

Supporting Information

Holland et al. 10.1073/pnas.1210882109

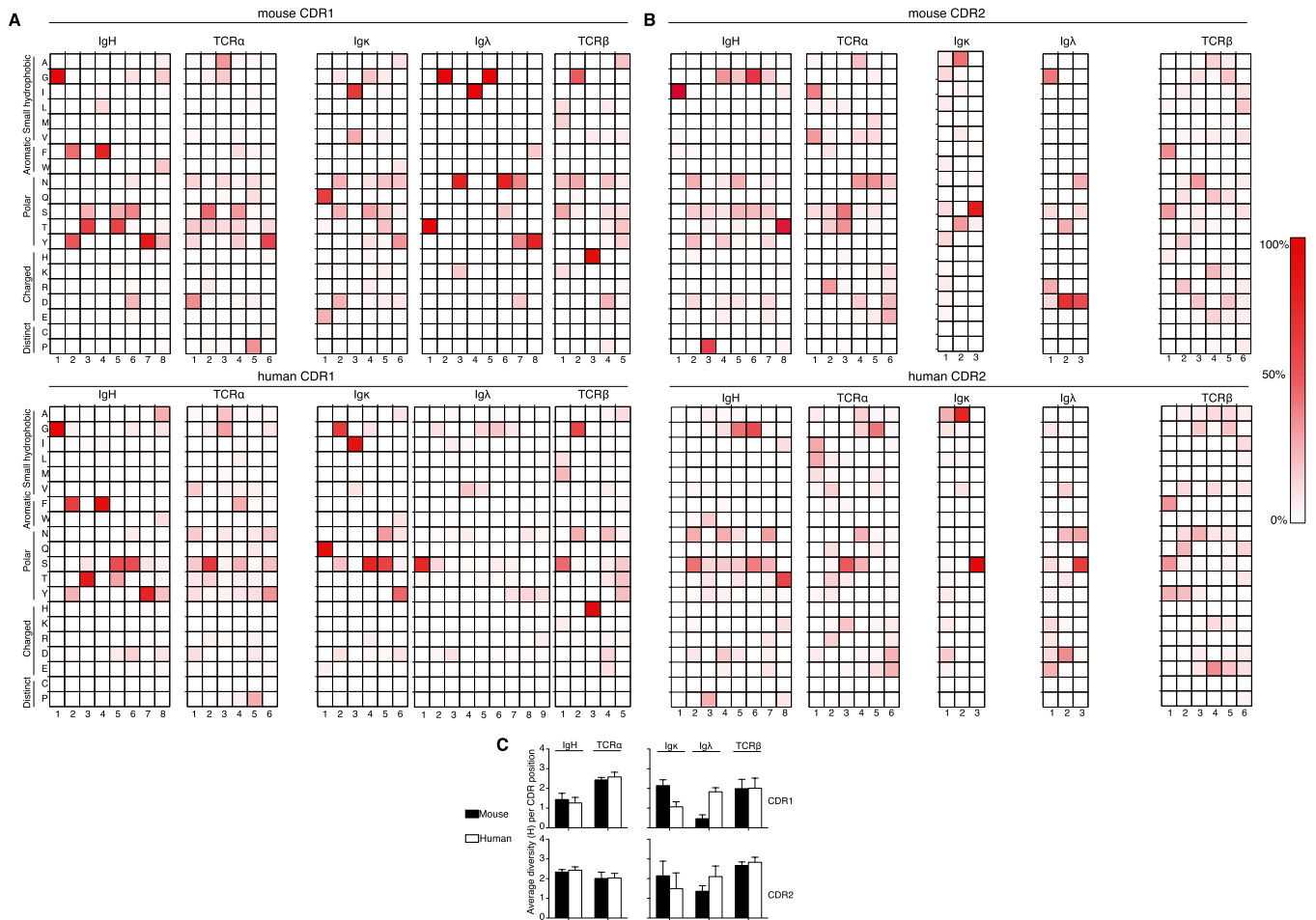


Fig. S1. Amino acid use within germ-line complementarity determining regions (CDRs) of mouse and human T-cell receptor (TCR) and Ig variable gene segments. CDR1 (A) and CDR2 (B) are shown. The percentage use of each amino acid at each position is represented by color intensity. Amino acids are grouped according to their chemical properties as indicated on the y axis. Variable (V)-segments that encode CDRs of the most common length out of the total number present in the genome were analyzed. (C) Diversity (H) values for each CDR (using the most common length) were calculated using Shannon entropy analysis in conjunction with the Protein Variability Server for each position. Error bars represent SEM. Sequences are from ImMunoGeneTics (www.imgt.org).

Table S1. Cont.

