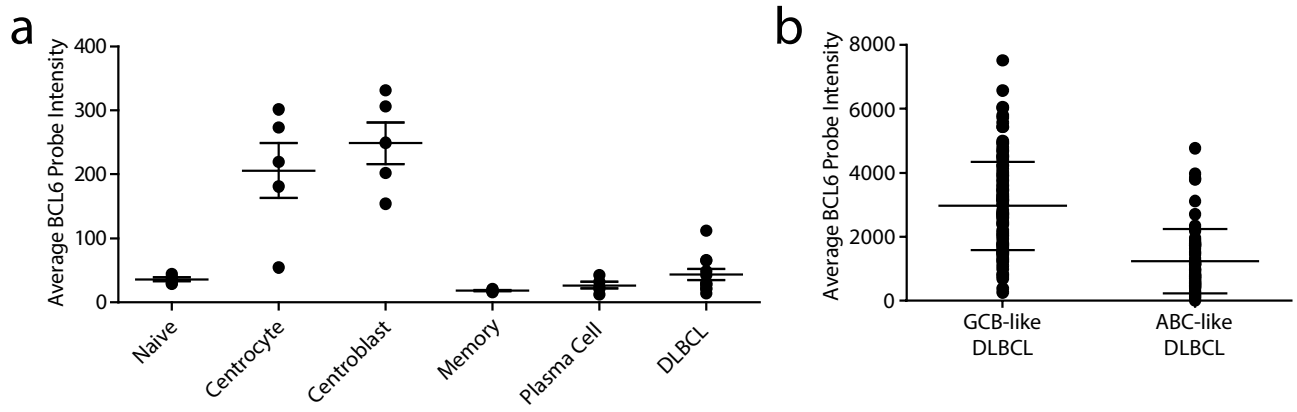
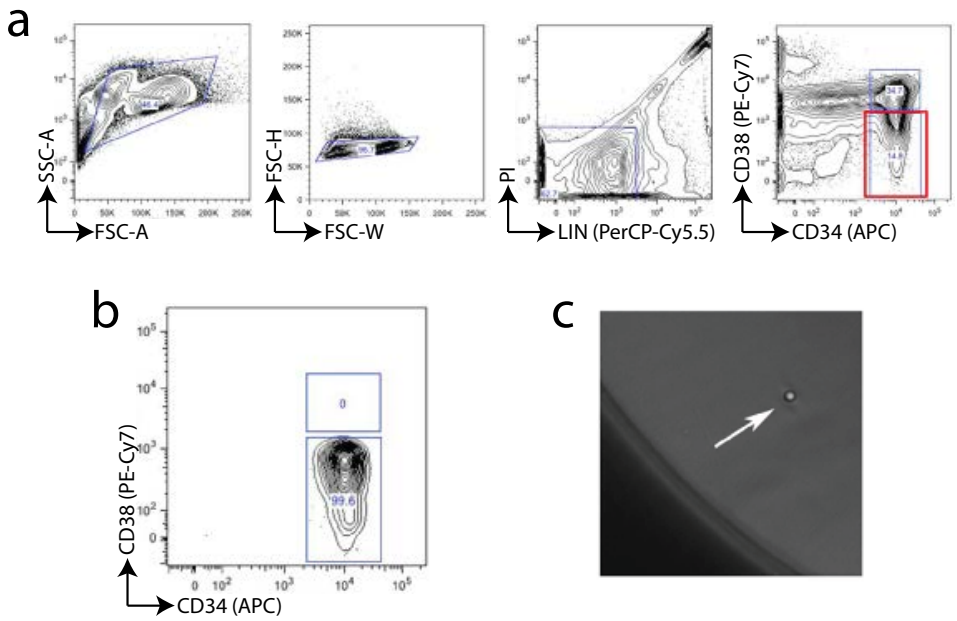
Supplementary Information for Green and Vicente-Dueñas *et al.*

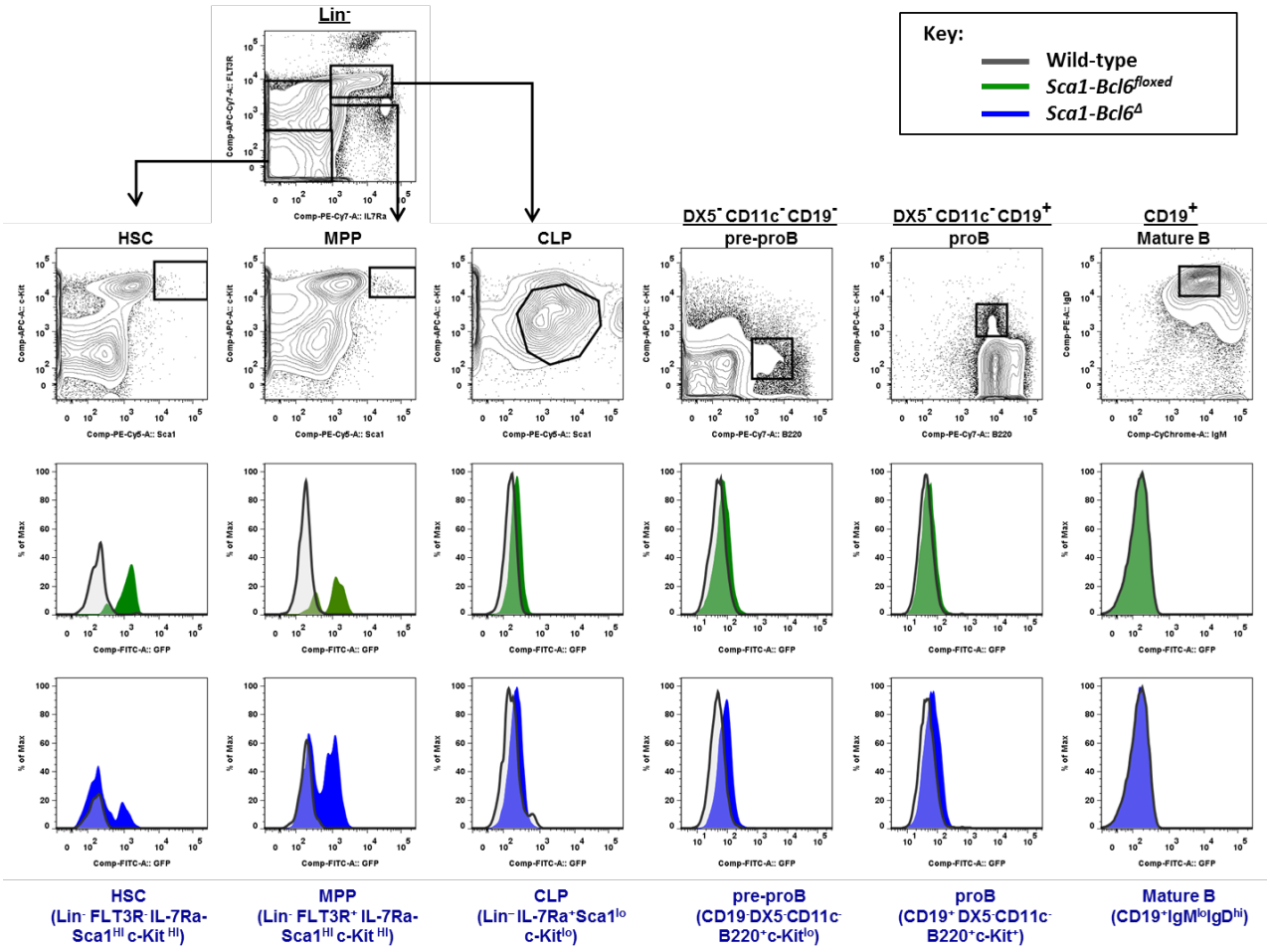
Contents:



