Inversion of a topological domain leads to restricted changes in its gene expression and affects interdomain communication

Rafael Galupa, Christel Picard, Nicolas Servant, Elphège P Nora, Yinxiu Zhan, Joke G van Bemmel, Fatima El Marjou, Colin Johanneau, Maud Borensztein, Katia Ancelin, et al.

▶ **To cite this version:**

Rafael Galupa, Christel Picard, Nicolas Servant, Elphège P Nora, Yinxiu Zhan, et al.. Inversion of a topological domain leads to restricted changes in its gene expression and affects interdomain communication. Development (Cambridge, England), 2022, 149 (9), pp.dev200568. 10.1242/dev.200568 . hal-03746685

HAL Id: hal-03746685
https://hal.science/hal-03746685

Submitted on 8 Aug 2022

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

© 2022. Published by The Company of Biologists Ltd | Development (2022) 149, dev200568. doi:10.1242/dev.200568

RESEARCH ARTICLE

Inversion of a topological domain leads to restricted changes in its gene expression and affects interdomain communication

Rafael Galupa[1,*,¶,], Christel Picard[1,§,¶], Nicolas Servant[2,3], Elphège P. Nora[1,‡], Yinxiu Zhan[4,5], Joke G. van Bemmel[1], Fatima El Marjou[6], Colin Johanneau[6], Maud Borensztein[1,§], Katia Ancelin[1], Luca Giorgetti[4] and Edith Heard[1,7,*,**]**

ABSTRACT

The interplay between the topological organization of the genome and the regulation of gene expression remains unclear. Depletion of molecular factors (e.g. CTCF) underlying topologically associating domains (TADs) leads to modest alterations in gene expression, whereas genomic rearrangements involving TAD boundaries disrupt normal gene expression and can lead to pathological phenotypes. Here, we targeted the TAD neighboring that of the noncoding transcript *Xist*, which controls X-chromosome inactivation. Inverting 245 kb within the TAD led to expected rearrangement of CTCF-based contacts but revealed heterogeneity in the 'contact' potential of different CTCF sites. Expression of most genes therein remained unaffected in mouse embryonic stem cells and during differentiation. Interestingly, expression of *Xist* was ectopically upregulated. The same inversion in mouse embryos led to biased *Xist* expression. Smaller inversions and deletions of CTCF clusters led to similar results: rearrangement of contacts and limited changes in local gene expression, but significant changes in *Xist* expression in embryos. Our study suggests that the wiring of regulatory interactions within a TAD can influence the expression of genes in neighboring TADs, highlighting the existence of mechanisms of inter-TAD communication.

KEY WORDS: TADs, Gene expression, X-inactivation, Xist, Genomic engineering

[1]Mammalian Developmental Epigenetics Group, Genetics and Developmental Biology Unit, Institut Curie, PSL Research University, CNRS UMR3215, INSERM U934, Paris 75005, France. [2]Bioinformatics, Biostatistics, Epidemiology and Computational Systems Unit, Institut Curie, PSL Research University, INSERM U900, Paris 75005, France. [3]MINES ParisTech, PSL Research University, CBIO-Centre for Computational Biology, Paris 75006, France. [4]Friedrich Miescher Institute for Biomedical Research, Basel 4058, Switzerland. [5]University of Basel, Basel 4001, Switzerland. [6]Transgenesis Facility, Institut Curie, Paris 75005, France. [7]Collège de France, Paris 75231, France.
*Present address: European Molecular Biology Laboratory, Heidelberg 69117, Germany. ‡Present address: Cardiovascular Research Institute, University of California San Francisco, San Francisco, CA 94158, USA; Department of Biochemistry and Biophysics, University of California San Francisco, San Francisco, CA 94158, USA. §Present address: IGMM, Univ Montpellier, CNRS, Montpellier 34090, France.
¶These authors contributed equally to this work

**Authors for correspondence (rafael.galupa@embl.de; edith.heard@embl.org)

R.G., 0000-0001-7319-043X; J.G.v.B., 0000-0001-7034-3664; M.B., 0000-0002-4378-5018; K.A., 0000-0002-2117-9754; E.H., 0000-0001-8052-7117

This is an Open Access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0), which permits unrestricted use, distribution and reproduction in any medium provided that the original work is properly attributed.

Handling Editor: James Briscoe
Received 21 January 2022; Accepted 28 February 2022

INTRODUCTION

The three-dimensional folding of the genome has been increasingly recognized as an essential component for our understanding of gene regulation (Dekker and Mirny, 2016; McCord et al., 2020). Chromosome conformation capture techniques (Denker and de Laat, 2016) have unraveled a complex hierarchy of structural layers that organize mammalian chromosomes, composed of domains of high-frequency contacts (Zhan et al., 2017). At the sub-megabase level, these domains are generally designated topologically associating domains (TADs) (Dixon et al., 2012; Nora et al., 2012) and are well conserved across species and invariant across cell types (Dekker and Heard, 2015). The dynamics of the formation and maintenance of TADs and their boundaries during development and each cell cycle remains under investigation (Szabo et al., 2019) but appears to depend on the interplay between the architectural proteins cohesin and the zinc finger protein CTCF (Fudenberg et al., 2016; Haarhuis et al., 2017; Nora et al., 2017; Rao et al., 2017; Sanborn et al., 2015; Schwarzer et al., 2017; Wutz et al., 2017). Enriched at boundaries between TADs (Dixon et al., 2012; Phillips-Cremins et al., 2013), CTCF is required for chromatin loops observed between CTCF sites and for the organization and insulation of most TADs (Nora et al., 2017). Remarkably, CTCF-mediated contacts mainly occur between CTCF sites in which the CTCF motifs lie in a convergent orientation (Rao et al., 2014; Tang et al., 2015), and the contacts depend on the orientation of the motifs: altering the orientation of a CTCF site can disrupt a loop and lead to the formation of new ones (de Wit et al., 2015; Guo et al., 2015; Sanborn et al., 2015).

TADs are thought to instruct gene regulatory landscapes, allowing promoters and their regulatory elements to meet often and lead to a more efficient transcriptional output (Symmons et al., 2016). Accordingly, TADs represent the folding scale at which promoter-enhancer interactions and gene co-regulation are maximized (Zhan et al., 2017). The communication between promoters and enhancers is generally assumed to rely on chromatin looping, and long-range contacts within TADs can be dynamic during processes that involve rewiring of the regulatory networks, such as differentiation (Dixon et al., 2015). However, the interplay between such topological organization and the regulation of gene expression remains unclear. Loss of TADs upon depletion of CTCF or cohesin leads to relatively small effects on gene expression (Nora et al., 2017; Rao et al., 2017; Schwarzer et al., 2017; Wutz et al., 2017), and genomic rearrangements involving mammalian TADs and their boundaries can have either very modest effects (Amândio et al., 2020; Despang et al., 2019; Paliou et al., 2019; Rodríguez-Carballo et al., 2017; Williamson et al., 2019) or disrupt normal gene expression and underlie pathological phenotypes (Flavahan et al., 2016; Franke et al., 2016; Hnisz et al., 2016; Lupiáñez et al., 2015).

Development (2022) 149, dev200568. doi:10.1242/dev.200568

Here, we explored the relationship between TAD organization and transcriptional regulation in a crucial developmental regulatory landscape, the mouse X-inactivation center (*Xic*). The *Xic* is the master regulator of the initiation of X-chromosome inactivation in female placental mammals (Augui et al., 2011; Rastan and Brown, 1990), harboring the noncoding RNA *Xist* locus and the regulatory elements necessary for its female-specific developmental control. *Xist* is repressed in mouse embryonic stem cells (mESCs) and becomes upregulated from one of the two X-chromosomes in females upon exit of the pluripotent state, leading to random X-inactivation. This upregulation depends on the *Xist cis*-regulatory landscape (Heard et al., 1999), the full extent of which is still undefined; however, it is partitioned in at least two TADs, with the *Xist* locus lying close to the boundary between them (Fig. 1A) (Nora et al., 2012). The TAD in which the *Xist* promoter is included (referred to here as the Xist-TAD) contains some *Xist*-positive regulators (Augui et al., 2007; Barakat et al., 2011, 2014; Furlan et al., 2018; Gontan et al., 2012; Jonkers et al., 2009; Tian et al., 2010), whereas the adjacent TAD (referred to here as the Tsix-TAD) contains the promoter of *Tsix*, the antisense transcription unit to *Xist* that blocks its upregulation (Lee and Lu, 1999; Luikenhuis et al., 2001; Stavropoulos et al., 2001), as well as other elements that act as a *cis*-repressor of *Xist* [such as *Linx* and *Xite* (*Rr18*); see below].

To investigate how the topological organization of the Tsix-TAD impacts the regulation of genes both therein and in the neighboring Xist-TAD, we generated a mutant allele in mESCs and in mice with an inversion of almost the entire Tsix-TAD (245 kb out of 300 kb). We found that rewiring the Tsix-TAD structural landscape led to the formation of new chromatin contacts within the TAD, generally following the folding principles determined by the orientations of CTCF motifs. These topological alterations were accompanied by changes in gene expression of two out of seven genes within the TAD in differentiating mESCs. Interestingly, we found that the expression of *Xist* in the neighboring TAD was ectopically upregulated, suggesting that inter-TAD communication was affected.

RESULTS

Generating a genomic inversion encompassing the Tsix-TAD (245 kb-INV)

The Tsix-TAD harbors three hotspots of physical contacts (Fig. 1A), involving three different loci: (1) the *Xite* element, a proximal enhancer of *Tsix* (Ogawa and Lee, 2003), also involved in the position and insulation of the boundary between the Tsix- and Xist-TADs (van Bemmel et al., 2019); (2) the noncoding *Linx* locus, which harbors two *cis*-regulatory elements involved in controlling *Xist* expression (Galupa et al., 2020); and (3) *Chic1*, previously

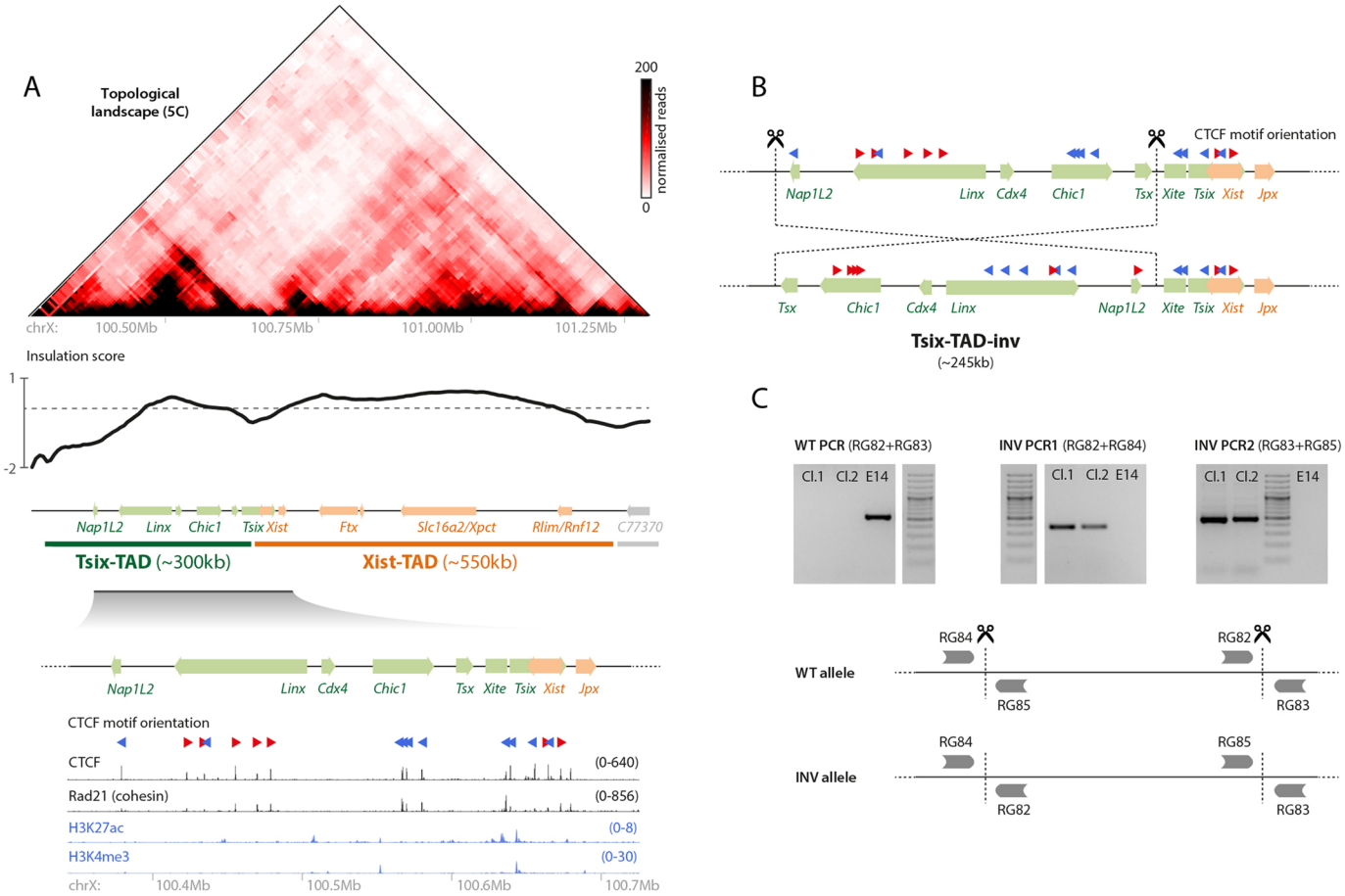


Fig. 1. Strategy for inverting almost the entire Tsix-TAD. (A) Topological organization of the *Xic* (top) and chromatin ChIP-seq profiles (bottom; see Materials and Methods for sources); the *Xist*/*Tsix* locus lies at the boundary between two TADs. The red and blue arrowheads indicate the orientation of the CTCF motif (orientated left or right, respectively). (B) Targeting strategy for inverting the ~245 kb region comprising most of the Tsix-TAD, except *Tsix* and its known regulator *Xite*, and leaving the boundaries intact. (C) PCR strategy (bottom) and gel results (top) for detecting the inversion events. E14 is the wild-type (WT) parental cell line. Cl.1 and Cl.2 are the two clones that were generated and analyzed throughout the study.

