Genome-wide Translocation Sequencing Reveals Mechanisms of Chromosome Breaks and Rearrangements in B Cells

Roberto Chiarle,[1,2,7] Yu Zhang,[1,7,*] Richard L. Frock,[1,7] Susanna M. Lewis,[1,7] Benoit Molinie,[3] Yu-Jui Ho,[1] Darienne R. Myers,[1] Vivian W. Choi,[1] Mara Compagno,[1,2] Daniel J. Malkin,[1] Donna Neuberg,[4] Stefano Monti,[5,6] Cosmas C. Giallourakis,[3,*] Monica Gostissa,[1,*] and Frederick W. Alt[1,*]

[1]Howard Hughes Medical Institute, Immune Disease Institute, Program in Cellular and Molecular Medicine, Children's Hospital Boston and Departments of Genetics and Pediatrics, Harvard Medical School, Boston, MA 02115, USA
[2]Department of Biomedical Sciences and Human Oncology and CERMS, University of Torino, 10126 Turin, Italy
[3]Gastrointestinal Unit, Center for Study of Inflammatory Bowel Disease, Massachusetts General Hospital, Boston, MA 02114, USA
[4]Dana-Farber Cancer Institute, Harvard Medical School, Boston, MA 02215, USA
[5]Broad Institute, 5 Cambridge Center, Cambridge, MA 02142, USA
[6]Section of Computational Biomedicine, Boston University School of Medicine, Boston, MA 02118, USA
[7]These authors contributed equally to this work
*Correspondence: zhang@idi.harvard.edu (Y.Z.), cgiallourakis@partners.org (C.C.G.), gostissa@idi.harvard.edu (M.G.), alt@enders.tch.harvard.edu (F.W.A.)
DOI 10.1016/j.cell.2011.07.049

SUMMARY

Whereas chromosomal translocations are common pathogenetic events in cancer, mechanisms that promote them are poorly understood. To elucidate translocation mechanisms in mammalian cells, we developed high-throughput, genome-wide translocation sequencing (HTGTS). We employed HTGTS to identify tens of thousands of independent translocation junctions involving fixed I-SceI meganuclease-generated DNA double-strand breaks (DSBs) within the *c-myc* oncogene or *IgH* locus of B lymphocytes induced for activation-induced cytidine deaminase (AID)-dependent *IgH* class switching. DSBs translocated widely across the genome but were preferentially targeted to transcribed chromosomal regions. Additionally, numerous AID-dependent and AID-independent hot spots were targeted, with the latter comprising mainly cryptic I-SceI targets. Comparison of translocation junctions with genome-wide nuclear run-ons revealed a marked association between transcription start sites and translocation targeting. The majority of translocation junctions were formed via end-joining with short microhomologies. Our findings have implications for diverse fields, including gene therapy and cancer genomics.

INTRODUCTION

Recurrent oncogenic translocations are common in hematopoietic malignancies including lymphomas (Küppers and Dalla-Favera, 2001) and also occur frequently in solid tumors such as prostate and lung cancers (Shaffer and Pandolfi, 2006). DNA double-strand breaks (DSBs) are common intermediates of these genomic aberrations (Stratton et al., 2009). DSBs are generated by normal metabolic processes, by genotoxic agents including some cancer therapeutics, and by V(D)J and immunoglobulin (Ig) heavy (H) chain (IgH) class switch recombination (CSR) in lymphocytes (Zhang et al., 2010). Highly conserved pathways repair DSBs to preserve genome integrity (Lieber, 2010). Nevertheless, repair can fail, resulting in unresolved DSBs and translocations. Recurrent translocations in tumors usually arise as low-frequency events that are selected during oncogenesis. However, other factors influence the appearance of recurrent translocations including chromosomal location of oncogenes (Gostissa et al., 2009). Chromosomal environment likely affects translocation frequency by influencing mechanistic factors, including DSB frequency at translocation targets, factors that contribute to juxtaposition of broken loci for joining, and mechanisms that circumvent repair functions that promote intrachromosomal DSB joining (Zhang et al., 2010).

IgH CSR is initiated by DSBs that result from transcription-targeted AID-cytidine deamination activity within *IgH* switch (S) regions that lie just 5′ of various sets of C_H exons. DSBs within the donor S_μ region and a downstream acceptor S region are fused via end-joining to complete CSR and allow expression of a different antibody class (Chaudhuri et al., 2007). Clonal translocations in human and mouse B cell lymphomas often involve *IgH* S regions and an oncogene, such as *c-myc* (Küppers and Dalla-Favera, 2001; Gostissa et al., 2011). In this regard, AID-generated *IgH* S region DSBs directly participate in translocations to *c-myc* and other genes (Franco et al., 2006; Ramiro et al., 2006; Wang et al., 2009). Through its role in somatic hypermutation (SHM) of *IgH* and Ig light (*IgL*) variable region exons, AID theoretically might generate lower frequency DSBs in Ig loci that serve as translocation intermediates (Liu and Schatz, 2009). In addition, AID mutates many non-Ig genes in activated

B cells at far lower levels than Ig genes (Liu et al., 2008); such off-target AID activity also may contribute to translocations of non-Ig genes (Robbiani et al., 2008). Indeed, AID even has been suggested to initiate lesions leading to translocations in nonlymphoid cancers, including prostate cancer (Lin et al., 2009). However, potential roles of AID in generating DSBs genome-wide have not been addressed. In this regard, other sources of translocation-initiating DSBs could include intrinsic factors, such as oxidative metabolism, replication stress, and chromosome fragile sites, or extrinsic factors such as ionizing radiation or chemotherapeutics (Zhang et al., 2010).

DSBs lead to damage response foci formation over 100 kb or larger flanking regions, promoting DSB joining and suppressing translocations (Zhang et al., 2010; Nussenzweig and Nussenzweig, 2010). IgH class switching in activated B cells can be mediated by yeast I-SceI endonuclease-generated DSBs without AID or S regions, suggesting general mechanisms promote efficient intrachromosomal DSB joining over at least 100 kb (Zarrin et al., 2007). In somatic cells, classical nonhomologous end-joining (C-NHEJ) repairs many DSBs (Zhang et al., 2010). C-NHEJ suppresses translocations by preferentially joining DSBs intrachromosomally (Ferguson et al., 2000). Deficiency for C-NHEJ leads to frequent translocations, demonstrating that other pathways fuse DSBs into translocations (Zhang et al., 2010). Correspondingly, an alternative end-joining pathway (A-EJ), that prefers ends with short microhomologies (MHs), supports CSR in the absence of C-NHEJ (Yan et al., 2007) and joins CSR DSBs to other DSBs to generate translocations (Zhang et al., 2010). Indeed, C-NHEJ suppresses p53-deficient lymphomas with recurrent IgH/c-myc translocations catalyzed by A-EJ (Zhu et al., 2002). Various evidence suggests that A-EJ may be translocation prone (e.g., Simsek and Jasin, 2010).

The mammalian nucleus is occupied by nonrandomly positioned genes and chromosomes (Meaburn et al., 2007). Fusion of DSBs to generate translocations requires physical proximity; thus, spatial disposition of chromosomes might impact translocation patterns (Zhang et al., 2010). Cytogenetic studies revealed that certain loci involved in oncogenic translocations are spatially proximal (Meaburn et al., 2007). Studies of recurrent translocations in mouse B cell lymphomas suggested that aspects of particular chromosomal regions, as opposed to broader territories, might promote proximity and influence translocation frequency (Wang et al., 2009). Nonrandom position of genes and chromosomes in the nucleus led to two general models for translocation initiation. "Contact-first" poses translocations to be restricted to proximally positioned chromosomal regions, whereas "breakage-first" poses that distant DSBs can be juxtaposed (Meaburn et al., 2007). In-depth evaluation of how chromosomal organization influences translocations requires a genome-wide approach.

To elucidate translocation mechanisms, we developed approaches that identify genome-wide translocations arising from a specific DSB in vivo. Our studies isolate large numbers of translocations from primary B cells, which were activated for CSR, and provide a comprehensive analysis of the relationships among particular classes of DSBs, transcription, chromosome domains, and translocation events.

RESULTS

Development of High-Throughput Genomic Translocation Sequencing

We developed high-throughput genomic translocation (HTGTS) to isolate junctions between a chromosomal DSB introduced at a fixed site and other sequences genome-wide. Such junctions, other than those involving breaksite resection, mostly should result from end-joining of introduced DSBs to other genomic DSBs. Thus, HTGTS will identify other genomic DSBs capable of joining to the test DSBs. With HTGTS, we isolated from primary mouse B cells junctions that fused IgH or c-myc DSBs to sequences distributed widely across the genome (Figures 1A and 1B). We chose c-myc and IgH as targets because they participate in recurrent oncogenic translocations in B cell lymphomas. To generate c-myc- or IgH-specific DSBs, we employed an 18 bp canonical I-SceI meganuclease target sequence, which is absent in mouse genomes (Jasin, 1996). One c-myc target was a cassette with 25 tandem I-SceI sites, to increase cutting efficiency, within c-myc intron 1 on chromosome (chr)15 (termed c-myc$^{25xI-SceI}$; Figure 1C; Wang et al., 2009). For comparison, we employed an allele with a single I-SceI site in the same position (termed c-myc$^{1xI-SceI}$) (Figure 1C; see Figures S1A–S1C available online). For IgH, we employed an allele with two I-SceI sites in place of endogenous Sγ1 (termed ΔSγ1$^{2xI-SceI}$) on chr12 (Zarrin et al., 2007). As a cellular model, we used primary splenic B cells activated in culture with αCD40 plus IL4 to induce AID, transcription, DSBs and CSR at Sγ1 (IgG1) and Sε (IgE), during days 2–4 of activation. At 24 hr, we infected B cells with I-SceI-expressing retrovirus to induce DSBs at I-SceI targets (Zarrin et al., 2007). Cells were processed at day 4 to minimize doublings and potential cellular selection. As high-titer retroviral infection can impair C-NHEJ (Wang et al., 2009), we also assayed B cells that express from their Rosa26 locus an I-SceI-glucocorticoid receptor fusion protein (I-SceI-GR) that can be activated via triamcinolone acetonide (TA) (Figure 1D; Figures S1D–S1F). The c-myc$^{25xI-SceI}$ cassette was frequently cut in TA-treated c-myc$^{25xI-SceI}$/ROSA$^{I-SceI-GR}$ B cells (Figure S1G).

