Molecular Immunology 59 (2014) 71–78

Contents lists available at ScienceDirect

Molecular Immunology

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

Fundamental characteristics of the expressed immunoglobulin VH and VL repertoire in different canine breeds in comparison with those of humans and mice

Sebastian C.J. Steiniger[a,*], William E. Dunkle[a], Gary F. Bammert[a], Thomas L. Wilson[a], Abhiram Krishnan[a,b], Steven A. Dunham[a], Gregory C. Ippolito[b], Graeme Bainbridge[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

ARTICLE INFO

Article history:
Received 28 September 2013
Received in revised form 12 January 2014
Accepted 13 January 2014
Available online 4 February 2014

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

ABSTRACT

Complementarity determining regions (CDR) are responsible for binding antigen and provide substantial diversity to the antibody repertoire, with VH CDR3 of the immunoglobulin variable heavy (VH) domain playing a dominant role. In this study, we examined 1200 unique canine VH and 500 unique variable light (VL) sequences of large and small canine breeds derived from peripheral B cells. Unlike the human and murine repertoire, the canine repertoire is heavily dominated by the *Canis lupus familiaris* IGHV1 subgroup, evolutionarily closest to the human IGHV3 subgroup. Our studies clearly show that the productive canine repertoire of all analyzed breeds shows similarities to both human and mouse; however, there are distinct differences in terms of VH CDR3 length and amino acid paratope composition. In comparison with the human and murine antibody repertoire, canine VH CDR3 regions are shorter in length than the human counterparts, but longer than the murine VH CDR3. Similar to corresponding human and mouse VH CDR3, the amino acids at the base of the VH CDR3 loop are strictly conserved. For identical CDR positions, there were significant changes in chemical paratope composition. Similar to human and mouse repertoires, the neutral amino acids tyrosine, glycine and serine dominate the canine VH CDR3 interval (comprising 35%) although the interval is nonetheless relatively depleted of tyrosine when compared to human and mouse. Furthermore, canine VH CDR3 displays an overrepresentation of the neutral amino acid threonine and the negatively charged aspartic acid while proline content is similar to that in the human repertoire. In general, the canine repertoire shows a bias towards small, negatively charged amino acids. Overall, this analysis suggests that functional canine therapeutic antibodies can be obtained from human and mouse sequences by methods of speciation and affinity maturation.

1. Introduction

Immunoglobulin (IG) is an essential component of the adaptive humoral immune system in vertebrates. IG is composed of two heavy (H) and two light (L) chains, both of which contain variable and constant regions (Kirkham and Schroeder, 1994; Lefranc and Lefranc, 2001). In most mammals, H chains are encoded by a single locus of IGH genes, whereas L chains are encoded by a separate kappa or lambda locus (Kirkham and Schroeder, 1994). The diversity of the variable regions in mammalian IG is responsible for antigen recognition. IG diversity can be generated by six different processes: (1) germline diversity; (2) V(D)J rearrangement; (3) the insertion and deletion of nucleotides; (4) the process of somatic hypermutation; (5) recombinatorial diversity by VH:VL pairing and (6) gene conversion (Rajewsky, 1996; Tonegawa, 1983; Jenne et al., 2003; Becker and Knight, 1990; Weinstein et al., 1994). The primary immune response uses V(D)J rearrangement to generate sequence and structural diversity and yields antibodies with low affinity to a broad spectrum of antigens. The secondary immune response uses somatic hypermutation confined to germinal centre reactions resulting in high affinity antibodies with singular specificity (Tonegawa, 1983; Kim et al., 1981; Neuberger and Milstein, 1995). For canines, the IG locus has been described previously, and consists of 80 germline IGHV genes from three gene families, with the human VH3 analogue being the most prevalent (Bao et al., 2010; Braganza et al., 2011).

Abbreviations: VH, variable heavy (for domain); VL, variable light (for domain); V-KAPPA, variable kappa domain; V-LAMBDA, variable lambda domain; CL, constant light (for domain); C-KAPPA, constant kappa domain; AA, amino acid; VH CDR3, third complementarity-determing region of the variable heavy chain; RIN, RNA integrity value.

* Corresponding author. Tel.: +1 269 833 2846; fax: +1 269 833 4078.
E-mail addresses: sebastian.steiniger@zoetis.com, sebastian.steiniger@pfizer.com (S.C.J. Steiniger).

