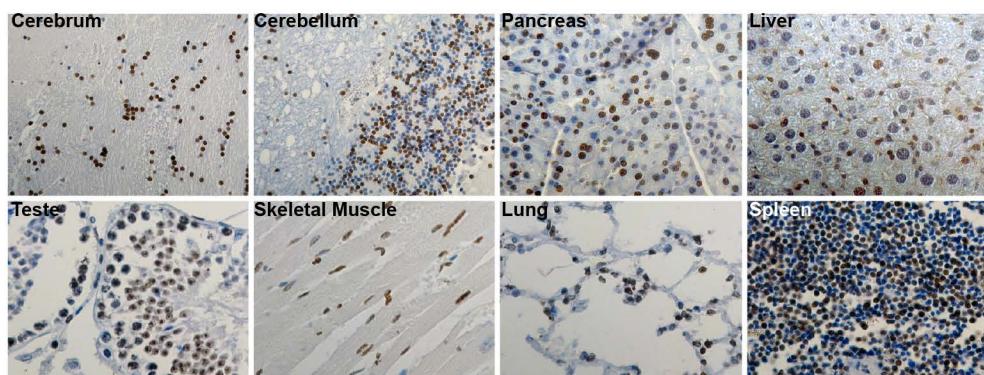
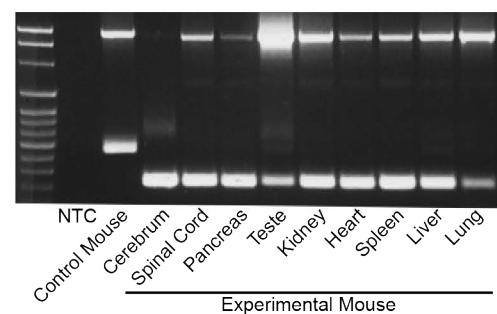


Supplemental Figure 1. Characterization of Cnp-Cre driven SB-mutagenized mouse model