We employed two HTGTS methods. For the adaptor-PCR approach (Figure 1E, Siebert et al., 1995), genomic DNA was fragmented with a frequently cutting restriction enzyme, ligated to an asymmetric adaptor, and further digested to block amplification of germline or unrearranged target alleles. We then performed nested PCR with adaptor- and locus-specific primers. Depending on the locus-specific PCR primers, one or the other side of the I-SceI DSB provides the "bait" translocation partner (Figure 1C), with the "prey" provided by DSBs at other genomic sites. As a second approach, we employed circularization PCR (Figure 1E) (Mahowald et al., 2009), in which enzymatically fragmented DNA was intramolecularly ligated and digested with blocking enzymes and nested PCR was performed with locus-specific primers. Following sequencing of PCR products, we aligned HTGTS junctions to reference genomes and scripted filters to remove artifacts from aligned databases. We experimentally controlled for potential background by generating HTGTS libraries from mixtures of human DNA and mouse DNA from activated I-SceI-infected c-myc$^{25xI-SceI}$ or ΔSγ1$^{2xI-SceI}$ B
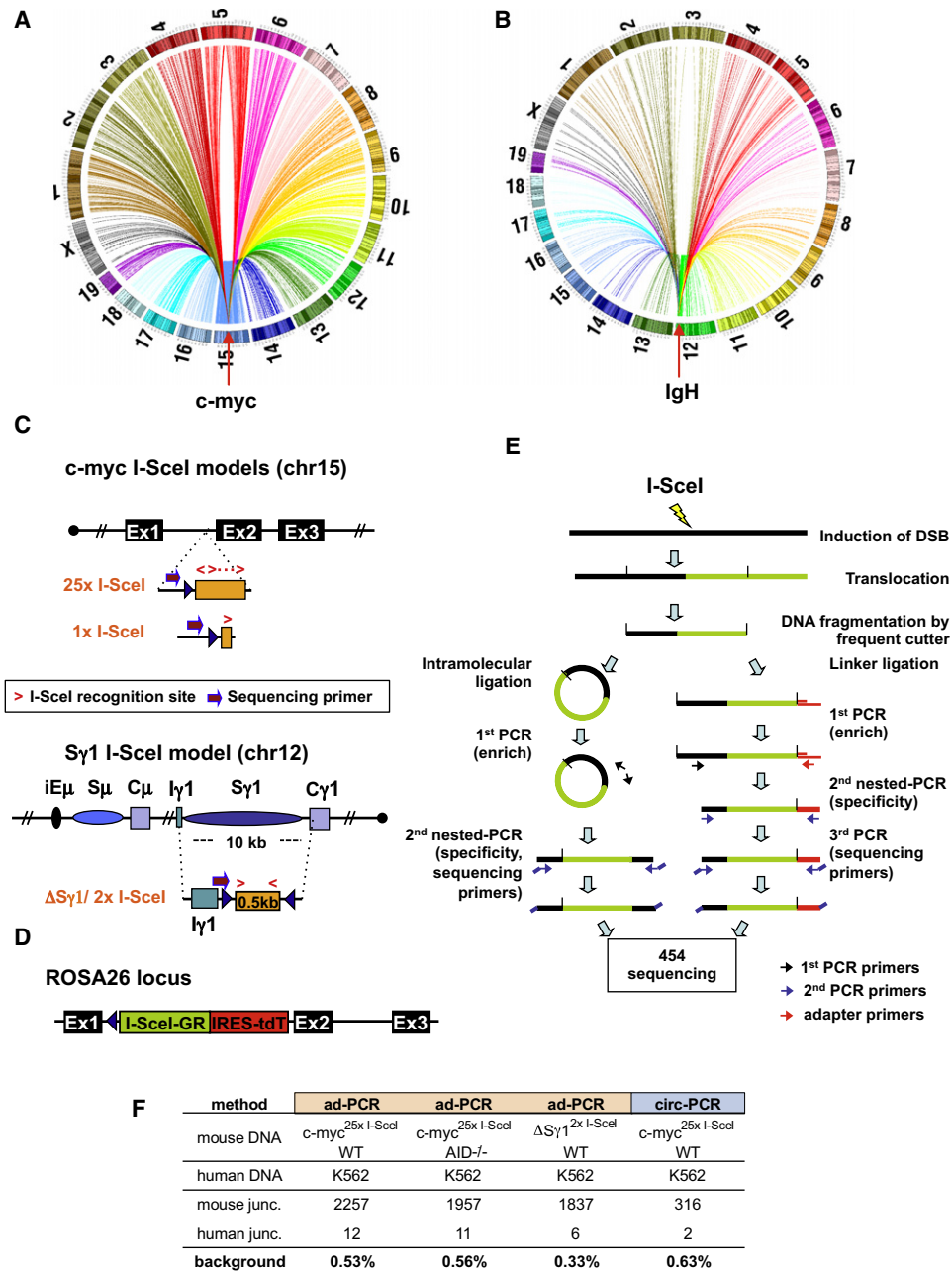
Figure 1. High-Throughput Genomic Translocation Sequencing

(A and B) Circos plots of genome-wide translocation landscape of representative *c-myc* (A) or *IgH* (B) HTGTS libraries. Chromosome ideograms comprise the circumference. Individual translocations are represented as arcs originating from specific I-SceI breaks and terminating at partner site.

(C) Top: a cassette containing either 25 or one I-SceI target(s) was inserted into intron 1 of *c-myc* (see Figures S1A–S1C). Bottom: a cassette composed of a 0.5 kb spacer flanked by I-SceI target replaced the *IgH* Sγ1 region. Relative orientation of I-SceI sites is indicated by red arrows. Position of primers for generation and sequencing HTGTS libraries is shown.

(D) An expression cassette for I-SceI fused to a glucocorticoid receptor (I-SceI-GR) was targeted into *Rosa26* (see Figures S1D–S1G). The red fluorescent protein Tomato (tdT) is coexpressed via an IRES.

(E) Schematic representation of HTGTS methods; left: circularization-PCR, right: adaptor-PCR. See text for details.

(F) Background for HTGTS approaches, calculated as percent of artifactual human:mouse hybrid junctions when human DNA was mixed 1:1 with mouse DNA from indicated samples.

cells; junctions fusing mouse and human sequences were less than 1% of the total (Figure 1F). We identified nearly 150,000 independent junctions from numerous libraries from different mice (Table S1). Resulting genome-wide junction maps are shown either as colored dot plots of overall distribution of translocation numbers in selected size bins (useful for visualizing hot spots) or bar plots that compress hot spots and illustrate translocation site density. HTGTS yields an average of 1 unique junction/5 ng of DNA, corresponding to about one junction/1000 genomes. Major findings were reproduced with both HTGTS methods (e.g., Figure S2A). Moreover, while the largest portion of data was obtained with c-myc$^{25xI\text{-}SceI}$ alleles cut via retroviral I-SceI, major findings were reproduced via HTGTS from the c-myc$^{25xI\text{-}SceI}$ allele cleaved by I-SceI-GR and the c-myc$^{1xI\text{-}SceI}$ allele cleaved by retroviral I-SceI (Figures S2C and S2D).

Analysis of Genome-wide Translocations from *c-myc* DSBs

For HTGTS of c-myc$^{25xI\text{-}SceI}$ or c-myc$^{1xI\text{-}SceI}$ alleles, we used primers about 200 bp centromeric to the cassette (Figure 1C) to detect junctions involving broken ends (BEs) on the 5′ side of *c-myc* I-SceI DSBs ("5′ *c-myc*-I-SceI BEs"). Based on convention, prey sequences joined to 5′ c-myc-I-SceI BEs are in (+) orientation if read from the junction in centromere to telomere direction and in (−) orientation if read in the opposite direction (Figures S3A–S3D). Joins in which 5′ c-myc-I-SceI BEs are fused to resected 3′ c-myc-I-SceI BEs would be (+) (Figure S3A). Intrachromosomal joins to DSBs centromeric or telomeric to 5′ c-myc-I-SceI BEs would be (+) or (−) depending on the side of the second DSB to which they were joined, with potential outcomes including deletions, inversions, and extrachromosomal circles (Figures S3B and S3C). Junctions to DSBs on different chromosomes could be (+) or (−) and derivative chromosomes centric or dicentric (Figure S3D). Analyses of over 100,000 independent junctions from 5′ c-myc-I-SceI BEs from WT and AID$^{-/-}$ backgrounds revealed prey to be distributed widely throughout the genome with similar general distribution patterns (Figure 2; Figures S2B, S2E, and S2F). Other than 200 kb downstream of the bait DSB, intrachromosomal and interchromosomal junctions were evenly distributed into (+) and (−) orientation (Figure 2; Figure S3I). This finding implies that extrachromosomal circles and acentric fragments are represented similarly to other translocation classes, suggesting little impact of cellular selection on junction distribution. The junctions of 5′ c-myc-I-SceI BE from c-myc$^{25xI\text{-}SceI}$, c-myc$^{1xI\text{-}SceI}$, and c-myc$^{25xI\text{-}SceI}$/ROSA$^{I\text{-}SceI\text{-}GR}$ models were all consistent with end-joining, and most (75%–90%) had short junctional MHs (Table S1).

