



**Supplemental Material to:**

**Sreekumar Kodangattil, Christine Huard, Cindy Ross, Jian Li, Huilan Gao, Alessandro Mascioni, Santosh Hodawadekar, Snehal Naik, Jessica Min-debartolo, Alberto Visintin, and Juan C Almagro**

**The functional repertoire of rabbit antibodies and antibody discovery via next-generation sequencing**

**2014; mAbs 6(3)**

**<http://dx.doi.org/10.4161/mabs.28059>**

**<https://www.landesbioscience.com/journals/mabs/article/28059/>**

**Table 1.** Roche/454 sequencing statistics.

A. Number and average length of sequence reads that satisfied the QC metrics as described in Material and Methods.

Source	VH		VK	
	Number of reads	Average Length	Number of Reads	Average Length
<b>BM</b>	235,336	411	28,874	375
<b>SP</b>	289,888	409	44,039	375
<b>IM</b>	286,904	405	29,763	376

B. Number of unique protein sequences without frameshifts and at least as long as their corresponding germline V gene segment. These data sets were used to estimate the diversity of the repertoire and calculate the somatic mutations.

Source	VH			VK		
	Total	Unique	%	Total	Unique	%
<b>BM</b>	224,535	163,795	73%	24,544	19,550	80%
<b>SP</b>	267,364	219,237	82%	37,813	31,763	84%
<b>IM</b>	264,437	171,815	65%	25,452	18,767	74%

**Table 2.** Germline genes used in the IGVH/K and IGJH/K assignment. The nomenclature and sequences are reported in IMGT (<http://www.imgt.org/>)

**IGVH:**

IGHV1S33\*01, IGHV1S34\*01, IGHV1S36\*01, IGHV1S40\*01, IGHV1S43\*01, IGHV1S44\*01, IGHV1S45\*01,  
IGHV1S47\*01, IGHV1S49\*01, IGHV1S50\*01, IGHV1S51\*01, IGHV1S52\*01, IGHV1S53\*01, IGHV1S54\*01,  
IGHV1S55\*01, IGHV1S56\*01, IGHV1S57\*01, IGHV1S58\*01, IGHV1S59\*01, IGHV1S60\*01, IGHV1S61\*01,  
IGHV1S62\*01, IGHV1S63\*01, IGHV1S64\*01, IGHV1S65\*01, IGHV1S66\*01, IGHV1S67\*01, IGHV1S68\*01,  
IGHV1S69\*01, IGHV1S7\*01,  
IGHV1S8\*01

**IGVK:**

IGKV1S1\*01, IGKV1S10\*01, IGKV1S11\*01, IGKV1S12\*01, IGKV1S14\*01, IGKV1S15\*01, IGKV1S16\*01,  
IGKV1S17\*01, IGKV1S18\*01, IGKV1S19\*01, IGKV1S2\*01, IGKV1S20\*01, IGKV1S21\*01, IGKV1S22\*01,  
IGKV1S23\*01, IGKV1S24\*01, IGKV1S25\*01, IGKV1S26\*01, IGKV1S27\*01, IGKV1S28\*01, IGKV1S29\*01,  
IGKV1S3\*01, IGKV1S30\*01, IGKV1S31\*01, IGKV1S32\*01, IGKV1S33\*01, IGKV1S34\*01, IGKV1S35\*01,  
IGKV1S36\*01, IGKV1S37\*01, IGKV1S39\*01, IGKV1S4\*01, IGKV1S40\*01, IGKV1S41\*01, IGKV1S42\*01,  
IGKV1S44\*01, IGKV1S45\*01, IGKV1S46\*01, IGKV1S47\*01, IGKV1S48\*01, IGKV1S49\*01, IGKV1S5\*01,  
IGKV1S50\*01, IGKV1S51\*01, IGKV1S52\*01, IGKV1S53\*01, IGKV1S54\*01, IGKV1S55\*01, IGKV1S56\*01,  
IGKV1S57\*01, IGKV1S58\*01, IGKV1S59\*01, IGKV1S6\*01, IGKV1S60\*01, IGKV1S61\*01, IGKV1S62\*01,  
IGKV1S63\*01, IGKV1S64\*01, IGKV1S65\*01, IGKV1S66\*01, IGKV1S67\*01, IGKV1S68\*01, IGKV1S7\*01,  
IGKV1S8\*01 IGKV1S9\*01

