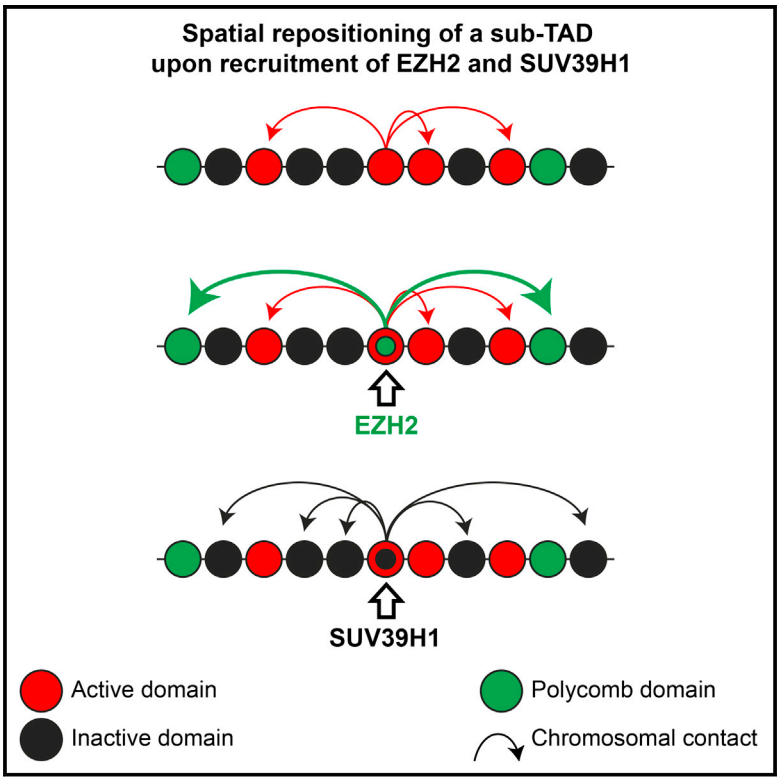
Article

Molecular Cell

Cause and Consequence of Tethering a SubTAD to Different Nuclear Compartments

Graphical Abstract



Authors

Patrick J. Wijchers, Peter H.L. Krijger, Geert Geeven, ..., Lisette C.M. Anink-Groenen, Pernette J. Verschure, Wouter de Laat

Correspondence

w.delaat@hubrecht.eu

In Brief

Wijchers et al. combine 4C technology with *lacO*/lacR binding platforms to show that a given genomic locus can adopt multiple nuclear positions depending on the recruited protein. Focal protein recruitment can move an entire subTAD, but repositioning appeared largely unrelated to changes in its chromatin composition and activity.

Highlights

- 4C technology reveals nuclear repositioning of genomic loci upon protein recruitment
- Given subTAD can adopt multiple nuclear positions depending on the recruited factor
- Switching between nuclear compartments can be uncoupled from transcriptional changes
- Histone-modifying SUV39H1 relies on its chromodomain for locus repositioning

Accession Numbers

GSE76174

Wijchers et al., 2016, Molecular Cell *61*, 461–473
February 4, 2016
http://dx.doi.org/10.1016/j.molcel.2016.01.001

CellPress

Molecular Cell
Article

Cause and Consequence of Tethering a SubTAD to Different Nuclear Compartments

Patrick J. Wijchers,[1,3] Peter H.L. Krijger,[1,3] Geert Geeven,[1,3] Yun Zhu,[1,3] Annette Denker,[1] Marjon J.A.M. Verstegen,[1] Christian Valdes-Quezada,[1] Carlo Vermeulen,[1] Mark Janssen,[1] Hans Teunissen,[1] Lisette C.M. Anink-Groenen,[2] Pernette J. Verschure,[2] and Wouter de Laat[1,*]

[1]Hubrecht Institute-KNAW & University Medical Center Utrecht, Uppsalalaan 8, 3584 CT Utrecht, the Netherlands
[2]Synthetic Systems Biology and Nuclear Organization Group, Swammerdam Institute for Life Sciences, University of Amsterdam, 1098 XH Amsterdam, the Netherlands
[3]Co-first author
*Correspondence: w.delaat@hubrecht.eu
http://dx.doi.org/10.1016/j.molcel.2016.01.001

SUMMARY

Detailed genomic contact maps have revealed that chromosomes are structurally organized in megabase-sized topologically associated domains (TADs) that encompass smaller subTADs. These domains segregate in the nuclear space to form active and inactive nuclear compartments, but cause and consequence of compartmentalization are largely unknown. Here, we combined *lacO*/lacR binding platforms with allele-specific 4C technologies to track their precise position in the three-dimensional genome upon recruitment of NANOG, SUV39H1, or EZH2. We observed locked genomic loci resistant to spatial repositioning and unlocked loci that could be repositioned to different nuclear subcompartments with distinct chromatin signatures. Focal protein recruitment caused the entire subTAD, but not surrounding regions, to engage in new genomic contacts. Compartment switching was found uncoupled from transcription changes, and the enzymatic modification of histones per se was insufficient for repositioning. Collectively, this suggests that trans-associated factors influence three-dimensional compartmentalization independent of their cis effect on local chromatin composition and activity.

INTRODUCTION

Nuclear organization of interphase chromosomes is considered to be an important regulator of genome function. Chromosomes occupy nonrandom positions in the nucleus, and the position of a gene relative to nuclear compartments is thought to influence the probability of expression. Microscopy studies and high-throughput chromosome conformation capture (3C) techniques (de Wit and de Laat, 2012) have greatly increased our knowledge of the three-dimensional configuration of the genome in the nucleus (Bickmore and van Steensel, 2013). However, the mechanisms that govern the spatial arrangement of chromosomes and gene positioning in the nucleus are still poorly understood.

Individual chromosomes occupy visually discrete volumes within the nucleus (chromosome territories) (Bolzer et al., 2005). Hi-C and 5C studies revealed that chromosomes are partitioned into hundreds of distinct chromatin interaction domains called topologically associating domains (TADs) (Dixon et al., 2012; Nora et al., 2012). Most physical contacts between gene promoters and their regulatory elements are thought to take place within such domains (Dixon et al., 2015; Shen et al., 2012). TADs vary in their transcriptional activity, DNA sequence, and chromatin composition (Tanay and Cavalli, 2013). In Hi-C data, this segregation of different chromatin types in the nucleus is apparent as a biphasic state with an A and B compartment (Lieberman-Aiden et al., 2009), which have recently been subdivided into A1–2 and B1–4, each with their own chromatin signature (Rao et al., 2014). The A compartment represents colocalization of transcriptionally active or permissive chromatin and B the spatial crowding of inactive chromatin at the nuclear periphery (Guelen et al., 2008), nucleoli (Németh et al., 2010; van Koningsbruggen et al., 2010), and chromocenters (Wijchers et al., 2015). Since this spatial segregation is linked to gene expression, the composition of the A and B compartments is cell-type specific (Dixon et al., 2015). Although this suggests a role for nuclear compartmentalization in gene regulation, to what extent cell-type-specific compartment switches are a cause or consequence of gene activity is largely unclear.

Cell-type-specific nuclear arrangements imply that genomic loci must possess an intrinsic flexibility to allow repositioning in response to differentiation signals. Live-cell microscopy studies demonstrated that gene dynamics are generally constrained during interphase (Chubb et al., 2002), and that with a few exceptions (Chuang et al., 2006), gene repositioning generally requires mitosis before new locations can be adopted in daughter cells (Kumaran and Spector, 2008). Genes are physical parts of much larger chromosomes and will therefore rely on neighboring chromosomal segments for their freedom to move. Neuronal differentiation-induced activation of Mash1 is accompanied by visual relocation toward the nuclear center, dragging along immediately flanking genes that nevertheless remain transcriptionally silent (Williams et al., 2006). Repositioning of genomic regions relative to the

nuclear lamina occurs on both single-gene and multi-gene scales (Peric-Hupkes et al., 2010). The factors that influence the scale at which nuclear repositioning is regulated, and the consequences of locus repositioning for surrounding regions, are still unknown.

There are numerous examples of endogenous genes that change nuclear positions during lineage commitment, as judged from fluorescence in situ hybridization (FISH) studies. When silenced, they tend to locate to more peripheral nuclear locations in the nucleus (Peric-Hupkes et al., 2010) or closer to chromocenters (Brown et al., 1997), whereas they adopt a more internal nuclear position when activated (Peric-Hupkes et al., 2010). Repositioning of genes to the nuclear periphery through physical tethering leads to repression of associated genes in many, but not all, cases (Kumaran and Spector, 2008). Peripheral positioning therefore correlates with, but does not strictly imply, gene silencing.

So far, studies designed to experimentally alter the protein composition of a given gene locus and follow its nuclear positioning in a given cell type relied on microscopic techniques that measure location relative to visibly distinct nuclear structures such as the nuclear periphery, chromocenters, nucleoli, Polycomb bodies, speckles, transcription factories, or other gene loci. These nuclear landmarks all contain different DNA- and chromatin-binding proteins. The general perception from these studies is that tethering of compartment-specific proteins leads to locus recruitment to the corresponding nuclear entity. But how do we reconcile this with the compartmentalized genome conformation appreciated from Hi-C studies?

Few studies used 3C-based methods to monitor conformational changes that follow upon binding of regulatory proteins to a locus. 4C technology is highly suited for this, as it unbiasedly assays the genome-wide contacts of a locus of choice (Simonis et al., 2006). Blocking transcription and depleting RNA polymerase II from active gene loci did not cause any major changes in their genome-wide 4C contact maps (Palstra et al., 2008), nor did the recruitment of the glucocorticoid receptor to its target genes (Hakim et al., 2011). The introduction of the β-globin locus control region (a "super enhancer") into an already active chromosomal region did not induce new contacts but specifically strengthened those already formed with regions targeted by shared transcription factors (Noordermeer et al., 2011). A similarly subtle change in preferred contacts with cognate chromosomal regions was seen when the pluripotency factor NANOG was recruited to an artificial lac operator cassette (de Wit et al., 2013). These latter observations suggest that genes can reside in more than just two (A and B) genomic environments, a notion that receives support from data suggesting that various types of chromosomal regions with similar chromatin signatures preferentially cluster in the nuclear space (Bantignies et al., 2011; Rao et al., 2014; Tolhuis et al., 2011). However, these findings are generally based on genome-wide correlations, but little evidence from experimental locus repositioning is available to support these concepts.

To start exploring the flexibility of different genomic regions in response to the recruitment of various chromatin regulators, we followed chromosomally integrated *lacO*/lacR platforms and surrounding sequences by 4C technology, and we assessed how repositioning relates to transcriptional changes and epigenetic features.

RESULTS

Experimental System to Manipulate Locus Positioning in the Nucleus

To manipulate the nuclear positioning of defined genomic loci, we sequentially integrated a recruitment platform into two genomic regions of the same embryonic stem cells (ESCs) (Figure 1A). These ESCs were F1s derived from 129/Sv and C57BL/6 mouse strains, with only the alleles from the latter carrying the recruitment platform. This platform consisted of a kanamycin resistance gene flanked on either side by an array of 120 lac operator (*lacO*) repeats that serve as high-affinity binding sites for the *Escherichia coli* Lac repressor (lacR) (Figure 1A). After integration of the first array on chromosome 8 (previously used in de Wit et al., 2013), the additional floxed *Neomycin* resistance gene was removed to allow a second integration on chromosome 11 in the same cells. Correct and single integrations were validated by Southern blot (Figure S1A, available online).

We then expressed three different transcription and chromatin factors fused to an enhanced green fluorescent protein (EGFP)-lacR moiety to direct these proteins to the *lacO* cassettes. These factors included EZH2, the component of the Polycomb Repressive Complex 2 (PRC2) that is responsible for H3K27 trimethylation; SUV39H1, the histone methyltransferase factor that deposits H3K9me3 in constitutive heterochromatin; and NANOG, a pluripotency transcription factor that can have an activating as well as repressive effect on transcription (Liang et al., 2008). Viral transduction, followed by 10 days of cellular selection and expansion to obtain sufficient cells with a large enough (>70%) proportion of GFP-positive cells for subsequent analysis, led to expression of fusion proteins of the expected size but at varying levels (Figure 1B). EGFP-lacR showed a uniform nuclear distribution, except for two bright spots marking the *lacO* arrays (Figure 1C). These two bright foci were also visible in EGFP-lacR-NANOG cells, which otherwise displayed a more grainy pattern than EGFP-lacR alone, consistent with the thousands of NANOG binding sites across the genome. EGFP-lacR-EZH2 was found throughout the nucleus with multiple bright foci, presumably highlighting the presence of Polycomb bodies (Cheutin and Cavalli, 2014). EGFP-lacR-SUV39H1 showed the typical enrichment in DAPI-dense pericentromeric heterochromatin foci (PCH) (Aagaard et al., 1999). We checked the expression of a number of endogenous target genes in cells ectopically expressing EGFP-lacR-EZH2 and EGFP-lacR-NANOG and found this essentially unaltered or slightly reduced (in case of EZH2 expression) (Figure S1B). Binding of the fusion proteins to the *lacO* arrays was further verified by chromatin immunoprecipitation (ChIP) using antibodies against GFP, confirming again that all proteins bound to the array, albeit with different efficiencies (Figure 1D). Thus, all fusion proteins bound to the lacO arrays and showed the expected nuclear distribution.

Locus Susceptibility to Spatial Repositioning Depends on Genomic Location and Associated Factors

To explore the ability of the regulatory proteins to change the nuclear position of the *lacO* loci, we applied 4C-seq (Splinter et al., 2012), a 3C-based technology which probes

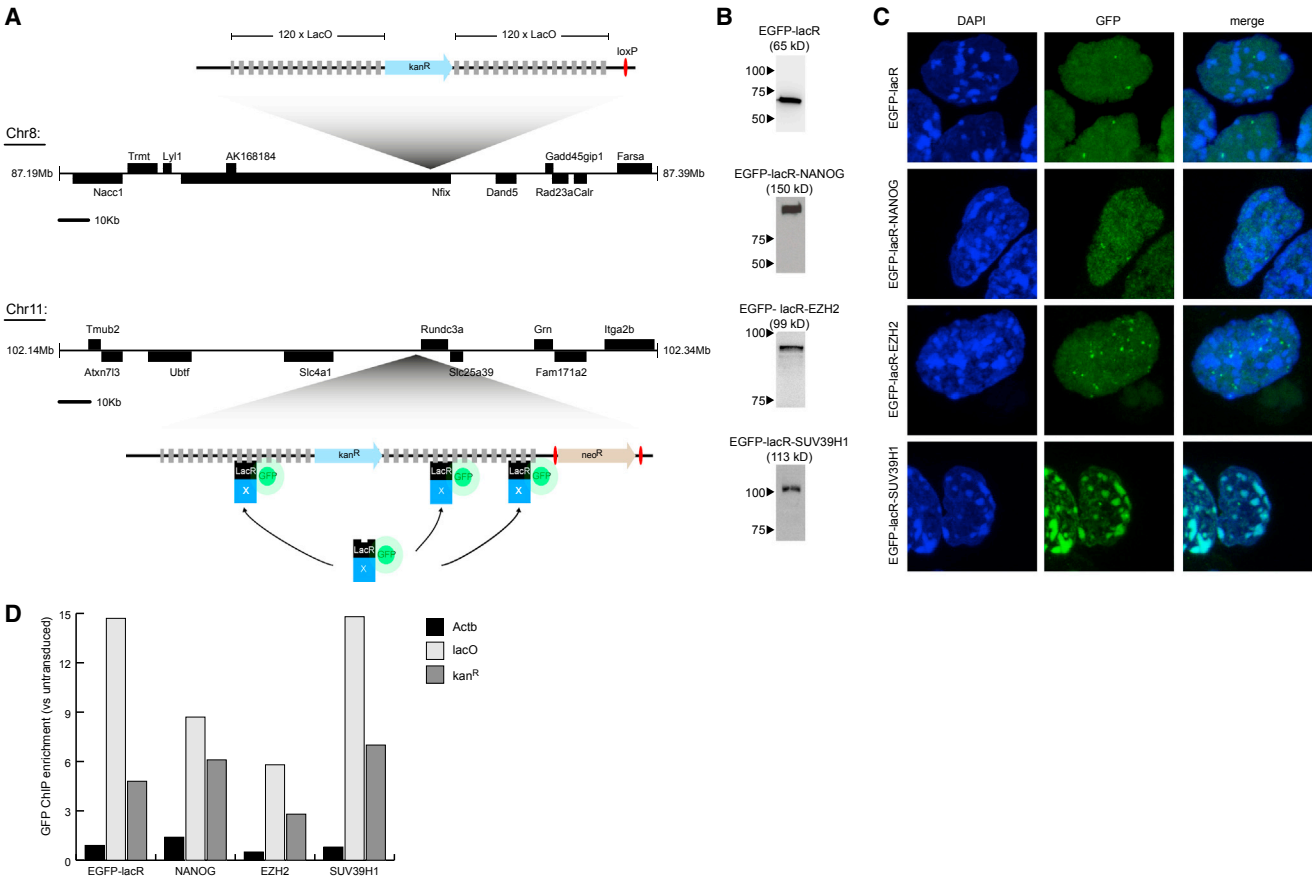


Figure 1. A *lacO*/lacR Recruitment Platform to Induce Nuclear Repositioning

(A) Schematic view of *lacO* integration sites on chr8 and chr11. *trans*-acting factors are recruited to the *lacO* array when fused to EGFP-lacR. See also Figure S1.

