

1 **SUPPLEMENTARY MATERIALS for**

2

3 **Human APOBEC3B promotes tumor heterogeneity *in vivo* including signature mutations**  
4 **and metastases**

5

6 Cameron Durfee<sup>1</sup>, Nuri Alpay Temiz<sup>2,3,\*</sup>, Rena Levin-Klein<sup>3,\*</sup>, Prokopios P. Argyris<sup>4,\*</sup>, Lene  
7 Alsøe<sup>5,6</sup>, Sergio Carracedo<sup>5</sup>, Alicia Alonso de la Vega<sup>7,8</sup>, Joshua Proehl<sup>1</sup>, Anna M. Holzhauer<sup>3</sup>,  
8 Zachary J. Seeman<sup>3</sup>, Yu-Hsiu T. Lin<sup>1</sup>, Rachel I. Vogel<sup>3,9</sup>, Rocio Sotillo<sup>7,8</sup>, Hilde Nilsen<sup>5,6</sup>, and  
9 Reuben S. Harris<sup>1,10</sup>,

10

11 <sup>1</sup> *Department of Biochemistry and Structural Biology, University of Texas Health San Antonio,*  
12 *San Antonio, Texas, USA, 78229*

13 <sup>2</sup> *Institute for Health Informatics, University of Minnesota, Minneapolis, Minnesota, USA, 55455*

14 <sup>3</sup> *Masonic Cancer Center, University of Minnesota, Minneapolis, Minnesota, USA, 55455*

15 <sup>4</sup> *Division of Oral and Maxillofacial Pathology, College of Dentistry, Ohio State University,*  
16 *Columbus, Ohio, USA, 43210*

17 <sup>5</sup> *Department of Clinical Molecular Biology, University of Oslo, 0318, Oslo, Norway*

18 <sup>6</sup> *Department of Microbiology, Oslo University Hospital, N-0424 Oslo, Norway*

19 <sup>7</sup> *Division of Molecular Thoracic Oncology, German Cancer Research Center (DKFZ), Im*  
20 *Neuenheimer Feld 280, 69120 Heidelberg, Germany*

21 <sup>8</sup> *Translational Lung Research Center Heidelberg (TRLG), German Center for Lung Research*  
22 *(DZL)*

23 <sup>9</sup> *Department of Obstetrics, Gynecology, and Women's Health, University of Minnesota,*  
24 *Minneapolis, Minnesota, USA*

25 <sup>10</sup> *Howard Hughes Medical Institute, University of Texas Health San Antonio, San Antonio, Texas,*  
26 *USA, 78229*

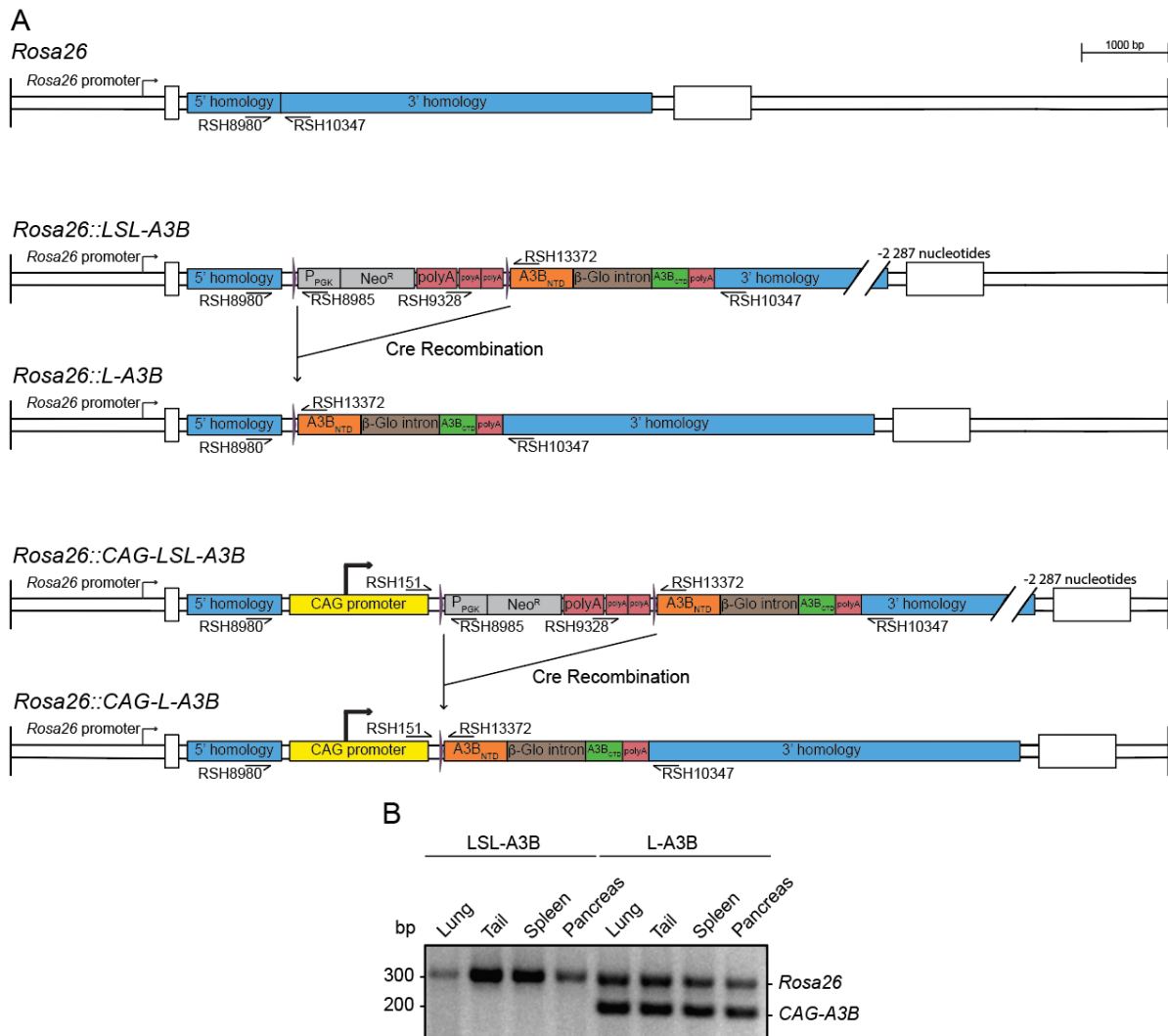
27 \* Equal secondary contributions.

28

29 Corresponding author: [rsh@uthscsa.edu](mailto:rsh@uthscsa.edu)

30

31 Supplementary Materials: Figures S1-S12 and Table S1



33

34 **Figure S1. Construction and validation of mice expressing APOBEC3B, related to Figure 1**

35 (A) To-scale schematics of the WT *Rosa26* locus and, following gene targeting, the fully intact

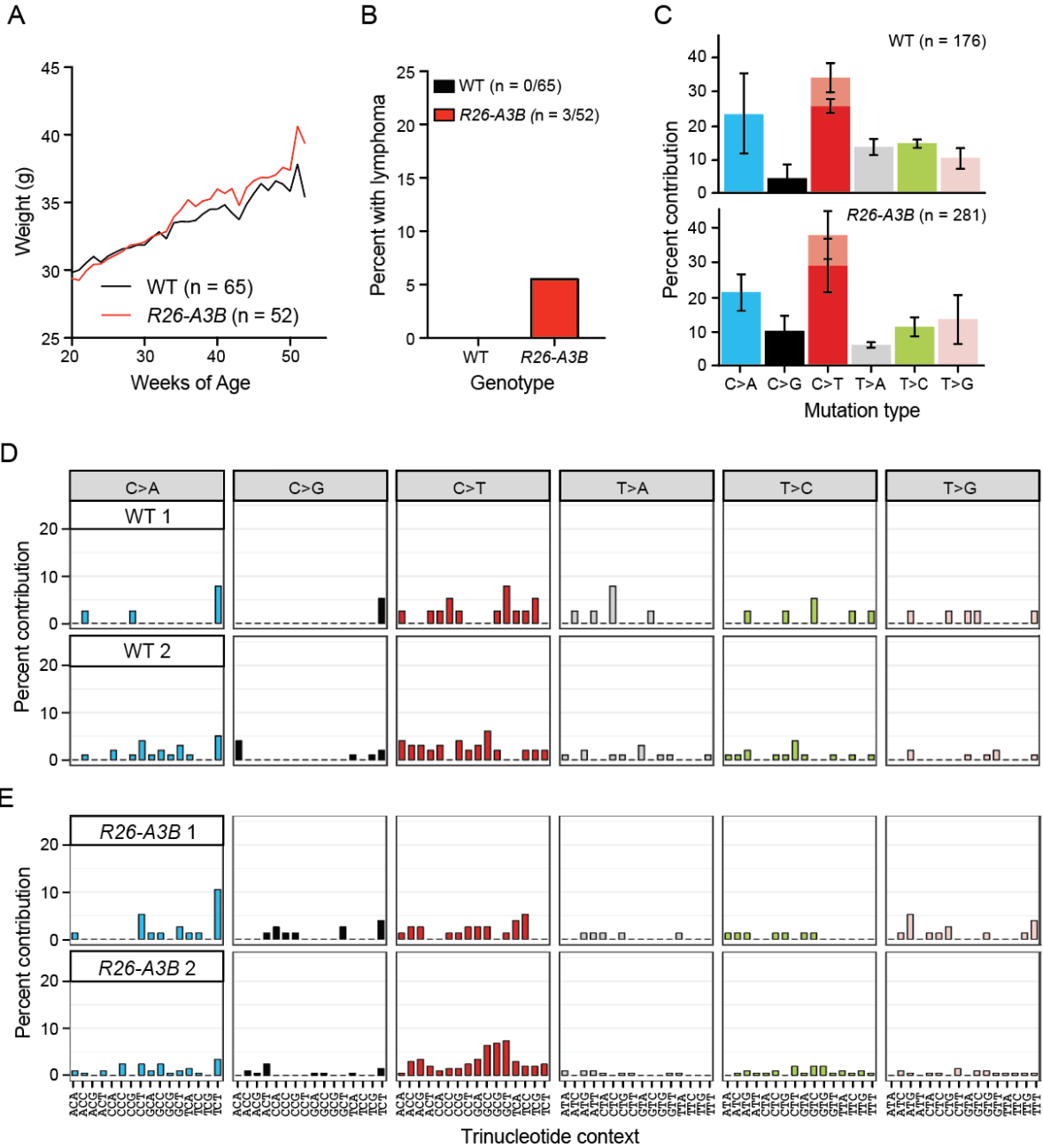
36 *Rosa26::LSL-A3B* and *Rosa26::CAG-LSL-A3B* minigene alleles. Cre-mediated recombination

