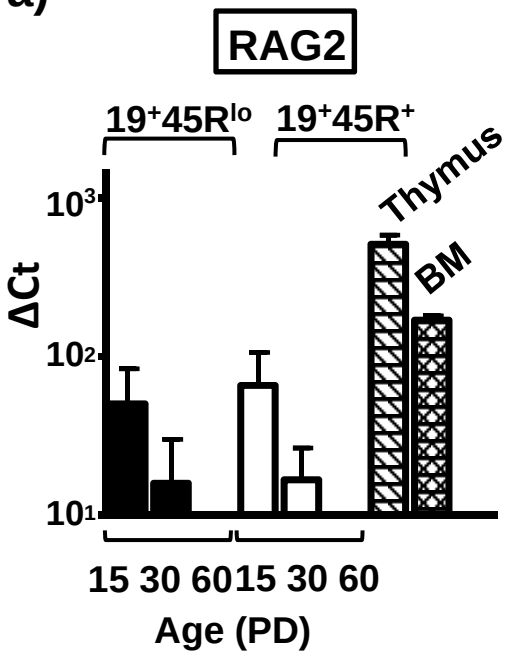
a)



b)

Frequency analysis of B-cell progenitors in FL(14dpc), adult bone marrow and spleen

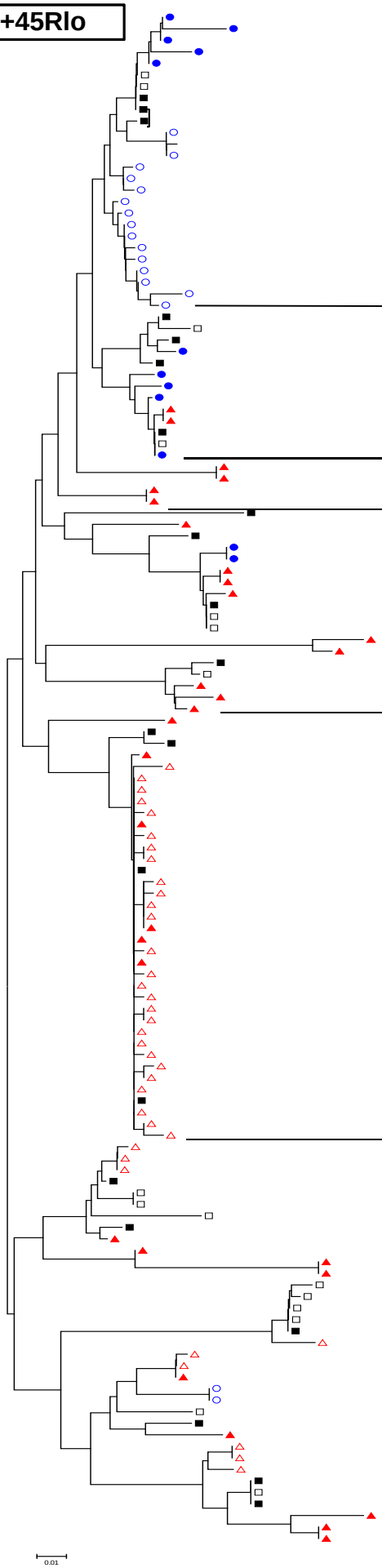
ORIGIN	POPULATION	EXP1	EXP2	EXP3	EXP4	MEAN
FETAL LIVER	$19^{+}45R^{lo}$	1/44	1/12	1/12		1/22
	$19^{+}45R^{+}$	1/83	1/41	1/54		1/59
BONE MARROW	$19^{+}45R^{lo}$	1/32	1/57	1/14	1/50	1/38
	$19^{+}45R^{+}$	1/54	1/73	1/88	1/91	1/76
	$19^{-}45R^{+}$	1/167	1/472			1/319
ADULT SPLEEN	$19^{+}45R^{-}$	0	0			$<1/10^{5}$
	$19^{+}45R^{+}$	0	0			$<1/10^{4}$

Supplemental Figure 1: RAG2 expression and quantification of clonable B cells from different fetal and adult samples. a) qRT-PCR RAG2 expression on splenic $19^{+}45R^{lo}$ and $19^{+}45R^{+}$ sorted cells from PD15, PD30 and PD60 samples. The results are presented as in Fig. 1 (mean ± SEM, n=4 performed in duplicates). Samples obtained from adult thymus and bone marrow (BM) were used as positive controls. b) Frequency analysis in limiting dilution conditions (*Lefkovits, I. et al.* Immunol. Today 1984 5:265-268) of different B cell samples from fetal liver (gestational age 14), bone marrow and spleen (adult PD60). Cells were plated at 1000, 600, 300, 100 and 30 cels/well. A total number of 10^{5} cells were plated in the presence of the stromal cell line ST2 and 30 ng/ml of rIL7 (Peprotech, Rocky Hill, NJ) as reported (*Martinez-M. JA. et al.* Blood 2001 98:1862-1871). The frequency was calculated as described (*Rolink, A. et al.* Blood 1993 81:2290-2300).

