ORIGINAL PAPER

Sequencing of *Sylvilagus VDJ* genes reveals a new *VHa* allelic lineage and shows that ancient *VH* lineages were retained differently in leporids

Ana Pinheiro · José Melo-Ferreira · Joana Abrantes · Nicola Martinelli · Antonio Lavazza · Paulo C. Alves · Christian Gortázar · Pedro J. Esteves

Received: 30 June 2014/Accepted: 16 September 2014/Published online: 30 September 2014 © Springer-Verlag Berlin Heidelberg 2014

Abstract Antigen recognition by immunoglobulins depends upon initial rearrangements of heavy chain *V*, *D*, and *J* genes. In leporids, a unique system exists for the *VH* genes usage that exhibit highly divergent lineages: the *VH*a allotypes, the *Lepus* sL lineage and the *VH*n genes. For the European rabbit (*Oryctolagus cuniculus*), four *VH*a lineages have been described, the a1, a2, a3 and a4. For hares (*Lepus* sp.), one *VH*a lineage was described, the a2L, as well as a more ancient sL lineage. Both genera use the *VH*n genes in a low frequency of their *VDJ* rearrangements. To address the hypothesis that the *VH* specificities could be associated with different environments, we sequenced *VDJ* genes from a third leporid

A. Pinheiro · J. Melo-Ferreira · J. Abrantes · P. C. Alves · P. J. Esteves (⊠)

CIBIO Centro de Investigação em Biodiversidade e Recursos Genéticos, InBio Laboratório Associado, Universidade do Porto, Campus Agrário de Vairão, 4485-661 Vairão, Portugal e-mail: pjesteves@cibio.up.pt

A. Pinheiro · P. C. Alves · P. J. Esteves Departamento de Biologia, Faculdade de Ciências, Universidade do Porto, 4169-007 Porto, Portugal

A. Pinheiro · C. Gortázar SaBio - IREC (CSIC-UCLM-JCCM), Ronda de Toledo s/n, 13071 Ciudad Real, Spain

N. Martinelli · A. Lavazza

Istituto Zooprofilattico Sperimentale della Lombardia e dell'Emilia Romagna (IZSLER), Via Antonio Bianchi 7/9, 25124 Brescia, Italy

P. C. Alves

Wildlife Biology Program, College of Forestry and Conservation, University of Montana, Missoula, MT 59812, USA

P. J. Esteves

Centro de Investigação em Tecnologias da Saúde, IPSN, CESPU, 4585-116 Gandra, Portugal

genus, *Sylvilagus*. We found a fifth and equally divergent *VH*a lineage, the a5, and an ancient lineage, the sS, related to the hares' sL, but failed to obtain *VH*n genes. These results show that the studied leporids employ different *VH* lineages in the generation of the antibody repertoire, suggesting that the leporid *VH* genes are subject to strong selective pressure likely imposed by specific pathogens.

Keywords $IGHV \cdot$ Leporids \cdot Antibody diversification \cdot VH allotypes \cdot Evolution

Introduction

For the generation of the primary antibody repertoire, most mammals express a varying number of rearranged VDJ genes (Flajnik 2002). The European rabbit (Oryctolagus cuniculus) has a unique system in that, despite having more than 200 VH genes (Ros et al. 2004), it predominantly expresses just one IGHV, the most D-proximal VH1 gene, to generate 80-90 % of its antibody repertoire (Knight and Becker 1990; Knight 1992). The immunoglobulin (Ig) molecules derived from the VH1 express the so-called VHa allotypic markers (e.g. Kindt 1975; Margolies et al. 1977), with three serological lineages having been described in the domestic rabbit, the a1, a2 and a3. These are highly divergent (± 20 % amino acid sequence differences), with the allelic specificities of a1 and a2 being correlated with several amino acid differences in framework regions (FR) 1 and 3 (Tonnelle et al. 1983; Mage et al. 1984; Knight and Becker 1990). A fourth equally divergent allotypic lineage, the a4, was described in wild European rabbit Iberian populations (Esteves et al. 2004). The remaining 10-20 % of Ig molecules are encoded by the VHn genes, VHx, VHy and VHz (Kim and Dray 1973; Horng et al. 1976; Roux 1981), that

map at least 100 Kb upstream of *VH1* (Mage et al. 2006) and do not express *VH*a allotype-specific determinants.

Studies on the IGHV locus diversity for other leporids are limited to a few data obtained for some Lepus species. Serological analysis conducted for Lepus americanus showed cross-reaction with rabbit anti-al, anti-a2 and anti-a3 antisera (van der Loo 1987). As for Lepus europaeus and Lepus granatensis, a serological analysis of several populations revealed two phenotypes: partial reaction to anti-a2 antisera and no reaction to any rabbit antiserum (Esteves 2003). Sequencing of rearranged VH genes for Lepus specimens, some of which expressed a2-cross-reacting serum proteins, proved that the a2 polymorphism is trans-specific. The Lepus a2L lineage showed some of the most outstanding characteristics of the rabbit VHa genes such as the presence of five out of 11 amino acid residues that characterize the allotype VHa2 (Mage et al. 1984), and being more closely related to the rabbit VH1-a2 allele than VH1-a2 was to its allelic counterparts VH1-a1 and VH1-a3 (Esteves et al. 2005). A second lineage, the sL, apparently a Lepus-specific lineage, groups distinctively apart from all other VH lineages present in rabbit, showing as distinctive characters the ancestral $_{70}$ SVK₇₂ motif and residues A10 and K95 (IMGT numbering, Lefranc et al. 2003) (Esteves et al. 2005). More recently, it was shown that, like rabbits, Lepus uses the VHn genes in 5-10 % of its VDJ rearrangements revealing that the antibody repertoire is subject to selective pressure and that the low-frequency usage of VHn genes could be a remnant of an ancient leporid immunologic response to pathogens (Pinheiro et al. 2011, 2013).

Considering that evidence suggests that the leporid VH lineages used to generate the antibody repertoire are subject to selective pressure, in this study, we addressed the hypothesis that the VH lineages represented in the antibody repertoire could be different in species that inhabit different environments due to exposure to different pathogens. Family Leporidae comprises 11 genera, among which Sylvilagus occupies an intermediate position between Oryctolagus and Lepus (Alves and Hackländer 2008), having diverged from Oryctolagus 10 million years ago and from Lepus 12 million years ago (Matthee et al. 2004). Oryctolagus and the two Lepus species, L. europaeus and L. granatensis, for which VDJ genes nucleotide sequences have been obtained, inhabit the Eurasian continent while Sylvilagus is a Native American leporid (Chapman et al. 1980; Flux and Angermann 1990). Starting in 1966, eastern cottontails were introduced in northern Italy through a series of massive releases for hunting purposes; nowadays, eastern cottontails are widespread in this region (Vidus-Rosin et al. 2008). Herein, we sequenced Sylvilagus floridanus VDJ-rearranged genes.

Material and methods

Samples, amplification and sequencing of rearranged VDJ genes

Spleen samples were collected from four eastern cottontail (*S. floridanus*) specimens collected in Pistoia Province, Italy. Total RNA was extracted using RNeasy Mini Kit (Qiagen, Hilden, Germany), following first-strand complementary (c)DNA synthesis using the iScript cDNA Kit (Bio-Rad, Hercules, CA, USA). *VDJ*-rearranged genes were PCR amplified using primers VH (5'GGAGACTGGGCTGCGCTG GCTTCTCCTGGT3'; Esteves et al. 2005) and JH2 (5'TGAG GAGACGGTGACCAGGGTGCCT3'; Pinheiro et al. 2013). PCR amplification conditions were as follows: 4 min at 95 °C followed by 35 cycles at 95 °C (45 s), 62 °C (45 s) and 72 °C (60 s), with a final extension at 72 °C (5 min). PCR products were purified (MinElute PCR Purification Kit, Qiagen, Hilden, Germany) and cloned into the pGEM-T Easy vector system II (Promega, Madison, WI, USA).

Phylogenetic analyses

To perform phylogenetic analyses, available sequences for leporid *VDJ* genes were taken from GenBank. Used sequences include rabbit germ line *VH1* and *VH4* genes of a1, a2 and a3 allotypes; rabbit cDNA *VH* gene sequences representative of allotypes a1, a2, a3, a4.1 and a4.2; rabbit germ line and rabbit and *Lepus* cDNA *VH* gene sequences of *VHx*, *VHy* and *VHz* and *Lepus* cDNA *VH* gene sequences representative of lineages a2L, sLe and sLg. *VH* genes of the human *VH3* family were used as an outgroup. These were aligned with sequences obtained in this study using CLUSTAL W (Thompson et al. 1994) as implemented in BioEdit software (Hall 1999) and the amino acid sequences inferred. Accession numbers for all sequences are given in Table 1.

The computer program jModelTest v2.1.1 (Darriba et al. 2012) was used to assess the fit of our dataset to 88 models of sequence evolution considering the corrected Akaike information criterion (AICc). The phylogenetic relationships were analysed in a Bayesian inference (BI) framework, specifying as optimal mutation model K80+G, retrieved from jModelTest. BI analyses were performed using BEAST v1.7.4 (Drummond and Rambaut 2007). The BEAUti application, part of BEAST package, was used to create the input file for *BEAST. Posterior phylogenies were determined using an uncorrelated lognormal relaxed clock (Drummond et al. 2006) and the Yule tree prior. All parameter priors were set to the defaults. Three independent runs of 50×10^6 generations were performed, sampling trees and parameters every 5×10^3 generations, and concatenated using LogCombiner.

