Genome-wide mapping of long-range contacts unveils clustering of DNA double-strand breaks at damaged active genes

François Aymard1,6, Marion Aguirrebengoa1,6, Emmanuelle Guillou1,6, Biola M Javierre2, Beatrix Bugler1, Coline Arnould1, Vincent Rocher1, Jason S Iacovoni3, Anna Biernacka4, Magdalena Skrzypczak4, Krzysztof Ginalski4, Maga Rowicka5, Peter Fraser2 & Gaëlle Legube1

The ability of DNA double-strand breaks (DSBs) to cluster in mammalian cells has been a subject of intense debate in recent years. Here we used a high-throughput chromosome conformation capture assay (capture Hi-C) to investigate clustering of DSBs induced at defined loci in the human genome. The results unambiguously demonstrated that DSBs cluster, but only when they are induced within transcriptionally active genes. Clustering of damaged genes occurs primarily during the G1 cell-cycle phase and coincides with delayed repair. Moreover, DSB clustering depends on the MRN complex as well as the Formin 2 (FMN2) nuclear actin organizer and the linker of nuclear and cytoplasmic skeleton (LINC) complex, thus suggesting that active mechanisms promote clustering. This work reveals that, when damaged, active genes, compared with the rest of the genome, exhibit a distinctive behavior, remaining largely unrepaired and clustered in G1, and being repaired via homologous recombination in postreplicative cells.

Translocations, which occur when two DNA DSBs are abnormally rejoined, are highly deleterious genome rearrangements favoring cancer initiation and progression. However, the mechanisms that drive their formation remain poorly understood. One prerequisite for translocation is the juxtaposition of two distant DSBs, an event that would be favored if DSBs cluster, i.e. are brought together in spatial proximity within the nucleus. In budding yeast, a lacO- or TetO-tagged genomic locus exhibits increased motion when it is damaged with the I-SceI endonuclease. Intriguingly, this phenomenon, referred to as ‘local mobility’, is accompanied by a ‘global mobility’ process whereby the other chromosomes also explore wider volumes in the nuclear space, although to a lesser extent than the damaged locus itself1–3. Moreover, induction of two DSBs by the HO and I-SceI endonucleases triggers formation of a single Rad52 focus, thus suggesting that DSBs can indeed coalesce in yeast4.

However, whether DSBs cluster in higher eukaryotes has been a subject of controversy over the past decade, owing to conflicting results obtained through microscopy-based methods (refs. 5–13; reviewed in ref. 14). On the one hand, DSBs induced on a linear tract by α-particles, etoposide or high doses of radiation can move over long distances and cluster over time5–8. In addition, DSBs induced at subtelomeric regions in cells that use alternative lengthening of telomeres for telomere maintenance or that are induced by a restriction enzyme (AsiSI) undergo clustering15,16. Finally, DSB tracking in thousands of cells by using high-throughput time-lapse microscopy has revealed that two damaged loci can frequently be juxtaposed and lead to translocation17. However, in other studies, X-ray or heavy-ion-induced DSBs have not been found to exhibit this distinctive behavior, and dynamic analysis of a single broken chromosome end has shown only small-scale local motion9–11, thus arguing against clustering.

DSBs are repaired primarily by homologous recombination (HR) and nonhomologous end joining (NHEJ), both of which rely on profoundly different mechanisms. NHEJ directly reseals the two DNA ends in a process called resection, which is initiated by the MRN complex and is followed by nucleoprotein-fragment assembly, homology search, strand invasion and DNA synthesis (reviewed in ref. 18). Importantly this pathway is strongly suppressed during the G1 cell-cycle phase, most probably to minimize the use of an illegitimate copy as a template19. Beyond the tight control exerted by the cell-cycle stage on repair-pathway usage, a growing body of evidence suggests that repair is also regulated by chromatin and varies across the genome (reviewed in ref. 20). High-resolution genome-wide-scale profiles of RAD51 and XRCC4 (which are involved in HR and NHEJ, respectively), after induction of multiple sequence-specific DSBs by a restriction enzyme (AsiSI), have revealed that, in postreplicative cells, DSBs in intergenic regions or silent genes are

1LBCMCP, Centre de Biologie Intégrative (CBI), CNRS, Université de Toulouse, UT3, Toulouse, France. 2Nuclear Dynamics Programme, Babraham Institute, Cambridge, UK. 3Bioinformatics Platform I2M, INSERM and University of Toulouse, Toulouse, France. 4Laboratory of Bioinformatics and Systems Biology, Centre of New Technologies, University of Warsaw, Warsaw, Poland. 5Department of Biochemistry and Molecular Biology, University of Texas Medical Branch at Galveston, Galveston, Texas, USA. 6These authors contributed equally to this work. Correspondence should be addressed to G.L. (gaelle.legube@univ-tlse3.fr).

Received 2 September 2016; accepted 7 February 2017; published online 6 March 2017; doi:10.1038/nsmb.3387
repaired primarily by NHEJ, whereas those occurring in active genes are channeled to HR\textsuperscript{21}. Trimethylated K36 of histone H3 (H3K36me3), a histone mark previously linked to transcriptional elongation, contributes to target HR at transcriptionally active loci\textsuperscript{21–24}.

Because the initial chromatin context controls the repair-pathway choice and because the ability of a locus to roam in the nucleus has been proposed to be regulated by the chromatin-compaction state (e.g., in ref. 25), we hypothesized that the propensity of DSBs to cluster may also depend on the properties of the broken loci, i.e., their genomic positions. Here, to assess this hypothesis, we examined clustering among AsiSI-induced DSBs scattered throughout the genome and located in both transcriptionally active and inactive regions, by using a chromosome conformation capture (3C)-derivative technique (capture Hi-C\textsuperscript{26}). We found that DSB clustering occurs primarily among breaks induced in active genes, during G1, in a manner dependent on the nuclear actin organizer FMN2 and the LINC complex. Furthermore, clustering correlates with delayed repair in G1; hence, instead of promoting repair, clustering may contribute to ‘sequestration’ of persistent DSBs.

RESULTS

Capture Hi-C experimental design in DIvA cells

The DIvA (DSB inducible via AsiSI) human cell line allows for the induction of approximately 100 DSBs within euchromatin (because AsiSI activity is inhibited by DNA methylation) after treatment with 4-hydroxytamoxifen (4-OHT)\textsuperscript{21,27}. Using time-lapse microscopy, we have recently found that AsiSI-induced DSBs exhibit clustering in an ATM-dependent manner\textsuperscript{16}. However, the identities of DSBs able to coalesce, among the 100 breaks induced by AsiSI, were unknown. To identify those DSBs, we performed Hi-C followed by sequence capture and high-throughput sequencing (capture Hi-C)\textsuperscript{26} (Fig. 1a).