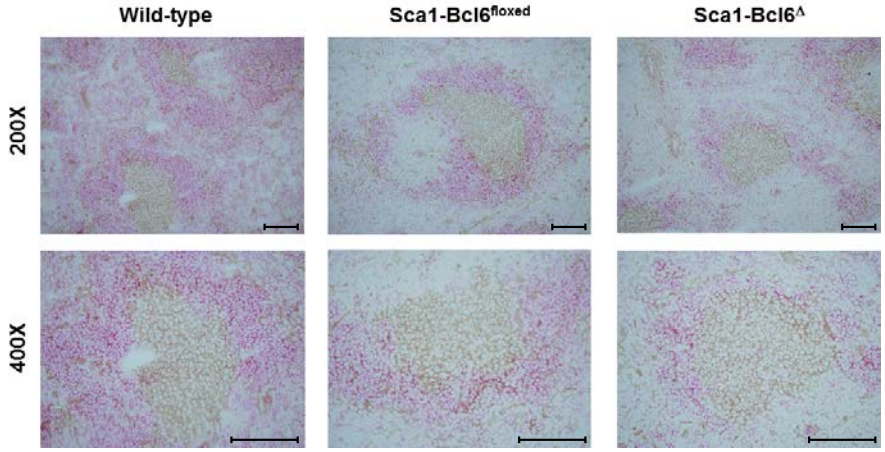
Supplementary Fig. 1: Expression of BCL6 in normal B-cell subsets compared to DLBCL tumors. a) Using publicly available gene expression microarray data, we observed BCL6 transcript abundance to be elevated in normal centrocytes and centroblasts. However, BCL6 transcript abundance was comparably low in DLBCL tumor specimens. This suggests that the lack of difference in *BCL6* transcript abundance between cases with *BCL6* DNA copy number gain or translocation compared to those without is not the result of uniformly high expression of BCL6 in all DLBCL tumors via alternate mechanisms of BCL6 up-regulation in tumors without genetic perturbation of the *BCL6* gene. Error bars represent the mean +/- the standard deviation. **b)** *BCL6* transcript abundance is shown for DLBCL tumors classed as GCB-like or ABC-like using the Wright algorithm, as shown in Figure 1. Transcript abundance is significantly higher in the GCB-like subtype compared to the ABC-like subtype (T-test P-value < 0.001) despite *BCL6* DNA copy number gains and translocations being more prevalent in the ABC-like subtype. Note that probe intensities are not comparable between panels *a* and *b* because of a significant batch effect between these experiments. Error bars represent the mean +/- the standard deviation.



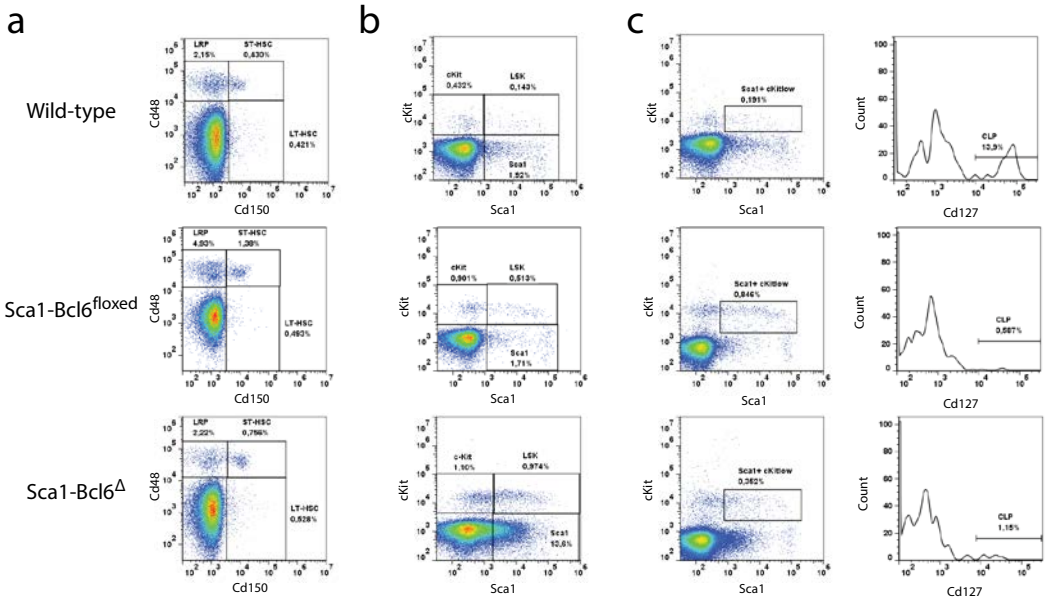
Supplementary Fig. 2: Single cell sorting of human bone marrow hematopoietic stem/progenitor cells (HSPCs). a) Human HS/PCs were gated (from left to right) as intact cells, singlets, lineage-marker-negative, and $CD34^{+}$ $CD38^{-}$. **b)** Bulk sorting and post-sort analysis showed high fidelity of the sort. **c)** Bright-field microscopy shows an example of a single cell deposited in a well of a terisaki plate.



