Supporting Information Figure Legends

Supporting Information Fig 1. Deletion of YY1 in *Igκ AID yy1*$^{f/f}$ splenic B cells affects CSR and nuclear AID levels. Splenic B cells isolated from *Igκ AID yy1*$^{f/f}$ mice were either mock treated, or treated with TAT-CRE to delete YY1, then grown in LPS plus IL4 for 3 to 4 days. (A) CSR was evaluated by FACS with anti-IgG1 antibody. A representative FACS plot is shown. This experiment was repeated over six times. (B) Quantitation of CSR data in B cells from six *Igκ AID yy1*$^{f/f}$ mice either mock treated, or treated with TAT-CRE. Error bars show the standard deviation of the mean, and triple asterisks indicate $p<0.001$ in a two tailed T-test. (C) Nuclear protein levels in response to YY1 deletion were assayed by immunoblotting with anti-AID, anti-YY1 and anti-TBP antibodies. A representative blot from 3 experiments is shown. (D). Quantitation of three experiments showing reduction of AID/TBP ratios after TAT-CRE treatment. Error bars represent standard deviation of the mean and the asterisk represents $p<0.05$ in a two tailed T-test.

Supporting Information Fig. 2. Mutation frequencies at Sμ in splenic B cells from six individual *Igκ AID yy1*$^{f/f}$ mice are presented. Isolated splenic B cells were either mock treated, or treated with TAT-CRE to delete the *yy1* gene, then induced with LPS plus IL4 in culture for 4 days. DNA was isolated, the Sμ region was amplified by PCR, cloned, and subjected to Sanger sequencing. Each bar represents the average of the mutation frequencies of each sequenced clone.

The number of clones, base pairs sequenced, and mutations observed are presented in Supporting Information Table 1. Asterisks represent $p < 0.05$ in a two tailed T-test.

Supporting Information Fig. 3. Variant allele frequency (VAF) at the Sμ locus is shown in mock and TAT-CRE treated samples compared to the AID knockout background. BAM files from 6 independent NGS runs were merged. Line graphs show VAFs for all sequenced positions. Mock samples are red, TAT-CRE samples are green and AID KO samples are black. P values were calculated using the Wilcoxon test. The box plot shows VAF distribution in mock, TAT-CRE and AID KO samples. Graphs depict minimum and maximum values (whiskers), first and third quartile, and the median. Dots represent outliers. P values were calculated using the Mann-Whitney test.