37 excises the *loxP*-STOP-*loxP* (LSL) transcriptional stop cassette and allows the *Rosa26* promoter

38 or the *CAG* promoter to drive human *A3B* expression (R26-A3B and CAG-A3B, respectively).

39 (B) PCR genotyping distinguishes the wildtype *Rosa26* locus from each knock-in allele using the

40 indicated primer sets.



41

42 **Figure S2. Low levels of human A3B have limited mutagenic or tumorigenic capacity in mice,**  
 43 **related to Figure 3**

44 (A) Average of weights in grams taken weekly from 20-52 weeks of age for each indicated  
 45 genotype. 65 WT mice and 52 *Rosa26-A3B* mice are represented in this group.

46 (B) Percent of mice between 11 and 23 months that had a lymphoma upon euthanasia and necropsy

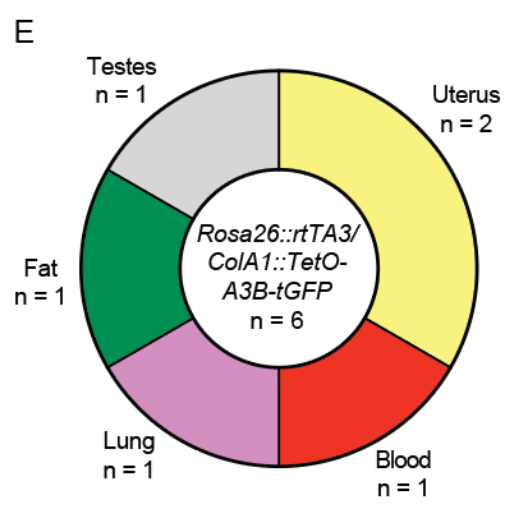
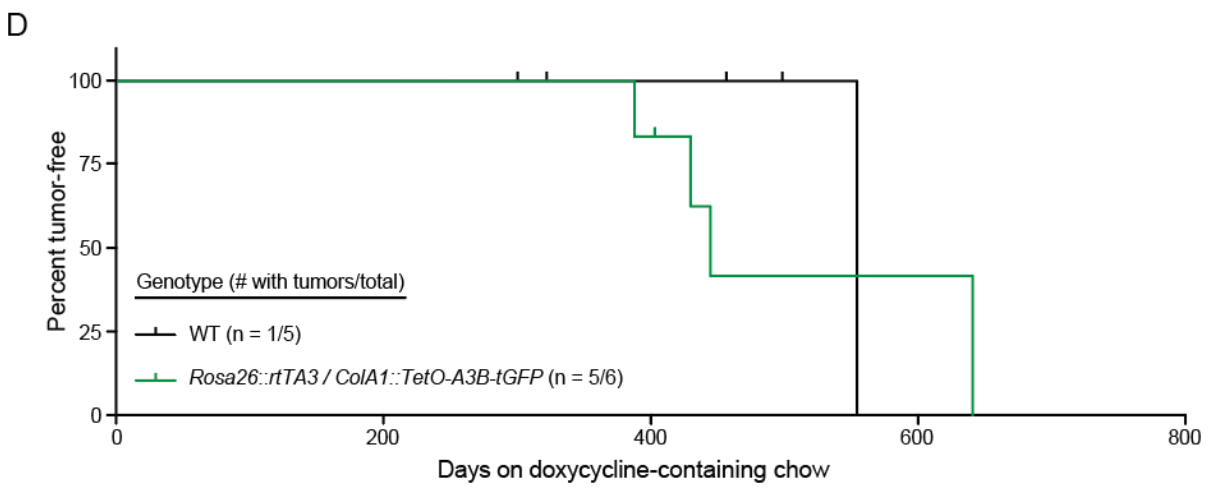
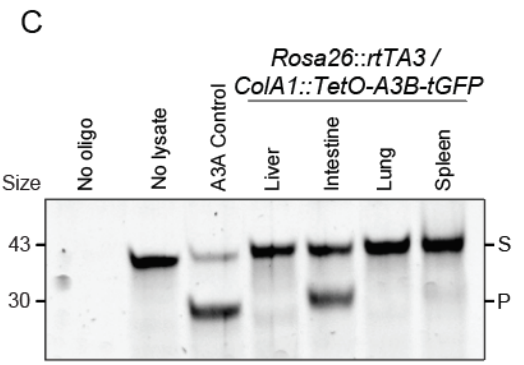
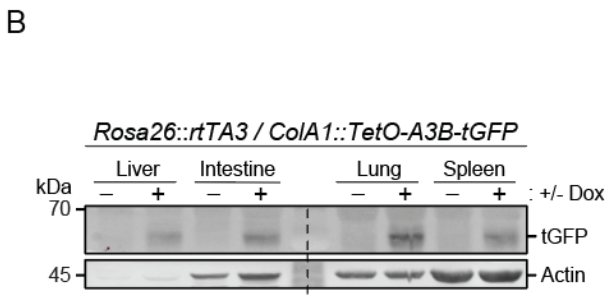
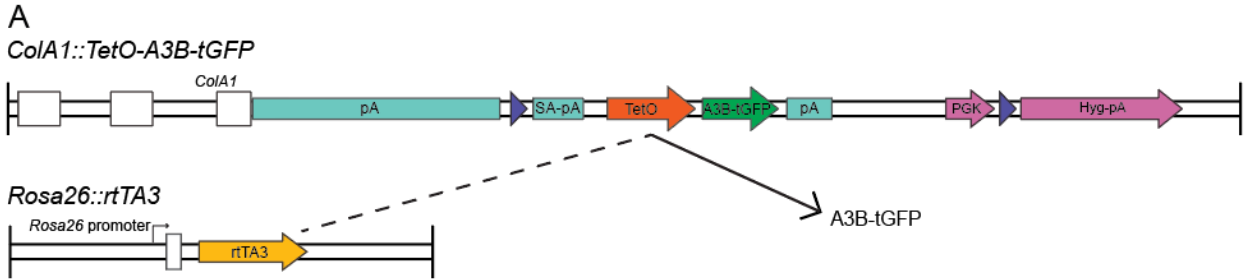
47 (p>0.99 by Fisher's exact test).

48 (C) Bar plot showing relative proportion of each type of SBS mutation in spleens from 2 WT (12  
49 months) and 2 *R26-A3B* mice (15 and 18 months), from whole-exome sequencing (WES) at 30x  
50 depth. The light red portion of C-to-T mutations represent mutations in CG motifs.

51 (D-E) Trinucleotide mutation profiles of the total SBS mutations in the exomes of WT mice and  
52 *CAG-A3B* mice, respectively.

53

54



56 **Figure S3. A3B expressed from the ColA1 locus is modestly tumorigenic, related to Figure 3**

57 (A) Schematic of the engineered *ColA1::TetO-A3B-tGFP* locus and *Rosa26::rtTA3* locus, which  
58 enables expression of A3B-tGFP following doxycycline (Dox) administration.

