

Genomic cloning and sequence analysis of Interleukin-10 from *Labeo rohita*

Sweta Karan¹, Himani Kaushik¹, Nipun Saini¹, Pramod K Sahoo², Aparna Dixit³ & Lalit C Garg^{1*}

Received September 26, 2014; Accepted October 01, 2014; Published October 30, 2014

¹Gene Regulation Laboratory, National Institute of Immunology, Aruna Asaf Ali Marg, New Delhi - 110 067, India; ²Fish Health Management Division, Central Institute of Freshwater Aquaculture, Kausalyaganga, Bhubaneswar - 751 002, India; ³School of Biotechnology, Jawaharlal Nehru University, New Delhi - 110 067, India; Lalit C Garg - Email: lalit@nii.ac.in; lalitcarg@gmail.com; Phone: +91 11 26703652; Fax: +91 11 26742125; *Corresponding author

Abstract:

Interleukin-10 (IL-10) is a pleiotropic cytokine and plays an important role in inflammation, immunoregulation and the pathogenesis of various diseases. Therefore, it is our interest to isolate, clone, sequence and characterize IL-10 gene from the fish *Labeo rohita* (*Lr*). The gene was amplified using genomic DNA isolated from head kidney with primers designed on conserved sequence homologues of fishes belonging to Cyprinidae family. The gLrIL-10 is 1467 nucleotides long with five exons and four introns sharing the same organization as of mammalian IL-10 genes. An open reading frame of 537 bp was found to encode a putative 179 amino acid protein with a signal peptide of 22 amino acids with conserved signature sequence motif. Sequence analysis showed similarity with the IL-10 from most fresh water fishes of Cyprinidae family. LrIL-10 has 27.2 % identity and 54.95 % similarity with the human IL-10. Sequence analysis followed by phylogenetic studies showed highest identity with *Catla catla* (98%) followed by *Cyprinus carpio* (93%), *Hypophthalmichthys molitrix* (89%) and is distantly related to human, rhesus monkey and frog. These data from primary sequence characterization may be used to further understand transcriptional regulation and functional characterization of LrIL-10 in relation to species-specific molecular immunology.

Keywords: Interleukin-10, *Labeo rohita*, cloning.

Abbreviations: IL-10, Interleukin-10, *Lr*, *Labeo rohita*; nt, nucleotides.

Background:

Interleukin-10 (IL-10) was earlier described as a multifunctional Type II/ T helper 2 cytokine [1]. Subsequently, it has been shown to be broadly expressed cytokine produced by many cells of innate and adaptive immune system including, activated T cells, B cells, keratinocytes, mast cells, monocytes/macrophages, NK cells, neutrophils, DCs, eosinophils, mesangial cells, epithelial cells, tumor cells and other cell types [2, 3, 4]. This complex cytokine controls inflammatory responses and primarily inhibit the synthesis of a number of cytokines, including interferon γ , tumor necrosis factor, IL-1, IL-2, IL-3, IL-6, IL-8 and IL-12 [5] and hence was initially named as Cytokine synthesis inhibitory factor [6]. It is a potent inhibitor of antigen presentation as they are responsible for inhibiting major

histocompatibility complex class II expression and also inhibit the differentiation and maturation of DCs [7]. Apart from its potent anti-inflammatory role, this interleukin also has immunostimulatory activity toward the growth and differentiation of activated B cells, cytokine activated T cells, mast cells and potentiates cytokine production by natural killer cells [8, 9]. IL-10 has important therapeutic applications for the treatment of acute and chronic inflammation, autoimmunity, cancer and infectious diseases [10].

The induction of the anti-inflammatory response is mediated through the IL-10 receptor (IL-10R) and activation of signal transducer and activator of transcription 3 (STAT3) [11, 12]. The transcription of this gene has been studied in various cells. IL-10

gene expression is regulated by constitutively and ubiquitously expressed SP1 and SP3 as key transcription factors, hence the transcription of IL-10 continues even when other cytokine genes are inactive [13]. Many transcription factors regulate the expression of IL-10 both in antigen presenting cells and in CD4+T cells. Transcriptions of IL-10 from different species have been reported to be differentially regulated under different conditions and a number of signalling molecules have been reported to regulate IL-10 regulation [4]. The IL-10 gene has been extensively studied from a number of mammalian species and is reported to have five exons and four introns and several 'ATTTA'- mRNA instability motifs at 3'- untranslated regions [14]. The gene encoding human IL-10 is located on chromosome 1 and a large number of single nucleotide polymorphisms (SNPs) have been identified in the IL-10 gene promoter. The gene encodes a 178 amino acids long protein, which is secreted after cleavage of signal peptide comprising of 18 amino acids [15]. IL-10 is a key immunoregulatory cytokine during infection and henceforth, it is required for optimal pathogen clearance and ameliorates immunopathology. Therapeutic strategies can augment IL-10 to reduce host injury during infection. In this study, interleukin-10 gene was cloned and genomic sequence was characterized from *Labeo rohita*, an economically important major fresh water carp of Indian subcontinent. The results from this study will help to understand the biological functions and structure-function relation of the IL-10 of *L. rohita*.

GenBank Acc. No AY887900.1), and *Hypophthalmichthys molitrix* (Silver carp, GenBank Acc. No. DQ058296.1) were taken from nucleotide sequence database (<http://www.ncbi.nlm.nih.gov/nucleotide>) [16]. Identity and similarity percentage among the three species were calculated using Emboss water (http://www.ebi.ac.uk/Tools/psa/emboss_water/) [17].

