Published in final edited form as:
Cell. 2012 March 2; 148(5): 908–921. doi:10.1016/j.cell.2012.02.002.

Chromosomal translocations are guided by the spatial organization of the genome

Yu Zhang[1,*], **Rachel Patton McCord**[2,*], **Yu-Jui Ho**[1], **Bryan R. Lajoie**[2], **Dominic G. Hildebrand**[1], **Aline C. Simon**[1], **Michael S. Becker**[1], **Frederick W. Alt**[1], and **Job Dekker**[2]

[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, Massachusetts 02115, USA

[2]Programs in Systems Biology and Gene Function and Expression, Department of Biochemistry and Molecular Pharmacology, University of Massachusetts Medical School, Worcester, MA 01605-0103, USA

SUMMARY

The extent to which the three dimensional organization of the genome contributes to chromosomal translocations is an important question in cancer genomics. We now have generated a high resolution Hi-C spatial organization map of the G1-arrested mouse pro-B cell genome and mapped translocations from target DNA double strand breaks (DSBs) within it via high throughput genome-wide translocation sequencing. RAG endonuclease-cleaved antigen-receptor loci are dominant translocation partners for target DSBs regardless of genomic position, reflecting high frequency DSBs at these loci and their co-localization in a fraction of cells. To directly assess spatial proximity contributions, we normalized genomic DSBs via ionizing-radiation. Under these conditions, translocations were highly enriched *in cis* along single chromosomes containing target DSBs and within other chromosomes and sub-chromosomal domains in a manner directly related to pre-existing spatial proximity. Our studies reveal the power of combining two high-throughput genomic methods to address long-standing questions in cancer biology.

Keywords

Translocations; 3D nuclear organization; DNA double-strand breaks; genome stability

INTRODUCTION

Recurrent chromosomal translocations are hallmarks of many cancers (Zhang et al., 2010; Gostissa et al., 2011). These translocations often represent rare events in tumor progenitors that are highly selected for contribution to oncogenicity. Thus, since cancer genome studies examine only the final outcome of a multi-step transformation, the complex genomic alterations observed are often difficult to fully interpret (Stratton, 2011). Notably, factors that mechanistically enhance translocations can promote recurrent non-oncogenic

Address Correspondence to: Frederick W. Alt (alt@enders.tch.harvard.edu), Job Dekker (job.dekker@umassmed.edu).
*Equal Contribution

translocations in tumors and also promote particular oncogenic translocations (Wang et al., 2009; Gostissa et al., 2009a). Most translocations are formed by end-joining of two DSBs (Stephens et al., 2009). Thus, mechanistic factors that could influence translocations include DSB frequency at translocating loci, factors that promote physical juxtaposition ("synapsis") of DSBs, and factors that circumvent cellular DSB repair pathways that promote normal rejoining (Zhang et al., 2010)

DSBs are major drivers of translocations (Tsai and Lieber, 2010; Chiarle et al., 2011; Klein et al., 2011). Lymphoid malignancies commonly have recurrent translocations involving programmed DSBs in antigen receptor loci as one partner. In developing lymphocytes, such DSBs are introduced into both Immunoglobulin (Ig) and T cell receptor (TCR) loci by the RAG endonuclease during V(D)J recombination.(Schatz and Swanson, 2011). In mature B cells, assembled Ig loci can be broken downstream of activation induced cytidine deaminase (AID) activity during Ig heavy chain (IgH) class switch recombination (CSR) (Liu and Schatz, 2009b). Recurrent antigen receptor locus translocations in lymphoid malignancies reflects both high frequency programmed DSBs within these loci, as well as ability of their powerful enhancers to activate translocated oncogenes (Gostissa et al., 2009b).

Sources of DSBs in non-Ig translocation partners are less well-characterized (Tsai and Lieber, 2010). Mouse studies implicated AID-initiated *c-myc* DSBs in translocations (Robbiani et al., 2008); and human tumor analyses suggested that collaborations between AID and RAG generate DSBs leading to oncogene translocations (Tsai et al., 2008). High throughput genome wide translocation sequencing ("HTGTS") studies (Chiarle et al., 2011) and studies using a similar approach (Klein et al., 2011) showed that I-SceI DSBs within the *IgH* or *c-myc* translocate to other DSBs widely across the genome. In these studies, endogenous translocation hotspots were generated by AID-induced DSBs within off-target genes; while a broader set of genome wide translocations were associated with transcription start sites ("TSSs") (Chiarle et al., 2011; Klein et al., 2011). Beyond transcription, cell intrinsic factors, such as oxidative metabolism, replication stress, or cell extrinsic factors such as ionizing radiation (IR) or chemotherapeutics may generate DSBs (Tsai and Lieber, 2010). In general, DSBs outside of antigen receptor loci, for example in *c-myc*, are likely rate-limiting for translocations (Wang et al., 2009). In this regard, recurrent translocations in non-lymphoid tumors, for example in myeloid leukemias or prostate cancers, may be driven by oncogenic selection for rare events and/or reflect mechanistic drivers other than DSBs (Kurzrock et al., 2003; Lin et al., 2009; Mani et al., 2009).

Fusion of two DSBs to generate translocations requires their synapsis. Chromosomes occupy distinct nuclear territories (Cremer and Cremer, 2010) and are further organized into open and closed chromatin domains within different nuclear compartments (Gilbert et al., 2004; Lieberman-Aiden et al., 2009). Due to the substantial heterogeneity in spatial genome organization, most transcribed genes co-localize in at least a small subset of cells (Simonis et al., 2006; Lieberman-Aiden et al., 2009). The non-random nuclear position of genes and chromosomes has led to the notion that close spatial proximity of two genomic sequences will promote their preferential translocations (Meaburn et al., 2007). Accordingly, cytogenetic studies showed several genes that contribute to translocations were, on average, in relatively close proximity (e.g. Neves et al., 1999; Roix et al., 2003; Osborne et al., 2007; Wang et al., 2009), and live cell imaging studies showed DSBs to be relatively immobile (Soutoglou et al., 2007). However, analysis of the overall influence of genome organization on translocations requires a comprehensive characterization of 3D genome structure and measurement of the landscape of potential translocations in the same cellular system. We now directly assess the impact of DSB location within the 3D genome on translocation potential by mapping the genome structure of a G1-arrested mouse pro-B cell line with Hi-C and identifying the translocations that form with an induced DSB in these cells.

RESULTS

G1-Arrested Pro-B cells as a Model System to Study Translocation Mechanisms

HTGTS and related studies of translocations in activated mouse B cells revealed enrichment of translocations along chromosome 15 on which the target I-SceI site in *c-myc* was located, suggesting, among other possibilities, potential effects related to chromosome organization (Chiarle et al., 2011; Klein et al., 2011). However, these studies did not determine whether enrichment was only on the true "*cis*" chromosome or on both chromosome 15 copies. Further studies to determine effects of chromosomal organization on translocations in such activated B cells are limited by potentially confounding factors. First, actively transcribed AID translocation hotspots, as well collectively abundant transcription-associated translocations, may be driven both by increased DSB frequency, as well as by propensity of actively transcribed genes to associate (Gilbert et al., 2004; Simonis et al., 2006; Lieberman-Aiden et al., 2009). Second, in activated cycling B cells, there are changes in genome organization and various cellular selections imposed as cells traverse the cell cycle. Finally, analysis of multiple I-SceI integration sites on different chromosomes is required to fully assess genome organization effects. Therefore, we developed a cell line system that overcomes these potential caveats.

We employed Bcl-2 transgenic Abelson Murine Leukemia Virus (A-MuLV) transformed pro-B cell lines that, in response to treatment with *v-abl* inhibitor STI571, arrest in the G1 cell cycle phase (Bredemeyer et al., 2006) (Fig. 1A). The STI571-mediated G1 arrest also induces RAG, which cleaves the endogenous *Igκ*, providing a positive control for frequent endogenous DSBs. For introduction of targeted DSBs at defined loci, we randomly integrated a retroviral substrate containing 26 I-SceI target sites (to increase cutting efficiency) and isolated different pro-B clones that each had a unique, single genomic substrate integration. To generate DSBs at these sites, we stably expressed an I-SceI-GR fusion protein (Chiarle et al., 2011), which allowed ligand-inducible cleavage of I-SceI substrates following G1-arrest (Fig. 1A).

Pro-B cells display little cytogenetic instability, even when deficient for non-homologous end-joining DSB repair (Sekiguchi et al., 2001). Correspondingly, when HTGTS libraries were prepared from wild type (WT) G1-arrested A-MuLV transformants with an I-SceI substrate in chromosome (chr) 15, we failed to observe translocations to Igκ (chr 6) and also found an increased background of artifactual junctions compared to what we found in activated B cells (Chiarle et al., 2011) (Fig. S1A and B). These findings suggested a low translocation frequency in WT pro-B lines that reduced our ability to generate large numbers of translocations junctions. DSBs activate the Ataxia-Telangiectasia mutated (ATM) kinase which then phosphorylates multiple substrates. We previously showed that ATM-deficiency results in dramatic cytogenetic instability in pro-B cells (Sekiguchi et al., 2001). In the absence of ATM, V(D)J recombination DSBs in A-MuLV transformants progress to chromosome breaks and translocations (Bredemeyer et al., 2006). Correspondingly, upon treating chr 15-integrated and chr 2-integrated WT pro-B lines, respectively, with ATM inhibitor (Ku55933 or "ATMi") before generation of HTGTS libraries, we found dramatic *Igκ* hotspots for both, as well as multiple translocations into other RAG targets including *IgH*, *TCRγ* and *TCRα* loci (Fig. S1A; Table S1). Therefore, we focused subsequent studies on genetically ATM-deficient ($ATM^{-/-}$) pro-B cell lines.