Prado et al.

VH1 sequences from 19+45Rlo

- IgM PD15
- IgM PD60
- IgG1 PD15
- IgG1 PD60
- IgA PD15
- IgA PD60

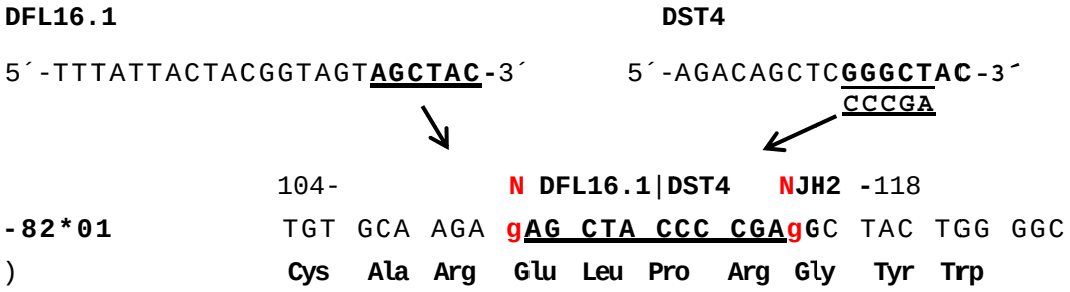


PD15/PD60		
IgM	IgG1	IgA
2/3	0	15/5
2/4	0/2	0/5
0	0/4	0
3/4	0/9	0/2
0/4	27/6	0
9/6	8/9	2/0

Supplemental Figure 2: Phylogenetic tree of VH1 sequences from $19^{+}45R^{lo}$. Each symbol represents a single sequence, as indicated by the code in the figure. In the right are shown the numbers summarizing PD15 and PD60 IgM, IgG1 and IgA sequences in each branch/clonotype, indicated by the horizontal lines.

Prado et al.

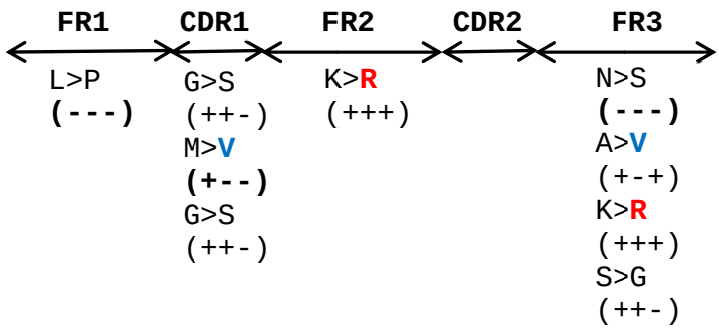
a)