Cloning of genomic LrIL-10

Genomic DNA was isolated from *L. rohita* head kidney by the method described by Upadhyaya *et al.* [18] and used as a template for amplification of the genomic IL-10 gene of *L. rohita*. Primers were designed on the basis of multiple sequence alignment using ClustalW of nucleotide sequences of three fish species namely *Cyprinus carpio* (JX524551.1), *Danio rerio* (AY887900.1), *Hypophthalmichthys molitrix* (Silver carp, DQ058296.1) The gLrIL-10 was PCR amplified using the forward primer 5'-GTTTTTCAGTGGAGTCATCCTT-3' and the reverse primer 5'-TAAGTGCTGACTCTCTTTGA-3' using Taq polymerase (Takara Bio Inc., Japan). The reaction mixture consisted of 100 ng of *L. rohita* genomic DNA, dNTP mix (0.2 mM each), 10 pmole of each primer, 1 × Taq buffer, 3 mM MgCl₂ and 5 % DMSO and 1 U of ExTaq DNA polymerase (Takara Bio Inc, Japan), in a final volume of 50 µl. Amplification was carried out for 30 cycles, each including denaturation at 95 °C for 30 sec, annealing at 52 °C for 30 sec and extension at 72 °C for 30 sec. Initial denaturation was done for 1 min and the final extension at 72 °C for 7 min. The amplified product was gel purified using Qiagen gel extraction kit as per manufacturer's instructions and ligated to pGEM-T Easy vector (Promega, USA). Competent *Escherichia coli* DH5α cells (Novagen, USA) were transformed with the ligation mix and the transformants were selected on Luria-Bertani-agar medium (Difco, USA) containing ampicillin (100 µg/ml). The positive clones were confirmed by colony PCR using the same primers and PCR conditions as described for gene amplification and release of the insert upon restriction digestion with *SalI* (New England Biolabs, USA). DNA sequence of the cloned insert was determined by automated DNA sequencing (DNA Sequencing Facility, University of Delhi South Campus, New Delhi).

Sequence analysis of genomic LrIL-10

The gLrIL-10 amino acid sequences were deduced using the ExPasy translate tool (<http://www.expasy.org>) [19]. The nucleotide sequences were analyzed by the BLAST program (www.ncbi.nlm.nih.gov/BLAST) [20]. Theoretical molecular weight and isoelectric point was analysed using Compute pI/Mw tool at the ExPasy server [21]. Signal peptide search was performed with SignalP 3.0. (www.cbs.dtu.dk/services/SignalP) [22]. Multiple sequence alignment was created with the ClustalO program (<http://www.ebi.ac.uk/Tools/msa/clustalo>) [23]. IL-10 protein sequences were obtained from Genbank. Different species selected for multiple sequence alignment with their common names and GenBank accession numbers are given in the legend to **Supplementary Figure 1** (see supplementary material).

Phylogenetic relationship analysis

The evolutionary relationship of the protein encoded by the cloned rLrIL-10 sequence with IL-10 from other species was inferred using the Neighbor-Joining method [24] after multiple sequence alignment. Phylogenetic analyses were conducted in MEGA6 with 500 replications to obtain Bootstrap values [25].

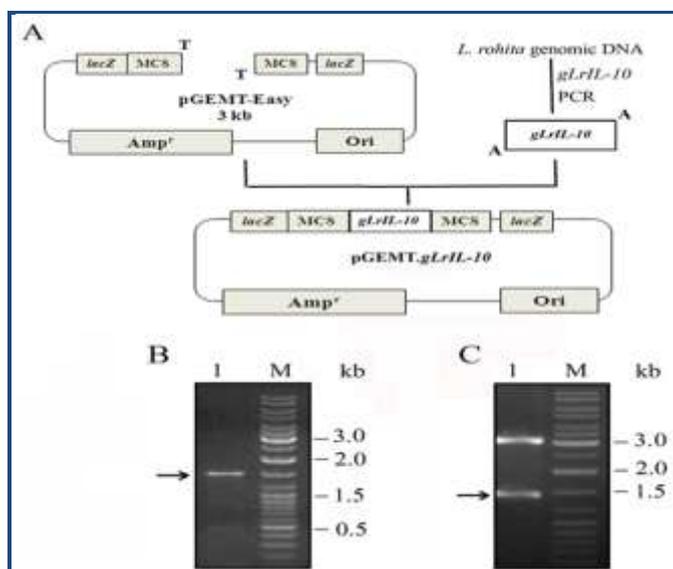


Figure 1: Cloning of the genomic IL-10 of *L. rohita*. **A)** Strategy for cloning genomic IL-10 of *L. rohita*; **B)** PCR amplification of genomic IL-10. Lane 1 shows the PCR amplified product of ~1.5 kb (indicated by arrow). "M" indicates DNA molecular weight markers; **C)** Restriction digestion of the recombinant plasmid pGEMT-gLrIL-10. Lane 1 shows the recombinant plasmid digested with *SalI*. The arrow points to the released insert of ~1.5 kb. "M" indicates DNA molecular weight markers.