Development (2022) 149, dev200568. doi:10.1242/dev.200568

implicated in the maintenance of the organization of the Tsix-TAD (Giorgetti et al., 2014). Each of these loci harbors a set of CTCF sites involved in mediating the observed physical contacts and, within each locus, most CTCF motifs present the same orientation (Fig. 1A). Sites within *Linx* are 'convergently oriented' toward those within *Chic1* or *Xite*, the preferred orientation to form chromatin loops (Rao et al., 2014; Tang et al., 2015). Contacts between *Chic1* and *Xite* (the CTCF motifs of which occur in a 'tandem') are also observed (Fig. 1A). The contacts between these three loci might occur in pairwise fashion and/or simultaneously; physical modeling suggests that all conformations are possible (Giorgetti et al., 2014) and deletions of the CTCF-binding sites in either *Xite* (van Bemmel et al., 2019) or *Linx* (Galupa et al., 2020) show that contacts between the two remaining loci still occur.

We investigated whether this complex topological organization might be crucial for correct communication between the surrounding *cis*-regulatory elements (such as those within *Xite* and *Linx*) and their targets, therefore ensuring appropriate gene expression of *Tsix* and *Xist* and correct patterns of X-inactivation. Using a CRISPR/Cas9 editing approach in male mESCs, which carry a single X chromosome, we targeted a ∼245 kb region encompassing all loci within the Tsix-TAD, including the CTCF clusters within *Linx* and *Chic1*, but excluding *Xite* and *Tsix* (Fig. 1B). We decided not to include *Xite* in the inversion because: (1) *Xite* is already known to influence *Xist* expression (via *Tsix*); and (2) if *Xite* was inverted along with the rest of the TAD, the relative CTCF orientations between *Xite*, *Linx* and *Chic1* would not have changed. The targeted region does not involve either of the two boundaries of the TAD. We successfully generated two clones harboring an inversion allele (245 kb-INV) (Fig. 1C). This genomic inversion swaps the orientations of all CTCF motifs therein relative to those outside of the inverted region, in particular for *Linx* and *Chic1* (Fig. 1B) and, therefore, is expected to lead to the formation of new contacts within the TAD.

245 kb-INV leads to rearrangement of contacts within the TAD and increased insulation with neighboring TAD

To assess the topological organization of the 245 kb-INV allele, we performed carbon-copy chromosome conformation capture (5C) on the *Xic* (Dostie et al., 2006; Nora et al., 2012) for mutant and control mESCs (Fig. 2A). 5C analysis revealed that three hotspots of contacts can still be observed in the Tsix-TAD on the 245 kb-INV allele (Fig. 2B; please note that the 5C map is shown after 'correction' of the new genomic sequences in the inverted allele). These involve the same three loci as in control cells: in its new position, the *Chic1* CTCF cluster is able to establish contacts with *Linx* and with *Xite* (Fig. 2B); *Linx* and *Xite*, with CTCF sites in 'tandem' orientation in the 245 kb-INV allele, also interact (as *Chic1* and *Xite* do in control cells) (Fig. 2B). Therefore, inverting the *Linx* and *Chic1* CTCF clusters simultaneously appears to lead to new but similar hotspots of physical contacts within the Tsix-TAD compared with control. This might have been expected given that the overall distribution and orientation of CTCF sites within the TAD remain similar between the wild-type and the inverted alleles (Fig. 1B). In other words, the *Chic1* CTCF cluster on the inverted allele occupies an equivalent position to *Linx* on the wild-type allele, and vice versa. Therefore, the 245 kb inversion can lead to the formation of similar loops within the Tsix-TAD compared with the wild type, although involving different *cis*-regulatory elements.

Nevertheless, we also noticed some significant differences in the topology of the 'inverted' Tsix-TAD. Increased contacts were observed upstream of the inverted region, corresponding to contacts stemming from the *Linx* CTCF cluster in its new position (Fig. 2B, bottom, black arrow, red region in the differential map; this region shows no particular chromatin signatures, such as CTCF-binding or active chromatin marks). This suggests a different 'strength' for the *Linx* and *Chic1* CTCF clusters: in the inverted allele, the *Linx* CTCF cluster strongly interacts with regions upstream of *Chic1* (Fig. 2B, bottom, black arrow), whereas, in the wild-type configuration, the *Chic1* CTCF cluster does not form such strong contacts with regions upstream of *Linx* (Fig. 2B, top, black arrow). Conversely, we also observed a strong localized reduction in contacts (Fig. 2B, differential map, blue arrow) associated with the switch in positions between *Linx* and *Chic1*: the *Linx* CTCF cluster at its original position was able to form long-range contacts beyond *Chic1* and *Xite*, namely with elements within the Xist-TAD (Fig. 2B, top, blue arrow). These contacts were lost (or strongly reduced) in the 245 kb-INV cells (Fig. 2B, bottom and differential map, blue arrows), indicating that the *Chic1* CTCF cluster did not establish long-range contacts with the Xist-TAD when placed in the *Linx* CTCF cluster position. In fact, this loss of contacts across the boundary extended along the whole Xist-TAD (Fig. 2C). Again, this suggests a stronger potential for the *Linx* CTCF cluster to form contacts compared with the *Chic1* CTCF cluster.

We also evaluated the extent to which these topological changes had an impact on the overall insulation of the TADs. Insulation score analysis (see Materials and Methods) revealed a clear gain of insulation across the boundary between the Tsix-TAD and the Xist-TAD (Fig. 2D; lower insulation scores are reflective of increased insulation). The loss of *Linx*-mediated contacts across the boundary probably accounts, at least partially, for this increased insulation between the TADs. In summary, the 245 kb inversion repositions CTCF clusters within the Tsix-TAD, leading to reconfiguration of specific intra- and inter-TAD contacts accompanied by stronger insulation with the neighboring Xist-TAD.

245 kb-INV leads to gene expression changes within the Tsix-TAD and across the boundary

We next set out to determine whether similar interaction patterns, but different wiring of sequences within the Tsix-TAD, led to any transcriptional changes. To this end, we profiled transcript levels across the *Xic* using digital gene expression analysis (NanoString) (Geiss et al., 2008) in control and mutant cells in the pluripotent state (d0) and during early differentiation (d0.5-d2.5) (Fig. 3A). Expression of most genes within the Tsix-TAD and the Xist-TAD was unaffected in 245 kb-INV cells (Fig. 3B), including that of the three loci involved in the topological alterations: *Linx*, *Chic1* and *Xite*. This suggests no or limited effect of the structural alterations on the transcriptional regulation of these loci. Expression of *Tsix* was significantly reduced in mutant cells in the pluripotent state (d0) (Fig. 3B), but this effect did not persist consistently during differentiation. The deletion of the same region that we inverted here also led to downregulation of *Tsix* in mESCs (Galupa et al., 2020); together with the current data, this suggests that the region contains important sequences for *Tsix* regulation, that this regulation depends on the orientation of the region as a whole and might also depend on the orientation of individual regulatory sequences.

However, we did notice consistent changes during differentiation in mutant cells for two genes within the Tsix-TAD: *Nap1l2*, which was significantly upregulated at all time points (Fig. 3B) and *Tsx*, which was significantly downregulated (Fig. 3B). Interestingly, both genes lie at the extremities of the inverted region, and switch

Development (2022) 149, dev200568. doi:10.1242/dev.200568

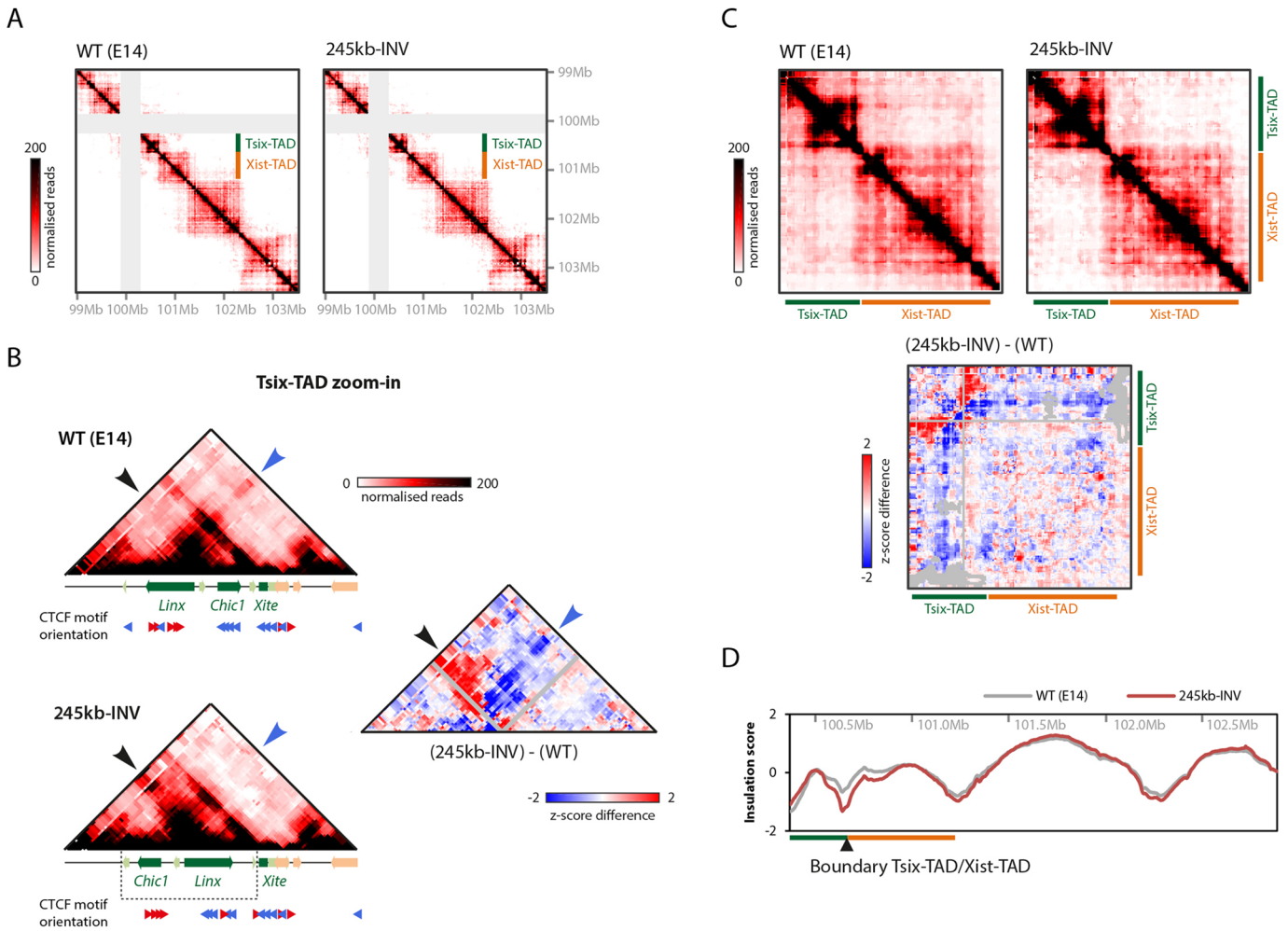


Fig. 2. Rearrangement of contacts within the TAD and increased insulation with neighboring TAD following 245 kb intra-TAD inversion. (A) 5C profiles of wild-type (WT; two replicates pooled) and 245 kb-INV mutant (two clones pooled) mESCs. The mutant map is corrected for inversion and gray pixels represent filtered contacts (see Materials and Methods). (B) Detailed view of the Tsix-TAD. The small red and blue arrowheads indicate the orientation of the CTCF motif (orientated left and right, respectively). The large black and large blue arrowheads indicate specific contacts that are gained or lost in the mutant, respectively (more details in the text). Gray pixels represent filtered contacts. (C) View of the Tsix- and Xist-TADs, and differential map representing the subtraction of *z*-scores calculated for WT and 245kb INV mutant maps separately. Gray pixels represent filtered contacts. (D) Insulation scores across the *Xic* TADs and downstream TADs based on 5C profiles for WT and 245 kb-INV mutant mESCs. The 'troughs' represent TAD boundaries.

their relative positions in the TAD between the wild-type and mutant configurations. It is likely that changes in their gene expression are associated with altered proximity to the *Xite* enhancer element. Given that deletion of *Xite* leads to downregulation of *Tsx* (van Bemmel et al., 2019), moving *Tsx* away from *Xite* on the 245 kb-INV allele could lead to its observed downregulation. Conversely, the increased linear proximity of *Nap1l2* to *Xite* could underlie *Nap1l2* upregulation. Changes in interaction frequencies between *Xite* and these two elements in the 245 kb-INV allele support this hypothesis, because they reflect the changes in their genomic distances (increased for *Xite-Nap1l2* and decreased for *Xite-Tsx*, compared with control; Fig. S1).

