SUPPLEMENTAL INFORMATION

Accelerated progression of chronic lymphocytic leukemia in Eμ-TCL1 mice expressing catalytically inactive RAG1

Vincent K. Nganga, Victoria L. Palmer, Hina Naushad, Michele D. Kassmeier, Dirk K. Anderson, Greg A. Perry, Nathan M. Schabla, and Patrick C. Swanson

SUPPLEMENTAL METHODS

Mice

Genotyping for the dnRAG1 and Eμ-TCL1 transgenes was performed as described [1,2]. Cohorts were either sacrificed at predetermined ages (6, 12, 24, and 36 weeks), or when they became moribund. All mice were housed in individually ventilated microisolator cages in an AAALAC certified animal facility in accordance with university and federal guidelines, and mouse protocols were approved by the Creighton University Institutional Animal Care and Use Committee. Survival data was obtained by monitoring a cohort of 11 Eμ-TCL1 and 14 DTG mice until the animals became visibly ill and necessitated euthanasia. This data was used to generate Kaplan-Meier plots to compare survival characteristics between the genotypes and was further analyzed for significance using the log-rank test.

Flow cytometry and cell sorting

Single-cell suspensions prepared from spleen, bone marrow, lymph nodes, and peripheral blood were depleted of red blood cells by hypotonic lysis and stained with fluorochrome-conjugated antibodies as described earlier [3]. The following antibodies were used: BD Biosciences (San Jose, CA) anti-B220-PE-TXRD (RA3-6B2), anti-CD19-APC-Cy7 (ID3), anti-CD5-biotin (53-7.3), anti-CD21/CD35-PE (7G6), anti-CD11b-PE (M1/70), anti-CD23-Biotin (B3B4), anti-CD4-APC-Cy7 (GK1.5), and anti-CD25-PE-Cy7 (PC61), and eBioscience (San Diego, CA) anti-CD5-PE (53-7.3), anti-IgM-APC (II/41), anti-IgD-FITC (11-26c), anti-CD3-APC (145-2C11), anti-CD8-A700 (53-6.7), anti-TCRβ-FITC (H57-597). Samples stained with biotinylated antibodies were detected using streptavidin-Qdot585 (Invitrogen, Carlsbad, CA). Data collection and cell sorting was performed using a FACSAria flow cytometer (BD Biosciences). Data was analyzed using the FlowJo software (Tree Star, Inc. Ashland, OR).

Ig gene analysis

To analyze clonality and V(D)J recombination status, genomic DNA was isolated from spleen tissue using the Wizard Genomic DNA Purification Kit (Promega, Madison, WI), and 10-15μg was digested overnight with EcoR1, separated on a 0.8% agarose gel, transferred to Amersham Hybond-N+ nylon membrane (GE Healthcare, Piscataway, NJ) and hybridized with a ^{32}P-labeled J_H probe. Phosphor images were acquired using a Typhoon 9410 variable mode imager (GE Healthcare).

To analyze Ig gene usage and mutational status, RNA was isolated from spleen tissue obtained from three ill Eμ-TCL1 and three ill DTG mice using the Ambion RiboPure Kit (Novagen) according to manufacture instructions. cDNA was prepared from 1μg of RNA in a 20μL reaction using First Strand cDNA Synthesis Kit (Novagen) with oligo(dT) primers according to manufacturer instructions. Ig genes were amplified from cDNA (1-2μL) by PCR using Platinum *Taq* DNA Polymerase High Fidelity (Invitrogen) and MuIgV_H5′- A to -F forward and MuIgMV_H3′-1 reverse primers to amplify heavy chain sequences (Mouse Ig-Primer Sets, Novagen), and mouse universal 5′ Mκ forward and 3′ κC reverse primers[4] to amplify light chain sequences.

PCR products were gel-isolated using the GeneJET Gel Extraction Kit (Fermentas, Thermo Scientific, Waltham, MA) and cloned into the PCR 2.1 TOPO TA vector (Invitrogen). Isolated colonies were grown in minicultures overnight and plasmid DNA was purified using the E.Z.N.A Plasmid Mini Kit I (Omega Bio-Tek, Norcross, GA). For both heavy and light chain genes, at least 10 independent clones per mouse were sequenced using a commercial vendor (ACGT, Inc., Wheeling, IL). Sequences, with primer sites omitted, were analyzed using IMGT/V-QUEST tool (http://www.imgt.org) to identify Ig gene usage, mutations, and CDR3 composition and isoelectric point.

Microarray

Total RNA was isolated from sorted splenic $CD19^{+}B220^{hi}CD5^{-}$ B cells obtained from WT mice or splenic $CD19^{+}CD5^{+}$ B cells obtained from dnRAG1, Eμ-TCL1 and DTG mice using the Ribopure kit (Ambion, Austin, TX) according to manufacturer instructions. Biotin-end labeled cDNA was prepared from total RNA (100-200 ng) using whole transcript labeling kits from either Affymetrix (Affymetrix, Santa Clara, CA) or Ambion per manufacturer instructions. Resultant cDNA was hybridized overnight to Mouse Gene 1.0 ST Arrays and washed, stained, and scanned using the Affymetrix GeneChip system with a 3000 7G Affymetrix scanner at the University Nebraska Medical Center Microarray Core Facility. All procedures were conducted following Affymetrix suggested protocols. Array data sets were normalized using the Robust Multichip Average (RMA) algorithm included in the Affymetrix Expression Console software. Further analyses were performed using dChip (http://biosun1.harvard.edu/complab/dchip/)[5].

Data sets obtained from 12 week-old animals and those obtained from splenic and peritoneal B1a B cells, and splenic developing (transitional T1-T3), mature (marginal zone and mature follicular), and activated (germinal center) subsets available within the NCBI GEO series GSE15907 through the Immunological Genome Project Consortium (ImmGen)[6] were normalized using the RMA algorithm in the Affymetrix Expression Console. Batch effects were adjusted using ComBat[7]. A hierarchical cluster was generated using dChip using 102 genes with the greatest variation that were identified by the filtering criteria which set the standard deviation for logged data between 1.10 and 1000, and an expression level at greater than or equal to 5.65 in 25% of samples. Principal component analysis was also performed by dChip using the filtered gene set.

Real-time quantitative PCR

Primer sets used for qPCR are as follows:

Prl2a1
Forward: 5′-GGAAAAGAGCAATGGACTCCTGG-′3; Reverse: 5′-CAGTCTCTGACTTCAAGGATGCC-′3

Il10
Forward: 5′-CGGGAAGACAATAACTGCACCC-′3; Reverse: 5′-CGGTTAGCAGTATGTTGTCCAGC-′3

Sox4
Forward: 5′-GCCTCCATCTTCGTACAACC-′3; Reverse: 5′-AGTGAAGCGCGTCTACCTGT-′3

Rgs13
Forward: 5′-CTACATCCAGCCACAGTCTCCT-′3; Reverse: 5′-TGAGCTTCTTCAAAGCATGTTTGAG-′3

Sell
Forward: 5′-ATGGTGAGCATCCCAGCCTA -′3; Reverse: 5′-CCCCTTCCAGCATTCCATCA-′3

Actin beta
Forward: 5′-AGAGGGAAATCGTGCGTGAC -′3; Reverse: 5′-CAATAGTGACCTGGCCGT -′3