Supplementary Fig. 3: GFP expression from the Bcl6-IRES-GFP cassette in the hematopoietic stem/progenitor (HSPC) compartment of *Sca1-Bcl6*^{floxed} and *Sca1-Bcl6*$^{\Delta}$ transgenic mice. FACS analysis was performed to detect the expression of GFP in hematopoietic stem cells (HSC), multipotential progenitors (MPP), common lymphoid progenitors (CLP) and in the earliest stages of B cell development in the BM (pre-proB and proB cells), defined as described in the methods section. The histograms depict the GFP expression levels in the corresponding populations in control (red curve) versus transgenic (green and blue curves) mice. GFP expression is detectable in the earliest developmental stages, according to their Sca1^{HIGH} expression levels, and starts to decline already in the CLP stage. GFP expression is barely detectable in pre-proB cells and is gone at the proB cell stage. No GFP expression can be detected during the rest of B cell development (splenic mature B cells are shown as an example) or in other hematopoietic lineages (data not shown). The data are representative FACS plots from one out of 10 mice analyzed.



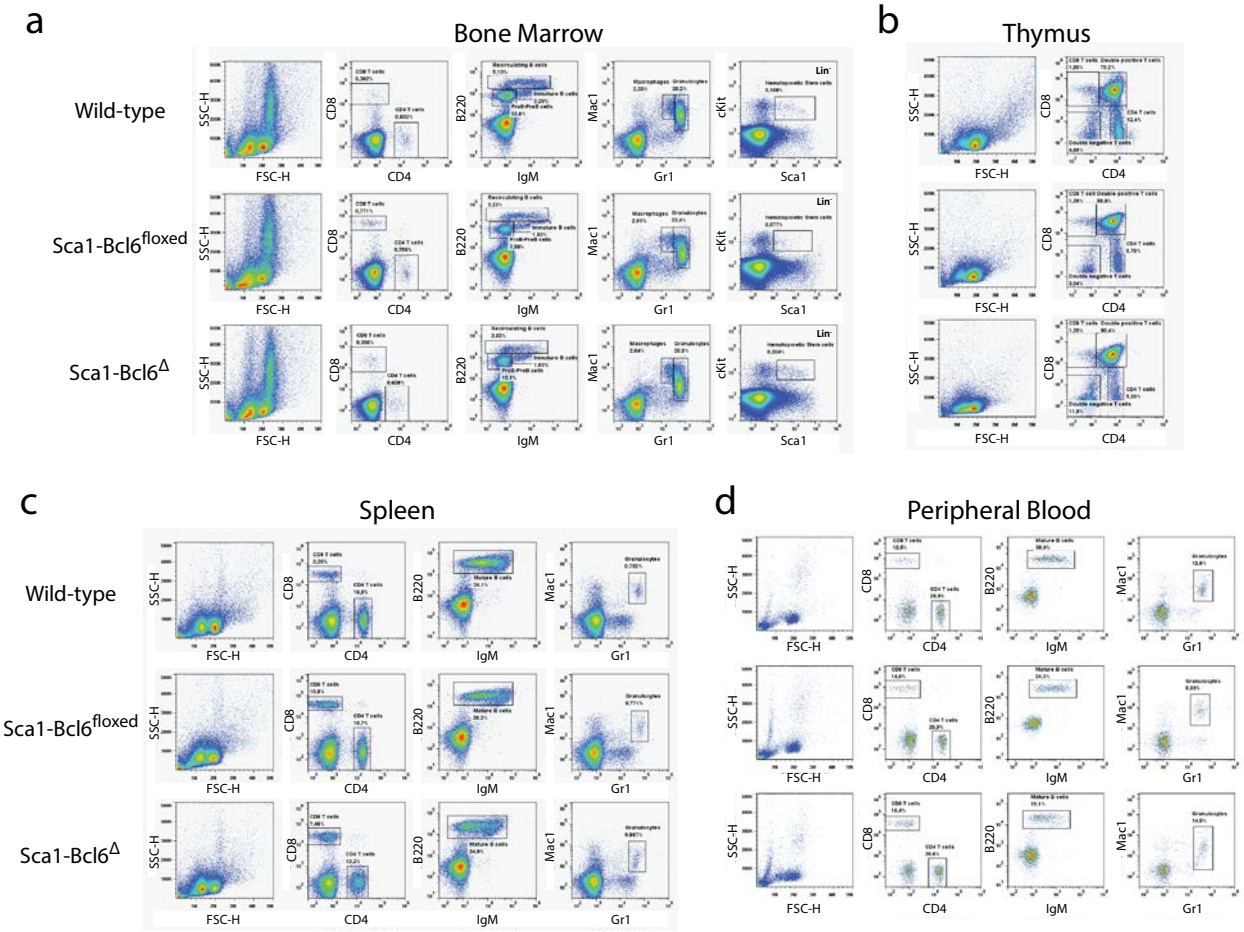
Supplementary Fig. 4: Formation of germinal centres in *Sca1-Bcl6^{floxed}* and *Sca1-Bcl6^{Δ}* (= *Sca1-Bcl6^{floxed}, mb1-Cre*) mice. Cryosections of the spleen from mice of the indicated genotypes ten days after immunization with sheep red blood cells were stained with fluorescein isothiocyanate (FITC)–anti-IgD antibody (pink) and biotinylated peanut agglutinin (PNA, brown) followed by detection with an alkaline phosphatase-coupled anti-FITC antibody (visualized with Fast Red) and horseradish peroxidase-conjugated streptavidin (visualized with diaminobenzidine), respectively, and are representative of ≥3 replicate experiments. Scale bars represent 100μm.



d

		Mean ± SEM		
Subset	**Markers**	**Wild-Type (n=8)**	**Sca1-Bcl6^{floxed} (n=4)**	**Sca1-Bcl6$^{\Delta}$ (n=4)**
Long-term HSC (a)	Lin$^-$CD48$^-$CD150$^+$	0.2534 ± 0.04413	0.2950 ± 0.07461	0.4058 ± 0.04824
Short-term HSC (a)	Lin$^-$CD48$^+$CD150$^+$	0.2800 ± 0.06309	1.075 ± 0.2349	0.7703 ± 0.1739
Lineage-restricted progenitor (a)	Lin$^-$CD48$^+$CD150$^-$	3.785 ± 0.4606	4.783 ± 0.1042	2.228 ± 0.1983
LSK (b)	Lin$^-$Sca1$^+$cKit$^+$	0.1596 ± 0.02816	0.5580 ± 0.1761	0.6803 ± 0.2111
Common lymphoid progenitor (c)	Lin$^-$Sca1$^+$cKit$^+$Cd127$^+$	8.740 ± 1.808	0.4947 ± 0.2586	1.540 ± 0.1778