We also observed changes in expression of *Xist*, the long noncoding RNA locus that is regulated by the *Xic* to trigger the initiation of X-chromosome inactivation. Normally very low in male cells, *Xist* expression was slightly upregulated in the mutants at later differentiation time points (~twofold at d1.5; Fig. 3B). In female cells undergoing X-inactivation, upregulation of *Xist* is accompanied by local accumulation of its RNA in *cis*, forming a so-called 'Xist cloud', which can readily be detected by RNA fluorescence *in situ* hybridization (FISH) (Augui et al., 2011). RNA FISH revealed the formation of *Xist* clouds in ~4-7% of mutant male cells upon differentiation, which was never observed in wild-type male cells (Fig. 3C). Thus, the inversion of 245 kb within the Tsix-TAD leads to ectopic expression of *Xist*, the promoter of which is located in the neighboring TAD.

Female embryos with a 245 kb-INV allele show higher *Xist* allelic imbalance

Given the impact of the 245 kb inversion on *Xist* expression in male cells, we investigated whether this was also the case in female embryos at post-implantation stages, when random X-inactivation is known to have already occurred (Rastan, 1982). To this end, we generated an equivalent 245 kb-INV allele in mice (see Materials and Methods) and collected post-implantation heterozygous embryos. These embryos were derived from crosses between polymorphic mouse strains (Fig. 4A,D), which allowed us to distinguish the allelic origin of transcripts. Analysis of RNA allelic ratios for *Atp7a*, an X-linked gene, revealed no preferential gene silencing for one or the other allele (Fig. 4B,E).

Development (2022) 149, dev200568. doi:10.1242/dev.200568

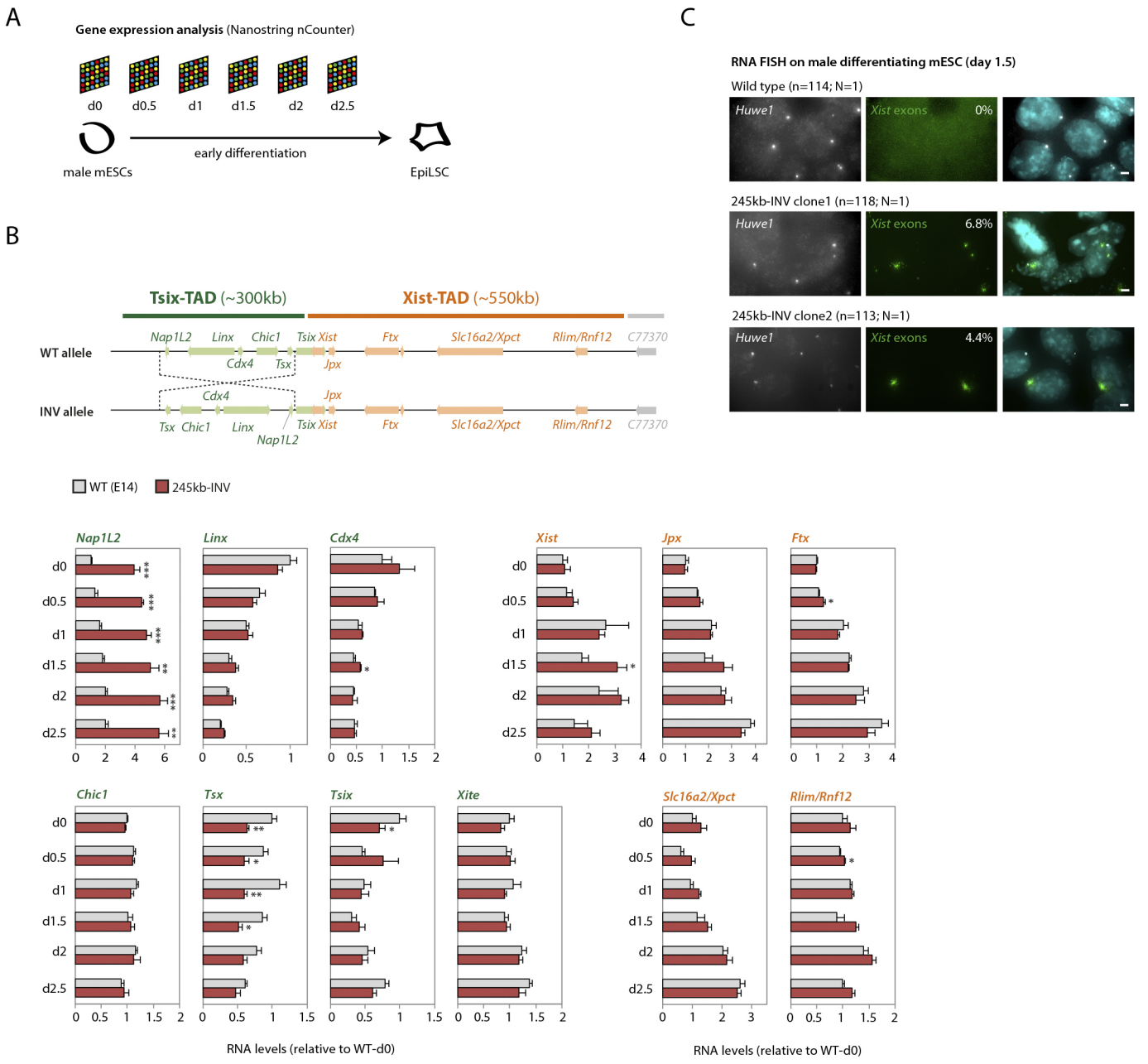


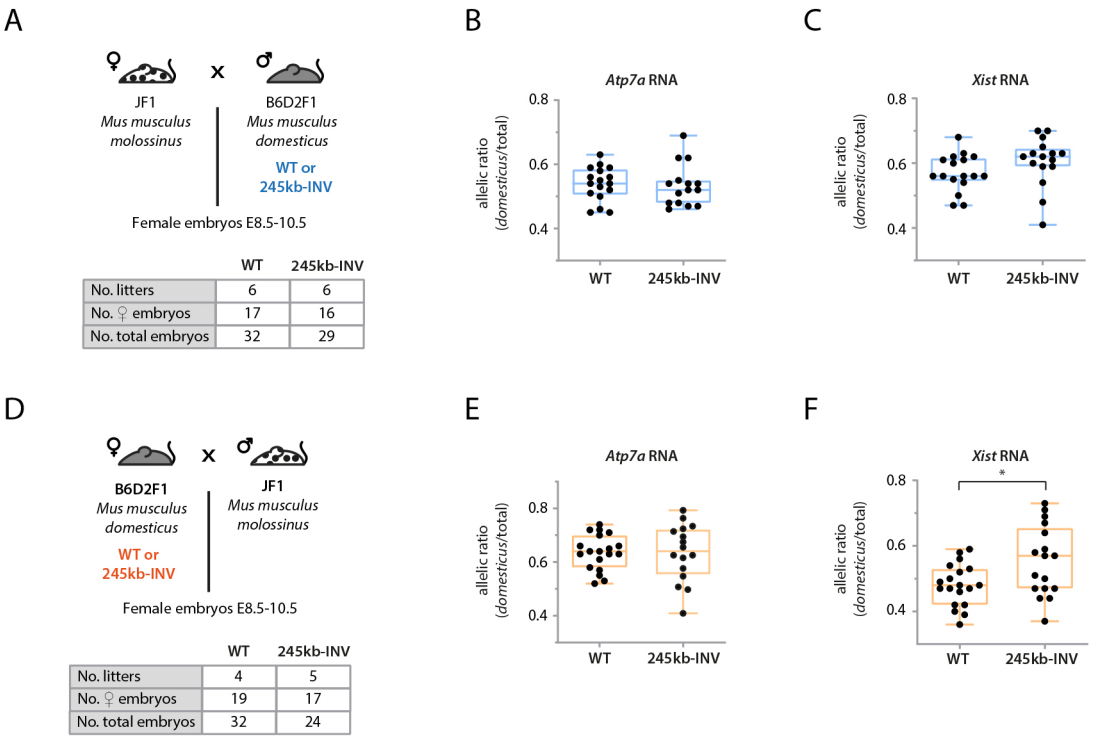
Fig. 3. Inversion leads to transcriptional changes of specific genes within the TAD and of *Xist* across the TAD boundary. (A) Schematic of the mESC-to-epiblast-like stem cell (EpiLSC) differentiation and time points analyzed by Nanostring nCounter (see Materials and Methods). (B) Gene expression analysis during differentiation (d0-d2.5). Data are normalized to wt-d0 for each gene, and represent the mean±s.d. of two biological replicates (wild type) or of two independent clones (mutant). Data were analyzed with a two-tailed paired Student's *t*-test (*P<0.05; **P<0.01; ***P<0.001). (C) RNA FISH for *Huwe1* (X-linked gene outside of the *Xic*) and *Xist* (exonic probe) on mESCs differentiated to d1.5. The percentage of cells with *Xist* RNA accumulation is indicated and represents the mean from two independent experiments, where *N* indicates the number of experiments and *n* indicates the number of nuclei counted. Scale bars: 2 µm.

However, analysis of *Xist* RNA allelic ratios between mutant and control embryos showed slightly higher *Xist* allelic ratios in the mutant embryos, regardless of whether the mutant allele was inherited paternally (Fig. 4C) or maternally (Fig. 4F); this increase was statistically significant for maternal transmission (P<0.05). These results are consistent with the upregulation of *Xist* that we observed in mutant cells (Fig. 3B,C). Thus, the 245 kb inversion leads to higher *Xist* levels in *cis*, but this does not result in skewed patterns of X-inactivation. Of note, litter size appeared to be reduced upon maternal transmission of the 245 kb-INV allele, with a skewed sex ratio (71% females in 245 kb-INV versus 59% in control), suggesting that the inversion may have more phenotypic consequences.

Mutating clusters of CTCF sites within *Linx* and *Chic1* lead to changes in *Xist* expression

To further explore the link between the topological organization of the Tsix-TAD and *Xist* regulation, we generated alleles with deletions and/or inversions of the clusters of CTCF sites within *Linx* and within *Chic1*. We previously deleted a large intronic interval containing three *Linx* CTCF sites from male ESCs (~51 kb) and from mice (~25 kb) (Galupa et al., 2020), which led to some