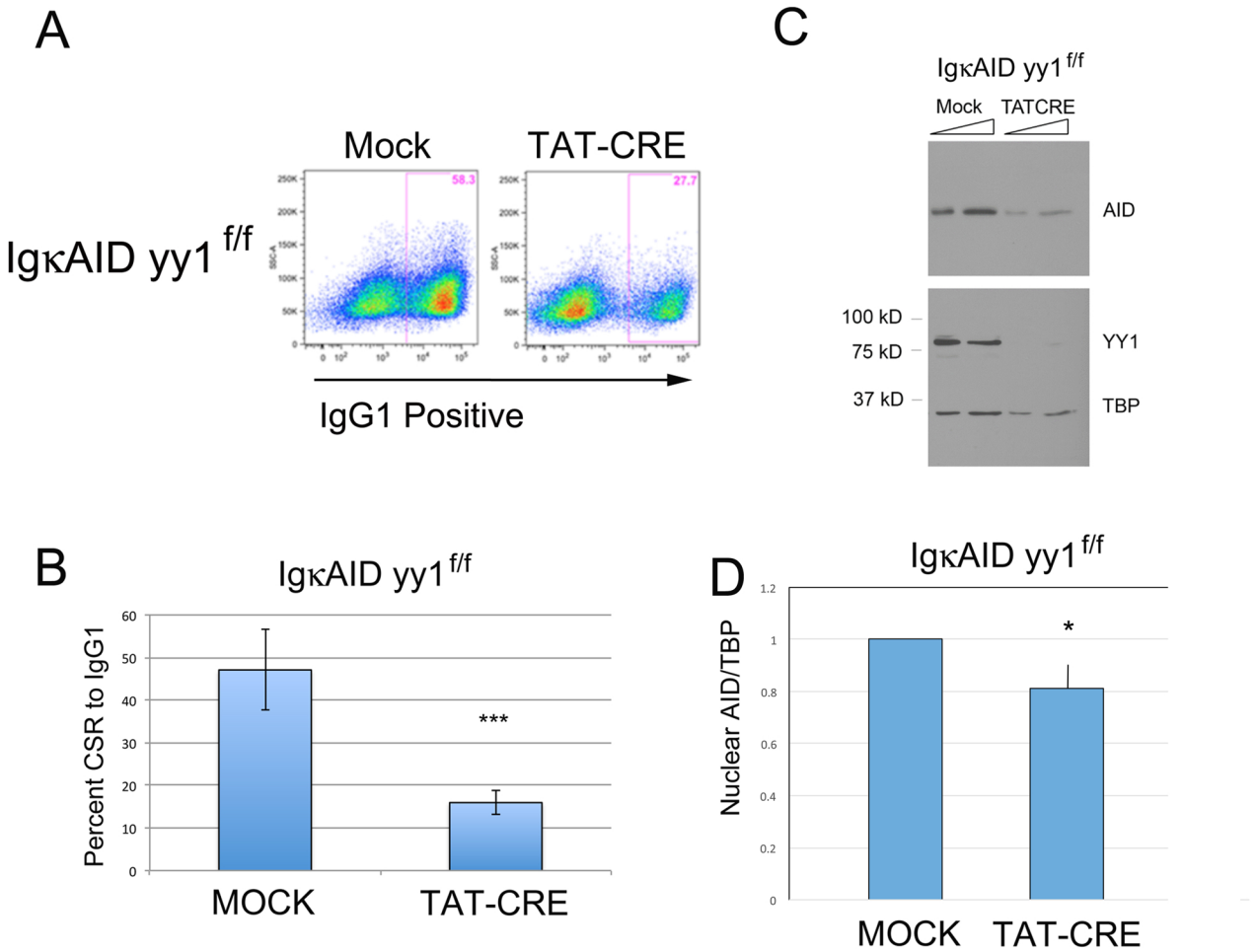
Supporting Information Fig. 4. Variant allele frequency (VAF) at the Jh4 locus is shown in mock and TAT-CRE treated samples compared to the AID knockout background. BAM files from 6 independent NGS runs were merged. Line graphs show VAFs for all sequenced positions. Mock samples are red, TAT-CRE samples are green and AID KO samples are black. P values were calculated using the Wilcoxon test. The box plot shows VAF distribution in mock, TAT-CRE and AID KO samples. Graphs depict minimum and maximum values (whiskers), first and third quartile, and the median. Dots represent outliers. P values were calculated using the Mann-Whitney test.

Supporting Information Fig. 5.

