Activation-Induced Cytidine Deaminase Targets DNA at Sites of RNA Polymerase II Stalling by Interaction with Spt5

Rushad Pavri,[1] Anna Gazumyan,[1,2] Mila Jankovic,[1] Michela Di Virgilio,[1] Isaac Klein,[1] Camilo Ansarah-Sobrinho,[3] Wolfgang Resch,[3] Arito Yamane,[3] Bernardo Reina San-Martin,[1,4] Vasco Barreto,[1,5] Thomas J. Nieland,[6] David E. Root,[6] Rafael Casellas,[3,*] and Michel C. Nussenzweig[1,2,*]

[1]Laboratory of Molecular Immunology
[2]Howard Hughes Medical Institute
The Rockefeller University, New York, New York 10065, USA
[3]Genomics and Immunity, The National Institute of Arthritis and Musculoskeletal and Skin Diseases (NIAMS), and Center for Cancer Research, National Cancer Institute (NCI), National Institutes of Health, Bethesda, MD 20892, USA
[4]Institut de Génétique et de Biologie Moléculaire et Cellulaire (IGBMC), INSERM U964 / CNRS UMR7104 / Université de Strasbourg, 67404, Illkirch, France
[5]Laboratory of Epigenetics and Soma, Instituto Gulbenkian de Ciência, P-2780-156 Oeiras Portugal
[6]RNAi Platform, The Broad Institute of MIT and Harvard, Cambridge, MA 02142, USA
*Correspondence: nussen@mail.rockefeller.edu (M.C.N.), casellar@mail.nih.gov (R.C.)
DOI 10.1016/j.cell.2010.09.017

SUMMARY

Activation-induced cytidine deaminase (AID) initiates antibody gene diversification by creating U:G mismatches. However, AID is not specific for antibody genes; Off-target lesions can activate oncogenes or cause chromosome translocations. Despite its importance in these transactions little is known about how AID finds its targets. We performed an shRNA screen to identify factors required for class switch recombination (CSR) of antibody loci. We found that Spt5, a factor associated with stalled RNA polymerase II (Pol II) and single stranded DNA (ssDNA), is required for CSR. Spt5 interacts with AID, it facilitates association between AID and Pol II, and AID recruitment to its *Ig* and non-*Ig* targets. ChIP-seq experiments reveal that Spt5 colocalizes with AID and stalled Pol II. Further, Spt5 accumulation at sites of Pol II stalling is predictive of AID-induced mutation. We propose that AID is targeted to sites of Pol II stalling in part via its association with Spt5.

INTRODUCTION

AID is a cytidine deaminase that initiates immunoglobulin somatic hypermutation (SHM) and class switch recombination (CSR) (Muramatsu et al., 2000, 1999; Revy et al., 2000). It does so by deaminating cytidine residues in ssDNA (Bransteitter et al., 2003; Chaudhuri et al., 2003; Dickerson et al., 2003; Pham et al., 2003; Ramiro et al., 2003; Sohail et al., 2003). The resulting U:G mismatches can be processed by several different

DNA repair pathways to produce mutations or DNA double-strand breaks (Di Noia and Neuberger, 2007; Peled et al., 2008).

In addition to diversifying the antibody repertoire by SHM and CSR, AID also contributes to malignant transformation by initiating chromosome translocations (Ramiro et al., 2006; Ramiro et al., 2004; Robbiani et al., 2008; Nussenzweig and Nussenzweig, 2010) and by producing mutations in non-*Ig* genes such as *Bcl-6* (Pasqualucci et al., 1998, 2001; Shen et al., 1998). Although the comparative frequency of mutation at non-*Ig* genes is low, AID mutates 25% of the genes transcribed in germinal center B cells, where it is normally expressed (Liu et al., 2008). Furthermore, even low levels of mutation are sufficient to produce substrates for translocation (Robbiani et al., 2008; Robbiani et al., 2009). Consistent with the breadth of genes found mutated by AID in germinal center B cells, AID overexpression in transgenic mice leads to extensive translocation of non-*Ig* genes and cancer (Robbiani et al., 2009). In addition, AID deregulation has been associated with *H. pylori* infection and gastric cancer (Matsumoto et al., 2007), and with translocation in prostate malignancy (Lin et al., 2009). Finally, AID is also of interest because it has been implicated as a cytosine demethylase involved in reprogramming pluripotent cells (Bhutani et al., 2010; Morgan et al., 2004; Popp et al., 2010; Rai et al., 2008).

Although the precise mechanism which targets AID to *Ig* genes is unknown, AID-induced mutations are associated with transcription and are most prevalent in a 2 kb region beginning downstream of the promoter (Di Noia and Neuberger, 2007; Peled et al., 2008; Stavnezer et al., 2008; Storb et al., 2007). Transcription is also required for CSR, suggesting that RNA polymerase II (Pol II) might facilitate AID access to target DNA (Di Noia and Neuberger, 2007; Peled et al., 2008; Stavnezer-Nordgren and Sirlin, 1986; Stavnezer et al., 2008; Storb et al., 2007; Yancopoulos et al., 1986). This idea was confirmed by the observation that transcriptional regulatory elements are essential to both hypermutation and CSR (reviewed in (Di Noia

and Neuberger, 2007; Peled et al., 2008; Stavnezer et al., 2008; Storb et al., 2007)). Consistent with these findings, AID is associated with Pol II (Nambu et al., 2003). In *E. coli* and in in vitro assays, transcription liberates ssDNA, the substrate for AID (Bransteitter et al., 2003; Chaudhuri et al., 2003; Dickerson et al., 2003; Pham et al., 2003; Ramiro et al., 2003; Sohail et al., 2003). In more complex systems, transcription is also required for AID to access chromatinized substrates (Shen et al., 2009); however, the role of transcription in SHM and CSR is not completely understood.

AID is a relatively small enzyme composed of 198 amino acids (Muramatsu et al., 1999). It preferentially deaminates cytosine residues embedded in WRCY consensus sequences (where W = adenosine/thymine, R = purine, and Y = pyrimidine) (Rogozin and Kolchanov, 1992). This preference is dictated in part by the composition of the active site (Wang et al., 2010). However, WRCY motifs are present throughout the genome and cannot fully account for AID target choice. While several AID cofactors have been reported, including replication protein A (RPA), protein kinase-Ar1α, and CTNNBL1, none of these are known to impart specificity to AID (Basu et al., 2005; Chaudhuri et al., 2004; Conticello et al., 2008; McBride et al., 2006; Pasqualucci et al., 2006).

Here we report that Spt5, a factor normally associated with stalled or paused Pol II, is required for CSR. Spt5 is required for AID recruitment to switch regions, for switch region mutation, and for AID association with Pol II. Furthermore, genes that accumulate Spt5 also accumulate AID and suffer AID-dependent mutations.

RESULTS

shRNA Screen for CSR in CH12 Cells

To identify factors required for CSR, we developed a lentiviral-based shRNA screening strategy using the murine B cell line, CH12. This cell line expresses AID and undergoes CSR to IgA in response to stimulation with interleukin 4 (IL-4), CD40 ligation and transforming growth factor β (TGFβ) (Nakamura et al., 1996). AID is limiting for CSR in these cells because its knockdown by specific shRNA results in reduction of CSR in a manner consistent with the decrease in AID protein levels (Figures 1A and 1B). In addition, shRNA-induced knockdown of other known regulators of the reaction result in the expected decrease in CSR (Figure 1C). Therefore, the level of CSR in CH12 cells is limited by the amount of AID and its cofactors suggesting that CSR can be used as an assay for additional factors that might be required for AID function in these cells. To screen for such factors, we developed an shRNA screen for CSR in CH12 cells.

We assembled an shRNA lentiviral library containing 8797 hairpins representing 1745 genes selected primarily on the basis of their expression in CH12 cells (Table S1A available online) and germinal center B cells (Klein et al., 2003; Moffat et al., 2006; Root et al., 2006) (Table S1B). Factors directly involved in transcription, or in cotranscriptional and posttranscriptional events, such as mRNA processing, turnover and export, and DNA repair factors, kinases and phosphatases were preferentially retained (reviewed in (Di Noia and Neuberger, 2007; Peled et al., 2008; Stavnezer et al., 2008; Storb et al., 2007)). Finally, we added

several DNA repair factors and transcription-associated factors that were not selected based on their expression, but that might be required based on the literature (Table S1B).

The recombinant lentiviruses were prepared and screened in a 96-well format in triplicate (Moffat et al., 2006; Root et al., 2006), and CSR and viability for each sample was evaluated by flow cytometry (Figure 1D). Each plate contained three negative control shRNAs (shLacZ, shGFP and shRFP) and a positive control AID shRNA (Figure 1E, shAID). Positive hits were defined as viable shRNA-expressing clones that exhibited at least 50% reduction in CSR compared to the controls (Figures 1E). Positive hits were rearrayed and rescreened in triplicate. The screen uncovered 181 hits of which 28 were previously shown to be involved directly or indirectly in CSR (Figure 1F and Table S2).

We tested the candidate hairpins for knockdown of the target mRNA and their effects on AID mRNA, and μ-and α-germline transcripts (GLTs). We focused on those genes that did not alter AID mRNA or μ-and α-GLTs and assayed for association of the corresponding protein with AID by coimmunoprecipitation.

Spt5 Is Required for CSR in CH12 and Primary B Cells

Suppressor of Ty 5 homolog (Spt5), a transcription elongation factor associated with paused Pol II, was selected for further analysis (reviewed in Gilmour, 2009; Lis, 2007; Peterlin and Price, 2006). Two unique shRNAs targeting Spt5 decreased CSR (Figure 2A), and the decrease was specific as determined by complementation with an Spt5 cDNA lacking the sequence targeted by shSpt5-1 (Spt5Δ), but not by a cDNA with intact target sites (Figure S1). Spt5 knockdown also decreased switch region hypermutation (Figure 2B), but did not alter the steady state levels of AID, or μ- or α-germline mRNA (Figure 2C), or cell division as measured by CFSE dye dilution (Figure S2A). Finally, CH12 cells expressing these shRNAs showed decreased Spt5 protein, whereas AID protein levels were unaltered (Figure 2D).

