iScience, Volume *23*

Supplemental Information

Loss of H3K36 Methyltransferase SETD2 Impairs V(D)J Recombination during Lymphoid Development

S. Haihua Chu, Jonathan R. Chabon, Chloe N. Matovina, Janna C. Minehart, Bo-Ruei Chen, Jian Zhang, Vipul Kumar, Yijun Xiong, Elsa Callen, Putzer J. Hung, Zhaohui Feng, Richard P. Koche, X. Shirley Liu, Jayanta Chaudhuri, Andre Nussenzweig, Barry P. Sleckman, and Scott A. Armstrong

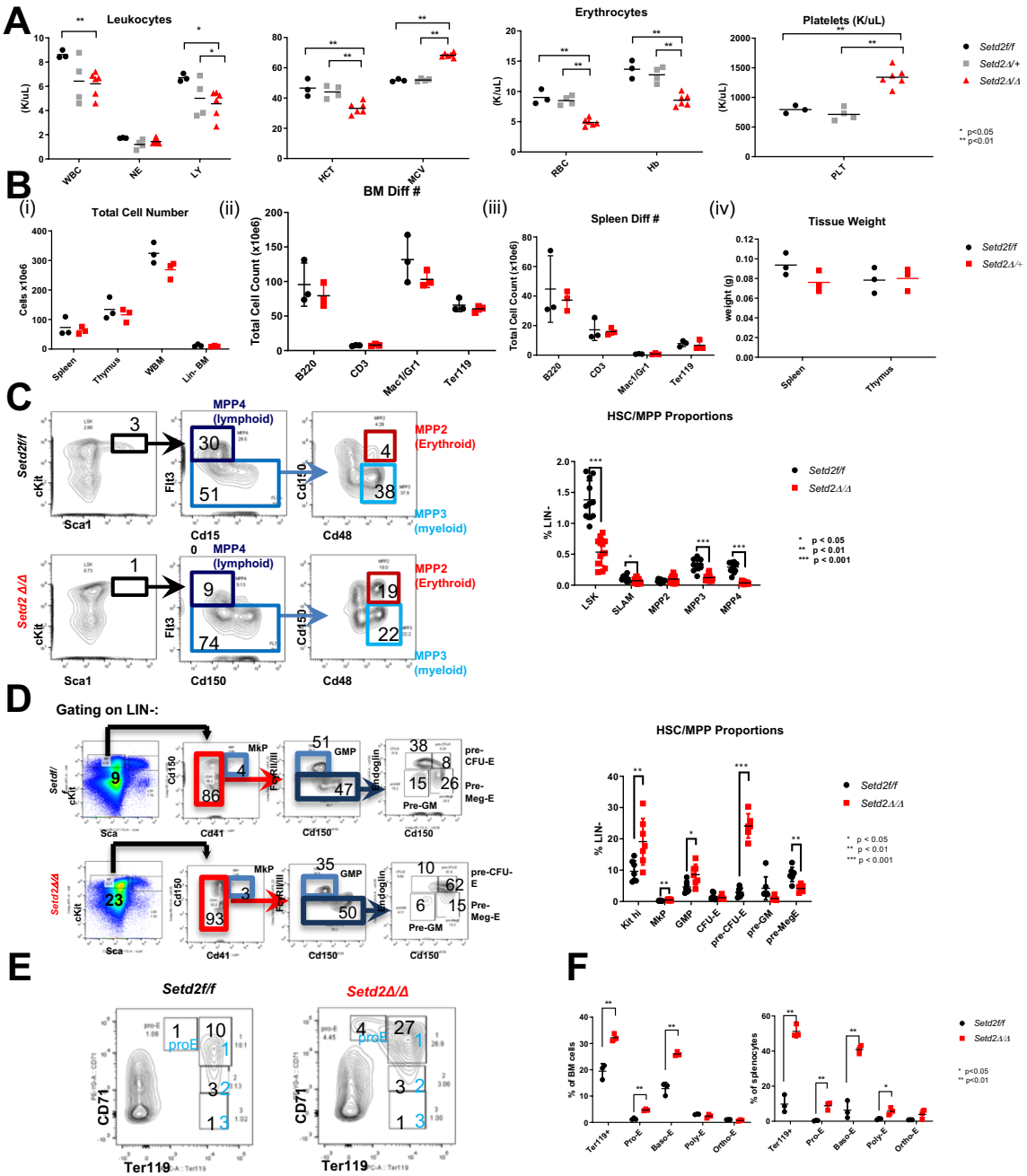


Figure S1. Loss of *Setd2*/H3K36me3 disrupts normal hematopoiesis, Related to Figure 1

(A) CBC and differentials from *Setd2Δ/Δ, Setd2Δ/+* and controls (*Setd2f/f*) (n≥3 for all groups) *, p<0.05 **, p<0.01 ***, p<0.001 **(B)** (i) total cellularity of spleen, thymus, whole bone marrow and lineage negative BM, total cellularity of cells positive for differentiated markers B220, Cd3, Mac1, Gr1, or Ter119 in (ii) bone marrow or (iii) spleen and (iv) spleen and thymus weights for heterozygous *Setd2* mice and controls (n= 3 for all groups). All values were non-significant. **(C)** Representative flow cytometry plot of HSC fractions, multipotent primed progenitors (MPP1-4) and percent composition of lineage negative cells of HSC stem cell and progenitor populations (n= 10 for all groups). **(D)** Representative flow cytometry plots for early myeloid and erythroid progenitor populations in bone marrow. MkP: Kit^{hi}Sca1^{-}Cd41^{+}Cd150^{+}, GMP: Kit^{hi}Sca1^{-}Cd41^{-}Cd150^{-}FcγRII/III^{+}, pre-GM: Kit^{hi}Sca1^{-}Cd41^{-}FcγRII/III^{-}Cd150^{-}Endoglin^{-}, pre-CFU-E: Kit^{hi}Sca1^{-}Cd41^{-}FcγRII/III^{-}Cd150^{+}Endoglin^{+}, pre-MegE: Kit^{hi}Sca1^{-}Cd41^{-}FcγRII/III^{-}Cd150^{+}Endoglin^{-} and proportion of HSC and erythroid progenitor populations of control and *Setd2Δ/Δ* (n = 7 for all groups). **(E)** Representative flow cytometry of erythroid progenitor populations. proE, 1-Basophilic erythroblasts (Baso-E), 2-late Baso-E and chromatophilic erythroblasts (Poly-E), and 3-orthochromatophilic erythroblasts (Ortho-E) in control and *Setd2Δ/Δ* bone marrow. **(F)** Percent composition of different erythroid progenitor populations in bone marrow and spleen of control and *Setd2Δ/Δ* mice (n = 3). *, p<0.05 **, p<0.01 ***, p<0.001, error bars represent SD.

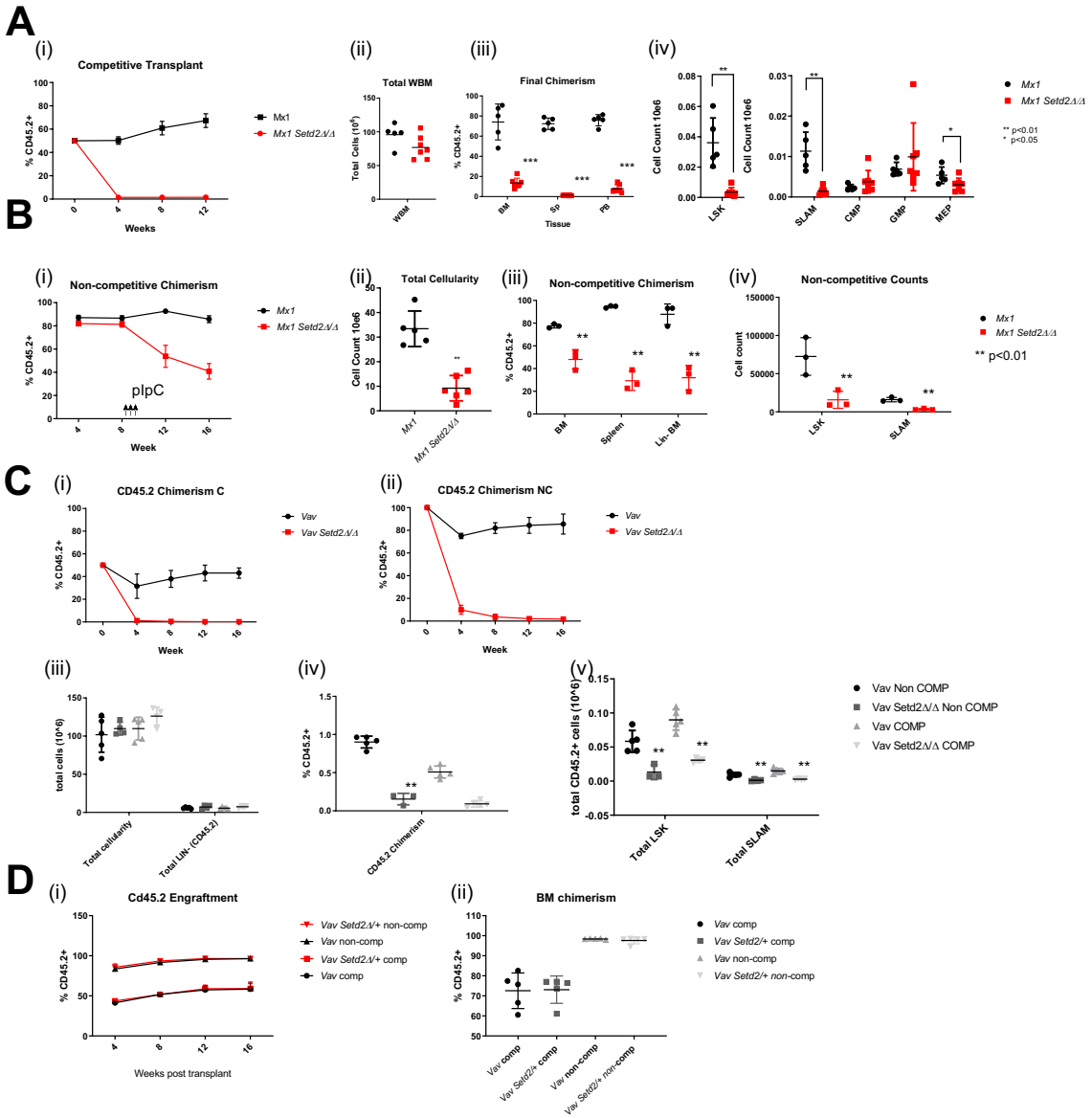


Figure S2. *Setd2Δ/Δ* HSCs are defective in competitive and non-competitive bone marrow (BM) reconstitution assays, Related to Figure 1

(A) (i) Peripheral blood engraftment, as measured by Cd45.2$^+$ of competitive reconstitution of lethally irradiated (9cGy) recipient mice with 1x10^6 WBM *Mx1 Setd2Δ/Δ* and control bone marrow with 1x10^6 WBM from syngeneic Cd45.1$^+$ mice after induction of Mx1-cre with polyI:polyC (pIpC). (ii) Total BM cellularity of mice 16 weeks after induction with pIpC (iii) percent Cd45.2 involvement in BM, spleen and peripheral blood, (iv) total cell counts of Cd45.2$^+$ donor HSC stem and progenitor populations. **(B)** Non-competitive reconstitution assay as conducted in (A) without 1x10^6 Cd45.1$^+$ wildtype competitor WBM cells. Arrows indicate when pIpC treatment was initiated. (ii-iv) graphs were generated from mice 28 days post pIpC induction as *Setd2Δ/Δ* mice were moribund due to complete BM failure. **(C)** Competitive and non-competitive reconstitution of lethally irradiated (9cGy) recipient mice with 1x10^6 WBM *Vav-cre Setd2Δ/Δ* and control BM with 1x10^6 WBM from syngeneic Cd45.1$^+$ mice. Overall peripheral blood Cd45.2 chimerism in (i) competitive and (ii) non-competitive reconstitution assays. (iii) total cellularity of Cd45.2$^+$ whole and lineage negative BM (iv) Cd45.2$^+$ chimerism (v) percentage and total numbers of Cd45.2$^+$ donor derived HSC of competitive and non-competitive transplants 16 weeks post-transplant **(D)** Competitive and non-competitive reconstitution of lethally irradiated recipient mice with 1x10^6 WBM *Vav-cre Setd2Δ/Δ* and *Vav-cre Setd2Δ/+* control BM with 1x10^6 WBM from syngeneic Cd45.1$^+$ mice. (i) Overall peripheral blood Cd45.2$^+$ chimerism and (ii) BM Cd45.2$^+$ chimerism in competitive and non-competitive reconstitution assays 16 weeks post-transplant. All data is representative of at least 3 independent transplant experiments. n = 5 for each transplant group. *, p<0.05 **, p<0.01 ***, p<0.001, error bars represent SD.

