Supplementary Materials and Methods

Setdb1 genotyping

To detect floxed and wild type *Setdb1* alleles primers spanning the floxed region were used (5′-TGCCCCCACCAACTTTATAC-3′ (F), 5′-AAACACTCCCCCACAGACAG-3′ (R)). PCR amplification generated 2 fragments of 600 bp and 500 bp which corresponded to *Setdb1* floxed and *Setdb1* wild type alleles, respectively. To identify presence of the *Setdb1* delta allele the following primers were used: 5′-TGCCCCCACCAACTTTATAC-3′ (F) and 5′-AGTAAATCTTTGAGCCAGAGCAAGC-3′ (R), resulting in a 350 bp PCR amplicon. The *Mb1-Cre* allele was detected using the following primer pair: 5′-CCCTGTGGATGCCACCTC-3′ (F) and 5′-GTCCTGGCATCTGTCAGAG-3′ (R). *Vav-bcl2* was detected using primers 5′-ACGGTGGTGGAGGAGCTCTTC-3′ (F) and 5′-AAAACCTCCCACACCTCCCCCTGAA-3′ (R).

Chromatin immunoprecipitation

For ChIP-qPCR $1x10^6$ sorted cells were fixed using 1% formaldehyde. Fixation was stopped by adding glycine (final concentration 0.125M). Fixed cells were washed twice by rotation using PBS 10% FCS. After the last wash, pellets were flash frozen.

Cells were lysed by adding buffer B (50 mM Tris-HCl pH 8.0, 10 mM EDTA, 1% SDS). Lysates were then transferred to AFA Fiber microtubes with Snap-Cap and sonicated using Covaris E220 for 20 minutes at 4°C.

For the immunoprecipitation step α-H3K9me3 (Activ Motif, #39161.39162, lot #13509002) and α-H3K4me3 (Diagenode, #pAB-003-050, lot #A49-001 and #CS-003-100, lot #A5051-001P) antibodies were bound to magnetic Dynabeads (Applied technology). Next, sheared chromatin was diluted in buffer A (10 mM Tris HCl pH 7.5, 1 mM EDTA, 0.5 mM EGTA, 1% Triton 100 X, 0.1% SDS, 0.1% Na-deoxycholate, 140 mM NaCl) and incubated with the appropriate antibody-conjugated beads for 4 hours at 4°C on a rotator. Immunoprecipitated material was first washed 3 times with buffer A and then once with buffer C (10 mM Tris-HCl pH 8.0, 10 mM EDTA). Buffer A, B and C were all provided with protease inhibitor cocktail (Roche). Chromatin was then resuspended in elution buffer (50 mM Tris-HCl pH 8.0, 10 mM EDTA, 1% SDS) and incubated at 65°C on a mixer for 20 minutes. Eluted chromatin was next incubated overnight at 65°C for reverse cross-linking. Chromatin samples were treated with

RNase A and proteinase K and cleaned by phenol/chloroform purification. This material was subsequently used for qRT-PCR quantification with SYBR Green.

RNA-seq analysis

RNA-seq reads were mapped to the mouse genome (mm10) using tophat2 (Kim et al., 2013). Expression of genes in RPKM was calculated with cuffdiff and cummerbund (Trapnell et al., 2013). Dot plots were generated with ggplot2 (REF). GO term analysis of differentially expressed genes was performed with GOstat (Beissbarth and Speed, 2004). For gene set enrichment analysis (Mootha et al., 2003; Subramanian et al., 2005) of differentially expressed genes custom gene lists for fraction A and fraction B/C specific genes were generated based on expression data from the ImmGen consortium.

For expression analysis of individual repeat elements, RNA-seq reads were mapped to the genome using bowtie (Langmead et al., 2009). Coverage across repeat families and individual repeats was analyzed using analyzeRepeats (Heinz et al., 2010). To identify individual MLV repeats we extracted all MLV elements from the rmsk database (UCSC) and quantified normalized RNA-seq read coverage using analyzeRepeats.pl. The four top-regulated MLV repeats display a >10 fold change in expression (control vs. *Setdb1*^{Mb1}; see Supplementary Table S3).

DNA methylation analysis

Genomic DNA from control and *Setdb1*^{Mb1} pro-B cells was subjected to bisulfite conversion using the EpiTect Bisulfite Kit™ (Qiagen) according to the manufacturer's protocol. Target regions were amplified by PCR, subcloned into pBluescript-SK2+ (Stratagene) and analyzed by Sanger sequencing. Methylation analysis of sequencing data was performed using BiQ Analyzer (Bock et al., 2005).

ChIP-seq analysis

Paired end ChIP-seq reads were mapped to the mouse genome (mm10) using bowtie (Langmead et al., 2009). Reads mapping to multiple locations were discarded. Setdb1 peaks were identified using MACS (Zhang et al., 2008) and PeakRanger (Feng et al., 2011). Coverage across Setdb1 peaks was calculated using homer (Heinz et al., 2010). Cluster

analysis was performed using cluster3 (de Hoon et al., 2004) and visualized with Java TreeView (Saldanha, 2004).

V-DJ recombination

Pro-B cells (Kit^+ $CD19^+$ IgM^- IgD^-) were isolated from bone marrow of *control* (n=3), *Setdb1*^{Mb1} (n=3), B*cl2* (n=1) and *Setdb1*^{Mb1}*; Bcl2* (n=1) mice using FACS Aria cell sorter. Genomic DNA was prepared from these cells using standard procedures. The analysis of V-DJ recombination was done by a PCR based assay as described (Fuxa et al., 2004).

Definition of hematopoietic cell types

For better definition of the hematopoietic populations cell doublets were excluded. Cell population analysis of transplanted mice was performed using the same marker scheme shown above with the implementation of CD45.1 and CD45.2 markers to discriminate donor bone marrows. Lineage negative cells (lin^-) were detected using the following PE labeled antibody cocktail: CD45R/B220, CD5, CD19, CD11c, CD8a, CD4, Ly-6G (Gr-1), CD3e, CD19, CD11b(Mac-1).

For cell sorting, we used the marker combination to detect pro-B cells. Alternatively we sorted $CD43^+$ $CD19^+$ cell population (pro-B and pre-B cells), to perform ChIP-qPCR. IgD^+ IgM^+ cells were also sorted from *Setdb1*^{Mb1}*; Bcl2* spleen to check the deletion rate in the peripheral mature B cells.

Envelope protein and Fcγr2b staining

To detect the MLV envelope protein on pro-B cells, $4x10^6$ bone marrow cells were resuspended in a volume of 50 µl and incubated 30 minutes at RT with rat α-env (83A25) diluted 1:5. Subsequently cells were incubated for 1 hour at RT with an anti-rat secondary antibody conjugated Alexa-647. Bone marrow cells were then stained with IgD, IgM, CD19 and c-kit to discriminate the pro-B cell population (IgD^-, IgM^-, $CD19^+$, $c\text{-}kit^+$). Since all antibodies used for the pro-B cell staining were rat antibodies, env staining had to be performed in the absence of Fc-block and before the pro-B cell staining to avoid non-specific binding of the secondary α-rat Alexa-647.

Detection of Fcγr2b was performed in the absence of Fc-block which would also recognize Fcγr2b. Pro-B cells were stained with the same marker combination described above together with α-Fcgr2b conjugated with PE-Cy7. After every incubation step, cells were washed with FACS buffer to remove the excess of antibody before analysis with FACS Canto.

DNA damage analysis - γH2A.X foci enumeration

For γH2A.X foci analysis, bone marrow cells were first labeled for B220-PE (eBiosciences) and then enriched by magnetic sorting using anti-PE MicroBeads (Miltenyi). Sorted cells were stained for CD19-APC (eBiosciences), fixed and permeabilized using a Foxp3 staining kit (eBioscience) according to the manufacturer's instructions. Subsequently, cells were incubated with rabbit anti-mouse phospho-histone H2A.X (Ser139) primary antibody (Cat# 2577, Cell Signaling) at the dilution of 1:1000 in blocking solution at 4°C overnight. Secondary antibody staining was performed with an Alexa Fluor 488-conjugated goat anti-rabbit IgG antibody diluted 1:1000 in blocking solution for 60 min at room temperature. DAPI was used to visualize nuclei.

A sample without primary antibody served as a negative control. For positive control, cells isolated from lethally irradiated mouse were subjected for the analysis as described above.

Images were captured from 100,000 cells at 60X magnification using the next generation imaging flow cytometry with the Amnis ImagestreamX Mark II (Millipore). The acquired data was analyzed with the IDEAS v6 software. γH2A.X foci were enumerated using the spot count wizard according to the detailed protocol described before (Bourton et al., 2012; Parris et al., 2015).

Table S1. Regulated genes in *Setdb1*^{Mb1} pro-B cells.