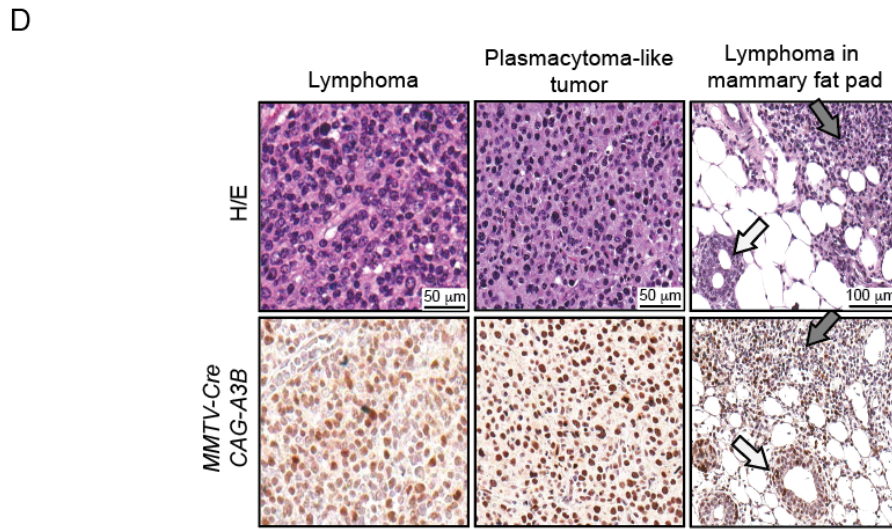
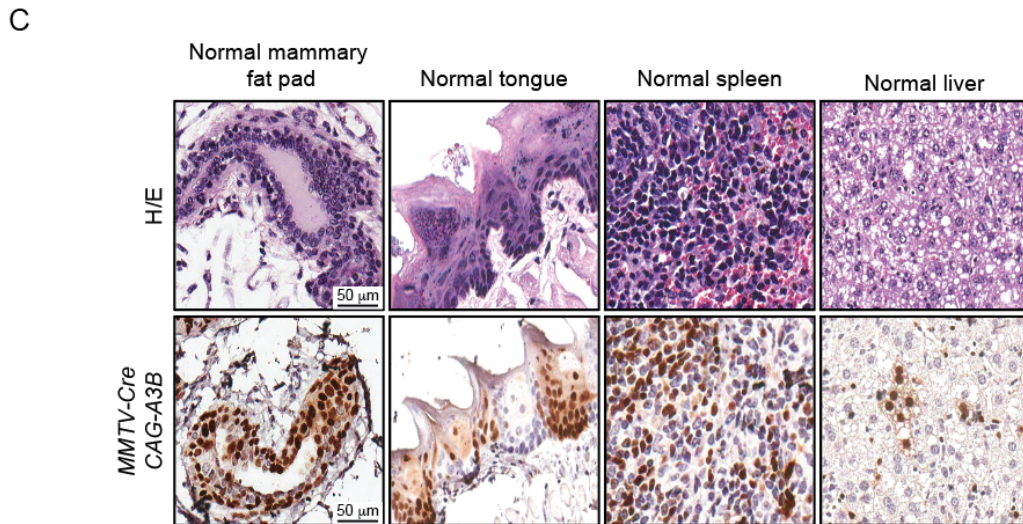
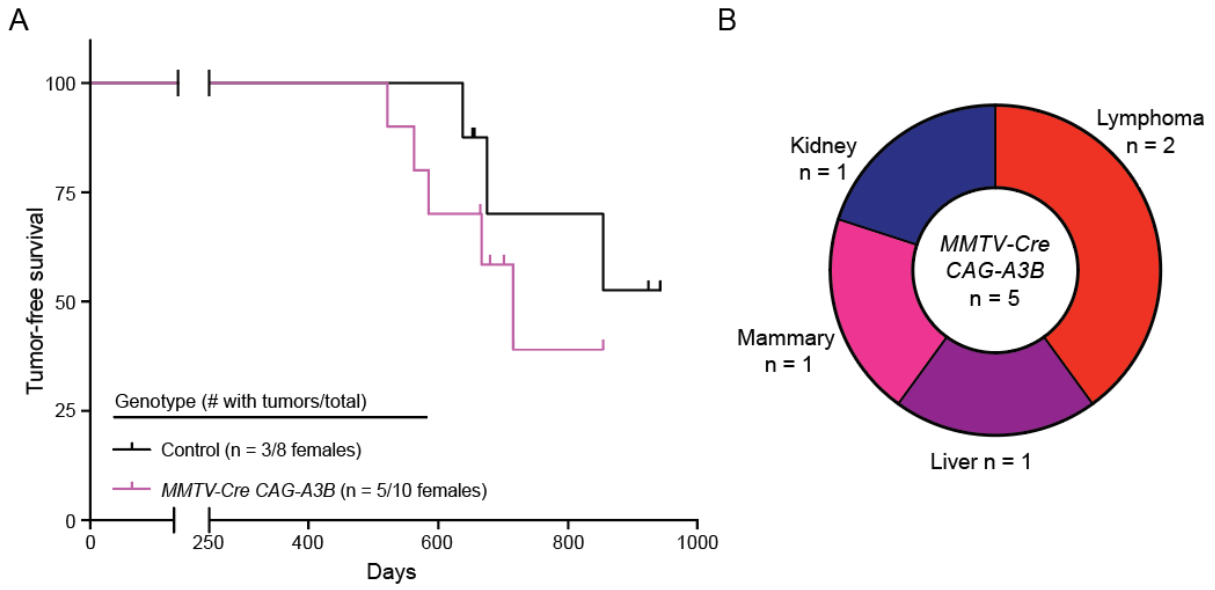
59 (B-C) Immunoblot and ssDNA deaminase activity of A3B-tGFP expressed in the indicated tissues  
60 from *Rosa26::rtTA3/ColA1::TetO-A3B-tGFP* mice fed with (+) or without (-) Dox-containing  
61 chow for 15 days. Recombinant A3A is a positive control for activity (S, substrate; P, product),  
62 and actin provides a loading control.

63 (D) Kaplan-Meier curves comparing tumor-free survival of WT (n=5) and  
64 *Rosa26::rtTA3/ColA1::TetO-A3B-tGFP* (n=6) animals fed with Dox-containing chow for 1-2  
65 years from birth (p=0.37 by log-rank Mantel-Cox test).

66 (E) Pie chart showing the total number of each tumor type found in *Rosa26::rtTA3/ColA1::TetO-*  
67 *A3B-tGFP* mice.

68

69



71 **Figure S4. *CAG-A3B* expression mediated by *MMTV-Cre* has limited tumor phenotype,**  
72 **related to Figure 3**

73 (A) Kaplan-Meier curve showing tumor-free survival of *MMTV-Cre* control (n=8 females) and  
74 *MMTV-Cre CAG-A3B* (n=10 females) mice (p=0.35 by log-rank Mantel-Cox test).

75 (B) Pie chart showing the total number of primary tumor locations that developed in *MMTV-Cre*  
76 *CAG-A3B* females.

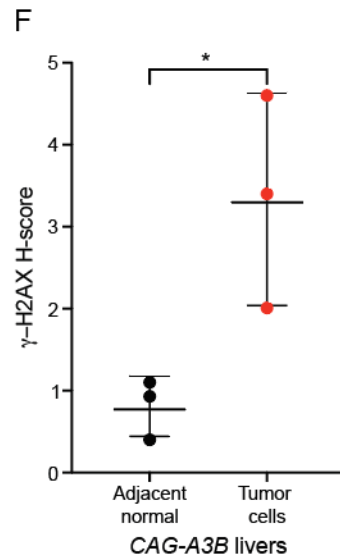
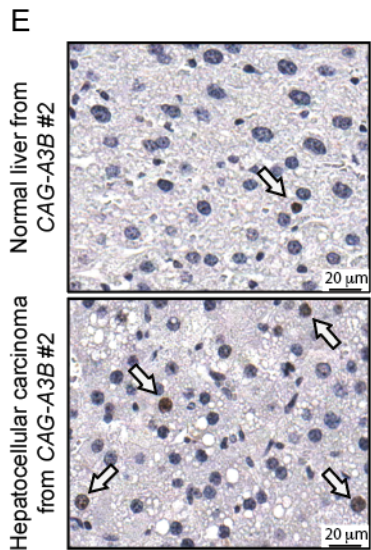
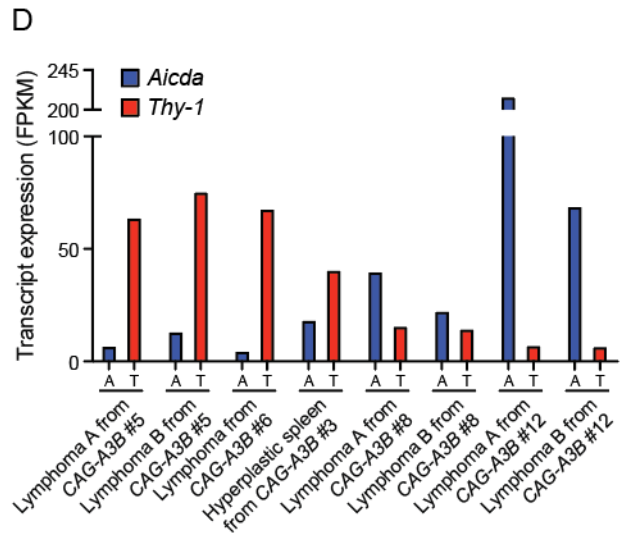
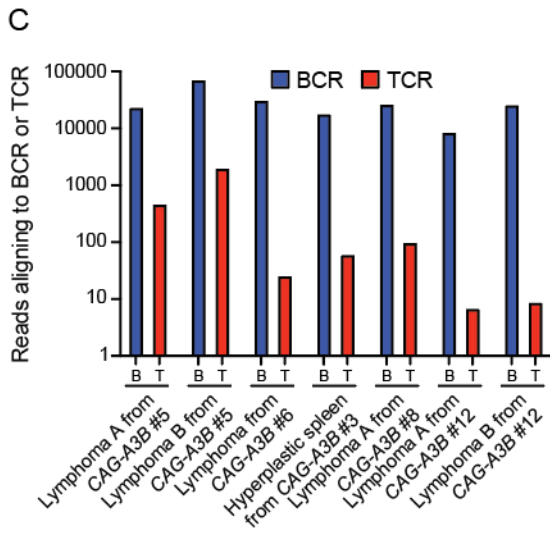
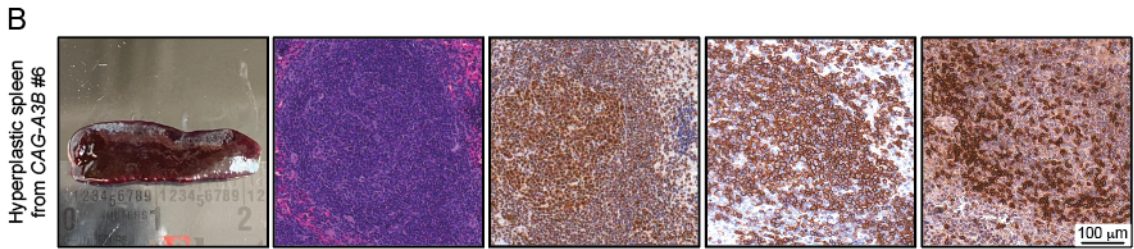
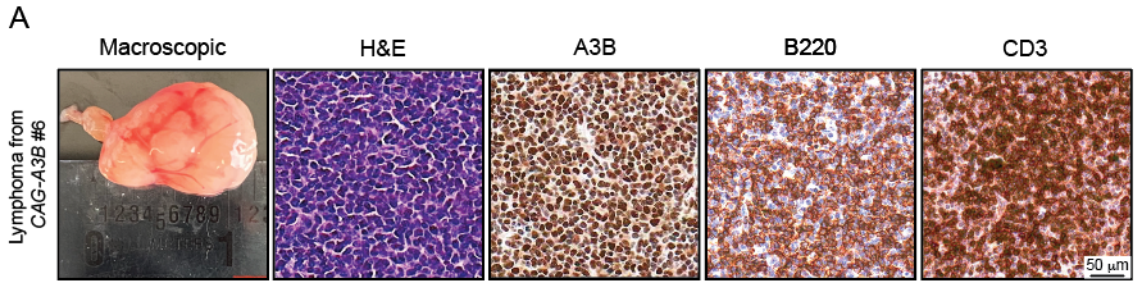
77 (C) H&E and anti-A3B IHC of indicated normal tissues in *MMTV-Cre CAG-A3B* mice. These  
78 immunohistological analyses indicate partial *MMTV-Cre* penetrance in multiple tissues beyond  
79 the mammary glands.