Development (2022) 149, dev200568. doi:10.1242/dev.200568



	WT	245kb-INV
No. litters	6	6
No. ♀ embryos	17	16
No. total embryos	32	29

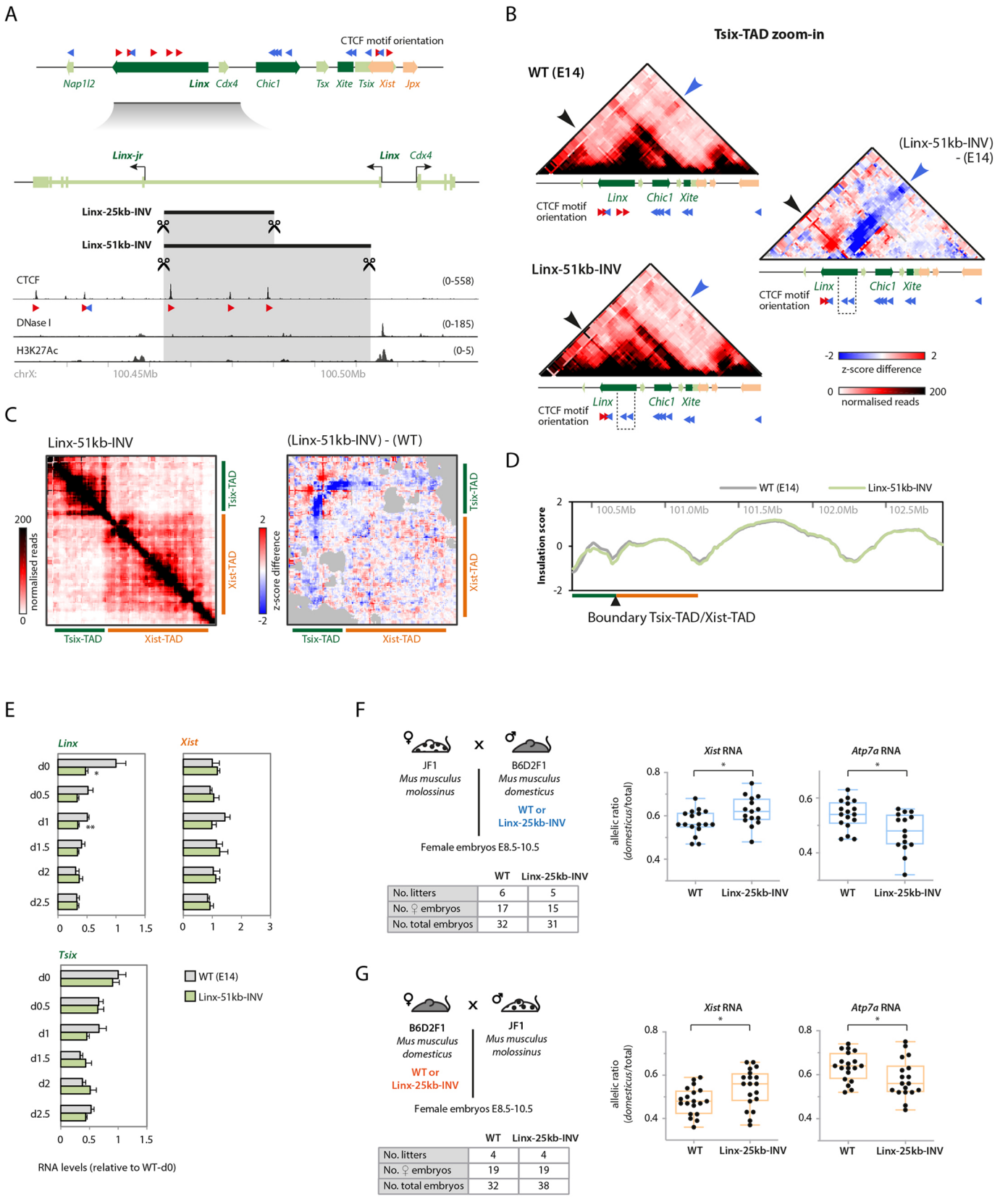
	WT	245kb-INV
No. litters	4	5
No. ♀ embryos	19	17
No. total embryos	32	24

Fig. 4. Female embryos with the 245 kb-INV allele show a bias in *Xist* expression. (A,D) Crosses used for the analysis of RNA allelic ratios in female hybrid embryos inheriting the *Mus musculus domesticus* allele either paternally (A; blue) or maternally (D; red). Tables summarize the number of embryos collected. (B,C,E,F) RNA allelic ratios for the X-linked gene *Atp7a* (B,E) and *Xist* (C,F). Each black dot corresponds to a single female embryo. Box-and-whisker plots indicate median, interquartile range and min/max values, respectively, with blue and red plots indicating paternally or maternally inherited alleles, respectively. Data were analyzed using the Mann–Whitney U test (*P<0.05).

alterations in the topological organization of the two *Xic* TADs but no changes in *Xist* expression in female embryos. Thus, we tested the impact of inversions of exactly the same regions in male mESCs (Linx-51 kb-INV) and in mice (Linx-25 kb-INV) (Fig. 5A). 5C analysis of the Linx-51 kb-INV allele revealed higher frequency of contacts between the now-inverted *Linx* locus and regions immediately upstream (Fig. 5B, black arrowhead), and lower frequency of contacts between (inverted) *Linx* and *Chic1* and between *Xite* and elements within the Xist-TAD (Fig. 5B,C, blue arrowhead), in agreement with the change in orientation of the three *Linx* CTCF sites. These results are reminiscent of what we observed for the 245 kb-INV allele (Fig. 2B,C), and support the hypothesis that loss of contacts with the Xist-TAD in the 245 kb-INV allele is associated with inversion of the CTCF sites within *Linx*. Consistently, analysis of insulation scores across the TADs revealed a gain of insulation across the boundary between the Tsix-TAD and the Xist-TAD (Fig. 5D), although less pronounced than that observed for the 245 kb-INV allele (Fig. 2D). We next analyzed gene expression across the *Xic* for the Linx-51 kb-INV mESCs in the pluripotent state (d0) and during early differentiation (d0.5-d2.5); expression of *Linx* was significantly downregulated at some time points (Fig. 5E) but no changes were observed for *Xist* or *Tsix* (Fig. 5E) or for any other locus across the *Xic* (Fig. S2). However, when we analyzed *Xist* expression in female embryos carrying an heterozygous *Linx*-25 kb-INV allele, we observed significantly higher expression of *Xist* for the inverted allele, for both paternal and maternal transmission (Fig. 5F,G), and also a corresponding decrease in expression of the X-linked gene *Atp7a* (Fig. 5F,G), suggestive of skewed X chromosome inactivation compared with control. Overall, the inversion of the *Linx* CTCF cluster led to similar phenotypes compared with the large 245 kb inversion, namely a decrease in contact frequency between *Linx* and the Xist-TADs, a concomitant gain of insulation between them, and increased *Xist* expression in *cis* in female embryos.

We previously generated in male mESCs a ∼4 kb deletion within *Chic1* (Giorgetti et al., 2014) that encompassed two of the three CTCF-binding sites present in the locus (Fig. 6A), but we did not study its impact on chromosome conformation or on *Xist* expression, which we set out to do here. Differential 5C analysis between this *Chic1*-4 kbΔ allele and wild type showed a reduction in contacts between *Chic1* and *Linx* and also between *Chic1* and *Xite* (Fig. 6B), consistent with loss of the *Chic1* CTCF sites. We also noted an apparent increase in contact frequency between *Linx* and *Xite* (Fig. 6B; Fig. S1), which would be consistent with a model of competition between *Chic1* and *Xite* CTCF sites to form loops with the CTCF sites within *Linx*. However, these differences in contact frequencies overall remained rather close to the 'noise' levels of the 5C map. We wondered whether these effects would be more pronounced if the remaining CTCF-binding site was also removed; thus, we generated, in male mESCs, a larger deletion (*Chic1*-14 kbΔ) encompassing all three CTCF sites within *Chic1* (Fig. 6A). We observed more-pronounced contact rearrangements within the Tsix-TAD as for *Chic1*-4 kbΔ (Fig. 6C), suggesting that it is the loss of the CTCF sites that underlies the observed topological differences. To study the impact of these deletions on gene expression across the *Xic*, we profiled transcript levels, as performed previously, in the pluripotent state (d0) and during early differentiation. Expression of *Chic1* itself was consistently upregulated in both *Chic1*-4 kbΔ and *Chic1*-14 kbΔ (Fig. 6D,E); it is intriguing to think that this could be linked to its now shorter length, because shorter genes have been associated with higher levels of expression (Castillo-Davis et al., 2002; Chiaromonte et al.,

Development (2022) 149, dev200568. doi:10.1242/dev.200568



	WT	Linx-25kb-INV
No. litters	6	5
No. ♀ embryos	17	15
No. total embryos	32	31

	WT	Linx-25kb-INV
No. litters	4	4
No. ♀ embryos	19	19
No. total embryos	32	38

Fig. 5. See next page for legend.

2003). We also observed higher expression of *Cdx4*, the gene upstream of *Chic1*: interestingly, the effects appeared to scale up with the larger deletion: in *Chic1*-4 kbΔ mESCs, there was a slight increase in *Cdx4* levels across time points but this was not statistically significant, whereas, in *Chic1*-14 kbΔ, the increase was more pronounced and statistically significant for some time points.

Development (2022) 149, dev200568. doi:10.1242/dev.200568

Fig. 5. Inversion of *Linx* cluster of CTCF sites leads to Xist upregulation in *cis*. (A) The *Linx* locus, CTCF binding, and orientation of CTCF motifs associated with CTCF ChIP-seq peaks. The red and blue arrowheads indicate the orientation of the CTCF motif (orientated left or right, respectively). The targeted inversions *Linx*-25 kb-INV and *Linx*-51 kb-INV are indicated. (B) 5C profiles (Tsix-TAD zoom-in) of wild-type (WT; two replicates pooled) and *Linx*-51 kb-INV (two clones pooled) mESCs, and 5C differential map, representing the subtraction of *z*-scores calculated for WT and Linx-51 kb-INV maps. The large black and large blue arrowheads indicate specific contacts that are gained or lost in the mutant, respectively (more details in the text). (C) (Left) 5C profile of Linx-51 kb-INV mESCs (two clones pooled); the map is corrected for inversion and gray pixels represent filtered contacts (see Materials and Methods); (Right) 5C differential map, representing the subtraction of *z*-scores calculated for WT and *Linx*-51 kb-INV maps separately. (D) Insulation scores across the *Xic* TADs and downstream TADs based on 5C profiles for WT and *Linx*-51 kb-INV mESCs. The 'troughs' represent TAD boundaries. (E) Gene expression analysis during differentiation (d0-d2.5). Data are normalized to wt-d0 for each gene, and represent the mean±s.d. of two biological replicates (WT) or of two independent clones (mutant). Data were analyzed with a two-tailed paired Student's *t*-test. (F,G) (Left) Crosses used for analysis of RNA allelic ratios in female hybrid embryos inheriting the *Mus musculus domesticus* allele either paternally (F; blue) or maternally (G; red). Tables summarize the number of embryos collected. (Right) RNA allelic ratios for *Xist* and the X-linked gene *Atp7a*. Each black dot corresponds to a single female embryo. Box-and-whisker plots indicate median, interquartile range and min/max values, with blue and red plots indicating paternally or maternally inherited alleles, respectively. Data were analyzed using the Mann–Whitney U test (**P*<0.05; ***P*<0.01).

This effect could be connected to the removal of all CTCF sites from within the *Chic1* locus, which could 'shield', or insulate, *Cdx4* from activating influences downstream of the CTCF sites. *Xist* expression was also more affected in mESCs containing the larger deletion: we observed a mostly consistent downregulation across all time points, but this effect was not statistically significant in this context. However, in female embryos, we did observe a statistically significant decrease in *Xist* expression from the deletion alleles (Fig. 6D,E), which was more pronounced for the *Chic1*-14 kbΔ allele. This suggests that the *Chic1* CTCF cluster might operate to favor *Xist* expression in *cis*. These results potentially also illustrate how loss of one additional CTCF-binding site might be enough to cause stronger changes in chromosome conformation and gene expression.

Together, our results on inverting or deleting *Linx* and *Chic1* CTCF clusters highlight the rather complex regulatory landscape within the *Xic*. Similar to the 245 kb inversion, these mutant alleles reveal how *Xist* is sensitive to changes involving CTCF-binding sites within the neighboring Tsix-TAD. These results also suggest that the phenotypes observed in the 245 kb-INV mESCs are likely a combination of effects from changing different elements within the Tsix-TAD.

DISCUSSION

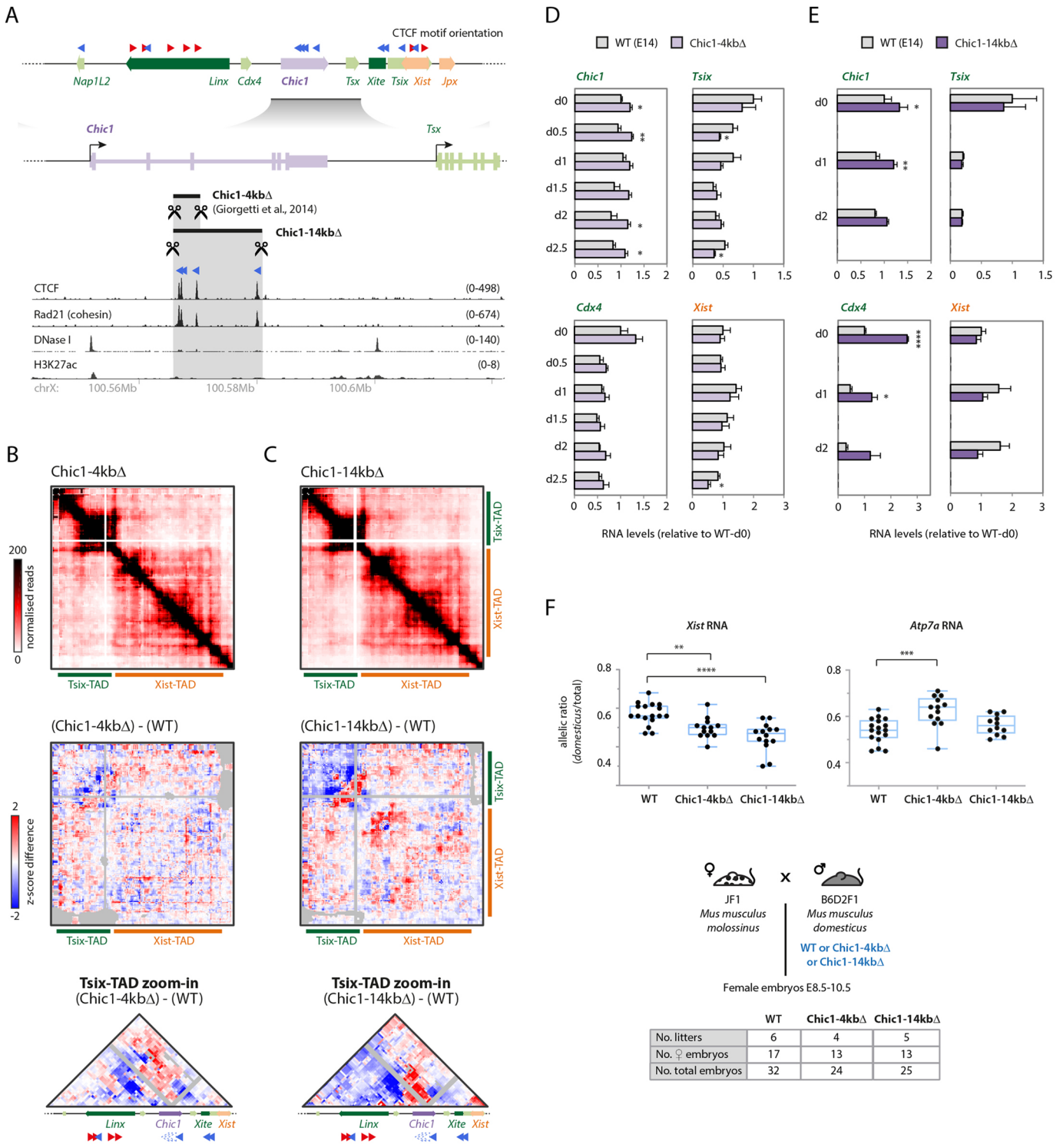
Here, we explored the structural and transcriptional consequences of inverting a large genomic region encompassing almost an entire TAD (80%; 245 kb out of 300 kb). We found that this inversion led to rearrangement of contacts and to changes in expression of some genes within the TAD. We also observed increased contact insulation with the neighboring TAD and ectopic upregulation of a gene in that TAD, the noncoding RNA *Xist* locus.

