Published in final edited form as:

Nat Immunol. 2021 February ; 22(2): 240–253. doi:10.1038/s41590-020-00827-8.

Smc3 dosage regulates B cell transit through germinal centers and restricts their malignant transformation

Martín A. Rivas[1], **Cem Meydan**[2,3], **Christopher R. Chin**[1,2,3], **Matt F. Challman**[1], **Daleum Kim**[1], **Bhavneet Bhinder**[4], **Andreas Kloetgen**[5], **Aaron D. Viny**[6,7], **Matt R. Teater**[1], **Dylan R. McNally**[1,4], **Ashley S. Doane**[1,4], **Wendy Béguelin**[1], **María Teresa Calvo Fernández**[1], **Hao Shen**[1], **Xiang Wang**[1], **Ross L. Levine**[6,7,8], **Zhengming Chen**[9], **Aristotelis Tsirigos**[5,10,11], **Olivier Elemento**[2,3,4], **Christopher E. Mason**[2,3,12,13], **Ari M. Melnick**[1,*]

[1]Division of Hematology and Medical Oncology, Department of Medicine, Weill Cornell Medicine, New York, NY, USA.

[2]Department of Physiology and Biophysics, Weill Cornell Medicine, New York, NY, USA.

[3]The HRH Prince Alwaleed Bin Talal Bin Abdulaziz Al-Saud Institute for Computational Biomedicine, Weill Cornell Medicine, New York, NY, USA.

[4]Caryl and Israel Englander Institute for Precision Medicine, Weill Cornell Medicine, New York, NY, USA.

[5]Department of Pathology, New York University School of Medicine, New York, NY, USA.

[6]Human Oncology & Pathogenesis Program, Memorial Sloan-Kettering Cancer Center, New York, NY, USA.

[7]Leukemia Service, Department of Medicine, Memorial Sloan-Kettering Cancer Center, New York, NY, USA.

[8]Center for Epigenetics Research, Memorial Sloan-Kettering Cancer Center, New York, NY, USA.

[9]Division of Biostatistics and Epidemiology, Department of Population Health Sciences, Weill Cornell Medicine, New York, NY, USA.

[10]Institute for Computational Medicine, New York University School of Medicine, New York, NY, USA.

[11]Applied Bioinformatics Laboratories, New York University School of Medicine, New York, NY, USA.

[*]**Correspondence:** Ari M. Melnick, MD, Division of Hematology and Medical Oncology, Department of Medicine, Weill Cornell Medicine, New York, NY 10021, USA, Tel: 212-746-7643, amm2014@med.cornell.edu.

AUTHOR CONTRIBUTIONS

M.A.R. has conceptually designed and performed most experiments with the help of M.F.C. and D.K., analyzed and interpreted data, drafted the manuscript and supervised the study. C.M. has analyzed and interpreted RNA-seq, Hi-C, Mint-ChIP data and C.R.C. has analyzed single-cell RNA-seq data with the help of M.R.T., X.W., A.K., B.B. and A.S.D. A.D.V. and R.L.L. have provided the Smc3 animal model and contributed to interpretation of results. D.M.N. has helped with execution of experiments and analysis of data. W.B., M.T.C.F. and H.S. have provided technical expertise and helped with interpretation of results. Z.C. has analyzed the clinical data. A.T., O.E. and C.E.M. have provided expertise and resources for data analysis. A.M.M. participated in the conceptual design, interpretation of data, draft of the manuscript, acquisition of funds and supervision of the project. All authors have read and approved the submitted version.

[12]The WorldQuant Initiative for Quantitative Prediction, Weill Cornell Medicine, New York, NY, USA.

[13]The Feil Family Brain and Mind Research Institute, Weill Cornell Medicine, New York, NY, USA

Abstract

During the germinal center (GC) reaction, B cells undergo extensive redistribution of cohesin complex and 3D reorganization of their genomes. Yet, the significance of cohesin and architectural programming in the humoral immune response is unknown. Herein we report that homozygous deletion of *Smc3* encoding the cohesin ATPase subunit abrogated GC formation, yet in marked contrast *Smc3* haploinsufficiency induced GC hyperplasia, skewing of GC polarity and impaired plasma cell differentiation. Genome-wide chromosomal conformation and transcriptional profiling revealed defects in GC B cell terminal differentiation programs controlled by lymphoma epigenetic tumor suppressors *Tet2* and *Kmt2d*, and failure of $Smc3^{wt/-}$ GC B cells to switch from B cell to plasma cell defining transcription factors. *Smc3* haploinsufficiency preferentially impaired connectivity of enhancer elements controlling various lymphoma tumor suppressor genes, and accordingly *Smc3* haploinsufficiency accelerated lymphomagenesis in mice with constitutive *Bcl6* expression. Collectively, our data indicate a dose-dependent function for cohesin in humoral immunity to facilitate the B cell to plasma cell phenotypic switch, while restricting their malignant transformation.

INTRODUCTION

Germinal centers (GCs) are transient structures that arise upon T cell-dependent antigen encounter during the humoral immune response[1]. B cells entering the GC reaction undergo waves of transcriptional reprogramming, which instruct their various phenotypic transitions[2]. GC B cells that generate high-affinity immunoglobulin are selected through T cell help for terminal differentiation into antibody-secreting plasma cells. The transition to plasma cell involves further phenotypic and transcriptional changes, including downregulation of B cell lineage transcription factors such as *Pax5* and *Ebf1*, and upregulation of factors that encode the plasma cell phenotype such as *Zbtb20* and *Prdm1*[3,4]. How these mechanisms are coordinated at the genome-scale level remains poorly understood. The mammalian genome is spatially organized into boundary delimited, topologically associated domains (TADs) that segregate genes into functionally discrete 3D units[5,6]. Within TADs gene expression is coordinated by enhancers that form contacts with variable numbers of gene promoters located up to hundreds of kilobases away[7,8]. Recent studies have revealed profound differences of genomic 3D architectural features in naïve and GC B cells, such as shifts in TAD boundaries, and differential interactions of promoters and enhancers regulating, suggesting that the coordinated GC transcriptional response is dependent on architectural remodeling of the genome.

Architectural proteins such as cohesin play a vital role in TAD and gene enhancer interactions[9]. Indeed, cohesin binding was observed to redistribute to regulatory elements or boundaries that gained interactivity in GC B cells, suggesting a potential critical role for cohesin in coordinating transcriptional programs during the humoral immune response.

Cohesin is a highly conserved multimeric protein complex also involved in chromosomal condensation and sister chromatid cohesion[10]. The canonical cohesin complex consists of four major subunits: Smc3, Smc1, Stag1/Stag2 and Rad21, which form a ring-shaped structure around chromatin fibers[11]. The structural maintenance of chromosome 3 (Smc3) protein is the sole ATPase motor subunit[12,13], and plays a critical role in forming chromatin domains through DNA loop-extrusion. Homozygous deletion of *Smc3* in mice conferred embryonic lethality[14] while mice heterozygous for *Smc3* displayed developmental defects such as abnormal craniofacial morphology[14]. These manifestations are also seen in humans with congenital cohesin mutations such as in Cornelia de Lange syndrome[15]. In B cells, the absence of cohesin impairs immunoglobulin class switch recombination, indicating a role for this complex in organizing immunoglobulin loci. Given the massive architectural changes observed in GC B cells and their relation to cohesin binding, we embarked on an effort to determine a putative role of cohesin complex in controlling cellular phenotypes during the humoral immune response, focusing on the role of its essential ATPase subunit Smc3.

RESULTS

Smc3 haploinsufficiency induces GC hyperplasia

Cohesin subunit genes are upregulated in GC B cells as compared to naive B cells or plasma cells (Extended Data Fig 1a–b). To define the role of Smc3 during the humoral immune response, we crossed mice bearing a floxed *Smc3* allele with the Cγ1-cre strain, expressing cre recombinase in GC B cells (Extended Data Fig 1c–d)[16,17]. Upon reaching immunological maturity, $C\gamma1^{wt/cre};Smc3^{-/-}$ ($Smc3^{-/-}$), $C\gamma1^{wt/cre};Smc3^{wt/-}$ ($Smc3^{wt/-}$) and $C\gamma1^{wt/cre};Smc3^{wt/wt}$ ($Smc3^{wt/wt}$) control mice were injected with sheep red blood cells (SRBCs) to induce GCs. Mice were sacrificed eight days later, at the GC reaction peak. Histology studies revealed intact splenic architecture in $Smc3^{-/-}$ and $Smc3^{wt/-}$ animals (Figure 1a). However, staining with the GC specific lectin peanut agglutinin (PNA) revealed loss of GC formation in $Smc3^{-/-}$ mice, also noted by absence of $Ki67^{+}$ proliferative B cell clusters (Figure 1a–c).

Although we anticipated that *Smc3* heterozygosity might yield a hypomorphic phenotype, $Smc3^{wt/-}$ animals manifested significantly larger and more abundant GCs (**$p<0.01$, Figure 1a–c). Accordingly, flow cytometry analysis showed severe reduction of GC B cells in $Smc3^{-/-}$ mice, and significant increase in $Smc3^{wt/-}$ animals ($p<0.01$, Figure 1d–f, Extended Data Fig 1e). The residual GC B cells in $Smc3^{-/-}$ mice had failed to excise the targeted allele (Extended Data Fig 1f). Analysis of other splenic B cell populations showed the expected proportion of B cells, as well as marginal zone and follicular B cell populations in the $Smc3^{-/-}$ and $Smc3^{wt/-}$ animals (Extended Data Fig 1g). Increased GC B cells were also detected at four days ($P<0.01$), and 15 post immunization in $Smc3^{wt/-}$ mice ($P<0.05$, Extended Data Fig 1h–i). We confirmed reduction of Smc3 mRNA in $Smc3^{wt/-}$ GC and post-GC B cells by qPCR (Extended Data Fig 1j) and decreased Smc3 protein by capillary immunoblot (Extended Data Fig 1k). There was no reduction of Smc3 in naïve B cells or of cohesin proteins Smc1a and Rad21 (Extended Data Fig 1k). Collectively, we find that reduced dosage of Smc3 paradoxically yields a gain of function rather than hypomorphic phenotype.

Smc3$^{wt/-}$ GC B cells manifest greater proliferation

To determine whether increased GC cellularity was due to increase proliferation, we immunized *Smc3*$^{wt/-}$ or *Smc3*$^{wt/wt}$ mice with SRBCs and injected them eight days later with 5-ethynyl-2'-deoxyuridine (EdU). One hour later, mice were sacrificed and splenic GC B cells stained for DNA content using DAPI (Extended Data Fig 2a). Flow-cytometry analysis showed a significantly higher percentage of GC B cells in S-phase in *Smc3*$^{wt/-}$ mice ($p<0.05$, Figure 2a–b). Given that GC B cells are quite heterogeneous, we next confirmed this finding using the more controlled environment of GC 3D organoids (Extended Data Fig 2b)[18]. Organoids from *Smc3*$^{wt/-}$ mice manifested a four-fold increase in the numbers of GC-like B cells compared to *Smc3*$^{wt/wt}$ (Figure 2c–d). Administering BrdU one hour before flow cytometry revealed significantly greater proportions of GC-like cells in S phase in *Smc3*$^{wt/-}$ organoids ($p<0.01$, Figure 2e). The *Smc3* allele was efficiently excised in organoid GC B cells (Figure 2f). However, we did not observe any difference in the proportion of Caspase-3^{+} GC B cells in immunized *Smc3*$^{wt/-}$ mice (Figure 2g and Extended Data Fig 2d). Collectively, *Smc3* haploinsufficiency results in more proliferative GC B cells, without affecting the relative fraction of cells undergoing apoptosis.

Smc3 haploinsufficiency primarily affects centrocytes

GC B cells include proliferative centroblasts (B220^{+}FAS^{+}CD38^{-}CXCR4^{+}CD86^{-}) forming the GC dark zone (DZ) and generally non-replicative centrocytes (B220^{+}FAS^{+}CD38^{-}CXCR4^{-}CD86^{+}) that form the GC light zone (LZ) (Extended Data Fig 3a–b). Strikingly, we observed significantly increased centrocytes but not centroblasts in *Smc3*$^{wt/-}$ ($p<0.01$, Figure 3a–c). We then used a *Rosa26* lox-stop-lox yellow fluorescent protein (YFP) reporter strain to observe GC structures in vivo (Extended Data Fig 3d). *Rosa26*^{YFP};*Smc3*$^{wt/-}$ or *Rosa26*^{YFP};*Smc3*$^{wt/wt}$ mice were immunized with SRBCs, and injected with CD35 antibody conjugated to BV421 24 hours before sacrifice to stain follicular dendritic cells (FDCs), which are exclusive to the LZ. We observed a significant proportional increase in the LZ area as composed to DZ in *Smc3*$^{wt/-}$ GCs (Figure 3c–d), as further confirmed with flow cytometry of YFP^{+} GCB cells in these animals (Extended Data Fig 3e–f). Therefore, the primary GC impact of *Smc3* haploinsufficiency manifests among centrocytes in the GC LZ.

Smc3 haploinsufficiency impairs plasma cell differentiation

B cells might accumulate in the LZ if they are unable to exit the GC reaction to become antibody-secreting plasma cells (PCs). To address this question, we immunized *Smc3*$^{wt/-}$ and *Smc3*$^{wt/wt}$ mice with NP-KLH, followed by a NP-CGG boost 21 days later and evaluated cognate immunoglobulin production after 14, 26, 35 and 70 days (Extended Data Fig 4a). Serum ELISA studies revealed significant reduction of low- $NP_{(30)}$ and high-affinity $NP_{(8)}$ anti-NP IgG1, IgG2b, IgG3 and Igλ at most time points in *Smc3*$^{wt/-}$ mice (Figure 4a). However, we did not observe significant changes in the $NP_{(8)}/NP_{(30)}$ binding ratio, suggesting that *Smc3* haploinsufficiency does not affect immunoglobulin affinity maturation and might instead impair plasma cell differentiation. Indeed, measuring the abundance of bone marrow resident NP-specific antibody-secreting cells by ELISpot revealed significantly fewer NP-reactive plasma cells in *Smc3*$^{wt/-}$ mice, but no impairment in their affinity to NP. There was also significant decrease in bone marrow B220^{-}CD138^{+} BMPCs and long-lived

NP specific BMPCs ($B220^{-}CD138^{+}IgG1cyt^{+}NPcyt^{+}$) in $Smc3^{wt/-}$ mice, ($p<0.05$, Figure 4c–d and Extended Data Fig 4b). Although total loss of cohesin was reported to impair class switch recombination[19,20], we did not detect reduction in IgG1 switched relative to IgM^{+} GC B cells in $Smc3^{wt/-}$ mice (Extended Data Fig 4c–d).

To more directly investigate plasma cell differentiation, we cultured B220 enriched splenocytes derived from $Rosa26^{YFP}$;$Smc3^{wt/-}$ and $Rosa26^{YFP}$;$Smc3^{wt/wt}$ mice in media containing IL-4, LPS and CD40 agonist antibodies for eight days[21,22] (Extended Data Fig 4e). Flow cytometry was performed each day, gating on YFP^{+} cells to select for those with cre activation and absence of DAPI staining for viability. We defined $DAPI^{-}YFP^{+}B220^{+}CD138^{+}$ cells as viable plasmablasts, and $DAPI^{-}YFP^{+}B220^{-}CD138^{+}$ cells as viable PCs (Figure 4e). Although there was equivalent plasmablast formation at early time points, these cells subsequently plateaued and then actually decreased in $Smc3^{wt/-}$ cells. Progressive increase in PCs began by day five in $Smc3^{wt/wt}$ B cells, whereas PC differentiation was profoundly impaired among $Smc3^{wt/-}$ B cells (Figure 4e–f). This was most likely due to differentiation blockade rather than cell death, since there was no increase of Caspase-3 activation in $Smc3^{wt/-}$ plasmablasts (Extended Data Fig 4f–g). PC differentiation requires multiple rounds of cell division[23], and we indeed observed reduced incorporation of EdU in the subpopulation of $IRF4^{hi}$ GC B cells that are believed to be committed to becoming PCs[24] (Extended Data Fig 4h–j).

PC differentiation requires downregulation of B cell lineage transcription factors such as *Pax5* and *SpiB*, and induction of *Irf4*, *Xbp1,* and *Prdm1*[3]. We examined expression of these genes by qPCR from day four to day six in sorted ex vivo PBs. In contrast to wild-type B cells, $Smc3^{wt/-}$ PBs failed to downregulate the key B cell transcription factor *Spib* (Figure 4g). *Irf4*, which normally increases as B cells transition to plasma cell was induced as expected, consistent with the initially normal levels of plasmablasts in $Smc3^{wt/-}$ cells. Along these lines, flow cytometry analysis of primary GC B cells revealed equivalent numbers of $B220^{+}FAS^{+}GL7^{+}IRF4^{+}$ GC B cells, suggesting that the initial phases of GC exit remain intact in $Smc3^{wt/-}$ mice (Extended Data Fig 4h–i). Although *Xbp1* expression was unaffected, there was significant impairment in the induction of *Prdm1*, in accordance with impaired PC differentiation (Figure 4g). Hence full *Smc3* expression is specifically required for the transition where B cells downregulate key B cell defining factors to fully differentiate into plasma cells.

Aberrant transcriptional programming in $Smc3^{wt/-}$ B cells

To further understand this phenotype we performed RNA-seq in centrocytes and observed a $Smc3^{wt/-}$ transcriptional signature skewed towards gene repression, with 104 downregulated and 46 upregulated genes (Figure 5a, Supplementary Table 1). Gene set enrichment analysis (GSEA) revealed significant downregulation of genes that are normally induced in centrocytes[25] and PCs[26] (Figure 5b, Supplementary Table 2). Consistent with the higher EdU incorporation in $Smc3^{wt/-}$ GC B cells, there was significant upregulation of proliferation gene sets [27,28,29,30,31] (Figure 5b), including genes such as *CcnB2*, *Plk1, Ret and Aurka* (Figure 5c). From the mechanistic perspective there was downregulation of genes repressed in GC B cells due to loss of lymphoma tumor suppressors *Tet2* and *Kmt2d* (Figure

5b and Extended Data Fig 5a). Likewise, there was downregulation of genes that lose the 5'hydroxymethylcytosine mark (5hmc) in *Tet2*$^{-/-}$ GC B cells, and genes that lose H3K4me2 mark in *Kmt2d* deficient DLBCL cells. This is notable in that both *Tet2* and *Kmt2d* deficient mice manifest defective exit of B cells from the GC reaction[21,32]. Indeed, *Tet2* and *Kmt2d* were both relatively downregulated in *Smc3*$^{wt/-}$ GC B cells at the transcript level, as were several other putative lymphoma tumor suppressors such as *Dusp4*, *Dusp2, Kmt2c*, and *Setd2* among others (Figure 5c). *Smc3* haploinsufficiency thus impairs GC exit transcriptional programs, mimicking the effect of lymphoma tumor suppressors.

Impaired LZ to plasma cell transition in *Smc3*$^{wt/-}$ mice

To further delineate how GC exit might be perturbed in vivo, we performed single-cell RNA-sequencing of 490 canonical immune and lymphoma relevant transcripts (Supplementary Table 3) in sorted YFP^{+}IgD^{-} cells from *Rosa26*^{YFP};*Smc3*$^{wt/-}$ and *Rosa26*^{YFP};*Smc3*$^{wt/wt}$ mice to enrich for GC and post-GC cells, with spike-in of 5% sorted CD4^{+}CXCR5^{+}PD1^{+} TFH cells (Extended Data Fig 5b). We projected canonical signatures and genes from GC B-cell subpopulations[33,34] and plasmablasts[26], onto the single-cell RNA-seq UMAP representations (Figure 5d–f, Supplementary Table 4). GC B cells expressed *Aicda* and *Bcl6,* DZ B cells expressed high levels of *Ki67*, LZ B cells expressed centrocyte genes *Cd86* and *Cd83*. A small subpopulation of *Myc* expressing cells reflected GC B cells transitioning from LZ to DZ[35]. Plasmablasts expressed *Xbp1,* and T cells expressed *Cd4* and *Trac*. To determine how the GC to plasma cell transition might be altered by *Smc3* haploinsufficiency, we performed a pseudotime analysis including centroblasts, centrocytes and plasmablasts (Figure 5g). Distribution of cells in the PCA space is shown in Figure 5h and their membership in the centroblast, centrocyte and plasmablastic/plasma cell compartments in Figure 5i. Plotting the density of DZ, LZ and plasmablastic/plasma cells according to their pseudotime distribution (Figure 5j–k) confirmed the increase of GC B cells (especially LZ) and reduced abundance of cells making the transition from centrocyte to plasmablast (Figure 5k). These data suggest a cell transition bottleneck that may contribute to accumulation of centrocytes that upon exiting the GC reaction fail to transition towards the plasma cell fate.