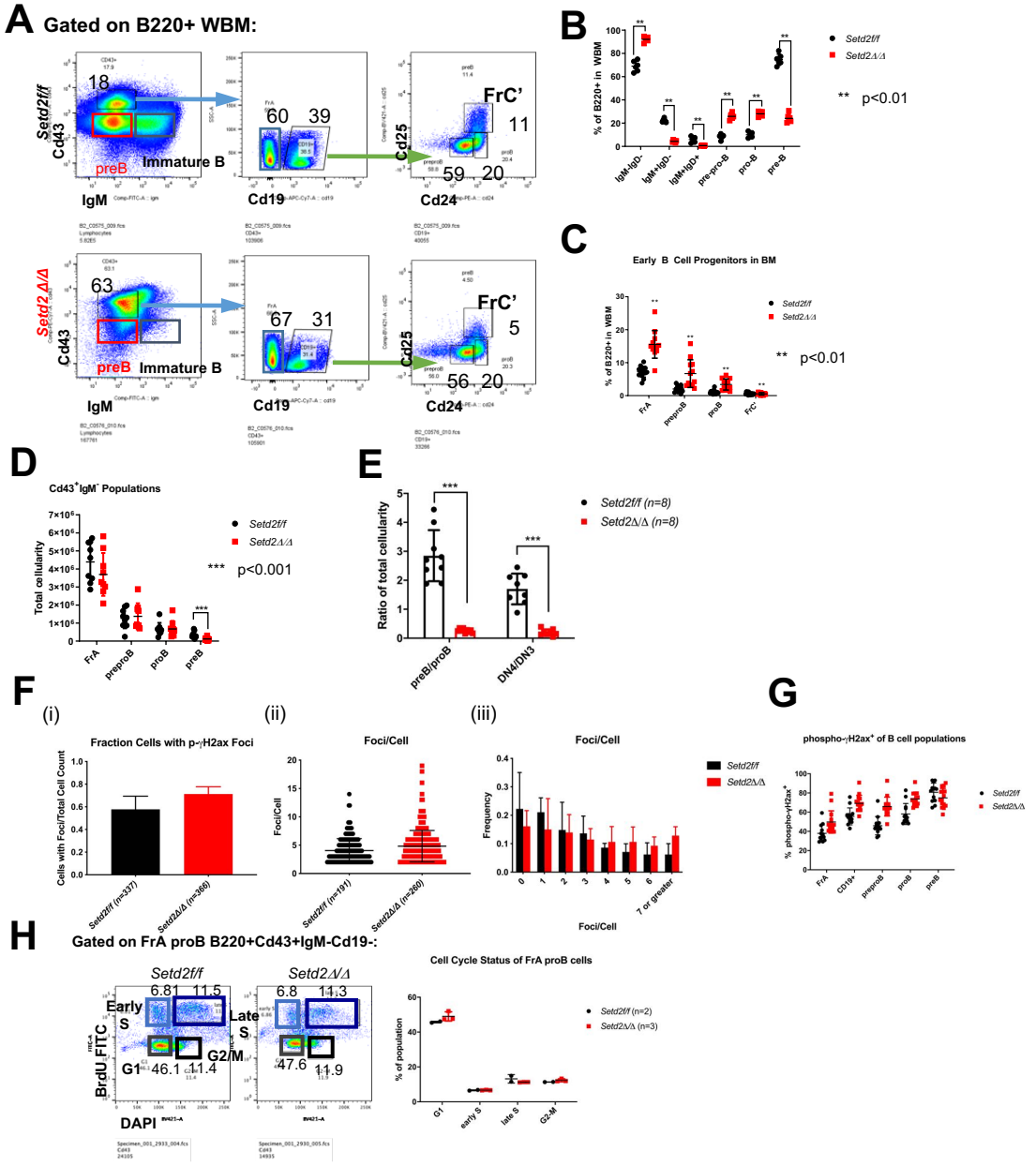


Figure S3. Loss of *Setd2*/H3K36me3 arrests development at a proB cell stage, Related to Figure 3. **(A)** Representative flow cytometry plots with gating strategy for different B cell progenitor populations for controls and *Setd2Δ/Δ* B220$^+$ BM cells. Fraction A (FrA): B220$^+$IgM$^-$Cd43$^+$Cd19$^-$Cd24^{lo}Cd25; pre-proB: B220$^+$IgM$^-$Cd43$^+$Cd19$^+$Cd24^{lo}Cd25$^-$, proB: B220$^+$IgM$^-$Cd43$^+$Cd19$^+$Cd24^{hi}Cd25$^-$; FrC': B220$^+$IgM$^-$Cd43$^+$Cd19$^+$Cd24^{hi}Cd25$^+$; preB: B220$^+$IgM$^-$Cd43$^-$; immature B: B220$^+$Cd43$^-$IgM$^+$. **(B)** Percent composition of B220$^+$ WBM of different B cell population in the bone marrow (n = 6). **(C)** B cell progenitor proportions in B220$^+$ BM of *Setd2Δ/Δ* and controls (n= 14 for all groups). **(D)** Total cellularity of B cell progenitor proportions in B220$^+$Cd43$^+$IgM$^-$ BM of *Setd2Δ/Δ* and controls. n= 9 for all groups. **(E)** Ratio of preB to proB compartment and DN4 to DN3 compartment total cellularity from BM of *Setd2Δ/Δ* and controls (n= 8 for all groups, from Fig. 3A-B) **(F)** Quantification of phospho-γH2ax foci by immunofluorescence from sorted FrA proB cells from *Setd2Δ/Δ* (n=8 mice) and controls (n=7 mice). (i) Fraction of total cells containing foci (n=366 for *Setd2Δ/Δ*, n= 337 for controls). (ii) Foci/cell in cells containing phospho-γH2ax foci (n=260 for *Setd2Δ/Δ*, n= 191 for controls). (iii) Frequency of foci/cell for all cells containing phospho-γH2ax foci for *Setd2Δ/Δ* and controls. **(G)** percent of different B cell populations positive for phospho-γH2ax by flow cytometry (n=15 for *Setd2Δ/Δ,* n=13 for controls). All groups were significant (p<0.01) except for preB population. **(H)** Representative flow cytometry plot of cell cycle status indicated by co-staining with DAPI and BrdU incorporation of sorted FrA proB cells from *Setd2Δ/Δ* (n=3) and controls (n= 2) and summary graph of each cell cycle stage. All values for cell cycle status were non-significant. **, p<0.01 ***, p<0.001, error bars represent SD.

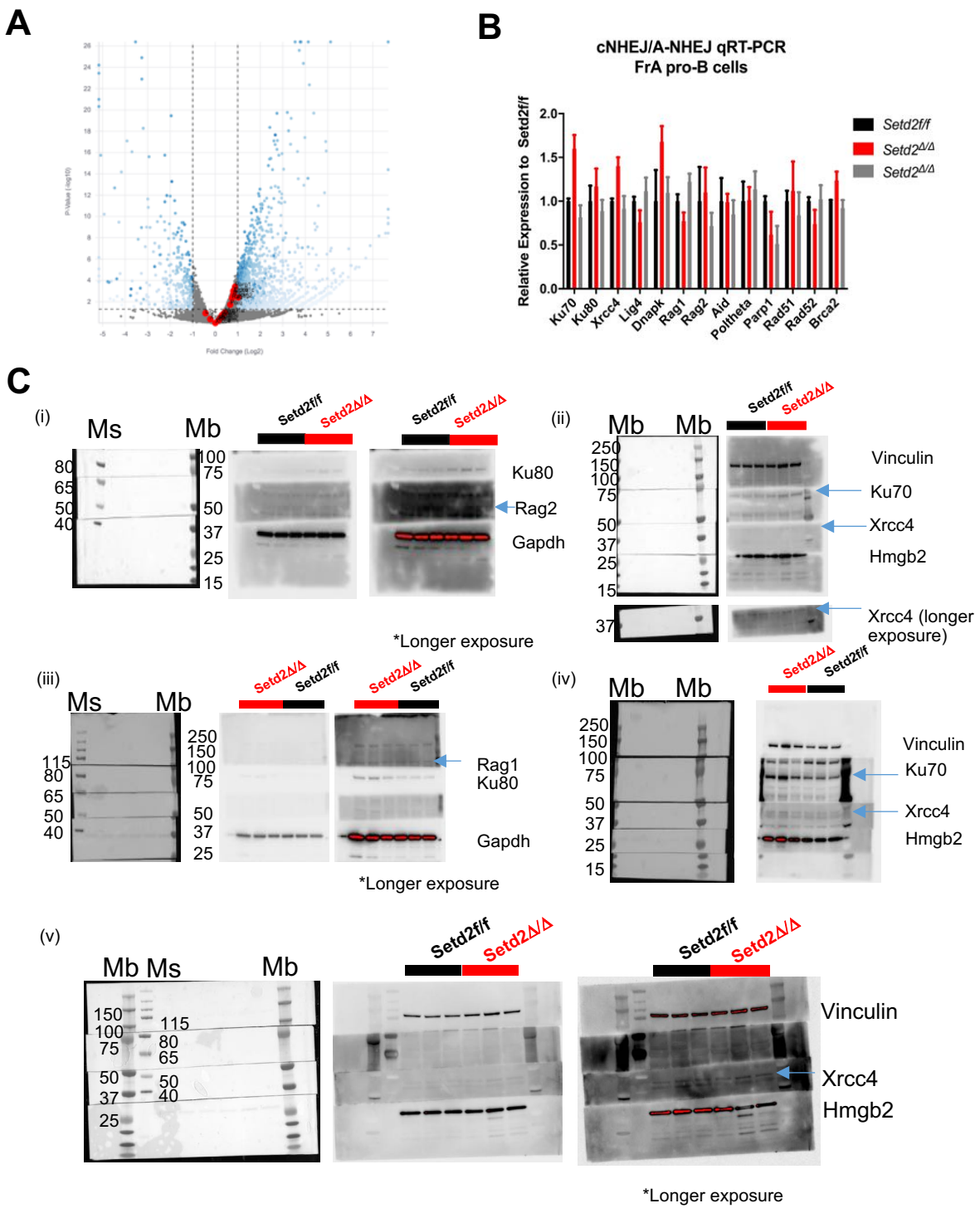


Figure S4. Loss of *Setd2*/H3K36me3 does not alter expression of C-NHEJ and A-NHEJ proteins, Related to Figure 3.

