were scanned for putative IgL

transcripts (Supplementary File 1). In order to delineate concordant IGL loci to each EST, alignments were performed using the Lastz program (http://main.g2.bx.psu.edu/) (Goecks et al. 2010). Resultant sequence alignment hits were further characterized using the Tablet—Next Generation Sequence Assembly Visualization software (http://bioinf.scri.ac.uk/tablet/) (Milne et al. 2010). EST clones were assigned to concordant C_L if a threshold nucleotide sequence identity above 99 % was met.

Junctional diversity and somatic hypermutation analyses

Junctional diversity and tests for somatic hypermutation were carried out for ESTs whose constant regions were of 100 % identity with germline C_L. This stringent fitting criterion was employed due to the potential existence of additional C_L present within possible genomic gaps. ESTs with productive open reading frames were aligned to germline IgL using ClustalW to identify regions of gene expression. Alignments of medaka EST sequences were further refined using the IMGT/V-QUEST tool (Giudicelli et al. 2004) available in the IMGT database (the international ImMuno-GeneTics information system®) (http://imgt.cines.fr) (Lefranc 2011).

Numbers and positions of base pair differences between genomic sequences and concordant ESTs were determined for each V_L segment, including the FR1, FR2, FR3, CDR1, and CDR2 regions. The CDR3 was not included in hypermutation analyses due to unknown variability introduced during V_L and J_L joining. To assess if antigen selection pressure might be acting on medaka IgL, the multinomial distribution model proposed by Lossos et al. (2000; available at http://stat.stanford.edu/immunoglobulin) was used to discern the probability that the number of replacement mutations were not due to chance. The level of significance for probable selection was set at a threshold of $P \leq 0.05$.

Phylogenetic studies

Comparative phylogenetic studies were carried out with using the program MEGA5 (Kumar et al. 2008). ClustalW and MUSCLE alignments were used in the neighbor-joining and minimum evolution models to plot the phylogenetic trees (pairwise deletion, Jones–Taylor–Thornton matrix) and enter range activated sites (gamma number 2.5). The veracity of these trees was evaluated using the above-mentioned method and by executing 1,000 replicate bootstrapping events.

Results

Medaka IGL genomic organization

A total of 137 IgL gene segments were found to be localized to a single medaka chromosome (chromosome 11). IgL were confined to five different genomic scaffolds (designated 379, 157, 1550, 92, and 355). Within these clusters, three distinct IgL regions were localized named IGL1, IGL2, and IGL3 (Fig. 1 and Supplementary File 2). In each of the three IgL, regions clusters of V_L, J_L, and C_L segments are found (Fig. 2 and Supplementary File 2). Collectively, 63 V_L, 42 J_L, and 32 C_L medaka gene segments were identified.

Medaka IgL are of the kappa (κ) isotype

Sequence comparisons of identified medaka IgL segments with those of other teleosts revealed medaka IgL to be most similar to the kappa (κ) isotype (Fig. 3 and Supplementary File 3). Similar to other teleosts (Daggfeldt et al. 1993; Ghaffari and Lobb 1993, 1997; Bao et al. 2010; Edholm et al. 2009, 2011; Hikima et al. 2011; Zimmerman et al. 2008, 2011), the medaka V_L segments are positioned in opposite transcriptional polarity to J_L and C_L segments (Fig. 2 and Supplementary File 2). In addition, the RSS flanking each medaka V_L present a 12-bp spacer consistent with classification as the κ isotype.

The medaka C_L sequences branch into three groups that correspond to each of the three IGL1, IGL2, and IGL3 genomic regions (Fig. 4). Although medaka C_L segments belong to the κ isotype, their sequences harbor distinct differences from one another. For example, C_L members of the IGL1 can be differentiated by the lack of Trp at position 51, while IGL2 C_L segments harbor specific triplets, Leu-Asp-Ser at 85–86–87 positions and Gly-Ser-Gln at 94–95–96 positions (Fig. 4). The degree of amino acid similarity between the C_L segments from the three genomic regions was also determined. We found that the C_L segments belonging to IGL1 and IGL3 are the most similar (79–89 % identity), whereas, IGL2 members share 67– 73 % and 62–69 % amino acid identity with IGL1 and IGL3 members, respectively.

Comparisons of the medaka V_L revealed four distinct families of V_L gene segments (Fig. 5). The medaka V_L family 1 comprises the highest number of V_L and includes all V_L from the IGL1 region and the majority of V_L segments in the genomic IGL3 region. Families 2 and 3 are IGL2 locus specific, and finally, four V_L segments belonging to the IGL3 locus are clustered into the fourth family designation. These findings support the idea that the three IgL regions were generated by duplication and each evolved independently.

IGL loci functionality

In total, 227 medaka EST sequences were identified from the NBRP medaka database (www.shigen.nig.ac.jp/medaka) (Sasado et al. 2010). Alignment of in-frame ESTs to concordant germ line sequences revealed that all IGL1, IGL2, and IGL3 gene regions in medaka are expressed. In total, 15 ESTs were found to be concordant to IGL1, 72 to IGL2, and 43 to IGL3. More than 60 % of these ESTs were found to lack a V_L segment while leader sequences and $V_LJ_LC_L$ segments were present in 29 of the ESTs (Supplementary Files 4 and 5). The identity of the EST sequences with germline V_L and C_L segments is 94–100 % for the IGL2 regions and 100 % for the IGL3 regions suggesting that assignment of EST sequences to concordant germline is an appropriate method by which to ascertain that all three of the IGL regions in medaka are expressed.

Junctional diversity and somatic hypermutation

Alignments of medaka ESTs to concordant germline segments revealed no apparent N or P nucleotide addition, however, in all ESTs deletions of several bases from the ends of the V_L and or J_L could be detected (Supplementary File 6). Although it is not clear if such deletions are due to imprecision in rag-mediated double strand breakage or possible exonuclease activity prior to joining, these deletions suggest that nucleotide elimination contributes to

genetic diversity in the medaka immunoglobulin repertoire. Alignments of ESTs to concordant germline sequences revealed several mutations concentrated on V_L segments (Fig. 5). Overall, the mutational frequency was found to be highest in CDR1 and CDR2 gene regions when compared to those corresponding to framework (FR) regions (Fig. 6). Theoretical probabilities of an excess or scarcity of mutations occurring in CDR or FR regions occurring by chance (Table 1) revealed higher numbers of replacement mutations in CDR regions. The presence of a greater number of replacements over substitution mutations would appear to indicate antigen selection of variants with improved binding properties (Sablitzky et al. 1985).

