

# Immunoglobulin heavy chain diversity in Pteropid bats: evidence for a diverse and highly specific antigen binding repertoire

Michelle L. Baker · Mary Tachedjian · Lin-Fa Wang

Received: 4 August 2009 / Accepted: 22 January 2010 / Published online: 17 February 2010  
© Springer-Verlag 2010

**Abstract** Bats are the natural host reservoir for range of emerging and re-emerging viruses, many of which cause significant morbidity and mortality in other mammals, yet appear to result in no clinical consequences for bats. The ability of bats to coexist with a variety of viruses presents an interesting immunological problem that has not been examined in any detail but which could provide significant insights into the evolution of antiviral mechanisms in mammals. Towards a better understanding of the bat immune system, we analysed the expressed heavy chain variable (VH) regions of antibodies from the black flying fox, *Pteropus alecto*. The germline repertoire of the closely related Pteropid bat, *Pteropus vampyrus*, whose genome has been sequenced was also examined for comparative purposes. Representative VH genes were found in all three mammalian VH clans (I, II and III) in both the expressed *P. alecto* VH repertoire and the germline *P. vampyrus* VH repertoire. Evidence for the use of multiple heavy chain diversity (DH) and joining (JH) segments for the generation of diverse VDJ rearrangements was also present in the expressed antibody repertoire of *P. alecto*. The long period of co-evolutionary history of bats

with viruses may have resulted in a variety of highly specific VH segments being hardwired into the genomes of bats and may have implications for their ability to successfully cope with a diversity of viral antigens.

**Keywords** Immunoglobulin · Evolution · Bats

## Introduction

Immunoglobulin genes are assembled by recombination of germline-encoded gene segments: variable (V), diversity (D) and joining (J) for heavy (H) chains and V and J for light (L) chains. Variation in the amino acid residues at the N terminal ends that are encoded by the V regions of both H and L chains contributes to antibody diversity and establishes antibody specificity. There are three mechanisms for generating antibody diversity that are used to varying degrees by different species: germline diversity, somatic mutation and gene conversion. Germline diversity refers to the presence of a significant number of different germline-encoded segments that contribute to generating diversity in the primary antibody repertoire through the random combination of V, (D) and J segments. Humans and mice each have a large number of diverse germline-encoded VH segments that generate a highly diverse antibody repertoire. However, this seems to be more the exception rather than the rule, with most mammals having germline VH segments of high sequence similarity to each other and relying on post-V(D)J recombination mechanisms. In these species, additional diversity may be contributed by more extensive addition of N and P nucleotides during the recombination process, the addition of single base pair mutations through somatic hypermutations or gene conversion mechanisms (Butler 1997).

---

M. L. Baker (✉) · M. Tachedjian · L.-F. Wang  
CSIRO Livestock Industries,  
Australian Animal Health Laboratory,  
PO Bag 24, Geelong, Victoria 3220, Australia  
e-mail: Michelle.Baker@csiro.au

M. L. Baker  
Center for Evolutionary and Theoretical Immunology,  
Department of Biology, The University of New Mexico,  
Albuquerque, NM 87131, USA

L.-F. Wang  
Australian Biosecurity Cooperative Research Centre for Emerging  
Infectious Diseases,  
Geelong, Victoria 3220, Australia

Vertebrate VH genes are grouped into families based on the criteria of  $\geq 75\%$  nucleotide identity, and these families are further grouped into clans (Brodeur and Riblet 1984). Mammalian VH genes can be classified into groups A, B and C (or clans I, II and III) based on their sequence similarity (Kirkham et al. 1992). Species such as primates and rodents have multiple VH families with representatives of all three clans (van Dijk et al. 1993; Mainville et al. 1996). In contrast, most other mammals (e.g., rabbits, cattle, swine, sheep, platypus and opossum) and chickens have limited germline VH diversity. In these cases, the germline VH within a species typically share  $\sim 80\%$  nucleotide identity and may be only a single VH family or a set of recently derived VH families (Johansson et al. 2002; Butler 1997; Baker et al. 2005). The numbers of functional VH genes vary between species. In humans and mice, there are approximately 44 and 96 functional VH genes, respectively, in addition to a large number of VH pseudogenes (Das et al. 2008). In contrast, rabbits have  $\geq 100$  VH genes present in the genome, but only one VH segment is used for generating 70–90% of the expressed antibody repertoire (Knight 1992). A more extreme example is the chicken which has only one functional VH gene and a number of VH pseudogenes (Reynaud et al. 1989; McCormack et al. 1991).

Co-evolution with viruses and other pathogens plays an important role in shaping the diversity and specificity of the antibody repertoire and can act at both the somatic and genomic level (Ota and Nei 1994; Rajewsky et al. 1987; Kirkham et al. 1992; Roost et al. 1995). At the somatic level, hypermutation results in the introduction of single base pair mutations that increase antibody specificity, thus shaping the secondary antibody response. Positive selection on germline-encoded V segments is also believed to result in certain V segments being fixed in the germline to ensure efficient protection against common pathogens encountered over a long period of host pathogen co-evolutionary history (Langman and Cohn 1987).

Bats have been identified as a natural reservoir for a variety of viruses with RNA viruses accounting for the overwhelming majority of emerging pathogens (Calisher et al. 2006; Wong et al. 2006). Although bats may be persistently infected with many viruses, they rarely display clinical symptoms (Williamson et al. 1998, 2000; Leroy et al. 2005; Swanepoel et al. 1996; Leroy et al. 2009; Sulkin et al. 1966). The only viruses that have been demonstrated to cause clinical symptoms of disease in bats are rabies virus and the closely related Australian bat lyssavirus (Field et al. 1999; McKoll et al. 2002). However, results of experimental infections are inconsistent, with only a small proportion of bats succumbing to infection (McKoll et al. 2002). Viruses such as rabies have a long history of association with bats. However, as humans and other animals continue to encroach on bat habitats, more viruses

are being discovered, many of which, such as SARS coronavirus and Hendra and Nipah viruses, lead to lethal consequences in humans and other animals (Calisher et al. 2006; Wong et al. 2006). Whilst being the second most species-rich mammalian group after rodents, they are also the least studied of any group (Simmons 2005). How bats remain asymptomatic is not known, and few studies have examined the immune response of any bat species. Evidence that antibody responses in bats are both qualitatively and quantitatively lower than other mammals has been reported by a number of investigators (Hatten et al. 1968; Wellehan et al. 2009; Chakraborty and Chakravarty 1984). In addition, experimental virus infections have demonstrated the simultaneous presence of virus and antibody in bat blood. Sulkin et al. (1966) speculated that this could be due to the formation of a virus–antibody complex which dissociates readily or antibody which combines with virus but fails to neutralise infectivity. However, the nature of the antibody repertoire in bats has not been investigated, and no information currently exists on other antiviral mechanisms that may be responsible for inhibiting viral replication in bats. In this paper, we report the first molecular characterisation of the expressed antibody repertoire of a member of the Chiroptera family, the black flying fox, *Pteropus alecto*, with comparison to the germline VH repertoire of the closely related Pteropid bat, *Pteropus vampyrus*. Our results demonstrate that similar to primates and rodents, Pteropid bats have VH genes distributed in all three mammalian VH clans, consistent with the maintenance of V genes with the capacity to recognise a range of antigenic epitopes. The antigen-binding region of these genes also displays characteristics typically associated with weak antigen binding but high antigen specificity.

## Materials and methods

### Animals and RNA extraction

The *P. alecto* used in this study was a wild-caught individual from Queensland, Australia. The spleen from this adult male was collected in RNAlater (Ambion, Austin, TX, USA). Total RNA, with the concomitant removal of genomic DNA with DNaseI, was extracted using RNeasy mini kit following the manufacturer's recommended protocols (Qiagen, Valencia, CA, USA). Tissue homogenisation and disruption was carried out using a modification of the method described by Bossart et al. (2008) with homogenisation performed for  $2 \times 20$  s at 6,800 rpm with a Precellys 24 tissue homogenizer (Bertin Technologies, Carlsbad, CA, USA). All experiments were approved by the Australian Animal Health Laboratories animal ethics committee.