Methodology:

Analysis of sequence conservation in IL-10 homologues from fish belonging to Cyprinidae family

Multiple nucleotide and amino acid sequence alignment of IL-10 from homologues of fishes belonging to *Cyprinidae* family was performed. IL-10 sequences of *Cyprinus carpio* (Common carp, GenBank Acc. No. JX524551.1), *Danio rerio* (Zebra fish,

Thirty five IL-10 amino acid sequences (corresponding to Interleukin-10 of different species) were selected. The confidence probability (multiplied by 100) that the interior branch length is greater than 0, as estimated using the bootstrap test (500) replicates is shown next to the branches [26]. The tree is drawn to scale, with branch lengths in the same units as those

of the evolutionary distances used to infer the phylogenetic tree. The evolutionary distances were computed using the Poisson correction method [27] and are in the units of the number of amino acid substitutions per site. All positions containing gaps and missing data were eliminated. There were a total of 139 positions in the final dataset.

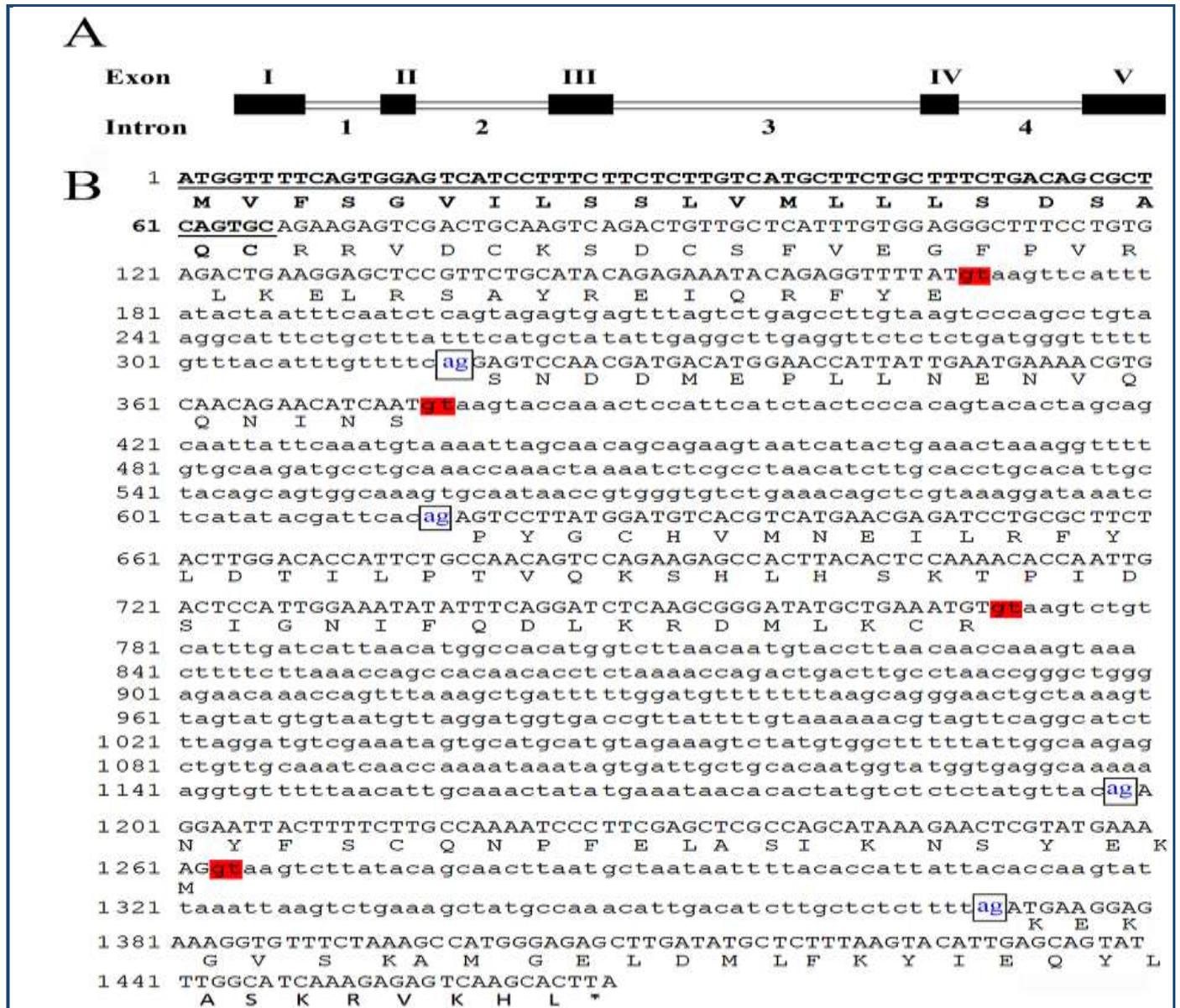


Figure 2: A) Schematic representation of genomic organization of IL-10 gene of *L. rohita*. The filled black boxes represent the five exons (I-V) whereas the double line in between numbered as 1-4 represent the 4 introns spanning 169...318, 376...617, 771...1199 and 1263...1371, respectively; B) The *L. rohita* IL-10 gene sequence and deduced amino acid sequence. Intronic regions are shown in lower case with the intron splicing consensus 5'-(gt) and 3'-(ag), shown in red boxed letters and blue fonts, respectively. The deduced amino acid sequence is given below the exons. Underlined nucleotide sequence shows the region encoding the signal peptide (amino acid residues 1-22). The stop codon is marked with an asterisk (*).

