Orientation-specific joining of AID-initiated DNA breaks promotes antibody class switching

Junchao Dong1*, Rohit A. Panchakshari1*, Tingting Zhang1++, Yu Zhang1+, Jiazi Hu1, Sabrina A. Volpi2, Robin M. Meyers1, Yu-Jui Ho1,†, Zhou Du1, Davide F. Robbiani3, Feilong Meng1, Monica Gostissa1,†, Michel C. Nussenzweig1,†, John P. Manis2 & Frederick W. Alt1

During B-cell development, RAG endonuclease cleaves immunoglobulin heavy chain (IgH) V, D, and J gene segments and orchestrates their fusion as deleitional events that assemble a V(D)J exon in the same transcriptional orientation as adjacent Cμ constant region exons1,2. In mice, six additional sets of constant region exons (Cμs) lie 100–200 kilobases downstream in the same transcriptional orientation as V(D)J and Cμ exons. Long repetitive switch (S) regions precede Cμ and downstream Cμs. In mature B cells, class switch recombination (CSR) generates different antibody classes by replacing Cμ with a downstream Cμ (ref. 2).

Activation-induced cytidine deaminase (AID) initiates CSR by promoting deamination lesions within Sμ and a downstream acceptor S region3; these lesions are converted into DNA double-strand breaks (DSBs) by general DNA repair factors4. Productive CSR must occur in a deleitional orientation by joining the upstream end of an Sμ DSB to the downstream end of an acceptor S-region DSB. However, the relative frequency of deleional to inversionsal CSR junctions has not been measured. Thus, whether orientation-specific joining is a programmed mechanistic feature of CSR as it is for V(D)J recombination and, if so, how this is achieved is unknown. To address this question, we adapt high-throughput genome-wide translocation sequencing into a highly sensitive DSB end-joining assay and apply it to endogenous AID-initiated S-region DSBs in mouse B cells. We show that CSR is programmed to occur in a productive deleitional orientation and does so via an unprecedented mechanism that involves in cis IgH organizational features in combination with frequent S-region DSBs initiated by AID. We further implicate ATM-dependent DSB-response factors in enforcing this mechanism and provide an explanation of why CSR is so reliant on the 53BP1 DSB-response factor.

Most chromosomal DSBs end join to segments of DSBs genome-wide without orientation (end) specificity4,5. Similarly, non-productive ‘inversional’ CSR junctions have been found in transformed B cells6, suggesting CSR also may not be orientation-specific4,6 (Fig. 1a). To address this possibility, we employed digestion–circularization PCR (DC–PCR, Extended Data Fig. 1a) to identify the orientation of CSR junctions between Sμ and S1 in purified mouse B cells stimulated with anti-CD40 plus IL4 to activate AID-targeting to S1 and Sc, and class-switching to IgG1 (and IgG2a). Most Sμ to S1 junctions identified by this semi-quantitative approach were deleitional (Extended Data Fig. 1b).

To confirm DC–PCR findings and analyse potential mechanisms, we used high-throughput genome-wide translocation sequencing (HTGTS), an unbiased genome-wide approach that identifies ‘prey’ DSB junctions to a fixed ‘bait’ DSB with nucleotide resolution4 (Extended Data Fig. 1c). We refer to broken ends of bait IgH DSBs as 5′- and 3′-broken ends; specific primers allow use of each as bait4 (Fig. 1b, c). Prey junctions are denoted + if prey is read from the junction in a centromere-to-telomere direction and − if in the opposite direction4 (Fig. 1b, c). The + and − outcomes for intrachromosomal joining of broken ends of different DSBs on the same chromosome include rejoining of a DSB subsequent to resection, or joining the broken ends of two separate DSBs to form intrachromosomal inversions, deletions, or excision circles4,5 (Fig. 1b, c).

To assess the relative frequency at which non-AID-initiated IgH DSBs join in deleitional versus inversional orientation, we expressed I-Sce1 endonuclease in anti-CD40/IL4-activated AID-deficient B cells in which I-Sce1 targets were inserted upstream of Sμ and downstream of S1 (IghI-96k allele1; Extended Data Fig. 1d, e), or in AID-sufficient B cells in which S1 and Sμ were replaced with I-Sce1 targets (ASμI-96k/AS1-I-25k allele1; Fig. 1d and Extended Data Fig. 1f). HTGTS with primers that captured junctions involving 3′- or 5′-broken ends of I-Sce1 bait DSBs in the S1 locale revealed that a major class of recovered junctions were re-joins of bait DSBs following resection (Fig. 1d and Extended Data Fig. 1f–j). A second major class of bait junctions in the S1 locale involved intact or resected 3′- or 5′-broken ends of I-Sce1-generated DSB in the Sμ locale, which comprised relatively similar numbers of deleional (+) and inversional (−) junctions for bait 3′-broken ends (Extended Data Fig. 1e, f) and similar numbers of excision circle (−) versus inversional (+) junctions for bait 5′-broken ends (Extended Data Fig. 1e, f). As expected4, bait 3′- and 5′-broken ends from the S1 locale recovered similar levels of + and − junctions genome-wide (Extended Data Fig. 2a–d). We conclude that joining between two I-Sce1 DSBs in different IgH S-region locations in CSR-activated B cells lacks any notable preference for or against inversional versus deleitional joins.

In AID-deficient IgHI-96k B cells, I-Sce1 5′- and 3′-broken end baits downstream of S1 did not capture IgH DSB hotspots beyond I-Sce1-generated broken ends upstream of Sμ (Extended Data Fig. 1d, e). In contrast, I-Sce1 5′- and 3′-broken end baits from the ASμI-96k/AS1-I-25k allele in AID-sufficient B cells joined frequently to AID-initiated Sc DSBs 60 kilobases (kb) downstream (Fig. 1d and Extended Data Fig. 1f), with the majority (~80%) of 3′ and 5′ ASμI-96k/AS1-I-25k broken end joins distributed across the 4-kb Sc in orientations that generate, respectively, excision circles (Fig. 1d) or deletions (Extended Data Fig. 1f). We also performed HTGTS on activated, I-Sce1-expressing B cells in which only S1 was replaced by an I-Sce1 cassette (AS1-I-25k allele2; Fig. 1e). Beyond break-site junctions, major IgH hotspot regions of 3′ AS1-I-25k broken ends were Sμ and Sc (Fig. 1e and Extended Data Fig. 2j). Junctions occurred broadly across Sμ, with 80% in a deleitional orientation; while 90% of Sc junctions were in the reciprocal excision circle orientation (Fig. 1e; Extended Data Fig. 2j). CH12F3 B

1Howard Hughes Medical Institute, Program in Cellular and Molecular Medicine, Boston Children’s Hospital, and Department of Genetics, Harvard Medical School, Boston, Massachusetts 02115, USA.
2Boston Children’s Hospital and Joint Program in Transfusion Medicine, Harvard Medical School, Boston, Massachusetts 02115, USA.
3Howard Hughes Medical Institute, Laboratory of Molecular Immunology, The Rockefeller University, New York, New York 10065, USA. Present address: Eli Lilly and Company, Alexandria Center for Life Sciences, 450 East 29th Street, New York, New York 10016, USA (Y.Z.); National Institute of Biological Sciences, Beijing 102206, Beijing, China (Y.Z.); Cold Spring Harbor Laboratory, Watson School of Biological Sciences, Cold Spring Harbor, New York 11724, USA (Y.-J.H.); 121Bio, 700 Main Street, Cambridge, Massachusetts 02139, USA (M.G.).
*These authors contributed equally to this work.