### Rapid amplification of cDNA ends

Full-length coding sequences for IgM, IgG and IgA were obtained using 5' and 3' rapid amplification of cDNA ends (RACE) polymerase chain reactions (PCRs) using the GeneRacer Kit (Invitrogen, Carlsbad, CA, USA) with GoTaq polymerase (Promega, Madison, WI, USA) and the manufacturer's recommended PCR conditions on total spleen RNA extracted from an adult male *P. alecto*. Primers were designed based on sequences obtained from the *P. vampyrus* whole genome sequence deposited in the Ensembl Genome Browser (Hubbard et al. 2009). 5' RACE was performed as a single-step PCR using IgM, IgG and IgA constant region-specific primers (5'-GCAGGAGGTTCTCACAGGAGACAA-3', 5'-GACCGAGGGTCCCTCCAGGAGAT-3' and 5'-GAGAAGTTCTCTCCGCGGTTCCAT-3'). 3' RACE was performed as a single step with IgM, IgG and IgA variable region-specific primers (5'-GTCCACCGGTAAACCCACCCTGTA-3', 5'-GACAGGACTTCACGTGTACGGTGAT-3' and 5'-GT CATGGAAGAGTGGGGACAGCTA-3'). To obtain additional sequences for the 5' end which contains the VH region, multiple clones were sequenced from RACE PCRs performed using 5' IgM- and IgG-specific primers.

### PCR amplification of germline JH segments

*P. alecto* germline JH segments were amplified using a primer specific for the recombination signal sequence (5'-GGG TTTTGTGAGGGAGA-3') of a JH segment identified in the *P. vampyrus* genome in combination with JH-specific primers (5'-TGAGGAGACGGTGACCACGG-3' and 5'-TGAGGAGACGGTGACTGGCG-3'). *P. alecto* genomic DNA was used as a target in PCR with an annealing temperature of 50°C using GoTaq Hotstart Polymerase (Promega).

### Sequencing and phylogenetic analysis

PCR and RACE-PCR products were cloned into the pCR4-TOPO vector using the TOPO TA Cloning Kit for sequencing (Invitrogen). Primers for the T3 and T7 promoters were employed for sequencing using BigDye Terminator Cycle Sequencing Kit v3.1 (Applied Biosystems, Foster City, CA, USA) with non-isotopic dye terminators in 10- $\mu$ l reactions, according to the manufacturer's instructions, and analysed on an Applied Biosystems 3130 XL Genetic Analyser. Chromatograms were edited manually using the Sequencher 4.6 software (Gene Codes Corporation, Ann Arbor, MI, USA) and were compared with sequences in the GenBank database using the BLAST algorithm (Altschul et al. 1990). All sequences were aligned using the ClustalX programme (Thompson et al. 1994). Nucleotide sequences were aligned and gapped manually using the Bioedit programme based on the protein alignment to retain codon positions. Alignments

were made using sequences corresponding to the framework regions (FR) 1–3 of the V domains, inclusive of complementary determining regions (CDR) 1 and 2. Based on the nucleotide alignments, phylogenetic trees were constructed by the neighbour joining (NJ) method of Saitou and Nei (1987), maximum parsimony (MP) and minimum evolution (ME) using the MEGA4 programme (Kumar et al. 2004). NJ and MP analyses using MEGA4 were performed using 1,000 bootstrap replicates and the Kimura two-parameter and mini heuristic search, respectively.

### VH sequences from other species

Previously published VH sequences from other species are as follows: Human, *Homo sapiens* (Hosa), VH sequences were obtained from the VBASE database (Tomlinson et al. 1996). Mouse, *Mus musculus* (Mumu), VH family representatives were as follows: 7183, U04227; 3660, K01569; 3609 N, X55935; DNA4, M20829; J558, Z37145; J606, X03398; Q52, M27021; S107, J00538; SM7, M31285; VH11, Y00743; and X24, X00163. Cow, *Bos taurus* (Bota) VH was AF015505. Sheep, *Ovis aries* (Ovar) VH was Z49180. Pig, *Sus scrofa* (Susc) VH was U15194. Possum, *Trichosurus vulpecula* (Trvu) VH was AAL87470. Opossum, *Monodelphis domestica* VH was AF012122. Platypus, *Ornithorhynchus anatinus* (Oran) VH was AF381291. Echidna, *Tachyglossus aculeatus* (Taac) VH was AAM61760, AAM60783, AY101439, AY101442 and AY101445. Horned shark, *Heterodontus franciscii* (Hefri) VH, X13449.

### Analysis of the *P. vampyrus* whole genome sequence

For comparative purposes, VH segments were identified in the whole genome sequence of the Malaysian flying fox *P. vampyrus* available in the Ensembl database (assembly pteVam1, 2.63X coverage, July 2008) using the BLAT, BLASTN and BLASTX algorithms.

## Results

### Isolation of *P. alecto* VH cDNA sequences

The constant regions of IgG, IgA and IgM were identified in the whole genome sequence of *P. vampyrus* and used to design oligonucleotide primers to amplify full-length sequences from *P. alecto* spleen cDNA using a combination of RACE and RT-PCR. Full-length IgG, IgM and IgA sequences containing the VH, DH, JH and CH regions homologous to sequences from other mammals were obtained. These sequences have been submitted to Genbank (accession nos. GQ427150–GQ427152). To sample the expressed VH repertoire in *P. alecto*, additional 5' RACE

PCRs were performed using the C $\gamma$  and C $\mu$  primers paired with the forward Generacer adaptor primer so as not to bias the VH sequences amplified. An additional 26 clones containing full VH sequences were isolated and sequenced which contained 20 unique VH sequences (accession nos. GQ427153–GQ427172). Unique sequences were identified based on the presence of differences in FR1–FR3. A total of 23 unique sequences including these 20 sequences along with the VH regions obtained from full-length IgG, IgM and IgA were included in further analyses.

#### Identification of multiple VH families in the expressed antibody repertoire of *P. alecto*

An alignment of the deduced amino acid sequences of the 23 unique *P. alecto* VH cDNAs is shown in Fig. 1, with sequences aligned based on sequence similarity in the V region. Pairwise analysis of the 23 *P. alecto* VH sequences revealed that these sequences shared 50–96% nucleotide identity to each other (Fig. 2). The presence of five distinct families designated VH1 (clones 19g, 20g, 21g, 26g, 30g, 31g, 35g, 36g, 1m and 2m); VH2 (IgA, 23g, 25g, 32g, 34g and 18g); VH3 (13g, 12g, 28g and 33g); VH4 (IgG and 10g) and VH5 (IgM) was apparent based on sharing  $\geq 75\%$  nucleotide identity within a family (Fig. 2). The same analysis performed excluding CDRs to limit the impact of somatic mutations on family assignment did not change the assignment of sequences to particular families (not shown).

Conserved alanine and arginine residues are present in the CDR3 region of VH genes of most species and form the base of the antigen-binding site (Schelonka et al. 2007). These residues were not conserved in the *P. alecto* group 3 VH sequences (Fig. 1). It is possible that the lack of conservation of these residues may result in the formation of a different range of structures in this subset of antibodies.

