

Published in final edited form as:

Nat Genet. 2019 June ; 51(6): 1024–1034. doi:10.1038/s41588-019-0412-0.

The bipartite TAD organization of the X-inactivation center ensures opposing developmental regulation of *Tsix* and *Xist*

Joke G. van Bommel^{†,*,1,2}, Rafael Galupa^{†,*,1}, Chris Gard¹, Nicolas Servant^{3,4}, Christel Picard¹, James Davies⁵, Anthony James Szempruch⁶, Yinxiu Zhan^{7,8}, Jan J. ylicz^{1,9}, Elphège P. Nora¹⁰, Sonia Lameiras¹¹, Elzo de Wit¹³, David Gentien¹², Sylvain Baulande¹¹, Luca Giorgetti⁷, Mitchell Guttman⁶, Jim R. Hughes⁵, Douglas R. Higgs⁵, Joost Gribnau², Edith Heard^{*,1}

¹Institut Curie, CNRS UMR3215, INSERM U934, 26 rue d'Ulm, Paris F-75248, France ²Erasmus MC, Rotterdam, The Netherlands ³Institut Curie, PSL Research University, INSERM, U900, F-75005 Paris, France ⁴MINES ParisTech, PSL Research University, CBIO-Centre for Computational Biology, F-75006 Paris, France ⁵Medical Research Council, Molecular Haematology Unit, Weatherall Institute of Molecular Medicine, University of Oxford, John Radcliffe Hospital, Headington, Oxford, OX3 9DS, UK ⁶Division of Biology and Biological Engineering, California Institute of Technology, Pasadena, CA 91125, USA ⁷Friedrich Miescher Institute for Biomedical Research, Maulbeerstrasse 66, 4058 Basel, Switzerland ⁸University of Basel, CH-4003 Basel, Switzerland ⁹University of Cambridge, Department of Physiology, Development and Neuroscience, Downing Street, Cambridge CB2 3EG, UK ¹⁰Gladstone Institute of Cardiovascular Disease and Roddenberry Center for Stem Cell Biology and Medicine at Gladstone, San Francisco, CA 94158, USA ¹¹Institut Curie Genomics of Excellence (ICGex) Platform, Institut Curie Research Center, Paris, France ¹²Institut Curie, PSL Research University,

Users may view, print, copy, and download text and data-mine the content in such documents, for the purposes of academic research, subject always to the full Conditions of use:http://www.nature.com/authors/editorial_policies/license.html#terms

*to whom correspondence should be addressed: edith.heard@curie.fr and jokevbemmel@gmail.com.

[†]J.G.vB. and R.G. contributed equally to this work

[#]Present addresses: Gladstone Institute of Cardiovascular Diseases, San Francisco, CA 94158, USA (J.G.vB.); European Molecular Biology Laboratory, Heidelberg, 69117, Germany (R.G.)

Reporting Summary

Further information on research design is available in the Nature Research Life Sciences Reporting Summary linked to this article.

Data availability

Data have been deposited in the NCBI GEO under the accession number GSE111205. Reagents, cell lines and other data supporting the findings of this study are available from the corresponding author upon request.

Code availability

Our custom pipeline for 5C data processing, 5C-Pro, is available at <https://github.com/bioinfo-pf-curie/5C-Pro>. Custom codes used in this study will be provided upon request.

Author Contributions

Following CRediT (Contributor Role Taxonomy) guidelines, <http://docs.casrai.org/CRediT>

Conceptualization: J.G.vB. and E.H. (lead); R.G., L.G. and E.P.N. (supporting); J.G. and E.H. (supervision). Formal Analysis: J.G.vB. and N.S. (lead), R.G., A.J.S., Y.Z., E.dW., L.G. (equal). Investigation: J.G.vB. and R.G. (lead), C.G., A.J.S., C.P., E.P.N., J.J.Z. and S.L. (supporting). Resources: J.D., Y.Z., L.G., J.D., J.R.H. and D.R.H. Writing – Original Draft Preparation: J.G.vB. and E.H. (lead), E.P.N., R.G. and C.G. (supporting), with input from all authors. Writing – Review & Editing: R.G. and E.H. (lead), J.G.vB. (supporting). Visualization: J.G.vB. and R.G. Supervision: J.G.vB., R.G. and E.H. (lead), J.D., D.G., S.B., M.G., J.R.H., D.R.H. and J.G. (supporting). Funding Acquisition: E.H., J.G.vB. and J.G.

Competing Interests Statement:

The authors declare no competing interests.

Translational Research Department, Genomics platform, Paris, F-75248, France ¹³Oncode Institute and Division of Gene Regulation, the Netherlands Cancer Institute, Plesmanlaan 121, 1066 CX Amsterdam, the Netherlands

Abstract

The mouse X-inactivation center (*Xic*) locus represents a powerful model for understanding the links between genome architecture and gene regulation, with the non-coding genes *Xist* and *Tsix* showing opposite developmental expression patterns while being organized as an overlapping sense/antisense unit. The *Xic* is organized into two topologically associating domains (TADs) but the role of this architecture in orchestrating *cis*-regulatory information remains elusive. To explore this, we generated genomic inversions that swap the *Xist/Tsix* transcriptional unit and place their promoters in each other's TAD. We found that this led to a switch in their expression dynamics: *Xist* became precociously and ectopically up-regulated, both in male and female pluripotent cells, while *Tsix* expression aberrantly persisted during differentiation. The topological partitioning of the *Xic* is thus critical to ensure proper developmental timing of X inactivation. Our study illustrates how the genomic architecture of *cis*-regulatory landscapes can affect the regulation of mammalian developmental processes.

Introduction

Spatiotemporal regulation of gene expression during mammalian development often involves distal *cis*-regulatory elements, which can be located tens to hundreds of kilobases away from their target genes ¹. Dynamic interactions between enhancers and their target promoters preferentially occur within 200kb-1Mb topologically associating domains (TADs) ^{2,3}, within which smaller self-interacting domains can form ^{4,5}. TADs represent a functionally privileged scale in the hierarchical folding of chromosomes ⁶, and are largely conserved across cell types and species ^{2,7}, in contrast to smaller self-interacting domains. Transcription within TADs is often co-regulated ^{3,8}, and most enhancer-promoter pairs reside within the same TAD ^{6,9,10}. TADs are frequently delimited by genomic sites bound by CTCF (CCCTC-binding factor) and cohesin ^{2,4,6,11}, both of which have been shown to contribute to TAD organization ^{4,5,12–14}. Genetic alterations of CTCF binding sites (CBSs) have shown that the genomic location and orientation of CBSs can determine the directionality of long-range interactions and TAD organization ^{15–21}. However, the functional relationship between TADs, their boundaries and transcriptional regulation remains unclear.

Here we used the X-inactivation center (*Xic*), a paradigm of developmentally regulated loci with complex *cis*-regulatory landscapes, to understand how genomic architecture might affect transcriptional regulation. The *Xic* is required for the initiation of X-chromosome inactivation (XCI) in female mammals ^{22–25}. It harbors both *Xist*, the key long noncoding RNA (lncRNA) that triggers XCI, and its antisense transcription unit, *Tsix*. *Xist* becomes monoallelically upregulated at the onset of XCI and produces a lncRNA that coats the future inactive X chromosome in *cis* and triggers its silencing. *Tsix* is transcribed antisense to *Xist*

and represses *Xist* expression during differentiation (reviewed in 26). The developmental regulation of *Xist* and *Tsix* during random XCI can be explored using mouse embryonic stem cells (mESCs). In pluripotent male and female mESCs, *Tsix* is robustly expressed, while *Xist* is barely transcribed 27,28,29. Differentiation of mESCs is associated with downregulation of *Tsix* and the activation of *Xist* expression, which is very transient in males 29, and more robust and long-lasting in females, presumably due to a double dose of X-linked factors (reviewed in 26). *Tsix* expression during early differentiation is believed to contribute to the monoallelic regulation of *Xist* in female differentiating mESCs 30. Initially, *Tsix* is highly expressed from both X chromosomes; during differentiation it becomes repressed on the future inactive X (expressing *Xist*) and remains transcribed only from the future active X 31–33. In summary, *Xist* and *Tsix* adopt opposite transcriptional fates during mESC differentiation, despite their overlapping transcriptional units. Interestingly, their promoters lie in separate, adjacent TADs 3 (Fig. 1A-C). The TAD containing the *Xist* promoter (named TAD-E 3, ~550 kb) also includes several coding (e.g. *Rnf12*) and non-coding sequences (e.g. *Jpx*, *Ftx*) reported to regulate *Xist* 34–36; likewise, the TAD harboring the *Tsix* promoter (named TAD-D 3, ~250 kb) also contains putative *Tsix* cis-regulators (e.g. *Xite*, *Linx*) – reviewed in 26,37–39. Importantly, the opposite transcriptional fates of *Xist* and *Tsix* during differentiation are coordinated with those of the other loci within their respective TADs 3. This bipartite organization of the *Xic* into two TADs has thus been proposed to separate the *Xist/Tsix* regulatory landscapes and to promote coordinated expression of the genes within each TAD. However, whether appropriate *Xist* and *Tsix* regulation does require such partitioning of the promoters from each other within separate TADs remains unknown. Here we explore the extent to which TAD environments can have an impact on accurate gene regulation by generating genomic inversions that swap the *Xist* and *Tsix* promoters between neighboring TADs and assessing the degree to which three-dimensional (3D) organization, appropriate gene regulation and XCI timing are affected.