We generated large scale translocation libraries from three $ATM^{-/-}$ pro-B lines in which I-SceI substrates were integrated, respectively, into chr 18, chr 2, or chr 7 (Table S1). Consistent with ATMi-treated WT line results, predominant translocations in $ATM^{-/-}$ lines were within endogenous Igκ and other RAG target loci (Fig. 1B, Fig. 2, Fig. 3 and Fig. S2), with up to 25% of total junctions occurring within these loci. Junctions that were not to

antigen receptor loci tended to be spread across the genome. In addition, we observed increased junctions near the I-SceI breaksite, which were substantially comprised of resections (Fig. S3B; Chiarle et al., 2011). Even when junctions within 2Mb of the breaksite were excluded, however, we observed a mild preference for junctions along the length of the same chromosome on which the I-SceI sequence was targeted (Fig. 1C, Fig. 2, Fig. 3 and Fig. S2). Finally, we observed dominant junctions to *Igκ* for two additional integrations on chr 2, as well as integrations on chr 3 and chr 13 (Table S1).

IR Treatment of G1-arrested ATM-deficient Pro-B Lines Normalizes DSB Frequency Genome Wide

The finding that I-SceI DSBs in 8 different chromosomal sites form frequent translocations with 5 different antigen receptor loci indicates that high frequency DSBs can dominate translocation patterns. In this regard, we propose that the rate at which particular translocations are formed should be influenced by the product of at least three variables; namely the frequencies of DSBs at each partner site and the frequency at which these sites are synapsed. As illustrated by the extremes of such a model, highly frequent DSBs at one site would not drive translocations to a frequently synapsed site that was not broken; while frequent DSBs at two sites would not drive translocations between them if they were never proximal. Further tests of this model require normalization of DSB frequency across the genome so that proximity is an overall rate limiting factor. For this purpose, we γ-irradiated G1-arrested pro-B lines, a dosage that generates several hundred random DSBs per genome (Rothkamm and Lobrich, 2003). We reasoned that such ectopically introduced DSBs would obviate dominant endogenous DSBs as the main driving force for translocations. Accordingly, in IR-treated cells, the percentage of junctions within Ig and TCR loci decreased substantially (Fig. 1B and Table S1). However, Igκ still represented a weaker translocation hotspot after IR (Fig. 1B and Fig. S3A), suggesting that IR-introduced random DSB frequency remained lower than that of the endogenous *Igκ* DSBs. In addition, we did not observe enrichment of translocation junctions around TSSs in the γ-irradiated pro-B lines (Fig. 1D–F), supporting the notion that IR treatment helped normalize a broader distribution of DSBs.

IR-treatment generated a dramatic effect on the distribution of translocations across pro-B cell chromosomes. Thus, we observed a tremendous increase in the proportion of junctions along the full length of the (numerical) chromosome harboring the target I-SceI DSBs and not just in the area proximal to the breaksite (Fig. 1C, 2, 3, S2). Indeed, much of the chromosome harboring individual breaksites, in effect, became a translocation hotspot, with the overall level of junctions ranging from 25–40% of the total. This phenomenon was observed for all 8 separate I-SceI integrations. Based on studies of chromosome territories in cycling human cells (Bolzer et al., 2005; Lieberman-Aiden et al., 2009), we speculated that enriched breaksite chromosome translocations subsequent to IR may reflect spatial proximity of sequences along individual chromosomes.

Allele Specific Translocation Analysis of cis-Translocations

To test whether or not the high level of translocations on the breaksite-containing chromosome represented true *cis*-translocations, we generated A-MuLV transformed pro-B cell lines from 129/BALB/c F1 mice and introduced into them a single-copy retroviral construct containing a V(D)J recombination substrate (Bredemeyer et al., 2006). We used this substrate because of its high efficiency of generating DSBs and translocations in the ATM-deficient pro-B lines (Mahowald et al., 2009). We treated two cell lines which, respectively, contained a substrate integrated into the BALB/c-derived chr 7 and into the 129-derived chr 9 with STI571 and ATMi to promote unrepaired DSBs within integrated V(D)J substrates and also with IR to generate genome-wide DSBs (Fig. 4). Translocation

patterns of RAG-generated construct DSBs strongly resembled those generated from I-SceI-initiated DSBs after IR treatment, with libraries from both chr7 and chr9 integrations being enriched for junctions at antigen receptor loci and all along the breaksite chromosome (Fig S4A,B). Thus, our general findings with I-SceI DSBs can be generalized to more than one type of target DSB and, more specifically, to a known class of physiological DSBs. We extracted and plotted 129 versus BALB/c allele-specific translocations based on available SNPs (Yalcin et al., 2011) (Fig. 4A,B and Fig. S4C,D). After removal of junctions lying within 2Mb of the breaksite to remove resections, we observed enriched breaksite chromosome translocations only on the BALB/c allele for the chr. 7 integration and only on the 129 allele for chr. 9 integration, indicating that they were, indeed, enriched only on the *cis* chromosome (Fig. 4 and Fig. S4E–J). Translocations to other chromosomes showed no allele-specific bias.

Hi-C analysis reveals the spatial organization G1-arrested mouse pro-B Cell Genome

To determine whether the spectrum of translocations observed by HTGTS was related to pre-existing spatial juxtaposition frequencies of translocation partners, we performed Hi-C (Lieberman-Aiden et al. 2009) to map the spatial organization of the complete genome in G1-arrested $ATM^{-/-}$ and WT pro-B cells prior to induction of DSBs and irradiation. Hi-C combines Chromosome Conformation Capture (3C, Dekker et al., 2002) with purification of ligation junctions followed by deep sequencing to detect and quantify chromosomal interactions throughout the genome. The resulting Hi-C data provide insights into the folding of the mouse genome in the $ATM^{-/-}$ cells that were used for analysis of translocations (Fig. 5) as well as in WT pro-B cells (Supp. Fig. S5). A two-dimensional heat map of interactions throughout the genome, where each pixel represents all interactions between two 10 Mb regions, shows that intra-chromosomal interactions are much more frequent than inter-chromosomal interactions (Fig. 5A), reflecting “chromosome territories” that have been observed previously by Hi-C and cell imaging (Lieberman-Aiden, 2009; Mayer et al. 2005). When we analyzed interactions within single chromosomes at higher resolution (1 Mb), we found a decreasing interaction frequency with increasing genomic distance (Fig. 5B,D), which is characteristic of the flexible polymer nature of the chromatin fiber (Dekker et al., 2002).

The intra-chromosomal interaction maps are also characterized by reproducible "plaid" patterns of alternating regions with high and low interaction frequency (Fig. 5B and S5D). This pattern becomes even more obvious upon normalizing Hi-C signals for genomic distance and displaying the correlations between interaction profiles of loci (Fig. 5C). Previous work showed that the first principle component of this correlation map captures this plaid pattern, with positive values corresponding to open and transcriptionally active domains (A-domains), and negative values corresponding to closed and inactive chromatin (B-domains) (Lieberman-Aiden et al., 2009). In these G1-arrested $ATM^{-/-}$ and WT cells, we found that this domain structure correlated with gene density, showing that the G1 mouse genome is compartmentalized in A- and B-domains, similar to human (Fig. 5C, S5E).

In cycling human cells, intra-chromosomal contact probability ($P(s)$) has been shown to decay rapidly as a function of genomic distance (s), following a power law with exponent −1 ($P(s) \sim s^{-1}$) for loci separated by 0.5–7 Mb. This represents a fractal globule conformation, a polymer state characterized by dense packing in the absence of topological entanglements (Lieberman-Aiden, 2009, Mirny 2011). When we plotted the Hi-C data for mouse G1-arrested cells in a similar fashion we find a power law scaling with exponent −1 for loci separated by 0.5 up to at least 5 Mb, indicating that mouse chromatin in G1 cells also folds according to a fractal globule (Fig. 5D).

Whole chromosomes within the mammalian nucleus often occupy preferred positions relative to one another (Boyle et al., 2001; Lieberman-Aiden, 2009; Sengupta et al., 2008). Indeed, in the G1-arrested mouse pro-B cells, we also find that the longest chromosomes (chr 1, 2, 3, etc.) interact with each other more frequently than with the shortest chromosomes (chr 17, 18, 19 etc.) and *vice versa* (Fig. 5E). However, intermediate length chromosomes have a less well-defined position relative to other chromosomes (Suppl. Fig. 5E; see below).

Overall, G1-arrested WT cells and ATM−/− cells showed similar folding principles (Pearson correlation between WT and ATM−/− around 0.98), indicating that, as judged at the current resolution of Hi-C, ATM deficiency by itself does not affect formation and relative positioning of chromosome territories, chromosome compartmentalization in A- and B-domains, and folding of chromatin according to a fractal globule (Supp. Fig. S5). We observed changes in the precise genomic locations of A- and B-compartments in 8% of the genome of WT as compared to $ATM^{-/-}$ cells; as expected given that these cells may display some differences in gene expression (Bredemeyer et al., 2008), which impacts the localization of genes in the A- or B-compartment (Lieberman-Aiden et al. 2009).

Intra-chromosomal translocation frequency is correlated with Hi-C contact probability

We compared the features of 3D spatial genome organization in the ATM−/− G1-arrested pro-B lines before translocations with the patterns of translocations after I-SceI induction and IR treatment. The most obvious similarity between spatial organization and translocation frequency is the enrichment of both chromatin interactions and translocations within the chromosome harboring the I-SceI DSBs (Figs. 2,3,5). This finding suggests that enrichment for translocations on the breaksite chromosome is driven by spatial proximity of regions of the same chromosome. To investigate whether translocations correlate with spatial proximity at a finer scale within the breaksite chromosome, we compared translocation frequency with interaction frequency between the chromosomal region containing the I-SceI site and 1 Mb bins along the rest of the breaksite chromosome (Fig. 6A, 6B). The translocation profile looks strikingly similar to and correlates strongly with the Hi-C interaction frequency between the I-SceI site and other regions along the breaksite chromosome for I-SceI sites on chr18, chr2, and chr7 (Fig 6, Supp. Fig. 6D–E). Thus, when acceptor DSBs are widely available along all chromosomes (e.g. after IR), our findings show that the spatial conformation of the chromosome containing the breaksite is a strong predictor of translocation frequency along that chromosome.