#### Evidence for multiple DH and JH segments in the *P. alecto* antibody repertoire

As expected, most of the variation in the *P. alecto* VH sequences was found in the CDRs, particularly the CDR3 region (Fig. 1). Sequences encoding the heavy chain CDR3 consist of the DH segment, N and P nucleotides and the 5' end of the JH segment. The CDR3 region plays an important role in determining the fine specificity of the antibody binding site (Xu and Davis 2000). The *P. alecto* CDR3 sequences ranged from 18 to 54 bp in length, corresponding to 6–18 amino acid residues (Figs. 1 and 3). The diversity of the *P. alecto* CDR3 sequences is consistent with the use of multiple different DH segments during recombination. Short stretches of similarity can be recognised between some CDR3 regions (20g and 30g; 36g and 33g) which may correspond to common germline DH

segments, or closely related ones. However, most of the DH regions appear quite different from each other, consistent with their origin from a diverse DH gene pool. Also evident in these sequences is a high GC content which often reflects extensive N nucleotide addition which in turn may contribute to changes in the amino acid composition within the antigen-binding site (Fig. 3).

The JH segment encodes the 3' part of the CDR3 and the whole of the FR4 region, and in most species, there are multiple JH segments in the germline that can be randomly recombined during V(D)J recombination (Butler 1997). A nucleotide alignment of the CDR3 to FR4 regions of 23 unique VH segments is shown in Fig. 3. Based on distinct patterns shared by multiple sequences, we identified at least five JH segments among the 23 sequences analysed, consistent with the presence of multiple germline JH segments (Fig. 3). Included in the alignment are two germline JH segments which correspond to JHI and JHII of the *P. alecto* cDNAs (accession no. GU354017). Comparison of the cDNA sequences and germline segments provides evidence for nucleotide substitutions, some of which may be the result of allelic variation, but some of these changes are also likely to be a consequence of somatic mutation. Additional germline JH segments were also identified which did not correspond to any of the cDNAs (not shown), thus providing additional evidence that *P. alecto* may have an even more diverse JH repertoire than is represented by the cDNAs sequenced here.

As the amino acid composition of the CDR3 region plays an important role in antigen specificity, the predicted residues in the CDR3 regions of the *P. alecto* VH genes were compared to those of mouse CDR3 regions (Kabat et al. 1991). The per cent amino acid residue composition of the CDR3 regions of *P. alecto* and mice, excluding the three C-terminal positions which are highly conserved in most species, is shown in Table 1. The most abundant amino acid residues in the *P. alecto* CDR3 regions were alanine (A, 14.9%), glycine (G, 13.6%), arginine (R, 10.9%), tyrosine (Y, 10.3%) and serine (S, 7.5%). Compared with the mouse sequences, A and R made up a larger proportion of the amino acid residues in the *P. alecto* CDR3 regions, whereas Y residues made up a smaller percentage (Table 1). In all species examined, from shark to human, the germline composition of the DH segment is enriched for Y, G and S in one of the reading frames. The amino terminus end of the JH segment also contributes to the CDR3 region, and Y is the most highly used amino acid in JH segments in other species (Ivanov et al. 2002). Identification of several germline JH sequences of *P. alecto* revealed that most sequences contained few Y residues similar to that observed in the *P. alecto* cDNAs (Fig. 3 and data not shown). DH segments identified in the genome of the closely related Pteropid bat, *P. vampyrus*, also appeared to display a paucity of Y residues

	Leader	FR1 20	40	CDR1	FR2 60
19g	MRLWGLLLCLVTAPQGVLS	QVQLQETGPGLVTPSQTLSLTCTVS		GFPLTS. .FS	VHWARQAPGKGLEWVGV
20g	-----	-----K-----		--S--N. .YA	---T-----P-----
21g	-----	-----K-----		--S--T. .YA	---T-----P-----
26g	-----	-----K-----		--S-- . .NA	---T-----N-----A
30g	-----D---	-A---S---K-----A--		--S--SN. .YG	I--I-----E--Q---Y
31g	-----	-P--R-S---IK-----		D-S--SN. .HG	---V-----AY
35g	-----	-----S-----K-----		--S--N. .YG	L--T-----A-Y
36g	-----	--E-L-----K-----		--S--IN. .Y-	---T-----I
1m	-----M-----	-----K-----		--S-- . .YA	---T-----A
2m	-----	-----K-----		-VS-M- . .SG	---T-----N
IgA	-AFGLSWVF--AVLR--QC	E---V-S-GDV-R-GES-R--AAT		--DFS. .YG	MG-V---QA-----SS
23g	-AFGLSWVF--AVLR--QC	E---V-S-GDV-R-GGS-R-S-AA-		--TFS- . .YA	MN-V--V-----SY
32g	-AFGLSWVF--AVLR--QC	E---V-S-GDV-R-GGS-R-S-AA-		--TFSN. .YG	MN-V--V-----SY
34g	-AFGLSWVF--AVLR--QC	E---V-S-GDV-R-GGS-R--VA-		--TFSN. .NW	MS-V-----SY
18g	-AFGLSWVF--AVLR--QC	E---V-S-GDV-R-GGS-R--AA-		--AFS- . .YG	MS-V---Q-R-----SS
25g	-AFGLSWVF--AVLR--HC	E---V-S-GAV-R-GDS-R--AA-		--VFGD. .YA	MT-V---Q-----SS
12g	-DTLCSTVL-L-V-HW-P-	--T-K-S---MR-T--T--SF-		--S--STADMG	-S-V--P---A---LAT
13g	-DTLCSTVL-L-V-HW---	--T-K-S---MR-T--T--SF-		--S--TTNMG	-S-V--P---A---LAT
28g	-DTLCSTVL-L-V-HW---	--T-K-S---LR-T--T--SF-		--S--TTDMG	-S-V--P---A---LAS
33g	-DTLCSTVLR-L-V-HW---	--T-K-S---R-T--T--SF-		--S--TTNLG	-S-V-----ET---LAT
10g	-DWSWSV-L--AVAT-AHA	----VQS-AEVVK-GASVKVS-KA-		-YTF-N. .YI	-N-V---R-Q---M-A
IgG	-DRSWRV-L--AVAT--HA	----VQS-AEVVK-GASVKIS-KA-		-YTFSD. .YR	M--V-----Q---M-D
IgM	-----	-----S-----K-----A--		-GSV--S. .YY	WS-I--R-----M- .
	CDR2 80	FR3 100	CDR3 120	FR4 140	
19g	VAKT.GNIYKFLKSRVTV	TRDISKSQVFLTVNSVTSEDTAIYYC	AREEFPWSGS. . . . .	FDG WGQGAVTVSS	
20g	IGYS. --T---SA-----RI	---T-----KL-----V---	--DPN-EVRPGTN.NYAM-V	--R--L-----	
21g	MTYS. -ST--NPA----LSI	---T-----SL-----M---	--IG.AGT- . . . . .	--R-S-----	
26g	IWGG. -STN-NPA----LSI	---T-----KL-----V---	-GG. .TGVAEY. .YYGM-A	--R--L-----	
30g	SWSD. -DTV-NPI----LSL	---T--G--S-KLS---V---V---	--DGYSSTWYPS. .YFGM-A	--R--L-----	
31g	IWND. -ETS-NSA-R--LSI	---T-----NL---A---L-F-	--DLG-RVD. . . . .WMHA	--R-VL-----	
35g	IWSD. -DTS-TPA----LSI	---T-----KL-----V---	---G. .VE- . . . .LYN---	---T-----	
36g	M-TS. -RT--NST---LNI	---T--N--FRL-----V-F-	--APGYST- . . . .YYGM-A	--R--L-----	
1m	IGSS. -RT--NP---LSI	-K-T-----KL-----M---	--DGESC--TCYPDYR-	---T-----	
2m	IWSD. --TK-NP--QA-LSI	---T-----ML---DK---V---	VI-DIGRT- . . . . .ELGL	---T-----	
IgA	ISASGEIR--GP-V-G-F-I	S--N--KTLY-EMRN-KF----V-H-	TQ-SYGS-RNDRF. .YGM-V	--R-VL-S--P	
23g	IRYDGSST--AA-V-G-F-I	S--NA-NTLY-QMS-LKP---V---	-KTSIARA-TRAD. .SYV-V	--R-SP-----	
32g	IRNDGSST--AA-V-A-F--	S--NA-NTLYVQMS-LQP---V---	-AA. . . . .W--S	---L-----	
34g	IRSDDSSSTH-AD-V-G-F-I	ST-NAINT-Y-QLS-LKAD---V---	TTSTAVTRY. . . . .GM-A	--R--L-----	
18g	ISISG-ST--PA-V-G-F-M	S--N--NTLY-QMS-LKP---L---	T--NYAYRTP. . . . .L-L	---TG-----	
25g	IRDNGTSA--TA-VRG-F-I	S--NFRQMA--QLSRLKI--SG---	YPGLG-D-E. . . . .--R	---T--S---	
12g	IWWD. DDKN-SP-----L-M	-K-T--N--L--LTNTDPR---T---	--TG.LKV-YS. . . . .YV-V	---SP-----	
13g	IWWD. DDK--SP-----L-I	-K-T--N--V--MS-TDPA---T-F-	V-MG.DNWNLYP. . . .S---	--R-T-----	
28g	IWWD. DEK--SP-----L-I	-K-T--N--V--MSNTNPA-----	--M. .DPFN-DY. . . .Y-S	---TP---P-	
33g	-WWD. DDKH-RA-----L-I	-G-T-RN--V--LTTADPS-A-T-F-	--IPPDG-YWSR. . . .P--L	--P-I--G---	
10g	-HPDD-VTD-AQ-FRG---I	-A-R-TNT-Y-DLK-LGP--A-T---	V-AVA. . . . .S-Y	---T-----	
IgG	-SPKT-VTDHAQTFEG---F	-A-K-TRTA-MDLI-LRP---T---	--SYG. . . . .--L	---T-----	
IgM	YWS. .-STR-NPAFQG-ISI	-A-T-RN-FS-QL-----V---	-- . . .GIA- . . . .DNL--	---T-----	