Discussion

The medaka (*O. latipes*) represents an important model vertebrate organism in many fields of biology including developmental biology, genetics, evolution, toxicology, and immunology (Takeda 2008; Takeda and Shimada 2010; Pham et al. 2011). In the present study, we identified three IgL kappa regions which are localized to a single medaka chromosome (chromosome 11). In addition, through alignments of ESTs to concordant germline sequences, we have shown each of the three identified IgL regions to be functionally expressed. The genomic configuration of medaka IgL was also found to be in contrast to that of other teleosts (Partula et al. 1996; Haire et al. 2000; Bengtén et al. 2006; Edholm et al. 2011) in which multiple IgL isotypes have been found to be partitioned over multiple chromosomes.

The IgL kappa regions in medaka are arranged in a compact multi-cluster similar to that found in cartilaginous and teleost fishes (Edholm et al. 2011). Phylogenetic similarities indicate the medaka IgL may have been generated by cluster duplications with each resultant IgL region evolving rather independently. The presence of V_L segments from two different IgL regions in a same family however suggests a plausible situation of convergent evolution wherein retained IgL might be explained by a complex model of birth-and-death evolution (Nei and Rooney 2005). It is possible that the four medaka V_L families were represented early in evolutionary history and once the duplication of ancestral locus occurred, some of the V_L families disappeared from the duplicated loci. It seems that a mechanism for genetic material exchange between IgL regions is unlikely as the constant regions of individual IGL clusters grouped by IgL region. It seems improbable that genetic exchange between IGL regions would include only the V_L segments while excluding the C_L exons.

The genomic organization of medaka IgL was found to harbor clusters of V_L segments in opposite transcriptional polarity to J_L and C_L segments. This configuration implies that primary gene rearrangements would be generated by inversion. An inversional rearrangement preserves any intervening V_L and J_L gene segments that are between the V_L and J_L segments to be joined. Configurations of V_L and J_L in opposite transcriptional orientations are thought to maximize secondary rearrangements at IgL loci and serve as an important mechanism for receptor editing (continued gene rearrangement to replace self-reactive B cells during B cell development) (Hsu and Criscitiello 2006).

The presence of multiple V_L-J_L-C_L clusters in medaka also suggests that rearrangements may involve V_L joining from one cluster to another cluster similar to the rearrangement patterns described in zebrafish IgL (Zimmerman et al. 2008). Mechanisms to generate IgL appear to have undergone major transitions during the evolutionary history of vertebrates. Differences in genomic configuration facilitate several different strategies that affect to the size of the genomic repertoire and the expression process (Danilova and Amemiya 2009). The presence of multiple V_L, J_L, and C_L segments allows for increased combinatorial diversity which can be enhanced through nucleotide deletion and/or nucleotide addition at the junctions at which IgL segments are joined (Benedict et al. 2000; Market and Papavasiliou 2003). Our analyses of ESTs revealed that junctional diversity can originate by nucleotide deletions in medaka. This is reminiscent of the homology-directed V(D)J recombination process observed in newborn mice and humans, which occurs at sites of short sequence homology wherein no N nucleotides have been also been observed (Feeney 1992; Bauer et al. 2007). The ESTs encoding medaka $V_LJ_LC_L$ regions were also found to lack N or P nucleotides. It remains to be seen however if N and P addition can also play a factor in contributing to IgL diversity in this species.

The analysis of cDNA libraries revealed that all loci are functional; however, many of the ESTs lacked a corresponding V_L segment. This phenomenon has also been described in zebrafish (Haire et al. 2000) and Atlantic cod (Daggfeldt et al. 1993; Hsu and Criscitiello 2006) and it may be related to a low efficiency in mechanisms which eliminate aberrant immunoglobulin light chain transcripts (Chemin et al. 2010). Nevertheless, the medaka JCκ transcripts may be translated using the alternative translational initiation codon CUG (Gerashchenko et al. 2010) and could originate a functional surrogate JCκ protein. A surrogate JCκ protein might play a role in B cell development during the expression of a pre-B cell receptor (pre-BCR) at the pre-B cell stage (Zhang et al. 2004). This pre-BCR consists of a homodimer of μ heavy chains associated with surrogate light chains and the transmembrane signal molecules (Igλ and Igβ). In human and mouse, the surrogate light chain is composed of VpreB (homologous to V λ) and lambda 5 (λ5, homologous to Jλ-Cλ genes) (Mårtensson et al. 2007). Different VpreB gene lineages (VpreB1, VpreB2, and VpreB3) have been identified (Mårtensson et al. 2007; Wang et al. 2012); however, only VpreB3 counterparts appear to be conserved during the evolution of non-mammalian vertebrates (Rosnet et al. 2004). To date, homologues of the VpreB1, VpreB2, and λ5 genes have only been found in eutherian mammals suggesting that other evolutionary lineages may use an alternative pre-BCR form that lacks this VpreB/λ5 surrogate light chain. In humans, germline Vκ and JCκ transcripts encoding proteins that functionally substitute for VpreB and λ5 components have been described (Francés et al. 1994; Rangel et al. 2005). Therefore, it seems plausible that the JCκ transcripts found in medaka may play an important role in the development and maturation of B cells.

Following B cell development and antigen stimulation, it is well established that rearranged immunoglobulin genes can be further modified by somatic hypermutation (SHM). Somatic hypermutation can facilitate an increased affinity of antibodies to antigen during the progression of adaptive immune responses (Teng and Papavasiliou 2007). Although teleosts appear to lack the typical germinal centers for B cell selection and expansion found in mammalian models, emerging experimental evidence suggests that AID-induced

mutagenesis and antigenic selection is active in teleosts (Yang et al. 2006; Barreto and Magor 2011; Marianes and Zimmerman 2011). The presence of three distinct expressed IGL regions each exhibiting evidence of SHM positions the medaka as a valuable research model for understanding antibody diversification in an organism with considerable importance in the fields of developmental biology, toxicology, immunology, and comparative evolution.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

References

Bao Y, Wang T, Guo Y, Zhao Z, Li N, Zhao Y, et al. The immunoglobulin gene loci in the teleost *Gasterosteus aculeatus* . Fish Shellfish Immunol. 2010; 28(1):40–48. [PubMed: 19782140]

Barreto VM, Magor BG. Activation-induced cytidine deaminase structure and functions: a species comparative view. Dev Comp Immunol. 2011; 35(9):991–1007. [PubMed: 21349283]

Bauer K, Hummel M, Berek C, Paar C, Rosenberger C, Kerzel S, Versmold H, Zemlin M. Homology-directed recombination in IgH variable region genes from human neonates, infants and adults: implications for junctional diversity. Mol Immunol. 2007; 44(11):2969–2977. [PubMed: 17292963]