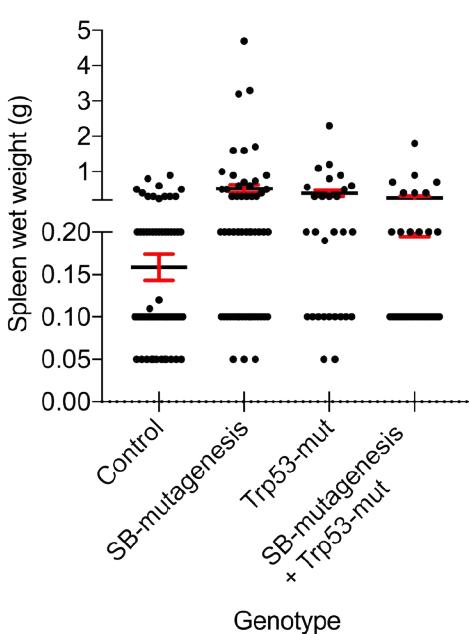
A



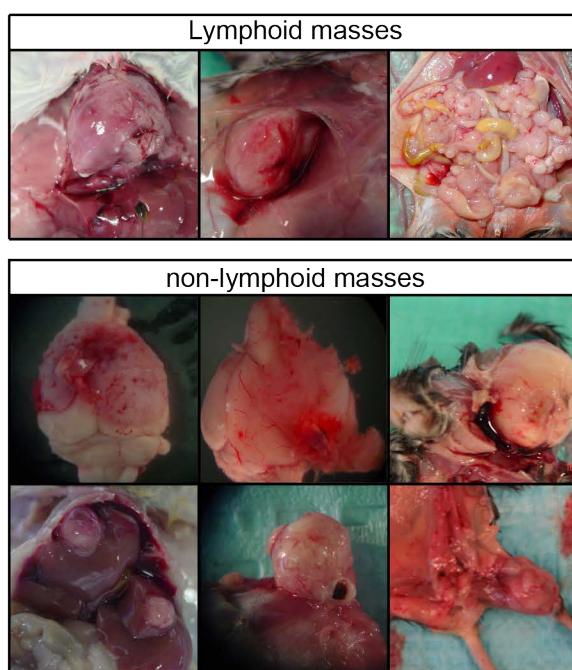
B



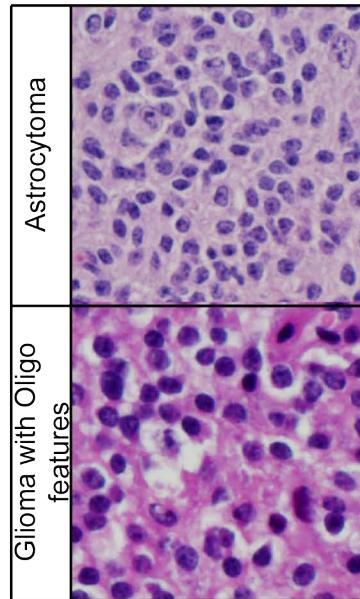
C



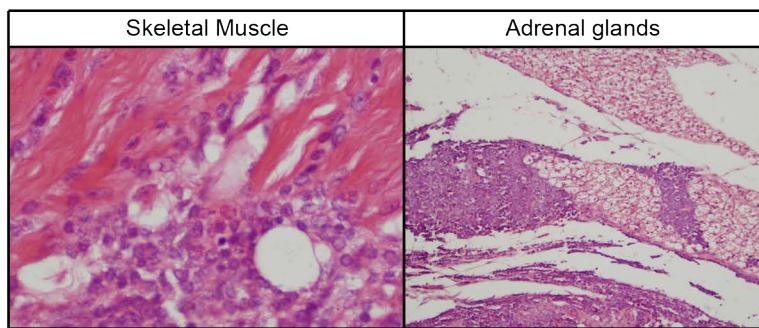
D



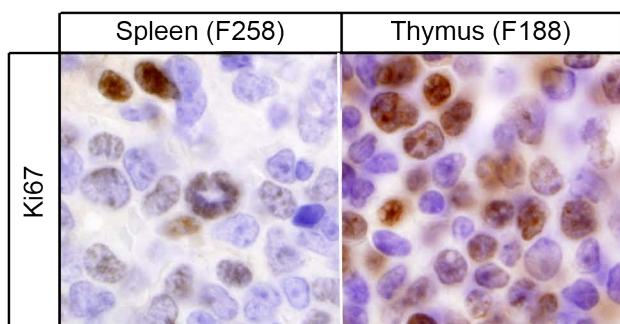
E



F



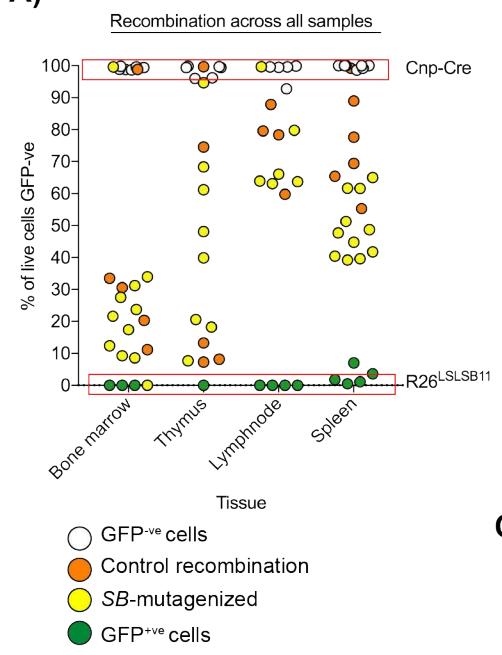
G



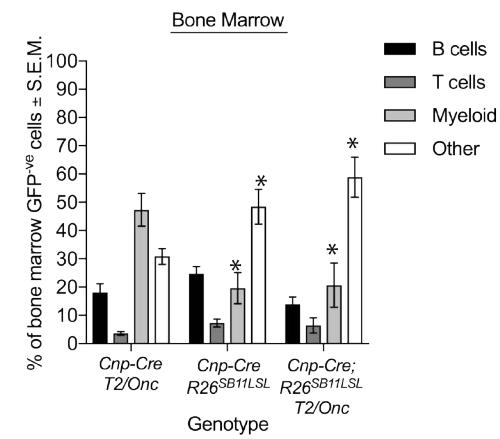
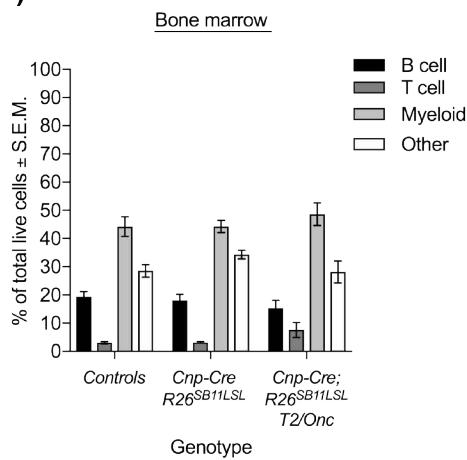
Supplemental Figure 1. Characterization of Cnp-Cre driven SB-mutagenized mouse model (A) IHC for SB expression in a tissue panel from a *Cnp-Cre;R26^{LSLSB11}* mouse: cerebrum, cerebellum, pancreas, liver, teste, skeletal muscle, lung, spleen. (B) Agarose gel image of a PCR-based excision assay. Upper band indicates the donor T2/Onc concatemer. The lower band indicates the footprint after T2/Onc excision from donor concatemer. Experimental mouse = *Cnp-Cre;R26^{LSLSB11};T2/Onc*. (C) Bar graph depicts spleen wet-weight of control versus experimental cohorts. Unpaired student t-test p<0.0001. (D) Images from experimental mice at time of necropsy. (E) H&E stain of two brain tumors identified in the screen. (F) H&E stain of skeletal muscle and adrenal glands displaying infiltrative lymphoma. (G) IHC for Ki67 on SB-mutagenized spleen and thymus samples.

Supplemental Figure 2. Characterization of Cnp⁺ cells in lymphoid compartments

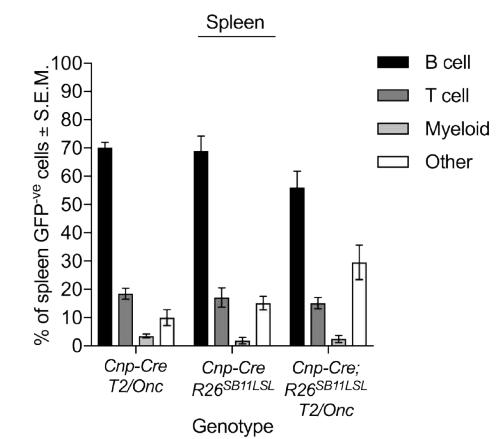
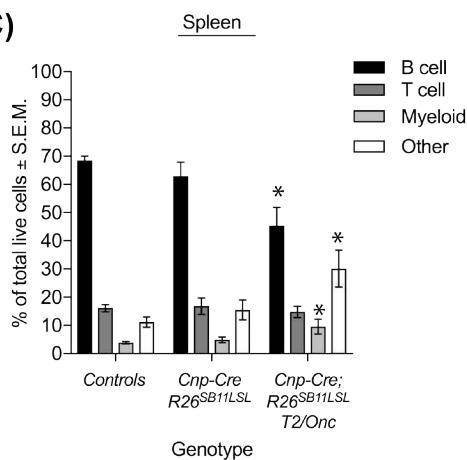
A)