WT and AID$^{-/-}$ HTGTS maps for 5′ c-myc-I-SceI BEs had other common features. First, the majority of junctions (75%) arose from joining 5′ c-myc-I-SceI BEs to sequences within 10 kb, with most lying 3′ of the breaksite (Figure 3A; Figure S4A). The density of joins remained relatively high within a region 200 kb telomeric to the breaksite (Figure 3A; Figure S4A). Notably, most junctions within this 200 kb region, but not beyond, were in the (+) orientation, consistent with joining to resected 3′ c-myc-I-SceI BEs (Figure 3A; Figure S4A). About 15% of junctions occurred within the region 100 kb centromeric to the breaksite. As these could not have resulted from resection (due

to primer removal), they may reflect the known propensity for joining intrachromosomal DSBs separated at such distances (Zarrin et al., 2007). Compared with other chromosomes, chr15 had a markedly high density of translocations along its 50 Mb telomeric portion and also a high density along its centromeric portion (Figure 2). Many chromosomes had smaller regions of relatively high or low translocation density, with such overall patterns conserved between WT and AID$^{-/-}$ backgrounds (Figure 2; Figures S2A–S2F). Finally, although the majority of hot spots were WT specific, a number were shared between WT and AID$^{-/-}$ backgrounds (see below).

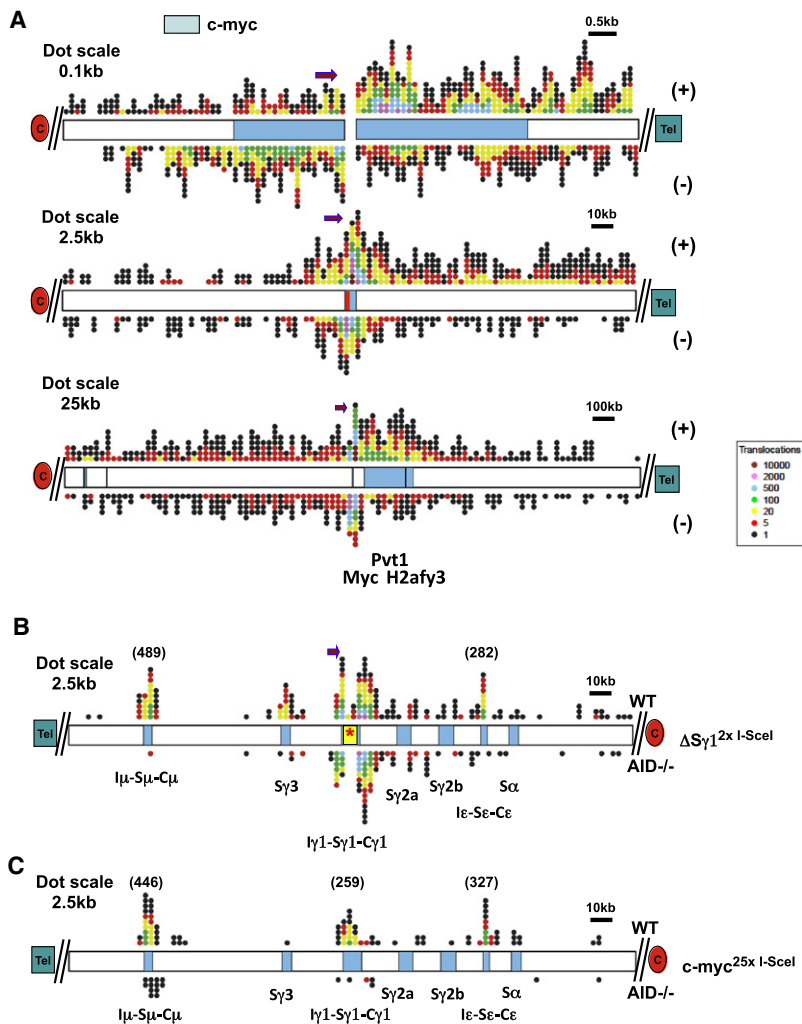### Analysis of HTGTS Libraries from *IgH* DSBs

For HTGTS of the ΔSγ1$^{2xI\text{-}SceI}$ alleles, we used primers about 200 bp telomeric to the I-SceI cassette (Figure 1C), allowing detection of junctions involving BEs on the 5′ side of Sγ1 I-SceI DSBs ("5′ Sγ1-I-SceI BEs"). Intra- and interchromosomal joins involving 5′ Sγ1-I-SceI BEs result in (+) or (−) junctions with the range of potential chromosomal outcomes including deletions, inversions, extrachromosomal circles, and acentrics (Figures S3E–S3H). We isolated and analyzed approximately 9000 and 8000 5′ Sγ1-I-SceI BE junctions from WT and AID$^{-/-}$ libraries, respectively (Figures S2G and S2H). Reminiscent of the 5′ c-myc-I-SceI junctions, about 75% of these junctions were within 10 kb of the breaksite, with a larger proportion on the 3′ side and predominantly in the (−) orientation, consistent with joining to resected 3′ Sγ1-I-SceI BEs (Figures S4B–S4D). Outside the breaksite region, the general 5′ Sγ1-I-SceI BE translocation patterns resembled those observed for 5′ c-myc-I-SceI BEs, with both (+) and (−) translocations occurring on all chromosomes (Figures S3J and S2G). Though we analyzed more limited numbers of 5′ Sγ1-I-SceI BE junctions (Table S2 and Figures S2G and S2H), the broader telomeric region of chr12 had a notably large number of hits, and within this region, there were *IgH* hot spots in WT, but not AID$^{-/-}$, libraries (Figure 3B).

Sμ and Sε are major targets of AID-initiated DSBs in B cells activated with αCD40/IL4. Correspondingly, substantial numbers of 5′ Sγ1-I-SceI BE junctions from WT, but not AID$^{-/-}$, B cells joined to either Sμ or to Sε, which, respectively, lie approximately 100 kb upstream and downstream of the ΔSγ1$^{2xI\text{-}SceI}$ cassette (Figure 3B; Figures S4B–S4D). These findings support the notion that DSBs separated by 100–200 kb can be joined at high frequency by general repair mechanisms (Zarrin et al., 2007). We also observed frequent junctions from WT libraries specifically within Sγ3, which lies about 20 kb upstream of the breaksite, a finding of interest as joining Sγ3 to donor Sμ DSBs during CSR in αCD40/IL4-activated B cells occurs at low levels (see below). Notably, in WT, but not in AID$^{-/-}$ libraries, we found numerous junctions within Sγ1 (Figure S4D), which is also targeted by AID in αCD40/IL4-activated B cells. As Sγ1 is present only on the non-targeted chr12 homolog due to the ΔSγ1$^{2xI\text{-}SceI}$ replacement, these findings demonstrate robust translocation of 5′ Sγ1-I-SceI BEs to AID-dependent Sγ1 DSBs on the homologous chromosome, consistent with trans-CSR (Reynaud et al., 2005). Finally, while AID deficiency greatly reduced junctions into S regions, we observed a focal cluster of five 5′ Sγ1-I-SceI BE junctions in or near Sμ in AID$^{-/-}$ ΔSγ1$^{2xI\text{-}SceI}$ libraries (Figure 3B; Figure S4C).
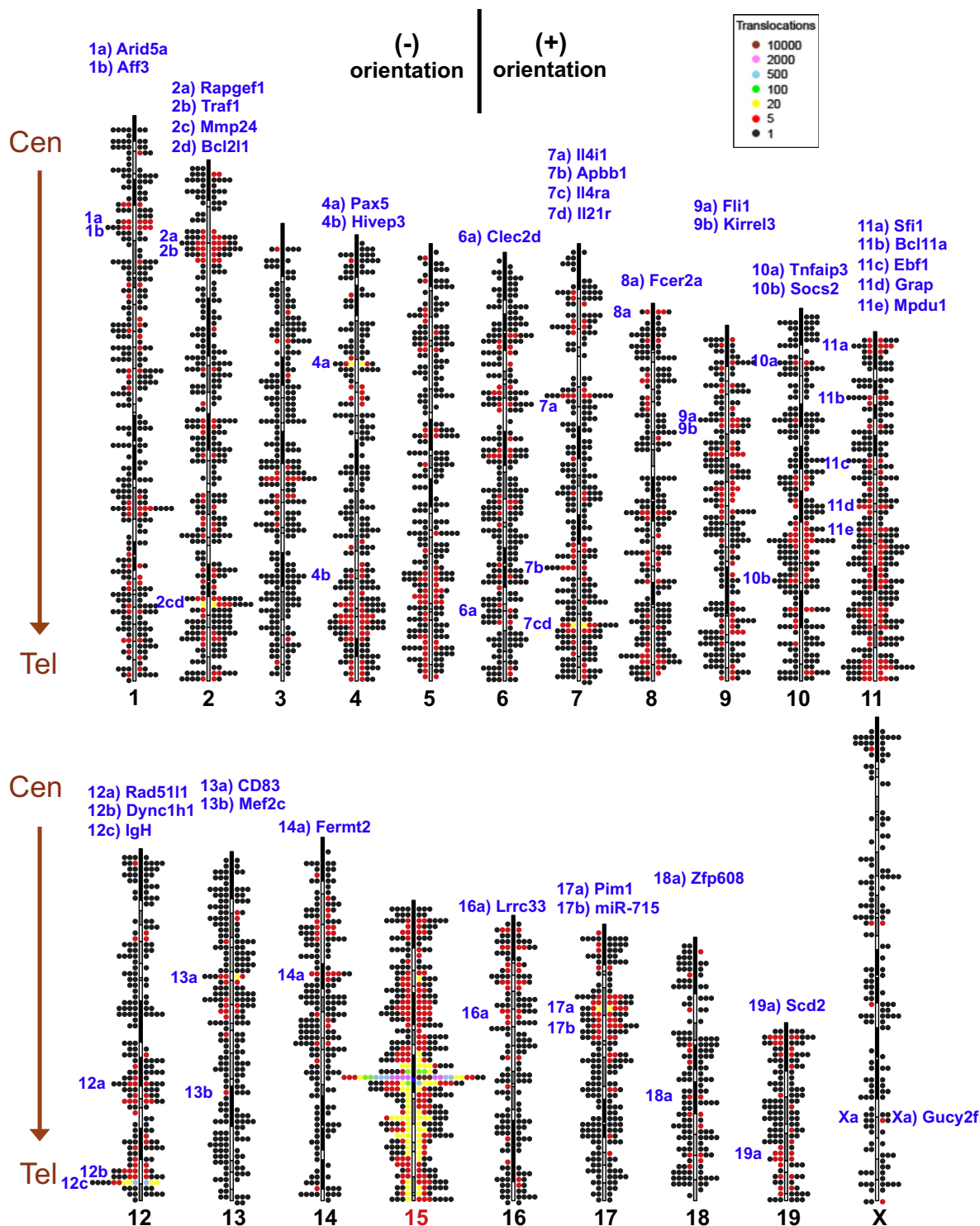
Figure 2. Genome-wide Distribution and Orientation of Translocations from *c-myc* DSBs