Capture probes were designed throughout γ-H2AX domains established after DSB induction in DIvA cells\textsuperscript{21} and surrounding the 100 most cleaved AsiSI sites \textit{in vivo} (details in Supplementary Note). Three control regions of equivalent size (2 Mb) but lacking DSBs were included. An average of 50 probes were designed per domain with a high coverage of the immediate vicinity of the DSB in a ±10-kb window (one probe per HindIII fragment, i.e., per 5 kb) and with lower

Figure 1 Capture Hi-C reveals DSB-induced local changes in chromosome folding. (a) Diagram of the experimental pipeline: DSBs are introduced at known positions throughout the human genome after 4-hydroxytamoxifen (4-OHT) treatment. Capture Hi-C experiments are conducted both before and after DSB induction to identify DSBs that interact within the nucleus after damage (as detailed in Supplementary Note). 100-kb-resolution contact maps determine changes in \textit{cis} interaction within γ-H2AX domains, whereas 2-Mb-resolution interaction maps allow for identification of domains capable of interacting (i.e., clustering) after damage. (b) 100-kb-resolution differential interaction map of damaged versus undamaged, shown for two domains (Dom) captured on chromosome 1. Data are expressed as the natural log of the differential interaction count (normalized to the number of captured reads and probe number, as outlined in Online Methods). Arrows indicate DSBR positions. The γ-H2AX profiles obtained by ChIP-seq at the same chromosomal locations\textsuperscript{21} after damage are shown below the maps. (c) Averaged interaction heat map between damaged versus undamaged cells at 100-kb resolution within 2-Mb captured domains, around the three control domains lacking DSBs (bottom). Adjusted P values were computed with Edge R, on the basis of capture Hi-C analyses from two independent cell cultures (as outlined in Online Methods). –log\textsubscript{10} (P) values are indicated, with negative fold changes (damaged < undamaged) in blue and positive fold changes (damaged > undamaged) in yellow.
coverage of the rest of the domain (one probe per ten HindIII fragments, i.e., per approximately 50 kb) (Supplementary Tables 1 and 2). Capture Hi-C was performed in undamaged or damaged cells, in two biological replicates from independent cell cultures. Ditag analyses behaved as expected, thus indicating that both 3C libraries in two biological replicates from independent cell cultures. Ditag and capture steps worked efficiently (Supplementary Note). Both replicates exhibited a strong correlation (Supplementary Fig. 1a–c). Both replicates exhibited a strong correlation (Supplementary Fig. 1d,e).

**Cis interactions are increased within γ-H2AX domains after damage**

To visualize the interaction data, differential contact matrices between untreated and treated samples were calculated at 100-kb resolution on captured regions. Interestingly, we observed changes in cis interactions within γ-H2AX domains after damage, in both independent biological replicates (Fig. 1b and Supplementary Fig. 2). On average, cis interactions were reproducibly enhanced around DSBs after damage, in contrast to control regions (Fig. 1c). An increased contact frequency was observed on both sides of the break, over a region spanning approximately 1 Mb (500 kb on each side), a size reminiscent of the average size of γ-H2AX domains. The interaction pattern, observed as a cross centered on the DSB (Fig. 1c), indicated that the regions surrounding the break frequently come into contact with the DSB and further supported our previous hypothesis that H2AX phosphorylation is sustained through direct contact of loci in cis with the DSB.

![Figure 2 Clustering occurs for DSBs induced on different chromosomes and correlates with γ-H2AX level.](image)

(a) Differential maps representing the differential interaction (damaged versus undamaged) between each 2-Mb domain captured on the human genome, computed from capture Hi-C experiments from two independent cell cultures (indicated as biological replicate (BR) 1 and BR 2). Data are expressed as the natural log of the number of differential interaction counts. DSBs are ordered according to their positions on the genome, and chromosomes (chr) are indicated by colored bars. The three control regions lacking DSBs are positioned at the end (black line). Each square depicts the differential interaction obtained between one DSB containing a captured domain and another (as described in Supplementary Note). (b) The numbers of interactions between domains were measured, and adjusted P values between damaged and undamaged samples were calculated on the basis of the Capture Hi-C data sets obtained from two independent cell cultures, by using Edge R (outlined in Online Methods). −log_{10} (P) values are indicated, with negative fold changes (FC <0, damaged < undamaged) in blue and positive fold changes (FC >0, damaged > undamaged) in yellow. DSBs are ordered according to their positions in the genome, and chromosomes are indicated by colored bars. The three control regions lacking DSBs are positioned at the end, as in a. (c) Circos plots showing the statistically significant (adjusted P <0.05) interactions induced after 4OHT treatment for two selected DSB-containing captured domains. Connecting lines are colored according to the log_{2} fold change between damaged and undamaged cells. Adjusted P values were calculated from damaged and undamaged samples from capture Hi-C experiments on two independent cell cultures by using EdgeR (Online Methods). (d) Interaction heat map between damaged versus undamaged cells at a 2-Mb domain resolution, sorted on the basis of γ-H2AX level. Adjusted P values between damaged and undamaged samples were calculated on the basis of a capture Hi-C data set obtained from two independent cell cultures, by using Edge R (outlined in Online Methods). −log_{10} (P) values are indicated, with negative fold changes (damaged < undamaged) in blue and positive fold changes (damaged > undamaged) in yellow. DSBs are sorted on the basis of their γ-H2AX level, as analyzed by ChIP-seq. Controls are indicated on the left side of the matrix. Dom, domain.
Clustering occurs between DSBs induced on the same or different chromosomes

To identify DSBs capable of interacting (i.e., clustering), we next created differential contact matrices (damaged versus undamaged) for all 2-Mb domain resolution for DSBs defined as HR prone or NHEJ prone21 (as detailed in Online Methods). P values were computed as described in Figure 2b. –log10 (P) values are indicated, with negative fold changes (damaged vs undamaged) in blue and positive fold changes (damaged > undamaged) in yellow. (b) Box plot showing the distribution of DSB-induced interactions between each HR-prone DSB with any other HR-prone DSB (HR/HR), or each NHEJ-prone DSB with any other NHEJ-prone DSB (NHEJ/NHEJ). The difference between the two classes is highly significant (two-sided Wilcoxon Mann–Whitney test, HR/HR, n = 506 pairs; NHEJ/NHEJ, n = 420 pairs). Center line, median; box limits, first and third quartiles; whiskers, minimum and maximum without outliers; points, outliers (Online Methods). (c) DSBs were classified according to their ability to cluster (ranked as high, medium, and low). RNA polymerase II occupancy (left) and H3K36me3 enrichment (right), analyzed in DiVA cells by ChiP–seq (ref. 21 and Online Methods) were averaged over the closest genes. Highly clustered DSBs also exhibited the highest RNA polymerase II and H3K36me3 levels (two-sided Wilcoxon Mann–Whitney test, low, n = 20 DSBs; medium, n = 17 DSBs; high, n = 20 DSBs). Center line, median; box limits, first and third quartiles; whiskers, minimum and maximum without outliers; points, outliers (Online Methods). Source data are available online.