**IGJH/K:**

IGHJ1\*01, IGHJ2\*01, IGHJ3\*01, IGHJ4\*01, IGHJ5\*01, IGHJ6\*01 and IGKJ1-2\*01

**Table 3.** Full list of IGVH germline gene usage observed in the samples analyzed.

Gene	BM		SP		IM	
	Counts	%	Counts	%	Counts	%
1S40	150266	58.19	180059	58.87	97162	32.32
1S45	85033	32.93	109673	35.86	68549	22.80
1S43	8261	3.20	7946	2.60	3047	1.01
1S47	7815	3.03	6048	1.98	2767	0.92
1S7	5814	2.25	883	0.29	1047	0.35
1S44	494	0.19	225	0.07	6172	2.05
1S8	203	0.08	44	0.01	50	0.02
1S33	125	0.05	386	0.13	41	0.01
1S28	90	0.03	302	0.10	82	0.03
1S69	37	0.01	156	0.05	121193	40.31
1S36	34	0.01	46	0.02	203	0.07
1S25	25	0.01	19	0.01	213	0.07
1S17	11	0.00	11	0.00	50	0.02
1S54	7	0.00	20	0.01	33	0.01
1S24	5	0.00	2	0.00	4	0.00
1S64	2	0.00	4	0.00	18	0.01
1S50	1	0.00	0	0.00	0	0.00
1S1	0	0.00	0	0.00	1	0.00
1S13	0	0.00	2	0.00	18	0.01
1S26	0	0.00	0	0.00	9	0.00
1S31	0	0.00	4	0.00	5	0.00
1S34	0	0.00	12	0.00	3	0.00
1S53	0	0.00	3	0.00	1	0.00
1S58	0	0.00	1	0.00	1	0.00

**Table 4.** Full list of IGVK germline gene usage observed in the samples analyzed.

Gene	BM		SP		IM	
	Counts	%	Counts	%	Counts	%
1S10	5548	19.21	6947	15.77	4157	13.97
1S36	3699	12.81	4028	9.15	3620	12.16
1S1	3463	11.99	5273	11.97	1824	6.13
1S17	2014	6.98	1687	3.83	1414	4.75
1S34	1746	6.05	2268	5.15	1458	4.90
1S15	1410	4.88	3193	7.25	1739	5.84
1S42	1317	4.56	1777	4.04	980	3.29
1S52	1252	4.34	3538	8.03	3138	10.54
1S37	1066	3.69	987	2.24	1201	4.04
1S23	829	2.87	2021	4.59	1396	4.69
1S4	762	2.64	1289	2.93	676	2.27
1S32	469	1.62	341	0.77	491	1.65
1S2	432	1.50	642	1.46	458	1.54
1S18	426	1.48	941	2.14	929	3.12
1S49	409	1.42	664	1.51	400	1.34
1S64	327	1.13	845	1.92	291	0.98
1S19	315	1.09	439	1.00	357	1.20
1S40	310	1.07	893	2.03	748	2.51
1S54	292	1.01	454	1.03	225	0.76
1S3	256	0.89	300	0.68	194	0.65
1S12	219	0.76	697	1.58	167	0.56
1S33	174	0.60	350	0.79	83	0.28
1S62	141	0.49	398	0.90	547	1.84
1S30	137	0.47	298	0.68	248	0.83
1S63	131	0.45	323	0.73	205	0.69
1S67	123	0.43	330	0.75	153	0.51
1S5	119	0.41	215	0.49	53	0.18
1S16	117	0.41	122	0.28	369	1.24
1S50	117	0.41	118	0.27	57	0.19
1S29	112	0.39	247	0.56	201	0.68
1S39	112	0.39	128	0.29	75	0.25
1S22	105	0.36	172	0.39	77	0.26
1S7	102	0.35	93	0.21	82	0.28
1S47	93	0.32	101	0.23	99	0.33
1S56	93	0.32	323	0.73	151	0.51
1S46	79	0.27	281	0.64	545	1.83
1S59	73	0.25	151	0.34	88	0.30
1S60	73	0.25	129	0.29	80	0.27
1S9	52	0.18	115	0.26	81	0.27