Click here to Download Tables S1

Table S2. Expression of repetitive elements in *Setdb1*^{Mb1} pro-B cells.

Click here to Download Tables S2

Table S3. Expression of individual MLV elements in *Setdb1*^{Mb1} pro-B cells.

Click here to Download Tables S3

Table S4. Antibodies

FACS antibodies

Reactivity	Clone	Fluorochrome	Provider
BP-1	6C3	PE	eBioscience
CD117	2B8	PE	Pharmingen
CD11b (Mac-1)	M1/70	PE	Pharmingen
CD11c	HL3	PE	Pharmingen
CD127(IL7Rα)	A7R34	PE-Cy5	eBioscience
CD16/32	93	PE-Cy7	eBioscience
CD19	1D3	PE	Pharmingen
CD19	1D3	APC	Pharmingen
CD19	1D3	APC-Cy7	Pharmingen
CD24 (HSA)	M1/69	FITC	Pharmingen
CD25	PC61	PE-Cy5	eBioscience
CD34	RAM34	Alexa-Fluor 647	eBioscience
CD34	RAM34	eFluor660	eBioscience
CD3e	145-2C11	PE	Pharmingen
CD4	(L3T4)(PM4-5)	PE	Pharmingen
CD43	S7	APC	Pharmingen
CD45.1	A20	PE-Cy7	eBioscience
CD45.2	104	APC	eBioscience
CD45R (B220)	RA3-6B2	PE	Pharmingen
CD45R (B220)	RA3-SB2	Alexa-Fluor750	eBioscience
CD5	53-7.3	PE	Pharmingen
CD8a	53-6.7	PE	Pharmingen
DX-5	DX5	PE	eBioscience
Env	82A25	N/A	Frank Malik
Fcblock(CD16/CD32)	2.4G2	N/A	Pharmingen
IgD	11-26c.2a	FITC	Pharmingen
IgM	II/41	FITC	Pharmingen
Ly-6G (Gr-1)	RB6-8C5	PE	eBioscience
Rat IgG	polyclonal	Alexa647	Life Technologies
Sca-1	D7	FITC	Pharmingen
Ter119	Ter-119	PE	eBioscience

ChIP antibodies

Epitope	Company	Catalog number	LOT	Technique
H3K9me3	Activ Motif	#39161.39162	#13509002	ChIP-qPCR
H3K4me3	Diagenode	#pAB-003-050	#A49-001	ChIP-qPCR
H3K4me3	Diagenode	#CS-003-100	A5051-001P	ChIP-qPCR
Setdb1	Santa Cruz	SC-66884X		ChIP-Seq
Setdb1	Themo Scientific	PA5-30334		ChIP-seq
H3K9me3	Diagenode	pAb-056-050	A1675-001P	ChIP-Seq
H3K9ac	Millipore	#07-352	DAM1813175	ChIP-Seq

Table S5. Definition of hematopoietic cell types for FACS analysis and FACS sorting

Cell population	Gating strategy
Immature B	living cells, lymph, IgD-, B220+, IgM+
Mature B	living cells, lymph, IgM-, B220+, IgD+
pro-B	living cells, lymph, CD19+, IgD-, IgM-, c-kit+ CD25-
pre-B	living cells, lymph, CD19+, IgD-, IgM-, c-kit- CD25+
Fr. A	living cells, lymph, CD43+, B220+, HSA/CD24-/ low, BP-1-
Fr. B	living cells, lymph, CD43+, B220+, HSA/CD24+/high, BP-1-
Fr. C	living cells, lymph, CD43+, B220+, HSA/CD24 high, BP-1+
LSK	living cells, lin-, Sca+, c-kit+
CLP	living cells, lin-, IL7rα+, Sca low, c-kit low
CMP	living cells, lin-, IL7rα-, c-kit high, Sca-, CD34+, CD16/32-/ low
GMP	living cells, lin-, IL7rα-, c-kit high, Sca-, CD34+, CD16/32+
MEP	living cells, lin-, IL7rα-, c-kit high, Sca-, CD34-, CD16/32-
Pre-pro	living cells, lymph, lin-, IL7rα+, c-kit low, CD43low, B220+, CD93+
B cells	living cells, lymph, CD19+, B220+

Table S6. Primers for bisulfite PCR analysis

target	internal ID		sequence (5' to 3' direction)	taken from reference
IAP GAG region (478 bp)	GS2672	fw	aggttagtttgttgattggttttag	(Sadic et al., 2015)
	GS2673	rw	aatcaacaaaataaactccctaacc	
MLV group1 (256bp) (MLV4)	GS3758	fw	GTTTTTAAAATTTTTTAAAGATAAGATTAA	(Collins et al., 2015)
	GS3759	rw	TTATAATAAAATCTTTCATTCCCCC	
Emv2 (272 bp) (MLV8)	GS3764	fw	TTAGGGTTAGATTTAGAGGGGTGGT	(Collins et al., 2015)
	GS3765	rw	CTAAATAACCCAATCAATAAATCC	

Table S7. Primer pairs for qRT-PCR and ChIP-qPCR analyses

Primer	Sequence	Experiment
MLV1-gag F	TCTTGGCCACCGTAGTTACAG	qRT-PCR/ChIP-qPCR
MLV1-gag R	CCAGTGTCCCTTTTCTTTGCAG	qRT-PCR/ChIP-qPCR
MLV5-gag F	AGCTCCAAAGAATCCGAAACG	qRT-PCR/ChIP-qPCR
MLV5-gag R	ATCTGTATCTGGCGGTTCCG	qRT-PCR/ChIP-qPCR
MLV8-gag F	TGACCCAGCGTCTCTTCTTG	qRT-PCR/ChIP-qPCR
MLV8-gag R	GGACCGCTTCTAAAAACATGGG	qRT-PCR/ChIP-qPCR
AI506816 F	CCTGCTATGAAGGGGACAAAG	qRT-PCR
AI506816 R	ATCTTCGGAAGAGCAGTCAGTG	qRT-PCR
Fcgr2b F	GGAAGGACACTGCACCAGTC	qRT-PCR
Fcgr2b R	CCAGTGACAGCAGCCACAAT	qRT-PCR
Tubb3 F	GGCAACTATGTAGGGGACTCAG	qRT-PCR
Tubb3 R	ATGGTTCCAGGTTCCAAGTC	qRT-PCR
Gapdh F	TCAAGAAGGTGGTGAAGCAG	qRT-PCR
Gapdh R	GTTGAAGTCGCAGGAGACAA	qRT-PCR
Hprt F	ATGAGCGCAAGTTGAATCTG	qRT-PCR
Hprt R	CAGATGGCCACAGGACTAGA	qRT-PCR
Xbp1u F	GACTATGTGCACCTCTGCAG	qRT-PCR
Xbp1u R	CTGGGAGTTCCTCCAGACTA	qRT-PCR
Xbp1s F	GAGTCCGCAGCAGGTG	qRT-PCR
Xbp1s R	GTGTCAGAGTCCATGGGA	qRT-PCR
Hspa5 F	TGCAGCAGGACATCAAGTTC	qRT-PCR
Hspa5 R	TTCTGGGGCAAATGTCTTGG	qRT-PCR
Bcl2l11 F	GCTGTGTTCCACTTGGATTCAC	qRT-PCR
Bcl2l11 R	AAGGTTGCTTGGCCATTTGG	qRT-PCR
Pdia6 F	TGGTGGGTACAGTTCTGGAAAG	qRT-PCR
Pdia6 R	CACACCACGGAGCATAAAACTC	qRT-PCR
IAPs-gag F	AGCAGGTGAAGCCACTG	ChIP-qPCR
IAPs-gag R	CTTGCCACACTTAGAGC	ChIP-qPCR
IAPs-global F	CGGGTCGCGGTAATAAAGGT	ChIP-qPCR

IAPs-global R	ACTCTCGTTCCCCAGCTGAA	ChIP-qPCR
Tubb3-intron1 F	TTCTGACTCGCATTCCCATCC	ChIP-qPCR
Tubb3-intron2 R	GGCTTAAGTGGCAACCTCAAAG	ChIP-qPCR
Def8-intron1 F	TGAGCCTTCGGTTTCACAAC	ChIP-qPCR
Def8-intron2 R	CAAAGCGCACCTCACATTTC	ChIP-qPCR
H19 F	AGCTTTGAGTACCCCAGGTTCA	ChIP-qPCR
H19 R	GCCTCTGCTTTTATGGCTATGG	ChIP-qPCR
Gapdh F	CCATCCCACGGCTCTGCAC	ChIP-qPCR
Gapdh R	GCAAGGCTTCCGTGCTCTCG	ChIP-qPCR