LETTER

doi:10.1038/nature14970

©2015 Macmillan Publishers Limited. All rights reserved
lymphoma cells in which Sz was replaced with an I-Sce1 site had a similar orientation bias of Sz 1-Sce1 3'-broken end joining to Sz DSBS (Extended Data Fig. 2n–q). Joining of the 5'-broken ends of ΔSzU2ΔI (on the ΔSzU2ΔI/ΔSzU2ΔI allele) to AID-initiated DSBS in Sy3, Sy2b and Sy2a in lipopolysaccharide plus anti-IgD-dextran-activated B cells were similarly orientation-biased (Extended Data Fig. 3a–c). However, joining of the 5'-broken ends of ΔSzU2ΔI across an array of 28 × 1-Sce1 sites replacing Sy113 was not orientation-biased (Extended Data Fig. 3d, e). Together, these findings suggest that orientation-specific CSR joining requires an S-region sequence and/or unique aspects of S-region DSBS.

Mammalian S regions are G-rich on the non-template strand, giving AID-initiated 5′ and 3′ S-region broken ends a potential end-sequence bias. Also, when transcribed in the sense direction, S regions generate stable R-loops, which could differentially affect 5′ and 3′ S-region broken end structure. To test the potential roles of S regions in orientation-specific CSR, we used a Cas9/gRNA approach to invert Sz on the productive allele of CH12F3 B cells, which modestly reduced CSR (Extended Data Fig. 3f–h). We then assayed CH12F3 cells in which Szx was replaced with an I-Sce1 site and Sz was in a normal or inverted orientation. These assays revealed that inversions of I-Sce1-generated 3'-broken ends at the Szx locale to Sz DSBS were similarly biased for deletional junctions independent of Sz orientation (Fig. 2a–c).

Consistent with low-level trans CSR, HTGTS libraries from activated ΔSzU2ΔI/ΔSzU2ΔI B cells contained numerous junctions from ΔSzU2ΔI 3'-broken ends across the trans Sz, which, in contrast to cis ΔSzU2ΔI 3'. broken end Sz junctions, occurred in + and − orientations at a similar frequency (Fig. 2d). Likewise, bait 3'-broken ends from the ΔSzU2ΔI IgH allele identified approximately equal numbers of (+) versus (−) junctions to AID off-target DSBS in Igha on chromosome 7 (Extended Data Fig. 2e). Finally, translocations between bait 5′ I-Sce1 DSB broken ends in c-myc and prey AID-initiated Sz and Sz broken ends in CSR-activated B cells lacked orientation bias (Fig. 2e). We conclude that orientation-dependent CSR joining does not require orientation-associated features of Sz sequence, transcription, or transcripts. Moreover, AID-initiated DSBS per se are not sufficient to promote orientation specificity, as demonstrated by orientation-independence of DSBS joining to them in trans. Thus, beyond S-region sequences and/or high frequency AID-initiated DSBS within them, aspects of IgH locus organization in cis must play a critical role in promoting orientation-dependent CSR joining.

We tested whether joining between two sets of endogenous AID-initiated S-region DSBS is orientation-dependent. Use of core S-region DSBS as HTGTS bait is confounded by their highly repetitive nature. Therefore, we used as bait a 150-base-pair (bp) sequence at the 5′ end of Sz (5′Sz), which retains 14 of approximately 500 Sz AID-target
motifs (Fig. 3a, left panel). HTGTS of anti-CD40/IL4-stimulated B cells with the 5’Sµ broken end primer revealed break-site junctions, as well as Sy1 and Sc junctions (Fig. 3b, c). Consistent with AID-initiation, bait junctions were enriched at AID-targets within the 5’Sµ bait (Fig. 3a, right panel). 5’Sµ broken end junctions spread broadly over prey S regions, with up to 95% in a deletional orientation (Fig. 3c). For comparison, we tested a 150-bp 5’Sµ remnant of Sµ (rSµ; Extended Data Fig. 4a, left panel), retained when the rest of Sµ was deleted.7 B cells homozygous for rSµ have reduced IgG1 CSR but nearly normal IgE CSR.18 HTGTS with either 5’Sµ rSµ or 3’Sµ rSµ broken end primers of anti-CD40/IL4- and lipopolysaccharide/anti-IgD-dextran-stimulated B cells, respectively, revealed junctions to Sy1 and Sc and to Sy3, Sy2b, and Sy2a (Extended Data Fig. 4). 5’Sµ broken end junctions spread over target S regions, with >90% in a deletional orientation (Extended Data Fig. 4b, f); while >90% of 3’Sµ broken end junctions were in the complementary excision circle orientation (Extended Data Fig. 4c, g). Within the bait rSµ, junctions again were enriched at AID targets (Extended Data Fig. 4a). Consistent with IgH class-switching patterns, rSµ HTGTS junctions occurred more frequently to Sµ than those from the 5’Sµ bait in the context of full-length Sµ (Extended Data Fig. 4b). Analyses of rSµ-mutant CH12F3 cells gave similar results (Extended Data Fig. 5a–c). Thus, AID-initiated Sµ DSB joining to all downstream acceptor S regions is strongly biased towards the deletional orientation.

CSFR DSBs generate a DSB response (DSBR) in which ATM activates histone H2AX and 53BP1 in chromatin flanking DSBs, thereby contributing to end-joining.19-21 ATM or H2AX deficiency moderately reduces CSR (Extended Data Fig. 6a).21,19,23,24 However, 53BP1 deficiency causes a more drastic reduction (Extended Data Fig. 6a), suggesting specialized CSR roles,21,19,22 such as promoting S-region synopsis or protecting S-region DSBs from resection.21,23-25 To elucidate influences on orientation-specific CSR, we employed HTGTS to assay joining of AID-initiated 5’Sµ broken ends to AID-initiated Sy1 and Sµ DSBs in anti-CD40/IL4-activated ATM-, H2AX-, and 53BP1-deficient B cells, as well as in B cells deficient for Rif1-1, a 53BP1-associated factor that mediates resection blocking.26,27 ATM-, H2AX-, and Rif1-deficient B cells had reduced Sy1 and Sµ IgG1 CSR but were not deficient in IgG1 CSR25 to Sµ DSBs in anti-CD40/IL4-activated ATM-, H2AX-, and 53BP1-deficient B cells, as well as in B cells deficient for Rif1-1, a 53BP1-associated factor that mediates resection blocking.26,27 ATM-, H2AX-, and Rif1-deficient B cells had reduced Sy1 and Sµ IgG1 CSR but were not deficient in IgG1 CSR25 compared to wild type; 53BP1-deficient B cells had a greater reduction, with most localizing to the break-site region (Fig. 3d, e and Extended Data Fig. 6f–d). Most break-site junctions were resections, which were longest (up to about 6 kb) for 53BP1 deficiency (Extended Data Fig. 6e, f; see extended discussion in Supplementary Information for Extended Data Fig. 6f). Compared to wild type, bait 5’Sµ junctions to Sy1 and Sµ DSBs in different DSBR-deficient backgrounds had varying decreases in orientation specificity, with H2AX deficiency having the smallest and 53BP1 deficiency the largest (Figs 3d, e and 4a; Extended Data Fig. 6c, d and Extended Data Table 1a, b). Indeed, residual junctions of 5’Sµ to Sy1 and Sµ locales in 53BP1-deficient B cells showed relatively normalized inversion:deletion ratios (Fig. 4a), a finding confirmed by DC–PCR (Extended Data Fig. 1b). Finally, 53BP1-deficiency did not impact joining orientation of 5’Sµ and 3’Sµ I-SceI-generated broken ends in AID-deficient Ighm-906 B cells (Extended Data Fig. 1g).