 Table 1
 Vh gene sequence accession numbers for sequences used in the phylogenetic analysis

Sequence identification	Accession number
Rabbit germline sequences	
VH1a1, VH4a1	M93171, M93181
VH1a2, VH4a2	M93172, M93182
VH1a3, VH4a3	M93173, L03846
Vhx, Vhx2	L03846, M19706
Vhy, Vhy2	L03890, L03874
Vhz	AF264469
Rabbit cDNA sequences	
al_Ocun1-Ocun4	AF029933, AF029934, AF029938, AF029940
a2_Ocun1-Ocun4	L03849, L03851, L03853, L03856
a3_Ocun1-Ocun4	AF029923, AF098235, AF264434, AF264435
a4.1_I- IV, a4.2_I- III	AY207979- AY207982, AY208042, AY208047, AY207967
Vhx_Oca, Vhx_Oca2	AY207986, AY208045
Vhy_Oca	AY208006
Hare cDNA sequences	
sLe1_397, sLe1_400, sLe2_408, sLe2_412, Le8.1, Le8.12	AY288451, AY288453, AY288459, AY288462, KF460076, KF460085
sLg15_640-642, sLg15_644	AY288464- AY288467
a2_Le1.396, a2_Le1.401, a2_Lg19C83, a2_Lg19C86, Le5.13, Le6.1, Le8.16	AY288450, AY288454, AY288491, AY288494, KF460042, KF460056, KF460087
VHx Le5.2, Le5.17,	KF460034, KF460046
VHy-VHz Le6.17, Le6.24, Le8.11	KF460070, KF460075, KF460084,
Human VH3 family sequences	
Hsap1- Hsap4	M99666, M99672, M99679, M99682
Current study	
Sfl1	KM275500- KM275522
Sfl2	KM275523- KM275542
Sfl3	KM275543- KM275568
Sfl4	KM275569- KM275585

Convergence was checked using Tracer v1.5 (Rambaut and Drummond 2007), and summary trees were generated with TreeAnnotator v1.5.4, part of the BEAST package.

Genetic distances between VH genes among species were calculated using the "compute net average distance between groups" option of MEGA5.0 software (Tamura et al. 2011). This option corrects for variance due to differences within groups, as expected in VH-rearranged genes which are somatically diversified.

Results

A total of 86 VDJ gene sequences were obtained for the four studied S. floridanus individuals. Phylogenies were estimated with the whole VH sequence and excluding the CDRs, yielding similar clustering of the analysed sequences. Thus, we present here the phylogeny obtained with the whole VH sequence (Fig. 1). Clusters of the previously described rabbit and hare lineages are shown with good support on the phylogenetic tree: the VHn lineage (1.00 posterior probability), the sL lineage (1.00 posterior probability) and the VHa cluster (0.96 posterior probability) that includes the VHa1, VHa2, VHa2L, VHa3 and VH4 lineages (1.00, 1.00, 0.99, 1.00 and 1.00 posterior probabilities, respectively). The S. floridanus sequences obtained in this study fall into two groups that include only Sylvilagus VH sequences: (i) a cluster that is not related to any previously described leporid lineages, the sS lineage (1.00 posterior probability), and (ii) a second group that clusters with the rabbit VHa lineages, designated here as a5 lineage (1.00 posterior probability). This a5 lineage is further subdivided into two groups, the a5.1 and a5.2 variants (0.98 and 0.95 posterior probabilities, respectively).

Of the 86 obtained sequences, 46 were classified as a5.1, 26 were classified as a5.2 and 14 sequences were classified as sS. These three groups appear to have an allelic distribution as each one of the analysed individuals has sequences belonging to two of the groups: the Sf11 and Sf14 specimens have sequences that were classified as a5.1 and sS (13–9 and 11–5) and the Sf12 and Sf13 specimens have sequences that were classified as a5.1 and a5.2 (2–17 and 17–9).

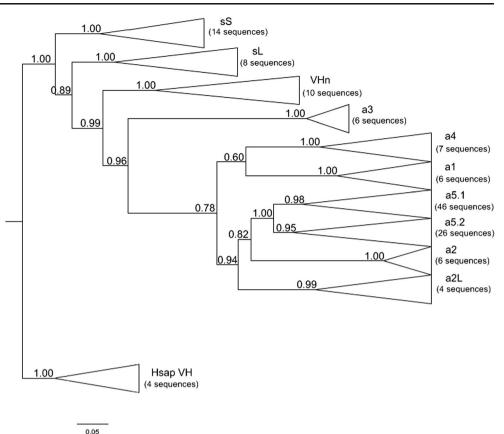
Genetic distances between rabbit, hares and cottontail VH genes

The genetic distances between *S. floridanus* sS lineage and the a1–a4 lineages (0.12–0.16, Table 2) are identical to those between the hares sL lineage and the *VH*a1–*VH*a4 lineages (0.10–0.16, Table 2) or between the *VH*n and the *VH*a1–*VH*a4 lineages (0.08–0.15, Table 2). The *S. floridanus* a5 lineage genetic distances to other *VH*a lineages (0.06–0.10, Table 2) are identical to the genetic distances observed between *VH*a1 and *VH*a4 lineages (0.07–0.14, Table 2). The genetic distance between the a5.1 and a5.2 variants is lower than the genetic distance between a4.1 and a4.2 variants (0.03 and 0.06, respectively).

Characterization of expressed VH in S. floridanus

The inferred amino acid sequences were grouped according to the phylogenetic clustering and compared to *Lepus* cDNA and rabbit germ line and cDNA *VH* genes representative of *VH*n and *VH*a allotypes as well as sequences representative of lineages a2L and sL (Fig. 2).

Fig. 1 Phylogenetic tree of leporid VH genes. Human VH clan 3 genes were used to root the tree. Groups were collapsed for simplification; the number of sequences in each group is indicated. The a1, a2, a3 and a4 groups include only European rabbit sequences, the a2L and sL groups include only hare sequences and the a5.1, a5.2 and sS groups include only S. floridanus sequences. The VHn group includes European rabbit and hares sequences. BI posterior probabilities are depicted in front of each node



The rabbit VHa sequences are characterized by the 19LTLTCT24 of FR1 and 70WAK72 of FR3 motifs as well as showing a deletion at position 2 (Mage et al. 1984). The 72 sequences obtained in this study that were classified as VHa have the deletion at position 2, and the majority either has the characteristic 70 WAK72 or a 70 WAQ72 motif. The 19 LTLTCT24 of FR1 is also present though the majority of the sequences have a D_{20} substitution. These sequences further have as characteristics the motif 83SAG85 and residue D96, while also sharing the VHa2 I/F13 and T79 residues. Thus, these sequences have the rabbit VHa hallmark residues but have evolved some particular features which substantiate their inclusion in a new VHa lineage that we designate as a5. According to the phylogenetic tree, this a5 lineage has two variants, the a5.1 and a5.2, and amino acid differences exist between them that support their distinction. The 19LTLTCT24 motif is more conserved in the sequences assigned as a5.2 while those that were classified as a 5.1 have the D_{20} as well as K₂₄ substitutions. In FR3, the ₈₃SAG₈₅ D₉₆ motif is more conserved in the sequences that were classified as a5.1 while those identified as a5.2 have K₈₀, L₈₇, G₉₃ and E₉₇ residues that, although not a characteristic to this variant, are not present in the a5.1 variant.

The 14 sequences that were classified as a new lineage (sS) have the ancestral $_{70}$ SVK $_{72}$, also present in the hares sL lineage, and the $_2$ EQ₃ motif of rabbit *VHx* peptides. The

FR3 amino acid sequence between residues 83 and 101 has nine substitutions that are characteristics to this lineage: ⁸³DA₈₄, ⁸⁷LE₈₈, D₉₃, ⁹⁵KTT₉₇ and I₁₀₁.

Discussion

The generation of antibody repertoire in the European rabbit has been widely studied, and several particularities have been found. It seems to rely on a reduced number of VH genes: the VH1, which expresses the so-called VHa allotypic markers, and the VHx, VHy and VHz genes, or VHn genes, which do not have the VHa allotype specificities (Mage et al. 1984; Knight and Becker 1990). For the European rabbit, four highly divergent allelic lineages have been described for the VHa allotypes: a1, a2, a3 and a4 (Mage et al. 1984; Knight and Becker 1990; Esteves et al. 2004). In hares (Lepus sp.), the existence of a VHa lineage, a2L, with nucleotide and amino acid sequence similarity with the rabbit a2 lineage and which shows cross-reaction with sera against the rabbit a2 lineage has showed the trans-specific nature of the VHa2 polymorphism (Esteves et al. 2005). A more ancestral lineage, sL, presumably also derived from the VH1 gene, was found to be expressed in hares (Esteves et al. 2005). More recently, Pinheiro and co-workers (2013) showed that hares and rabbits

Groups

VHn sL sS VHa

a1 a2 a2L a3 a4

a5

	VHn	sL	sS	VHa									
				al	a2	a2L	a3	a4			a5		
								a4.1	a4.2		a5.1	a5.2	
		0.01	0.01	0.02	0.02	0.01	0.01	0.02	0.02	0.02	0.01	0.01	0.01
	0.07		0.01	0.02	0.02	0.02	0.01	0.02	0.02	0.02	0.02	0.02	0.01
	0.09	0.05		0.02	0.02	0.02	0.02	0.02	0.02	0.02	0.01	0.02	0.01
	0.10	0.11	0.12		0.02	0.01	0.01	0.01	0.01	0.02	0.01	0.01	0.01
	0.15	0.16	0.16	0.11		0.01	0.02	0.02	0.02	0.02	0.01	0.01	0.01
	0.08	0.10	0.14	0.07	0.08		0.02	0.01	0.02	0.02	0.01	0.01	0.01
	0.08	0.10	0.13	0.08	0.14	0.10		0.02	0.02	0.02	0.01	0.01	0.01
	0.12	0.11	0.14	0.08	0.13	0.09	0.11		_	_	0.01	_	-

0.10

0.15

0.10

0.11

0.10

0.09

_

Genetic distances between groups of sequences, calculated using the p-distance, are shown below the diagonal. Standard error estimates are shown above the diagonal in italics. All ambiguous positions were removed for each sequence pair

0.09

0.11

0.06

0.07

0.06

0.13

0.16

0.08

0.09

0.08

0.08

0.12

0.07

0.09

0.07

share the low-frequency usage of the VHn genes in the generation of the antibody repertoire.