**Fig. 1** Alignment of deduced amino acid sequences of 23 *P. alecto* VHs. The designation following the clone number refers to isolation by RACE-PCR using C $\mu$  (m) or C $\gamma$  (g) specific primers. IgM, IgG

and IgA clones refer to VH regions obtained from full-length clones. Sequences are aligned based on sequence similarity in the VH region. Dashes indicate similarity and dots indicate gaps

(data not shown). This result is consistent with the lack of Y in the *P. alecto* sequences being the result of the amino acid composition of the DH and JH regions rather than a consequence of somatic mutations.

*P. vampyrus* has a diverse germline VH repertoire

Although no genome sequence is available from *P. alecto* as yet, a low coverage whole genome sequence is publicly

VH1 (10)	75-94			
VH2 (6)	57-66	76-96		
VH3 (4)	60-66	53-59	85-96	
VH4 (2)	54-58	55-65	49-52	80
VH5 (1)	67-73	54-61	56-57	52
	VH1	VH2	VH3	VH4

**Fig. 2** Range of per cent nucleotide identity between 23 *P. alecto* VH sequences based on an alignment of the region corresponding to FR1 to FR3 inclusive of CDRs. Numbers in parentheses indicate number of clones isolated from each family

available from a closely related Pteropid bat, *P. vampyrus*. VH sequences from *P. alecto* and other mammals were used to search the *P. vampyrus* whole genome sequence in the publicly available Ensembl database to examine the germline VH repertoire of this species. Thirty-three scaffolds

ranging from 1.22 to 54.06 Kb contained VH segments, resulting in the identification of 74 unique *P. vampyrus* VH sequences. Of these, 11 contained stop codons and are presumably pseudogenes. Fifty-one sequences had a leader peptide, open reading frame (ORF) and an identifiable recombination signal sequence (RSS) and are potentially functional VH segments. Of the remaining 12 VHs with ORFs, no leader was identified in five and no RSS was found in seven. Pairwise analysis was performed using 61 of the most complete *P. vampyrus* sequences with ORFs (51 potentially functional VHs, five with an RSS but no signal sequence and five with a signal sequence but no identifiable RSS). The results of this analysis demonstrated that these sequences share 49–99% nucleotide identify and belong to five different VH families based on sharing  $\geq 75\%$  nucleotide identity within a family (Fig. 4). However, due to the fragmentary nature of the *P. vampyrus* assembly, it is likely

	CDR3												FR4																										
<b>I</b>																																							
Genomic3-2													TAC	TCG	TAC	TTC	GGT	GTC	TGG	GGC	CCG	GGC	TCG	CCA	GTC	ACC	GTC	TCC	TCA										
12g													Y	S	Y	F	G	V	W	G	P	G	S	P	V	T	V	S	S										
23g	ACT	AGT	AT-	GCT	CGA	-CT	GG-	-CC	AG-	-C-	G-	-	-	-	G-	-	D	-	-	-	G-	-	C	-	A	-	-	-	-	-	-	-	-	-					
	T	S	I	A	R	A	G	T	R	A	D	-	-	-	-	-	-	-	-	R	-	-	-	-	-	-	-	-	-	-	-	-	-						
<b>II</b>																																							
Genomic1													AAC	TTT	GAT	TTC	TGG	GGC	AAG	GGA	ACC	GTG	GTC	ACC	GTC	TCC	TCA												
10g													N	F	D	F	W	G	K	G	T	V	V	T	V	S	S												
IgM													GCG	GTG	GC-	AG-	-C	-	AT	-	-	-	-	-	-	-	-	-	-	-									
19g													A	V	A	S	-	Y	-	-	-	-	-	-	-	-	-	-	-	-									
35g													G	I	A	G	D	-	L	-	-	-	-	Q	-	-	-	-	-	-									
13g	ATG	GGT	GAT	AAC	T-G	-AC	TAC	CT-	CCA	TCT	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-							
1m	GAT	GGG	GAA	AGT	TGC	TCG	GGT	TCA	ACT	TGT	TAC	CCC	GAC	TAT	CG-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-						
	D	G	E	S	C	S	G	S	T	C	Y	P	D	Y	R	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-						
<b>III</b>																																							
32g													GCG	TGG	TTT	GAC	TCC	TGG	GGC	CAG	GGA	GCC	CTG	GTC	ACC	GTC	TCC	TCG											
21g													A	W	F	D	S	W	G	Q	G	A	L	V	T	V	S	S											
25g													I	G	A	G	T	G	-	F	-	-	R	-	S	V	-	-	-										
31g													G	L	G	P	D	S	E	-	R	-	-	-	-	T	V	-	S	-	-								
34g													D	L	G	P	R	V	D	-	M	H	A	-	R	-	V	-	-	-	-								
36g													S	T	A	V	T	R	Y	Y	G	M	-	A	-	R	-	-	-	-	-								
26g													GCC	CCA	GGG	TAC	AGT	<u>ACT</u>	<u>GGT</u>	<u>TAT</u>	<u>TAC</u>	G-C	A-G	-	-	-	-	-	-	-	-								
30g													A	P	G	Y	S	T	G	Y	G	M	-	A	-	R	-	-	-	-	-								
20g													GGT	ACT	ACT	GGA	GTA	<u>GCT</u>	<u>GAG</u>	<u>TAT</u>	<u>TAC</u>	TAC	G-C	A-G	-	-	-	-	-	-	-								
IgA	GAG	TCG	TAT	GGG	TCG	AGC	CGG	AAC	GAT	CGT	TTC	TAC	GGC	ATG	GAC	GTC	TGG	GGT	CGG	GGA	GTC	TTG	GTC	AGC	GTC	TCC	CCA												
	E	S	Y	G	S	S	R	N	D	R	F	Y	G	M	D	V	W	G	R	G	V	L	V	S	V	S	P												

**Fig. 3** Nucleotide and amino acid translation from CD3 to FR4 of 23 different *P. alecto* VH sequences and two germline JH segments (germline3-2 and germline1). Roman numerals indicate the five

different groups of JH segments. Regions of homology within the CDR3 region of different clones are shown *underlined in bold*. Dashes indicate similarity

**Table 1** Percentage of amino acids in the CDR3 of 23 *P. alecto* VH genes and from mouse VH genes

		Bat	Mouse <sup>a</sup>
Small hydrophilic	C	0.51	0.16
	S	7.5	9.28
	T	5.25	1.6
	P	5.4	2.24
	A	14.9	4.16
Acid, acid amine, hydrophilic	G	13.6	14.97
	N	3.3	2.24
	D	5.4	10.72
	E	4.7	1.52
	Q	0.28	0.08
Basic	H	0	1.92
	R	10.9	4.72
	K	0.67	0.4
Small hydrophobic	M	0.75	0.72
	I	2.6	0.88
	L	2.1	4.32
	V	4.7	1.6
Aromatic	F	1.4	4.32
	Y	10.3	29.7
	W	2.8	4.72

<sup>a</sup>From Kabat et al. 1991

that additional VH segments exist in the genome of this species. Since *P. alecto* and *P. vampyrus* are closely related species of Pteropid bats, we might expect *P. alecto* to have similar numbers of germline VH segments in its genome.