Owing to the potential difficulty in measuring relative resection of recurrent re-joins at or near the break-site, we focused on prey S-region broken end resections (Extended Data Figs 1 and 6; see extended discussion in Supplementary Information for Extended Data Figs 1d, e and 6f). Because S regions are long and AID-initiated DSB locations within them are diverse, we estimated relative resection...
by quantifying bait broken end to prey broken end junctions downstream of S-region positions where the incidence of wild-type junctions decreases to background (Fig. 3b–e). Based on this ‘long’ S-region resection assay, ATM- and H2AX-deficient cells had modest resection decreases, Rif1-deficient cells slightly greater increases, and 53BP1-deficient versus ATM-deficient B cells that are not revealed by our long resection assay. Another possibility would involve a putative specialized role for 53BP1 in stabilizing synapsed S regions.

We demonstrate that CSR is mechanistically programmed to occur in a productive deletional orientation. Based on our findings, we propose a working model for orientation-specific CSR, in which a key component is the organization of S regions within topologically-associated domains (TADs) that promote their frequent S-region synopsis via Langevin motion (Fig. 4c). Within such TADs, we implicate additional Igh-specific organizational features, not yet fully elucidated, in playing a fundamental role in mediating synopsis in an orientation that promotes deletional joining (Fig. 4c). We find that functions of such organizational features are complemented by S regions, potentially associated with their ability to promote AID-initiated DSBs, multiple frequent DSBs, or both. Our studies also implicate DSBR factors in enforcing this mechanism (Fig. 4d). The broader DSBR probably contributes by tethering un-synapsed S-region DSBs for efficient re-joining, keeping them from separating into chromosomal breaks that could frequently translocate with orientation independence to S-region broken ends within the TADs; this function would also allow subsequent AID-initiated breakage and joining to a synapsed S region (Fig. 4c).
also prevent long end-resections that could cause S-region broken ends to linger in resection complexes, preventing synopsis with other S-region broken ends and/or diminishing ability to be joined by classical non-homologous end joining (Fig. 4d). Different DSBR factors have differential impact in tethering versus resection inhibition and, thus, may impact orientation dependence via different routes. For example, ATM deficiency inhibits resection by impairing CIP activation57, but promotes resection via other nuclease by impairing inhibitory activities of H2AX, 53BP1 and, indirectly, Rif126,27 (Fig. 4d). 53BP1-deficiency is unique in that it both impairs tethering for rejoining and activates resection of un-joined ends by failure to activate Rif1, leading to extreme resections and the greatest impairment of CSR and orientation-dependent joining (Fig. 4d). As common and unique impacts of 53BP1 deficiency markedly affect both donor and acceptor S regions, they would be multiplicative and, thereby, explain the profound impact of 53BP1-deficiency on CSR.

**Online Content** Methods, along with any additional Extended Data display items and Source Data, are available in the online version of the paper; references unique to these sections appear only in the online paper.

**Received 7 April; accepted 21 July 2015.**

**Published online 26 August 2015.**

1. Schatz, D. G. & Swanson, P. C. V(D)J recombination: mechanisms of initiation. *Annu. Rev. Genet.* **45**, 167–202 (2011).
2. Alt, F. W., Zhang, Y., Meng, F. L., Guo, C. & Schwer, B. Mechanisms of programmed DNA lesions and genomic instability in the immune system. *Cell* **152**, 417–429 (2013).
3. Di Noia, J. M. & Neuberger, M. S. Molecular mechanisms of antibody somatic hypermutation. *Annu. Rev. Biochem.* **76**, 1–22 (2007).
4. Chiarello, R. et al. Genome-wide translocation sequencing reveals mechanisms of chromosome breaks and rearrangements in B cells. *Cell* **147**, 107–119 (2011).
5. Frock, R. L. et al. Genome-wide detection of DNA double-stranded breaks induced by engineered nuclease. *Nature Biotechnol.* **33**, 179–186 (2015).
6. Yancopoulos, G. D. et al. Secondary genomic rearrangement events in pre-B cells: VHDJH replacement by a LINE-1 sequence and directed class switching. *EMBO J.* **5**, 3259–3266 (1986).
7. Jack, H. M. et al. Looping out and deletion mechanism for the immunoglobulin heavy-chain class switch. *Proc. Natl Acad. Sci. USA* **85**, 1581–1585 (1988).
8. Vaandragrter, J. W. et al. DNA fiber fluorescence in situ hybridization analysis of immunoglobulin class switching in B-cell neoplasia: aberrant CH gene rearrangements in follicle-center-cell lymphoma. *Blood* **92**, 2871–2878 (1998).
9. Lenz, G. et al. Aberrant immunoglobulin class switch recombination and switch translocations in activated B cell-like diffuse large B cell lymphoma. *J. Exp. Med.* **204**, 633–643 (2007).
10. Harriman, W., Volk, H., Defranoux, N. & Wabl, M. Immunoglobulin class switch recombination. *Annu. Rev. Immunol.* **11**, 361–384 (1993).
11. Bothmer, A. et al. 53BP1 regulates DNA resection and the choice between classical and alternative end joining during class switch recombination. *J. Exp. Med.* **207**, 885–896 (2010).
12. Zarrin, A. A. et al. Antibody class switching mediated by yeast endonuclease-generated DNA breaks. *Science* **315**, 377–381 (2007).
13. Gostissa, M. et al. IgH class switch excising a general property of two DNA breaks to be joined in cis over long chromosomal distances. *Proc. Natl Acad. Sci. USA* **111**, 2644–2649 (2014).
14. Tian, M. & Alt, F. W. Transcription-induced cleavage of immunoglobulin switch regions by nucleotide excision repair nucleases in *vitro*. *J. Biol. Chem.* **275**, 24163–24172 (2000).
15. Yu, K., Chedin, F., Hsieh, C. L., Wilson, T. E. & Lieber, M. R. R-loops at immunoglobulin class switch regions in the chromosomes of stimulated B cells. *Nature Immunol.* **4**, 442–451 (2003).
16. Reynaud, S. et al. Interallelic class switch recombination contributes significantly to class switching in mouse B cells. *J. Immunol.* **174**, 6176–6183 (2005).
17. Khamlichi, A. A. et al. Immunoglobulin class-switch recombination in mice devoid of any Sμ tandem repeat. *Blood* **103**, 3828–3836 (2004).
18. Zhang, T. et al. Downstream class switching leads to IgE antibody production by B lymphocytes lacking IgM switch regions. *Proc. Natl Acad. Sci. USA* **107**, 3040–3045 (2010).
19. Nussenzweig, A. & Nussenzweig, M. C. Origin of chromosomal translocations in lymphoid cancer. *Cell* **141**, 27–38 (2010).
20. Franco, S. et al. H2AX prevents DNA breaks from progressing to chromosome breaks and translocations. *Mol. Cell* **21**, 201–214 (2006).
21. Bredemeyer, A. L. et al. ATM stabilizes DNA double-strand-break complexes during V(D)J recombination. *Nature* **442**, 466–470 (2006).
22. Daniel, J. A. & Nussenzweig, A. The AID-induced DNA damage response in chromatin. *Mol. Cell* **50**, 309–321 (2013).
23. Reina-San-Martin, B., Chen, J., Nussenzweig, A. & Nussenzweig, M. C. Enhanced intra-switch region recombination during immunoglobulin class switch recombination in 53BP1−/− B cells. *Eur. J. Immunol.* **37**, 235–239 (2007).
24. Bassing, C. H. & Alt, F. W. The cellular response to general and programmed DNA double-strand breaks. *DNA Repair (Amst.)* **3**, 781–796 (2004).
25. Yamane, A. et al. RPA accumulation during class switch recombination represents 5′→3′ DNA-end resection during the S-G2/M phase of the cell cycle. *Cell Rep.* **3**, 138–147 (2013).
26. Panier, S. & Boulton, S. J. Double-strand break repair: 53BP1 comes into focus. *Nature Rev. Mol. Cell Biol.* **15**, 7–18 (2014).
27. Zimmermann, M. & de Lange, T. 53BP1: pro choice in DNA repair. *Trends Cell Biol.* **24**, 108–117 (2014).
28. Lucas, J. S., Zhang, Y., Dudko, O. K. & Murre, C. 3D trajectories adopted by coding and regulatory DNA elements: first-passage times for genomic interactions. *Cell* **158**, 339–352 (2014).
29. Helmink, B. A. et al. H2AX prevents CtIP-mediated DNA end resection and aberrant repair in G1-phase lymphocytes. *Nature* **469**, 245–249 (2011).