Visualization of Hi-C data with circos plots clearly indicated that DSB-induced long-range interactions were not restricted to the same chromosome and that clustering also occurred between DSBs induced on different chromosomes (Fig. 2c and Supplementary Fig. 3b). In addition, DSBs clustering was favored between loci that were initially in proximity (Supplementary Fig. 3a–c), but DSBs induced on physically close domains (belonging to the same chromosome, Supplementary Fig. 3c, left) did not necessarily cluster, thus indicating that spatial proximity is probably required but is not sufficient for DSB clustering. Finally, when DSBs were sorted on the basis of damage-induced γ-H2AX levels obtained by chromatin immunoprecipitation (ChiP)–seq in DiVA cells23, the DSBs exhibiting higher γ-H2AX levels demonstrated a higher ability to cluster (Fig. 2d). Altogether, these data indicated that DSBs in the human genome can undergo clustering even when they are induced on different chromosomes, in a manner correlated with the level of γ-H2AX.

Clustering is enhanced between DSBs induced in transcriptionally active regions

In a previous study, we discovered that all DSBs induced in the genome are not necessarily repaired by HR in postreplicative cells, even though the sister chromatid is available. Indeed, on the basis of...
RAD51 and XRCC4 ChIP–seq data, we determined the repair pathway used at each AsiSI-induced DSB and found that in G2, DSBs induced in active gene bodies or promoters were repaired primarily by HR (hereafter designated as HR prone); in contrast, DSBs induced in silenced regions were unable to recruit RAD51 and instead underwent NHEJ (hereafter designated as NHEJ prone)21. We thus wondered whether DSBS clustering might depend on the pathway used for repair. Strikingly, sorting the captured domains on the basis of RAD51 ChIP–seq signal revealed that DSBs with high levels of RAD51 recruitment also showed statistically significantly increased clustering (Supplementary Fig. 4a). Enhanced clustering was observed at HR-prone DSBs compared with NHEJ-prone DSBs (Fig. 3a), and interactions among the HR-prone foci at damage were significantly higher than interactions among NHEJ-prone DSBs (Fig. 3b). Importantly, analysis of DSBs levels in DivA cells confirmed that the cleavage efficiency was similar in both categories (Supplementary Fig. 4b), thus indicating that the differential behavior of HR-prone and NHEJ-prone DSBs toward clustering was not due to DSB induction efficiency.

Because HR-prone DSBs occurred primarily within active genes, we further investigated the correlation between clustering ability and the transcriptional status of the damaged locus. RNA polymerase II occupancy and H3K36me3 enrichment (indicators of transcriptional activity) previously analyzed by ChIP–seq in DivA cells21 were calculated at DSB regions exhibiting low, medium or high clustering (Online Methods). DSBs that exhibited high clustering were located in regions with the highest transcriptional activity (Fig. 3c). Altogether, these data indicated that in human cells, DSBs are able to cluster, but this distinctive behavior is favored at breaks that are induced in transcriptionally active chromatin and are repaired by HR in postreplicative cells.

**Clustering is enhanced during the G1 cell-cycle phase**

The preferential occurrence of clustering at DSBs repaired by HR may indicate that clustering is somehow linked to homology search, a process required for HR and probably involving enhanced chromatin mobility2. However, previous studies have suggested that clustering is enhanced in the G1 phase of the cell cycle5, during which HR is strongly inhibited19. This apparent contradiction prompted us to investigate whether DSB clustering was also enhanced in G1 in DivA cells. Importantly, we have previously found by ChIP–microarray (ChIP–chip) analysis that both the number of DSBs induced by AsiSI and the extent of γ-H2AX spreading are identical between G1 and G2 (ref. 27). Hence, clustering can also be inferred by measuring the numbers and sizes of γ-H2AX foci detected within the nuclei (Fig. 4a). Larger foci were observed in G1 compared with G2 cells (Fig. 4b). To systematically measure the number and size of γ-H2AX foci in a large number of G1 and G2 damaged nuclei, we performed high-throughput microscopy (Online Methods). Cell-cycle stages were determined with Hoechst staining (Supplementary Fig. 5a). For each cell, the number of γ-H2AX foci was plotted against the average γ-H2AX foci detected within the nuclei (Fig. 4a), and interactions among γ-H2AX foci were statistically significantly increased clustering (Supplementary Fig. 4a). Enhanced clustering was observed at HR-prone DSBs compared with NHEJ-prone DSBs (Fig. 3a), and interactions among the HR-prone foci at damage were significantly higher than interactions among NHEJ-prone DSBs (Fig. 3b). Importantly, analysis of DSBs levels in DivA cells confirmed that the cleavage efficiency was similar in both categories (Supplementary Fig. 4b), thus indicating that the differential behavior of HR-prone and NHEJ-prone DSBs toward clustering was not due to DSB induction efficiency.

Because HR-prone DSBs occurred primarily within active genes, we further investigated the correlation between clustering ability and the transcriptional status of the damaged locus. RNA polymerase II occupancy and H3K36me3 enrichment (indicators of transcriptional activity) previously analyzed by ChIP–seq in DivA cells21 were calculated at DSB regions exhibiting low, medium or high clustering (Online Methods). DSBs that exhibited high clustering were located in regions with the highest transcriptional activity (Fig. 3c). Altogether, these data indicated that in human cells, DSBs are able to cluster, but this distinctive behavior is favored at breaks that are induced in transcriptionally active chromatin and are repaired by HR in postreplicative cells.
the number of foci was higher on average (Supplementary Fig. 5b)).

Altogether, these data indicated enhanced DSB clustering in G1.

Clustered, HR-prone DSBs exhibit delayed repair

Notably, long-range mobility of DSBs has often been associated with persisting breaks or delayed repair. For example, unrepairable DSBs in yeast move toward the nuclear periphery. Similarly, in higher eukaryotes, DSBs induced in heterochromatin or in ribosomal DNA also exhibit relocalization to the periphery of the heterochromatin focus or the nucleolus, respectively—two subnuclear compartments that have been suggested to be refractory to efficient repair. Therefore, the clustering of damaged active genes observed in G1 may reflect delayed repair. Hence, we measured repair kinetics at clustered and unclustered DSBs in G1-arrested cells compared with cycling cells. Briefly, we used a previously developed assay (Online Methods) that allows purification and qPCR-mediated quantification of DSB levels in damaged cells (4OHT) or at different time points during repair (after auxin addition), (Supplementary Fig. 5c and Supplementary Table 3). Importantly, lovastatin treatment performed to arrest the cells in G1 (Supplementary Fig. 5d) did not alter the auxin-induced degradation of AsiSI (Supplementary Fig. 5e). Remarkably, three selected DSBs that undergo clustering in G1 (Fig. 5a, right) exhibited delayed repair in G1-arrested cells compared with normally cycling cells, even 8 h after auxin addition. In contrast, unclustered DSBs were rapidly repaired in both cycling or G1-arrested cells (Fig. 5a, left).