Results & Discussion:

Sequence analysis of IL-10 homologues of fishes belonging to Cyprinidae family

In order to clone the *L. rohita* genomic IL-10, primers were designed on the basis of identity and similarity shared among the nucleotide and amino acid sequences of known IL-10 from fishes of *Cyprinidae* family. Nucleotide and amino acid sequence alignment of the three selected species are shown in

Supplementary Figure 2 & Supplementary Figure 3, (see supplementary material) respectively. Nucleotides sequence analysis using Emboss water tool revealed that *H. molitrix* and *C. carpio* shared 74.4% identity while their amino acid showed 87% identity and 95% similarity. Amino acid and nucleotide sequence of *H. molitrix* and *Danio rerio* showed 50.3% and 50.2% similarity, respectively.

Cloning of genomic *LrIL-10*

Cloning strategy for cloning the genomic IL-10 from *L. rohita* is shown in **Figure 1A**. Primers designed on the basis of conserved regions of IL-10 homologues from three fishes of *Cyprinidae* family namely *C. carpio*, *H. molitrix*, and *D. rerio* were used to amplify genomic IL-10 from *Labeo rohita* head kidney. An amplified product of ~1.5 kb (**Figure 1B**) was ligated to pGEM-T Easy vector. Colony PCR using primers used for PCR amplification confirmed cloning of the PCR amplified product in to pGEM-T Easy vector. Release of ~1.5 kb from the recombinant plasmid upon digestion with *Sall* further confirmed cloning of the amplified product (**Figure 1C**). Final confirmation of the positive clones was made by automated sequencing. The nucleotide sequence of genomic IL-10 (*gLrIL-10*) gene has been deposited to GenBank database with accession number HM363517.1.

Genomic structure, nucleotide (nt) sequence and organization of the cloned *gLrIL-10* are shown in **Figure 2**. The cloned *gLrIL-10* is of 1467 nucleotide length. Analysis of the DNA sequence revealed the *gLrIL-10* to be made up of five exons and four introns with distinct well conserved intron-exon junctions observed in IL-10 homologues from other species. Schematic representation of the cloned *rLrIL-10* is shown in **Figure 2A**. The five exons (indicated as I,II, III, IV and V) are of 168 bp, 57 bp, 153 bp, 63 bp, and 96 bp, spanning nucleotide positions 1-168, 319-375, 618-770, 1200-1262 and 1372-1467, respectively (**Figure 2B**). The four introns numbered as 1 to 4 stretch between the exons and are of 150 bp, 242 bp, 429 bp and 109 bp in length, respectively (shown in lower case in **Figure 2B**). Exon I and intron 3 are the longest stretch while exon II and intron 4 are the shortest. Human genomic *IL-10* is also made up of five exons and 4 introns although introns of *gLrIL-10* are much smaller than those of human *IL-10*. Exon I is of longest stretch in both human and *gLrIL-10*, but instead of intron 3 as in *gLrIL-10*, intron 4 is longest in human *IL-10* gene (Accession no. U16720.1). Amongst *Cyprinidae* family, *IL-10* genes from *C. carpio* (Accession no. JX524551.1), *H. molitrix* [Acc. No. DQ058296.1], *C. catla* [Acc. No. 312386552] and *D. rerio* [Acc. No. AY887900.1] have also been reported with 5 exons and 4 introns. Typical intron splice motif was observed at the 5' (GT, shown in red box) and 3' (AG, shown in blue fonts) ends of each intron. The mature *gLrIL-10* encodes for 157 amino acid residues, with a theoretical molecular weight of 18.53 kDa and isoelectric point of 8.5. The first 22 amino acid residues were predicted to be a signal peptide by the SignalP program with cleavage site present between Cys22 and Arg23.

Blast P of the *gLrIL-10* amino acid sequence showed that it shared 26-98% identity with known IL-10 sequences of other species. The *gLrIL-10* amino acid sequence showed highest homology (98%) with *C. catla* IL-10, while least similarity was observed with the IL-10 of *Ochotona princeps* (American pica). Percentage identity and similarity of pairwise *LrIL-10* sequence alignment of *LrIL-10* with human is 27.2 % and 54.95 %, respectively. BlastP of the amino acid sequence of *LrIL-10* showed that among the fish species, percentage identities of the *LrIL-10* with the IL-10 of *C. catla*, *C. carpio*, *H. molitrix*, *D. rerio*, *C. auratus*, *C. idella*, *T. rubripes*, and *D. labrax*, were 98%, 93%, 89%, 80%, 87%, 88%, 47%, 46%, respectively. Despite the low amino acid identity between IL-10 from human and other mammalian species, there is a high degree conservation of residues which

may play important role for stabilization of the structural core of IL-10. Multiple sequence alignment of the *gLrIL-10* with the IL-10 of other species, shown in **Supplementary Figure 1** revealed that although there are significant differences among the IL-10 from different species, the signature sequences of the IL-10 super family remain conserved. Two typical IL-10 signature sequence motifs were identified in *gLrIL-10*, L⁴³ to L⁷³ and K¹³³ to E¹⁴⁶ (LLNENVQQNI-NSPYGCHVMNEILRFYLDITL and_KAMGELDMLFKYIE). The putative mature peptides of *LrIL-10* contain four conserved cysteine residues (Cys31, Cys80, Cys128 and Cys134), which are known to play critical role in tertiary structure [28]. The mature peptide also contains two extra cysteine residues (Cys27 and Cys32), which are significantly conserved only in *Cyprinidae* family, however their role in IL-10 structure and function is not yet investigated.