Genome-wide map of translocations originating from the c-*myc*[25xI-SceI] cassette (chr15) in αCD40/IL4-activated and I-SceI-infected B cells. Single junctions are represented by dots located at corresponding chromosomal position. The dot scale is 2 Mb. Clusters of translocations are indicated with color codes, as shown in legend. (+) and (−) orientation junctions (see Figure S3) are plotted on right and left side of each ideogram, respectively. Hot spots (see Figure 4A) are listed in blue on top, with notation on the left side of chromosomes to indicate position. Data are from HTGTS libraries from seven different mice. Centromere (Cen) and telomere (Tel) positions are indicated. See also Figure S2.

A

c-myc

Dot scale
0.1kb

0.5kb

(+)

(−)

Dot scale
2.5kb

10kb

(+)

(−)

Dot scale
25kb

100kb

(+)

(−)

Pvt1
Myc H2afy3

Translocations
● 10000
● 2000
● 500
● 100
● 20
● 5
● 1

B

Dot scale
2.5kb

(489)

(282)

10kb

WT

Iμ-Sμ-Cμ

Sγ3

Sγ2a

Sγ2b

Sα

Iε-Sε-Cε

Iγ1-Sγ1-Cγ1

ΔSγ1²ˣ ᴵ⁻ˢᶜᵉᴵ

AID-/-

C

Dot scale
2.5kb

(446)

(259)

(327)

10kb

WT

Iμ-Sμ-Cμ

Sγ3

Sγ2a

Sγ2b

Sα

Iγ1-Sγ1-Cγ1

Iε-Sε-Cε

c-myc²⁵ˣ ᴵ⁻ˢᶜᵉᴵ

AID-/-

Figure 3. Distribution of *IgH*- and *c-myc*-Proximal Junctions

(A) Distribution of junctions around chr15 breaksite in the pooled c-*myc*[25xI-SceI] HTGTS library. Top: 10 kb around breaksite (represented as a split). Middle: 250 kb around breaksite (represented by red bar); Bottom: 2.5 Mb around breaksite. (+) and (−)-oriented junctions are plotted on top and bottom of chromosome diagrams, respectively.

(B and C) Distribution of translocation junctions at *IgH* in the pooled ΔSγ1[2xI-SceI] (B) or c-*myc*[25xI-SceI] (C) HTGTS libraries. Translocations in WT (top) and *AID*[−/−] (bottom) B cells are shown. Positions of S regions within the 250 kb *IgH* C_H region are indicated. Color codes are as in Figure 2. Dot size, position of centromere (red oval) and telomere (green rectangle), and orientation of the sequencing primer are indicated. See also Figure S4.

Most *c-myc* Translocation Hot Spots Are Targeted by AID

To identify 5′ c-*myc*-I-SceI BE translocation hot spots in an unbiased manner, we separated the genome into 250 kb bins and identified bins containing a statistically significant enrichment of translocations (Extended Experimental Procedures). This approach identified 55 hot spots in WT libraries and 15 in *AID*[−/−] libraries (Table S3; Figure 4A). Among the 43 most significant hot spots, 39 were in genes and 4 were in intergenic regions. Of these 43 hot spots, 21 were present at significantly greater levels in WT versus *AID*[−/−] backgrounds, and, therefore, classified as AID dependent; while 9 more were enriched (from 3- to 6-fold) in the WT background and were potentially AID dependent (Table S3; Figure 4A). The other 13 were equally represented between WT and *AID*[−/−] backgrounds (Table S3; Figure 4A). Of these 13, two exist in multiple copies (*Sfi1 and miR-715*), which may have contributed to their classification as hot spots (Quinlan et al., 2010; Ira Hall, personal communication); five reached hot spot significance in only one of the two backgrounds (Table S3; Figure 4A).

The Sμ, Sγ1, and Sε regions, which are targeted for CSR DSBs by αCD40/IL4 treatment, were by far the strongest AID hot spots for 5′ c-*myc*-I-SceI BEs, with other non-*IgH* AID-dependent hot spots ranging from 1% to 10% of Sμ levels (Figure 4A). Translocation specificity to these three S regions, which together comprise less than 20 kb, was striking; there were only a few junctions in the remainder of the C_H locus, which includes 4 other S regions not substantially activated by αCD40/IL4 (Figure 3C). Notably, there was only one 5′ c-*myc*-I-SceI BE junction with Sγ3, even though Sγ3 was a marked hot spot for 5′ Sγ1-I-SceI BEs. In this regard, while AID-dependent DSBs in Sγ3 likely are much less frequent than in Sμ, Sγ1, and Sε under αCD40/IL4 stimulation conditions, Sγ3 DSBs may be favored targets of 5′ Sγ1-I-SceI BEs because of linear proximity. Finally, translocations occurred in Sμ and Sγ1 in *AID*[−/−] B cells at much lower levels than in WT, but frequently enough to qualify them as AID-independent hot spots (Figure 4A).

Several top AID SHM or binding targets in activated B cells (Liu et al., 2008; Yamane et al., 2011) were translocation hot spots for 5′ c-*myc*-I-SceI BEs, including our top 3 non-*IgH* hot spots (*Il4ra*, *CD83*, and *Pim1*) and probable AID-dependent translocation targets (e.g., *Pax5* and *Rapgef1*) (Figure 4A; Table S3). We also identified other AID-dependent translocation hot spots including the *Aff3*, *Il21r*, and *Socs2* genes, and a nonannotated intergenic transcript on chr4 (Gm12493, Figure 4A; Table S3). We confirmed the ability of such hot spots to translocate to the c-*myc*[25xI-SceI] cassette by direct PCR (Table S4). We conclude that AID not only binds and mutates numerous non-Ig target genes but also acts on them to cause DSBs and translocations.

Translocations Genome-wide Frequently Occur Near Active Transcription Start Sites

To quantify transcription genome-wide, we applied unbiased global run-on sequencing (GRO-seq; Core et al., 2008) to αCD40/IL4-activated, I-SceI-infected B cells. GRO-seq measures

elongating Pol II activity and distinguishes transcription on both strands. For all analyses, we excluded junctions within 1 Mb of the *c-myc* breaksite to avoid biases from this dominant class of junctions. To analyze remaining junctions from WT and *AID*$^{-/-}$ backgrounds, we determined nearest transcription start sites (TSSs) and divided translocations based on whether or not the TSS had promoter proximal activity based on GRO-seq (Extended Experimental Procedures). Strikingly, both WT and *AID*$^{-/-}$ junctions, when dominant *IgH* translocations were excluded, showed a distinct peak that reached a maximum about 300–600 bp on the sense side of the active TSSs and spanned from about 600 bp on the antisense side to about 1 kb on the sense side (Figures 4B and 4C). Translocation hot spot genes, including *Il4ra*, *CD83*, *Gm12493*, *Pim1*, as well as potential hot spots including *Pax5* and *Bcl11a*, had a substantial proportion of their translocations within 1–2 kb regions starting 200–400 bp in the sense direction from their bidirectional TSSs (Figures 5A and 5B). In one striking example of TSS-proximal translocation targeting, there were distinct translocation peaks downstream of the TSSs of *Il4ra* and *Il21r*, which lies just 20 kb downstream; yet, there were no detected translocations into the 3′ portion of *Il4ra* even though it was highly transcribed (Figure 5A). While lower level translocations into some AID hot spot genes in *AID*$^{-/-}$ mice had less correlation with TSS proximity (Figures 5A and 5B), the overall correlation of translocations and active TSS appeared similar in WT and *AID*$^{-/-}$ mice (Figures 4B and 4C; Figures S5A and S5B). Together, our findings indicate a relationship between active TSSs and AID-dependent and independent translocations genome-wide. In this context, we did not find a marked TSS correlation for translocations into nontranscribed genes (Figures 4B and 4C).