80 (D) H&E and anti-A3B IHC of representative lymphoproliferative tumors in *MMTV-Cre CAG-*  
81 *A3B* mice. Gray arrows point toward lymphomas, and white arrows to mammary glands.

82

83





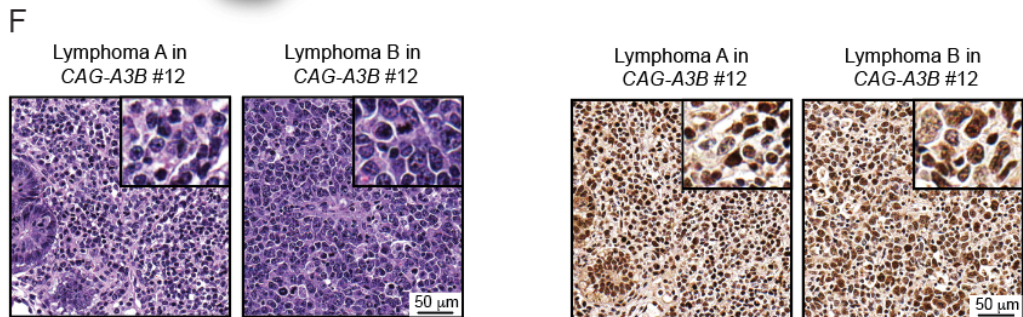
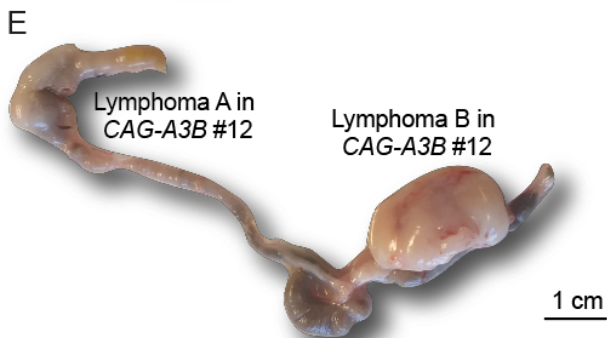
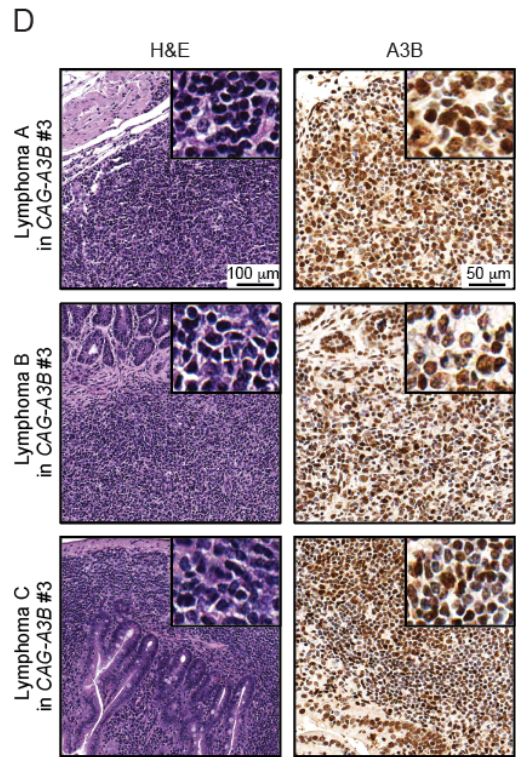
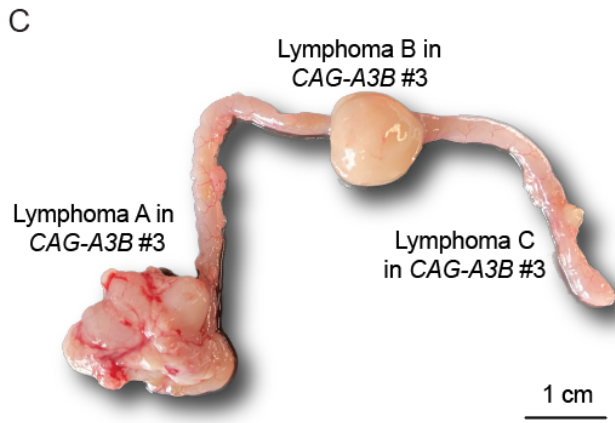
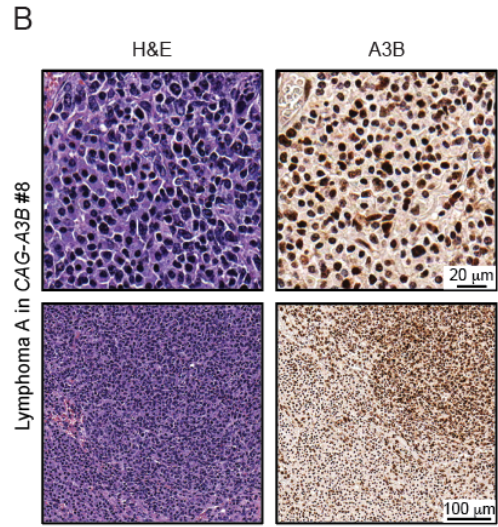
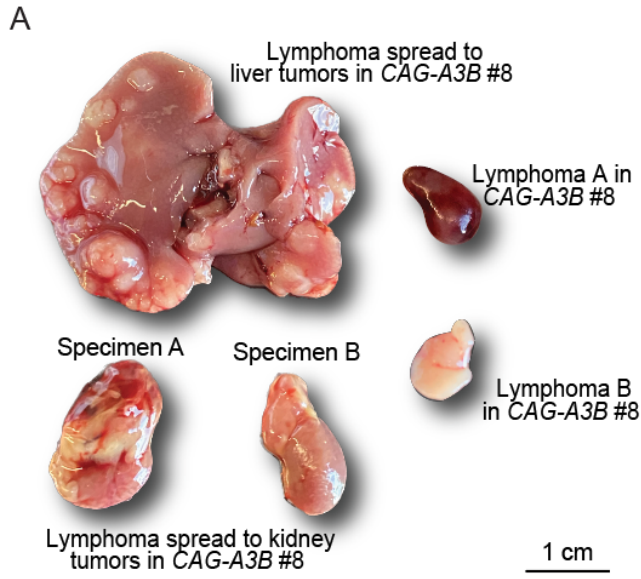
85 **Figure S5. Phenotypic diversity of *CAG-A3B* tumors, related to Figure 4**

86 **(A-B)** Macroscopic photograph and photomicrographs of H&E, anti-A3B, anti-B220, and anti-  
87 CD3 IHC of the lymphoma from *CAG-A3B* #6 and hyperplastic spleen from *CAG-A3B* #6,  
88 respectively. These lesions consist of a mixed B-cell and T-cell population, as shown by high  
89 levels of both B220 and CD3 immunostaining. High A3B levels are expressed within the nuclei  
90 of most cells. The macroscopic view of the lymphoma is also presented in Figure 4C.

91 **(C)** A histogram showing the number of *CAG-A3B* lymphoma RNAseq reads that align to  
92 recombined B-cell receptor (BCR; B) or T-cell receptor (TCR; T) genes using the TRUST4  
93 algorithm. Each histogram bar represents the largest single contig for each recombined gene. The  
94 fact that the BCR rearrangement is clonal and represented by >100-fold more reads demonstrates  
95 that most lymphomas have a B cell origin.