To assess the repair kinetics in G1 on a global scale, we further performed direct in situ breaks labeling, enrichment on streptavidin and next-generation sequencing (BLESS), a technique developed to capture and identify DSB positions at the genome-wide level (Supplementary Fig. 6a,b). Analyses of BLESS data indicated that DSBs were induced in G1 after 4OHT addition and were partially repaired after 2 h of auxin treatment (example in Supplementary Fig. 6c; averaged signal at AsiSI sites in Supplementary Fig. 6d). We found that, whereas NHEJ-prone DSBs exhibited almost complete repair 2 h after auxin addition, HR-prone DSBs remained largely unrepaired (Fig. 5b). Altogether, these data indicated that DSBs induced during G1 in transcriptionally active genes are not fully proficient for repair and are prone to clustering.

MRN, FMN2 and the SUN2 component of the LINC complex sustain clustering

To gain insights into the mechanisms that might mediate clustering, we further used high-throughput microscopy combined with short interfering RNA (siRNA)-mediated depletion of selected candidates (Supplementary Fig. 7 and Supplementary Table 4). Previous studies have implicated the MRN complex in promoting DSB mobility and clustering and in favoring translocation. Interestingly, depletion of both MRE11 and NBS1 led to a decrease in the number of cluster-positive cells and in the clustering index (Fig. 6a,b and Supplementary Fig. 8), thus further confirming the role of the MRN complex in clustering. Depletion of known components of the DNA-damage response (RNF8,
and unprotected telomere mobility within the nucleus39. On the other hand, disruption of the LINC complex, which is embedded in the nuclear envelope, can influence DSB mobility and/or repair. On the one hand, actin organizer40. We found that both pathways contribute to DSB clustering after induction of DNA damage, in a manner dependent on the FMN2 protein.

Figure 8a supports this hypothesis, although it remains to be established whether MRN contributes to clustering through promoting resection.

Given the increased risk of translocation associated with DSB clustering, one may wonder why this potentially deleterious mechanism has been selected during evolution. Studies in yeast have suggested that DSB clustering promote ‘repair center’ assembly, thereby concentrating and favoring recruitment of repair machineries. However, we find here that DSB clustering coincided with delayed repair, a finding that does not support the idea of repair centers in human cells. Instead, we favor the hypothesis that DSB clustering sequesters and/or prepares DSBs for faithful repair. Indeed, using the sole available approach, we found here that DSB clustering coincided with delayed repair, a finding that does not support the idea of repair centers in human cells. Instead, we favor the hypothesis that DSB clustering sequesters and/or prepares DSBs for faithful repair. Indeed, using the sole available approach, we found here that DSB clustering coincided with delayed repair, a finding that does not support the idea of repair centers in human cells. Instead, we favor the hypothesis that DSB clustering sequesters and/or prepares DSBs for faithful repair.

DISCUSSION

The present analysis of long-range contacts after DSB induction enabled us to uncover a number of important features of the DNA-damage response. First, damage-induced changes within γ-H2AX domains revealed a cross-shaped pattern, thus indicating that regions located in the ±500-kb region surrounding the DSB frequently come into contact with the break. We have previously reported that ATM, the main H2AX kinase, is recruited to restricted regions (~10 kb), whereas γ-H2AX spans windows of over 1–2 Mb (ref. 16). In addition, in budding yeast, the ATM and ATR homologs Tel1 and Mec1 are able to phosphorylate H2A in trans41. Therefore, we have previously proposed that γ-H2AX spreading is likely to occur through the dynamics of the chromatin fiber around DSBs, thus bringing distant H2AX-containing nucleosomes into the physical proximity of ATM bound in the vicinity of DNA ends14,28. The observation that increased contacts between loci in cis and the break itself occurred in a region reminiscent of the γ-H2AX domain (1 Mb on average; Fig. 1c) lent further support to this hypothesis.

Moreover, we unambiguously established that DSBs can undergo clustering even between different chromosomes. Furthermore, we showed that the ability to cluster is not an intrinsic property of DSBs but instead depends on the nature of the damaged locus, because DSBs cluster only when they are induced in transcriptionally active genes. This observation probably accounts for the discrepancies previously reported in other studies, because various DNA-damaging agents do not necessarily induce DSBs at equivalent genomic locations. Notably, it has previously been shown that translocations occur primarily in active genes42–45. Although this phenomenon has been attributed to an increased susceptibility of active genes to breakage, our results suggest that increased clustering of damaged active genes may also contribute to the translocation pattern.

Importantly, our study also revealed that DSBs induced in active genes, although they are repaired by HR in postreplicative cells, are refractory to repair in G1. We propose that these persisting breaks display large-scale motion and further coalesce within foci, as suggested by the observed enhanced clustering in G1 compared with G2 (Fig. 7).

Interestingly, this behavior is reminiscent of that of other large-scale DSB-mobility events that have been almost systematically associated with persistent or ‘difficult’ DSBs30,32,35,46 (reviewed in ref. 14). The reasons underlying repair deficiency at active genes in G1 remain unknown. Although clustering was altered by DRB-mediated transcriptional inhibition (Supplementary Fig. 8), it was not decreased after treatment with other inhibitors (such as triptolide; data not shown). This finding suggests that the delayed repair observed at active genes should be attributed to chromatin and secondary DNA structures assembled at transcribed regions rather than to transcription itself. RNA polymerase II–machinery occupancy, supercoiling or RNA–DNA hybrids that form at transcribing loci47 may delay resolving of broken ends, thus resulting in minimal resection, which would subsequently impede classical NEJ. In postreplicative cells, such processed ends would be further processed by the HR machinery. In contrast, in G1, these DSBs would persist, owing to the downregulation of HR activity (Fig. 7). The involvement of MRN, whose activity sustains the initiation of resection, in clustering (Fig. 6a,b) further supports this hypothesis, although it remains to be established whether MRN contributes to clustering through promoting resection.



Figure 7 Model for DSB clustering. Whereas DSBs in intergenic regions and inactive genes are repaired primarily by NHEJ throughout the cell cycle, the repair of DSBs induced in active genes is dependent on the cell-cycle phase. In postreplicative cells, DSBs in transcribed regions are prone to HR repair, whereas in G1, those breaks persist longer and coalesce with one another, thus forming clusters of unrepaired active genes. Txn, transcription.