Translocations between chromosomes are correlated with Hi-C contact probability

To examine whether the correlation between translocations and spatial interactions extends to inter-chromosomal events genome-wide, we compared the frequency of interaction between *trans* chromosomal regions and the I-SceI breaksite to their frequency of forming translocations. We excluded bins that contained translocation hotspots related to Ig and TCR loci. We then calculated the frequency of interaction between 5 Mb bins on *trans* chromosomes and the I-SceI site to their frequency of forming translocations with the I-SceI break. We found that bins with higher translocation counts (>=2 translocations per 1,000 in the dataset) also had a significantly higher distribution of Hi-C interaction scores with the induced breaksite than bins with few or no translocations (< 2 translocations per 1,000 in the dataset; Fig. 7A). This result is statistically significant over a broad range of translocation thresholds (Suppl. Fig. S7A). Thus, even though interchromosomal interactions are weaker than intrachromosomal interactions, these weak preferences affect the likelihood of translocations between interchromosomal regions. This significant correlation between translocation frequency and Hi-C score is also evident when the average Hi-C profile is

plotted around bins that display frequent and infrequent inter-chromosomal translocations with I-SceI sites (Suppl. Fig. S7B).

To evaluate the impact of whole chromosome nuclear positioning on translocation frequency, we compared interaction patterns of whole chromosomes (Fig. 5E) with translocation frequency across whole chromosomes. We observed a high correlation between inter-chromosomal translocation frequencies and whole chromosome nuclear position for both the chr18 and the chr2 I-SceI target sites (Fig. 7B). Strikingly, the chr2 and chr18 I-SceI target sites showed completely different inter-chromosomal translocation distributions, which correlated with their own specific inter-chromosomal nuclear interaction profiles (Fig 7B). Thus, while DSB frequency is presumably similar across all chromosomes in these two experiments, the frequency of translocations differs dramatically depending on the relative proximity of other chromosomes to the I-SceI breaksite chromosome. While we noted an occasional discordance between whole chromosome Hi-C interaction levels and translocation frequency (such as the chr2 I-SceI site with chr9), the correlation in these cases between translocation frequency and spatial proximity was still significant at a 5 Mb resolution (Supp. Fig. S7D). Finally, like other medium-length chromosomes, chr 7 did not exhibit a strong positioning relative to other whole chromosomes (Supp Fig. S7C). Correspondingly, while we did not find a significant correlation between chr7 I-SceI DSB translocations and whole chromosome position (Supp. Fig. S7E), we did find a positive correlation (R=0.53) when inter-chromosomal translocation frequency was compared to the nuclear position of a smaller region of chr 7 (1Mb) around the I-SceI target site (Supp. Fig. S7E).

DISCUSSION

Our combined Hi-C and HTGTS analyses show definitively that the 3D genome organization and spatial proximity among loci strongly influence patterns of chromosomal rearrangements and translocations genome wide. Nuclear organization and chromosome conformation have been implicated in gene regulation (Cremer and Cremer, 2010), and also have been shown to influence translocations based on cytogenetic studies (Neves et al., 1999; Roix et al., 2003; Meaburn et al., 2007; Wang et al., 2009). However, such cytogenetic studies have certain limitations, including only being done for a single pair of genomic loci and in a limited number of cells. In addition, the actual definition of two proximal loci in cytogenetic studies has been somewhat arbitrary (Neves et al., 1999; Mathas et al., 2009; Mani et al., 2009; Lin et al., 2009; Wang et al., 2009). Our approach avoids such caveats, allowing clear demonstration that spatial proximity can guide joining of DSBs genome wide to form translocations. Thus, in IR-treated G1-arrested pro-B lines, where DSBs are not limiting, the position of a DSB can vastly change the landscape of its predominant translocation partners (Fig. 7B). In addition, our study further shows that cellular heterogeneity in the spatial conformation of the genome allows high frequency DSBs to dominate the translocation landscape, even if they are not proximal in a large proportion of the cells.

Our studies reveal the power of combining two high-throughput genomic methods to answer long-standing questions regarding the influence of nuclear organization on important biological processes. The only previous mammalian Hi-C dataset studied genome organization in cycling human cells (Lieberman-Aiden et al., 2009). Our current study demonstrates the broad applicability of Hi-C by applying it to analyze the genome organization of G1-arrested mouse pro-B lines. Application of HTGTS and related methods to activated primary B cells previously demonstrated that AID-initiated DSBs are the major endogenous sources of translocations to *c-myc* and *IgH* target I-SceI DSBs in that cell type (Chiarle et al., 2011; Klein et al., 2011). Our current studies show an even stronger

contribution of RAG-initiated DSBs in generating high frequency translocations with target I-SceI DSBs in multiple different chromosomal locations in $ATM^{-/-}$ A-MuLV transformed pro-B lines. Therefore, endogenous DSBs generated in the context of antigen receptor locus diversification dominate the translocation landscape in both pro-B cells and mature B cells. Notably, we found frequent translocations to multiple known RAG targets including various T cell receptor loci in the pro-B lines, even though some of these loci are not thought to be highly accessible to RAG in pro-B cells, demonstrating the sensitivity of HTGTS for revealing "off-target" RAG-initiated DSBs.

Hi-C analyses of Genome Organization in G1-arrested Mouse Pro-B Cells

Our Hi-C data show that basic features of chromosome and nuclear organization previously described for human cells can be extended to mouse cells. As in human cells, mouse chromosomes are compartmentalized in A- and B-compartments and chromatin at the Mb scale folds according to a fractal globule. In addition, chromosomes display preferred relative positions, with longer chromosomes preferentially associating with each other, as do the shorter chromosomes. These similarities are striking, given that mouse chromosomes are fewer in number than human chromosomes and differ in their centromere placement (acrocentric). Together with the recent characterization of Drosophila genome organization (Sexton et al. 2012), this mouse Hi-C dataset reveals functional organization of chromatin that is common across species. In addition, whereas previous Hi-C analyses were complicated by the use of cycling cells, resulting in interaction maps representing the average of all chromosome conformations throughout the cell cycle (Lieberman-Aiden et al., 2009), our current Hi-C map of G1-arrested cells represents a far more homogeneous interphase organization. Similar features between G1 arrested cells and cycling cells likely arise from the large fraction of cells being in G1 in a population of cycling cells. Though cell type specific differences in genome organization are expected, the mouse Hi-C data generated here can inform a broad range of future studies of the implications of genome folding in cell function in this important model organism.

Our Hi-C data provide the highest-resolution analysis of the spatial organization of a mammalian genome to date. The Hi-C libraries were sequenced approximately 5–10 times more deeply than in previous studies (Supp. Fig. S6; Supp. Table 2). We estimate that the resolution of intra-chromosomal interaction detection is around 100 Kb, whereas inter-chromosomal interactions can be assessed at about 1 Mb resolution (Supp. Fig. S6). Previous Hi-C maps for human cells allowed a resolution of about 1 Mb in *cis*, and 5 Mb in *trans*. In our study we find strong correlations between intra-chromosomal translocation frequencies and Hi-C interaction frequencies when we compare these phenomena at 1 Mb resolution.

Interestingly, when we analyze our Hi-C data at the highest level of resolution (i.e. 100 Kb), we find the correlation between spatial proximity with translocation frequency is reduced, though it remains high (Fig S6C). This somewhat lower correlation at a very local level may reflect the local movement of DSBs within a distance of a few hundred Kb, consistent with previously described Brownian motion of chromatin over distances up to 0.5–1 μm (Marshall et al.,1997). This DSB movement on a very local scale may also contribute to synapsis of sequences that are spatially proximal.

Heterogeneity in Genome Organization Allows Recurrent DSBs to Drive Translocations

The view of genome organization provided by Hi-C reflects cellular heterogeneity. Thus, the observed number of molecules in a Hi-C experiment that represents a certain interaction corresponds to the fraction of cells in the experiment in which those genomic regions were proximal. Our Hi-C studies show that 5 antigen receptor locus translocation "hotspots" (*Igh,*

Igκ, Igλ, TCRα/δ, TCRγ) have a finite probability of contacting 8 different randomly integrated I-SceI loci, indicating that these 40 different pairs of loci are in close spatial proximity in a sub-population of cells (Supp. Fig. S3C). Thus, this genomic spatial heterogeneity in the G1-arrested pro-B cells allows high frequency antigen receptor locus DSBs and frequent I-SceI DSBs to translocate in the fraction of cells where their contact occurs. Thus, our findings support the proposed model in which translocation frequency is a function of DSB frequency at two sites and the fraction of individual cells in a population in which the sites are juxtaposed. The latter will in part reflect known heterogeneity in the genomic spatial organization, as well as known factors that cause certain classes of genes, (e.g. genes that are actively transcribed or existing on similar sized chromosomes) to have an increased chance of being proximal (Gilbert et al., 2004; Lieberman-Aiden et al., 2009).

Impact of Spatial Proximity Within the Genome on Cancer Genomes

By saturating the genome with DSBs, we generated a situation in which high frequency DSBs in a limited number of locations no longer totally dominates translocations. Under these conditions, spatial proximity of two DSBs is a dominant factor in determining the landscape of translocations genome wide. By extension, our findings on the ability of spatial proximity to guide translocations indicates that if a gene was broken very infrequently, it would translocate to a potential target DSB, whether it occurred at high or low frequency, and at a rate that would be strongly influenced by spatial proximity. This finding has great relevance to translocations in cancer; since it is likely that, other than major RAG- or AID-initiated DSBs in antigen receptor loci, DSBs in many partner translocation targets in lymphoid cells and translocation-initiating DSBs in other cell types likely occur at low frequency. Indeed, in many non-lymphoid cancers, factors other than DSB generation, such as spatial proximity, may play a much more important role than in lymphoid cancers (Lin et al., 2009; Mani et al., 2009). In addition, formation of translocations between randomly generated DSBs, such as those induced by chemotherapies and radiotherapies, will likely reflect a strong influence of spatial proximity like that we observe upon IR-treatment of pro-B lines. Our findings also suggest that spatial proximity may be a major driving force for the activation of certain oncogenes via translocation to a wide range of recurrent partners (Ohno et al., 2006; Barreca et al., 2011; Liu et al., 2009a).