Table S8. Plasmids

ID	plasmid name	fragments	cloning primers fw (5' to 3') or fragments	cloning primers rw (5' to 3') or fragments	origin	comment and usage
183	psPAX2	-	-	-	Didier Trono (Addgene: 12260)	Used for lentiviral packaging Fig. 7C, 7D
655	pLKO2mod/EGFP-WPRE	-	-	-	(Kuhn et al., 2010)	Used for GFP over-expression Fig. 7D
802	pLKO1mod	-	-	-	(Kuhn et al., 2010)	shRNA cloning Fig. 7C
811	pLP-eco-env	-	-	-	(Dambacher et al., 2012)	Used for lentiviral packaging Fig. 7C, 7D
849	pLKO1mod/shSCRAMBLED	-	-	-	(Dambacher et al., 2012)	Used for knockdown control Fig. 7C
963	pLenti6/EF1a-3FLAG-IRES-PURO	-	-	-	(Sadic et al., 2015)	Used as backbone for env overexpression Fig. 7D
1483	pLKO1mod/shBcl2l11-1	annealed oligos	CGCGTCCGGGACGAGTTCAACGAAACTTACCTCGAGGTAAGTTTCGTTGAACTCGTCTTTTTGGAAA	CCGGTTTCCAAAAAGACGAGTTCAACGAAACTTACCTCGAGGTAAGTTTCGTTGAACTCGTCCCGGA	shRNA TRCN0000231244 (RNAi Consortium)	Used for knockdown of Bcl2l11 Fig. 7C
		backbone	Xho1	Spe1	pLKO1mod(#802)	
1502	pLenti6/EF1a-env1	env chr1 (CDS + genomic region downstream)	AGGTGTCGTGACTAGTTTGGATCCCACCATGGAAGGTCCAGCGTTCT	ACCAAGAACAAACCCCAGCT	Pre-amplification from genomic DNA of wild type (C57Bl6/J) pro B cells	Used for env over-expression Fig. 7D
		env1 chr1 (CDS)	AGGTGTCGTGACTAGTTTGGATCCCACCATGGAAGGTCCAGCGTTCT	GTAATCCAGAGGTTGATTGAATTCTTATTCACGCGATTCTACTTCT	Amplification from gel-purified pre-amplification	
		backbone	BamHI digest	EcoRI digest	pLenti6/EF1a-3FLAG-IRES-PURO (#963)	

Supplementary references

Beissbarth, T. and Speed, T. P. (2004). GOstat: find statistically overrepresented Gene Ontologies within a group of genes. *Bioinformatics* **20**, 1464-1465.

Bock, C., Reither, S., Mikeska, T., Paulsen, M., Walter, J. and Lengauer, T. (2005). BiQ Analyzer: visualization and quality control for DNA methylation data from bisulfite sequencing. *Bioinformatics* **21**, 4067-4068.

Bourton, E. C., Plowman, P. N., Zahir, S. A., Senguloglu, G. U., Serrai, H., Bottley, G. and Parris, C. N. (2012). Multispectral imaging flow cytometry reveals distinct frequencies of gamma-H2AX foci induction in DNA double strand break repair defective human cell lines. *Cytometry A* **81**, 130-137.

Collins, P. L., Kyle, K. E., Egawa, T., Shinkai, Y. and Oltz, E. M. (2015). The histone methyltransferase SETDB1 represses endogenous and exogenous retroviruses in B lymphocytes. *Proceedings of the National Academy of Sciences of the United States of America* **112**, 8367-8372.

Dambacher, S., Deng, W., Hahn, M., Sadic, D., Frohlich, J., Nuber, A., Hoischen, C., Diekmann, S., Leonhardt, H. and Schotta, G. (2012). CENP-C facilitates the recruitment of M18BP1 to centromeric chromatin. *Nucleus* **3**, 101-110.

de Hoon, M. J., Imoto, S., Nolan, J. and Miyano, S. (2004). Open source clustering software. *Bioinformatics* **20**, 1453-1454.

Feng, X., Grossman, R. and Stein, L. (2011). PeakRanger: a cloud-enabled peak caller for ChIP-seq data. *BMC bioinformatics* **12**, 139.

Fuxa, M., Skok, J., Souabni, A., Salvagiotto, G., Roldan, E. and Busslinger, M. (2004). Pax5 induces V-to-DJ rearrangements and locus contraction of the immunoglobulin heavy-chain gene. *Genes & development* **18**, 411-422.

Heinz, S., Benner, C., Spann, N., Bertolino, E., Lin, Y. C., Laslo, P., Cheng, J. X., Murre, C., Singh, H. and Glass, C. K. (2010). Simple combinations of lineage-determining transcription factors prime cis-regulatory elements required for macrophage and B cell identities. *Molecular cell* **38**, 576-589.

Kim, D., Pertea, G., Trapnell, C., Pimentel, H., Kelley, R. and Salzberg, S. L. (2013). TopHat2: accurate alignment of transcriptomes in the presence of insertions, deletions and gene fusions. *Genome biology* **14**, R36.

Kuhn, P. H., Wang, H., Dislich, B., Colombo, A., Zeitschel, U., Ellwart, J. W., Kremmer, E., Rossner, S. and Lichtenthaler, S. F. (2010). ADAM10 is the physiologically relevant, constitutive alpha-secretase of the amyloid precursor protein in primary neurons. *The EMBO journal* **29**, 3020-3032.

Langmead, B., Trapnell, C., Pop, M. and Salzberg, S. L. (2009). Ultrafast and memory-efficient alignment of short DNA sequences to the human genome. *Genome biology* **10**, R25.

Mootha, V. K., Lindgren, C. M., Eriksson, K. F., Subramanian, A., Sihag, S., Lehar, J., Puigserver, P., Carlsson, E., Ridderstrale, M., Laurila, E., et al. (2003). PGC-1alpha-responsive genes involved in oxidative phosphorylation are coordinately downregulated in human diabetes. *Nature genetics* **34**, 267-273.

Parris, C. N., Adam Zahir, S., Al-Ali, H., Bourton, E. C., Plowman, C. and Plowman, P. N. (2015). Enhanced gamma-H2AX DNA damage foci detection using multimagnification and extended depth of field in imaging flow cytometry. *Cytometry A* **87**, 717-723.

Sadic, D., Schmidt, K., Groh, S., Kondofersky, I., Ellwart, J., Fuchs, C., Theis, F. J. and Schotta, G. (2015). Atrx promotes heterochromatin formation at retrotransposons. *EMBO Rep* **16**, 836-850.

Saldanha, A. J. (2004). Java Treeview--extensible visualization of microarray data. *Bioinformatics* **20**, 3246-3248.

Subramanian, A., Tamayo, P., Mootha, V. K., Mukherjee, S., Ebert, B. L., Gillette, M. A., Paulovich, A., Pomeroy, S. L., Golub, T. R., Lander, E. S., et al. (2005). Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles. *Proceedings of the National Academy of Sciences of the United States of America* **102**, 15545-15550.

Trapnell, C., Hendrickson, D. G., Sauvageau, M., Goff, L., Rinn, J. L. and Pachter, L. (2013). Differential analysis of gene regulation at transcript resolution with RNA-seq. *Nature biotechnology* **31**, 46-53.

Zhang, Y., Liu, T., Meyer, C. A., Eeckhoute, J., Johnson, D. S., Bernstein, B. E., Nusbaum, C., Myers, R. M., Brown, M., Li, W., et al. (2008). Model-based analysis of ChIP-Seq (MACS). *Genome biology* **9**, R137.