(B) Western blots confirming fusion proteins are of the correct size. Protein markers (left) and expected size of fusion protein (below) are indicated.

(C) Nuclear distributions of the fusion proteins used.

(D) ChIP with GFP antibody to confirm binding of each fusion protein to the lacO cassette.

for chromosomal regions spatially juxtaposed to a genomic site of interest. For the *lacO* locus on chromosome 11 (chr11), we used a 4C viewpoint in the *Neomycin* resistance (*Neo*) gene. To follow the *lacO* cassette on chromosome 8 (chr8), which no longer had a *Neo* gene, we employed an allele-specific 4C-seq (Splinter et al., 2011), taking advantage of a SNP in the same gene (*Nfix*) but 30 kilobases (kb) upstream of the *lacO* array. The SNP allowed distinguishing contacts made by the *lacO*-containing BL/6 allele and the wild-type 129/Sv allele, and we confirmed that it produced almost identical profiles to those of the *Neo* viewpoint in single-targeted cells (Figure S2A). Interchromosomal contacts are not nearly as abundant as contacts within a given chromosome, and their robust detection by 4C would require deeper sequencing of more complex 4C libraries than those analyzed here. Therefore, as in most Hi-C studies, we limit ourselves to the analysis of intrachromosomal contacts, which, in contrast, can readily be identified.

We found that binding of each chromatin protein had little impact on the genomic contacts made by the *lacO* locus on chr8 (Figure S2B). Very few, mostly quantitative, contact changes occurred within the otherwise unaltered genomic environment. Only upon EZH2 recruitment was one prominent new contact seen, with an H3K27me3-rich chromosomal region located at ~127 Mb of chr8 (Figure S2B; data not shown). In these same cells, all three chromatin proteins, but not EGFP-lacR alone (Figure 2A), had a much more pronounced impact on the contact profiles of the *lacO* locus on chr11 (Figures 2B–2D). NANOG and EZH2 recruitment each led to new contacts with regions not seen by the untargeted, or LacR-bound, array (Figures 2B and 2C); these changes were relatively subtle in the case of NANOG, whereas EZH2 recruitment induced prominent new contacts across a number of loci. Even more dramatic changes in contacts were seen after binding of SUV39H1, which led to massive loss of 4C signals across nearly all commonly contacted loci and a concomitant gain in 4C signals at normally ignored intervening chromosomal parts (Figure 2D). Thus, it appears that the susceptibility of a locus to change its 3D genomic neighborhood depends on genomic location. Since newly juxtaposed regions were different between NANOG-, EZH2-, and SUV39H1-transduced cells, the direction and degree of repositioning appears to be determined by associated *trans*-acting factors.

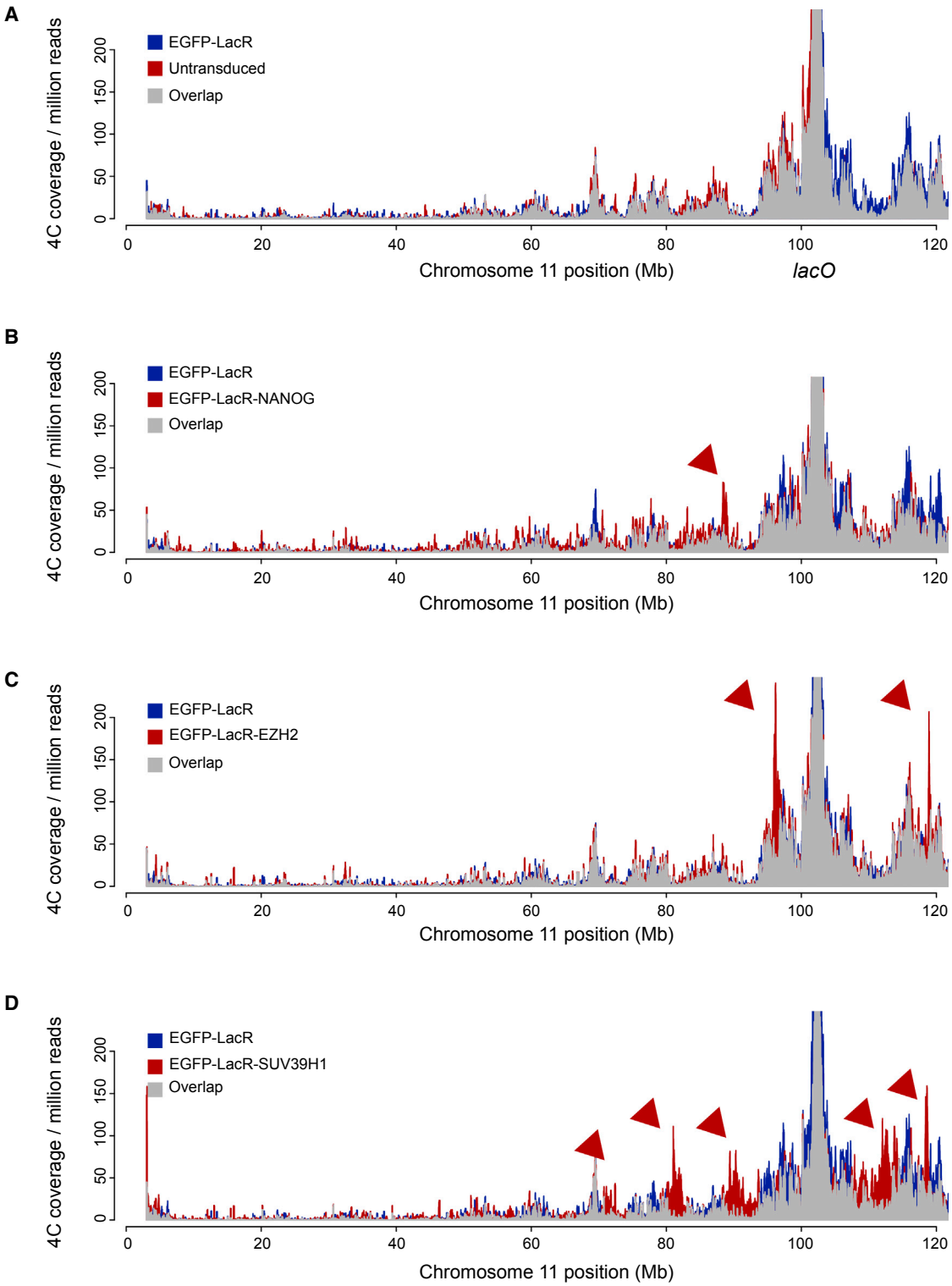


Figure 2. Susceptibility to Spatial Repositioning Depends on Genomic Location and *trans*-Acting Factors

(A–D) 4C profiles comparing chromosome-wide contacts of the *lacO* cassette on chr11 in (A) untransduced and EGFP-lacR transduced cells, (B) EGFP-lacR and EGFP-lacR-NANOG transduced cells, (C) EGFP-lacR and EGFP-lacR-EZH2 transduced cells, and (D) EGFP-lacR and EGFP-lacR-SUV39H1 transduced cells. The position of the *lacO* array is indicated. Arrowheads point at clear examples of differentially contacted regions. See also Figure S2.

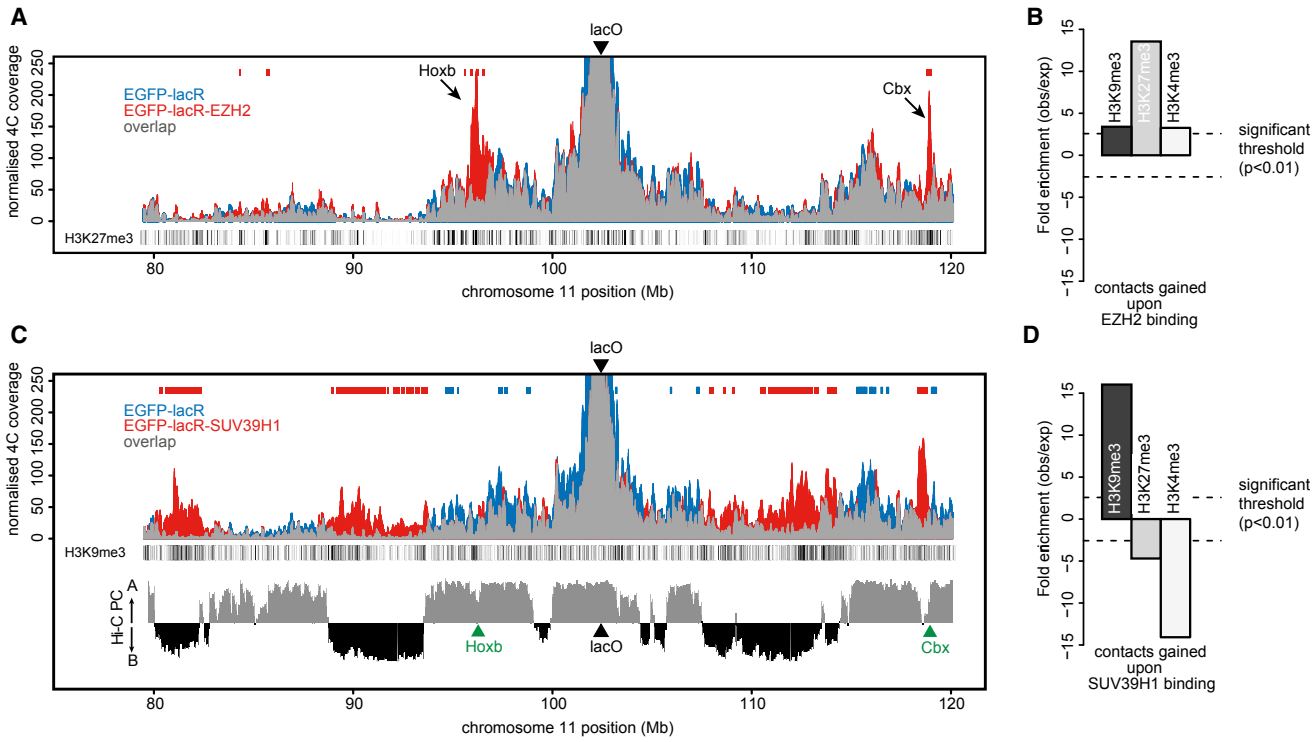


Figure 3. Local Chromatin Signature Influences Direction of Nuclear Repositioning

(A) 4C profiles of the *lacO* array viewpoint from EGFP-lacR (blue) and EGFP-lacR-EZH2 (red) transduced cells. Density of H3K27me3 sites is indicated below profiles. Red and blue bars above profiles indicate windows where 4C signal was significantly enriched in EGFP-lacR-EZH2 versus EGFP-lacR. See also Figure S4.

(B) Enrichment values of indicated histone modifications in regions significantly enriched upon EZH2 binding to the lacO (red bars in A). See also Figure S3.

(C) Same as (A), but for EGFP-lacR (blue) and EGFP-lacR-SUV39H1 (red). Density of H3K9me3 sites is indicated. Bottom panel displays Hi-C principle component values, corresponding to A and B compartments.

(D) As in (B), but for regions enriched upon SUV39H1 binding to the lacO.

Factor-Dependent Switching between Subcompartments with Distinct Chromatin Signatures

To further explore factor-dependent repositioning, we quantitatively compared normalized 4C profiles of transduced and untransduced cells and characterized the differentially contacted chromosomal segments. The targeting of NANOG induced only small shifts in the *LacO*-contacted regions, but the chromosomal parts that gained contacts were enriched in binding sites for NANOG (Figure S3), as seen before (de Wit et al., 2013). EGFP-lacR-Ezh2 profiles also globally matched that of EGFP-lacR, except at two regions that showed a very prominent accumulation of contacts, one being the *HoxB* cluster and the other being a cluster of *Cbx* genes (Figure 3A). These loci efficiently recruit both the Polycomb PRC2 complex protein EZH2 and the PRC1 subunit Ring1B (Ku et al., 2008) and are highly enriched for H3K27me3 (Figure 3A), a modification for which EZH2 is responsible (Simon and Kingston, 2009). Systematic analysis indeed showed that EZH2 recruitment to *LacO* promoted contacts with other chromosomal regions enriched for the H3K27me3 mark (Figure 3B). This shows that EZH2 promotes spatial crowding of genomic regions with a Polycomb chromatin signature.

EGFP-lacR-SUV39H1 binding to the *lacO* led to massive changes in contact frequencies across large chromosomal intervals (Figure 3C). This time, the newly contacted domains were depleted for the H3K27me3 chromatin mark and lacked H3K4me3, but were strongly enriched for H3K9me3, the histone mark deposited by SUV39H1 (Figure 3D). Thus, like EZH2 and NANOG, SUV39H1 binding causes the locus to change its preferred genomic neighbors in nuclear space. The three factors, however, induce different topological changes, each promoting locus movement to distinct nuclear environments occupied by genomic regions with the corresponding associated factors and chromatin signatures. To validate the 4C results, we performed DNA FISH using BAC probes against the integration site, two SUV39H1-contacted regions, and one EZH2-contacted region (the HoxB locus), plus a probe against the LacO repeats. Whereas the LacO-targeted site and its corresponding untargeted site on the homologous chromosome showed very similar contact frequencies with each of the three tested regions, we found significantly increased contact frequencies (p value < 0.001; one-sided test) with the HoxB locus exclusively upon EZH2 recruitment to the LacO-targeted site. Similarly, upon SUV39H1 recruitment the LacO site showed significantly increased FISH contacts (p value < 0.01 for both) with the two