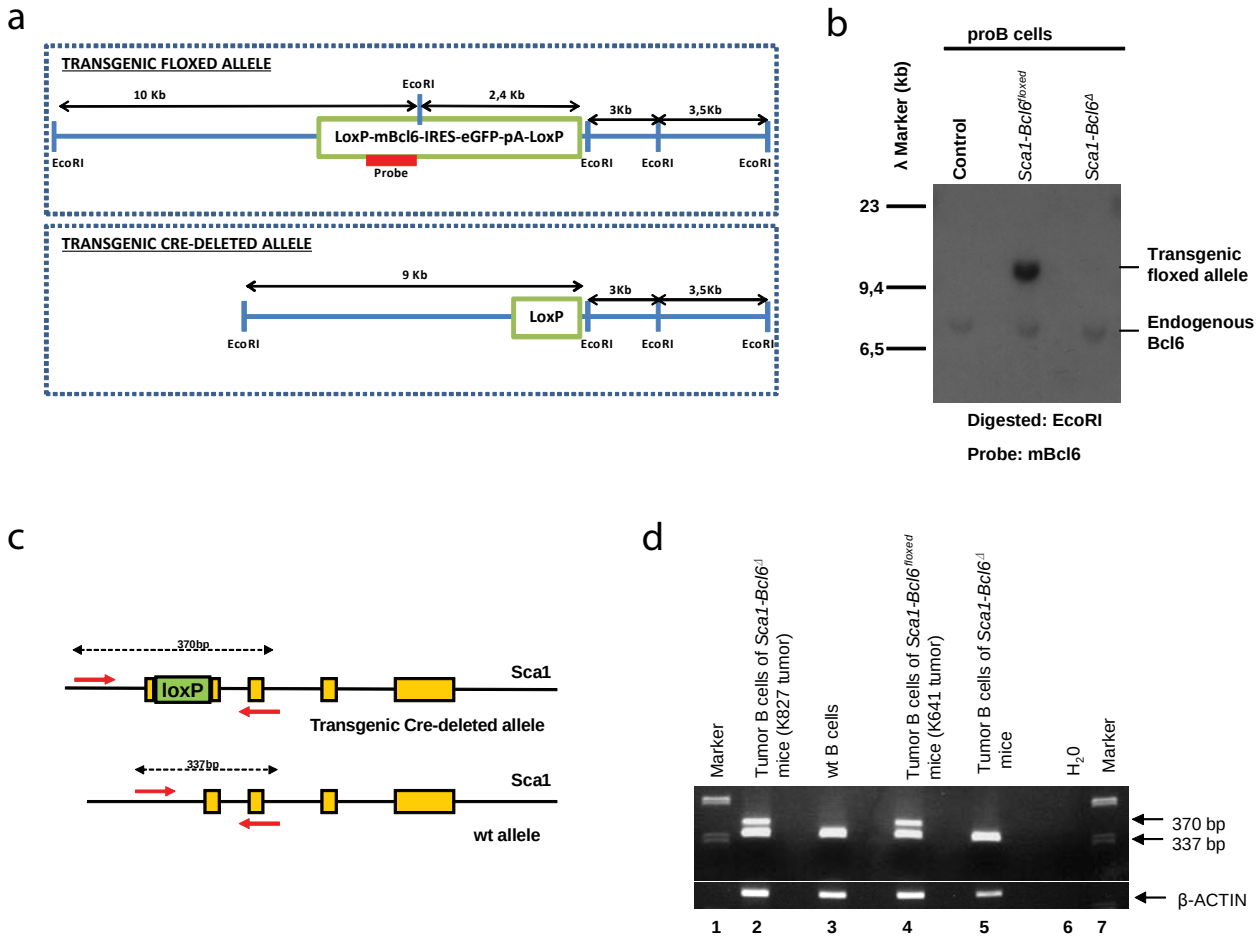
Supplementary Fig. 5: Flow cytometric analysis of hematopoietic precursor subsets in 4 week old tumor-free mice. a) Increases can be seen in long-term (LT-HSC) and short term hematopoietic stem cells (ST-HSC), and lineage-restricted progenitor (LRP) cells within *Sca1-Bcl6^{floxed}* and *Sca1-Bcl6$^{\Delta}$* (= *Sca1-Bcl6^{floxed}, mb1-Cre*) mice compared to wild-type mice. **b)** *Sca1-Bcl6^{floxed}* and *Sca1-Bcl6$^{\Delta}$* mice show increased numbers of Lin$^-$Sca1$^+$cKit$^+$ (LSK) cells compared to wild-type mice. **c)** Decreased numbers of common lymphoid progenitors (CLP) are found in both *Sca1-Bcl6^{floxed}* and *Sca1-Bcl6$^{\Delta}$* mice in comparison to wild-type mice. **d)** A summary of the gate statistics for the data in parts a-c. All data have been gated on negative expression of lineage markers, and are representative of at least 3 replicate experiments.