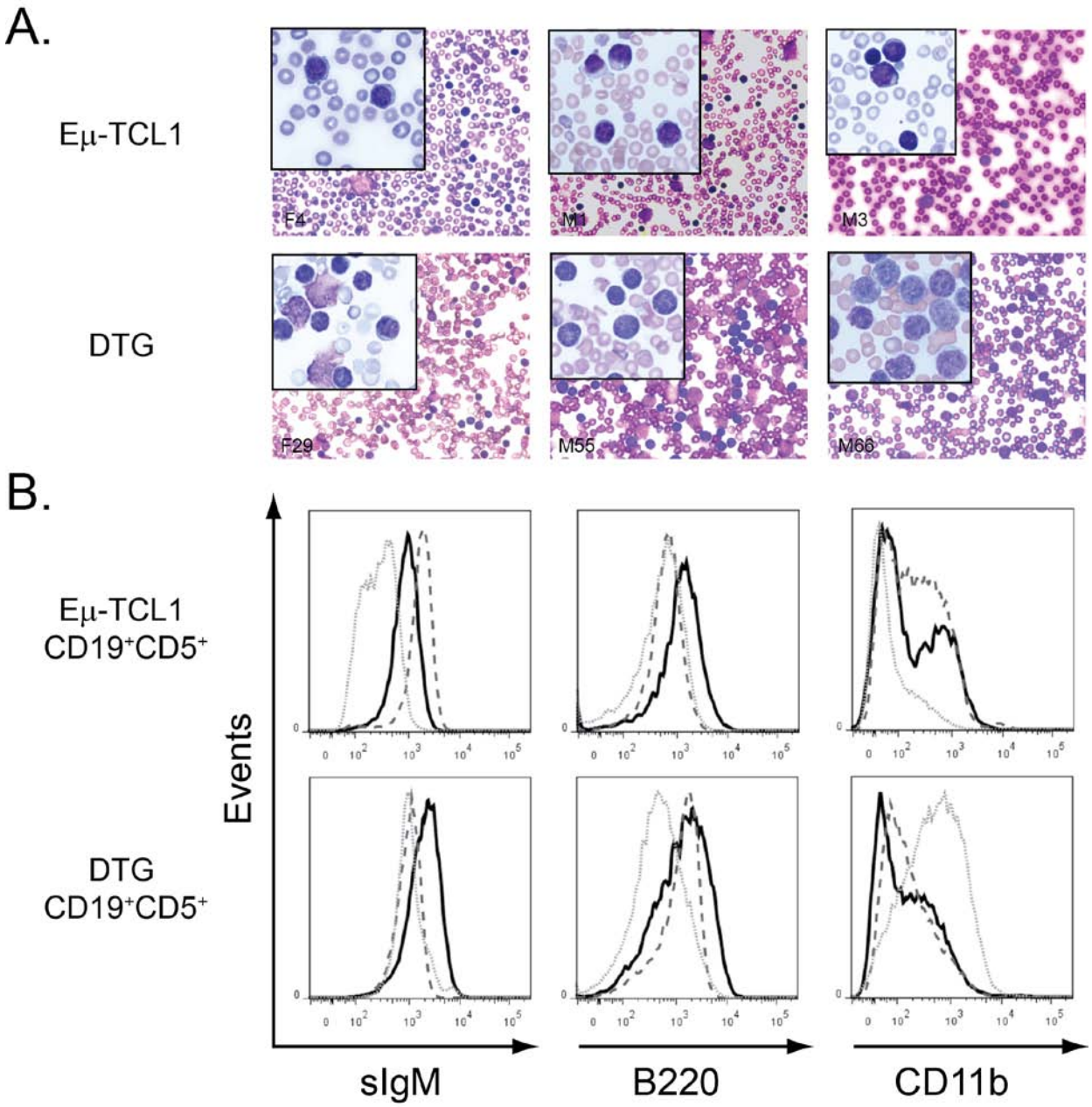


Figure S1. Histological and flow cytometric evaluation of ill Eμ-TCL1 mice and DTG mice.

(A) Peripheral blood smears of ill Eμ-TCL1 and DTG mice at end point were stained with Wright-Giemsa and imaged as in Fig. 2A (400x). Magnified images of representative cells are shown in the inset. (B) Flow cytometry was used to analyze splenic lymphocytes from three different ill Eμ-TCL1 mice (top row) and DTG mice (bottom row) for the expression of sIgM, B220, and CD11b on $CD19^+CD5^+$ B cells. Histograms for the three different mice are shown as solid, dashed, and dotted lines.

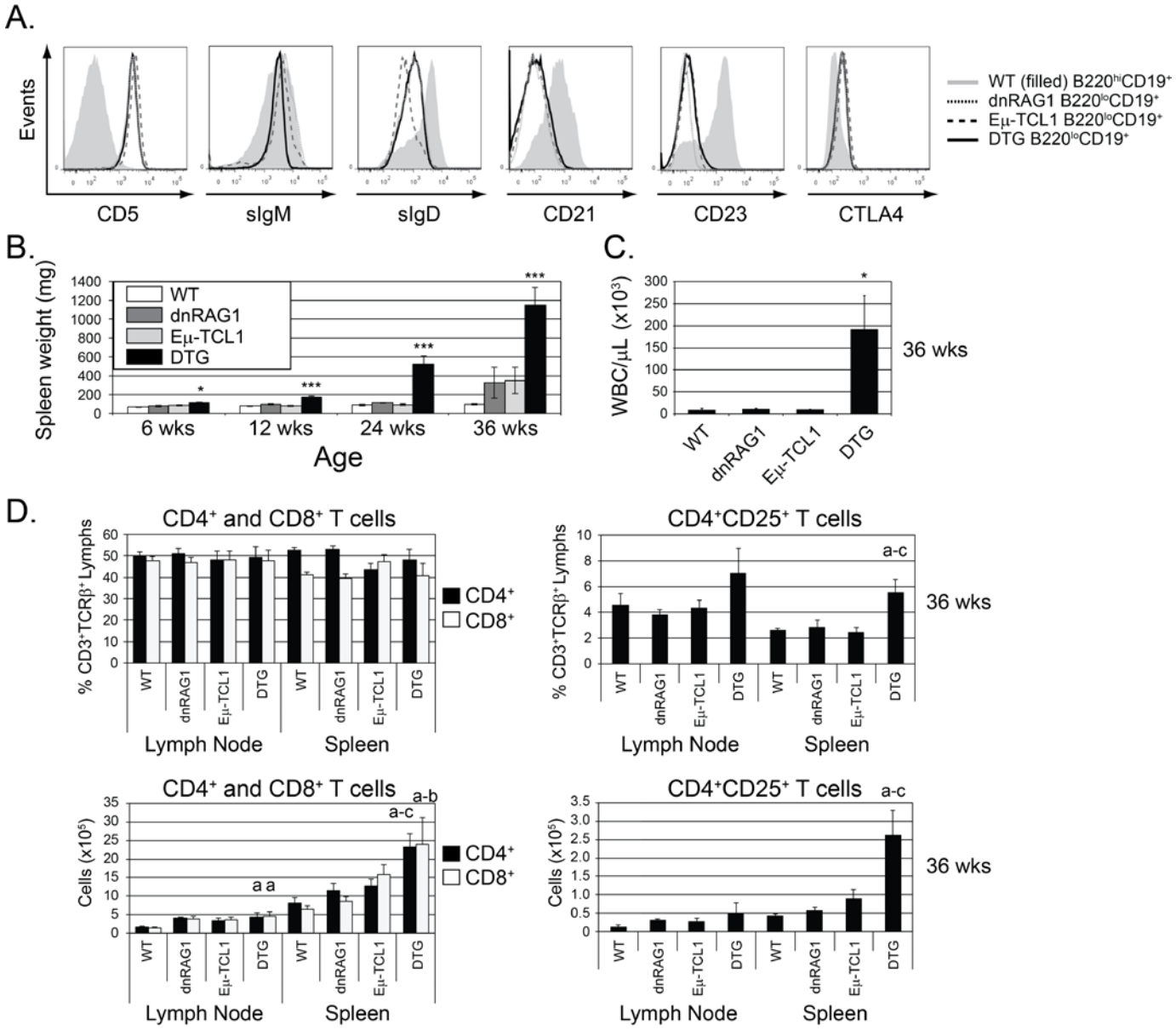


Figure S2. Characterization of CLL-like disease progression in DTG mice.

(A) Flow cytometry was used to compare the expression of CD5, sIgM, sIgD, CD21, CD23, and CTLA4 expression on splenic WT $CD19^{+}B220^{hi}$ B cells to splenic $CD19^{+}B220^{lo}$ B cells from dnRAG1, Eμ-TCL1 mice and DTG mice. Spleen weights (B) and white blood cell (WBC) counts (C) were compared for WT, dnRAG1, Eμ-TCL1 mice and DTG mice at either 6, 12, 24, and 36 weeks of age (B) or at 36 weeks of age only (C). Error bars represent the standard error of the mean. 5-6 animals of each genotype were analyzed at each time point. Values obtained for DTG mice are significantly different from those obtained for WT, dnRAG1, or Eμ-TCL1 mice (*, $p<0.05$; ***, $p< 0.001$). (D) The percentages of gated $CD3^{+}TCR\beta^{+}$ lymphocytes that are $CD4^{+}$, $CD8^{+}$, or $CD4^{+}CD25^{+}$ T cells (top) and the absolute numbers of these T cell populations (bottom) in the lymph nodes and spleen of 36 week-old WT, dnRAG1, Eμ-TCL1, and DTG mice (n=8-14 per genotype) as determined by flow cytometry are shown in bar graph format. Error bars represent the standard error of the mean. Statistically significant differences ($p<0.05$) between values obtained for DTG mice relative to WT (***a***), dnRAG1 (***b***), or Eμ-TCL1 (***c***) mice are indicated.