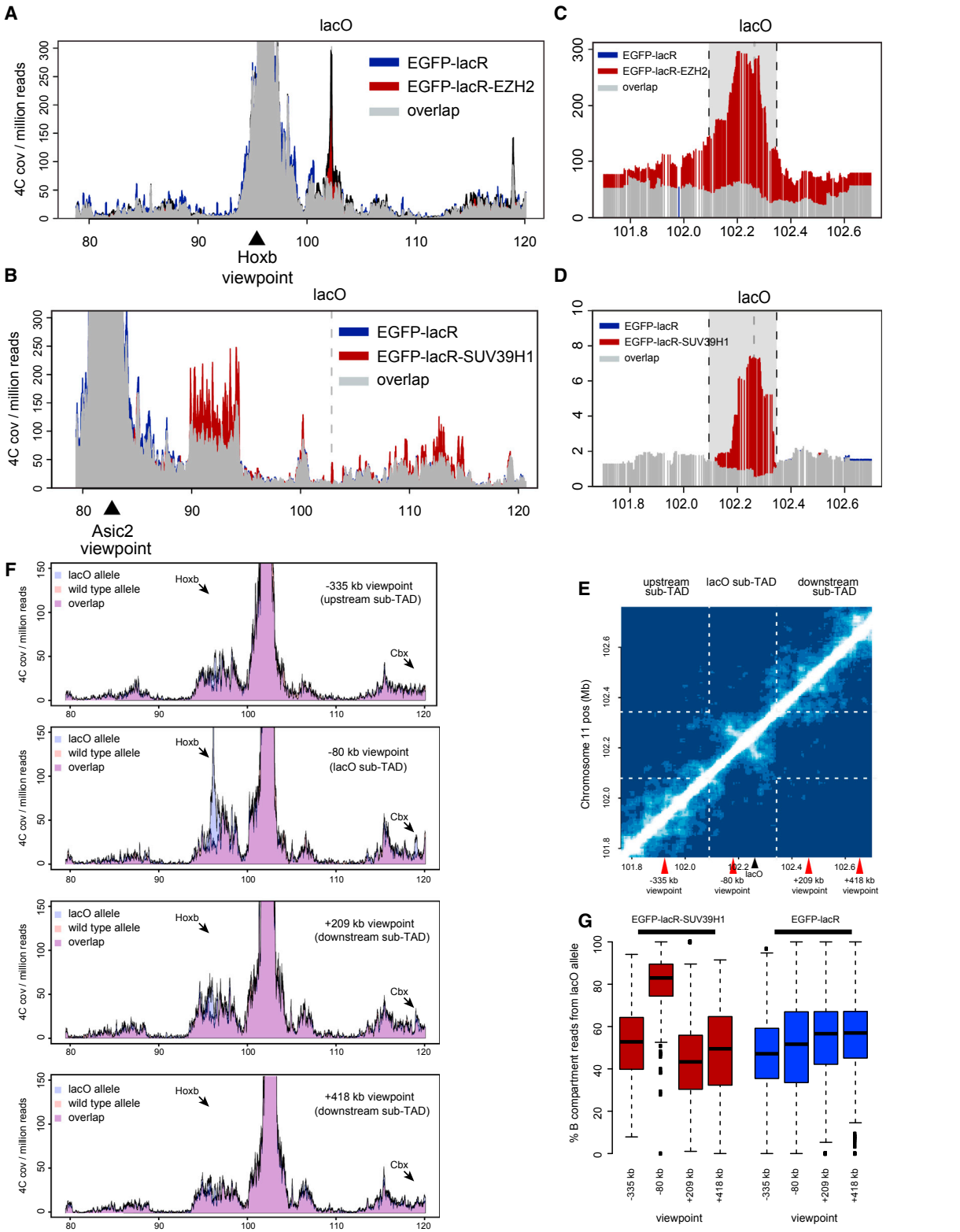


Figure 4. A Repositioned Locus Drags along Its Associated SubTAD

(A) 4C profiles looking from the EZH2-induced contact at the *HoxB* cluster (indicated by arrowhead) in EGFP-lacR (blue) and EGFP-lacR-EZH2 (red) transgenic cells. Location of the *lacO* integration site is indicated by dashed line. See also Figure S4.

(legend continued on next page)

4C-predicted SUV39H1 target regions (Figure S4). Thus, as extensively shown before (see, e.g., Simonis et al., 2006), FISH confirms that our 4C approach faithfully identifies long-range chromosomal contacts.

Since the locus on chr11 can be forced to move to at least three different genomic environments, our data confirm recent evidence for the existence of multiple subcompartments (Rao et al., 2014) within the "active" A and "inactive" B compartments (Lieberman-Aiden et al., 2009). We intersected the 4C contact profiles with Hi-C data generated from ESCs to investigate how the gained and lost contacts compared to regions belonging to the originally defined A and B compartments (Lieberman-Aiden et al., 2009). The untargeted locus on chr11 resides in the active A compartment (Figure 3C). This is true also for the *HoxB* and the *Cbx* gene clusters that are nevertheless ignored by the locus unless EZH2 is recruited. The same is true for NANOG-dense regions: they belong to the A compartment but are only selected by *lacO* as preferred genomic neighbors upon NANOG recruitment. As shown before, NANOG-dense chromosomal regions and Polycomb-associated regions each preferentially cluster among themselves (de Wit et al., 2013; Denholtz et al., 2013; Vieux-Rochas et al., 2015). Thus, we independently corroborate recent high-resolution Hi-C data showing that the A compartment can be further subdivided into subcompartments in which similarly typed chromosomal regions specifically accumulate (Rao et al., 2014). In ESCs, the cell-type-specific pluripotency factors establish one such subcompartment, and Polycomb group proteins establish another A-type subcompartment. Depending on its associated proteins, an integrated *LacO* platform can switch between these subcompartments.

Interestingly, SUV39H1 recruitment caused the locus to switch from the A to the B compartment. Regions originally surrounding the untargeted *lacO* locus that massively lost contacts upon SUV39H1 binding all belonged to the A compartment, whereas the chromosomal parts that accumulated 4C signals showed a striking correlation with those belonging to the B compartment. In such instances, new contacts were observed across the entire B compartment domains extending to their boundaries, beyond which a loss in contacts was seen. This shows that in the same cell type a locus can switch environments not only within a compartment but also from one compartment to another.

Spatial Repositioning Is Regulated at the SubTAD Level

To validate these results, we performed reciprocal 4C-seq experiments with viewpoints located in newly identified EZH2- or SUV39H1-specific contact domains. 4C profiles from the *HoxB* locus (~6 Mb away) confirmed prominent new contacts around the *lacO* integration site in EGFP-lacR-Ezh2 transgenic cells that were absent in untransduced cells (Figure 4A). Similarly, looking from a locus containing the *Asic2* gene (>20 Mb away), i.e., one of the regions contacted specifically upon SUV39H1 recruitment to *LacO*, revealed a clear SUV39H1-specific peak at the position of the *lacO* transgene (Figure 4B). This confirms the chromosome conformational changes induced by the local recruitment of these chromatin factors. Looking more closely at the reciprocal 4C contact data, we noticed that the segment around the *lacO* integration site with the highest differential signal spanned a similar-sized region for both EGFP-lacR-Ezh2- and EGFP-lacR-Suv39h1-transduced cells (Figures 4C and 4D). The size and the borders of this region corresponded to a structural domain appreciable from a Hi-C contact matrix (Figure 4E) that we produced for mouse ESCs (based on the analysis of 26 million valid Hi-C tags with a *cis*/total enrichment of 74%; Geeven et al., 2015). This region measures approximately 200 kb, which is much smaller than the originally defined TAD at this chromosomal location (spanning ~840 kb, from chr11:101480001–102320000) (Dixon et al., 2012) and would therefore better qualify as a subTAD (Phillips-Cremins et al., 2013). It therefore appears that focal recruitment of *trans*-acting factors can cause an entire subTAD to get engaged in new contacts.

To further investigate the extent to which the 3D positioning of flanking sequences is influenced by protein recruitment to *LacO*, we performed allele-specific 4C from viewpoints that lie up- and downstream of the *lacO* array and compared profiles between the *lacO*-transgenic and untargeted allele. Upon EZH2 binding to the *lacO* locus, a viewpoint in the same subTAD but 80 kb upstream of *LacO* showed the same switch in chromosomal contacting partners as seen for the *lacO* array itself (Figure 4F), making many new distal contacts with the upstream *HoxB* locus and the downstream *Cbx* cluster. Viewpoints further upstream (−335 kb) and downstream (+209 and +418 kb) of *lacO*, both located in directly neighboring subTADs (Figure 4E), hardly showed these differential contacts upon EZH2 and SUV39H1 binding (Figure 4F). Similarly, only the viewpoint located at −80 kb in the same subTAD as *lacO* followed *lacO* in its A → B compartment switch upon SUV39H1 binding, while viewpoints in neighboring subTADs appeared resistant to this repositioning (Figures 4G and S5). We conclude that the 3D positioning of chromosomal segments can be controlled at the level of subTADs, i.e., at the submegabase level involving chromosomal

(B) Same as (A), but for EGFP-lacR (blue) and EGFP-lacR-SUV39H1 (red), looking from a SUV39H1-induced contact in the *Asic2* gene. Note that the 4C is not allele specific, implying that contacts come from both alleles, which effectively dilutes the signal at the *lacO* integration site.
(C) Same as (A), zoomed in on the *lacO* locus. The gray area highlights the region with the highest differential 4C signal.
(D) Same as (B), zoomed in on the *lacO* locus. The gray area highlights the region with the highest differential 4C signal.
(E) Hi-C contact matrix of region surrounding the *lacO* locus on chr11 (data from ESCs). Dashed lines correspond to the borders of the gray area in (C) and (D). Note that the borders coincide with edges of a block of locally increased contact frequencies (*lacO* subTAD). Red arrowheads indicate positions of viewpoints used in (F).
(F) Allele-specific 4C profiles comparing the lacO and wild-type alleles upon binding of EZH2 for a viewpoint 335 kb upstream of the lacO in a neighboring subTAD, 80 kb upstream within the same subTAD, and 209 kb and 418kb downstream of the LacO in neighboring subTADs.
(G) Boxplots showing the percentage of reads coming from the *lacO*-transgenic allele (versus wild-type allele) across regions mapping to the B compartment upon EGFP-lacR-SUV39H1 (red) or EGFP-lacR only (blue) binding. See also Figure S5.

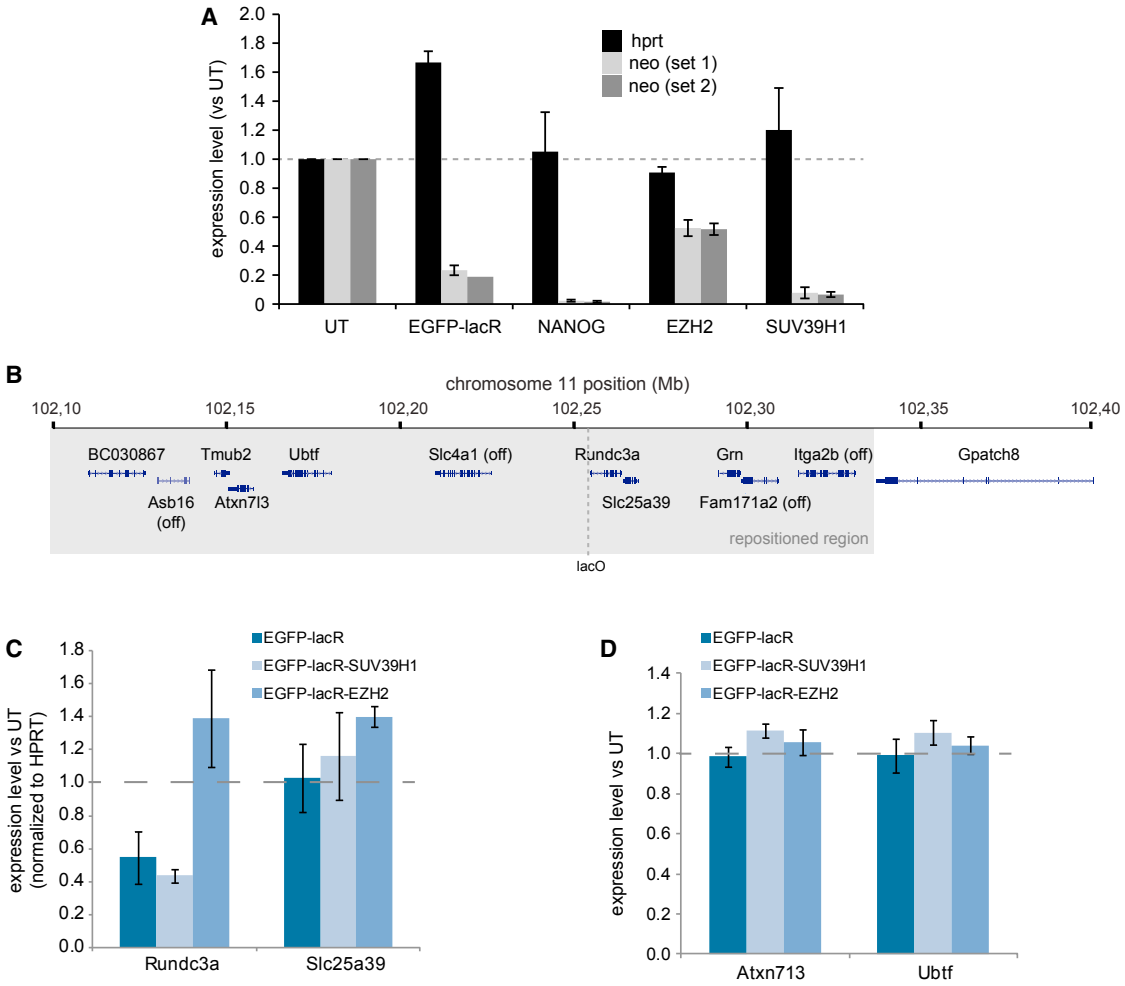


Figure 5. Spatial Repositioning Does Not Drive Gene Expression Changes

(A) Quantitative RT-PCR analysis of *Neomycin* (two primer sets) and control gene (*Hprt*). Data was normalized to *Actb* and expressed as fold-change over untransduced (UT) cells. Error bars represent SD.

(B) Schematic overview of genes surrounding the *lacO* integration site. Genes that are transcriptionally inactive in ES cells are indicated by "off." The gray area corresponds to the subTAD that is repositioned upon SUV39H1 binding (see also Figure 4)

(C) Quantitative RT-PCR on the two genes most nearby the *lacO* integration site. Note that the downregulated *Rundc3a* gene is only 36 bp downstream of the *lacO* transgene. Error bars represent SD.

(D) SNP-based expression analysis of the more distal *Atxn713* and *Ubtf* genes. Data plotted are the fold-changes in the fraction of reads coming from the *lacO*-targeted allele. Values of all replicates were normalized to the first untransduced (UT) replicate. Error bars represent SD.

regions that measure only a few hundreds of kilobases in size. Sequences belonging to the same subTAD experience the same genomic environment, while those belonging to flanking subTADs may occupy another nuclear subcompartment.

Nuclear Compartment Switching and Gene Expression Are Not Causally Related

Having established a system where we can manipulate the spatial genomic neighborhood of a subTAD, we then wished to understand the relationship with gene expression. We first asked whether transcriptional changes could be a driving force behind repositioning, focusing on the expression of the *Neo* gene that was cointegrated with and immediately flanking the *lacO* array on chr11. *Neo* expression was reduced by the binding of the *trans*-acting factors (Figure 5A), consistent with their roles as transcriptional repressors (Liang et al., 2008). EGFP-LacR binding by itself was also sufficient to repress *Neo* expression (Figure 5A), a phenomenon seen before and attributed to the tight binding of lacR to DNA (Dubarry et al., 2011). In the case of EGFP-lacR alone, repression occurred without nuclear repositioning of the locus; for NANOG and EZH2 this was accompanied by engagement with a few new but mutually exclusive contacting regions, and for SUV39H1 this occurred concomitant with a seemingly complete switch toward the B compartment. We conclude therefore that a change in transcriptional output is insufficient for nuclear repositioning of the subTAD.

Vice versa, repositioning may induce expression changes of the genes that comigrate with *lacO* to new compartments

upon factor tethering. To investigate whether and to what extent this occurs, we focused on the genes immediately surrounding *LacO* in the same subTAD of chr11 (Figure 5B). In SUV39H1-transduced cells these genes follow *LacO* (Figure 4G) and migrate from the A to the B compartment in a large proportion of the cells (Figure 3C). We analyzed their expression by quantitative RT-PCR (Figure 5C) or by next-generation sequencing of PCR-amplified cDNA products (Figure 5D), using SNPs that allowed discrimination of the *lacO*-transgenic allele. From this, we found that only the immediately flanking Rundc3a gene, with its promoter 36 bp downstream of the integration site, displayed decreased expression levels. It did so, however, not only upon SUV39H1 binding but also upon EGFP-lacR binding (Figure 5C), which does not induce spatial repositioning of the domain. Rundc3a expression was unaffected (or even slightly upregulated) upon EZH2 recruitment, despite this factor causing spatial repositioning to Polycomb bodies. Altered Rundc3 gene expression therefore cannot be attributed to nuclear repositioning. The expression of other genes across the repositioned subTAD was unaffected upon SUV39H1 and EZH2 recruitment (Figures 5C and 5D). We therefore conclude that nuclear repositioning of the subTAD does not drive changes in its gene expression, and vice versa. In summary, genome-wide datasets clearly show that transcriptional control and genomic compartmentalization correlate, but consistent with other findings (Dixon et al., 2015; Palstra et al., 2008; Therizols et al., 2014) our results emphasize that they are often not causally related.