**IGVK**  
**(continued)**

Gene	BM		SP		IM	
	Counts	%	Counts	%	Counts	%
1S55	35	0.12	68	0.15	16	0.05
1S53	34	0.12	128	0.29	48	0.16
1S58	32	0.11	61	0.14	33	0.11
1S31	30	0.10	55	0.12	15	0.05
1S6	29	0.10	73	0.17	24	0.08
1S44	27	0.09	52	0.12	38	0.13
1S68	27	0.09	100	0.23	94	0.32
1S27	21	0.07	36	0.08	54	0.18
1S48	15	0.05	50	0.11	67	0.23
1S41	14	0.05	36	0.08	35	0.12
1S14	13	0.05	40	0.09	19	0.06
1S26	12	0.04	5	0.01	1	0.00
1S20	6	0.02	60	0.14	15	0.05
1S51	6	0.02	3	0.01	1	0.00
1S8	4	0.01	8	0.02	7	0.02
1S25	2	0.01	2	0.00	6	0.02
1S65	2	0.01	2	0.00	0	0.00
1S28	1	0.00	0	0.00	0	0.00
1S21	0	0.00	0	0.00	17	0.06
1S24	0	0.00	1	0.00	1	0.00
1S35	0	0.00	0	0.00	2	0.01
1S66	0	0.00	3	0.01	1	0.00

**Table 5.** Ranking of CDR-H3 sequences based on the frequency of the germline gene – CDR-H3 sequence combination. Highlighted in orange is the germline gene – CDR-H3 combination identical to that of a hybridoma derived antibody.

<b>Ranking</b>	<b>Germline gene – CDR-H3</b>	<b>BM</b>	<b>SP</b>	<b>IM</b>
1	IGHV1S40_CARGGDLW	1	1	3665
2	IGHV1S69_CARKAVASGYSFDIW			3660
3	IGHV1S69_CASGIDGYRNIW			3297
4	IGHV1S45_CARKNNYNYYPYNFW			3292
5	IGHV1S69_CARGYTDYLHIW			2375
6	IGHV1S69_CGRGVPGGNDGDGIW			2242
7	IGHV1S69_CARKAVASGYTFDIW			2147
8	IGHV1S69_CARAYDDGDADFDPW			1899
9	IGHV1S45_CAKSAAYVDGGSPYATDGFDPW			1744
<b>10</b>	<b>IGHV1S40_CARNVGYNTWDLW</b>			<b>1559</b>
11	IGHV1S45_CARKNNYNYYPYNLW			1539
12	IGHV1S40_CARSYNDYDGFYFNLW			1486
13	IGHV1S69_CARGSHIKSAMDIA			1454
14	IGHV1S69_CARYSNDYMGFNIW			1367
15	IGHV1S45_CARDQYGVSSGYHLYYFDLW			1267

**Table 6.** Primers used to amplify the VH and VL repertoire of the BM, SP and IM samples. The primers were designed and validated by Rader et al., 2000. In red semi-degenerated positions – see code below.