96 **(D)** Transcript expression levels of B-cell specific transcript *Aicda* mRNA and pan T-cell marker  
97 *Thy-1* in lymphomas from *CAG-A3B* animals (additional analyses of the same RNAseq data sets  
98 as in panel C).

99 **(E-F)** Representative images and quantification of  $\gamma$ -H2AX immunostaining of HCCs and normal  
100 adjacent liver parenchyma from *CAG-A3B* animals, respectively. H-score is determined by the  
101 extent of nuclear reactivity in terms of percent cells and is equal to the sum of strong staining = %  
102 x 3, medium staining = % x 2, weak staining = % x 1, no staining = % x 0 (n=3 per condition;  
103 mean +/- SD shown; p= 0.0314 by unpaired t-test).



105 **Figure S6. Disseminated lymphomas in *CAG-A3B* mice, related to Figure 4**

106 (A) Macroscopic pictures of lymphoma metastasis to the liver, kidney, and multiple lymph nodes

107 in *CAG-A3B* mouse #8.

108 (B) H/E (left) and anti-A3B IHC (right) of lymphoma A in *CAG-A3B* mouse #8. Representative

109 photomicrographs from the lymphoma spread to liver and kidney can be found in Figures 4H-I.

110 (C) Macroscopic picture of a diffuse lymphoma spread to the intestinal mucosa and Peyer's

111 patches in *CAG-A3B* mouse #3.

112 (D) H/E (left) and anti-A3B IHC (right) of lymphomas A-C in *CAG-A3B* mouse #3.

113 (E) Macroscopic picture of a lymphoma involving a mesenteric lymph node and the intestinal

114 mucosa/Peyer's patch in *CAG-A3B* mouse #12.

115 (F) H/E (left) and anti-A3B IHC (right) of lymphomas A and B in *CAG-A3B* mouse #12. All boxed

116 images in panels B, D, and F are representative areas of the corresponding tissues at 4x the

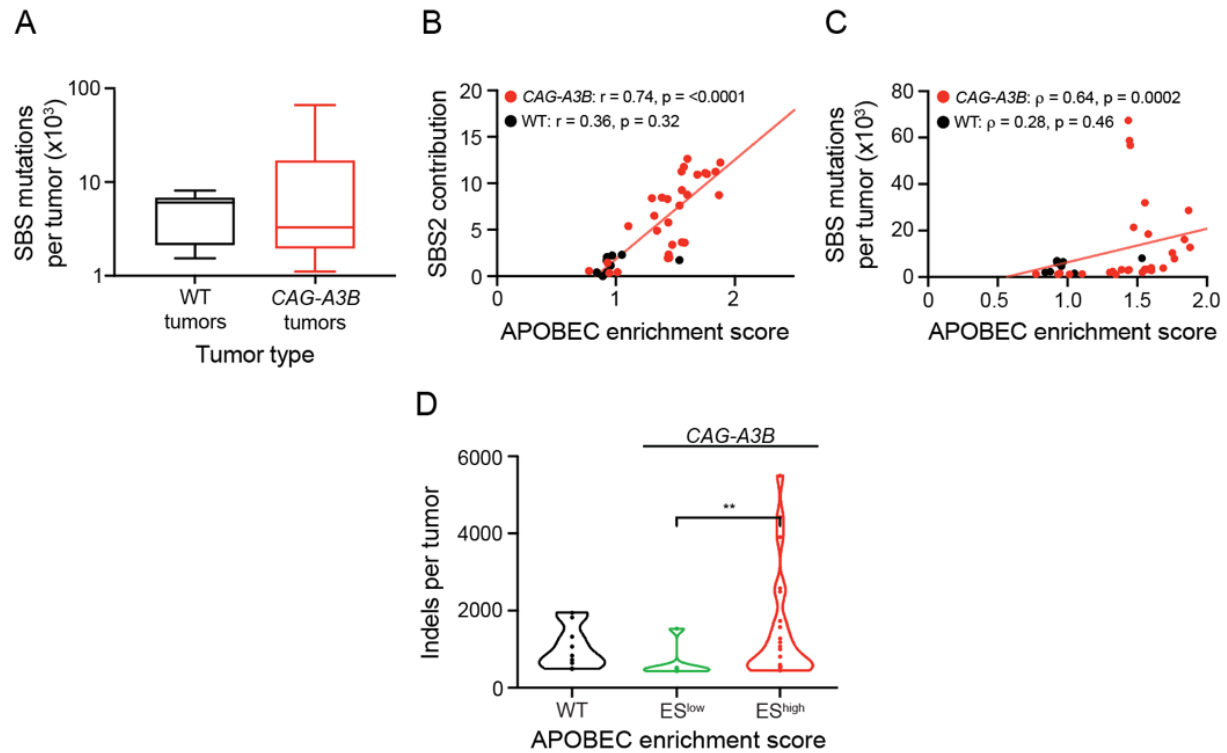
117 magnification. Note that the macroscopic views of the lymphomas in Figure S6C and S6E are also

118 shown in Figure 4C, and shown again here with extended views to more broadly represent all

119 features of the disseminated lymphomas.

120

121



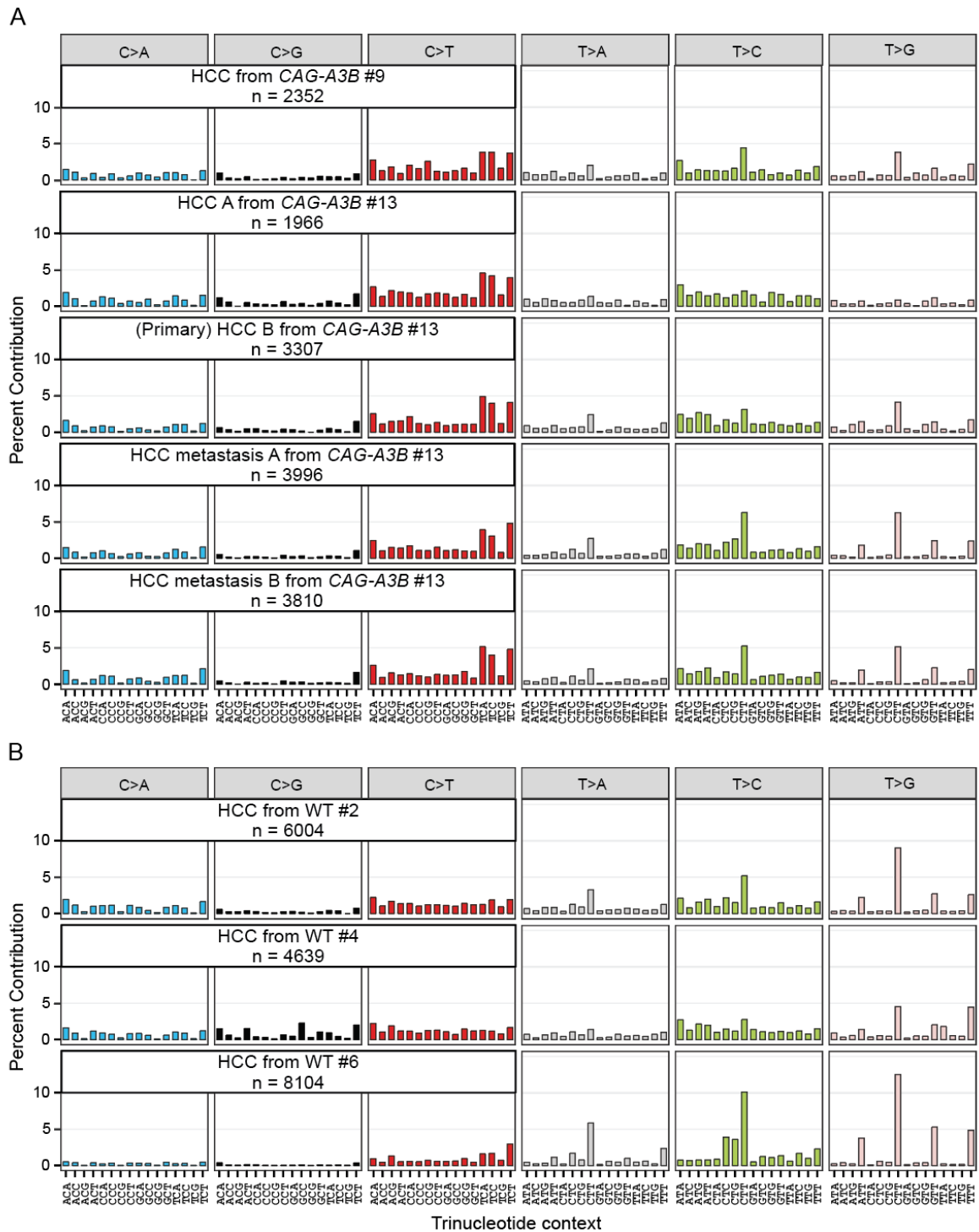
122

123 **Figure S7. Genomic mutations in WT and *CAG-A3B* mice, related to Figure 5**

124 (A) Box and whisker plots of the number of SBS mutations in tumors from WT and *CAG-A3B*  
 125 mice. The middle horizontal line is the median, the outer horizontal lines are the upper and lower  
 126 quartiles, and the whiskers outside the box represent the maximum and minimum ( $p > 0.99$  by  
 127 Mann-Whitney).