Results

Genomic inversions involving the *Xist/Tsix* transcriptional unit place their promoters into each other's TADs

We designed and generated a genomic inversion of the *Xist/Tsix* locus (~40 kb), including their respective promoters, in male mESCs (see Methods for detailed description; Fig. 1D, Supplementary Fig. 1A). To investigate whether this inversion switches the topological environment of the *Xist* and *Tsix* promoters, we used Capture-C 40,41 to obtain high-resolution interaction profiles for the two promoters in both wild type and inverted alleles (Fig. 1E for clone #1, Supplementary Fig. 2C for clone #2, Fig. 1C for Capture-C viewpoints). In wild type (WT) cells, the interaction profiles clearly reflected the presence of the two TADs: the *Tsix* promoter preferentially interacted with sequences within TAD-D, while the *Xist* promoter preferentially interacted with sequences within TAD-E. In cells harboring the 40 kb [*Tsix-Xist*] inversion, the interaction profiles of the promoters switched: the *Tsix* promoter preferentially interacted with sequences within TAD-E (70% in INV vs 30% in WT, averages), while the *Xist* promoter preferentially interacted within TAD-D (71% in INV vs. 40% in WT, averages) (Fig. 1E, Supplementary Fig. 2C, and

Supplementary Fig. 2F for visual comparison of the wildtype and inverted alleles). Modeling of the background contact probabilities at either side of the viewpoints, before and after inversion, confirmed that the promoters switched their interaction preferences from one TAD to the other (Supplementary Fig. 2G). To evaluate the structure of the TADs in the inverted allele, we then performed 5C (chromosome conformation capture carbon-copy 42) across a 4.5-Mb region centered on *Xist* (as before 3). The 5C interaction profile of the inverted allele was very similar to that of the wild type, based on differential maps and insulation scores (Fig. 1F-H for clone #1, Supplementary Fig. 3C-E for clone #2). We thus successfully placed the *Xist* and *Tsix* promoters in each other's TADs, without compromising the insulation between the TADs or changing the overall TAD structure. The overall absence of changes in TAD structure is consistent with models of CTCF-mediated loops (reviewed in 43), given that the CBS configuration between the wild type and inverted alleles is equivalent (compare Fig. 1C with 1D). Together, our results reveal that the promoters of *Xist* and *Tsix*, when inverted, switch their interaction preference, losing to some extent their original interactions, and adopting interactions predominantly within the new TADs in which they lie. Our 40-kb inversion therefore allows us to investigate the consequences of changing the topological and *cis*-regulatory environment of a promoter.

We also designed and generated a second inversion in male mESCs spanning a 70-kb region including not only *Xist* and *Tsix* but also their immediately adjacent loci, *Jpx* and *Xite* (see Methods for detailed description; Fig. 2A, Supplementary Fig. 1). *Xite* is a reported enhancer of *Tsix* 44,45, while *Jpx* is co-expressed with *Xist* during differentiation 34, and has been reported to activate *Xist*, either in *cis* 46, in *trans* 34,47 or both 48. We performed Capture-C for the TSSs of *Xist* and *Tsix* in male ESCs harboring the 70-kb [*Xite-Jpx*] inversion and found that, similarly to the 40-kb [*Tsix-Xist*] inversion, the interaction profiles for both *Tsix* and *Xist* promoters were switched when compared to wild type (Fig. 2B for clone #1, Supplementary Fig. 2D for clone #2). To investigate the new interactomes in the 70-kb inversion in more detail, we performed Capture-C for *Xite* (overlapping CBS #2) or *Jpx* TSS (Fig. 2C for clone #1, Supplementary Fig. 2E for clone #2, Fig. 1C for Capture-C viewpoints). In wild type cells, *Xite*'s interactions are predominantly within TAD-D (85%, average), showing particularly strong interactions with *Linx* and *Chic1*, as previously reported 3,49. The TSS of *Jpx* has an interaction profile mostly restricted to TAD-E (72%, average). In cells harboring the 70-kb [*Xite-Jpx*] inversion, both *Xite* and *Jpx* showed preferential interactions within the new TAD each now lies in (83% and 72%, respectively). We found that *Xite* no longer interacted with *Linx* and *Chic1* and instead displayed strong interactions within TAD-E, whereas *Jpx* showed a fairly flat profile of interactions, nevertheless restricted to TAD-D. To evaluate the structure of the TADs in the 70-kb inverted allele, we performed 5C and observed significant changes (Fig. 2E-G for clone #1, Supplementary Fig. 3C-E for clone #2). In particular, there was a localized loss of structure within TAD-D, when comparing the 70-kb inversion to wild type (black arrowheads in Fig. 2E-F). This corresponded to interactions established by *Xite* in the WT, and by *Jpx* in the inverted allele, indicating that *Jpx* is unable to replace *Xite* original interactions, as already suggested by Capture-C (Fig. 2C-D). We also observed a localized gain of interactions (a "flame") within TAD-E (green arrowheads in Fig. 2E-F). These are ectopic interactions between *Xite* and sequences within TAD-E in the inverted allele, confirming what we

observed with the Capture-C for *Xite* (Fig. 2C; Supplementary Note 1). Thus the 70-kb inversion revealed similarities and differences compared to the 40-kb inversion. The promoters of *Xist* and *Tsix* switched interaction profiles in both cases, and the bipartite TAD structure was conserved. However, significant changes occurred within the TADs in the 70-kb inversion, apparently associated with loss of *Xite* within TAD-D and its new position within TAD-E. Notably, compared to the 40-kb inversion, the 70-kb inversion displaced an additional pair of CBSs found within *Xite* (CBSs #1-2); therefore, the final configuration of CBSs in the inverted allele is not equivalent to wild type (compare Fig. 1C to Fig. 2A). The structural phenotypes observed in the 70-kb inversion, where the orientation of the two *Xite* CBSs become convergent with those present within TAD-E, are consistent with recent models for CTCF-mediated loops, whereby CBSs in convergent orientation are more likely to interact with each other, forming the apex of cohesin-extruded loops 15–21.

***Xite* contributes to insulating the *Xist/Tsix* TADs and establishes a TAD boundary in an orientation-dependent manner**

The above results suggested that the CBSs within *Xite* have a strong influence on TAD structure and organization. To test this more directly, we generated deletions and inversions of the *Xite* locus (~18 kb) (Fig. 3A and 3F, Supplementary Fig. 1A). 5C analysis of the *Xite* male mESCs revealed a significant increase (45%) in contacts between TAD-D and TAD-E (Fig. 3B-C for clone #1, Supplementary Fig. 3C,D for clone #1, Supplementary Fig. 3F), suggesting that *Xite* is important for topological insulation between these TADs, also indicated by the insulation score profile (Fig. 3D, Supplementary Fig. 3E). The *Xite* locus includes an enhancer element (*XiteE*, ~12 kb) 44, and a structural element (*XiteC*, ~6 kb) 49 with two CBSs (CBSs #1-2) in the same orientation. Transcription factor binding to enhancers as well as CBSs have been proposed to play roles in looping and long-range interactions. To distinguish between the contributions of the enhancer versus the structural element, *XiteE* and *XiteC* were each deleted alone (Supplementary Fig. 1). 5C analysis of *XiteC* and *XiteE* mESCs (Fig. 3E, Supplementary Note 2) revealed an increase in contacts between TAD-D and TAD-E only in *XiteC* mESCs (Fig. 3C), suggesting that it is the structural element within *Xite*, harboring two CBSs, that contributes to insulating TAD-D and TAD-E from each other.

We also investigated the topological consequences of inverting *Xite* (Fig. 3F). In *Xite*-INV cells, we observed a localized gain of interactions between *Xite* (inverted) and sequences along TAD-E (Fig. 3G, Supplementary Figs. 3C-E and 4A), accompanied by a localized loss of structure within TAD-D. This suggests that inverting *Xite* switched its interacting preference from TAD-D to TAD-E, consistent with the new orientations of the two CBSs present within *Xite* (compare Fig. 3F to Fig. 1C). These results also support a role for *Xite* in the structural changes observed in the 70-kb inversion (Fig. 2E-F). Importantly, the insulation score profile of the *Xite*-INV allele revealed that the minimum value between the two TADs was shifted ~60 kb compared to WT (Fig. 3H, Supplementary Fig. 3E). This indicates that the position of the TAD boundary was changed, also corroborated by an increase in size of TAD-E (compare arrowheads in Fig. 3G and Fig. 3I). To distinguish between the contributions of the enhancer versus the structural element in this altered architecture, we also generated inversions of *XiteE* and *XiteC* alone (Supplementary Fig. 1).

Once again, inversion of *XiteC*, but not of *XiteE*, led to the same changes as observed with *Xite*-INV (Fig. 3J; Supplementary Fig. 4A). Based on the analysis of this collection of *Xite* deletions and inversions, we conclude that the *Xite* structural element plays an important role in shaping the topological landscape of TAD-D and TAD-E and in promoting insulation between the two TADs.