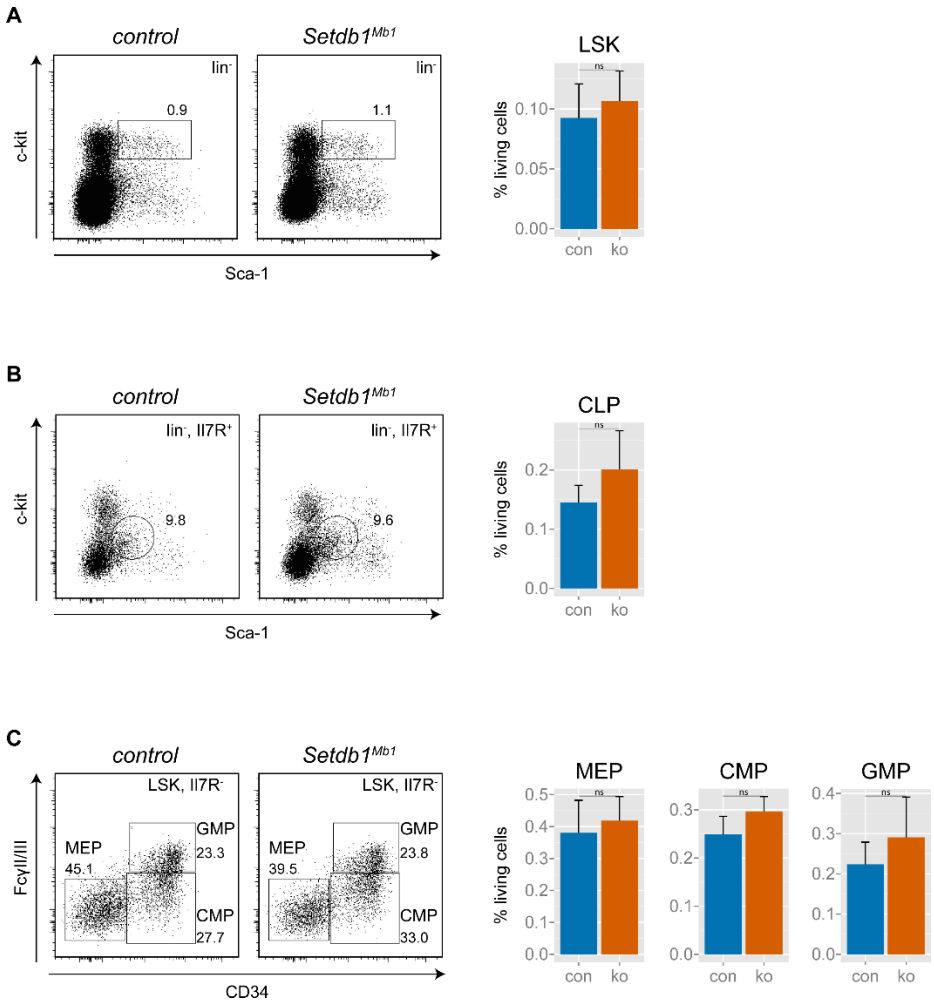

Figure S1. Normal frequencies of hematopoietic progenitors in *Setdb1*^{Mb1} mice

(A) FACS analyses of LSK cells (Lin$^-$ Sca$^+$ Kit$^+$) in *control* (con) and *Setdb1*^{Mb1} (ko) bone marrow. Bargraph depicts average cell numbers as percentage of living cells from 6 mice per genotype. NS, not significant (unpaired two-tailed Student's t-test).

(B) FACS analyses of CLPs (Lin$^-$ IL7Rα$^+$ Sca^{low} Kit^{low}) in *control* (con) and *Setdb1*^{Mb1} (ko) bone marrow. Bargraph depicts average cell numbers as percentage of living cells from 6 mice per genotype. NS, not significant (unpaired two-tailed Student's t-test).

(C) FACS analyses of myeloid progenitors: CMPs (Lin$^-$ IL7Rα$^-$ Kit^{high} Sca$^-$ CD34$^+$ and CD16/32$^-$); GMPs (Lin$^-$ IL7Rα$^-$ Kit^{high} Sca$^-$ CD34$^+$ and CD16/32$^+$) and MEPs (Lin$^-$ IL7Rα$^-$ Kit^{high} Sca$^-$ CD34$^-$ and CD16/32$^-$) in *control* (con) and *Setdb1*^{Mb1} (ko) bone marrow. Bargraph depicts average cell numbers as percentage of living cells from 6 mice per genotype. NS, not significant (unpaired two-tailed Student's t-test).

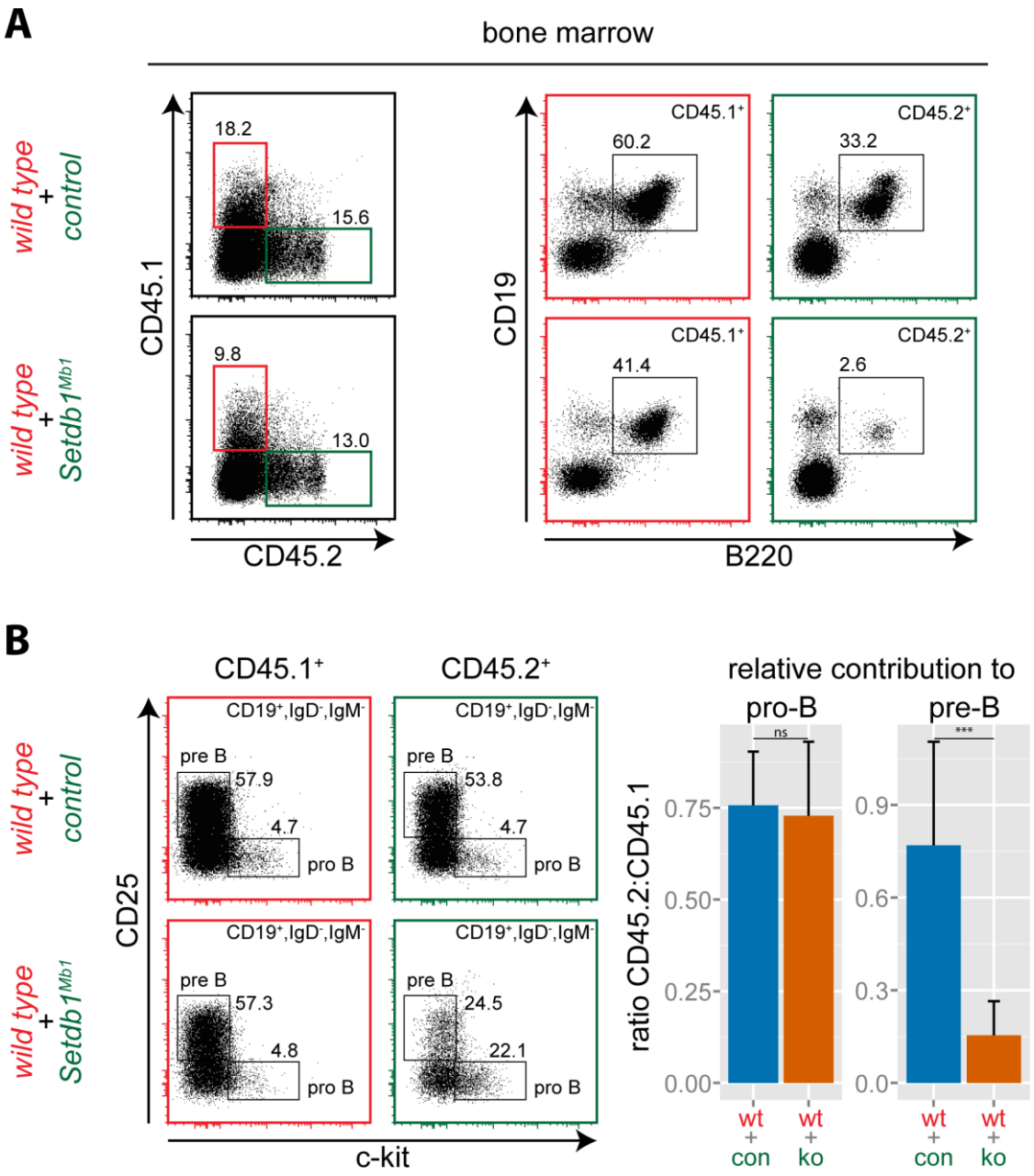

Figure S2. Setdb1 has cell-intrinsic functions for B cell development.

(A) Representative FACS plots showing the relative contribution to the B cell lineage ($B220^+$ $CD19^+$) of wild type vs. *control* or $Setdb1^{Mb1}$ donor bone marrow.

(B) Representative FACS plots showing contribution of donor bone marrow (wild type and *control* or wild type and $Setdb1^{Mb1}$) to pro-B and pre-B cell populations in recipient mice. Bargraph shows the quantification (n=3) of pro-B and pre-B cells in recipient mice as ratio between *control* (con) or $Setdb1^{Mb1}$ (ko) to wild type (wt). ***$P < 0.001$ (unpaired two-tailed Student's t-test).