Belov K, Zenger KR, Hellman L, Cooper DW. Echidna IgA supports mammalian unity and traditional Therian relationship. Mamm Genome. 2002; 13(11):656–663. [PubMed: 12461652]

Benedict CL, Gilfillan S, Thai TH, Kearney JF. Terminal deoxynucleotidyl transferase and repertoire development. Immunol Rev. 2000; 175:150–157. [PubMed: 10933600]

Bengtén E, Clem LW, Miller NW, Warr GW, Wilson M. Channel catfish immunoglobulins: repertoire and expression. Dev Comp Immunol. 2006; 30(1–2):77–92. [PubMed: 16153707]

Chemin G, Tinguely A, Sirac C, Lechouane F, Duchez S, Cogne M, Delpy L. Multiple RNA surveillance mechanisms cooperate to reduce the amount of nonfunctional Ig kappa transcripts. J Immunol. 2010; 184(9):5009–5017. [PubMed: 20357261]

Criscitiello MF, Flajnik MF. Four primordial immunoglobulin light chain isotypes, including lambda and kappa, identified in the most primitive living jawed vertebrates. Eur J Immunol. 2007; 37(10): 2683–2694. [PubMed: 17899545]

Daggfeldt A, Bengtén E, Pilström L. 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 cDNAs and hybridization analysis. Immunogenetics. 1993; 38(3):199–209. [PubMed: 8505063]

Danilova N, Bussmann J, Jekosch K, Steiner LA. The immunoglobulin heavy-chain locus in zebrafish: identification and expression of a previously unknown isotype, immunoglobulin Z. Nat Immunol. 2005; 6(3):295–302. [PubMed: 15685175]

Danilova N, Amemiya CT. Going adaptive: the saga of antibodies. Ann N Y Acad Sci. 2009; 1168:130–155. [PubMed: 19566706]

Das S, Mohamedy U, Hirano M, Nei M, Nikolaidis N. Analysis of the immunoglobulin light chain genes in zebra finch: evolutionary implications. Mol Biol Evol. 2010; 27(1):113–120. [PubMed: 19744999]

Das S, Nikolaidis N, Klein J, Nei M. Evolutionary redefinition of immunoglobulin light chain isotypes in tetrapods using molecular markers. Proc Natl Acad Sci U S A. 2008; 105(43):16647–16652. [PubMed: 18940927]

Edelman G, Gally J. The nature of Bence-Jones proteins. Chemical similarities to polypeptide chains of myeloma globulins and normal gammaglobulins. Med Aug. 1962; 1116(1):207–227.

Edholm ES, Wilson M, Bengten E. Immunoglobulin light (IgL) chains in ectothermic vertebrates. Dev Comp Immunol. 2011; 35(9):906–915. [PubMed: 21256861]

Edholm ES, Wilson M, Sahoo M, Miller NW, Pilström L, Wermenstam NE, et al. Identification of Igsigma and Iglambda in channel catfish, *Ictalurus punctatus*, and Iglambda in Atlantic cod, *Gadus morhua* . Immunogenetics. 2009; 61(5):353–370. [PubMed: 19333591]

Feeney AJ. Predominance of VH-D-JH junctions occurring at sites of short sequence homology results in limited junctional diversity in neonatal antibodies. J Immunol. 1992; 149:222–229. [PubMed: 1607655]

Flajnik MF. Comparative analyses of immunoglobulin genes: surprises and portents. Nat Rev Immunol. 2002; 2(9):688–698. [PubMed: 12209137]

Francés V, et al. A surrogate 15 kDa JC kappa protein is expressed in combination with mu heavy chain by human B cell precursors. EMBO J. 1994; 24:5937–5943. [PubMed: 7813432]

Gambón-Deza F, Sánchez-Espinel C, Magadán-Mompó S. Presence of an unique IgT on the IGH locus in three-spined stickleback fish (*Gasterosteus aculeatus*) and the very recent generation of a repertoire of VH genes. Dev Comp Immunol. 2010; 34(2):114–122. [PubMed: 19733587]

Gambón-Deza F, Sánchez-Espinel C, Mirete-Bachiller S, Magadán-Mompó S. Snakes antibodies. Dev Comp Immunol. 2012; 38:1–9. [PubMed: 22426516]

Gerashchenko V, Su D, Gladyshev N. CUG start codon generates thioredoxin/glutathione reductase isoforms in mouse testes. J Biol Chem. 2010; 285(7):4595–602. [PubMed: 20018845]

Gerdes T, Wabl M. Physical map of the mouse lambda light chain and related loci. Immunogenetics. 2002; 54(1):62–5. [PubMed: 11976793]

Ghaffari SH, Lobb CJ. Structure and genomic organization of immunoglobulin light chain in the channel catfish. An unusual genomic organizational pattern of segmental genes. J Immunol. 1993; 151(12):6900–6912. [PubMed: 8258698]

Ghaffari SH, Lobb CJ. Structure and genomic organization of a second class of immunoglobulin light chain genes in the channel catfish. J Immunol. 1997; 159(1):250–258. [PubMed: 9200461]

Giudicelli V, Chaume D, Lefranc MP. IMGT/V-QUEST, an integrated software program for immunoglobulin and T cell receptor V-J and V-D-J rearrangement analysis. Nucleic Acids Res. 2004; 32:W435–W440. Web Server issue. [PubMed: 15215425]

Goecks J, Nekrutenko A, Taylor J, Team G. Galaxy: a comprehensive approach for supporting accessible, reproducible, and transparent computational research in the life sciences. Genome Biol. 2010; 11(8):R86. [PubMed: 20738864]

Haire RN, Rast JP, Litman RT, Litman GW. Characterization of three isotypes of immunoglobulin light chains and T-cell antigen receptor alpha in zebrafish. Immunogenetics. 2000; 51(11):915–923. [PubMed: 11003385]

Hansen JD, Landis ED, Phillips RB. Discovery of a unique Ig heavy-chain isotype (IgT) in rainbow trout: implications for a distinctive B cell developmental pathway in teleost fish. Proc Natl Acad Sci USA. 2005; 102(19):6919–6924. [PubMed: 15863615]

Hikima JI, Jung TS, Aoki T. Immunoglobulin genes and their transcriptional control in teleosts. Dev Comp Immunol. 2011; 35(9):924–936. [PubMed: 21078341]

Hsu E, Criscitiello MF. Diverse immunoglobulin light chain organizations in fish retain potential to revise B cell receptor specificities. J Immunol. 2006; 177(4):2452–2462. [PubMed: 16888007]

Klein J. Are invertebrates capable of anticipatory immune responses? Scand J Immunol. 1989; 29(5): 499–505. [PubMed: 2658012]

Kumar S, Nei M, Dudley J, Tamura K. MEGA: a biologist-centric software for evolutionary analysis of DNA and protein sequences. Brief Bioinform. 2008; 9(4):299–306. [PubMed: 18417537]

Lefranc MP. IMGT, the International ImMunoGeneTics Information System. Cold Spring Harb protocol. 2011; 6:595–603.