The fact that loss of *Xite* did not result in a collapse or merging of the two TADs suggests that insulation between the TADs can be maintained by other elements (see Discussion for more details). Presumably, the numerous different CBSs present within both TAD-D and TAD-E, and at the boundary between them, can act redundantly to provide some degree of insulation between these TADs, even when the *Xite* structural element is removed. Deleting any of those CBSs alone may have little effect; indeed, we deleted CBS#3 (or RS1450), which was previously postulated to be a putative boundary, 50,51 and observed no loss of TAD insulation (Supplementary Note 3 and Supplementary Fig. 5).

Xist* is aberrantly up-regulated in male mESCs when placed in the regulatory landscape of *Tsix

We next addressed whether expression of *Xist* and *Tsix* was affected in the 40-kb and 70-kb inverted alleles, in which their respective promoters are exposed to each other's topological and regulatory environment (Fig. 1-2). We assessed the expression status of several genes, including those within the *Xic*, in wild type and mutant male mESCs. In both the 40-kb [*Tsix-Xist*] and the 70-kb [*Xite-Jpx*] inversions, expression of *Tsix* was not significantly affected, despite its new genomic location (Fig. 4A, Supplementary Fig. 6A for second clone). On the other hand, *Xist* became significantly upregulated when placed within TAD-D: 38-fold and 18-fold in two independent 40-kb clones; 4-fold and 2.6-fold in two independent 70-kb clones (Fig. 4A, Supplementary Fig. 6A). Upregulation of *Xist* does not normally occur in male mESCs, presumably due to *Xist*-repressive mechanisms operating in mESCs (e.g. pluripotency, *Tsix*) and to the absence of a double dose of X-linked *Xist* activators. Relocation of the *Xist* promoter into TAD-D and its *cis*-regulatory environment in our inversions appears to override such mechanisms/requirements. In addition, transcriptional analysis in our collection of *Xite* inversions revealed that *Xist* expression is sensitive to the orientation of *Xite* structural element (*XiteC*) (see Supplementary Note 4 and Supplementary Fig. 4B-E). Taken together, our results highlight how critical the *cis*-regulatory landscape is for appropriate *Xist* regulation.

We also observed that, for the 40-kb [*Tsix-Xist*] inversion, *Cdx4* (in TAD-D) and *Ftx*, *Xpct* and *Rnf12* (in TAD-E) were all downregulated 2-fold or more (red bars in Supplementary Fig. 6A). To address whether this could be due to ectopically expressed *Xist* RNA leading to some degree of gene silencing, we performed RNA fluorescence *in situ* hybridization (FISH), RNA antisense purification (RAP) and mRNA sequencing on our male mutants and controls. *Xist* expression is mostly absent in wild type male mESCs as measured by RNA FISH (<6% of cells); however, in cells harboring the 40-kb [*Tsix-Xist*] inversion, we observed that ~70%/40% of cells expressed *Xist* (in the two clones, respectively), with a significant percentage of cells (~30%/13%) exhibiting *Xist* accumulation or even larger cloud formation (red bars in Fig. 4B for clone #1 and Supplementary Fig. 6C for clone #2).

Concomitantly, we observed significant *Xist* RNA coating of the X chromosome by RAP (Fig. 4C-D, in red), following a pattern similar to the *Xist*-coated inactive X in mouse fibroblasts 52 (dotted line Fig. 4D; $r = 0.75$). In the 70-kb [*Xite-Jpx*] inversion, *Xist* RNA showed abnormal accumulation by RNA FISH in 9-14% of cells (blue bars in Fig. 4B for clone #1 and Supplementary Fig. 6C for clone #2), but no substantial X-chromosome coating by RAP (Fig. 4C-D, in blue). This is consistent with the lesser degree of *Xist* up-regulation observed in the 70-kb inversion compared to the 40-kb inversion (Fig. 4A, Supplementary Fig. 6A). Finally, we determined mRNA transcription levels of autosomal and X-linked genes by RNA-seq. RNA-seq analyses confirmed the changes in gene expression within the *Xic* described above (Supplementary Fig. 6B). For a chromosome-wide analysis of X-linked silencing, we calculated X/Autosome (X/A) expression ratios as a measure of dosage compensation 53; in wild type male mESCs, this ratio was 0.97 (Fig. 4E). The X/A ratio in the two 40-kb [*Tsix-Xist*] clones decreased mildly but significantly to 0.91 and 0.92 (Wilcoxon Rank Sum test, $P = 2.2 \times 10^{-16}$) (Fig. 4E, Supplementary Fig. 6D, in red). This mild silencing was specific to genes enriched for *Xist* RNA at their TSS (Fig. 4F, in red). Limited X-linked gene silencing was significant in one 70-kb [*Xite-Jpx*] inversion clone but not in the other (Fig. 4E, Supplementary Fig. 6D, in blue), consistent with the other results.

Taken together, our results show that relocating the *Xist* promoter from TAD-E into TAD-D, with or without its local regulator *Jpx*, leads to ectopic *Xist* activation in male mESCs (summarized in Fig. 4G; see Supplementary Note 5). This underlines the importance of the TAD environment for the appropriate transcriptional regulation of *Xist*. It also highlights the relevance of the spatial partitioning between TAD-D and TAD-E and their *cis*-regulatory landscapes, as the *Xist/Tsix* TAD boundary seems to physically insulate the *Xist* promoter from the activating influence of elements within the *Tsix* TAD.

Perturbed initiation of X-chromosome inactivation when inverting the genomic organization of *Xist* and *Tsix*

To investigate the consequences of our genomic inversions on the regulation of *Xist* and *Tsix* during X- inactivation, we generated the 40-kb and 70-kb inversions in female mESCs with polymorphic X chromosomes (PGK12.1-derived, see Methods). We obtained a heterozygous 40-kb [*Tsix-Xist*] inversion (40 kb INV/WT), a heterozygous 70-kb [*Xite-Jpx*] inversion (70 kb WT/INV), and a homozygous 40-kb [*Tsix-Xist*] inversion (40 kb INV/INV) (Fig. 5A). To trigger XCI, we differentiated wild type and mutant female mESCs towards epiblast-like stem cells (EpiLSCs) 54,55 (see Methods and Fig. 5B) and analyzed the kinetics of *Xist* RNA accumulation using RNA FISH at days 0, 2 and 4. We also analyzed day 10 EpiLSCs that were independently generated; at this stage, in the majority of cells, *Xist* is expressed from the inactive X, whereas *Tsix* is repressed on the inactive X and has also started to be silenced on the active X. In undifferentiated mESCs harboring the heterozygous or homozygous 40-kb [*Tsix-Xist*] inversion, we observed abnormal *Xist* RNA accumulation in ~40% or ~55% of the cells, respectively (Fig. 5C, second and third panel, Day 0); 16% of homozygous-inversion cells showed ectopic *Xist* RNA accumulation on both chromosomes. In female mESCs harboring the heterozygous 70-kb [*Xite-Jpx*] inversion, ectopic *Xist* RNA accumulation was also observed, though in fewer cells (~10%) (Fig. 5C, second panel, Day

0). Allelic quantification of Xist RNA by pyrosequencing confirmed that the ectopic Xist RNA accumulation originated from the inverted allele in both 40-kb and 70-kb inversions (Fig. 5D, Day 0). These results demonstrate that in female mESCs, similarly to male mESCs, *Xist trans*-repression mechanisms operating in mESCs are not sufficient to prevent *cis*-activation of the *Xist* promoter when it is placed within *Tsix*'s regulatory landscape.

During differentiation, the initial differences between wild type and mutant cells decreased over time (Fig. 5C, Day 2 and Day 4); at day 4, the proportion of cells accumulating Xist RNA on one or two X chromosomes was equivalent across all cell lines (~70%) (Fig. 5C, Day 4). Furthermore, in day 10 EpiLSCs cells harboring an inversion, the percentage of cells with accumulation of *Xist* was significantly reduced to ~60%, compared to ~90% in wild type (Fig. 5C, Day 10). Notably, all cell lines harboring inversions exhibited normal differentiation hallmarks (Supplementary Fig. 7C-D), and did not suffer X-chromosome loss. In addition, pyrosequencing confirmed that throughout differentiation the fraction of Xist RNA from the inverted allele decreased significantly over time (Fig. 5D). This suggests that *Xist* expression and/or accumulation is impaired during differentiation of female mESC harboring the inversions (see Supplementary Notes 6 and 7).

***Tsix* silencing is impaired when its promoter is exposed to *Xist*'s regulatory landscape**

We then evaluated *Tsix* nascent transcription by RNA FISH in day 10 EpiLSCs harboring the 40-kb and 70-kb inversions. During differentiation of female mESCs, *Tsix* transcription becomes rapidly repressed on the chromosome that expresses *Xist* 31–33. This transition from bi-allelic to mono-allelic *Tsix* expression is almost complete in day 10 EpiLSCs, with less than 5% of cells with a *Xist* accumulation showing *Tsix* expression from both X chromosomes (Fig. 5E, grey bars). Interestingly, in both inversions, we found that global *Tsix* levels were increased (Supplementary Fig. 7G) and observed a significantly increased proportion of *Xist* positive cells that expressed *Tsix* from the two X chromosomes (~10-20% Chi-squared test, $P = 0.001$) (Fig. 5E). We conclude that exposing *Tsix* to the regulatory environment that normally belongs to *Xist* impairs *Tsix* silencing during differentiation, and leads to the co-occurrence of Xist RNA accumulation and *Tsix* expression from the same allele.