128 (B-C) Scatterplots comparing APOBEC mutation signature enrichment scores to the percentages  
 129 of SBS2 or to the full base substitution loads in tumors from WT and *CAG-A3B* animals,  
 130 respectively (B, Pearson correlation coefficient, C, Spearman's rank correlation coefficient, and  
 131 p-values indicated).

132 (D) Violin plots of the total number of indels in tumors from WT mice in comparison to tumors  
 133 from *CAG-A3B* animals with low or high APOBEC enrichment scores (ES;  $p = 0.0076$  for ES<sup>low</sup> vs  
 134 ES<sup>high</sup> groups by Mann-Whitney U-test).



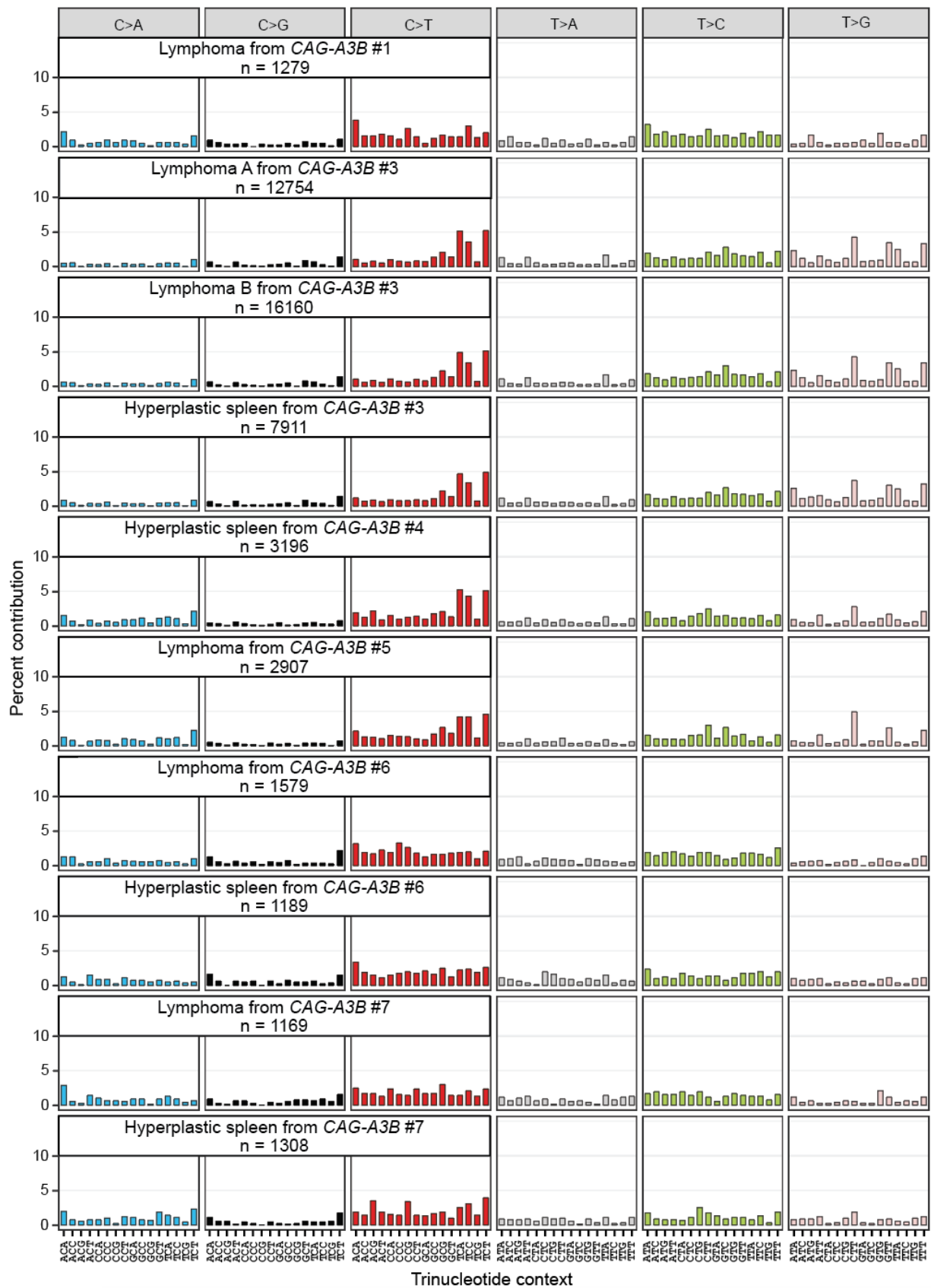
135

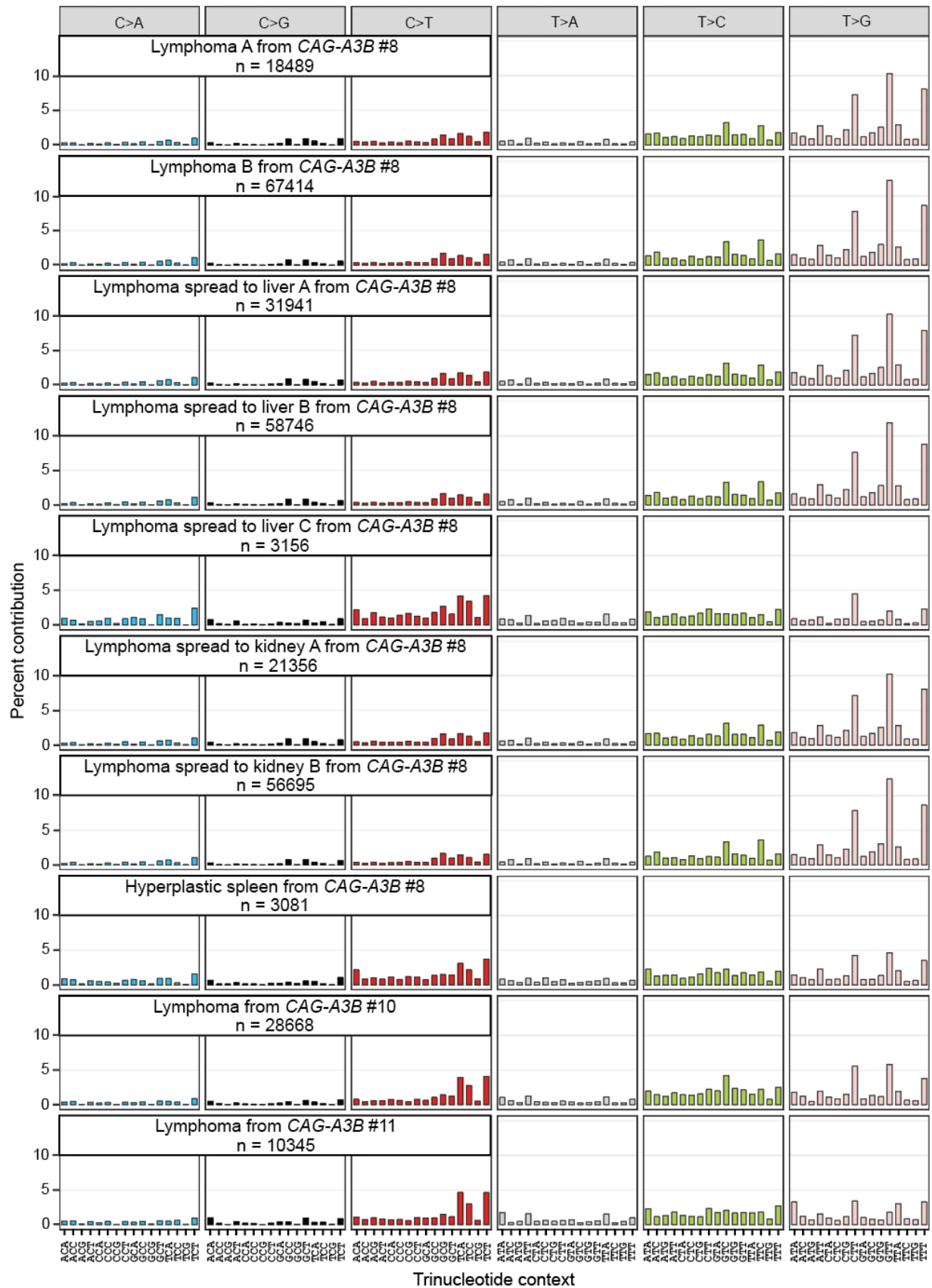
136 **Figure S8. Mutational spectra of liver tumors, related to Figure 5**

137 (A-B) Trinucleotide mutation profiles of all whole-genome sequenced liver tumors in *CAG-A3B*

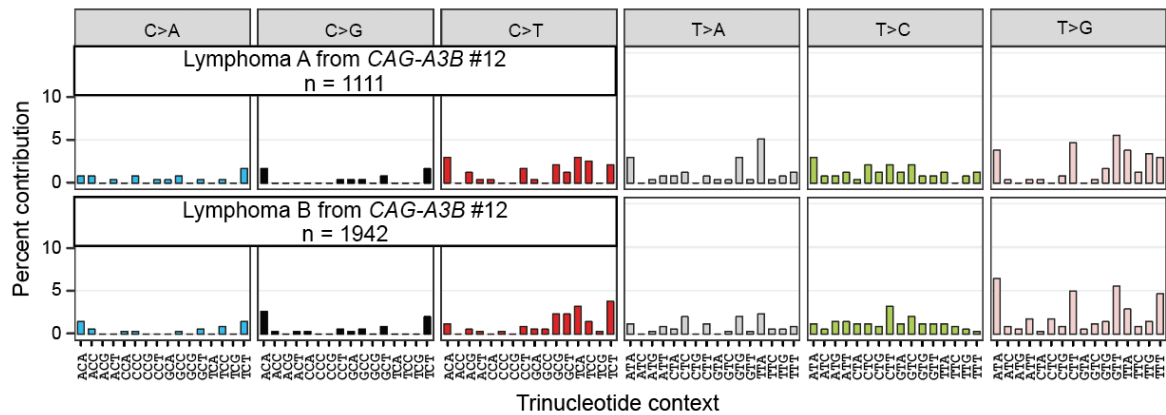
138 and WT mice not shown in Figure 5 (n-values = total SBS mutations in each tumor).