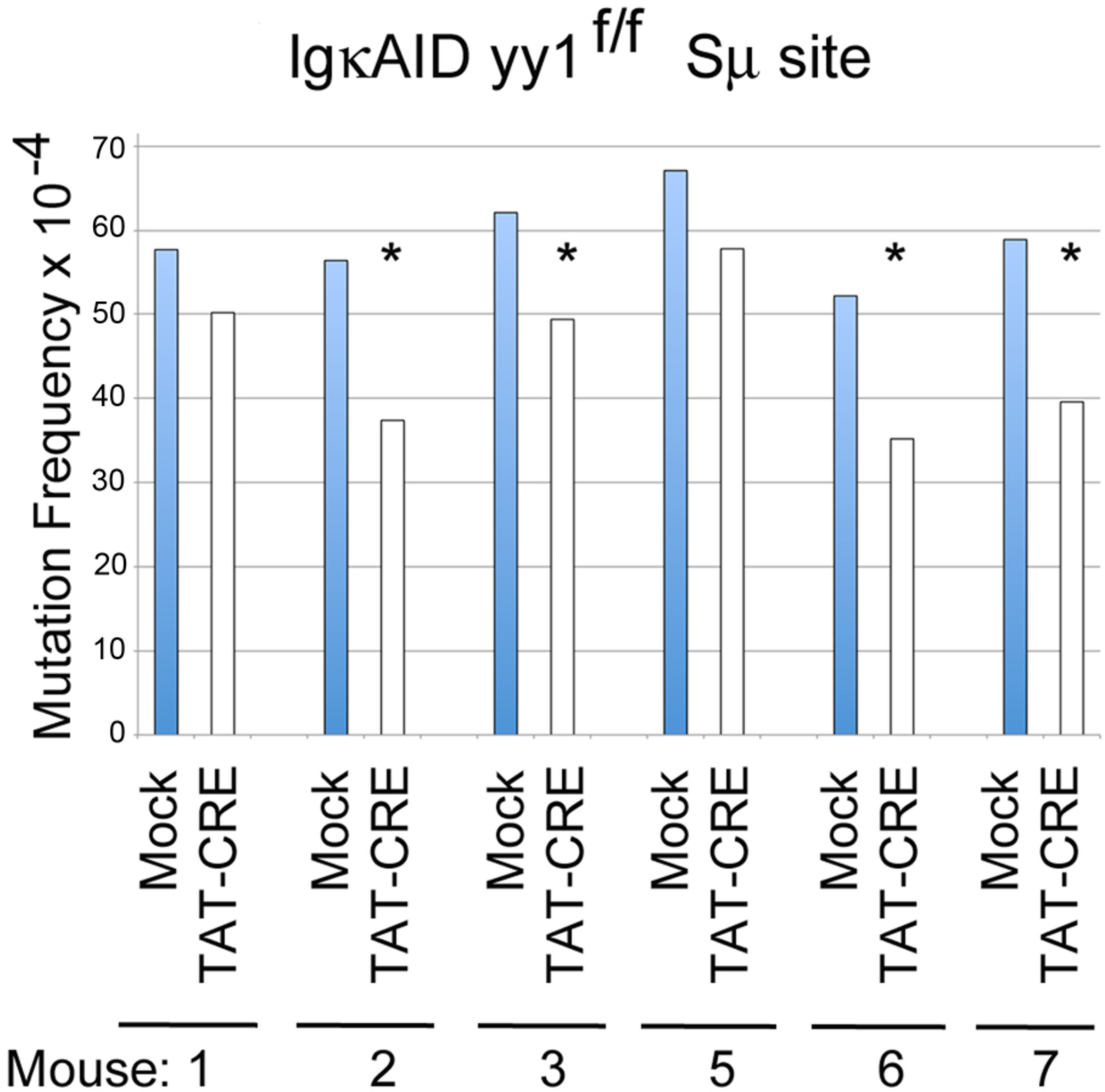
Variant allele frequency (VAF) at the Cd83 locus is shown in mock and TAT-CRE treated samples compared to the AID knockout background. BAM files from 6 independent NGS runs were merged. Line graphs show VAFs for all sequenced positions. Mock samples are red, TAT-CRE samples are green and AID KO samples are black. P values were calculated using the Wilcoxon test. The box plot shows VAF distribution in mock, TAT-CRE and AID KO samples. Graphs depict minimum and maximum values (whiskers), first and third quartile, and the median. Dots represent outliers. P values were calculated using the Mann-Whitney test.

Supporting Information Fig. 6. Evaluation of the effect of wild-type YY1 and YY1 mutants on the level of AID ubiquitination. HEK293T cells were transfected with CMV-FlagAID plus either GALYY1 or various YY1 mutants. Two days after transfection cells were treated with MG132 to inhibit proteosomal degradation, nuclear extracts were prepared, heated to dissociate protein interactions, immunoprecipitated with anti-Flag, then subjected to western blot with anti-ubiquitin antibody. Blots were subsequently stripped and probed with anti-Flag to indicate the level of AID. Level of ubiquitination (percent or fold difference) was calculated using densiometric measurement of ubiquitin signal normalized to the level of immunoprecipitated AID. (A) Nine independent experiments with wild

type YY1 fused to the GAL4 DNA binding domain. Blue bars show the level of AID ubiquitination without co-transfected YY1 (defined as 100%), and red bars show the relative level in the presence of YY1. (B) Seven independent experiments with the GALYY1 1-200 mutant containing YY1 amino acids 1-200. Blue bars show the level of AID ubiquitination without co-transfected YY1 1-200 (defined as 100%), and red bars show the relative level in the presence of YY1 1-200. (C) Experiments with GALYY1 1-200 compared to GALYY1 1-200Δ16-80. The level of AID ubiquitination observed with GALYY1 1-200 was defined as 1, and the level of AID ubiquitination with GALYY1 1-200Δ16-80 was calculated relative to this value.