Local Modification of Histones Is Not Sufficient for Nuclear Repositioning

We next wished to investigate the role of chromatin composition in the spatial repositioning. Consistent with their enzymatic activity, EZH2 and SUV39H1 recruitment led to local deposition of H3K27me3 and H3K9me3, respectively, at the *lacO* locus on chr11 (Figures 6A–6C). Although we did not extensively analyze the corecruitment of other PcG proteins, we did notice that EZH2 (a PRC2 complex member) attracted Ring1B, a PRC1 component (Figure S6A). The H3K27me3 on chr11 did not spread far into adjacent regions: relative levels increased somewhat across the subTAD but remained low in absolute terms, with the exception of a site at +60 kb that is naturally rich in H3K27me3 (Figures 6A and 6B). This is consistent with recent findings that EZH2 spreading upon induced recruitment of Polycomb group proteins to a *tetO* transgene is usually limited to <10 kb (Blackledge et al., 2014). Similarly, upon SUV39H1 recruitment to *lacO*, there was a modest increase in H3K9me3 levels at the surrounding sequences, but their absolute levels remained marginal compared to the array itself (Figure 6C).

The deposition of these chromatin marks and the concomitant repositioning of the locus to nuclear subcompartments of similar chromatin signature (Figure 3) raise the possibility that histone modifications play an active role in the engagement of specific long-range contacts and establishment of nuclear subcompartments. To further investigate this, we constructed a mutant SUV39H1 that lacked the chromodomain (CD) but retained the histone methyltransferase domains (Figure 6D). The CD is required for binding to H3K9me3, and, correspondingly, this SUV39H1$^{\Delta CD}$ mutant no longer localized to PCH (Figure 6E). It still bound to the *lacO* array and efficiently deposited H3K9me3 (Figures 6E,6F, and S6B). The SUV39H1$^{\Delta CD}$ mutant was also fully capable of repressing *Neo* gene expression to levels lower than that already induced by EGFP-LacR alone and even slightly lower than that induced by wild-type SUV39H1(Figure 6G). The latter seems to correspond to the somewhat higher levels of H3K9me3 deposited by the SUV39H1$^{\Delta CD}$ mutant (Figure 6F). Despite this evident activity, however, SUV39H1$^{\Delta CD}$ binding, in contrast to SUV39H1, did not result in locus repositioning. We no longer observed distinct interactions with genomic regions in the B compartment (Figure 6H), but instead, SUV39H1$^{\Delta CD}$ gave similar 4C profiles to those of EGFP-LacR transduced cells (Figure 6H). This implies that local deposition of H3K9 methylation may control gene silencing but is insufficient to direct spatial repositioning.

DISCUSSION

The link between nuclear localization, chromatin composition, and gene expression is primarily based on correlations extracted from genome-wide datasets. Microscopy studies tracing the location of individual loci in general confirm established relationships but also reveal exceptions (Kumaran and Spector, 2008) and show that there is large cell-to-cell variability in nuclear positioning of given loci (Kind et al., 2013). Tracing a locus by microscopy relies on measuring distances relative to visible landmarks such as the nuclear periphery, chromocenters, or selected other loci, but how such movement relates to the TADs, compartments, and other topological features appreciable from high-throughput 3C data is unclear. With the aim to bridge this gap and establish cause and consequence in the three-way relationship between nuclear organization, gene expression, and chromatin composition, we combined *lacO*/LacR recruitment platforms with allele-specific 4C technology. We recruited a selection of *trans*-acting proteins that could modulate these parameters at defined genomic locations in the same cell type. Although endogenous genes are usually not targeted by so many copies of a given protein, repressive H3K27me3 and H3K9me3 marks can span genomic regions of tens of kb (Hawkins et al., 2010), not unlike the ~10 kb *lacO* array integrated here.

Susceptibility to Spatial Repositioning Depends on Genomic Location

All three *trans*-acting factors were able to change the spatial positioning of the *lacO* locus when integrated upstream of *Rundc3a* on chr11, albeit to different extents. However, the same *lacO* array integrated in the *Nfix* gene on chr8 seemed almost impervious to these large-scale repositioning influences, showing only minor small-scale changes (Figure S2; de Wit et al., 2013). As data from both integration sites comes from the same cell with identical amounts of *lacO* repeats in each transgene, these differential responses cannot be explained by technical differences. This proves, therefore, that genomic context is an important determinant of the flexibility of a locus to roam the nucleus (Krijger and de Laat, 2013). What elements in the genomic

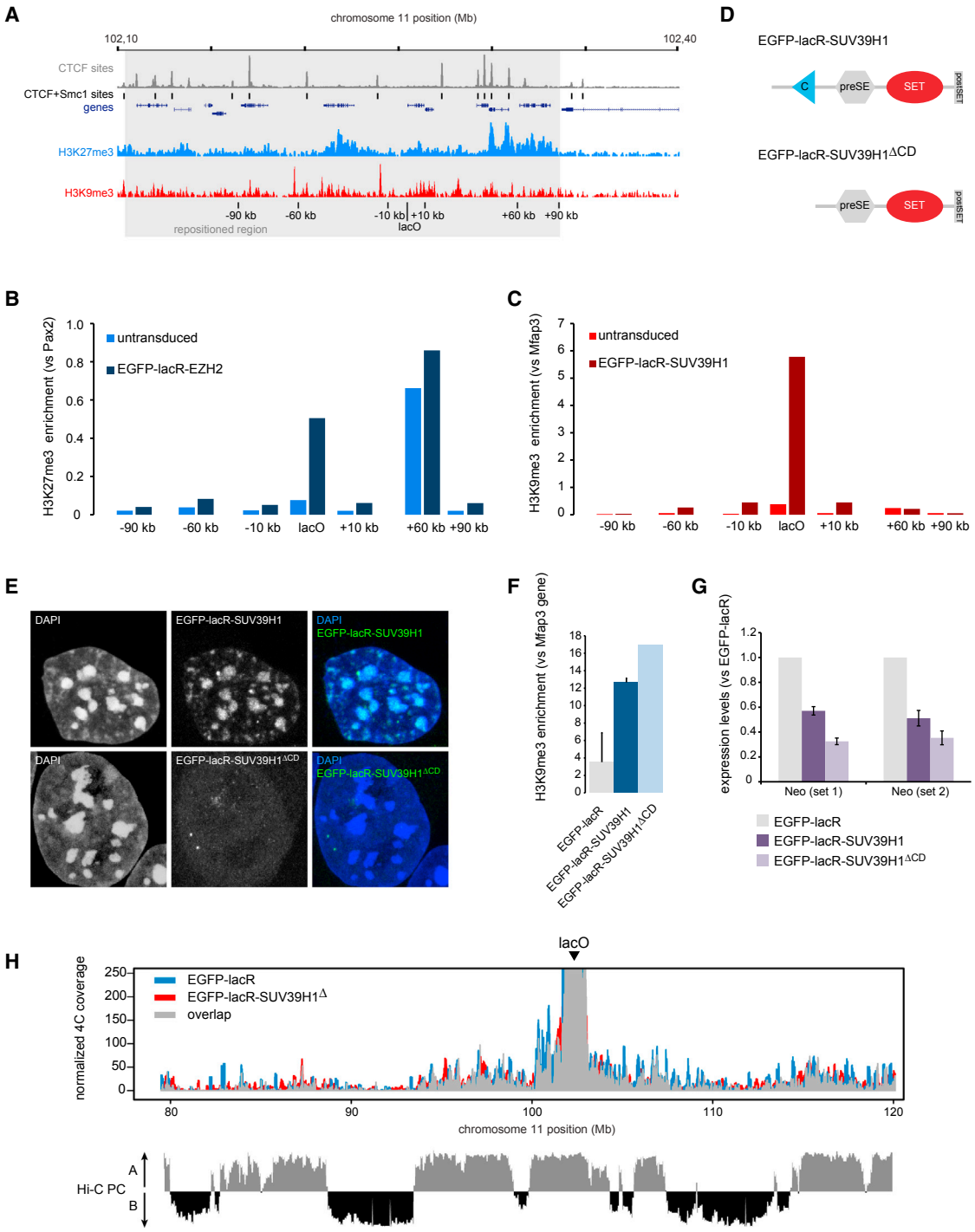


Figure 6. The Role of Chromatin in Spatial Repositioning

(A) Schematic overview of *lacO* locus on chr11, with respect to CTCF binding sites (gray), shared CTCF-cohesin sites (blocks), H3K27me3 (blue), and H3K9me3 (red) histone modifications. Locations of ChIP primers are indicated. Grey area corresponds to the subTAD that is repositioned upon EZH2 or SUV39H1 binding (Figure 4).

(B) ChIP for H3K27me3 in untransduced (light blue) and EGFP-lacR-EZH2 (dark blue) cells for the genomic sites indicated in (A). Input-normalized data are expressed as enrichment over the (H3K27me3-rich) *Pax2* promoter. See also Figure S6.

(C) As in (B) but for H3K9me3 in untransduced (light red) and EGFP-lacR-SUV39H1 (dark red) transduced cells. Input-normalized data is expressed as enrichment over the (H3K9me3-rich) *Mfap3* gene.

(D) Schematic display of the wild-type and mutant SUV39H1 proteins. See also Figure S6.

(legend continued on next page)

habitat surrounding these *lacO* arrays are responsible for the differential capacity to relocalize is still unclear. Differences in gene density, transcriptional activity, and distance to telomere and centromere may play a role (Brown et al., 2008), but most of these parameters are not markedly different between the *lacO* loci on chr8 and chr11. For example, both loci localize to the A (active) compartment in ESCs. The integration site at 102.2 Mb on chr11 (122 Mb in size) is considerably closer to the telomere, though, and therefore possibly more flexible than the integration site at 87.5 Mb on chr8 (129 Mb in size). Studies using more systematic approaches such as TRIP (Akhtar et al., 2014) could perhaps help to further dissect the parameters dictating locus mobility in the future.

Changes in Nuclear Position Are Controlled by Topological Constraints

As each locus is physically connected to the rest of its chromosome, its repositioning will by default carry along adjacent chromosomal regions. Differentiation-induced changes in association with the nuclear lamina and A/B compartments usually correspond to one or multiple neighboring TADs (Dixon et al., 2015; Peric-Hupkes et al., 2010). However, the causes and consequences of these movements are unclear; do these represent the coordinated movement of the complete TADs, or are they mostly the result of passive movement following the regulated repositioning of a single gene?

Here, we establish that localized chromatin modification of a relatively small chromosomal segment induces new contacts with chromosomal regions elsewhere on the same chromosome. Our reciprocal 4C analysis reveals that the remainder of the associated subTAD, rather than the megabase-scaled TAD, also engages in the same contacts. Thus, the nuclear position of a locus can be controlled at the level of subTADs, which hints at a hitherto unknown flexibility of subTADs to adopt a different nuclear position from the remainder of its associated TAD. Overall, we expect that spatial contact partners of a given subTAD will be selected based on their chromatin composition and linear proximity on the chromosome, but that with additional constraints imposed by the remaining segments of the chromosome, each will have their own preference for a spatial neighborhood.

trans-Acting Factors Form Nuclear Subcompartments

What determines where a locus is repositioned to? Although binding of all three *trans*-acting factors had a repressive effect on gene expression, they guided their target locus in different directions or kept it at the same location. This suggests that loci can have a remarkable flexibility in their nuclear localization and that loci with the same transcriptional status can occupy different subcompartments in the nucleus. It also shows that the repositioning to different subcompartments cannot be the mere consequence of transcriptional changes.

Our data suggest that specific nuclear subcompartments form as a consequence of affinities between proteins and/or modifications associated with the participating chromosomal regions. All three *trans*-acting factors showed a strong preference to move the target locus to genomic neighborhoods with a chromatin signature corresponding to that deposited by these factors at the *lacO* array. How these proteins impose the observed repositioning is not entirely clear. Given the limited freedom of chromatin to roam the interphase nucleus, the substantial repositioning that we measure probably depends on one or multiple rounds of mitosis, as also seen in other systems (Kumaran and Spector, 2008). Polycomb group proteins have the intrinsic property to aggregate to establish a compact chromatin structure when bound to nearby nucleosomes (Francis et al., 2004). Likewise, the deposition of H3K9me3 by SUV39H1 serves as a docking platform for a whole range of heterochromatic proteins, including heterochromatin protein 1, which is thought to oligomerize to promote chromatin compaction (Grewal and Jia, 2007). Similar to experimentally induced chromatin loops through spatial interaction between two dimerization domains at an enhancer and promoter (Deng et al., 2012), the inclination of such protein complexes to self-associate may therefore increase the probability of the *lacO* locus to stabilize contacts with genomic sites that are bound by the same factors.

Despite the strong association of chromatin modifications with topological organization of the genome, the failure of our mutant SUV39H1 to reposition the *lacO* locus despite its normal methyltransferase activity implies that focal histone modifications are not sufficient to drive spatial repositioning. The mutant SUV39H1 version lacked the chromodomain, a protein module responsible for binding to H3K9me3. We previously demonstrated that recruitment of only the chromodomain of CBX1/HP1beta was sufficient to relocate and tether a *lacO* array to PCH, presumably through the simultaneous binding of lacR to the lacO repeats and the chromodomain to the H3K9me3 decorations abundant in PCH (Wijchers et al., 2015). Our current finding that SUV39H1 requires its chromodomain for repositioning seems fully consistent with this. Then why is the H3K9me3 deposition on *LacO* by itself not sufficient for PCH-associated chromodomains to grab and reposition the locus? A possible explanation may be that there are not sufficient H3K9me3 molecules per *lacO* allele to efficiently compete for stable associations with chromodomain-containing proteins that are accumulated at PCH, or that the presence of a mutant SUV39H1 acts as steric hindrance and shields H3K9me3 from chromodomains. Alternatively, the N-terminal part of SUV39H1 that was removed is also responsible for interactions with HP1 and required for recruitment of the complete set of classical heterochromatin proteins (Muramatsu et al., 2013). The heterochromatin established at the lacO array by the mutant SUV39H1 may

(E) Nuclear distributions of the fusion proteins used.

(F) ChIP data confirming H3K9me3 deposition upon binding of both wild-type and mutant SUV39H1. Error bars represent SD.

(G) Quantitative RT-PCR analysis of *Neomycin* (two primer sets). Data were normalized to *Actb* and expressed as fold-change over EGFP-lacR cells. Error bars represent SD.

(H) 4C profiles of the *lacO* array viewpoint (arrowhead) from EGFP-lacR (blue) and EGFP-lacR-SUV39H1$^{\Delta CD}$ (red) cells. Bottom panel displays Hi-C principle component values, with positive and negative corresponding to A and B compartments, respectively.

therefore lack particular modules that are involved in nuclear repositioning.

Altogether, our data show that subTADs differ in their susceptibility to adopt new nuclear positions. We identify nuclear subcompartments where subTADs with similar epigenetic features can be recruited. Associated *trans*-acting factors drive this nuclear compartmentalization independent of their effect on local chromatin composition and activity. While there seems to be no causal relationship between compartmentalization and gene expression, this 3D organization will cause spatial crowding of transcription and chromatin factors, which is likely to contribute to maintenance of transcriptional states.

EXPERIMENTAL PROCEDURES

lacO Targeting and LacR-Fusion Transduction

The *lacO* array was inserted in chr8 and chr11 via site-specific homologous recombination, as described (de Wit et al., 2013). *LacO*-transgenic cells were transduced with EGFP-lacR fusions using lentivirus, selected with puromycin for approximately 10 days when cells had reached sufficient numbers for collection, and tested for purity by flow cytometry (minimum 70% GFP positive). Further details can be found in the Supplemental Information.