We also assessed whether Xist RNA clouds were affected in those cells in which *Xist* expression co-occurred with bi-allelic *Tsix* expression. We selected *Xist*-positive cells, and compared the *Xist* clouds in cells with bi-allelic *Tsix* expression to cells with no or mono-allelic *Tsix* expression (Fig. 5F). In cells that express *Tsix* from one or neither of the X chromosomes, we observed typical Xist RNA cloud formation in the majority of cells (86-95%). In cells that express *Tsix* from both X chromosomes, this proportion was significantly reduced (29-52%) and a significant proportion of these cells (48-71%) showed a reduced or dispersed Xist RNA accumulation. Despite a low total number of cells with prolonged bi-allelic *Tsix* expression, the correlation between bi-allelic *Tsix* expression and reduced/dispersed Xist RNA clouds is highly significant (Fig. 5F). These results suggest that prolonged *Tsix* expression might contribute to the reduced levels of Xist RNA during differentiation in the inversions (see also Supplementary Note 6).

Taken together, our results indicate that the original configuration prevents the *Xist* and *Tsix* promoters from frequently interacting with each other's regulatory environment, and that this is critical for appropriate *Xist* expression and timely *Tsix* repression during differentiation. We propose that the partitioning into two TADs ensures correct XCI by insulating *Xist* and *Tsix* regulatory environments from each other.

Discussion

In this study, we addressed the extent to which topological and regulatory environments are critical for the appropriate developmental regulation of *Xist* and *Tsix*. We found that the promoters of *Xist* and *Tsix*, when relocated into each other's TAD, lost their original chromosomal interactions to some extent, and adopted novel interactions mostly restricted to, and guided by, the new TAD within which they lay, consistent with similar findings at other loci in the genome 16,56,57. This allowed us to specifically address the functional impact of switching the TAD environment of two promoters at once.

Interestingly, the *Tsix* and *Xist* promoters exhibited quite different behaviors when moved to a new TAD environment. While *Tsix* steady-state expression levels seemed unaffected in mESCs (discussed below), we observed aberrant activation of *Xist* upon relocation of its promoter to TAD-D. This led to *Xist* RNA accumulation, X chromosome coating and partial gene silencing (Fig. 4). This ectopic upregulation of *Xist* could be due to the loss of its native *cis*-regulatory landscape and/or gain of a new environment. Given that a deletion spanning most of *Xist* TAD, from *Jpx* up to *Rnf12*, did not lead to aberrant upregulation of *Xist* expression 46, we favor the second hypothesis: exposure to TAD-D and the *cis*-regulatory environment of *Tsix*, which is normally active in mESCs 3, led to *Xist* upregulation. This ectopic upregulation is milder in the 70-kb inversion than in the 40-kb inversion probably because the TAD-D landscape is a less potent "activator" without *Xite*.

In contrast, *Tsix* expression in male mESCs seems unchanged when its promoter is placed in the *Xist* TAD (Fig. 4), suggesting that *Tsix* is much less sensitive to its *cis*-regulatory environment in the pluripotent state (see also Supplementary Note 8). However, during female differentiation, we showed that displacing the promoter of *Tsix* into the *Xist* TAD led to co-occurrence of *Tsix* expression and *Xist* accumulation on the same allele (Fig. 5). This abnormal *Tsix* expression is probably sustained or reignited by the *cis*-regulatory environment of TAD-E, in which other loci are normally able to "escape" *Xist*-mediated silencing (e.g. *Jpx*, *Ftx*) 58,59. Taken together, our results highlight how the sensitivity of a promoter to its *cis*-environment can depend on distinct cell states, and that in some instances the *cis*-regulatory environment can have a dominant effect over *trans*-acting mechanisms.

Our study also provides important insights into the *Xist/Tsix* TAD boundary, a rather typical TAD boundary on the X chromosome in terms of its insulation strength (Supplementary Fig. 3G). Inversion of the 40-kb region led to a switch in the *Xist* and *Tsix* interaction profiles, which implies that elements within the inverted region are able to restrict and direct the interactions of the *Xist* and *Tsix* promoters. Moreover, we identified *Xite* as a critical element for TAD boundary position and insulation (Fig. 3). *Xite* not only lies very close to the *Xist/Tsix* boundary but also harbors a pair of CBSs that mediate strong intra-TAD

interactions. Our data thus provide support for previous proposals that the interactions within TADs can contribute to defining TAD boundaries, by preventing interactions with neighboring TADs and creating a boundary by default 49,60. The formation of the *Xist* and *Tsix* TADs likely arises from a combination of both intra-TAD scaffolding, and insulator elements.

In summary, the appropriate kinetics and allelic regulation of *Xist* and *Tsix* during XCI require *Xic*'s *cis*-regulatory landscape, but also its spatial partitioning, which prevents exposure of *Xist* to the *cis*-regulatory elements within the *Tsix* TAD and vice versa. Identifying the exact sequences responsible for this bipartite TAD structure (and its boundary), as well as all the sequences that regulate *Xist* and *Tsix* in a TAD-specific manner, are interesting avenues of future research. In light of our results, we propose that the spatial partitioning of the *Xic* is an integral part of its definition as the minimal locus necessary and sufficient to trigger XCI 22–25. The two *Xic* TADs serve to (i) insulate *Xist* physically from frequent *cis*-regulatory influences of the *Tsix* TAD in ES cells, (ii) provide the required *cis*-regulatory landscape for *Xist* to be upregulated during differentiation, and (iii) shield *Tsix* from the regulatory influences exerted by the *Xist* TAD during early differentiation. Our study further validates the *Xic* as a genetic paradigm to understand how 3D genome organization can affect the regulation of transcriptional dynamics during developmental processes.

Methods

Design of inversions

The 40-kb [*Tsix-Xist*] inversion starts 315 bp upstream of the *Tsix* TSS, including the major *Tsix* promoter 45, the H3K27ac, H3K4me3, DNaseI and p300 enriched region and the minisatellite region DxPas34, which is characterized as a component of the enhancer of *Tsix* 45. The 40-kb [*Tsix-Xist*] inversion ends ~2 kb upstream of *Xist*, including the *Xist* promoter 63, which is enriched in H3K27ac and corresponds to a DNase I hypersensitive site (Supplementary Fig. 1A). The 70-kb [*Xite-Jpx*] inversion starts ~23 kb upstream of *Tsix*, including the minor *Tsix* promoter 45, the *Xite* enhancer region (see below for details on exact annotation) 44,45, the H3K27ac, H3K4me1, p300 enriched regions, and the DNase I hypersensitive site, and it ends 53 bp prior to the end of the *Jpx* transcript (Supplementary Fig. 1A). Notably, these genomic inversions do not disrupt the sequence motifs of any of the known binding sites for pluripotency factors (Supplementary Fig. 1A), including REX1, the target of RNF12 64, and do not affect CTCF binding in the region (Supplementary Fig. 1B).

In UCSC RefSeq annotation 61, *Tsix* is annotated with the minor upstream promoter, while transcriptome analysis shows *Tsix* to be transcribed from the major downstream promoter only. *Xite* is not annotated at all in UCSC RefSeq annotation. Originally, *Xite* was annotated based on a deletion (L) covering chrX:100621540-100634048 mm9 44. *Xite* has often been annotated as the region in between the minor and major *Tsix* promoter. This annotation overlaps with the downstream part of the original annotation which has been shown to contain enhancer potential, while the upstream part has not 44. This annotation runs up to the major *Tsix* promoter and includes a pair of CBSs, while the original annotation ends just before this pair of CBSs. Both annotations include the XhoI-StuI fragment that is required

for the enhancer potential of *Xite* 45. In this study, *Xite* has been depicted from the minor to the major *Tsix* promoter, including the pair of CBSs (*XiteC* in this study) and the enhancer activity containing XhoI-StuI fragment (*XiteE* in this study).

mESC culture and differentiation

Feeder-independent mouse ESCs (male: E14Tg2a.4; female: Pgk#106) were cultured on gelatin-coated flasks. Male (E14Tg2a.4) cells were grown in Glasgow medium supplemented with 2 mM L-Glutamine, 0.1 mM NEAA and 1 mM sodium pyruvate (GIBCO). Female (Pgk#106) cells were grown in DMEM (GIBCO). Both media contained 15% FBS (GIBCO), 10 M β -mercaptoethanol (Sigma) and 1,000 U/ml of leukemia inhibitory factor (LIF, Chemicon). Cells were cultivated in 8% CO₂. Male E14Tg2a.4 cells were obtained from the Sanger Institute. E14Tg2a.4 is a subclone of E14Tg2a. E14Tg2a cells were derived from the 129P2 mouse strain and Hprt mutated 65. Female Pgk#106 were derived from Pgk12.1 66 and contain a heterozygote tetO knock-in allele, as described previously 67.