Impaired gene enhancer looping in *Smc3*$^{wt/-}$ centrocytes

Given the role of cohesin in genome architecture, we next performed in situ Hi-C[36] in sorted naïve B cells, centroblasts, and centrocytes from *Smc3*$^{wt/-}$ and *Smc3*$^{wt/wt}$ mice (Extended Data Fig 6a), and murine plasma cell lines. A global survey of nuclear topology at low resolution showed little impact of *Smc3* haploinsufficiency (Figure 6a). Moreover, comparing and contrasting *Smc3*$^{wt/-}$ vs *Smc3*$^{wt/wt}$ insulation scores revealed minimal TAD boundary disruption, whereas insulation boundaries were massively changed in the transition from naïve B cell to GC B cells or from GC B cell to plasma cell (Figure 6b). Even the relatively minor transition between GC centroblast and centrocyte involved a greater shifting of insulation scores than that caused by *Smc3* haploinsufficiency. However, we observed abundant reduction of intra-TAD vs. inter-TAD interactivity in *Smc3*$^{wt/-}$ centrocytes ($p<2.2\times10^{-16}$, Figure 6c and Extended Data Fig 6b–c), reflected by an overall reduction in TAD boundary strength (Extended Data Fig 6d). Plotting differential interactivity across chromosomal regions further illustrated the intra-TAD connectivity

reduction (Figure 6d), also be seen by collapsing these domains into a single meta-TAD (Figure 6e).

The most dramatic interactivity loss occurred in contacts that were <50 kb apart ($p<2.2\times10^{-16}$, Figure 6f), with still significant in the 50–250 kb range ($p<2.2\times10^{-16}$) and no defect in longer range interactions. These defects reflected significant reduction in connectivity of promoters and enhancers ($p<2.2\times10^{-16}$ and $p<2.2\times10^{-16}$ respectively) as compared to background differential intra-TAD interactions ($p<1.2\times10^{-32}$, Figure 6g). GSEA showed significant loss of promoter connectivity among genes downregulated in *Smc3*$^{wt/-}$ centrocytes (Figure 6h), and association between loss of connectivity and transcriptional repression ($p<1.6\times10^{-7}$, Extended Data Fig 6e). Putative enhancers showed more severe loss of interactions than promoters ($P=1.2\times10^{-32}$, Figure 6i), with highly localized loss of connectivity. Exploring specific gene loops at 5kb resolution showed that lymphoma tumor suppressors such as *Dusp4, Dusp2, Zeb2* and others were among the top quartile of genes with disrupted promoter-enhancer loops (Figure 6j). Taken together, the data suggest that **i)** full dosage of Smc3 is required for the proper coordination between the many hundreds of enhancers that are engaged in the B cell to plasma cell transition[4]; and **ii)** *Smc3* haploinsufficiency impairs multiple tumor suppressor genes, suggesting a potential role for *Smc3* in restricting lymphomagenesis.

Smc3 haploinsufficiency does not impair GC entry

To determine whether other major B cell phenotype transitions linked to extensive architectural reprogramming (such as naïve B cell to GC B cells – Figure 6b) are also dependent on *Smc3* dosage we used *Cd19*^{cre} to delete *Smc3* in immature B cells (Extended Data Fig 7a–b)[37]. As before, we observed that *Smc3* homozygous deletion led to complete ablation of B cells (Figure 7a and Extended Data Fig 7c–e). However, analysis of *Cd19*$^{wt/cre}$;*Smc3*$^{wt/-}$ mice yielded no impairment of B cell differentiation, with normal abundance of pre-B, immature B cells, follicular B cells and naïve B cells (Extended Data Fig 7d–f). Most importantly, upon immunization, instead of seeing an impairment in GC B cells, we observed significant GC expansion and significant increase in the CC-to-CB ratio reminiscent of the *Cγ1*^{cre};*Smc3*$^{wt/-}$ phenotype ($p<0.0001$, Figure 7a–f). Likewise, GC B cells from *Cd19*$^{wt/cre}$;*Smc3*$^{wt/-}$ mice incorporated more EdU than control GC B cells (Figure 7g), and showed no changes in apoptotic markers (Figure 7h). We speculate that perhaps only the most extreme transcriptional and phenotypic transitions require full cohesin dosage. Indeed, phylogenetic analysis of architectural shifts occurring during the humoral immune response indicate that the transition to plasma cell is more profound than the transition from NB cell to GC B cells (Figure 7i). Smc3 dosage is thus exclusively required for the shift from B cell to plasma cell phenotype during B cell development and the humoral immune response.

Smc3 restricts malignant transformation of GC B cells

Our data suggest that *Smc3* is required to maintain tumor suppressor gene (e.g. *Tet2*) expression and function in GC B cells. *Tet2* deficiency in GC B cells was shown to accelerate DLBCL lymphomagenesis in *IμBcl6* mice[32]. Therefore we crossed our Smc3 floxed mice with the *IμBcl6* strain to generate: *IμBcl6*;*Cγ1*$^{wt/cre}$;*Smc3*$^{wt/wt}$ (Bcl6),

IμBcl6;*Cγ1*$^{wt/cre}$;*Smc3*$^{wt/–}$ (Bcl6/Smc3), *Cγ1*$^{wt/cre}$;*Smc3*$^{wt/–}$ (Smc3) and control *Cγ1*$^{wt/cre}$;*Smc3*$^{wt/wt}$ animals. Cohorts of 38–55 mice per genotype were generated by bone marrow transplant of CD45.2 donor cells into wild-type CD45.1 recipients (Extended Data Fig 8a). Three hundred days after transplant, five healthy individuals from each group were randomly selected for histological analysis. Splenomegaly was observed in all Bcl6 and Bcl6/Smc3 animals, but was more severe in those with *Smc3* haploinsufficiency (Figure 8a). Histologic analysis revealed profound disruption of spleen and lymph node tissue organization in Bcl6/Smc3 mice, with disorganized follicular structures and diffusely infiltrating B cells (Figure 8b–c). In contrast, lymphoid follicular structures were still evident in Bcl6 mice although partially distorted and enlarged. Examination of other organs such as liver and lung showed presence of abnormal B cell infiltrates in Bcl6/Smc3 mice consistent with systemic dissemination of lymphoma (Figure 8d–e). This was less prominent in Bcl6 mice and Smc3 animals, and absent from controls. These cells were confirmed to be derived from donor mice (Extended Data Figure 8b), and *Smc3* haploinsufficiency was confirmed by PCR (Extended Data Fig 8c). All five Bcl6/Smc3 mice manifested evidence of multi-organ systemic lymphoma at this time point whereas only 1 out of 5 of the Bcl6 mice presented in this manner (Figure 8f).

Among remaining animals, disease-free survival rate was significantly shorter in Bcl6/Smc3 mice compared to all the other groups (P<0.0001, log-rank Mantel-Cox), with a mean survival of 400 days vs. 483 days for Bcl6 animals (Figure 8g). Furthermore, there was significantly greater frequency of systemic lymphoma Bcl6/Smc3 at day 500 (p<0.0001, Figure 8h). *Smc3* haploinsufficiency alone was sufficient for development of lymphomas in 13% of mice without the *IμBcl6* allele, which was greater than controls (p=0.003). This prompted us to explore the significance of *Smc3* expression in primary human DLBCLs, which only rarely manifest *SMC3* somatic mutations. Notably, Smc3 mRNA abundance inversely correlated with overall survival in a univariate Cox analysis in two independent cohorts of 322 and 757 DLBCL patients respectively (Figure 8i)[38,39,40,41,42]. This effect was still observed in a Multivariable Cox analysis that included age, gender, and DLBCL subtype (Figure 8j). Hence, although cohesin mutations are rare in DLBCL, reduced expression of this gene is associated with more aggressive and clinically unfavorable disease, consistent with lymphomagenesis in *Smc3* haploinsufficient mice.

DISCUSSION

Herein we show that cohesin ATPase subunit Smc3 plays a central role in B cell to plasma cell transitions and restricts malignant transformation of GC B cells. *Smc3* haploinsufficiency manifested as GC hyperplasia with expansion of centrocytes having enhanced proliferative potential and impaired plasma cell differentiation. However, reduction in *Smc3* had no impact in the ability of B cells to differentiate from the pre-B to mature B cell stage[43], or to transition into the GC reaction[44]. Architectural patterning within the B cell lineage is mediated in part by Pax5, which is downregulated during plasma cell differentiation[45]. PC differentiation is linked to massive changes in transcriptional programming as these cells become highly specialized for protein production and secretion[46]. We show that the architectural changes that accompany the B cell to plasma cell transition are even more extensive than those occurring in the GC reaction. Hence there may

be a quantitative threshold for cohesin requirement for switching between major transcriptional programs, and/or a qualitatively different function for *Smc3* to mediate this process. Either way, this concept is consistent with studies of *Smc3* in the hematopoietic system, where its loss of function caused significant impairment of progenitor cells switching into erythroid or lymphoid lineages[17,47,48]. These findings also align with the requirement for Smc3 ATPase activity in chromatin loop extrusion[49,50], which is evidently critical for reprogramming genomic architectural patterning in plasma cells.

Smc3 haploinsufficiency yielded selective impairment in expression of genes that are normally activated in centrocytes in response to immune signals. Similar sets of GC exit genes are also impaired in the absence of the *Tet2* or *Kmt2d* tumor suppressors[21,32]. *Tet2* mediates its effects in GC B cells though cytosine hydroxymethylation of gene enhancers and its loss of function leads to GC hyperplasia with LZ expansion and impaired plasma cell differentiation[32]. Kmt2d functions to induce H3K4 monomethylation of enhancers and its deficiency leads to a proliferation advantage, impaired GC exit and repression of tumor suppressor genes[21,51]. Both *Tet2* and *Kmt2d* deficiency results in formation of B cell lymphomas in mice[21,32,51]. *Smc3* haploinsufficient GC B cells reflect features of both of these phenotypes and aberrant downregulation of many genes that are similarly affected by loss of *Tet2* and *Kmt2d*.

The causal relationship between enhancer promoter contacts and gene expression was established for example by studies showing that proximity of enhancers and promoters in the *Hbb* loci can induce gene expression[52]. Pathological enhancer promoter looping can cause aberrant oncogene expression and malignant transformation[53]. S*mc3* haploinsufficiency in centrocytes caused extensive reduction in chromatin connectivity, consistent with studies showing that full cohesin knockout abolishes long and short range looping[54]. Reduced dosage of *Smc3* had a preferential effect on short range interactions, as predicted by computational algorithms modeling cohesin dosage effects on loop extrusion[55]. This defect was observed throughout the genome, but was especially concentrated at gene enhancers including those for several tumor suppressor genes. Notably, Cuartero et.al. observed that *Rad21* loss of function had a highly selective impact on macrophage enhancers[56]. Although there was very little effect on the macrophage transcriptome, loss of *Rad21* selectively impaired responsiveness of a subset of activation inducible enhancers to external signals. Likewise, we find that *Smc3* deficiency impaired only a subset of centrocyte induced genes, whereas others including those associated with *Myc*, anaplerotic metabolism and nucleotide synthesis which are normally induced by T cell help, were upregulated in *Smc3* haploinsufficient GC B cells. This may at least partially explain the proliferative advantage of these cells. Although we don’t understand why these genes are induced, it has been suggested that cohesin can also mediate repressive functions[57]. These results suggest that maintaining enhancer connectivity or perhaps creating new enhancer-promoter contacts might be particularly taxing for the cohesin complex. There is considerable debate regarding the precise role of cohesin in driving transcription through looping effects, vs being secondarily involved downstream of other regulatory proteins. Both models are of interest and relevance to our findings and may occur simultaneously in GC B cells and plasmablasts.

Our data nominate *Smc3* as a putative tumor suppressor for GC B cells, given its cooperation with Bcl6 in accelerating lymphomagenesis. This effect was reminiscent of how *Tet2* deficiency cooperates with *Bcl6* to form lymphomas[32]. Oncogenesis in the *Smc3* deficient setting is likely multifactorial, and possibly linked to its driving GC hyperplasia, impairment of tumor suppressor expression, differentiation blockade and induction of Myc/proliferative signaling. *Smc3* and other cohesin complex encoding genes are frequently mutated in cancers, but are extremely rare in B cell lymphomas[58]. This is puzzling in light of our results showing that it exerts powerful tumor suppressor effects in B cells, but in accordance with the inferior survival among DLBCL patients with low *SMC3* expression. Given that *Smc3* deficiency phenocopies loss of *Kmt2d* and *Tet2*, we postulate that mutation of these genes may offer a path of lower resistance to malignant transformation than loss of *Smc3* alleles. For example, it is possible that *Smc3* loss of function could also have growth suppressive effects that detract from its oncogenicity. Collectively, our data point to novel functions for cohesin in architectural programming of the humoral immune response, including an essential role in specifying the plasma cell phenotype, as well as in dynamic regulation of enhancers involved in exit from the GC reaction; and show that cohesin subunits can serve as new class of “architectural tumor suppressors” for B cells.

METHODS

Generation of Conditional *Smc3*-deficient Mice

The Research Animal Resource Center of the Weill Cornell Medical College approved all mouse procedures. The *Smc3* allele was deleted by targeting exon 4 in a construct obtained from the EUCOMM consortium (*Smc3*$^{tm1a(EUCOMM)Wtsi}$)[17]. The generated mice (*Smc3*$^{fl/fl}$) were crossed to the B6.129P2(Cg)-*Ighg1*$^{tm1(cre)Cgn}$/J mice (*Cγ1*^{cre}, The Jackson Laboratory[16]) or to B6.129P2(C)-*Cd19*$^{tm1(cre)Cgn}$/J mice (*CD19*^{cre}, The Jackson Laboratory[37]). *Cγ1*$^{cre/cre}$;*Smc3*$^{wt/–}$ were also crossed to the reporter strain B6.129X1-*Gt(ROSA)26Sor*$^{tm1(EYFP)Cos}$/J (*Rosa26*^{YFP}, The Jackson Laboratory[59]). For lymphomagenesis experiments, *Cγ1*$^{cre/cre}$;*Smc3*$^{wt/–}$ were crossed to *Ighm*$^{wt/tm1(Bcl6)Rdf}$ (*IμBcl6*)[60].

Patient data

For univariable and multivariable Cox analysis, we used publicly available gene expression data from 322 DLBCL patients (cohort 1 from British Columbia Cancer Agency)[38] and from 757 DLBCL patients (cohort 2, an independent cohort from our institution)[39,40,41,42].

Germinal Center Assessment in Mice

Age- and sex-matched mice were immunized intraperitoneally (i.p.) at 8–12 weeks of age with either 0.5 ml of a 2% sheep red blood cell (SRBC, Cocalico Biologicals) suspension in PBS or 100 μg of highly substituted NP-KLH (NP-32 Keyhole Limpet Hemocyanin, Biosearch Technologies) in alum (Thermo Scientific), or 100 μg of highly substituted NP-CGG (NP-28 Chicken Gamma Globulin, Biosearch Technologies), and were sacrificed after 4–70 days, depending the experiment.

Bone Marrow Transplantation

6 to 8 weeks C57BL6/J or B6.SJL-Ptprca Pepcb/BoyJ (CD45.1) mice were lethally irradiated with two doses of 450 rad each the day previous to the injection. Bone marrow cells were isolated from *Cγ1*$^{wt/cre}$;*Smc3*$^{wt/wt}$, *Cγ1*$^{wt/cre}$;*Smc3*$^{wt/-}$, *IμBcl6*;*Cγ1*$^{wt/cre}$;*Smc3*$^{wt/wt}$ and *IμBcl6*;*Cγ1*$^{wt/cre}$;*Smc3*$^{wt/-}$ CD45.2 donors, and transplanted through tail vein injection into CD45.1 mice (1.0×10^6 per mouse). After engraftment, mice were immunized with SRBCs every 3 weeks for the duration of the experiment, to favor the continuous production of germinal centers. With the exception of mice euthanized at specific time points, all mice were followed until any one of several criteria for euthanizing were met, including severe lethargy, more than 10% body weight loss, and palpable splenomegaly that extended across the midline, in accordance with Weill Cornell Medical College and Memorial Sloan-Kettering Cancer Center Institutional Animal Care and Use Committee approved animal protocols. Animal care was in strict compliance with institutional guidelines established by the Weill Cornell Medical College, the Memorial Sloan-Kettering Cancer Center, the Guide for the Care and Use of Laboratory Animals[61], and the Association for Assessment and Accreditation of Laboratory Animal Care International.

Flow cytometry Analyses

Single-cell suspensions from mouse spleens and 3D B cell follicular organoids were stained using the following fluorescent-labeled anti-mouse antibodies: from BD Biosciences: FITC anti-GL7 (553666, dilution 1:500), FITC anti-mouse IgG1 (562026, dilution 1:250), PE anti-CD95/Fas (554258, dilution 1:500), PE anti-IgD (558597, dilution 1:500), PE-Cy7 anti-CD86 (560582, dilution 1:400), PE-Cy7 anti-CD95/Fas (557653, dilution 1:500), PerCP-Cy5.5 anti-BrdU (560809, dilution 1:200), AlexaFluor®647 anti-GL7 (561529, dilution 1:750), AlexaFluor®647 anti-Active Caspase-3 (560626, dilution1:200), APC anti-mouse IgG1 (560089, dilution 1:200), BV421 anti-mouse IgG1 (562580, dilution 1:250), and BUV737 anti-CD138 (564430, dilution 1:250); from Thermo Fisher Scientific:FITC anti-CD21/CD35 (11–0212-81, dilution 1:500), PE-Cy7 anti-B220 (25–0452, dilution 1:750), PE anti-CD23 (12–0232-85, dilution 1:500), PE anti-CD184/CXCR4 (12–9991-81, dilution 1:250), PerCP-Cy5.5 anti-CD45.1 (45–0453-80, dilution 1:500), PerCP-Cy5.5 anti-CD45.2 (45–0454-80, dilution 1:500), APC anti-CD38 (17–0381-81, dilution 1:500), and APC anti-IgM (17–5790-82, dilution 1:500); from eBioscience: eFluor506 anti B220 (RA3–6B2, dilution 1:500); from BioLegend: PE anti-IRF4 (646404, dilution1:50), PerCP-Cy5.5 anti-GL7 (144610, dilution 1:500), PerCP-Cy5.5 anti-Fas (152610, dilution 1:500), APC anti-CD3 (100236, dilution 1:500), APC-Cy7 anti-B220 (103224, dilution 1:750), APC-Cy7 anti-CD38 (102728, dilution 1:500), and BV421 anti-CD138 (142507, dilution 1:400); DAPI was used for the exclusion of dead cells and to determine DNA content in cell cycle analysis. Ghost Dye™ Violet 510 Viability Dye was used for the exclusion of dead cells. For internal markers, cells were fixed and permeabilized with BD Cytofix/Cytoperm Fixation/Permeabilization Solution Kit (BD Biosciences). Invitrogen™ Click-iT™ EdU Alexa Fluor™ 488 Flow Cytometry Assay Kit was used to detect EdU incorporation by flow cytometry. Data was acquired on BD FACSCantoII flow cytometer (BD Biosciences) or BD LSRFortessa and analyzed using FlowJo software package (BD Biosciences).

3D B Cell Follicular Organoid

For B cell purification, spleens were freshly obtained from $C\gamma1^{wt/cre};Smc3^{wt/wt}$ and $C\gamma1^{wt/cre};Smc3^{wt/-}$ mice. Splenocytes were isolated by a combination of mechanical and gradient separation methods (Fico/Lite-LM, Atlanta Biologicals). B cells were obtained from splenocytes through negative selection using EasySepTM Mouse B cell Isolation Kit (Stem Cell Technologies) in accordance with manufacturer's protocol (yield ~ 90% $B220^+$). Organoids were performed as previously described[18]. Organoids were fabricated in 96-well plates by first adding 10 μl of 3% hydrogel silicate nanoparticle (SiNP) followed by injecting 10 μl cell-containing gelatin solution into the initial SiNP droplet, and then mixing the entire hydrogel through repeated pipetting. Each organoid contained 50,000 B cells and 80,000 40LB cells. Organoids were cured for about 10 min before the addition of RPMI media with 10% FBS and penicillin G/streptomycin, containing 50 ng/ml recombinant mouse IL-4 (R&D Systems) and 25 ng/ml IL-21 (R&D Systems) and were incubated at 37 °C with 5% CO_2.

2D B cell cultures

For LPS-IL-4-anti-CD40 cultures, primary B cells were purified by negative selection using anti-CD43 MicroBeads (Miltenyi Biotec) and cultured in complete RPMI supplemented with 10% FBS, 50 mM of 2-β-Mercaptoethanol (Gibco), 20 mM HEPES pH 7.2–7.5 (Gibco), 25 ng/ml of IL-4 (R&D), 25 μg/ml lipopolysaccharide (LPS, Sigma-Aldrich) and 1 μg/ml functional-grade anti-CD40 (ThermoFisher Scientific).