When the dominant *IgH* hot spots were included in the translocation/transcription analyses, the translocation peak shifted from about 300–600 bp to about 1.5 kb downstream of the TSS in the sense direction (compare Figures 4B and 4C to Figures S5C and S5D). In B cells, transcription through Sμ initiates from the V(D)J exon and Iμ exon promoters upstream of Sμ. B cell activation with αCD40/IL4 stimulates CSR between Sμ and Sγ1 or Sε by inducing AID and by activating Iγ1 and Iε promoters upstream of Sγ1 and Sε. Indeed, most translocations into germline C$_H$ genes in WT αCD40/IL4-activated B cell were tightly clustered 1-2 kb downstream in the 5′ portion of Sμ, Sγ1, and Sε, consistent with transcription robustly targeting AID to S regions (Figure 5C). Finally, AID-independent *IgH* translocations were scattered more broadly through S and C regions, suggesting that DSBs that initiate them arise by a different, AID-independent mechanism of S region instability (Figure 5C).

For 5′ *c-myc*-I-SceI BEs (outside the breaksite region), 55% of translocations were within genes, whereas genes account for only 36% of the genome (Table S5). Therefore, we asked whether translocations from 5′ *c-myc*-I-SceI BEs varied with gene density. For this purpose, we compared translocation densities to available gene density maps and to our GRO-seq transcription maps of all genes (Figure 6; Figures S6 and S7). Strikingly, translocation distribution was highly correlated with gene density and transcription level. In general, chromosomal regions with highest transcriptional activity had highest translocation density. In contrast, regions with very low or undetectable transcription gener-

ally were very low in translocations (Figure 6; Figures S6 and S7). Notably, we found no obvious regions with high overall transcription and low translocation levels, supporting a direct relationship between active transcription and translocation targeting genome-wide. In this context, we observed several robust AID-independent hot spot peaks that were relatively distant to the TSS and/or occurred in nonactive genes (Figures 4B and 4C, asterisks); these hot spots were generated by I-SceI activity at cryptic endogenous I-SceI sites as discussed next.

HTGTS Libraries Reveal Numerous Cryptic Genomic I-SceI Target Sites

Eleven AID-independent translocation targets for 5′ *c-myc*-I-SceI BEs were in genes and two were in intergenic regions (Table S3). Eight of these hot spot regions, in which junctions were tightly clustered, contained potential I-SceI-related sites, many of which were very near (within 50 bp) or actually contributed to translocation junctions. These putative cryptic I-SceI sites had from 1 to 5 divergent nucleotides with respect to the canonical 18 bp target site (Figure 7A). We scanned the mouse genome for potential cryptic I-SceI sites that diverged up to three positions and identified ten additional sites that map within 400 bp of one or more 5′ *c-myc*-I-SceI BE translocation junctions (Figure 7A). In vitro I-SceI digestion of PCR-amplified genomic fragments demonstrated that all eight putative I-SceI targets at hot spots, and six of seven tested additional putative I-SceI targets, were bona fide I-SceI substrates (Figures 7A and 7B). We performed direct translocation PCRs with three selected cryptic I-SceI sites and confirmed I-SceI-dependent translocation to the *c-myc*$^{25xI-SceI}$ cassette (Figure 7C). Finally, GRO-seq analyses showed that five of eight cryptic I-SceI translocation hot spots were in transcriptionally silent areas and that two I-SceI-generated hot spots in transcribed genes were distant from the TSS (Figures 4B and 4C, asterisks; Figures 7D and 7E), highlighting the distinction between the I-SceI-generated hot spots and most other genomic translocation hot spots.

DISCUSSION

With HTGTS, we have identified the genome-wide translocations that emanate from DSBs introduced into *c-myc* or *IgH* in activated B cells. A substantial percentage of these translocations (80%–90%) join introduced DSBs to sequences on the same chromosome proximal to the breaksite, likely reflecting the strong preference for C-NHEJ to join DSBs intrachromosomally (Ferguson et al., 2000; Zarrin et al., 2007; Mahowald et al., 2009). The remaining 10%–20% translocate broadly across all chromosomes, with translocation density correlating with transcribed gene density. Translocations are most often near TSSs within individual genes. Despite *c-myc* and *IgH* DSBs translocating broadly, there are translocation hot spots, with the majority being generated by cellular AID activity and most of the rest by ectopically expressed I-SceI activity at cryptic genomic I-SceI target sequences. Notably, targeted DSBs join at similar levels to both (+) and (−) orientations of hot spot sequences, arguing against a role for cellular selection in their appearance. This finding also suggests that both sides of hot spot DSBs have similar opportunity to translocate to a DSB on another chromosome.
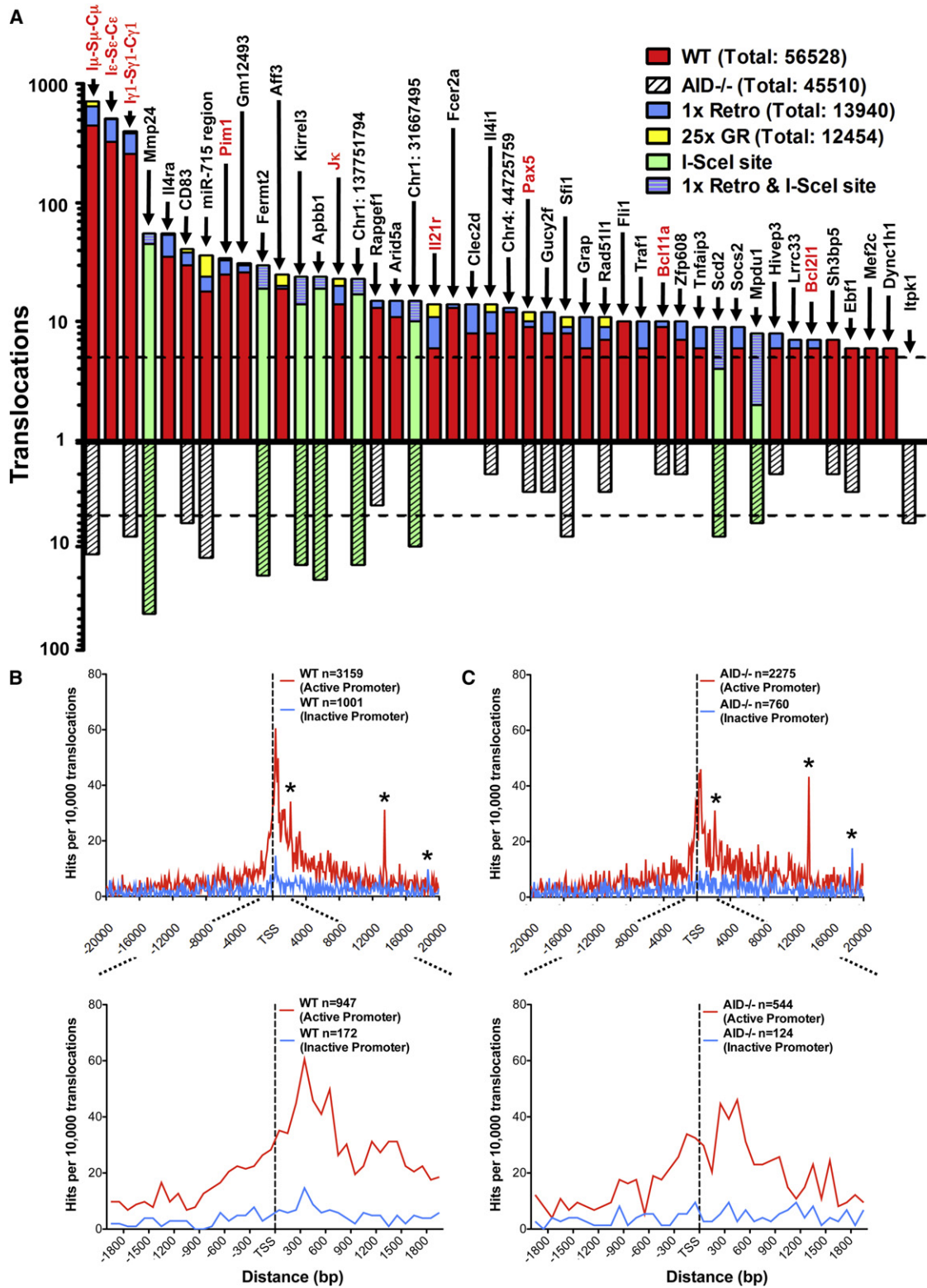
Figure 4. Identification of Specific and General Translocation Hot Spots