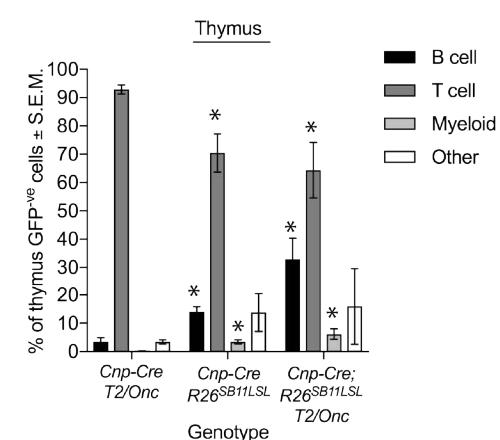
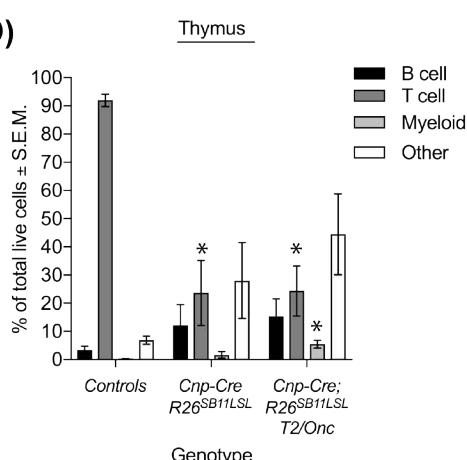
B)



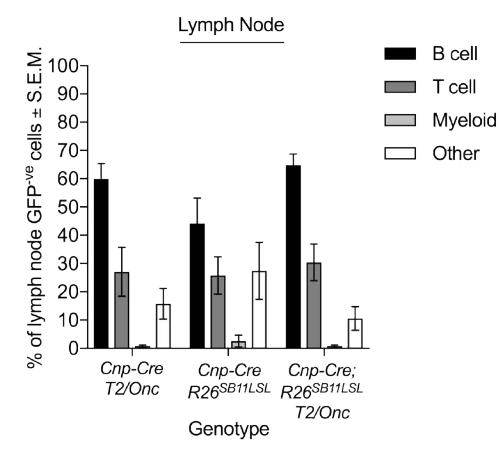
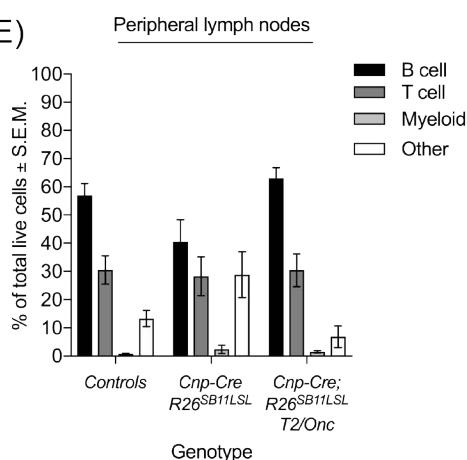
C)



D)

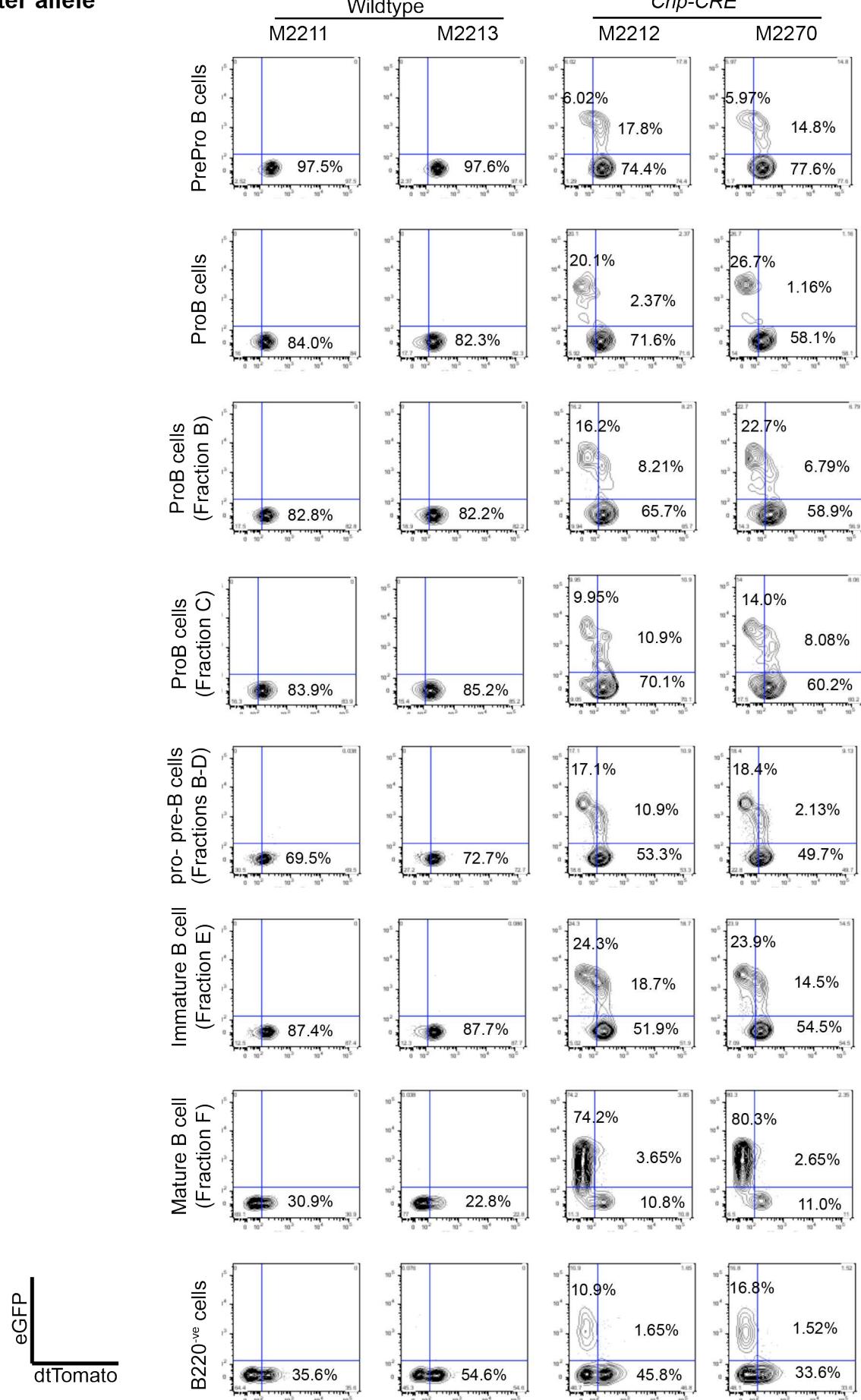


E)