0.12

0.16

0.10

0.11

0.11

0.11

0.15

0.11

0.12

0.11

0.14

0.19

0.12

0.13

0.12

a4.1

a4.2

a5.1

a5.2

In this study, we observed VH-expressed genes for S. floridanus that have the rabbit VHa signature, lineage a5, as well as VH-expressed genes that belong to a more ancestral lineage, sS, that seems to be S. floridanus specific. In the phylogenetic tree, the S. floridanus a5 sequences group with rabbit and hare VHa lineages, clearly apart from the S. floridanus sS sequences, and the genetic distances also indicate that the a5 lineage is more similar to the VHa lineages than to the sS lineage. These results suggest that the VHa allotypes have an origin that pre-dates the leporid radiation. Some VHa lineages originated before the species split and are shared between species, such as the a2 and a2L lineages, whereas others apparently are species specific. As Esteves et al. (2004) argued, it is likely that evolutionary modes, and mutational rates, vary among VHa lineages having increased in the a1, a2 and a4 lineages, which would explain the large inter-lineage distances despite a more recent origin. An interesting aspect resides in the differences in number of VHa lineages found in each leporid species studied so far. Apparently, the rabbit has retained a greater diversity of VHa lineages, with at least four lineages (Mage et al. 1984; Knight and Becker 1990; Esteves et al. 2004), while hares and cottontail have just one lineage each (Esteves et al. 2005; Pinheiro et al. 2013; present study). This higher number of VHa lineages found in rabbit can be the result of a compensatory mechanism due to the loss of a more ancient lineage of VH genes, which would be the rabbit equivalent of the hares sL and cottontail sS. These ancient lineages, characterized by the motif $_{70}$ SVK $_{72}$ and more highly divergent from the VHa lineages, can importantly contribute to the diversity of the antibody repertoire. Its loss may have imposed the diversification of the *VHa* lineages in rabbit. Contrarily to what is widely accepted, increasing evidence suggests the *VH* lineages which are rearranged in different leporid species are under different selective pressures, either driven by pathogens or by the self, though it is not evident how this selection operates (Pinheiro et al. 2013). It has also been shown that in rabbit, the diversification of the antibody repertoire requires the interaction with exogenous factors such as the gut microbiota (Lanning et al. 2000a, b; Sehgal et al. 2002), which may also exert selective pressure on the expressed *VH* lineages.

0.01

0.13

0.13

0.06

0.11

0.09

0.02

0.02

0.03

0.02

0.02

0.01

Sequencing of random germ line VH genes has shown that VHn genes are present in S. floridanus genome (Esteves 2003); thus, we hypothesized that, similarly to what was found for rabbit and hares (Pinheiro et al. 2013), S. floridanus would use these genes in a low percentage of VDJ rearrangements. Interestingly, despite the identical sequencing effort made for Lepus and Sylvilagus rearranged VH (circa 25 sequences per individual; Pinheiro et al. 2013; present study), we failed to recover rearranged VDJ genes with characteristics of rabbit and hares VHn genes in the studied S. floridanus specimens. S. floridanus may not use the VHn genes in the generation of the antibody repertoire or may do so in such a low percentage (under 4 %) that the sequencing effort would have to be greatly increased to allow its detection. The usage of VHn genes in low frequency both in rabbit (Kim and Dray 1973; Roux 1981) and hares shows that, in leporids, the VH lineages used to generate the antibody repertoire are subject to selective pressure (Pinheiro et al. 2013). The rabbit and hare species for which the expression