Pteropid bats have members of all three mammalian VH clans

Phylogenetic analysis was performed using a nucleotide alignment of the *P. alecto* VH sequences with VH sequences from several other mammalian species. Included in the analysis were 11 mouse, six human, one pig, one sheep, one cow, one opossum, one possum, one platypus, five echidna and five *P. vampyrus* VH sequences. Horned shark, *H. francisci*, was used as an outgroup to the mammalian VH because elasmobranch VH genes have been shown to be the most ancient of the vertebrate VH genes (Sitnikova and Su 1998). The tree shown in Fig. 5 was reconstructed using the NJ method; however, identical results were found when MP and ME were used (not shown). The mammalian VH cluster into three groups (I, II and III in Fig. 5), sometimes referred to as A, B and C, respectively (Nei et al. 1997). The *P. alecto* and *P. vampyrus* VH are clustered within clans I, II and III (Fig. 5). The clustering of sequences is consistent with the pairwise analysis described above and demonstrates that the

VH1 (12)	94-97				
VH2 (29)	54-67	76-99			
VH3 (2)	67-70	49-61	98		
VH4 (16)	55-52	56-69	50-56	75-96	
VH5 (2)	69-73	53-58	70-73	50-56	93
	VH1	VH2	VH3	VH4	VH5

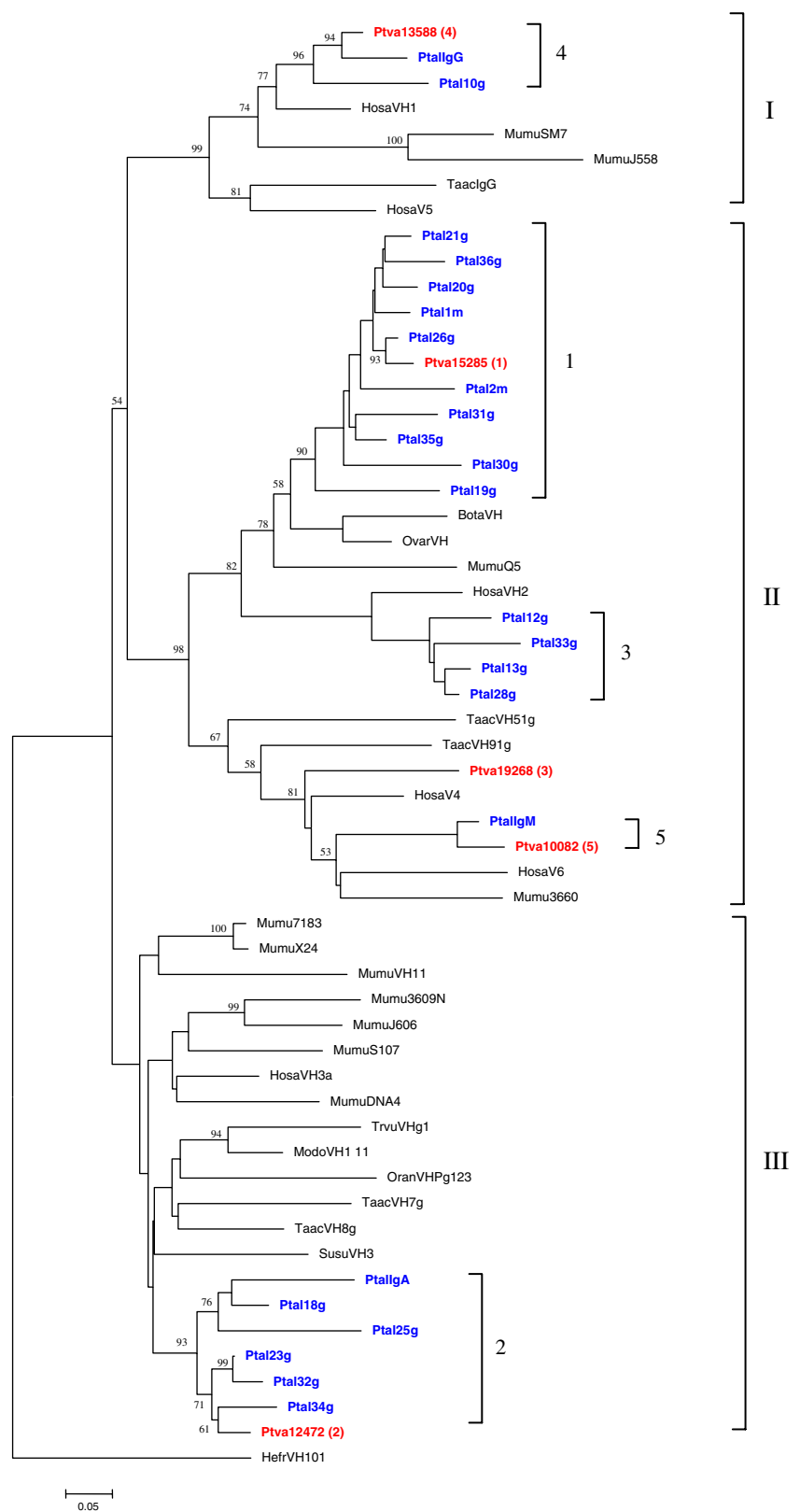
**Fig. 4** Range of per cent nucleotide identity between 61 *P. vampyrus* VH sequences based on an alignment of the region corresponding to FR1 to FR3 inclusive of CDRs. Numbers in parentheses indicate number of sequences identified from each family

*P. alecto* sequences that are members of the same VH family cluster in a species-specific manner within the phylogenetic tree. Phylogenetic analysis was also performed using 61 *P. vampyrus* VH sequences with the most complete ORFs (not shown). As with *P. alecto*, the *P. vampyrus* VH segments clustered in all three known mammalian VH clans, (16 in group I, 16 in group II and 29 in group III). Of these 61 *P. vampyrus* sequences, only one representative of each of the five VH families identified in the pairwise analysis described above is shown in Fig. 5. The *P. vampyrus* VH sequences were interspersed with the *P. alecto* sequences, with the exception of members of family 3 which formed their own clade in each species within the VHII clan. The clustering of *P. alecto* and *P. vampyrus* sequences together in the phylogenetic tree is consistent with the duplication of these genes prior to the divergence of Pteropid bats from a common ancestor. Overall, our phylogenetic and pairwise analyses described above demonstrate that *P. alecto* and *P. vampyrus* have a diverse set of VH genes distributed in multiple families across each of the mammalian VH clans.

