Biologicals 46 (2017) 81–87

Contents lists available at ScienceDirect

Biologicals

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

# Comparative analysis of the feline immunoglobulin repertoire

Sebastian C.J. Steiniger [a, *], Jacob Glanville [c], Douglas W. Harris [a], Thomas L. Wilson [a], Gregory C. Ippolito [b], Steven A. Dunham [a]

[a] Zoetis VMRD Global Therapeutics Research, 333 Portage Street, Kalamazoo, MI 49001, USA
[b] Department of Molecular Biosciences, University of Texas at Austin, Austin, TX, USA
[c] Distributed Bio Inc, San Francisco, CA, USA

ARTICLE INFO

*Article history:*
Received 1 September 2016
Received in revised form
2 November 2016
Accepted 9 January 2017
Available online 26 January 2017

*Keywords:*
Amino acid composition
Antibody engineering
Feline antibody repertoire
Immunoglobulin heavy chain
complementary-determining region 3

ABSTRACT

Next-Generation Sequencing combined with bioinformatics is a powerful tool for analyzing the large number of DNA sequences present in the expressed antibody repertoire and these data sets can be used to advance a number of research areas including antibody discovery and engineering. The accurate measurement of the immune repertoire sequence composition, diversity and abundance is important for understanding the repertoire response in infections, vaccinations and cancer immunology and could also be useful for elucidating novel molecular targets. In this study 4 individual domestic cats (*Felis catus*) were subjected to antibody repertoire sequencing with total number of sequences generated 1079863 for VH for IgG, 1050824 VH for IgM, 569518 for VK and 450195 for VL. Our analysis suggests that a similar VDJ expression patterns exists across all cats. Similar to the canine repertoire, the feline repertoire is dominated by a single subgroup, namely VH3. The antibody paratope of felines showed similar amino acid variation when compared to human, mouse and canine counterparts. All animals show a similarly skewed VH CDR-H3 profile and, when compared to canine, human and mouse, distinct differences are observed. Our study represents the first attempt to characterize sequence diversity in the expressed feline antibody repertoire and this demonstrates the utility of using NGS to elucidate entire antibody repertoires from individual animals. These data provide significant insight into understanding the feline immune system function.


## 1. Introduction

The development of High-Throughput Sequencing or Next Generation Sequencing (NGS) has made antibody repertoire sequencing feasible and affordable. The profiling of entire antibody repertoires, also referred to as Rep-Seq or BCR-Seq, offers new possibilities for studying whole antibody repertoires of individuals. These sequencing approaches, combined with bioinformatics, allow us to study and compare individuals and their antibody response in health and disease quantitatively and in a systems-based approach [2,7,8,16,18,27,45]. Recent examples in the search for potent neutralizing antibodies for Influenza or HIV and vaccine profiling demonstrate the power of this technique [18]. It also allows the precise determination of entire antibody lineages. Overall, NGS Sequencing allows us to generate an unprecedented depth to analyze antibody diversity [20].

The last decades have brought increasing interest in the cat as a companion animal as well as a model for specific diseases [21,23,32,41]. The feline immune system, similar to canine or the human immune system, achieves antibody diversity by V(D)J recombination, a process where a diverse assortment of variable genes (V), diversity (D) genes and joining (J) genes are recombined to create an antibody variable domain. Further post-recombination processes like insertion and deletion of nucleotides and somatic hypermutation can considerably increase the diversity of the variable region [6,24,33,43,46].

Felines show the canonical mammalian organization of the IGH genome and the IG locus has been described previously [12]; it consists of 24 IGHV genes, 13 IGKV and 47 IGLV genes [31]. Domestic cats (*Felis catus*) have been studied in detail immunologically; however, the actual, expressed feline antibody repertoire has not been characterized [4,25,26,36]. *Felis catus* has the same IGHV gene structure as *Canis familiaris*, with a similar IGLV gene expansion [26,31]. However, despite canine and feline being closely

* Corresponding author.
*E-mail address:* sebastian.steiniger@zoetis.com (S.C.J. Steiniger).

http://dx.doi.org/10.1016/j.biologicals.2017.01.004
1045-1056/

related, the feline IGHV locus is largely devoid of orthologs.

Recent studies have shown that, despite similarities in antibody structure, variable region diversification for both, heavy (VH) and light (VL) chain varies significantly between species. Furthermore, the relative role of VH and VL chains in antigen recognition depends on the species involved [11,14]. There are differences between species, specifically in the complementarity determining region 3 of the VH CDR-H3, in terms of length, structure and amino acid composition [40]. Because of its location in the center of the epitope binding site, and the regulation of B cell development and B cell immune responses [22,50], the exceptionally diverse VH CDR-H3 is a preferred spot for the introduction of *in vivo* and *in vitro* mutations for the affinity maturation of antibodies. For subsequent antibody engineering and speciation efforts a thorough understanding of interspecies differences in the amino acid bias of the CDR-H3 loop is paramount [15,30,47].