We show that spatial proximity leads to a tremendous preference for DSBs to translocate to sites all along the true *cis* chromosome on which they reside and not just near the breaksite. Such notions have been speculated (Chiarle et al., 2011; Klein et al., 2011 and Mahowald et al., 2009), but never before tested. This finding should facilitate interpretation of cancer genomes and the evolution of increasingly complex karyotypes. As one example, preferential joining of *cis* DSBs along a chromosome might have mechanistic relevance for chromothripsis, a chromosome catastrophe phenomenon that results in extensive intra-chromosomal rearrangements in various cancer genomes (Stephens et al, 2011; Liu et al., 2011; Rausch et al. Cell 2012). Thus, chromothripsis may be promoted by the intrinsic spatial organization of given chromosomes such that DSBs within them are highly likely to be re-ligated to another DSB on the same chromosome. C*is*-translocation preferences may also contribute to the predisposition of DSBs to be fused intra-chromosomally by C-NHEJ (Ferguson et al., 2000). Finally, spatial organization of chromosomes in cancer progenitors may contribute to particular initial oncogenic rearrangements, with changes in spatial organization during subsequent stages of cancer progression influencing accumulation of additional genomic rearrangements.

EXPERIMENTAL PROCEDURES

Cell lines

A-MuLV transformed cell lines were generated as described (Bredemeyer et al., 2006: Li et al., 2008). A retroviral construct containing 26X I-SceI sites was randomly integrated in the genome. F1 mice were generated by crossing BABL/cJ and 129S6.For F1 cells, a retroviral construct containing DEL-CJ substrate was used (Bredemeyer et al., 2006). Integration sites were cloned as described (Mahowald et al., 2009).

Generation and analysis of HTGTS libraries

Cells were treated with 3μM STI571 (Novartis) and 100nM TA (Sigma) for 4 days. In some experiments, 5Gy IR was applied at day 2. ATM inhibitor (Ku55933, Tocris) was used at 15μM. HTGTS was performed as described (Chiarle et al., 2011). (see Extended Experimental Procedures). Data are available from NCBI/SRA (SRA049021).

Generation and analysis of Hi-C libraries

Hi-C was performed on A-MuLV transformants after 2 days of treatment with 3μM STI571 as described (Lieberman-Aiden et al., 2009) with some modifications. Hi-C data were coverage-corrected (to account for restriction fragment size and mappability) according to the total number of sequence reads in each bin. (see Extended Experimental Procedures). Data are available at GEO accession no. GSEXXXXX and can be interactively visualized at http://hic.umassmed.edu.

Allele specific analysis of translocations

SNP data were downloaded from the Wellcome Trust Sanger Institute (Yalcin et al., 2011). We further extracted translocation coordinates that have SNPs with “Above Threshold Genotype” tags (ATG) between the 129S5 and BALB/cJ strains. Translocations from six individual F1 libraries that contain those SNPs were further analyzed and plotted.

Highlights

A high resolution Hi-C spatial organization map of the G1-arrested mouse cell genome

Genomic spatial heterogeneity contributes to recurrent chromosomal translocations

Spatial organization of chromosomes promotes dominant translocations *in cis*

Translocations are guided genome-wide by pre-existing spatial genome organization

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

This work was supported by NIH grant ARRA Supplement 3P01CA092625-09S1 and a Leukemia and Lymphoma Society of America (LLS) SCOR grant to F.W.A and by NIH grant R01 HG003143 and a W.M. Keck Foundation Distinguished Young Scholar Award to J.D. Y.Z. was supported by CRI postdoctoral fellowship. D.G.H., A.C.S., and M.G.B. were supported by the German National Merit Foundation. F.W.A. is an Investigator of the Howard Hughes Medical Institute.

REFERENCES

Barreca A, Lasorsa E, Riera L, Machiorlatti R, Piva R, Ponzoni M, Kwee I, Bertoni F, Piccaluga PP, Pileri SA, et al. Anaplastic lymphoma kinase in human cancer. Journal of molecular endocrinology. 2011; 47:R11–R23. [PubMed: 21502284]

Bolzer A, Kreth G, Solovei I, Koehler D, Saracoglu K, Fauth C, Muller S, Eils R, Cremer C, Speicher MR, et al. Three-dimensional maps of all chromosomes in human male fibroblast nuclei and prometaphase rosettes. PLoS biology. 2005; 3:e157. [PubMed: 15839726]

Boyle S, Gilchrist S, Bridger JM, Mahy NL, Ellis JA, Bickmore WA. The spatial organization of human chromosomes within the nuclei of normal and emerin-mutant cells. Human molecular genetics. 2001; 10:211–219. [PubMed: 11159939]

Bredemeyer AL, Helmink BA, Innes CL, Calderon B, McGinnis LM, Mahowald GK, Gapud EJ, Walker LM, Collins JB, Weaver BK, et al. DNA double-strand breaks activate a multi-functional genetic program in developing lymphocytes. Nature. 2008; 456:819–823. [PubMed: 18849970]

Bredemeyer AL, Sharma GG, Huang CY, Helmink BA, Walker LM, Khor KC, Nuskey B, Sullivan KE, Pandita TK, Bassing CH, et al. ATM stabilizes DNA double-strand-break complexes during V(D)J recombination. Nature. 2006; 442:466–470. [PubMed: 16799570]

Chiarle R, Zhang Y, Frock RL, Lewis SM, Molinie B, Ho YJ, Myers DR, Choi VW, Compagno M, Malkin DJ, et al. Genome-wide translocation sequencing reveals mechanisms of chromosome breaks and rearrangements in B cells. Cell. 2011; 147:107–119. [PubMed: 21962511]

Cremer T, Cremer M. Chromosome territories. Cold Spring Harbor perspectives in biology. 2010; 2 a003889.

Dekker J, Rippe K, Dekker M, Kleckner N. Capturing chromosome conformation. Science. 2002; 295:1306–1311. [PubMed: 11847345]

Ferguson DO, Sekiguchi JM, Chang S, Frank KM, Gao Y, DePinho RA, Alt FW. The nonhomologous end-joining pathway of DNA repair is required for genomic stability and the suppression of translocations. Proceedings of the National Academy of Sciences of the United States of America. 2000; 97:6630–6633. [PubMed: 10823907]

Gilbert N, Boyle S, Fiegler H, Woodfine K, Carter NP, Bickmore WA. Chromatin architecture of the human genome: gene-rich domains are enriched in open chromatin fibers. Cell. 2004; 118:555–566. [PubMed: 15339661]

Gostissa M, Alt FW, Chiarle R. Mechanisms that promote and suppress chromosomal translocations in lymphocytes. Annual review of immunology. 2011; 29:319–350.

Gostissa M, Ranganath S, Bianco JM, Alt FW. Chromosomal location targets different MYC family gene members for oncogenic translocations. Proceedings of the National Academy of Sciences of the United States of America. 2009a; 106:2265–2270. [PubMed: 19174520]

Gostissa M, Yan CT, Bianco JM, Cogne M, Pinaud E, Alt FW. Long-range MANUSCRIPT ACCEPTED oncogenic activation of Igh-c-myc translocations by the Igh 3' regulatory region. Nature. 2009b; 462:803–807. [PubMed: 20010689]

Klein IA, Resch W, Jankovic M, Oliveira T, Yamane A, Nakahashi H, Di Virgilio M, Bothmer A, Nussenzweig A, Robbiani DF, et al. Translocation-capture sequencing reveals the extent and nature of chromosomal rearrangements in B lymphocytes. Cell. 2011; 147:95–106. [PubMed: 21962510]

Kurzrock R, Kantarjian HM, Druker BJ, Talpaz M. Philadelphia chromosome-positive leukemias: from basic mechanisms to molecular therapeutics. Annals of internal medicine. 2003; 138:819–830. [PubMed: 12755554]

Li G, Alt FW, Cheng HL, Brush JW, Goff PH, Murphy MM, Franco S, Zhang Y, Zha S. Lymphocyte-specific compensation for XLF/cernunnos end-joining functions in V(D)J recombination. Molecular cell. 2008; 31:631–640. [PubMed: 18775323]

Lieberman-Aiden E, van Berkum NL, Williams L, Imakaev M, Ragoczy T, Telling A, Amit I, Lajoie BR, Sabo PJ, Dorschner MO, et al. Comprehensive mapping of long-range interactions reveals folding principles of the human genome. Science. 2009; 326:289–293. [PubMed: 19815776]

Lin C, Yang L, Tanasa B, Hutt K, Ju BG, Ohgi K, Zhang J, Rose DW, Fu XD, Glass CK, et al. Nuclear receptor-induced chromosomal proximity and DNA breaks underlie specific translocations in cancer. Cell. 2009; 139:1069–1083. [PubMed: 19962179]

Liu H, Cheng EH, Hsieh JJ. MLL fusions: pathways to leukemia. Cancer biology & therapy. 2009a; 8:1204–1211. [PubMed: 19729989]

Liu M, Schatz DG. Balancing AID and DNA repair during somatic hypermutation. Trends in immunology. 2009b; 30:173–181. [PubMed: 19303358]

Liu P, Erez A, Nagamani SC, Dhar SU, Kolodziejska KE, Dharmadhikari AV, Cooper ML, Wiszniewska J, Zhang F, Withers MA, et al. Chromosome catastrophes involve replication mechanisms generating complex genomic rearrangements. Cell. 2011; 146:889–903. [PubMed: 21925314]

Mahowald GK, Baron JM, Mahowald MA, Kulkarni S, Bredemeyer AL, Bassing CH, Sleckman BP. Aberrantly resolved RAG-mediated DNA breaks in Atm-deficient lymphocytes target chromosomal breakpoints in cis. Proceedings of the National Academy of Sciences of the United States of America. 2009; 106:18339–18344. [PubMed: 19820166]

Mani RS, Tomlins SA, Callahan K, Ghosh A, Nyati MK, Varambally S, Palanisamy N, Chinnaiyan AM. Induced chromosomal proximity and gene fusions in prostate cancer. Science. 2009; 326:1230. [PubMed: 19933109]