Before mESC to EpiLSC induction, female mESCs were grown for two days in medium as described above, supplemented with 3 μ M GSK3 inhibitor (CT-99021) and 1 μ M MEK1 inhibitor (PD-0325901). To induce mESCs into EpiLSCs 54,55, cells were plated on fibronectin coated 6-well plates (10 μ g/ml in PBS, Chemicon FC010) and grown in N2B27 medium (Stem Cell Sciences) supplemented with 10 ng/ml Fgf2 (R&D) and 20 ng/ml Activin A (R&D) at a cell density of 200,000 cells/cm². Cells were grown for 2 or 4 days, while refreshing the medium every day. For day 10 EpiLSCs, cells were plated as described above. At day 2, cells were incubated with JAK inhibitor (Calbiochem) for 24 hours. Cells were split at day 3, day 5 and day 8 and collected for analysis at day 10.

Plasmids

sgRNAs were designed at each side of the region to be deleted or inverted, and cloned by annealing oligo pairs in pX330 for Cas9 nuclease, pX335 for Cas9 nickase and pX459 for Cas9 nuclease with Puromycin selection marker, according to the protocol described in 68. pX330, pX335 and pX459 plasmids were a gift from Feng Zhang, respectively Addgene #42230 68, Addgene #42335 68 and Addgene #48139 69.

For *Xite* mutants a pair of TALENs was designed on each side of the region to be deleted or inverted. For TALEN construction, the TALE Toolbox kit (Addgene) and the protocol described in Sanjana et al. (2012) were used. TALEN backbones were modified to contain a CAGGS promoter instead of the default CMV promoter 49.

sgRNA and TALEN sequences and Cas9 plasmids for each cell line were used as listed in Supplementary Table 1.

Generation of mutant mESC lines

mESCs were transfected with CRISPR plasmids or TALENs by means of the Amaxa 4D Nucleofector™ system (Lonza) using the P3 Primary Cell 4D-Nucleofector X Kit (V4XP-3024) and the CG-104 (for E14Tg2a) or CG-110 (for Pgk#106) programs. 5 million

cells were transfected with 5 µg of each CRISPR plasmid, resuspended and seeded at serial dilutions in 10 cm dishes to ensure optimal density for colony picking. Transfected cells were selected with puromycin for 48 h, and grown for 8-10 days. Single colonies or pools of colonies were picked into 96-well plates. Genomic DNA was isolated in 96-well plates and used for PCR-based screening. The strategy was inspired on the Epigenesys protocol by Nora and Heard, 2012, described here: <https://www.epigenesys.eu/en/protocols/genome-engineering/816-engineering-genomic-deletions-and-inversions-in-mouse-es-cells-using-custom-designed-nucleases>. Positive clones were subsequently re-seeded at single-cell dilution in 96-well plates, followed by PCR screening to obtain mono-clonal colonies. The PCR product of the selected clones was Sanger-sequenced to determine the exact location and allele of the respective deletion or inversion. Wild type alleles were also sequenced, to ensure their integrity. Karyotypes of selected clones were checked by metaphase chromosome spreads. Coordinates (mm9) of the deletions and inversions, can be found in Supplementary Table 1 and sequencing results of the deletions and inversion boundaries can be found in Supplementary Figure 8A-B.

5C - Chromosome Conformation Capture

5C was performed as described in Nora et al. (2017) (with some modifications, see below) 14, which incorporates *in situ* ligation 5 and adopts a single-PCR strategy to construct 5C-sequencing libraries from the 3C template.

Cross-linking was performed as described previously 14 using 2% formaldehyde. For 3C, 10 million cells were lysed in 1 ml 10 mM Tris-HCl pH8.0, 10 mM NaCl 0.2% NP40 for 15 min, pelleted at 4°C and washed once with 1 ml lysis buffer. Cells were then resuspended in a 1.5 ml tube in 100 µl 0.5% SDS in water, incubated at 62°C for 10 min, and immediately supplemented with 50 µl 10% Triton X-100 and 290 µl water, followed by incubation at 37°C for 15 min. For restriction digestion, 50 µl 10× NEB2 buffer was added to the samples. Next, 50 µl were taken as an undigested control before adding 1,000 U of HindIII (high-concentration, NEB) for overnight incubation in a thermomixer at 1,400 rpm. The next morning, cells were incubated for 20 min at 65°C, cooled at room temperature and 20 µl were taken as a digestion control. Ligation buffer and 10U T4 ligase (ThermoFisher cat 15224) were added and after 4 h incubation at 25°C in a thermomixer at 1,000 rpm, nuclei were centrifuged at 2,000 rpm, resuspended in 240 µl of 5% SDS and 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 hours. DNA was then purified by adding 500 µl isopropanol, incubated at -80°C overnight, centrifuged at 12,000 rpm at 4°C, washed with 70% ethanol, air dried and resuspended in 100 µl water, followed by incubation with RNase A at 37°C. 3C templates were quantified using Qubit DNA Broad-Range (ThermoFisher) and diluted to 100 ng/µl.

For 5C we used the set of oligonucleotides described in 3. Four 10 µl 5C annealing reactions were assembled in parallel, each using 500 ng of 3C template, 1 µg Salmon Sperm (ThermoFisher), 10 fmol of each 5C oligonucleotide in 1× NEBuffer™ 4. Samples were denatured at 95°C for 5 min and incubated at 48°C for 16-18 h. 10 µl of 1× Taq ligase buffer with 5U Taq ligase were added to each annealing reaction followed by incubation at 48°C for 4h and 65°C for 10 min. Negative controls (no ligase, no template, 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 14 that anneal to the universal T3/T7 portion of the 5C oligonucleotides. For this, each 5C ligation reaction was used as the template for three parallel PCRs (12 PCRs total), using per reaction 6 μ l of 5C ligation with 1.125 U AmpliTaq Gold (ThermoFisher) in 1 \times PCR buffer II, 1.8 mM MgCl₂, 0.2 mM dNTPs, 1.25 mM primers in 25 ml total. Cycling conditions were 95°C for 9 min, 25 cycles of 95°C for 30 sec, 60°C for 30 sec, 72°C for 30 sec 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 (ThermoFisher), pooled and sequenced using 12 pM for the 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). Between 7.2 and 28 million reads were sequenced per library (see Supplementary Table 2).

5C analysis

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 70. 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 the primer resolution. Based on our previous experiments, inefficient primers were discarded from downstream analysis. Quality controls of the experiments were then performed using the HiTC BioConductor package 71.

Data from biological replicates were pooled (summed) and binned using a running median (window = 30 kb, final resolution = 6 kb). We normalized 5C interactions for the total number of reads and filtered out outlier probes and singletons, as previously described 3,72,73. Insulation scores calculated according to Crane et al 62. We also developed a novel method to exclude noisy interactions 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 those 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 operatively assess the similarity of a given interaction with neighboring interactions, we calculated the coefficient of variation (CV) of interactions (pixels in the 5C map) in a 10 \times 10 square centered on every interaction. We then set out to discard pixels for which the corresponding coefficient of variation was bigger than a threshold.

Given that the distribution of the coefficient of variation of all 5C samples in this study is bimodal around CV = 1 (Supplementary Fig. 2B), we set the CV threshold to 1. Discarded

interactions appear as grey pixels in the differential 5C maps. For differential analysis between two samples of interest (generally wild type versus mutant), we calculated the \log_2 ratio between the respective filtered maps, and/or determined the difference between Z-scores calculated for each individual map 73. 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.

Capture-C

Capture probes were designed using CapSequm 40. Sequences and genomic coordinates of capture probes can be found in Supplementary Table 3. Capture-C was performed as described in 41. In short, cells were fixed in 2% Formaldehyde. 10 million cells were lysed in 10 mM Tris-HCl, pH 8, 10 mM NaCl, 0.2% NP-40, 1x complete protease inhibitor cocktail (Roche) and homogenized using a Dounce homogenizer. To prepare 3C templates, chromatin was reconstituted in 664 μ l water, and 50 μ l of sample were taken as an undigested control. To prepare for digestion, 80 μ l of restriction buffer and 10 μ l 20% SDS was added, and incubated for 1h at 37°C in a Thermomixer at 1,400 rpm. Next, 66 μ l 20% Triton X 100 was added and incubation was continued for one hour. Digestion was performed overnight by adding 10 μ l DpnII (NEB, 50,000 U/ml) for 3 hours, followed by 10 μ l DpnII overnight, and another 10 μ l DpnII the next morning for 3-4 hours, at 37°C in a Thermomixer. 100 μ l were taken as an unligated control. Digests were heat inactivated, diluted by adding 500 μ l water and ligated overnight with 8 μ l T4 Ligase (30 U/ μ l EL0013 ThermoScientific) and 133 μ l Ligation buffer at 16°C in a Thermomixer at 1,400 rpm. DNA was then purified by overnight Proteinase K (ThermoScientific ELO0491) treatment at 65°C, followed by phenol/chloroform and chloroform extraction, and ethanol precipitation at -80°C for 4 hours.

Sequencing libraries were prepared from 6 μ g of 3C library by sonication using an S220 focused ultrasonicator (Covaris) to 200 bp and indexed using NEBNext® reagents (New England Biolabs) according to the manufacturer's protocol. Subsequently, 2 μ g indexed sequencing library was enriched for fragments of interest by incubation with 13 pmol of biotinylated oligonucleotides (Integrated DNA Technologies, Supplementary Table 3) using the SeqCap EZ system (#06953212001, Roche/NimbleGen) according to manufacturer's instructions. Two rounds of captures of respectively 72 and 24 hours were performed. In the case of multiple captures for the same viewpoints, capture reactions were pooled and volumes multiplied by the number of reactions. Library size was confirmed using TapeStation D1000 (Agilent), and DNA concentrations were determined using Qubit (Thermo Fisher Scientific).