**RHFabVk1 (Primer 21)**

5'GGGCCAGGCGGCCGAGCTCGTGMTGACCCAGACTCCA3'

**RHFabVk2 (Primer 22)**

5'GGGCCAGGCGGCCGAGCTCGATMTGACCCAGACTCCA3'

**rhfabvk3 (Primer 23)**

5'AGATGGTGCAGCCACAGTTCGTTTGATTCCACATTGGTGCC3'

**rhfabvk4 (Primer 24)**

5'AGATGGTGCAGCCACAGTTCGTAGGATCTCCAGCTCGGTCCC3'

**rhfabvk5 (Primer 25)**

5'AGATGGTGCAGCCACAGTTCGTTTGACACCACCTCGGTCCC3'

**RHFabVH1 (Primers 26-27)**

5'GCTGCCAAC CAGCCATGGCCAGTCGGTGGAGGAGTCC**R**GG3'

**RHFabVH2 (Primer 28)**

5'GCTGCCAACACCAGCCATGGCCAGTCGGTGAAGGAGTCCGAG3'

**RHFabVH3 (Primers 29-30)**

5'GCTGCCAACACCAGCCATGGCCAGTCG**Y**TGGAGGAGTCCGGG3'

**RHFabVH4 (Primers 31-32-33-34)**

5'GCTGCCAACACCAGCCATGGCCAG**S**AGCAGCTG**R**TGGAGTCCGG3'

**rhfabvh5 (Primers 35-36-37-38)**

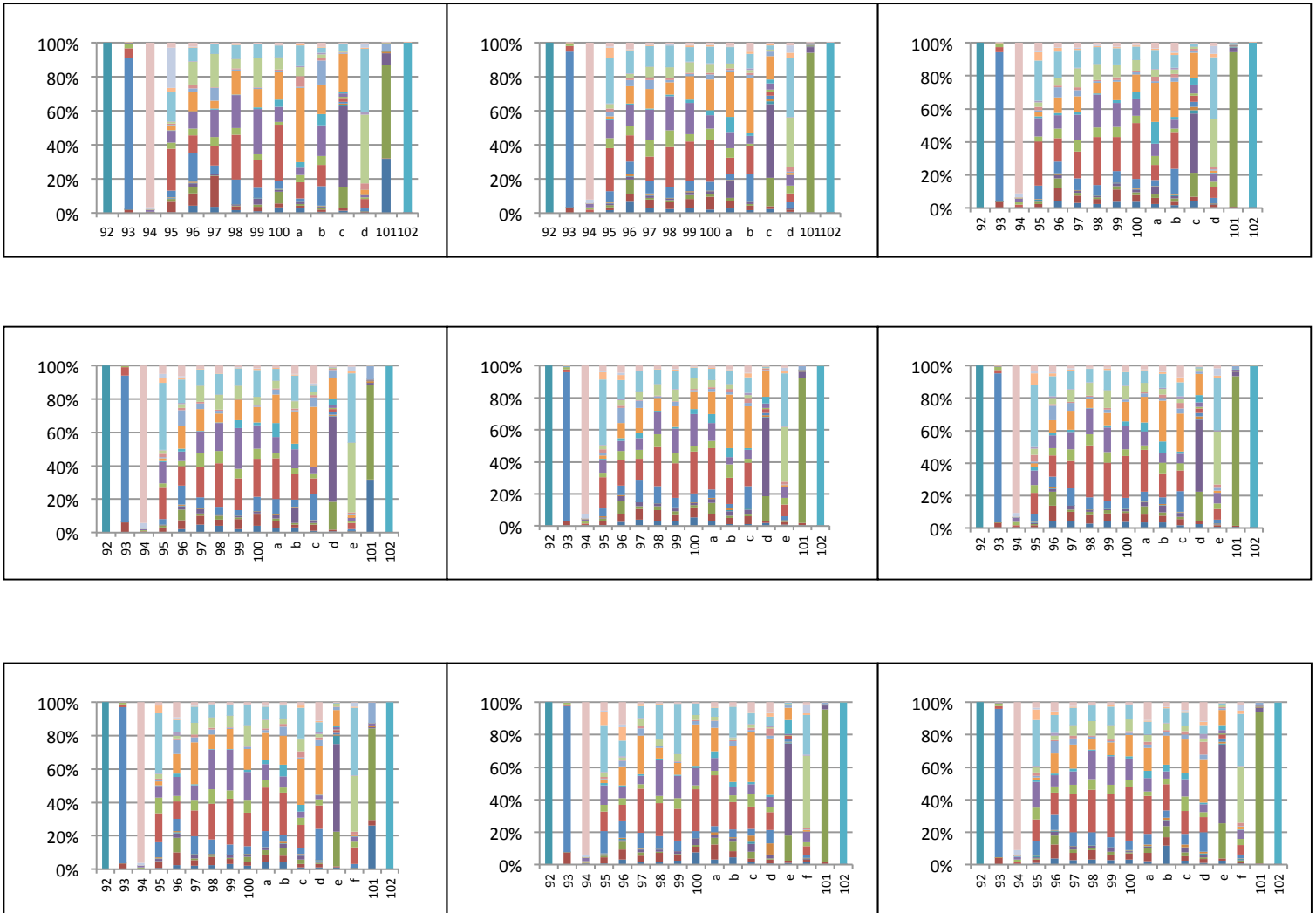
5'CGATGGGCCCTTGGTGGAGGCTG**R**GAG**A**YGGTGACCAAGGGTGCC3'