Phylogenetic analysis

The optimal tree with the sum of branch length = 5.12655948 was obtained (**Figure 3**). The phylogenetic tree was generated by multiple sequence alignment of IL-10 amino acid sequences from 35 species (**Figure 3**). The tree topology shows that *LrIL-10* is very closely related to fishes of *Cyprinidae* family i.e. *C. catla* (*catla*), *C. carpio* (common carp), *H. molitrix* (silver carp) and *C. idella* (grass carp). As evident from the tree, other vertebrates such as human (*H. sapiens*), cattle (*B. Taurus*), elephant (*E. maximus*), deer (*C. elaphus*), rhesus monkey (*M. mulatta*) and sheep (*O. aries*), and virus are distantly related to *LrIL-10*.

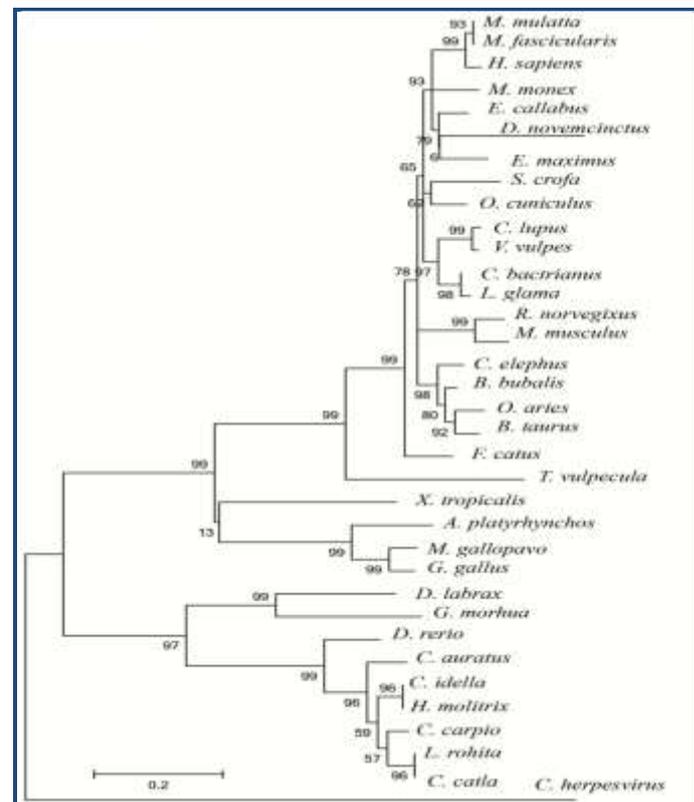


Figure 3: Rooted phylogenetic tree of the deduced amino acid sequence of *LrIL-10* of *L. rohita* and IL-10 from other species. Amino acid sequences of 35 selected species were retrieved from NCBI database. Multiple sequence alignment was performed using ClustalO. Sequence divergence is represented by the distances from the nodes i.e. the branch length (denoted by numbers).

Conclusion:

The cloning, sequencing and characterization of the genomic IL-10 gene from the fish *L. rohita* is of interest. Our analysis on its gene structure, sequence conservation and phology show that *gLrIL-10* is an IL-10 homologue. Analysis also shows that the gene contains two typical IL-10 signature sequence motifs. The gene is phylogenetically closer to IL-10 of Cyprinidae family and very distantly related to the IL-10 of human, elephant, deer and cattle. The genomic structure of IL-10 gene elucidated from *Labeo rohita* will help in understanding its function and role.

Acknowledgement:

Department of Biotechnology, Ministry of Science and Technology, New Delhi, India is acknowledged for providing financial support.

References:

- [1] Florentino DF *et al.* *J Exp Med.* 1989 **170**: 2081 [PMID: 2531194]
- [2] Moore KW *et al.* *Annu Rev Immunol.* 1993 **11**: 165 [PMID: 8386517]
- [3] Zdanov A, *Curr Pharm Des.* 2004 **10**: 3873 [PMID: 15579076]
- [4] Saraiva M & O'Garra A, *Nat Rev Immunol.* 2010 **10**: 170 [PMID: 20154735]
- [5] D'Andrea A *et al.* *J Exp Med.* 1993 **178**: 1041 [PMID: 8102388]
- [6] Moore KW *et al.* *J Immunol.* 2012 **189**: 2072 [PMID: 22908364]
- [7] Akdis CA & Blaser K, *Immunology* 2001 **103**: 131 [PMID: 11412299]
- [8] Rousset F *et al.* *Proc Natl Acad Sci USA.* 1992 **89**: 1890 [PMID: 1371884]
- [9] Carson WE *et al.* *Blood* 1995 **85**: 3577 [PMID: 7540068]
- [10] Asadullah K *et al.* *Pharmacol Rev.* 2003 **55**: 241 [PMID: 12773629]
- [11] Moore KW *et al.* *Ann Rev Immun.* 2001 **19**: 683 [PMID: 11244051]
- [12] Murray PJ, *Curr Opin Pharmacol.* 2006 **6**: 379 [PMID: 16713356]
- [13] Tone M *et al.* *J Immunol.* 2000 **165**: 286 [PMID: 10861063]
- [14] Brow CY *et al.* *J Biol chem.* 1996 **271**: 20108 [PMID: 8702732]
- [15] Sabat R *et al.* *Cytokine and Growth Factor Rev.* 2010 **21**: 331 [PMID: 21115385]
- [16] <http://www.ncbi.nlm.nih.gov/nuccore>.
- [17] http://www.ebi.ac.uk/Tools/psa/emboss_water/
- [18] Upadhyaya T *et al.* *DNA Seq.* 2007 **18**: 302 [PMID: 17541836]
- [19] <http://www.expasy.org>
- [20] www.ncbi.nlm.nih.gov/BLAST.
- [21] <http://web.expasy.org/computepep/>
- [22] www.cbs.dtu.dk/services/SignalP.
- [23] <http://www.ebi.ac.uk/Tools/msa/clustalo>.
- [24] Saitou N & Nei M, *Mol Biol Evol.* 1987 **4**: 406 [PMID: 3447015]
- [25] Tamura K *et al.* *Mol Biol Evol.* 2013 **30**: 2725 [PMID: 24132122]
- [26] Dopazo J, *J Mol Evol.* 1994 **38**: 300 [PMID: 8006997]
- [27] Zuckerkandl E & Pauling L, *J Theor Biol.* 1965 **8**: 357 [PMID: 5876245]
- [28] Walter MR & Nagabhusan TL, *Biochemistry* 1995 **34**: 12118 [PMID: 7547951]

Edited by P Kanguane

Citation: Karan *et al.* *Bioinformatics* 10(10): 623-629 (2014)

License statement: This is an open-access article, which permits unrestricted use, distribution, and reproduction in any medium, for non-commercial purposes, provided the original author and source are credited

Supplementary material:

Supplementary Figure 1: Multiple sequence alignment of amino acid sequence of IL-10 homologues from different species using CLUSTAL O (1.2.1). Signal sequence is underlined. Boxed sequences show IL-10 signature sequences. Four cysteine residues are shown. Common names and accession numbers for IL-10 homologues from the selected species are as follows: *L. rohita* (Rohu, HM363517.1), *Cyprinus carpio* (common carp, JX524551.1), *Ctenopharyngodon idella* (grass carp, HQ388294.1), *H. molitrix* (silver carp, DQ058296.1), *Danio rerio* (Zebra fish, AY887900.1), *C. catla* (Catla, HQ221997.1), *Dicentrarchus labrax* (European seabass, AM268529.1), *Gadus morhua* (Atlantic cod, EU004087.1), *Meleagris gallopavo* (turkey, XM_003212953.1), *Oryctolagus cuniculus* (rabbit, AF068058.1), *Macaca mulatta* (Rhesus monkey, NM_001044727.1), *Carassius auratus* (gold fish, HQ259106.1), *Rattus norvegicus* (Norway rat, NM_012854.2), *Felis catus* (domestic cat, AF060520.1), *Canis lupus familiaris* (dog, U33843.1), *Sus scrofa* (pig, L20001.1), *Ovis aries* (sheep, NM_001009327.1), *Bos taurus* (cattle, NM_174088.1), *Anas platyrhynchos* (Pekin duck, JN786941.1), *Gallus gallus* (chicken, AJ621254.1), *Mus musculus* (mouse, NM_010548.2), *Xenopus tropicalis* (Western clawed frog, NM_001171929.1), *Camelus bactrianus* (camel, AB246674.1), *Cervus elaphus* (deer, U11767.1), *Bubalus bubalis* (buffalo, AB246351.1), *Vulpes vulpes* (red fox, AJ621190.1), *Elephas maximus* (elephant, KJ572794.1), *Equus caballus* (horse, U38200.1), *Macaca fascicularis* (crab eating macaque, AB000514.1), *Marmota monax* (woodchuck, AF012909.1), *Trichosurus vulpecula* (common brushtail, AF026277.1), *Dasyurus novemcinctus* (armadillo, EF551061.1), *Lama glama* (llama, AB107649.1), *Homo sapiens* (human, NM_000572.2), *Cyprinid herpesvirus 3* (herpes virus, HM467225.1). "*" indicates single fully conserved residue, ":" indicates conservation of strong groups while "." indicates Conservation of weak groups. The "-" indicates no consensus".