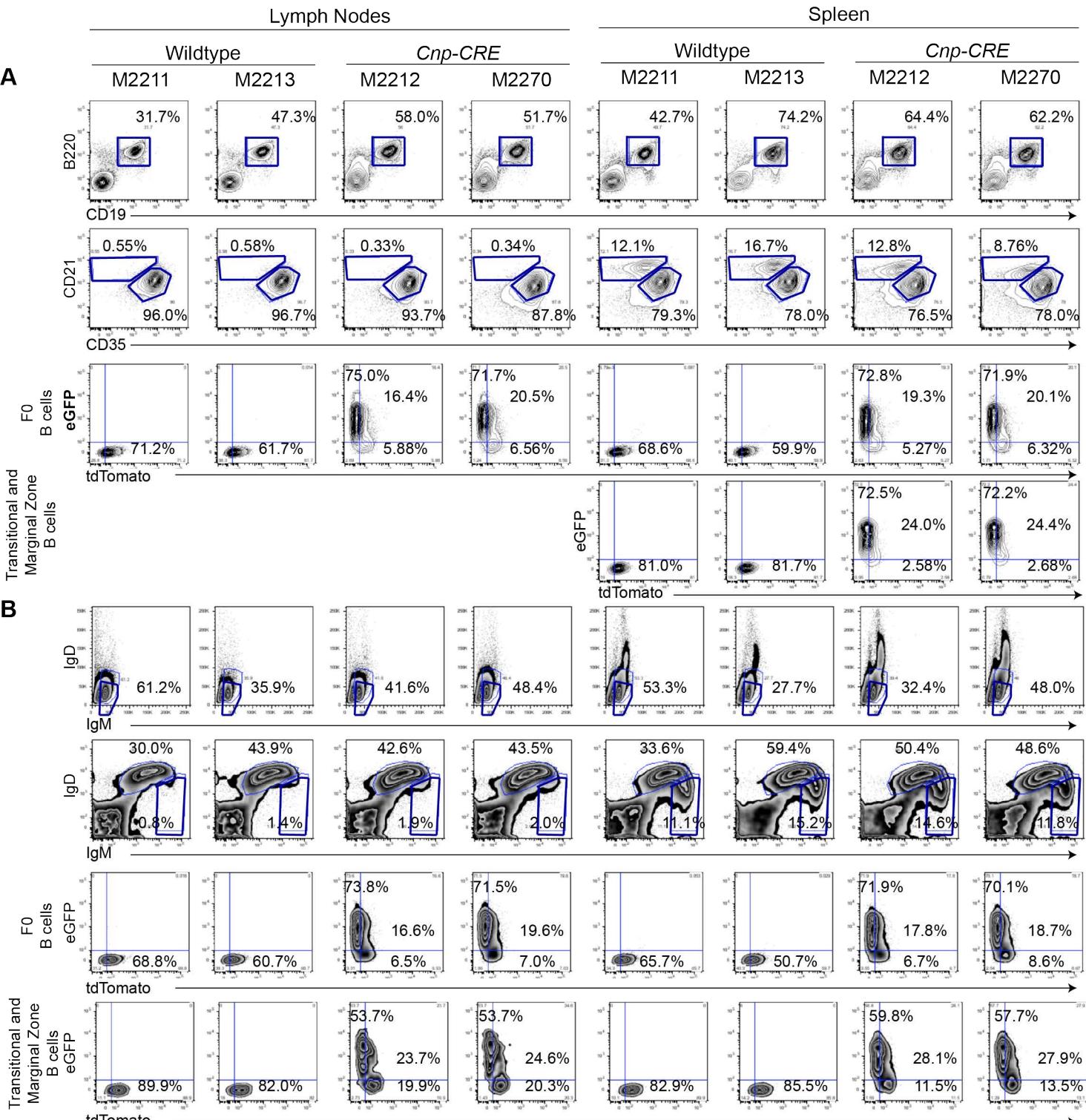
Supplemental Figure 2. Characterization of Cnp⁺ cells in lymphoid compartments. (A) Scatter plot indicating the percentage of GFP-ve cells in each of the lymphoid tissues. Color of dot represents the genetic background of the sample. Each sample was further characterized for recombination based on B cell, T cell, myeloid or ‘other’ in the bone marrow (B), spleen (C) and thymus (D). Bar graphs on the left indicate the distribution of each cell type from total live cell populations while the graphs on the right indicate the percentage of GFP-ve cells for each of the cell type.

Supplemental Figure 3: Bone marrow phenotyping of wild type and Cnp-Cre mice on a conditional reporter allele



Supplemental Figure 3. Bone marrow phenotyping of wildtype and Cnp-Cre mice possessing a conditional reporter allele. Flow cytometry analysis of bone marrow from 4 mice (n=2 reporter, n=2 *Cnp-cre*; reporter). If *Cnp-Cre* is expressed, cells will become EGFP positive. Hardy fractions were performed on bone marrow samples: PreProB, ProB, Fraction B, Fraction C, Fractions B-D, Fraction E, Fraction F and B220 negative cells.

Supplemental Figure 4. Immunophenotyping of lymph nodes and spleens



Supplemental Figure 4. Immunophenotyping of lymph nodes and spleens. Lymph nodes and spleens were harvested from 4 mice (n=2 reporter mice, n=2 *Cnp-Cre*; reporter mice). A) Gating strategy used to identify B cell populations. Top row is B220 by CD19 (two B cell markers), the next row is CD21 by CD35. This staining is for Marginal zone and FO B cells. Underneath you can see the eGFP expression – similar to what was done in the bone marrow. B) This stain is IgD by IgM – one gate IgM^{high}IgD^{low} is for FO B cells. The IgM^{high}IgD^{low} is for Transitional and MZ B cells. There is a breakdown of the GFP expression in these subsets.