(A) Volcano plot of fold expression of a FrA proB cell compartment sorted from n=3 independent *Setd2Δ/Δ* and littermate control mice and subjected to by RNA-Sequencing. Genes highlighted include: *Rag1, Rag2, Xrcc4, Lig4, Xrcc5, Xrcc6, Xrcc7, Prkdc, Hmgb2, Mre11a, Dclre1c, Atm, Parp1, Polq, Brca2, Ctbp1, Rad52, Mrnip.* **(B)** Real-time PCR of C-NHEJ and A-NHEJ genes of FrA proB cells sorted from n=2 *Setd2Δ/Δ* mice and a littermate control. Expression was normalized to *Gapdh* expression and calculated relative to the wildtype control by a ΔΔCt method. **(C)** (i-v) Immunoblotting of FrA proB cells sorted from n=3 *Setd2Δ/Δ* mice and n=3 littermate controls (*Setd2f/f*) for C-NHEJ proteins Ku70, Ku80, Xrcc4, Rag1, Rag2, Hmgb2 with loading controls Vinculin and Gapdh. Mb – low molecular weight ladder, Ms – high molecular weight ladder.

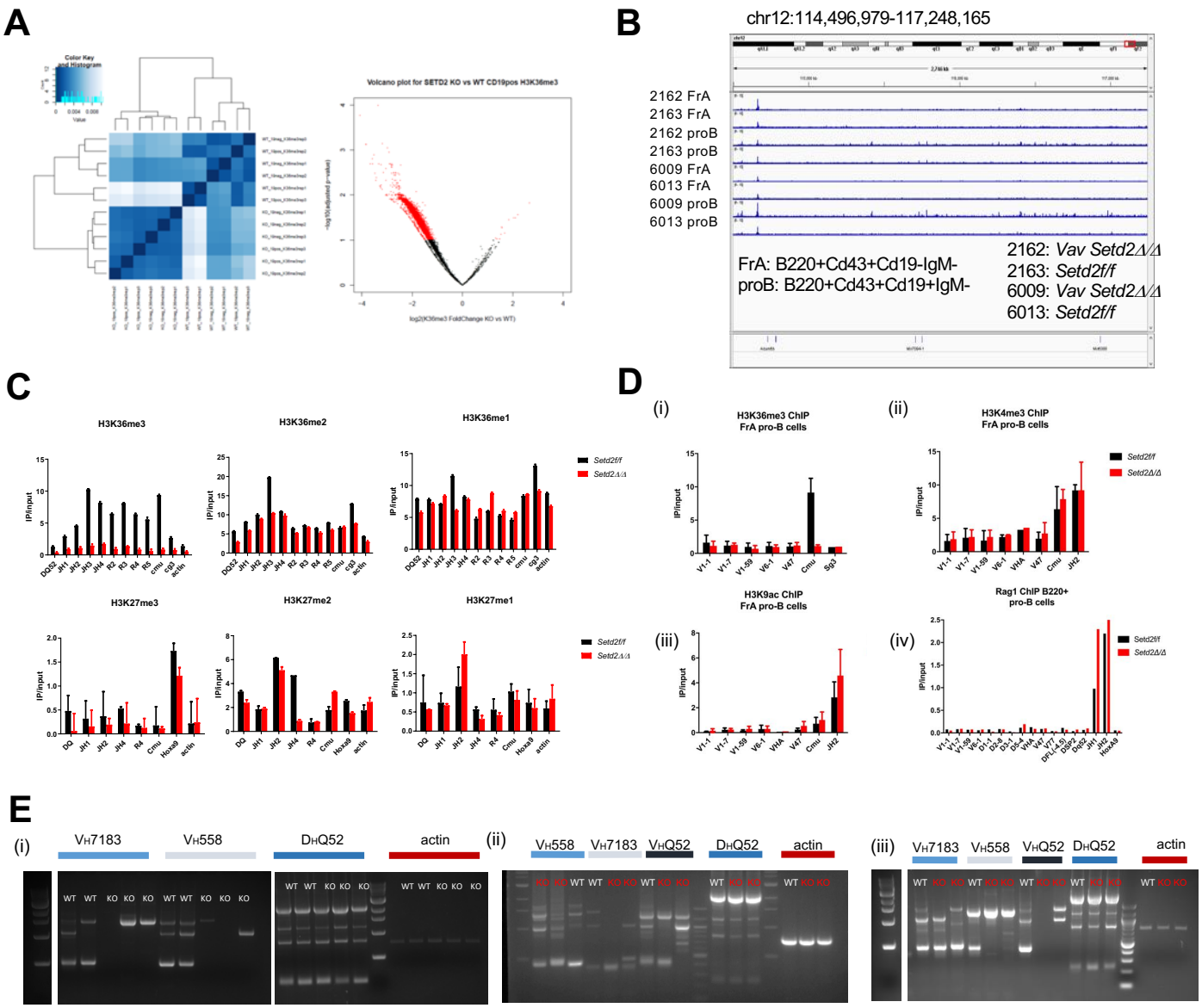


Figure S5. Loss of H3K36me3 does alter local chromatin architecture or accessibility of the *Igh* locus but leads to aberrant V(D)J recombination, Related to Figure 4.

(A) Composite Heatmap and volcano plot of H3K36me3 ChIP-Seq signal of Cd19$^-$ and Cd19$^+$ control (WT) and *Setd2Δ/Δ* (KO) proB cells (B220$^+$Cd43$^+$IgM$^-$) (n=3 for each genotype). **(B)** ATAC-sequencing tracks from Cd19$^-$ and Cd19$^+$ control (WT) and *Setd2Δ/Δ* (KO) proB cells (B220$^+$Cd43$^+$IgM$^-$) (n=2 for each genotype) across the Igh locus. **(C)** Representative ChIP-PCR of relative abundance of H3K36me1-3 and H3K27me1-3 at regulatory region of Igh locus. Regions R1-R5 were regions with H3K36me signal in between the Eμ enhancer binding site and Cμ gene region with HoxA9, Cγ3 and actin as controls. Data is representative of n = 2 independent experiments. **(D)** (i-iii) H3K36me3, H3K4me3, and H3K9ac ChIP-PCR of variable heavy chain gene families and from n=4 independent ChIP experiments and (iv) Rag1 occupancy by ChIP-PCR variable gene family genes with primers from (Ji et al 2019 and (Ji et al., 2010; Hauser et al., 2014; Subrahmanyam et al., 2012; Chakraborty et al., 2009; Hesslein, et al, 2003). (n=1 ChIP). **(E)** (i-iii) replicates of PCR assay to detect V(D)J recombination products of rearrangement shown in Figure 2E of the Igh locus of different V_H families from independently sorted proB cells from *Setd2Δ/Δ* (KO) and controls (WT). No product meant germline non-rearrangement.

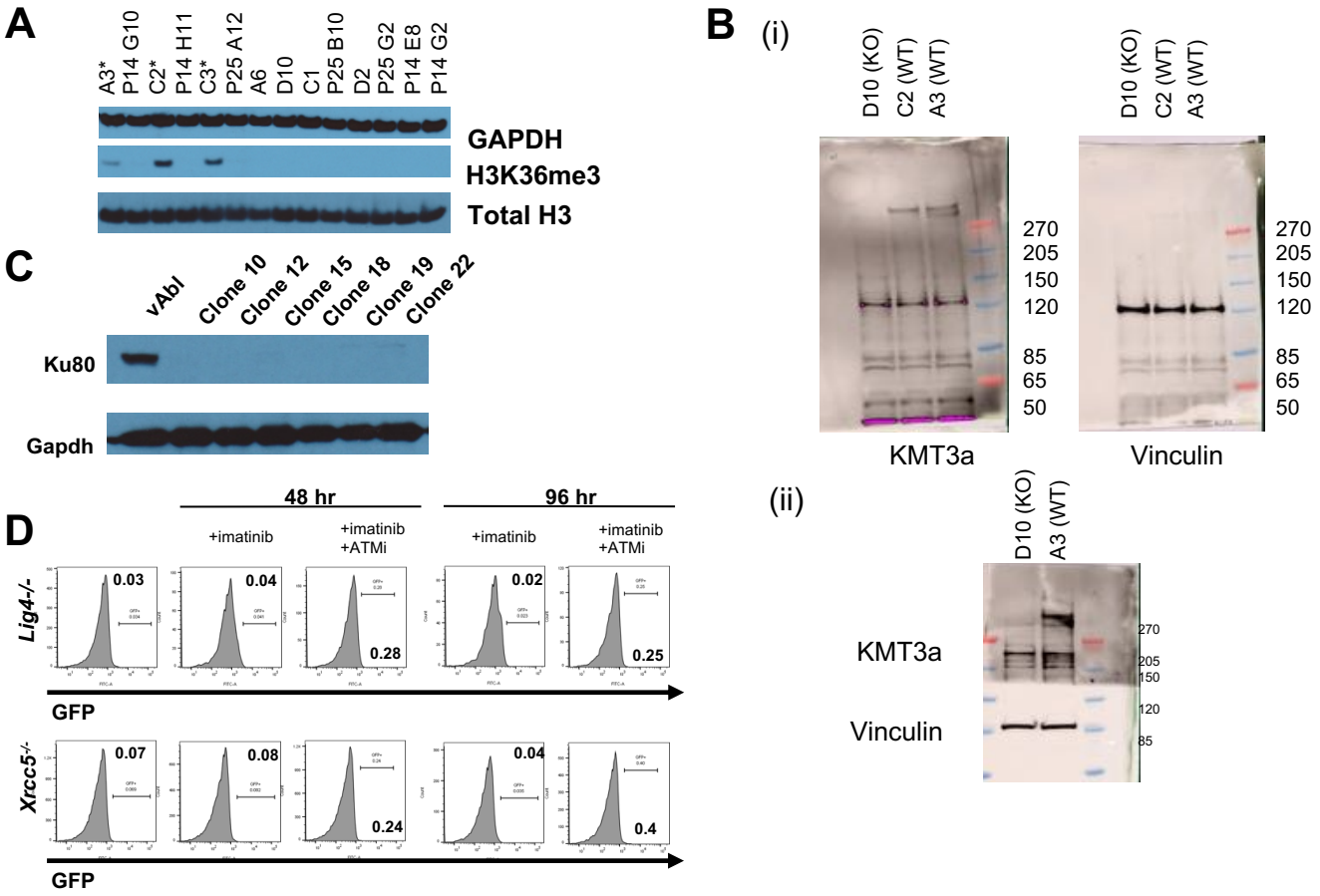
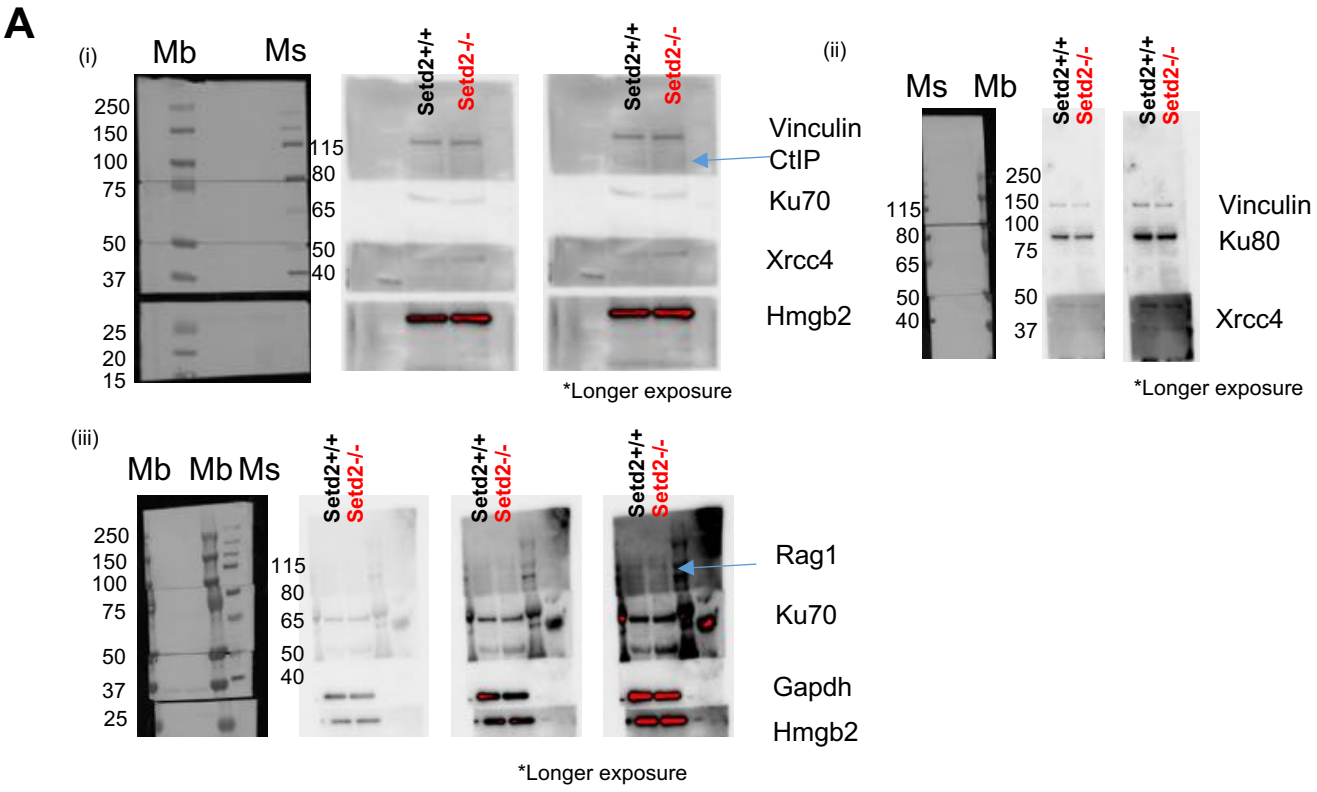