Construct	Peptide	CDR length	Nucleotide (no. of sequences)
	FRSGAP	6	TTTCGATCGGGAGCTCCT (1)
	FRSQAP	6	TTTCGAAGCCAAGCTCCT (1)
			TTTCGAAGTCAAGCTCCT (3)
			TTTCGATCACAAGCTCCT (1)
			TTTCGATCTCAAGCTCCT (1)
			TTTCGTAGTCAAGCTCCT (1)
	FRTGAP	6	TTTCGAACGGGGGCTCCT (1)
	FRVQAP	6	TTTCGGGTGCAAGCTCCT (13)
	FRVRAP	6	TTTCGGGTCCGGGCTCCT (1)
	FRYQAP	6	TTTCGGTATCAAGCTCCT (3)
	FTWQAP	6	TTTACGTGGCAAGCTCCT (1)
	HLRGAP	6	CACCTCAGGGGCGCTCCT (1)
	LRNRAP	6	CTTCGAAATCGGGCTCCT (1)
	STLQAP	6	TCGACCTTGCAAGCTCCT (1)
	FQTQAP	7	TTCCAACGCAAGCTCCT (1)
	FRGMQAP	7	TTTCGAGGGATGCAAGCTCCT (1)
	FRKGQAP	7	TTTCGAAAGGGCAAGCTCCT (1)
	FRKKKAP	7	TTTCGAAAAAAAAAAGCTCCT (1)
	FRNIGAP	7	TTTCGAAATATAGGGGCTCCT (1)
	FRNSQAP	7	TTTCGTAAGTCTCAAGCTCCT (2)
	FRNSRTP	7	TTTCGAAATAGTAGGACTCCT (1)
	FRNTGAP	7	TTTCGAAATACGGGAGCTCCT (3)
	FRRGKAP	7	TTTCGAAGGGGAAAGCTCCT (3)
	FRTDGAP	7	TTTCGAACAGACGGAGCTCCT (2)
	FRIHPQAP	8	TTTCGAATCCACCCTCAAGCTCCT (1)
	FRIPGAP	8	TTTCGAATCCGGGGGGGCTCCT (1)
	FRNGLQAP	8	TTTCGAAATGGGGTTCGAAGCTCCT (1)
	FRNIPGAP	8	TTTCGAAATATCCAGGGGCTCCT (1)
	FRNIRGAP	8	TTTCGAAATATAAGGGGAGCTCCT (2)
	FRNISKAP	8	TTTCGAAATATTTCCAAGCTCCT (2)
	FRNIVGAP	8	TTTCGAAATATCGTAGGGGCTCCT (3)
	FRNPLRAP	8	TTTCGAAATCCCTTGGGAGCTCCT (1)
	FRTPLQAP	8	TTTCGAACCCCTTTGCAAGCTCCT (2)
	FRTSLQAP	8	TTTCGAACCTCCCTCCAAGCTCCT (2)
	FLTSTLQAP	9	TTCCCTGACTTCGACCTTGCAAGCTCCT (1)
	FRKRPLQAP	10	TTTCGAAAGAGACCCCTCCAACAAGCTCCT (6)
CDR1 WT	SGHHV	5	TCAGGACACCACGTT (5)
	SGHKV	5	TCAGGACATAAGGTT (2)
	SGHNV	5	TCAGGACATAATGTT (2)
	SGHRV	5	TCAGGACATAGAGTT (3)
			TCAGGACATAGGGTT (6)
	SGHSR	5	TCAGGACATAGTCCG (1)
	SGHSV	5	TCAGGACATAGTGTT (5)
	SGHTV	5	TCAGGACATACCGTT (1)
	SGLHV	5	TCAGGACTCCACGTT (1)
	SGQAV	5	TCAGGACAAGCTGTT (2)
			TCAGGACAGGCTGTT (1)
	SGQNV	5	TCAGGACAAAATGTT (2)
	SGRHA	5	TCAGGACGCCATGCT (1)
	SRRAV	5	TCACGTCGAGCTGTT (1)
	CIIVAV	6	TGTATCATTGTGGCTGTT (1)
	CNIVAV	6	TGTAACATTGTGGCTGTT (1)
	SGHFAV	6	AGTGGACATTTTGTCTGTT (1)
	SGHHAV	6	TCAGGACACCACGCTGTT (3)
	SGHNGV	6	TCAGGACATAACGGTGT (2)
	SGHNHV	6	TCAGGACATAACCATGTT (1)
	SGHPAV	6	TCAGGACATCCCGCTGTT (1)
	SGHRAV	6	TCAGGACATAGAGCTGTT (1)
			TCAGGACATAGAGCTGTT (22)
			TCAGGACATAGGGCTGTT (6)
	SGHRDV	6	TCAGGACATAGAGATGTT (1)
			TCAGGACATAGGGATGTT (1)
	SGHRGV	6	TCAGGACATAGAGGTGTT (2)
	SGHRV	6	TCAGGACATAGAAGTGT (2)

Table S1. Cont.