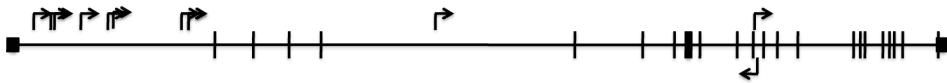
Supplemental Figure 5: T2/Onc insertion schematics

A) Putative proto-oncogenes

Bach2



NfkB1

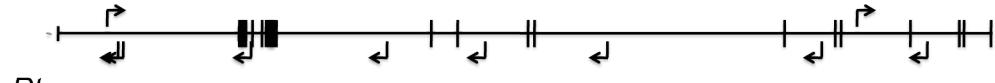


Map3k8

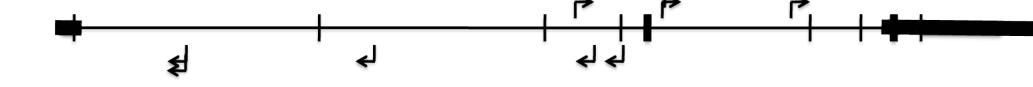


B) Putative tumor suppressor genes

Ambra1

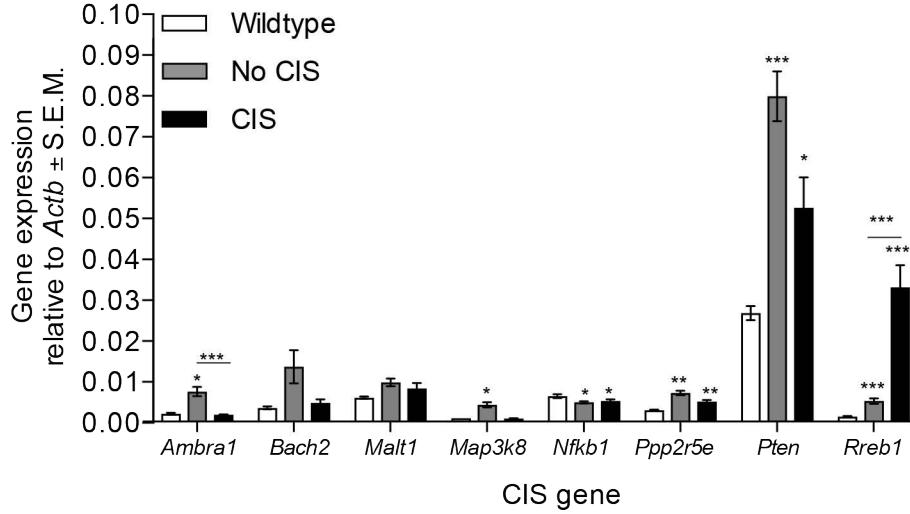


Pten



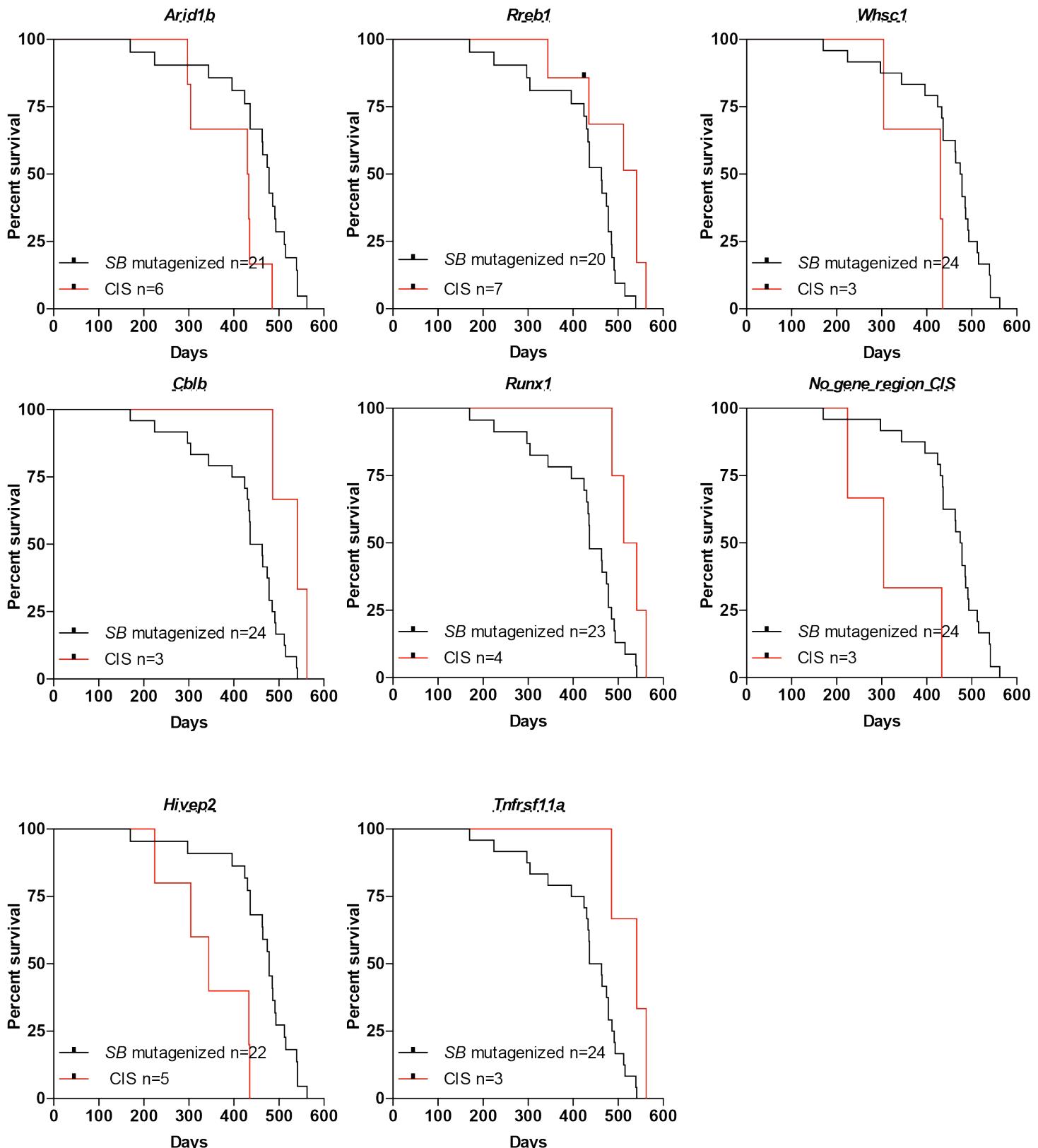
C)