**Supplementary Information** is available in the online version of the paper.

**Acknowledgements** We thank K. Yu for providing the CH12F3-RMCE (1F7) cell line and exchange cassette plasmid. This work was supported by National Institute of Health grants AI077595 to F.W.A., CA133781 to J.M., AI112602 to D.F.R., and AI037526 and AI072529 to M.C.N. S.V. was supported by NIH training grant T32HL66987. F.W.A. and M.C.N. are investigators of the Howard Hughes Medical Institute. J.H. is supported by a Robertson Foundation/Cancer Research Institute Irvington Fellowship. F.M. is a Lymphoma Research Foundation postdoctoral fellow and was a Cancer Research Institute postdoctoral fellow.

**Author Contributions** J.D., R.P., T.Z., J.P.M. and F.W.A. designed the study; J.D., R.P., T.Z., J.H. and S.V. performed experiments; Y.H. and R.M. designed bioinformatics pipelines; R.M., J.D., R.P. and Z.D. performed computational analyses of sequencing data; J.D., R.P. and F.W.A. wrote the paper. Other authors provided reagents, designed or performed certain experiments, and helped revise the paper.

**Author Information** HTGTS sequencing data has been deposited in the GEO database under the accession number GSE71005. Reprints and permissions information is available at www.nature.com/reprints. The authors declare no competing financial interests. Readers are welcome to comment on the online version of the paper. Correspondence and requests for materials should be addressed to F.W.A. (alt@enders.tch.harvard.edu) or to J.P.M. (manis@enders.tch.harvard.edu).
METHODS

No statistical methods were used to predetermine sample size. The experiments were not randomized and the investigators were not blinded to allocation during experiments and outcome assessment.

Mice. IgG4-305 AID−/− (ref. 11), A53T+/−;IgG4-305 I−/− chimaera, A53T−/+I−/−-Sce (ref. 12), SphK1−/− (ref. 17), c-myc-250-IRES (ref. 4), ATM−/− (ref. 30), H2AX−/− (ref. 31), S3BP1−/− (ref. 32), and Rffl/FD199 (ref. 33) lines have been reported previously. Mouse work was performed under protocols approved by the Boston Children’s Hospital and the Rockefeller University Institutional Animal Care and Use Committees.

Plasmids and oligonucleotides. Oligonucleotides for gRNAs for CRISPR/Cas9-mediated targeting of various IgH regions were cloned into pX330 vector (Addgene plasmid ID 42230) as described. The target sequences of Cas9 constructs are listed in the DNA oligonucleotides table in the Supplementary Information. Exchange vector (pHL28) with heterologous loxp sites was obtained from K. Yu. A 200-bp GFP-derived sequence was amplified and ligated to an I-SceI recognition sequence and subsequently introduced into the pHL28 vector to make the pHL-I-Sce I-exchange vector. To obtain the I-SceI expression plasmid for transducing CH12 cell lines, I-SceI-IRES-GFP fragment was shuttled from a retroviral construct (pMX-I-SceI-IRES-GFP) into pcDNA3.0 (Invitrogen) vector.

B-cell culture, transduction and FACs analysis. Mature splenic B cells isolated using a CD19 negative selection kit (MACS) were cultured in lymphocyte medium R15 (RPMI1640, 15% FBS, 1-glutamate, 1% penicillin and streptomycin). B-cell stimulation was performed with anti-CD40 (1 μg/ml), anti-CD40L (20 ng/ml, Tocris) was added to stimulated cells at day 1 post-stimulation to a final concentration of 10 ng/ml and was maintained during the course of the experiment until collection of the cells for FACs and HTGTS libraries.

Cells line and nucleofection. CH12F3 cell line stimulation to IgA was performed as described. CH12F3 cells with recombine-mediated cassette exchange (RMCE) in place of the endogenous 5′ region, referred to as 1F7 cells, were maintained at 37 °C, 5% CO2 and cultured in RPMI media with 10% FCS, 0.5% penicillin/streptomycin, 50 μM β-mercaptoethanol. Exchange vector with heterologous loxp sites containing 1×-1-SceI site embedded in 200bp of GFP-derived sequence was cloned. RMCE was performed as previously described. Exchanged ASx−/− clones were verified by PCR, Sanger sequencing and Southern blotting. ASx−/− cells were then stimulated with anti-CD40, IL4 and TGFB for 15 h followed by nucleofection with pcDNA-I-SceI-IRES-GFP expression vector using 4D-nucleofector X (Lonza, solution SF, protocol CA-137) with the gRNA vectors to excise the sequences between Iγ4 intron and −130 bp downstream of Cγ polyadenylation on the non-coding allele that has already switched to Sx. Single-cell subclones were seeded into 96-well plates 12 h post-nucleofection, and the resulting clones were screened by PCR and Southern blot. One confirmed positive clone was further modified by qDNA vectors targeted at the 5′ Sx1 and 3′ Sx1 regions to invert the Sx (−4 kb) sequence. Initial screening for positive clones was performed by PCR, followed by Southern blotting and Sanger sequencing for the inversion junction. The resultant clones were stimulated with anti-CD40, IL4 and TGFB. IgA CSR was measured by FACs on days 2 and 3 post-stimulation. ASx−/− IgA (inv/inv) clones were obtained by targeting the aforementioned 1×-I-SceI-RMCE-positive cells with gRNA targeting 5′ Sx2 and 3′ Sx for inverting the Sx sequence same as above. The resultant positive clones were verified by PCR, Southern blotting and Sanger sequencing for the inversion junction. To make rSx−/−CH12F3 cells, the aforementioned CH12F3 (non-productive allele ASx−/−-Sx) cells were used to further truncate Sx sequences on the coding allele with gRNA targeting 5′ Sx2 and 3′ Sx. Single-cell deletion subclones were screened and confirmed by PCR and Southern blot. The resultant rSx−/−CH12F3 cells were stimulated with anti-CD40, IL4 and TGFB and harvested on days 2 and 3 for DNA isolation for HTGTS library preparation.