Similarly, primary B cells treated with LPS and IL-4 and infected with retroviruses directing the synthesis of shRNAs specific for Spt5 showed decreased Spt5 protein (Figure S2B) and a concomitant decrease in CSR to IgG1 (Figure 2E). We conclude that Spt5 is required for CSR in primary B cells.

Spt5 Associates with AID in Fibroblasts and Primary B Cells

Since both Spt5 (Wada et al., 1998) and AID (Nambu et al., 2003) associate with Pol II, we asked if Spt5 is also associated with AID. Endogenous Spt5 was coprecipitated from 293T cells transfected with Flag-tagged AID (F-AID) using anti-Flag antibodies (Figure 3A). Conversely, F-AID was coprecipitated by anti-Spt5 antibodies from the same cells under identical conditions (Figure 3B). In contrast, APOBEC-2, a closely related deaminase, did not coprecipitate with Spt5 in either direction (Figures 3A and 3B). Finally, endogenous Spt5 was also coimmunoprecipitated with F-AID from activated B cells isolated from F-AID knockin mice (AID$^{F/F}$ mice) that express physiological levels of AID, and undergo near-normal levels of CSR (Figure 3C and Figure S3). DNA or RNA was not required for the Spt5-AID interaction since the extracts were treated with Benzonase, a nuclease that digests all nucleic acids. We conclude that
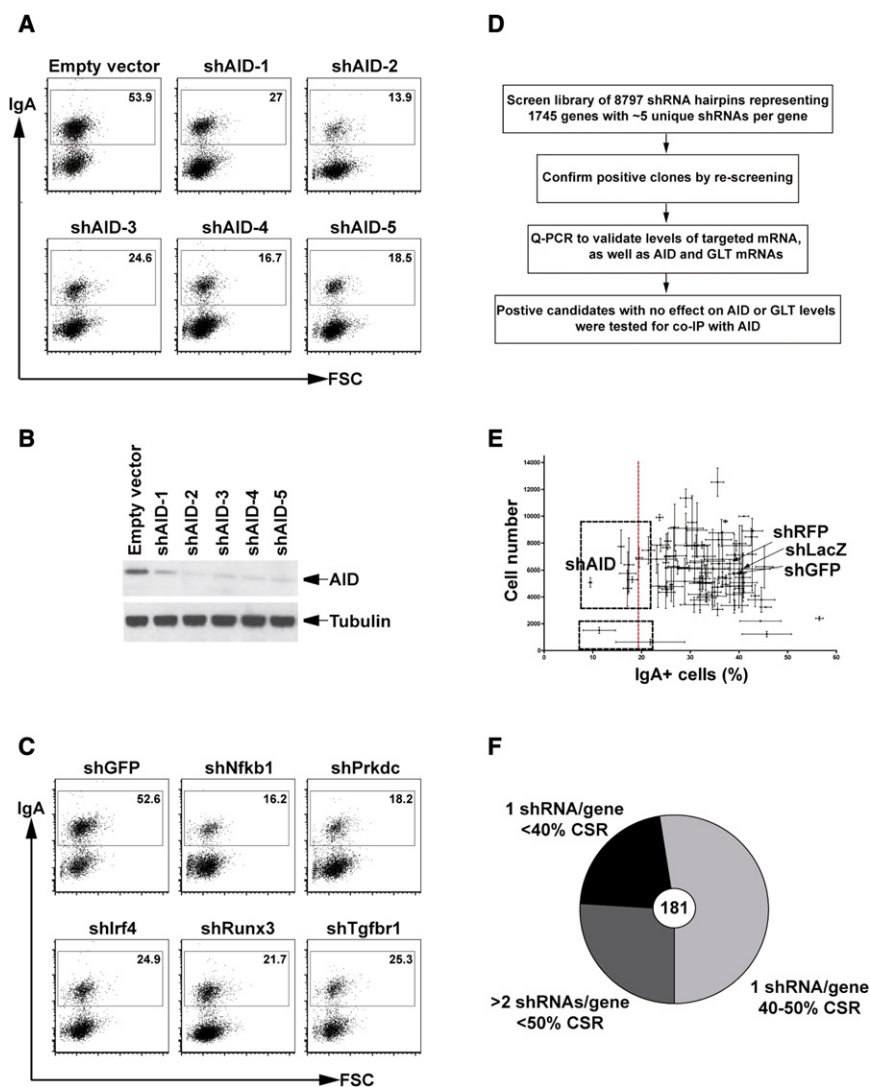
Figure 1. Lentiviral-Based shRNA Screen in CH12 Cells

(A) CSR is sensitive to AID depletion. Flow cytometry plots of CH12 cells infected with five unique shRNAs to AID (shAID1-5) and empty vector control. Numbers indicate the percentage of IgA positive cells.

(B) AID protein levels in whole cells extracts from the same cells shown in (A). Western blots were probed with an anti-AID antibody and anti-tubulin as a loading control.

(C) Representative flow cytometry plots of CH12 cells infected with shRNAs against genes involved in CSR: *Nfkb1* (NFκB p50 subunit), *Prkdc* (DNAPKcs catalytic subunit), *Irf4*, *Runx3*, and *Tgfbr1* (TFGβ receptor 1) (Table S3).

(D) Schematic of the experimental approach used for the screen.

(E) Representative data from a single plate of shRNAs analyzed in triplicate. Error bars show the standard deviation obtained from the three replicate plates for %IgA+ cells (*x* axis) and cell numbers (*y* axis). Negative (LacZ, GFP and RFP) and positive (AID) controls shRNAs are indicated. The dotted red line shows the position corresponding to 50% of the averaged negative control CSR value. Two sets of clones with < 50% CSR are boxed. The upper box consists of viable clones that are considered as positive hits. The lower box contains clones that were discarded due to poor viability.

(F) Pie chart showing the distribution of 181 selected hits as a function of the number of shRNAs per gene and their effect on CSR as calculated based on the percentage reduction of CSR compared to the averaged negative control values as shown in (E).

See also Table S1 and Table S2.

Spt5 is associated with AID in transfected fibroblasts and activated B cells.

AID and Spt5 Can Associate In Vitro

Since Spt5 can directly associate with Pol II in vitro (Yamaguchi et al., 1999a), we asked if this was the case for the interaction between Spt5 and AID. To test this idea, bacterially expressed GST-AID was captured on glutathione sepharose beads and incubated with purified recombinant Spt5. Only a fraction of the Spt5 was specifically bound to GST-AID, but there was no binding to GST-APOBEC2 or GST (Figure 3D). Thus, AID and Spt5 can interact in vitro, but the association is weak and other factors or posttranslational modifications likely facilitate this association in vivo. Consistent with this idea, extracts prepared in the presence of phosphatase inhibitors showed increased AID-Spt5 association (Figure S4A). Although AID activity in vivo is enhanced by phosphorylation of serine 38 (S38) or threonine 140 (T140) (Chaudhuri et al., 2004; McBride et al., 2006, 2008),

neither S38A nor T140A mutations alter the interaction of AID with Spt5 (Figure S4B).

Pol II Association with AID Is Dependent on Spt5

Spt5 binds to Pol II and induces stalling in vitro (Yamaguchi et al., 1999a) and in vivo (Lis, 2007; Rahl et al., 2010). In addition, Spt5 also functions as an adaptor that links several cotranscriptional activities to the Pol II machinery (see Discussion). To determine whether Spt5 is required for AID association with Pol II, we depleted Spt5 from CH12 cells expressing F-AID and examined the effects on the association between AID and Pol II. Whereas both Pol II and Spt5 are normally coprecipitated with F-AID, the association between AID and Pol II was decreased in Spt5-depleted cells when compared to the shLacZ control, suggesting that the AID-Pol II interaction (Nambu et al., 2003) is dependent on Spt5 (Figure 3E). In contrast, Pol II depletion did not alter the AID-Spt5 interaction suggesting that Pol II is not essential for this association
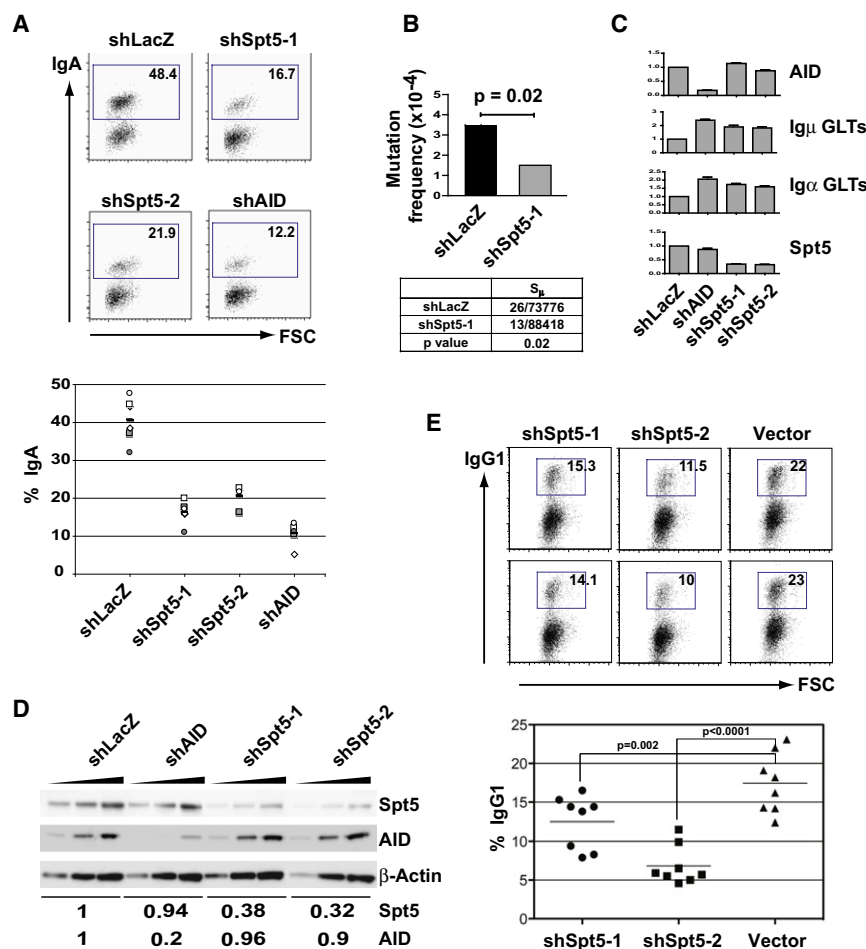
Figure 2. Spt5 Is Required for CSR and Switch Region Mutation in CH12 Cells

(A) Upper panel shows representative flow cytometry plots of CH12 infected with two unique shRNAs to Spt5 (shSpt5-1 and shSpt5-2) and controls (shLacZ and shAID) and stimulated to undergo CSR. Numbers indicate percentage of IgA positive cells. The graph in the lower panel summarizes the data from four to six independent experiments.