Construct	Peptide	CDR length	Nucleotide (no. of sequences)
CDR1 Δ			TCAGGACATAGATCTGTT (5)
			TCAGGACATAGGTCTGTT (1)
	SGHRTV	6	TCAGGACATAGAAGTCTGTT (1)
	SGHRWV	6	TCAGGACATAGATGGGTT (1)
	SGHSAV	6	TCAGGACATAGCGCTGTT (2)
			TCAGGACATTCCGCTGTT (1)
	SGHSGV	6	TCAGGACATAGTGGGTT (4)
	SGHSHV	6	TCAGGACATTCCCATGTT (1)
	SGHSKV	6	TCAGGACATAGTAAAGTT (1)
	SGHSNV	6	TCAGGACATAGTAATGTT (1)
	SGHSRV	6	TCAGGACATAGCCGGGTT (1)
			TCAGGACATAGTAGGGTT (4)
			TCAGGACATAGTCGCGTT (4)
	SGHSSV	6	TCAGGACATAGTAGTGT (1)
	SGHSTV	6	TCAGGACATAGTACCGTT (4)
			TCAGGACATAGTACTGTT (6)
	SGHTAV	6	TCAGGACATACCGCTGTT (2)
	SGLCVF	6	TCAGGACTTTGTGTTTTC (1)
	SGQIRV	6	TCAGGACAAATCCGTT (2)
	SGQNEV	6	TCAGGACAGAACGAGTT (1)
	SGQTSV	6	TCAGGACAAACCTCTGTT (2)
	SGSPDV	6	TCAGGATCCCCTGACGTT (3)
	SGSRAV	6	TCAGGATCTCGCGCTGTT (1)
	SADGV	5	TCGGCAGACGGAGTT (1)
	SAFGV	5	TCAGCATTGGAGTT (1)
	SAGRV	5	TCGGCAGGTCGAGTT (1)
	LAEGGV	6	TTGGCAGAGGGAGGAGTT (1)
	LAGSVV	6	TTGGCAGGATCAGTAGTT (1)
	SAEGGV	6	TCGGCAGAGGGAGGAGTT (12)
	SAESGV	6	TCGGCAGAATCAGGAGTT (2)
	SAEVGV	6	TCGGCAGAGGTAGGAGTT (1)
	SAFFRF	6	TCGGCATTTCAGAGTT (1)
	SAFLGV	6	TCGGCATTTCAGGAGTT (1)
	SAFSGF	6	TCTGCATTTTCAGGATTT (1)
	SAFSVF	6	TCGGCATTTCAGTATTT (1)
	SAGAGV	6	TCGGCAGGGCAGGAGTT (3)
	SAGFGV	6	TCGGCAGGATTTGGGTT (2)
	SAGLKF	6	TCGGCAGGTTTAAAATTT (1)
	SAGLKV	6	TCGGCAGGTTTAAAAGTT (1)
	SAGLRV	6	TCGGCAGGATGAGAGTT (1)
	SAGRGV	6	TCGGCAGGAAGGGGAGTT (3)
	SAGSGV	6	TCGGCAGGATCAGGAGTT (7)
			TCGGCAGGATCCGGAGTT (3)
			TCGGCAGGATCGGGAGTT (7)
			TCGGCAGGATCTGGAGTT (5)
			TCGGCAGGATCTGGGTT (2)
			TCTGCAGGATCAGGAGTT (1)
			TCTGCAGGATCTGGAGTT (1)
	SAGSKL	6	TCGGCAGGATCAAAACTT (1)
	SAGSKV	6	TCGGCAGGATCAAAAAGTT (8)
			TCGGCAGGATCGAAAAGTT (3)
	SAGSRV	6	TCGGCAGGATCAAGAGTT (7)
		TCGGCAGGATCCAGAGTT (4)	
		TCGGCAGGATCCCGGTT (3)	
		TCGGCAGGATCGAGAGTT (2)	
		TCGGCAGGATCGAGGGTT (1)	
		TCGGCAGGATCTCGAGTT (1)	
SAGSSV	6	TCGGCAGGATCAAGTGT (2)	
SAGSVV	6	TCGGCAGGATCAGTAGTT (16)	
SALFRV	6	TCGGCATTATTCAGAGTT (1)	
SAVGGV	6	TCGGCAGTTGGAGGAGTT (1)	
SGSPGV	6	TCGGGGAGTCCGGGAGTT (1)	
VADVGV	6	GTCGCAGATGTAGGAGTT (1)	

Sequences shaded in gray represent regeneration of the original complementarity determining region (CDR) template sequence.