(A) Graph representing translocation numbers in frequently hit genes and non-annotated chromosomal regions. Only hot spots with more than five hits are shown and are ordered based on frequency of translocations in the pooled c-$myc^{25xI\text{-}Scel}$/WT HTGTS library (top bars). Respective frequencies of translocations in the pooled c-$myc^{25xI\text{-}Scel}$/$AID^{-/-}$ HTGTS library are displayed underneath (bottom bars). Green bars represent frequent hits involving cryptic I-Scel sites. Blue and
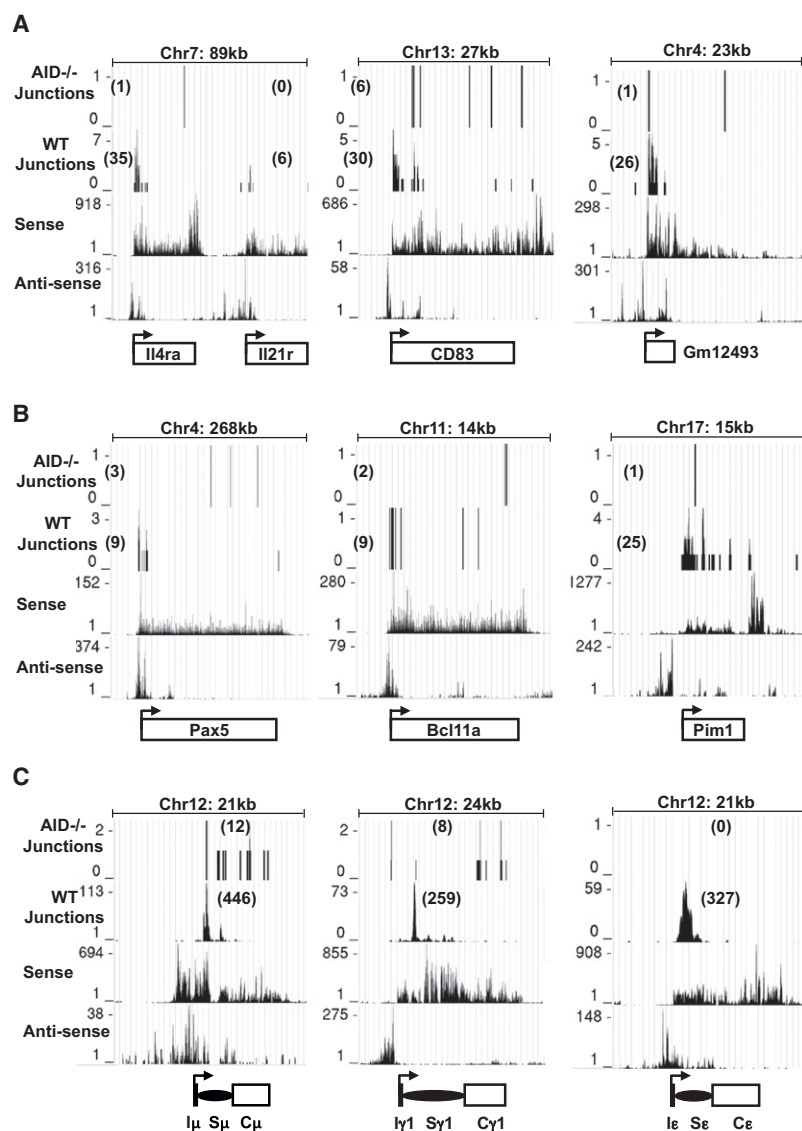
Figure 5. Translocations Preferentially Occur Near TSSs

WT and $AID^{-/-}$ c-$myc^{25xI\text{-}SceI}$ HTGTS libraries were analyzed. In each panel, translocation junctions are in the first and second rows (WT and $AID^{-/-}$ as indicated). The third and fourth rows represent sense and antisense nascent RNA signals from GRO-seq. The IgH μ, γ1, ε genes are shown in (C), the next most frequently hit hot spots in (A) and three selected oncogene hot spots in (B). The transcriptional start site (arrow) is at the bottom of each panel. The size of each genomic region and number of junctions in each are shown.

et al., 2009). HTGTS also provides a method to discover recurrent genomic DSBs, as evidenced by ability of HTGTS to find known DSBs, such as AID-initiated DSBs in S regions, and previously unrecognized genomic I-SceI targets. HTGTS should be readily applicable for genome-wide screens for translocations and recurrent DSBs in a wide range of cell types.

AID Has a Dominant Role in Targeting Recurrent Translocations Genome-wide

Prior studies demonstrated that AID binds to and mutates non-Ig genes (Pasqualucci et al., 2001; Liu et al., 2008; Yamane et al., 2011). We find that AID also induces DSBs and translocations in non-Ig genes with the peak of translocation junctions spanning the region of the TSS. Thus, processes closely associated with transcription and, potentially, transcriptional initiation may attract AID activity to these non-Ig gene targets, consistent with ectopically expressed AID mutating yeast promoter regions (Gómez-González and Aguilera, 2007). IgH translocation junctions mostly fall 1.5–2 kb downstream of the activated I region TSSs within S regions, which are known to be specialized AID targets. Thus, transcription through S regions attracts and focuses AID activity, at least in part via pausing mechanisms and by generating appropriate DNA substrates, such as R-loops, for this single-strand DNA-specific cytidine deaminase (Yu et al., 2003; Pavri and Nussenzweig, 2011; Chaudhuri et al., 2007). Notably, S regions still qualified as translocation hot spots for 5' c-myc-I-SceI BEs in $AID^{-/-}$ B cells, supporting suggestions that these regions, perhaps via transcription, may be intrinsically prone to DSBs (Dudley et al., 2002; Kovalchuk et al., 2007; Unniraman et al., 2004). Given

The majority of HTGTS junctions from the c-myc I-SceI DSBs are mediated by end-joining and contain short MHs, reminiscent of joins in cancer genomes (Stratton et al., 2009) and consistent with roles for either (or both) C-NHEJ or A-EJ (Zhang et al., 2010). Recurrence of translocations in cancer genomes is a characteristic used to consider them as potential oncogenic "drivers." Our HTGTS studies establish that many recurrent translocations form in the absence of selection and, thus, are caused by factors intrinsic to the translocation mechanism (Wang et al., 2009; Lin

yellow portions of top bars represent translocations found in c-$myc^{1xI\text{-}SceI}$ and c-$myc^{25xI\text{-}SceI}$/ROSA$^{I\text{-}SceI\text{-}GR}$ libraries, respectively. Genes translocated in human and mouse lymphoma or leukemia are in red. The dashed line represents the cutoff for significance over random occurrence for each of the two groups (see Table S3).

(B and C) Genome-wide distribution of translocations relative to TSSs. Junctions from c-$myc^{25xI\text{-}SceI}$/WT (B) or c-$myc^{25xI\text{-}SceI}$/$AID^{-/-}$ (C) libraries (excluding 2 Mb around chr15 breaksite and IgH S regions) are assigned a distance to the nearest TSS and separated into "active" and "inactive" promoters as determined by GRO-seq. Translocation junctions are binned at 100 bp intervals. n represents the number of junctions within 20 kb (upper panels) or 2 kb (lower panels) of TSS. Asterisks indicate cryptic genomic I-SceI sites. See also Figure S5.
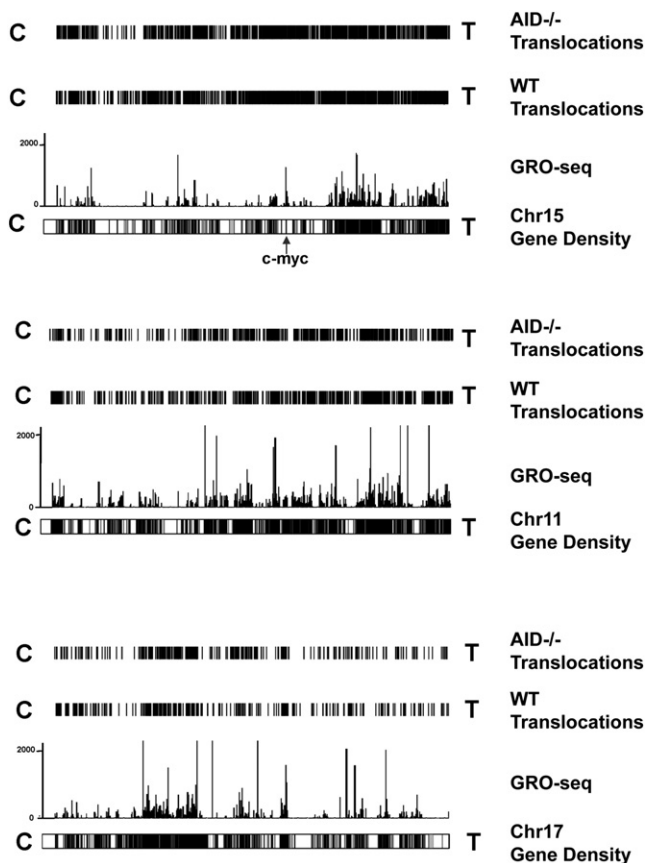
Figure 6. Translocations Cluster to Transcribed Regions
Translocation density maps from pooled c-*myc*[25xI-SceI]/WT and c-*myc*[25xI-SceI]/*AID*[−/−] HTGTS libraries are aligned with combined sense and antisense nascent RNA signals for chr 15, 11, and 17 using the UCSC genome browser. Chromosome gene densities are displayed below GRO-seq traces. Chromosomal orientation from left to right is centromere (C) to telomere (T). See also Figures S6 and S7.

the differential targeting of CSR and SHM (Liu and Schatz, 2009), application of HTGTS to germinal center (GC) B cells, in which AID initiates SHM within variable region exons, may reveal novel AID genomic targets not observed in B cells activated for *IgH* CSR in culture, potentially including genes that could contribute to GC B cell lymphoma (Küppers and Dalla-Favera, 2001).

A General Role for Transcription and Transcription Initiation in Targeting Translocations

We find a remarkable genome-wide correlation between transcription and translocations even in *AID*[−/−] cells, with a peak of translocation junctions lying near active TSSs. In this context, while the majority of junctions were located in the sense transcriptional direction, junctions also occurred at increased levels close to the TSS on the antisense side (e.g., Figures 4B and 4C; Figure 5), correlating with focal antisense transcription in the immediate vicinity of active promoters (Core et al., 2008) (Figure 5). Notably, we observed a number of regions genome-wide that were quite low in or devoid of translocations and transcription, but few, if any, that were low in translocations but high in transcription

(Figure 6). On the other hand, we found that transcription is not required for high-frequency translocations, since many I-SceI-dependent hot spots are in nontranscribed regions. Together, our observations are consistent with transcription mechanistically promoting translocations by promoting DSBs. Thus, our findings strongly support the long-standing notion of a mechanistic link between transcription, DSBs, and genomic instability (Aguilera, 2002; Haffner et al., 2011; Li and Manley, 2006).