qRT-PCR analysis: CIS genes



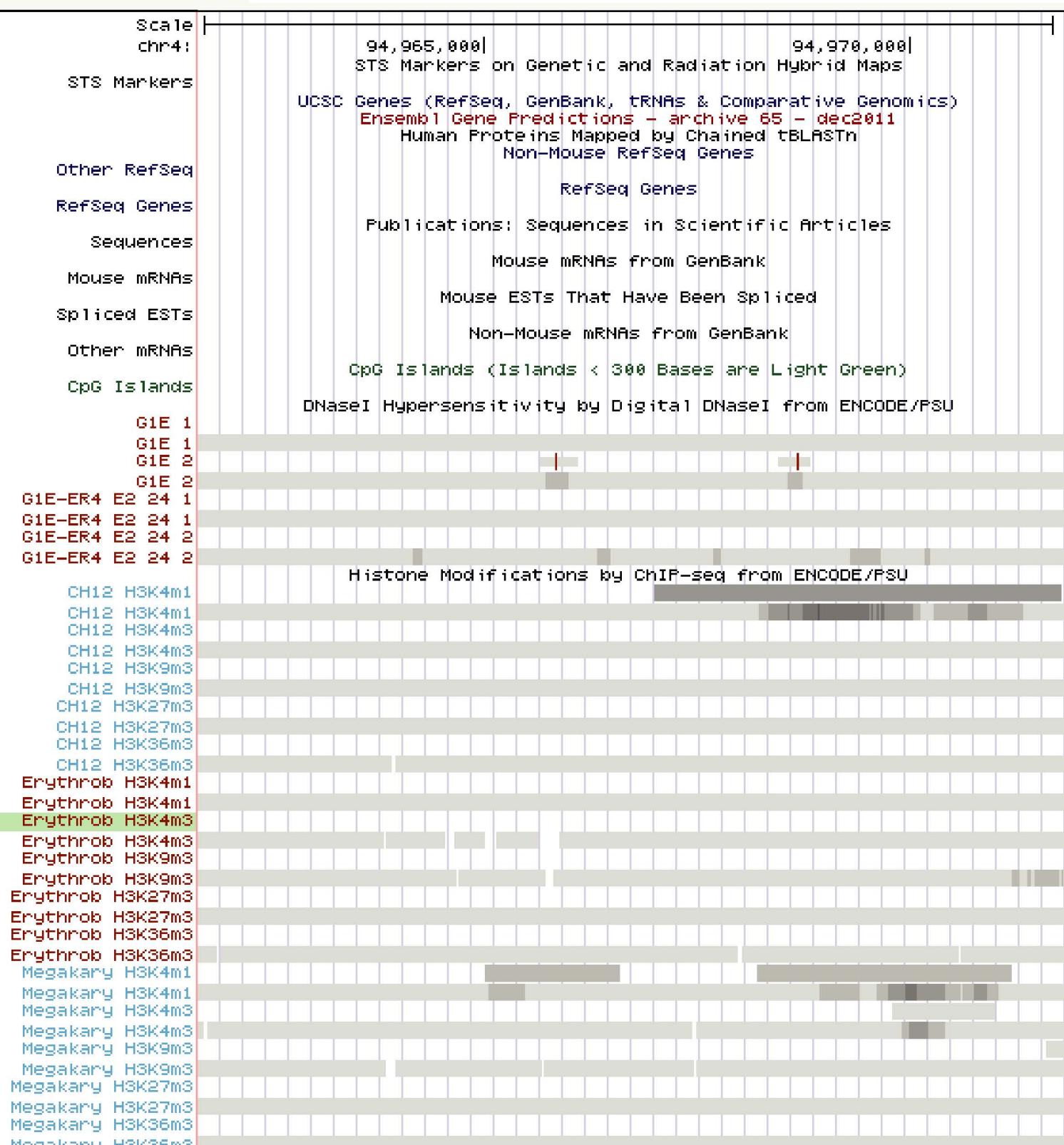
Supplemental Figure 5. Schematics of T2/Onc insertions within putative proto-oncogenes and TSGs. (A) Schematics representing 3 putative proto-oncogenes that were CIS-associated genes mutated by T2/Onc: *Bach2*, *NfkB1* and *Map3k8*. Arrows represent the direction in which the MSCV promoter in T2/Onc is orientated compared to the direction of gene transcription (left to right). Arrows pointing toward the right would be predicted to drive transcription. Arrows pointing toward the left would be predicted to disrupt transcription. T2/Onc insertions can be used to predict gene function when looking at all the insertions as a whole. Arrows pointing in the same direction of transcription and clustered together within or upstream of a gene predict the gene to function as a proto-oncogene (A). Arrows with no bias in directionality or preference for positioning within a gene predict the gene to function as a TSG (B). (B) Schematic representing 2 putative TSGs (*Ambra1*, *Pten*). Images based on IGV 2.0 browser. (C) Bar graph depicts results from qRT-PCR assessment of eight CIS-associated genes expression in wildtype spleen samples and splenomegaly samples with and without T2/Onc insertions within the gene. Statistics were performed using Prism6 Graphpad student t-Test corrected for multiple comparisons (Holm-Sidak, $\alpha=5.0\%$). Wildtype spleens $n=3$, Splenomegaly samples ranged from $n=3$ to $n=10$ depending on CIS gene for No CIS and CIS cohorts. * $p<0.05$, ** $p<0.001$, *** $p<0.0001$

Supplemental Figure 6: Kaplan-Meier survival curves of CIS and Co-CIS-associated genes.