## Discussion

Bats are asymptomatic carriers of a variety of viruses but have been poorly studied in terms of their immune response to viral infection. The few studies that have examined immune responses of bats have reported lower antibody responses compared to other mammals and evidence for inconsistent formation of neutralising antibody in response to viral infection (Sulkin et al. 1966; Hatten et al. 1968; Wellehan et al. 2009; Chakraborty and Chakravarty 1984). Although it is possible that other immune mechanisms such as those involved in innate immunity are responsible for controlling viral replication prior to the development of the antibody response, little information is available on any aspect of bat immunology. Defining the antibody repertoire of this unique group of mammals is an important step in understanding the nature of antiviral responses and the mechanisms involved in the asymptomatic nature of viral infections in bats. Towards understanding the antibody

**Fig. 5** Phylogenetic tree based on alignments of VH sequences from *P. alecto*, *P. vampyrus* and representatives of other mammals and one non-mammal. All *P. alecto* sequences are shown in blue and *P. vampyrus* sequences are shown red in the tree. The tree was constructed using the NJ method (Saitou and Nei 1987). The numbers on the branch nodes indicate bootstrap values based on 1,000 replicates. Roman numerals on the right indicate the three VH clans (Tutter and Riblet 1989). Numbers within the three clans indicate the *P. alecto* VH family identified by pairwise analysis (Fig. 2). Species names are: *B. taurus* (Bota), *H. sapiens* (Hosa), *H. fransciscii* (Hefr), *M. domestica* (Modo), *M. musculus* (Mumu), *O. aries* (Ovar), *O. anatinus* (Oran), *P. alecto* (Ptal), *P. vampyrus* (Ptva) *S. scrofa* (Susc), *T. vulpecula* (Trvu), *T. aculeatus* (Taac). Numbers after the *P. vampyrus* sequences correspond to scaffolds in the *P. vampyrus* whole genome assembly, and numbers in parentheses indicate the five *P. alecto* families identified by pairwise analysis (Fig. 2)





responses of bats and the evolution of the antibody repertoire in Chiroptera, we analysed the level of expressed VH diversity in the black flying fox, *P. alecto*, and compared that to the level of germline VH diversity in the closely related flying fox, *P. vampyrus*. Both *P. alecto* and *P. vampyrus* are old-world fruit bats and members of the Pteropodidae family which forms a monophyletic clade within the Megachiropteran order (Jones et al. 2002). *P. alecto* has a geographic distribution that extends from Australia into Papua New Guinea and Indonesia, whilst *P. vampyrus* is found predominantly in Malaysia, Indonesia, Thailand and the Philippines (Nowak, 1994). Although their geographic range is not the same, they are host reservoirs of a variety of closely related viruses, the most important of which include the Henipaviruses: Hendra in *P. alecto* and Nipah in *P. vampyrus* (Wong et al. 2006).

#### Evolution of Chiroptera VH genes

VH genes evolve by the birth and death process with the duplication and deletion of VH segments, resulting in some species having retained a high level of germline diversity whereas others having lost it (Ota and Nei 1994; Butler 1997; Das et al. 2008). The VH genes of a range of tetrapods are distributed in three main clans (I, II and III). VH segments corresponding to all three VH clans are found in the amphibian *Xenopus laevis*, consistent with the appearance of VH groups I, II, and III early in tetrapod evolution (Haire et al. 1990). Species such as sheep and cattle have deleted much of the ancestral repertoire and only have representatives of clan II, whereas pigs only have clan III genes (Butler 1997). Among mammalian lineages, primates and rodents have the most diverse VH repertoires, with a large number of VH genes distributed in all three clans (Butler, 2006; Das et al. 2008). A recent analysis of the genomes of cats and dogs has revealed that these species also have representative VH genes in all three mammalian VH clans, but the majority of their VH genes belong to clan III and they have only limited diversity in clans I and II (cats have three VHI genes and a single VHII gene and dogs have only a single member of both the VHI and VHII clans; Das et al. 2008). The monotreme, the echidna, also has VH genes in all three mammalian clades, but the platypus has only VHIII genes (Baker et al. 2005). The expressed VH repertoire of *P. alecto* and the germline repertoire of *P. vampyrus* provide evidence for a diverse VH repertoire within the Pteropid bats, with the expressed antibody diversity of *P. alecto* being derived from all three mammalian VH clans. The two members of the Chiroptera family described here represent the only eutherian species other than primates and rodents that are known to have retained a high level of the ancestral VH diversity (Das et al. 2008). As the H chain has been demonstrated to make

the most important contribution to antigen binding in humans and mice, evidence of a diverse VH repertoire in bats may be consistent with this also being the case in bats (Wilson and Stanfield 1993).

Although the fossil record for bat evolution is incomplete, bats are believed to have originated approximately 50 million years ago (Teeling et al. 2005). The correspondingly ancient origin of many of the viruses maintained in bats is consistent with a long period of co-evolution of bats and viruses (reviewed in Calisher et al. 2006). The presence of a large pool of V segments from which to recombine the expressed VH repertoire might be expected to provide bats with a large number of antigenic specificities that may be important in a species exposed to a variety of viral antigens. The co-evolution of hosts and pathogens has the ability to shape antibody diversity, resulting in some V region genes being selectively retained in the germline due to their capacity to accommodate particular antigens (Ota and Nei 1994; Rajewsky et al. 1987; Kirkham et al. 1992; Roost et al. 1995). It is possible that the high VH diversity seen in bats is a consequence of the co-evolution of bats with viruses, resulting in certain V regions from each of the three mammalian VH clans being hardwired into the genome to maintain important binding specificities.

#### *P. alecto* VH genes display characteristics associated with high antigen specificity

In addition to a diverse repertoire of VH segments, the expressed VH genes of *P. alecto* also displayed evidence of being derived from a number of different germline-encoded DH and JH segments. The DH segment is encoded within the CDR3 region which is responsible for much of the fine specificity of the antibody binding site (Xu and Davis 2000). Long CDR3 regions are found in species with only a single VH family such as platypus, cattle and chickens whose CDR3 regions are up to 19, 27 and 30 amino acid residues, respectively, and may compensate for limited germline diversity (Berens et al. 1997; Mansikka and Toivanen 1991; Johansson et al. 2002). The *P. alecto* CDR3 regions ranged from 6 to 18 amino acid residues in length and are therefore comparable to those of humans and mice (Stenvik et al. 2000). The Pteropid bats therefore appear to fit the same pattern as humans and mice, with high germline diversity reducing the need to diversify the CDR3 regions. The JH regions identified in the expressed VH repertoire also appeared to correspond to multiple germline segments which displayed some evidence for somatic mutation. The number of functional JH segments in other species ranges from only one in chickens to six in humans (Reynaud et al. 1989; Ravetch et al. 1981). The five different JH regions that were identified in our cDNAs are therefore comparable to that of humans. Overall, our

data are consistent with *P. alecto* deriving its antibody repertoire from a large pool of VH, DH and JH segments.

A curious aspect of *P. alecto* VH sequences was evident in the amino acid composition of the CDR3 regions. The amino acid composition of the CDR3 region plays an important role in the ability of antibodies to bind antigen, affecting both the specificity of the antigen–antibody interaction and the ability of antibodies to accommodate different antigens. Antibodies from a variety of species ranging from amphibians to humans and mice have CDR3 regions that are rich in tyrosines and glycines (Berens et al. 1997; Golub et al. 1997; Reynaud et al. 1989; Friedman et al. 1994). The CDR3 regions of polyreactive antibodies are rich in aromatic residues such as tyrosine and tryptophan which interact with structurally diverse antigens and form a “sticky” antigen-binding site (Padlan 1994; Mian et al. 1991). The CDR3 regions of the *P. alecto* VH repertoire contained fewer tyrosine residues compared with other mammals and a greater proportion of arginine and alanine residues. Tyrosine residues appear to be directly involved in antigen binding as they can interact with relatively distant structures by hydrogen bonding through their aromatic ring or through polar atoms in their side chain (Mian et al. 1991). The presence of tyrosines in the CDR3 also confers structural diversity, allowing for many forms of intermolecular and intramolecular bonding which in turn can lead to greater polyreactivity (Zemlin et al. 2003). In contrast, arginine residues have been reported to be detrimental to antigen binding even in small doses and may even contribute to self-reactivity (Birtalan et al. 2008; Radic et al. 1993). The combination of fewer tyrosines and a higher number of arginines in the bat CDR3 regions compared to other species may have resulted in the evolution of antibodies with lower polyreactivity that form only a weak association with antigen. Although such antibodies would be unable to bind with the same avidity as those with a “sticky” antigen-binding site, they may be less promiscuous and therefore have greater specificity for particular antigens once they do bind. Observations that bats fail to produce neutralising antibody following infection with certain viruses may also be due to the weak association of antigen–antibody complexes formed (Sulkin et al. 1966). However, whether this is really the case remains to be experimentally determined as do the other immune mechanisms responsible for inhibiting viral replication in bats.