(B) Decreased switch region mutation after Spt5 knockdown in CH12 cells. Upper panel represents the mutation frequency and corresponding p value from control (shLacZ) and shSpt5-1 infected cells stimulated to undergo CSR for 48 hr. The table in the lower panel summarizes the mutation analysis (represented as unique mutations/nucleotides sequenced).

(C) Graphs show Q-PCR analysis for Spt5, AID, Igα and Igμ germline (GLT) mRNA levels in activated CH12 cells infected as in (A) with the indicated shRNAs. The data summarizes three independent experiments with standard deviation indicated as error bars. In all cases, shLacZ was assigned an arbitrary value of 1.0.

(D) Western blot analysis of Spt5 and AID protein levels in WCEs from activated CH12 cells infected with the indicated shRNAs. Threefold serial dilutions of WCEs were loaded. β-Actin was used as a loading control. Numbers below the blots represent normalized band intensities for Spt5 and AID with the shLacZ lanes assigned an arbitrary value of 1.

(E) Spt5 is required for CSR in primary B cells. Representative flow cytometry plots of B cells stimulated with LPS + IL-4 and infected with retroviruses expressing shSpt5-1, shSpt5-2 or LMP vector alone. Efficiency of switching was determined by gating on GFP-positive cells. Numbers indicate percentage of IgG1 positive cells. The graph in the lower panel summarizes the data from three independent experiments.

See also Figure S1 and Figure S2.

(Figure S4C). We conclude that Spt5 serves as an adaptor that recruits AID to Pol II.

AID Recruitment to *Ig* Switch Regions Is Dependent on Spt5

To determine whether AID recruitment to the *Ig* switch region is dependent on Spt5, we performed quantitative PCR-based ChIP analysis with two different anti-AID antibodies (Chaudhuri et al., 2004; McBride et al., 2006). Spt5 depletion resulted in significant reduction of AID occupancy in the switch region (Figure 3F, $p = 1.4 \times 10^{-6}$). We conclude that Spt5 is required for AID recruitment to the *Ig* switch region in B cells undergoing CSR.

Spt5 Is Associated with Stalled Pol II in B Cells

AID mutates *Ig* genes and up to 25% of the expressed genes in germinal center B cells (Liu et al., 2008; Pasqualucci et al., 1998, 2001; Shen et al., 1998). To determine whether Spt5 localization in the genome of activated B cells coincides with AID-dependent mutation, we performed genome-wide chromatin immunoprecipitation and sequencing (ChIP-seq) with antibodies against

Spt5 and Pol II. Spt5 was found throughout the genome of activated B cells undergoing CSR (Figure 4 and Table S3A). As in other cell types that have been assayed for Spt5 localization, this protein was also concentrated at promoter regions coincident with Pol II peaks in activated B cells (Gilmour, 2009; Lis, 2007; Peterlin and Price, 2006; Rahl et al., 2010) (Figures 4A–4C).

Spt5 is a stalling factor in vitro (Wada et al., 1998; Yamaguchi et al., 1999b) and associated with stalled Pol II in various cell types in vivo (Rahl et al., 2010; Zeitlinger et al., 2007). The amount of Pol II stalling can be quantitated by calculating a stalling or traveling index (*Is*), which is a ratio of the Pol II density at promoter regions compared to the gene body (Zeitlinger et al., 2007; see Experimental Procedures). Genes with *Is* > 3 are considered stalled whereas those with *Is* < 1 are considered elongating genes (Zeitlinger et al., 2007). Stalling is widespread in the B cell genome (5594 genes, 61%, Table S3A), and in addition, the Pol II and Spt5 stalling indices were significantly correlated (Spearman's correlation coefficient, $\rho = 0.8$), consistent with previous observations in other cell types (Nechaev et al., 2010; Rahl et al., 2010) (Figure 4D, and A.Y. and R.C., unpublished data, accession
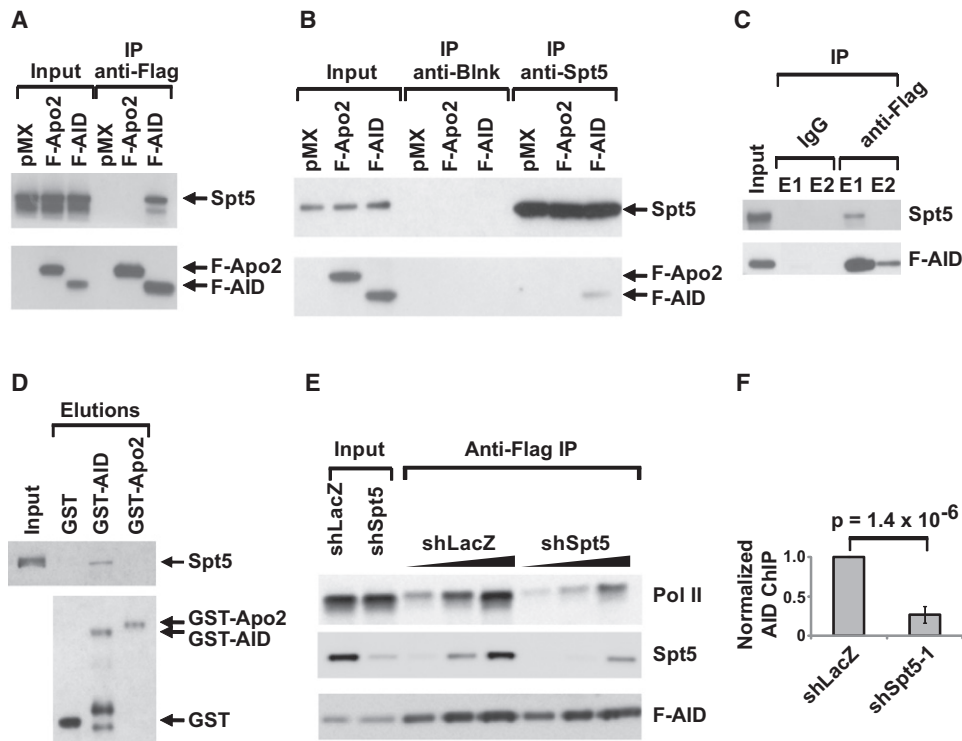
Figure 3. Spt5 Interacts with AID in Fibroblasts and Primary B Cells

(A) Anti-Flag immunoprecipitates from whole cell extracts (WCEs) from 293T cells transfected with Flag-tagged AID (F-AID), or Flag-tagged Apobec2 (F-Apo2) or pMX vector probed with anti-Flag or anti-Spt5 antibodies as indicated.

(B) Anti-Spt5 immunoprecipitates from WCEs from 293T cells transfected as in (A). Blots were probed as in (A). Anti-Blnk was used as an isotype control.

(C) Anti-Flag immunoprecipitates from WCEs from cultured splenic AID$^{F/F}$ B cells. Blots were probed as in (A). E1 and E2 represent first and second elutions with Flag peptide respectively.

(D) Bacterially expressed GST-AID, GST-APOBEC2 (GST-Apo2) or GST alone were bound to glutathione sepharose beads and incubated with purified recombinant Spt5-Spt4 heterodimer (DSIF). Bound material was eluted and analyzed by SDS-PAGE and blotted using antibodies against Spt5 and GST. The input lane for DSIF represents 1% of the amount used in the reaction.

(E) Anti-Flag immunoprecipitates from WCEs of CH12 cells transfected with F-AID and depleted of Spt5 by shSpt5-1. shLacZ is used as a control. Blots were probed as in (A) and with anti-Pol II.

(F) ChIP analysis for AID occupancy in Sμ regions of CH12 cells infected with shSpt5-1 or shLacZ control. Data represents a total of 7 experiments using two different anti-AID antibodies (Chaudhuri et al., 2004; McBride et al., 2006). For each experiment, shLacZ was assigned an arbitrary value of 1. The p value is indicated.

See also Figure S3 and Figure S4.

number GSE24178). Most strikingly, AID occupancy in activated B cells is also tightly correlated with Spt5 (see below and A.Y. and R.C., unpublished data, accession number GSE24178).

To determine how Spt5 accumulation relates to mRNA levels, we compared the density of Spt5 sequence reads to B cell mRNA-seq levels (both measured as reads per kbp per million sequences (RPKM) (Figure 4E, and (Kuchen et al., 2010)). Although there was some correlation between Spt5 and mRNA levels (Figure 4E, ρ = 0.55), there was a 1- to 2-log variation in mRNA levels for genes accumulating similar levels of Spt5. Thus, in B cells, as in other cells (Nechaev et al., 2010; Rahl et al., 2010), Spt5 (or Pol II) accumulation is not necessarily equivalent to cellular mRNA levels.