Cell Sorting

Naïve and GC B cells were sorted from the spleens of 8-day SRBC immunized mice. Briefly, single-cell suspensions were stained with anti-B220, anti-CD95/Fas and anti-GL7 antibodies. Centrocytes and centroblasts were stained with anti-B220, anti-CD95/Fas, anti-GL7 antibodies, anti CXCR4 and anti-CD86 antibodies. DAPI was used for the exclusion of dead cells. Cell sorting was performed in BD Influx in the Weill Cornell Medicine Flow Cytometry Core Facility.

ELISA

Murine serum samples were collected before NP-KLH immunization and 14-, 21-, 26-, 35- and 70-days post NP-KLH, and immunoglobulin titers were analyzed by ELISA. Sera were tested for the binding of NP-specific IgG1, IgG2b, IgG3, IgM and Igλ antibodies (SouthernBiotech, dilution 1:500) to $NP_{(8)}$-BSA and $NP_{(30)}$-BSA coated plates (Biosearch Technologies). Optical density (OD) at 405 nm was measured in a plate reader (BioTek) and the absorbance ratio was calculated by dividing the mean OD in $NP_{(8)}$-BSA coated wells by the mean OD in $NP_{(30)}$-BSA coated wells.

ELISPOT

Murine bone marrow cells were collected 70 days after NP-KLH immunization, and IgG1 production was analyzed by ELISPOT. ELISPOT plates (Millipore) were activated by incubation with 35% ethanol. Cells were tested for the binding of NP-specific IgG1 antibody (SouthernBiotech, dilution 1:500) to $NP_{(8)}$-BSA and $NP_{(30)}$-BSA coated plates (Biosearch

Technologies). The plates were scanned, and the number of spots were analyzed by ZellNet Consulting Inc., NJ. The ratio of the number of spots was calculated by dividing the mean number of spots in $NP_{(8)}$-BSA coated wells by the mean number of spots in $NP_{(30)}$-BSA coated wells.

Immunohistology

Mice organs were fixed in 4% formaldehyde and embedded in paraffin. Deparaffinized slides were antigen retrieved in citrate buffer pH 6.4 and endogenous peroxidase (HRP) activity was blocked by treating the sections with 3% hydrogen peroxide in methanol. Indirect immunohistochemistry was performed with anti species-specific biotinylated secondary antibodies followed by avidin–horseradish peroxidase or avidin-AP, and developed by Vector Blue or DAB color substrates (Vector Laboratories). Sections were counterstained with hematoxylin if necessary. The following antibodies were used: biotin-conjugated anti-PNA (Vector Laboratories), biotin-conjugated anti-B220 (Invitrogen), and Ki67 (Vector VP-K451). Photomicrographs were examined using a Zeiss Axioskop imaging microscope. ImageJ 1.48v software (NIH) was used to quantify germinal center areas.

Immunoblotting

Lysates from sorted NBC, CC or CB cells were prepared using 20 mM Tris, pH 8, 135 mM NaCl, 1% NP-40 and 10% glycerol lysis buffer. Protein lysates were resolved using the fully automated capillary immnoblot Jess instrument (ProteinSimple), and probed with primary antibodies against Smc3, Smc1a, Rad21 (Cell Signaling) or β-actin (Sigma Aldrich). Peroxidase-conjugated correspondent secondary antibody and enhanced chemiluminescence were used for detection. Densitometry values were obtained from the Compass for SimpleWestern 5.0 software.

RT-qPCR

RNA was prepared using TRIzol extraction (Invitrogen). cDNA was prepared using Verso cDNA Synthesis kit (Thermo Fisher Scientific) and detected by Fast SYBR™ Green (Thermo Fisher Scientific) on QuantStudio 6 Flex Real-Time PCR System (Thermo Fisher Scientific). We normalized gene expression to Hprt1 and expressed values relative to control using the ΔΔCt method. Results were represented as fold expression with the standard deviation for 2 series of triplicates.

qPCR primers

The following primers were used in qPCR experiments:

SpiB_F: 5′-CTGGCCATTCTACACGGTTT-3′,

SpiB_R: 5′-ATCTGGCCAACAAAACAAGG-3′,

Prdm1_F: 5′-GACGGGGGTACTTCTGTTCA-3′,

Prdm1_R: 5′-GGCATTCTTGGGAACTGTGT-3′,

Irf4_F: 5′-GCAGCTCACTTTGGATGACA-3′,

Irf4_R: 5′-CCAAACGTCACAGGACATTG-3′,

Xbp1_F: 5′-TGAATGGCCCTTAGCATTTC-3′,

Xbp1_R: 5′-CACAGAACAGGACGCTGTGT-3′,

Pax5_F: 5′-AACTTGCCCATCAAGGTGTC-3′,

Pax5_R: 5′-CTGATCTCCCAGGCAAACAT-3′,

Sdc1_F: 5′-AGGCAGCTGTCACATCTCAT-3′,

Sdc1_R: 5′-GTCTTGTTCTCCAGAGCCCT-3′,

Hprt1_F: 5′-AGGACCTCTCGAAGTGTTGG-3′,

Hprt1_R: 5′-TTGCAGATTCAACTTGCGCT-3′.

Genotyping PCR

Smc3$^{fl/fl}$, *Smc3*$^{wt/fl}$, and *Smc3*$^{wt/wt}$ littermate mice were genotyped by PCR with primers Smc3-FloxF (5′-TGTTGCCTCCCTGTGTTCTCAGGC-3′), Smc3-CommonF (5′-TCTTCGTCCAGAGCAGCGATTGGC-3′), and Smc3-CommonR (5′-TGAGCCATCATGTGGATGCT-3′) using the following parameters: 95°C for 4 min, followed by 35 cycles of 95°C for 45 s, 56°C for 45 s, and 72°C for 1 min, and then 72°C for 5 min. The wild-type allele was detected as a band at 287 bp, whereas the floxed allele was detected as a band of 313 bp. Excision after Cre recombination was confirmed by PCR with primers to detect a band at 349 bp[17].

RNA-sequencing

mRNA-seq Library Preparation and RNA-seq libraries were prepared using the Illumina TruSeq RNA sample kits, according to the manufacturer. Libraries were validated using the Agilent Technologies 2100 Bioanalyzer and Quant-iT dsDNA HS Assay (Life Technologies) and 8–10 pM sequenced on HiSeq2000 sequencer. RNA-seq data were aligned to mm10 and Gencode M12[62] transcripts using STAR (v2.5.3)[63]. Gene expression quantified by featureCounts[64] to counts and normalized to Transcripts per Million (TPM)[65]. Differentially expressed genes between *Smc3*$^{wt/-}$ and *Smc3*$^{wt/wt}$ were identified using count data with a negative binomial model with the DESeq2 package[66]. Pathway enrichment was calculated by using GSEA[67] and FGSEA[68] on the $\log_2$ fold change ranking results from DESeq2 output[66] with MSigDB[69] and other gene signature databases from literature, using murine and human orthologs of genes as necessary.

In situ Hi-C

1.5×10^6 flow sorted mouse GC B cells from *Cγ1*$^{wt/cre}$;*Smc3*$^{wt/wt}$ (n=3) and *Cγ1*$^{wt/cre}$;*Smc3*$^{wt/-}$ (n=3) were fixed in 1% formaldehyde for 10 min. Fixation was quenched by the addition of 0.125 M glycine for 10 min. In situ Hi-C was performed as described[36]. Briefly, nuclei were permeabilized and DNA was digested overnight with 100 U DpnII (New England BioLabs). The ends of the restriction fragments were labeled using

biotin-14-dATP and ligated in 1 mL final volume. After reversal of crosslinks, ligated DNA was purified and sheared to a length of ~400 bp, at which point ligation junctions were pulled down with streptavidin beads, DNA fragments repaired, dA-tailed and Illumina adapters ligated. Library was produced by 6–10 cycles of PCR amplification. Sequencing was performed in a HiSeq2500 Illumina Sequencer, pair-end 50 bp, in the Weill Cornell Medicine Epigenomics Core. Hi-C data was processed using HiC-bench and HiCExplorer with bowtie aligner[70,71,72] to mouse genome (mm10) and filtering multi-mapped, single-sided, duplicated, low quality, self-ligated reads. All downstream analyses were conducted by accepted intra-chromosomal read pairs. Contact matrices were built by summing the contacts in 5 kb resolution and merging bins for lower-resolution representations as needed. Hi-C data was normalized using either Iterative correction (ICE)[73] or total sum scaling, where each bin was normalized by the total number of reads for a specific chromosome and scaled to 100 million reads. For downstream analyses generally ICE normalized matrices were used with the exception of cases where different bins were compared against each other (such as ranking total connectivity of different promoters against each other), in which case total sum scaling was preferred over ICE. A/B compartments were identified using CScoreTool on 100 kb resolution ICE normalized matrices, splitting each chromosome into two group based on positive and negative C-scores, and designating the A compartment as the side with most overlap with the H3K27ac peaks and reads[74]. Topologically associating domains (TADs) were calculated using insulation square method[75] on ICE normalized matrices. Virtual 3C figures were plotted by taking contacts between two anchor points. Anchors for enhancer-promoter loops were defined as follows. Putative enhancer peaks defined from H3K27ac peaks were associated to all of the genes with transcription start sites within 150 kb of the enhancer peak (many-to-many enhancer-gene mapping). Any enhancer that was not associated with a gene so far were linked only to the closest gene transcription start site that was within 500 kb. Genes that contained enhancers upstream of their TSS were limited to these by removing downstream enhancer associations, however genes with no upstream enhancers kept all their downstream enhancer associations. All Hi-C 5 kb bin or bins that contains the enhancer region were taken, any enhancers that are in consecutive bins (within 5 kb) of another were merged. A similar approach was used to merge bins for genes with multiple TSS for different transcripts that resided in consecutive Hi-C bins. After that the contact scores between these anchor points were taken and plotted as reads per 100 million, and $\log_2$ fold change was calculated by taking the median of all pairwise comparisons between *Smc3*$^{wt/-}$ and *Smc3*$^{wt/wt}$ replicates including a pseudocount of 1. For GSEA analysis all enhancer contacts for a given gene promoter were summed together, and genes were ranked by the delta promoter-enhancer contact score between *Smc3* haploinsufficient and wild-type mice. To find genes that were affected by downregulation and loss of enhancer-promoter connectivity both RNA-seq expression data and the Hi-C contacts were utilized. For this purpose, expression for a specific gene in centrocytes was standardized to *z*-scores, and the same process was also repeated for all enhancer-promoter contacts for a given gene. These standardized scores were fed into an ANoVA model with covariates for genotype and controlling for assay type (as RNA-seq or Hi-C) and were filtered to genes with significant terms for the genotype covariate ($P < 0.01$).

Mint-ChIP Sequencing

Mint-ChIP was performed using H3K27Ac antibody (Cat#39133, lot 31814008) from Active Motif as described[76]. Briefly, 5×10^4 FACS-sorted centrocytes, in triplicates, were processed with MNase to fragment native chromatin. Barcoded adapters were ligated to every sample and samples were multiplexed for chromatin immunoprecipitation (ChIP). After ChIP, material was linearly amplified by RNA in vitro transcription in the presence of RNAse inhibitor RNAseOUT (Invitrogen 10777019). Fragments were reverse-transcribed and amplified as a library. Sequencing was performed in a NextSeq500. The sequencing files were first de-multiplexed and trimmed based the 8 nucleotides barcode using Je demultiplex allowing 1 mismatch on read 1[77]. The demultiplexed reads were then aligned to mm10 genome using Burrows-Wheeler Aligner (BWA)[78]. The duplication reads were then removed from the alignment fils using picard MarkDuplicates with REMOVE_DUPLICATES ("Picard Toolkit." 2019. Broad Institute, GitHub Repository. http://broadinstitute.github.io/picard/; Broad Institute). The H3K27ac peaks were called using Genrich comparing all replicates to their respective H3 total as input and combining p-values of each bin across all $C\gamma1^{wt/cre}$;$Smc3^{wt/wt}$ and $C\gamma1^{wt/cre}$;$Smc3^{wt/-}$ replicates (https://github.com/jsh58/Genrich). Peaks were filtered to an adjusted p value of 0.01. Putative enhancers were defined as H3K27ac peaks that were at least 5 kb away from a transcription start site.

Single-Cell RNA Sequencing

BD Rhapsody data was generated according to the manufacturer's protocol and sequenced on an Illumina NextSeq500 using 75bp × 2 Paired-end Read Sequencing. We used the BDRhapsody™ Immune Response Panel Mm (633753, Doc ID 55673. Rev 2.0) containing primers for 397 genes and a custom gene panel designed with help from BD (see Supplementary Table 2 for probe list) for additional 93 genes. Fastq files were uploaded to Seven Bridges Genomics. Data was demultiplexed and sequences analyzed with BD's Rhapsody pipeline (BD Rhapsody Analysis Pipeline 1.4 Beta) on Seven Bridges (www.sevenbridges.com). Data was then loaded into Seurat (version 3.0.2) for analysis[79]. This generated a sparse matrix file of features by barcodes. This sparse matrix data was then read into R using the R package Seurat 3.0.2, and standard quality control was run to remove cells with few genes. Data was then scaled and normalized. Linear dimensional reduction was performed by calculation of PCA from the most variable genes. Cells were then clustered using a resolution value of 0.5 and visualized by UMAP. Module scores were calculated using the AddModuleScore function with a control value of 5. Individual genes and module scores were projected and used to identify appropriate classification of clusters. Single cells were assigned to cell populations based on cluster score for DZ, LZ or PC gene markers and module scores.

Single-Cell Trajectory Analysis

RNA trajectory analysis was performed using the R package Slingshot version 1.2.0[80]. This package was used to create a "pseudotime" based on a combination of PCA 1 and 2 calculated by Seurat, using the cells identified as DZ as the anchor point. Linage plots were generated by projecting pseudotime onto cells mapped by PCAs 1 and 2. Pseudotime density

plots were generated by genotype and zone using the ggplot2 version 3.2.1 geom_density function. Difference in density was determined by using the density function in R on the pseudotime data for each genotype, then subtracting the wild-type density Y values from the Smc3Het density Y values. These differences were then plotted by ggplot2 geom_col function. Density distribution was determined to be significant by Wilcoxon rank sum test on the pseudotime values for each genotype.

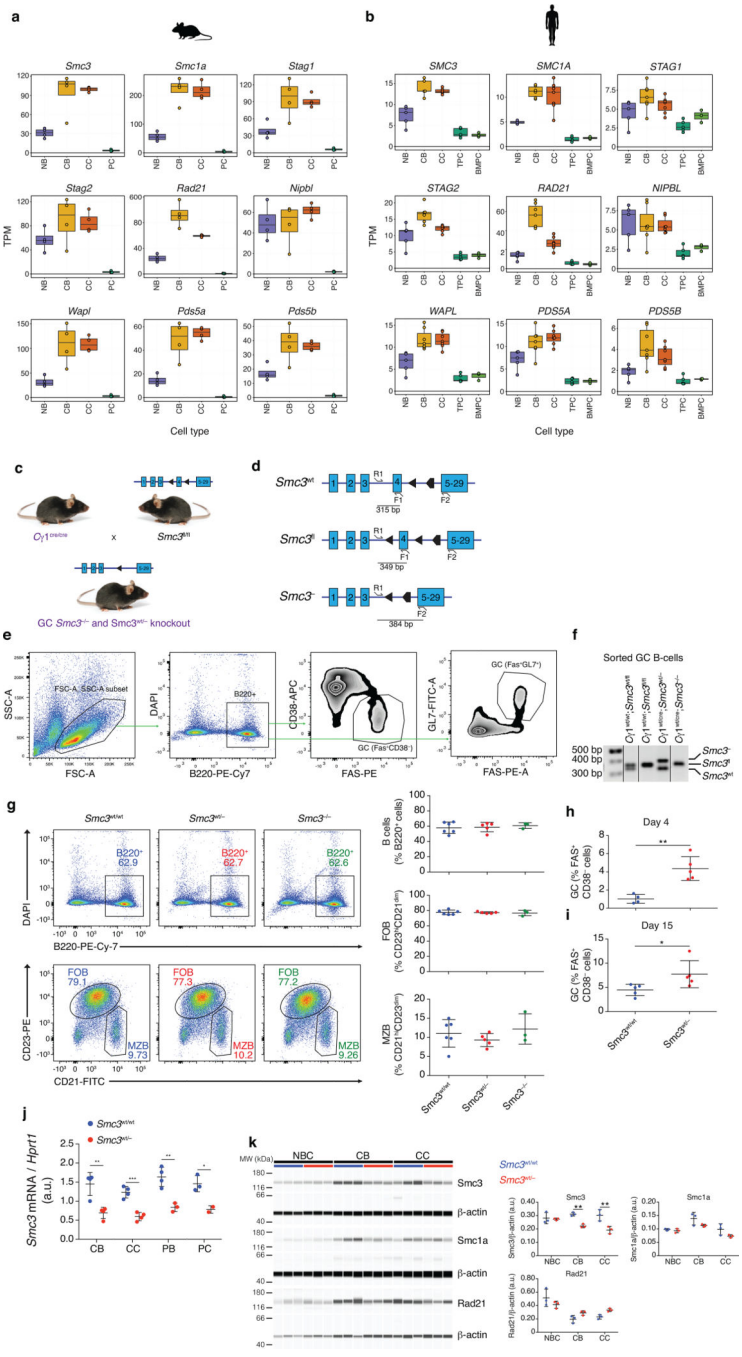
Quantification and statistical analysis

Statistical analysis was performed with Prism software (GraphPad Software). p-values were determined using Student's t tests (two-tailed). For multiple comparisons, One-Way ANoVA followed by Tukey test was applied. Differences between groups were considered significant for P values < 0.05 (*p<0.05; **p<0.01; ***p<0.001; ****p<0.0001). Survival curves were analyzed through Kaplan-Meier analysis and the Log-rank (Mantel-Cox) test was applied. No repeated measures were performed unless specifically stated.

DATA AVAILABILITY STATEMENT

Datasets generated in this manuscript were deposited in the NCBI's Gene Expression Omnibus under the accession number GSE143852.

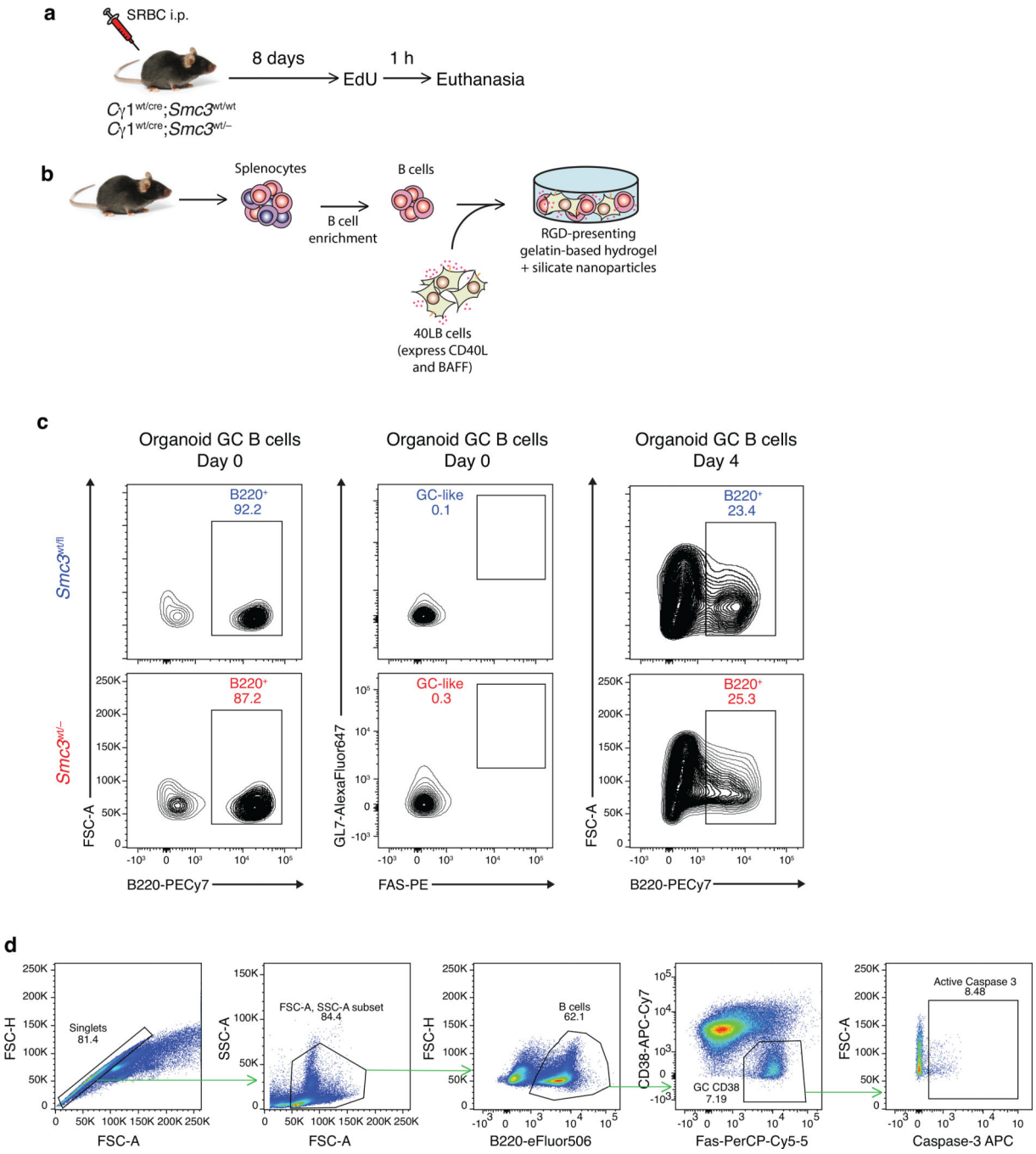
Extended Data



Extended Data Fig. 1. *Smc3* knockout abrogates GC formation whereas *Smc3* haploinsufficiency induces GC Hyperplasia.