The rearrangement of contacts within the Tsix-TAD upon inverting a 245 kb region occurred largely as expected based on the 'rules' associated with the orientation of CTCF motifs within the TAD (de Wit et al., 2015; Guo et al., 2015; Rao et al., 2014; Sanborn et al., 2015; Tang et al., 2015). We found that the three loci involved in the strong contacts observed in the wild-type Tsix-TAD were still able to form strong contacts with each other in the 'inverted' Tsix-TAD (Fig. 2). Yet, these elements could not fully replace each other in their new positions, despite similar composition in terms of number of CTCF sites and levels of CTCF binding based on published chromatin immunoprecipitation-sequencing (ChIP-seq) data (Fig. 1). In particular, the *Linx* CTCF cluster appears to have a stronger potential to form contacts compared with the *Chic1* CTCF cluster. At the same relative position within the TAD, and with the same CTCF motif orientation, these CTCF clusters show a different range of interactions, as described above (Fig. 2). These differences suggest that not all CTCF-bound sites are equally capable of mediating the same type of interactions. Little is known about what determines which CTCF sites contact with each other, and whether there could be specific affinities between sites depending, for instance, on which other protein complexes are bound at each site or nearby. The ChIP-seq signal for CTCF is comparable between CTCF sites within *Linx* and *Chic1*, and there are the same number of CTCF sites within each locus. One difference in the organization of these sites is the spacing between them: CTCF sites within *Chic1* are more clustered than are those within *Linx*; this could play a role in orchestrating which contacts are formed, and how. 'Loop extrusion' has been proposed as a mechanism to form TADs and chromatin contacts (Fudenberg et al., 2016; Goloborodko et al., 2016; Sanborn et al., 2015), by which an 'extruding factor' (such as cohesin) engulfs two DNA chains and moves along them, extruding DNA until it reaches 'stalling factors' (such as CTCF), which block its progression; thus, a chromatin loop would be formed and stabilized. Could the length of the intervals between CTCF sites influence the likely point at which the cohesin complex (or the extruding factor) gets stalled? Perhaps more-distributed sites (as at the *Linx* locus) rather than more-clustered sites (as at the *Chic1* locus) provide more opportunities for stalling cohesin, given the very fast rate at which CTCF binds and unbinds chromatin (Hansen et al., 2017) and the rate of extrusion by cohesin (Davidson et al., 2019; Kim et al., 2019). Another potential explanation (not mutually exclusive) is that the differences in 'contact potential' depend on the different sequences flanking the consensus CTCF motifs within *Linx* and *Chic1*, as suggested by a recent study of CTCF sites as transcriptional insulators (Huang et al., 2021).

Transcriptional changes in mutant 245 kb-INV mESCs and during differentiation were observed for two genes (*Nap1l2* and *Tsx*) within the Tsix-TAD when compared with controls (Fig. 3). As discussed previously, we suggest that these changes are associated with (genomic) proximity to the enhancer element *Xite* and not necessarily with the new topological structure of the inverted Tsix-TAD. Perhaps more interesting is the fact that most other genes within the Tsix-TAD do not show changes in expression, especially the *Linx*, *Chic1* and *Xite* loci, which are involved in the topological changes observed for the 245 kb-INV allele. This could have a number of explanations: (i) the expression of these genes might not be particularly reliant on *cis*-regulation and, therefore, might be impervious to topological changes; (ii) interactions between these genes and their *cis*-regulatory elements might not depend on topological organization and, therefore, still occur regardless of the topological changes; or (iii) interactions between these genes and their *cis*-regulatory elements might depend on the topological organization; the new contacts allow these interactions to occur as efficiently as in wild type and, therefore, no changes in expression are observed. Further genetic exploration of these loci will be crucial to exclude hypotheses.

Surprisingly, expression of *Xist*, which lies outside of the Tsix-TAD, in the neighboring TAD, was mildly upregulated, to an extent that we could detect accumulation of Xist RNA in 'clouds' in mutant

Development (2022) 149, dev200568. doi:10.1242/dev.200568



	WT	Chic1-4kbΔ	Chic1-14kbΔ
No. litters	6	4	5
No. ♀ embryos	17	13	13
No. total embryos	32	24	25

Fig. 6. Deletion of *Chic1* cluster of CTCF sites leads to *Xist* downregulation in *cis*. (A) The *Chic1* locus, CTCF binding and orientation of CTCF motifs associated with CTCF ChIP-seq peaks. The targeted deletions *Chic1*-4 kbΔ and *Chic1*-14 kbΔ are indicated. The red and blue arrowheads indicate the orientation of the CTCF motif (orientated left or right, respectively). (B,C) (Top) 5C profiles of *Chic1*-4 kbΔ (B, two clones pooled) and *Chic1*-14 kbΔ (C, one clone, two replicates pooled). (Middle) 5C differential maps, representing the subtraction of *z*-scores calculated for wild-type (WT) and deletion maps. Gray pixels represent filtered contacts (see Materials and Methods). (Bottom) Differential maps for the Tsix-TAD. (D,E) Gene expression analysis during differentiation (d0-d2.5). Data are normalized to wt-d0 for each gene, and represent the mean±s.d. of two biological replicates (WT and *Chic1*-14 kbΔ) or of two independent clones (*Chic1*-4 kbΔ). Data were analyzed with a two-tailed paired Student's *t*-test (*P<0.05; **P<0.01; ****P<0.0001). (F) (Top) RNA allelic ratios for *Xist* and the X-linked gene *Atp7a*. Each black dot corresponds to a single female embryo. Box-and-whisker plots indicate median, interquartile range and min/max values. Data were analyzed with a Mann–Whitney U test (**P<0.01; ***P<0.001; ****P<0.0001). (Bottom) Crosses used for the analysis of RNA allelic ratios in female hybrid embryos inheriting the *Mus musculus domesticus* allele paternally. Tables summarize the number of embryos collected.

Development (2022) 149, dev200568. doi:10.1242/dev.200568

male cells, which we never observe(d) in wild-type male cells (Fig. 3). This upregulation could be associated with one or more of the other alterations observed on the 245 kb-INV allele, either structural, or transcriptional, or both. For instance, we observed reduced expression of *Tsix*, the antisense *cis*-repressor of *Xist*, in the pluripotent state, which could impact *Xist* regulation; during differentiation, however, when *Xist* is upregulated, we did not detect differences in *Tsix* expression. Further research will be needed to clarify the involvement of *Tsix* in the *Xist* phenotype observed here.

Could *Xist* upregulation be a consequence of *Nap1l2* upregulation or *Tsx* downregulation? Genetic studies with *Nap1l2* (Attia et al., 2007; Rogner et al., 2000) did not report any effects on *Xist* expression or sex-specific phenotypes; thus, upregulation of *Nap1l2* is unlikely to cause *Xist* upregulation, although this cannot be formally excluded. By contrast, knockout studies of *Tsx* (Anguera et al., 2011) reported *Xist* RNA clouds in a small percentage of differentiating male mutant mESCs; the authors proposed that *Xist* expression was upregulated because of its negative *cis*-regulators *Tsix* and *Xite* being downregulated during differentiation. This *Xist* phenotype is identical to the one we observed (Fig. 3), although, in 245 kb-INV mutant cells, there is still some *Tsx* expression (contrary to the *Tsx* knockout) and we did not observe changes in *Tsix* or *Xite* expression during differentiation. Thus, downregulation of *Tsx* in 245 kb-INV mutant cells might account, partially or perhaps even completely, for ectopic *Xist* upregulation. This raises interesting questions of how such inter-TAD regulation/communication between *Tsx* and *Xist* could occur. Similarly, we recently reported that another locus within the Tsix-TAD, *Linx*, contains sequences that affect expression of *Xist* in the neighboring TAD in a *Tsix*-independent manner (Galupa et al., 2020). A slight increase in *Xist* expression in *cis* was also observed in 245 kb-INV heterozygous embryos, but it was not statistically significant and did not result in skewed patterns of X-inactivation (Fig. 4). These results underlie the importance of verifying whether changes in gene expression result in differences in the phenotypes they mediate: in many studies, it often remains an open question whether the changes observed in gene expression, especially when modest, do matter for the processes in which those genes are involved.

In agreement with previous studies, our study illustrates that the relationship between chromosome structure and gene expression is rather complex. The almost 'dogmatic' view that TADs restrict gene *cis*-regulation (Finn and Misteli, 2019; Koch, 2019) is at odds with a growing amount of evidence that mechanisms of inter-TAD communication exist, albeit potentially subject to modulation by TADs and their boundaries. Here, we showed that, on the one hand, expression of genes within a TAD can be tolerant to changes in contacts within that TAD, whereas, on the other hand, inversion of a large region within a TAD affected the expression of a gene in the neighboring TAD, potentially because of accompanying changes in topological organization and topological insulation. Further investigations are warranted for a more-complete understanding of the relationship between the topological organization of the genome and the transcriptional regulation of its genes.

MATERIALS AND METHODS

All the materials and methods described below have also been published previously (Galupa et al., 2020).

Tissue culture conditions

The E14 mESC line and clones derived from it were grown in flasks or on dishes coated with 0.1% (wt/vol) gelatin. Culture media for mESCs comprised Glasgow medium (Gibco) supplemented with 2 mM L-glutamine, 0.1 mM nonessential amino acids, 1 mM sodium pyruvate, 15% fetal bovine serum (FBS) (Gibco), 0.1 mM b-mercaptoethanol (Sigma) and 1000 U/ml leukemia inhibitory factor (LIF) (Chemicon). All lines were cultivated at 37°C under 8% CO_2 and passaged according to their confluency, generally every other day. Medium was refreshed daily. For early EpiLSC differentiation assays, mESCs were washed with 1× PBS, incubated with trypsin at 37°C for 20 min and resuspended in ES medium without LIF. After cell counting, the desired number of cells was resuspended in differentiation medium and 8×10^5 cells per well were seeded on a fibronectin-coated (10 µg/ml, Millipore) six-well plate in differentiation medium. Differentiation medium comprised N2B27 medium, 20 ng/ml activin A (R&D Systems) and 12 ng/ml FGF-basic (R&D Systems). Differentiation medium was changed daily and cells were washed in PBS before collection to remove dead cells. Cells were routinely checked for mycoplasma contamination.

Mouse experimentation

Animal care and use for this study were performed in accordance with the recommendations of the European Community (2010/63/UE) for the care and use of laboratory animals. Experimental procedures, including genomic engineering (see below), were in compliance with international guidelines and were specifically approved by the ethics committee of the Institut Curie CEEA-IC #118 and given authorization by the French national authorities (references: APAFIS##13962-2018030717538778-v2 and APAFIS#8812-2017020611033784-v2).

Postimplantation embryos were collected at the embryonic day (E) 8.5-10.5 stage, assuming plugging at midnight. Females with a vaginal plug were weighed every other day and only taken for dissection if a significant increase in weight was observed (~2 g for B6D2F1 mice, ~1 g for JF1 mice) at the expected point of E8.5-E10.5 development. Extra-embryonic tissues were taken for sexing the embryos. Whole embryos were washed three times in 1× PBS before being frozen for allelic expression analysis.