DC–PCr. The DC–PCR assay was performed as described previously. Briefly, genomic DNA was isolated and subsequently purified by phenol chloroform extraction from day 4 anti-CD40/IL4 stimulated B cells. Five micrograms of genomic DNA was digested overnight with 20 U of EcoRI (Roche). Ligations were performed under dimethyl conditions to promote circularization. Digested DNA was ligated overnight at 16°C with a concentration of 1.8–9 ng/μl in a total volume of 100 μl per reaction. Three to four ligation reactions were pooled, column purified, concentrated and serially diluted at a 1:5 ratio. PCR was performed in 50 μl per reaction using 2.5 U Taq (Qiagen) with serially diluted DNA starting from ~50–150 ng. Primers were designed to amplify the Sx-Sy1 rearrangements that occur during CSR to IgG1 in direct chromosomal joining of Sx-Sy1 with excision of circular DNA or inversion of sequences between broken ends of Sx and Sy1. As a control for EcoRI digestion and circularization of input DNA, amplification of an EcoRI fragment of nicotinic acetylcholine receptor subunit gene (Chnrn1) was performed, which, after EcoRI digestion and circularization, generates a 753-bp DC–PCR product. To quantify the amount of direct or inversion joins amplified by PCR, DC–PCR products of direct or inversion joins were cloned into the pCR2.1 Topo TA vector. Precise placasmids were obtained and a standard curve was generated ranging from 4 to 10,000 copies per reaction. After running on 1% agarose gel, PCR fragments were transferred to nitrocellulose membrane and hybridized to a 3′ Sy1 probe according to standard Southern blotting procedures. Primers for direct joining PCR: forward, 5′-CAT GAGAGTTGTGACTGATGTTGCCTCTTGCACAC-3′; reverse, 5′-ACCTAGCTGCTAGTTGAGTATGGTTTGACAAACGG-3′; Primers for inversion joining PCR: forward, 5′-CAG TACAGAGAAGACTGACGTTGAG-3′; reverse, 5′-CCATAGCAGTGG TCAATCTTGTCTCC-3′. Primers for control Chnrn1 DC–PCR: forward, 5′-GGCCACATGACGGACTGTTGGGTTCCACCCAG-3′; reverse, 5′-GGCC GTGCACACCGGAGTGAGTACACCTATAG-3′. Oligonucleotide probe for the detection of both deletional and inversional CSR joining products: SYT1 CCGTGGTACGTTACGAAAGGCT.

High-throughput genome-wide translocation sequencing (HTGTS). HTGTS libraries were generated by emulsion-mediated PCR (EM–PCR) and linear-amplification-mediated PCR (LAM–PCR) methods as described in ref. 5. In brief, bwa-mem (Bioruptor, Diagenode) gDNA was subjected to LAM–PCR using 1 μl Taq polymerase (Qiagen) per reaction with a single biotinylated primer for 50 cycles of 94°C for 18 s, 94°C for 30 s; 58°C for 30 s; 72°C for 90 s. One more unit of Taq polymerase was added to the reaction mixture to execute PCR for an additional 35 cycles. Biotinylated DNA fragments were captured with Dynabeads MyOne streptavidin C1 beads (Invitrogen) at room temperature for 1 h, followed by on-bead ligation at 25°C for 2 h with bridge adapters in the presence of 15% PEG-8000 (Sigma) and 1 mM hexamine cobalt chloride (Sigma). After washing beads with B&W buffer as described by the manufacturer, ligated products were subjected to 15 cycles of on-bead PCR with Phusion polymerase (Fisher), locus-specific and adapter primer followed by blocking digestion with appropriate restriction enzymes to remove uncut germline gDNA. A third round of tagging PCR to add Illumina Miseq-compatible adapters at 5′ and 3′ ends of the second-round PCR product was carried out for another 10 cycles with Phusion polymerase. PCR products were size-fractionated for DNA fragments between 300–1000 bp on a 1% agarose gel, column purified (Qiagen) before loading onto Illumina Miseq machine for sequencing.

Data analyses. Data analysis of MiSeq sequencing reads has been described in ref. 5. In brief, de-multiplexing for the MiSeq reads was performed using the fastq-multiplex tool from ea-utils (https://github.com/pa/ea-utils/) and adapter sequence trimming was performed using the SeqPrep utility from ea-utils (https://github.com/jstjohn/SeqPrep). Reads were mapped using Bowtie2 (http://bowtie-bio.sourceforge.net/bowtie2/manual.shtml) to either mm9 (for libraries generated with Riffl knockout cells and CH12F3-derived cells) or modified mm9 reference genome (for all other genotypes) containing the 176-kb IgH constant region of 129S5 mice, in which the region between chr12:114493849–114665808 of mm9 was replaced with DNA sequence ranging from 1416975 to 1593283 on the 129S5 IgH reference sequences AJ851863. In cases where necessary, for instance when aligning reads to the Sx+locus on the IgH−allele and other circumstances, we further modified the custom 129S5_IgH genome to include the cassette of wild-type locus on the IgH−allele. A BEST PATH algorithm was used for the analysis of genomic information before aligning MiSeq reads by Bowtie2. CH12F3 clone was derived from CH12LX lymphoma cell line23, CH12.LX cells were subcloned from the original CH12 lymphoma cell line4, which originated from a C57BL/10 mouse substrain double congenic for H-2a H-2b (ref. 39). C57BL/10 and C57BL/6 are both substrains of C57BL and thus we use BL/6 (mm9) as reference genome when running our HTGTS data analyses pipeline on libraries made with CH12F3 cells. To reflect HTGTS data analyses pipeline on libraries made with CH12F3 cells. To reflect HTGTS data analyses pipeline on libraries made with CH12F3 cells. To reflect HTGTS data analyses pipeline on libraries made with CH12F3 cells. To reflect HTGTS data analyses pipeline on libraries made with CH12F3 cells. To reflect HTGTS data analyses pipeline on libraries made with CH12F3 cells.
filtered on following criteria: (1) reads must include both a bait alignment and a prey alignment; and (2) the bait alignment cannot extend more than 10 nt beyond the targeted site. For reads mapped to the repetitive low-mappability regions, multiple competing alignments with identical or similar scores exist and the coordinates for best alignment are randomly chosen among the competing ones. For junctions mapped to each individual repetitive S region, there are no competing alignments from outside of that region as shown by simulation (see details below), although the exact junction coordinate within the region could not be identified. We also applied filter to remove duplicates (referred to as ‘de-dup’ hereafter) wherein the coordinates of the end of the bait alignment were compared to the start of the prey alignment across all reads. A read is marked as a duplicate if it has bait and prey alignment coordinates within 2 nt of another read’s bait and prey alignments. To plot all the S-region junctions, we took the ones filtered by a mappability filter but unequivocally mapped to S regions and removed the reads through the de-dup program mentioned above, before combining with ‘good’ reads passing both the mappability and de-dup filters. A grey box over S regions (for example, S1 and S71) in the figures is used to denote the repetitive regions in these S sequences wherein the randomly assigned mappability-filtered reads were included. Additionally, we applied post-filtering stringencies to remove junctions mapped to simple sequence repeats, telomere repeats and reads with excessive microhomology >20 nt and insertions >30 nt before further analysis. In the end, the combined and cleaned junctions were then plotted genome-wide or onto desired S regions by using the PlotRegion tool (for details see section below).

Scripts and details of pipeline parameters are available upon request.

**Pipeline simulation for S-region mappability.** Results of the S-region mappability simulations are available in the Supplementary Information.

**S-region junction plotting.** As described above, junctions filtered by the mappability filter are retrieved and de-duped before combining with normal junctions. To plot junction coordinates onto individual S regions or the entire Igk constant region, combined junctions are binned using the PlotRegion tool into 100 bins (bin size varies depending on the length of target region that libraries are plotted to) on the basis of the junction coordinates and orientation of joining. The bincount file (histogram information for junction distribution in both joining orientation) generated by the PlotRegion tool is used to calculate the percentage of junctions in each bin in either + or − orientation of the total number of junctions mapped to the region of interest. The results were then plotted as linear graphs by the Prism software. Note that the scale on top of each graph indicates the size of region plotted and is fixed as 1/10 of the size of the plotted region, thus is always 10× bin size.