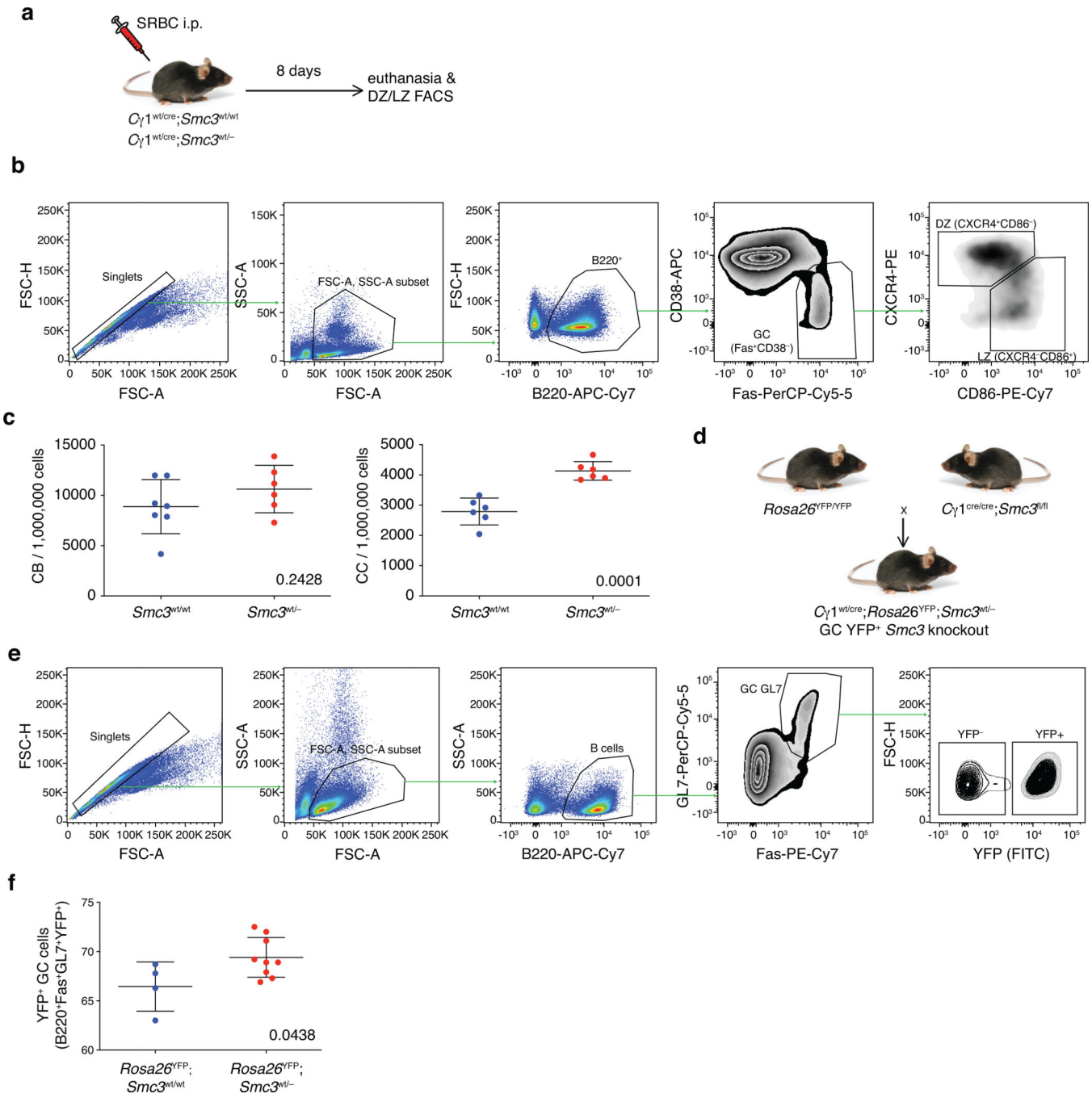
(a) Box plots showing transcripts per million (TPM) of cohesin core subunits (*Smc3*, *Smc1a*, *Stag1*, *Stag2* and *Rad21*) and regulator proteins (*Nipbl*, *Wapl*, *Pds5a*, *Pds5b*) in mouse and human **(b)**. NB: naive B cells, CB: centroblasts, CC: centrocytes, PC: splenic plasma cells, TPC: tonsillar plasma cells, BMPC: bone marrow plasma cells. **(c)** Scheme depicting *Smc3* conditional knockout model breeding strategy and **(d)** genotyping strategy.

(e) Gating strategy used for flow cytometric analysis of GC B cells. **(f)** Genotyping of sorted GC B cells. Representative gel of one experiment out of four performed. **(g)** Representative flow cytometry plots showing total B cells ($B220^+$ cells) and follicular B cells ($B220^+CD23^{hi}CD21^{lo}$ cells) and marginal zone B cell populations ($B220^+CD23^{lo}CD21^{hi}$ cells) and quantitative data for one representative experiment out of 4 performed in $Smc3^{wt/wt}$ (n=6), $Smc3^{wt/-}$ (n=4) and $Smc3^{-/-}$ (n=3) as shown (right). One-way ANoVA followed by Tukey test for multiple comparisons was used. **(h-i)** Quantification of GC B cells, defined as $B220^+Fas^+CD38^-$, 4 days **(h**, p=0.002) or 15 days **(I**, p=0.0434) after SRBC immunization. Two-tail unpaired *t*-test was used. **(j)** RT-qPCR comparing Smc3 mRNA expression in GC B cells (centroblasts, CB and centrocytes, CC), plasmablasts (PB) and plasma cells (PC) in $C\gamma1^{wt/cre};Smc3^{wt/wt}$ and $C\gamma1^{wt/cre};Smc3^{wt/-}$ mice as indicated. Two-tail unpaired *t*-test was used. CB, p=0.0039, CC: p=0.0005, PB: p=0.0041 and PC: p=0.0271. **(k)** Capillary immunoblot analyses comparing Smc3, Smc1a, Rad21 and Sag2 protein levels in centroblasts (CB), centrocytes (CC) and naïve B cells (NBC) from $C\gamma1^{wt/cre};Smc3^{wt/wt}$ (n=3) and $C\gamma1^{wt/cre};Smc3^{wt/-}$ (n=3) mice as indicated. Representative gel our of two experiments performed. Right panels show quantifications of the cohesin complex subunits using β-actin protein as loading control. Two-way ANoVA was used. **p<0.01. In **(a)** and **(b)** box plots show the median as center, first and third quartiles as the box hinges, and whiskers extend to the smallest and largest value no further than the 1.5 × interquartile range (IQR) away from the hinges. In **(g)**, **(h)**, **(i), (j)** and **(k)** data are presented as mean +/− SD.



Extended Data Fig. 2. *Smc3* haploinsufficiency confers proliferative advantage to GC B cells without chromosomal instability.

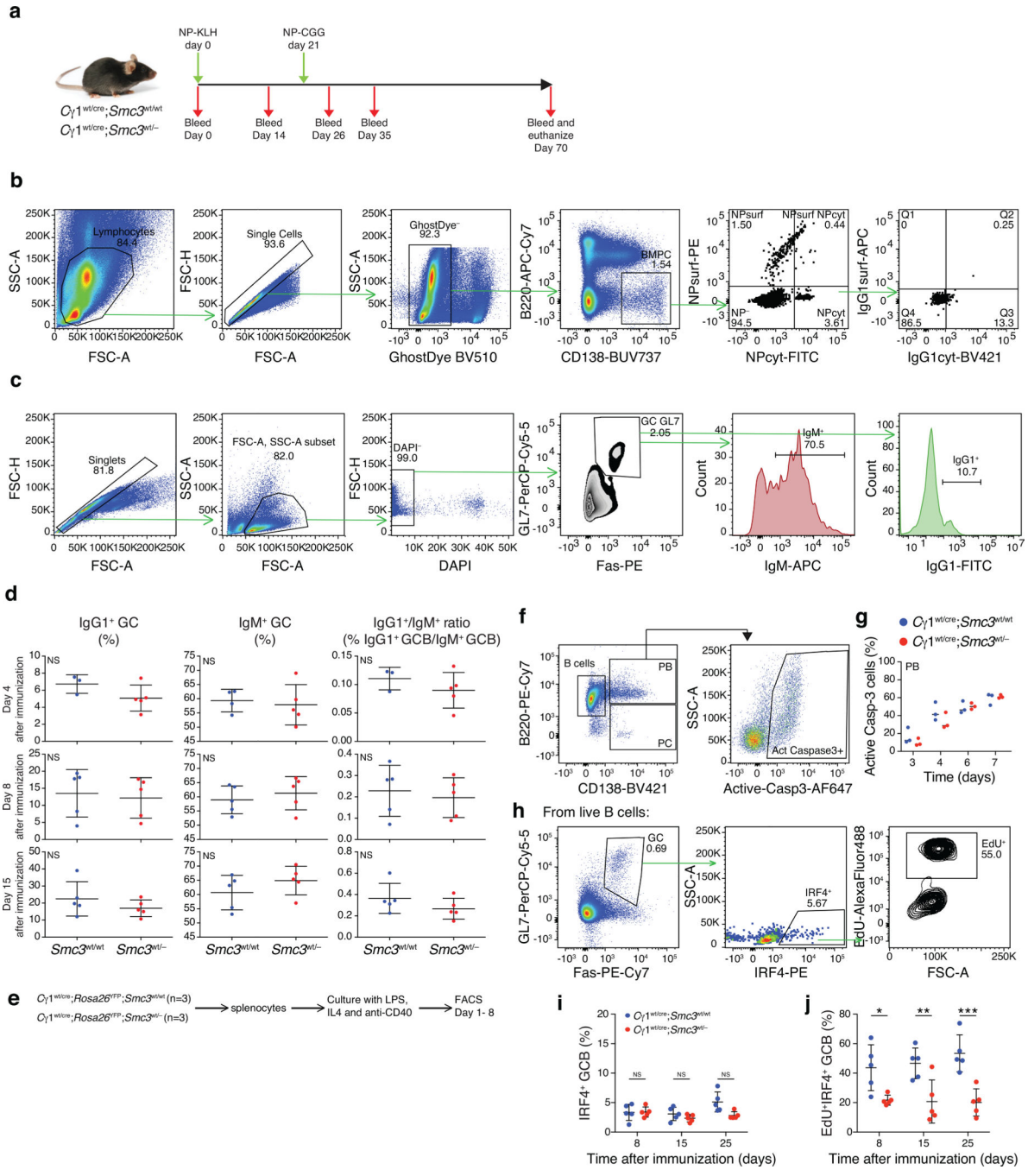
(a) $C\gamma1^{wt/cre};Smc3^{wt/wt}$ (n=5) and $C\gamma1^{wt/cre};Smc3^{wt/-}$ (n=5) were immunized with SRBCs to induce GC formation. One hour prior to euthanasia, mice were injected with 50 mg/kg i.v. 5-ethynyl-2'-deoxyuridine (EdU). **(b)** GC organoids were performed with splenocytes from $C\gamma1^{wt/cre};Smc3^{wt/wt}$ and $C\gamma1^{wt/cre};Smc3^{wt/-}$ mice. **(c)** On day 0 and day 4 after organoid plating, the presence of B cells and GC-like cells ($B220^{+}Fas^{+}GL7^{+}$) was determined by flow cytometry. **(d)** Gating strategy followed to detect apoptotic active Caspase-3^{+} GC B cells.



Extended Data Fig. 3. GC hyperplasia induced by *Smc3* haploinsufficiency is due to expansion of centrocytes.

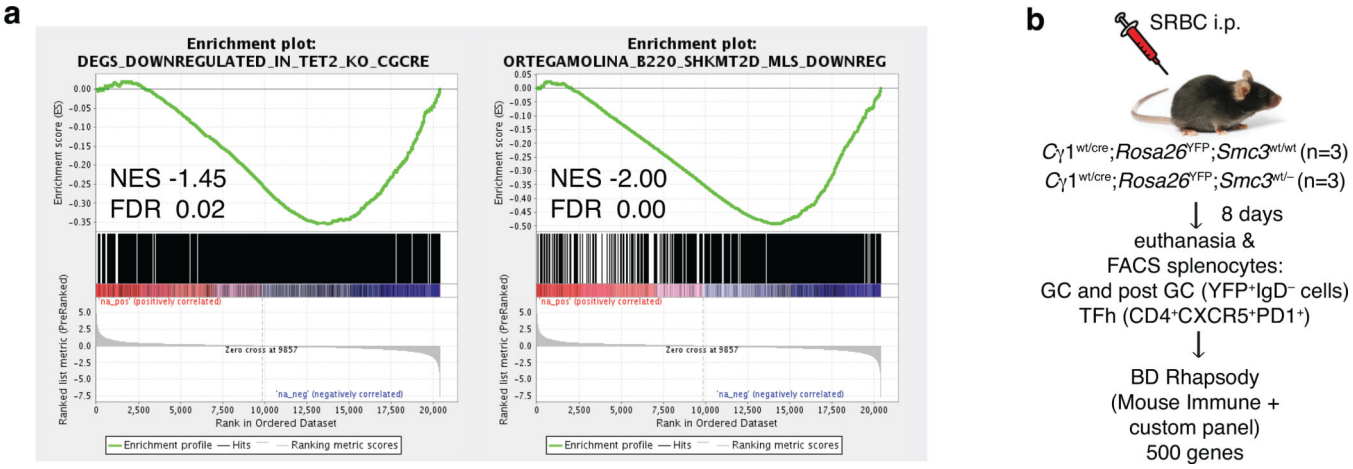
(a) Mice were immunized with sheep red blood cells (SRBCs) to induce GC formation and euthanized 8 days after. **(b)** Gating strategy followed to identify DZ ($B220^{+}Fas^{+}CD38^{-}CXCR4^{+}CD86^{-}$) and LZ ($B220^{+}Fas^{+}CD38^{-}CXCR4^{-}CD86^{+}$) GC B cells. **(c)** Quantification of absolute numbers of DZ centroblasts ($C\gamma1^{wt/cre};Smc3^{wt/wt}$ (n=7) and $C\gamma1^{wt/cre};Smc3^{wt/-}$ (n=6) mice) or LZ centrocytes ($C\gamma1^{wt/cre};Smc3^{wt/wt}$ (n=6) and $C\gamma1^{wt/cre};Smc3^{wt/-}$ (n=6) mice). **(d)** $C\gamma1^{wt/cre};Smc3^{fl/fl}$ were crossed to $Rosa26^{YFP/YFP}$ to produce mice that express

the yellow fluorescent protein when expressing the cre recombinase. **(e)** Gating strategy used in $C\gamma1^{wt/cre};Rosa26^{YFP};Smc3^{wt/-}$ and $C\gamma1^{wt/cre};Rosa26^{YFP};Smc3^{wt/wt}$ mice to identify YFP $^{+}$ GC B cells. **(f)** Plot showing YFP^{+} GC B cells in the reporter $C\gamma1^{wt/cre};Rosa26^{YFP};Smc3^{wt/-}$ (n=4) and $C\gamma1^{wt/cre};Rosa26^{YFP};Smc3^{wt/wt}$ (n=9) mice. In **(c)** and **(f)** data are presented as mean +/− SD.



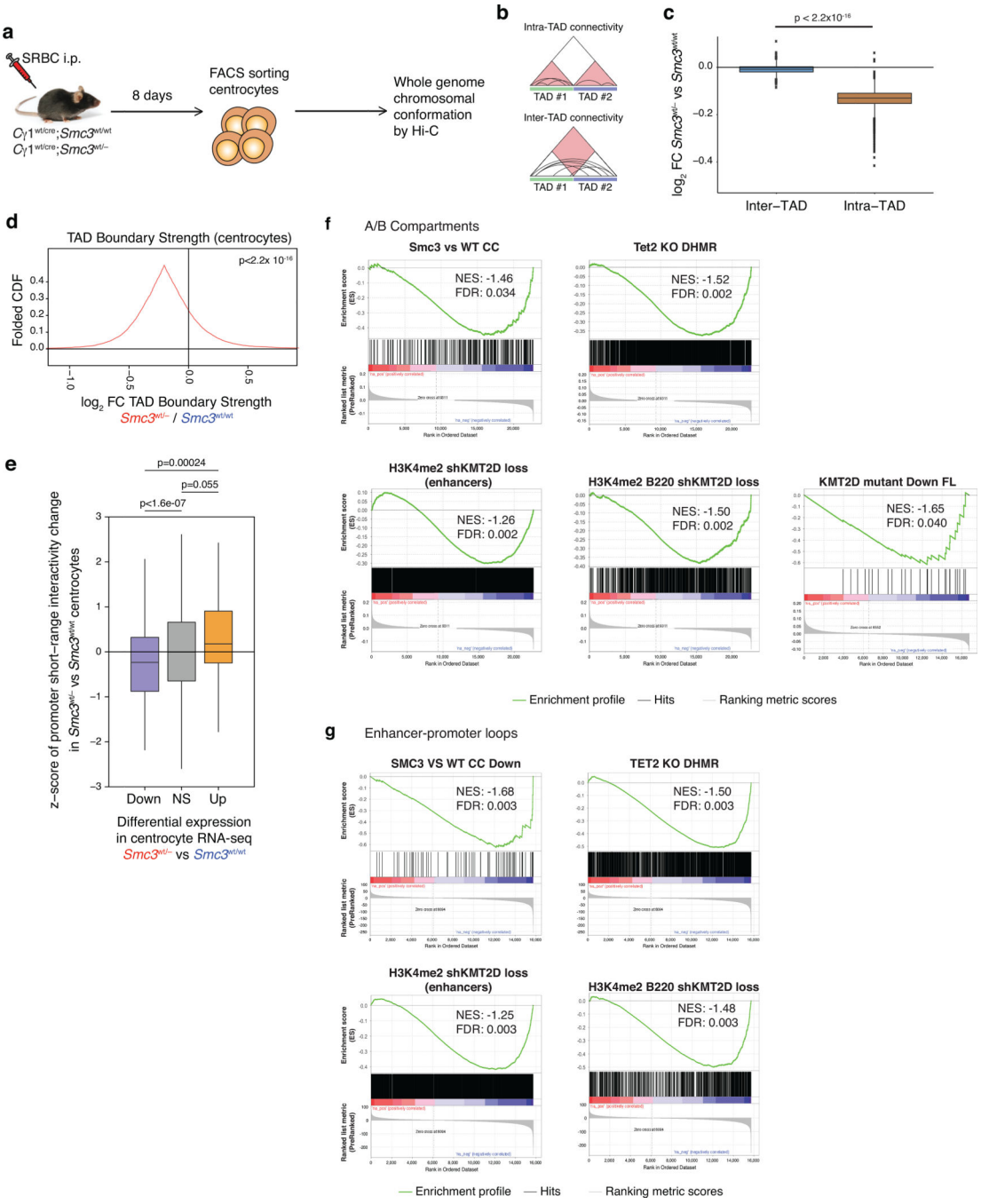
Extended Data Fig. 4. *Smc3* haploinsufficient mice have impaired terminal differentiation

(a) Mice were immunized on day 0 with 100 μg NP-KLH:Alum 1:1 i.p., and received a boost on day 21 of 100 μg NP-CGG: Alum 1:1 i.p. Mice were bleed on days 0, 14, 26, 35 and 70, at which time were euthanized. **(b)** Representative flow cytometry gating strategy followed to determine the presence of bone marrow plasma cells (BMPC, $CD138^{+}B220^{-}$ cells) and long-lived BMPCs ($NPcyt^{+}IgG1cyt^{+}CD138^{+}B220^{-}$ cells). **(c)** Representative flow cytometry gating strategy followed to determine the percentage of IgM- and IgG1-expressing GC B cells. **(d)** Frequency of $IgG1^{+}$, IgM^{+}, or $IgG1^{+}/IgM^{+}$ GC B cells in 4, 8 or 15 day SRBC immunized $C\gamma1^{wt/cre}$;$Smc3^{wt/wt}$ (n=5) or $C\gamma1^{wt/cre}$; $Smc3^{wt/-}$ (n=5) mice. Two tail unpaired *t*-test. All differences are statistically non-significant. **(e)** Splenocytes from $C\gamma1^{wt/cre}$;$Rosa26^{YFP}$;$Smc3^{wt/wt}$ (n=5) and $C\gamma1^{wt/cre}$;$Rosa26^{YFP}$;$Smc3^{wt/-}$ (n=5) mice were cultured in the presence of IL-4, LPS and anti-CD40 antibody and PB/PC followed by flow cytometry and **(f)** levels of active Cspase-3 were detected by flow cytometry using the presented gating strategy. **(g)** Active Caspase-3 levels in B cells, PBs and PCs from ex vivo cultures. **(h-i)** 8, 15 and 25-day immunized mice were sacrificed and splenocytes stained for GC markers. Nuclear staining of Irf4 was done afterwards. **(h)** Representative flow cytometry plots of $IRF4^{+}$ GC B cells and **(i)** plots depicting the percentage of $IRF4^{+}$ GC B cells in $C\gamma1^{wt/cre}$;$Smc3^{wt/wt}$ (n=5) and $C\gamma1^{wt/cre}$;$Smc3^{wt/-}$ (n=5) mice (at every time point). **(j)** At each time point, mice were injected with 50 μg/kg EdU i.p. and sacrificed one hour later. Quantification of $IRF4^{+}$ GC B cells in S-phase in the spleens of SRBC immunized mice for the indicated times in $C\gamma1^{wt/cre}$;$Smc3^{wt/wt}$ (n=5) and $C\gamma1^{wt/cre}$;$Smc3^{wt/-}$ (n=5) mice at every time point. **p<0.01, ***p<0.001 and ****p<0.0001. In **(d)**, **(i)** and **(j)** data are presented as mean +/− SD. **(i)** and **(j)** two-way ANoVA and Sidak multi-comparison test was used.