Capture-C libraries were sequenced on a MiSeq instrument (Illumina) using 150 base paired-end reads (PE150). Between 2.0 and 9.4 million read pairs were sequenced per library (see Supplementary Table 2). Raw reads were first trimmed using the Trim Galore! pipeline (http://www.bioinformatics.babraham.ac.uk/projects/trim_galore/), and then processed using the HiC-Pro pipeline 74 (v2.8.0), until the detection of valid interaction products. Interaction products including the viewpoint of choice were selected using the

make Viewpoint HiC-Pro utility. For plotting, interactions within 1 kb around the viewpoint were excluded, after which interaction frequencies were normalized to the number of interactions per DpnII fragment per 10,000 total interactions within the analyzed region (chrX:100214149-101420149), followed by a running mean with a window size of 7 DpnII fragments.

Directional Capture-C signals were calculated as the percentage of interactions occurring in the 5' or 3' landscapes at either side of the respective viewpoint, relative to the total number of interactions within the visualized window (chrX: 100214149-101420149|mm9) after excluding the interactions within a 20kb window around the viewpoint and normalized to the size of the 5' or 3' landscapes.

RNA expression analysis

For gene expression profiling, cells were lysed by using Trizol (Invitrogen), then RNA was isolated using Silica column purification (Qiagen RNeasy Mini kit) including DNase treatment. Nanostring nCounter[®] quantification 75 was performed using 500 ng of total RNA per sample on a custom expression Codeset (Supplementary Table 4). Positive controls were used for scaling of the raw data and the housekeeping genes *Actb*, *Rrm2* and *Sdha* were used for normalization. Expression data for the different mutants is always compared to a wild type that was processed on the same nCounter[®] run.

For allelic expression analysis of *Xist*, cDNA was prepared using random priming with Superscript III (Life Technologies), followed by PCR amplification using primers with the following sequences: [B_{tn}]AGAGAGCCCAAAGGGACAAA and TGTATAGGCTGCTGGCAGTCC. Successfully amplified PCR products from cDNA were purified and annealed with the pyrosequencing primer, GCTGGCAGTCCTGA, and analyzed for the G/A SNP at mm9 position 100665998 using the PyroMark Q24 (Qiagen). PCR and sequencing primers were designed using the PyroMark Assay Design software.

RNA-seq libraries were prepared from 500 ng of DNase-treated total RNA (RIN > 7) with the TruSeq Stranded mRNA kit (Illumina). Sequencing was performed in using 100 base paired-end (PE100) reads using the Illumina HiSeq2000 platform (Supplementary Table 2). RNA sequencing reads have been aligned on the Mouse reference genome (mm9) using the STAR mapper (v2.5.2b) 76 and the GENCODE (vM1) annotations 77, allowing up to 6 mismatches and reporting only unique hits. Reads counts per gene were also generated with STAR and combined across samples to generate the raw counts table. Gene counts were filtered to be > 1 in at least two samples and normalized by the trimmed mean of M values (TMM) using the edgeR package 78,79. Differential expression was determined using the EdgeR package 78,79. X:A expression ratio as the level of expression of X-linked genes divided by the global level of expression of the autosomal genes was calculated as described in 80. Only genes with CPM > 1 in the wild type were used for this analysis.

RNA FISH

RNA FISH was performed as previously described 81, using probes specified in Supplementary Table 5. 3D image stacks were analyzed using custom-made ImageJ macros.

In brief, cells were dissociated using accutase and subsequently attached to poly-Lysine (0.1% w.v. in water) coated coverslips by 15 min incubation. Attached cells were washed in PBS and fixed in 3% paraformaldehyde for 10 min. Cells were permeabilized on ice for 5 min in 1× PBS, 0.5% Triton X-100 and 2 mM ribonucleoside vanadyl complex (NEB) and washed and stored in 70% ethanol at -20°C. Prior to RNA FISH, cells were dehydrated through washes with 80%, 95%, 100% and 100% ethanol. Nascent *Tsix* transcripts were detected with the pLG10 plasmid as probe 49, using 1 µg plasmid for nick translation. Genomic coordinates of sequences recognized by the pLG10 probe are chrX: 100641751-100646253 (mm9). Xist RNA was detected using a custom designed strand-specific probe that covers all exons with ~75-bp-long oligonucleotides (chrX: 100656573-100678635, mm9), which are end-labeled with the Alexa488 fluorophore (Roche). Per coverslip, 50 ng of labeled pLG10 and/or 100 ng of end-labeled oligonucleotides were ethanol-precipitated and (co)hybridized in FISH hybridization buffer (50% Formamide, 20% Dextran sulfate, 2× SSC, 1 µg/µl BSA, 10 mM ribonucleoside vanadyl complex) overnight at 37°C. The next day, cells were washed at 42°C three times for 5 min with 50% formamide / 2× SSC at pH7.3 and three times for 5 min with 2× SSC. Samples were counterstained with 0.2 mg/ml DAPI, washed two times for 5 min at RT with 2× SSC, and mounted with 90% glycerol, 0.1× PBS and 0.1% p-phenylenediamine (Sigma). Images were acquired on a DeltaVision widefield system, a minimum of 100 cells per clone per replicate were counted (Supplementary Table #6)

RNA antisense purification (RAP) coupled with DNA sequencing

5 million mouse ES cells (E14, 40-kb or 70-kb inversion) were crosslinked with 2 mM disuccinimidyl glutarate and 3% formaldehyde. Cells were lysed and Xist RNA was purified as previously described 52. Xist RNA was captured using antisense 5' biotinylated 90-mer DNA probes, which spanned the *Xist* sequence. Captures were carried out as described previously 52. Briefly, lysates were precleared with streptavidin magnetic beads for 30 min at 37°C. After preclearing, supernatant was mixed with 1 µg of *Xist* probes and incubated at 37°C for 2 h. A portion of this sample was retained as the input+probe control. Probe-RNA hybrids were captured with streptavidin magnetic beads at 37°C for 30 min. Beads were washed four times with GuSCN hybridization buffer (20 mM Tris-HCl pH 7.5, 7 mM EDTA, 3 mM EGTA, 150 mM LiCl, 1% NP-40, 0.2% N-lauroylsarcosine, 0.1% sodium deoxycholate, 3 M guanidine thiocyanate, 2.5 mM TCEP) followed by two washes with GuSCN wash buffer (20 mM Tris-HCl pH 7.5, 10 mM EDTA, 1% NP-40, 0.2% N-lauroylsarcosine, 0.1% sodium deoxycholate, 3 M guanidine thiocyanate, 2.5 mM TCEP). Captured DNA was eluted with 15 U RNase H in 20 µl RNase H buffer at 37°C for 1 h, as previously described 52. After elution, protein crosslinking was reversed for both capture and input+probe samples, and DNA was prepared for Illumina sequencing using a NEBNext® ChIP-Seq Library Prep kit. Capture efficiency was assessed by qPCR for the Xist RNA in the elution relative to the input+probe samples.

Sequencing alignment and analysis was carried out as previously described 52,82,83. In brief, libraries were sequenced using 35-bp paired-end reads. Between 5.8 and 921.2 million read pairs were sequenced per library (see Supplementary Table 2). After sequencing, adapter indexes were trimmed from all reads and samples were aligned to the mouse genome

(build: mm9). Duplicates were marked and removed using picard and this data was used for enrichment calculations across the X chromosome at 10-kb window sizes. All elution counts were normalized against input+probe controls from their respective samples. In addition, mouse lung fibroblast data from GSE46918 was processed as above and used for enrichment calculations.

Chromatin immunoprecipitation

Cells were crosslinked as previously described 84. Briefly, cells were initially fixed in 1.5 mM ethylene glycolbis[succinimidyl succinate] (EGS), next 1% formaldehyde was added and further incubated for 10 min. After rinsing twice in PBS, cells were scraped and pellets snap frozen. Fixed pellets were lysed in 2 ml Nuclear Lysis Buffer (0.5% Triton X-100, 0.1 M sucrose, 5 mM MgCl₂, 1 mM EDTA, 10 mM Tris-HCl pH 8.0, 1× Protease Inhibitors) for 10 min on ice and douce-homogenized. Nuclei were pelleted and resuspended in 0.3 ml Lysis Buffer (1 mM EDTA, 0.5 mM EGTA, 10 mM Tris pH 8.0, 0.5% N-Lauroylsarcosine, 1× Protease Inhibitors). Chromatin was sheered using a BioruptorPlus (Diagenode) set to high for 30 cycles. Unsonicated chromatin was removed by centrifugation. Supernatant was diluted with 1.2 ml of Dilution Buffer (1.25% Triton, 0.125% sodium deoxycholate, 6 mM EDTA, 10mM Tris-HCl pH8.0, 1× Protease Inhibitors) and antibody coated beads were added (per IP: 8 µl of anti-CTCF antibody [Active Motif; #61311] and 40 µl of Protein-A dynabeads). Samples were left rotating overnight at 4°C. Following day, beads were magnet-separated and washed 4 times with Low Salt Buffer (0.1% SDS; 1% TritonX-100; 2 mM EDTA; 20 mM Tris-HCl, pH 8.1; 150 mM NaCl; 0.1% sodium deoxycholate), 2 times with High Salt Buffer (0.1% SDS; 1% TritonX-100; 2 mM EDTA; 20 mM Tris-HCl, pH 8.1; 360 mM NaCl; 0.1% sodium deoxycholate) and 2 times with LiCl buffer (0.25 M LiCl; 1% NP40; 1.1% sodium deoxycholate; 1 mM EDTA; 10 mM Tris-HCl pH 8.1). Prior to elution all samples were rinsed once in TE. ChIP-DNA was eluted in ProtK-Digestion Buffer (20 mM HEPES; 1 mM EDTA; 0.5% SDS; 0.8 mg/ml Proteinase K) for 15 min at 56 °C. Beads were separated and the supernatant was further digested for another 2 h at 56 °C and then decrosslinked for 4 h at 68 °C. DNA was isolated using AMPure XP beads. Samples were quantified in triplicates using real-time qPCR and enrichments were normalized to the ChIP input sample as well as a positive enrichment control CBS region. Primer sequences are available in Supplementary Table 7.