Spt5 Genomic Occupancy Is Predictive of AID-Dependent Mutation

Upon genome-wide analysis of Spt5 occupancy in the promoter proximal region (−1–2 kb relative to the transcriptional start site

[TSS]), we found that Iμ bore the greatest tag count (Figure 5A and Table S3B). The IgV_H region could not be mapped because each B cell has a unique rearrangement; however, a strong Spt5 signal was found from the IgH enhancer region through the switch region (Figure 5A). Mir142, a robust AID target (Robbiani et al., 2009), is also embedded in a region of high Spt5 accumulation (Figure 5B and Table S3C). In contrast, Taci, Whsc1, H2Eα, A20, Anxa4, and Wdfy3, all of which are expressed in activated B cells (Kuchen et al., 2010), but are not mutated (Liu et al., 2008; Robbiani et al., 2009), do not accumulate Spt5 (Figure 5C and Tables S3A and S3B).

To determine whether Spt5 accumulation is predictive of mutations, we sequenced 10 genes that ranked within the top 5% of genes analyzed for Spt5 tag density (Spt5^{hi}), measured as the density of sequence tags or reads per million base-pairs (TPM), in the promoter-proximal region (Table S3B, Figure 5B and Figure 6, and Kuchen et al., 2010). As controls,
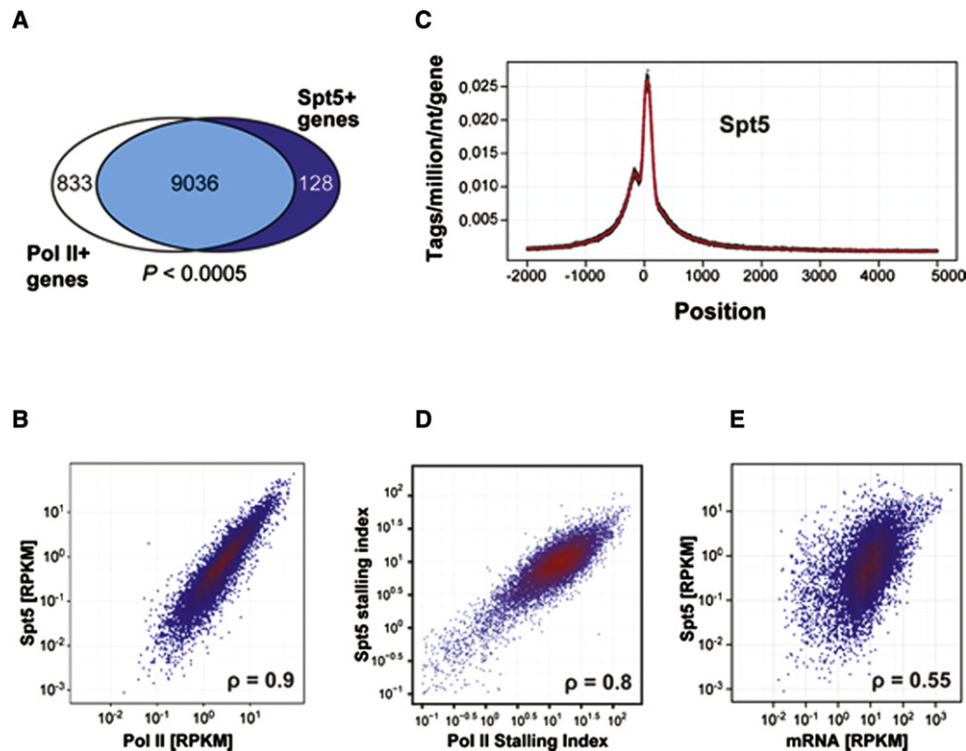
Figure 4. ChIP-Seq Analysis of Spt5 Genomic Occupancy

(A) Venn diagram showing overlap between genes recruiting Spt5 and Pol II using ChIP-Seq data from LPS+IL4 activated B cells (Table S4). There is a significant association between the presence of Spt5 and Pol II at genes (Pearson's Chi-square test; p < 0.0005).

(B) Correlation between Spt5 and Pol II density per gene. For each gene that recruited above-background amounts of Pol II and Spt5, the number of sequence tags aligning between −1 Kb upstream of the transcriptional start site to its transcriptional termination site were normalized per gene length (in Kb), per million aligned reads (reads per Kb per million, RPKM) and shown as a hexagonal binning plot. Spearman's correlation coefficient (ρ) is indicated.

(C) Spt5 profile at all Spt5+ genes from −2 Kb to +5 Kb of the TSS. Data was normalized as reads per million per nucleotide. Dots represent densities at individual nucleotides and the line a 10 nucleotide moving average.

(C) Correlation between the stalling index calculated based on Pol II or Spt5 occupancy (see Experimental Procedures). Spearman's correlation coefficient (ρ) is indicated.

(E) Comparative analysis of transcript levels (determined by mRNA-Seq, [Kuchen et al., 2010]) and Spt5 recruitment at all Spt5+ genes. Spearman's correlation coefficient (ρ) is indicated.

See also Table S3.

we sequenced 8 highly expressed genes (Kuchen et al., 2010) that had a ~4- to 6-fold lower Spt5 tag density (Spt5[lo]) in the same region (Figure 5C and Figure 6 and Table S3B). For each selected gene, a region starting around the TSS, corresponding to the peak of Spt5, and extending ~500–600 bp downstream was sequenced (Figure 5, Figure 6, and Figure S5). Because the rate of mutation at non-*Ig* genes is normally very low unless repair is impaired (Liu et al., 2008; Pasqualucci et al., 1998, 2001; Shen et al., 1998), we used B cells derived from transgenic mice overexpressing AID from the *Igκ* promoter (*Igκ*AID) (Robbiani et al., 2009). These mice display elevated levels of AID protein with concomitant increases in CSR and somatic mutation; nevertheless, they retain AID targeting specificity (Robbiani et al., 2009). All 10 Spt5[hi] genes (Table S3B) were mutated with frequencies from 4.6×10^{-4} for miR142 to 0.8×10^{-4} for H3f3b (Figure 6A and Figure S5). In contrast, none of the eight Spt5[lo] genes (Table S5) were mutated above background levels (Figure 6A and Figure S5).

To determine whether genes occupied by Spt5 correspond to sites of AID recruitment, we compared Spt5 and AID ChIP-seq data (Figure 6B and A.Y. and R.C., unpublished data, accession number GSE24178). Strikingly, the tag density for AID per gene (measured as reads per kilobase per million [RPKM]) was uniformly and directly proportional to the tag density of Spt5 ($\rho = 0.75$, Figure 6B and A.Y. and R.C., unpublished data, accession number GSE24178). To determine if AID recruitment to non-*Ig* genes was dependent on Spt5, we performed ChIP for AID localization at the *Gas5* gene, a stalled gene (Table S3A) which accumulates AID-mediated mutation (Figure 6A). As shown in Figure 6C, AID recruitment to *Gas5* is impaired upon Spt5 depletion. We conclude that Spt5 and AID accumulation coincide genome-wide and that high density Spt5 occupancy is predictive of AID-mediated mutation.

DISCUSSION

Genetic and biochemical evidence indicate that AID initiates SHM, CSR and chromosome translocation by deaminating
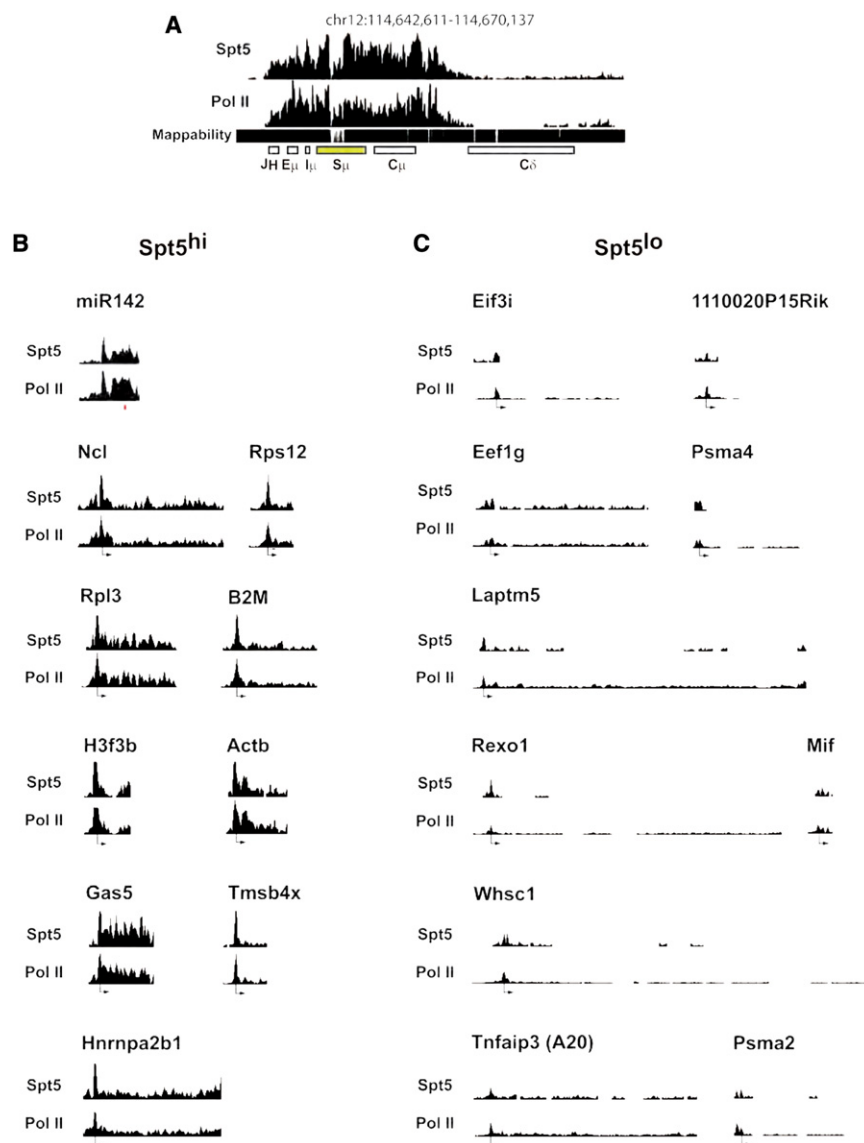
Figure 5. ChIP-seq Profiles of Spt5 on Selected Genes