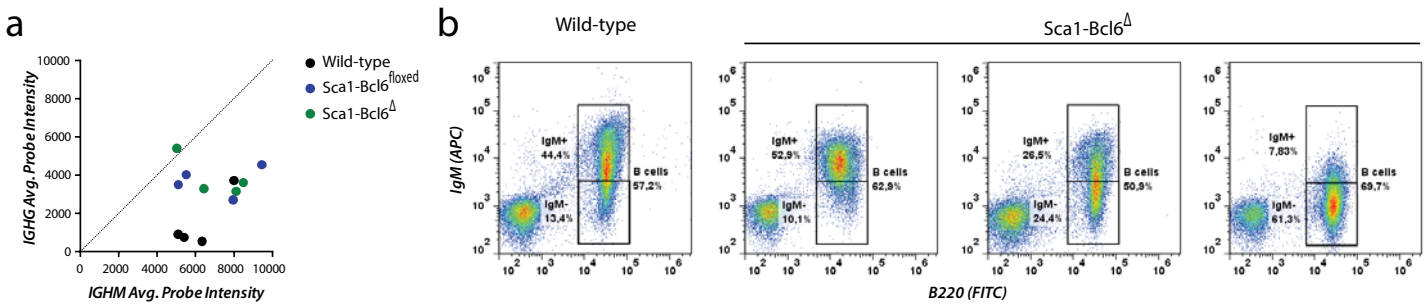
e

			Percentage		
Compartment	**Subset**	**Markers**	**Wild-Type**	**Sca1-Bcl6^{floxed}**	**Sca1-Bcl6Δ**
Bone Marrow (a)	Recirculating B cells	IgM+B220lo	7.91	5.24	3.83
	Immature B cells	IgM+B220hi	11	8.82	6.45
	Pre-B and Pro-B cells	IgM-B220lo	13.8	7.58	10.1
	CD8 T cells	CD8+	0.302	0.771	0.396
	CD4 T cells	CD4+	0.803	0.756	0.608
	Granulocytes	Gr1+Mac1+	20.2	23.4	25.9
	Hematopoietic stem cells	Sca1+ cKit- Lin-	0.108	0.077	0.204
Thymus (b)	Double positive T cells	CD4+ CD8+	79.2	88	80.4
	CD8 T cells	CD8+	1.8	1.39	1.4
	CD4 T cells	CD4+	12.4	5.78	5.29
	Double negative T cells	CD4- CD8-	4.68	3.94	11.9
Spleen (c)	CD8 T cells	CD8+	3.35	10.8	7.48
	CD4 T cells	CD4+	18.9	18.7	13.3
	Mature B cells	IgM+/-B220+	34.1	26.3	24.9
	Granulocytes	Gr1+Mac1+	0.792	0.771	0.887
Peripheral blood (d)	CD8 T cells	CD8+	12.5	14.6	16.4
	CD4 T cells	CD4+	28.9	25.9	30.6
	Mature B cells	IgM+/-B220+	30	24.3	19.1
	Granulocytes	Gr1+Mac1+	12.6	8.59	14.5

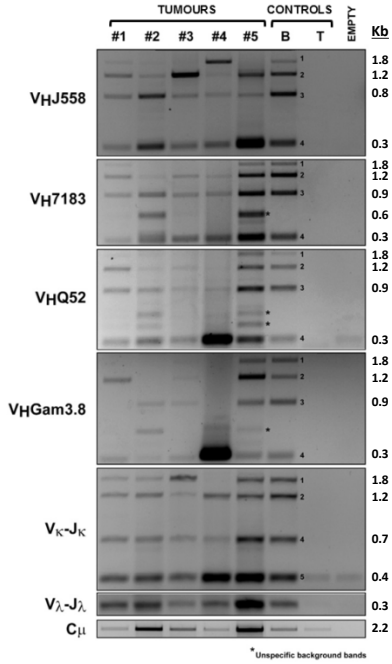
Supplementary Fig. 6: Flow cytometric analysis of hematopoietic subsets in young tumor-free mice. Representative plots of cell subsets from the bone marrow **(a)**, thymus **(b)**, spleen **(c)** and peripheral blood **(d)** are shown and summarized **(e)**. These show altered numbers of lymphocyte subsets in the main compartments of *Sca1-Bcl6^{floxed}* and *Sca1-Bcl6$^{\Delta}$* (= *Sca1-Bcl6^{floxed}, mb1-Cre*) mice compared to wild-type mice.



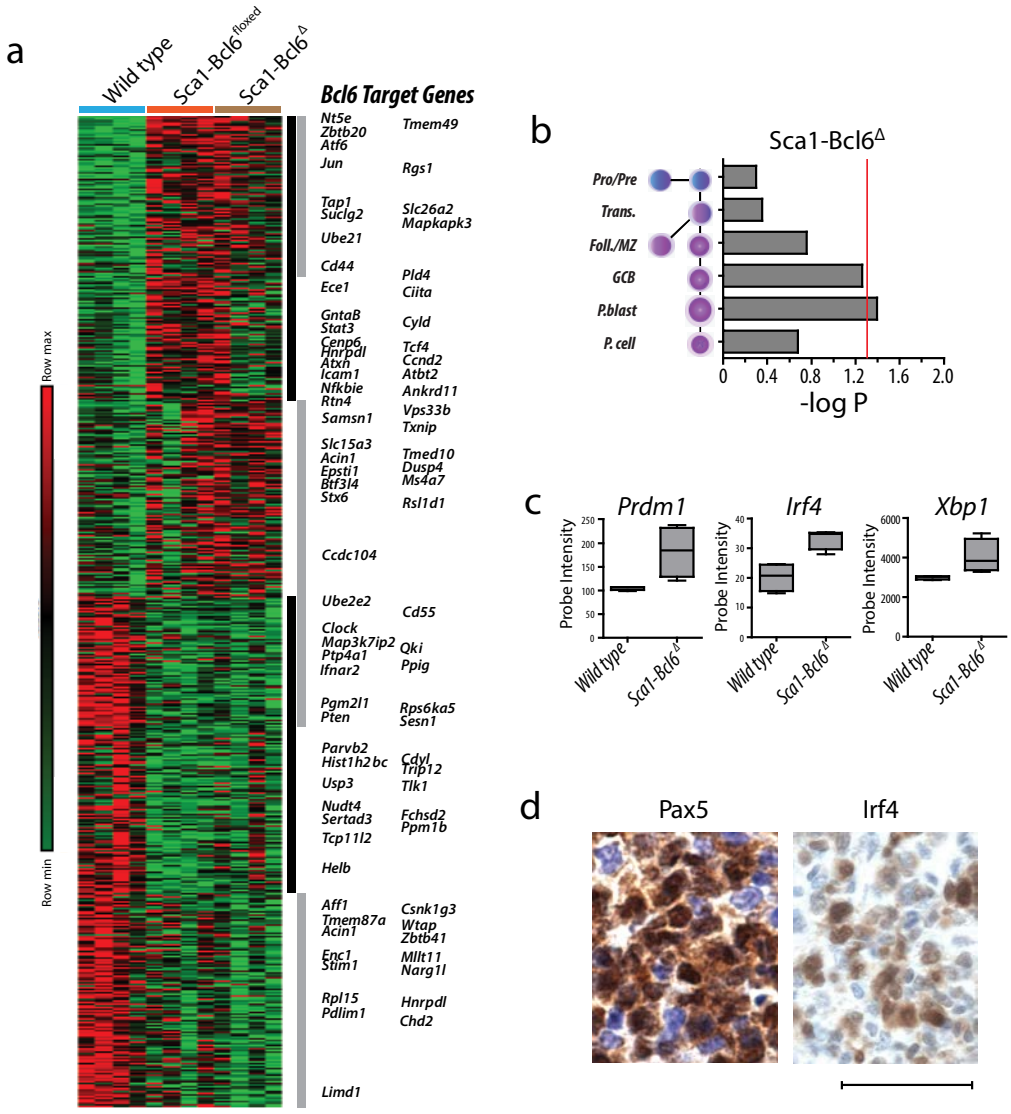
Supplementary Fig. 7: Efficient deletion of the transgenic Bcl6 allele in Pro-B cells. a) Diagrammatic representation of the EcoRI restriction sites within the transgenic floxed and cre-deleted alleles. **b)** Southern blot analysis of EcoRI digested DNA from pro-B cells of wild-type (control), *Sca1-Bcl6*^{floxed} and *Sca1-Bcl6*$^{\Delta}$ (= *Sca1-Bcl6*^{floxed}*, mb1-Cre*) mice, using a probe specific for murine *Bcl6*, shows efficient deletion of the floxed allele within pro-B cells of *Sca1-Bcl6*$^{\Delta}$ mice. B220^{+}ckit^{+} pro-B cells were obtained from 1 week culture in the presence of IL7 and ST2 feeder cells. **c)** Diagrammatic representation of PCR primer locations (red arrow) with reference to the transgenic cre-deleted allele and wild-type *Sca1* allele, showing a larger product that is generated from the cre-deleted allele. **d)** PCR analysis shows the presence of the larger bands in two representative *Sca1-Bcl6*$^{\Delta}$ mice (lanes 2 and 4), confirming cre-mediated deletion of the transgenic *Bcl6* allele within the tumor cells of *Sca1-Bcl6*$^{\Delta}$ mice.