Extended Data Fig. 5. $Smc3^{wt/-}$ centrocytes fail to upregulate GC exit genes and transition towards the plasma cell lineage.

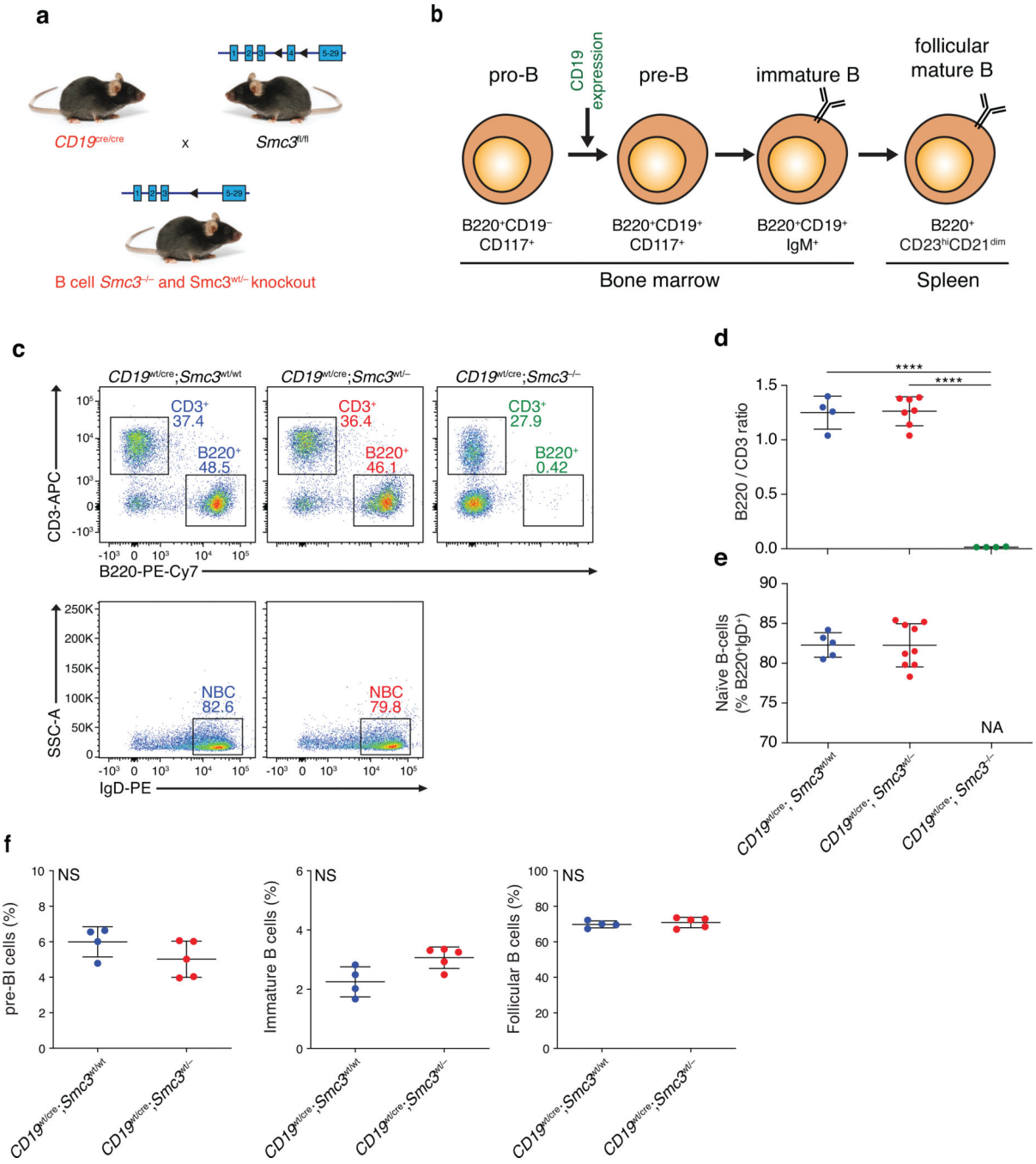
(a) Gene set enrichment analysis (GSEA) plots for TET2_KO and shKMT2D_DOWN. **(b)** Sorting strategy used for single cell RNA-seq by BDRhapsody.



Extended Data Fig. 6. *Smc3* haploinsufficiency disrupts intra-TAD interactions and GC exit gene looping

(a) Genome wide chromosomal interactivity was mapped by in situ Hi-C in $C\gamma1^{wt/cre};Smc3^{wt/wt}$ (n=2) or $C\gamma1^{wt/cre};Smc3^{wt/-}$ (n=2) centrocytes ($B220^{+}Fas^{+}CD38^{-}CXCR4^{-}CD86^{+}$) from 8 day SRBC immunized mice. **(b)** Cartoon depicting the definition of intra- and inter-TAD connectivity. **(c)** Log_2 fold change of inter- and intra-TAD interactions in *Smc3* haploinsufficient versus wild-type centrocytes. p-value expresses significant difference between inter- and intra-TAD interactions. Wilcoxon rank-sum test was used,

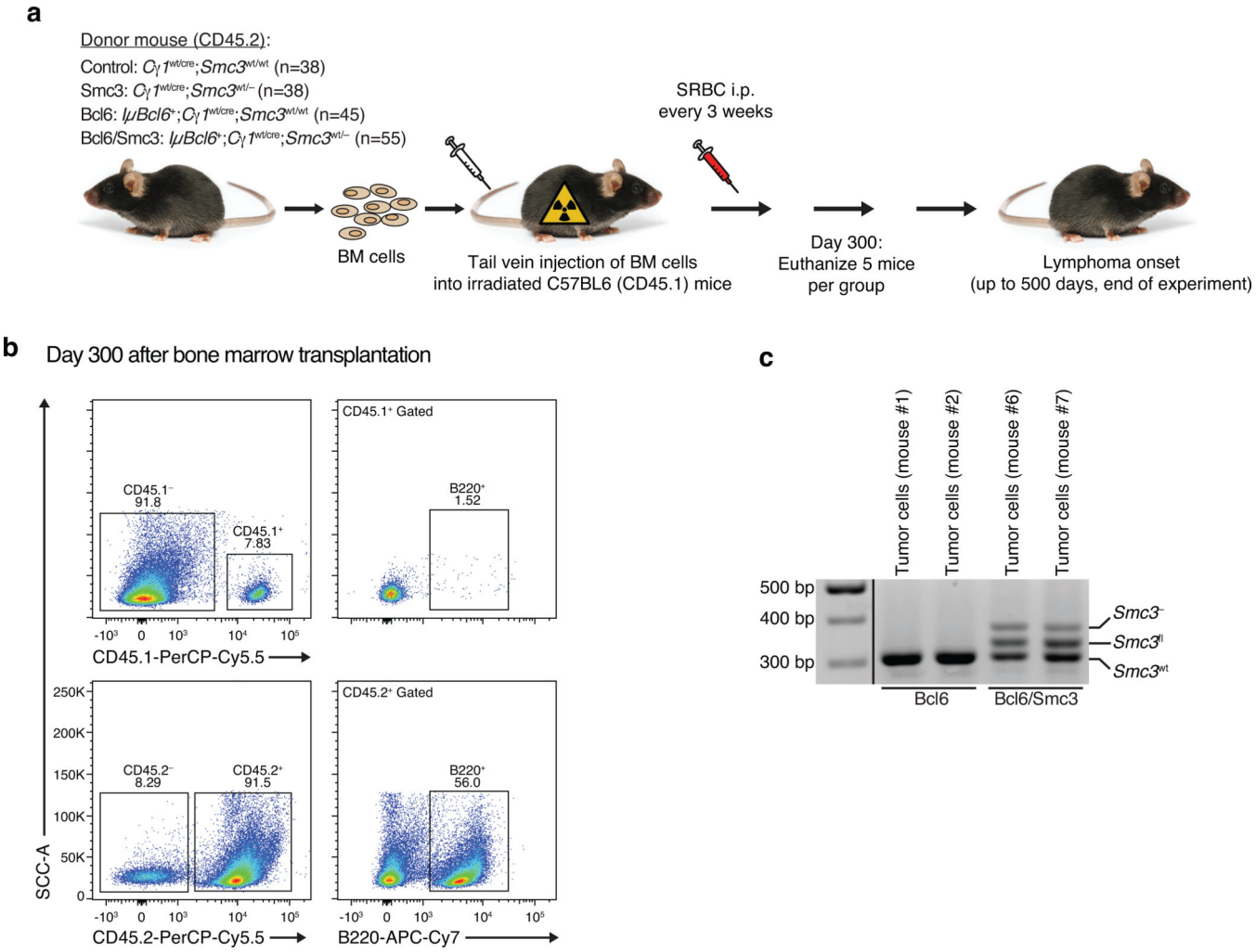
$p<2.2\times10^{-16}$. (**d**) TAD boundary strength for *Smc3* haploinsufficient centrocytes versus wild-type centrocytes by Hi-C. Wilcoxon rank-sum test was used, $p<2.2\times10^{-16}$. **(e)** z-score of promoter short-range interactivity change plotted for genes which are downregulated, non-changing or upregulated in centrocyte RNA-seq of *Smc3* haploinsufficient cells versus control. Wilcoxon rank-sum test was used, Down vs NS $p<1.6\times10^{-7}$, NS vs Up p=0.055 and Down vs Up p=0.00024. **(f)** Gene set enrichment of compartment changes in *Smc3* haploinsufficient vs wild-type centrocytes. Compartment C-scores are calculated in 100 kb intervals and annotated to genes based on the location of their TSS within these bins. Genes are ranked by Δ c-score and analyzed by GSEA. **(g)** Gene set enrichment of enhancer-promoter interactivity changes in *Smc3* haploinsufficient vs wild-type centrocytes. Enhancers are annotated to the nearest TSS (as explained in Methods), normalized contacts between a gene and all its enhancers are summed to get an aggregate promoter-enhancer looping score. Genes are ranked by Δ promoter-enhancer score between *Smc3* haploinsufficient and wild-type mice and analyzed by GSEA. In **(c)** and **(e)** box plots show the median as center, first and third quartiles as the box hinges, and whiskers extend to the smallest and largest value no further than the 1.5 × interquartile range (IQR) away from the hinges.



Extended Data Fig. 7. *Smc3* dosage does not impair B cell differentiation and is only required for GC exit.

(a) Scheme depicting *Smc3* conditional knockout model breeding strategy. **(b)** Early B cell differentiation stages and markers used for identification. **(c)** Representative flow cytometry plots of splenic CD3$^+$ T cells, B220$^+$ B cells and naïve B cells in *CD19*$^{wt/cre}$;*Smc3*$^{wt/wt}$ (n=4), *C19*$^{wt/cre}$;*Smc3*$^{wt/-}$ (n=7) and *C19*$^{wt/cre}$;*Smc3*$^{-/-}$ (n=4) and **(d)** quantification of ratio between splenic B220$^+$/CD3$^+$ cells in *Cd19*$^{wt/cre}$;*Smc3*$^{wt/wt}$ (n=4), *Cd19*$^{wt/cre}$;*Smc3*$^{wt/-}$ (n=7) and *Cd19*$^{wt/cre}$;*Smc3*$^{-/-}$ (n=4) (****p<0.0001) or **(e)** naïve B cells in

Cd19$^{wt/cre}$;*Smc3*$^{wt/wt}$ (n=5) and *Cd19*$^{wt/cre}$;*Smc3*$^{wt/-}$ (n=9). **(f)** Quantification of the percentage of bone marrow pre-BI cells, immature B cells, and splenic follicular B cells in *Cd19*$^{wt/cre}$;*Smc3*$^{wt/wt}$ (n=4) and *Cd19*$^{wt/cre}$;*Smc3*$^{wt/-}$ (n=5). NS: non-significant differences. In **(d)**, **(e)**, and **(f)** data are presented as mean +/− SD. In **(d)** and **(e)** one-way ANoVA and Tukey multi-comparison test was used, and two-way unpaired *t*-test in **(f)**.



Extended Data Fig. 8. Haploinsufficiency of *Smc3* cooperates with *Bcl6* to induce lymphoma.
(a) Bone marrow transplantation experimental design. Bone marrow cells were isolated from *Cγ1*$^{wt/cre}$;*Smc3*$^{wt/wt}$, *Cγ1*$^{wt/cre}$;*Smc3*$^{wt/-}$, *IμBcl6*;*Cγ1*$^{wt/cre}$;*Smc3*$^{wt/w}$ and *IμBcl6*;*Cγ1*$^{wt/cre}$;*Smc3*$^{wt/-}$ CD45.2 donors, and transplanted through tail vein injection into lethally irradiated CD45.1 mice. After engraftment, mice were immunized with SRBCs every 3 weeks for the duration of the experiment. Mice were euthanized when presented with signs of disease. **(b)** Representative flow cytometry plots performed 300 days after engraftment showing that all B220$^+$ are CD45.2$^+$ and CD45.1$^-$, demonstrating that all B cells in these mice come from the donor and not from the recipient. **(c)** Representative allele occurrence was showed by performing PCR in DNA from tumors from

IμBcl6;*Cγ1*$^{wt/cre}$;*Smc3*$^{wt/w}$ and *IμBcl6*;*Cγ1*$^{wt/cre}$;*Smc3*$^{wt/–}$ mice. One representative gel is shown out of 2 experiments performed.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

ACKNOWLEDGEMENTS

The authors thank J. Yan and B. Ren (UCSD) for their help in setting up the situ Hi-C protocol. Sequencing was performed in the Epigenomics Core and the Genomics Core, Weill Cornell Medicine. The authors thank J. McCormick and T. Baumgartner from the Flow Cytometry Core Facility at Weill Cornell Medicine and G. Nanjangud from the Molecular Genetics Core at MSKCC for expert assistant with protocols. A.M.M. is supported by NCI/NIH R35 CA220499, NCI/NIH P01 CA229086-01A1 (together with A.T.), LLS-SCOR 7012-16, LLS TRP 6572-19, the Samuel Waxman Cancer Research Foundation, The Follicular Lymphoma Consortium and the Chemotherapy Foundation. M.A.R. was a recipient of a Postdoctoral Fellowship Grant from the Lymphoma Research Foundation and is funded by NHLBI 5 T32 HL 135465-3 training grant. O.E. is supported by NIH grants UL1TR002384, R01CA194547, LLS SCOR grants 180078-02, 7021-20. R.L.L. was supported by NCI P30 CA 008748, by NCI R35 CA197594-01A1, NCI R01 CA216421, and NCI PS-OC U54 CA143869-05. A.D.V. is supported by NCI K08 CA215317, the William Raveis Charitable Fund Fellowship of the Damon Runyon Cancer Research Foundation (DRG 117-15), and an Evans MDS Young Investigator grant from the Edward P. Evans Foundation. C.E.M. would like to thank the Scientific Computing Unit (SCU), XSEDE Supercomputing Resources, as well as the Starr Cancer Consortium: I7-A765, I9-A9-071, I13-0052 and funding from the WorldQuant Foundation, The Pershing Square Sohn Cancer Research Alliance, the NIH (grants R01CA249054, R01AI151059, P01CA214274, R01 AI125416-03, and R21 AI129851-02), and the Leukemia and Lymphoma Society (LLS) grants LLS 9238-16, and LLS-MCL-982. A.T. is funded by The American Cancer Society RSG-15-189-01-RMC, St. Baldrick's Scholar Award, NCI/NIH Cancer Center Support Grant P30CA016087 and NCI/NIH P01CA229086-01A1 (together with A.M.M.).

COMPETING INTERESTS STATEMENT

A.M.M. receives research funding from Janssen Pharmaceuticals, Sanofi, and Daiichi Sankyo, has consulted for Epizyme and Constellation and is on the scientific advisory board of KDAC pharmaceuticals. R.L.L. is on the Supervisory Board of Qiagen, was on the scientific advisory board of Loxo, Imago, Mana, Auron, C4 Therapeutics and Isoplexis which include equity interest. R.L.L. receives research support from and consulted for Celgene and Roche and consults for Incyte, Lilly, Janssen, Astellas, Morphosys and Novartis. R.L.L. receives research support from Prelude and has received honoraria from AstraZeneca, Roche, Lilly and Amgen for invited lectures and from Gilead for grant reviews. O.E. is supported by Janssen and Eli Lilly research grants. O.E. is scientific advisor and equity holder in Freenome, Owkin, Volastra Therapeutics and One Three Biotech. C.E.M. is a cofounder and board member for Biotia and Onegevity Health, as well as an advisor or compensated speaker for Abbvie, Acuamark Diagnostics, ArcBio, BioRad, DNA Genotek, Genialis, Genpro, Karius, Illumina, New England Biolabs, QIAGEN, Whole Biome, and Zymo Research. CM is a consultant and analyst for Onegevity Health. M.A.R., C.R.C., M.F.C., D.K., B.B., A.K., A.D.V., M.R.T., D.M.N., A.S.D., X.W., W.B., M.T.C.F., H.S., Z.C. and A.T. have no relevant competing interests to declare.