Marshall WF, Straight A, Marko JF, Swedlow J, Dernburg A, Belmont A, Murray AW, Agard DA, Sedat JW. Interphase chromosomes undergo constrained diffusional motion in living cells. Current biology : CB. 1997; 7:930–939. [PubMed: 9382846]

Mathas S, Kreher S, Meaburn KJ, Johrens K, Lamprecht B, Assaf C, Sterry W, Kadin ME, Daibata M, Joos S, et al. Gene deregulation and spatial genome reorganization near breakpoints prior to formation of translocations in anaplastic large cell lymphoma. Proceedings of the National Academy of Sciences of the United States of America. 2009; 106:5831–5836. [PubMed: 19321746]

Mayer R, Brero A, von Hase J, Schroeder T, Cremer T, Dietzel S. Common themes and cell type specific variations of higher order chromatin arrangements in the mouse. BMC cell biology. 2005; 6:44. [PubMed: 16336643]

Meaburn KJ, Misteli T, Soutoglou E. Spatial genome organization in the formation of chromosomal translocations. Seminars in cancer biology. 2007; 17:80–90. [PubMed: 17137790]

Mirny LA. The fractal globule as a model of chromatin architecture in the cell. Chromosome research. 2011; 19:37–51. [PubMed: 21274616]

Neves H, Ramos C, da Silva MG, Parreira A, Parreira L. The nuclear topography of ABL, BCR, PML, RARalpha genes: evidence for gene proximity in specific phases of the cell cycle and stages of hematopoietic differentiation. Blood. 1999; 93:1197–1207. [PubMed: 9949162]

Ohno H. Pathogenetic and clinical implications of non-immunoglobulin ; BCL6 translocations in B-cell non-Hodgkin's lymphoma. Journal of clinical and experimental hematopathology : JCEH. 2006; 46:43–53. [PubMed: 17142954]

Osborne CS, Chakalova L, Mitchell JA, Horton A, Wood AL, Bolland DJ, Corcoran AE, Fraser P. Myc dynamically and preferentially relocates to a transcription factory occupied by Igh. PLoS biology. 2007; 5:e192. [PubMed: 17622196]

Rausch T, Jones DT, Zapatka M, Stutz AM, Zichner T, Weischenfeldt J, Jager N, Remke M, Shih D, Northcott PA, et al. Genome Sequencing of Pediatric Medulloblastoma Links Catastrophic DNA Rearrangements with TP53 Mutations. Cell. 2012; 148:59–71. [PubMed: 22265402]

Robbiani DF, Bothmer A, Callen E, Reina-San-Martin B, Dorsett Y, Difilippantonio S, Bolland DJ, Chen HT, Corcoran AE, Nussenzweig A, et al. AID is required for the chromosomal breaks in c-myc that lead to c-myc/IgH translocations. Cell. 2008; 135:1028–1038. [PubMed: 19070574]

Roix JJ, McQueen PG, Munson PJ, Parada LA, Misteli T. Spatial proximity of translocation-prone gene loci in human lymphomas. Nature genetics. 2003; 34:287–291. [PubMed: 12808455]

Rothkamm K, Lobrich M. Evidence for a lack of DNA double-strand break repair in human cells exposed to very low x-ray doses. Proceedings of the National Academy of Sciences of the United States of America. 2003; 100:5057–5062. [PubMed: 12679524]

Schatz DG, Swanson PC. V(D)J recombination: mechanisms of initiation. Annual review of genetics. 2011; 45:167–202.

Sekiguchi J, Ferguson DO, Chen HT, Yang EM, Earle J, Frank K, Whitlow S, Gu Y, Xu Y, Nussenzweig A, et al. Genetic interactions between ATM and the nonhomologous end-joining factors in genomic stability and development. Proceedings of the National Academy of Sciences of the United States of America. 2001; 98:3243–3248. [PubMed: 11248063]

Sengupta K, Camps J, Mathews P, Barenboim-Stapleton L, Nguyen QT, Difilippantonio MJ, Ried T. Position of human chromosomes is conserved in mouse nuclei indicating a species-independent mechanism for maintaining genome organization. Chromosoma. 2008; 117:499–509. [PubMed: 18563425]

Sexton T, Yaffe E, Kenigsberg E, Bantignies F, Leblanc B, Hoichman M, Parrinello H, Tanay A, Cavalli G. Three-Dimensional Folding and Functional Organization Principles of the Drosophila Genome. Cell. 2012 Online.

Simonis M, Klous P, Splinter E, Moshkin Y, Willemsen R, de Wit E, van Steensel B, de Laat W. Nuclear organization of active and inactive chromatin domains uncovered by chromosome conformation capture-on-chip (4C). Nature genetics. 2006; 38:1348–1354. [PubMed: 17033623]

Soutoglou E, Dorn JF, Sengupta K, Jasin M, Nussenzweig A, Ried T, Danuser G, Misteli T. Positional stability of single double-strand breaks in mammalian cells. Nature cell biology. 2007; 9:675–682.

Stephens PJ, Greenman CD, Fu B, Yang F, Bignell GR, Mudie LJ, Pleasance ED, Lau KW, Beare D, Stebbings LA, et al. Massive genomic rearrangement acquired in a single catastrophic event during cancer development. Cell. 2011; 144:27–40. [PubMed: 21215367]

Stephens PJ, McBride DJ, Lin ML, Varela I, Pleasance ED, Simpson JT, Stebbings LA, Leroy C, Edkins S, Mudie LJ, et al. Complex landscapes of somatic rearrangement in human breast cancer genomes. Nature. 2009; 462:1005–1010. [PubMed: 20033038]

Stratton MR. Exploring the genomes of cancer cells: progress and promise. Science. 2011; 331:1553–1558. [PubMed: 21436442]

Tsai AG, Lieber MR. Mechanisms of chromosomal rearrangement in the human genome. BMC genomics. 2010; 11(Suppl 1):S1. [PubMed: 20158866]

Tsai AG, Lu H, Raghavan SC, Muschen M, Hsieh CL, Lieber MR. Human chromosomal translocations at CpG sites and a theoretical basis for their lineage and stage specificity. Cell. 2008; 135:1130–1142. [PubMed: 19070581]

Wang JH, Gostissa M, Yan CT, Goff P, Hickernell T, Hansen E, Difilippantonio S, Wesemann DR, Zarrin AA, Rajewsky K, et al. Mechanisms promoting translocations in editing and switching peripheral B cells. Nature. 2009; 460:231–236. [PubMed: 19587764]

Yalcin B, Wong K, Agam A, Goodson M, Keane TM, Gan X, Nellaker C, Goodstadt L, Nicod J, Bhomra A, et al. Sequence-based characterization of structural variation in the mouse genome. Nature. 2011; 477:326–329. [PubMed: 21921916]

Zhang Y, Gostissa M, Hildebrand DG, Becker MS, Boboila C, Chiarle R, Lewis S, Alt FW. The role of mechanistic factors in promoting chromosomal translocations found in lymphoid and other cancers. Advances in immunology. 2010; 106:93–133. [PubMed: 20728025]

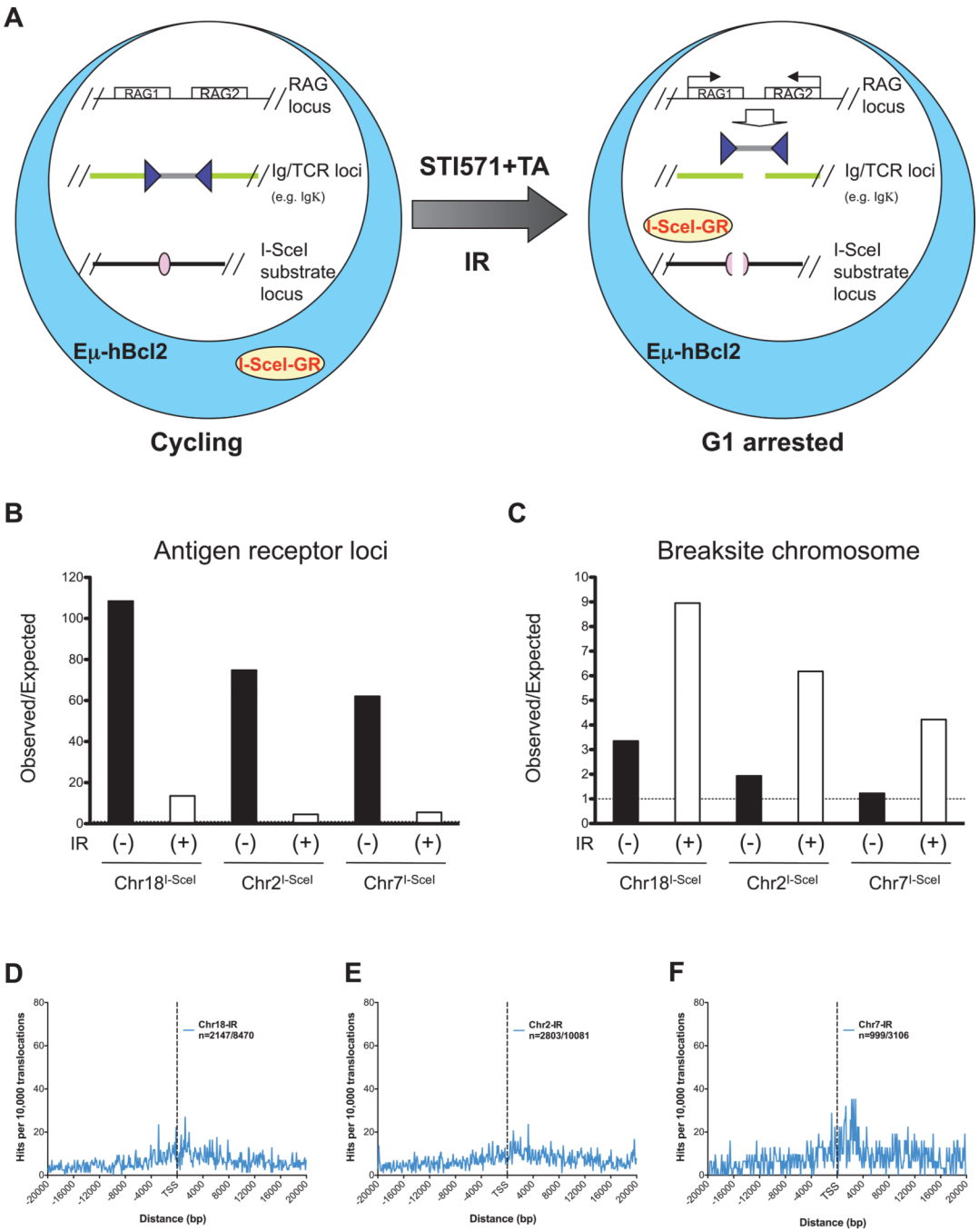

Figure 1. HTGTS to study the role of genome organization in promoting translocations