In mice, distinct populations of B cells exist within different compartments of the spleen. During development, B cells with CDR3 regions that are enriched in arginine and hydrophobic amino acids are sorted away from the follicles but remain abundant in the marginal zone (Schelonka et al. 2007). As unsorted spleen cells were used in the present study, it is possible that bats have a larger marginal zone population that may have biased the population of VH regions to those

enriched in arginine and alanine. Other factors that shape the antibody repertoire include age and prior antigen exposure. The sequences obtained in the present study were from the spleen from a mature wild-caught male flying fox and therefore represent the adult antibody repertoire. However, as the history of antigen exposure of this individual is unknown, it is possible that previous viral infections have contributed to the selection of antibodies that encode particular amino acids. Clearly, more work remains to be performed to examine the nature of antibody development in bats and the influence of antigen exposure. However, the identification of sequences with a similar bias in amino acid composition in the genome of the closely related Pteropid bat, *P. vampyrus*, is consistent with bats having an unusual antibody repertoire.

In summary, this study represents the first analysis of the antibody repertoire in any Chiroptera species. Our results demonstrate that at least in Pteropid bats, there is evidence for a diverse antigen-binding repertoire with antibodies that may form weak but highly specific antigen–antibody complexes. The evolution of antibodies of this nature is likely due to the long co-evolutionary history of bats with viruses, providing a mechanism to rapidly recognise common viruses without the need for antibodies to undergo extensive further hypermutation to increase specificity. Further work will examine the nature of the antibody response to viral infection, the antibody repertoire of infected bats and determine whether post-V(D)J recombination mechanisms also play a role in generating antibody diversity in bats.

**Acknowledgements** The authors wish to acknowledge support from a CSIRO CEO Science Leaders award to L-F.W. We thank Craig Smith, Hume Field, Gary Crameri and Deborah Middleton for provision of bat spleen material for this study.

## References

- Altschul SF, Gish W, Miller W, Myers EM, Lipman DJ (1990) Basic local alignment search tool. *J Mol Evol* 215:403–410
- Baker ML, Belov K, Miller RD (2005) Unusually similar patterns of antibody V segment diversity in distantly related marsupials. *J Immunol* 174:5665–5671
- Berens SJ, Wylie DE, Lopez OJ (1997) Use of a single VH family and long CDR3s in the variable region of cattle Ig heavy chains. *Int Immunol* 9:189–199
- Birtalan S, Zhang Y, Fellouse FA, Shao L, Schaefer G, Sidhu SS (2008) The intrinsic contributions of tyrosine, serine, glycine and arginine to the affinity and specificity of antibodies. *J Mol Biol* 377:1518–1528
- Bossart KN, Tachedjian M, McEachern JA, Crameri G, Zhu Z, Dimitrov DS, Broder CC, Wang LF (2008) Functional studies of host-specific ephrin-B ligands as Henipavirus receptors. *Virology* 72:357–371
- Brodeur PH, Riblet R (1984) The immunoglobulin heavy chain variable region (Igh-V) locus in the mouse. I. One hundred Igh-V