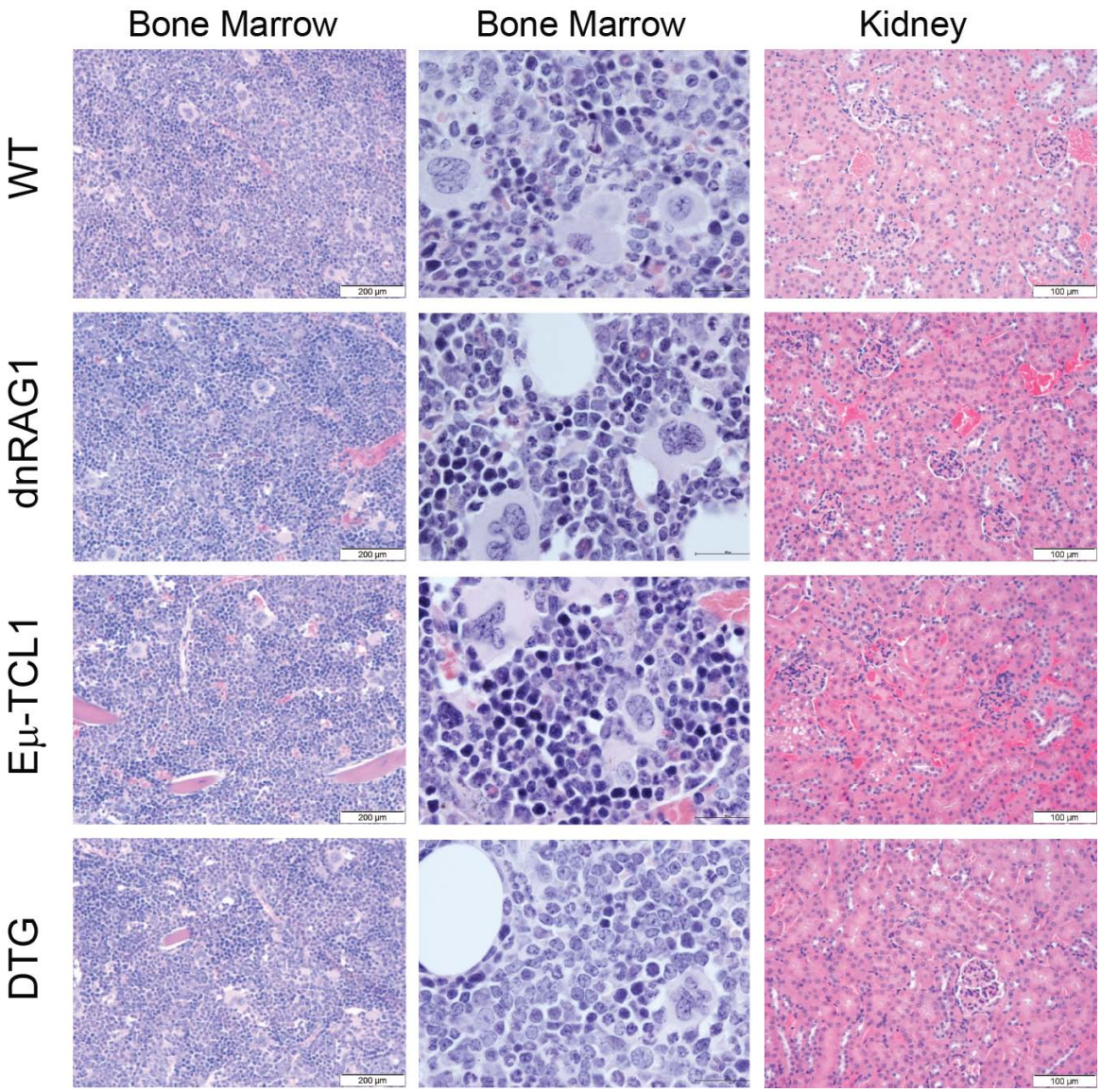


Figure S3. Histology of bone marrow and kidney sections of 36 week-old WT, dnRAG1, Eμ-TCL1, and DTG mice.

Paraffin-embedded bone marrow and kidney sections were developed with hematoxylin and eosin. Images in columns 1 and 3 (100x for bone marrow, and 200x for kidney) were acquired using a Nikon i80 microscope and DigiFire camera running ImageSys digital imaging software (Soft Imaging Systems GmBH, Munster, Germany). Bone marrow images in column 2 (1000x) were acquired using a Nikon i80 microscope and Nikon Digital Sight DS-F1 camera running the NIS-Elements Imaging software version 2.33. Bone marrow and kidney show little or no abnormal infiltration at this time point. Scale bars: bone marrow, 200 μM; kidney, 100 μM.

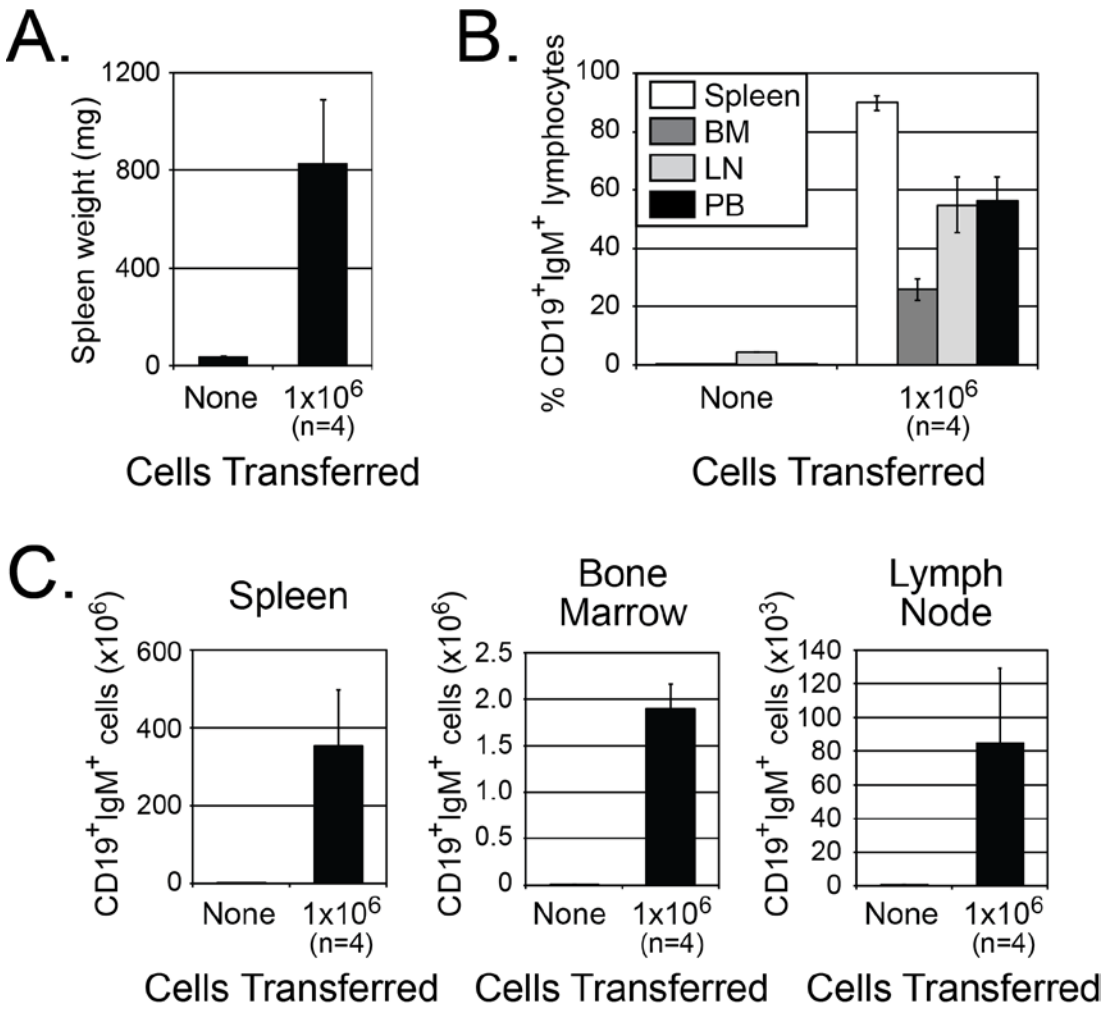


Figure S4. Analysis of SCID mice engrafted with leukemic B cells from DTG mice.