53BP1 and XRCC4 did not significantly alter clustering (Supplementary Fig. 8a,b). In addition, two recent studies have suggested roles for cytoskeleton proteins in DSB mobility and/or repair. On the one hand, disruption of the LINC complex, which is embedded in the nuclear envelope that connects the cytoskeleton to the nucleoskeleton, decreases DSB and unprotected telomere mobility within the nucleus39. On the other hand, actin has been found to polymerize within the nucleus shortly after induction of DNA damage, in a manner dependent on the FMN2 actin organizer40. We found that both pathways contribute to DSB clustering. Indeed, depletion of SUN2 (but not of SUN1) and FMN2 led to a decrease in cluster-positive cells and in the clustering index (Fig. 6c,d and Supplementary Fig. 8a,b). These data indicated that DSBs cluster in G1 in a manner dependent on the MRN complex and also on microtubule- and actin-related networks, thus suggesting that active mechanisms may be at work in this process. Finally, because clustering primarily involves DSBs induced in transcriptionally active regions, we investigated whether transcription inhibition would affect clustering. A pretreatment with 5,6-dichloro-1-β-D-ribofuranosylbenzimidazole (DRB), a well-known transcription inhibitor, indeed led to decreased clustering (Supplementary Fig. 8c,d). However, such a decrease in clustering was not observed with another transcription inhibitor (tripotolide), thus suggesting that transcription per se is not responsible for the induced clustering of damaged active genes (as discussed further below).
A mechanism for the suppression of homologous recombination at DNA double-strand breaks.

Clouaire, T. & Legube, G. Positional stability of damaged chromatin domains along radiation tracks in mammalian cells. Radiat. Res. 171, 405–418 (2009).

Kruhlik, M.J. et al. Changes in chromatin structure and mobility in living cells at sites of DNA double-strand breaks. Curr. Biol. 17, 823–834 (2006).

Roukos, V. et al. Spatial dynamics of chromosome translocations in living cells. Science 341, 660–664 (2013).

Mladenov, E., Magin, S., Soni, A. & Hilakos, G. DNA double-strand-break repair in higher eukaryotes and its role in genomic instability and cancer: cell cycle and proliferation-dependent regulation. Semin. Cancer Biol. 37-38, 51–64 (2016).

Orthwein, A. et al. A mechanism for the suppression of homologous recombination in G1 cells. Nature 528, 422–426 (2015).

Clouaire, T. & Legube, G. DNA double strand break repair pathway choice: a chromatin based decision? Nucleus 6, 107–113 (2015).

Ayward, F. et al. Transcriptionally active chromatin recruits homologous recombination at DNA double-strand breaks. Nat. Struct. Mol. Biol. 21, 366–374 (2014).

Pflister, S.X. et al. SETD2-dependent histone H3K36 trimethylation is required for homologous recombination repair and genome stability. Cell Rep. 7, 2006–2018 (2014).

Carvalho, S. et al. SETD2 is required for DNA double-strand break repair and activation of the p53-mediated checkpoint. eLife 3, e02482 (2014).

Daugschild, M. et al. LEDGF/p75 promotes DNA-end resection and homologous recombination. Nat. Struct. Mol. Biol. 19, 803–810 (2012).

Hajjoul, H. et al. High-throughput chromatin motion tracking in living yeast reveals the flexibility of the fiber throughout the genome. Genome Res. 23, 1829–1838 (2013).

Schoenfelder, S. et al. The pluripotent regulatory circuitry connecting promoters to their long-range interacting elements. Genome Res. 25, 582–597 (2015).

Iacovoni, J.S. et al. High-resolution profiling of gammah2ax around DNA double-strand breaks in the mammalian genome. EMBO J. 29, 1446–1457 (2010).

Ayward, F. & Legube, G. A TAD closer to AT. Mol. Cell. Oncol. 3, e134411 (2016).

Bozler, A. et al. Three-dimensional maps of all chromosomes in human male fibroblast nuclei and prometaphase rosettes. PLoS Biol. 3, e157 (2005).

Nagi, S. et al. Functional targeting of DNA damage to a nuclear pore-associated SUMO-dependent ubiquitin ligase. Science 322, 597–602 (2008).