Figure S6. Generation of v-Abl transformed *Setd2, Xrcc5* and *Lig4* deficient cell lines, Related to Figure 5.

(A) H3K36me3 levels of several individual v-Abl clones (used in Figure 5) that were either *Setd2* wildtype (wt, indicated with asterisk) or *Setd2* deficient lines (*Setd2*$^{-/-}$ or KO). Total H3 and Gapdh were used as loading controls. **(B)** (i-ii) immunoblots of *Setd2* wildtype or *Setd2*$^{-/-}$ v-Abl lines for Kmt3a (Setd2) and Vinculin loading controls. **(C)** Immunoblot of v-Abl lines for *Xrcc5* (*Ku80*) null lines. Gapdh was used as a loading control. **(D)** Representative flow cytometric analysis of GFP expression in *Lig4* and *Xrcc5*$^{-/-}$ pMG-INV v-Abl cells treated with Abl kinase inhibitor imatinib (STI-571) and ATM kinase inhibitor (ATMi, KU55933) for 48 and 96 hours. Representative of 4 independent induction experiments.



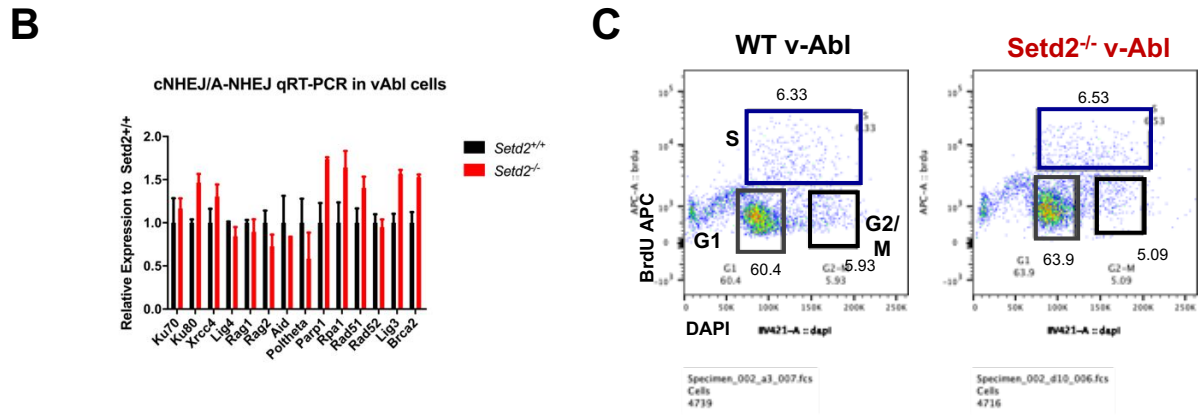


Figure S7. Deletion of *Setd2* in v-Abl cells does not alter expression of C-NHEJ and A-NHEJ proteins or cell cycle status, Related to Figure 5.

(A) (i-iii) Immunoblotting of WT and *Setd2*$^{-/-}$ v-Abl clones for C-NHEJ proteins Ku70, Ku80, Xrcc4, Rag1, Rag2, Hmgb2 with loading controls Vinculin and Gapdh. Mb – low molecular weight ladder, Ms – high molecular weight ladder. **(B)** Real-time PCR of C-NHEJ and A-NHEJ genes of *Setd2* deficient and WT v-Abl transformed cells. Expression was normalized to *Gapdh* expression and calculated relative to the wildtype control by a ΔΔCt method. **(C)** Representative flow cytometry plot of cell cycle status indicated by co-staining with DAPI and BrdU incorporation of *Setd2* deficient and WT v-Abl cells. All data representative of n=2 independent experiments. Error bars represent SD.

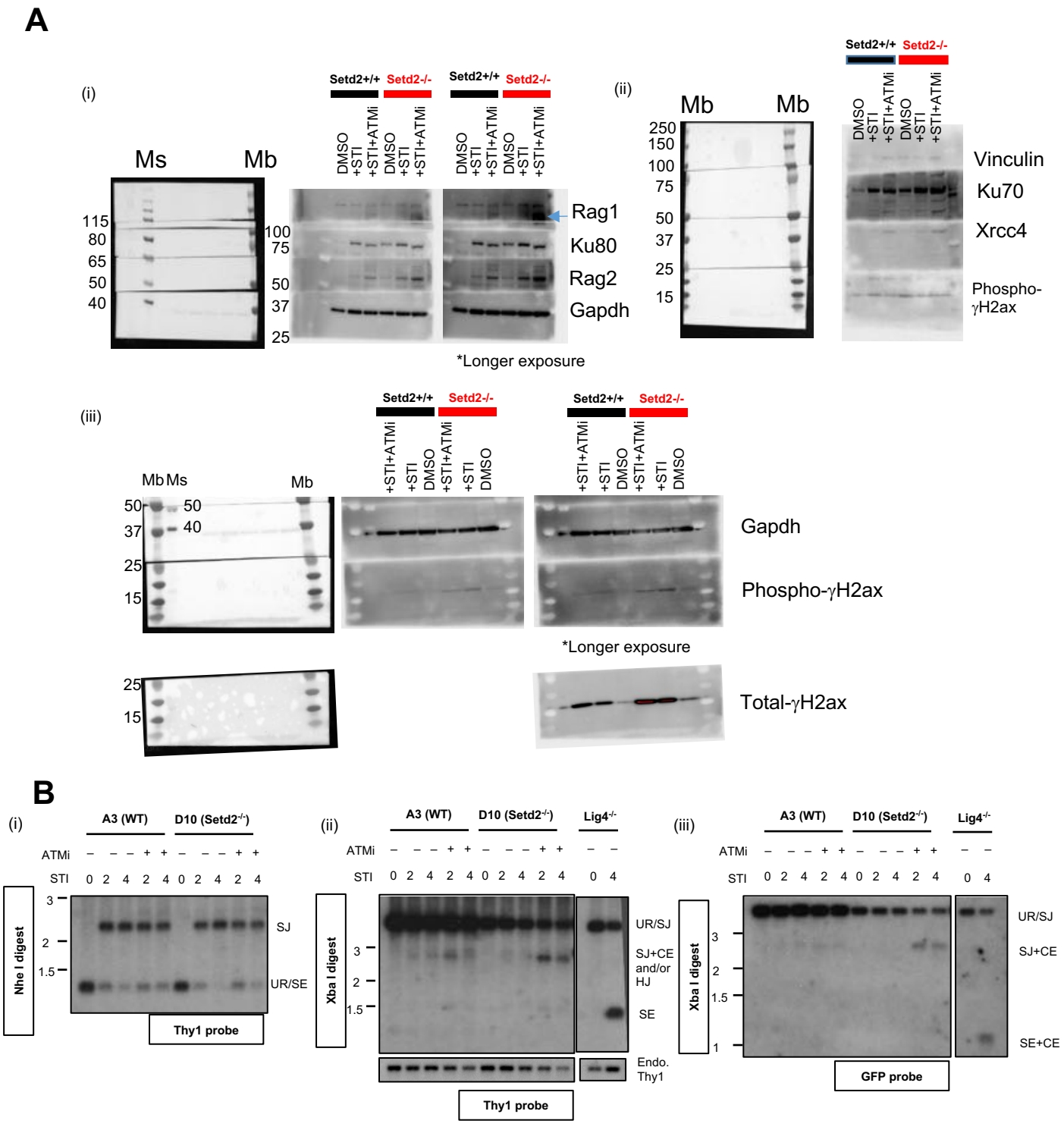


Figure S8. Dual loss of *Setd2*/H3K36me3 and Atm activity leads to abnormal recombination and end-joining repair, Related to Figure 5.

(A) (i-iii) Immunoblotting for NHEJ proteins and phospho- and total γH2ax of WT (*Setd2*$^{+/+}$) and *Setd2* deficient (*Setd2*$^{-/-}$) v-Abl cells treated for 48 hours with STI or STI+ATMi. Data representative of 3 independent induction experiments. Mb – low molecular weight ladder, Ms – high molecular weight ladder. **(B)** Replicate Southern blot analysis of genomic DNA from *Setd2*$^{-/-}$ clone (D10) and control line that were digested with NheI (left) and XbaI (middle and right panels) and hybridized with Thy1 probe (middle, right panels) or GFP probe (left panel). Bands indicated are for URs, SJs, CJs, SEs, CJ+SEs or SJ+CEs. Hybrid Joins (HJ) indicated as well. Southern blots were generated from an independent induction experiment of v-Abl clones with imatinib and ATMi shown in Figure 5.

