Developmental and Comparative Immunology 34 (2010) 579–589

Contents lists available at ScienceDirect

Developmental and Comparative Immunology

journal homepage: www.elsevier.com/locate/dci

Phylogeny, genomic organization and expression of λ and κ immunoglobulin light chain genes in a reptile, *Anolis carolinensis*☆

Qian Wu [a,1], Zhiguo Wei [a,1], Zhi Yang [a], Tao Wang [a], Liming Ren [a], Xiaoxiang Hu [a], Qingyong Meng [a], Ying Guo [a], Qinghong Zhu [a], Jacques Robert [b], Lennart Hammarström [c], Ning Li [a,**], Yaofeng Zhao [a,*]

[a] State Key Laboratory of Agrobiotechnology, College of Biological Sciences, China Agricultural University, Beijing 100094, PR China
[b] Department of Microbiology & Immunology, Box 672, University of Rochester, Medical Center, Rochester, NY 14642, USA
[c] Division of Clinical Immunology, Department of Laboratory Medicine, Karolinska University Hospital Huddinge, SE-141 86, Stockholm, Sweden

ARTICLE INFO

Article history:
Received 22 December 2008
Received in revised form 23 December 2009
Accepted 30 December 2009
Available online 20 January 2010

Keywords:
Antibody
Evolution
Immunoglobulin light chain
Reptile

ABSTRACT

The reptiles are the last major taxon of jawed vertebrates in which immunoglobulin light chain isotypes have not been well characterized. Using the recently released genome sequencing data, we show in this study that the reptile *Anolis carolinensis* expresses both λ and κ light chain genes. The genomic organization of both gene loci is structurally similar to their respective counterparts in mammals. The identified λ locus contains three constant region genes each preceded by a joining gene segment, and a total of 37 variable gene segments. In contrast, the κ locus contains only a single constant region gene, and two joining gene segments with a single family of 14 variable gene segments located upstream. Analysis of junctions of the recombined VJ transcripts reveals a paucity of N and P nucleotides in both expressed λ and κ sequences. These results help us to understand the generation of the immunoglobulin repertoire in reptiles and immunoglobulin evolution in vertebrates.

1. Introduction

Immunoglobulin (Ig) light (L) chains contribute considerably to the combinatorial antibody diversity through pairing with the heavy chains [1]. Two types of IgL chains, λ and κ, are expressed in mammals. The λ gene locus often contains several pairs of J_λ–C_λ (J: Joining, C: Constant region) with multiple V (variable) genes located upstream, whereas the κ locus contains only a single C gene at the 3′ end and a few Jκ and multiple Vκ gene segments upstream [2]. The two L chains are expressed at different ratios in different species [2], and both chains can be diversified through imprecise VJ recombination and somatic hypermutation (SHM) [3].

Although the IgL isotypes ρ and type III, corresponding to the mammalian κ and λ respectively, have previously been identified in amphibians, the most primitive tetrapods [4–6], birds possess only λ chains [7]. The diversification of L chain (as well as the heavy chain) in birds is mainly dependent on gene conversion, that utilizes upstream pseudo-V genes as donors to modify the most 3′ functional V gene, which is the only one that is subject to VJ recombination [8,9]. Surprisingly, little is known yet at the genomic level regarding the number and IgL gene isotypes in reptiles, a taxon phylogenetically close to birds.

L chains in teleost fish are divided into three types (Type 1–3) [10], where Type 1 and Type 3 have recently been proposed to be κ orthologs [11]. Type 2 is most closely related to *Xenopus* σ, and is thought to have no orthologs in higher vertebrates [11,12]. A λ ortholog was just recently identified in bony fish [13]. The genomic organization of L chain genes has recently been thoroughly characterized in zebrafish [12,14], and found to differ slightly from those in other teleost, and cartilaginous fish [2,15], where the L genes are organized in multiple, independently rearranging miniloci containing few gene segments (one to three V segments, one or two J and one C gene) [2,15–18]. In Type 2 and Type 3 loci of the zebrafish IgL genes, there are evidently more V segments for each C gene.

Three types of L chain genes, Type I (NS5), Type II (NS3) and Type III (NS4) have previously been defined in cartilaginous fish such as skates and sharks [19–21]. Recently, Flajnik and colleagues identified a fourth type of IgL gene in sharks, which is thought to be orthologous to the *Xenopus* L chain σ [11]. Their phylogenetic analysis also suggests that the Type I gene in cartilaginous fish is closely related to, but distinct from, the σ type (thus regarded as σ-cart), and establishes orthology between the cartilaginous Type II and λ, Type III and κ [11].