Oza, P., Jaspersen, S.L., Miele, A., Dekker, J. & Peterson, C.L. Mechanisms that regulate localization of a DNA double-strand break to the nuclear periphery. Genes Dev. 23, 912–927 (2009).
32. Chiolo, I. et al. Double-strand breaks in heterochromatin move outside of a dynamic HP1α domain to complete recombinational repair. Cell 144, 732–744 (2011).
33. Jakob, B. et al. DNA double-strand breaks in heterochromatin elicit fast repair protein recruitment, histone H2AX phosphorylation and relocation to euchromatin. Nucleic Acids Res. 39, 6489–6499 (2011).
34. Harding, S.M., Boiarsky, J.A. & Greenberg, R.A. ATM dependent silencing links nucleolar chromatin reorganization to DNA damage recognition. Cell Rep. 13, 251–259 (2015).
35. van Sluis, M. & McStay, B. A localized nucleolar DNA damage response facilitates recruitment of the homology-directed repair machinery independent of cell cycle stage. Genes Dev. 29, 1151–1163 (2015).
36. Chailleux, C. et al. Quantifying DNA double-strand breaks induced by site-specific endonucleases in living cells by ligation-mediated purification. Nat. Protoc. 9, 517–528 (2014).
37. Crosseto, N. et al. Nucleotide-resolution DNA double-strand break mapping by next-generation sequencing. Nat. Methods 10, 361–365 (2013).
38. Mitra, A., Skrzypczak, M., Ginalski, K. & Rowicka, M. Strategies for achieving high sequencing accuracy for low diversity samples and avoiding sample bleeding using illumina platform. PLoS One 10, e0120520 (2015).
39. Lottersberger, F., Karssemeijer, R.A., Dimitrova, N. & de Lange, T. 53BP1 and the LINC complex promote microtubule-dependent DSB mobility and DNA repair. Cell 163, 880–893 (2015).
40. Belin, B.J., Lee, T. & Mullins, R.D. DNA damage induces nuclear actin filament assembly by Formin-2 and Spire-1/2 that promotes efficient DNA repair. eLife 4, e07735 (2015).
41. Lee, C.S., Lee, K., Legube, G. & Haber, J.E. Dynamics of yeast histone H2A and H2B phosphorylation in response to a double-strand break. Nat. Struct. Mol. Biol. 21, 103–109 (2014).
42. Chiarle, R. et al. Genome-wide translocation sequencing reveals mechanisms of chromosome breaks and rearrangements in B cells. Cell 147, 107–119 (2011).
43. Klein, I.A. et al. Translocation-capture sequencing reveals the extent and nature of chromosomal rearrangements in B lymphocytes. Cell 147, 95–106 (2011).
44. Wei, P.C. et al. Long neural genes harbor recurrent DNA break clusters in neural stem/progenitor cells. Cell 164, 644–655 (2016).
45. Schwer, B. et al. Transcription-associated processes cause DNA double-strand breaks and translocations in neural stem/progenitor cells. Proc. Natl. Acad. Sci. USA 113, 2258–2263 (2016).
46. Tsouroula, K. et al. Temporal and spatial uncoupling of DNA double strand break repair pathways within mammalian heterochromatin. Mol. Cell 63, 293–305 (2016).
47. Ginno, P.A., Lim, Y.W., Lott, P.L., Korf, I. & Chêdin, F. GC skew at the 5’ and 3’ ends of human genes links R-loop formation to epigenetic regulation and transcription termination. Genome Res. 23, 1590–1600 (2013).
48. Harrigan, J.A. et al. Replication stress induces 53BP1-containing OPT domains in G1 cells. J. Cell Biol. 193, 97–108 (2011).
49. Lukas, C. et al. 53BP1 nuclear bodies form around DNA lesions generated by mitotic transmission of chromosomes under replication stress. Nat. Cell Biol. 13, 243–253 (2011).
50. Durkin, S.G. & Glover, T.W. Chromosome fragile sites. Annu. Rev. Genet. 41, 169–192 (2007).
51. Le Tallec, B. et al. Common fragile site profiling in epithelial and erythroid cells reveals that most recurrent cancer deletions lie in fragile sites hosting large genes. Cell Rep. 4, 420–428 (2013).
52. Kalocsay, M., Hiller, N.J. & Jentsch, S. Chromosome-wide Rad51 spreading and SUMO-H2A.Z-dependent chromosome fixation in response to a persistent DNA double-strand break. Mol. Cell 33, 335–343 (2009).
53. Ryu, T. et al. Heterochromatic breaks move to the nuclear periphery to continue recombinational repair. Nat. Cell Biol. 17, 1401–1411 (2015).
54. Matsuoka, S. et al. ATM and ATR substrate analysis reveals extensive protein networks responsive to DNA damage. Science 316, 1160–1166 (2007).
55. Yamada, K., Ono, M., Perkins, N.D., Rocha, S. & Lamond, A.I. Identification and functional characterization of FMN2, a regulator of the cyclin-dependent kinase inhibitor p21. Mol. Cell 49, 922–933 (2013).
ONLINE METHODS
Biotinylated-RNA bait library design. A genomic restriction map was created to generate a list of all HindIII fragments in the genome, and this list was subsequently filtered against 2-Mb regions around the 100 best cleaved AssiSI sites\textsuperscript{21} and three control regions of 2 Mb lacking AssiSI recognition sites (chr 1, 185500000–185750000; chr 17, 70000000–72000000; and chr 5, 111000000–113000000). Probes were designed on every HindIII fragment in a ±10-kb region around the AssiSI sites or around the centers of control regions. When two windows of 20 kb overlapped, each probe on the overlapping region was synthesized only once. Over the rest of the 2-Mb window, baits were designed every ten HindIII fragments, i.e. ~50 kb. When two 2-Mb windows overlapped, thus leading to a HindIII fragment being captured twice, the fragment just downstream was selected, to capture each fragment only once. For each selected restriction fragment, two 120-bp capture probes were designed, one for each end of the fragment. Because of the size-selection step after sonication, the probes had to fall entirely within a region no more than 500 bp from the end of the fragment. Each probe was required to have no more than three consecutive bases masked by repeat masker, and they had to have a GC content of 25% to 65%, to match the efficient capture range of the SureSelect target-enrichment system (Agilent Technologies). When multiple probes passed these criteria, the probe nearest the end of the restriction fragment was chosen. The full list of HindIII sites near which RNA capture probes were designed is shown in Supplementary Table 1. Coordinates and names of captured domains as well as the sites of the gaps between captured regions are shown in Supplementary Table 2.

Capture Hi-C. Generation of a capture Hi-C library was carried out as previously described\textsuperscript{56}. Briefly, 30 × 10\textsuperscript{6} cells were fixed in 2% formaldehyde for 10 min. After overnight digestion with HindIII, DNA ends were labeled with biotin-14–dATP by using the Klenow fragment of DNA polymerase I and were religated in preserved nuclei. 50 µg of DNA was sheared to an average size of 450 bp (Covaris). Sheared DNA was then end-repaired, adenine-tailed and size-selected twice from 250 to 600 bp (AMPure XP beads). Chimeras marked by biotin were pulled down with MyOne Streptavidin C1 DynaBeads (Invitrogen) and ligated to PE adaptors (Illumina). The Hi-C libraries (retained on the beads) were amplified with PE primers (Illumina) with ten PCR amplification cycles. Capture was next carried out on the Hi-C library, with SureSelect target enrichment, with the custom-designed biotinylated-RNA bait library and custom paired-end blockers, according to the manufacturer’s instructions (Agilent Technologies). After library enrichment by capture, a postcapture PCR amplification step was carried out with PE PCR 1.0 and PE PCR 2.0 primers, with seven PCR amplification cycles. Capture Hi-C libraries were performed for damaged (4OHT treated) or undamaged (untreated) samples in biological duplicates (i.e., independent experiments) (BR 1 and BR 2) and were sequenced on an Illumina HiSeq 1000 platform. High-throughput sequencing data have been deposited in Array Express under accession number E-MTAB-4846.

Sample sequencing, mapping and quality control. Samples were sequenced with a HiSeq 1000 platform, thus generating 51-bp paired-end reads, at the Babraham Institute in Cambridge. The quality of each raw sequencing file (fastq) was validated with FastQC. After quality control, fastq files were aligned against the reference human genome hg19 with the HiCUP pipeline\textsuperscript{57}, this pipeline identifies a HiC junctions (hicup\_truncator) in the fastq file, generates a paired-end read alignment file in .BAM format (hicup\_mapper) by mapping each read pair corresponding to a Hi-C ditag (purified Hi-C fragment) and filters out artificial read pairs (hicup\_filter and hicup\_deduplicator). The pipeline also provides a report summarizing quality control during the pipeline, including the number of too-short and truncated reads, number of unmapped reads, number of filtered wrong pairs (invalid, same fragment, replication), ditag length and duplication summary (Supplementary Fig. 1b). Capture efficiency was shown to be efficient and reproducible for each sample (Supplementary Fig. 1c,d).

Interaction counts. Seqmonk (http://www.bioinformatics.babraham.ac.uk/projects/seqmonk/) was used to visualize aligned bam files and to compute the number of interactions among our regions of interest. Briefly, the number of interactions was calculated over the 2-Mb windows, where probes were designed, and around the 100 best cleaved AssiSI sites and three control regions. When sites were less than 1 Mb apart, their 2-Mb regions were merged, thus leading to 88 domains (85 sites and 3 controls). The number of ditags corresponding to any possible pair of domains was calculated with Seqmonk, and interaction reports containing the 3,228 pairs of domains were used for further visualization and analysis. When indicated (Fig. 1b and Supplementary Fig. 2), each domain was subdivided into 100-kb windows, thus leading to a Seqmonk interaction report containing the number of interactions among 1,574,425 pairs of subdomains.