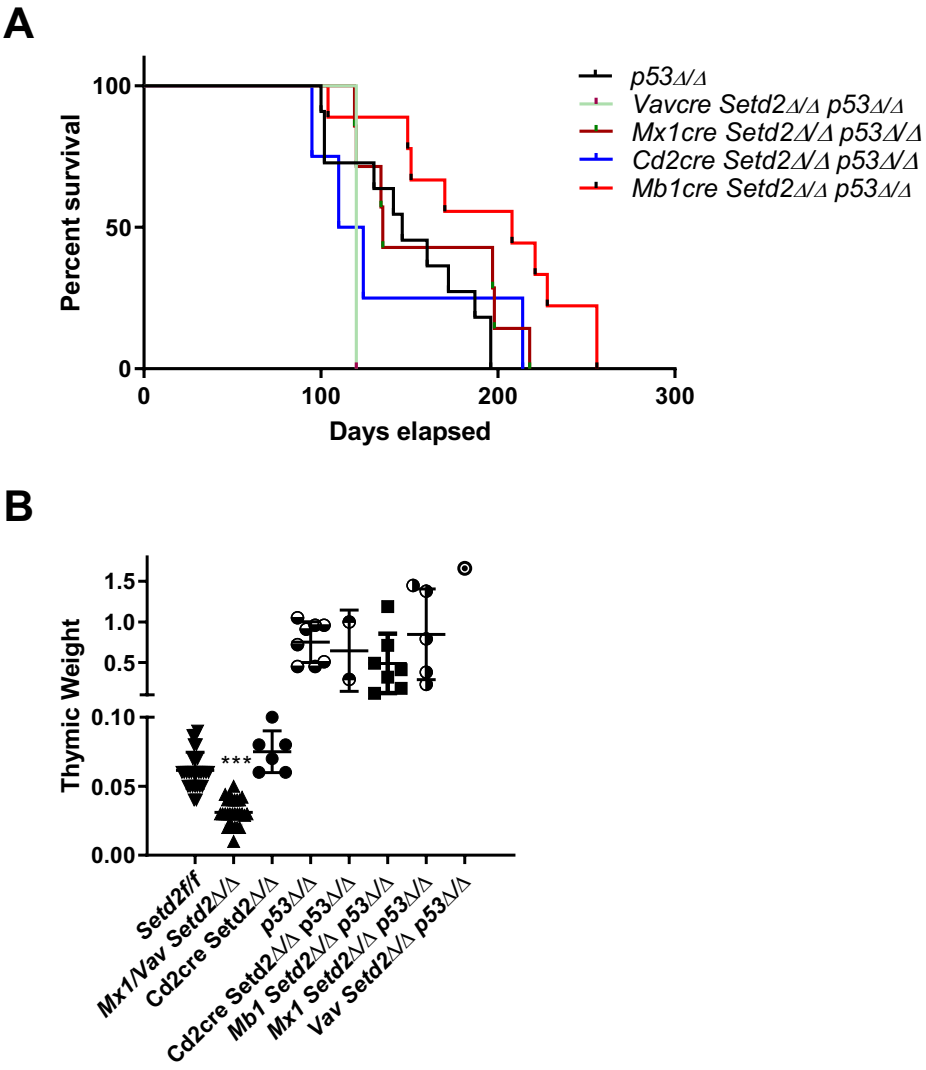


Figure S9. Dual loss of *Setd2*/H3K36me3 and p53 does not lead to the development of pro-B lymphomas, Related to Figure 5.

(A) Kaplan-Meier survival of double knockout *Vav* (n= 1), *Mx1* (n=7), *Cd2* (n=4) , *Mb1cre* (n=9) *Setd2Δ/Δ p53Δ/Δ* mice and controls (n = 11). All mice developed thymic lymphomas. **(B)** Weight of thymus in mouse thymic lymphomas from double knockout mice and controls. ***, $p<0.001$ calculated to control (*Setd2f/f* $p53^{+/+}$). Error bars represent SD.

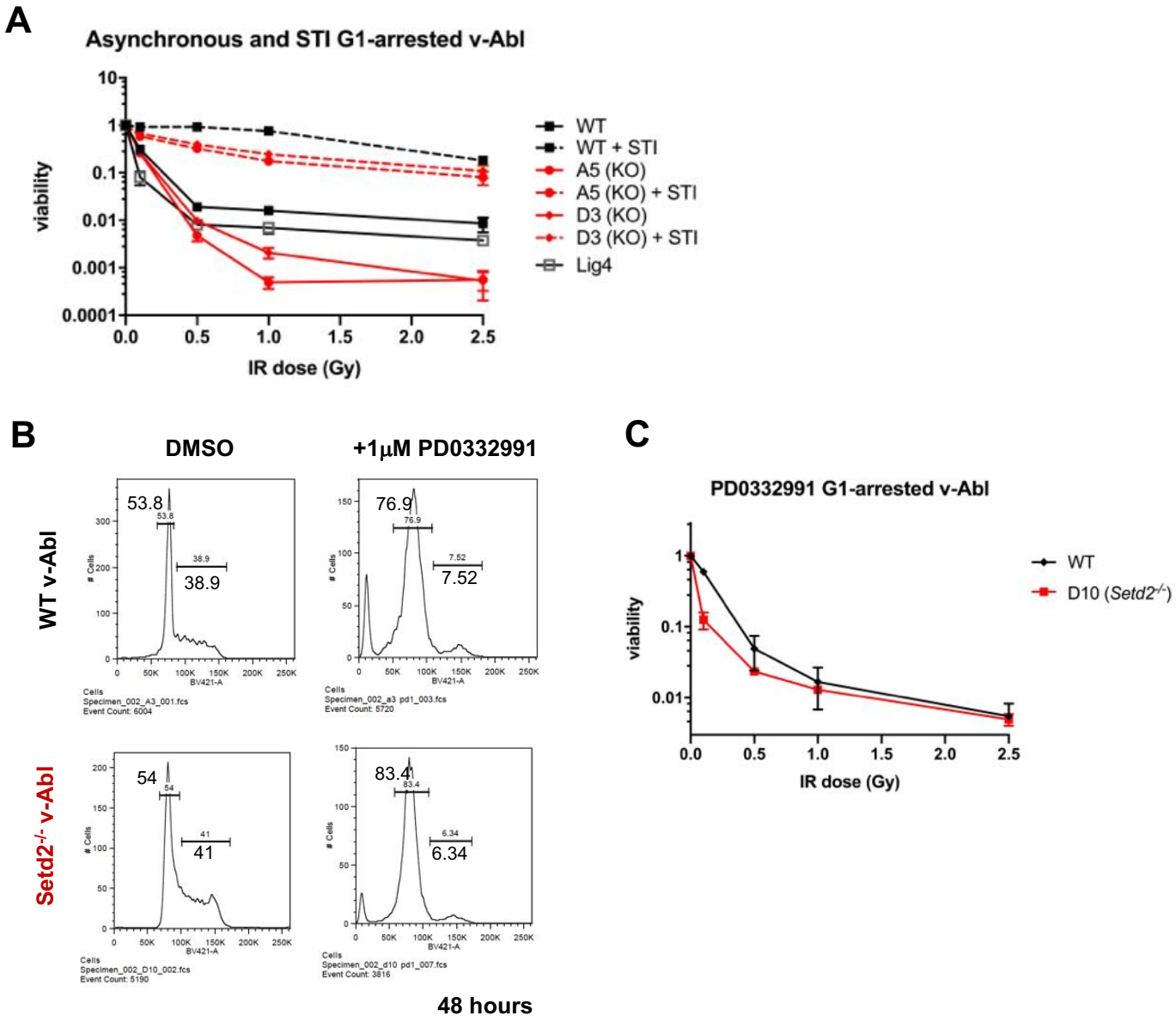


Figure S10. Asynchronous and synchronous *Setd2* deficient v-Abl cells exhibit increased sensitivity to ionizing radiation, Related to Figure 7.

(A) Asynchronous and STI G1-arrested WT, *Lig4*$^{-/-}$ and 2 different *Setd2*$^{-/-}$ v-Abl cells subjected to different doses of ionizing radiation and serially diluted 5 times in triplicate to assess survival. Values plotted as a mean of each dose as a percent of non-irradiated controls, bars represent standard deviation. Data is representative for 3 different independent treatment experiments. Viability was assessed at 72 hours. **(B)** Representative flow cytometric analysis of cell cycle status as determined by DNA content with DAPI staining 48 hours after treatment with 1mM Cdk4/6 inhibitor PD0332991. Data representative of n=3 independent experiments. **(C)** PD0332991 G1-arrested WT *Setd2*$^{-/-}$ v-Abl cells subjected to different doses of ionizing radiation and serially diluted 5 times in triplicate to assess survival. Values plotted as a mean of each dose as a percent of non-irradiated controls, bars represent standard deviation. Data is representative for 3 different independent treatment experiments. Viability was assessed at 72 hours.