REFERENCES

1. Mesin L, Ersching J & Victora GD Germinal Center B Cell Dynamics. Immunity 45, 471–482 (2016). [PubMed: 27653600]
2. Mlynarczyk C, Fontan L & Melnick A Germinal center-derived lymphomas: The darkest side of humoral immunity. Immunol Rev 288, 214–239 (2019). [PubMed: 30874354]
3. Shapiro-Shelef M & Calame K Regulation of plasma-cell development. Nat Rev Immunol 5, 230–242 (2005). [PubMed: 15738953]
4. Nutt SL, Taubenheim N, Hasbold J, Corcoran LM & Hodgkin PD The genetic network controlling plasma cell differentiation. Semin Immunol 23, 341–349 (2011). [PubMed: 21924923]
5. Dekker J, Marti-Renom MA & Mirny LA Exploring the three-dimensional organization of genomes: interpreting chromatin interaction data. Nat Rev Genet 14, 390–403 (2013). [PubMed: 23657480]
6. Yu M & Ren B The Three-Dimensional Organization of Mammalian Genomes. Annu Rev Cell Dev Biol 33, 265–289 (2017). [PubMed: 28783961]

7. Dixon JR et al. Topological domains in mammalian genomes identified by analysis of chromatin interactions. Nature 485, 376–380 (2012). [PubMed: 22495300]
8. Jin F et al. A high-resolution map of the three-dimensional chromatin interactome in human cells. Nature 503, 290–294 (2013). [PubMed: 24141950]
9. Kagey MH et al. Mediator and cohesin connect gene expression and chromatin architecture. Nature 467, 430–435 (2010). [PubMed: 20720539]
10. Losada A Cohesin in cancer: chromosome segregation and beyond. Nat Rev Cancer 14, 389–393 (2014). [PubMed: 24854081]
11. Peters JM, Tedeschi A & Schmitz J The cohesin complex and its roles in chromosome biology. Genes Dev 22, 3089–3114 (2008). [PubMed: 19056890]
12. Ladurner R et al. Cohesin's ATPase activity couples cohesin loading onto DNA with Smc3 acetylation. Curr Biol 24, 2228–2237 (2014). [PubMed: 25220052]
13. Hu B et al. ATP hydrolysis is required for relocating cohesin from sites occupied by its Scc2/4 loading complex. Curr Biol 21, 12–24 (2011). [PubMed: 21185190]
14. White JK et al. Genome-wide generation and systematic phenotyping of knockout mice reveals new roles for many genes. Cell 154, 452–464 (2013). [PubMed: 23870131]
15. Liu J & Krantz ID Cornelia de Lange syndrome, cohesin, and beyond. Clin Genet 76, 303–314 (2009). [PubMed: 19793304]
16. Casola S et al. Tracking germinal center B cells expressing germ-line immunoglobulin gamma1 transcripts by conditional gene targeting. Proc Natl Acad Sci U S A 103, 7396–7401 (2006). [PubMed: 16651521]
17. Viny AD et al. Dose-dependent role of the cohesin complex in normal and malignant hematopoiesis. J Exp Med 212, 1819–1832 (2015). [PubMed: 26438361]
18. Beguelin W et al. EZH2 enables germinal centre formation through epigenetic silencing of CDKN1A and an Rb-E2F1 feedback loop. Nat Commun 8, 877 (2017). [PubMed: 29026085]
19. Thomas-Claudepierre AS et al. The cohesin complex regulates immunoglobulin class switch recombination. J Exp Med 210, 2495–2502 (2013). [PubMed: 24145512]
20. Zhang X et al. Fundamental roles of chromatin loop extrusion in antibody class switching. Nature 575, 385–389 (2019). [PubMed: 31666703]
21. Ortega-Molina A et al. The histone lysine methyltransferase KMT2D sustains a gene expression program that represses B cell lymphoma development. Nat Med 21, 1199–1208 (2015). [PubMed: 26366710]
22. Soro PG et al. Differential involvement of the transcription factor Blimp-1 in T cell-independent and -dependent B cell differentiation to plasma cells. J Immunol 163, 611–617 (1999). [PubMed: 10395648]
23. Scharer CD, Barwick BG, Guo M, Bally APR & Boss JM Plasma cell differentiation is controlled by multiple cell division-coupled epigenetic programs. Nat Commun 9, 1698 (2018). [PubMed: 29703886]
24. Klein U et al. Transcription factor IRF4 controls plasma cell differentiation and class-switch recombination. Nat Immunol 7, 773–782 (2006). [PubMed: 16767092]
25. Victora GD et al. Identification of human germinal center light and dark zone cells and their relationship to human B-cell lymphomas. Blood 120, 2240–2248 (2012). [PubMed: 22740445]
26. Agirre X et al. Long non-coding RNAs discriminate the stages and gene regulatory states of human humoral immune response. Nat Commun 10, 821 (2019). [PubMed: 30778059]
27. Cho RJ et al. Transcriptional regulation and function during the human cell cycle. Nat Genet 27, 48–54 (2001). [PubMed: 11137997]
28. Whitfield ML et al. Identification of genes periodically expressed in the human cell cycle and their expression in tumors. Mol Biol Cell 13, 1977–2000 (2002). [PubMed: 12058064]
29. Su AI et al. A gene atlas of the mouse and human protein-encoding transcriptomes. Proc Natl Acad Sci U S A 101, 6062–6067 (2004). [PubMed: 15075390]
30. Rosenwald A et al. The use of molecular profiling to predict survival after chemotherapy for diffuse large-B-cell lymphoma. N Engl J Med 346, 1937–1947 (2002). [PubMed: 12075054]

31. Zeller KI et al. Global mapping of c-Myc binding sites and target gene networks in human B cells. Proc Natl Acad Sci U S A 103, 17834–17839 (2006). [PubMed: 17093053]
32. Dominguez PM et al. TET2 Deficiency Causes Germinal Center Hyperplasia, Impairs Plasma Cell Differentiation, and Promotes B-cell Lymphomagenesis. Cancer Discov 8, 1632–1653 (2018). [PubMed: 30274972]
33. Victora GD et al. Germinal center dynamics revealed by multiphoton microscopy with a photoactivatable fluorescent reporter. Cell 143, 592–605 (2010). [PubMed: 21074050]
34. Ersching J et al. Germinal Center Selection and Affinity Maturation Require Dynamic Regulation of mTORC1 Kinase. Immunity 46, 1045–1058 e1046 (2017). [PubMed: 28636954]
35. Dominguez-Sola D et al. The proto-oncogene MYC is required for selection in the germinal center and cyclic reentry. Nat Immunol 13, 1083–1091 (2012). [PubMed: 23001145]
36. Rao SS et al. A 3D map of the human genome at kilobase resolution reveals principles of chromatin looping. Cell 159, 1665–1680 (2014). [PubMed: 25497547]
37. Rickert RC, Roes J & Rajewsky K B lymphocyte-specific, Cre-mediated mutagenesis in mice. Nucleic Acids Res 25, 1317–1318 (1997). [PubMed: 9092650]
38. Jiang Y et al. CREBBP Inactivation Promotes the Development of HDAC3-Dependent Lymphomas. Cancer Discov 7, 38–53 (2017). [PubMed: 27733359]
39. Hummel M et al. A biologic definition of Burkitt's lymphoma from transcriptional and genomic profiling. N Engl J Med 354, 2419–2430 (2006). [PubMed: 16760442]
40. Jais JP et al. The expression of 16 genes related to the cell of origin and immune response predicts survival in elderly patients with diffuse large B-cell lymphoma treated with CHOP and rituximab. Leukemia 22, 1917–1924 (2008). [PubMed: 18615101]
41. Lenz G et al. Stromal gene signatures in large-B-cell lymphomas. N Engl J Med 359, 2313–2323 (2008). [PubMed: 19038878]
42. Shaknovich R et al. DNA methylation signatures define molecular subtypes of diffuse large B-cell lymphoma. Blood 116, e81–89 (2010). [PubMed: 20610814]
43. Lin YC et al. Global changes in the nuclear positioning of genes and intra- and interdomain genomic interactions that orchestrate B cell fate. Nat Immunol 13, 1196–1204 (2012). [PubMed: 23064439]
44. Bunting KL et al. Multi-tiered Reorganization of the Genome during B Cell Affinity Maturation Anchored by a Germinal Center-Specific Locus Control Region. Immunity 45, 497–512 (2016). [PubMed: 27637145]
45. Johanson TM et al. Transcription-factor-mediated supervision of global genome architecture maintains B cell identity. Nat Immunol 19, 1257–1264 (2018). [PubMed: 30323344]
46. Ise W & Kurosaki T Plasma cell differentiation during the germinal center reaction. Immunol Rev 288, 64–74 (2019). [PubMed: 30874351]
47. Mazumdar C et al. Leukemia-Associated Cohesin Mutants Dominantly Enforce Stem Cell Programs and Impair Human Hematopoietic Progenitor Differentiation. Cell Stem Cell 17, 675–688 (2015). [PubMed: 26607380]
48. Mullenders J et al. Cohesin loss alters adult hematopoietic stem cell homeostasis, leading to myeloproliferative neoplasms. J Exp Med 212, 1833–1850 (2015). [PubMed: 26438359]
49. Davidson IF et al. DNA loop extrusion by human cohesin. Science 366, 1338–1345 (2019). [PubMed: 31753851]
50. Kim Y, Shi Z, Zhang H, Finkelstein IJ & Yu H Human cohesin compacts DNA by loop extrusion. Science 366, 1345–1349 (2019). [PubMed: 31780627]
51. Zhang J et al. Disruption of KMT2D perturbs germinal center B cell development and promotes lymphomagenesis. Nat Med 21, 1190–1198 (2015). [PubMed: 26366712]
52. Deng W et al. Controlling long-range genomic interactions at a native locus by targeted tethering of a looping factor. Cell 149, 1233–1244 (2012). [PubMed: 22682246]
53. Groschel S et al. A single oncogenic enhancer rearrangement causes concomitant EVI1 and GATA2 deregulation in leukemia. Cell 157, 369–381 (2014). [PubMed: 24703711]
54. Rao SSP et al. Cohesin Loss Eliminates All Loop Domains. Cell 171, 305–320 e324 (2017). [PubMed: 28985562]

55. Fudenberg G, Abdennur N, Imakaev M, Goloborodko A & Mirny LA Emerging Evidence of Chromosome Folding by Loop Extrusion. Cold Spring Harb Symp Quant Biol 82, 45–55 (2017). [PubMed: 29728444]

56. Cuartero S et al. Control of inducible gene expression links cohesin to hematopoietic progenitor self-renewal and differentiation. Nat Immunol 19, 932–941 (2018). [PubMed: 30127433]

57. Sasca D et al. Cohesin-dependent regulation of gene expression during differentiation is lost in cohesin-mutated myeloid malignancies. Blood 134, 2195–2208 (2019). [PubMed: 31515253]

58. De Koninck M & Losada A Cohesin Mutations in Cancer. Cold Spring Harb Perspect Med 6, a026476 (2016). [PubMed: 27742736]

METHODS-ONLY REFERENCES

59. Srinivas S et al. Cre reporter strains produced by targeted insertion of EYFP and ECFP into the ROSA26 locus. BMC Dev Biol 1, 4 (2001). [PubMed: 11299042]

60. Cattoretti G et al. Deregulated BCL6 expression recapitulates the pathogenesis of human diffuse large B cell lymphomas in mice. Cancer Cell 7, 445–455 (2005). [PubMed: 15894265]

61. Silverman J, Suckow MA & Murthy S The IACUC handbook, Third edition. edn. CRC Press, Taylor & Francis Group: Boca Raton, 2014.

62. Harrow J et al. GENCODE: the reference human genome annotation for The ENCODE Project. Genome Res 22, 1760–1774 (2012). [PubMed: 22955987]

63. Dobin A et al. STAR: ultrafast universal RNA-seq aligner. Bioinformatics 29, 15–21 (2013). [PubMed: 23104886]

64. Liao Y, Smyth GK & Shi W featureCounts: an efficient general purpose program for assigning sequence reads to genomic features. Bioinformatics 30, 923–930 (2014). [PubMed: 24227677]

65. Li B, Ruotti V, Stewart RM, Thomson JA & Dewey CN RNA-Seq gene expression estimation with read mapping uncertainty. Bioinformatics 26, 493–500 (2010). [PubMed: 20022975]

66. Love MI, Huber W & Anders S Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. Genome Biol 15, 550 (2014). [PubMed: 25516281]

67. Subramanian A et al. Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles. Proc Natl Acad Sci U S A 102, 15545–15550 (2005). [PubMed: 16199517]

68. Sergushichev AA An algorithm for fast preranked gene set enrichment analysis using cumulative statistic calculation. bioRxiv (2016).

69. Liberzon A et al. Molecular signatures database (MSigDB) 3.0. Bioinformatics 27, 1739–1740 (2011). [PubMed: 21546393]

70. Langmead B & Salzberg SL Fast gapped-read alignment with Bowtie 2. Nat Methods 9, 357–359 (2012). [PubMed: 22388286]

71. Lazaris C, Kelly S, Ntziachristos P, Aifantis I & Tsirigos A HiC-bench: comprehensive and reproducible Hi-C data analysis designed for parameter exploration and benchmarking. BMC Genomics 18, 22 (2017). [PubMed: 28056762]

72. Wolff J et al. Galaxy HiCExplorer: a web server for reproducible Hi-C data analysis, quality control and visualization. Nucleic Acids Res 46, W11–W16 (2018). [PubMed: 29901812]

73. Imakaev M et al. Iterative correction of Hi-C data reveals hallmarks of chromosome organization. Nat Methods 9, 999–1003 (2012). [PubMed: 22941365]

74. Zheng X & Zheng Y CscoreTool: fast Hi-C compartment analysis at high resolution. Bioinformatics 34, 1568–1570 (2018). [PubMed: 29244056]

75. Crane E et al. Condensin-driven remodelling of X chromosome topology during dosage compensation. Nature 523, 240–244 (2015). [PubMed: 26030525]

76. van Galen P et al. A Multiplexed System for Quantitative Comparisons of Chromatin Landscapes. Mol Cell 61, 170–180 (2016). [PubMed: 26687680]

77. Girardot C, Scholtalbers J, Sauer S, Su SY & Furlong EE Je, a versatile suite to handle multiplexed NGS libraries with unique molecular identifiers. BMC Bioinformatics 17, 419 (2016). [PubMed: 27717304]

78. Li H & Durbin R Fast and accurate short read alignment with Burrows-Wheeler transform. Bioinformatics 25, 1754–1760 (2009). [PubMed: 19451168]

79. Stuart T et al. Comprehensive Integration of Single-Cell Data. Cell 177, 1888–1902 e1821 (2019). [PubMed: 31178118]

80. Street K et al. Slingshot: cell lineage and pseudotime inference for single-cell transcriptomics. BMC Genomics 19, 477 (2018). [PubMed: 29914354]

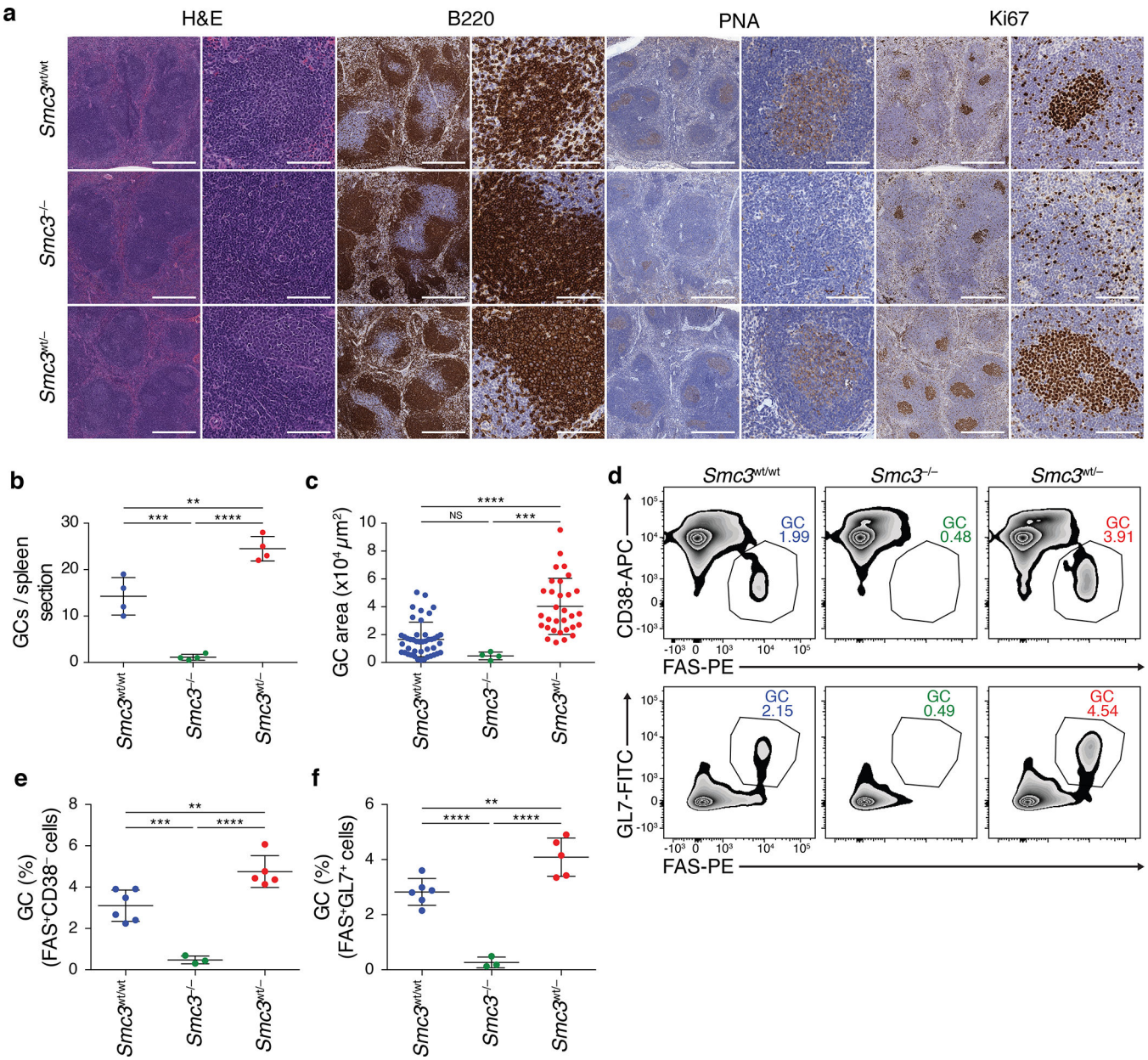


Figure 1. *Smc3* knockout abrogates GC formation whereas *Smc3* haploinsufficiency induces GC Hyperplasia.

(a-f) $C\gamma1^{wt/cre}$;$Smc3^{wt/wt}$ (n=6), $C\gamma1^{wt/cre}$;$Smc3^{wt/-}$ (n=5) and $C\gamma1^{wt/cre}$;$Smc3^{-/-}$ (n=3) were immunized with sheep red blood cells (SRBCs) to induce GC formation. **(a)** Immunohistochemistry (IHC) of GC using peanut-agglutinin (PNA), the proliferation marker Ki67, the B cell marker B220 or hematoxylin and eosin (H&E) in representative tissue sections in $Smc3^{wt/wt}$, $Smc3^{wt/-}$ and $Smc3^{-/-}$ mice as indicated. For pictures in the first, third, fifth and seventh column bar measure is 500 μm, and for pictures in the second, fourth, sixth and eighth column bar measure is 100 μm. **(b)** Quantitative data of the number of GC per spleen area from IHC images and **(c)** quantitative data of the GC area from IHC images. **(d)** Representative flow cytometry plots of GC B cells, defined as $B220^{+}Fas^{+}CD38^{-}$

(top row) or B220^{+}Fas^{+}GL7^{+} (lower row) in *Smc3*$^{wt/wt}$, *Smc3*$^{wt/-}$ and *Smc3*$^{-/-}$ mice. **(e-f)** Quantitative GC B cell data for one representative experiment out of 4 performed as shown. **p<0.01, ***p<0.001 and ****p<0.0001. In **(b)**, **(c)**, **(e)** and **(f)** data are presented as mean +/− SD. In all four cases, one-way ANoVA and Tukey multi-comparison test was used.