☆ The sequences reported in this study have been deposited in the NCBI GenBank with following accession numbers: EU407219–EU407221, EU419646–EU419702, GU338709–GU338771.

* Corresponding author. Tel.: +86 10 62734945; fax: +86 10 62733904.
** Corresponding author. Tel.: +86 10 62731146; fax: +86 10 62733904.
E-mail addresses: ninglcau@cau.edu.cn (N. Li), yaofengzhao@cau.edu.cn (Y. Zhao).
[1] These authors contributed equally to this work.

0145-305X/$ – see front matter © 2010 Elsevier Ltd. All rights reserved.
doi:10.1016/j.dci.2009.12.019

In all jawed vertebrates, V(D)J recombination is a site-specific somatic process involving RAG1 and RAG2 that introduce DNA double-strand break (DSB) at the border of the coding segments and recombination signal sequences (RSS) [22]. Although the re-ligation of the signal (non-coding) ends is precise, the joining of the coding ends is carried out by non-homologous end joining (NHEJ) pathway with removal or addition of a random number of nucleotides [23]. The added nucleotides are classified into P (palindromic) and N (added by TdT) nucleotides according to their origin. This imprecise re-joining of the coding ends amplifies the V(D)J combinatorial diversity. In some cases, however, the re-joining process utilizes short homologous stretches (microhomology) shared by two coding ends, leading to generation of coding junctions that lack P and N nucleotides [24–26].

IgL isotypes and their modes of diversification have been extensively investigated in many jawed vertebrates including cartilaginous and bony fish, amphibians, birds and mammals, but not reptiles [2,15]. In this study, we report the analysis of IgL chain isotypes in the green anole lizard (*Anolis carolinensis*) and its evolutionary implications (while our paper was prepared, another group also reported the presence of both λ and κ genes in the green anole lizard [27]).

2. Materials and methods

2.1. Animals, RNA isolations and reverse transcriptions

About 3-year-old green anole lizards were purchased from a local Beijing pet market. Total RNA was prepared using a TRNzol kit (TIANGEN BIOTECH, Beijing). Reverse-transcription was conducted using M-MLV reverse transcriptase following the manufacture's instructions (Invitrogen, Beijing).

2.2. Cloning of the expressed lizard Igλ, Igκ light chain genes at the cDNA level

Expressed lizard Igλ light chains were amplified using Lizard-IgLVIs I (5′-TGT CCA TTC TCA AGG AAC GTT GAC TCA-3′), Lizard-IgLVIs II (5′-TGT CCA TTC ACA GCG GAC AGT GAC TCA-3′) as sense primers, and Lizard-IgLCas (5′-TTT CAT GGG TCA CCT TGC AGG TGT A-3′) (conserved among the lizard Cλ genes) as an anti-sense primer. Expressed lizard Igκ light chains were amplified using Lizard-VKs I (5′-GAA TCY AGT GGR CAR ATT GTC ATC ACT CA-3′), Lizard-VKs II (5′-GAG TCA AGT GGG GAM ATT GTK GTC ACT CA-3′), Lizard-VKss1 (5′-GAA TCC AGT GGA CAA ATT GT-3′), Lizard-VKss2 (5′-GAA TCC AGT GGG CAG ATT GT-3′), Lizard-VKss3 (5′-GCA TCA ACT GGG CAG ATT GT-3′) as sense primers, and CK-5′RACE3 (5′-AGT TGC TCC TGC GAG GGT GG-3′) as an anti-sense primer.

All the PCR amplifications were conducted using a proofreading enzyme *Pyrobest* DNA Polymerase (TaKaRa, Dalian). The resulting PCR products were cloned into the pMD-19 T vector (TaKaRa, Dalian) and sequenced.

2.3. Cloning of rearranged VJ sequence from genomic DNA

Genomic DNA was used to amplify germline rearranged Igλ VJ sequence. The sense primers used were LV12U (5′-AGG CTC TCT GTT TGC TTG TG-3′), LV15U (5′-TGC TGT TCT TGG TGC TTC TC-3′), LV16U (5′-GCC AAC AAC AGG AAC ATC TC-3′), LV19U (5′-CAC TTT CCA TAG GCA CAA GC-3′), LV25U (5′-TGG CTT CCT TGC TGT TCT TG-3′), and the anti-sense primers (derived from each J–C intron) were LC1JD (5′-TAA GTC TGG AGT AGG TTT GG-3′), LC2JD (5′-ATG GCA CTG GAA CAC TCT AC-3′), LC3JD (5′-GCA TGA ACC CCA ATT AGA CC-3′). All the sense primers were used in different combinations with the three anti-sense primers. The resulting PCR products were cloned and sequenced as above.