Nucleotide code: R; A or G / S; G or C / Y; C or T

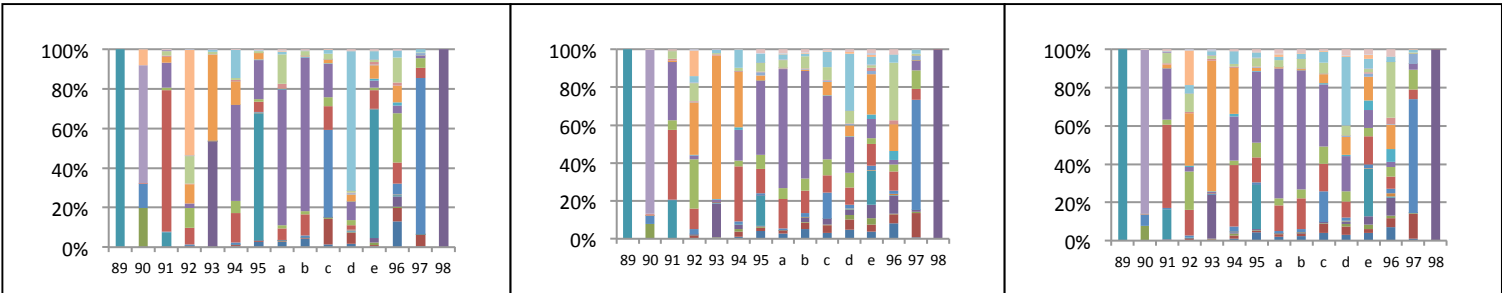
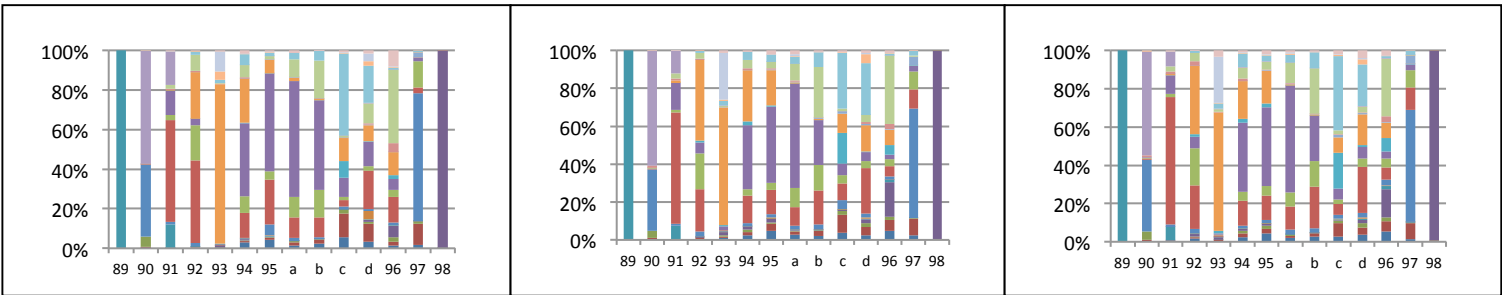
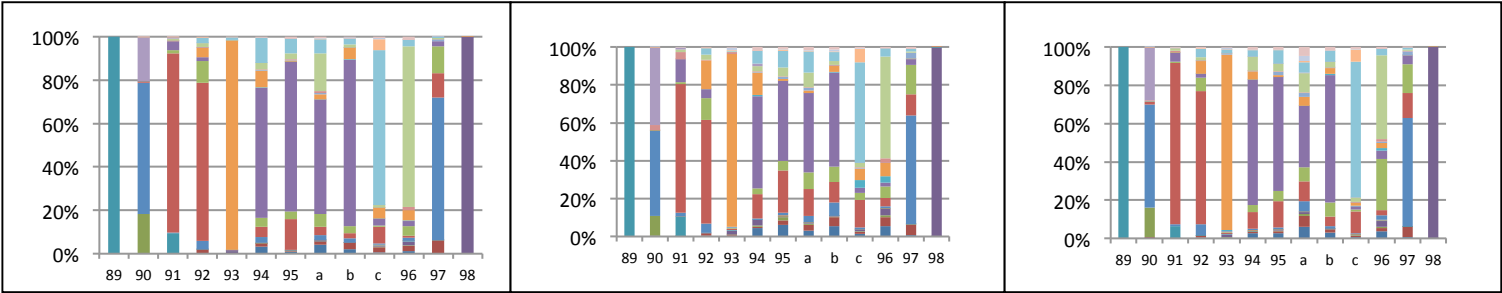