**Calculation of joining orientation bias and acceptor S-region resection.** For simplicity, joining from 5’Sµ to downstream Sγ1 and Sc breaks are used for the explanation of orientation bias and resection of acceptor S-region DSBS. Junctions mapped to Sγ1 and Sc can be divided into six regions (denoted by a–f) in either + or − orientation:

\[
\begin{align*}
\text{b} & \mid \text{c} \mid (\text{−}) \\
\text{d} & \mid \text{e} \mid \text{f} \mid (\text{+})
\end{align*}
\]

Junctions encompassing core Sγ1/Sc are illustrated as b and e regions for − and + junctions respectively, c region (deletional joining, − orientation) or d region (inversional joining, + orientation) represent joining of bait DSB broken ends to resected acceptor Sγ1/Sc DSBS. Junctions falling into regions a or f represent joining to non-AID-generated de novo breaks of unknown source and are often very small in number, and thus were omitted from the calculation of both resection and orientation bias. Since in most genetic backgrounds other than 53BP1−/− inversional joins are much rarer than deletions, the level of resection junctions into the d region fluctuates much more than resection junctions into the c region. We thus chose the c region for calculating resection in all genotypes as follows:

\[
\text{resection ratio} = \frac{c}{b+c} \times 100
\]

The degree of orientation bias, for the purpose of positively correlating with the level of resection, is calculated as the ratio of inversional joins versus deletional joins as below:

\[
\text{bias ratio} = \frac{d + e}{b + c} \times 100
\]

To make a bar graph for comparison of orientation bias degree and resection levels in the CSR junctions obtained from libraries with different genetic backgrounds, individual replicate HTGTS libraries were first size-normalized to the one with smallest junction number in the region of interest among the replicates; resection and bias ratio values from individual experiments were calculated separately and averaged were used for statistical analysis with unpaired two-tailed t-tests. Experiments for each genotype were performed at least three times.

30. Borghesani, P. R. et al. Abnormal development of Purkinje cells and lymphocytes in Atm mutant mice. *Proc. Natl Acad. Sci. USA* **97**, 3336–3341 (2000).
31. Bassing, C. H. et al. Histone H2AX: a dosage-dependent suppressor of oncogenic translocations and tumors. *Cell* **114**, 359–370 (2003).
32. Morales, J. C. et al. Role for the BRCA1 C-terminal repeats (BRCT) protein 53BP1 in maintaining genomic stability. *J. Biol. Chem.* **278**, 14971–14977 (2003).
33. Di Virgilio, M. et al. Rif1 prevents resection of DNA breaks and promotes immunoglobulin class switching. *Science* **339**, 711–715 (2013).
34. Cong, L. et al. Multiplex genome engineering using CRISPR/Cas systems. *Science* **339**, 819–823 (2013).
35. Han, L., Masani, S. & Yu, K. Overlapping activation-induced cytidine deaminase hotspot motifs in Ig class-switch recombination. *Proc. Natl Acad. Sci. USA* **108**, 11584–11589 (2011).
36. Chu, C. C., Paul, W. E. & Max, E. E. Quantitation of immunoglobulin Mγ1 heavy chain switch region recombination by a digestion-circularization polymerase chain reaction method. *Proc. Natl Acad. Sci. USA* **89**, 6978–6982 (1992).
37. Nakamura, M. et al. High frequency class switching of an IgM−/− B lymphoma clone CH12F3 to IgA+ cells. *Int. Immunol.* **8**, 193–201 (1996).
38. Bishop, G. A. & Haughton, G. Induced differentiation of a transformed clone of Ly-1 B cells by clonal T cells and antigen. *Proc. Natl Acad. Sci. USA* **83**, 7410–7414 (1986).
39. Haughton, G., Arnold, L. W., Bishop, G. A. & Mercolino, T. J. The CH series of murine B cell lymphomas: neoplastic analogues of Ly-1 B cells. *Immunol. Rev.* **33**, 35–52 (1986).
40. Faust, G. G. & Hall, I. M. YAHA: fast and flexible long-read alignment with optimal breakpoint detection. *Bioinformatics* **28**, 2417–2424 (2012).

©2015 Macmillan Publishers Limited. All rights reserved
Extended Data Figure 1 | Deletional CSR in in vitro activated B cells by DC-PCR; I-Sce1 DSBs within the Igh constant region locus in activated B cells join with orientation-independence. a, Schematic representation of DC–PCR assay. b, DC–PCR results from anti-CD40/IL4-activated wild-type and 53BP1−/− B cells. c, Schematic representation of the HTGTS method. d, e, HTGTS libraries analyses of anti-CD40/IL4-stimulated IgH−/− B cells with 3′-broken end (d, red arrow, n = 3) or 5′-broken end (e, blue arrow, n = 3) primers. BE, broken end. f, HTGTS libraries with 5′-broken end primer (blue arrow, n = 3) from ΔScl+/−/ΔScl−/− B cells stimulated with anti-CD40/IL4. g, Bar graph depicting deletion:inversion and excision-circle:inversion ratios between two I-Sce1 sites and between I-Sce1 and S region in wild-type versus 53BP1−/− backgrounds. For detailed legends and further discussion, refer to the Supplementary Information.
Extended Data Figure 2 | Genome-wide translocation junctions lack orientation bias; statistical analyses for experimental replicates orientation-biased joining between I-SceI break in place of S orientation bias; statistical analyses for experimental replicates orientation bias. a, b, Circos plots for translocation junctions across the whole genome from 3' -broken end (a, n = 4) or 5' -broken end (b, n = 3) HTGTS with anti-CD40/IL4 stimulated ASμ12xI/ASγ12xI B cells. c, d, Bar graphs depicting genome-wide percentage of junctions from pooled 3' - and 5' -broken end libraries plotted separately in — or + orientations. Error bars are s.d. e, Joining from ASγ12xI 3' -broken end to AID off-target DSBs in Igγ gene on chromosome 7. f, Bar graph showing the number of junctions (average ± s.d.) recovered from Igλ1.6kb AID2/− 3' -broken end HTGTS libraries (n = 3) at the break site and the upstream Sμ12xI prey break as a percentage of the total number of junctions mapped to the 200 kb Igγ constant region. Right panel shows the percentage of junctions mapping at Sγ12xI (average ± s.d.) over the total Igλ junctions that are mapped in the deletion (Del) or inversional (Inv) orientation. The numbers above the bar graph (average ± s.d.) denote the ratio of deletional to inversional junctions. g, Percentage of junctions (average ± s.d.) recovered from the Igκλ1.6kb AID−/− 5' -broken end HTGTS libraries (n = 3). h, Percentage of junctions (average ± s.d.) recovered from the ASμ12xI/ASγ12xI 5' -broken end libraries (n = 4). i, Percentage of junctions (average ± s.d.) recovered from the ASμ12xI/ASγ12xI 5' -broken end libraries (n = 3). j, Percentage of junctions (average ± s.d.) recovered from the wild-type ASγ12xI 3' -broken end libraries (n = 3). k, Percentage of junctions (average ± s.d.) recovered from the ASμ12xI CH12F3 3' -broken end libraries (n = 3) and ASγ12xI Sμ(INV) CH12F3 cells 3' -broken end libraries (n = 3). l, Bar graphs depicting percentage of trans junctions mapping to Sμ in — and + orientations from libraries of ASμ12xI/ASγ12xI B cells (n = 3) cloning from ASγ12xI 3' -broken ends. m, Bar graphs depicting percentage of trans junctions mapping to Sμ in — and + orientations and to Sc in — and + orientations from libraries of c-myc25×I 5' -broken ends (n = 3). n, HTGTS library analyses of ASγ12xI CH12F3 cells stimulated with anti-CD40, IL4 and TGFβ and nucleofected with I-Sce1 expression plasmid. Cells were harvested on day 3 post-stimulation for 3' -broken ends (n = 6) and 5' -broken ends (n = 6) libraries. o, 3' - and 5' -broken end libraries are normalized with 'symmetric junctions' (see Supplementary Information). q, Bar graph showing percentage of junctions from ASγ12xI CH12F3 cells (n = 6) from 3' - and 5' -broken end primers. For detailed legends and further discussion refer to the Supplementary Information.
Extended Data Figure 3 | Joining between I-SceI break at Sµ and AID-initiated S-region breaks in lipopolysaccharide (LPS)-activated ASp2x1/ASy12x1 B cells and clustered I-SceI breaks in ASµ2x1/ASy12x1 B cells in place of Sγ1; inverted Sµ in CH12F3 cells support robust IgA CSR. a, Diagram of Igh locus organization in ASµ2x1/ASy12x1 B cells highlighting AID-initiated breaks in Sγ3, Sγ2b and Sγ2a regions upon LPS stimulation and potential outcomes of CSR in the form of deletion (red, −) and inversional joining (blue, +). b, Plots showing enlarged distribution of pooled prey junctions in a 20-kb region flanking Sγ3 and Sγ2b and Sγ2a from HTGTS libraries of ASµ2x1/ASy12x1 B cells (n = 3) stimulated with LPS and anti-IgD-dextran and infected with I-SceI-expressing retrovirus. c, Bar graph from three independent ASµ2x1/ASy12x1 5′-broken end libraries showing the percentage of junctions mapped at different S regions. d, Diagram of Igh locus organization in ASµ2x1/ASy12x1 B cells highlighting joining outcomes of I-SceI-mediated bait DSBs at ASµ2x1 to clustered I-SceI DSBs at ASy12x1 in the form of deletion (red, −) and inversional joining (blue, +). e, Pooled prey junctions from independent ASµ2x1/ASy12x1 B cell libraries (n = 2, left panel, emulsion-mediated PCR; n = 2, right panel, linear-amplification-mediated HTGTS). f, Southern blot for Sµ inversion on the productive allele of CH12F3 cells with non-productive allele deleted. g, IgA CSR on day 3 for CH12F3 (non-productive ASµ-Sα, productive allele Sµ(INV)) cells stimulated with anti-CD40, IL4 and TGFβ. h, IgA CSR on CH12F3 (productive allele Sµ(INV), non-productive allele ASµ-Sα) cells stimulated with anti-CD40, IL4 and TGFβ. Two independent clones of CH12F3 (Sµ(INV), non-productive ASµ-Sα). For detailed legends and further discussion refer to the Supplementary Information.
Extended Data Figure 4 | Orientation-biased joining between AID-initiated rSμ and downstream AID-initiated S-region breaks in anti-CD40/IL4-activated and LPS-activated Sμ-truncated B cells.  