(A, B, and C) Pol II and Spt5 reads per million plotted in 100 bp windows across (A) the *Igμ* locus, (B) Spt5^{hi}, and (C) Spt5^{lo} genes. The axes scales are identical for all histograms. Tag mappability (shown below) was calculated based on the percentage of 36 nt sequences that uniquely aligned to the genomic site with a 10 bp window resolution. Only windows with a significant enrichment compared to a random background model are shown. The location of the TSS for each gene is indicated. The histograms cover the length of the gene. *Whsc1* and *Tnfaip3* were previously sequenced (Robbiani et al., 2009). All profiles were generated using the UCSC genome browser. See also Table S3.

cytidine residues in ssDNA that are exposed during transcription (Chaudhuri and Alt, 2004; Di Noia and Neuberger, 2007; Nussenzweig and Nussenzweig, 2010; Peled et al., 2008; Stavnezer et al., 2008). AID initiated processes are therefore limited by regulators of transcription initiation such as PTIP, which facilitates Pol II access to specific switch regions by regulating their H3K4 methylation (Daniel et al., 2010). However, active transcription is not sufficient to allow AID access to DNA, and cannot explain why AID-mediated lesions are found primarily in the promoter proximal region of only some transcribed genes. Since Pol II stalling is a feature of promoter-proximal regions, the observation that Spt5, a stalling factor, associates with AID and is required for AID localization to target genes, provides a molecular explanation for the pattern of mutation.

Inducible transcription of genes carrying paused Pol II is an important mechanism for regulating gene expression (Gilmour, 2009; Lis, 2007; Peterlin and Price, 2006; Bai et al., 2010; Core

et al., 2008; Guenther et al., 2007; Lefebvre et al., 2002; Muse et al., 2007; Zeitlinger et al., 2007; Bentley and Groudine, 1986; Krumm et al., 1992;Raschke et al., 1999;Kao et al., 1987). Pausing is typically found downstream of promoters and is associated with permanganate sensitivity, which is indicative of the presence of ssDNA (Giardina et al., 1992).

**Spt5 Is Required for Pol II Stalling In Vitro and In Vivo**

Spt5 was originally identified as an elongation factor in a yeast suppressor screen (Swanson et al., 1991). It was subsequently purified biochemically as a heterodimeric complex with Spt4 called 5,6-dichloro-1-β-d-ribofuranosyl-benzimidazole (DRB) sensitivity inducing factor (DSIF) (Wada et al., 1998; Yamaguchi et al., 1999b). DSIF, in association with negative elongation factor (NELF), binds to Pol II and induces pausing in vitro (Wada et al., 1998; Yamaguchi et al., 1999a). Genomewide ChIP studies have established a strong correlation between Spt5 and Pol II stalling in vivo (Rahl et al., 2010). These and related studies showed that the presence of Pol II in promoter regions does not necessarily correlate with transcription (Bai et al., 2010; Gilmour, 2009; Lefebvre et al., 2002; Lis, 2007; Nechaev et al., 2010; Peterlin and Price, 2006; Rahl et al., 2010). Consistent with these studies, we find only a partial correlation between Spt5 or Pol II occupancy and mRNA levels in activated B cells (Figure 4E), and importantly, that shRNA knockdown of Spt5 did not decrease AID mRNA, or Igμ or Igα sterile transcripts (Figures 2C and 2D).

Current models suggest that the stalled Pol II complex is reactivated by inductive signals that recruit the P-TEFb kinase, which phosphorylates Pol II and Spt5, thereby releasing NELF from the complex and activating transcription (Kim and Sharp, 2001; Marshall et al., 1996; Marshall and Price, 1995; Wada et al.,
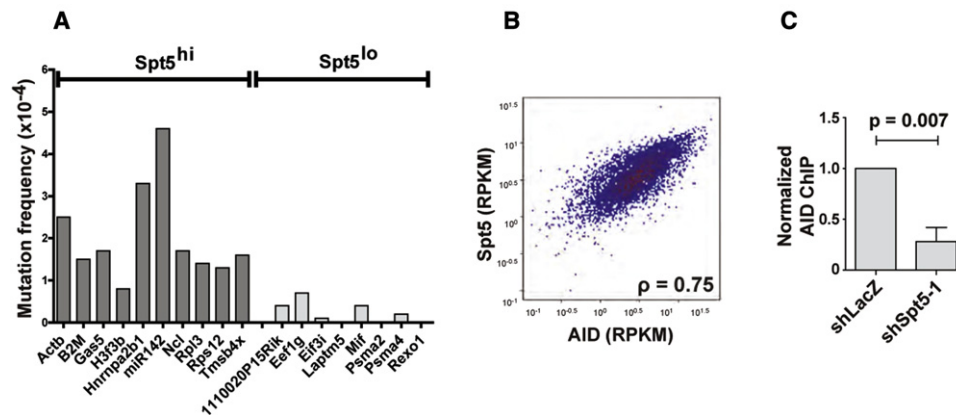
Figure 6. Spt5 Occupancy Is Predictive of AID-Dependent Somatic Mutations

(A) Graphical representation of somatic mutation analysis for Spt5hi and Spt5lo genes from IgκAID and AID$^{-/-}$ splenic B cells (see Figures 5B and 5C). Mutations in the AID$^{-/-}$ control is subtracted in each case (see Figure S5) and mutation frequencies indicated.

(B) Correlation between Spt5 and AID read density per gene. For each gene that recruited above-background amounts of AID and Spt5, the number of sequence tags aligning between −1 Kb upstream of the transcriptional start site to its transcriptional termination site were normalized per gene length (in Kb), per million aligned reads (reads per Kb per million, RPKM) and shown as a hexagonal binning plot. The Spearman's correlation coefficient (ρ) is indicated.

(C) ChIP analysis for AID occupancy at the Gas5 gene in CH12 cells infected with shSpt5-1 or shLacZ control. Data represents a total of 4 experiments using two different anti-AID antibodies. For each experiment, shLacZ was assigned an arbitrary value of 1. The p value is indicated.

See also Figure S5.

1998; Yamada et al., 2006). Phosphorylated Spt5 remains associated with Pol II throughout the elongation phase. Spt5 also engages in interactions with various cotranscriptional factors thereby serving as an adaptor linking these factors to the transcriptional machinery. Spt5 links Pol II to splicing factors (Pei and Shuman, 2002), capping enzyme (Wen and Shatkin, 1999), the exosome complex (Andrulis et al., 2002), transcription coupled repair factors (Ding et al., 2010), NFkB, and E-box proteins (Amir-Zilberstein and Dikstein, 2008). Our data suggest that Spt5 also facilitates the interaction of AID with Pol II (Figure 3E) and thereby targets this enzyme to genomic loci accumulating paused Pol II (Figure 3F, Figure 4D, Figure 6B, and A.Y. and R.C., unpublished data, accession number GSE24178).

Stalled Pol II in the Ig Locus

In activated B cells, the Ig locus is unique in having a large domain of densely packed Spt5 and Pol II molecules extending several kilobases (Figure 5A and Tables S3A and S3B). The idea that Pol II pausing might be linked to mutation (Peters and Storb, 1996) was proposed based on the characteristics of Ig hypermutation, and the position of hypermutation relative to transcriptional start sites (reviewed in Di Noia and Neuberger, 2007; Peled et al., 2008; Stavnezer et al., 2008; Storb et al., 2007). A mutator factor, MuF, was hypothesized to associate with Pol II and generate mutations when Pol II is paused during elongation (Peters and Storb, 1996). More recently, detailed analyses of transcription and Pol II occupancy in the switch regions have confirmed that transcription is indeed impeded throughout the switch regions, most likely due to the presence of G-rich repetitive sequence elements that facilitate DNA distortion and formation of R loops (Daniels and Lieber, 1995; Rajagopal et al., 2009; Ronai et al., 2007; Tian and Alt, 2000; Wang et al., 2009; Yu et al., 2003). Altogether, this makes the Ig locus an ideal

substrate for targeted mutation by AID because: (1) Spt5 facilitates association between AID and Pol II, (2) the stalled Pol II molecules provide an abundance of ssDNA for AID, and (3) the reduced rate of elongation provides AID with increased time of residence at the target.

Finally, in addition to the switch region, several genes mutated by AID were already known to have paused Pol II at sites corresponding to regions that are somatically mutated including c-myc (Bentley and Groudine, 1986; Krumm et al., 1992), Pim1 (Rohwer et al., 1996), and Igκ (Raschke et al., 1999). Our experiments provide a mechanistic explanation for the association between Pol II stalling and AID-mediated somatic mutation. In addition, they reveal the full spectrum of AID targets, including genes such as Gas5, which also undergoes reciprocal translocation in B cell lymphomas (Nakamura et al., 2008).