Statistics

Statistical analysis of inter-TAD interactions was performed using two-sample, two-tailed heteroscedastic *t* test. Statistical analysis of Xist RNA FISH experiments in male cell lines was performed using one sided Welch's *t* tests, making use of *t.test* from the stats R package. *P* values were adjusted for multiple testing by Benjamini-Hochberg correction, using *p.adjust* from the stats R package. Statistical differences between mutant and wildtype Xist RAP enrichment or X-linked gene silencing were determined by a two-sided Wilcoxon Rank Sum tests, using *wilcox.test* from the stats R package. Statistical analysis of Xist RNA FISH experiments in female cell lines was performed on independent experiments by two-sided Fisher's exact tests, making use of *fisher.test* from the stats R package. Bi-allelic Tsix expression measured by RNA-FISH in the female cell lines was tested for statistical significance on independent experiments by Pearson's Chi-squared tests with Yates'

continuity correction, using `chisq.test` from the stats R package. Statistical analysis of reduced and dispersed Xist RNA cloud formation in mutant female cell lines that expresses Tsix RNA from one or neither X chromosome versus both X chromosomes was performed on each independent experiment using two-sided Fisher's exact test, employing `fisher.test` from the stats R package. *P* values for differentially expressed genes measured by RNA sequencing were calculated with gene-wise exact tests for differences in the means between two groups of negative-binomially distributed counts, using `exactTest` from the R package R. *P* values were corrected for multiple testing by a Benjamini-Hochberg correction, using `p.adjust` from the stats R package.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgements

We would like to thank the Heard laboratory for their technical input and critical discussions; Daan Noordermeer for critical discussion and advice on Capture-C data analysis and interpretation; Amy Chow from the Guttman laboratory for RAP-DNA cell culture; members of the Bourc'his laboratory; Cécile Reyes and Audrey Rapinat from the Nanostring platform; Jelena M Telenius, Marieke Oudelaar and Damien Downes from the Hughes and Higgs laboratories. This work was supported by ERC Advanced Investigator award (ERC-2014-AdG no. 671027), Labelisation La Ligue, FRM (grant DEI20151234398), ANR DoseX 2017, Labex DEEP (ANR-11-LBX-0044), part of the IDEX Idex PSL (ANR-10-IDEX-0001-02 PSL), and ABS4NGS (ANR-11-BINF-0001) to E.H.; NWO-ALW Rubicon (825.13.002) and Veni (863.15.016) fellowships to J.G.vB.; Région Ile-de-France (DIM Biothérapies) and Fondation pour la Recherche Médicale (FDT20160435295) fellowships to R.G.; Sir Henry Wellcome Postdoctoral Fellowship (201369/Z/16/Z) to J.J.Z.; MRC Clinician Scientist Fellowship (ref. MR/R008108) to J.D.; Wellcome Trust Strategic Award (reference 106130/Z/14/Z) to J.R.H.; New York Stem Cell Foundation and California Institute of Technology funds to M.G.; M.G. is a New York Stem Cell Foundation – Robertson Investigator; Research Foundation and ERC Starting Grant (ERC-2017-STG no. 759366) to L.G.; LabEx and EquipEx (ANR-10-IDEX-0001-02 PSL, ANR-11-LBX-0044, and « INCa-DGOS- 4654 » SIRIC11-002) to the Nanostring platform of Institut Curie; Equipex (ANR-10-EQPX-03), France Génomique Consortium from the Agence Nationale de la Recherche ("Investissements d'Avenir" program) (ANR-10-INBS-09-08), and Canceropole Ile-de-France and by the SiRIC-Curie program - SiRIC Grant (INCa-DGOS- 4654) to the ICGex Next Generation Sequencing platform of the Institut Curie.

References

1. de Laat W, Duboule D. Topology of mammalian developmental enhancers and their regulatory landscapes. *Nature*. 2013; 502:499–506. [PubMed: 24153303]
2. Dixon JR, et al. Topological domains in mammalian genomes identified by analysis of chromatin interactions. *Nature*. 2012; 485:376–80. [PubMed: 22495300]
3. Nora EP, et al. Spatial partitioning of the regulatory landscape of the X-inactivation centre. *Nature*. 2012; 485:381–5. [PubMed: 22495304]
4. Phillips-Cremins JE, et al. Architectural protein subclasses shape 3D organization of genomes during lineage commitment. *Cell*. 2013; 153:1281–95. [PubMed: 23706625]
5. Rao SS, et al. A 3D map of the human genome at kilobase resolution reveals principles of chromatin looping. *Cell*. 2014; 159:1665–80. [PubMed: 25497547]
6. Zhan Y, et al. 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*. 2017; 27:479–490. [PubMed: 28057745]
7. Vietri Rudan M, et al. Comparative Hi-C reveals that CTCF underlies evolution of chromosomal domain architecture. *Cell Rep*. 2015; 10:1297–309. [PubMed: 25732821]
8. Le Dily F, et al. Distinct structural transitions of chromatin topological domains correlate with coordinated hormone-induced gene regulation. *Genes Dev*. 2014; 28:2151–62. [PubMed: 25274727]

9. Shen Y, et al. A map of the cis-regulatory sequences in the mouse genome. *Nature*. 2012; 488:116–20. [PubMed: 22763441]
10. Symmons O, et al. Functional and topological characteristics of mammalian regulatory domains. *Genome Res*. 2014; 24:390–400. [PubMed: 24398455]
11. Van Bortle K, et al. Insulator function and topological domain border strength scale with architectural protein occupancy. *Genome Biol*. 2014; 15:R82. [PubMed: 24981874]
12. Li Y, et al. Characterization of constitutive CTCF/cohesin loci: a possible role in establishing topological domains in mammalian genomes. *BMC Genomics*. 2013; 14:553. [PubMed: 23945083]
13. Sofueva S, et al. Cohesin-mediated interactions organize chromosomal domain architecture. *EMBO J*. 2013; 32:3119–29. [PubMed: 24185899]
14. Nora EP, et al. Targeted Degradation of CTCF Decouples Local Insulation of Chromosome Domains from Genomic Compartmentalization. *Cell*. 2017; 169:930–944 e22. [PubMed: 28525758]
15. Guo Y, et al. CRISPR Inversion of CTCF Sites Alters Genome Topology and Enhancer/Promoter Function. *Cell*. 2015; 162:900–10. [PubMed: 26276636]
16. Lupianez DG, et al. Disruptions of topological chromatin domains cause pathogenic rewiring of gene-enhancer interactions. *Cell*. 2015; 161:1012–1025. [PubMed: 25959774]
17. Sanborn AL, et al. Chromatin extrusion explains key features of loop and domain formation in wild-type and engineered genomes. *Proc Natl Acad Sci U S A*. 2015; 112:E6456–65. [PubMed: 26499245]
18. de Wit E, et al. CTCF Binding Polarity Determines Chromatin Looping. *Mol Cell*. 2015; 60:676–84. [PubMed: 26527277]
19. Narendra V, et al. CTCF establishes discrete functional chromatin domains at the Hox clusters during differentiation. *Science*. 2015; 347:1017–21. [PubMed: 25722416]
20. Tang Z, et al. CTCF-Mediated Human 3D Genome Architecture Reveals Chromatin Topology for Transcription. *Cell*. 2015; 163:1611–27. [PubMed: 26686651]
21. Lupianez DG, Spielmann M, Mundlos S. Breaking TADs: How Alterations of Chromatin Domains Result in Disease. *Trends Genet*. 2016; 32:225–37. [PubMed: 26862051]
22. Rastan S. Non-random X-chromosome inactivation in mouse X-autosome translocation embryos--location of the inactivation centre. *J Embryol Exp Morphol*. 1983; 78:1–22. [PubMed: 6198418]
23. Rastan S, Robertson EJ. X-chromosome deletions in embryo-derived (EK) cell lines associated with lack of X-chromosome inactivation. *J Embryol Exp Morphol*. 1985; 90:379–88. [PubMed: 3834036]
24. Heard E, Avner P, Rothstein R. Creation of a deletion series of mouse YACs covering a 500 kb region around Xist. *Nucleic Acids Res*. 1994; 22:1830–7. [PubMed: 8208607]
25. Lee JT, Strauss WM, Dausman JA, Jaenisch R. A 450 kb transgene displays properties of the mammalian X-inactivation center. *Cell*. 1996; 86:83–94. [PubMed: 8689690]
26. Galupa R, Heard E. X-chromosome inactivation: new insights into cis and trans regulation. *Curr Opin Genet Dev*. 2015; 31:57–66. [PubMed: 26004255]
27. Nesterova TB, et al. Pluripotency factor binding and Tsix expression act synergistically to repress Xist in undifferentiated embryonic stem cells. *Epigenetics Chromatin*. 2011; 4:17. [PubMed: 21982142]
28. Navarro P, et al. Molecular coupling of Xist regulation and pluripotency. *Science*. 2008; 321:1693–5. [PubMed: 18802003]
29. Sousa EJ, et al. Exit from Naive Pluripotency Induces a Transient X Chromosome Inactivation-like State in Males. *Cell Stem Cell*. 2018; 22:919–928 e6. [PubMed: 29804891]
30. Lee JT. Regulation of X-chromosome counting by Tsix and Xite sequences. *Science*. 2005; 309:768–71. [PubMed: 16051795]
31. Debrand E, Chureau C, Arnaud D, Avner P, Heard E. Functional analysis of the DXPas34 locus, a 3' regulator of Xist expression. *Mol Cell Biol*. 1999; 19:8513–25. [PubMed: 10567576]
32. Lee JT, Davidow LS, Warshawsky D. Tsix, a gene antisense to Xist at the X-inactivation centre. *Nat Genet*. 1999; 21:400–4. [PubMed: 10192391]