	1	2	
<i>L. rohita</i>	<u>MVPSGVILSSIVMMLI</u> —SDSAQCRRVDCKSDCCSFVEGFPVRLKELRSATREIQRFYESNDD—MEP	LLMEVQQNI—NSPYGCHVMNEILRFYLDITLPT—VQKS—HL—HSKTPIDS	111
<i>D. rerio</i>	<u>MTPEGVILSALVTLII</u> —CICAQSRVECKTDCCSFVEGFPPLRLRLRSAYKEIQRFYESNDD—LEP	LLNEDIKHNI—NSPYGCHVMNEILHFYLETILPTALQKN—PLKHSSTPIDS	113
<i>C. auratus</i>	<u>MVPTGVILSALVWFLI</u> —SDSAQCRRVNCCKSECCSFVEGFPVRLKELRSATREIQRFYESNDD—LAP	LLDEVQQHI—NSPYGCHVMNEILRFYLDITLPTAVQKD—HL—HSKTPINS	112
<i>C. idella</i>	<u>MTPSRVVPSAIVMMLI</u> —SESAQCRRVDCKSECCSFVEGFPVRLKELRSATREIQRFYESNDD—LEP	LLMEVQQNI—NSPYGCHVMNEILRFYLETILPTAVQKN—HL—HPKTPIDS	112
<i>H. molitrix</i>	<u>MTPSRVVPSAIVMMLI</u> —SESAQCRRVDCKSECCSFVEGFPVRLKELRSATREIQRFYESNDD—LEP	LLMEVQQNI—NSPYGCHVMNEILRFYLETILPTAVQKN—HL—HPKTPIDS	112
<i>C. carpio</i>	<u>MVPSGVILSALVWFLI</u> —SDSAQCRRVDCKTDCCSFVEGFPVRLKELRPAITREIQRFYESNDD—MEP	LLDEVQQNI—NSPYGCHVMNEILRFYLDITLPTAVQKD—HL—HSKTPINS	112
<i>C. catla</i>	<u>MTPTGVILSSIVMMLI</u> —SDSAQCRRVDCKSDCCSFVEGFPVRLKELRSATREIQRFYESNDD—MEP	LLMEVQQNI—NSPYGCHVMNEILRFYLDITLPTAVQKS—HL—HSKTPIDS	112
<i>D. labrax</i>	<u>MTPSRILLSTIVVLSF</u> —FCVVMCS—PMCHWQCCRFVEGFPGLRQLRADPTEIQDPTAEAND—LDA	LLDQTVEDTL—KTPFACHAINSLIFPYLSTVLP—TAMAGVTEDTKDLKPHMES	115
<i>G. morhua</i>	--- <u>MWTLFLLVYVLEI</u> —CDVMACT—PTCHWQCCRFVEGSPVRLRRLRQDPAVIRDYEAEDD—LEI	LLDQSLIAHSL—RSPFGCHAMKNIIDPYLKTVLP—TAIADMTTEENKVKPHIES	112
<i>C. Herpesvirus</i>	— <u>MLAVLILPATIYF</u> —----FARGAPATPKDSCVYLIGQTPQLRQLMAVQAIIGADGSGVDEDDM	PIYPSDVMNELASTSVACDAIKKVLTMNI—GILP—NVVTAAY—---PDKKSEVDE	109
<i>T. vulpecula</i>	— <u>MTMMLRCLLQVLT</u> —SS—---NLSSALEDNCKTFTSTLPHMLRELRAAFSVKTYPPQTRK—LET	LIDKSLLEHL—KSYLGCQALSEMIKFLYEEVMP—PRAEENE—---LDVKEVDS	107
<i>D. novemcinctus</i>	— <u>MTSAILCCIVVFLA</u> —GVAASRHSQSTLSEDDCTHFPASLPHMLRELRAAFSVKTYPPQMKDQ—LDN	LLTRSLLEDF—KGYLGCQALSEMIQFYLEEVM—PQAEVHS—---PDIKQVSS	111
<i>R. norvegicus</i>	— <u>MTSAILCCIVVFLA</u> —GVAASRHSQSTLSEDDCTHFPASLPHMLRELRAAFSVKTYPPQMKDQ—LDN	LLTRSLLEDF—KGYLGCQALSEMIKFLYEEVMP—PQAEVHS—---PDIKEHLS	111
<i>M. musculus</i>	— <u>MTSAILCCIVVFLA</u> —GVAASRHSQSTLSEDDCTHFPASLPHMLRELRAAFSVKTYPPQMKDQ—LDN	LLTRSLLEDF—KGYLGCQALSEMIQFYLEEVM—PQAEVHS—---PDIKEHLS	111
<i>M. monax</i>	— <u>MTSAILCCIVVFLA</u> —GVAASRHSQSTLSEDDCTHFPASLPHMLRELRAAFSVKTYPPQMKDQ—LDN	LLTRSLLEDF—KGYLGCQALSEMIQFYLEEVM—PQAEVHS—---PDIKEHLS	111
<i>S. scrofa</i>	— <u>MTSAILCCIVVFLA</u> —GVAASRHSQSTLSEDDCTHFPASLPHMLRELRAAFSVKTYPPQMKDQ—LDN	LLTRSLLEDF—KGYLGCQALSEMIQFYLEEVM—PQAEVHS—---PDIKEHLS	107
<i>O. cuniculus</i>	— <u>MTSAILCCIVVFLA</u> —GVAASRHSQSTLSEDDCTHFPASLPHMLRELRAAFSVKTYPPQMKDQ—LDN	LLTRSLLEDF—KGYLGCQALSEMIQFYLEEVM—PQAEVHS—---PDIKEHLS	111
<i>F. catus</i>	— <u>MTSAILCCIVVFLA</u> —GVAASRHSQSTLSEDDCTHFPASLPHMLRELRAAFSVKTYPPQMKDQ—LDN	LLTRSLLEDF—KGYLGCQALSEMIQFYLEEVM—PQAEVHS—---PDIKEHLS	111