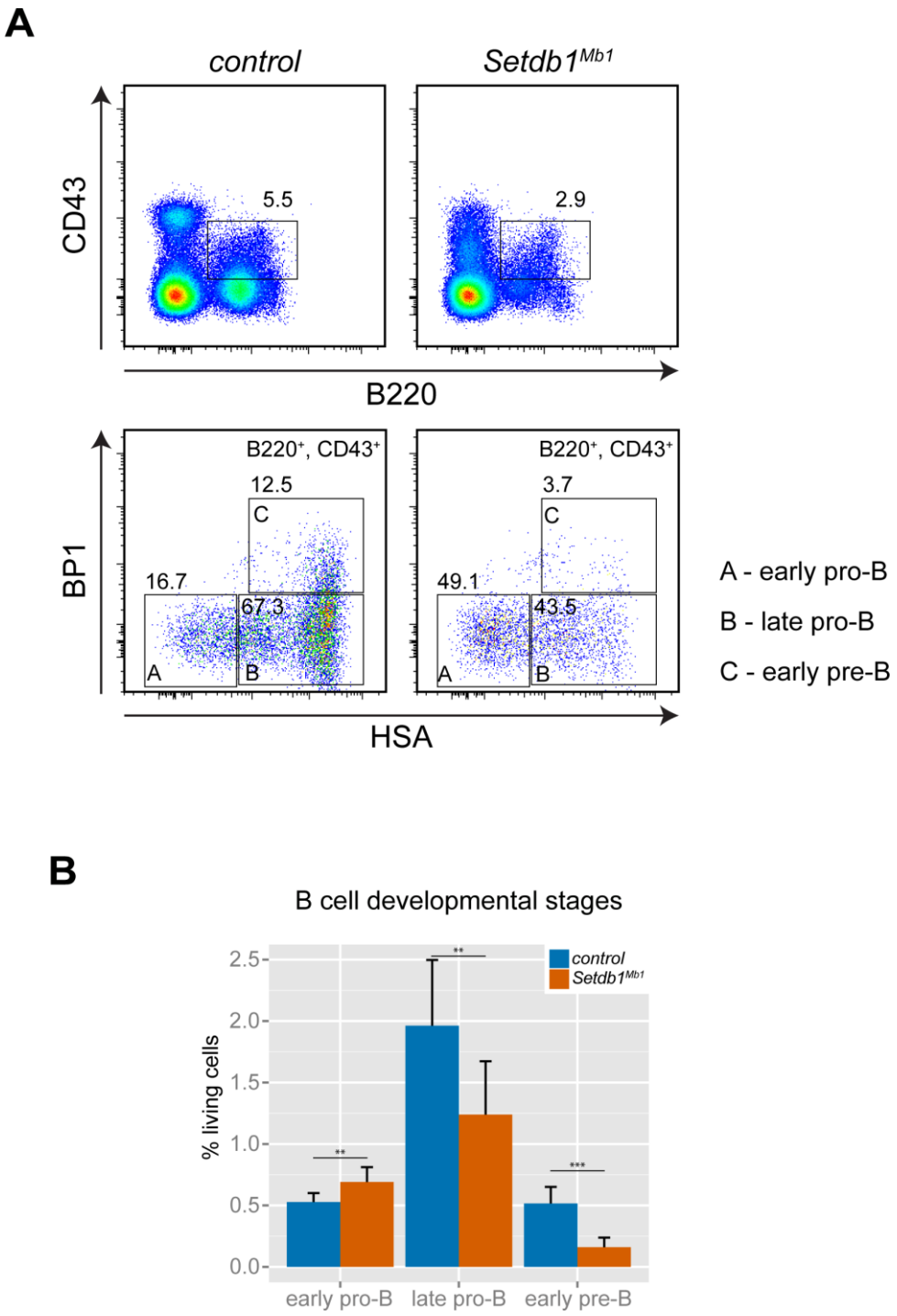

Figure S3. Block in pro-B to pre-B cell transition in *Setdb1*^{Mb1} mice.

(A) Distinct stages of B cell development were determined by FACS analysis of *control* and *Setdb1*^{Mb1} bone marrow according to Hardy, 1991: Fr.A (B220^{+} CD43^{+} HSA^{low} BP-1^{-}), Fr.B (B220^{+} CD43^{+} HSA^{high} BP-1^{-}) and Fr.C (B220^{+} CD43^{+} HSA^{high} BP-1^{+}).

(B) Bargraph depicts average cell numbers of Hardy cell stages as percentage of living cells from 6 mice per genotype. ***P < 0.001, **P < 0.01 (unpaired two-tailed Student's t-test).

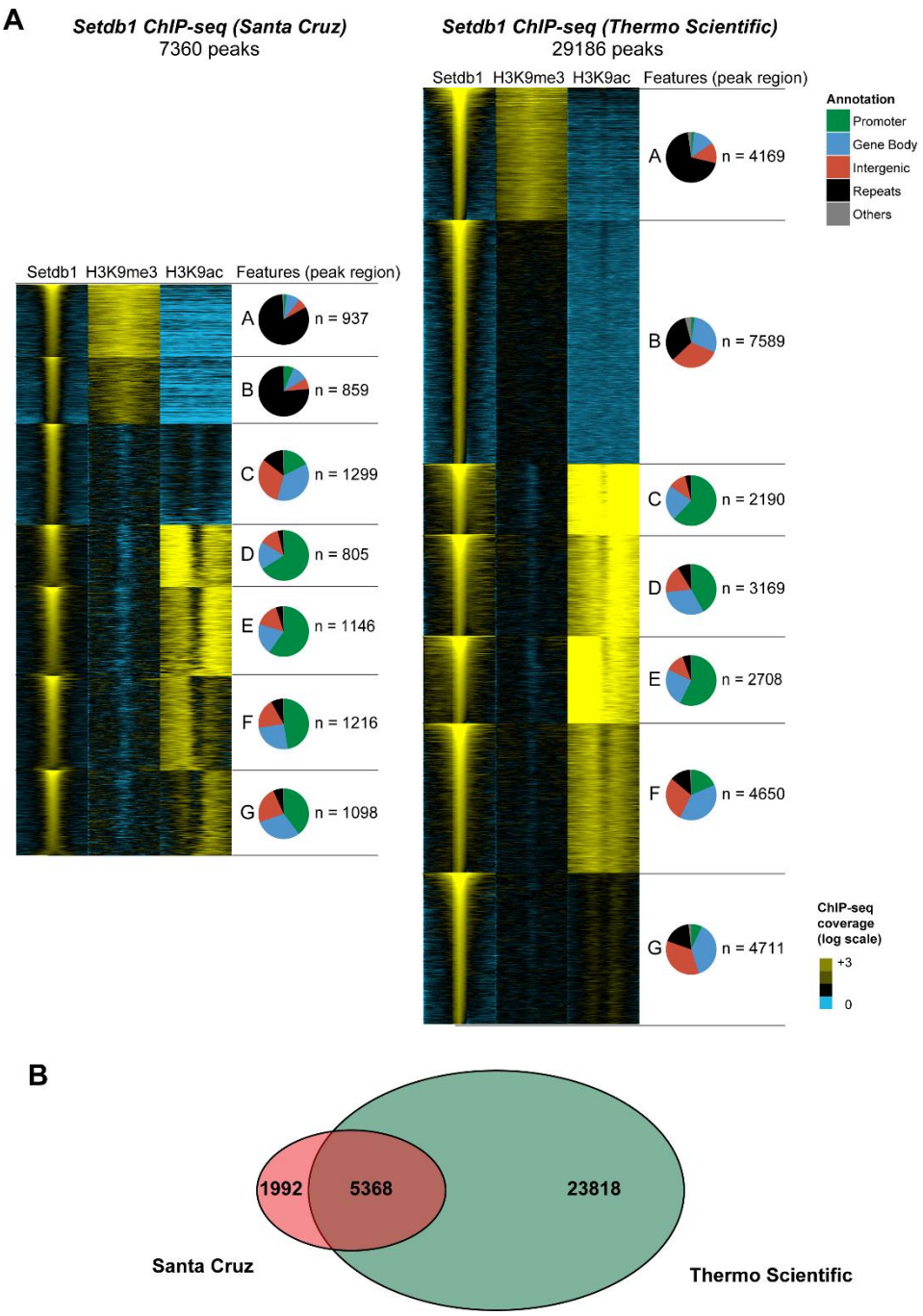

Figure S4. Comparative analysis of two independent Setdb1 ChIP-seq datasets.

(A) ChIP-Seq analysis for Setdb1 (Santa Cruz antibody and Thermo Scientific antibody), H3K9me3 and H3K9ac in short-term cultured *Rag2*$^{-/-}$ pro-B cells. Heatmaps shows log-transformed read coverage for Setdb1 and H3K9 modifications 1500 bp across all Setdb1 binding sites identified in each dataset. Peak clusters were generated based on H3K9me3/H3K9ac occupancy using Cluster3 software. Pie charts depict the frequency of genomic features at Setdb1 peaks in each cluster.

(B) Venn diagram depicting overlap between Setdb1 peaks identified in the Santa Cruz vs. Thermo Scientific dataset.