•	FR1-I (1-2		CDR1-IMGT (27-38)	FR2-IMGT (39-55)	CDR2-IMGT (56-85)		FR3-IMGT (66-104)
a Phylogenetic	A (1-15)	B (16-26)	BC (27-38) —	(39-46) C' (39-46) (47-5	C'C"	C" D (66-74) (75-84)	(85-96) E F (85-96) (97-104)
groups	1 10 	15 16 23 2 			55 56 65 66	74 75 	84 85 89 96 97 104
VH1a1 VH1a2 a2L	K E G .	VTPGTPLTLTCTV FK.TDT	SGFSLS SYAMS N . I .		A . G A		
VH1a3 a4.1	L D .	. К А S А . К G T	FSY.C	 ЕЕК	A C . Y A G S S G		ST.QML.AA T.T.ALEPSG
a4.2 a5 a5.1	K D .	. K G T	V.	. I VP . EER . A	YYYG DAA		T.TNLQPSG
SfI1-2 SfI1-4 SfI1-6		FK.EET.DK. FK.EET.DK.			Y.F.N.GA Y.F.NAGA Y.AY.NTGA		S A G T M . D L . D A
Sfl1-7 Sfl1-11 Sfl1-13		IKGT.D.SA IKGT.DP. FK.EET.DK		Q'	Y. L.NGGGA.	N S . T TRGT Q T TRST A ARGT	SAG
Sfl1-17 Sfl1-18	TG.	IKGSM IKGT.D		- ⁻ -	Y Y TG A Y FGG . D F C .	Q. A. TRST N	SAGTML.DA
Sfl1-21 Sfl1-25 Sfl1-26		IKGT.D FK.EET.DK. IKGT.D			Y Y . D I G A Y Y Y F . T Y R A D		SAGTML.DA
Sfl1-27 Sfl1-29 Sfl1-30		I K GT . D . S A FK . EET . D K . FK . EET . D R			Y. Y. DIVH Y. F. TNGY. WF.	Q T TRST S. T ARST RST	SAGTML.DA SAGTML.DA AGTML.DA
Sfl2-11 Sfl2-24	L T	I K GT . D I K GT		.	YF.N.DNA FF.N.DNAN	T	S A G T . T M L . V A
Sfl2-26 Sfl3-1 Sfl3-3		FK.EET.DK. FK.EET.DK.			F . H T D A V G C V W . G . G N N . Y A	Q T TRST 	S A G T M L . D A
Sfl3-4 Sfl3-5 Sfl3-8	EQL	IK.GT.D.S.A IK.GT.D.S.A IK.GT.D.P.A		F	Y . F . G G A Y . F . G G A Y . F . G G A		S A G T M L . D A A S A G T M L . D A S A G T M L . D A
Sfl3-9 Sfl3-10 Sfl3-13		IKGT.D IKGT.D			Y		SAGTMQDA SAGTML.DA SAGTM.GL.DV
Sfl3-16 Sfl3-17	L	IK.EET.DR. FK.EET.DK.		. I	Y F . T . G A D R V . T . G R	R. TQ T TRST	S A G T M L . D A
Sfl3-19 Sfl3-20 Sfl3-21	L G . T G . L G .	FK.EET.DK. FK.EET.DK. FK.EET.DK.			Y F . N I G A Y F . D R N P H Y V . F A G T I		S A G T M L . D A . A
Sf13-22 Sf13-24 Sf13-28	L G . . E Q L G . R E Q L G .	FK.EET.DK. IK.GT.D.S.A IK.GT.D.S.A	I D F N H S T G V A T G V A	. T	Y V . F . G V G R Y F . G G A Y F . G G A		SAG
Sfl3-29 Sfl4-1 Sfl4-4	. EQL	IKGT.D.SA FK.EET.D FK.EKT.DK.		. F	Y F . G G A Y F A D Y F . G . A G . R . S	PQ T TRST Q T TRST 	S A G T M L . D A
Sf14-7 Sf14-8 Sf14-9	L T G . 	IKGT FK.EET.DK. FK.EET.DK.	I I T . S I A I D N . I .		Y.F.N.GN.		S A G T M L . D A V
Sfl4-12 Sfl4-17		FK.EET.DK. FK.EET.DK.	N Y		Y F . G . A G . W . S Y F . G . A G . W . S	T T R N T :	S A G T M L . D A
Sfl4-18 Sfl4-27 Sfl4-28		. QT . GT . D A FK . EET . D K FK . EET . D K .	NKHY 	·	Y F . R D D W . Y F . G . A G . W . S Y . N Y A		S A G T V L . D A
Sfl4-29 Sfl4-30		. QT . GT . D A . QT . GT <u>.</u> D A	NKHY	· [•	YF.RDDW YF.RDDW.	T T TRST T	S A G T M L . D A
b		MGT	CDR1-IMGT	FR2-IMGT	CDR2-IMGT		FR3-IMGT
	(1- A (1-15)	IMGT 26) B (16-26)	CDR1-IMGT (27-38) BC (27-38)	FR2-IMGT (39-55) C C (39-46) (47-5	(56-65) C'C"	C" D (66-74) (75-84)	FR3-IMGT (66-104) E F (85-96) (97-104)
Phylogenetic groups	(1- A (1-15) 1 10	26) B	(27-38) BC (27-38) →	(39-55) C C (39-46) (47-5	(56-65) C'C"		(66-104) E F
Phylogenetic groups VHa VH1a1 a5.2	(1- A (1-15) 1 10 1	26) B (16-26) 15 16 23 2	(27-38) BC (27-38) — ≥6 27 38 39 ° ° .	(39-55) C C' (39-46) (47-5)	(56-65) C'C" 55) (56-65)	(66-74) (75-84) 74 75 A SWAKGRFTISK	(66-104) E (85-96) 84 85 89 96 97 104
Phylogenetic groups VHa VH1a1 a5.2 Sfl2-4 Sfl2-5 Sfl2-6	(1- A (1-15) 1 10 1	26) B (16-26) 15 16 23 :	(27-38) BC (27-38) — ≥6 27 38 39 ° ° .	(39-55) C C' (39-46) (47-5)	(56-65) CCC (55) (56-65) 55 56 65 66 ° ° .	(66-74) (75-84) 74 75 A SWAK G R F T I S K Q A T . S T L A TR S T Q A T . S T	(66-104) E F (85-96) 84 85 89 96 97 104
Phylogenetic groups VHa VH1a1 a5.2 Sfl2-4 Sfl2-5	(1- A (1-15) 1 10 1	26) B (16-26) 15 16 23 : 	(27-38) BC (27-38) — ≥6 27 38 39 ° ° .	(39-55) C C' (39-46) (47-5)	(56-65) CC° 55) (56-65) 55 56 65 66 ° ° . WI GI I SSSG ST Y Y. Y. V. AT T. F. GL DR D LSS.	(66-74) 74 75 0	(66-104) E (85-96) B 4 45 B 4 45 B 4 45 B 4 45 B 4 45 B 7 104 C
Phylogenetic groups VHa VH1a1 a5.2 Sfl2-4 Sfl2-5 Sfl2-6 Sfl2-12 Sfl2-12 Sfl2-13 Sfl2-16 Sfl2-16 Sfl2-17	(1- A (1-15) 1 10 Q SVEESGGR LG L	B (16-20) 15 16 23 : LV T P G T P LT L T C T V F P . N T IK . N T F K . N T . K . N T . K . N T . K . N T . K . N T	(27-38) BC (27-38) BC (27-38) SG F S L S S Y A M S N Y - N Y - 	C (39-46) (39-46) (41-46 47 - 1	(56-65) (56-65) 55 56 65 66 ° ° . WI G I I S S G G S T Y Y Y . V . A T T . F GL D R D L S S Y . Y G A G I M T A N Y N G G A N Y G G A N Y D G G A M T . G G	(66-74) 74 75 A SWAK G R F T I S K A SWAK G R F T I S K C A A - T S T - Q A - T S T - Q A - T S T - S T S T - S T S T S T S T S T S T - S T S T S T S T S T S T S T S T S T S	E F 84 85 89 96 97 104 ST V L - RM 6L - ST V L - RM GL - - SA G T. RM L A - - - ST V L - RM GL A - - - ST V L - RM GL A - <
Phylogenetik groups VHa VH1a1 Sf12-4 Sf12-6 Sf12-13 Sf12-13 Sf12-13 Sf12-13 Sf12-13 Sf12-13 Sf12-13 Sf12-13 Sf12-13 Sf12-23 Sf12-22 Sf12-22 Sf12-22	(1- A (1-15) 1 10 0 - SVEESGGR - L	B (16-26) 15 16 23	(27-38) BC (27-38) 26 27 38 39 1 ° ° 1 S G F S L S S Y A M S N Y N Y N Y N Y N Y 	C (39-46) (39-46) (41-4 41 46 47 	(56-65) (CC" 55) (56-65) 55 56 65 66 ° ° . WI G I I S S G S T Y Y. Y V . A T T. F. 	(66-74) 74 75 75 74 75 76 75 77 75 76 75 77 75 77 75 77 75 78 77 75 79 75 70 74 75 70 75 71 75 71 75 71 75 71 75 74 75 74 75 74 75 74 75 74 75 75 75 75 75 75 75 76 75 76 75 77 75 77 75 77 75 77 75 78 75	(66-104) E F (85-96) (97-104) 84 65 89 96 97 104 I I I TSTTVDLK ITSPTTEDTATYFCA STV.LRM. GL.A
Phylogenetik groups VHa VH1a1 3 Sil2-4 Sil2-4 Sil2-2 Sil2-3 Sil2-2 Sil2-3 Sil	(1- A (1-15) 1 10 Q SVEESGGR 	B (16-26) 15 16 23 : U T PG T P L L L T C T V F P . N T I K . N T K . N T I K . G T I K . N T I K . G T I K . N T I K . G	(27-38) BC (27-38) I	C (39-55) C (39-46) (41-4) (47-4)	(56-65) CCC 55) (56-65) (66-65) VIGISSGSTYY V.V.ATTF (GDRDLSS V.Y.GAGI V.Y.GAGI V.Y.GG V.TG V.Y.GG V.TG V.Y.GG V.TG V.Y.GG V.	(66-74) 74 75 74 75 74 75 75 74 75 74 75 75 74 75 74 75 7	(66-104) E F (85-96) (97-104) 84 85 89 96 97 104 I I I T S TTV DLK ITS PTTE DT A TY FCA STV L RM . GL . A STV L RM . GL . A SAG STV L M . GL . A STV STV L - M . GL . A STU L - M . GL . A STV L - M . GL . A STV L - M . GL . A STV L - M . GL . A
Phylogenetik groups VHa VH1a1 a5.2 Sfl2-4 Sfl2-5 Sfl2-15 Sfl2-15 Sfl2-16 Sfl2-16 Sfl2-16 Sfl2-17 Sfl2-20 Sfl2-22 Sfl2-23 Sfl2-25 Sfl2-28	(1- A (1-15) 1 10 	B (16-20) 15 16 23 : U T P G T P L T L T C T V F P . N T IK . N T K . N T F K . N T K . N T F P . N T	(27-38) BC (27-38) BC (27-38) BC SG F S L S SYA MS SG F S L S SYA MS PI I D SG F S L S SYA MS PI I D SS I PI I D SS I PI I D SS I PI I D SS I PI D SS I PI D SS I PI D SS I PI D SS I PI	C (39-46) (39-46) (47-4) (4	(56-65) (56-65) (55) (56-65) (55) (56-65) (56) (56-65) (57) (56) (56) (56) (56) (56) (56) (56) (56	(66-74) 74 75 A SWAK G R F T I S K A SWAK G R F T I S K C A A T R S T - A C A A T R S T - S A T R S T - S A T R S T - S A T R S T - C A A	(66-104) E F (85-96) (97-104) 84 85 89 96 97 104 I I I I I TSTTVDLK ITSPTTEDTATYFCA I I I STV.LRM.GL.A. STV.L-A. STV.L-A. STV.L-STTEDTATYFCA STV.LM.GL.A. STV.L-A. STV.L-A. STV.L-STTEDTATYFCA STV.LM.GL.A. STV.L-A. STV.L-A. STV.L-M.GL.A. STV.L-A. STV.L-A. STV.L-M.GL.A. STV.L-A. STV.L-A. STV.L-M.GL.A. STV.L-A. STV.L-A.
Phylogenetik groups VHa 45.2 Sil2-4 Sil2-4 Sil2-13 Sil2-16 Sil2-16 Sil2-16 Sil2-16 Sil2-16 Sil2-17 Sil2-20 Sil2-23 Sil2-23 Sil2-23 Sil2-23 Sil2-23 Sil2-23 Sil2-23 Sil2-31 Sil2-32 Sil2-32 Sil2-32 Sil2-32 Sil2-32 Sil2-32 Sil2-32 Sil2-32 Sil2-32 Sil2-32 Sil2-32 Sil2-32 Sil2-32 Sil2-32 Sil2-32 Sil2-32 Sil2-32 Sil2-33 Sil3-34 Sil	(1- A (1-15) 1 10 0 - SVEESGGR - LG - L - P G - L - P G - L G G G G G G G G G G 	B (16-20) 15 16 23 : U T P G T P L T L T C T V F P . N T IK . N T K . N T F K . N T K . N T F P . N T	(27-38) BC (27-38) BC (27-38) BC SG F S L S SYA MS SG F S L S SYA MS BC SG F S L S SYA MS BC SG F S L S SYA MS BC SG F S L S SYA MS SC S S S S S S S S S S S S S S S S S S	C (39-46) (39-46) (47-4) (4	(56-65) (56-65) (55) (56-65) (55) (56-65) (56) (56-65) (57) (56) (56) (56) (56) (56) (56) (56) (56	(66-74) 74 75 A SWAK G R F T I S K A SWAK G R F T I S K C A A T R S T - A C A A T R S T - S A T R S T - S A T R S T - S A T R S T - C A A	(66-104) E F (85-96) (97-104) 84 85 89 96 97 104 I I I I I TS TTV DLK ITS PTT EDTATY FCA STV I I STV L RM GL A STV L M GL A STU L M GL A STV L M GL A STV L M GL A STV L M GL A STU