SCID mice receiving either no cells or $1x10^6$ sorted $CD19^+CD5^+$ B cells from one of two different 36 week-old DTG mice (cells from each DTG mouse were transferred into two SCID recipients; n=4 total) were sacrificed at 3 months post-transfer. Spleen weights were measured (A), and the percentage (B) and/or absolute number (C) of $CD19^+IgM^+$ B cells were determined in spleen, bone marrow (BM), lymph node (LN), and peripheral blood (PB).

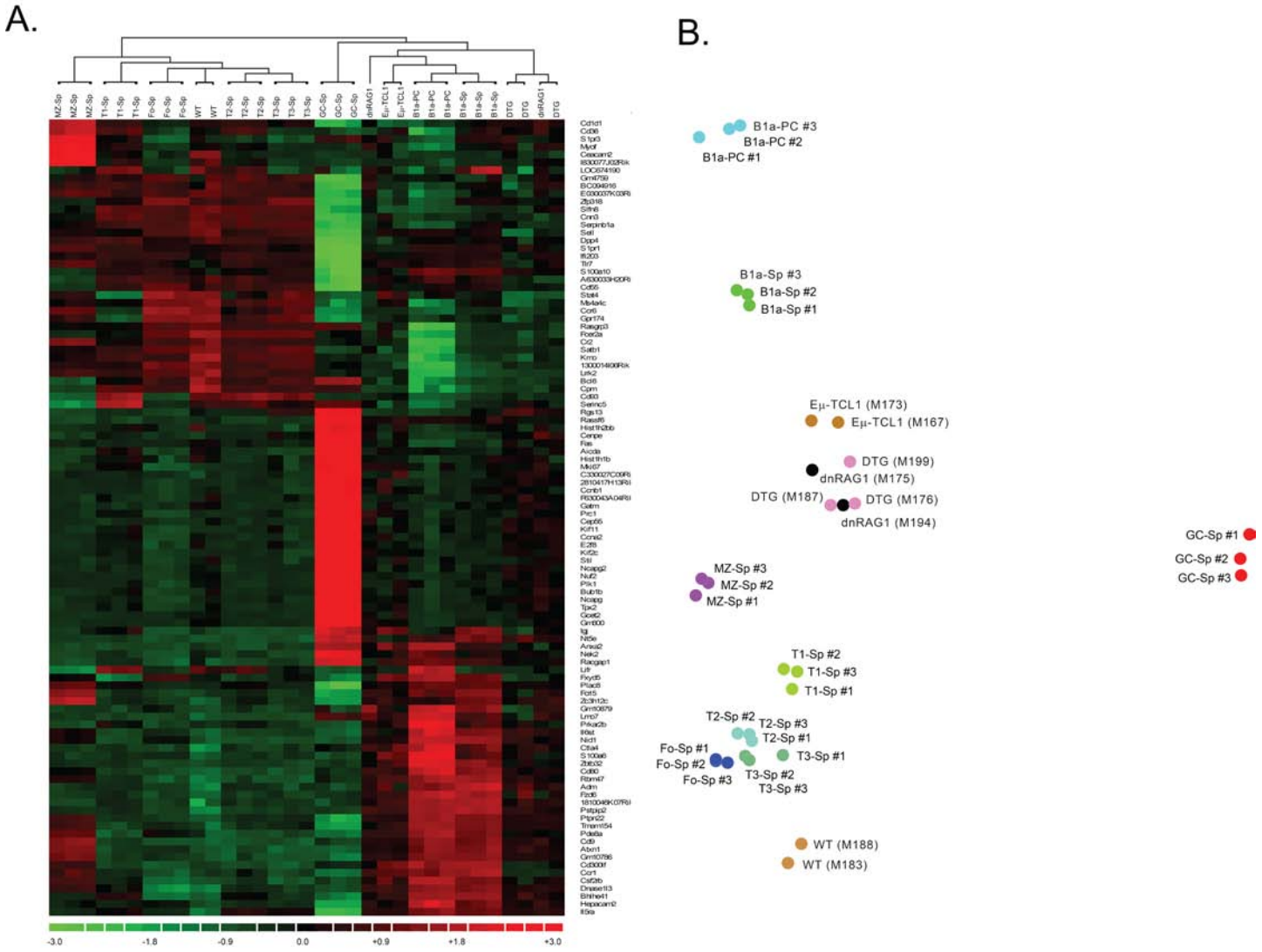


Figure S5. CD5^{+} B cells from dnRAG1, Eμ-TCL1, and DTG mice are most similar to normal B1a B cells.

(A) The gene expression profiles of sorted B cells from 12 week-old mice in this study were compared to those from normal developing, mature, and activated B cell subsets by unsupervised hierarchical clustering analysis. Clustering was performed by dChip on batch effect-corrected $\log_2$ expression values for 102 genes (rows) and 33 arrays (columns), which includes those from sorted splenic WT CD19^{+}B220^{hi}CD5^{-} (WT) and transgenic CD19^{+}B220^{lo}CD5^{+} (dnRAG1, Eμ-TCL1, and DTG) B cells from 12 week-old mice reported here, and those obtained from sorted splenic and peritoneal B1a cells (B1a-Sp and B1a-PC), splenic transitional B cells (T1-T3-Sp), splenic marginal zone (MZ-Sp) and follicular B cells (Fo-Sp), and splenic germinal center (GC-Sp) B cells available through the ImmGen database[6]. Red and green intensities indicate high to low gene expression, respectively. The dendrogram shows the relationships between the populations. (B) Principal component analysis. PCA was performed on 102 genes identified in (A) by dChip.

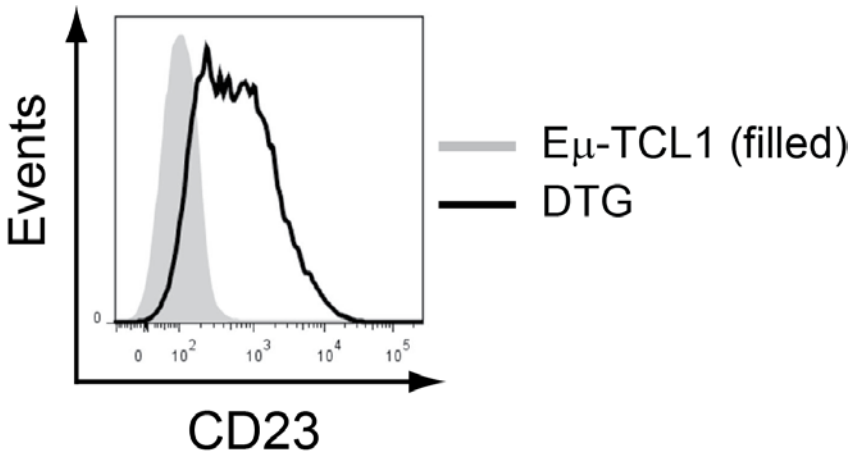
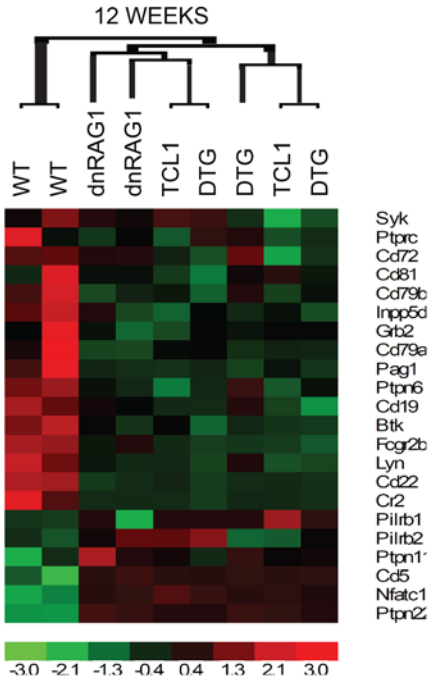


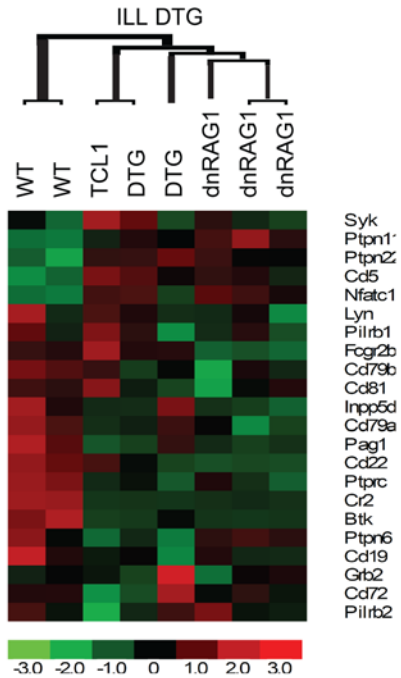
Figure S6. CD23 expression on leukemic cells from an ill DTG mouse.