Litman GW, Anderson MK, Rast JP. Evolution of antigen binding receptors. Annu Rev Immunol. 1999; 17:109–117. [PubMed: 10358755]

Lossos IS, Tibshirani R, Narasimhan B, Levy R. The inference of antigen selection on Ig genes. J Immunol. 2000; 165(9):5122–5126. [PubMed: 11046043]

Lu G, Moriyama E. Vector NTI, a balanced all-in-one sequence analysis suite. Brief Bioinform. 2004; 5(4):378–388. [PubMed: 15606974]

Lundqvist ML, Middleton DL, Radford C, Warr GW, Magor KE. Immunoglobulins of the non-galliform birds: antibody expression and repertoire in the duck. Dev Comp Immunol. 2006; 30(1–2):93–100. [PubMed: 16150486]

Magadán-Mompó S, Sáanchez-Espinel C, Gambón-Deza F. Immunoglobulin heavy chains in medaka (*Oryzias latipes*). BMC Evol Biol. 2011; 11:165. [PubMed: 21676244]

Marchalonis JJ, Schluter SF. On the relevance of invertebrate recognition and defence mechanisms to the emergence of the immune response of vertebrates. Scand J Immunol. 1990; 32(1):13–20. [PubMed: 2382105]

Marchalonis JJ, Kaveri S, Lacroix-Desmazes S, Kazatchkine MD. Natural recognition repertoire and the evolutionary emergence of the combinatorial immune system. FASEB J. 2002; 16(8):842–848. [PubMed: 12039866]

Marianes AE, Zimmerman AM. Targets of somatic hypermutation within immunoglobulin light chain genes in zebrafish. Immunology. 2011; 132(2):240–255. [PubMed: 21070232]

Market E, Papavasiliou FN. V(D)J recombination and the evolution of the adaptive immune system. PLoS Biol. 2003; 1(1):E16. [PubMed: 14551913]

Mårtensson L, Keenan A, Licence S. The pre-B-cell receptor. Curr Opin Immunol. 2007; 2:137–142. [PubMed: 17306522]

Milne I, Bayer M, Cardle L, Shaw P, Stephen G, Wright F, et al. Tablet—next generation sequence assembly visualization. Bioinformatics. 2010; 26(3):401–402. [PubMed: 19965881]

Nei M, Rooney AP. Concerted and birth-and-death evolution of multigene families. Annu Rev Genet. 2005; 39:121–152. [PubMed: 16285855]

Partula S, Schwager J, Timmusk S, Pilström L, Charlemagne J. A second immunoglobulin light chain isotype in the rainbow trout. Immunogenetics. 1996; 45(1):44–51. [PubMed: 8881036]

Pham CH, Park KS, Kim BC, Kim HN, Gu MB. Construction and characterization of Japanese medaka (*Oryzias latipes*) hepatic cDNA library and its implementation to biomarker screening in aquatic toxicology. Aquat Toxicol. 2011; 105(3–4):569–575. [PubMed: 21937010]

Pilström L. The mysterious immunoglobulin light chain. Dev Comp Immunol. 2002; 26(2):207–215. [PubMed: 11696386]

Rangel R, et al. Assembly of the kappa preB receptor requires a V kappa-like protein encoded by a germline transcript. J Biol Chem. 2005; 18:17807–17814. [PubMed: 15757909]

Rast JP, Smith LC, Loza-Coll M, Hibino T, Litman GW. Genomic insights into the immune system of the sea urchin. Science. 2006; 314(5801):952–956. [PubMed: 17095692]

Rosnet O, Blanco-Betancourt C, Grivel K, Richter K, Schiff C. Binding of free immunoglobulin light chains to VpreB3 inhibits their maturation and secretion in chicken B cells. J Biol Chem. 2004; 11:10228–10236. [PubMed: 14670953]

Sablitzky F, Wildner G, Rajewsky K. Somatic mutation and clonal expansion of B cells in an antigen-driven immune response. EMBO J. 1985; 4(2):345–3450. [PubMed: 3926481]

Sasado T, Tanaka M, Kobayashi K, Sato T, Sakaizumi M, Naruse K. The National BioResource Project Medaka (NBRP Medaka): an integrated bioresource for biological and biomedical sciences. Exp Anim. 2010; 59(1):13–23. [PubMed: 20224166]

Smith LC, Davidson EH. The echinoid immune system and the phylogenetic occurrence of immune mechanisms in deuterostomes. Immunol Today. 1992; 13(9):356–362. [PubMed: 1466753]

Solovyev V, Kosarev P, Seledsov I, Vorobyev D. Automatic annotation of eukaryotic genes, pseudogenes promoters. Genome Biol. 2006; 7(Suppl1):S10. [PubMed: 16925832]

Stabenau A, McVicker G, Melsopp C, Proctor G, Clamp M, Birney E. The Ensembl core software libraries. Genome Res. 2004; 14(5):929–933. [PubMed: 15123588]

Stanke M, Steinkamp R, Waack S, Morgenstern B. AUGUSTUS: a web server for gene finding in eukaryotes. Nucleic Acids Res. 2004; 32:W309–W312. Web Server issue. [PubMed: 15215400]

Takeda H. Draft genome of the medaka fish: a comprehensive resource for medaka developmental genetics and vertebrate evolutionary biology. Dev Growth Differ. 2008; 50(Suppl 1):S157–S166. [PubMed: 18430160]

Takeda H, Shimada A. The art of medaka genetics and genomics: what makes them so unique? Annu Rev Genet. 2010; 44:217–241. [PubMed: 20731603]

Teng G, Papavasiliou FN. Immunoglobulin somatic hypermutation. Annu Rev Genet. 2007; 41:107–120. [PubMed: 17576170]

Wang X, Parra E, Miller D. A VpreB3 homologue in a marsupial, the gray short-tailed opossum, *Monodelphis domestica* . Immunogenetics. 2012; 8:647–652. [PubMed: 22684248]