Phylogenetic groups VHe S12-4 S12-4 S12-4 S12-2 S12-15 S12-15 S12-15 S12-15 S12-15 S12-25 S12	(1- A (1-15) 1 10 Q SVEESGG R - L	B (16-26) 15 16 23 1 1 1 1 1 1 V V TP GT P LIT CT V FP N T IK N K N T K N T K N T K N T K G T K G T K G T FP N T FP G T FP N T FP G T FP O T FP O T FP T FP N T IK N T FP N T <td>(27-38) BC (27-38) 26-27 38-39 </td> <td>C (39-55) C (39-46) 41 46 47 </td> <td>(66-65) CCC 55) (66-65) (66-65) (66-65) (66-65) (1</td> <td>(66-74) 74 75 74 75 74 75 75 75 76 75 76 75 76 75 77 75 77 75 70 7 70 70 7 70 70 7 70 7 70 7 70 70 70 70 70 70 70 70 70 70 70 70 70 7</td> <td>(66-104) E F (85-96) (97-104) 84 45 89 96 97 1 TSTTVDLKITSPTTEDTATYFCA STV.LRM.GLA. SAG.T.RM.LA. SAG.T.RM.LA. SAG.T.RM.LA. SAG.T.M.GLA. SAG.T.M.L.DA. STV.LM.GLA. STV.LM.GLA. SAG.T.M.L.DA. STV.LM.GLA. STV.LM.GLA. STV.LM.GLA. STV.LM.GLA. STV.LM.GLA. STV.LM.GLA. STV.LM.GLA. STV.LM.GLA. MGLA. STV.LM.GLA. STV.LM.GLA. MGLA. STV.LM.GLA. MGLA. TV.LM.GLA. MM.LM.GLA. TV.LM.GLA. MM.L.M.GLA. TV.LM.GLA. MGLA. TV.LM.GLA. TV.LM.GLA. TV.LM.GLA. TV.LM.GLA. TV.LM.GLA.</td>	(27-38) BC (27-38) 26-27 38-39 	C (39-55) C (39-46) 41 46 47 	(66-65) CCC 55) (66-65) (66-65) (66-65) (66-65) (1	(66-74) 74 75 74 75 74 75 75 75 76 75 76 75 76 75 77 75 77 75 70 7 70 70 7 70 70 7 70 7 70 7 70 70 70 70 70 70 70 70 70 70 70 70 70 7	(66-104) E F (85-96) (97-104) 84 45 89 96 97 1 TSTTVDLKITSPTTEDTATYFCA STV.LRM.GLA. SAG.T.RM.LA. SAG.T.RM.LA. SAG.T.RM.LA. SAG.T.M.GLA. SAG.T.M.L.DA. STV.LM.GLA. STV.LM.GLA. SAG.T.M.L.DA. STV.LM.GLA. STV.LM.GLA. STV.LM.GLA. STV.LM.GLA. STV.LM.GLA. STV.LM.GLA. STV.LM.GLA. STV.LM.GLA. MGLA. STV.LM.GLA. STV.LM.GLA. MGLA. STV.LM.GLA. MGLA. TV.LM.GLA. MM.LM.GLA. TV.LM.GLA. MM.L.M.GLA. TV.LM.GLA. MGLA. TV.LM.GLA. TV.LM.GLA. TV.LM.GLA. TV.LM.GLA. TV.LM.GLA.
Phylogenetik groups VH a 5.2 Si2-4 Si2-5 Si2-5 Si2-15 Si2-15 Si2-15 Si2-15 Si2-15 Si2-15 Si2-15 Si2-22 Si2-25 Si2-25 Si2-22 Si2-25 Si2-23 Si2-25 Si2-23 Si2-23 Si2-23 Si3-24 Si3-31 Si3-31 Si3-12 Si3-18 Si3-18 Si3-18 Si3-18 Si3-25	(1- A (1-15) 1 10 2 SVEESGGR - LG - L	B (16-26) 15 16 23 1 1 1 1 1 1 V V TP GT P LIT CT V FP N T IK N K N T K N T K N T K N T K G T K G T K G T FP N T FP G T FP N T FP G T FP O T FP O T FP T FP N T IK N T FP N T <th>(27-38) BC (27-38) 26-27 38-39 </th> <th>C (39-55) C (39-46) 41 46 47 </th> <th>(66-65) CCC 55) (66-65) (66-65) (66-65) (66-65) (1</th> <th>(66-74) 74 75 74 75 74 75 75 75 76 75 76 75 76 75 77 75 77 75 70 7 70 70 7 70 70 7 70 7 70 7 70 70 70 70 70 70 70 70 70 70 70 70 70 7</th> <th>(66-104) E F (85-96) (97-104) 84 45 89 96 97 1 TSTTVDLKITSPTTEDTATYFCA STV.LRM.GLA. SAG.T.RM.LA. SAG.T.RM.LA. SAG.T.RM.LA. SAG.T.M.GLA. SAG.T.M.L.DA. STV.LM.GLA. STV.LM.GLA. SAG.T.M.L.DA. STV.LM.GLA. STV.LM.GLA. STV.LM.GLA. STV.LM.GLA. STV.LM.GLA. STV.LM.GLA. STV.LM.GLA. STV.LM.GLA. MGLA. STV.LM.GLA. STV.LM.GLA. MGLA. STV.LM.GLA. MGLA. TV.LM.GLA. MM.LM.GLA. TV.LM.GLA. MM.L.M.GLA. TV.LM.GLA. MGLA. TV.LM.GLA. TV.LM.GLA. TV.LM.GLA. TV.LM.GLA. TV.LM.GLA.</th>	(27-38) BC (27-38) 26-27 38-39 	C (39-55) C (39-46) 41 46 47 	(66-65) CCC 55) (66-65) (66-65) (66-65) (66-65) (1	(66-74) 74 75 74 75 74 75 75 75 76 75 76 75 76 75 77 75 77 75 70 7 70 70 7 70 70 7 70 7 70 7 70 70 70 70 70 70 70 70 70 70 70 70 70 7	(66-104) E F (85-96) (97-104) 84 45 89 96 97 1 TSTTVDLKITSPTTEDTATYFCA STV.LRM.GLA. SAG.T.RM.LA. SAG.T.RM.LA. SAG.T.RM.LA. SAG.T.M.GLA. SAG.T.M.L.DA. STV.LM.GLA. STV.LM.GLA. SAG.T.M.L.DA. STV.LM.GLA. STV.LM.GLA. STV.LM.GLA. STV.LM.GLA. STV.LM.GLA. STV.LM.GLA. STV.LM.GLA. STV.LM.GLA. MGLA. STV.LM.GLA. STV.LM.GLA. MGLA. STV.LM.GLA. MGLA. TV.LM.GLA. MM.LM.GLA. TV.LM.GLA. MM.L.M.GLA. TV.LM.GLA. MGLA. TV.LM.GLA. TV.LM.GLA. TV.LM.GLA. TV.LM.GLA. TV.LM.GLA.
Phylogenetic groups VHe S12-4 S12-4 S12-4 S12-6 S12-15 S12-15 S12-15 S12-15 S12-25 S12-25 S12-25 S12-25 S12-29 S12-20 S12-20 S12-20 S12-21 S12-22 S12-25 S12-21 S12-25 S12-21 S12-25 S13-25 S12-25 S12	(1- A (1-15) 1 10 Q SVEESGGR 	B (16-26) 15 16 23 : V T PG TP LT LT C TV FP NT IK NT ME GT D K. FP NT IK NT FP NT	(27-38) BC (27-38) BC (27-38) BC SG F S L S	(39-46) (39-46) (47-4) (47-	(56-65) (56-65) (55) (55) (56-65) (56-65) (56-65) (56-65) (51) (55) (56-65) (56-65) (56-65) (57	(66-74) (75-84) 74 75 A SWAK G R F T I S K A SWAK G R F T I S K A SWAK G R F T I S K N L A TRS T - Q A T RS T - S T T S T - S S T S S T - S S T S S T - S S S S S S S S S S S S S S S S S S S	(66-104) E F (85-96) (97-104) 84 85 89 96 97 1 1 1 1 1 <t<< th=""></t<<>
Phylogenetic groups VHe S122 S122 S122 S122 S122 S122 S122 S122	(1- A (1-15) 1 10 Q SVEESGGR - L	B (16-26) 15 16 23 : V T PG TP LT LT C TV FP NT IK NT ME GT D FF NT IK NT FP NT	(27-38) BC (27-38) 26-27 38-39 ° ° SG F S L S SY A MS PI (D SG F S L S SY A MS PI (D SG F S L S SY A MS PI (D SG F S L S SY A MS PI (D SG F S L S SY A MS PI (D SG F S L S SY A MS PI (D SG F S L S SY A MS PI (PI	C (39-46) (39-46) (41-4647 	(66-65) CCC 55) (66-65) 55 56 65 66 1 ° ° . WI G I I S S G S T Y Y Y . V . A T T . F G L DR D L S S 	(66-74) 74 75 74 75 A SWAK G R F T I S K A SWAK G R F T I S K N L A T R S T N L A T R S T N L A T R S T S T T S T S T T S T S T T S T C A A T	(66-104) E F (85-96) (97-104) 84 85 89 96 97 104 I I I I I TS TTV DLK ITS PTTE DT ATY FCA STV LRM.GL.A STV STV.LRM.GL.A STV.L-A STV STV.LM.GL.A STV.L-A STV STV.LM.GL.A STV.L-A STV.L-A STV.LM.GL.A STV.L-A STV.L-A SAG.T.TM.L.AA STV.L-A STV.L-A SAG.T.TM.GL.A STV.L-A STV.L-A STV.L-M.GL.A STV.L-A STV.L-A SAG.T.M.GL.A STV.L-A
Phylogenetic groups VHe S12-4 S12-4 S12-4 S12-4 S12-13 S12-16 S12-16 S12-17 S12-22 S12-23 S12-23 S12-23 S12-23 S12-23 S12-23 S12-23 S12-23 S12-23 S12-23 S12-23 S12-23 S13-24 S13-14 S13-14 S13-14 S13-14 S13-14 S13-14 S13-14 S13-14 S13-14 S13-14 S13-14 S13-15 S11-15 S11-15 S11-15	(1- A (1-15) 1 10 0 - SVEESGGR - L	B (16-26) 15 16 23 : 1 <th>(27-38) BC (27-38) 26-27 38-39 </th> <th>C (39-55) C (39-46) 41 4647 </th> <th>(66-65) CCC 55) (66-65) (66-65) (66-65) (66-65) (66-65) (7) WIGIISSGSTY Y.V.ATT.F. GLDRDLSS V.YGAGI. METAN Y.GGA. (7) V.GGA. (7) V.GGA. (7) V.GGA. (7) V.GGA. (7) V.GGA. (7) V.GGA. (7) V.GGA. (7) V.GGA. (7) V.GGA. (7) V.DRVF. (6) N.YN. (7) V.DRVA. (7) V.DRVA. (7) V.DRVA. (7) V.DRVA. (7) V.DRVA. (7) V.DRVA. (7) V.DRVA. (7) V.DRVA. (7) V.DRVA. (7) V.TGA. (7) V.DRVA. (7) V.DRVA. (7) V.DRVA. (7) V.DRVA. (7) V.DRV</th> <th>(66-74) 74 75 75 84) 74 75 74 75 75 84) 74 75 74 75 74 75 75 84 74 75 74 75 74 75 74 75 75 84 75 75 75 75 75 75 75 75</th> <th>(66-104) E F (85-56) (97-104) 84 45 89 96 97 1 . . 1 . . 1 . . 1 . . 1 . . 1 . . 1 . . 1 . . 1 . . 1 . . 1 . . 2 . . 3 . . 3 . . 8 . . 8 . . 5 . . 5 . . 8 . . 8 . . 8 . . 8 . . 1 . . 1 . <td< th=""></td<></th>	(27-38) BC (27-38) 26-27 38-39 	C (39-55) C (39-46) 41 4647 	(66-65) CCC 55) (66-65) (66-65) (66-65) (66-65) (66-65) (7) WIGIISSGSTY Y.V.ATT.F. GLDRDLSS V.YGAGI. METAN Y.GGA. (7) V.GGA. (7) V.GGA. (7) V.GGA. (7) V.GGA. (7) V.GGA. (7) V.GGA. (7) V.GGA. (7) V.GGA. (7) V.GGA. (7) V.DRVF. (6) N.YN. (7) V.DRVA. (7) V.DRVA. (7) V.DRVA. (7) V.DRVA. (7) V.DRVA. (7) V.DRVA. (7) V.DRVA. (7) V.DRVA. (7) V.DRVA. (7) V.TGA. (7) V.DRVA. (7) V.DRVA. (7) V.DRVA. (7) V.DRVA. (7) V.DRV	(66-74) 74 75 75 84) 74 75 74 75 75 84) 74 75 74 75 74 75 75 84 74 75 74 75 74 75 74 75 75 84 75 75 75 75 75 75	(66-104) E F (85-56) (97-104) 84 45 89 96 97 1 . . 1 . . 1 . . 1 . . 1 . . 1 . . 1 . . 1 . . 1 . . 1 . . 1 . . 2 . . 3 . . 3 . . 8 . . 8 . . 5 . . 5 . . 8 . . 8 . . 8 . . 8 . . 1 . . 1 . <td< th=""></td<>