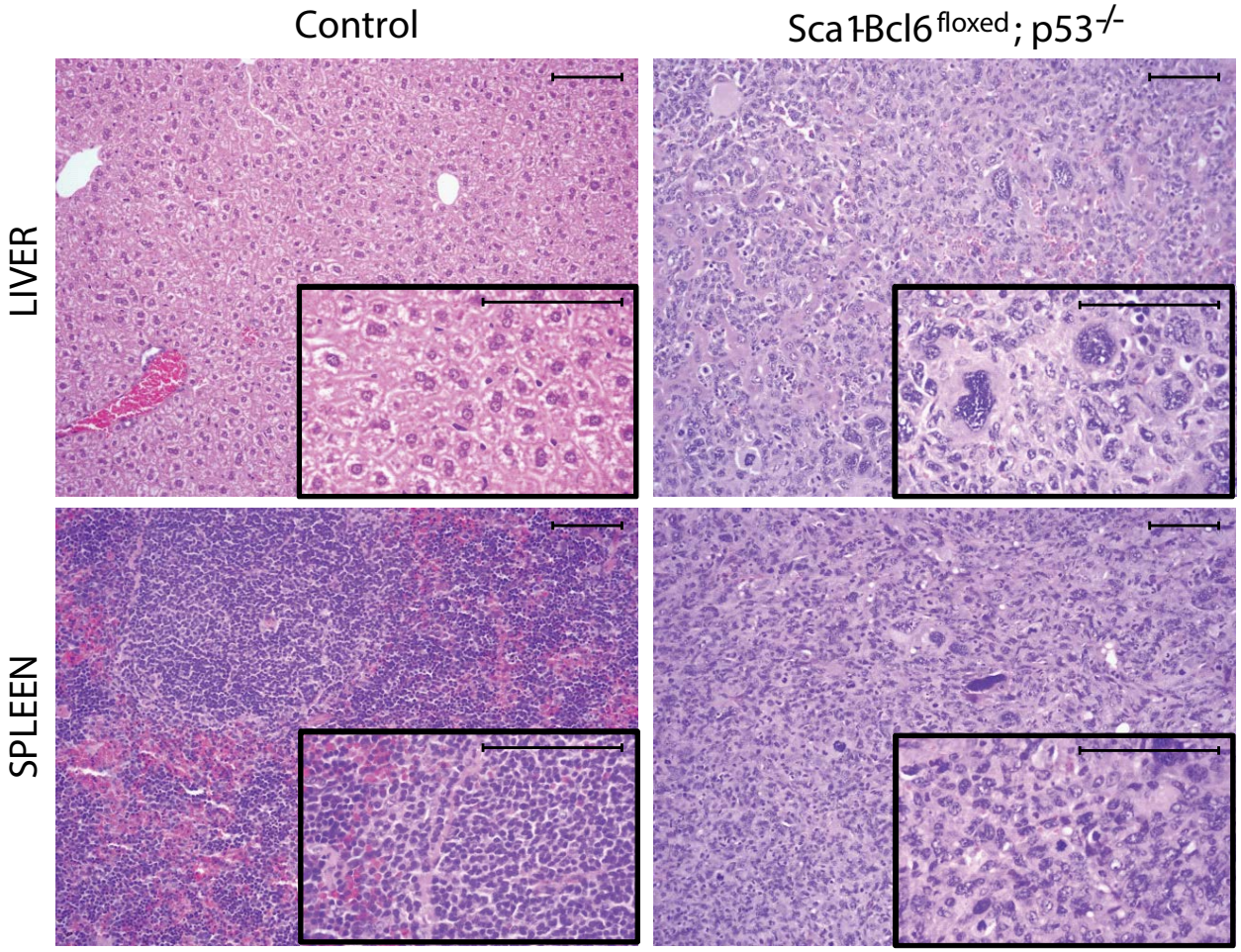
Supplementary Fig. 8: Immunoglobulin heavy-chain isotype expression on murine tumors. a) Average probe intensity values for IGHM (μ heavy-chain) and IGHG (γ heavy-chain) are shown for wild-type spleens and murine tumors. All show expression of IGHM, but there is some expression of IGHG within murine tumors. b) Flow cytometric analysis of cell surface IgM expression are shown for B-cells from a representative wild-type spleen and three representative *Sca1-Bcl6*$^{\Delta}$ tumors. Variable expression of IgM can be seen on *Sca1-Bcl6*$^{\Delta}$ tumors indicating that a subset of cells/tumors undergo class switch recombination.