Wheeler DL, Church DM, Lash AE, Leipe DD, Madden TL, Pontius JU, Schuler GD, Schriml LM, Tatusova TA, Wagner L, Rapp BA. Database resources of the National Center for Biotechnology Information. Nucleic Acids Res. 2001; 29(1):11–16. [PubMed: 11125038]

Yang F, Waldbieser GC, Lobb CJ. The nucleotide targets of somatic mutation and the role of selection in immunoglobulin heavy chains of a teleost fish. J Immunol. 2006; 176(3):1655–1667. [PubMed: 16424195]

Zhang M, Srivastava G, Lum L. The pre-B cell receptor and its function during B cell development. Cell Mol Immunol. 2004; 2:89–94. [PubMed: 16212894]

Zimmerman AM, Romanowski KE, Maddox BJ. Targeted annotation of immunoglobulin light chain (IgL) genes in zebrafish from BAC clones reveals kappa-like recombining/deleting elements within IgL constant regions. Fish Shellfish Immunol. 2011; 31(5):697–703. [PubMed: 20933599]

Zimmerman AM, Yeo G, Howe K, Maddox BJ, Steiner LA. Immunoglobulin light chain (IgL) genes in zebrafish: Genomic configurations and inversional rearrangements between (V(L)-J (L)-C(L)) gene clusters. Dev Comp Immunol. 2008; 32(4):421–434. [PubMed: 18022691]

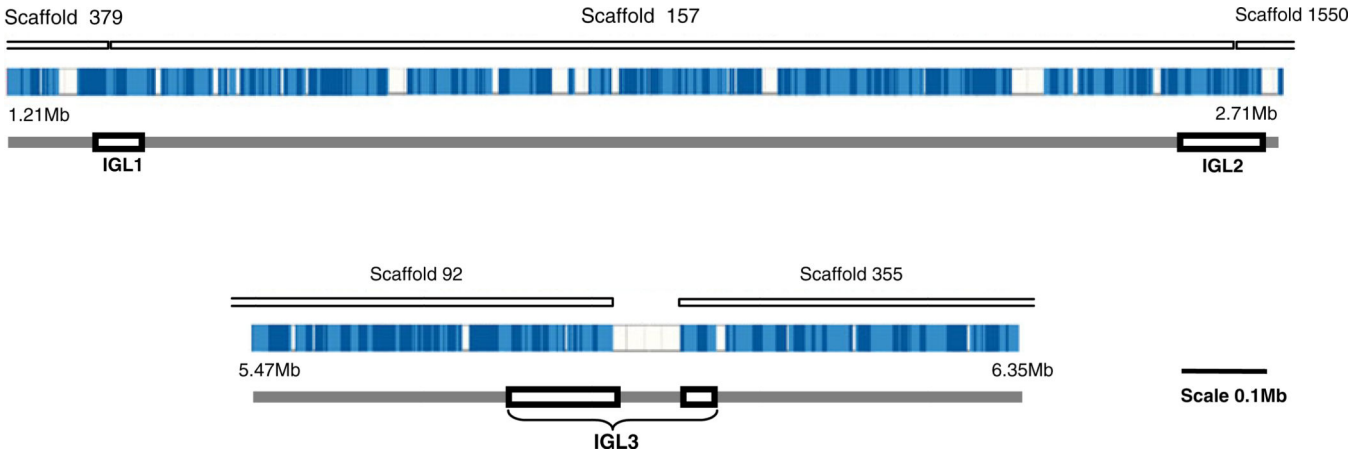

Fig. 1.

Chromosomal localization of medaka IgL loci. The figure depicts location of scaffolds encoding immunoglobulin light chain genes on chromosome 11. Three loci (IGL1, IGL2, and IGL3) can be deduced in which V_L-J_L-C_L clusters were identified. Contigs are represented in *blue* whereas *white segments* show gaps in the genomic sequence

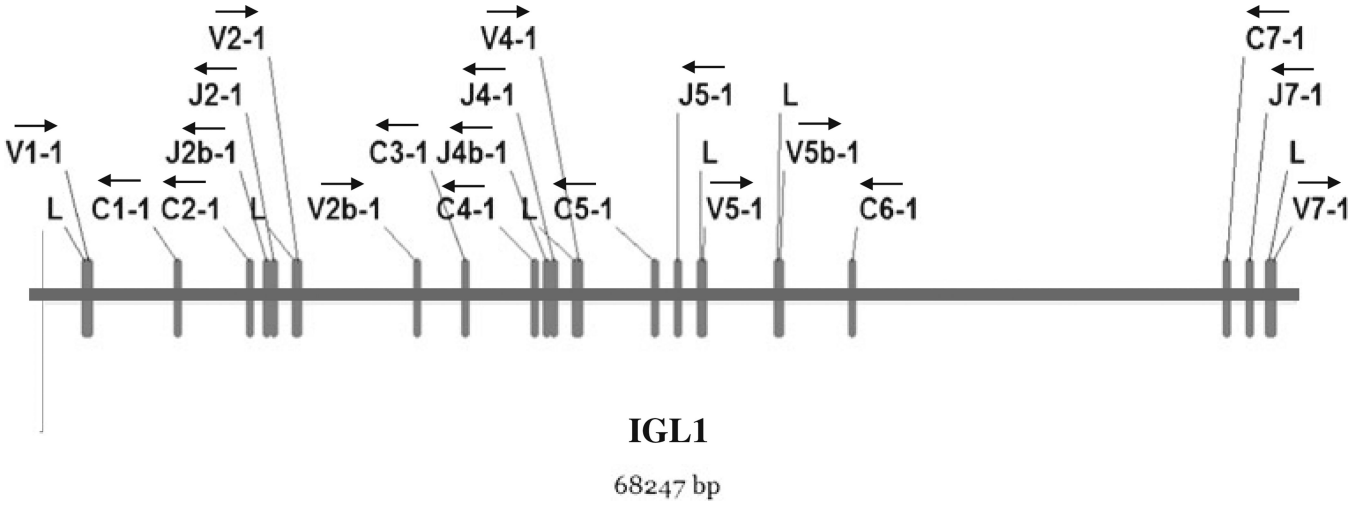

Fig. 2.