Genomic engineering of mice and mESCs

Inversion 245 kb-INV and deletion *Chic1*-14 kbΔ were generated using CRISPR-Cas9 (mESCs and mice) technologies, using the process described below. Inversions within the *Linx* locus (*Linx*-25 kb-INV and *Linx*-51 kb-INV) were generated using the same constructs and primers as the equivalent deletions, described in Galupa et al. (2020). The *Chic1*-4 kbΔ deletion had been generated previously (Giorgetti et al., 2014).

We designed short guide (sg) RNAs to flank the region of interest:

- For 245 kb-INV: CR30 (ACTGGTTCAGCCACTCACCG) and CR32 (CTGAGCTGGTTCATACAGGT).
- For *Chic1*-14 kbΔ: CR21 (AAAGATCGTTTCTATCTAGC) and CR16R (CGCCAAACTTCCAAAATGGC).

For cloning sgRNAs, we used pX459-v2 (Addgene 62988) and a protocol from the Zhang lab (https://media.addgene.org/cms/filer_public/e6/5a/e65a9ef8-c8ac-4f88-98da-3b7d7960394c/zhang-lab-general-cloning-protocol.pdf). sgRNA constructs were amplified upon transformation of DH5α competent cells (Takara) grown at 37°C, and then sequenced to verify that the cloning was correct. Midipreps for all constructs were prepared at a final concentration >1 mg/ml using the NucleoBond Xtra Midi Plus kit (Macherey-Nagel).

mESCs were transfected with sgRNA constructs using the P3 Primary Cell 4D-Nucleofector X Kit (Lonza) and the Amaxa 4D Nucleofector™ system (Lonza), with the transfection program CG-104. Each transfection included 5 million cells resuspended in the nucleofection mix (prepared according to the manufacturer's instructions) containing 5 µg of each sgRNA (two constructs). As a transfection control, 10 µg of pmaxGFP (Lonza) was used, for which the nucleofection efficiency was around 90%. Cells were immediately resuspended in pre-warmed culture medium after nucleofection and seeded at three serial 10× dilutions in 10-cm dishes to ensure optimal density for colony picking. Transfected cells were selected

Development (2022) 149, dev200568. doi:10.1242/dev.200568

with puromycin for 48 h, and grown for 8-10 days. Single colonies were picked into 96-well plates. Genomic DNA was isolated in 96-well plates for PCR-based screening of inversions. The genotyping primers used were as follows:

- For 245 kb-INV: RG82 (CAATCACTCTTGCCTTACCAATT), RG83 (CCCAAACCAACCCTTGACTG), RG84 (GTTGGGACC-TAAACTCTAGTACA), RG85 (AGTGGACTAGCTTTGCCTCA).
- For *Chic1*-14 kbΔ: EN118 (GCCTGCAGTCTTACCAGGAG), EN119 (TAATCTGCAGCGTGTTGAGG), RG123 (TCCTCCC-TTACCAGTCTCCT), RG124 (CAGAATCCCGGATGTGAGGA).

We sequenced the PCR products from the inversion alleles to determine the breakpoints:

- For 245 kb-INV: clone1, chrX-100377328 and chrX-100622017; clone2, chrX-100377337 and chrX-100622025 (coordinates in mm9).
- For *Chic1*-14 kbΔ: clone1, chrX-103370850 and chrX-103384956 (coordinates mm10).

The mouse mutant lines were generated following the strategy described by Wang et al. (2013) with minor modifications. Cas9 mRNA was transcribed *in vitro* from a T7-Cas9 pCR2.1-XL plasmid (Greenberg et al., 2017) using the mMESSAGE mMACHINE T7 ULTRA kit (Life Technologies) and purified with the RNeasy Mini kit (Qiagen), or bought from Tebu-bio (L-7206). The sgRNAs were amplified by PCR with primers containing a 5′ T7 promoter sequence from the plasmids used for mESC transfection. After gel purification, the T7-sgRNA PCR products were used as the template for *in vitro* transcription with the MEGAshortscript T7 kit (Life Technologies) and the products were purified using the MEGAclear kit (Life Technologies). Cas9 mRNA and the sgRNAs were eluted in DEPC-treated RNase-free water, and their quality was assessed by electrophoresis on an agarose gel after incubation at 95°C for 3 min with the denaturing agent provided with the *in vitro* transcription kits. Cas9 mRNA and sgRNAs (at 100 ng/µl and 50 ng/µl, respectively) were injected into the cytoplasm of mouse B6D2F1 zygotes from 8-week-old superovulated B6D2F1 (C57BL/6J×DBA2) females mated to stud males of the same background. Zygotes with well-recognized pronuclei were collected in M2 medium (Sigma) at E0.5. Injected embryos were cultured in M16 medium (Sigma) at 37°C under 5% CO_2, until transfer at the one-cell stage the same day or at the two-cell stage the following day to the infundibulum of the oviduct of a pseudogestant CD1 female at E0.5 (25-30 embryos were transferred per female). All weaned mice (N0) were genotyped for the presence of inversion alleles using the same genotyping primers as for mESC mutant lines. Mice carrying inversion alleles were crossed to B6D2F1 mice and their progeny screened again for the presence of the inversion allele. The F1 mice were considered the 'founders' and bred to B6D2F1 mice; their progeny was then intercrossed to generate homozygous mice and lines were kept in homozygosity.

RNA fluorescence *in situ* hybridization

RNA FISH was performed as described previously with minor modifications (Chaumeil et al., 2008). Briefly, differentiating mESCs were dissociated using accutase (Invitrogen) and adsorbed onto poly-L-lysine (Sigma)-coated coverslips #1.5 (1 mm) for 5 min. Cells were fixed with 3% paraformaldehyde in PBS for 10 min at room temperature and permeabilized for 5 min on ice in PBS containing 0.5% Triton X-100 and 2 mM ribonucleoside-vanadyl complex (New England Biolabs). Coverslips were preserved in 70% ethanol (EtOH) at −20°C. To begin the FISH experiments, coverslips were dehydrated through an EtOH series (80%, 95% and twice at 100%) and air dried quickly, then lowered onto a drop of the probe/hybridization buffer mix [50% formamide, 20% dextran sulfate, 2× saline-sodium citrate (SSC), 1 µg/µl BSA, 10 mM ribonucleoside-vanadyl complex] and incubated overnight at 37°C. The next day, coverslips were washed three times at 42°C in 50% formamide in 2× SSC (pH 7.2-7.4) and three times at 42°C in 2× SSC. Nuclei were counterstained with DAPI (0.2 mg/ml), coverslips were mounted [90% glycerol, 0.1× PBS, 0.1% p-phenylenediamine at pH 9], and cells were imaged using a wide-field DeltaVision Core microscope (Applied Precision).

The probes used were a *Huwe1* bacterial artificial chromosome (BAC; BACPAC Resources Center, RP24-157H12) and oligos (~75 nucleotides long) covering all *Xist* exons (Roche, custom design). The BAC was labeled using the Nick Translation kit (Abbot) following the manufacturer's instructions. Oligos were end-labeled with an Alexa488 fluorophore (Abbot). Probes were either EtOH precipitated (BAC) or vacuum dried (oligos) and resuspended in formamide with shaking at 37°C. BAC was coprecipitated with mouse Cot-1 DNA (Invitrogen), and competition to block repetitive sequences was performed for at least 20 min at 37°C, and after denaturation (75°C, 10 min). Probes were then mixed with one volume of 2× hybridization buffer.

Gene expression analysis (mESCs)

Cells were collected for gene expression analysis at 0 h, 12 h, 24 h, 36 h, 48 h and 60 h of EpiLSC differentiation. Cells were lysed with Trizol (Invitrogen), and RNA was isolated using the RNAeasy Mini kit (Qiagen), including DNase treatment. RNA samples were systematically run on an agarose gel to check their integrity. For reverse transcription, cDNA was synthesized from 0.5 µg of RNA using SuperScript™ III Reverse Transcriptase and random primers (both Invitrogen) according to the manufacturer's recommendations. Two independent reverse transcription experiments were carried out for each sample, pooled at the end and diluted 25-fold prior to qPCR or allelic expression analysis. No-reverse transcription controls were processed in parallel. The NanoString nCounter gene expression system (Geiss et al., 2008) was used to characterize transcriptional differences in wild-type and mutant mESCs systematically, before or during differentiation. We used 500 ng of total RNA from each sample for each nCounter hybridization round. We designed a customized probe codeset (van Bemmel et al., 2019) to identify nearly a hundred transcripts from *Xic* genes, other X-linked genes, pluripotency factors, differentiation markers, proliferation markers and normalization genes. Standard positive controls included in the kit were used for scaling the raw data. The genes *Actb*, *Rrm2* and *Sdha* were used for normalization. Differential expression was always calculated for samples run on the same nCounter hybridization.

Allelic expression analysis (mouse embryos)

Embryos were lysed in RLT buffer (Qiagen) supplemented with 0.01% 2-mercaptoethanol. After two rounds of vortexing (15 s each), lysates were applied directly to a QIAshredder spin column (Qiagen) and centrifuged for 3 min at >15,000 ***g***. RNA was extracted using the RNAeasy Mini kit, including DNase treatment, and following the manufacturer's instructions. RNA samples were systematically run on an agarose gel to check their integrity. cDNA was prepared as described above for the gene expression analysis of mESCs, and was then PCR amplified with biotinylated primers and pyrosequenced for allele quantification on a Pyromark Q24 system (Qiagen). The same PCR approach was performed on no-reverse transcription control samples to confirm the absence of genomic DNA contamination. The primers used were designed with PyroMark Assay Design software and validated on XX polymorphic genomic DNA at a ratio of 50:50% (±4%). A list of primers and SNPs used for allele quantification can be found in Galupa et al. (2020).

Chromosome conformation capture

3C libraries were prepared based on previous protocols (Nora et al., 2017; Rao et al., 2014), with some modifications. Crosslinked cells (in 2% formaldehyde; 10 million for each sample) were lysed in 10 mM Tris-HCl (pH 8), 10 mM NaCl, 0.2% NP-40, 1× complete protease inhibitor cocktail (Roche) for 15 min on ice. Nuclei were resuspended in 100 µl 0.5% SDS, incubated at 62°C for 10 min and quenched with 50 µl 10% Triton X-100 and 290 µl water at 37°C for 15 min. Digestion was performed overnight by adding 50 µl HindIII (New England Biolabs) buffer and 10 µl high-concentration HindIII and incubating the samples at 37°C in a thermomixer. Before this step, an aliquot was taken from each sample as an undigested control. Digests were heat inactivated for 20 min at 65°C and an aliquot was taken from each sample as a digested (unligated) control. Samples were cooled at room temperature for 10 min before adding the ligation cocktail. 3C libraries were ligated for 4 h at 25°C with 10U T4 ligase and ligation buffer (Thermo Fisher Scientific) in a thermomixer at 100 ***g***. Ligated samples were then centrifuged at 300 ***g***, resuspended in 240 µl 5% SDS and

Development (2022) 149, dev200568. doi:10.1242/dev.200568

1 mg Proteinase K, incubated at 55°C for 30 min, supplemented with 50 µl 5 M NaCl and incubated at 65°C for 4 h. DNA was then purified by adding 500 µl isopropanol, incubated at −80°C overnight, centrifuged at 13,000 ***g*** at 4°C, washed with 70% EtOH, air dried and resuspended in 100 µl water, followed by incubation with RNase A at 37°C for 1 h. 3C templates were quantified using Qubit DNA Broad-Range (Thermo Fisher Scientific) and diluted to 100 ng/µl. Libraries and respective controls (undigested and digested aliquots) were verified on a gel.

5C was performed using the method described by Nora et al. (2017), which adopts a single-PCR strategy to construct 5C-sequencing libraries from the 3C template. Briefly, four 10 µl 5C annealing reactions were assembled in parallel, each using 500 ng of 3C template, 1 µg salmon sperm (Thermo Fisher Scientific) and 10 fmol of each 5C oligonucleotide in 1× NEBuffer 4 [5C set of oligonucleotides described by Nora et al. (2012)]. Samples were denatured at 95°C for 5 min and incubated at 48°C for 16-18 h. Then, 10 µl of 1× Taq ligase buffer with 5U Taq ligase was added to each annealing reaction followed by incubation at 48°C for 4 h and 65°C for 10 min. Negative controls (no ligase, no template or no 5C oligonucleotide) were included during each experiment to ensure the absence of contamination. To attach Illumina-compatible sequences, 5C libraries were directly PCR amplified with primers harboring 50-mer tails containing Illumina sequences that anneal to the universal T3/T7 portion of the 5C oligonucleotides (Nora et al., 2017). For this, each 5C ligation reaction was used as the template for three parallel PCRs (12 PCRs in total), using 6 µl 5C ligation with 1.125 U AmpliTaq Gold (Thermo Fisher Scientific) per reaction in 1× PCR Buffer II, 1.8 mM $MgCl_2$, 0.2 mM dNTPs and 1.25 mM primers in a total of 25 ml. Cycling conditions were: 95°C for 9 min, 25 cycles of 95°C for 30 s, 60°C for 30 s, 72°C for 30 s followed by 72°C for 8 min. PCR products from the same 3C sample were pooled and run on a 2.0% agarose electrophoresis gel. 5C libraries (231 bp) were then excised and purified with the MinElute Gel Extraction kit (QIAGEN). Library concentrations were estimated using TapeStation (Agilent) and Qubit (Thermo Fisher Scientific), pooled and sequenced using 12 pM for loading on rapid flow cells using the HiSeq 2500 system (Illumina). Sequencing mode was set as 20 dark cycles followed by 80 bases in single-end reads (SR80).