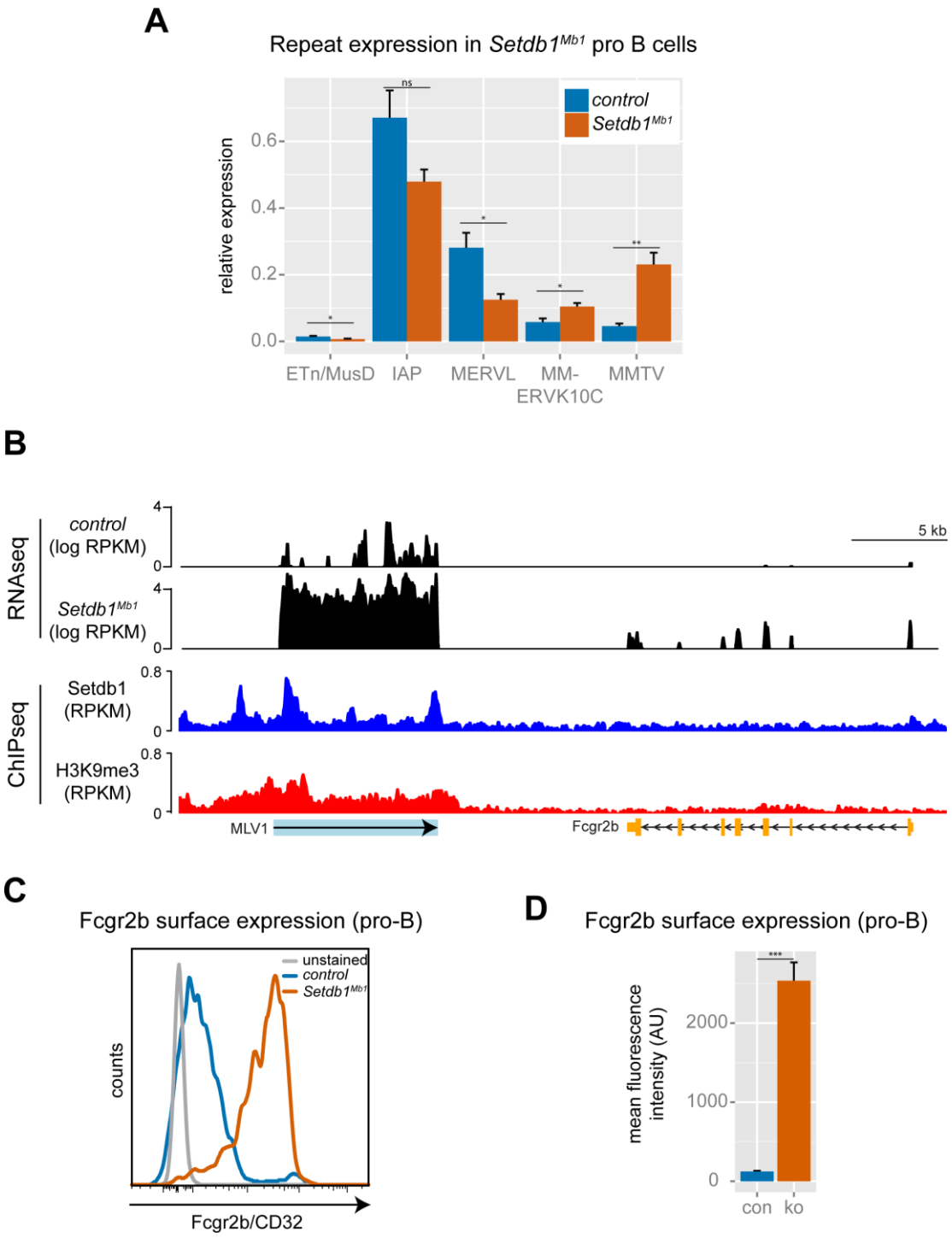

Figure S5. Retrotransposon expression; derepression of MLV1 and upregulation of Fcgr2b.

(A) Quantitative RT-PCR of retrotransposon classes in control vs. *Setdb1*^{Mb1} pro-B cells. Expression was calculated as relative expression to housekeeping genes from 6 biological replicates. NS, not significant; *P < 0.05 and **P < 0.01 (unpaired two-tailed Student's t-test).

(B) Coverage plot of normalized RNA-seq (*control* vs. *Setdb1*^{Mb1} pro-B cells) and ChIP-seq (short-term cultured *Rag2*$^{-/-}$ pro-B cells) coverage across the genomic region of MLV1.

(C) Fcgr2b protein expression detected by FACS on control and *Setdb1*^{Mb1} pro-B cells (CD19^{+} IgM^{-} IgD^{-} CD25^{-} Kit^{+}).

(D) Bargraph depicts the average Fcgr2b expression in *control* (con) and *Setdb1*^{Mb1} (ko) pro-B cells calculated as mean fluorescence intensity (MFI) from 6 mice per genotype. ***$P < 0.001$ (unpaired two-tailed Student's t-test).

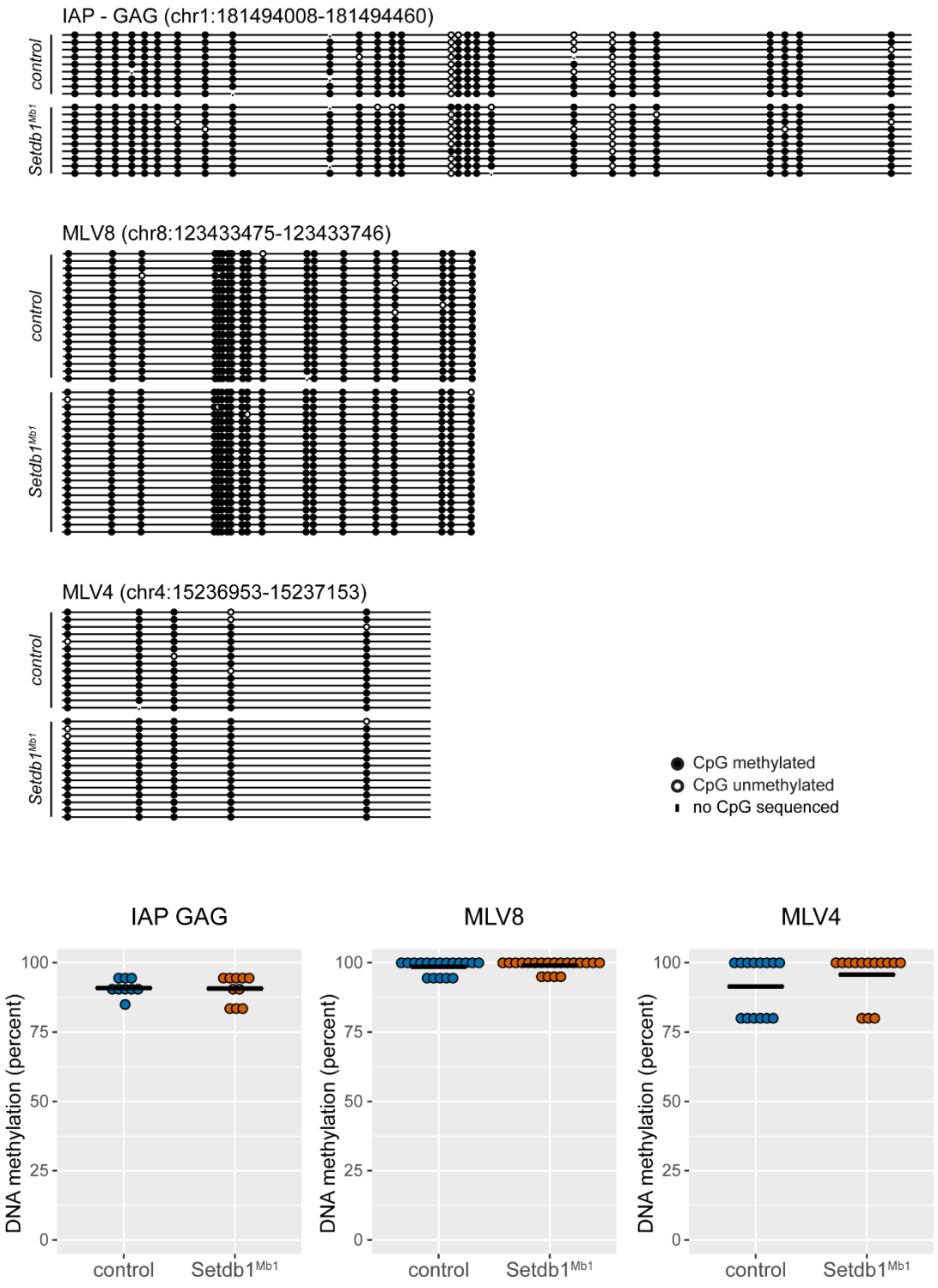

Figure S6. DNA methylation analysis of retrotransposons in pro-B cells.

DNA methylation of IAP GAG (no transcriptional change), MLV8 (derepressed) and MLV4 (no transcriptional change) was analyzed by bisulfite sequencing in control and *Setdb1*^{Mb1} pro-B cells.