4C-Seq Analysis

4C sample preparation was performed as described (Splinter et al., 2012); data mapping and analysis was as described (de Wit et al., 2013). To allow direct comparisons of 4C profiles of different samples, mapped reads are normalized for sequencing depth by multiplying by a factor such that the total sum of mapped reads in *cis* (i.e., on the bait chromosome), after discarding the 4C fragend with the highest multiplicity, equals one million. For visualization, we compute the average 4C signal across windows of 51 consecutive fragends (centered on the 25th) per experiment. Per condition we plot the average of this signal as taken from two independent biological replicates.

FISH

FISH was performed as described previously and detailed in the Supplemental Information. ImageJ (Image5D plugin) software was used to score touching (no unstained pixel in between FISH signals) and nontouching (with unstained pixel(s) in between) alleles.

ACCESSION NUMBERS

The accession number for the 4C data reported in this paper is GEO: GSE76174.

SUPPLEMENTAL INFORMATION

Supplemental Information includes Supplemental Experimental Procedures and six figures and can be found with this article online at http://dx.doi.org/10.1016/j.molcel.2016.01.001.

AUTHOR CONTRIBUTIONS

P.J.W., P.H.L.K., and Y.Z. performed all experiments with help from A.D., M.J.A.M.V., M.J., C.V.-Q., C.V., and H.T. All computational and statistical analyses were performed by G.G. L.C.M.A.-G. and P.J.V. provided reagents. P.J.W., P.H.L.K., and W.d.L. designed and supervised the experiments, and the manuscript was written by P.J.W. and W.d.L.

ACKNOWLEDGMENTS

We thank Daan Noordermeer, Erik Splinter, Petra Klous, and Harmen van de Werken for technical support; Niels Geijsen for sharing phage2 constructs; Stefan van der Elst for FACS sorting; Stieneke van den Brink for ESC culture; and Anko de Graaff and the Hubrecht Imaging Center for supporting the imaging. This work was supported by a Long-Term Postdoctoral Fellowship of the Human Frontier Science Program (LT000709/2013) to A.D., an NWO/CW TOP grant (714.012.002), an NWO VICI grant 724.012.003, a NanoNextNL grant, an EU grant 2010-259743 (MODHEP), and a European Research Council Starting Grant (209700, "4C") to W.d.L.

Received: June 12, 2015
Revised: October 30, 2015
Accepted: December 22, 2015
Published: January 28, 2016

REFERENCES

Aagaard, L., Laible, G., Selenko, P., Schmid, M., Dorn, R., Schotta, G., Kuhfittig, S., Wolf, A., Lebersorger, A., Singh, P.B., et al. (1999). Functional mammalian homologues of the Drosophila PEV-modifier Su(var)3-9 encode centromere-associated proteins which complex with the heterochromatin component M31. EMBO J. *18*, 1923–1938.

Akhtar, W., Pindyurin, A.V., de Jong, J., Pagie, L., Ten Hoeve, J., Berns, A., Wessels, L.F., van Steensel, B., and van Lohuizen, M. (2014). Using TRIP for genome-wide position effect analysis in cultured cells. Nat. Protoc. *9*, 1255–1281.

Bantignies, F., Roure, V., Comet, I., Leblanc, B., Schuettengruber, B., Bonnet, J., Tixier, V., Mas, A., and Cavalli, G. (2011). Polycomb-dependent regulatory contacts between distant Hox loci in Drosophila. Cell *144*, 214–226.

Bickmore, W.A., and van Steensel, B. (2013). Genome architecture: domain organization of interphase chromosomes. Cell *152*, 1270–1284.

Blackledge, N.P., Farcas, A.M., Kondo, T., King, H.W., McGouran, J.F., Hanssen, L.L., Ito, S., Cooper, S., Kondo, K., Koseki, Y., et al. (2014). Variant PRC1 complex-dependent H2A ubiquitylation drives PRC2 recruitment and polycomb domain formation. Cell *157*, 1445–1459.

Bolzer, A., Kreth, G., Solovei, I., Koehler, D., Saracoglu, K., Fauth, C., Müller, S., Eils, R., Cremer, C., Speicher, M.R., and Cremer, T. (2005). Three-dimensional maps of all chromosomes in human male fibroblast nuclei and prometaphase rosettes. PLoS Biol. *3*, e157.

Brown, K.E., Guest, S.S., Smale, S.T., Hahm, K., Merkenschlager, M., and Fisher, A.G. (1997). Association of transcriptionally silent genes with Ikaros complexes at centromeric heterochromatin. Cell *91*, 845–854.

Brown, J.M., Green, J., das Neves, R.P., Wallace, H.A., Smith, A.J., Hughes, J., Gray, N., Taylor, S., Wood, W.G., Higgs, D.R., et al. (2008). Association between active genes occurs at nuclear speckles and is modulated by chromatin environment. J. Cell Biol. *182*, 1083–1097.

Cheutin, T., and Cavalli, G. (2014). Polycomb silencing: from linear chromatin domains to 3D chromosome folding. Curr. Opin. Genet. Dev. *25*, 30–37.

Chuang, C.H., Carpenter, A.E., Fuchsova, B., Johnson, T., de Lanerolle, P., and Belmont, A.S. (2006). Long-range directional movement of an interphase chromosome site. Curr. Biol. *16*, 825–831.

Chubb, J.R., Boyle, S., Perry, P., and Bickmore, W.A. (2002). Chromatin motion is constrained by association with nuclear compartments in human cells. Curr. Biol. *12*, 439–445.

de Wit, E., and de Laat, W. (2012). A decade of 3C technologies: insights into nuclear organization. Genes Dev. *26*, 11–24.

de Wit, E., Bouwman, B.A., Zhu, Y., Klous, P., Splinter, E., Verstegen, M.J., Krijger, P.H., Festuccia, N., Nora, E.P., Welling, M., et al. (2013). The pluripotent genome in three dimensions is shaped around pluripotency factors. Nature *501*, 227–231.

Deng, W., Lee, J., Wang, H., Miller, J., Reik, A., Gregory, P.D., Dean, A., and Blobel, G.A. (2012). Controlling long-range genomic interactions at a native locus by targeted tethering of a looping factor. Cell *149*, 1233–1244.

Denholtz, M., Bonora, G., Chronis, C., Splinter, E., de Laat, W., Ernst, J., Pellegrini, M., and Plath, K. (2013). Long-range chromatin contacts in embryonic stem cells reveal a role for pluripotency factors and polycomb proteins in genome organization. Cell Stem Cell *13*, 602–616.

Dixon, J.R., Selvaraj, S., Yue, F., Kim, A., Li, Y., Shen, Y., Hu, M., Liu, J.S., and Ren, B. (2012). Topological domains in mammalian genomes identified by analysis of chromatin interactions. Nature *485*, 376–380.

Dixon, J.R., Jung, I., Selvaraj, S., Shen, Y., Antosiewicz-Bourget, J.E., Lee, A.Y., Ye, Z., Kim, A., Rajagopal, N., Xie, W., et al. (2015). Chromatin architecture reorganization during stem cell differentiation. Nature *518*, 331–336.

Dubarry, M., Loïodice, I., Chen, C.L., Thermes, C., and Taddei, A. (2011). Tight protein-DNA interactions favor gene silencing. Genes Dev. *25*, 1365–1370.

Francis, N.J., Kingston, R.E., and Woodcock, C.L. (2004). Chromatin compaction by a polycomb group protein complex. Science *306*, 1574–1577.

Geeven, G., Zhu, Y., Kim, B.J., Bartholdy, B.A., Yang, S.M., Macfarlan, T.S., Gifford, W.D., Pfaff, S.L., Verstegen, M.J., Pinto, H., et al. (2015). Local compartment changes and regulatory landscape alterations in histone H1-depleted cells. Genome Biol. *16*, 289.

Grewal, S.I., and Jia, S. (2007). Heterochromatin revisited. Nat. Rev. Genet. *8*, 35–46.

Guelen, L., Pagie, L., Brasset, E., Meuleman, W., Faza, M.B., Talhout, W., Eussen, B.H., de Klein, A., Wessels, L., de Laat, W., and van Steensel, B. (2008). Domain organization of human chromosomes revealed by mapping of nuclear lamina interactions. Nature *453*, 948–951.

Hakim, O., Sung, M.H., Voss, T.C., Splinter, E., John, S., Sabo, P.J., Thurman, R.E., Stamatoyannopoulos, J.A., de Laat, W., and Hager, G.L. (2011). Diverse gene reprogramming events occur in the same spatial clusters of distal regulatory elements. Genome Res. *21*, 697–706.

Hawkins, R.D., Hon, G.C., Lee, L.K., Ngo, Q., Lister, R., Pelizzola, M., Edsall, L.E., Kuan, S., Luu, Y., Klugman, S., et al. (2010). Distinct epigenomic landscapes of pluripotent and lineage-committed human cells. Cell Stem Cell *6*, 479–491.

Kind, J., Pagie, L., Ortabozkoyun, H., Boyle, S., de Vries, S.S., Janssen, H., Amendola, M., Nolen, L.D., Bickmore, W.A., and van Steensel, B. (2013). Single-cell dynamics of genome-nuclear lamina interactions. Cell *153*, 178–192.

Krijger, P.H., and de Laat, W. (2013). Identical cells with different 3D genomes; cause and consequences? Curr. Opin. Genet. Dev. *23*, 191–196.

Ku, M., Koche, R.P., Rheinbay, E., Mendenhall, E.M., Endoh, M., Mikkelsen, T.S., Presser, A., Nusbaum, C., Xie, X., Chi, A.S., et al. (2008). Genomewide analysis of PRC1 and PRC2 occupancy identifies two classes of bivalent domains. PLoS Genet. *4*, e1000242.

Kumaran, R.I., and Spector, D.L. (2008). A genetic locus targeted to the nuclear periphery in living cells maintains its transcriptional competence. J. Cell Biol. *180*, 51–65.

Liang, J., Wan, M., Zhang, Y., Gu, P., Xin, H., Jung, S.Y., Qin, J., Wong, J., Cooney, A.J., Liu, D., and Songyang, Z. (2008). Nanog and Oct4 associate with unique transcriptional repression complexes in embryonic stem cells. Nat. Cell Biol. *10*, 731–739.

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

Muramatsu, D., Singh, P.B., Kimura, H., Tachibana, M., and Shinkai, Y. (2013). Pericentric heterochromatin generated by HP1 protein interaction-defective histone methyltransferase Suv39h1. J. Biol. Chem. *288*, 25285–25296.

Németh, A., Conesa, A., Santoyo-Lopez, J., Medina, I., Montaner, D., Péterfia, B., Solovei, I., Cremer, T., Dopazo, J., and Längst, G. (2010). Initial genomics of the human nucleolus. PLoS Genet. *6*, e1000889.

Noordermeer, D., de Wit, E., Klous, P., van de Werken, H., Simonis, M., Lopez-Jones, M., Eussen, B., de Klein, A., Singer, R.H., and de Laat, W. (2011). Variegated gene expression caused by cell-specific long-range DNA interactions. Nat. Cell Biol. *13*, 944–951.

Nora, E.P., Lajoie, B.R., Schulz, E.G., Giorgetti, L., Okamoto, I., Servant, N., Piolot, T., van Berkum, N.L., Meisig, J., Sedat, J., et al. (2012). Spatial partitioning of the regulatory landscape of the X-inactivation centre. Nature *485*, 381–385.

Palstra, R.J., Simonis, M., Klous, P., Brasset, E., Eijkelkamp, B., and de Laat, W. (2008). Maintenance of long-range DNA interactions after inhibition of ongoing RNA polymerase II transcription. PLoS One *3*, e1661.

Peric-Hupkes, D., Meuleman, W., Pagie, L., Bruggeman, S.W., Solovei, I., Brugman, W., Gräf, S., Flicek, P., Kerkhoven, R.M., van Lohuizen, M., et al. (2010). Molecular maps of the reorganization of genome-nuclear lamina interactions during differentiation. Mol. Cell *38*, 603–613.

Phillips-Cremins, J.E., Sauria, M.E., Sanyal, A., Gerasimova, T.I., Lajoie, B.R., Bell, J.S., Ong, C.T., Hookway, T.A., Guo, C., Sun, Y., et al. (2013). Architectural protein subclasses shape 3D organization of genomes during lineage commitment. Cell *153*, 1281–1295.

Rao, S.S., Huntley, M.H., Durand, N.C., Stamenova, E.K., Bochkov, I.D., Robinson, J.T., Sanborn, A.L., Machol, I., Omer, A.D., Lander, E.S., and Aiden, E.L. (2014). A 3D map of the human genome at kilobase resolution reveals principles of chromatin looping. Cell *159*, 1665–1680.

Shen, Y., Yue, F., McCleary, D.F., Ye, Z., Edsall, L., Kuan, S., Wagner, U., Dixon, J., Lee, L., Lobanenkov, V.V., and Ren, B. (2012). A map of the cis-regulatory sequences in the mouse genome. Nature *488*, 116–120.

Simon, J.A., and Kingston, R.E. (2009). Mechanisms of polycomb gene silencing: knowns and unknowns. Nat. Rev. Mol. Cell Biol. *10*, 697–708.

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

Splinter, E., de Wit, E., Nora, E.P., Klous, P., van de Werken, H.J., Zhu, Y., Kaaij, L.J., van Ijcken, W., Gribnau, J., Heard, E., and de Laat, W. (2011). The inactive X chromosome adopts a unique three-dimensional conformation that is dependent on Xist RNA. Genes Dev. *25*, 1371–1383.

Splinter, E., de Wit, E., van de Werken, H.J., Klous, P., and de Laat, W. (2012). Determining long-range chromatin interactions for selected genomic sites using 4C-seq technology: from fixation to computation. Methods *58*, 221–230.

Tanay, A., and Cavalli, G. (2013). Chromosomal domains: epigenetic contexts and functional implications of genomic compartmentalization. Curr. Opin. Genet. Dev. *23*, 197–203.

Therizols, P., Illingworth, R.S., Courilleau, C., Boyle, S., Wood, A.J., and Bickmore, W.A. (2014). Chromatin decondensation is sufficient to alter nuclear organization in embryonic stem cells. Science *346*, 1238–1242.

Tolhuis, B., Blom, M., Kerkhoven, R.M., Pagie, L., Teunissen, H., Nieuwland, M., Simonis, M., de Laat, W., van Lohuizen, M., and van Steensel, B. (2011). Interactions among Polycomb domains are guided by chromosome architecture. PLoS Genet. *7*, e1001343.

van Koningsbruggen, S., Gierlinski, M., Schofield, P., Martin, D., Barton, G.J., Ariyurek, Y., den Dunnen, J.T., and Lamond, A.I. (2010). High-resolution whole-genome sequencing reveals that specific chromatin domains from most human chromosomes associate with nucleoli. Mol. Biol. Cell *21*, 3735–3748.

Vieux-Rochas, M., Fabre, P.J., Leleu, M., Duboule, D., and Noordermeer, D. (2015). Clustering of mammalian Hox genes with other H3K27me3 targets within an active nuclear domain. Proc. Natl. Acad. Sci. USA *112*, 4672–4677.

Wijchers, P.J., Geeven, G., Eyres, M., Bergsma, A.J., Janssen, M., Verstegen, M., Zhu, Y., Schell, Y., Vermeulen, C., de Wit, E., and de Laat, W. (2015). Characterization and dynamics of pericentromere-associated domains in mice. Genome Res. *25*, 958–969.

Williams, R.R., Azuara, V., Perry, P., Sauer, S., Dvorkina, M., Jørgensen, H., Roix, J., McQueen, P., Misteli, T., Merkenschlager, M., and Fisher, A.G. (2006). Neural induction promotes large-scale chromatin reorganisation of the Mash1 locus. J. Cell Sci. *119*, 132–140.