Potential Influences of Genome Organization on Translocations

The high level of translocations of 5′ c-*myc*-I-SceI BEs to other sequences along much the length of chr15, while generally correlated with transcription, likely may be further promoted by high relative proximity of many intrachromosomal regions (Lieberman-Aiden et al., 2009). Proximity might also contribute to the apparently increased frequency of 5′ c-*myc*-I-SceI BEs to certain regions of various chromosomes (e.g., Figure 2). In this regard, the relative frequency of chr15 5′ c-*myc*-I-SceI BE translocations to the Sμ and Sε regions on chr12 were only 5 and 7 fold less, respectively, than levels of intra-*IgH* 5′ Sγ1-I-SceI BE joins to Sμ and Sε (Figure 3C). Thus, even though DSBs are rare in *c-myc*, their translocation to *IgH* when they do occur is driven at a high rate by other mechanistic aspects, most likely proximal position (Wang et al., 2009). However, we also note that sequences lying in regions across all chromosomes translocate to DSBs in *c-myc* on chr15 and *IgH* in chr12, suggesting the possibility that, in some cases, DSBs might move into proximity before joining, perhaps during the cell cycle or via other mechanisms (e.g., Dimitrova et al., 2008).

HTGTS Reveals an Unexpectedly Large Number of Genomic I-SceI Targets

Our HTGTS studies revealed 18 cryptic genomic I-SceI sites as translocation targets. There could potentially be more cryptic I-SceI sites; to find the full spectrum, bait sequences may need to be introduced into a variety of chromosomal locations to neutralize position effects. Beyond I-SceI, the HTGTS approach could readily be extended through the use Zinc finger nucleases (Händel and Cathomen, 2011), meganucleases (Arnould et al., 2011), or TALENs (Christian et al., 2010) designed to cleave specific endogenous sites, thereby obviating the need to introduce a cutting site and greatly facilitating the process. The above three classes of endonucleases are being developed for targeted gene correction of human mutations in stem cells for gene therapy. One major concern with such nucleases is relative activity on the specific target versus off-target activity, with the latter being difficult to assess. HTGTS provides a means for identifying off-target DSBs generated by such enzymes, for assessing ability of such off-target DSBs to translocate, and for identifying the sequences to which they translocate.

EXPERIMENTAL PROCEDURES

Mouse Strains Utilized

ΔSγ1[2xI-SceI], c-*myc*[25xI-SceI] and *AID*[−/−] mice were described (Zarrin et al., 2007; Wang et al., 2009; Muramatsu et al., 2000). c-*myc*[1xI-SceI] mice were generated similarly to c-*myc*[25xI-SceI] mice (see Extended Experimental Procedures). ROSA[I-SceI-GR] mice were generated by targeting an I-SceI-GR/IRES-tdTomato expression cassette into *Rosa26* (Extended Experimental
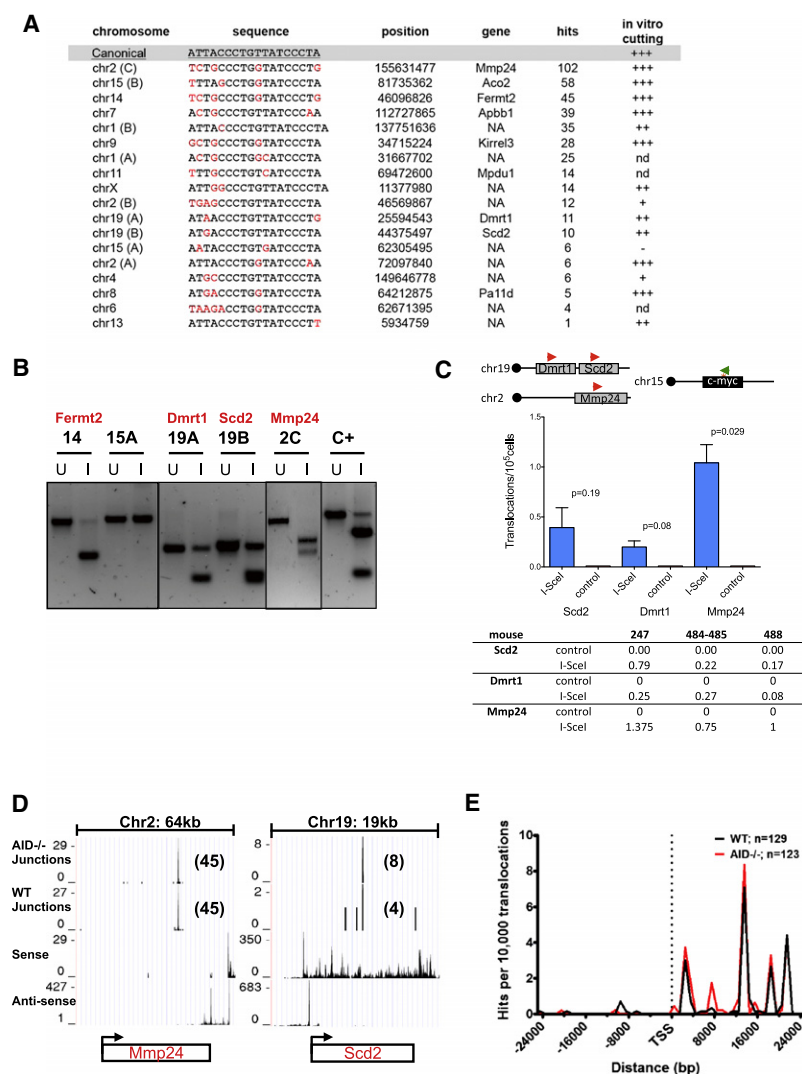
Figure 7. Identification of Cryptic I-SceI Sites in the Mouse Genome by HTGTS

(A) Cryptic I-SceI site translocation targets. The canonical I-SceI recognition sequence is on top; nucleotides divergent from the consensus are in red. Chromosomal position and gene location of each cryptic site are indicated. "Hits" represent total number of unique junctions in a 4 kb region centered around each site in the pool of all HTGTS libraries (see also Table S6). In vitro cutting efficiency, evaluated as in Extended Experimental Procedures, is indicated. NA, intergenic or not annotated; nd, not determined.

(B) In vitro cutting of PCR products encompassing indicated cryptic I-SceI sites. C+, positive control: PCR fragment containing a canonical I-SceI site. U, uncut; I, I-SceI-digested.

(C) PCR to detect translocations between c-$myc^{25xI\text{-}SceI}$ and cryptic I-SceI sites in $Scd2$, $Dmrt1$, and $Mmp24$ genes. (Top) Position of primers used for PCR amplification. (Middle) Average frequency of translocations ± SEM (Bottom) Number of translocations/10^5 cells from three independent c-$myc^{25xI\text{-}SceI}$ WT mice.

(D) Transcription in genes containing I-SceI sites determined by GRO-seq. Translocation junctions are in the first ($AID^{-/-}$) and second (WT) rows; sense and antisense nascent RNA signals are in the third and fourth rows.

(E) Distance of cryptic I-SceI hot spots from the nearest TSS in pooled HTGTS libraries from WT and $AID^{-/-}$ c-$myc^{25xI\text{-}SceI}$ B cells.

of the I-SceI cassette. Libraries were sequenced by Roche-454. See Extended Experimental Procedures for details.

Data Analysis

Alignment and Filtering

Sequences were aligned to the mouse reference genome (NCBI37/mm9) with the BLAT program. Custom filters were used to purge PCR repeats and multiple types of artifacts including those caused by in vitro ligation and PCR mispriming.

Hot Spot Identification

Translocations from WT or $AID^{-/-}$ libraries minus those on chr15 or the IgH locus were pooled. The adjusted genome was then divided into 250 kb bins and bins containing five or more hits constituted a hot spot (details in Extended Experimental Procedures).

In Vitro Testing of Putative Cryptic I-SceI Sites

A genomic region encompassing each candidate I-SceI site was PCR-amplified and 500 ng of purified products were incubated with 5 units of I-SceI for 3 hr. Reactions were separated on agarose gel and relative intensity of uncut and I-SceI-digested bands calculated with the FluorchemSP program (Alpha Innotech) (see Extended Experimental Procedures).

PCR Detection of Translocations between c-myc and Cryptic I-SceI Sites

Translocation junctions between c-myc and cryptic I-SceI targets were PCR-amplified according to the standard protocol (Wang et al., 2009). Primers and PCR conditions are detailed in Extended Experimental Procedures.

GRO-Seq

Nuclei were isolated from day 4 αCD40/IL4-stimulated and I-SceI-infected c-$myc^{25xI\text{-}SceI}$ B cells as described (Giallourakis et al., 2010). GRO-seq libraries were prepared from 5×10^6 cells from two independent mice using a described protocol (Core et al., 2008). Both libraries were sequenced on the Hi-Seq 2000

Procedures). All mice used were heterozygous for modified alleles containing I-SceI cassettes. The Institutional Animal Care and Use Committee of Children's Hospital, Boston approved all animal work.

Splenic B Cell Purification, Activation in Culture, and Retroviral Infection

All procedures were performed as previously described (Wang et al., 2009). c-$myc^{25xI\text{-}SceI}$/ROSA$^{I\text{-}SceI\text{-}GR}$ B cells were cultured in medium containing charcoal-stripped serum and I-SceI-GR was activated with 10 μM triamcinolone acetate (TA, Sigma).

Generation of HTGTS Libraries

Genomic DNA was digested with HaeIII for c-$myc^{25xI\text{-}SceI}$ samples or MspI for ΔSγ1$^{2xI\text{-}SceI}$ samples. For adaptor-PCR libraries, an asymmetric adaptor was ligated to cleaved genomic DNA. Ligation products were incubated with restriction enzymes chosen to reduce background from germline and unrearranged targeted alleles. Three rounds of nested PCR were performed with adaptor- and locus-specific primers. For circularization-PCR libraries, HaeIII- or MspI-digested genomic DNA was incubated at 1.6 ng/μl to favor intramolecular ligation and samples treated with blocking enzymes as above. Two rounds of nested PCR were performed with primers specific for sequences upstream

platform with single-end reads and analyzed as described (see Extended Experimental Procedures). After filtering and alignment, we obtained 34,212,717 reads for library 1 and 15,913,244 reads for library 2. As results between libraries were highly correlated, we show results only from replicate 1.