Flow cytometry was used to compare CD23 expression on $CD19^{+}B220^{lo}$ B cells from the Eμ-TCL1 mouse (M52) and the ill DTG mouse (F57) used for the comparative gene expression analysis found in Table S5, which showed that Fcer2a (CD23) was upregulated ~32-fold on $CD19^{+}CD5^{+}$ B cells from the DTG mouse relative to those from the Eμ-TCL1 mouse.



Expression of B cell receptor signaling genes in 12 week-old mice

	Fold change			
Gene Symbol	**dnRAG1/WT**	**TCL1/WT**	**DTG/WT**	**Gene Description**
Ptpn22	3.15	2.92	2.91	protein tyrosine phosphatase, non-receptor type 22 (lymphoid)
Nfatc1	2.32	2.46	2.29	nuclear factor of activated T-cells, cytoplasmic, calcineurin-dependent 1
Cd5	1.91	1.85	1.86	CD5 antigen
Ptpn11	1.44	1.26	1.26	protein tyrosine phosphatase, non-receptor type 11
Pilrb2	1.15	1.08	1.09	paired immunoglobin-like type 2 receptor beta 2
Pilrb1	-1.01	1.07	1.05	paired immunoglobin-like type 2 receptor beta 1
Syk	-1.06	-1.17	-1.12	spleen tyrosine kinase
Grb2	-1.26	-1.22	-1.16	growth factor receptor bound protein 2
Ptprc	-1.25	-1.42	-1.17	protein tyrosine phosphatase, receptor type, C
Ptpn6	-1.22	-1.40	-1.17	protein tyrosine phosphatase, non-receptor type 6
Cd81	-1.13	-1.13	-1.19	CD81 antigen
Cd79a	-1.30	-1.19	-1.21	CD79A antigen (immunoglobulin-associated alpha)
Cd79b	-1.26	-1.25	-1.23	CD79B antigen
Cd19	-1.16	-1.25	-1.26	CD19 antigen
Cd72	-1.17	-1.75	-1.30	CD72 antigen
Inpp5d	-1.27	-1.35	-1.30	inositol polyphosphate-5-phosphatase D
Lyn	-1.31	-1.40	-1.35	Yamaguchi sarcoma viral (v-yes-1) oncogene homolog
Btk	-1.42	-1.43	-1.59	Bruton agammaglobulinemia tyrosine kinase
Fcgr2b	-1.33	-1.52	-1.60	Fc receptor, IgG, low affinity IIb
Pag1	-2.00	-1.93	-2.15	phosphoprotein associated with glycosphingolipid microdomains 1
Cd22	-3.05	-2.60	-3.04	CD22 antigen
Cr2 (CD21)	-4.49	-4.50	-5.39	complement receptor 2



Expression of B cell receptor signaling genes in older mice

	Fold change			
Gene Symbol	**dnRAG1/WT**	**TCL1/WT**	**DTG/WT**	**Gene Description**
Ptpn22	2.55	2.82	3.12	protein tyrosine phosphatase, non-receptor type 22 (lymphoid)
Cd5	2.34	3.29	2.24	CD5 antigen
Nfatc1	2.62	2.72	1.64	nuclear factor of activated T-cells, cytoplasmic, calcineurin-dependent 1
Ptpn11	1.82	1.40	1.46	protein tyrosine phosphatase, non-receptor type 11
Cd81	-1.15	1.11	1.17	CD81 antigen
Grb2	-1.03	1.00	1.13	growth factor receptor bound protein 2
Syk	1.11	1.61	1.06	spleen tyrosine kinase
Cd79a	-1.12	-1.10	1.01	CD79A antigen (immunoglobulin-associated alpha)
Fcgr2b	-1.53	1.27	1.00	Fc receptor, IgG, low affinity IIb
Cd79b	-1.10	-1.02	-1.02	CD79B antigen
Pilrb2	1.00	-1.18	-1.03	paired immunoglobin-like type 2 receptor beta 2
Cd72	-1.06	-2.86	-1.12	CD72 antigen
Lyn	-1.12	1.01	-1.16	Yamaguchi sarcoma viral (v-yes-1) oncogene homolog
Inpp5d	-1.59	-1.42	-1.16	inositol polyphosphate-5-phosphatase D
Pilrb1	-1.08	1.10	-1.18	paired immunoglobin-like type 2 receptor beta 1
Ptpn6	-1.02	-1.30	-1.20	protein tyrosine phosphatase, non-receptor type 6
Btk	-1.28	-1.30	-1.36	Bruton agammaglobulinemia tyrosine kinase
Cd19	-1.12	-1.12	-1.40	CD19 antigen
Ptprc	-1.35	-1.33	-1.50	protein tyrosine phosphatase, receptor type, C
Pag1	-3.24	-6.53	-1.82	phosphoprotein associated with glycosphingolipid microdomains 1
Cd22	-3.38	-1.20	-2.99	CD22 antigen
Cr2 (CD21)	-12.25	-26.54	-10.40	complement receptor 2

Figure S7. Hierarchical clustering analysis of data sets from 12 week-old and older WT and transgenic animals using genes involved in B cell receptor signaling.

A supervised hierarchical clustering analysis was performed using a set of 22 genes known to play a role in B cell receptor signaling pathways. The relative expression of these genes in WT $CD19^{+}B220^{hi}CD5^{-}$ B cells compared to transgenic $CD19^{+}CD5^{+}$ B cells is shown at right for 12 week old mice (top panel) and older animals (bottom panel).

Table S1. Analysis of IgVH sequences from ill Eμ-TCL1 and DTG mice

Mouse ID	Clone #(s)	VH gene †	Identity to germline	D gene	J gene	CDR3 ‡	pI	VH gene reported reactivity ¥
Eμ-TCL1 M1	A01,A11	5-12*01	100%	1-1*01	4*01	CARKRAGYGSREYAMDYW	8.97	
	A06	5-12*01	99.65%	1-1*01	4*01	CARKRAGYGSREYAMDYW	8.97	
	A02,A07	5-12*01	100%	4-1*01	1*03	CARLNWGVGYFDVW	6.15	
	A03	5-4*01	99.65%	1-1*01	4*01	CARELGDYGREVYAMDYW	4.42	
	A05	5-6*01	100%	1-1*01	4*01	CARHGSNYDHYAMDYW	6.48	
	A08	7-3*01	100%	1-2*01	3*01	CARCSPYDAWFAYW	6.10	Sm[8]
	B01	1-22*01	100%	3-1*01	1*03	CARSGAGYYWYFDVW	6.14	
	B05,B07	1-52*01	100%	2-5*01	1*03	CARYYSNYWYFDVW	6.14	
	B11	1-52*01	99.65%	2-5*01	1*03	CARYYSNYWYFDVW	6.14	
	B02	1-52*01	99.31%	2-5*01	1*03	CARYYSNYWYFDVW	6.14	
	B03-2	1-52*01	100%	2-14*01	2*01	CARISANLLEDYW	4.44	
	B04	1-81*01	100%	1-1*01	2*01	CARADYYGSSYGGDYW	4.44	Sm[8], DNA[9], ssDNA[10]
	B09	14-3*01	100%	2-5*01	1*03	CATYYSNYVRFAYW	8.19	
	B06	1-62-2*01	99.65%	1-1*01	1*03	CARHEDYYYGSSWYFDVW	4.17	
	B10	1-78*01	98.61%	2-4*01	3*01	CARGIYYDYDVFYW	4.44	Thymocyte[11]
Eμ-TCL1 F13	B01,B02, B03, B04-2, B05-2, B06, B08, E01, E02, E03, E04, E05, E06, E07	1-52*01	100%	2-5*01	1*03	CARYYSNYWYFDVW	6.14	
	B07, B09-2, E08,	1-52*01	99.65%	2-5*01	1*03	CARYYSNYWYFDVW	6.14	
Eμ-TCL1 F26	B04, B06, B07, B08 B09, B11, E01, E08, E09	1-55*01	100%	1-1*01	2*01	CASIYYYGSSYYFDYW	3.75	DNA[12]
	B01, B02, B03, B10	1-85*01 (unproductive - stop codons, out of frame) 1-72*01 (unproductive - stop codons,	100%	2-5*01	2*01	CAR**#DYW	NR	
	E02, E03, E04, E05 E06, E07, E10, E08	out of frame)	100%	2-3*01	1*03	CA*MVT#WYFDVW	NR	