0161-5890/$ – see front matter © 2014 Elsevier Ltd. All rights reserved.
http://dx.doi.org/10.1016/j.molimm.2014.01.010

Antibody binding sites comprise six regions of sequence hypervariability (three for the VH domain and three for the VL domain), which are also referred to as complementarity determining regions (CDR). Antibody CDR contain extensive sequence diversity, even among germline genes, and the most important and diverse CDR in the VH domain is VH CDR3. It establishes more contacts with antigen than any other CDR (Schroeder et al., 1998). In all analyzed species, VH CDR3 is observed as the most diverse of all CDR (Glanville et al., 2009; Schroeder, 2006; Popkov et al., 2003; Butler and Wertz, 2012; Wang et al., 2013; Schroeder et al., 2010; Steiniger et al., 2007; Staflin et al., 2010; Fukuchi et al., 2010). Despite similarities in antibody structure, there are differences between species, specifically in VH CDR3 length, structure and amino acid composition (Finlay and Almagro, 2012; Ivanov et al., 2002). Because of its importance in determining antibody affinity, specificity, and the regulation of B cell development and B cell immune responses (Xu and Davis, 2000; Ippolito et al., 2006), VH CDR3 is a preferred spot for the introduction of *in vivo* and *in vitro* mutations for the affinity maturation of antibodies. Therefore, a thorough understanding of interspecies differences in the amino acid bias of the VH CDR3 loop is paramount for subsequent antibody engineering and speciation efforts (Mahon et al., 2013).

Monoclonal antibodies as therapeutics in veterinary medicine are a growing field and are expected to play a great role in the future of disease treatment. The ability to generate recombinant monoclonal antibodies against pathogens and therapeutic targets is an area of growing importance for livestock and companion animals alike (Gearing et al., 2013). For canines, treatments for pain, inflammatory diseases and cancer are potential areas of application; however it is currently not possible to obtain true canine antibodies through traditional hybridoma technology or B cell immortalization, which makes repertoire cloning an attractive alternative. A possible route to obtain true canine antibodies as well as the optimization of lead molecules lies in the generation of native canine antibody libraries and the subsequent screening against targets of interest (Braganza et al., 2011). Together with structural studies, antibody repertoire analyses have been highly informative approaches to understand how different species create diverse antibody repertoires (Finlay and Almagro, 2012), although at present little is known about the true breadth of IG diversity and structure of the functional canine immune response.

Antibody based therapies for the treatment of canine diseases will require the generation of high-affinity, fully canine, caninized or speciated antibodies. Like humanized antibodies, caninized or speciated molecules still possess the potential to be immunogenic and can be challenging in terms of development or stability. In order to support our antibody engineering efforts and to gain further insight into the structure of the productive canine VH and VL repertoire, we analyzed the diversity of a set of canine antibody sequences, including gene subgroup usage, length distribution, and amino acid composition of the VH CDR3 loops for different breeds. We show that the canine repertoire is similar among different breeds, and, even though it does exhibit global features evolutionary conserved among vertebrates, it is distinctly different from its human and mouse mammalian counterparts.

2. Materials and methods

2.1. Primer design

An in-house database, consisting of publically available whole genome DNA sequences of boxer and poodle breeds, DNA sequences obtained internally from dog chromosome 8, and libraries generated from amplification of antibody variable domains was generated and used for BLAST searching. Protein sequences for known human and mouse variable IgG domains were used to query the nucleotide database using the tBLASTn program. The aligned portion of each canine database hit was retrieved along with its corresponding nucleotide coordinates. A customized Perl script was used for batch retrieval of the nucleotide sequences encoding the translated canine hits. Identified nucleotide sequences were aligned using CLUSTALW (Vector NTI, Invitrogen) and used for primer design. In order to keep primer degeneracy low, 22 primers for VH FR1 were designed. For light chain amplification, 34 VL FR1 and 1 CL constant region primer were designed for the lambda repertoire, and 14 VK FR1 and 1 CK constant region primer were generated for the kappa repertoire. For the amplification of the IG genes, a reverse primer in CH1 was used, and for the amplification of IgM sequences a reverse primer specific for the IgM constant region was used (Bao et al., 2010; Braganza et al., 2011; Tang et al., 2001; Wasserman and Capra, 1978). The nucleotide sequences of all primers designed are listed in Table 2.

2.2. Isolation and quantitation of total RNA

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. Animals were grouped into two categories according to the size of the breed: small and large. Small breeds (under 20 kg) included Yorkshire Terrier, Dachshund and Chihuahuas; whereas large breeds (over 20 kg) included German Shepherds, Dobermans, Rottweilers and Golden Retrievers among others. 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 canine IGH and IGL sequences

All of the RNA samples with RIN values above 5 were used for cDNA synthesis. We used the Life Technologies One step RT-PCR Kit with Platinum Taq 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. PCR fragments were TA cloned into pCR2.1 (Life Technologies) and sequenced (Holton and Graham, 1991). For sequence analysis in-house bioinformatics tools, IMGT/V-Quest (Brochet et al., 2008; Giudicelli et al., 2011; Alamyar et al., 2012; Lefranc et al., 1999, 2012) and IGAT software were used (Rogosch et al., 2012).

3. Results

We focused our analysis on three different breeds, namely small breeds, large breeds and laboratory Beagles. RNA isolated from canine spleens and PBMCs was used to generate IgG and IgM antibody libraries. Overall, approximately 1200 unique VH and 500 unique VL sequences were used to perform sequence alignments, identification of framework region (FR) and CDR analysis of subgroup usage. For our VH CDR3 analysis we used the IMGT numbering system and IMGT/V-Quest (Brochet et al., 2008; Giudicelli et al., 2011; Alamyar et al., 2012; Lefranc et al., 1999, 2012) together with IgAT (Rogosch et al., 2012).