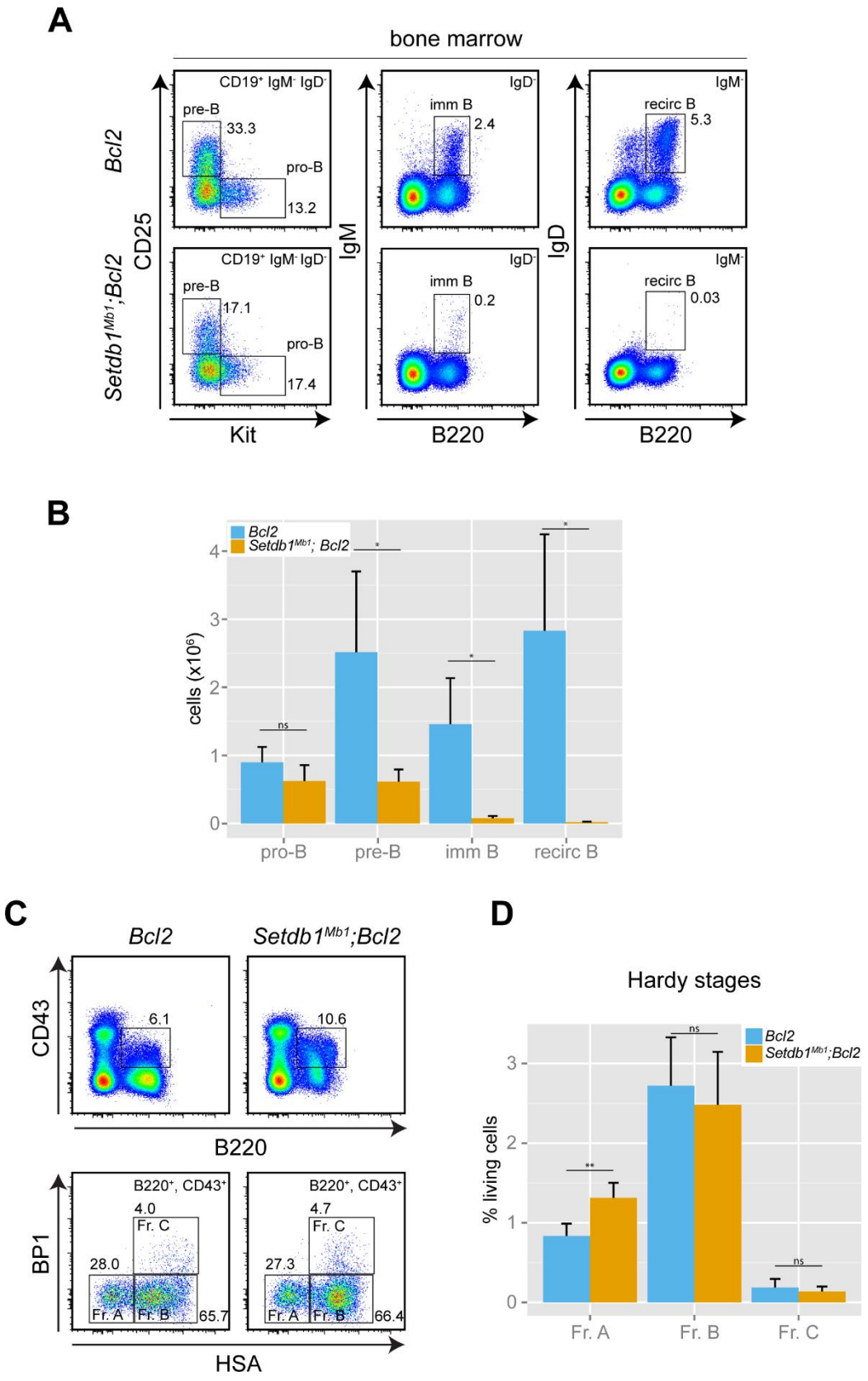

Figure S7. Forced expression of pro-survival Bcl2 partially rescues B cell development.

(A) Representative FACS plots showing different stages of B cell development in the bone marrow of *Bcl2* and *Setdb1*^{Mb1}*; Bcl2* mice.

(B) Bargraph showing average total cell numbers of B cell developmental stages in bone marrow from *Bcl2* and *Setdb1*^{Mb1}*; Bcl2* mice (n=6). NS, not significant; *P < 0.05, **P < 0.01 and ***P < 0.001 (unpaired two-tailed Student's t-test).

(C) Distinct stages of B cell development were determined by FACS analysis of *Bcl2* and *Setdb1*^{Mb1}*; Bcl2* bone marrow according to Hardy, 1991: Fr.A (B220^{+} CD43^{+} HSA^{low} BP-1^{-}), Fr.B (B220^{+} CD43^{+} HSA^{high} BP-1^{-}) and Fr.C (B220^{+} CD43^{+} HSA^{high} BP-1^{+}).

(D) Bargraph depicts average cell numbers of Hardy cell stages as percentage of living cells from 6 mice per genotype. NS, not significant; **P < 0.01 (unpaired two-tailed Student's t-test).

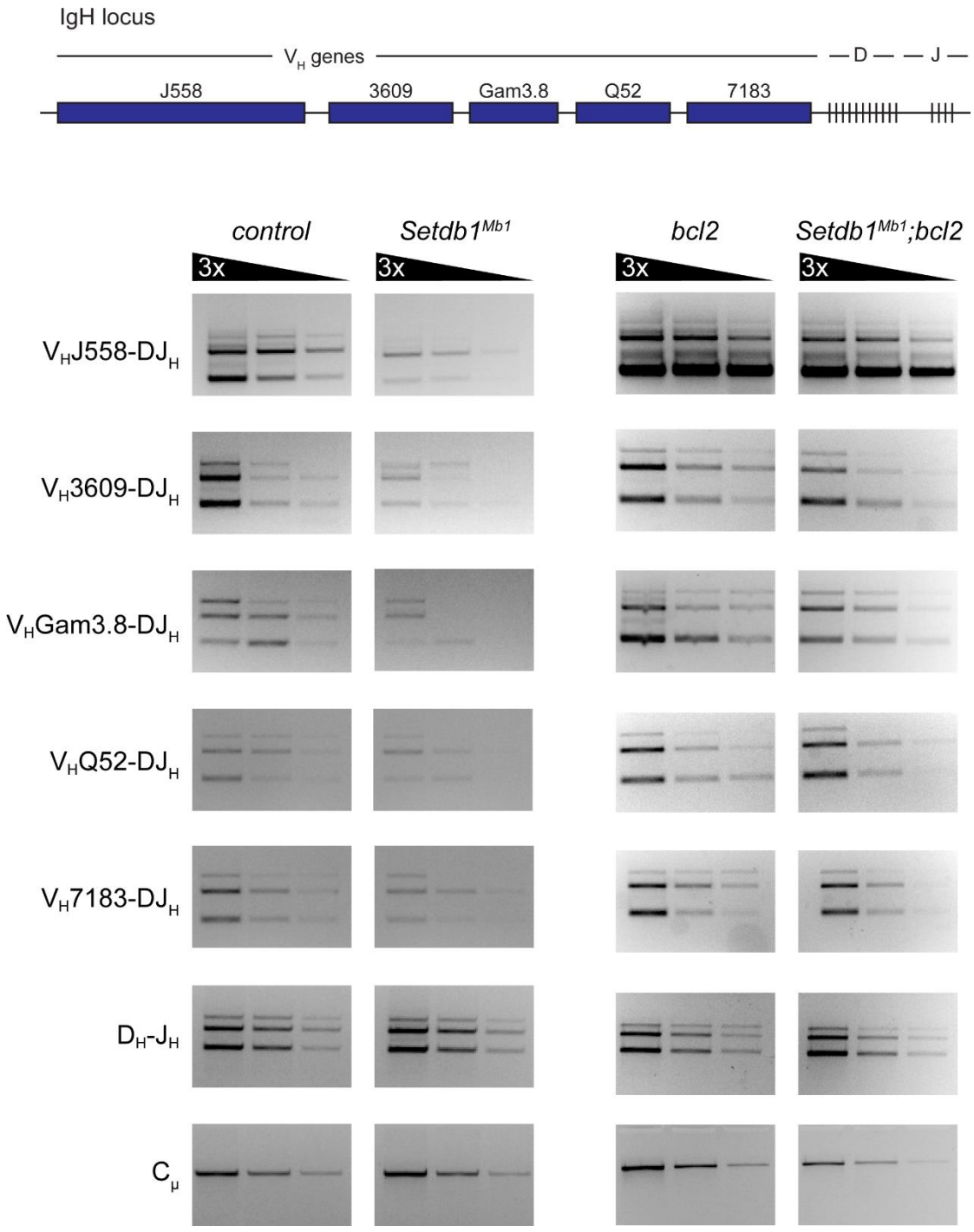

Figure S8. V-DJ recombination is not affected by loss of *Setdb1* in pro-B cells.