Detailed representation of $V_LJ_LC_L$ clusters in the medaka IGL1 locus. Leader, V_L, and J_L segments and constant (C) exons are depicted as *rectangles*. Delineation of exons was discerned using the computer software (FGNESH and Augustus) and alignment with available ESTs. *Arrows* indicate transcriptional polarity of V_L segments, J_L segments, and C_L exons

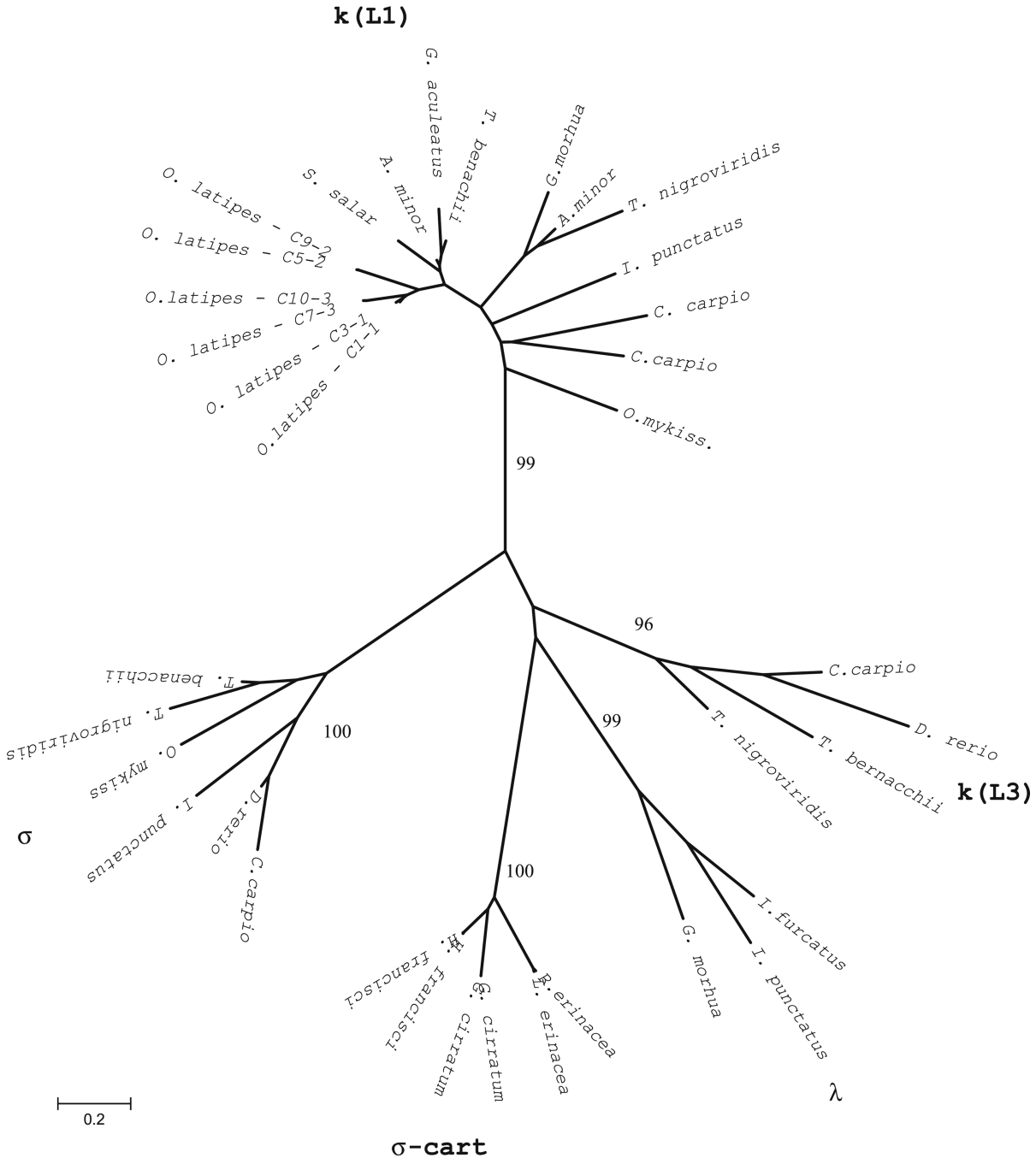

Fig. 3.

Phylogenetic analysis of medaka C_L genes. The unrooted tree was constructed using two representative amino acid C_L sequences from each medaka IGL locus and other identified fish C_L sequences. The phylogenetic tree was constructed using the pairwise deletion algorithm, JTT matrix and differences by sites activated with gamma parameter 2.5. Sequences from other teleosts include GenBank accession numbers: kappa/Ll (*G. aculeatus* AY278356, *T. bernacchii* DQ842627, *A. minor* AF137397 and AF137398, *S. salar* AF406958, *T. nigroviridis* AJ575806, *G morhua* X68514, *I. punctatus* L25533, *C. carpio*

AB015905 and X65260, *O. mykiss* AB035729), kappa/L3 (*C. carpio* AB035730, *D. rerio* AF246193, *T. bernacchii* DQ842626, *T. nigroviridis* U25705), lambda (*I. furcatus* CK403484; *I. punctatus* EU872022, *G. morhua* AJ293807), sigma (*C. carpio* AB091118, *D. rerio* AF246162, *I. punctatus* EU872021, *T. nigroviridis* AJ575637; *T. bernacchii* EF114785), sigma-cart (*G. cirratum* AAV34681 and AAV34682; N. Shark NS5; *H. francisci* XI5315 and B33937; *L. erinacea* L25568; *E erinacea* AAA87170)