Molecular Cell

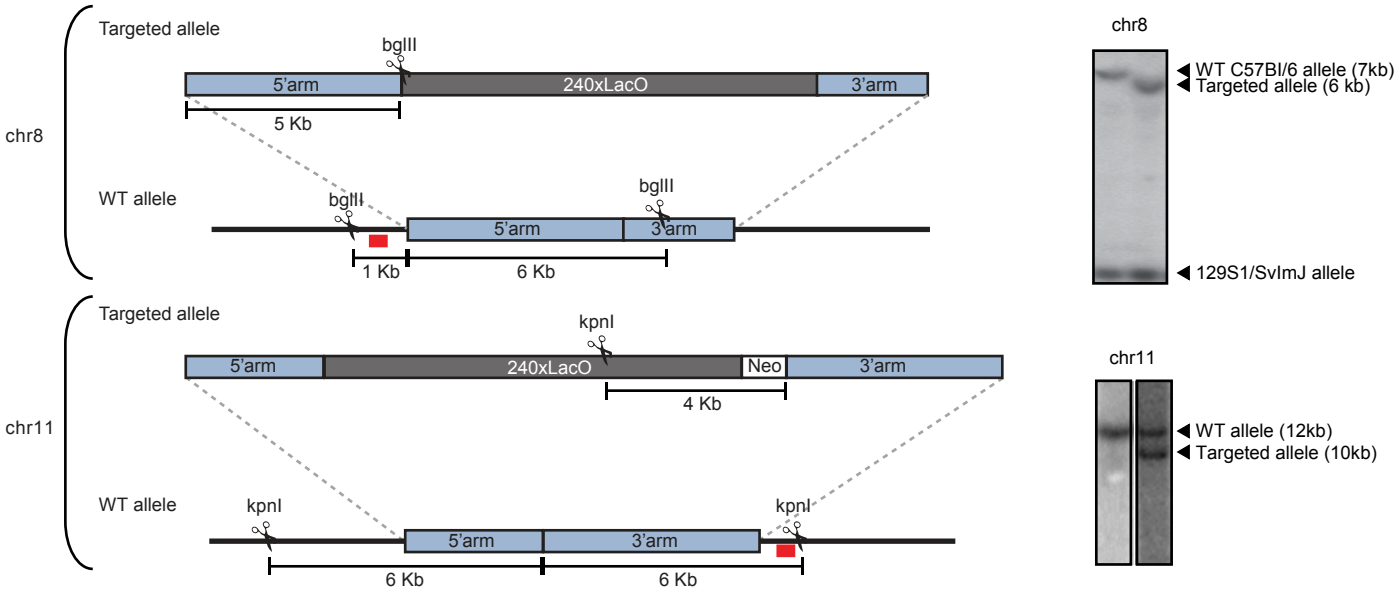
Supplemental Information

Cause and Consequence of Tethering a SubTAD to Different Nuclear Compartments

Patrick J. Wijchers, Peter H. L. Krijger, Geert Geeven, Yun Zhu, Annette Denker, Marjon J. A. M. Verstegen, Christian Valdes-Quezada, Carlo Vermeulen, Mark Janssen, Hans Teunissen, Lisette C. M. Anink-Groenen, Pernette J. Verschure, and Wouter de Laat

Figure S1

A


Targeted allele
bglII
5'arm
240xLacO
3'arm
5 Kb
chr8
WT allele
bglII
bglII
5'arm
3'arm
1 Kb
6 Kb
chr8
WT C57Bl/6 allele (7kb)
Targeted allele (6 kb)
129S1/SvImJ allele
Targeted allele
kpnI
5'arm
240xLacO
Neo
3'arm
4 Kb
chr11
WT allele
kpnI
kpnI
5'arm
3'arm
6 Kb
6 Kb
chr11
WT allele (12kb)
Targeted allele (10kb)

B

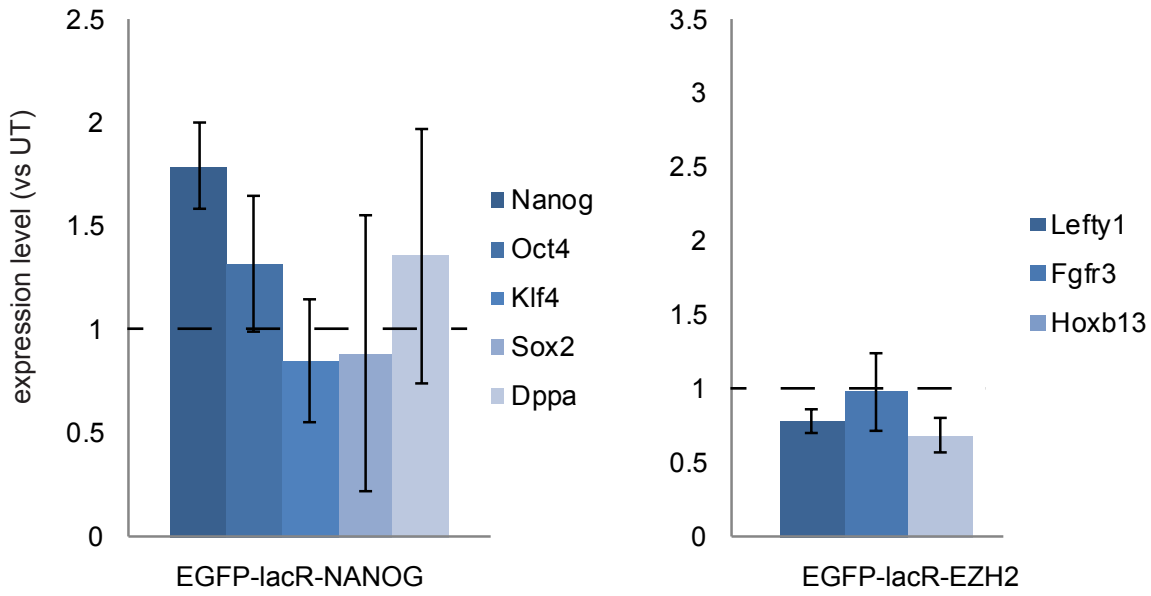

expression level (vs UT)
2.5
2
1.5
1
0.5
0
Nanog
Oct4
Klf4
Sox2
Dppa
EGFP-lacR-NANOG
3.5
3
2.5
2
1.5
1
0.5
0
Lefty1
Fgfr3
Hoxb13
EGFP-lacR-EZH2

Figure S1, related to Figure 1. A lacO/lacR recruitment platform to induce nuclear repositioning.

A) Strategy and data from Southern blot analysis confirming correct integration of the *lacO* cassette on chr8 (top) and chr11 (bottom). Red blocks indicate probes used for hybridization.

B) Quantitative expression analysis (qPCR on cDNA) for a number of endogenous target genes of Nanog and of Ezh2. The analysis reveals that the ectopic expression of the corresponding LacR fusion genes has little impact on the expression of the presumed target genes. Note that the primers used for Nanog also analyze transgene expression, showing that there is less than 2-fold overexpression. Error bars indicate standard deviation (SD).

Figure S2

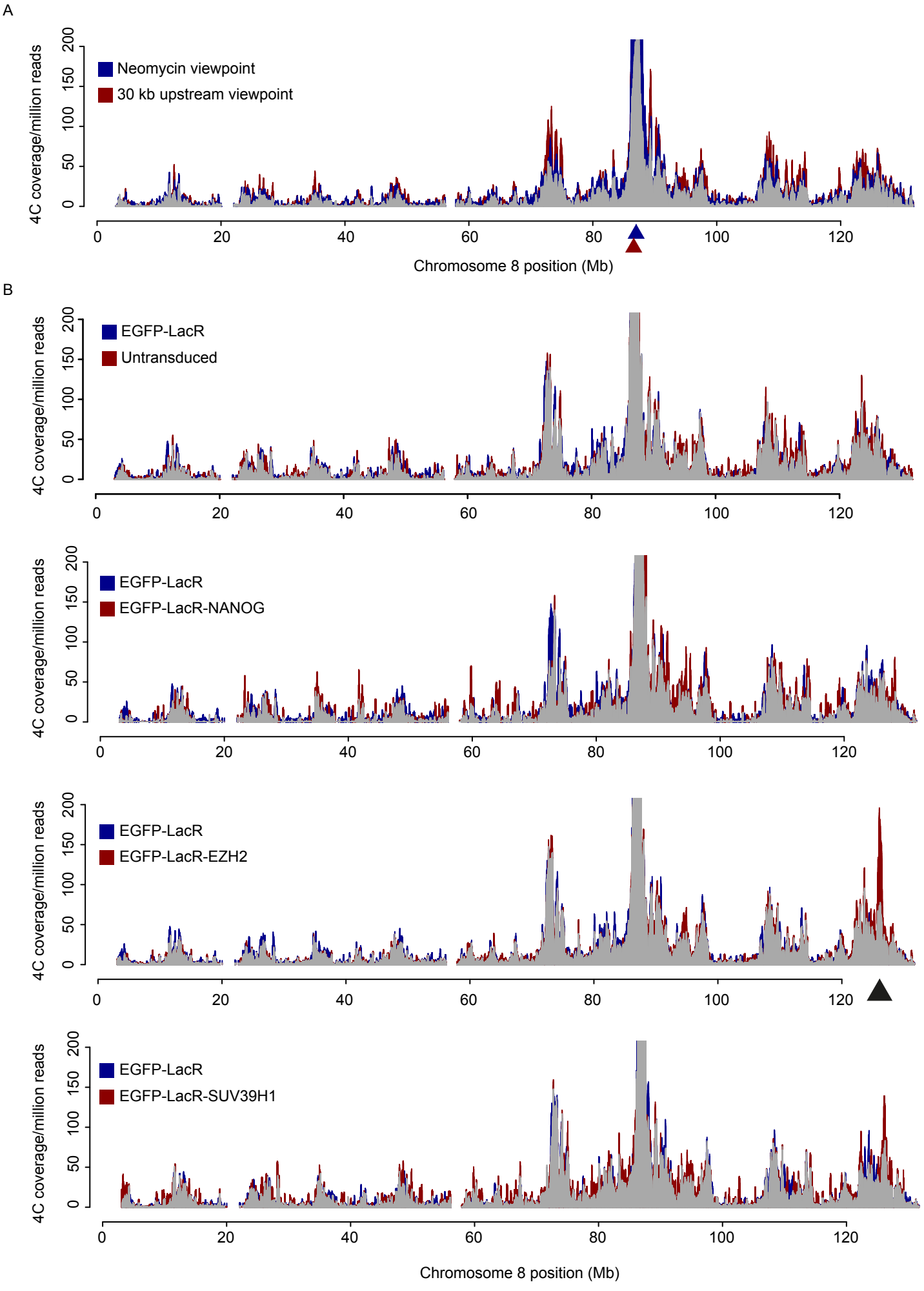

A
4C coverage/million reads
Neomycin viewpoint
30 kb upstream viewpoint
Chromosome 8 position (Mb)
B
EGFP-LacR
Untransduced
EGFP-LacR-NANOG
EGFP-LacR-EZH2
EGFP-LacR-SUV39H1

Figure S2, related to Figure 2. Susceptibility to spatial repositioning depends on genomic location and trans-acting factors.

A) Overlays of 4C profiles on chromosome 8, comparing contacts made by the '*Neomycin*' gene in untransduced ES cells carrying only a *lacO-Neomycin* transgene on chr8, *versus* contacts made by a viewpoint '30 kb upstream' of the integration site. Viewpoints are indicated by arrowheads. Note that the profiles show some quantitative differences in contact frequencies, but only at the chromosomal sites that are contacted by both viewpoints (no 'new' interactions).

B) As in A) but for the double-transgenic *lacO* cells transduced with the different EGFP-lacR fusions. Note here that all contact profiles are highly similar and that only upon EZH2 recruitment a single 'new' contact is seen, near the telomere of chromosome 8 (indicated with an arrowhead).

Figure S3

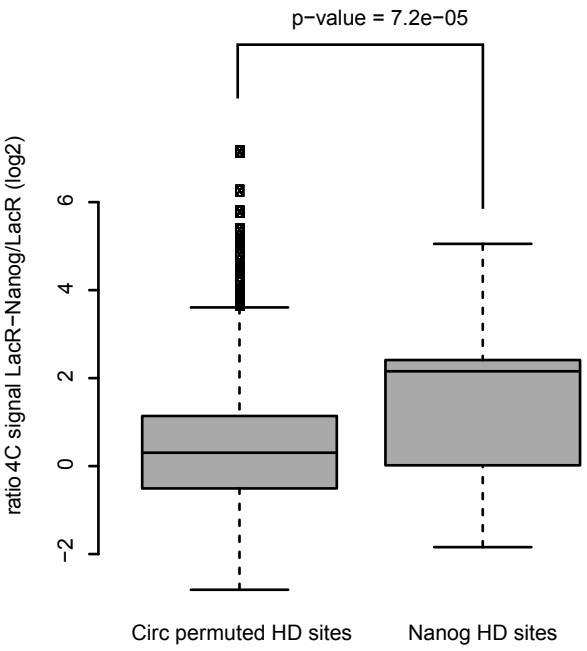
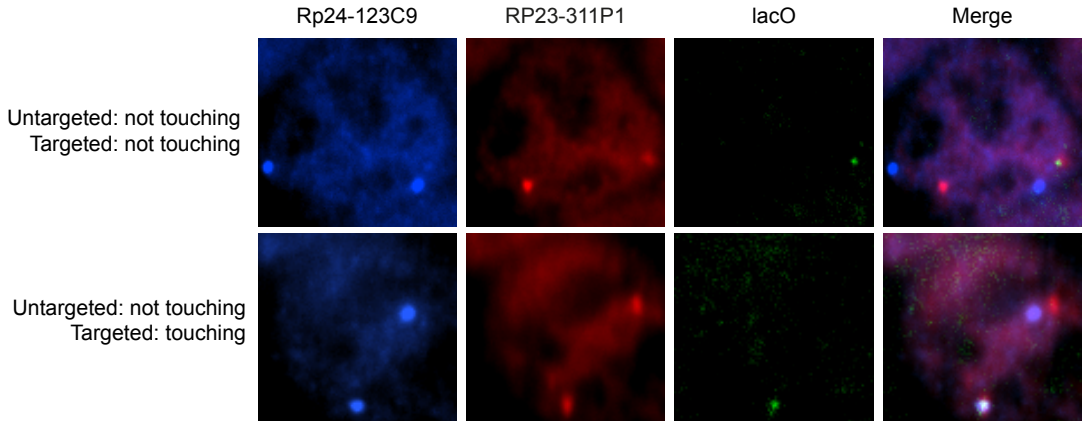

p−value = 7.2e−05
ratio 4C signal LacR−Nanog/LacR (log2)
−2
0
2
4
6
Circ permuted HD sites
Nanog HD sites

Figure S3, related to Figure 3. Local chromatin signature influences direction of nuclear repositioning.

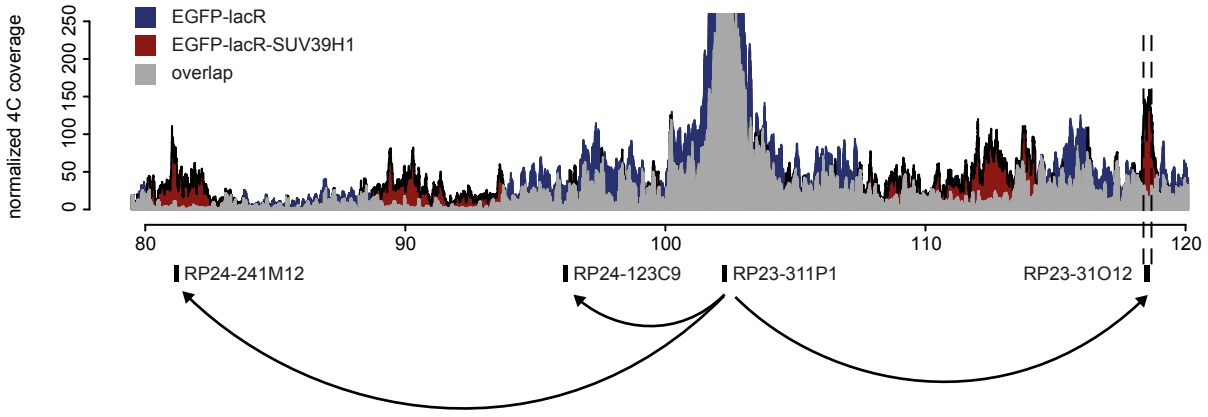
Boxplot showing the ratio in 4C signals of EGFP-lacR-NANOG and EGFP-lacR cells across high-density NANOG binding sites and upon circular permutation of these positions. Note that NANOG-binding significantly increased contact frequency on high-density NANOG binding sites.