R analysis and normalization. Each count matrix was then loaded in R software for further analysis. Manipulation and analyses were performed with the Bioconductor GenomicRanges R package\textsuperscript{58} available from the Bioconductor software project (http://www.bioconductor.org/). To compare different samples, the number of interactions was normalized to the total number of aligned read pairs captured on at least one end. The number of interactions between 2-Mb or 100-kb domain pairs was further normalized against the total number of probes used to capture those specific windows, to account for the variation in probe density across each domain. In other words, for each pair of domains (100 kb or 2 Mb), the number of interactions was normalized against the sum of probes designed in each of the two domains.

Differential heat maps of damaged versus undamaged cells. The numbers of interactions for each pair of domains (2 Mb or 100 kb) measured in the +4OHT (damaged) and in −4OHT (undamaged) samples were independently subtracted for both pairs of replicate samples. Differential heat maps were drawn with a natural log (ln) scale (in absolute values, then multiplied by direction of change), as follows: sign (+4OHT − −4OHT) × ln[+4OHT − −4OHT]. When indicated, domains were ordered on the basis of RAD51 or γ-H2AX levels (averaged over ±2 kb or ±40 kb, respectively), as analyzed by ChIP–seq\textsuperscript{21}.

Statistical heat maps. To evaluate the significance of the differences observed between +4OHT and −4OHT, P-values were computed for each pair of domains, taking into account both replicates, with EdgeR (Empirical analysis of Digital Gene Expression data in R) available through the Bioconductor EdgeR R package\textsuperscript{59}. For a symmetrical visualization taking into account the direction of the fold change (in which up and down represent enhanced and decreased interaction levels between samples), P values were next transformed as follows: plotted value = sign(log fold change) × −log\textsubscript{10} (P value), and represented as heat maps. Significant negative changes (decreased interaction after damage) are plotted in blue, whereas significant positive changes (increased interaction after damage) are plotted in yellow. Plotted values were considered statistically significant when the values were >1.3 and <−1.3 (P value < 0.05).

Determination of HR-prone and NHEJ-prone DSBs. HR/NHEJ categories were computed on the basis of BLESS data in DlvA cells (T. Clouaire, M.A., V.R., A.B., M.S. et al., unpublished data) and ChIP–seq against XRCC4 and Rad51 (ref. 21). Sites were ordered on the basis of their average level of BLESS in a window of 1 kb centered on the AssiSI site. The 80 most cleaved sites were selected. A ratio between the level of RAD51 in a 4-kb window and the level of XRCC4 in a window of 500 bp around each DSB was computed (as performed in ref. 21). We deemed the 30 sites with the highest ratio to be HR-prone sites and the 30 sites with the lowest ratio to be NHEJ-prone sites.

Average interaction maps. For each 2-Mb domain, vectors containing the number of interactions between pairs of 100-kb windows were generated for each replicate damaged and undamaged sample. A mean vector was computed on the basis of all vectors for all 100 DSBs and the three control regions. EdgeR was then used on these mean vectors to identify significant differences in the number of interactions between pairs of 100-kb windows. The results are shown as heat maps, after P-value transformation, as described above.

Circos plots. Circos plots represent interactions between a captured domain and all other domains in the genome. Only significantly increased interactions (EdgeR P value <0.05; log fold change > 0) between undamaged and damaged conditions were drawn with the Bioconductor OmiCircos R package\textsuperscript{60}. Connecting lines are colored on the basis of the log\textsubscript{2} fold change.

Correlation with RNA pol II and H3K36me3 occupancy. The domains were separated into three equal categories, on the basis of the number of differential
interactions of each domain with the others (low, medium or high clustering ability). Transcriptional activity around AsiSI sites was determined with PolII-S2P and H3K36me3 ChIP–seq average levels from the closest gene (from −500 bp from the TSS to the TTS) neighboring each AsiSI site.

Cell lines and cell culture. DlVA (AsiSI-ER-U20S) and AID-DlVA (AID-AsiSI-ER-U20S) cells were previously developed with U2-OS from ATCC. They were cultured in DMEM supplemented with antibiotics, 10% FCS (Invitrogen) and either 1 µg/mL puromycin (DlVA cells) or 800 µg/mL G418 (AID-DlVA cells) at 37 °C under a humidified atmosphere with 5% CO2. The cell lines were regularly tested for mycoplasma contamination.

Treatment and siRNA transfection. For AsiSI-dependent DSB induction, cells were treated with 300 nM 4OHT (Sigma, H7904) for 4 h. For transcriptional inhibition, cells were pretreated with DRB (20 µM or 100 µM, as indicated) for 1 h, which was left on the cells during a subsequent 4-h treatment with 4OHT (Supplementary Fig. 8d). For cell synchronization in G1, cells were incubated with 2 mM thymidine for 16 h, released for 4 h and subjected to the second thymidine treatment for 19 h. G1 and G2 cells were treated with 4OHT after 11 h and 4 h, respectively, and fixed 4 h later (Fig. 4b). For BLESS experiments (Fig. 5b), cells were treated with 4OHT (4 h) 12 h after release, then subjected to a 2-h auxin treatment before harvesting. For cell arrest in G1 (Fig. 5a), cells were treated with 40 µM lovastatin for 48 h. siRNA transfections were performed with a Cell Line Nucleofector kit V (Amaxa) according to the manufacturer’s instructions or with interferin (PolyPlus). Sequences of siRNAs are shown in Supplementary Table 3.

High-throughput microscopy. DlVA cells were plated in 96-well plates (2,500 cells per plate) after transfection with siRNAs with a Nucleofector Kit V (Amaxa). 48 h later, cells were treated with 4OHT (300 nM) for 4 h. After fixation, permeabilization and saturation steps, γH2AX was stained overnight with a mouse primary and anti–γH2AX monoclonal antibody (Millipore, BW301), validated on the manufacturer’s website) and the secondary antibody anti-mouse Alexa 647 (Invitrogen) and ligated to a distal linker (Sigma). The resulting DNA fragments were then linearized by I-SceI (NEB) digestion and amplified by PCR. Libraries were prepared with a TruSeq DNA LT Sample Prep Kit (Illumina), then subjected to 2 × 61-bp and 2 × 70-bp next-generation sequencing on an Illumina HiSeq 2500. Raw data were filtered with a BLESS adaptor to keep only the end proximal to the cut from the paired read. These proximal reads were aligned with bwa-ali, and PCR duplicates were cleaned with samtools. The coverage and the normalization by the total number of reads was then generated with R. For each site, we computed the average level of normalized counts in a window of ±500 bp around sites.