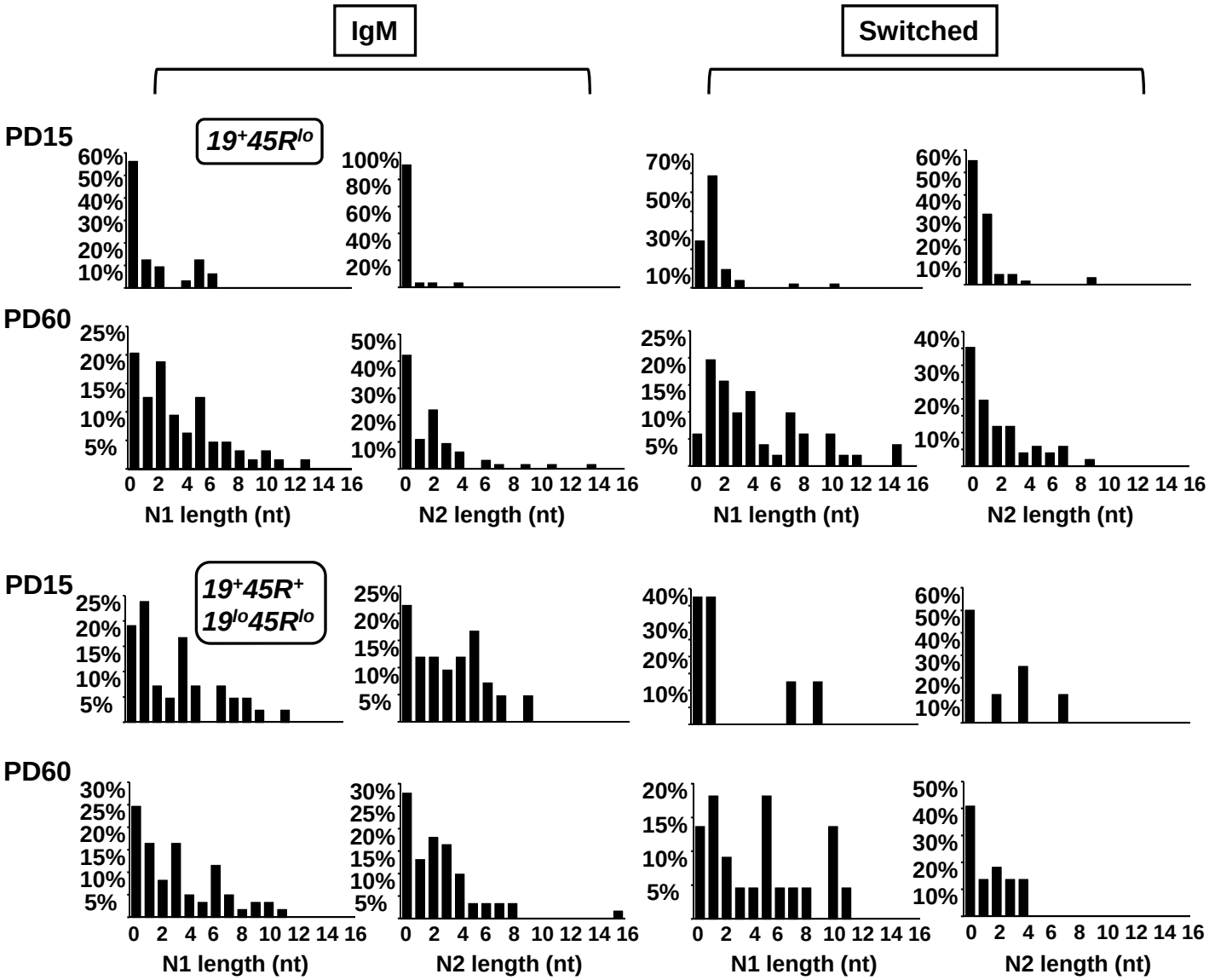
b)

	Nucleotide	Amino acid
Codon Number	----------11	-------1
(IMGT)	12344477788899900	12347890
	27957829905703500	27982530
Position in codon	21123222331332323	21122122
IGHV1-82*01	TGAGAAACTAACAGGCC	LGMKKSSA
#1	C.	P. . . . .
#2	. . G.	. . V. . . .
#3	. A.	. S. . . . .
#4	. A. . . . G. . . G.	. S. . RG. .
#5	. . . A. . . . . G.	. . . . .
#6	. . . . . G. . G. . . .	. . . . .
#7	. . . . . G.	. . . . .
#8	. . . . . T.	. . . . .
#9	. . . . . A. . .	. . . . . N.
#10	. . . . . C.	. . . . .
#11	. . . . . G.	. . . R. . .
#12	. . . . G.	. . . . .
#13	. . . . . C. T	. . . . .
#14	. . . . . T.	. . . . .
#15	. . . . . T.	 V

c)



Supplemental Figure 3: DH to DH gene rearrangement present in a sequence clonotype from switched $19^{+}45R^{lo}$ cells. a) Schematic representation of the DH to DH rearrangement. Above are shown the germline sequences of the DFL16.1, DST4 and the complementary for the DST4 gene segment. Bold letters indicated the motifs found in sequences from switched PD15 $19^{+}45R^{lo}$ cells. In lower case, N nt insertion. b) Nucleotide (left) and amino acid (right) sequence variations at the VH region according IMGT numbering. Codon numbers are indicated vertically above the sequences. Position 1, 2, and 3 refer to the first, second and third nucleotides in the codon, respectively. c) Schematically depicted amino acid changes on the VH region; in red and blue are hydrophilic and hydrophobic amino acid, respectively. In parenthesis, code of amino acid changes defined by IMGT: (+++) very similar; (++-) (+-+) similar; (+--) bold, dissimilar; (---) bold, very dissimilar.



Supplemental Figure 4: Distribution of N additions (N1 and N2) expressed as percentage of all unique sequences.

Prado et al

Supplemental Table I: Comparative analyses of CDR-H3 length (nt) from the sequences with DFL16.1, other D region gene segments and sequences with no identifiable D region.