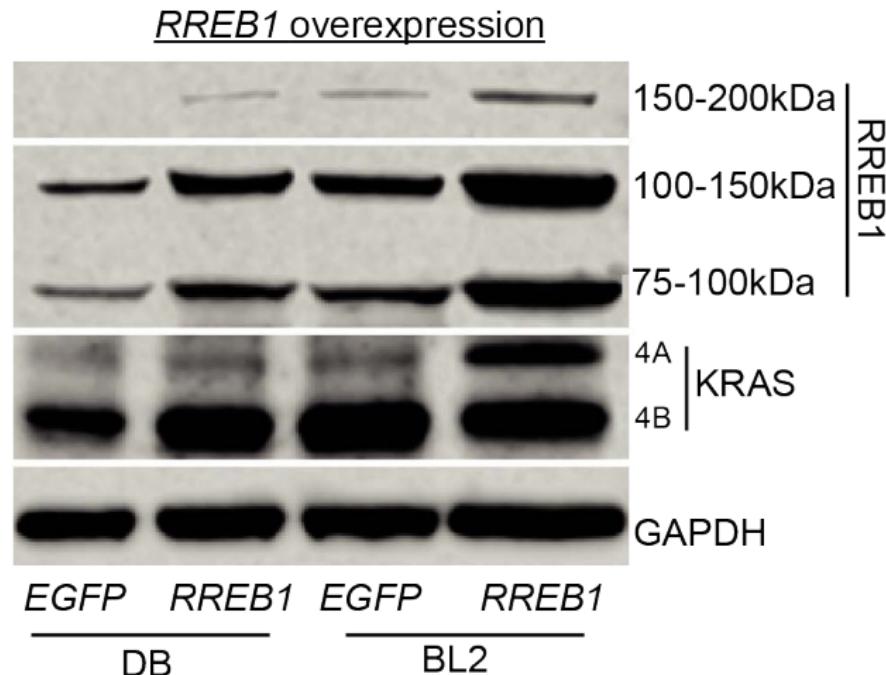
Supplemental Figure 6: Kaplan-Meier survival curves of CIS and Co-CIS-associated genes.
 Kaplan-Meier survival curve comparing animal cohorts containing the respective CIS-associated gene versus animals that do not have the CIS. Statistics were performed with Prism GraphPad 6: Log-rank Mantel Cox test. p-values are listed on the graphs. Median survival and N values for each cohort are listed in Supplemental Table 5.

Supplemental Figure 7: UCSC genome browser view chr4: 94961700-94971700



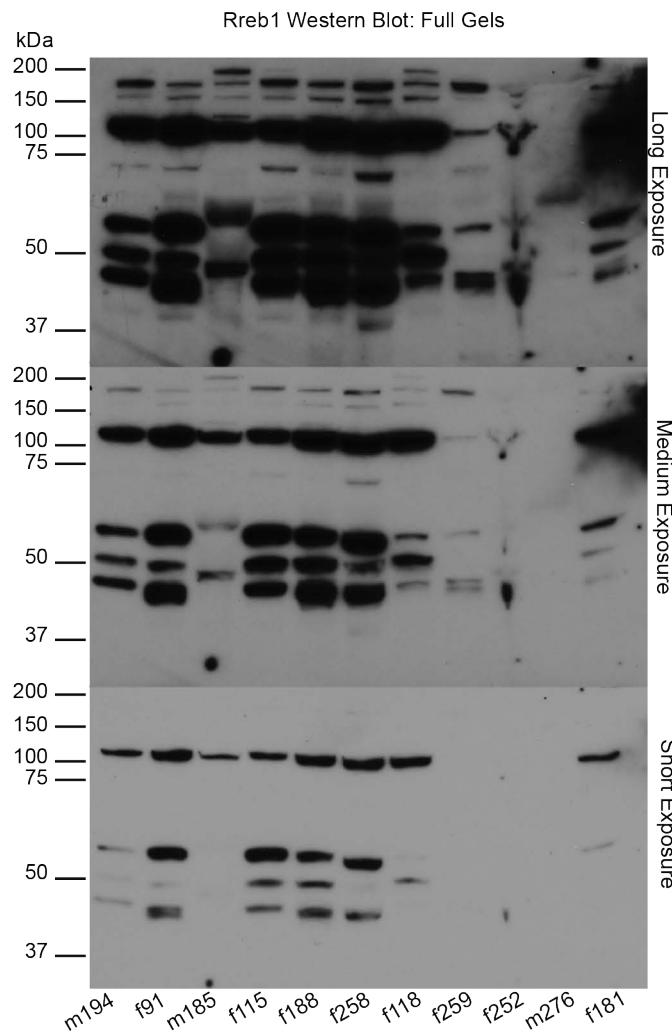
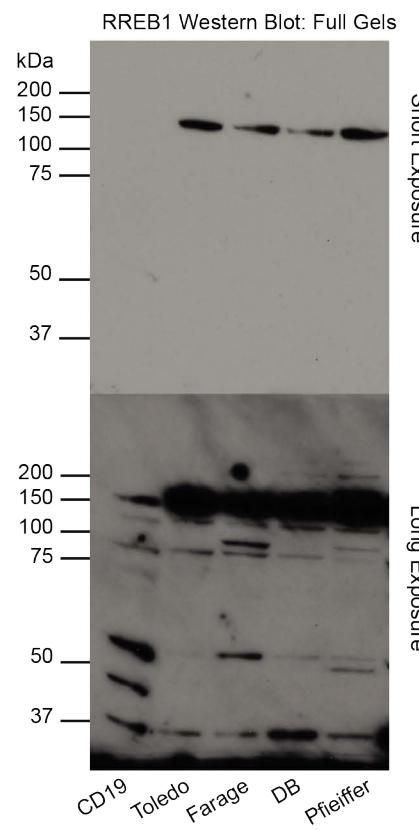
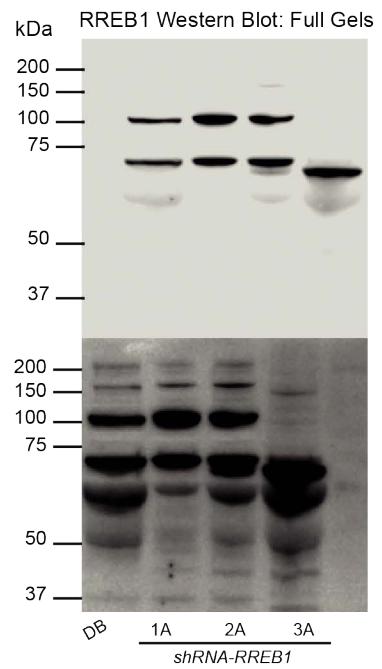
Supplemental Figure 7: UCSC genome browser view chr4: 94961700-94971700. Image depicts a screen shot of the UCSC genome browser at the CIS location site spanning 10kb. On the left is listed the ENCODE methylation analysis at various histone marks in the region. The dark grey boxes to the left indicate regions that have altered methylation.