A

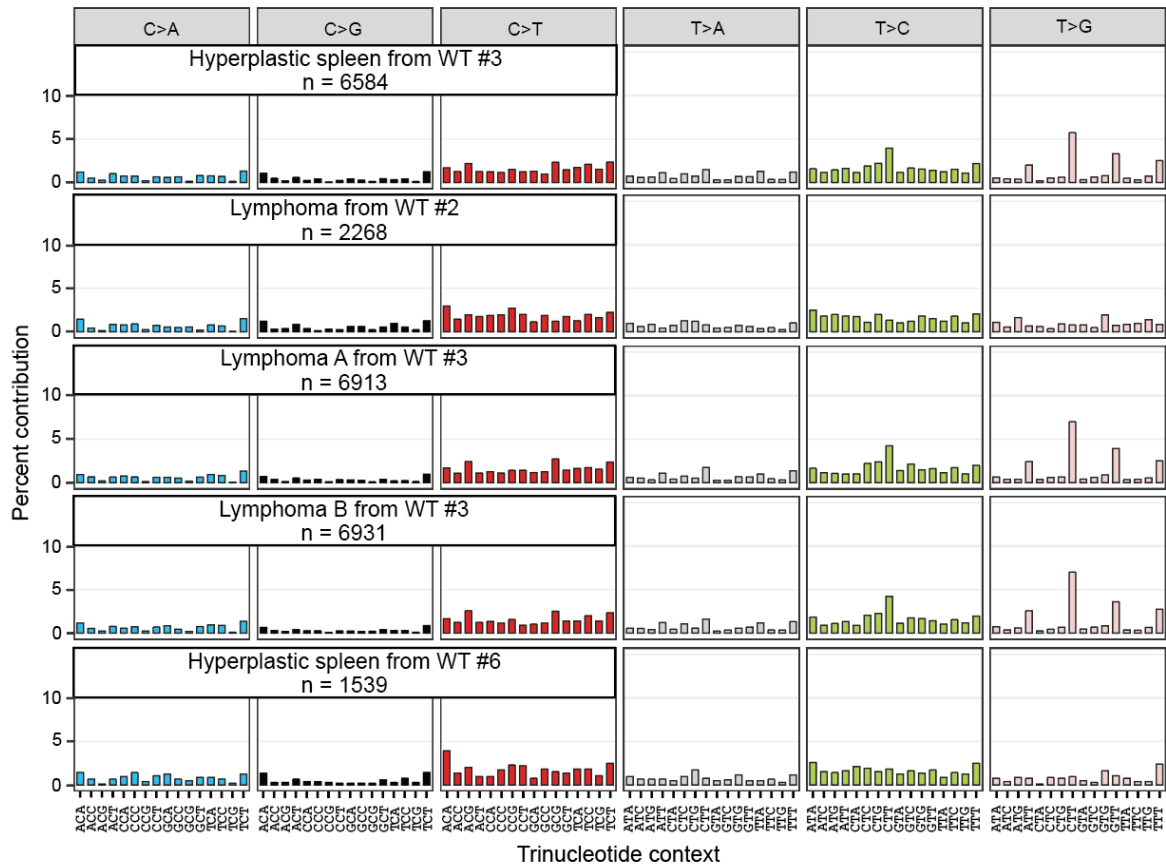








B

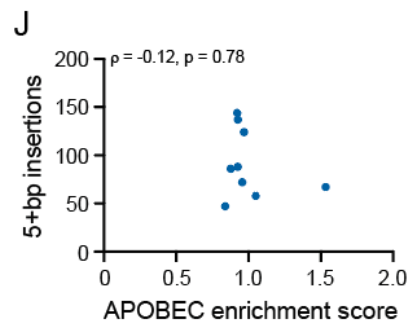
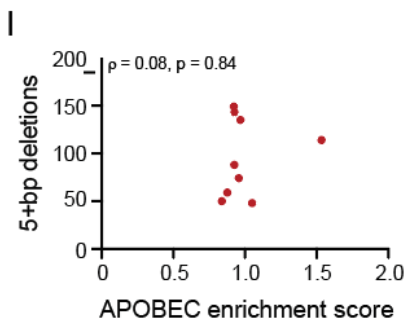
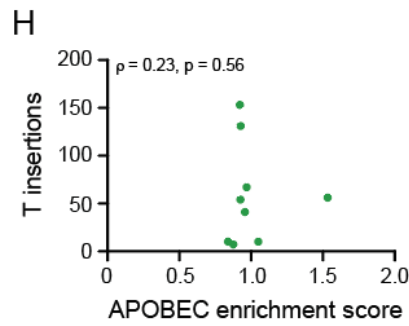
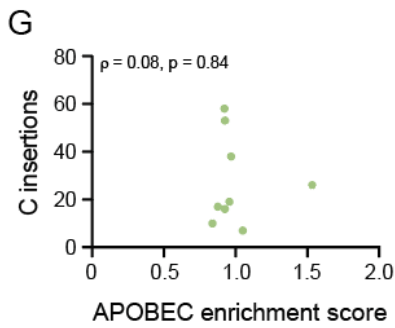
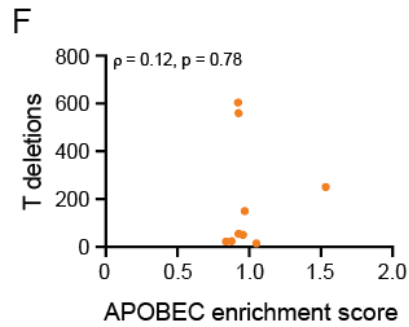
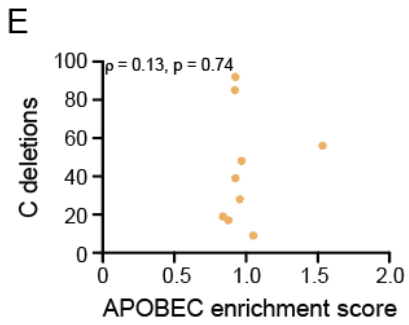
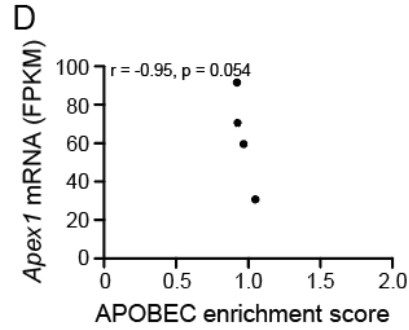
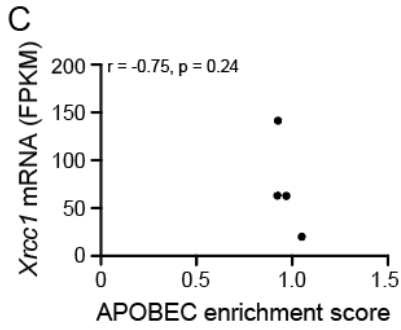
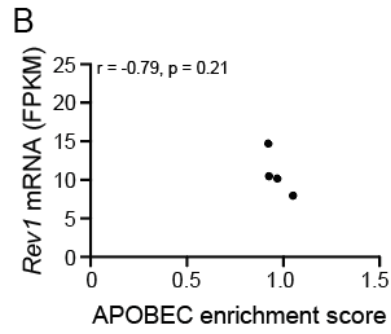
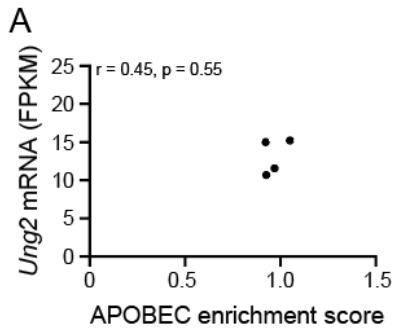


142

143 **Figure S9. Mutational spectra of blood tumors, related to Figure 5**

144 (A-B) Trinucleotide mutation profiles of all whole-genome sequenced blood tumors in *CAG-A3B*

145 and WT mice not shown in Figure 5 (n-values = total SBS mutations in each tumor).