Concluding Remarks

Although our findings demonstrate a mechanism by which AID gains access to the promoter proximal region of genes, several questions remain about how antibody diversification is mediated. In particular, AID recruitment is only the first of several steps required to bring about CSR and SHM. Following its recruitment to DNA, AID must gain access to target DNA. Although Spt5 acts as an adaptor for AID, localizing it to paused Pol II and associated ssDNA, this may not be sufficient. AID mutates both DNA strands, and paused Pol II exposes only the non-transcribed strand (Giardina et al., 1992; Gilmour, 2009; Lis, 2007; Peterlin and Price, 2006). In addition, the association between AID and paused Pol II does not explain why repair differs between Ig and non-Ig genes, and between different non-Ig AID targets (Liu et al., 2008). Hence, the mechanisms governing post-AID recruitment events required for CSR and SHM remain to be elucidated. AID and Spt5 can interact directly

in vitro but the interaction is weak suggesting that additional factors or posttranslational modifications may be required. Nevertheless, our data suggests that Spt5 links Pol II and AID, thereby providing a mechanistic explanation for the well-established correlation between AID and transcription. The association between Spt5 and AID also explains intrinsic features of hypermutation and CSR, including the enrichment of mutation in the promoter-proximal regions, which correspond to sites of Pol II stalling (Nechaev et al., 2010; Rahl et al., 2010; Zeitlinger et al., 2007).

In conclusion, we propose that AID utilizes the phenomenon of Pol II stalling, which is widespread in the B cell genome, and is particularly prominent on *Ig* loci, to gain access to its target genes across the genome.

EXPERIMENTAL PROCEDURES

Library Preparation
The lentiviral shRNA library (Table S1B) was prepared, titered, arrayed and validated as described (Moffat et al., 2006; Root et al., 2006; http://www.broadinstitute.org/rnai/trc/lib).

Library Screening
The lentiviral library was screened in a 96-well plate format in triplicate, starting from the infection stage through to flow cytometry analysis (schematically represented in Figure 1C). Each plate contained negative control viruses targeting LacZ, GFP and RFP and a positive control shRNA targeting AID. Cells were infected, selected, and stimulated to undergo CSR followed by FACS analysis (details in Supplemental Information).

shRNA Knockdown in Primary B Cells
The hairpin sequences for shSpt5-1 and shSpt5-2 (Table S1B) were cloned into the LMP retroviral vector (Open Biosystems) and transfected into BOSC23 cells to produce retrovirus (Robbiani et al., 2008). Primary B cells stimulated with LPS and IL-4 were cultured as described (Robbiani et al., 2008). After 24 hr in culture, B cells were infected with shRNA-expressing retroviral supernatants as described (Robbiani et al., 2008) and cultured for an additional 3 days with LPS and IL-4, followed by FACS analysis for IgG1 and Spt5 protein analysis by western blotting.

Immunoprecipitation
For Flag-IPs, 2 mg of WCE (prepared as described in Supplemental Information) was incubated with 20 μl Flag Agarose resin (Sigma) for 2 hr at 4°C in IP buffer (identical to WCE preparation buffers above adjusted to 150 mM NaCl for fibroblasts assays, and 200 mM NaCl for B cells and CH12 assays). This was followed by three washes in IP buffer and elution with 0.2 mg/ml Flag peptide (Sigma) for 1 hr at 4°C. Eluates were subjected to SDS-PAGE and western blot analysis. For anti-Spt5 IPs, 2 mgs of WCE were incubated with 3 μg of anti-Spt5 (Santa Cruz Biotechnology) for 2 hr at 4°C followed by capture of the immune complexes with 20 μl Protein A agarose (Roche) for 1 hr at 4°C. Beads were washed three times with IP buffer and bound material was extracted by boiling in 100 μl of Laemmli sample buffer. Eluted material was analyzed by SDS-PAGE and western blot. Antibodies used for probing western blots were as follows: Flag (Sigma), Spt5 (H300) (SantaCruz Biotechnology), Pol II (4H8) (Abcam) and Phospho-Ser PKC Substrate (Cell Signaling).

AID-Spt5 Interaction In Vitro
GST fusion proteins were expressed in *E. coli* and immobilized on Glutathione Sepharose 4 Fast Flow beads (GE Healthcare). Beads were incubated with 500 ng of purified DSIF (Spt5-Spt4) complex (a generous gift from Dr. Sohail Malik, The Rockefeller University) in 200 μl final volume of binding buffer (20 mM Tris [pH 7.5], 150 mM NaCl, 0.1% NP-40, 1 mM EDTA, Protease Inhibitor cocktail (Roche), 0.5 mM PMSF, 1 mM DTT, 0.5 mg/ml BSA) for 2 hr at 4°C with gentle rotation. After four washes with binding buffer, bound proteins were

eluted by boiling in NuPAGE LDS loading buffer (Invitrogen). Samples were then subjected to SDS-PAGE followed by western blot analysis.

Chromatin Immunoprecipitation and Sequencing
ChIP-seq was performed exactly as described (Kuchen et al., 2010). In brief, cells were fixed with 1% paraformaldehyde at 37°C for 10 min followed by sonication. Chromatin fragments were then immunoprecipitated with antibodies specific for Spt5 (Santa Cruz Biotechnology [H300] and BD Biosciences [anti-DSIF]), RNA Pol II (Abcam, [4H8]) or Ser5-phosphorylated RNA Pol II (Abcam, [phospho-S5]). Immunoprecipitates were processed following Illumina's protocol and sequenced on a Genome Analyzer. During analysis, short sequence tags were trimmed to 32 nts and aligned to the mouse genome (NCBI37/mm9) using Bowtie. Uniquely aligned reads were analyzed by SICER (Zang et al., 2009) using an expectation value E of 0.05 in a random background model. The requirement for unique alignment was not applied for IgSμ or IgSγ1 because of their high repetitive nature and low mappability (Figure 5A). Reads on significant islands as defined by SICER were normalized to the total number of reads on islands. Downstream analysis was carried out in R and Python.

Quantitative AID ChIP
CH12 cells were infected with shRNAs to Spt5 as above and subjected to ChIP analysis using two different anti-AID antibodies (Chaudhuri et al., 2004; McBride et al., 2006). Assays were performed as described (Vuong et al., 2009). The ChIP'd material was analyzed by Q-PCR and raw values were normalized to the input signals for each sample (Vuong et al., 2009). Reactions were performed in triplicate. Forward and reverse primers used for Sμ amplification were 5′ TAGTAAGCGAGGCTCTAAAAAGCAT 3′ and 5′ AGAACAGT CCAGTGTAGGCAGTAGA 3' respectively. Forward and reverse primers used for *Gas5* amplification were 5' TATGGCTTCGGGGCCTTGGA 3' and 5' CCTCCTAAAGTTTCCAGCTTGTGC 3' respectively.

Calculation of the Stalling Index
The stalling index was calculated based on Pol II ChIP-seq reads as described (Rahl et al., 2010; Zeitlinger et al., 2007). Briefly, the Pol II and Spt5 stalling indices are calculated in the same way and represent the ratio of read density at the promoter to the average gene body density. The promoter was defined as a 1 kb region extending from −0.5 kb to +0.5 kb relative to the TSS, and the gene body was defined as the region from +1kb downstream of the TSS up to the transcription termination site (TTS) (Rahl et al., 2010; Zeitlinger et al., 2007).

Additional experimental procedures can be found in the Supplemental Information.

SUPPLEMENTAL INFORMATION

Supplemental Information includes Extended Experimental Procedures, five figures, and four tables and can be found with this article online at doi:10.1016/j.cell.2010.09.017.

REFERENCES

Amir-Zilberstein, L., and Dikstein, R. (2008). Interplay between E-box and NF-kappaB in regulation of A20 gene by DRB sensitivity-inducing factor (DSIF). J. Biol. Chem. *283*, 1317–1323.

Andrulis, E.D., Werner, J., Nazarian, A., Erdjument-Bromage, H., Tempst, P., and Lis, J.T. (2002). The RNA processing exosome is linked to elongating RNA polymerase II in Drosophila. Nature *420*, 837–841.

Bai, X., Kim, J., Yang, Z., Jurynec, M., Akie, T., Lee, J., LeBlanc, J., Sessa, A., Jiang, H., DiBiase, A., et al. (2010). TIF1γ Controls Erythroid Cell Fate by Regulating Transcription Elongation. Cell *142*, 133–143.

Basu, U., Chaudhuri, J., Alpert, C., Dutt, S., Ranganath, S., Li, G., Schrum, J.P., Manis, J.P., and Alt, F.W. (2005). The AID antibody diversification enzyme is regulated by protein kinase A phosphorylation. Nature *438*, 508–511.

Bentley, D.L., and Groudine, M. (1986). A block to elongation is largely responsible for decreased transcription of c-myc in differentiated HL60 cells. Nature *321*, 702–706.

Bhutani, N., Brady, J.J., Damian, M., Sacco, A., Corbel, S.Y., and Blau, H.M. (2010). Reprogramming towards pluripotency requires AID-dependent DNA demethylation. Nature *463*, 1042–1047.

Bransteitter, R., Pham, P., Scharff, M.D., and Goodman, M.F. (2003). Activation-induced cytidine deaminase deaminates deoxycytidine on single-stranded DNA but requires the action of RNase. Proc. Natl. Acad. Sci. USA *100*, 4102–4107.

Chaudhuri, J., and Alt, F.W. (2004). Class-switch recombination: interplay of transcription, DNA deamination and DNA repair. Nat. Rev. Immunol. *4*, 541–552.

Chaudhuri, J., Khuong, C., and Alt, F.W. (2004). Replication protein A interacts with AID to promote deamination of somatic hypermutation targets. Nature *430*, 992–998.

Chaudhuri, J., Tian, M., Khuong, C., Chua, K., Pinaud, E., and Alt, F.W. (2003). Transcription-targeted DNA deamination by the AID antibody diversification enzyme. Nature *422*, 726–730.

Conticello, S.G., Ganesh, K., Xue, K., Lu, M., Rada, C., and Neuberger, M.S. (2008). Interaction between antibody-diversification enzyme AID and spliceosome-associated factor CTNNBL1. Mol. Cell *31*, 474–484.