Figure S4

A



B



	RP24-241M12	RP24-123C9	RP23-31O12
SUV39H1	44.8% (47/105)	29.3% (22/75)	46.7% (43/92)
Untargeted	25.6% (66/258)	32.4% (155/478)	32.2% (86/267)
Untransduced	21.2% (11/52)	34.1% (29/85)	27.8% (25/90)
EZH2	n.d.	72.5% (95/131)	n.d.

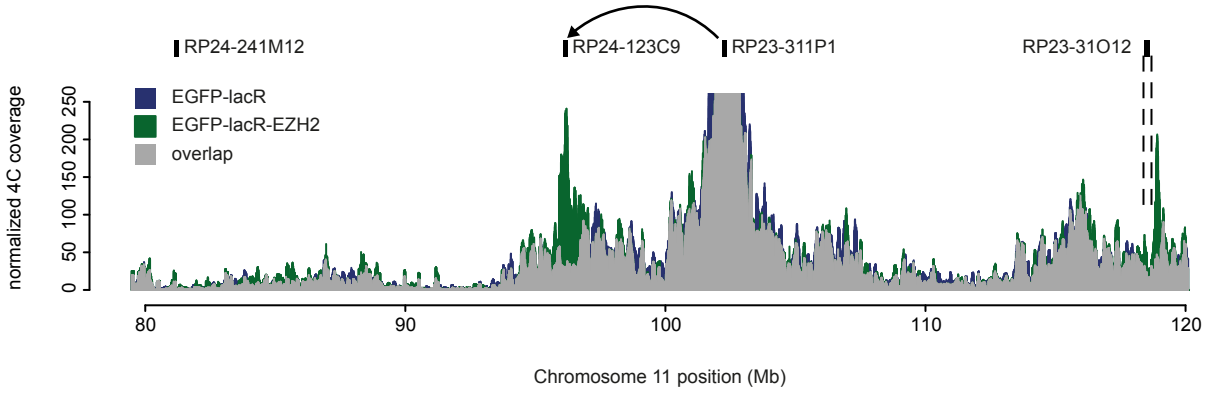


Figure S4. FISH validation of 4C contacts. Related to Figure 3 and Figure 4.

A) Examples of FISH results. Shown are (top) a cell in which, upon LacR-EZH2 expression, neither the untargeted nor the targeted site on chromosome 11 visually touches the *HoxB* locus 6 Mb upstream of *lacO* on chromosome 11, and (bottom) one in which only the *LacO* targeted site touches *HoxB* (bottom). Shown are z-stacks recording the signals of the different probes (three leftmost panels), and their merge (right panels).

B) Summary of the quantification of the FISH results. ImageJ (Image5D plug-in) software was used to count touching (no unstained pixel in between FISH signals) and non-touching (with unstained pixel(s) in between) alleles. Depending on the BAC probe used, the average diameter of the fluorescent signal varied between 0.8-1µm, implying that the centers of touching signals can be up to 1µm apart. Location and name of BAC is indicated, and the percentage of touching FISH signals is indicated for the 'untargeted' chromosome (i.e. the homologous chromosome that does not carry the *lacO* array), the *lacO* untransduced allele (which carries *lacO* without any associated LacR (fusion) protein), the LacR-EZH2 associated *lacO* allele and the LacR-SUV39H1 associated *LacO* allele. The plotted differential 4C contact profiles comparing LacR (blue) versus LacR-SUV39H1 (red) (top) and LacR (blue) versus LacR-EZH2 (green) (bottom) enable a direct comparison between differential contact frequencies measured by 4C and by FISH. Note that the most telomeric BAC probe hybridizes to the SUV39H1 contacted region, not to the *Cbx* locus contacted by EZH2. N.d.: not done.

Figure S5

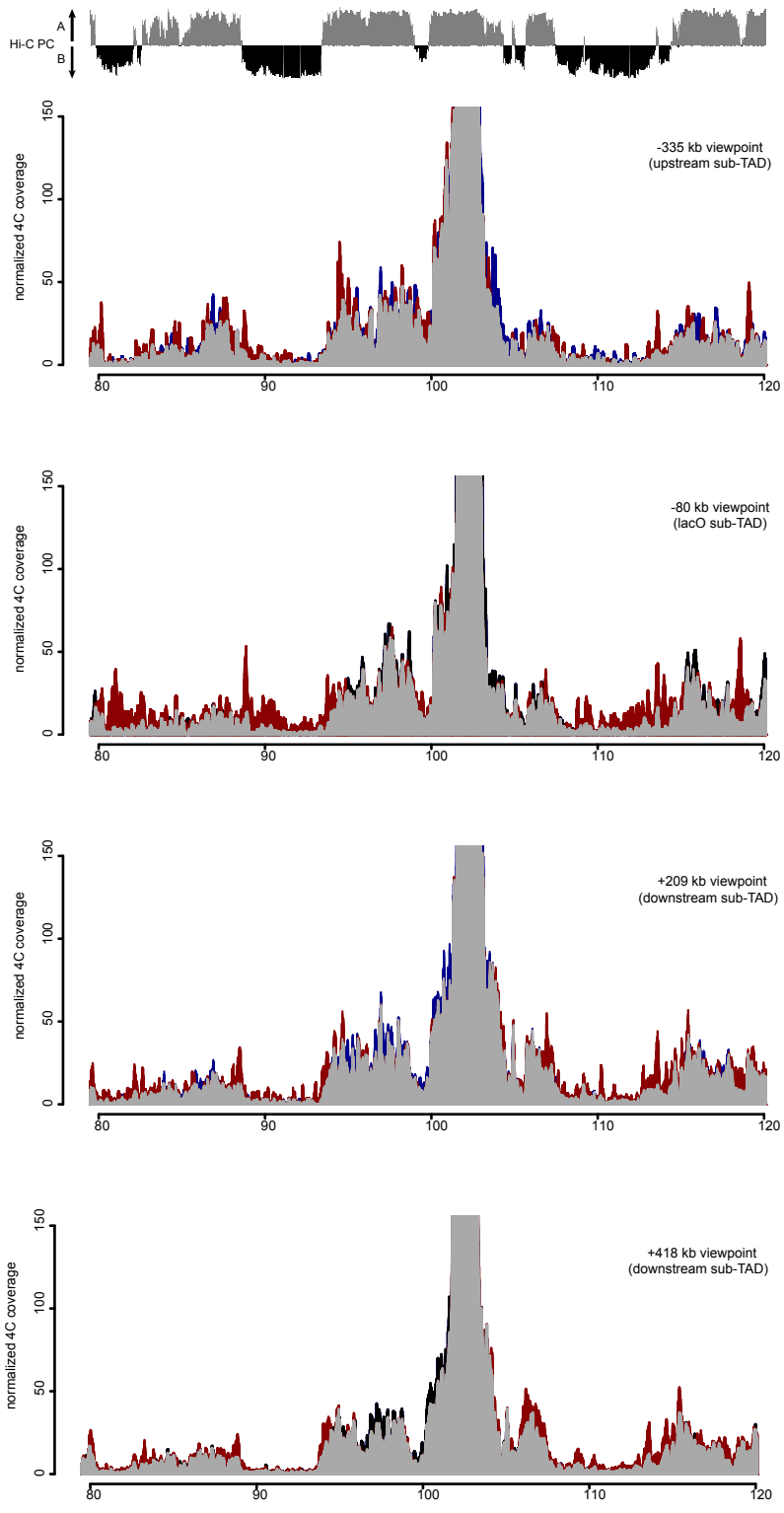
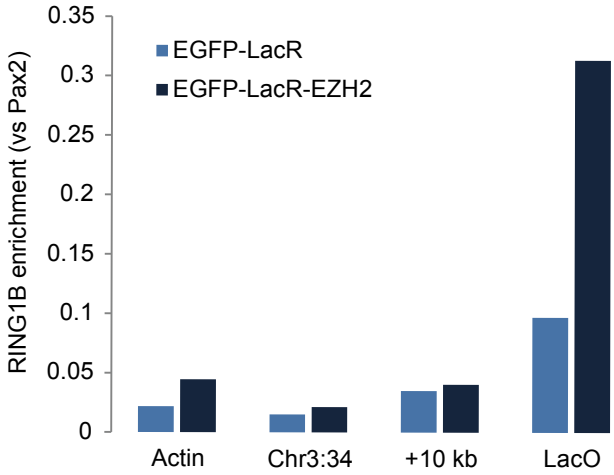

Hi-C PC
A
B
normalized 4C coverage
-335 kb viewpoint
(upstream sub-TAD)
-80 kb viewpoint
(lacO sub-TAD)
+209 kb viewpoint
(downstream sub-TAD)
+418 kb viewpoint
(downstream sub-TAD)
0
50
100
150
80
90
100
110
120

Figure S5, related to Figure 4. A repositioned locus drags along its associated sub-TAD.

Comparative allele-specific 4C profiles comparing the lacO and wild type alleles upon binding of SUV39H1, for a viewpoint 335 kb upstream of the lacO in the neighbouring sub-TAD, 80 kb upstream within the same sub-TAD and 209 and 418 kb downstream in immediately neighbouring sub-TADs. Note the increased contact frequency with regions in the B compartment (displayed on top) only for the viewpoint within the *lacO* sub-TAD.

Figure S6

A


0.35
0.3
0.25
0.2
0.15
0.1
0.05
0
RING1B enrichment (vs Pax2)
EGFP-LacR
EGFP-LacR-EZH2
Actin
Chr3:34
+10 kb
LacO

B

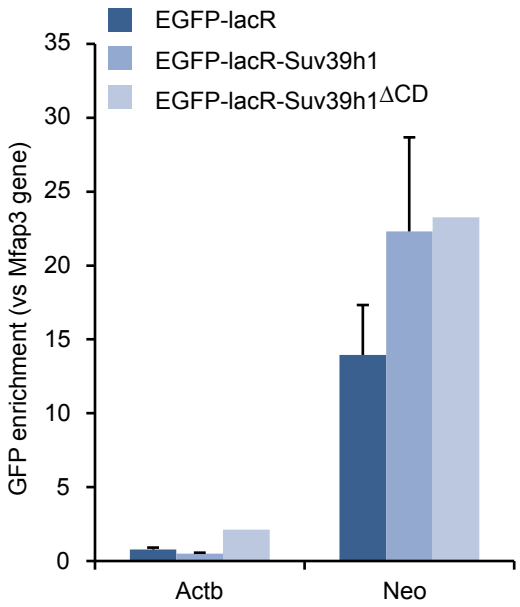

35
30
25
20
15
10
5
0
GFP enrichment (vs Mfap3 gene)
EGFP-lacR
EGFP-lacR-Suv39h1
EGFP-lacR-Suv39h1ΔCD
Actb
Neo

Figure S6, related to Figure 6. The role of chromatin in spatial repositioning.

A) ChIP data showing that EZH2 recruitment to LacO facilitates Ring1B association.

B) ChIP data confirming the binding of EGFP-lacR-SUV39H1 and EGFP-lacR- SUV39H1$^{\Delta CD}$ to the *Neo* gene on the recruitment platform. Error bars indicate standard deviation (SD).

Supplemental Experimental Procedures

Cell culture

Mouse embryonic stem cells C57Bl/6-129/Sv were cultured as previously described (de Wit et al., 2013). In brief, they were cultured on gelatin-coated plates in BRL (buffalo–rat liver cells)-conditioned DMEM (Hooper et al., 1987)(high glucose, Gibco) with 15% FBS, 1x non-essential amino acids (NEAA; Gibco), 1x penicillin–streptomycin (Gibco), 1:1000 b-mercaptoethanol (Invitrogen), 1x L-glutamine (Gibco) and 1000 U ml-1 leukaemia inhibitory factor (LIF; Gibco).

Generation of *lacO* targeted cell line

First, the *lacO* array (Lau et al., 2003) was inserted in chromosome 8 (chr8)(AatII site, mm9 chr8:87321244), as described (de Wit et al., 2013). Cre transfection (Splinter et al., 2006) was used to delete *Neomycin* at Chr8. Subcolonies were picked, and deletion was confirmed by PCR on genomic DNA from clonal cell lines. Then, the same *lacO* array was introduced into chromosome 11 (AatII site, mm9 chr11:102209489) using the same gene targeting protocol. Homology arms were excised with KpnI from bacterial artificial chromosome (BAC) RP23-311P1, and the *lacO* array inserted at the unique AatII site. The linearized targeting construct was introduced by electroporation. After 14 days of neomycin selection, positive colonies were screened by Southern blot (Figure S1). Probes for Southern blot were PCR-amplified from genomic DNA using the following primers:
chr8_forward: TGTGTGGTGATCATGTGTGC
chr8_reverse: TGCCACTCCTGTGTCTCAAG
chr11_forward: CCTCCTTTGGATACCTTCC
chr11_reverse: CTTTAAATCGGTGGCTGAGG

LacR-fusion constructs and transduction

Proteins of interest were cloned in frame downstream of EGFP-LacR . To transduce EGFP-LacR fusions under control of the EF1α promoter, they were introduced in place of the DsRed gene of the phage2-EF1α-DsRed-IRES-PURO vector (Wilson et al., 2008) by blunt cloning. The chromodomain of the EGFP-lacR-SUV39H1 fusion protein was removed using BstZ17I (NEB) and SphI (Roche), and re-ligated after Klenow fill in. *LacO*-transgenic cells were transduced and selected with Puromycin (P8833, SIGMA at 1µg/µl) for approximately 10 days when cells had reached sufficient numbers for collection and tested for purity by flow cytometry (minimum 70% GFP-positive).

Western blot

Nuclear extracts were made as described in (Andrews and Faller, 1991), except for EGFP-lacR-SUV39H1, where nuclei were boiled for 5 minutes in Laemmli buffer. Immunoblot analyses were performed with antibodies against GFP (ab290, Abcam) or NANOG (A300-397A, Bethyl Laboratories).

GFP distribution analysis

Stably transduced *lacO* cell lines were grown overnight on gelatin-coated coverslips. They were crosslinked with 4% paraformaldehyde (rT) (10 minutes), washed once with 0.125M glycine in PBS and PBS+0.1% Tween 20, and permeabilized with 0.2% Triton X-100 in PBS (5 minutes). After a final wash with PBS+0.1% Tween 20, VectaShield containing DAPI (Vector Labs) was applied and coverslips were sealed with nail polish. Images were taken with a Leica SPE confocal microscope and analyzed using ImageJ software. Maximum projections were made to allow simultaneous visualization of both *lacO* alleles.

Chromatin immunoprecipitation

Chromatin immunoprecipitation (ChIP) was done using Millipore's protocol with minor modifications. In brief, 5-10 million cells were fixed with 1% formaldehyde, then lysed in cell lysis buffer (10 mM Tris, pH 8.0; 10 mM NaCl; 0.2% NP-40; 10 mM Na-Butyrate; Proteinase inhibitor 1x) and nuclei lysis buffer (50 mM Tris, pH 8.0; 10 mM EDTA; 1% SDS; 10 mM Na-Butyrate; Proteinase inhibitor 1x). Isolated chromatin was sonicated to 500-1000 bp, processed on Covaris S2 as 30 seconds on and 30 seconds off, and then precipitated with GFP antibody overnight. After washing, chromatin was eluted, purified, and used for quantification analysis by qPCR, and data normalized to input and enrichment plotted over the untransduced cells. ChIP for histone modifications were performed based on the protocol in (Schmidt et al., 2009) with slight modifications. Briefly, cross-linked cells were lysed (50 mM Tris, pH 7.5; 150 mM NaCl; 5mM EDTA; 0.5% NP-40; 1% Triton X-100) and sonicated to 500-1000bp using a Bioruptor (Diagenode, 15 seconds on/15 seconds off) or a Covaris (20 cycles 30 seconds on/30 seconds off).