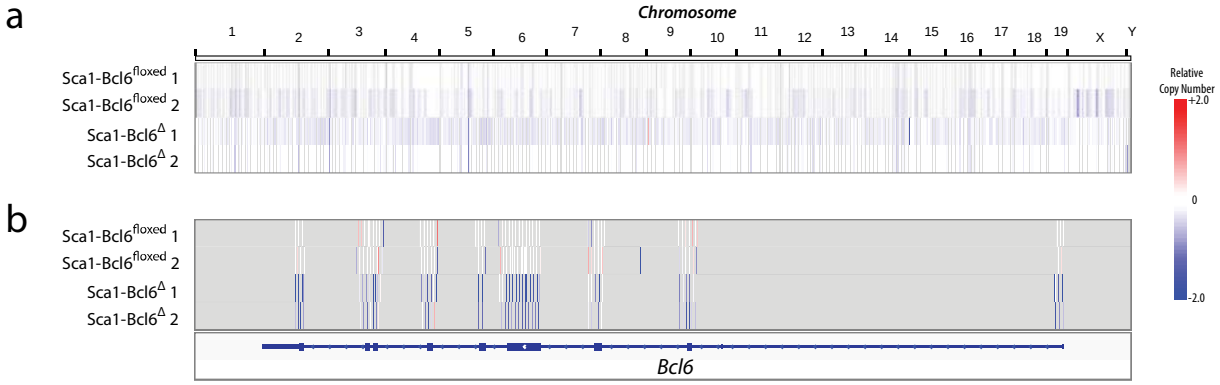
Supplementary Fig. 9: Increased immunoglobulin clonality in tumors from *Sca1-Bcl6$^{\Delta}$* mice. PCR analysis of immunoglobulin heavy- and light-chain gene rearrangements in cells from tumor bearing spleens of *Sca1-Bcl6$^{\Delta}$* (= *Sca1-Bcl6^{floxed}, mb1-Cre*) mice. Genomic DNA was extracted from the splenic tumors of 5 *Sca1-Bcl6$^{\Delta}$* mice. Thymocytes (T) were included as a negative control, and sorted CD19^{+} B-cells from the spleens of healthy mice were included as a control for polyclonal rearrangements (indicated by numbers, 1-5) within the mature B-cell population. PCRs were performed as described in Supplementary methods section, showing the presence of all the possible rearranged bands both for different distal and proximal VDJ heavy chain and Vκ light chain rearrangements in the transgenic B cells. It can be seen that tumors show increased clonality within their immunoglobulin repertoire. Locations of molecular weight markers are indicated on the right by their size in kilobases (Kb). *Non-specific background band.



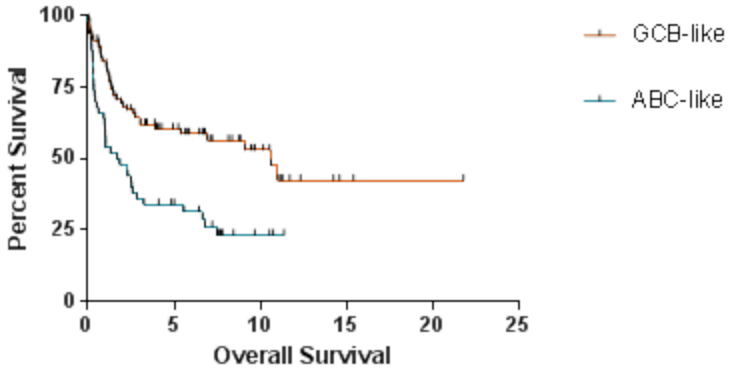
Supplementary Fig. 10: Similarity between tumors from Sca1-Bcl6floxed and Sca1-Bcl6Δ mice include transcriptional changes and expression of markers indicative of a post-germinal center stage of differentiation. a) Genes significantly increased or decreased in expression in the Sca1-Bcl6floxed (Black bar) and/or the Sca1-Bcl6Δ (= Sca1-Bcl6floxed, mb1-Cre) tumors (Grey bar) compared to spleens from wild-type mice are shown. Genes that are differentially expressed between each tumor and control spleens show significant overlap by hypergeometric probability ($p<0.001$). Bcl6 target genes are shown, and demonstrate that there is no significant skew in repression or induction of these genes. b) In line with our observation in Sca1-Bcl6floxed tumors, differentially expressed genes in Sca1-Bcl6Δ tumors showed a significant enrichment of the normal plasmablast signature (One-tailed P=0.041, FDR=0.165). c) Sca1-Bcl6Δ tumors showed increased expression of transcription factors indicative of a post-germinal center stage of differentiation. d) Immunohistochemical staining showed expression of markers indicative of B-cell identity (Pax5) and a post-germinal center stage of differentiation (Irf4). Scale bar represents 50µm.



Supplementary Fig. 11: Myeloid neoplasia in mice lacking p53. *Sca1-Bcl6*^{floxed} were crossed with *p53*$^{-/+}$ or *p53*$^{-/-}$ mice and the resulting strain monitored for tumor onset. Neither strain developed any B-cell lymphoma, however 25% (1/4) of *Sca1-Bcl6*^{floxed}*p53*$^{-/+}$ and 66.7% (2/3) of *Sca1-Bcl6*^{floxed}*p53*$^{-/-}$ mice developed Myeloid malignancies. Tumors were observed at 12.79 (±3.45) months and 5.55 (±0.44) months, respectively. Images are 20X, inset 40X, scale bar represents 100µm.



Supplementary Fig. 12: DNA copy number profiles of tumors from *Sca1-Bcl6^{floxed}* and *Sca1-Bcl6$^{\Delta}$* mice. a) A genome-wide view of relative DNA copy number in tumors from *Sca1-Bcl6^{floxed}* (n=2) and *Sca1-Bcl6$^{\Delta}$* (n=2) mice compared to age and gender-matched wild-type mice, showing absence of overt DNA copy number alterations. **b)** Absence of DNA copy number abnormalities was not the result of lack of sensitivity, as relative decrease of *Bcl6* in *Sca1-Bcl6$^{\Delta}$* could be readily detected. X-chromosome dosage could also be detected when comparing male and female mice (data not shown).



Supplementary Fig. 13: Cell of origin (COO) classification of DLBCL tumors. Classification of 249 tumors with available gene expression microarray data using the Wright 140 gene algorithm significantly stratified survival (Log-rank P-value < 0.001), with the ABC-like subtype having significantly worse overall survival as expected.