2.4. Southern blotting

Genomic DNA was digested with restriction endonuclease including Hind III, Nco I and Xba I, and was hybridized with individual Cκ, Vκ9, Cλ3, Vλ3, Vλ13, Vλ25 and Vλ35 cDNA probes which all were labeled using a PCR DIG probe synthesis kit (Roche). The hybridization and detection were conducted by following the manufacturer's instruction.

2.5. Construction of phylogenetic trees

Phylogenetic trees were made using MrBayes3.1.2 [28] and viewed in TREEVIEW [29]. Multiple sequence alignments were performed using *Clustalw*. Accession numbers of sequences used for variable regions were chicken λ (S65967); human λ (AC009286); mouse λ (AC140201); *X. laevis* type III V1 (L76575); *X. laevis* type III V2 (L76577); *X. laevis* type III V3 (L76586); *X. laevis* type III V4 (L76579); *X. laevis* type III V5 (L76576); *X. laevis* type III V6 (L76578); human κ (DQ101055); mouse κ (AY591701); *X. laevis* ρ (L15570); *X. laevis* σ (S78544); skate type I (L25568); hornshark type I (X15315); ratfish type II (L25549); nurse shark NS5 (AY720853); sandbar shark type II (M81314); skate type II (L25566); horn shark type II (L25560); hornshark type III (L25561); Salmon L3 (AF406956); fugu L1 (AB126061); carp L1b (AB073332); zebrafish L3 (AF246193); salmon L1 (AF273012); zebrafish L1 (AF246185); trout L1 (X65260); carp L3 (AB073335); trout L2 (U69987); carp L2 (AB091113); fugu L2 (DQ471453); zebrafish L2 (AF246183); Nurse shark σ (EF114759); horn shark σ (EF114760). Accession numbers of sequences used for constant regions were chicken λ (X04768); duck λ (X82069); platypus λ (AF525122); human λ (J00252); mouse λ (AC140201); *X. laevis* type III (BC082898); mouse κ (EF392842); human κ (AC210709); cow κ (BC122795); *X. laevis* ρ (BC068859); *X. laevis* σ (NM_001094414); zebrafish IGIC1 (AF246185); zebrafish IGIC2 (AF246162); zebrafish IGIC3 (AF246193); nurse shark NS5 (AY720857); nurse shark NS4 (L16765); nurse shark NS3 (Ref. [30]); cod IGIC1 (X68514); cod IGIC4 (AJ293808); salmon IGIC1 (AF273012); salmon IGIC2 (AF297518); salmon IGIC3 (AF406956); carp IGIC1 (AB015902); Carp IGIC2 (AB091120); carp IGIC3 (AB035730); skate type I (L25568); hornshark type I (X15316); skate type II (L25566); sandbar shark type II (M81314); hornshark type III (L25561); nurse shark σ (EF114766); horn shark σ (EF114760). All other sequences were either derived in this study or deduced from the genome database.

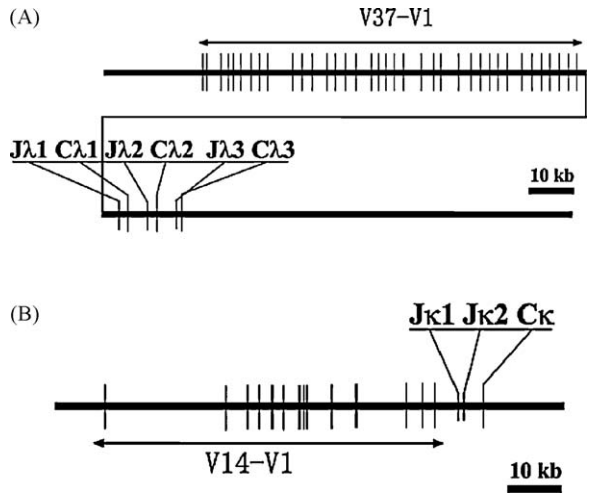

Fig. 1. Genomic organization of the lizard immunoglobulin light chain gene loci. (A) Genomic organization of the λ gene locus. (B) Genomic organization of the κ gene locus. V: variable gene segments; J: joining gene segments; C: constant region gene.