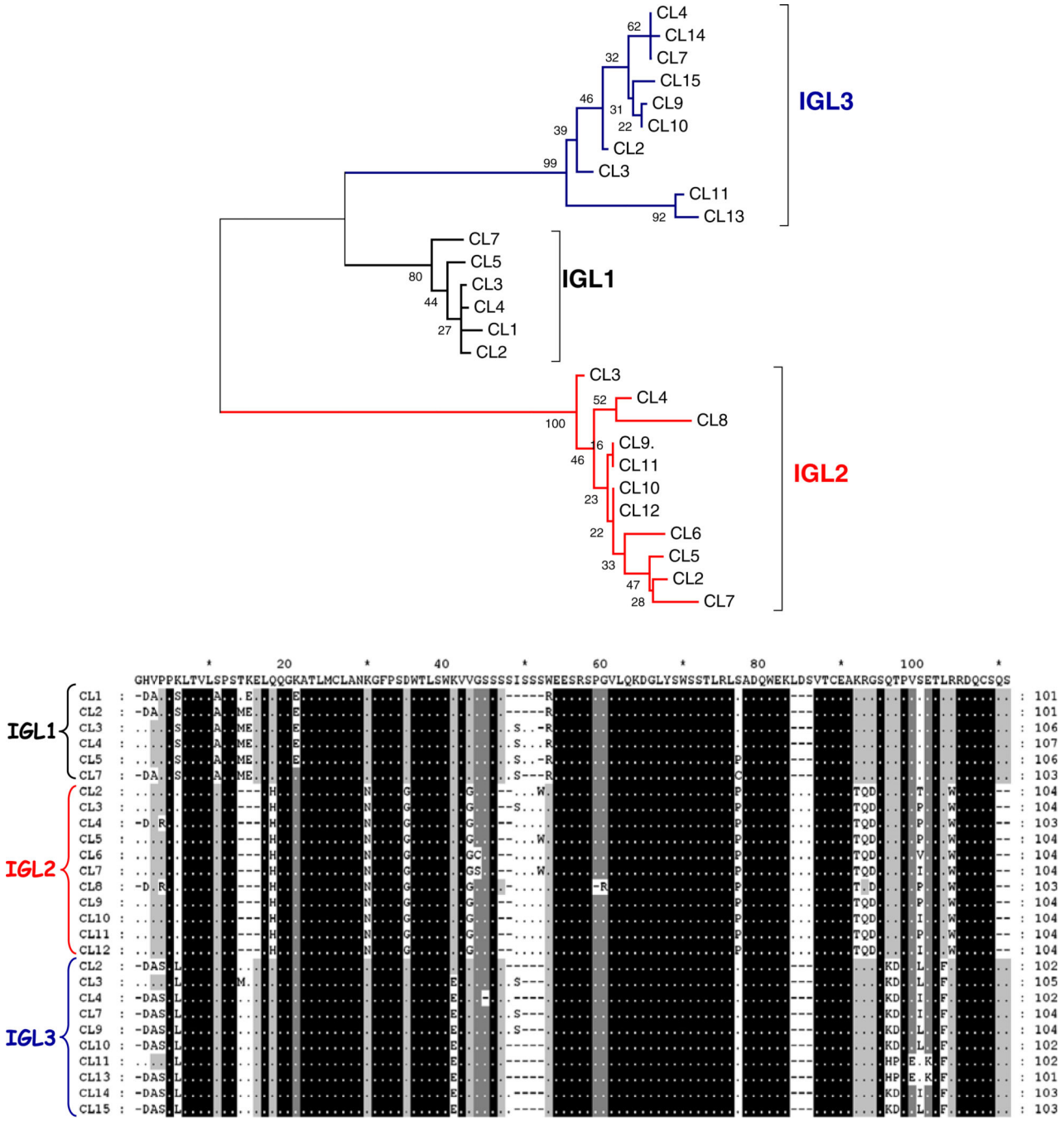

Fig. 4.

Comparisons of medaka C_L exons. **a** An unrooted tree of C_L amino acid sequences revealing C_L exons clustered into groups that correspond to each locus (IGL1, IGL2, and IGL3). The phylogenetic tree was constructed using the pairwise deletion algorithm, JTT matrix, and differences by sites activated with gamma parameter 2.5. **b** Different medaka C_L segments harbor regions of homology which could be utilized as molecular markers. *Arrows* indicate regions conserved in members from different C_L groups. Among polymorphisms (Ser at position 14; Gln or Glu at 17; a gap at position 20; Phe at 32; Pro at position 34; Asp, Glu,

Lys or Arg at position 60; Thr at 65; His at 91 and Phe at 102 position) that have also been utilized to identify the tetrapod C kappa chains (Das et al. 2008), only the Phe at position 32 appears conserved in medaka sequences

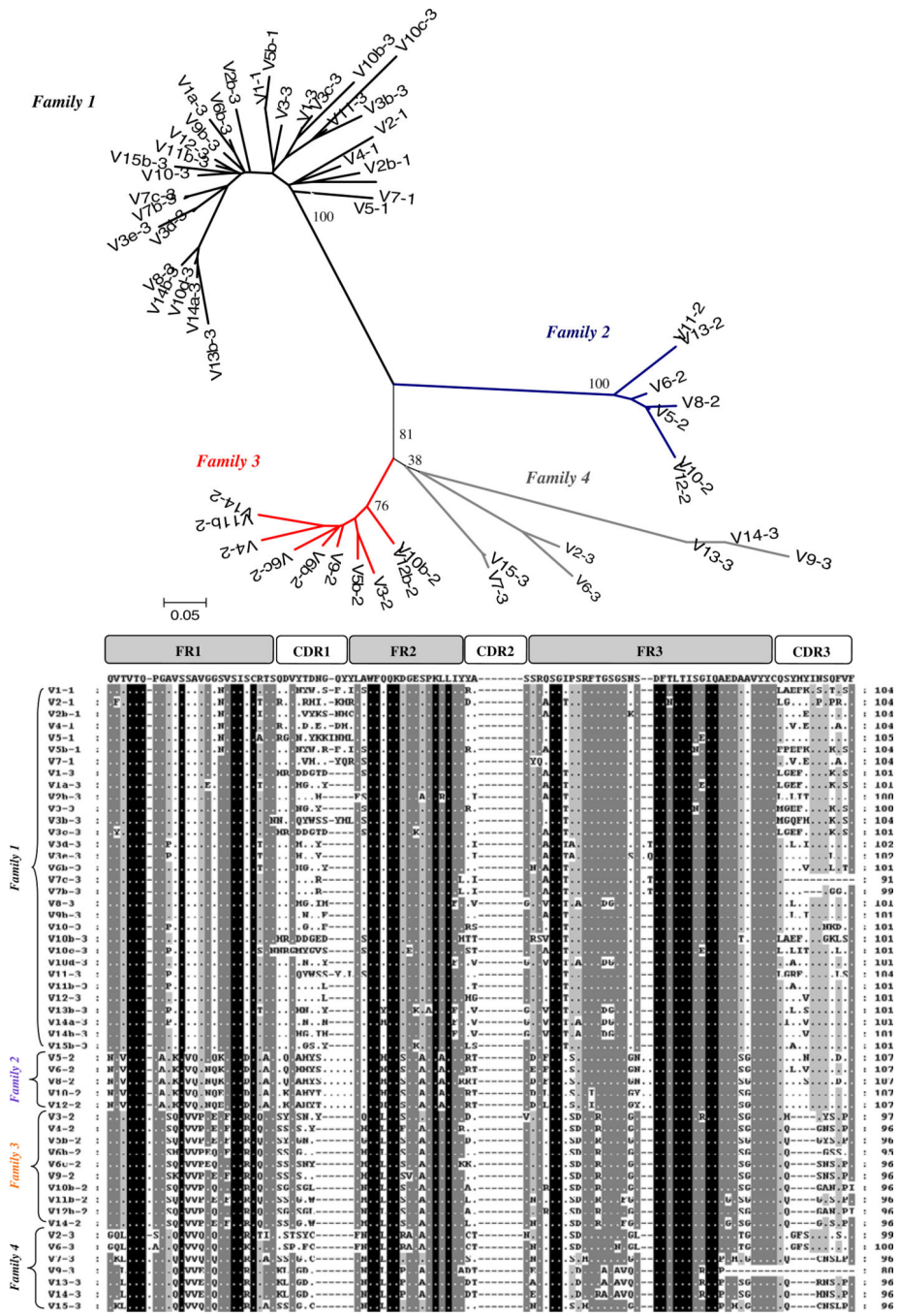

Fig. 5.