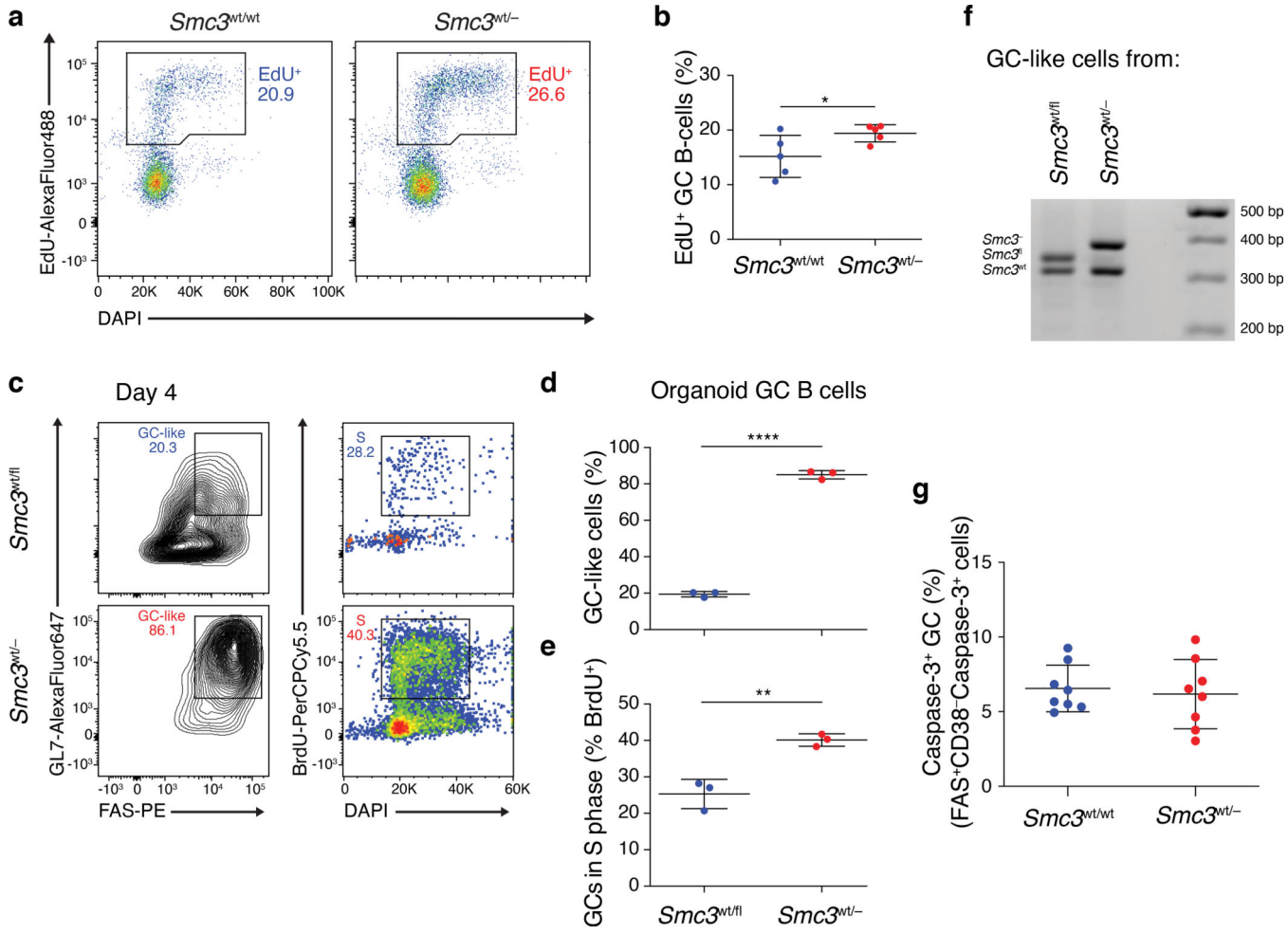


Figure 2. *Smc3* haploinsufficiency confers proliferative advantage to GC B cells without chromosomal instability.

(a-b) *Cγ1*$^{wt/cre}$;*Smc3*$^{wt/wt}$ (n=5) and *Cγ1*$^{wt/cre}$;*Smc3*$^{wt/–}$ (n=5) were immunized with SRBCs to induce GC formation. One hour prior to euthanasia, mice were injected with 50 mg/kg i.v. 5-ethynyl-2'-deoxyuridine (EdU). **(a)** Representative flow cytometry plots of cell cycle distribution of GC B cells and **(b)** quantitative data for one representative experiment out of 2 performed as shown, *p=0.0492. **(c)** Four days after plating, organoids were incubated 30 min in the presence of 3 μg/ml 5-bromo-2'-deoxyuridine (BrdU) and analyzed by flow cytometry for the presence of GC-like cells (B220$^+$Fas$^+$GL7$^+$). Percentage of cells in S-phase of the cell cycle was analyzed by BrdU staining. **(d)** Quantification of the frequency of GC-like cells, ***p<0.0001 and **(e)** S-phase GC-like cells in *Smc3*$^{wt/fl}$ (n=3) and *Smc3*$^{wt/–}$ (n=3), **p=0.0042. **(f)** Genotyping for the detection of *Smc3*^{wt}, *Smc3*^{fl} and *Smc3*$^{–}$ alleles from organoids in figure **(c)**. Genotyping was routinely performed to confirm excision of the *Smc3* allele. **(g)** Quantification of the percentage of Caspase-3$^+$ GC B cells in *Cγ1*$^{wt/cre}$;*Smc3*$^{wt/wt}$ (n=8) and *Cγ1*$^{wt/cre}$;*Smc3*$^{wt/–}$ (n=8), p=0.7079. In **(b)**, **(d)**, **(e)** and **(g)** data are presented as mean +/− SD. In all four cases, two tail unpaired *t*-test was used.

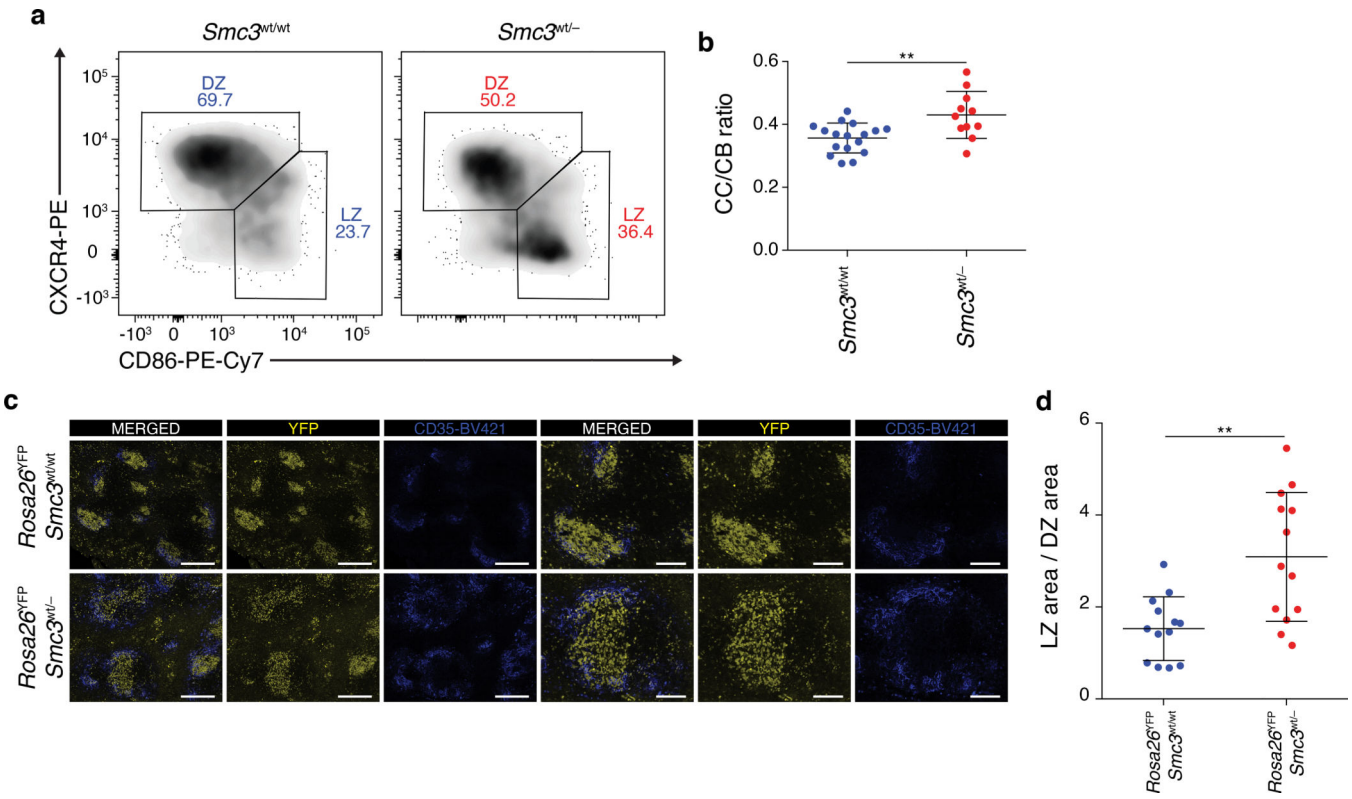


Figure 3. GC hyperplasia induced by *Smc3* haploinsufficiency is due to expansion of centrocytes.

(a-e) Mice were immunized with sheep red blood cells (SRBCs) to induce GC formation and euthanized 8 days after. **(a)** Representative flow cytometry plots of dark zone (DZ) and light zone (LZ) GC B cells, defined as $B220^{+}Fas^{+}CD38^{-}CXCR4^{+}CD86^{-}$ (DZ) or $B220^{+}Fas^{+}CD38^{-}CXCR4^{-}CD86^{+}$ (LZ) in $C\gamma1^{wt/cre};Smc3^{wt/wt}$ and $C\gamma1^{wt/cre};Smc3^{wt/-}$ mice. **(b)** LZ-to-DZ ratio calculated for $C\gamma1^{wt/cre};Smc3^{wt/wt}$ (n=17) and $C\gamma1^{wt/cre};Smc3^{wt/-}$ mice (n=11). **p<0.0036. **(c)** $C\gamma1^{wt/cre};Rosa26^{YFP};Smc3^{wt/wt}$ (n=3) and $C\gamma1^{wt/cre};Rosa26^{YFP};Smc3^{wt/-}$ (n=3) mice were injected with 2.5 μg anti-CD35-BV421 at day 7 after immunization, and sacrificed at day 8. Spleens were fixed in 4% paraformaldehyde overnight and cut at 70 μm in a vibratome, mounted and imaged in a Zeiss LSM 880 Airyscan High Resolution Detector confocal microscope. For pictures in the first to third column bar measure is 250 μm, and for pictures in the fourth to sixth column bar measure is 100 μm. **(d)** LZ to DZ area was calculated using Image J as $CD35\text{-}BV421^{+}YFP$ surface area / YFP surface area. **P<0.0014. In **(b)** and **(d)** data are presented as mean +/− SD. In both cases, two tail unpaired *t*-test was used.

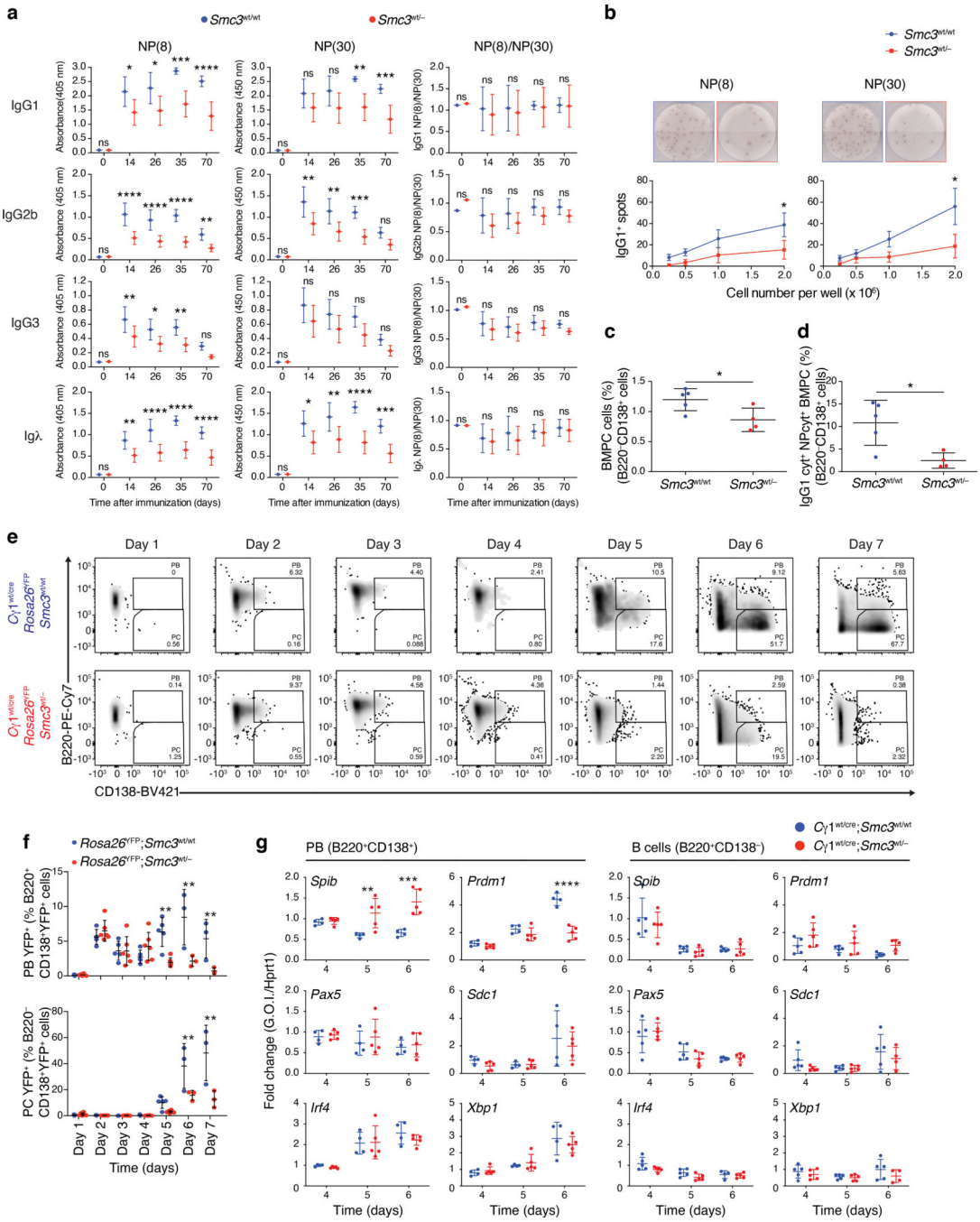


Figure 4. *Smc3* haploinsufficiency impairs plasma cell differentiation in vivo.

(a) Mice were immunized on day 0 with 100 μg NP-KLH:Alum 1:1 i.p., and received a boost on day 21 of 100 μg NP-CGG:Alum 1:1 i.p. Mice were bleed on days 0, 14, 26, 35 and 70, at which time were euthanized. ELISA of anti-$NP_{(8)}$ or $NP_{(30)}$ immunoglobulin isotypes (IgG1, IgG2b, IgG3 and Igλ) in serum of 14, 26, 35 and 70 day NP-KLH/NP-CGG-immunized $C\gamma1^{wt/cre}$;$Smc3^{wt/wt}$ (n=5) and $C\gamma1^{wt/cre}$;$Smc3^{wt/-}$ (n=5) mice. All comparisons versus $C\gamma1^{wt/cre}$;$Smc3^{wt/wt}$. ns: no significant statistical differences, *p<0.05, **p<0.01, ***p<0.001, ****p<0.0001. Two-way ANoVA and Sidak multi-comparison test

was used. **(b)** ELISpot analyses and quantification of $NP_{(8)}$ and $NP_{(30)}$ specific IgG1 in bone marrow cells isolated from $C\gamma1^{wt/cre};Smc3^{wt/wt}$ and $C\gamma1^{wt/cre};Smc3^{wt/-}$, as indicated. Two-way ANoVA and Sidak multi-comparison test was used. *p<0.05. **(c)** Quantification of bone marrow plasma cell (BMPC) from 70 day-immunized $C\gamma1^{wt/cre};Smc3^{wt/wt}$ (n=5) and $C\gamma1^{wt/cre};Smc3^{wt/-}$ (n=4) mice. *p=0.033. Two tail unpaired *t*-test was used. **(d)** Quantification of $IgG1^+$ cytoplasmatic NP^+ cytoplasmatic BMPC from 70 day-immunized $C\gamma1^{wt/cre};Smc3^{wt/wt}$ (n=5) and $C\gamma1^{wt/cre};Smc3^{wt/-}$ (n=4) mice. *p=0.0156. Two tail unpaired *t*-test was used. **(e)** Splenocytes from $C\gamma1^{wt/cre};Rosa26^{YFP};Smc3^{wt/wt}$ (n=5) and $C\gamma1^{wt/cre};Rosa26^{YFP};Smc3^{wt/-}$ (n=5) mice were cultured in the presence of IL-4, LPS and anti-CD40 antibody and plasmablast (PB) and plasma cell (PC) occurrence was assessed by flow cytometry. Day 1 through day 7 flow cytometry plots are shown in **(e)** and PB ($B220^+CD138^+$) and PC ($B220^-CD138^+$) quantification is shown on panel **(f).** Multiple *t*-test with Holm-Sidak correction for multiple comparison was used. Adjusted p-values for PB at days 5, 6 and 7: 0.000253, 0.000090 and 0.004262, and for PC at days 6 and 7: 0.000076 and <0.000001, respectively. **(g)** mRNA levels were determined by RT-qPCR performed in total RNA from $B220^+CD138^+$ PB from ex vivo cultures. Two-way ANoVA and Sidak multi-comparison test was used. **p<0.01, ***p<0.001, ****p<0.0001. In **(a)**, **(b)**, **(c), (d), (f)** and **(g)** data are presented as mean +/− SD.

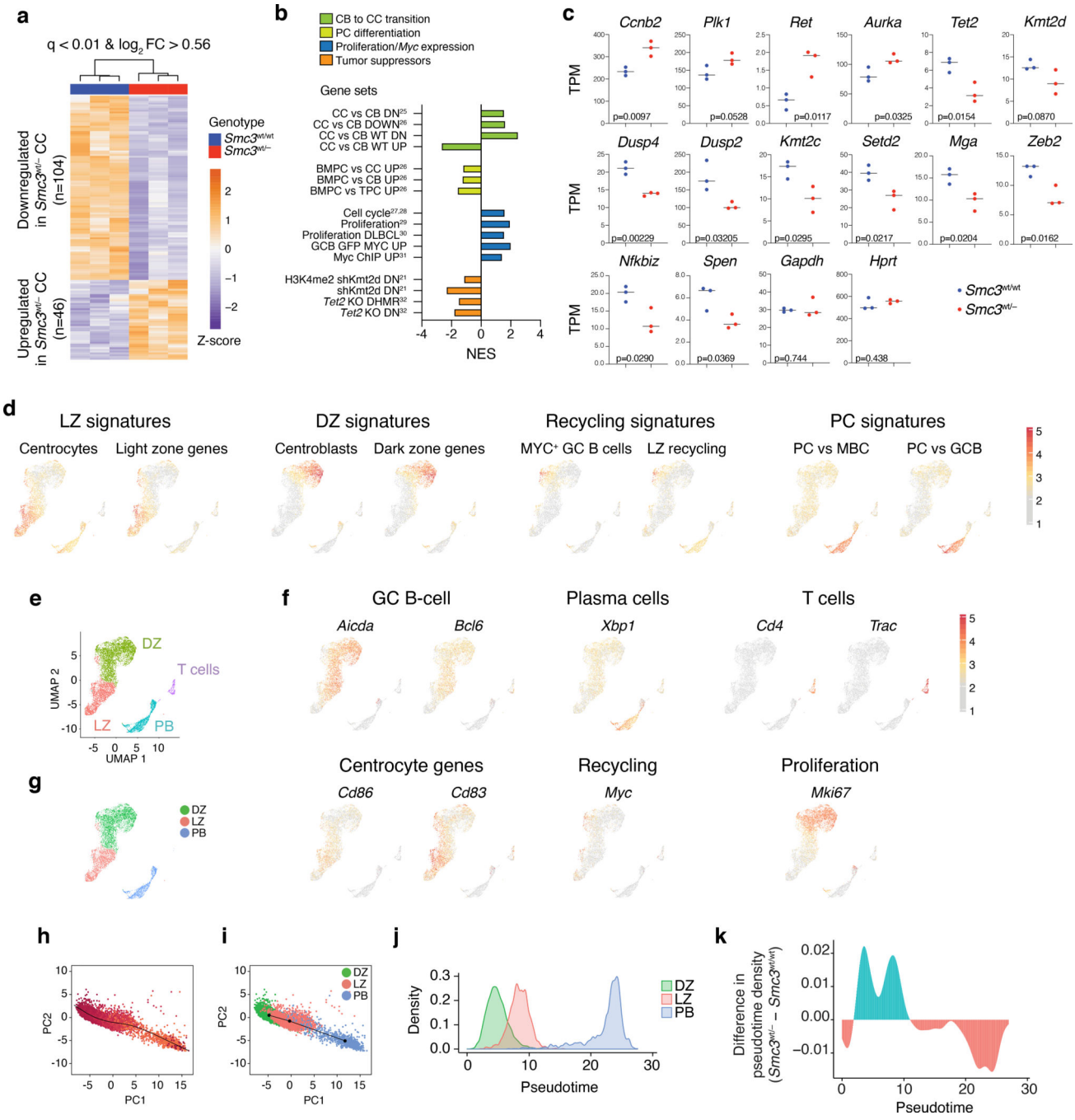


Figure 5. *Smc3*$^{wt/–}$ centrocytes fail to upregulate GC exit genes and transition towards the plasma cell lineage.