Mouse ID	Clone #(s)	VH gene †	Identity to germline	D gene	J gene	CDR3 ‡	pI	VH gene reported reactivity ¥
DTG M19	B01, B02	1-15*01	100%	1-1*01	2*01	CITTVVABYFDYW	5.45	
	B03, B14	1-15*01	99.65%	1-1*01	2*01	CITTVVABYFDYW	5.45	
	B05, B11	1-55*01	100%	1-1*01	2*01	CARGYYYGSSYFDYW	6.14	DNA[12]
	B07	1-78*01	99.65%	none	2*01	CARNW	8.25	Thymocyte[11]
	B04	5-17*01	100%	2-12*01	1*03	CAIYYRYFDVW	6.14	DNA[13]
	B08	5-17*01	100%	2-5*01	2*01	CAMAYYSNYFDYW	3.75	
	B15	5-17*01	100%	2-4*01	3*01	CAKLRRFAYW	10.02	
	C10	5-17*01	99.65%	1-1*01	2*01	CARDITTVVAPDYW	4.44	
	C01	9-3*01	99.65%	1-3*01	3*01	CARECPLTWSWFAYW	6.10	
	C02, C03, C04, C05, C06, C07, C09	2-5*01	100%	2-3*01	3*01	CAKEGYYAPFAYW	6.14	
DTG F20	A02, A03, A05, A06, A09	5-12*01	100%	3-2*02	4*01	CARHSSGYAMDYW	7.17	
	A07	5-12*01	99.65%	3-2*02	4*01	CARHSSGYAMDYW	7.17	
	A01, A10	5-12*01	99.65%	4-1*01	2*01	CARHGLDLDYW	5.61	
	A08	5-6*01	99.65%	1-1*01	1*03	CARHTTVVASYWYFDVW	7.17	
	B06	1-9*01	99.31%	1-1*01	3*01	CARDYYGSSYGGFAYW	6.14	Cytochrome C[14], cardiolipin[15]
	C01, C02, C05, C07, C08, C09-2	2-2*01	100%	2-4*01	2*01	CARYDYAAYYFDYW	4.44	
	C03, C04, C06	2-2*01	99.65%	2-4*01	2*01	CARYDYAAYYFDYW	4.44	
	B08, B10	1-70*01(ψ)	99.65-100%	1-1*01	1*03	CARFYYGSSWYFDVW	6.14-7.17	
DTG F29	A03, A07, A10, A11	5-6*01	100%	2-1*01	3*01	CARHYGNYWFAYW	8.22	
	A04, A08	5-6*01	99.65%	2-1*01	3*01	CARHYGNYWFAYW	8.22	
	A09	5-6*01	99.31%	2-1*01	3*01	CARHYGNYWFAYW	8.22	
	A01	5-6*01	100%	1-1*01	2*01	CARGNFDYW	6.15	
	A05	5-6-5*01 (unproductive - stop codons, out of frame)	82.98%	none	3*01	VQDTMVTTGLLT	NR	

† VH genes highlighted in yellow appear in both Eμ-TCL1 and DTG mice; VH genes in bold red font were also identified in sequences analyzed by Yan *et al.*[16]. Pseudogenes are indicated by (ψ).

‡ CDR3 sequences highlighted in green appear in both Eμ-TCL1 and DTG mice and were also identified in sequences analyzed by Yan *et al.*[16] Mutated amino acid residues are underlined. Frameshift mutations and in-frame stop codons are indicated by (#) and (*), respectively.

¥ Sm, Smith antigen

Table S2. Analysis of IgVL sequences from ill Eμ-TCL1 and DTG mice

Mouse ID	Clone #(s)	Vκ gene †	Identity to germline	Jκ gene	CDR3 ‡	pI	Vκ gene reported reactivity ¥
Eμ-TCL1 M1	01,12,03,05,06, 08, 09	4-91*01	100%	2*01	CQQGSSIPRTF	8.25	DNA[17,18], PtC[19]
	11	4-91*01	100%	4*01	CQQGSSIPRTF	8.25	DNA[17,18], PtC[19]
	07	4-91*01	99.61%	2*01	CQQGSSIPRTF	8.25	DNA[17,18], PtC[19]
	02	1-133*01	99.63	2*01	CVQGAHFPYTF	7.17	
Eμ-TCL1 F13	01,10,11,12,02,04,06,07	1-117*01	100%	4*01	CFQGSHVPLTF	7.17	Myeloperoxidase[20],
	08	1-117*01	99.63%	4*01	CFQGSHVPLTF	7.17	DNA[9,13,18]
	09	1-117*01	97.41%	4*01	CFQGSHVPLTF	7.17	phosphocholine[21]
Eμ-TCL1 F26	01,03,05,08,09,06	12-89*01	100%	2*01	CQNVLSTPPTF	5.49	
	10	12-89*01	99.61	2*01	CQNVLSTPPTF	5.49	
	02	14-126*01	99.22%	1*01	CLQHGESPWTF	5.45	Br-treated RBCs[22], PtC[19]
	04, 07	14-126*01	100%	1*01	CLQHGESPWTF	5.45	Br-treated RBCs[22], PtC[19]
DTG M19	04,05,07,09	4-91*01	100%	4*01	CQQGSSIPRTF	8.25	DNA[17,18], PtC[3]
	06,08	4-91*01	99.61%	4*01	CQQGSSIPRTF	8.25	DNA[17,18], PtC[3]
	11	4-91*01	100%	2*01	CQQGSSIPRTF	8.25	DNA[17,18], PtC[3]
	02	4-91*01	100%	4*01	CQQGSSIPRTF	8.25	DNA[17,18], PtC[3]
	03	3-2*01	100%	2*01	CQQSKEVPYTF	6.15	DNA[2,7]
	01	14-126*01	100%	2*01	CLQHGESPYTF	5.45	Br-treated RBCs[22], PtC[19]
DTG F20	10,11,13,14,04,07	14-126*01	100%	2*01	CLQHGESPYTF	5.45	Br-treated RBCs[22], PtC[19]
	12	14-126*01	99.61%	2*01	CLQHGESPYTF	5.45	Br-treated RBCs[22], PtC[19]
	08	14-126*01	99.61	2*01	CLRHGESPYTF	7.17	Br-treated RBCs[22], PtC[19]
	05	4-77*01 (ψ)	100%	4*01	CQQWSSSPP#F	NR	
	06	4-77*01 (ψ)	99.60%	4*01	CQQWSSSPP#F	NR	
DTG F29	14,08	14-126*01	100%	2*01	CLQHGESPYTF	5.45	Br-treated RBCs[22], PtC[19]
	02,05,06	14-126*01	99.61%	2*01	CLQHGESPYTF	5.45	Br-treated RBCs[22], PtC[19]
	07	14-126*01	100%	2*01	CLQHGESPYT	not identified	Br-treated RBCs[22], PtC[19]
	01	19-93*01	100%	1*01	CLQYDNLRTF	6.15	Sm[8], DNA[13]
	11	4-91*01	99.61%	5*01	CQQGSSIPLTF	5.49	DNA[17,18], PtC[19]
	09	4-68*01	100%	5*01	CQQWSSNPLTF	5.49	DNA[9]
	03	1-135*01 (unproductive, out of frame)	100%	1*01	CWQGTHFP##TF	NR	DNA[9,10,13,17]

† Vκ genes highlighted in yellow appear in both Eμ-TCL1 and DTG mice; Vκ genes in bold red font were also identified in sequences analyzed by Yan *et al.*[16]. Pseudogenes are indicated by (ψ).

‡ CDR3 sequences highlighted in green appear in both Eμ-TCL1 and DTG mice and were also identified in sequences analyzed by Yan *et al.*[16] Frameshift mutations are indicated by (#).