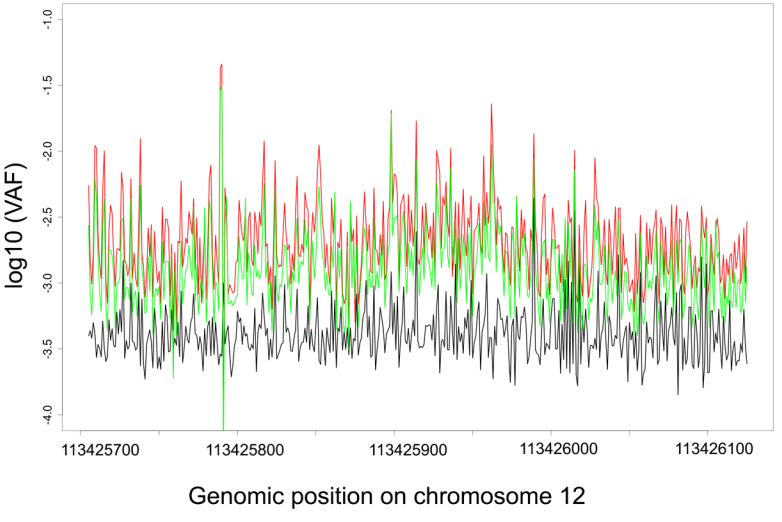


Supporting Information Fig. 1

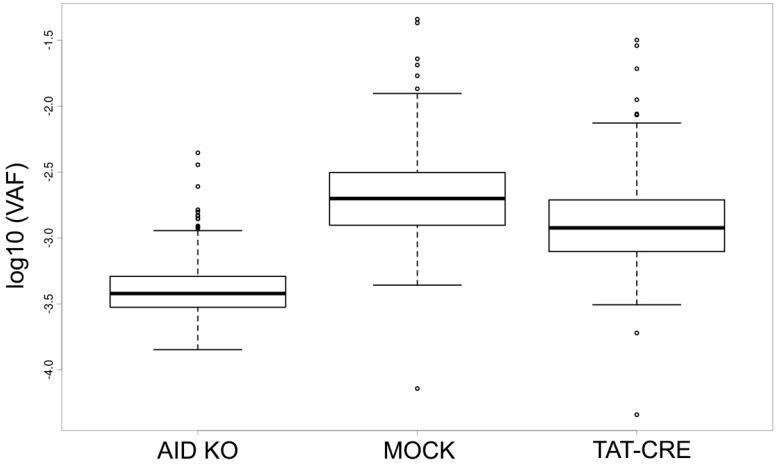


Supporting Information Fig. 2

Sμ locus



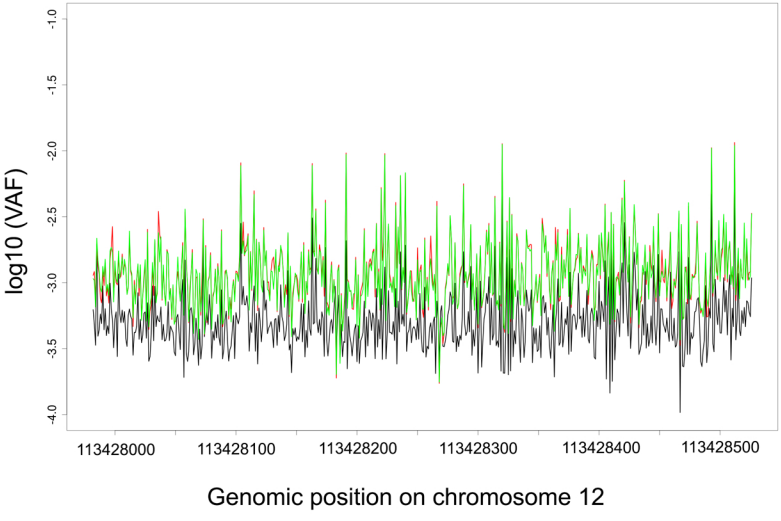
AID KO vs. MOCK p < 0.001
AID KO vs. TAT-CRE p < 0.001
MOCK vs. TAT-CRE p < 0.001