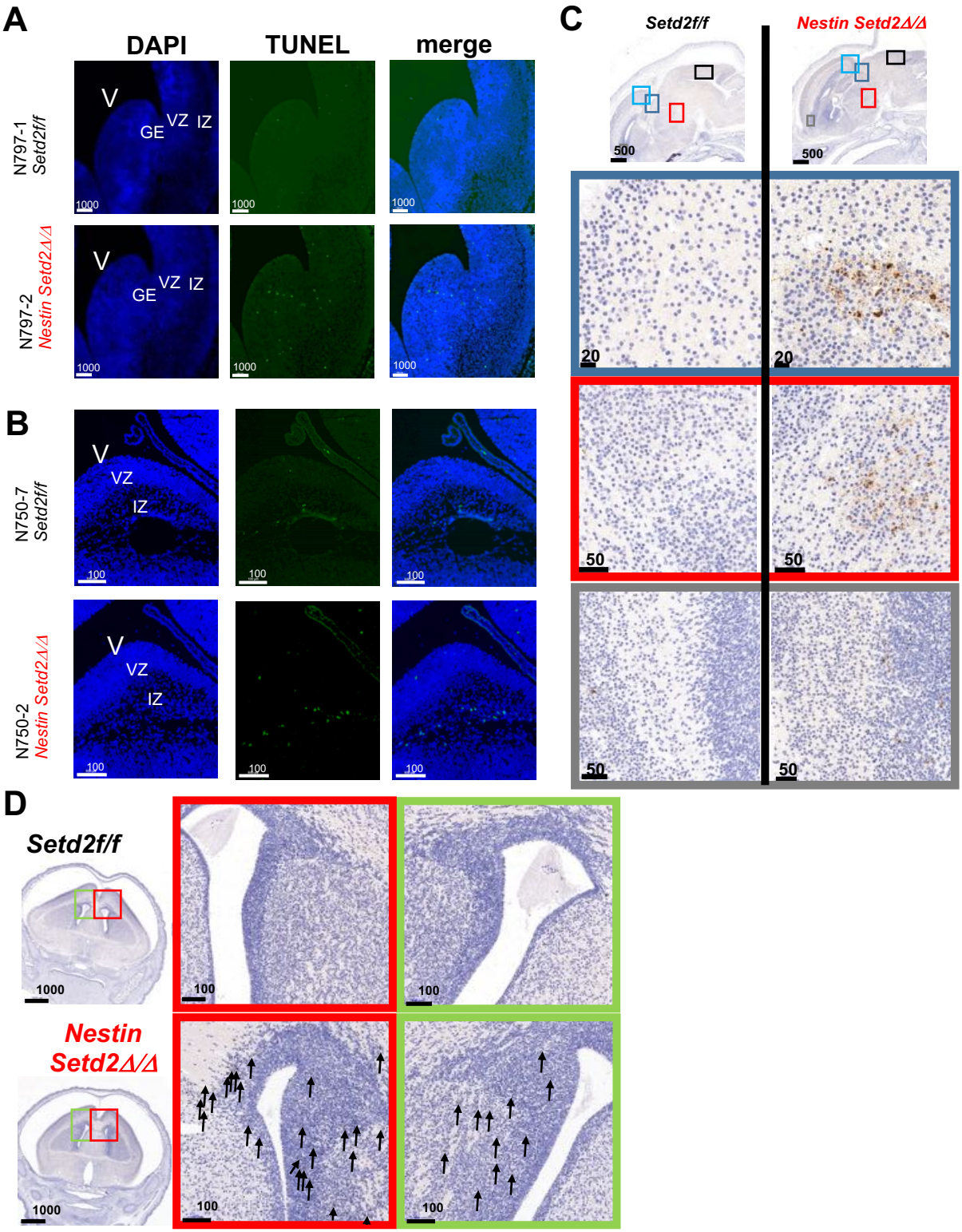


Figure S11. Post-mitotic neuronal apoptosis with loss of *Setd2*/H3K36me3, Related to Figure 7. **(A,B)** TUNEL assay of two independently matched control and *Nestin-cre Setd2Δ/Δ* E14.5 coronal sections of the lateral ventricle displaying the ganglionic eminence (GE), Lateral Ventricle (V), Ventricular zone and intermediate zone (IZ) of the telencephalon. Distances as indicated in microns. **(C)** Higher magnification of cleaved Caspase-3 immunohistochemistry of E18.5 sagittal regions of lateral ventricle of control and *Nestin-cre Setd2Δ/Δ* embryos from Figure 7D. Distances as indicated in microns. **(D)** Cleaved caspase-3 immunohistochemistry of coronal sections of the lateral ventricle of 2 hour post-partum control and *Nestin-cre Setd2Δ/Δ* pups from same samples as shown in Figure 7E. Magnifications indicated by corresponding red and green boxes. Arrows indicate cleaved caspase-3 staining and pyknotic nuclei. Distances as indicated in microns.

Transparent Methods

Mice and Isolation of primary B cells and v-Abl B cells

Strategy for generating *Setd2Δ/Δ* mice was previously described (Mar et al., 2018). *Setd2f/f* mice were bred to *Mx1, Vav1, Cd19, Mb1, hCD2-cre* mice to homozygosity. At 6-8 weeks of age, mice were sacrificed and bone marrow from femurs, tibias, hips, and spines were isolated. BM was then RBC lysed and stained with B220-biotin (BD Pharmingen) and subsequently stained with Anti-biotin microbeads (Miltenyi) and applied on a LS column (Miltenyi). Subsequent staining for cellular antigens was conducted. For V(D)J recombination product assessment, RNA-Seq, ChIP-Seq, ChIP-PCR, and qPCR were stained and sorted by FACS (BD-Aria) for $B220^{+}Cd43^{+}IgM^{-}Cd25^{-}Cd19^{-}$ proB cells and $B220^{+}Cd43^{-}IgM^{-}$ for preB. IgHelMD4 (Goodnow et al., 1988) and *Mb1cre* mice were generously provided by Jayanta Chaudhuri. All mouse experiments were approved by the Institutional Animal Care and Use Committees at Dana-Farber Cancer Institute and Memorial Sloan Kettering Cancer Center.

WT and *Lig4*$^{-/-}$ v-Abl B cell lines were generously provided by Andre Nussenzweig and Barry Sleckman (Bredemeyer et al., 2006). *Xrcc5* and *Setd2* knockout v-Abl cells were generated as before (Bredemeyer et al., 2006; Jacobsen et al., 2006). Briefly, CMV expression vectors containing Cas9 and a BFP tagged expression vector (Hung et al., 2018) with sgRNAs targeting gene of interest were nucleofected with Amaxa nucleofection Kit P4 as per manufacturer's instructions (Lonza). Less than 24 hours later, BFP^{+} cells were single cell sorted into 96 well plates and expanded and western blotting was conducted to confirm protein knockout. Single cell clones with confirmed knockout of protein were transduced with retrovirus containing pMG-INV (Hung et al., 2018) vector and subsequently sorted for expression of cell surface marker Thy1.2. The recombination substrate vector pMG-INV was generously provided by Barry Sleckman. To induce G1 arrest, cells were treated with 3μM imatinib (STI571) with or without 15 μM ATMi (KU55933) for up to 96 hours and assayed for GFP expression by flow cytometry and genomic DNA isolation for downstream PCR analysis. sgRNAs for CRISPR/Cas9 used were as follows: Xrcc5: sgXrcc5-1 (5'-GAATGATATCACTTCCGTAG-3'), sgXrcc5-2 (5'-GAGCTTGGTAAAGAAAAACG-3'), sgXrcc5-3 (5'-GTCATAAGCATATCGGACGA-3'), sgXrcc5-4 (5'-TGTCCTTAGAAGGCGAAGAC-3'), sgXrcc5-5 (5'-TGCACACAATCAGGGCGTCC-3'). Setd2: sgSetd2-1 (5'-GCATTCGCTTAATATCCCGG-3'), sgSetd2-2 (5'-GGAGTTCCCCTTATCGTGAG-3'), sgSetd2-3 (5'-TTGCTTATGATCGAATCCAA-3'), sgSetd2-3 (5'-TGCTCATGCTCAGAGTGACG-3'), sgSetd2-5 (5'-ATAATAGGGAGCCGACAGAC-3').

RNA-sequencing of LSK and proB cells

B cells were isolated as indicated above. LSKs were obtained by lineage depletion of WBM and conducted as per manufacturer's protocols with biotin-labeled antibodies for CD3, Gr1, Ter119, and B220 (BD Pharmingen) and subsequently subjected to magnetic depletion with anti-biotin microbeads and depletion on a LD column (Miltenyi). Lineage depleted cells were then stained with cKit and Sca1 and sorted on a FACS cell sorter (BD-SORP-AriaII). Qiagen RNA kits were used as per manufacturer's protocol for RNA isolation and purity was confirmed with RNA Tapestation (Illumina). RNA-seq libraries

were prepared with NEBNext UltraKits and for proB cells. For LSKs, RNA was amplified with SMARTer Ultra Low Input RNA Kit for Illumina Sequencing.

Hematopoietic Reconstitution assays
For competitive/non-competitive reconstitution assays, $1x10^6$ total, unfractionated whole BM cells of control and fully excised *Setd2* mice were injected into lethally irradiated mice (9 cGy) and bled every 4 weeks up to 16 weeks before these mice were sacrificed and hematopoietic reconstitution was assessed in the bone marrow. Non-excised *Mx1cre Setd2Δ/Δ* mice were treated with pI:pC after 8 weeks of engraftment in reconstitution assays.

BrdU incorporation assays
For *in vivo* assessment of BrdU incorporation, BrdU was injected intraperitoneally 4 hours before mice sacrificed and BrdU incorporation was assessed as per manufacturer's protocols (BD Biosciences). Recommended manufacturer's protocols were also followed for *in vitro* labelling of v-Abl cells. Cells were labelled for 2 hours before fixed and permeabilized.

Phospho-γH2ax Immunofluorescence

Cytospins of 20 000 sorted FrA proB cells were prepared at 500 rpm for 5 minutes, permeabilized and fixed in 4% paraformaldehyde/PBS at 4°C for 10 minutes. Phospho-γH2AX (Abcam) primary antibody was used at 1:500 and incubated at 4°C overnight. Secondary antibody Anti-rabbit Alexa647 (CST) was applied for 30 minutes at room temperature. DAPI (1 µg/ml) counterstaining was conducted for 5 minutes at room temperature and covered with Prolong Gold Antifade Reagent (ThermoFisher). Slides were analyzed by confocal microscopy (Leica TCS SP5) (Leica) and foci were quantified using Image J software (NIH).

Ionizing Radiation Sensitivity Assay
Ionizing radiation sensitivity assays were performed for v-Abl lines that were either asynchronous or G1 arrested with STI-571 (Selleckchem) or PD-0332991 (Sigma-Aldrich) before irradiation. Cells were treated with 0.1, 0.5, 1, or 2.5 Gy ionizing radiation with percent survival measured relative to a non-irradiated and non-treated control for each cell line assayed. Cells that were G1 arrested were treated 48 hours prior to irradiation before washing off medium containing 3μM STI571 or 1μM PD-0332991 and re-plated with fresh media. 50,000 cells were initially plated in 96 non-tissue culture treated plates and 1:1 serially diluted 5 times with 3 replicates for each condition and cultured at for 3 days. Viability was measured by staining with DAPI and measured by flow cytometry on BD Fortessa in HTS mode.

Genomic DNA isolation
Up to $5x10^6$ proB cells were harvested and genomic DNA extracted using PureLink Genomic DNA kit (Invitrogen) as per manufacturer's protocols.

PCR analyses

PCR of retroviral substrate coding joints (CJs) and hybrid joints (HJs): pMG-INV was generously provided by Barry Sleckman (Hung et al., 2018). Oligonucleotides CJ_F and CJ_HJ_R were used to amplify CJs in pMG-INV. Oligonucleotides HJ_F and CJ_HJ_R were used to amplify HJs. CJ_F: (5'-TCAGCCAGAAATTCAGTGGCA-3'); HJ_F: (5'-TTGTACACCCTAAGCCTCCG-3'); CJ_HJ_R: (5'-GCTTATCGATACCGTCGACCT-3'). All PCRs were done on genomic DNA from cells that had been treated with STI571 for 96 hours. The *Il2* gene was amplified using the IMR42 and IMR43 oligonucleotides (Bredemeyer et al., 2006). PCRs for the *Il2* gene, and all retroviral HJs and CJs, were carried out in 50µL with cycling conditions of 95ºC for 2 minutes followed by 30 cycles 94ºC 30s, 55ºC 30s, 72ºC 60s (Bredemeyer et al., 2006). Murine *Il-2*: forward (5'-CTAGGCCACAGAATTGAAAGATCT-3'); reverse (5'-GTAGGTGGAAATTCTAGCATGATGC-3')

PCR analyses of endogenous receptor gene rearrangements: PCR of Vκ6-23 HJs and CJs was carried out by amplifying 0.5µg of genomic DNA from B220 enriched splenocytes from *Setd2Δ/Δ* and *Setd2f/f* controls in 50µl with primers for amplification as follows: pκJa and pκ6a for HJ and pκJa and pκ6d for CJ. PCR conditions were 95ºC for 5 minutes followed by 17 cycles 94ºC 30s, 64ºC 30s, 72ºC 30s. Products from this reaction were amplified in 50µl using the above conditions with primers pκJa and pκ6b for HJ and pκJa and pκ6c for CJ and 25 amplification cycles oligonucleotides (Bredemeyer et al., 2006). pκ6a: (5'-TGCATGTCAGAGGGCACAACTG-3'); pκ6d: (5'-GAAATACATCAGACCAGCATGG-3'); pκ6b: (5'-CTACCAAACTTTGCAACACACAGGC-3'); pκ6c: (5'-ACATGTTGCTGTGGTTGTCTGGTG-3'); pκJa: (5'-GGAGAGTGCCAGAATCTGGTTTCAG-3').
PCR results were analyzed with high sensitivity D1000 TapeStation reagents (Agilent Technologies).