Chromatin equivalents of five million cells were used per IP, with antibodies against GFP (ab290, Abcam), H3K27me3 (ab6002, Abcam), H3K9me3 (ab8898, Abcam) and RING1B (#5694, Cell Signaling Tech). After immunoprecipitation, isolated DNA samples were used for quantification analysis by qPCR, and data normalized to input and enrichment calculated over plotted over the (transcriptionally active) actB promoter. Primers used:

genomic site	forward	reverse
Actb	GCAGGCCTAGTAACCGAGACA	AGTTTTGGCGATGGGTGCT
Pax2	CGGAGGGAATGAAGCAGGTT	GTTTCGAAGAGGTTCCCCGT
Mfap3	TCTGCAAGGAAGGCATCAGG	ACTCTTTCCCCCTCCCCTTT
lacO	TTCGATACCTTTATCCGCTCA	GCGGATAACAATTGCTGAAG
kan^{R}	TGATAATCCTGATATGAATAAATTGC	TTGGCACCTTTGCTAGATTAGAA
-90 kb	ATTTTGCCTGCTGTGTGCAG	TCCCTTCTCCACAGGGACAT
-60 kb	ACTATGTGGTCTTGGCTGGC	GGAGTGGCAGGAGAGAGGTA
-10 kb	GAGTAAGCCTGACGCCTGTT	AGACACACACTGTCCTGGTG
+10 kb	GACTTGCTCTCACCCCAACA	ATCTGGGTGAAAGGTGCCTG
+60 kb	GCCAGCACTGACTACACCTT	GCTGGGCTGAGCAGTGAATA
+90 kb	GGCTGATGGGGTGTGTTACC	ACCCACTTTCAGACGCAAGA
Chr3:34	GAACTGATCACAGGGACGTG	CAGGCCTGGTGAAGTATCTG

Gene expression analysis

RNA was isolated using TRIzol (Life technology) from cells trypsinized and then converted into cDNA using random primers (Promega) using standard manufacturer providing protocol (Promega). Quantitative PCR was performed and data normalized to the *Actb* gene. Allelic specific gene expression using high throughput sequencing was applied on cDNA synthesized in the same way. Primers used to detect gene transcripts flanked a SNP to discriminate C57Bl/6 allele from the 129S1/SvImJ allele. PCR product amplified from cDNA was sequenced on Illumina GAII. Sequencing reads were mapped to each allele to determine the fraction of reads coming from the *lacO*-transgenic BL/6 allele. Fold-changes were expressed relative to the BL/6 fraction in untransduced cells.

Primer used:

genomic site	forward	reverse
hprt	TCCTCCTCAGACCGCTTTT	CCTGGTTCATCATCGCTAATC
neo (set-1)	ATGCCTGCTTGCCGAATA	CCACAGTCGATGAATCCAGA
neo (set-2)	GCAGGATCTCCTGTCATCTCA	TAGCCGGATCAAGCGTATG
Grn	GCCCGTTCTCTAAGGGTGTG	ACAGCACCCAAGGGGTTATC
Slc25a39	AGGCAGTATCTTGGCCCCAT	GCACACGTACCCCAAGACA
mRundc3a	AAGGGCGAAGTTCTGGATGG	GGTAGTCGTAGCTTTGGGTGA
Atxn7l3	GGTTTGGGGCTCTGAGGAAA	TGGGAGGTGGGATACAGGTC
Ubtf	GCTCCTCTAACTGCTTGCCA	GGAGAGCCTACTTCCCACCT
Oct4	CTCCCTACAGCAGATCACTC	GAACCATACTCGAACCACAT
Nanog	CCATTCTGAACCTGAGCTAT	ACCATTGCTAGTCTTCAACC
Klf4	CAGGTACCCCTCTCTCTTCT	TGACAGCCATGTCAGACTC
Sox2	GGAGCAACGGCAGCTA	GTAGCGGTGCATCGGT
Dppa	ACGCCAGGACAGACTCGTAG	TGCTGCTCACTCGTTTCTTCT
Lefty	CAGCTCGATCAACCGCCAGT	GGCTGGCATGGCTGTGTT
Fgfr3	CTTAAGCGACAGGTGTCCT	CTGGATAGCTCCCACTTGG
HoxB13	CGTTTGCAGAGCCCAGTGTC	CTGCATACTCCCGCTCCAAC

4C-seq analysis

Standard 4C experiments were done as previously described (Splinter et al., 2012). For allele specific 4C we have used a paired-end 4C strategy (Holwerda et al., 2013), where the forward primer analyses the ligation product and the reverse primer flanks a SNP. After sequencing, this SNP is used to demultiplex the two alleles, to create two separate 4C profiles. Allelic specific 4C from the chr8 *lacO* viewpoint and reciprocal single-end 4C used a combination of HindIII and DpnII, single-end 4C from chr11 *lacO* viewpoint used a combination of HindIII and NlaIII. 4C data mapping and analysis was as described (de Wit et al., 2013). To allow direct comparisons of 4C profiles of different samples, data was normalized to equalize the number of reads along the chromosomes, and plotted as the 4C coverage per million mapped reads. Primers used:

	primer 1	primer 2
chr8	AATGATACGGCGACCACCGAG ATCTACACTCTTTCCCTACACG ACGCTCTTCCGATCTCTGGAAC TAAATGGAGGATC	CAAGCAGAAGACGGCATACGA GATCGGTCTCGGCATTCCTGCT GAACCGCTCTTCCGATCTCAA GCAGAAGACGGCATACGAGAT CGGTCTCGGCATTCCTGCTGAA CCGCTCTTCCGATCTTACCAGG ACCCCTGGGACCC
Neo	AATGATACGGCGACCACCGAA CACTCTTTCCCTACACGACGCT CTTCCGATCTCGAAGTTATCGA TCGAAGCTT	CAAGCAGAAGACGGCATACGA AGAAAAGCGGCCATTTTCCA
HoxB	AATGATACGGCGACCACCGAA CACTCTTTCCCTACACGACGCT CTTCCGATCTTCCCCTGGATGA GGAAGCTT	CAAGCAGAAGACGGCATACGA GAGCGGTTGACGCTGAGATC
Asci2	AATGATACGGCGACCACCGAA CACTCTTTCCCTACACGACGCT CTTCCGATCTTGTTAGGTGGCA CCAAGCTT	CAAGCAGAAGACGGCATACGA ACTAATGATGGGCACAGTTT
-335 kb viewpoint	AATGATACGGCGACCACCGAA CACTCTTTCCCTACACGACGCT CTTCCGATCTTGCACTGTCAGC CCAAGCTT	CAAGCAGAAGACGGCATACGA GATCGGTCTCGGCATTCCTGCT GAACCGCTCTTCCGATCTCAGT GCCACAGACTGCCC
-80 kb viewpoint	AATGATACGGCGACCACCGAA CACTCTTTCCCTACACGACGCT CTTCCGATCTAACTTGAGTGGA GAAAGCTT	CAAGCAGAAGACGGCATACGA GATCGGTCTCGGCATTCCTGCT GAACCGCTCTTCCGATCTCCTG GGTCTCTTGTCTACTCA
+209 kb viewpoint	AATGATACGGCGACCACCGAG ATCTACACTCTTTCCCTACACG ACGCTCTTCCGATCTGCTCCGC CTCCTAAAAGCTT	CAAGCAGAAGACGGCATACGA GATCGGTCTCGGCATTCCTGCT GAACCGCTCTTCCGATCTACGT CTTTGCAAGCTATTCGC
+418 kb viewpoint	AATGATACGGCGACCACCGAA CACTCTTTCCCTACACGACGCT CTTCCGATCTGTCTAAAACAGC CTAAGCTT	CAAGCAGAAGACGGCATACGA AAAGCATAGTCATTCAGGGA

Mapping and processing of 4C data

4C sequencing reads were aligned to the mouse reference genome (UCSC, mm9) using a mapping pipeline written in Perl as described in (van de Werken et al., 2012). In short, we first identify 4C captures by trimming primer

sequences for each bait specific 4C sequencing primer. Capture sequences are then aligned to a reduced genome consisting of sequences that flank restriction sites of the primary restriction enzyme (4C fragends). Non-unique 4C fragends were removed in subsequent analysis. Mapped reads are normalized for sequencing depth by multiplying by a factor such that the total sum of mapped reads in cis (i.e. on the bait chromosome), after discarding the 4C fragend with the highest multiplicity, equals one million. For visualization, we compute the average 4C signal across windows of 51 consecutive fragends (centered on the 25th) per experiment. We plot the average of this signal as taken from two independent biological replicates.

Statistical analysis of 4C data

For comparison of two 4C profiles we first binarize the data, such that at each 4C fragend we only record whether it was captured (at least 1 read mapping to the fragend) or not. We then create bins of 50 consecutive fragends and for each profile compute the number of captured fragends in each bin. We select bins in which at least 5 fragends are covered in at least 1 experiment for further analysis. For each bin we test the null hypothesis that the proportion of covered fragends in each experiment is equal with Fisher's exact test. We reject this null hypothesis for bins with a p-value < 0.001.

All statistical analyses were performed under R/Bioconductor (Gentleman et al., 2004). Manipulation with and computation of statistics on genomic intervals and domains was done using the GenomicRanges (Lawrence et al., 2013) package.

For comparative analysis, Hi-C and principle component analysis data from embryonic stem cells was taken from Geeven, Zhu et al., in prep., density maps of Nanog from (Marson et al., 2008), Smc1 ChIP-seq data from (Kagey et al., 2010), RNA-seq and CTCF, H3K4me3, H3K27me3 and H3K9me3 ChIP seq data from the ENCODE project (The ENCODE Project Consortium, 2011).

FISH

The FISH procedure was adapted from (Bienko et al., 2013; Solovei and Cremer, 2010). In short, cells were grown overnight on gelatin coated coverslips, fixed in 4% paraformaldehyde (441244, Sigma) for 10 minutes, permeabilised with 0.5% Triton X-100 for 10 minutes, and then subjected them to four freeze-thaw cycles in 20% glycerol/PBS. We incubated the cells in 0.1N HCl for 10 minutes, and subsequently kept them in 50% formamide 2xSSC until hybridization. Locus-specific FISH probes were prepared from BACs as previously described (Splinter et al., 2011). The following BACs were used: RP24-241M12 (chr11:81096552-81279234) (Asci2 locus), RP23-31O12 (chr11:118411048-118618579), RP24-123C9 (chr11:96069113-96245395) (HoxB locus) and RP23-311P1 (chr11:102184420-102354519) (integration site). We designed a custom LacO probe and ordered it as an HPLC purified 3' Quasar-670 linked oligomer from Stellaris (LacO-1: 5'- acctttatccgctcacaattcggttacccctgctagaggt-Quasar670 -3'). Per slide, 1 pM LacO probe was mixed with the appropriate BAC probes before dehydration. Probes were denatured with the cellular DNA for 5 minutes at 80°C and subsequently hybridized for 16 hours at 37 °C in a hybridization buffer consisting of 50% formamide, 2xSSC, 2.5x Denhardt's solution, 50mM EDTA and 10% Dextran Sulphate. After hybridization, cells were washed twice for 30 minutes in 25% formamide 2xSSC at 30°C, counter stained with DAPI and embedded on glass slides with Prolong Gold (P36934, Thermo Scientific) and imaged on a Leica MM-AF microscope. ImageJ (Image5D plug-in) software was used to score touching (no unstained pixel in between FISH signals) and non-touching (with unstained pixel(s) in between) alleles. Depending on the BAC probe used, the average diameter of the fluorescent signal varied between 0.8-1µm, implying that the centers of touching signals can be up to 1µm apart.

Supplemental References

Andrews, N.C., and Faller, D.V. (1991). A rapid micropreparation technique for extraction of DNA-binding proteins from limiting numbers of mammalian cells. Nucleic acids research *19*, 2499.

Bienko, M., Crosetto, N., Teytelman, L., Klemm, S., Itzkovitz, S., and van Oudenaarden, A. (2013). A versatile genome-scale PCR-based pipeline for high-definition DNA FISH. Nature methods *10*, 122-124.

de Wit, E., Bouwman, B.A., Zhu, Y., Klous, P., Splinter, E., Verstegen, M.J., Krijger, P.H., Festuccia, N., Nora, E.P., Welling, M.*, et al.* (2013). The pluripotent genome in three dimensions is shaped around pluripotency factors. Nature *501*, 227-231.

Gentleman, R.C., Carey, V.J., Bates, D.M., Bolstad, B., Dettling, M., Dudoit, S., Ellis, B., Gautier, L., Ge, Y., Gentry, J.*, et al.* (2004). Bioconductor: open software development for computational biology and bioinformatics. Genome biology *5*, R80.

Holwerda, S.J., van de Werken, H.J., Ribeiro de Almeida, C., Bergen, I.M., de Bruijn, M.J., Verstegen, M.J., Simonis, M., Splinter, E., Wijchers, P.J., Hendriks, R.W.*, et al.* (2013). Allelic exclusion of the immunoglobulin heavy chain locus is independent of its nuclear localization in mature B cells. Nucleic acids research *41*, 6905-6916.

Hooper, M., Hardy, K., Handyside, A., Hunter, S., and Monk, M. (1987). HPRT-deficient (Lesch-Nyhan) mouse embryos derived from germline colonization by cultured cells. Nature *326*, 292-295.

Kagey, M.H., Newman, J.J., Bilodeau, S., Zhan, Y., Orlando, D.A., van Berkum, N.L., Ebmeier, C.C., Goossens, J., Rahl, P.B., Levine, S.S.*, et al.* (2010). Mediator and cohesin connect gene expression and chromatin architecture. Nature *467*, 430-435.

Lau, I.F., Filipe, S.R., Soballe, B., Okstad, O.A., Barre, F.X., and Sherratt, D.J. (2003). Spatial and temporal organization of replicating Escherichia coli chromosomes. Molecular microbiology *49*, 731-743.

Lawrence, M., Huber, W., Pages, H., Aboyoun, P., Carlson, M., Gentleman, R., Morgan, M.T., and Carey, V.J. (2013). Software for computing and annotating genomic ranges. PLoS computational biology *9*, e1003118.

Marson, A., Levine, S.S., Cole, M.F., Frampton, G.M., Brambrink, T., Johnstone, S., Guenther, M.G., Johnston, W.K., Wernig, M., Newman, J.*, et al.* (2008). Connecting microRNA genes to the core transcriptional regulatory circuitry of embryonic stem cells. Cell *134*, 521-533.

Schmidt, D., Wilson, M.D., Spyrou, C., Brown, G.D., Hadfield, J., and Odom, D.T. (2009). ChIP-seq: using high-throughput sequencing to discover protein-DNA interactions. Methods (San Diego, Calif.) *48*, 240-248.

Solovei, I., and Cremer, M. (2010). 3D-FISH on cultured cells combined with immunostaining. Methods in molecular biology *659*, 117-126.

Splinter, E., de Wit, E., Nora, E.P., Klous, P., van de Werken, H.J., Zhu, Y., Kaaij, L.J., van Ijcken, W., Gribnau, J., Heard, E.*, et al.* (2011). The inactive X chromosome adopts a unique three-dimensional conformation that is dependent on Xist RNA. Genes & development *25*, 1371-1383.

Splinter, E., de Wit, E., van de Werken, H.J., Klous, P., and de Laat, W. (2012). Determining long-range chromatin interactions for selected genomic sites using 4C-seq technology: from fixation to computation. Methods (San Diego, Calif.) *58*, 221-230.

Splinter, E., Heath, H., Kooren, J., Palstra, R.J., Klous, P., Grosveld, F., Galjart, N., and de Laat, W. (2006). CTCF mediates long-range chromatin looping and local histone modification in the beta-globin locus. Genes & development *20*, 2349-2354.

The ENCODE Project Consortium (2011). A user's guide to the encyclopedia of DNA elements (ENCODE). PLoS biology *9*, e1001046.

van de Werken, H.J., Landan, G., Holwerda, S.J., Hoichman, M., Klous, P., Chachik, R., Splinter, E., Valdes-Quezada, C., Oz, Y., Bouwman, B.A.*, et al.* (2012). Robust 4C-seq data analysis to screen for regulatory DNA interactions. Nature methods *9*, 969-972.

Wilson, A.A., Kwok, L.W., Hovav, A.H., Ohle, S.J., Little, F.F., Fine, A., and Kotton, D.N. (2008). Sustained expression of alpha1-antitrypsin after transplantation of manipulated hematopoietic stem cells. American journal of respiratory cell and molecular biology *39*, 133-141.