Phylogenetic groups VHe S1224 S1224 S1225 S1225 S12212 S12215 S12215 S12215 S12215 S12225 S12222 S12222 S12223 S12225 S12229 S12231 S12225 S12229 S12231 S12225 S12232 S12232 S12232 S1324 S1324 S1324 S1325 S125 S1325 S13555 S13555 S13555 S135555 S135555555555	(1- A (1-15) 1 10 	B (16-26) 15 16 23 :: 1 1 1	(27-38) BC (27-38) (27-38) 26-27 36-39 1 * * 1 SG F S L S SY A MS R Y I I D G G T R Y I I D G G T H D S I N . I T F S S T F S S	C (39-55) C (39-46) (41 - 46 47 	(56-65) 55 (56-65) 55 (56-65) (56-65) (57-52) (56-65) (56-65) (57-52) (56-65) (56-65) (56-65) (56-65) (56-65) (57-62) (57-72) (57-72) </th <th>(66-74) (75-84) 74<75 1 A SWAAK G R F T I SK 3 A SWAAK G R T I SK 5 S S S T R ST 7 S S T T S T 5 G A A T R ST 7 G A A T ST 7 G A T ST 7 S V 7 S V 8 S V 8 S V 8 S V 8</th> <th>(66-104) E F (85-96) (97-104) 84 85 89 96 97 1 1 1 TS TTV DLK ITS PTTE DTATY FCA ST W. L RM. GL. A. ST Y. L M. GL. A. ST Y. L RM. GL. A. ST Y. L M. GL. D. ST Y. L M. GL. D. SAG. T. M. GL. D. SAG. T. M. GL. D. ST Y. L M. GL. D. SAG. T. M. L. D. </th>	(66-74) (75-84) 74<75 1 A SWAAK G R F T I SK 3 A SWAAK G R T I SK 5 S S S T R ST 7 S S T T S T 5 G A A T R ST 7 G A A T ST 7 G A T ST 7 S V 7 S V 8 S V 8 S V 8 S V 8	(66-104) E F (85-96) (97-104) 84 85 89 96 97 1 1 1 TS TTV DLK ITS PTTE DTATY FCA ST W. L RM. GL. A. ST Y. L M. GL. A. ST Y. L RM. GL. A. ST Y. L M. GL. D. ST Y. L M. GL. D. SAG. T. M. GL. D. SAG. T. M. GL. D. ST Y. L M. GL. D. SAG. T. M. L. D.
Phylogenetic groups VHe S12-4 S12-4 S12-4 S12-4 S12-15 S12-16 S12-16 S12-17 S12-20 S12-20 S12-20 S12-20 S12-20 S12-20 S12-20 S12-20 S12-20 S12-20 S12-20 S12-20 S12-20 S12-20 S12-20 S12-20 S13-20 S120 S12-20 S120 S120 S120 S120 S120 S120 S120 S1	(1- A (1-15) 1 10 1 Q SVEESGGR Q SVEESGGR Q SVEESGGR 	26) B (16-26) 15 15 16 23 1 1 1 <	(27-38) BC (27-38) (27-38) 26-27 SG FSLS R R R R N N N N N N N N N N N N N	(39-45) C (39-46) (41-4) (47-	(66-65) 55 (66-65) 55 (66-65) (1)	(66-74) (75-84) 74<75	(66-104) E F (85-96) (97-104) 84 45 89 96 97 1 1 1 STTV LL - RM GL A 1 SAG T. RM L A SAG T. EM GL A SAG T. M. GL A SAG T. M. GL A STV L - M GL A STV SAG T. M. GL A STV L - M. GL A M STV L - M. GL DA M S
Phylogenetic groups VHe St24 St24 St24 St215 St215 St216 St216 St216 St217 St227 St228 St227 St228 St227 St228 St229 St329 St3	(1- A (1-15) 1 10 	B (16-26) 15 16 23 : I I FP N FF N I K NT GT FF N I K NT F FK N IK GT K NT FF N IK E IK T FP NT IK E IK T FP NT FP NT <th>(27-38) BC (27-38) 26-27 36-39 1 </th> <th>C (39-55) C (39-46) (41-4) (41-46 47 </th> <th>(56-65) 55) (56-65) 55) (56-65) 9 (56-65) 9 (56-65) 9 (56-65) 9 (56-65) 9 (56-65) 9 (1) (50-65) 9 (1) (50-65) 9 (1) (1) 9 (1) (1) 9 (1) (1) 9 (1) (1) 9 (1) (1) 9 (1) (1) 9 (1) (1) 9 (1) (1) 9 (1) (1) 9 (1) (1) 9 (1) (1) 9 (1) (1) 9 (1) (1) 9 (1) (1) 9 (1) (1) 9 (1) (1) 9 (1) (1) 9</th> <th>(66-74) (75-84) 74 75 1 A SWAK G R F T I SK - N L A TR ST - S WAK G R F T I SK - N L A TR ST - S T ST - S T ST - S T S T - S T S T - S T S T - S T S T - S T S T - T S T - T S T - T S T - T S T - A T S T - Q A T - S T T R S T - Q A T S T - Q A T S T - Q A T S T - Q A T S T - Q A T S T - Q A T S T - Q A T S T - Q A T S T - S V N R</th> <th>(66-104) E F (85-96) (97-104) 84 85 89 96 97 104 I I TS TTV DLK ITS PTTE DTATY FCA ST V. L RM. GL.A ST V. L M. GL.A ST V. L M. GL.A ST V. L M. GL.A ST V. L M.GL.A ST V. L M.GL.A</th>	(27-38) BC (27-38) 26-27 36-39 1	C (39-55) C (39-46) (41-4) (41-46 47 	(56-65) 55) (56-65) 55) (56-65) 9 (56-65) 9 (56-65) 9 (56-65) 9 (56-65) 9 (56-65) 9 (1) (50-65) 9 (1) (50-65) 9 (1) (1) 9 (1) (1) 9 (1) (1) 9 (1) (1) 9 (1) (1) 9 (1) (1) 9 (1) (1) 9 (1) (1) 9 (1) (1) 9 (1) (1) 9 (1) (1) 9 (1) (1) 9 (1) (1) 9 (1) (1) 9 (1) (1) 9 (1) (1) 9 (1) (1) 9	(66-74) (75-84) 74 75 1 A SWAK G R F T I SK - N L A TR ST - S WAK G R F T I SK - N L A TR ST - S T ST - S T ST - S T S T - S T S T - S T S T - S T S T - S T S T - T S T - T S T - T S T - T S T - A T S T - Q A T - S T T R S T - Q A T S T - Q A T S T - Q A T S T - Q A T S T - Q A T S T - Q A T S T - Q A T S T - Q A T S T - S V N R	(66-104) E F (85-96) (97-104) 84 85 89 96 97 104 I I TS TTV DLK ITS PTTE DTATY FCA ST V. L RM. GL.A ST V. L M. GL.A ST V. L M. GL.A ST V. L M. GL.A ST V. L M.GL.A
Phylogenetic groups VHe S12-4 S12-4 S12-4 S12-4 S12-6 S12-13 S12-16 S12-16 S12-17 S12-20 S12-23 S12-23 S12-23 S12-23 S12-23 S12-23 S12-23 S12-23 S12-23 S12-23 S12-23 S13-24 S13-25 S13-24 S13-25 S13-24 S13-25 S13-24 S13-25 S13-24 S13-25 S13-24 S13-25 S13-24 S13-25 S13-24 S13-25 S13-24 S13-25 S13-24 S13-25 S13-24 S13-25 S13-24 S13-25 S13-	(1- A (1-15) 1 10 	B (16-26) 15 16 23 : I I FP N FF N I K NT GT FF N I K NT F FK N IK GT K NT FF N IK E IK T FP NT IK E IK T FP NT FP NT <th>(27-38) BC (27-38) 26-27 36-39 1 </th> <th>C (39-55) C (39-46) (41-4) (41-46 47 </th> <th>(56-65) 55) (56-65) 55) (56-65) 9 (56-65) 9 (56-65) 9 (56-65) 9 (56-65) 9 (56-65) 9 (1) (50-65) 9 (1) (50-65) 9 (1) (1) 9 (1) (1) 9 (1) (1) 9 (1) (1) 9 (1) (1) 9 (1) (1) 9 (1) (1) 9 (1) (1) 9 (1) (1) 9 (1) (1) 9 (1) (1) 9 (1) (1) 9 (1) (1) 9 (1) (1) 9 (1) (1) 9 (1) (1) 9 (1) (1) 9</th> <th>(66-74) (75-84) 74 75 1 A SWAK G R F T I SK - N L A TR ST - S WAK G R F T I SK - N L A TR ST - S T ST - S T ST - S T S T - S T S T - S T S T - S T S T - S T S T - T S T - T S T - T S T - T S T - A T S T - Q A T - S T T R S T - Q A T S T - Q A T S T - Q A T S T - Q A T S T - Q A T S T - Q A T S T - Q A T S T - Q A T S T - S V N R</th> <th>(66-104) E F (85-96) (97-104) 84 85 89 96 97 104 I I TS TTV DLK ITS PTTE DTATY FCA STV I STV L - RM GL SAG T. RM L A SAG T. EM L A SAG T. EM L A STV L - M. GL A SAG T. EM L A STV L - M. GL A SAG T. M. GL A S STV L - M. GL A SAG T. M. GL A STV L M. GL A STV L M. GL A </th>	(27-38) BC (27-38) 26-27 36-39 1	C (39-55) C (39-46) (41-4) (41-46 47 	(56-65) 55) (56-65) 55) (56-65) 9 (56-65) 9 (56-65) 9 (56-65) 9 (56-65) 9 (56-65) 9 (1) (50-65) 9 (1) (50-65) 9 (1) (1) 9 (1) (1) 9 (1) (1) 9 (1) (1) 9 (1) (1) 9 (1) (1) 9 (1) (1) 9 (1) (1) 9 (1) (1) 9 (1) (1) 9 (1) (1) 9 (1) (1) 9 (1) (1) 9 (1) (1) 9 (1) (1) 9 (1) (1) 9 (1) (1) 9	(66-74) (75-84) 74 75 1 A SWAK G R F T I SK - N L A TR ST - S WAK G R F T I SK - N L A TR ST - S T ST - S T ST - S T S T - S T S T - S T S T - S T S T - S T S T - T S T - T S T - T S T - T S T - A T S T - Q A T - S T T R S T - Q A T S T - Q A T S T - Q A T S T - Q A T S T - Q A T S T - Q A T S T - Q A T S T - Q A T S T - S V N R	(66-104) E F (85-96) (97-104) 84 85 89 96 97 104 I I TS TTV DLK ITS PTTE DTATY FCA STV I STV L - RM GL SAG T. RM L A SAG T. EM L A SAG T. EM L A STV L - M. GL A SAG T. EM L A STV L - M. GL A SAG T. M. GL A S STV L - M. GL A SAG T. M. GL A STV L M. GL A STV L M. GL A
Phylogenetic groups VHe VHe Sl2.4 Sl2.4 Sl2.4 Sl2.4 Sl2.2 Sl2.15 Sl2.12 Sl2.21 Sl2.25 Sl2.21 Sl2.22 Sl2.23 Sl3.24 Sl3.24 Sl3.24 Sl3.24 Sl3.24 Sl3.25	(1- A (1-15) 1 10 	B B (16-26) 15 15 16 23 1 1 . 1 1 . 1 1 . 1 1 . 1 1 . 1 1 . 1 1 . 1 1 . 1 1 . 1 . . 1 . . 1 <td< td=""><td>(27.38) BC (27.38) 26.27 36.99 1</td><td>C (39-55) C (39-46) (41 - 46 47 </td><td>(56-65) 55 (56-65) 55 (56-65) 55 (56-65) (56-65) (56-65) (56-65) (56-65) (51-55) (56-65) (57-70) (57-70) (57-70) (57-70) (57-70) (57-70)</td><td>(66-74) 74 75 74 75 74 75 8 WAAK GR FT I SK A SWAAK GR FT I SK A SWAAK GR FT I SK Q A T ST S T T ST S T T ST S T T ST C A T ST C A</td><td>(66-104) E F (85-96) (97-104) 84 85 89 96 97 104 I I TS TTV DLK ITS PTTE DTATY FCA ST V. L RM. GL.A ST V. L M. GL.A ST V. L M. GL.A ST V. L M. GL.A ST V. L M.GL.A ST V. L M.GL.A</td></td<>	(27.38) BC (27.38) 26.27 36.99 1	C (39-55) C (39-46) (41 - 46 47 	(56-65) 55 (56-65) 55 (56-65) 55 (56-65) (56-65) (56-65) (56-65) (56-65) (51-55) (56-65) (57-70) (57-70) (57-70) (57-70) (57-70) (57-70)	(66-74) 74 75 74 75 74 75 8 WAAK GR FT I SK A SWAAK GR FT I SK A SWAAK GR FT I SK Q A T ST S T T ST S T T ST S T T ST C A	(66-104) E F (85-96) (97-104) 84 85 89 96 97 104 I I TS TTV DLK ITS PTTE DTATY FCA ST V. L RM. GL.A ST V. L M. GL.A ST V. L M. GL.A ST V. L M. GL.A ST V. L M.GL.A