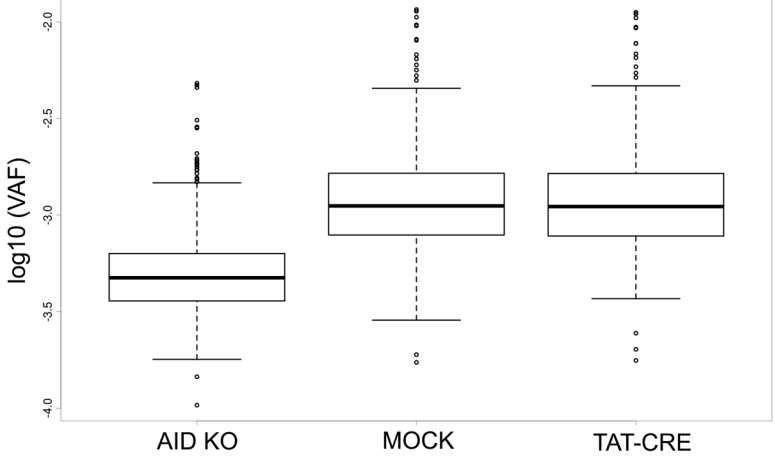
AID KO vs. MOCK p < 0.001
AID KO vs. TAT-CRE p < 0.001
MOCK vs. TAT-CRE p < 0.001

Supporting Information Fig. 3

Jh4 locus



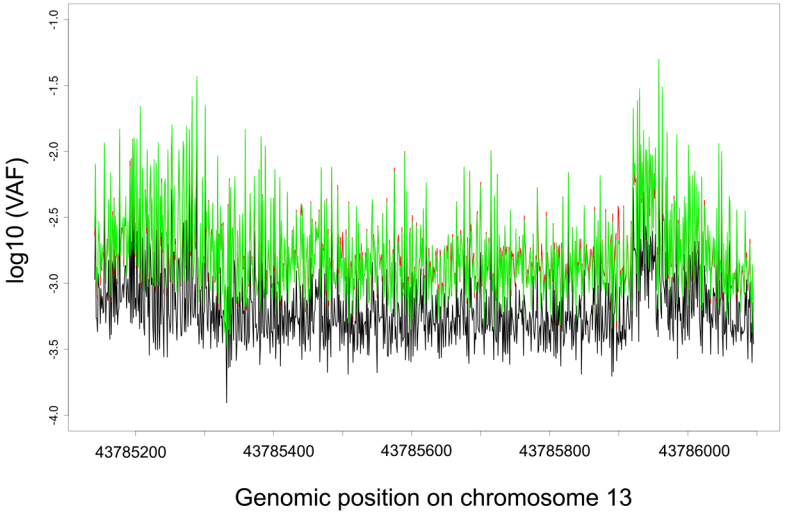
AID KO vs. MOCK p < 0.001
AID KO vs. TAT-CRE p < 0.001
MOCK vs. TAT-CRE p = 0.027



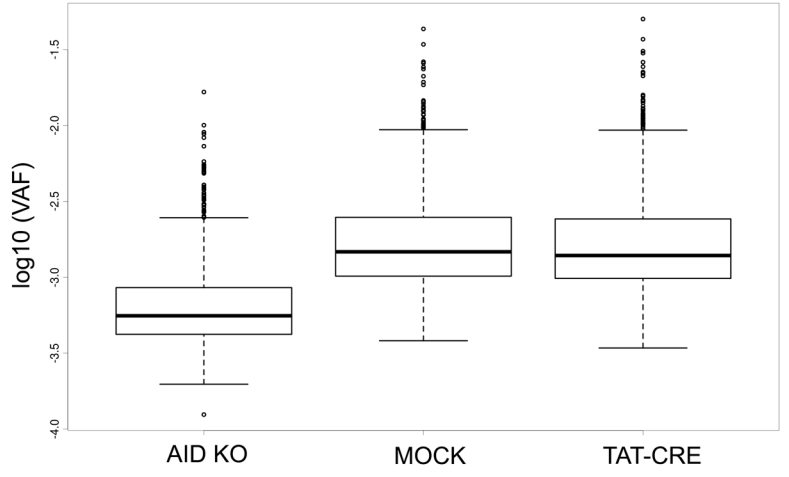
AID KO vs. MOCK p < 0.001
AID KO vs. TAT-CRE p < 0.001
MOCK vs. TAT-CRE p = 0.625