Sequencing data were processed using our custom pipeline, 5C-Pro, available at https://github.com/bioinfo-pf-curie/5C-Pro. Briefly, single-end sequencing reads were first trimmed to remove Illumina adapters and aligned on an *in silico* reference of all pairs of forward and reverse primers using the bowtie2 software (Langmead and Salzberg, 2012). Aligned reads were then directly used to infer the number of contacts between pairs of forward and reverse primers, thus providing a 5C map at primer resolution. Based on our previous experiments, inefficient primers were discarded from the downstream analysis. Quality controls of the experiments were then performed using the HiTC BioConductor package (Servant et al., 2012). Data from biological replicates were pooled (summed) and binned using a running median (window=30 kb, final resolution=6 kb). We normalized 5C contacts for the total number of reads and filtered out outlier probes and singletons, as previously described (Hnisz et al., 2016; Nora et al., 2012; Smith et al., 2016). We also developed a novel method to exclude noisy contacts in the 5C maps, called 'neighborhood coefficient of variation', available at https://github.com/zhanyinx/Coefficient_Variation. Considering that the chromatin fiber behaves as a polymer, the contact frequency of a given pair of genomic loci (e.g. i and j) cannot be very different from that of fragments $i\pm N$ and $j\pm N$ if N is smaller (or in the order of) than the persistence length of the chromatin fiber. Hence, a given pixel in the 5C map (which is proportional to the contact frequency between the two corresponding loci) can be defined as noisy if its numerical value is too different from those corresponding to neighboring interaction frequencies. To assess the similarity of a given interaction with neighboring contacts operatively, we calculated the coefficient of variation (CV) of contacts (pixels in the 5C map) in a 10×10 square centered on every contact. We then set out to discard pixels for which the corresponding CV was above a threshold. Given that the distribution of the CV of all 5C samples in this study was bimodal around CV=1, we set the CV threshold to 1. Discarded contacts appeared as gray pixels in the differential 5C maps. For differential analysis between two samples of interest, we calculated the difference between z-scores determined for each individual map (Smith et al., 2016). Samples corresponding to inversions of genomic regions were mapped to a virtually inverted map before analysis. Samples corresponding to deletions were corrected for the new distance between genomic elements; this distance adjustment was performed along with the z-score calculation. 5C data for the E14 cell line (used as control) have been published previously (Galupa et al., 2020) but control and mutant samples were collected and processed in parallel.

Statistical analysis

For RNA FISH, nCounter and allelic expression analysis, details of the statistical analyses used are provided in the figure legends, figures and/or Results, including the statistical tests used, exact value of n and what n represents.

Acknowledgements

We are grateful to Isabelle Grandjean for help and advice with animal management; to Patricia Diabangouaya for help with mouse genotyping; to Chris Gard for help with gene expression analysis; and to Denis Krndija for critical reading of the manuscript. We thank all members of the Heard lab for advice, support and helpful comments and discussions. We are also thankful to facilities at the Institut Curie, including the Mouse Facility, the BDD team of the PICT-IBiSA, the ICGex NGS platform (in particular Sónia Lameiras and Sylvain Baulande), the Genomics Platform (in particular David Gentien, Cécile Reyes, Audrey Rapinat and Benoit Albaud) and the Bioinformatics Platform. We acknowledge the Zhang lab for sharing plasmids, the ENCODE Consortium, and the Bruneau, Ren, Sharp, Stamatoyannopoulos and Young labs for generating datasets used in this study.

Competing interests

The authors declare no competing or financial interests.

Author contributions

Conceptualization: R.G., L.G., E.H.; Methodology: R.G., C.P., E.P.N., F.E.M., C.J., M.B., K.A.; Software: N.S., Y.Z., J.G.v.B., L.G.; Formal analysis: R.G., N.S., Y.Z., J.G.v.B.; Investigation: R.G., C.P., E.P.N.; Data curation: R.G., N.S., Y.Z., J.G.v.B., L.G.; Writing - original draft: R.G.; Writing - review & editing: R.G., K.A., L.G., E.H.; Visualization: R.G., Y.Z.; Supervision: R.G., M.B., K.A., L.G., E.H.; Project administration: R.G., E.H.; Funding acquisition: E.H.

Funding

This work was supported by fellowships from the Région Ile-de-France (DIM Biothérapies Fellowship) and Fondation pour la Recherche Médicale (FDT20160435295) to R.G. E.H. was supported by a European Research Council Advanced Investigator award (ERC-2014-AdG no. 671027), by La Ligue contre le cancer (Labélisation) by the Fondation pour la Recherche Médicale (DEI20151234398), by the Agence Nationale de la Recherche (DoseX 2017 and ANR-11-BINF-0001), by Labex DEEP (ANR-11-LBX-0044) and by part of the Initiative d' Excellence of the Université de Recherche Paris Sciences et Lettres (ANR-10-IDEX-0001-02 PSL). Open Access funding provided by the Heard lab. Deposited in PMC for immediate release.

Data availability

All next-generation sequencing data generated in this study have been deposited in GEO under the accession number GSE124596 (5C data for E14 cell line) and GSE180617 (5C data for 245 kb-INV cell line). 5C data for E14 cell line (used as control) has been published previously (Galupa et al., 2020) but control and mutant samples were collected and processed in parallel.

Peer review history

The peer review history is available online at http://journals.biologists.com/dev/article-lookup/doi/10.1242/dev.200568.

References

Amândio, A. R., Lopez-Delisle, L., Bolt, C. C., Mascrez, B. and Duboule, D. (2020). A complex regulatory landscape involved in the development of mammalian external genitals. *eLife* **9**, e52962. doi:10.7554/eLife.52962

Anguera, M. C., Ma, W., Clift, D., Namekawa, S., Kelleher, R. J. and Lee, J. T. (2011). Tsx produces a long noncoding RNA and has general functions in the germline, stem cells, and brain. *PLoS Genet.* **7**, e1002248. doi:10.1371/journal.pgen.1002248

Attia, M., Rachez, C., De Pauw, A., Avner, P. and Rogner, U. C. (2007). Nap1l2 promotes histone acetylation activity during neuronal differentiation. *Mol. Cell. Biol.* **27**, 6093-6102. doi:10.1128/MCB.00789-07

Augui, S., Filion, G. J., Huart, S., Nora, E., Guggiari, M., Maresca, M., Stewart, A. F. and Heard, E. (2007). Sensing X chromosome pairs before X inactivation via a novel X-pairing region of the Xic. *Science* **318**, 1632-1636. doi:10.1126/science.1149420

Augui, S., Nora, E. P. and Heard, E. (2011). Regulation of X-chromosome inactivation by the X-inactivation centre. *Nat. Rev. Genet.* **12**, 429-442. doi:10.1038/nrg2987

Barakat, T. S., Gunhanlar, N., Pardo, C. G., Achame, E. M., Ghazvini, M., Boers, R., Kenter, A., Rentmeester, E., Grootegoed, J. A. and Gribnau, J. (2011). RNF12 activates Xist and is essential for X chromosome inactivation. *PLoS Genet.* **7**, e1002001. doi:10.1371/journal.pgen.1002001

Barakat, T. S., Loos, F., van Staveren, S., Myronova, E., Ghazvini, M., Grootegoed, J. A. and Gribnau, J. (2014). The trans-activator RNF12 and cis-acting elements effectuate X chromosome inactivation independent of X-pairing. *Mol. Cell* **53**, 965-978. doi:10.1016/j.molcel.2014.02.006

Castillo-Davis, C. I., Mekhedov, S. L., Hartl, D. L., Koonin, E. v. and Kondrashov, F. A. (2002). Selection for short introns in highly expressed genes. *Nat. Genet.* **31**, 415-418. doi:10.1038/ng940

Chaumeil, J., Augui, S., Chow, J. C. and Heard, E. (2008). Combined immunofluorescence, RNA fluorescent in situ hybridization, and DNA fluorescent in situ hybridization to study chromatin changes, transcriptional activity, nuclear organization, and X-chromosome inactivation. In: *The Nucleus* (*Methods in Molecular Biology*), Vol. **463**, (ed. R. Hancock), pp. 297-308. Humana Press. doi:10.1007/978-1-59745-406-3_18

Chiaromonte, F., Miller, W. and Bouhassira, E. E. (2003). Gene length and proximity to neighbors affect genome-wide expression levels. *Genome Res.* **13**, 2602-2608. doi:10.1101/gr.1169203

Davidson, I. F., Bauer, B., Goetz, D., Tang, W., Wutz, G. and Peters, J.-M. (2019). DNA loop extrusion by human cohesin. *Science* **366**, 1338-1345. doi:10.1126/science.aaz3418

de Wit, E., Vos, E. S. M., Holwerda, S. J. B., Valdes-Quezada, C., Verstegen, M. J. A. M., Teunissen, H., Splinter, E., Wijchers, P. J., Krijger, P. H. L. and de Laat, W. (2015). CTCF binding polarity determines chromatin looping. *Mol. Cell* **60**, 676-684. doi:10.1016/j.molcel.2015.09.023

Dekker, J. and Heard, E. (2015). Structural and functional diversity of topologically associating domains. *FEBS Lett.* **589**, 2877-2884. doi:10.1016/j.febslet.2015.08.044

Dekker, J. and Mirny, L. (2016). The 3D genome as moderator of chromosomal communication. *Cell* **164**, 1110-1121. doi:10.1016/j.cell.2016.02.007

Denker, A. and de Laat, W. (2016). The second decade of 3C technologies: detailed insights into nuclear organization. *Genes Dev.* **30**, 1357-1382. doi:10.1101/gad.281964.116

Despang, A., Schöpflin, R., Franke, M., Ali, S., Jerković, I., Paliou, C., Chan, W.-L., Timmermann, B., Wittler, L., Vingron, M. et al. (2019). Functional dissection of the Sox9–Kcnj2 locus identifies nonessential and instructive roles of TAD architecture. *Nat. Genet.* **51**, 1263-1271. doi:10.1038/s41588-019-0466-z

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. doi:10.1038/nature11082

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. doi:10.1038/nature14222

Dostie, J., Richmond, T. A., Arnaout, R. A., Selzer, R. R., Lee, W. L., Honan, T. A., Rubio, E. D., Krumm, A., Lamb, J., Nusbaum, C. et al. (2006). Chromosome conformation capture carbon copy (5C): a massively parallel solution for mapping interactions between genomic elements. *Genome Res.* **16**, 1299-1309. doi:10.1101/gr.5571506

Finn, E. H. and Misteli, T. (2019). A genome disconnect. *Nat. Genet.* **51**, 1205-1206. doi:10.1038/s41588-019-0476-x

Flavahan, W. A., Drier, Y., Liau, B. B., Gillespie, S. M., Venteicher, A. S., Stemmer-Rachamimov, A. O., Suvà, M. L. and Bernstein, B. E. (2016). Insulator dysfunction and oncogene activation in IDH mutant gliomas. *Nature* **529**, 110-114. doi:10.1038/nature16490

Franke, M., Ibrahim, D. M., Andrey, G., Schwarzer, W., Heinrich, V., Schöpflin, R., Kraft, K., Kempfer, R., Jerković, I., Chan, W.-L. et al. (2016). Formation of new chromatin domains determines pathogenicity of genomic duplications. *Nature* **538**, 265-269. doi:10.1038/nature19800

Fudenberg, G., Imakaev, M., Lu, C., Goloborodko, A., Abdennur, N. and Mirny, L. (2016). Formation of chromosomal domains by loop extrusion. *Cell Rep.* **15**, 2038-2049. doi:10.1016/j.celrep.2016.04.085

Furlan, G., Hernandez, N. G., Huret, C., Galupa, R., van Bemmel, J. G., Romito, A., Heard, E., Morey, C. and Rougeulle, C. (2018). The Ftx noncoding locus controls X chromosome inactivation independently of its RNA products. *Mol. Cell* **70**, 462-472.e8. doi:10.1016/j.molcel.2018.03.024

Galupa, R., Nora, E. P., Worsley-Hunt, R., Picard, C., Gard, C., van Bemmel, J. G., Servant, N., Zhan, Y., el Marjou, F., Johanneau, C. et al. (2020). A conserved noncoding locus regulates random monoallelic Xist expression across a topological boundary. *Mol. Cell* **77**, 352-367.e8. doi:10.1016/j.molcel.2019.10.030

Geiss, G. K., Bumgarner, R. E., Birditt, B., Dahl, T., Dowidar, N., Dunaway, D. L., Fell, H. P., Ferree, S., George, R. D., Grogan, T. et al. (2008). Direct multiplexed measurement of gene expression with color-coded probe pairs. *Nat. Biotechnol.* **26**, 317-325. doi:10.1038/nbt1385

