Supplementary Table 1

Supplementary to Table 1. Raw mutation numbers for each dataset.

bases/sequence	523	124	114	158	127

MUTATION FROM:

UNG^{-}M^{-}
10 mice
368 seq's

MUTATION TO:	C	G	T	A
C	0	3	2	0
G	1	0	0	0
T	626	5	0	2
A	2	357	1	0
999	**629**	**365**	**3**	**2**

UNG^{+}M^{-}
11 mice
514 seq's

MUTATION TO:	C	G	T	A
C	0	43	2	1
G	8	0	0	5
T	257	17	0	9
A	6	179	0	0
527	**271**	**239**	**2**	**15**

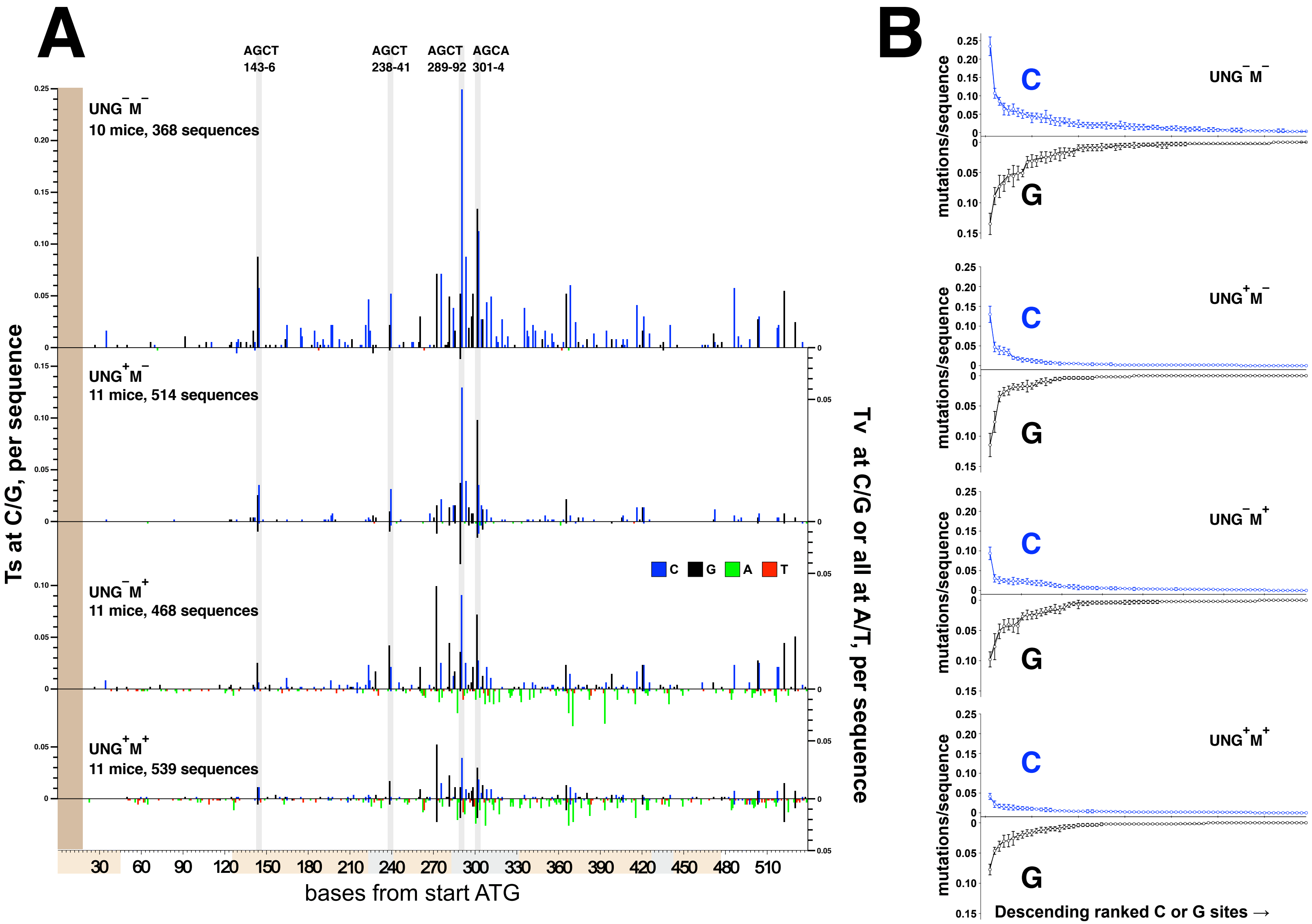
UNG^{-}M^{+}
11 mice
468 seq's

MUTATION TO:	C	G	T	A
C	0	5	33	62
G	6	0	15	99
T	270	7	0	97
A	1	324	29	0
948	**277**	**336**	**77**	**258**

UNG^{+}M^{+}
11 mice
539 seq's

MUTATION TO:	C	G	T	A
C	0	53	49	70
G	25	0	27	111
T	124	30	0	155
A	19	166	38	0
867	**168**	**249**	**114**	**336**

Supplementary Figure 1



Supplementary Figure 1 (previous page)

Supplementary to Figure 2. (A) Histograms of mutation frequencies at each base in the sequenced window of 523bp. Transitions at C (blue) or G (black) are shown above the x-axis. Transversions at C (blue) or G (black) and mutations at A (green) or T (red) are shown below the x-axis. The brown stripe indicates a 5’ region in which mutations were not collated because it over-lapped a secondary PCR primer. Grey dashed lines identify the four AGCW motifs present in the starting sequence (see Results and Discussion). Beige boxes at bottom indicate the L and VJ exons, while grey boxes indicate the three CDRs, according to IMGT (IMGT®, the international ImMunoGeneTics information system® http://www.imgt.org (founder and director: Marie-Paule Lefranc, Montpellier, France)). (B) Reproducibility of hotspot-targeting in the datasets. C (top, blue) or G (bottom, black) mutation sites were sorted as most to least mutated site (i.e. decendingly ranked) in each dataset. The variation between hosts in mean mutation at each sorted site is indicated by the error bars, which depict s.e.m.

Supplementary Figure 2 (below)

Supplementary to Figure 5. Mutation skylines for cells over-expressing (A) UNG2 or UNG2-GFP, or (B) SMUG1 or SMUG1-GFP. Data are presented as in Supplementary Figure 1A.