**Figure S1.** Amino acid composition of most frequently used CDR-H3 and CDR-L3 lengths. Only three representative lengths (10 – 12 amino acids; top to bottom) for VH and VL are shown. BM sample is shown on the left, SP in the middle and IM on the right. The composition for the remaining lengths will be provided upon request to the corresponding author.

**VH**



# VK



**Figure S2.** Alignment of 28-8 V regions with putative germline genes. Amino acid residues different from the germline are marked by vertical bars. An asterisk shows the mutation in the Cys residue at the CDR-L3.

## VH

**28-8** QSLEESGGDLVKPGASLTLTCTASGFSFSSNYMCWVRQAPGKGLEWIGCIYGGSSGST

**1S40/4** QSLEESGGDLVKPGASLTLTCTASGFSFSSSYMCWVRQAPGKGLEWIACIYAGSSGST

**28-8** YYASWAKGRFTISKISSTTVTLQVTSLSLAADTATYFCARNVGYNTWDLWGPATLVTVSS

**1S40/4** YYASWAKGRFTISKTSSTTVTLQMTSLTAADTATYFCAR-----YFNLWGPGLTIVTVSS

## VL

**1S41/1** AAVLTQTPSPVSVAVGGTVTINCQASQSVYNNNYLAWFQQKPGQPPKLLIYSASTL

**28-8** AAVLTQTPSPVSAAVGGTVTINCQASQSVYNNKNFAWYQQKPGQPPKLLIYEASTL

**1S18/1** AAVMTQTPSPVSVAVGGTVTINCQASQSVYNNNYLAWFQQKPGQPPKLLIYSASTL

**1S41/1** ASGVSSRFKGS GSGTQFTLTISGVQCDDAATYYCLGEFSCSSADC-AAFGGGTEVVVK

**28-8** ASGVPSRFKGS GSGTQFTLTISGVQCDDAATYYCLGEFTCSSADC-AAFGGGTEVVVK

**1S18/1** ASGVSSRFKGS GSGTQFTLTISGVQCDDAATYYCLGEFSSSSADCYNAFGGGTEVVVK