(a) Supervised hierarchical clustering analysis of centrocyte RNA-sequencing. Heatmap shows differentially expressed genes in *Cγ1*$^{wt/cre}$;*Smc3*$^{wt/–}$ (n=3) versus *Cγ1*$^{wt/cre}$;*Smc3*$^{wt/wt}$ (n=3). **(b)** Gene set enrichment analysis (GSEA) for the differentially expressed gene in *Cγ1*$^{wt/cre}$;*Smc3*$^{wt/–}$ (n=3) versus *Cγ1*$^{wt/cre}$;*Smc3*$^{wt/wt}$ (n=3). **(c)** Gene expression of representative genes involved in proliferation or tumor suppressor genes. **(d-f)** UMAP representation of single-cell RNA-seq by BDRhapsody Immune Response Targeted

Panel (mouse) + custom mouse panel (Supplementary Table 2). **(d)** LZ, DZ, PC and recycling signatures from RNA-seq analysis were projected into UMAP plots. **(e)** Single cells were identified as belonging to one of the following clusters, according to the gene signature expression: light zone (LZ) dark zone (DZ), T cells, or plasmablasts (PB). **(f)** Expression levels of germinal center (*Aicda*, *Bcl6*), plasma cell (*Xbp1*), centrocyte genes (*Cd86*, *Cd86*), T cell (*Cd4*, *Trac*), proliferation (*Mki67*) and recycling genes (Myc) were projected into UMAP plots. **(g)** Slingshot analysis was performed in GC B cells and PB. **(h-i)** Cells were projected in a PCA plot and pseudotimes were derived for the different populations in the GC. **(j)** DZ, LZ and PB density plotted along the differentiation pseudo-time and **(k)** difference in pseudo-time in *Smc3* haploinsufficient minus wild-type cells are shown.

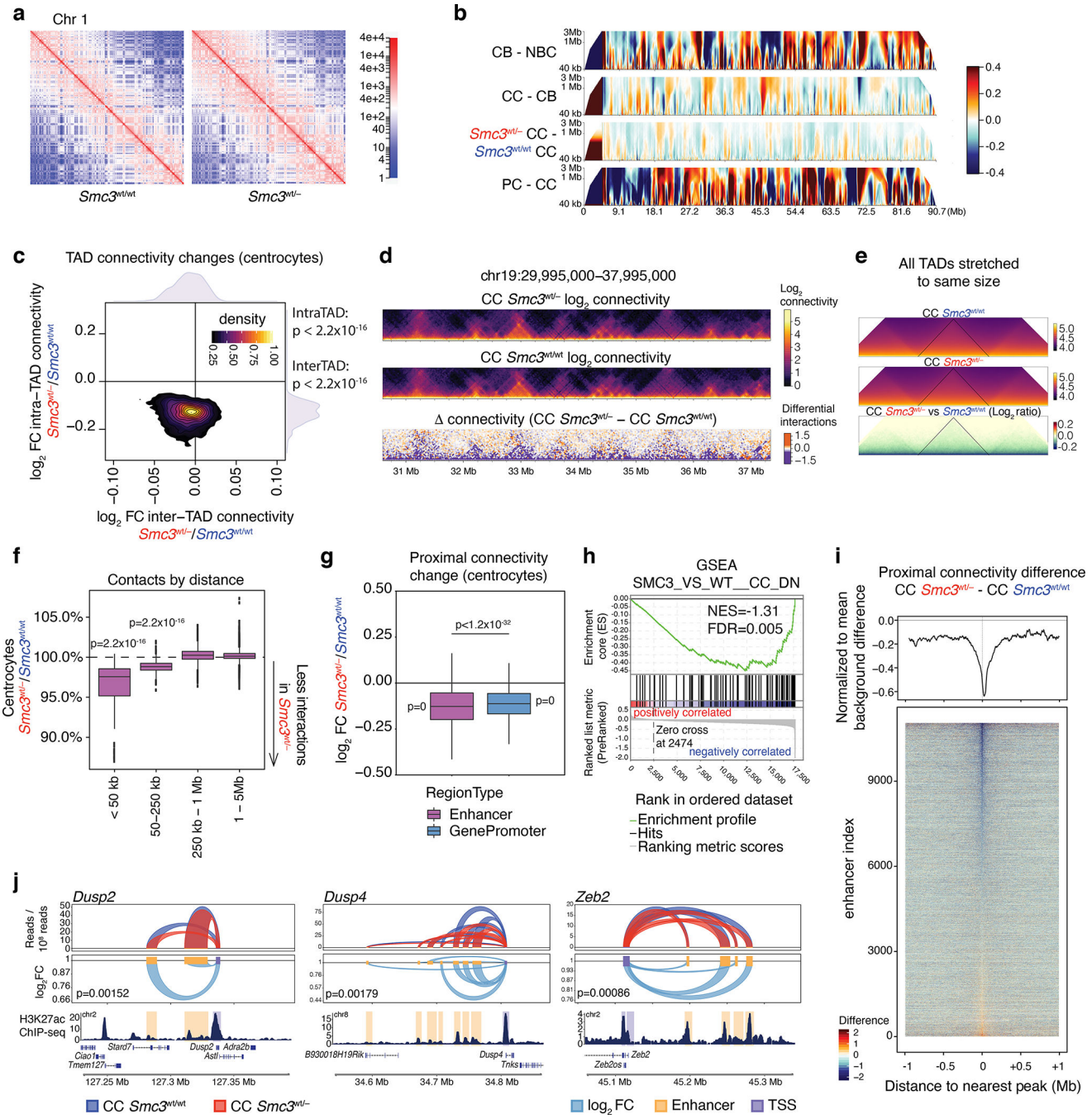


Figure 6. *Smc3* haploinsufficiency disrupts intra-TAD interactions and GC exit gene looping.

(a) (a) Normalized chromosomal contacts for chromosome 1 of combined *Cγ1*$^{wt/cre}$;*Smc3*$^{wt/wt}$ and *Cγ1*$^{wt/cre}$;*Smc3*$^{wt/–}$ centrocytes. **(b)** Insulation contact maps example for chr19 showing insulation for different cell transitions: wild-type centroblasts (CB) minus wild-type naïve B cells (NBC), wild-type centrocytes (CC) minus wild-type CB, *Smc3*-haploinsufficient CC minus wild-type CC, and wild-type PC minus wild-type CC. *x*-axis shows chromosomal location, *y*-axis shows different insulation square sizes. **(c)** $\log_2$ fold change plot showing intra-TAD vs inter-TAD connectivity. Wilcoxon rank-sum test was

used, intra TAD connectivity $p<2.2\times10^{-16}$ and inter TAD connectivity $p<2.2\times10^{-16}$. **(d)** Connectivity maps for a representative region of chr19 showing overall connectivity for *Smc3* haploinsufficient and wild-type CC and the difference between them. **(e)** Meta-TAD analysis summarizing results for all TADs stretched to the same size. Plots show the overall connectivity for centrocyte wild-type (upper), *Smc3* haploinsufficient (middle) and $\log_2$ fold change (lower). **(f)** Difference in genome-wide contacts by distance of centrocytes *Smc3* haploinsufficient normalized over wild-type. Wilcoxon rank-sum test was used, <50 kb $p<2.2\times10^{-16}$, 50–250 kb $p<2.2\times10^{-16}$. **(g)** Proximal connectivity changes by region type (gene promoter or enhancer). Wilcoxon rank-sum test was used, enhancer vs promoter $p<1.2\times10^{-32}$ and enhancers or promoters versus baseline p=0. **(h)** GSEA enrichment analysis for *Smc3* haploinsufficient gene expression signature is enriched among genes that lose interactions by Hi-C. **(i)** Proximal connectivity difference for enhancers in *Smc3* haploinsufficient versus wild-type centrocytes, normalized to mean background difference (top) and tornado plot showing ~12,000 individual enhancers. For each 5 kb bin the aggregate total proximal connectivity was taken by summing any contact within 100 kb of a bin. Delta proximal connectivity in $C\gamma1^{wt/cre};Smc3^{wt/-}$ versus $C\gamma1^{wt/cre};Smc3^{wt/wt}$ was calculated across each bin, and normalized to the background change by shifting the mean to zero. These differences were anchored around putative enhancers from H3K27ac peaks within centrocytes to show overall connectivity loss of centrocyte enhancers. **(j)** virtual 3C analysis of the Hi-C data for representative genes (*Dusp2*, *Dusp4* and *Zeb2*) that lose promoter-enhancer connectivity in *Smc3* haploinsufficient centrocytes. In **(f)** and **(g)** box plots show the median as center, first and third quartiles as the box hinges, and whiskers extend to the smallest and largest value no further than the 1.5 × interquartile range (IQR) away from the hinges.

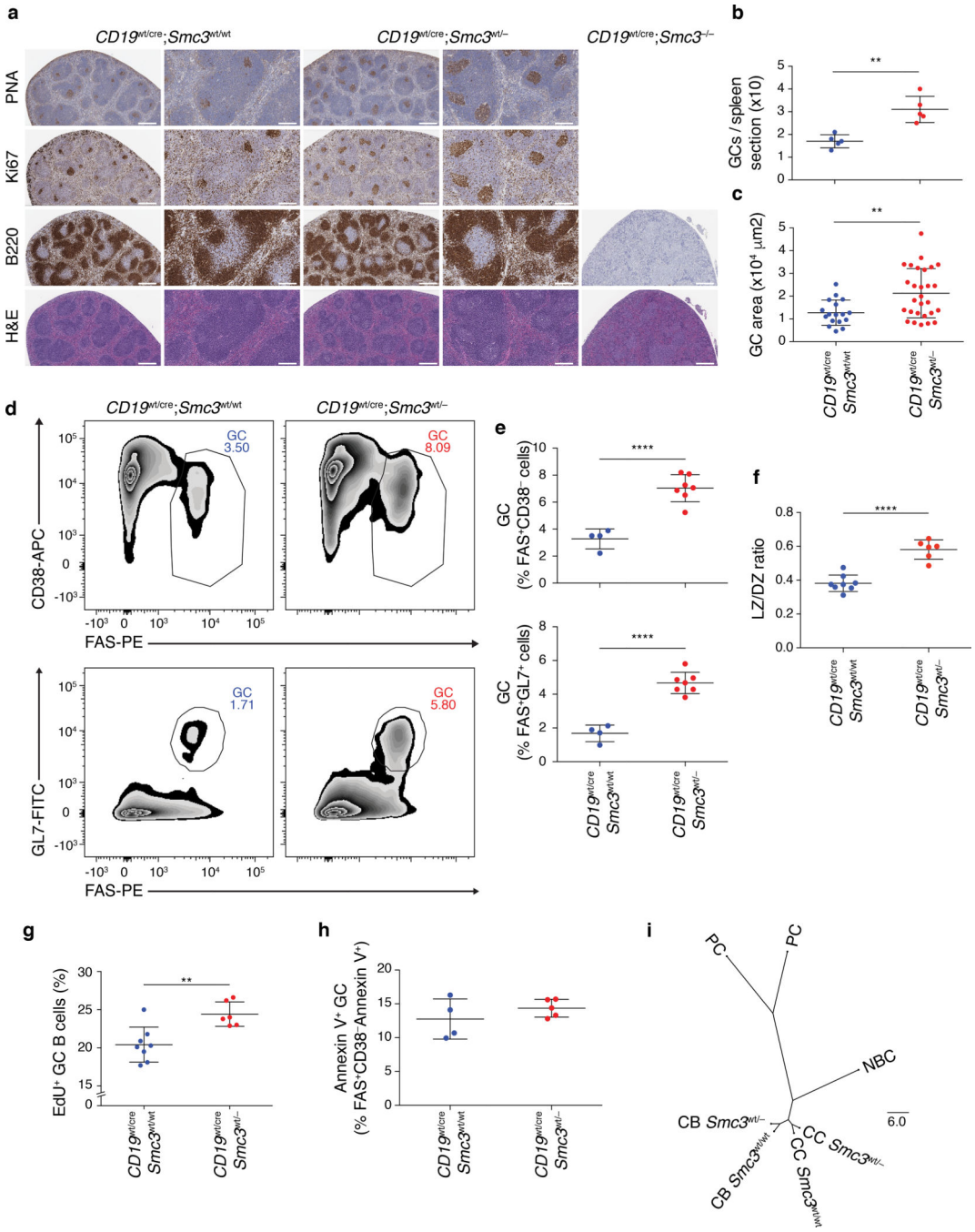


Figure 7. *Smc3* dosage does not impair B cell differentiation and is only required for GC exit.

(a-h) *Cd19*$^{wt/cre}$;*Smc3*$^{wt/wt}$, *Cd19*$^{wt/cre}$;*Smc3*$^{wt/–}$ and *Cd19*$^{wt/cre}$;*Smc3*$^{–/–}$ were immunized with SRBCs to induce GC formation. **(a)** IHC of GC using PNA, Ki67, B220 or H&E in representative tissue sections in *Smc3*$^{wt/wt}$, *Smc3*$^{wt/–}$ and *Smc3*$^{–/–}$ mice as indicated. For pictures in the first, third and fifth column bar measure is 500 µm, and for pictures in the second and fourth column bar measure is 200 µm. One representative experiment is shown out of 3 performed. **(b)** Quantitative data of the number of GC per spleen area from IHC images of *Cd19*$^{wt/cre}$;*Smc3*$^{wt/wt}$ (n=5) and *Cd19*$^{wt/cre}$;*Smc3*$^{wt/–}$ (n=5) and **(c)** quantitative

data of the GC area from IHC images. **p<0.01. **(d)** Representative flow cytometry plots of GC B cells, defined as $B220^{+}Fas^{+}CD38^{-}$ or $B220^{+}Fas^{+}GL7^{+}$ in $Cd19^{wt/cre};Smc3^{wt/wt}$ and $Cd19^{wt/cre};Smc3^{wt/-}$ mice. **(e)** Quantitative GC B cell data for one representative experiment with $Cd19^{wt/cre};Smc3^{wt/wt}$ (n=4) and $Cd19^{wt/cre};Smc3^{wt/-}$ (n=7) out of 3 performed as shown. ****p<0.0001. **(f)** LZ-to-DZ ratio calculated for $Cd19^{wt/cre};Smc3^{wt/wt}$ (n=8) and $Cd19^{wt/cre};Smc3^{wt/-}$ mice (n=6). ****p<0.0001. **(g)** $Cd19^{wt/cre};Smc3^{wt/wt}$ (n=8) and $Cd19^{wt/cre};Smc3^{wt/-}$ mice (n=6) were immunized with SRBCs to induce GC formation. One h prior to euthanasia, mice were injected with 50 mg/kg i.v. EdU. Quantitative data for one representative experiment out of 2 performed as shown (right). **p<0.01. **(h)** Apoptosis levels were determined by detection of Annexin V^{+} in gated GC B cells in $Cd19^{wt/cre};Smc3^{wt/wt}$ (n=4) and $Cd19^{wt/cre};Smc3^{wt/-}$ (n=5). **(i)** Phylogeny trees from compartments scores for plasma cell lines (THB-7 and MPC 11 OUAr), and FACS sorted naive B cells, centroblasts and centrocyte Hi-C. In **(b)**, **(c)**, **(e), (f), (g)** and **(h)** data are presented as mean +/− SD. In all cases, two tail unpaired *t*-test was used.

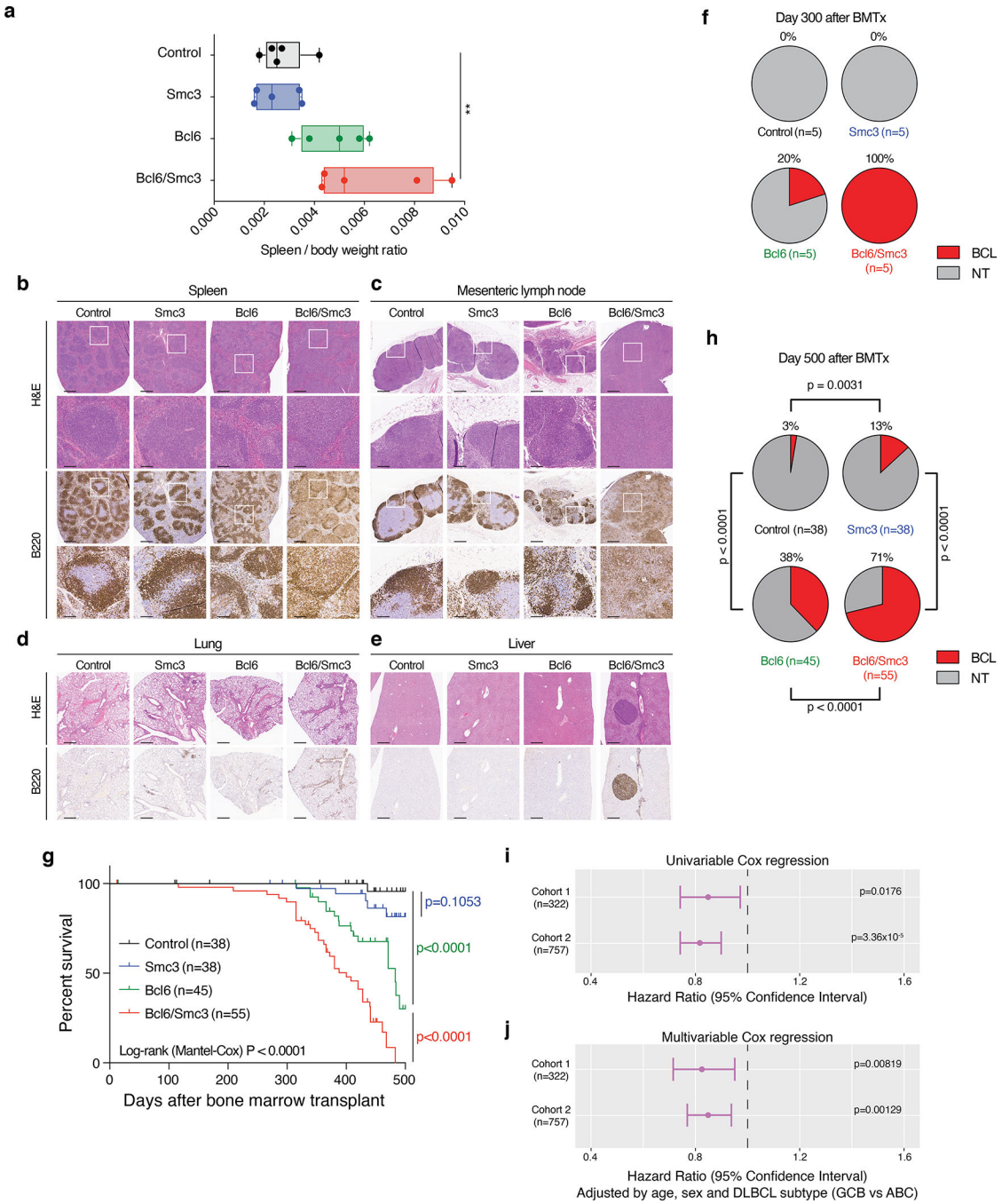


Figure 8. *Smc3* deficiency accelerates malignant transformation of GC B cells and is linked to inferior outcome of DLBCL patients.

(a) Spleen/body weight ratio of *IμBcl6*;*Cγ1*$^{wt/cre}$;*Smc3*$^{wt/-}$ (Bcl6/Smc3) mice compared to *Cγ1*$^{wt/cre}$;*Smc3*$^{wt/wt}$ (control), *Cγ1*$^{wt/cre}$;*Smc3*$^{wt/-}$ (Smc3) or *IμBcl6*;*Cγ1*$^{wt/cre}$;*Smc3*$^{wt/wt}$ (Bcl6) mice. Box plots show the median as center, first and third quartiles as the box hinges, and whiskers extend to the smallest and largest value no further than the 1.5 × interquartile range (IQR) away from the hinges. **p<0.01. **(b-e)** Representative images of histological analyses in **(b)** spleen, **(c)** mesenteric lymph node, **(d)** lung and **(e)** liver. Representative

images from one mouse out of >30 per group. Hematoxylin and eosin staining (H&E) and B220 immunohistochemistry is shown in samples from control, Smc3, Bcl6 and Bcl6/Smc3 mice. In **(b)** an **(c)**, for pictures in the first and third row bar measure is 1 mm, and for pictures in the second and fourth row bar measure is 250 μm. In **(d)** and **(e)**, bar measure is 1 mm. **(f)** Percentage of tumor bearing mice 300 days after bone marrow transplantation in the different cohorts of mice as indicated (n=5 per group). **(g)** Kaplan-Meier analysis showing percentage of mice surviving lymphoma-free in Bcl6/Smc3 mice compared to control, Smc3 and Bcl6 cohorts. Log-rank (Mantel-Cox) test was used. p-values shown are resulting from pairwise comparison of any two groups. Control vs Smc3: p=0.1053, Control vs Bcl6: p<0.0001 and Bcl6/Smc3 vs Bcl6: p<0.0001. **(h)** Proportion of mice that developed B cell lymphoma was calculated at day 500 after bone marrow transplant for animals in all four groups. Binomial test was used to calculate p-values. **(i)** Univariable Cox regression analysis and **(j)** multivariable Cox regression analysis were performed in two DLBCL cohorts (cohort 1: n=322 patients, and cohort 2: n=757 patients). In both cases, *Smc3* expression levels were used as a continuous variable. Multivariable analysis was adjusted by age, sex and DLBCL subtype of the patient. In **(i)** and **(j)** error bars represent 95% confidence intervals of hazard radio.