Schematic diagram of the V_H gene cluster of the *Igh* locus and the distal and proximal positions of the V_H gene families that were analyzed. DNA from control, *Setdb1*^{Mb1} as well as *Bcl2* and *Setdb1*^{Mb1}*; Bcl2* pro-B cells were analyzed for D_H–J_H and different V_H–DJ_H rearrangements by PCR of three-fold serial DNA dilutions. Input DNA was normalized by amplification of a PCR fragment from the IgH Cμ region.

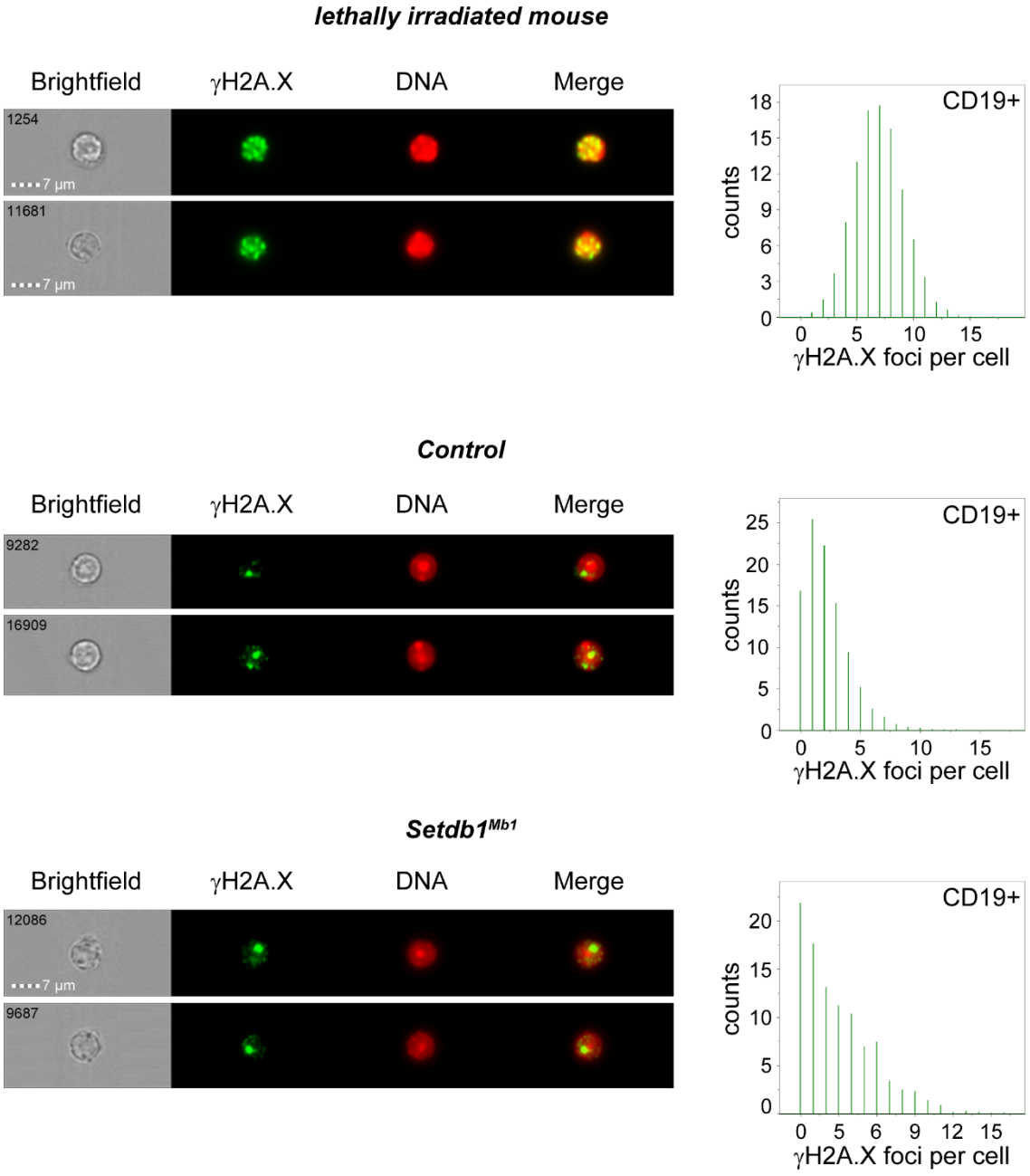

Figure S9. DNA damage analysis in B cells.

Representative images of CD19+ B cells stained with γH2A.X antibody. Histograms show the distribution of foci number per cell. Cells from lethally irradiated mice show in average high numbers of γH2A.X foci. Control and *Setdb1*^{Mb1} mutant cells display a comparable distribution of foci numbers per cell; the majority of cells show very few foci, which may stem from VD-J recombination events.

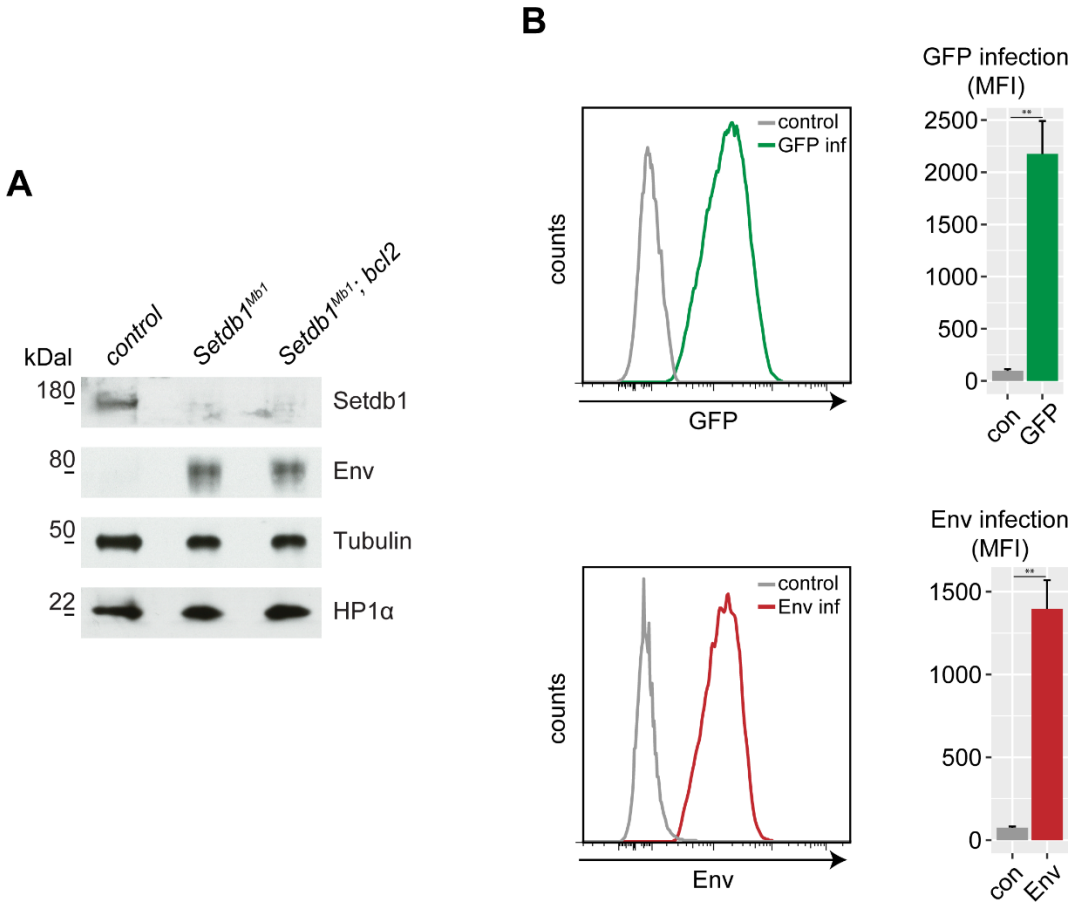

Figure S10. MLV Env overexpression in pro-B cells.

(A) Western blot analysis of sorted pro-B cells from control, *Setdb1*^{Mb1} and *Setdb1*^{Mb1}*; bcl2* mice. Both *Setdb1*^{Mb1} and *Setdb1*^{Mb1}*; bcl2* pro-B cells display loss of Setdb1 and strong expression of MLV Env protein. Tubulin and HP1α serve as loading controls.

(B) Lineage negative bone marrow cells (control - non infected, GFP infected and MLV Env infected) were differentiated into B cells and expression levels of GFP and Env were measured by FACS analysis. Bargraphs depict the average GFP or Env expression calculated as mean fluorescence intensity (MFI) from 3 control vs. infected samples. **P < 0.01 (unpaired two-tailed Student's t-test).