Biotherapeutics in veterinary medicine are now on the horizon and are expected to play a great role in the future of chronic disease treatment. Antibodies are multi-faceted proteins, capable of an extraordinary array of important functions. Not only major players in protection against invading pathogens, they also play critical roles in inflammatory and autoimmune diseases, their induction is central to many vaccine strategies, and they are an unrivalled platform for engineering highly effective diagnostic and therapeutic reagents for companion animals [17]. This study aims to present a characterization of the actual expressed feline antibody repertoire and is informative for the design of recombinant *de novo* feline antibody libraries since antibody based therapies for the treatment of feline diseases will require the generation of high-affinity, true feline, naturally occurring feline, feline derived, felinized or speciated antibodies. Like humanized antibodies, felinized or speciated molecules are less desirable due to the potential to be immunogenic and can be challenging in terms of expression or stability. In order to elucidate the diversity of the actual, productive feline VH and VL repertoire, we analyzed the successful antibody rearrangements to unknown antigens in four different domestic cats. All animals show a similarly skewed VH CDR-H3 profile, with very little sequence overlap between them. Also, we show that the feline repertoire is similar across different animals, and, even though it follows the canonical organization of the mammalian Ig locus and its similarity to the canine repertoire, it is distinctly different from its human, canine and mouse mammalian counterparts.

## 2. Materials and methods

### 2.1. Primer design

Protein sequences for known canine variable IgG domains were used to query the nucleotide database using the tBLASTn program. The aligned portion of each feline database hit was retrieved along with its corresponding nucleotide coordinates. Identified nucleotide sequences were aligned using CLUSTALW (Vector NTI, Invitrogen) and used for primer design. Degenerate primers for 10 VH FR1 were designed. For light chain amplification, 8 VL FR1 and 4 CL constant region primer (barcode) were designed for the lambda repertoire, and 4 VK FR1 and 4 CK region primer (barcode) were generated for the kappa repertoire. For the amplification of the IgG genes, 4 barcoded reverse primer in CH1 were used, and for the amplification of IgM sequences 4 barcoded reverse primer specific for the IgM constant region were used [5,9,42,44]. Primers were designed within the antibody constant regions such that the PCR product would result in approximately 400bp amplicons. The nucleotide sequences of all primers designed are listed in Table S1. The canine primers were designed and used as described in Ref. [40].

### 2.2. Isolation and quantitation of total RNA

Feline Blood samples were obtained from Bioreclamations (Bioreclamations Inc, NY, USA). All samples were collected from Domestic Short Hair Cats, which had not been medicated nor immunized. The animals are FELV and FIV negative and none of the donors were related. Blood was collected in PAXgene Blood RNA tubes (Becton Dickinson, NJ, USA) according to manufacturer's protocol and stored at −80 °C until further use. Before RNA purification, the PAXgene Blood RNA tubes were incubated for 2 h at room temperature. Thawed tubes were centrifuged at 5000 × *g* for 10 min, and the supernatant was discarded. The resulting pellet was dissolved in 300 μl Qiagen buffer BR1, and total RNA was isolated using the Qiagen QIAsymphony SP and PAXgene Blood RNA Kit (Qiagen, MA, USA). RNA quality was analyzed using an Agilent Bioanalyzer 2100 with RNA Nano Chip (Agilent Technologies, CA, USA).

### 2.3. cDNA synthesis and generation of feline IGH and IGL amplicons

The Life Technologies One step RT-PCR Kit with Platinum Taq was used for both cDNA synthesis and the subsequent PCR reactions. PCR cycling conditions were 50 °C for 30 min, followed by 95 °C for 4 min, followed by 40 cycles of 94 °C for 15 s, 56 °C for 30 s, and 68 °C for 90 s, and a final extension at 68 °C for 10 min. Independent amplicons pools were generated for IgG, IgM, Kappa and Lambda variable domains.

### 2.4. High-Throughput Sequencing of feline VH and VL gene repertoires

The PCR generated amplicons were processed for 2 × 250bp paired-end sequencing using the Illumina MiSeq platform according to the manufacturer's protocol. All datasets were processed using sequence quality and signal filters of the Illumina pipeline. Overlapping paired sequences were merged to generate a single full amplicon sequence using CLC Workbench (CLC Bio) default parameters with alignments scores costs of 2 for mismatch, 3 for gap, 0 for maximum unaligned end mismatches and 8 for minimum score. With this stringency, poorer quality sequences will not merge and all non-merging paired reads were discarded then duplicate amplicons were removed. The total number of unique amplicons obtained for analysis in this study were 1079863 for VH for IgG, 1050824 VH for IgM, 569518 for VK and 450195 for VL. The remaining sequences were then subjected to bioinformatics analysis using Abgenesis (DistributedBio) and IMGT/HighV-Quest [1,10,19,28,29] together with IgAT [35].

## 3. Results

We analyzed four individual domestic cats, and used RNA isolated from feline PBMCs to generate IgG and IgM antibody libraries. Overall, 3233336 sequence pairs were generated, and filtering based on quality score values, overlapping pairs merging and sequence uniqueness reduced the number to 1079863 unique VH for IgG, 1050824 unique VH for IgM, 569518 unique VK and 450195 unique VL sequences (Table 1) which were used to perform sequence alignments, identification of framework region (FR) and CDR analysis and subgroup usage. For the VH CDR-H3 analysis Distributed Bio's Abgenesis tools and IMGT/V-Quest [1,10,19,28,29] together with IgAT [35] were used. The unique clones and unique