¥ PtC, phosphatidylcholine; Br-treated RBCs, bromelain-treated red blood cells; Sm, Smith antigen;

Table S3. Top 50 differentially expressed genes between 12 week dnRAG1 and DTG mice

Gene symbol	DTG/dnRAG1	Gene description
Neto2	5.26	neuropilin (NRP) and tolloid (TLL)-like 2
Vash2	1.88	vasohibin 2
Tubb3	1.86	tubulin, beta 3
Crip1	1.69	cysteine-rich protein 1 (intestinal)
Pdcd1lg2	1.69	programmed cell death 1 ligand 2
Myadm	1.67	myeloid-associated differentiation marker
Fjx1	1.66	four jointed box 1 (Drosophila)
Tagln2	1.59	transgelin 2
Mid1	1.53	midline 1
Murc	1.48	muscle-related coiled-coil protein
Vim	1.48	vimentin
Olfr767	1.46	olfactory receptor 767
Tyrobp	1.45	TYRO protein tyrosine kinase binding protein
Cln3	1.44	ceroid lipofuscinosis, neuronal 3, juvenile (Batten, Spielmeyer-Vogt disease)
Lrrc8c	1.38	leucine rich repeat containing 8 family, member C
2010204K13Rik	1.34	RIKEN cDNA 2010204K13 gene
Nid1	1.34	nidogen 1
Ptk2	1.34	PTK2 protein tyrosine kinase 2
Gpr176	1.32	G protein-coupled receptor 176
Mad2l2	1.31	MAD2 mitotic arrest deficient-like 2 (yeast)
Rtn4rl1	1.31	reticulon 4 receptor-like 1
Ppp3ca	1.3	protein phosphatase 3, catalytic subunit, alpha isoform
Mirhg1	1.27	microRNA host gene 1 (non-protein coding)
Jup	1.26	junction plakoglobin
Vapb	1.23	vesicle-associated membrane protein, associated protein B and C
Slamf6	-1.33	SLAM family member 6
Cyb5b	-1.35	cytochrome b5 type B
Il2rb	-1.36	interleukin 2 receptor, beta chain
Trim7	-1.39	tripartite motif-containing 7
Hmgn3	-1.41	high mobility group nucleosomal binding domain 3
LOC635992	-1.44	similar to ubiquitin-conjugating enzyme E2 variant 2
Olfr99	-1.44	olfactory receptor 99
Rnu12	-1.45	RNA U12, small nuclear // RNA U12, small nuclear
Gm4383	-1.49	predicted gene 4383
4933404M02Rik	-1.52	RIKEN cDNA 4933404M02 gene
Pcp4	-1.6	Purkinje cell protein 4
Mpeg1	-1.63	macrophage expressed gene 1
Fcrl5	-1.7	Fc receptor-like 5
Gm7285	-1.71	predicted gene 7285
V165-D-J-C mu	-1.76	IgM variable region
Mctp2	-1.77	multiple C2 domains, transmembrane 2
Ipcef1	-1.78	interaction protein for cytohesin exchange factors 1
Clec2d	-1.83	C-type lectin domain family 2, member d
Gm9912	-1.89	predicted gene 9912
Zfp420	-1.92	zinc finger protein 420
Slc15a2	-2.06	solute carrier family 15 (H+/peptide transporter), member 2
EG665955	-2.25	predicted gene, EG665955
Gm5571	-2.27	predicted gene 5571 // predicted gene 5571
Igj	-2.35	immunoglobulin joining chain
Lmo7	-3.14	LIM domain only 7

Table S4. Top common differentially expressed genes in older transgenic mice relative to WT mice

		Fold change		
Gene Symbol	**dnRAG1/WT**	**EμTCL1/WT**	**DTG/WT**	**Gene Description**
1810046K07Rik	27.28	31.4	38.48	RIKEN cDNA 1810046K07 gene
Cyp11a1	11.31	14.22	26.34	cytochrome P450, family 11, subfamily a, polypeptide 1
Adm	36.41	25.5	23.38	adrenomedullin
Bmpr1a	10.04	17.82	19.52	bone morphogenetic protein receptor, type 1A
Pstpip2	25.95	53.52	18.61	proline-serine-threonine phosphatase-interacting protein 2
Bhlhe41	14.04	25.41	17.97	basic helix-loop-helix family, member e41
Ctla4	14.74	21.66	16.77	cytotoxic T-lymphocyte-associated protein 4
Vpreb3	-8.9	-13.56	-9.12	pre-B lymphocyte gene 3
Cr2	-12.25	-26.54	-10.4	complement receptor 2
Gm4955	-8.73	-12.58	-12.38	predicted gene 4955
Satb1	-9.08	-38.65	-13.07	special AT-rich sequence binding protein 1
Gpr171	-22.08	-24.98	-19.47	G protein-coupled receptor 171
Stat4	-25.5	-24.02	-21.45	signal transducer and activator of transcription 4
Lrrk2	-32.24	-37.39	-34.47	leucine-rich repeat kinase 2
Il10*	11.93	12.7	13.91	interleukin 10

*Il10 was not among the top 50 common differentially expressed genes in all pairwise comparisons.

Table S5. Top 50 differentially expressed genes between ill DTG mouse (M53) and Eμ-TCL1 mouse (M52)

Gene symbol*	DTG M53/ EμTCL1 M52 Fold change*	Gene description*
Prl2a1	**269.40**	**prolactin family 2, subfamily a, member 1**
Sox4	**42.75**	**SRY-box containing gene 4**
Prl8a2	**21.45**	**prolactin family 8, subfamily a, member 2**
Dpp4	18.78	dipeptidylpeptidase 4
Kcnj5	16.97	potassium inwardly-rectifying channel, subfamily J, member 5
1300014I06Rik	14.56	RIKEN cDNA 1300014I06 gene
Rpl39l	12.40	ribosomal protein L39-like
Rgs13	11.34	regulator of G-protein signaling 13
Xlr3b	11.27	X-linked lymphocyte-regulated 3B
Atp1b1	10.38	ATPase, Na+/K+ transporting, beta 1 polypeptide
Gm10561	10.21	predicted gene 10561
Tet1	10.18	tet oncogene 1
Scamp1	9.92	secretory carrier membrane protein 1
Gpr177	9.48	G protein-coupled receptor 177
Gm10384	9.15	predicted gene 10384
Wdfy3	9.00	WD repeat and FYVE domain containing 3
Atrnl1	8.60	attractin like 1
Dusp6	**8.36**	**dual specificity phosphatase 6**
Rgs18	8.26	regulator of G-protein signaling 18
1110032E23Rik	8.24	RIKEN cDNA 1110032E23 gene
Xlr3c	8.24	X-linked lymphocyte-regulated 3C
Xlr3a	7.85	X-linked lymphocyte-regulated 3A
F11r	7.68	F11 receptor
Tes	7.44	testis derived transcript
Ccdc125	7.25	coiled-coil domain containing 125
Zdhhc2	-8.43	zinc finger, DHHC domain containing 2
Cd68	-9.13	CD68 antigen
Cd3g	**-9.21**	**CD3 antigen, gamma polypeptide**
Gm410	-9.22	predicted gene 410
Kcnn4	-9.25	potassium intermediate/small conductance calcium-activated channel, subfamily N, member 4
Fcgrt	-10.94	Fc receptor, IgG, alpha chain transporter
Igk-V28	-12.16	immunoglobulin kappa chain variable 28 (V28)
Ighv1-72	-12.48	immunoglobulin heavy variable V1-72
Apobec2	-12.89	apolipoprotein B mRNA editing enzyme, catalytic polypeptide 2
Fxyd5	-13.16	FXYD domain-containing ion transport regulator 5
Prkar2b	**-13.75**	**protein kinase, cAMP dependent regulatory, type II beta**
Fbxw13	**-13.79**	**F-box and WD-40 domain protein 13**
Ccbp2	**-14.01**	**chemokine binding protein 2**
Emr1	**-14.46**	**EGF-like module containing, mucin-like, hormone receptor-like sequence 1**
Slamf9	-15.19	SLAM family member 9
Gm5486	-17.14	predicted gene 5486
Gstt1	**-17.57**	**glutathione S-transferase, theta 1**
Anxa2	-20.23	annexin A2
Fgl2	**-20.71**	**fibrinogen-like protein 2**
LOC100047053	-21.50	similar to monoclonal antibody kappa light chain
Pdlim1	-27.54	PDZ and LIM domain 1 (elfin)
LOC674190	**-30.31**	**similar to Ig heavy chain V region IR2 precursor**
Igh	**-33.42**	**immunoglobulin heavy chain complex**
LOC100046894	**-112.64**	**similar to Igk-C protein**
V165-D-J-C mu	**-317.51**	**IgM variable region**

* Items listed in bold are found in both Tables S5 and S6

Table S6. Top 50 differentially expressed genes between ill DTG mouse (F57) and Eμ-TCL1 mouse (M52)