Giorgetti, L., Galupa, R., Nora, E. P., Piolot, T., Lam, F., Dekker, J., Tiana, G. and Heard, E. (2014). Predictive polymer modeling reveals coupled fluctuations in chromosome conformation and transcription. *Cell* **157**, 950-963. doi:10.1016/j.cell.2014.03.025

Goloborodko, A., Marko, J. F. and Mirny, L. A. (2016). Chromosome compaction by active loop extrusion. *Biophys. J.* **110**, 2162-2168. doi:10.1016/j.bpj.2016.02.041

Gontan, C., Achame, E. M., Demmers, J., Barakat, T. S., Rentmeester, E., van IJcken, W., Grootegoed, J. A. and Gribnau, J. (2012). RNF12 initiates X-chromosome inactivation by targeting REX1 for degradation. *Nature* **485**, 386-390. doi:10.1038/nature11070

Greenberg, M. V. C., Glaser, J., Borsos, M., Marjou, F. E., Walter, M., Teissandier, A. and Bourc'his, D. (2017). Transient transcription in the early embryo sets an epigenetic state that programs postnatal growth. *Nat. Genet.* **49**, 110-118. doi:10.1038/ng.3718

Guo, Y., Xu, Q., Canzio, D., Shou, J., Li, J., Gorkin, D., Jung, I., Wu, H., Zhai, Y., Tang, Y. et al. (2015). CRISPR inversion of CTCF sites alters genome topology and enhancer/promoter function. *Cell* **162**, 900-910. doi:10.1016/j.cell.2015.07.038

Haarhuis, J. H. I., van der Weide, R. H., Blomen, V. A., Yáñez-Cuna, J. O., Amendola, M., van Ruiten, M. S., Krijger, P. H. L., Teunissen, H., Medema, R. H., van Steensel, B. et al. (2017). The cohesin release factor WAPL restricts chromatin loop extension. *Cell* **169**, 693-707.e14. doi:10.1016/j.cell.2017.04.013

Hansen, A. S., Pustova, I., Cattoglio, C., Tjian, R., Darzacq, X., Bernstein, B. E., Nusbaum, C., Myers, R. M., Brown, M., Li, W. et al. (2017). CTCF and cohesin regulate chromatin loop stability with distinct dynamics. *eLife* **6**, e25776. doi:10.7554/eLife.25776

Heard, E., Mongelard, F., Arnaud, D. and Avner, P. (1999). Xist yeast artificial chromosome transgenes function as X-inactivation centers only in multicopy arrays and not as single copies. *Mol. Cell. Biol.* **19**, 3156-3166. doi:10.1128/MCB.19.4.3156

Hnisz, D., Weintraub, A. S., Day, D. S., Valton, A.-L., Bak, R. O., Li, C. H., Goldmann, J., Lajoie, B. R., Fan, Z. P., Sigova, A. A. et al. (2016). Activation of proto-oncogenes by disruption of chromosome neighborhoods. *Science* **351**, 1454-1458. doi:10.1126/science.aad9024

Huang, H., Zhu, Q., Jussila, A., Han, Y., Bintu, B., Kern, C., Conte, M., Zhang, Y., Bianco, S., Chiariello, A. M. et al. (2021). CTCF mediates dosage- and sequence-context-dependent transcriptional insulation by forming local chromatin domains. *Nat. Genet.* **53**, 1064-1074. doi:10.1038/s41588-021-00863-6

Jonkers, I., Barakat, T. S., Achame, E. M., Monkhorst, K., Kenter, A., Rentmeester, E., Grosveld, F., Grootegoed, J. A. and Gribnau, J. (2009). RNF12 is an X-Encoded dose-dependent activator of X chromosome inactivation. *Cell* **139**, 999-1011. doi:10.1016/j.cell.2009.10.034

Kim, Y., Shi, Z., Zhang, H., Finkelstein, I. J. and Yu, H. (2019). Human cohesin compacts DNA by loop extrusion. *Science* **366**, 1345-1349. doi:10.1126/science.aaz4475

Koch, L. (2019). Toppling TAD tenets. *Nat. Rev. Genet.* **20**, 565. doi:10.1038/s41576-019-0164-9

Langmead, B. and Salzberg, S. L. (2012). Fast gapped-read alignment with Bowtie 2. *Nat. Methods* **9**, 357-359. doi:10.1038/nmeth.1923

Lee, J. T. and Lu, N. (1999). Targeted mutagenesis of Tsix leads to nonrandom X inactivation. *Cell* **99**, 47-57. doi:10.1016/S0092-8674(00)80061-6

Luikenhuis, S., Wutz, A. and Jaenisch, R. (2001). Antisense transcription through the Xist locus mediates Tsix function in embryonic stem cells. *Mol. Cell. Biol.* **21**, 8512-8520. doi:10.1128/MCB.21.24.8512-8520.2001

Lupiáñez, D. G., Kraft, K., Heinrich, V., Krawitz, P., Brancati, F., Klopocki, E., Horn, D., Kayserili, H., Opitz, J. M., Laxova, R. et al. (2015). Disruptions of topological chromatin domains cause pathogenic rewiring of gene-enhancer interactions. *Cell* **161**, 1012-1025. doi:10.1016/j.cell.2015.04.004

McCord, R. P., Kaplan, N. and Giorgetti, L. (2020). Chromosome conformation capture and beyond: toward an integrative view of chromosome structure and function. *Mol. Cell* **77**, 688-708. doi:10.1016/j.molcel.2019.12.021

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. doi:10.1038/nature11049

Nora, E. P., Goloborodko, A., Valton, A.-L., Gibcus, J. H., Uebersohn, A., Abdennur, N., Dekker, J., Mirny, L. A. and Bruneau, B. G. (2017). Targeted degradation of CTCF decouples local insulation of chromosome domains from genomic compartmentalization. *Cell* **169**, 930-944.e22. doi:10.1016/j.cell.2017.05.004

Ogawa, Y. and Lee, J. T. (2003). Xite, X-inactivation intergenic transcription elements that regulate the probability of choice. *Mol. Cell* **11**, 731-743. doi:10.1016/S1097-2765(03)00063-7

Paliou, C., Guckelberger, P., Schöpflin, R., Heinrich, V., Esposito, A., Chiariello, A. M., Bianco, S., Annunziatella, C., Helmuth, J., Haas, S. et al. (2019). Preformed chromatin topology assists transcriptional robustness of Shh during limb development. *Proc. Natl. Acad. Sci. USA* **116**, 12390-12399. doi:10.1073/pnas.1900672116

Phillips-Cremins, J. E., Sauria, M. E. G., Sanyal, A., Gerasimova, T. I., Lajoie, B. R., Bell, J. S. K., 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. doi:10.1016/j.cell.2013.04.053

Rao, S. S. P., 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. et al. (2014). A 3D map of the human genome at kilobase resolution reveals principles of chromatin looping. *Cell* **159**, 1665-1680. doi:10.1016/j.cell.2014.11.021

Rao, S. S. P., Huang, S.-C., Glenn St Hilaire, B., Engreitz, J. M., Perez, E. M., Kieffer-Kwon, K.-R., Sanborn, A. L., Johnstone, S. E., Bascom, G. D., Bochkov, I. D. (2017). Cohesin loss eliminates all loop domains. *Cell* **171**, 305-320.e24. doi:10.1016/j.cell.2017.09.026

Rastan, S. (1982). Timing of X-chromosome inactivation in postimplantation mouse embryos. *J. Embryol. Exp. Morphol.* **71**, 11-24. doi:10.1242/dev.71.1.11

Rastan, S. and Brown, S. D. (1990). The search for the mouse X-chromosome inactivation centre. *Genet. Res.* **56**, 99-106. doi:10.1017/S0016672300035163

Rodríguez-Carballo, E., Lopez-Delisle, L., Zhan, Y., Fabre, P. J., Beccari, L., El-Idrissi, I., Nguyen Huynh, T. H., Ozadam, H., Dekker, J. and Duboule, D. (2017). The HoxD cluster is a dynamic and resilient TAD boundary controlling the segregation of antagonistic regulatory landscapes. *Genes Dev.* **31**, 2264-2281. doi:10.1101/gad.307769.117

Rogner, U. C., Spyropoulos, D. D., Le Novère, N., Changeux, J. P. and Avner, P. (2000). Control of neurulation by the nucleosome assembly protein-1-like 2. *Nat. Genet.* **25**, 431-435. doi:10.1038/78124

Sanborn, A. L., Rao, S. S. P., Huang, S.-C., Durand, N. C., Huntley, M. H., Jewett, A. I., Bochkov, I. D., Chinnappan, D., Cutkosky, A., Li, J. et al. (2015). Chromatin extrusion explains key features of loop and domain formation in wild-type and engineered genomes. *Proc. Natl. Acad. Sci. USA* **112**, E6456-E6465. doi:10.1073/pnas.1518552112

Schwarzer, W., Abdennur, N., Goloborodko, A., Pekowska, A., Fudenberg, G., Loe-Mie, Y., Fonseca, N. A., Huber, W., H. Haering, C., Mirny, L. (2017). Two independent modes of chromatin organization revealed by cohesin removal. *Nature* **551**, 51-56. doi:10.1038/nature24281

Servant, N., Lajoie, B. R., Nora, E. P., Giorgetti, L., Chen, C.-J., Heard, E., Dekker, J. and Barillot, E. (2012). HiTC: exploration of high-throughput 'C' experiments. *Bioinformatics* **28**, 2843-2844. doi:10.1093/bioinformatics/bts521

Smith, E. M., Lajoie, B. R., Jain, G. and Dekker, J. (2016). Invariant TAD boundaries constrain cell-type-specific looping interactions between promoters and distal elements around the CFTR locus. *Am. J. Hum. Genet.* **98**, 185-201. doi:10.1016/j.ajhg.2015.12.002

Stavropoulos, N., Lu, N. and Lee, J. T. (2001). A functional role for Tsix transcription in blocking Xist RNA accumulation but not in X-chromosome choice. *Proc. Natl. Acad. Sci. USA* **98**, 10232-10237. doi:10.1073/pnas.171243598

Symmons, O., Pan, L., Remeseiro, S., Aktas, T., Klein, F., Huber, W. and Spitz, F. (2016). The Shh topological domain facilitates the action of remote enhancers by reducing the effects of genomic distances. *Dev. Cell* **39**, 529-543. doi:10.1016/j.devcel.2016.10.015

Szabo, Q., Bantignies, F. and Cavalli, G. (2019). Principles of genome folding into topologically associating domains. *Sci. Adv.* **5**, eaaw1668. doi:10.1126/sciadv.aaw1668

Tang, Z., Luo, O. J., Li, X., Zheng, M., Zhu, J. J., Szalaj, P., Trzaskoma, P., Magalska, A., Wlodarczyk, J., Ruszczycki, B. et al. (2015). CTCF-mediated human 3D genome architecture reveals chromatin topology for transcription. *Cell* **163**, 1611-1627. doi:10.1016/j.cell.2015.11.024

Tian, D., Sun, S. and Lee, J. T. (2010). The long noncoding RNA, Jpx, is a molecular switch for X chromosome inactivation. *Cell* **143**, 390-403. doi:10.1016/j.cell.2010.09.049

van Bemmel, J. G., Galupa, R., Gard, C., Servant, N., Picard, C., Davies, J., Szempruch, A. J., Zhan, Y., Żylicz, J. J., Nora, E. P. et al. (2019). The bipartite TAD organization of the X-inactivation center ensures opposing developmental regulation of Tsix and Xist. *Nat. Genet.* **51**, 1024-1034. doi:10.1038/s41588-019-0412-0

Wang, H., Yang, H., Shivalila, C. S., Dawlaty, M. M., Cheng, A. W., Zhang, F. and Jaenisch, R. (2013). One-step generation of mice carrying mutations in multiple genes by CRISPR/Cas-mediated genome engineering. *Cell* **153**, 910-918. doi:10.1016/j.cell.2013.04.025

Williamson, I., Kane, L., Devenney, P. S., Flyamer, I. M., Anderson, E., Kilanowski, F., Hill, R. E., Bickmore, W. A. and Lettice, L. A. (2019). Developmentally regulated Shh expression is robust to TAD perturbations. *Development* **146**, dev179523. doi:10.1242/dev.179523

Wutz, G., Várnai, C., Nagasaka, K., Cisneros, D. A., Stocsits, R. R., Tang, W., Schoenfelder, S., Jessberger, G., Muhar, M., Hossain, M. J. et al. (2017). Topologically associating domains and chromatin loops depend on cohesin and are regulated by CTCF, WAPL, and PDS5 proteins. *EMBO J.* **36**, 3573-3599. doi:10.15252/embj.201798004

Zhan, Y., Mariani, L., Barozzi, I., Schulz, E. G., Blüthgen, N., Stadler, M., Tiana, G. and Giorgetti, L. (2017). Reciprocal insulation analysis of Hi-C data shows that TADs represent a functionally but not structurally privileged scale in the hierarchical folding of chromosomes. *Genome Res.* **27**, 479-490. doi:10.1101/gr.212803.116