V(D)J recombination PCR: Assay was conducted as previously described (ten Boekel et al., 1995; Ehlich et al., 1994; Corcoran et al., 1998). Briefly, two rounds of PCR were conducted on 100-300ng of gDNA using forward primers amplifying V_H558, V_HQ52, V_H7183 family genes with nested primers located in the J_H4 gene segment. Products were visualized on a 1% agarose gel and the V-(D)J_H4 recombination product was gel extracted (Qiagen) and submitted for NGS sequencing. Vh7183_F1: (5'-CTCGCCATGGACTTCGGGTCAGTTGG-3'); Vh7183_F2: (5'-CAGCTGGTGGAGTCTGGGGGAGGC-3'): Vh558_F1: (5'-ACCATGGGATGGAGCTGKATCWTBC-3'): Vh558_F2: (5'-GTGARGCCTGGGRCTTCAGTGAAG-3'); VhQ52F: (5'-GCGAAGCTTCTCACAGAGCCTGTCCATCAC-3'): Vh7183: (5'-GCGAAGCTTGTGGAGTCTGGGGGAGGCTTA-3'); DhQ52_F1: (5'-CACAGAGAATTCTCCATAGTTGATAGCTCAG-3'); DhQ52I_F2: (5'-GCCTCAGAATTCCTGTGGTCTCTGACTGGT-3'); Jh4_R1: (5'-AGGCTCTGAGATCCCTAGACAG-3'); Jh4_R2: (5'-GGGTCTAGACTCTCAGCCGGCTCCCTCAGGG-3'); actin_738: forward (5'-GGTGTCATGGTAGGTATGGGT-3'), reverse (5'-CGCACAATCTCACGTTCAG-3').

Chromatin Immunoprecipitation (ChIP) and ChIP-sequencing

Chromatin immunoprecipitation was coupled with high-throughput sequencing (ChIP-seq). 0.5-2x10^6 primary sorted proB cells were crosslinked with 1% formaldehyde for 10 minutes followed by 0.125 M glycine for 5 minutes. Fixed cells were washed twice with ice-cold phosphate-buffered saline and resuspended in ChIP lysis buffer and sheared using a Covaris E220 ultrasonicator (Covaris). Sheared chromatin was incubated overnight at 4^{o}C with rabbit polyclonal anti-H3K36me3 (61101 pAb, Active Motif). Immune complexes were collected with protein A/G dynabeads (Invitrogen) and washed sequentially in low-salt wash buffer (20 mM Tris pH 8.0, 150 mM NaCl, 0.1% SDS, 1% Triton X-100, 2mM EDTA), high-salt wash buffer (20 mM Tris pH 8.0, 500 mM NaCl, 0.1% SDS, 1% Triton X-100, 2mM EDTA), LiCl wash buffer (10 mM Tris pH 8.0, 250 mM LiCl, 1% NP-40, 1% sodium deoxycholate, 1 mM EDTA), and TE. Chromatin was eluted buffer (1% SDS, 0.1 M NaHCO3), and then reverse cross-linked with 0.2 M NaCl at 65o C overnight. DNA was purified with a PCR purification kit (QIAGEN) and subjected to quantitative PCR or processed for ChIP-sequencing.

ChIP-PCR oligonucleotides used in this study were as follows (Ji et al., 2010; Hauser et al., 2014; Subrahmanyam et al., 2012; Chakraborty et al., 2009; Hesslein, et al, 2003):

Vha: forward (5'-CCTTCGCCCCAATCCACC-3'), reverse (5'-CAAGTAACCCTCAAGAGAATGGAGACTC-3'); Vh47: forward (5'-CTACAACCAGAAGTTCAAGGGCAA-3'), reverse (TCAGGCTGTGATTACAACACTGTGT-3'); Vh77: forward (5'-AAATCCTCCAGCACAGCCTA-3'), reverse (5'-TAGACCGCAGAGTCCTCAGA-3'); Dsp2: forward (5'- CAACAAAAACCCAGTATGCCCAG-3'), reverse (5'-GTGCTTTCACCTGTCTGTGGG-3'); Dfl-4.5: forward (5'-AGGCATCTCATCTCACTCTAAGC -3'), reverse (5'-TGTGTCCCTCTAAGACGAGTGAAT -3'); Dq52: forward (5'-CATTGGTCCCTGACTCAAGA-3'), reverse (5'-TCCCAGTTAGCACTGTGGTG-3'); Jh1: forward (5'-TGCTACTGGTACTTCGATGTCTG-3'), reverse (5'-GCCAGCTTACCTGAGGAGAC-3'); Jh2: forward(5'-CAGTCTCCTCAGGTGAGTCCT), reverse (5'-CCCAATGACCCTTTCTGACT-3'); Jh3: forward (5'-GCCTGGTTTGCTTACTGG-3'), reverse (5'-GACAAAGGGGTTGAATCT-3'); Jh4: forward (5'-CACCAGGAATTGGCATAA-3'), reverse (5'-CCTGAGGAGACGGTGACT-3'); Sμ: forward (5'-GCTAAACTGAGGTGATTACTCTGAGGT-3'), reverse (5'-GTTTAGCTTAGCGGCCCAGCTCATTCC-3'); Cμ: forward (5'-ATGTCTTCCCCCTCGTCTCC-3'), reverse (5'-TACTTGCCCCCTGTCCTCAG-3'); Sγ3: (5'-AATCTACAGAGAGCCAGGTGG-3'), reverse (5'-TGGTTTTCCATGTTCCCACTT-3');

Cγ: forward (5'-TGGACAAACAGAAGTAGACATGGGTC-3'), reverse (5'-GGGGTTTAGAGGAGAGAAGGCAC-3');

γ-actin: forward (5'-GACACCCAACCCCGTGACG-3'), reverse (5'-GCGGCCATCACATCCCAG-3');

IgHK36me3R1: forward (5'-TGGTTTCGGAGAGGTCCAGA-3'), reverse (5'-GTAGGCCTGGACTTTGGGTC-3'); IgHK36me3R2: forward (5'-CAAGCCCAGCTTTGCTTACC-3'), reverse (5'-CTGAGATGGGTGGGCTTCTC-3'); IgHK36me3R3: forward (5'-AGGGCTCTCAACCTTGTTCC-3'), reverse (5'-AGGTCGGCTGGACTAACTCT-3'); IgHK36me3R4: forward (5'-

TCTGGCTTACCATTTGCGGT-3'), reverse (5'-TCGGTGGCTTTGAAGGAACA-3'); IgHK36me3R5: forward (5'-TGGCAGAAGCCACAACCATA-3'), reverse (5'-CCCTCTGGCCCTGCTTATTG-3'); Hoxa9: forward (5'-GGAATAGGAGGAAAAAACAGAAGAGG-3'), reverse (5'TGTATGAACCGCTCTGGTATCCTT-3').
The following variable region ChIP-PCR primers were used for variable region were from (Ji et al 2019): V1-1: forward (5'- ACGTCACAGTGAGGATGTGACA-3'), reverse (5'- CTAGGCACATATCCTCCAGCAT-3'); V1-7: forward (5'-TCATCAAGCCTACAGGTTAGTC-3'), reverse (5'- AGACACAGTGGTGCAACCACAT-3'); V1-59: forward (5'- CATACTACACACCATCCTGGCT-3'), reverse (5'-AACCCTGGAGGAGTAGCAAACT-3'); V6-1: forward (5'-CTTCCTACACAAGCCATGGGTA-3'), reverse (5'- GCAACATGTTATGGAGGTTTGT-3'); D1-1: forward (5'- CTAGACTCAGTTTTTGGAGCTCAA-3'), reverse (5'-CTACGGTAGTAGCTACCACAGT-3'); D2-8: forward (5'-CTGTGGTAGTTACCATAGTAGAC-3'), reverse (5'-CTCTGGCCCCACCAGACAAT-3'); D3-1: forward (5'- AAAGCCAGAAAGGGAATAGGTCT-3'), reverse (5'-CTGTCACAGTGGGCACAGCT-3'); D5-4: (5'-CTGACTGGCTAAACACTGTAGA-3'), reverse (5'-CACAAGAGGTGGATTCTGTATGT-3'); J2: forward (5'-GAGGTTGTAAGGACTCACCTGA-3'), reverse: (5'-ACATTGTTAGGCTACATGGGTAGA-3'); J3: forward (5'-CTGCAGAGACAGTGACCAGAGT-3'), reverse (5'-TGGAGCCCTAGCCAAGGATCA-3')

For ChIP-Sequencing libraries were prepared using a ThruPLEX DNA-seq Kit (Rubicon Genomics) and validated using a TapeStation (Agilent Technologies) and Qubit 2.0 Flurometer (Thermo Fisher Scientific). Libraries were pooled and sequenced on a HiSeq2000 platform (Illumina).

Quantitative PCR
RNA isolated from sorted B cell populations and v-Abl cells were subjected to quantitative PCR with the following primers and normalized to *Gapdh* expression: From (Chakraborty et al., 2009): Dfl-4.5: forward (5'-AGGCATCTCATCTCACTCTAAGC-3'), reverse (5'-TGTGTCCCTCTAAGACGAGTGAAT-3'); Dq52: forward (5'-TGGTGCAAGGTTTTGACTAAGC-3'), reverse (CCAAACAGAGGGTTTTTGTTGAG-3'); Dsp2: forward (5'-TGTTACCTTACTTGGCAGGGATTT-3'), reverse (5'-TGGGTTTTTGTTGCTGGATATATC-3'); γ-actin: forward (5'-GGTGTCCGGAGGCACTCTT-3'), reverse (5'-TGAAAGTGGTCTCATGGATACCA-3'); Cμ: forward (5'- AGAGATCTGCATGTGCCCATT-3'), reverse (5'-TGGTGGGACGAACACATTTACA-3'); Eμ (5'): forward (5'-CTGACATTACTTAAAGTTTAACCGAGG-3'), reverse (5-CTCCAACTCAACATTGCTCAATTC-3');
Eμ(3'): forward (5'-ATTCAGCCGAAACTGGAGAGGTC-3'), reverse (5'-GGGGAAACTAGAACTACTCAAGC-3');
From Zan et al 2017: Aicda: forward (5'-AGAAAGTCACGCTGGAGACC-3'), reverse (5'-CTCCTCTTCACCACGTAGCA-3'); Rad52: forward (5'-

AGCCAGTATACAGCGGATGAA-3'), reverse (5'-GCCATGCGGCTGCTAATGTA-3'); Polθ: forward (5'-TGGCTATATGGGCAGCACCT-3' 5'-CAGAGCAATGCCCTTGGATTT-3'); Ku70: forward (5'-CACCAAGCGGTCTCTGACTT-3'), reverse (5'-AGAGAGGGCCTCAGGTAGTG-3'), Ku80: forward (5'-AGGCCCAGGAAGCTCTATCA-3'), reverse (5'-GCACTCTTGGATTCCCCACA-3'); Gapdh: forward (5′-TTCACCACCATGGAGAAGGC-3′), reverse (5′-GGCATGGACTGTGGTCATGA-3′)
Gapdh: forward (5'-CAAGCAGATGATGTTTCCTGTCC-3'), reverse (5'-AGAACTAAGGGTGGGTGGTGTAGC-3') (Wu et al., 2003).
Lig4: forward (5'-TCTGCCTTTAAGCCAATGCT-3'), reverse (5'-GTGAGAGAGCCTTCCTGTGG-3'); Xrcc4: forward (5'-TGTGTGAGTGCCAAAGAAGC-3'), reverse (5'-TCATCGGTGCTTCCATCATA-3') (Okamura et al., 2016); Lig3: forward (5'-CCTCTCCAAGCTCACCAAAG-3'), reverse (5'-TGCTCATTGTGAAGGACTCG-3');
Parp1: forward (5'-GCAGCGAGAGTATTCCCAAG-3'), reverse (5'-CCGTCTTCTTGACCTTCTGC-3') (Meador et al, 2008); Rag1: forward (5'–TGCAGACATTCTAGCACTCTGG–3'), reverse (5' ACATCTGCCTTCACGTCGAT–3');
Rag2: forward (5'-CACATCCACAAGCAGGAAGTACAC–3'), reverse (5'-CCCTCGACTATACACCACGTCAA–3') (Bender et al. 2004);
Rad51: (5'-CTCATGCGTCAACCACCAG-3'), reverse (5'-GCTTCAGGAAGACAGGGAGAG-3') (Pandit et al., 2012).