Core, L.J., Waterfall, J.J., and Lis, J.T. (2008). Nascent RNA sequencing reveals widespread pausing and divergent initiation at human promoters. Science *322*, 1845–1848.

Daniel, J.A., Santos, M.A., Wang, Z., Zang, C., Schwab, K.R., Jankovic, M., Filsuf, D., Chen, H., Gazumyan, A., Yamane, A., et al. (2010). PTIP Promotes Chromatin Changes Critical for Immunoglobulin Class Switch Recombination. Science *329*, 917–923.

Daniels, G.A., and Lieber, M.R. (1995). RNA:DNA complex formation upon transcription of immunoglobulin switch regions: implications for the mechanism and regulation of class switch recombination. Nucleic Acids Res. *23*, 5006–5011.

Di Noia, J.M., and Neuberger, M.S. (2007). Molecular mechanisms of antibody somatic hypermutation. Annu. Rev. Biochem. *76*, 1–22.

Dickerson, S.K., Market, E., Besmer, E., and Papavasiliou, F.N. (2003). AID mediates hypermutation by deaminating single stranded DNA. J. Exp. Med. *197*, 1291–1296.

Ding, B., LeJeune, D., and Li, S. (2010). The C-terminal repeat domain of Spt5 plays an important role in suppression of Rad26-independent transcription coupled repair. J. Biol. Chem. *285*, 5317–5326.

Giardina, C., Perez-Riba, M., and Lis, J.T. (1992). Promoter melting and TFIID complexes on Drosophila genes in vivo. Genes Dev. *6*, 2190–2200.

Gilmour, D.S. (2009). Promoter proximal pausing on genes in metazoans. Chromosoma *118*, 1–10.

Guenther, M.G., Levine, S.S., Boyer, L.A., Jaenisch, R., and Young, R.A. (2007). A chromatin landmark and transcription initiation at most promoters in human cells. Cell *130*, 77–88.

Kao, S.Y., Calman, A.F., Luciw, P.A., and Peterlin, B.M. (1987). Anti-termination of transcription within the long terminal repeat of HIV-1 by tat gene product. Nature *330*, 489–493.

Kim, J.B., and Sharp, P.A. (2001). Positive transcription elongation factor B phosphorylates hSPT5 and RNA polymerase II carboxyl-terminal domain independently of cyclin-dependent kinase-activating kinase. J. Biol. Chem. *276*, 12317–12323.

Klein, U., Tu, Y., Stolovitzky, G.A., Keller, J.L., Haddad, J., Jr., Miljkovic, V., Cattoretti, G., Califano, A., and Dalla-Favera, R. (2003). Transcriptional analysis of the B cell germinal center reaction. Proc. Natl. Acad. Sci. USA *100*, 2639–2644.

Krumm, A., Meulia, T., Brunvand, M., and Groudine, M. (1992). The block to transcriptional elongation within the human c-myc gene is determined in the promoter-proximal region. Genes Dev. *6*, 2201–2213.

Kuchen, S., Resch, W., Yamane, A., Kuo, N., Li, Z., Chakraborty, T., Wei, L., Laurrence, A., Yasuda, T., Peng, S., et al. (2010). Regulation of MicroRNA Expression and Abundance during Lymphopoesis. Immunity *32*, 828–839.

Lefebvre, B., Brand, C., Lefebvre, B., and Ozato, K. (2002). Chromosomal Integration of Retinoic Acid Response Elements Prevents Cooperative Transcriptional Activation by Retinoic Acid Receptor and Retinoid X Receptor. Mol. Cell. Biol. *22*, 1446–1459.

Lin, C., Yang, L., Tanasa, B., Hutt, K., Ju, B.G., Ohgi, K., Zhang, J., Rose, D.W., Fu, X.D., Glass, C.K., et al. (2009). Nuclear receptor-induced chromosomal proximity and DNA breaks underlie specific translocations in cancer. Cell *139*, 1069–1083.

Lis, J.T. (2007). Imaging Drosophila gene activation and polymerase pausing in vivo. Nature *450*, 198–202.

Liu, M., Duke, J.L., Richter, D.J., Vinuesa, C.G., Goodnow, C.C., Kleinstein, S.H., and Schatz, D.G. (2008). Two levels of protection for the B cell genome during somatic hypermutation. Nature *451*, 841–845.

Marshall, N.F., Peng, J., Xie, Z., and Price, D.H. (1996). Control of RNA polymerase II elongation potential by a novel carboxyl-terminal domain kinase. J. Biol. Chem. *271*, 27176–27183.

Marshall, N.F., and Price, D.H. (1995). Purification of P-TEFb, a transcription factor required for the transition into productive elongation. J. Biol. Chem. *270*, 12335–12338.

Matsumoto, Y., Marusawa, H., Kinoshita, K., Endo, Y., Kou, T., Morisawa, T., Azuma, T., Okazaki, I.M., Honjo, T., and Chiba, T. (2007). Helicobacter pylori infection triggers aberrant expression of activation-induced cytidine deaminase in gastric epithelium. Nat. Med. *13*, 470–476.

McBride, K.M., Gazumyan, A., Woo, E.M., Barreto, V.M., Robbiani, D.F., Chait, B.T., and Nussenzweig, M.C. (2006). Regulation of hypermutation by activation-induced cytidine deaminase phosphorylation. Proc. Natl. Acad. Sci. USA *103*, 8798–8803.

McBride, K., Gazumyan, A., Woo, E., Schwickert, T., Chait, B., and Nussenzweig, M. (2008). Regulation of class switch recombination and somatic mutation by AID phosphorylation. J. Exp. Med. *205*, 2585–2594.

Moffat, J., Grueneberg, D.A., Yang, X., Kim, S.Y., Kloepfer, A.M., Hinkle, G., Piqani, B., Eisenhaure, T.M., Luo, B., Grenier, J.K., et al. (2006). A lentiviral RNAi library for human and mouse genes applied to an arrayed viral high-content screen. Cell *124*, 1283–1298.

Morgan, H.D., Dean, W., Coker, H.A., Reik, W., and Petersen-Mahrt, S.K. (2004). Activation-induced cytidine deaminase deaminates 5-methylcytosine in DNA and is expressed in pluripotent tissues: implications for epigenetic reprogramming. J. Biol. Chem. *279*, 52353–52360.

Muramatsu, M., Kinoshita, K., Fagarasan, S., Yamada, S., Shinkai, Y., and Honjo, T. (2000). Class switch recombination and hypermutation require

activation-induced cytidine deaminase (AID), a potential RNA editing enzyme. Cell *102*, 553–563.

Muramatsu, M., Sankaranand, V.S., Anant, S., Sugai, M., Kinoshita, K., Davidson, N.O., and Honjo, T. (1999). Specific expression of activation-induced cytidine deaminase (AID), a novel member of the RNA-editing deaminase family in germinal center B cells. J. Biol. Chem. *274*, 18470–18476.

Muse, G.W., Gilchrist, D.A., Nechaev, S., Shah, R., Parker, J.S., Grissom, S.F., Zeitlinger, J., and Adelman, K. (2007). RNA polymerase is poised for activation across the genome. Nat. Genet. *39*, 1507–1511.

Nakamura, M., Kondo, S., Sugai, M., Nazarea, M., Imamura, S., and Honjo, T. (1996). High frequency class switching of an IgM+ B lymphoma clone CH12F3 to IgA+ cells. Int. Immunol. *8*, 193–201.

Nakamura, Y., Takahashi, N., Kakegawa, E., Yoshida, K., Ito, Y., Kayano, H., Niitsu, N., Jinnai, I., and Bessho, M. (2008). The GAS5 (growth arrest-specific transcript 5) gene fuses to BCL6 as a result of t(1;3)(q25;q27) in a patient with B-cell lymphoma. Cancer Genet. Cytogenet. *182*, 144–149.

Nambu, Y., Sugai, M., Gonda, H., Lee, C.G., Katakai, T., Agata, Y., Yokota, Y., and Shimizu, A. (2003). Transcription-coupled events associating with immunoglobulin switch region chromatin. Science *302*, 2137–2140.

Nechaev, S., Fargo, D., Santos, G., Liu, L., Gao, Y., and Adelman, K. (2010). Global Analysis of Short RNAs Reveals Widespread Promoter-Proximal Stalling and Arrest of Pol II in *Drosophila*. Science *327*, 335–338.

Nussenzweig, A., and Nussenzweig, M.C. (2010). Origin of chromosomal translocations in lymphoid cancer. Cell *141*, 27–38.

Pasqualucci, L., Kitaura, Y., Gu, H., and Dalla-Favera, R. (2006). PKA-mediated phosphorylation regulates the function of activation-induced deaminase (AID) in B cells. Proc. Natl. Acad. Sci. USA *103*, 395–400.

Pasqualucci, L., Migliazza, A., Fracchiolla, N., William, C., Neri, A., Baldini, L., Chaganti, R.S., Klein, U., Kuppers, R., Rajewsky, K., et al. (1998). BCL-6 mutations in normal germinal center B cells: evidence of somatic hypermutation acting outside Ig loci. Proc. Natl. Acad. Sci. USA *95*, 11816–11821.

Pasqualucci, L., Neumeister, P., Goossens, T., Nanjangud, G., Chaganti, R.S., Kuppers, R., and Dalla-Favera, R. (2001). Hypermutation of multiple proto-oncogenes in B-cell diffuse large-cell lymphomas. Nature *412*, 341–346.

Pei, Y., and Shuman, S. (2002). Interactions between fission yeast mRNA capping enzymes and elongation factor Spt5. J. Biol. Chem. *277*, 19639–19648.

Peled, J.U., Kuang, F.L., Iglesias-Ussel, M.D., Roa, S., Kalis, S.L., Goodman, M.F., and Scharff, M.D. (2008). The Biochemistry of Somatic Hypermutation. Annu. Rev. Immunol. *26*, 481–511.