<i>C. lupus</i>	— <u>MTSAILCCIVVFLA</u> —GVAASRHSQSTLSEDDCTHFPASLPHMLRELRAAFSVKTYPPQMKDQ—LDN	LLTRSLLEDF—KGYLGCQALSEMIQFYLEEVM—PQAEVHS—---PDIKEHLS	112
<i>V. vulpes</i>	— <u>MTSAILCCIVVFLA</u> —GVAASRHSQSTLSEDDCTHFPASLPHMLRELRAAFSVKTYPPQMKDQ—LDN	LLTRSLLEDF—KGYLGCQALSEMIQFYLEEVM—PQAEVHS—---PDIKEHLS	112
<i>C. elaphus</i>	— <u>MTSAILCCIVVFLA</u> —GVAASRHSQSTLSEDDCTHFPASLPHMLRELRAAFSVKTYPPQMKDQ—LDN	LLTRSLLEDF—KGYLGCQALSEMIQFYLEEVM—PQAEVHS—---PDIKEHLS	112
<i>O. aries</i>	— <u>MTSAILCCIVVFLA</u> —GVAASRHSQSTLSEDDCTHFPASLPHMLRELRAAFSVKTYPPQMKDQ—LDN	LLTRSLLEDF—KGYLGCQALSEMIQFYLEEVM—PQAEVHS—---PDIKEHLS	112
<i>B. Taurus</i>	— <u>MTSAILCCIVVFLA</u> —GVAASRHSQSTLSEDDCTHFPASLPHMLRELRAAFSVKTYPPQMKDQ—LDN	LLTRSLLEDF—KGYLGCQALSEMIQFYLEEVM—PQAEVHS—---PDIKEHLS	111
<i>B. bubalis</i>	— <u>MTSAILCCIVVFLA</u> —GVAASRHSQSTLSEDDCTHFPASLPHMLRELRAAFSVKTYPPQMKDQ—LDN	LLTRSLLEDF—KGYLGCQALSEMIQFYLEEVM—PQAEVHS—---PDIKEHLS	112
<i>E. maximus</i>	— <u>MTSAILCCIVVFLA</u> —GVAASRHSQSTLSEDDCTHFPASLPHMLRELRAAFSVKTYPPQMKDQ—LDN	LLTRSLLEDF—KGYLGCQALSEMIQFYLEEVM—PQAEVHS—---PDIKEHLS	111
<i>H. sapiens</i>	— <u>MTSAILCCIVVFLA</u> —GVAASRHSQSTLSEDDCTHFPASLPHMLRELRAAFSVKTYPPQMKDQ—LDN	LLTRSLLEDF—KGYLGCQALSEMIQFYLEEVM—PQAEVHS—---PDIKEHLS	111
<i>M. mulatta</i>	— <u>MTSAILCCIVVFLA</u> —GVAASRHSQSTLSEDDCTHFPASLPHMLRELRAAFSVKTYPPQMKDQ—LDN	LLTRSLLEDF—KGYLGCQALSEMIQFYLEEVM—PQAEVHS—---PDIKEHLS	111
<i>M. fascicularis</i>	— <u>MTSAILCCIVVFLA</u> —GVAASRHSQSTLSEDDCTHFPASLPHMLRELRAAFSVKTYPPQMKDQ—LDN	LLTRSLLEDF—KGYLGCQALSEMIQFYLEEVM—PQAEVHS—---PDIKEHLS	111
<i>E. caballus</i>	— <u>MTSAILCCIVVFLA</u> —GVAASRHSQSTLSEDDCTHFPASLPHMLRELRAAFSVKTYPPQMKDQ—LDN	LLTRSLLEDF—KGYLGCQALSEMIQFYLEEVM—PQAEVHS—---PDIKEHLS	111
<i>C. bactrianus</i>	— <u>MTSAILCCIVVFLA</u> —GVAASRHSQSTLSEDDCTHFPASLPHMLRELRAAFSVKTYPPQMKDQ—LDN	LLTRSLLEDF—KGYLGCQALSEMIQFYLEEVM—PQAEVHS—---PDIKEHLS	111
<i>L. glama</i>	— <u>MTSAILCCIVVFLA</u> —GVAASRHSQSTLSEDDCTHFPASLPHMLRELRAAFSVKTYPPQMKDQ—LDN	LLTRSLLEDF—KGYLGCQALSEMIQFYLEEVM—PQAEVHS—---PDIKEHLS	111
<i>X. tropicalis</i>	--- <u>MKPTLMLVTPPTKRVKVK</u> QSGDAEGNCSRVVWIFPAKLKELRATPQIKMFPQMKDMALEIV	LLQDQLLQHF—KGYLGCQVSEITIRFYLEEVLPQANH—-----YKMNVSF	106
<i>A. platyrhynchos</i>	--- <u>MTGCVVAMMLLAASTYPAKCL</u> LLTDPSCVHLSDLLPAKLKELRATPQIKMFPQMKDMALEIV	LLSSDLEDF—KGYLGCQVSEITIRFYLEEVLPQANH—-----YKMNVSF	110
<i>M. gallopavo</i>	--- <u>MTGCVVAMMLLAASTYPAKCL</u> LLTDPSCVHLSDLLPAKLKELRATPQIKMFPQMKDMALEIV	LLSSDLEDF—KGYLGCQVSEITIRFYLEEVLPQANH—-----YKMNVSF	108
<i>G. gallus</i>	--- <u>MTGCVVAMMLLAASTYPAKCL</u> LLTDPSCVHLSDLLPAKLKELRATPQIKMFPQMKDMALEIV	LLSSDLEDF—KGYLGCQVSEITIRFYLEEVLPQANH—-----YKMNVSF	108