 $\underline{\textcircled{O}}$ Springer

Fig. 2 Alignment of leporid VH protein sequences. a European rabbit and hares VHa lineages and S. floridanus a5.1 lineage. b S. floridanus a5.2 and sS lineages, hares sL lineage and European rabbit and hares VHn lineage. Dashes represent alignment gaps. Dots indicate identity with the master sequence except at gap positions. The "IMGT Protein display for V domain" header is shown (Lefranc et al. 2003). Sequences are grouped into phylogenetic groups according to the tree in Fig. 1. Shown are representatives of European rabbit VH1a1, VH1a2, VH1a3, Vhx, Vhy and Vhz germ line sequences (GenBank accession numbers M93171, M93172, M93173, L03846, L03890 and AF264469, respectively); European rabbit a4.1 and a4.2 cDNA sequences (GenBank accession numbers AY207980 and AY207967) and hares a2L and sL cDNA sequences (GenBank accession numbers AY288450, AY288465 and AY288459). Also shown are all S. floridanus cDNA sequences obtained in this study that were grouped into the a5.1, a5.2 and sS lineages (for GenBank accession numbers please see Table 1). Highlighted in grey are the VHa deletion at position 2 and motifs 19LTLTCT24 and 70WAK72, and sL/sS sequences 70SVK72 motif. Also highlighted are residue characteristics of a5 and sS lineages

of *VHn* has been shown, *L. europaeus*, are European leporids while *S. floridanus* is an American leporid (Chapman et al. 1980; Flux and Angermann 1990). Interestingly, two related caliciviruses affect the European rabbit and two hare species of Eurasian distribution (*L. europaeus* and *Lepus timidus*) but not cottontails (Abrantes et al. 2012; Lopes et al. 2013). Thus, we can put forward the hypothesis that the different continents' pathogen community may impose different selective pressures and favour the expression of *VH*n only in European species.