Immunofluorescence. Cells grown on glass coverslips were fixed with 4% paraformaldehyde, and immunofluorescence was carried out according to a standard protocol (described in ref. 27). A mouse monoclonal antibody (Upstate Biotechnology, JBW 301) was used to stain γ-H2AX (antibody validation provided on the manufacturer’s website). The secondary antibody was purchased from Molecular Probes (cat. no. A11032). Image acquisition was performed with MetaMorph on a wide-field microscope equipped with a cooled charge-coupled-device camera (CoolSNAP HQ2) with a 100× objective. Deconvolution was performed with the Richardson–Lucy algorithm with total variation (TV) regularization3. The algorithm was implemented in the image J deconvolution lab software (http://bigwww.epfl.ch/algorithms/deconvolutionlab/). For Supplementary Figure 3a, the 3D coordinates and the volume of each focus was obtained with eucDist = \( (x_{\text{focus}} - x_{\text{center}})^2 + (y_{\text{focus}} - y_{\text{center}})^2 + (z_{\text{focus}} - z_{\text{center}})^2 \) and correlated to the volume (x_focus × y_focus × z_focus) of each focus. Nineteen cells, from three independent experiments, were analyzed.

Repair assay. DlVA cells were treated with 4OHT for 4 h to induce damage and were then washed three times in prewarmed PBS and further incubated with 500 µg/mL auxin (Sigma, S1548) to allow repair. Cells were collected by trypsinization at different times after auxin treatment (0 h, 2, 8 h, 14 h), and their DNA was then extracted for use in repair assays. The full procedure for the repair assay has previously been described11,16. Briefly, DNA was extracted, ligated to a biotinylated double-stranded oligonucleotide, cohesive with AsiSI sites, and then fragmented by EcoRI digestion at 37 °C for 2 h and 4 h. DNA was pulled down with streptavidin beads (Sigma) at 4 °C overnight. Beads were resuspended in 100 µL of water and digested with HindIII at 37 °C for 4 h. After phenol/chloroform purification and precipitation, DNA was resuspended in 100 µL water. The proportion of unrepaired DSBs was then quantified by qPCR (SYBR premix, TaKaRa) in triplicate (primer sequences are provided in Supplementary Table 3).

Direct in situ breaks labeling, enrichment on streptavidin and next-generation sequencing (BESS). After treatment with 4OHT and or auxin in G1 synchronized cells, DSB labeling was carried out with the BLESS method, as previously described17. Briefly, cells were fixed with formaldehyde, lysed and mildly digested with proteinase K at 37 °C to purify intact nuclei, and then DSBs were blunted and 5'-phosphorylated with a Quick Blunting Kit (NEB). Subsequently, a biotinylated proximal linker (Sigma) was ligated to DSBs with T4 ligase (NEB). Next, DNA was extracted by precipitation with isopropanol and was sonicated with a Covaris S220 ultrasonicator to create fragments approximately 400 bp long. Labeled, biotinylated fragments were captured on streptavidin beads (Invitrogen) and ligated to a distal linker (Sigma). The resulting DNA fragments were then linearized by I-SceI (NEB) digestion and amplified by PCR. Libraries were prepared with a TruSeq DNA LT Sample Prep Kit (Illumina), then subjected to 2 × 61-bp and 2 × 70-bp next-generation sequencing on an Illumina HiSeq 2500. Raw data were filtered with a BLESS adaptor to keep only the end proximal to the cut from the paired read. These proximal reads were aligned with bwa-ali, and PCR duplicates were cleaned with samtools. The coverage and the normalization by the total number of reads was then generated with R. For each site, we computed the average level of normalized counts in a window of ±500 bp around sites.

Statistics. For all box-plot representations, outliers were defined as first quartile – 1.5 × (interquartile range) below the first quartile (Q1) or above the third quartile (Q3) in a data set, were excluded from graphical representations and further analysis. The number of nuclei in the top-left gate (cluster-negative) and the bottom-right gate (cluster-positive cells) were then counted, and the proportion of nuclei in these two gates for each condition is indicated in the scatter plot for representative experiments (Figs. 4c and 6a–d and Supplementary Fig. 8). The proportion of cluster-positive cells was calculated for at least three biological replicates in each condition, and the average was plotted together with the s.e.m. (Figs. 4c and 6a–d and Supplementary Fig. 8).

© 2017 Nature America, Inc., part of Springer Nature. All rights reserved.
the first to third quartiles, whiskers represent the maximum and minimum values without outliers, and points show outliers.

For each figure, statistical hypothesis tests were chosen according to the data distribution. For interaction counts of Hi-C data (Figs. 1c, 2b, d and 3a and Supplementary Fig. 4a), the EdgeR method was applied to compute P values for the number of Hi-C interactions for each pair of domains. Indeed, EdgeR implements statistical methodology based on the negative binomial distributions, which can be applied to differential signal analysis of genomic data that produce counts. The Fisher test (var.test function in R) was also applied to verify the homoscedasticity within groups: BR1 versus BR2 −4OHT, \( P = 0.745 \); BR1 versus BR2 +4OHT, \( P = 0.969 \); between-group BR1 −4OHT versus BR1 +4OHT, \( P = 0.694 \); BR1 −4OHT versus BR2 +4OHT, \( P = 0.666 \); BR2 −4OHT versus BR1 +4OHT, \( P = 0.946 \); and BR2 −4OHT versus BR2 +4OHT, \( P = 0.915 \). Thus, the variances were homogeneous within and among groups. For Figures 3b, c and 5b and Supplementary Figures 2d and 6d, Shapiro–Wilk tests (shapiro.test function in R) were applied to each sample to verify whether samples came from a normally distributed population. Because the hypothesis that data are normally distributed was rejected with more than 95% confidence, nonparametric unpaired Mann–Whitney–Wilcoxon tests (also called Wilcoxon rank-sum tests; wilcoxon.test function in R) were applied to test distribution differences between two samples.

**Data availability.** High-throughput sequencing data have been deposited in Array Express under accession number E-MTAB-4846. Source data for Figures 3b, c, 4c, 5a, b and 6a–d are available with the paper online. Other data and source codes are available upon request.

56. Nagano, T. et al. Comparison of Hi-C results using in-solution versus in-nucleus ligation. *Genome Biol.* **16**, 175 (2015).
57. Wingett, S. et al. HiCUP: pipeline for mapping and processing Hi-C data. *F1000Res.* **4**, 1310 (2015).
58. Lawrence, M. et al. Software for computing and annotating genomic ranges. *PLoS Comput. Biol.* **9**, e1003118 (2013).
59. Robinson, M.D., McCarthy, D.J. & Smyth, G.K. edgeR: a Bioconductor package for differential expression analysis of digital gene expression data. *Bioinformatics* **26**, 139–140 (2010).
60. Hu, Y. et al. OmicCircos: a simple-to-use R Package for the circular visualization of multidimensional omics data. *Cancer Inform.* **13**, 13–20 (2014).
61. Dey, N. et al. Richardson-Lucy algorithm with total variation regularization for 3D confocal microscope deconvolution. *Microsc. Res. Tech.* **69**, 260–266 (2006).