Peterlin, B.M., and Price, D.H. (2006). Controlling the elongation phase of transcription with P-TEFb. Mol. Cell *23*, 297–305.

Peters, A., and Storb, U. (1996). Somatic hypermutation of immunoglobulin genes is linked to transcription initiation. Immunity *4*, 57–65.

Pham, P., Bransteitter, R., Petruska, J., and Goodman, M.F. (2003). Processive AID-catalysed cytosine deamination on single-stranded DNA simulates somatic hypermutation. Nature *424*, 103–107.

Popp, C., Dean, W., Feng, S., Cokus, S.J., Andrews, S., Pellegrini, M., Jacobsen, S.E., and Reik, W. (2010). Genome-wide erasure of DNA methylation in mouse primordial germ cells is affected by AID deficiency. Nature *463*, 1101–1105.

Rahl, P.B., Lin, C.Y., Seila, A.C., Flynn, R.A., McCuine, S., Burge, C.B., Sharp, P.A., and Young, R.A. (2010). c-Myc Regulates Transcriptional Pause Release. Cell *141*, 432–445.

Rai, K., Huggins, I.J., James, S.R., Karpf, A.R., Jones, D.A., and Cairns, B.R. (2008). DNA demethylation in zebrafish involves the coupling of a deaminase, a glycosylase, and gadd45. Cell *135*, 1201–1212.

Rajagopal, D., Maul, R.W., Ghosh, A., Chakraborty, T., Khamlichi, A.A., Sen, R., and Gearhart, P.J. (2009). Immunoglobulin switch mu sequence causes RNA polymerase II accumulation and reduces dA hypermutation. J. Exp. Med. *206*, 1237–1244.

Ramiro, A.R., Jankovic, M., Callen, E., Difilippantonio, S., Chen, H.T., McBride, K.M., Eisenreich, T.R., Chen, J., Dickins, R.A., Lowe, S.W., et al. (2006). Role of genomic instability and p53 in AID-induced c-myc-Igh translocations. Nature *440*, 105–109.

Ramiro, A.R., Jankovic, M., Eisenreich, T., Difilippantonio, S., Chen-Kiang, S., Muramatsu, M., Honjo, T., Nussenzweig, A., and Nussenzweig, M.C. (2004). AID is required for c-myc/IgH chromosome translocations in vivo. Cell *118*, 431–438.

Ramiro, A.R., Stavropoulos, P., Jankovic, M., and Nussenzweig, M.C. (2003). Transcription enhances AID-mediated cytidine deamination by exposing single-stranded DNA on the nontemplate strand. Nat. Immunol. *4*, 452–456.

Raschke, E.E., Albert, T., and Eick, D. (1999). Transcriptional regulation of the Ig kappa gene by promoter-proximal pausing of RNA polymerase II. J. Immunol. *163*, 4375–4382.

Revy, P., Muto, T., Levy, Y., Geissmann, F., Plebani, A., Sanal, O., Catalan, N., Forveille, M., Dufourcq-Labelouse, R., Gennery, A., et al. (2000). Activation-induced cytidine deaminase (AID) deficiency causes the autosomal recessive form of the Hyper-IgM syndrome (HIGM2). Cell *102*, 565–575.

Robbiani, D.F., Bothmer, A., Callen, E., Reina-San-Martin, B., Dorsett, Y., Difilippantonio, S., Bolland, D.J., Chen, H.T., Corcoran, A.E., Nussenzweig, A., et al. (2008). AID is required for the chromosomal breaks in c-myc that lead to c-myc/IgH translocations. Cell *135*, 1028–1038.

Robbiani, D.F., Bunting, S., Feldhahn, N., Bothmer, A., Camps, J., Deroubaix, S., McBride, K.M., Klein, I.A., Stone, G., Eisenreich, T.R., et al. (2009). AID produces DNA double-strand breaks in non-Ig genes and mature B cell lymphomas with reciprocal chromosome translocations. Mol. Cell *36*, 631–641.

Rogozin, I.B., and Kolchanov, N.A. (1992). Somatic hypermutagenesis in immunoglobulin genes. II. Influence of neighbouring base sequences on mutagenesis. Biochim. Biophys. Acta *1171*, 11–18.

Rohwer, F., Todd, S., and McGuire, K.L. (1996). The effect of IL-2 treatment on transcriptional attenuation in proto-oncogenes pim-1 and c-myb in human thymic blast cells. J. Immunol. *157*, 643–649.

Ronai, D., Iglesias-Ussel, M.D., Fan, M., Li, Z., Martin, A., and Scharff, M.D. (2007). Detection of chromatin-associated single-stranded DNA in regions targeted for somatic hypermutation. J. Exp. Med. *204*, 181–190.

Root, D.E., Hacohen, N., Hahn, W.C., Lander, E.S., and Sabatini, D.M. (2006). Genome-scale loss-of-function screening with a lentiviral RNAi library. Nat. Methods *3*, 715–719.

Shen, H.M., Peters, A., Baron, B., Zhu, X., and Storb, U. (1998). Mutation of BCL-6 gene in normal B cells by the process of somatic hypermutation of Ig genes. Science *280*, 1750–1752.

Shen, H.M., Poirier, M.G., Allen, M.J., North, J., Lal, R., Widom, J., and Storb, U. (2009). The activation-induced cytidine deaminase (AID) efficiently targets DNA in nucleosomes but only during transcription. J. Exp. Med. *206*, 1057–1071.

Sohail, A., Klapacz, J., Samaranayake, M., Ullah, A., and Bhagwat, A.S. (2003). Human activation-induced cytidine deaminase causes transcription-dependent, strand-biased C to U deaminations. Nucleic Acids Res. *31*, 2990–2994.

Stavnezer-Nordgren, J., and Sirlin, S. (1986). Specificity of immunoglobulin heavy chain switch correlates with activity of germline heavy chain genes prior to switching. EMBO J. *5*, 95–102.

Stavnezer, J., Guikema, J.E., and Schrader, C.E. (2008). Mechanism and regulation of class switch recombination. Annu. Rev. Immunol. *26*, 261–292.

Storb, U., Shen, H.M., Longerich, S., Ratnam, S., Tanaka, A., Bozek, G., and Pylawka, S. (2007). Targeting of AID to immunoglobulin genes. Adv. Exp. Med. Biol. *596*, 83–91.

Swanson, M.S., Malone, E.A., and Winston, F. (1991). SPT5, an essential gene important for normal transcription in Saccharomyces cerevisiae, encodes an acidic nuclear protein with a carboxy-terminal repeat. Mol. Cell. Biol. *11*, 4286.

Tian, M., and Alt, F.W. (2000). Transcription-induced cleavage of immunoglobulin switch regions by nucleotide excision repair nucleases in vitro. J. Biol. Chem. *275*, 24163–24172.

Vuong, B., Lee, M., Kabir, S., Irimia, C., Macchiarulo, S., McKnight, G., and Chaudhuri, J. (2009). Specific recruitment of protein kinase A to the immunoglobulin locus regulates class-switch recombination. Nat. Immunol. *10*, 420–426.

Wada, T., Takagi, T., Yamaguchi, Y., Ferdous, A., Imai, T., Hirose, S., Sugimoto, S., Yano, K., Hartzog, G.A., Winston, F., et al. (1998). DSIF, a novel transcription elongation factor that regulates RNA polymerase II processivity, is composed of human Spt4 and Spt5 homologs. Genes Dev. *12*, 343–356.

Wang, L., Wuerffel, R., Feldman, S., Khamlichi, A.A., and Kenter, A.L. (2009). S region sequence, RNA polymerase II, and histone modifications create chromatin accessibility during class switch recombination. J. Exp. Med. *206*, 1817–1830.

Wang, M., Rada, C., and Neuberger, M.S. (2010). Altering the spectrum of immunoglobulin V gene somatic hypermutation by modifying the active site of AID. J Exp Med *207*, 141–153.

Wen, Y., and Shatkin, A.J. (1999). Transcription elongation factor hSPT5 stimulates mRNA capping. Genes Dev. *13*, 1774–1779.

Yamada, T., Yamaguchi, Y., Inukai, N., Okamoto, S., Mura, T., and Handa, H. (2006). P-TEFb-mediated phosphorylation of hSpt5 C-terminal repeats is critical for processive transcription elongation. Mol. Cell *21*, 227–237.

Yamaguchi, Y., Takagi, T., Wada, T., Yano, K., Furuya, A., Sugimoto, S., Hasegawa, J., and Handa, H. (1999a). NELF, a multisubunit complex containing RD, cooperates with DSIF to repress RNA polymerase II elongation. Cell *97*, 41–51.

Yamaguchi, Y., Wada, T., Watanabe, D., Takagi, T., Hasegawa, J., and Handa, H. (1999b). Structure and function of the human transcription elongation factor DSIF. J. Biol. Chem. *274*, 8085–8092.

Yancopoulos, G.D., DePinho, R.A., Zimmerman, K.A., Lutzker, S.G., Rosenberg, N., and Alt, F.W. (1986). Secondary genomic rearrangement events in pre-B cells: VHDJH replacement by a LINE-1 sequence and directed class switching. EMBO J. *5*, 3259–3266.

Yu, K., Chedin, F., Hsieh, C.L., Wilson, T.E., and Lieber, M.R. (2003). R-loops at immunoglobulin class switch regions in the chromosomes of stimulated B cells. Nat. Immunol. *4*, 442–451.

Zang, C., Schones, D.E., Zeng, C., Cui, K., Zhao, K., and Peng, W. (2009). A clustering approach for identification of enriched domains from histone modification ChIP-Seq data. Bioinformatics *25*, 1952–1958.

Zeitlinger, J., Stark, A., Kellis, M., Hong, J.W., Nechaev, S., Adelman, K., Levine, M., and Young, R.A. (2007). RNA polymerase stalling at developmental control genes in the Drosophila melanogaster embryo. Nat. Genet. *39*, 1512–1516.