Gene symbol	DTG F57/ EμTCL1 M52 Fold change	Gene description
Prl2a1	**200.97**	**prolactin family 2, subfamily a, member 1**
Xist	95.10	inactive X specific transcripts
Prl8a2	**66.42**	**prolactin family 8, subfamily a, member 2**
Sox4	**54.87**	**SRY-box containing gene 4**
Car2	53.97	carbonic anhydrase 2
Gpr126	39.55	G protein-coupled receptor 126
Pfn2	35.98	profilin 2
Cp	34.95	ceruloplasmin
Tnip3	32.26	TNFAIP3 interacting protein 3
Fcer2a	31.59	Fc receptor, IgE, low affinity II, alpha polypeptide
Cpm	30.20	carboxypeptidase M
Gm10879	29.01	predicted gene 10879
Anxa1	25.91	annexin A1
Dusp4	25.75	dual specificity phosphatase 4
Abp1	23.25	amiloride binding protein 1 (amine oxidase, copper-containing)
Id2	21.62	inhibitor of DNA binding 2
Rnf128	18.68	ring finger protein 128
Fabp7	18.40	fatty acid binding protein 7, brain
Dusp6	**17.93**	**dual specificity phosphatase 6**
Slc22a21	15.49	solute carrier family 22 (organic cation transporter), member 21
Akt3	14.84	thymoma viral proto-oncogene 3
Scin	14.54	scinderin
Azgp1	13.03	alpha-2-glycoprotein 1, zinc
Igh-6	12.75	immunoglobulin heavy chain 6 (heavy chain of IgM)
Parp8	12.59	poly (ADP-ribose) polymerase family, member 8
Cln8	-6.79	ceroid-lipofuscinosis, neuronal 8
Neurod4	-7.26	neurogenic differentiation 4
Lrrc25	-7.52	leucine rich repeat containing 25
Gp49a	-7.74	glycoprotein 49 A
Tmprss13	-7.96	transmembrane protease, serine 13
Prkar2b	**-8.03**	**protein kinase, cAMP dependent regulatory, type II beta**
Emr1	**-8.05**	**EGF-like module containing, mucin-like, hormone receptor-like sequence 1**
Kdm5d	-8.51	lysine (K)-specific demethylase 5D
Sell	-9.15	selectin, lymphocyte
Cd9	-9.97	CD9 antigen
Fgl2	**-10.09**	**fibrinogen-like protein 2**
Cd3g	**-10.35**	**CD3 antigen, gamma polypeptide**
Uty	-10.71	ubiquitously transcribed tetratricopeptide repeat gene, Y chromosome
Fbxw13	**-12.03**	**F-box and WD-40 domain protein 13**
Ccbp2	**-13.49**	**chemokine binding protein 2**
Gstt1	**-13.80**	**glutathione S-transferase, theta 1**
Hpse	-13.87	heparanase
Lgals1	-14.20	lectin, galactose binding, soluble 1
V165-D-J-C mu	**-14.31**	**IgM variable region**
Igh	**-16.44**	**immunoglobulin heavy chain complex**
Cd36	-16.52	CD36 antigen
Ddx3y	-24.47	DEAD (Asp-Glu-Ala-Asp) box polypeptide 3, Y-linked
Eif2s3y	-24.72	eukaryotic translation initiation factor 2, subunit 3, structural gene Y-linked
LOC674190	**-37.70**	**similar to Ig heavy chain V region IR2 precursor**
LOC100046894	**-48.43**	**similar to Igk-C protein**

* Items listed in bold are found in both Tables S5 and S6

SUPPLEMENTAL REFERENCES

1. Hassaballa AE, Palmer VL, Anderson DK, et al. Accumulation of B1-like B cells in transgenic mice over-expressing catalytically inactive RAG1 in the periphery. *Immunology*. 2011;134(4):469-486.
2. Bichi R, Shinton SA, Martin ES, et al. Human chronic lymphocytic leukemia modeled in mouse by targeted TCL1 expression. *Proc Natl Acad Sci U S A*. 2002;99(10):6955-6960.
3. Fusby JS, Kassmeier MD, Palmer VL, et al. Cigarette smoke-induced effects on bone marrow B-cell subsets and CD4(+):CD8(+) T-cell ratios are reversed by smoking cessation: Influence of bone mass on immune cell response to and recovery from smoke exposure. *Inhal Toxicol*. 2010; 22(9): 785-796.
4. Wang Z, Raifu M, Howard M, et al. Universal PCR amplification of mouse immunoglobulin gene variable regions: the design of degenerate primers and an assessment of the effect of DNA polymerase 3' to 5' exonuclease activity. *J Immunol Methods*. 2000;233(1-2):167-177.
5. Li C, Hung Wong W. Model-based analysis of oligonucleotide arrays: model validation, design issues and standard error application. *Genome Biol*. 2001;2(8):RESEARCH0032.
6. Heng TS, Painter MW. The Immunological Genome Project: networks of gene expression in immune cells. *Nat Immunol*. 2008;9(10):1091-1094.
7. Johnson WE, Li C, Rabinovic A. Adjusting batch effects in microarray expression data using empirical Bayes methods. *Biostatistics*. 2007;8(1):118-127.
8. Bloom DD, Davignon JL, Retter MW, et al. V region gene analysis of anti-Sm hybridomas from MRL/Mp-lpr/lpr mice. *J Immunol*. 1993;150(4):1591-1610.
9. Krishnan MR, Jou NT, Marion TN. Correlation between the amino acid position of arginine in VH-CDR3 and specificity for native DNA among autoimmune antibodies. *J Immunol*. 1996;157(6):2430-2439.
10. Kofler R, Strohal R, Balderas RS, et al. Immunoglobulin kappa light chain variable region gene complex organization and immunoglobulin genes encoding anti-DNA autoantibodies in lupus mice. *J Clin Invest*. 1988;82(3):852-860.
11. Lehuen A, Bartels J, Kearney JF. Characterization, specificity, and IgV gene usage of anti-lymphocyte monoclonal antibodies from perinatal mice. *Int Immunol*. 1992;4(10):1073-1084.
12. Kitagawa Y, Okuhara E. The separation of three antibody populations from anti-poly(A).poly(U) antibodies elicited in mice or rabbits and antigenic features of poly(A).poly(U)). *Mol Immunol*. 1982;19(2):257-266.
13. Tillman DM, Jou NT, Hill RJ, Marion TN. Both IgM and IgG anti-DNA antibodies are the products of clonally selective B cell stimulation in (NZB x NZW)F1 mice. *J Exp Med*. 1992;176(3):761-779.
14. Mueller CM, Minnerath JM, Jemmerson R. B lymphocyte recognition of the self antigen mouse cytochrome C in different mouse strains: targeting of the same dominant epitope by naturally-occurring cells expressing distinct VH genes. *Mol Immunol*. 1997;34(12-13):843-853.
15. Monestier M, Kandiah DA, Kouts S, et al. Monoclonal antibodies from NZW x BXSB F1 mice to beta2 glycoprotein I and cardiolipin. Species specificity and charge-dependent binding. *J Immunol*. 1996;156(7):2631-2641.
16. Yan XJ, Albesiano E, Zanesi N, et al. B cell receptors in TCL1 transgenic mice resemble those of aggressive, treatment-resistant human chronic lymphocytic leukemia. *Proc Natl Acad Sci U S A*. 2006;103(31):11713-11718.
17. Wloch MK, Alexander AL, Pippen AM, Pisetsky DS, Gilkeson GS. Differences in V kappa gene utilization and VH CDR3 sequence among anti-DNA from C3H-lpr mice and lupus mice with nephritis. *Eur J Immunol*. 1996;26(9):2225-2233.
18. Ibrahim SM, Weigert M, Basu C, Erikson J, Radic MZ. Light chain contribution to specificity in anti-DNA antibodies. *J Immunol*. 1995;155(6):3223-3233.
19. Pennell CA, Mercolino TJ, Grdina TA, Arnold LW, Haughton G, Clarke SH. Biased immunoglobulin variable region gene expression by Ly-1 B cells due to clonal selection. *Eur J Immunol*. 1989;19(7):1289-1295.
20. Jethwa HS, Clarke SH, Itoh-Lindstrom Y, Falk RJ, Jennette JC, Nachman PH. Restriction in V kappa gene use and antigen selection in anti-myeloperoxidase response in mice. *J Immunol*. 2000;165(7):3890-3897.
21. Guo WX, Burger AM, Fischer RT, Sieckmann DG, Longo DL, Kenny JJ. Sequence changes at the V-D junction of the VH1 heavy chain of anti-phosphocholine antibodies alter binding to and protection against Streptococcus pneumoniae. *Int Immunol*. 1997;9(5):665-677.
22. Reininger L, Ollier P, Poncet P, Kaushik A, Jaton JC. Novel V genes encode virtually identical variable regions of six murine monoclonal anti-bromelain-treated red blood cell autoantibodies. *J Immunol*. 1987;138(1):316-323.