Comparisons of medaka genomic V_L segments. The upper por-tion of this diagram depicts an unrooted phylogenetic tree based on amino acid sequences. Both the phylogenetic tree and corresponding alignment (lower portion of diagram agree with a classification of medaka V_L into four different families). Alignments were performed and scored based on recommendations of the IMGT (http://imgt.cines.fr)

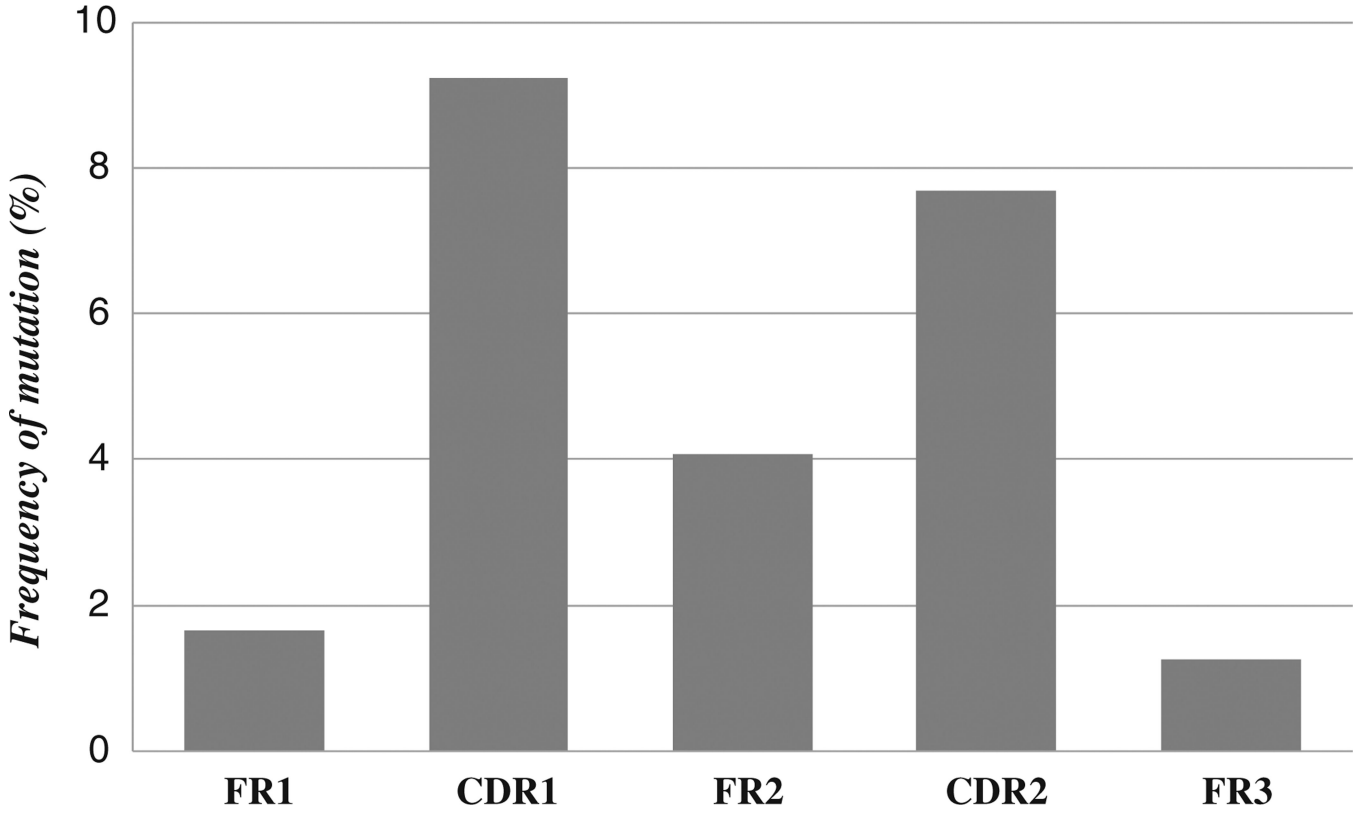

Fig. 6.

Somatic mutation in medaka V_L- The percentage of mutations in framework (FR) and complementary determining regions (CDR) from 13 representative ESTs are shown. The CDR3 was excluded from the analysis

Table 1

Analysis of mutations in VL segments of medaka immunoglobulin light chain coding ESTs

EST	Germline VL	Mutations			FR	FR-P_M[b]	Mutations		CDR	CDR-P_M[b]
		FR1	FR2	FR3	R:S[a]		CDR1	CDR2	R:S[a]	
OLOVA49021	V3-2	0	3	1	3:1	0.115	2	1	1:2	0.299
OLKI56H12	V5-2	3	0	1	2:2	0.000	5	0	4:1	0.001
OLSP40B17	V4-2	0	2	2	1:3	0.003	1	2	1:2	0.260
OLGI48J17	V4-2	1	7	2	5:5	0.002	2	0	2:0	0.145
OLSP42K18	V6b-2	4	3	1	2:6	0.000	5	4	4:5	0.019
OLKI20G22	V10-2	0	1	0	1:0	0.081	3	0	0:3	0.699
OLKI23P18	V10-2	0	2	0	2:0	0.164	3	0	0:3	0.735
OLKI35K02	V10-2	0	1	0	1:0	0.081	3	0	0:3	0.699
OLSP13L18	V7-3	2	0	0	1:1	0.055	2	0	1:1	0.141
OLKI53D04	V10d-3	2	5	5	3:9	0.003	0	0	0:0	0.860
OLKI61N09	V11-3	4	2	2	3:5	0.045	1	0	1:0	0.512
OLKI33H22	V12-3	0	1	2	1:2	0.001	4	2	5:1	0.001
OLKI30J18	V13-3	0	0	2	2:0	0.774	0	0	0:0	0.576

[a] Number of replacement (R) mutations vs. silent (S) mutations found in each VL subregion. The CDR3 and FR4 were not included in the analysis of mutations

[b] P_M value, probability calculated by multinomial distribution model that excess (for CDR) or scarcity (for FR) of mutations occurred by chance. The level of significance was $P \leq 0.05$