		DFL16.1	Other D	No D
PD15	**IgM $19^{+}45R^{lo}$**	34.5 ± 2.3[a] (9)	32.5 ± 1.4 (17)	20.4 ± 2[b] (6)
	IgM $19^{+}45R^{+}$	42.6 ± 3.2(11)	36.6 ± 1.5 (26)	25.8 ± 3[b] (5)
	Switched $19^{+}45R^{lo}$	28.4 ± 1(32)	30.5 ± 1.8 (22)	18.9 ± 1[b] (13)
	Switched $19^{+}45R^{+}$	30 ± 3.4(3)	41 ± 1 (3)	24 (2)
PD60	**IgM $19^{+}45R^{lo}$**	39.5 ± 2.1(12)	36.3 ± 0.7 (43)	15[b] (9)
	IgM $19^{+}45R^{+}$	37.8 ± 2.4(19)	35.2 ± 1.1 (36)	20.5 ± 2.1[b] (6)
	Switched $19^{+}45R^{lo}$	36.4 ± 3.8(11)	36.5 ± 1.0 (31)	27.4 ± 1.4 (9)
	Switched $19^{lo}45R^{lo}$	40.2 ± 2.2(7)	37.1 ± 1.3 (11)	36 (4)

[a]Length (nt) of CDR-H3 (number of sequences). Statistical significance of CDR-H3 length differences has been calculated by using unpaired two-tailed Student´s *t*-test. No significant differences were obtained when DFL16.1 *versus* other D regions were compared. The comparison of sequences with any identifiable D regions *versus* sequences with no D region was as follows: [b] $p >$ 0.001.

Supplemental Table II: Deconstruction of CDR-H3 on the sequences with DFL16.1 and other D regions.

	DFL16.1	V length (nt)	P-N1-P (nt)	D length (nt)	P-N2-P (nt)	J length (nt)
PD15	**IgM $19^{+}45R^{lo}$**	6.1 ± 0.4	0.8 ± 0.3	15.1 ± 1.6	0.1	12.1 ± 2.1
	IgM $19^{+}45R^{+}$	7 ± 1.8	3.8 ± 1.7	13.8 ± 2.6	3 ± 0.8	15 ± 2.1
	Switched $19^{+}45R^{lo}$	5.9 ± 0.1	1.1 ± 0.1	12.1± 0.1	0.9 ± 0.1	6.7 ± 0.8
	Switched $19^{+}45R^{+}$	6	0.6 ± 0.3	13.3 ± 3.7	2.3 ± 2	7.6 ± 2.7
PD60	**IgM $19^{+}45R^{lo}$**	5.6 ± 0.2	6 ± 1.1	14.8 ± 1.1	3.3 ± 1.5	9.6 ± 2.6
	IgM $19^{+}45R^{+}$	5.5 ± 0.3	4.2 ± 0.9	13.4 ± 1.2	2.7 ± 0.5	11.9 ± 1.1
	Switched $19^{+}45R^{lo}$	6.1 ± 0.3	2.7 ± 0.7	12.4 ±1.7	4.5 ± 1.3	10.5 ± 2.4
	Switched $19^{lo}45R^{lo}$	5.8 ± 0.5	3.4 ± 1.4	14.6 ± 1.1	1.8 ± 0.4	14.6 ±1.5

	Other D region	V length (nt)	P-N1-P (nt)	D length (nt)	P-N2-P (nt)	J length (nt)
PD15	**IgM $19^{+}45R^{lo}$**	5.5 ± 0.5	1.9 ± 0.7	9.5 ± 0.6	0.5 ± 0.4	15 ± 1.1
	IgM $19^{+}45R^{+}$	5.9 ± 0.3	3.5 ± 0.5	10.5 ± 0.8	3.6 ± 0.5	12.8 ± 0.6
	Switched $19^{+}45R^{lo}$	5.3 ± 0.4	2.1 ± 0.7	10.9 ± 0.9	2.4 ± 0.7	9.5 ± 1.16
	Switched $19^{+}45R^{+}$	5 ± 0.7	5 ± 1.9	15.75± 0.8	2.5 ± 0.9	10 ± 1.3
PD60	**IgM $19^{+}45R^{lo}$**	6.1 ± 0.2	3.1 ± 0.4	10.3 ± 0.4	2.3 ± 0.4	14.5 ± 0.6
	IgM $19^{+}45R^{+}$	5.4 ± 0.4	3.2 ± 0.5	10.1 ± 0.5	2.8 ± 0.5	13.5 ± 0.7
	Switched $19^{+}45R^{lo}$	5.4 ± 0.3	3.6 ± 0.5	10.1 ±0.5	2.3 ± 0.4	14.8 ± 0.7
	Switched $19^{+}45R^{+}$	5.6 ± 0.3	5.1 ± 1.0	10.1 ± 0.9	2.3 ± 0.6	14. ±0.9

Deconstruction studies were performed for the sequences presented in Supplemental Table I.

Prado et al