Supporting Information Fig. 4

Cd83 locus

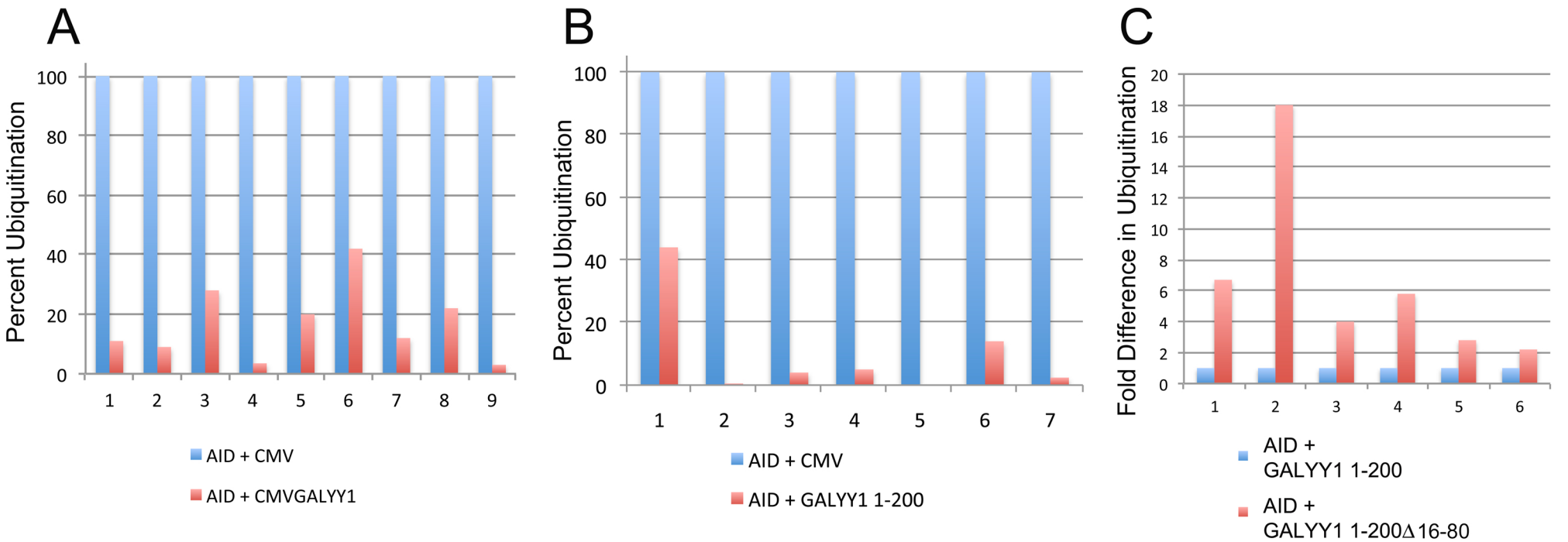


AID KO vs. MOCK $p < 0.001$
AID KO vs. TAT-CRE $p < 0.001$
MOCK vs. TAT-CRE $p < 0.001$



AID KO vs. MOCK $p < 0.001$
AID KO vs. TAT-CRE $p < 0.001$
MOCK vs. TAT-CRE $p = 0.260$

Supporting Information Fig. 5



Supporting Information Fig. 6

Supporting Information Table 1.
Mutation Frequency at Sμ in Individual Mice

Individual Mice	BP Sequenced	Mutation Numbers	Individual Clones
Mock 1	8,050	46	16
Mock 2	12,226	67	23
Mock 3	17,648	108	34
Mock 5	16,959	115	32
Mock 6	21,503	112	40
Mock 7	21,752	128	40
Total	98,138	576	185
Mutation Frequency	58.7×10^{-4}		