Supplementary Table 1: GISTIC peaks and association with overall survival (OS)

Peak Information				CHOP-treated (n=232)			R-CHOP-treated (n=196)		
Type	**Cytoband**	**Physical Position (hg18)**	**GISTIC Q-value**	**Frequency**	**OS Hazard**	**OS P**	**Frequency**	**OS Hazard**	**OS P**
Gain	1q23.2	chr1:158544936-158665465	0.003	24%	1.0099	0.9619	15%	0.934	0.865
Gain	1q32.1	chr1:199873896-200155294	0.003	28%	1.0063	0.9749	17%	1.685	0.116
Gain	2p15	chr2:58377658-61840812	0.032	14%	1.1602	0.5329	**13%**	**0.132**	**0.017**
Gain	3q27.2	chr3:187302125-187351166	0.054	**28%**	***1.872***	***0.001***	**12%**	***2.815***	***0.001***
Gain	6p21.1	chr6:41569374-41904794	0.053	20%	1.024	0.913	12%	0.431	0.144
Gain	7q21.3	chr7:93180209-93439104	0.046	31%	0.756	0.162	21%	0.805	0.536
Gain	11q23.2	chr11:112827130-112869367	<0.001	19%	1.072	0.754	13%	0.781	0.568
Gain	11q24.3	chr11:127889552-128156428	0.008	19%	0.985	0.945	15%	1.048	0.903
Gain	12q15	chr12:68632194-69005611	0.001	**27%**	***0.562***	***0.009***	16%	1.100	0.803
Gain	13q31.3	chr13:91650973-91679065	0.046	21%	1.176	0.449	16%	1.143	0.714
Gain	18q21.31	chr18:52775841-52777772	<0.001	**30%**	***1.591***	***0.011***	**13%**	**2.308**	**0.008**
Gain	19q13.41	chr19:57791450-58418753	0.045	**22%**	**1.648**	**0.012**	**7%**	***4.651***	***<0.001***
Loss	1p36.32	chr1:1-6689322	<0.001	21%	1.019	0.928	5%	1.924	0.199
Loss	1p35.2	chr1:30953302-32638082	0.015	8%	1.362	0.289	2%	0.941	0.952
Loss	1p13.3	chr1:110301699-111836039	0.056	6%	1.134	0.715	**2%**	**3.619**	**0.020**
Loss	2q22.3	chr2:144839311-145003028	0.002	8%	1.727	0.058	2%	3.539	0.061
Loss	3p26.1	chr3:6863179-7774863	0.018	4%	1.387	0.431	2%	0.823	0.847
Loss	3p21.1	chr3:49853083-53144064	0.001	4%	1.011	0.980	2%	1.086	0.935
Loss	4q21.22	chr4:83408656-84749491	0.006	7%	1.082	0.819	2%	2.343	0.223
Loss	6q21	chr6:107309779-107677173	<0.001	25%	0.818	0.346	12%	1.314	0.498
Loss	6q23.3	chr6:138174210-139048098	<0.001	22%	0.882	0.561	10%	1.804	0.139
Loss	8p23.1	chr8:3590569-12634489	<0.001	**11%**	**1.862**	**0.011**	4%	1.638	0.401
Loss	9p21.3	chr9:16096660-22474701	0.014	6%	0.653	0.304	3%	0.585	0.590
Loss	13q14.2	chr13:47704056-48229856	<0.001	6%	1.132	0.720	3%	2.577	0.058
Loss	13q34	chr13:109917656-114142980	<0.001	9%	1.520	0.149	2%	***3.403***	***0.070***
Loss	14q32.2	chr14:98679424-106368585	0.008	8%	1.336	0.356	2%	1.401	0.737
Loss	15q15.2	chr15:40533684-41963394	<0.001	10%	1.124	0.689	3%	2.468	0.115
Loss	16q12.1	chr16:49721911-51145363	0.001	7%	1.338	0.354	7%	1.088	0.871
Loss	17p13.2	chr17:1-6498734	<0.001	17%	0.904	0.679	4%	1.976	0.242
Loss	18q23	chr18:64487411-76117153	0.024	5%	0.767	0.559	2%	0.897	0.914

Supplementary Table 2: B-cell lymphoma is transplantable to secondary recipients

Strain	Tissue of Origin	Cell Subset	# Transplanted Cells	# Transplanted Animals	B-cell lymphoma incidence (%)	Disease Latency
Sca1-Bcl6$^{\Delta}$	Bone Marrow	HSC	1,000	5	100	323 ± 38
			5,000	5	100	274 ± 41
Wild-type	Bone Marrow	HSC	1,000	5	0	-
			5,000	5	0	-
Sca1-Bcl6$^{\Delta}$	Spleen	B220+	100,000	7	0	-
			1,000,000	7	0	-
Wild-type	Spleen	B220+	100,000	7	0	-
			1,000,000	7	0	-

Supplementary Table 3: Significantly enriched signatures in genes with conserved hypermethylation between HSPC and mature B cells of *Sca1-Bcl6*$^{\Delta}$ mice

Name	PubMed ID	Terms in Query	Terms in Genome	Genome Enrichment FDR	Terms in Promoter Analysis	Promoter Analysis P-value
Mouse StemCell_Schimmer06_1554genes	16484322	53	1291	<0.001	464	0.042
Human StemCell_Matushansky08_1453genes	18310505	52	1245	<0.001	376	0.005
Rat Hypothalamic_Mansuy10_1931genes	20937356	51	1251	<0.001	428	0.032
Human Embryo_Heim07_1650genes	17880687	50	1122	<0.001	346	0.002
Human Colon_Grade07_1102genes	17210682	37	833	<0.001	277	0.017
Genes down-regulated in B493-6 cells (B lymphocytes) upon serum stimulation but not affected by MYC.	15516975	32	695	<0.001	16	0
Mouse Lung_Rangasamy09_248genes	19286929	16	168	<0.001	74	0.002
Human Breast_Dai05_50genes	15899795	7	45	<0.001	14	<0.001
Top 50 down-regulated markers for DLBCL that distinguished between cured and fatal/refractory.	11786909	7	45	0.100	15	<0.001

Supplementary Table 4: Primer Sequences

Name	Sequence
VHJ558-F	CGAGCTCTCCARCACAGCCTWCATGCARCTCARC
VH7183-F	CGGTACCAAGAASAMCCTGTWCCTGCAAATGASC
VHQ52-F	CGGTACCAGACTGARCATCASCAAGGACAAYTCC
VHGam3.8-F	CAAGGGACGGTTTGCCTTCTCTTTGGAA
Vκ-F	GGCTGCAGSTTCAGTGGCAGTGGRTCWGGRAC
JH4-R	TCTCAGCCGGCTCCCTCAGGG
Jκ5-R	ATGCGACGTCAACTGATAATGAGCCCTCTCC
Cμ-F	TGGCCATGGGCTGCCTAGCCCGGGACTT
Cμ-B	GCCTGACTGAGCTCACACAAGGAGGA