A) Treatment of cycling A-MuLV transformed pro-B lines with STI 571 leads to G1 cell cycle arrest, RAG activation and V(D)J recombination at *Ig*κ. Treatment with TA (triamcinolone acetonide) leads to translocation of I-SceI-GR into the nucleus. Expression of a Bcl-2 transgene prevents apoptosis. Treatment of cells with 5Gy IR introduces random DSBs. **B)** Dominant RAG-initiated antigen receptor locus translocations. The observed/expected ratios of translocations within 5 antigen receptor loci (Igκ, IgH, Igλ, TCRγ, and TCRα) is shown for I-Sce-I targets on chr18, chr2, and chr7 based on pooled HTGTS data from Table S1. Filled bars represent values from indicated cells not treated with IR and open

bars represent values from cells treated with 5Gy IR. **C)** Dominant occurrence of translocations in I-SceI breaksite chromosome. After breaksite proximal junctions (±1Mb) were removed, observed/expected ratios of translocations (pooled data from Table S1) on breaksite chromosomes for different clones (i.e. chr18, chr2, and chr7, respectively) are shown. Filled bars represent samples from cells not IR-treated, while open bars represent samples from cells treated with 5Gy IR. **D–F)** Genome-wide distribution of translocations relative to TSSs. Junctions from I-SceI targets on chr2(D), chr7 (E), and chr18 (F) upon IR treatment are assigned a distance to the nearest TSS. Breaksite proximal junctions (±1Mb) as well as antigen receptor locus and two chr18 hotspots were removed. Translocation junctions are binned at 100 bp intervals. "n" represents the number of junctions within 20 kb of TSS. See also Fig. S1.

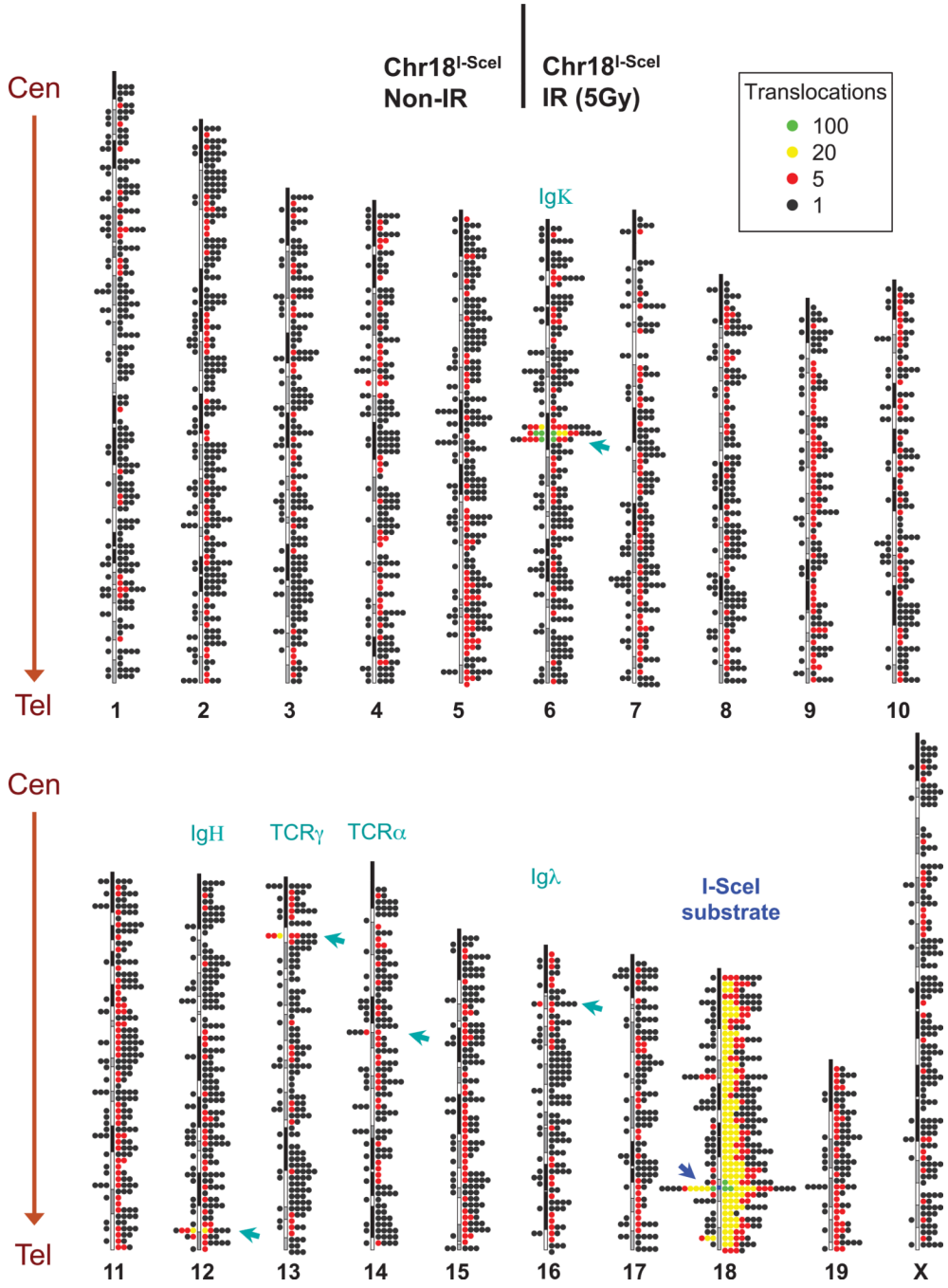

Figure 2. Genome-wide distribution of translocations from Chr18 I-SceI breaksite

Genome-wide map of translocations from the I-SceI cassette in chr18 (labeled by blue arrow) in A-MuLV transformed ATM$^{-/-}$ pro-B cells. The genome was divided into 2Mb bins and the number of unique translocations within each bin represented by colored dots with a black dot indicating one translocation, a red dot 5, a yellow dot 20 and a green dot 100, respectively. Junctions from STI571 and TA treated cells are plotted on left side of each chromosome ideogram, while translocations from STI571, TA, and 5Gy IR-treated cells are plotted on right side. Ig/TCR hotspots are indicated by green arrows. Centromere (Cen) and telomere (Tel) positions are indicated. Data are from pooled HTGTS libraries. See also Fig. S2.

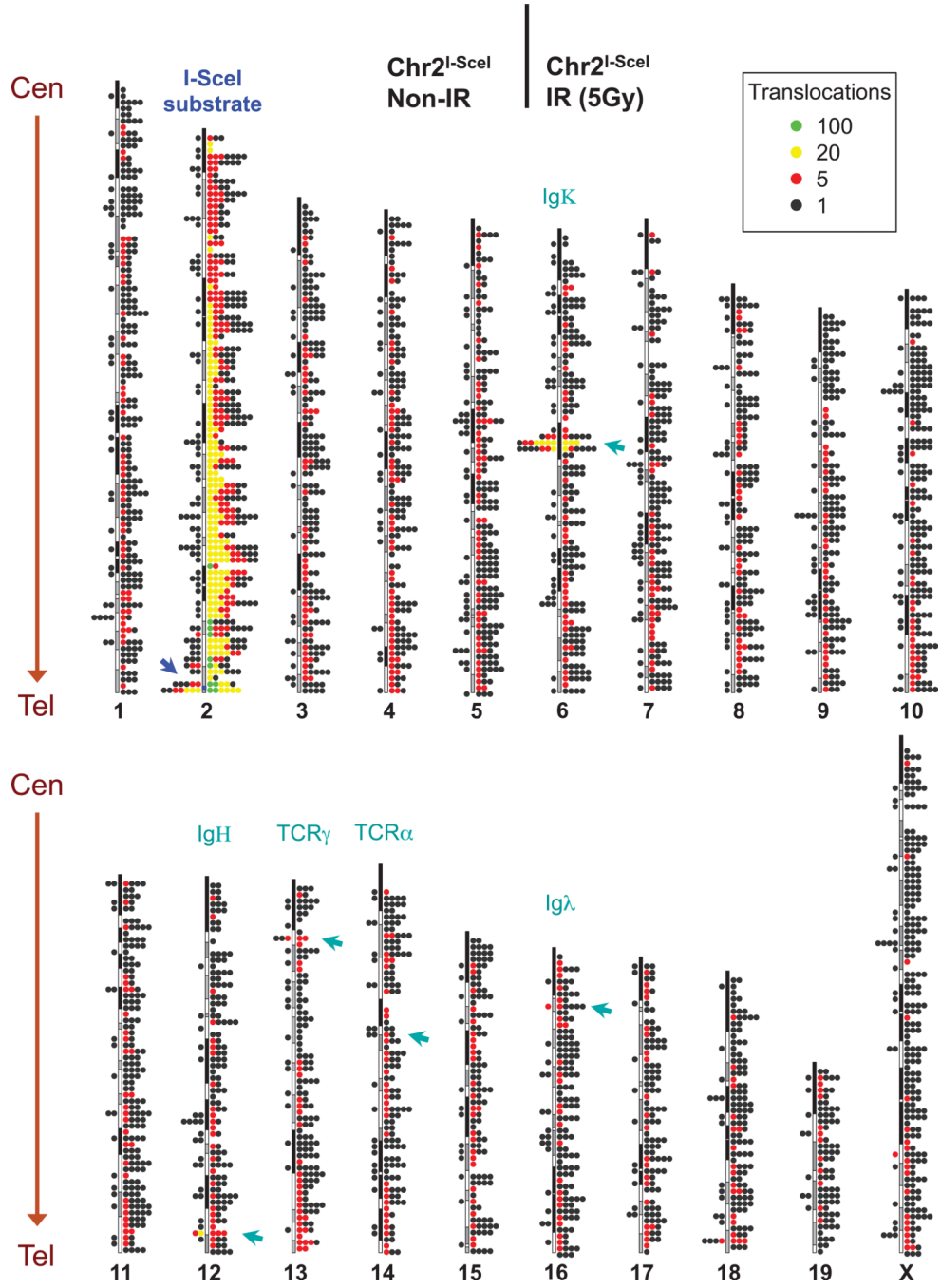

Figure 3. Genome-wide distribution of translocations from Chr2 I-SceI breaksite

Genome-wide map of translocations originating from the I-SceI cassette in chr2, labeled by blue arrow, from A-MuLV transformed ATM$^{-/-}$ mouse pro-B cells. Other details are as for Fig. 2. See also Fig. S3.