REFERENCES

Aguilera, A. (2002). The connection between transcription and genomic instability. EMBO J. 21, 195–201.

Arnould, S., Delenda, C., Grizot, S., Desseaux, C., Pâques, F., Silva, G.H., and Smith, J. (2011). The I-CreI meganuclease and its engineered derivatives: applications from cell modification to gene therapy. Protein Eng. Des. Sel. 24, 27–31.

Chaudhuri, J., Basu, U., Zarrin, A., Yan, C., Franco, S., Perlot, T., Vuong, B., Wang, J., Phan, R.T., Datta, A., et al. (2007). Evolution of the immunoglobulin heavy chain class switch recombination mechanism. Adv. Immunol. 94, 157–214.

Christian, M., Cermak, T., Doyle, E.L., Schmidt, C., Zhang, F., Hummel, A., Bogdanove, A.J., and Voytas, D.F. (2010). Targeting DNA double-strand breaks with TAL effector nucleases. Genetics 186, 757–761.

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.

Dimitrova, N., Chen, Y.C., Spector, D.L., and de Lange, T. (2008). 53BP1 promotes non-homologous end joining of telomeres by increasing chromatin mobility. Nature 456, 524–528.

Dudley, D.D., Manis, J.P., Zarrin, A.A., Kaylor, L., Tian, M., and Alt, F.W. (2002). Internal IgH class switch region deletions are position-independent and enhanced by AID expression. Proc. Natl. Acad. Sci. USA 99, 9984–9989.

Ferguson, D.O., Sekiguchi, J.M., Chang, S., Frank, K.M., Gao, Y., DePinho, R.A., and Alt, F.W. (2000). The nonhomologous end-joining pathway of DNA repair is required for genomic stability and the suppression of translocations. Proc. Natl. Acad. Sci. USA 97, 6630–6633.

Franco, S., Gostissa, M., Zha, S., Lombard, D.B., Murphy, M.M., Zarrin, A.A., Yan, C., Tepsuporn, S., Morales, J.C., Adams, M.M., et al. (2006). H2AX prevents DNA breaks from progressing to chromosome breaks and translocations. Mol. Cell 21, 201–214.

Giallourakis, C.C., Franklin, A., Guo, C., Cheng, H.L., Yoon, H.S., Gallagher, M., Perlot, T., Andzelm, M., Murphy, A.J., Macdonald, L.E., et al. (2010). Elements between the IgH variable (V) and diversity (D) clusters influence antisense transcription and lineage-specific V(D)J recombination. Proc. Natl. Acad. Sci. USA 107, 22207–22212.

Gómez-González, B., and Aguilera, A. (2007). Activation-induced cytidine deaminase action is strongly stimulated by mutations of the THO complex. Proc. Natl. Acad. Sci. USA 104, 8409–8414.

Gostissa, M., Alt, F.W., and Chiarle, R. (2011). Mechanisms that promote and suppress chromosomal translocations in lymphocytes. Annu. Rev. Immunol. 29, 319–350.

Gostissa, M., Ranganath, S., Bianco, J.M., and Alt, F.W. (2009). Chromosomal location targets different MYC family gene members for oncogenic translocations. Proc. Natl. Acad. Sci. USA 106, 2265–2270.

Haffner, M., De Marzo, A.M., Meeker, A.K., Nelson, W.G., and Yegnasubramanian, S. (2011). Transcription-induced DNA double strand breaks: both an oncogenic force and potential therapeutic target? Clin. Cancer Res. 17, 3858–3864.

Händel, E.M., and Cathomen, T. (2011). Zinc-finger nuclease based genome surgery: it's all about specificity. Curr. Gene Ther. 11, 28–37.

Jasin, M. (1996). Genetic manipulation of genomes with rare-cutting endonucleases. Trends Genet. 12, 224–228.

Kovalchuk, A.L., duBois, W., Mushinski, E., McNeil, N.E., Hirt, C., Qi, C.F., Li, Z., Janz, S., Honjo, T., Muramatsu, M., et al. (2007). AID-deficient Bcl-xL transgenic mice develop delayed atypical plasma cell tumors with unusual Ig/Myc chromosomal rearrangements. J. Exp. Med. 204, 2989–3001.

Küppers, R., and Dalla-Favera, R. (2001). Mechanisms of chromosomal translocations in B cell lymphomas. Oncogene 20, 5580–5594.

Lieber, M.R. (2010). The mechanism of double-strand DNA break repair by the nonhomologous DNA end-joining pathway. Annu. Rev. Biochem. 79, 181–211.

Li, X., and Manley, J.L. (2006). Cotranscriptional processes and their influence on genome stability. Genes Dev. 20, 1838–1847.

Lieberman-Aiden, E., van Berkum, N.L., Williams, L., Imakaev, M., Ragoczy, T., Telling, A., Amit, I., Lajoie, B.R., Sabo, P.J., Dorschner, M.O., et al. (2009). Comprehensive mapping of long-range interactions reveals folding principles of the human genome. Science 326, 289–293.

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

Liu, M., and Schatz, D.G. (2009). Balancing AID and DNA repair during somatic hypermutation. Trends Immunol. 30, 173–181.

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.

Mahowald, G.K., Baron, J.M., Mahowald, M.A., Kulkarni, S., Bredemeyer, A.L., Bassing, C.H., and Sleckman, B.P. (2009). Aberrantly resolved RAG-mediated DNA breaks in Atm-deficient lymphocytes target chromosomal breakpoints in cis. Proc. Natl. Acad. Sci. USA 106, 18339–18344.

Meaburn, K.J., Misteli, T., and Soutoglou, E. (2007). Spatial genome organization in the formation of chromosomal translocations. Semin. Cancer Biol. 17, 80–90.

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.

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

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

Pavri, R., and Nussenzweig, M.C. (2011). AID Targeting in Antibody Diversity. Adv. Immunol. *110*, 1–26.

Quinlan, A.R., Clark, R.A., Sokolova, S., Leibowitz, M.L., Zhang, Y., Hurles, M.E., Mell, J.C., and Hall, I.M. (2010). Genome-wide mapping and assembly of structural variant breakpoints in the mouse genome. Genome Res. *20*, 623–635.

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.

Reynaud, S., Delpy, L., Fleury, L., Dougier, H.L., Sirac, C., and Cogné, M. (2005). Interallelic class switch recombination contributes significantly to class switching in mouse B cells. J. Immunol. *174*, 6176–6183.

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., and Nussenzweig, M.C. (2008). AID is required for the chromosomal breaks in c-myc that lead to c-myc/IgH translocations. Cell *135*, 1028–1038.

Shaffer, D.R., and Pandolfi, P.P. (2006). Breaking the rules of cancer. Nat. Med. *12*, 14–15.

Siebert, P.D., Chenchik, A., Kellogg, D.E., Lukyanov, K.A., and Lukyanov, S.A. (1995). An improved PCR method for walking in uncloned genomic DNA. Nucleic Acids Res. *23*, 1087–1088.

Simsek, D., and Jasin, M. (2010). Alternative end-joining is suppressed by the canonical NHEJ component Xrcc4-ligase IV during chromosomal translocation formation. Nat. Struct. Mol. Biol. *17*, 410–416.

Stratton, M.R., Campbell, P.J., and Futreal, P.A. (2009). The cancer genome. Nature *458*, 719–724.

Unniraman, S., Zhou, S., and Schatz, D.G. (2004). Identification of an AID-independent pathway for chromosomal translocations between the Igh switch region and Myc. Nat. Immunol. *5*, 1117–1123.

Wang, J.H., Gostissa, M., Yan, C.T., Goff, P., Hickernell, T., Hansen, E., Difilippantonio, S., Wesemann, D.R., Zarrin, A.A., Rajewsky, K., et al. (2009). Mechanisms promoting translocations in editing and switching peripheral B cells. Nature *460*, 231–236.

Yamane, A., Resch, W., Kuo, N., Kuchen, S., Li, Z., Sun, H.W., Robbiani, D.F., McBride, K., Nussenzweig, M.C., and Casellas, R. (2011). Deep-sequencing identification of the genomic targets of the cytidine deaminase AID and its cofactor RPA in B lymphocytes. Nat. Immunol. *12*, 62–69.

Yan, C.T., Boboila, C., Souza, E.K., Franco, S., Hickernell, T.R., Murphy, M., Gumaste, S., Geyer, M., Zarrin, A.A., Manis, J.P., et al. (2007). IgH class switching and translocations use a robust non-classical end-joining pathway. Nature *449*, 478–482.

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*, 442–445.

Zarrin, A.A., Del Vecchio, C., Tseng, E., Gleason, M., Zarin, P., Tian, M., and Alt, F.W. (2007). Antibody class switching mediated by yeast endonuclease-generated DNA breaks. Science *315*, 377–381.

Zhang, Y., Gostissa, M., Hildebrand, D.G., Becker, M.S., Boboila, C., Chiarle, R., Lewis, S., and Alt, F.W. (2010). The role of mechanistic factors in promoting chromosomal translocations found in lymphoid and other cancers. Adv. Immunol. *106*, 93–133.

Zhu, C., Mills, K.D., Ferguson, D.O., Lee, C., Manis, J., Fleming, J., Gao, Y., Morton, C.C., and Alt, F.W. (2002). Unrepaired DNA breaks in p53-deficient cells lead to oncogenic gene amplification subsequent to translocations. Cell *109*, 811–821.