Supplemental Figure 8: RREB1 overexpression



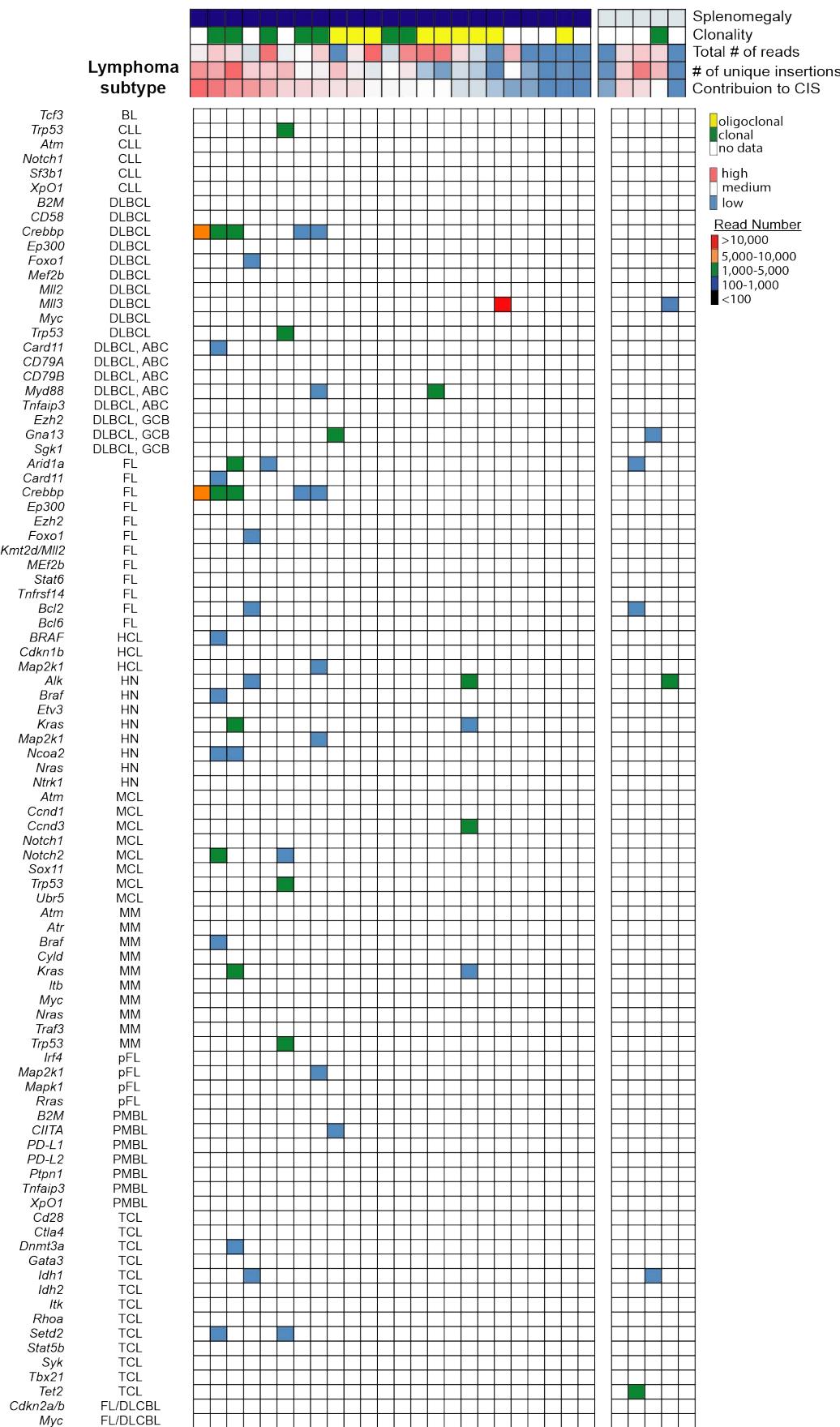
Supplemental Figure 8: RREB1 overexpression. Image of a Immunoblot for RREB1, KRAS and GAPDH on DLBCL cell line DB and Burkitt's lymphoma cell line BL2. The cell lines overexpress either *EGFP* (control) or *RREB1*.

Supplemental Figure 9: Full immunoblots

A**B****C**

Supplemental Figure 9: Full immunoblots. A) Immunoblot of Rreb1 on mouse spleen samples shown in Figure 4. B) Immunoblot of human DLBCL cell lines from Figure 6. C) Immunoblot of DB cells with shRNA targeting Rreb1 from Figure 6.

Supplemental Figure 10: Comparative analysis of CIS genes to known lymphoma genes



Supplemental Figure 10: Comparative analysis of CIS genes to known lymphoma genes. A list of annotated genes from WHO2016 classifications and the literature were queried for *T2/Onc* insertions in *SB*-mutagenized spleen samples. BL = Burkitt's lymphoma, MCL = mantle cell lymphoma, MM = multiple myeloma, pFL = pediatric follicular lymphoma, TCL = T cell lymphoma