- genes comprise seven families of homologous genes. *Eur J Immunol* 14:922–930
- Butler JE (1997) Immunoglobulin gene organization and the mechanism of repertoire development. *Scand J Immunol* 45:455–462
- Butler JE (2006) Why I agreed to do this. *Dev Comp Immunol* 30:1–17
- Calisher CH, Childs JE, Field HE, Holmes KV, Schountz T (2006) Bats: important reservoir hosts of emerging viruses. *Clin Microbiol Rev* 19:531–545
- Chakraborty AK, Chakravarty AK (1984) Antibody mediated immune response in the bat, *Pteropus giganteus*. *Dev Comp Immunol* 8:415–423
- Das S, Nozawa M, Klein J, Nei M (2008) Evolutionary dynamics of the immunoglobulin heavy chain variable region genes in vertebrates. *Immunogenetics* 60:47–55
- Field H, McCall B, Barret J (1999) Australian bat lyssavirus infection in a captive juvenile flying fox. *Emerg Infect Dis* 5:438–440
- Friedman ML, Tunyaplin C, Zhai SK, Knight KL (1994) Neonatal VH, D and JH gene usage in rabbit B lineage cells. *J Immunol* 152:632–641
- Golub R, Fellah JS, Charlemagne J (1997) Structure and diversity of the heavy chain VDJ junctions in the developing Mexican axolotl. *Immunogenetics* 46:402–409
- Haire RN, Amemiya CT, Suzuki D, Litman GW (1990) Eleven distinct VH gene families and additional patterns of sequence variation suggest a high degree of immunoglobulin gene complexity in a lower vertebrate, *Xenopus laevis*. *J Exp Med* 171:1721–1737
- Hatten BA, Allen R, Sulkin SE (1968) Immune response in Chiroptera to bacteriophage X174. *J Immunol* 101:141–150
- Hubbard TJ, P, Aken BL, Ayling S, Ballester B, Beal K, Bragin E, Brent S, Chen Y, Clapham P, Clarke L, Coates G, Fairley S, Fitzgerald S, Fernandez-Banet J, Gordon L, Gräf, S, Haider S, Hammond M, Holland R, Howe K, Jenkinson A, Johnson N, Kähäri A, Keefe D, Keenan S, Kinsella R, Kokocinski F, Kulesha E, Lawson D, Longden I, Megy K, Meidl P, Overduin B, Parker A, Pritchard B, Rios D, Schuster M, Slater G, Smedley D, Spooner W, Spudich G, Trevanion S, Vilella A, Vogel J, White S, Wilder S, Zadissa A, Birney E, Cunningham F, Curwen V, Durbin R, Fernandez-Suarez XM, Herrero J, Kasprzyk A, Proctor, G, Smith J, Searle S, Flicek, P (2009) Ensembl 2009. *Nuc Acids Res.* 37 Database issue:D690–D697. doi:10.1093/nar/gkn828
- Ivanov II, Link JM, Ippolito GC, Schroeder HW Jr (2002) Constraints on hydrophobicity and sequence composition of HCDR3 are conserved across evolution. In: Zanetti M Jr, Caprar JD (eds) *The antibodies*, vol 7. Taylor and Francis Group, London, pp 4–67
- Johansson J, Aveskogh M, Munday B, Hellman L (2002) Heavy chain V region diversity in the duck-billed platypus (*Ornithorhynchus anatinus*): long and highly variable complementary determining region 3 compensates for limited germline diversity. *J Immunol* 168:5155–5162
- Jones KE, Purvis A, MacLarnon A, Bininda-Emonds ORP, Simmons NB (2002) A phylogenetic supertree of the bats (Mammalia: Chiroptera). *Biol Rev* 77:223–259
- Kabat EA, Wu TT, Perry HM, Gottesman KS, Foeller C (1991) Sequences of proteins of immunological interest, 5th edn. USDHHS Public Health Service, National Institute of Health, New York
- Kirkham PM, Mortari F, Newton JA, Schroeder HW (1992) Immunoglobulin VH clan and family identity predicts variable domain structure and may influence antigen binding. *EMBO J* 11:603–609
- Knight KL (1992) Restricted VH gene usage and generation of antibody diversity in the rabbit. *Annu Rev Immunol* 10:593–616
- Kumar S, Tamura K, Nei M (2004) MEGA3: integrated software for molecular evolutionary genetics analysis and sequence alignment. *Brief Bioinform* 5:150–163
- Langman RE, Cohn M (1987) The E–T (elephant–tadpole) paradox necessitates the concept of a unit of B cell function: the protection. *Mol Immunol* 24:675–697
- Leroy EM, Kumulungui B, Pourrut X, Bouquet P, Hassanin A, Yaba P, Délicat A, Paweska JT, Gonzalez J-P, Swanepoel R (2005) Fruit bats as reservoirs of Ebola virus. *Nature* 438:575–576
- Leroy EM, Epelboin A, Mondonge V, Pourrut X, Gonzalez J-P, Muyembe-Tamfum J-J, Formenty P (2009) Human ebola outbreak resulting from direct exposure to fruit bats in Luebo, Democratic Republic of Congo, 2007. *Vector Borne Zoonotic Dis* 9:723–728. doi:10.1089= vbz.2008.0167
- Mainville CA, Sheehan KM, Klamann LD, Giorgetti CA, Press JL, Brodeur PH (1996) Deletional mapping of fifteen mouse VH gene families reveals a common organisation for three Igh haplotypes. *J Immunol* 156:1038–1046
- Mansikka A, Toivanen P (1991) D-D recombination diversifies the CDR3 region of chicken immunoglobulin heavy chains. *Scand J Immunol* 33:543–548
- McKoll KA, Chamberlain T, Lunt RA, Newberry KM, Middleton D, Westbury HA (2002) Pathogenesis studies with Australian bat lyssavirus in grey headed flying foxes (*Pteropus poliocephalus*). *Aust Vet J* 80:636–641
- McCormack WT, Tjoelker LW, Thompson CB (1991) Avian B cell development: generation of an immunoglobulin repertoire by gene conversion. *Annu Rev Immunol* 9:219–241
- Mian IS, Bradwell AR, Olson AJ (1991) Structure, function and properties of antibody binding sites. *J Mol Biol* 217:133–151
- Nei M, Gu X, Sitnikova T (1997) Evolution by the birth and death process in multigene families of the vertebrate immune system. *Proc Natl Acad Sci USA* 94:7799–7806
- Nowak RM (1994) Walker’s bats of the world. The Johns Hopkins University Press, Baltimore
- Ota T, Nei M (1994) Divergent evolution and evolution by the birth and death process in immunoglobulin VH gene families. *Mol Biol Evol* 11:469–482
- Padlan EA (1994) Anatomy of the antibody molecule. *Mol Immunol* 31:169–217
- Radic MZ, Mackle J, Erikson J, Mol C, Anderson WF, Weigert M (1993) Residues that mediate DNA binding of autoimmune antibodies. *J Immunol* 150:4966–4977
- Rajewsky K, Forster I, Cumano FA (1987) Evolutionary and somatic selection of the antibody repertoire in the mouse. *Science* 238:1088–1094
- Ravetch JV, Siebenlist U, Korsmeyer S, Waldmann T, Leder P (1981) Structure of the human immunoglobulin  $\mu$  locus: characterization of embryonic and rearranged J and D genes. *Cell* 27:583–591
- Reynaud C-A, Dahan A, Anquez V, Weill J-C (1989) Somatic hyperconversion diversifies the single VH gene of the chicken with a high incidence in the D region. *Cell* 59:171–183
- Roost H-P, Bachmann MF, Haag A, Kalinke U, Pliskat V, Hengartner H, Zinkernagel RM (1995) Early high-affinity neutralizing antiviral IgG responses without further overall improvements of affinity. *Proc Natl Acad Sci USA* 92:1257–1261
- Saitou N, Nei M (1987) The neighbor joining method: a new method for reconstructing phylogenetic trees. *Mol Biol Evol* 4:406–425
- Schelonka RL, Tanner J, Zhuang Y, Gartland GL, Zemlin M, Schroeder HW (2007) Categorical selection of the antibody repertoire in splenic B cells. *Eur J Immunol* 37:1010–1021
- Simmons NB (2005) Order chiroptera. In: Wilson DE, Reeder DM (eds) *Mammal species of the world: a taxonomic and geographic reference*, 3rd edn. John Hopkins University Press, Baltimore, pp 312–529
- Sitnikova T, Su C (1998) Coevolution of immunoglobulin heavy and light chain variable region gene families. *Mol Biol Evol* 15:617–625

- Stenvik J, Lundback AS, Jorgensen TO, Pilstrom L (2000) Variable region diversity of the atlantic cod (*Gadus morhua* L.) immunoglobulin heavy chain. *Immunogenetics* 51:670–680
- Sulkin SE, Allen R, Sims R, Singh KV (1966) Studies of arthropod-borne virus infections in Chiroptera. IV. The immune response of the big brown bat (*Eptesicus f. fuscus*) maintained at various environmental temperatures to experimental Japanese B encephalitis virus infections. *Am J Trop Med Hyg* 15:418–427
- Swanepoel R, Leman PA, Burt FJ, Zachariades NA, Braack LEO, Ksiazek TG, Rollin PE, Zaki SR, Peters CJ (1996) Experimental inoculation of plants and animals with Ebola virus. *Emerg Infect Dis* 2:321–325
- Teeling EC, Springer MS, Madsen O, Bates P, O'Brien SJ, Murphy WJ (2005) A molecular phylogeny for bats illuminates biogeography and the fossil record. *Science* 307:580–584
- Thompson JD, Higgins DG, Gibson TJ (1994) CLUSTAL W: improving the sensitivity of the progressive multiple sequence alignment through sequence weighting, position specific gap penalties and weight matrix choice. *Nucleic Acids Res* 22:4676–4680
- Tomlinson IM, Williams SC, Ignatovich O, Corbett SJ, Winter G (1996) VBASE sequence directory. MRC Centre for Protein Engineering, Cambridge, UK
- Tutter A, Riblet R (1989) Conservation of an immunoglobulin variable region gene family indicates a specific, noncoding function. *Proc Natl Acad Sci* 86:7460–7464
- van Dijk K, Mortari F, Kirkham PM, Schroeder HJ, Milner EC (1993) The human immunoglobulin VH7 gene family consists of a small, polymorphic group of six to eight gene segments dispersed throughout the VH locus. *Eur J Immunol* 23:832–839
- Wellehan JFX, Green LG, Duke DG, Bootorabi S, Heard DJ, Klein PA, Jacobson ER (2009) Detection of specific antibody responses to vaccination in variable flying foxes (*Pteropus hypomelanus*). *Comp Immunol Microbiol* 32:379–394
- Williamson MM, Hooper PT, Selleck PW, Gleeson LJ, Daniels PW, Westbury HA, Murray PK (1998) Transmission studies of Hendra virus (equine morbillivirus) in fruit bats, horses and cats. *Aus Vet J* 76:813–818
- Williamson MM, Hooper PT, Selleck PW, Westbury HA, Slocombe RF (2000) Experimental Hendra virus infection in pregnant guinea-pigs and fruit bats (*Pteropus poliocephalus*). *J Comp Pathol* 122:201–207
- Wilson IA, Stanfield RL (1993) Antibody–antigen interactions. *Curr Opin Struct Biol* 3:113–118
- Wong S, Lau S, Woo P, Yuen K-Y (2006) Bats as a continuing source of emerging infections in humans. *Rev Med Virol* 17:67–91
- Xu JL, Davis MM (2000) Diversity in the CDR3 region of VH is sufficient for most antibody specificities. *Immunity* 13:37–45
- Zemlin M, Klinger M, Link J, Zemlin C, Bauer K, Engler A, Schroeder HW, Kirkham PM (2003) Expressed murine and human CDR-H3 intervals of equal lengths exhibit distinct repertoires that differ in their amino acid composition and predicted range of structures. *J Mol Biol* 334:733–749