a, 150-bp rSμ sequence used as HTGTS bait with red arrow indicating rSμ 5’-broken end HTGTS primer; red and blue vertical lines indicate canonical AGCT or other RGYW AID-targeting motifs, respectively. Distribution of rSμ break points in junctions to downstream S regions recovered from anti-CD40/IL4-stimulated rSμ B cells. 
b, HTGTS analyses of anti-CD40/IL4-activated rSμ B cells, 5’ rSμ (red primer, n = 3) AID-initiated broken end junctions to AID-initiated DSBs in Sγ1 and Sγ2 which include deletional joining (orientation, red) or inversions (orientation, blue). 
c, HTGTS analyses of anti-CD40/IL4-activated rSμ B cells, 3’ rSμ (blue primer, n = 3) AID-initiated broken end junctions to AID-initiated DSBs in Sγ1 and Sγ2 which includes excision circle (orientation, blue) or inversions (orientation, red). 
d, Bar graph showing the percentage of junctions (average ± s.d.) from anti-CD40/IL4-activated rSμ 5’-broken end libraries mapped to Sγ1 and Sc. 

e, Bar graph showing the percentage of junctions (average ± s.d.) from anti-CD40/IL4-activated rSμ 3’-broken end libraries mapped to Sγ1 and Sc. 

f, HTGTS analyses of LPS-activated rSμ B cells, 5’ rSμ (red primer, n = 3) AID-initiated broken end junctions to AID-initiated DSBs in Sγ3, Sγ2b and Sγ2a which include deletional joining (orientation, red) or inversions (orientation, blue). 
g, HTGTS analyses of LPS-activated rSμ B cells, 3’ rSμ (blue primer, n = 3) AID-initiated broken end junctions to AID-initiated DSBs of above LPS-stimulated cells in Sγ3, Sγ2b and Sγ2a which includes excision circle (orientation, blue) or inversions (orientation, red). 
h, i, Percentage of junction distribution at Sγ3, Sγ2b and Sγ2a in both orientations from both 5’-broken end libraries (h) and 3’-broken end libraries (i) are shown as average ± s.d. from three independent experiments. For detailed legends and further discussion refer to the Supplementary Information.
Extended Data Figure 5 | Orientation-biased joining between rSμ and AID-induced Sα DSBs in CSR-activated CH12F3 cells. a, Diagram outlining potential junction outcomes from 5’ rSμ (red primer) or 3’ rSμ (blue primer) AID-initiated broken end junctions to AID-initiated DSBs in Sα upon anti-CD40, IL4 and TGFβ stimulation of ΔSμ CH12F3 cells. b, c, Top panel shows HTGTS libraries analyses of day 2 (b) and day 3 (c) stimulated CH12F3 (non-productive allele ΔSμ-Sα, productive allele ΔSα-Sα) cell cloning from 5’-broken end rSμ (red primer, n = 3), whereas lower panel shows HTGTS libraries cloning from 3’-broken end rSμ (blue primer, n = 3). d, Bar graph shows percentage of junctions (average ± s.d.) for 5’-broken end and 3’-broken end libraries indicated in b and c. e, Bar graph with percentage of junctions (average ± s.d.) from rSμ libraries mapped to prey Sα in the deletion (DEL) or inversion (INV) for 5’-broken end libraries and in excision circle (EC) or inversion orientation for 3’-broken end libraries. For detailed legends and further discussion refer to the Supplementary Information.
Day 4 CSR levels

| Treatment              | IgG1 (%) | IgE (%) |
|------------------------|----------|---------|
| WT (n=6)               | 45% ± 2.8| 14 ± 2  |
| ATM<sup>−/−</sup> (n=3) | 18.1 ± 0.6| N/A     |
| H2AX<sup>−/−</sup> Scy1<sup>−/−</sup>SceI<sup>−/−</sup> (n=3) | 8.3 ± 1.3| 3.8 ± 0.3|
| 53BP1<sup>−/−</sup> (n=8) | 3.2 ± 1.1| N/A     |
| Rif1<sup>−/−</sup> CD19<sup>Cre/Cre</sup> (n=3) | 6.6±1.6  | N/A     |

b) Graphs showing the distribution of IgG1 and IgE in different cell lines.

c) Diagrams illustrating the percentage of total events for different cell lines.

d) Diagrams showing the percentage of total events for different cell lines.

e) Diagram showing the process of deletion and resection.