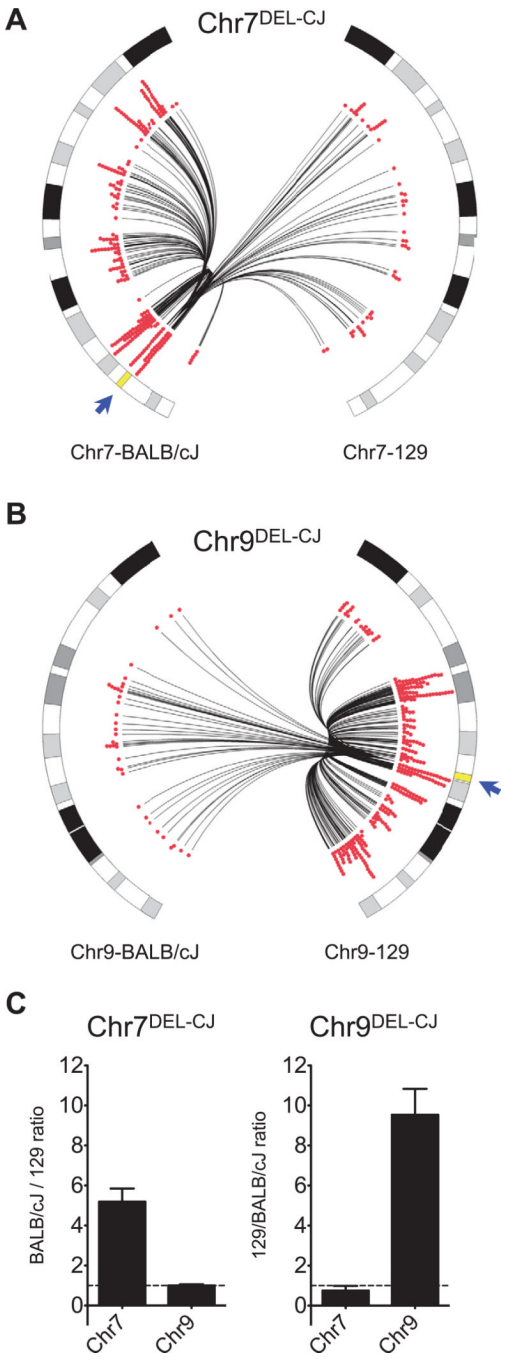

Figure 4. Allele specific distribution of chromosomal translocations in F1 A-MuLV Transformants

A, B) V(D)J recombination substrates ("DEL-CJ") were integrated into 129/BALB/cJ pro-B genome within chr7 BALB/cJ allele in one line **(A;** "Chr7$^{\mathbf{DEL\text{-}CJ}}$") and the chr9 129 allele in a different line (**B;** "Chr9$^{\mathbf{DEL\text{-}CJ}}$"). Translocation libraries from the V(D)J substrates were generated by HTGTS and translocations were mapped to the BALB/cJ versus 129 alleles based on available SNPs. The circos plots show the distribution of allelic specific junctions within the two copies of chr7 (**A**) and the two copies of chr9 (**B**) for the Chr7$^{\mathbf{DEL\text{-}CJ}}$ and Chr9$^{\mathbf{DEL\text{-}CJ}}$ lines, respectively. Individual translocations are represented as arcs originating from I-SceI DSBs and terminating at the partner site in BALB/cJ (left) or 129 (right)

chromosomes. The 2 Mb chromosomal region spanning substrate integration sites was omitted from analyses. **C)** Bar graphs show the relative allelic distribution of translocations on the BALB/cJ and 129 derived chromosomes 7 and 9 for V(D)J recombination substrates integrated into chr7 and chr9. The 2 Mb chromosomal region spanning the substrate integration site was omitted from analyses. See also Fig. S4.

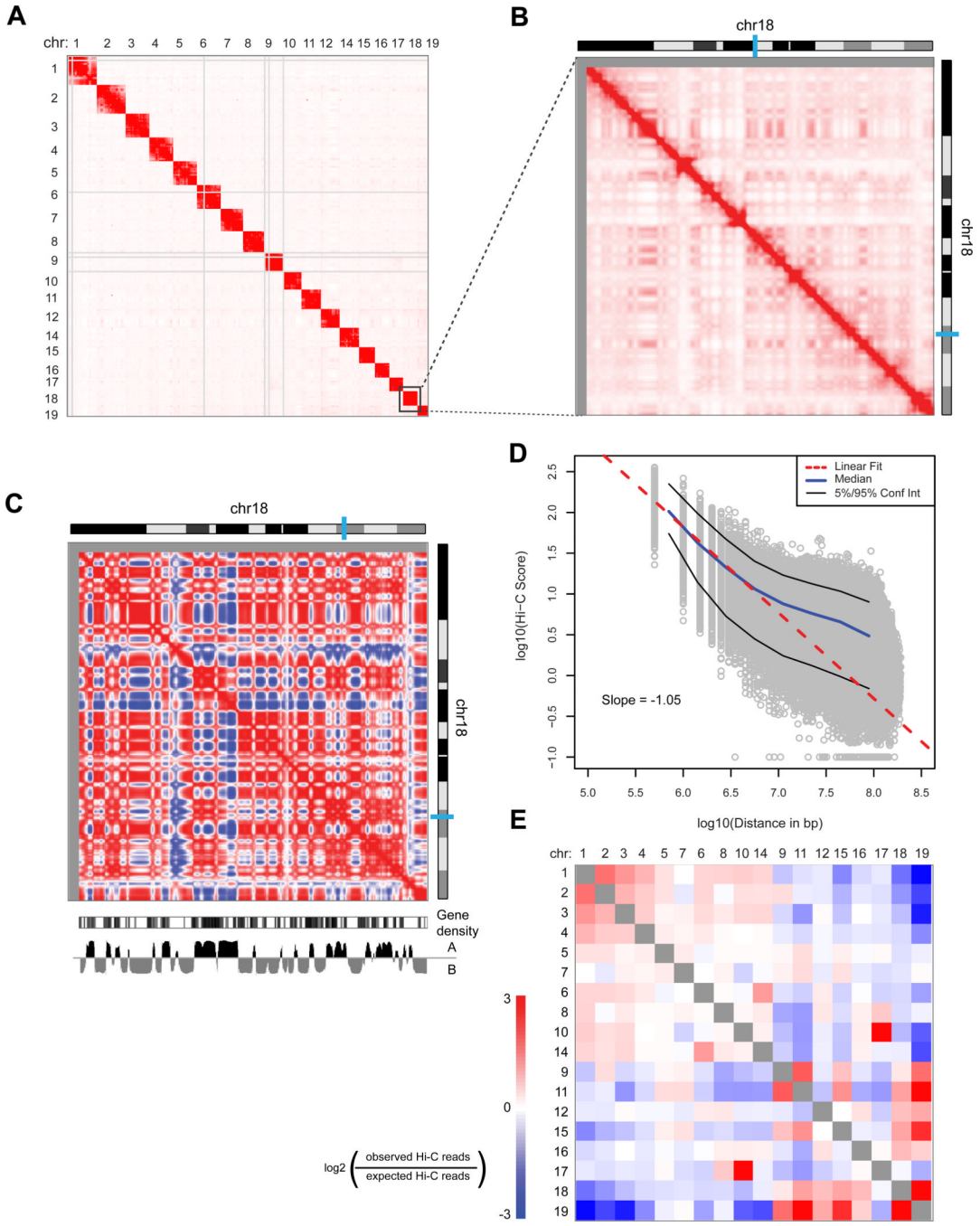

Figure 5. Hi-C analysis of G1 arrested mouse pro-B cell genome spatial organization

A) Heatmap representing the genome-wide chromatin interaction map at 10 Mb resolution. Color intensity indicates the corrected number of Hi-C sequencing reads from each pair of interacting fragments. Regions with many large restriction fragments (>100 kb) or no uniquely mapped reads are shown in gray. **B)** Higher resolution map (1 Mb bins smoothed with a 200 kb step size) of intra-chromosomal interactions along chr18. **C)** Correlation map of chr18 at 1 Mb resolution shows chromosome compartmentalization. The first principal component eigenvector (below heatmap: “A” and “B” refer to active and inactive chromatin states) identifies compartments and correlates with gene density. **D)** In log space, the

number of Hi-C contacts between two genomic loci scales linearly with the genomic distance separating the loci. The slope of −1 indicates a fractal globule polymer organization, and is observed for loci separated by 0.5 to at least 5 Mb. **E)** Observed/expected number of contacts between all pairs of whole chromosomes (sorted by length). Red indicates enrichment; blue indicates depletion. Chr13 is excluded due to a pre-existing translocation. All Hi-C data result from pooling all reads from 5 replicate experiments shown in Fig S5. Cell lines were A-MuLV transformed $ATM^{-/-}$ mouse pro-B cell lines arrested in G1 by STI571 treatment for 2 days. See Fig. S5 for WT results.