147 **Figure S10. Genomic correlations in WT mice, related to Figure 5**

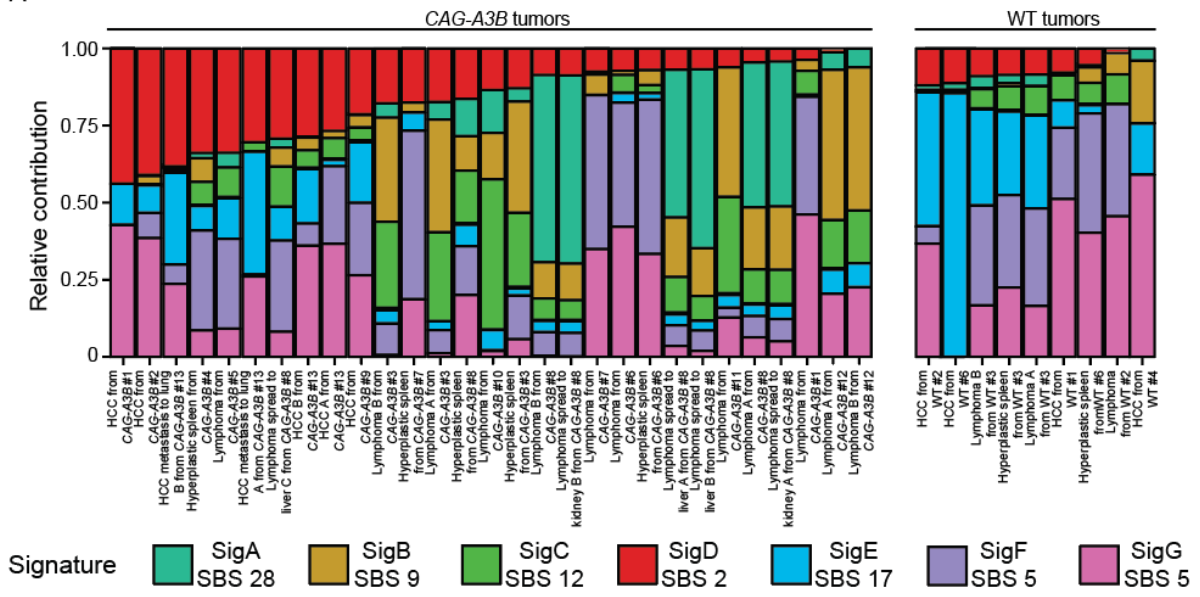
148 **(A-D)** Scatterplots of APOBEC enrichment score from WT lymphomas (n=4) compared to the  
149 mRNA levels of *Ung2*, *Apex1*, *Xrcc1*, and *Rev1*, respectively, from the same tumors (Pearson  
150 correlation coefficients and corresponding p-values indicated).

151 **(E-J)** Scatterplots showing relationships between APOBEC enrichment scores from WT tumors  
152 and the indicated indel types (Spearman's rank correlation coefficients and corresponding p-values  
153 indicated).

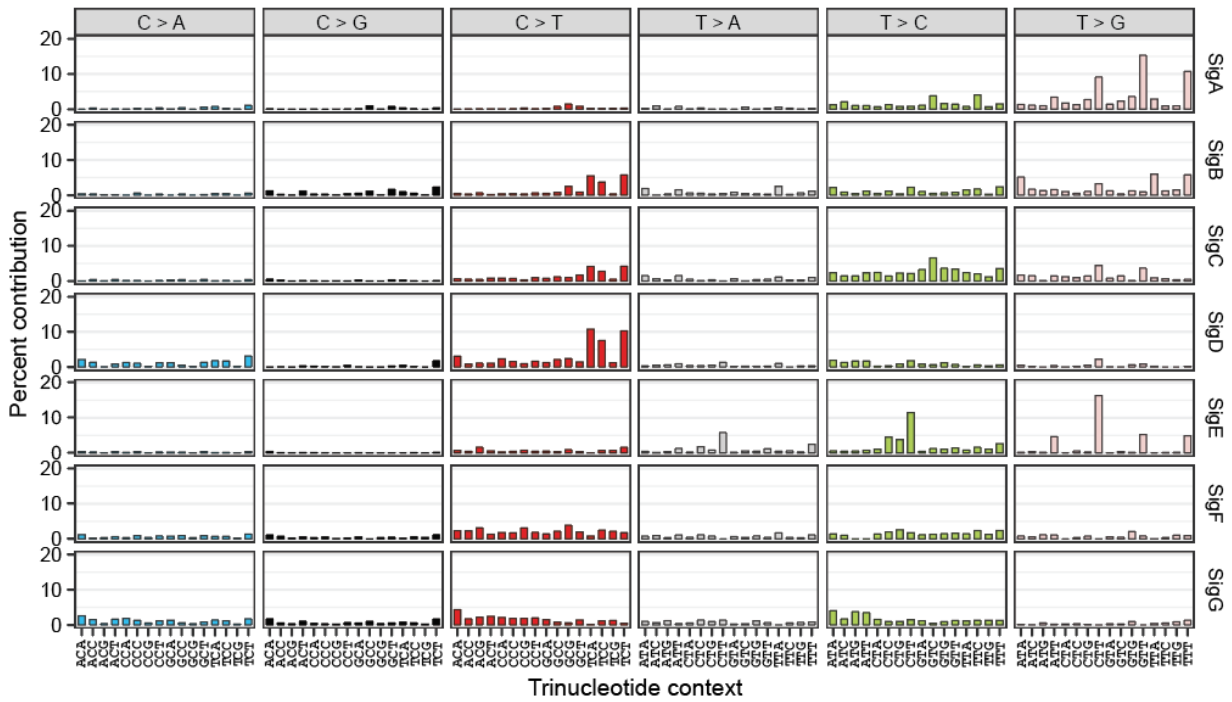
154

155

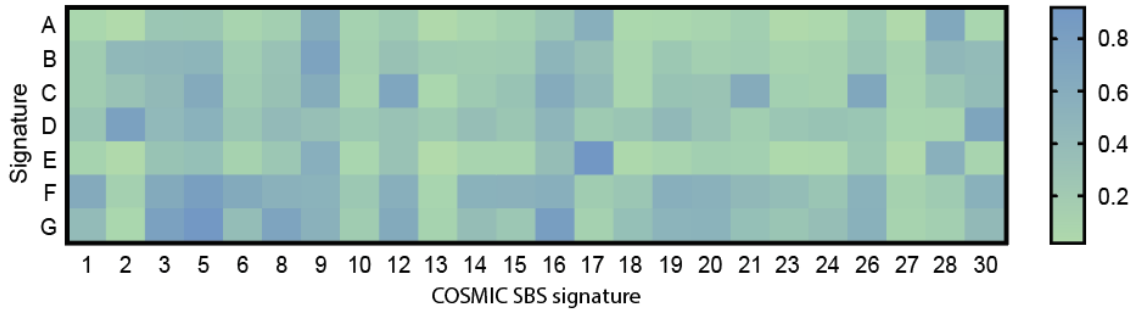
A



B



C



157 **Figure S11. Extraction of *de novo* signatures from all mouse tumors, related to Figure 5**

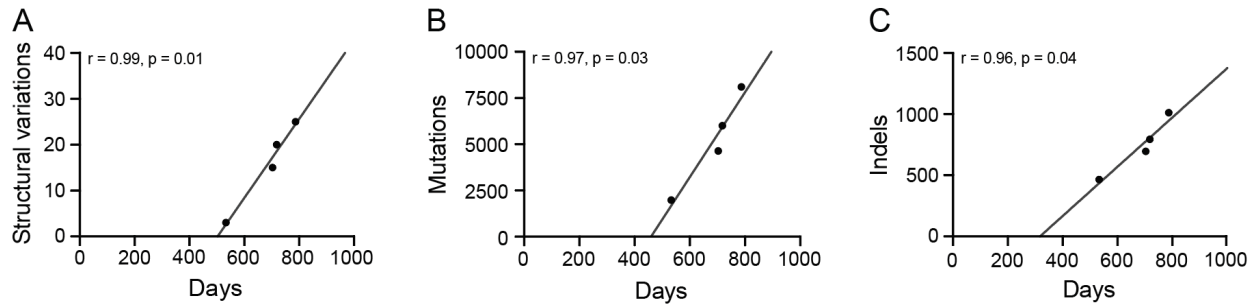
158 **(A)** Contribution of seven *de novo* signatures to the overall tumor mutation landscape based on  
159 non-negative matrix factorization (n=29 *CAG-A3B*; n=9 WT). Upon comparison to previously  
160 described COSMIC signatures, SigD is similar to SBS2 (cosine similarity, CS, = 0.82), SigA to  
161 SBS28 (CS = 0.72), SigB to SBS9 (CS = 0.78), SigC to SBS12 (CS =0.82), SigE to SBS17 (CS =  
162 0.93), SigF to SBS5 (CS = 0.83), and SigG to SBS5 (CS = 0.92).

163 **(B)** Trinucleotide mutation profiles of *de novo* signatures SigA to SigG extracted from tumors  
164 using non-negative matrix factorization.

165 **(C)** Heat map showing cosine similarities of the 7 mutation signatures in the tumors described here  
166 in comparison to established COSMIC SBS mutation signatures. COSMIC SBS mutation  
167 signatures 4, 7, 11, 25, and 19 are excluded due to lack of relevance to mouse models.

168

169



170

171 **Figure S12. Relationship between structural variations, SBS mutations, and indels and**  
 172 **animal age, related to Figure 5 and 6**

173 (A-C) Scatterplot of the number of structural variations, SBS mutations, and indels in HCCs (n=4)  
 174 from WT animals in comparison to age of the animal at the time of sacrifice (Pearson correlation  
 175 coefficients and corresponding p-values shown). These positive associations indicate a linear  
 176 relationship between mutation accumulation and age in tumors from WT animals.

177

178

179 **Table S1. Mice used in tumor-free survival and sequencing analyses**

180 Included as a separate file titled “Durfee-et-al-TableS1.xlsx”.

181