f) Graphs showing the distribution of IgG1 and IgE in different cell lines.
Extended Data Figure 6 | Level of junctions to downstream S regions in wild-type and DSBR-deficient 5’Sμ HTGTS libraries correlate with CSR levels; 5’Sμ break site undergoes variable degrees of resection from stimulated DSBR-deficient B cells. a, Table showing IgG1 and IgE CSR levels of splenic B cells from various genotypes (with number of replicates indicated) activated in vitro with anti-CD40 and IL4. FACS was performed on day 4 and values indicate average ± s.d. WT, wild type. b, Left panel shows bar graph for percentage of junctions (average ± s.d.) recovered from wild-type 5’Sμ 5’-broken end libraries mapped to Sμ, Sγ1 and Sc over the total number of junctions identified from the 200-kb Igh constant region. Remaining panels show the similar results from different DSBR-deficient backgrounds using the same 5’-broken end primer. c, d, 5’Sμ 5’-broken end HTGTS libraries analyses from H2AX−/− and RIF1fl/flCD19cre B cells are shown respectively. e, Diagram of potential junction outcomes from 5’Sμ AID-initiated 5’-broken end junctions to AID-initiated DSBs in Sγ1 and Sc. f, Data from HTGTS libraries mapped to the 20-kb region flanking 5’Sμ break site from B cells stimulated with anti-CD40/IL4 in wild-type and DSBR-deficient backgrounds. For detailed legends and further discussion refer to the Supplementary Information.
Extended Data Figure 7 | Orientation-biased joining between rS and AID-induced Sc and Se DSBs in wild-type, ATM-deficient, and 53BP1-deficient B cells. a, Diagram of potential junction outcomes from 5’ rSµ AID-initiated broken end junctions to AID-initiated DSBs in Sc and Se as described earlier. b–d, Linear plots of pooled junctions across the 200-kb IgH constant region (first panel), the 20-kb region flanking rSµ break site (second panel), the 20-kb region flanking downstream Sc1 (third panel) and Se (last panel) from three independent wild-type (b), ATM− (c), or 53BP1− (d) deficient 5’-broken end rSµ libraries. e, Bar graphs showing inversion:deletion (INV/DEL) bias ratios of HTGTS Sc1 (left panel) and Se (right panel) junctions in different genotypes, showing average ± s.d. from three independent libraries for each genotype. P values calculated by unpaired two-tailed t-tests. f, Bar graphs showing percentage of long resection of Sc1 (left) and Se (right) junctions in different genotypes as average ± s.d. n.s., not significant. P values calculated by unpaired two-tailed t-tests. g, Bar graphs showing the number of junctions (average ± s.d.) recovered from above experiments from 5’-broken end HTGTS libraries for the indicated genotypes (n = 3 for each) at the break site rSµ and downstream Sc1 and Se regions as a percentage of the total number of junctions mapped to the 200-kb IgH constant region. For detailed legends and further discussion refer to the Supplementary Information.
Extended Data Figure 8 | Orientation-biased joining of I-SceI DSBs at Sy1 to AID-induced S-region breaks in various DSBR-deficient backgrounds.

a, Schematic illustration of the ΔSy1^2xl allele and joining outcomes from 3'-broken end (red arrow) to AID-initiated upstream Sy and downstream Sc DSBs. b, Linear distribution of junctions between ΔSy1^2xl 3'-broken end to AID-induced Sy/Sc region DSBs in anti-CD40/IL4-stimulated wild-type (b, n = 4), ATM^-/- (c, n = 3), H2AX^-/- (d, n = 3) and 53BP1^-/- (e, n = 4) cells across the 200-kb Igh region (left panels), 10-kb Sy (middle panels) and Sc (right panels). f, Bar graphs (average ± s.d.) showing the percentage of junctions mapped at ΔSy1^2xl (break site), Sy and Sc over the total number of junctions in the 200-kb Igh constant region in different genotypes. g, Bar graphs (average ± s.d.) showing the percentage of junctions from above various genotypes of ΔSy1^2xl 3'-broken end libraries mapped to Sy and Sc as average ± s.d. h, Bar graph (average ± s.d.) showing comparison of percentages of junctions cloned using ΔSy1^2xl 3'-broken end involving resection of Sy (top panel) and Sc (bottom panel) breaks in indicated backgrounds. For detailed legends and further discussion refer to the Supplementary Information.
Extended Data Figure 9 | Inhibition of resection in 53BP1-deficient B cells by an ATM inhibitor (ATMi) does not rescue directional CSR joining to Sc1. a–d, Linear plots of pooled junctions across the 200-kb Igh constant region (left panels), the 20-kb region flanking downstream Sc1 (middle panels) and Sc (right panels) from wild type plus DMSO control (a, n = 2), wild type plus ATMi (b, n = 3), 53BP1^−/− plus DMSO (c, n = 3) and 53BP1^−/− plus ATMi (d, n = 3) libraries are shown as above. e, Bar graph showing the percentage of Sμ, Sc1 and Sc junctions (average ± s.d.) from wild type plus DMSO 5'Sμ libraries (left) and wild type plus ATMi 5'Sμ-broken end libraries (right, n = 3). f, Bar graph showing the percentage of Sμ, Sc1 and Sc junctions (average ± s.d.) from 53BP1^−/− plus DMSO (left) and 53BP1^−/− plus ATMi (right) 5'Sμ libraries. For detailed legends and further discussion refer to the Supplementary Information.
Extended Data Table 1 | Statistical comparison for orientation bias and resection in S\textsubscript{1} and S\textsubscript{c} junctions from wild-type and DSBR-deficient B cell libraries

|   | WT (n=5) | ATM\textsuperscript{a} (n=3) | H2AX\textsuperscript{a} (n=3) | 53BP1\textsuperscript{a} (n=8) | 53BP1\textsuperscript{a}+Al (n=3) | Ref\textsuperscript{a} (n=3) |
|---|----------|-------------------------------|--------------------------|----------------|------------------|----------------|
| a |          |                               |                           |                 |                  |                 |
|   | WT (n=5) | ATM\textsuperscript{a} (n=3) | H2AX\textsuperscript{a} (n=3) | 53BP1\textsuperscript{a} (n=8) | 53BP1\textsuperscript{a}+Al (n=3) | Ref\textsuperscript{a} (n=3) |
|   | P values calculated by unpaired two-tailed t-test for the degree of bias in the S\textsubscript{1} (a) and S\textsubscript{c} (b) junctions between wild-type and DSBR-deficient B cells with full-length S\textsubscript{m} for experiments described in Figs 3 and 4. Orientation bias was calculated as described in the Methods for individual libraries. Numbers in parenthesis indicate independent experiments performed for each genotype. N/A, not available. | 
|   |          |                               |                           |                 |                  |                 |
| b |          |                               |                           |                 |                  |                 |
|   | WT (n=5) | ATM\textsuperscript{a} (n=3) | H2AX\textsuperscript{a} (n=3) | 53BP1\textsuperscript{a} (n=8) | 53BP1\textsuperscript{a}+Al (n=3) | Ref\textsuperscript{a} (n=3) |
|   | P values calculated by unpaired two-tailed t-test for the level of resections in the S\textsubscript{1} (c) and S\textsubscript{c} (d) junctions for experiments described above. Percentage of resection was calculated as described in the Methods for individual libraries. Numbers in parenthesis indicate independent experiments performed for each genotype. | 
|   |          |                               |                           |                 |                  |                 |
| c |          |                               |                           |                 |                  |                 |
|   | WT (n=5) | ATM\textsuperscript{a} (n=3) | H2AX\textsuperscript{a} (n=3) | 53BP1\textsuperscript{a} (n=8) | 53BP1\textsuperscript{a}+Al (n=3) | Ref\textsuperscript{a} (n=3) |
|   | P values calculated by unpaired two-tailed t-test for the level of resections in the S\textsubscript{1} (d) and S\textsubscript{c} (d) junctions for experiments described above. Percentage of resection was calculated as described in the Methods for individual libraries. Numbers in parenthesis indicate independent experiments performed for each genotype. | 
|   |          |                               |                           |                 |                  |                 |
| d |          |                               |                           |                 |                  |                 |
|   | WT (n=5) | ATM\textsuperscript{a} (n=3) | H2AX\textsuperscript{a} (n=3) | 53BP1\textsuperscript{a} (n=8) | 53BP1\textsuperscript{a}+Al (n=3) | Ref\textsuperscript{a} (n=3) |
|   | P values calculated by unpaired two-tailed t-test for the level of resections in the S\textsubscript{1} (d) and S\textsubscript{c} (d) junctions for experiments described above. Percentage of resection was calculated as described in the Methods for individual libraries. Numbers in parenthesis indicate independent experiments performed for each genotype. | 
|   |          |                               |                           |                 |                  |                 |