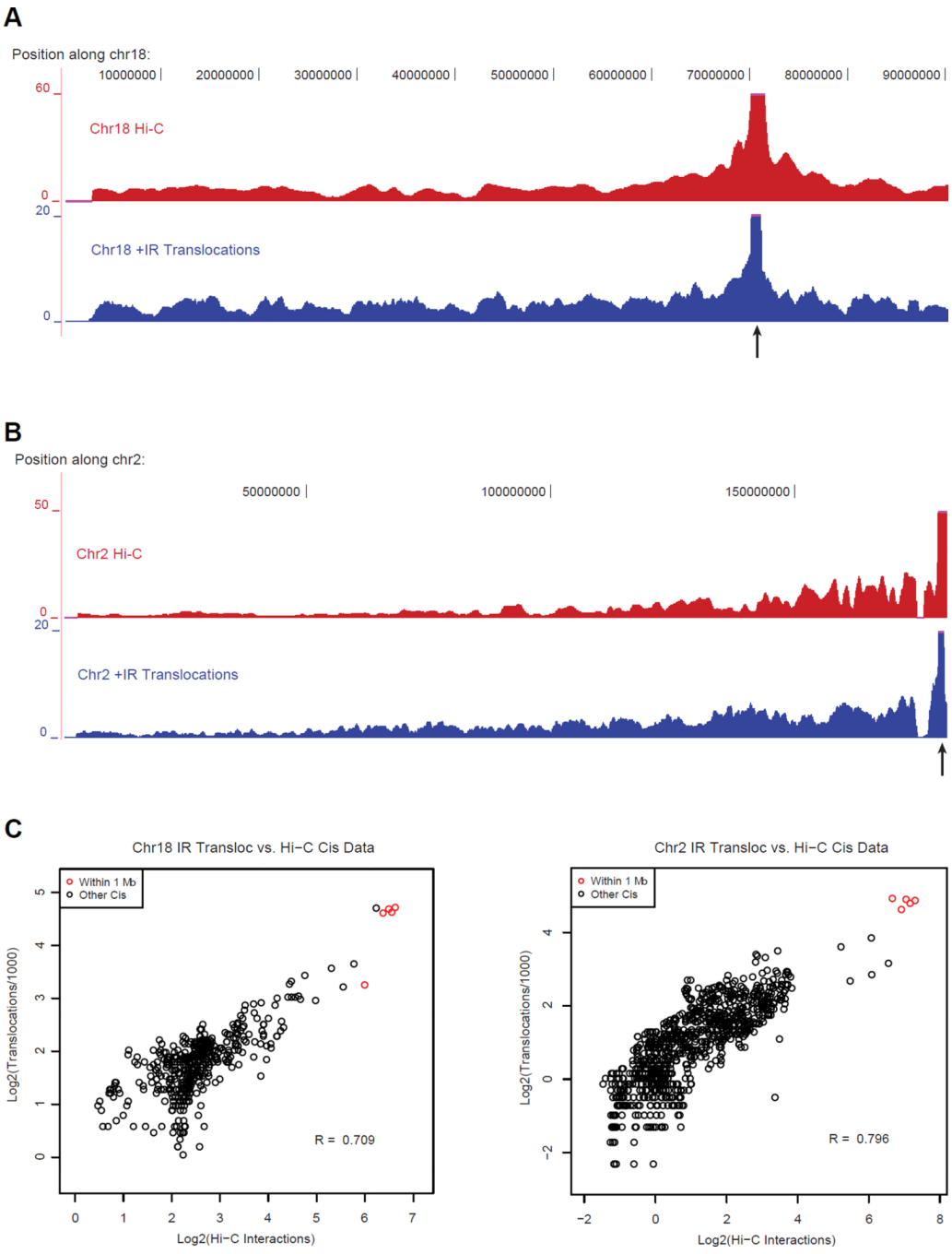

Figure 6. Spatial proximity correlates with translocation frequency along the chromosome containing the targeted I-SceI break

A) Top: Hi-C interactions between the 1 Mb bin containing the I-SceI site and other loci along the same chromosome (chr18; 1 Mb bins, 200 kb step). Bottom: Translocation frequency along chr18 after IR (1 Mb bins, 200 kb step). Arrow indicates I-SceI integration site. Some peak heights extend beyond the range shown (pink lines). **B)** Same graphs as A) for cells with the I-SceI site in chr2. **C)** Log-log plot of post-IR translocation frequency vs. Hi-C interaction frequency in each 1 Mb bin along chr18 for the chr18 I-SceI integration site (left) or chr2 for the chr2 I-SceI integration site (right). Translocation counts are normalized by the total number of translocations in the dataset, and Hi-C counts are corrected for

coverage as described in the Extended Experimental Procedures. Pearson correlations ("R") are shown. Points located within 1 Mb of the I-SceI site (red) are excluded from the correlation calculation. Similar results were obtained for the chr 7 integration site (Fig. S6).

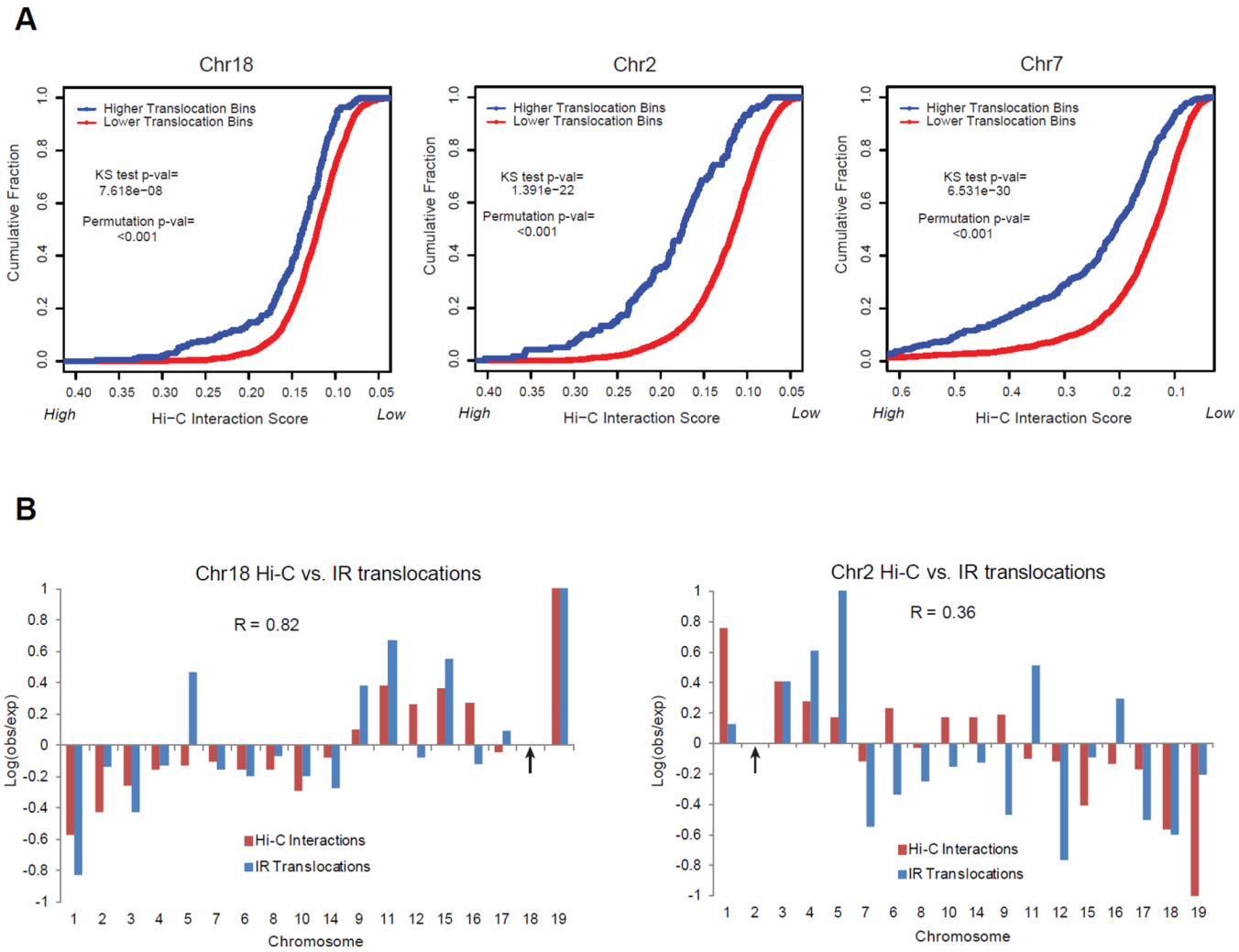

Figure 7. Translocations between chromosomes or sub-chromosomal regions are correlated with their relative spatial proximity

A) Cumulative distributions of Hi-C interaction frequencies are shown for *trans* chromosome 5 Mb bins that have either “high translocations” (blue; >=2 translocations/ 1,000 in dataset) or “low translocations” (red; < 2 translocations/1,000 in dataset) with the I-SceI site after irradiation. A 1 Mb fixed bin around the I-SceI site is used. Hi-C scores are displayed from high-to-low. The high translocation bins have significantly higher Hi-C scores than low translocation bins (one-tailed KS test); this difference is significant compared to 1,000 random permutations of the translocation dataset (“Permutation p-val”). This result is valid for a broad range of thresholds (Fig. S7A). **B)** Different I-SceI targeted break locations (indicated by arrows; left: chr18, right: chr2) display different whole chromosome translocation frequency profiles (blue) that correlate (R = Pearson correlation coefficient) with their different whole chromosome proximity profiles (red). Log ratios of observed/expected translocation frequencies and Hi-C interaction frequencies between whole chromosomes (calculated as in Fig 5E) are normalized to a maximum of 1. Chromosomes are sorted by length. All Ig and TCR hotspot loci are excluded from analyses in A) and B). See also Fig. S7.