Acknowledgments This work was partially funded by FEDER (Fundo Europeu de Desenvolvimento Regional) funds through the Programa Operacional Factores de Competitividade (COMPETE program; FCOMP-01-0124-FEDER-028286) and Portuguese national funds through FCT (Fundação para a Ciência e a Tecnologia; research project PTDC/BIO-ANM/3963/2012)—Quadro de Referência Estratégico Nacional (QREN) funds from the European Social Fund and Portuguese Ministério da Educação e Ciência. FCT also supported the doctoral grant of AP (ref.: SFRH/BD/71252/2010) and the FCT Investigator grant of JA (ref.: IF/01396/2013) and the Post-Doctoral grant of JMF (ref.: SFRH/BPD/43264/2008). "Genomics Applied to Genetic Resources" co-financed by North Portugal Regional Operational Programme 2007/2013 (ON.2 – O Novo Norte), under the National Strategic Reference Framework (NSRF), through the European Regional Development Fund (ERDF), also supported this work.

References

- Abrantes J, van der Loo W, Le Pendu J, Esteves PJ (2012) Rabbit haemorrhagic disease (RHD) and rabbit haemorrhagic disease virus (RHDV): a review. Vet Res 43(1):12
- Alves PC, Hackländer K (2008) Lagomorph species: geographical distribution and conservation status. In: Alves PC, Ferrand N, Hackländer K (eds) Lagomorph biology. Springer, Berlin, pp 395– 405
- Chapman JA, Hockman JG, Ojeda MM (1980) Sylvilagus floridanus. Mamm Species 136:1–8

- Darriba D, Taboada GL, Doallo R, Posada D (2012) jModelTest 2: more models, new heuristics and parallel computing. Nat Methods 9(8): 772
- Drummond AJ, Rambaut A (2007) BEAST: Bayesian evolutionary analysis by sampling trees. BMC Evol Biol 7:214
- Drummond AJ, Ho SY, Phillips MJ, Rambaut A (2006) Relaxed phylogenetics and dating with confidence. PLoS Biol 4(5):e88
- Esteves PJ (2003) Molecular and population genetic analysis of polymorphism at the antibody loci IgGCH2 and IgVH in lagomorphs. Faculdade de Ciências Universidade do Porto, Porto (http://hdl. handle.net/10216/10598)
- Esteves PJ, Lanning D, Ferrand N, Knight KL, Zhai SK, van der Loo W (2004) Allelic variation at the VHa locus in natural populations of rabbit (Oryctolagus cuniculus, L.). J Immunol 172(2):1044–1053
- Esteves PJ, Lanning D, Ferrand N, Knight KL, Zhai SK, van der Loo W (2005) The evolution of the immunoglobulin heavy chain variable region (IgVH) in Leporids: an unusual case of transspecies polymorphism. Immunogenetics 57(11):874–882
- Flajnik MF (2002) Comparative analyses of immunoglobulin genes: surprises and portents. Nat Rev Immunol 2(9):688–698
- Flux JEC, Angermann R (1990) The hares and jackrabbits. In: Chapman JA, Flux JEC (eds) Rabbits, hares and pikas: status survey and conservation action plan. The World Conservation Union, Gland, pp 61–94
- Hall TA (1999) BioEdit: a user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT. Nucleic Acids Symp Ser 41:95–98
- Horng WJ, Knight KL, Dray S (1976) Heavy chain variable region allotypic sub-specificities of rabbit immunoglobulins. I. Identification of three subpopulations of al IgG molecules. J Immunol 116(1):117–125
- Kim BS, Dray S (1973) Expression of the a, x, and y variable region genes of heavy chains among IgG, IgM, and IgA molecules of normal and a locus allotype-suppressed rabbits. J Immunol 111(3): 750–760
- Kindt TJ (1975) Rabbit immunoglobulin allotypes: structure, immunology, and genetics. Adv Immunol 21:35–86
- Knight KL (1992) Restricted VH gene usage and generation of antibody diversity in rabbit. Annu Rev Immunol 10:593–616
- Knight KL, Becker RS (1990) Molecular basis of the allelic inheritance of rabbit immunoglobulin VH allotypes: implications for the generation of antibody diversity. Cell 60(6):963–970
- Lanning D, Sethupathi P, Rhee KJ, Zhai SK, Knight KL (2000a) Intestinal microflora and diversification of the rabbit antibody repertoire. J Immunol 165(4):2012–2019
- Lanning D, Zhu X, Zhai SK, Knight KL (2000b) Development of the antibody repertoire in rabbit: gut-associated lymphoid tissue, microbes, and selection. Immunol Rev 175:214–228
- Lefranc MP, Pommié C, Ruiz M, Giudicelli V, Foulquier E, Truong L, Thouvenin-Contet V, Lefranc G (2003) IMGT unique numbering for immunoglobulin and T cell receptor variable domains and Ig superfamily V-like domains. Dev Comp Immunol 27(1):55–77
- Lopes AM, Gavier-Widén D, Le Gall-Reculé G, Esteves PJ, Abrantes J (2013) Complete coding sequences of European brown hare syndrome virus (EBHSV) strains isolated in 1982 in Sweden. Arch Virol 158(10):2193–6
- Mage RG, Bernstein KE, McCartney-Francis N, Alexander CB, Young-Cooper GO, Padlan EA, Cohen GH (1984) The structural and genetic basis for expression of normal and latent VHa allotypes of the rabbit. Mol Immunol 21(11):1067–1081
- Mage RG, Lanning D, Knight KL (2006) B cell and antibody repertoire development in rabbits: the requirement of gut-associated lymphoid tissues. Dev Comp Immunol 30(1–2):137–153
- Margolies MN, Cannon LE, Kindt TJ, Fraser B (1977) The structural basis of rabbit VH allotypes: serologic studies on a1 H chains with defined amino acid sequence. J Immunol 119(1):287–294

- Matthee CA, van Vuuren BJ, Bell D, Robinson TJ (2004) A molecular supermatrix of the rabbits and hares (Leporidae) allows for the identification of five intercontinental exchanges during the Miocene. Syst Biol 53(3):433–447
- Pinheiro A, Lanning D, Alves PC, Mage RG, Knight KL, van der Loo W, Esteves PJ (2011) Molecular bases of genetic diversity and evolution of the immunoglobulin heavy chain variable region (IGHV) gene locus in leporids. Immunogenetics 63(7):397–408
- Pinheiro A, de Mera IG, Alves PC, Gortázar C, de la Fuente J, Esteves PJ (2013) Sequencing of modern Lepus VDJ genes shows that the usage of VHn genes has been retained in both Oryctolagus and Lepus that diverged 12 million years ago. Immunogenetics 65(11): 777–84
- Rambaut A, Drummond AJ (2007) Tracer v1.5. Available from http:// beast.bio.ed.ac.uk/Tracer
- Ros F, Puels J, Reichenberger N, van Schooten W, Buelow R, Platzer J (2004) Sequence analysis of 0.5 Mb of the rabbit germline immunoglobulin heavy chain locus. Gene 330:49– 59
- Roux KH (1981) A fourth heavy chain variable region subgroup, w, with 2 variants defined by an induced auto-antiserum in the rabbit. J Immunol 127(2):626–632

- Sehgal D, Obiakor H, Mage RG (2002) Distinct clonal Ig diversification patterns in young appendix compared to antigen-specific splenic clones. J Immunol 168(11):5424–5433
- Tamura K, Peterson D, Peterson N, Stecher G, Nei M, Kumar S (2011) MEGA5: molecular evolutionary genetics analysis using maximum likelihood, evolutionary distance, and maximum parsimony methods. Mol Biol Evol 28:2731–2739
- Thompson JD, Higgins DG, Gibson TJ (1994) CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. Nucleic Acids Res 22:4673–4680
- Tonnelle C, Cazenave PA, Brezin C, Moinier D, Fougereau M (1983) Structural correlates to the rabbit immunoglobulin heavy chain a100 allotype. Mol Immunol 20(7):753–761
- van der Loo W (1987) Population genetical studies on the adaptive significance of the immunoglobulin allotypes in wild rabbit. In: Dubiski S (eds) The rabbit in contemporary immunological research. Longman Scientific & Technical, John Wiley, pp 101–126
- Vidus Rosin A, Gilio N, Meriggi A (2008) Introduced Lagomorphs as a threat to "native" Lagomorphs: the case of the Eastern cottontail (*Sylvilagus floridanus*) in northern Italy. In: Alves PC, Ferrand N, Hacklander K (eds) Lagomorph biology. Evolution, ecology and conservation. Springer, Eidenberg, pp 153–165