ATAC-sequencing

ATAC-seq was performed as previously described (Buenrostro et al., 2013). For each sample, cell nuclei were prepared from $5x10^4$ cells and incubated with 2.5 µL transposase (Illumina) in a 50 µL reaction for 30 minutes at 37°C. Following purification of transposase-fragmented DNA, the library was amplified by PCR and subjected to high-throughput sequencing on the HiSeq 2000 platform (Illumina).

NGS Data analysis and statistical methods

Reads from ChIP-seq and ATAC-seq libraries were trimmed for quality using 'trim_galore' and aligned to mouse genome assembly mm9 with bowtie2 using the default parameters and duplicates removed with the Picard tool MarkDuplicates (http://broadinstitute.github.io/picard/). Density profiles were created by extending each read to the average library fragment size for ChIP and 0 bp for ATAC, then computing density using the BEDTools suite (http://bedtools.readthedocs.io). Enriched regions were discovered using MACS (v1.4) and scored against matched input libraries (fold change > 2 and p-value < 1e-5). Genome browser tracks and read density tables were normalized to a sequencing depth of ten million mapped reads.

ChIP and immunoblotting antibodies

Whole cell extracts for immunoblotting and Chromatin Immunoprecipitations were conducted with H3K36me3 (61101 pAb, Active Motif), H3K36me3 (ab9050 Abcam), H327me1 (61015 Active Motif), H3K27me2 XP (D18C8, CST), H3K9ac (ab4441, Abcam), Hmgb2 (ab67282, Abcam), H3K27me3 (07-449, Millipore), H3K36me 1 (Ab9048, Abcam), H3K36me2 (ab9049, Abcam), H3K4me3 (Ab8580, Abcam), Rag1

(NBP1-74190, Novus), total H3 (Ab61251, Abcam), Ku80 (2753, CST), Gapdh (14C10, CST), Kmt3a (ab31358, Abcam), Xrcc4 (PA-76068, ThermoFisher), Rag2 (ab133609, Abcam), Ku70 (N3H10, ThermoFisher), Lig4 (ab26039, Abcam), phospho-γH2Ax (ab11174, Abcam), total H2Ax (7631 ,CST), Vinculin (4650, CST).

Southern Blot analyses

Southern blot analyses were conducted as described previously (Bredemeyer et al., 2006; Hung et al., 2018). Briefly, 10μg of genomic DNA from v-Abl B cells containing the pMG-INV recombination substrate were digested with XbaI or NheI and hybridized to a P^{32}-labeled probe for Thy1 or GFP. Thy1 and GFP probes were made from 800bp and 700bp cDNA fragments respectively. After incubating with Thy1 probe, XbaI digested Southern blots were stripped and re-probed with the GFP probe.

Detection of Igh CDR3 sequences from NGS of proB genomic DNA

MiXCR v2.1.11 (Bolotin et al., 2017) was used to detect Igh CDR3 sequences from the next generation sequencing data. Only the best matched V/D/J genes were kept for each CDR3 sequences. The unproductive CDR3s, which are out of frame or contain stop codons, were excluded from the downstream analysis. The number of N-nucleotide additions and total deletions were evaluated based on the refPoints column in MiXCR output file. Wilcoxon rank sum test was used to compare the difference for the number of CDR3 sequences, the length of CDR3 amino acid sequences, the number of N-nucleotide additions and total deletions between knockout (KO) and WT group. All the statistical tests were implemented using R.

RNA-seq data processing for TRUST

RNA-seq fastq files were aligned to mouse reference genome mm10 using STAR2 (Dobin et al., 2013). TRUST v3.0.2 (Li et al., 2017; Hu et al., 2018; Hu et al., 2019) was used to infer both partial and complete BCR CDR3 sequences and gene coverage from the aligned RNA-seq BAM files. For each RNA-seq sample, the B cell percentage was estimated by the number of reads mapped to Igh gene region divided by the number of total sequencing reads and the BCR diversity was evaluated by the normalized unique Igh CDR3 calls (Hu et al., 2019).

Generation, analysis and histology of embryos for neurogenesis

C57Bl6 *Setd2f/f* mice were bred to *Nestin-cre* mice to heterozygosity and timed matings between *Setd2f/f* and *Nestin Setd2Δ/+* were established to obtain E14.5, E16.5, E18.5 embryos and 2 hour post-partum pups. Tissues were fixed in 10% buffered formalin and embedded in paraffin and serially sectioned for sagittal and coronal sections (4 μm). TUNEL was performed on sections using TUNEL In Situ Cell Death Detection kit, POD (Roche) with Terminal Deoxynucleotidyl Transferase buffer (Takara Bio) and counterstained with DAPI. Immunostaining of cleaved caspase-3 was performed using anti-cleaved caspase-3 (Asp175) (5A1E) (CST 9664) with anti-rabbit conjugated to HRP secondary antibodies and visualized with DAB. Nuclei counterstained with hematoxylin. Whole tissues on slides were scanned by a digital slides scanner (3D Histech, MIDI) and viewed with Caseviewer (3D Histech).

Statistical analyses

Error bars in all data shown represent standard deviation. Unless otherwise indicated, determination of statistical significance and standard deviations were calculated using unpaired two-tailed Student's t test (when comparing two conditions e.g. control vs knockout) or one way ANOVA (when comparing across multiple conditions concurrently) using Prism 7 software (GraphPad).

Supplementary References

Bender, T.P., Kremer, C.S., Kraus, M., Buch T., and Rajewsky, K. (2004). Critical Functions for c-Myb at Three Checkpoints During Thymocyte Development. *Nat Immunol*. 5:721-9. doi: 10.1038/ni1085.

ten Boekel, E., Melchers, F., Rolink, A. (1995). The status of Ig loci rearrangements in single cells from different stages of B cell development. *International Immunology*. **7**, 1013-1019. doi:10.1093/intimm/7.6.1013.

Buenrostro, J.D., Giresi, P.G., Zaba, L.C., Chang, H.Y., and Greenleaf, H.Y. (2013). Transposition of native chromatin for fast and sensitive epigenomic profiling of open chromatin, DNA-binding proteins and nucleosome position. *Nat Methods*. doi: 10:1213-1218. doi:10.1038/nmeth.2688.

Corcoran, A., Riddell, A., Krooshoop, D., and Venkitaraman, A.R. (1998). Impaired immunoglobulin gene rearrangement in mice lacking the IL-7 receptor. *Nature*. 26:904-907. doi:10.1038/36122.

Deriano, L. and Roth, D.B. (2003). Modernizing the nonhomologous end-joining repertoire: alternative and classical NHEJ share the stage. *Annu Rev Genet.* 47:433–455. doi: 10.1146/annurev-genet-110711-155540.

Dobin, A., Davis, C.A. Schlesinger, F., Drenkow, J., Zaleski, C., Jha, S., Batut, P., Chaisson, M., and Gingeras, T.R. (2013). STAR: ultrafast universal RNA-seq aligner. *Bioinformatics.* 29:15–21. doi:10.1093/bioinformatics/bts635.

Ehlich, A., Martin, V., Müller, W., and Rajewsky, K. (1994). Analysis of the B-cell progenitor compartment at the level of single cells. *Curr Biol*. 4:573-583.

Goodnow, C.C., Crosbie, J., Adelstein, S., Lavoie, T.R., Smith-Gill, S.J., Brink, R.A., Pritchard-Briscoe, H., Wotherspoon, J.S., Loblay, R.H., Raphael, K., et al. (1988). Altered immunoglobulin expression and functional silencing of self-reactive B lymphocytes in transgenic mice. *Nature*. 334: 676-82. doi:10.1038/334676a0.

Hauser, J., Grundstrom, C., and Grundstrom, T. (2014). Allelic Exclusion of IgH through Inhibition of E2A in a VDJ Recombination Complex. *J Immunol*. 192:2460-2470. doi: 10.4049/jimmunol.1302216.

Hesslein, D.G.T., Pflugh, D.L., Chowdhury, D., Bothwell, A.L.M., Sen, R., and D.G. Schatz. (2003). Pax5 is required for recombination of transcribed, acetylated, 5′ IgH V gene segments. *Genes Dev.* 7: 37–42. doi: 10.1101/gad.1031403.

Meador, J.A., Zhao, M., Su, Y., Narayan, G., Geard, C.R., and A.S. Balajee. (2008).

Histone H2AX is a critical factor for cellular protection against DNA alkylating agents. *Oncogene*. 27:5662-5671. doi: 10.1038/onc.2008.187.

Okamura, K. and Nohara, K. (2016). Long-term Arsenite Exposure Induces Premature Senescence in B Cell Lymphoma A20 Cells. *Arch Toxicol.* 90: 793-803. doi: 10.1007/s00204-015-1500-2.

Pandit, S.K., Westendorp, B., Nantasanti, S., Liere, E., Tooten, P.C.J., Cornelissen, P.W.A., Toussaint, M.J.M., Lamers, W.H. and Bruin, A. (2012). E2F8 Is Essential for Polyploidization in Mammalian Cells. *Nat Cell Biol.* 14:1181-91. doi: 10.1038/ncb2585.

Subrahmanyam, R., Du, H., Ivanova, I., Chakraborty, T., Ji, Y., Zhang, Y., Alt, F.W., Schatz, D.G., and Sen, R. (2012). Localized epigenetic changes induced by DH recombination restricts recombinase to DJH junctions. *Nat. Immunol*. 13:1205–1212. doi: 10.1038/ni.2447.

Wu, K., Jiang, S., and Couch, F.J. (2003). p53 Mediates Repression of the BRCA2 Promoter and Down-Regulation of BRCA2 mRNA and Protein Levels in Response to DNA Damage. *J Biol Chem.* 278: 15652-60. doi: 10.1074/jbc.M211297200.

Zan, H., Tat, C., Qiu, Z., Taylor, J.R., Guerrero, J.A., Shen, T., and Casalib, P. (2017). Rad52 competes with Ku70/Ku86 for binding to S-region DSB ends to modulate antibody class-switch DNA recombination. *Nat Commun*. 8:14244. doi: 10.1038/ncomms14244.