Individual Mice	BP Sequenced	Mutation Numbers	Individual Clones
CRE 1	9,505	44	19
CRE 2	10,897	42	22
CRE 3	21,818	115	44
CRE 5	18,881	109	35
CRE 6	26,054	92	49
CRE 7	20,479	86	40
Total	107,479	488	209
Mutation Frequency	45.4×10^{-4}		

Supporting Information Table 2.
Mutation Frequency at the TACI gene in Individual Mice

Individual Mice	BP Sequenced	Mutation Numbers	Individual Clones
Mock 1	11,300	2	22
Mock 2	8,836	0	22
Mock 3	9,447	0	20
Total	29,583	2	64
Mutation Frequency	0.67×10^{-4}		

Individual Mice	BP Sequenced	Mutation Numbers	Individual Clones
CRE 1	4,582	1	8
CRE 2	16,228	1	38
CRE 3	11,266	0	25
Total	32,076	2	71
Mutation Frequency	0.62×10^{-4}		

Supporting Information Table 3.
Single Nucleotide Variant Frequency at Sμ, Jh4 and Cd83 loci in Individual Mice

Individual Mice – Sμ locus	BP Sequenced	SNV Numbers
Mock 0	99,778,879	79,239
Mock 1	120,356,081	107,799
Mock 2	57,237,341	87,514
Mock 3	198,973,861	183,599
Mock A	55,619,471	9,357
Mock B	56,915,133	6,275
Total	588,880,766	473,783
SNV frequency	8.05×10^{-4}	

Individual Mice – Sm locus	BP Sequenced	Mutation Numbers
CRE 0	211,355,005	64488
CRE 1	210,238,001	69563
CRE 2	242,835,784	104390
CRE 3	204,392,640	66222
CRE A	81,461,997	5186
CRE B	52,891,753	1846
Total	1,003,175,180	311,695
SNV frequency	3.11×10^{-4}	

Individual Mice – Jh4 locus	BP Sequenced	SNV Numbers
Mock 0	417,923,425	26,871
Mock 1	213,422,572	16,914
Mock 2	180,102,614	12,584
Mock 3	298,623,948	22,715
Mock A	156,629,498	11,421
Mock B	247,871,918	6,418
Total	1,514,573,975	96,923
SNV frequency	0.64×10^{-4}	

Individual Mice – Jh4 locus	BP Sequenced	Mutation Numbers
CRE 0	356,204,935	5,977
CRE 1	219,764,547	3,907
CRE 2	217,176,258	3,068
CRE 3	158,963,302	2,247
CRE A	129,448,294	812
CRE B	283,770,227	1,047
Total	1,365,327,563	17,058
SNV frequency	0.125×10^{-4}	

Individual Mice – Cd83 locus	BP Sequenced	SNV Numbers
Mock 0	55,364,207	1,084
Mock 1	108,289,364	3,717
Mock 2	83,037,139	3,241
Mock 3	88,228,089	4,651
Mock A	102,008,318	562
Mock B	143,518,757	364
Total	580,445,874	13,619
SNV frequency	0.235×10^{-4}	

Individual Mice – Cd83 locus	BP Sequenced	Mutation Numbers
CRE 0	45,070,336	374
CRE 1	98,460,695	1,489
CRE 2	77,610,020	1,275
CRE 3	83,801,858	1,631
CRE A	183,932,911	282
CRE B	149,781,465	113
Total	638,657,285	5,164
SNV frequency	0.081×10^{-4}	

Supporting Information Table 4.
Genomic regions analyzed. Regions with coverage higher than 10,000 reads/position were considered.

Gene	Chromosome	Start position	End position	Strand orientation	Median per base coverage	
					Mock	TATCRE
Sμ	chr12	113425705	113426125	reverse strand	148,566	296,069
Jh4	chr12	113427982	113428526	reverse strand	156,022	129,194
Cd83	chr13	43785141	43786095	forward strand	111,045	104,886

Reference genome GRCm38