33. Mise N, Goto Y, Nakajima N, Takagi N. Molecular cloning of antisense transcripts of the mouse Xist gene. *Biochem Biophys Res Commun.* 1999; 258:537–41. [PubMed: 10329421]
34. Tian D, Sun S, Lee JT. The long noncoding RNA, Jpx, is a molecular switch for X chromosome inactivation. *Cell.* 2010; 143:390–403. [PubMed: 21029862]
35. Barakat TS, et al. RNF12 activates Xist and is essential for X chromosome inactivation. *PLoS Genet.* 2011; 7:e1002001. [PubMed: 21298085]
36. Furlan G, et al. The Ftx Noncoding Locus Controls X Chromosome Inactivation Independently of Its RNA Products. *Mol Cell.* 2018; 70:462–472 e8. [PubMed: 29706539]
37. Augui S, Nora EP, Heard E. Regulation of X-chromosome inactivation by the X-inactivation centre. *Nat Rev Genet.* 2011; 12:429–42. [PubMed: 21587299]
38. Pollex T, Heard E. Recent advances in X-chromosome inactivation research. *Curr Opin Cell Biol.* 2012; 24:825–32. [PubMed: 23142477]
39. van Bommel JG, Mira-Bontenbal H, Gribnau J. Cis- and trans-regulation in X inactivation. *Chromosoma.* 2016; 125:41–50. [PubMed: 26198462]
40. Hughes JR, et al. Analysis of hundreds of cis-regulatory landscapes at high resolution in a single, high-throughput experiment. *Nat Genet.* 2014; 46:205–12. [PubMed: 24413732]
41. Davies JO, et al. Multiplexed analysis of chromosome conformation at vastly improved sensitivity. *Nat Methods.* 2016; 13:74–80. [PubMed: 26595209]
42. Dostie J, et al. Chromosome Conformation Capture Carbon Copy (5C): a massively parallel solution for mapping interactions between genomic elements. *Genome Res.* 2006; 16:1299–309. [PubMed: 16954542]
43. Merkenschlager M, Nora EP. CTCF and Cohesin in Genome Folding and Transcriptional Gene Regulation. *Annu Rev Genomics Hum Genet.* 2016; 17:17–43. [PubMed: 27089971]
44. Ogawa Y, Lee JT. Xite, X-inactivation intergenic transcription elements that regulate the probability of choice. *Mol Cell.* 2003; 11:731–43. [PubMed: 12667455]
45. Stavropoulos N, Rowntree RK, Lee JT. Identification of developmentally specific enhancers for Tsix in the regulation of X chromosome inactivation. *Mol Cell Biol.* 2005; 25:2757–69. [PubMed: 15767680]
46. Barakat TS, et al. The trans-activator RNF12 and cis-acting elements effectuate X chromosome inactivation independent of X-pairing. *Mol Cell.* 2014; 53:965–78. [PubMed: 24613346]
47. Sun S, et al. Jpx RNA activates Xist by evicting CTCF. *Cell.* 2013; 153:1537–51. [PubMed: 23791181]
48. Carmona S, Lin B, Chou T, Arroyo K, Sun S. LncRNA Jpx induces Xist expression in mice using both trans and cis mechanisms. *PLoS Genet.* 2018; 14:e1007378. [PubMed: 29734339]
49. Giorgetti L, et al. Predictive polymer modeling reveals coupled fluctuations in chromosome conformation and transcription. *Cell.* 2014; 157:950–63. [PubMed: 24813616]
50. Spencer RJ, et al. A boundary element between Tsix and Xist binds the chromatin insulator Ctcf and contributes to initiation of X-chromosome inactivation. *Genetics.* 2011; 189:441–54. [PubMed: 21840866]
51. Jegu T, Aeby E, Lee JT. The X chromosome in space. *Nat Rev Genet.* 2017; 18:377–389. [PubMed: 28479596]
52. Engreitz JM, et al. The Xist lncRNA exploits three-dimensional genome architecture to spread across the X chromosome. *Science.* 2013; 341:1237973. [PubMed: 23828888]
53. Brockdorff N, Turner BM. Dosage compensation in mammals. *Cold Spring Harb Perspect Biol.* 2015; 7:a019406. [PubMed: 25731764]
54. Brons IG, et al. Derivation of pluripotent epiblast stem cells from mammalian embryos. *Nature.* 2007; 448:191–5. [PubMed: 17597762]
55. Guo G, et al. Klf4 reverts developmentally programmed restriction of ground state pluripotency. *Development.* 2009; 136:1063–9. [PubMed: 19224983]
56. Franke M, et al. Formation of new chromatin domains determines pathogenicity of genomic duplications. *Nature.* 2016; 538:265–269. [PubMed: 27706140]

57. Rodriguez-Carballo E, et al. The HoxD cluster is a dynamic and resilient TAD boundary controlling the segregation of antagonistic regulatory landscapes. *Genes Dev.* 2017; 31:2264–2281. [PubMed: 29273679]
58. Johnston CM, Newall AE, Brockdorff N, Nesterova TB. Enox, a novel gene that maps 10 kb upstream of Xist and partially escapes X inactivation. *Genomics.* 2002; 80:236–44. [PubMed: 12160738]
59. Chureau C, et al. Ftx is a non-coding RNA which affects Xist expression and chromatin structure within the X-inactivation center region. *Hum Mol Genet.* 2011; 20:705–18. [PubMed: 21118898]
60. Hofmann A, Heermann DW. The role of loops on the order of eukaryotes and prokaryotes. *FEBS Lett.* 2015; 589:2958–65. [PubMed: 25912650]
61. Kent WJ, et al. The human genome browser at UCSC. *Genome Res.* 2002; 12:996–1006. [PubMed: 12045153]
62. Crane E, et al. Condensin-driven remodelling of X chromosome topology during dosage compensation. *Nature.* 2015; 523:240–4. [PubMed: 26030525]
63. Pillet N, Bonny C, Schorderet DF. Characterization of the promoter region of the mouse Xist gene. *Proc Natl Acad Sci U S A.* 1995; 92:12515–9. [PubMed: 8618932]
64. Gontan C, et al. RNF12 initiates X-chromosome inactivation by targeting REX1 for degradation. *Nature.* 2012; 485:386–90. [PubMed: 22596162]
65. Doetschman T, et al. Targetted correction of a mutant HPRT gene in mouse embryonic stem cells. *Nature.* 1987; 330:576–8. [PubMed: 3683574]
66. Norris DP, et al. Evidence that random and imprinted Xist expression is controlled by preemptive methylation. *Cell.* 1994; 77:41–51. [PubMed: 8156596]
67. Masui O, et al. Live-cell chromosome dynamics and outcome of X chromosome pairing events during ES cell differentiation. *Cell.* 2011; 145:447–58. [PubMed: 21529716]
68. Cong L, et al. Multiplex genome engineering using CRISPR/Cas systems. *Science.* 2013; 339:819–23. [PubMed: 23287718]
69. Ran FA, et al. Genome engineering using the CRISPR-Cas9 system. *Nat Protoc.* 2013; 8:2281–308. [PubMed: 24157548]
70. Langmead B, Salzberg SL, et al. Fast gapped-read alignment with Bowtie 2. *Nat Methods.* 2012; 9:357–9. [PubMed: 22388286]
71. Servant N, et al. HiTC: exploration of high-throughput 'C' experiments. *Bioinformatics.* 2012; 28:2843–4. [PubMed: 22923296]
72. Hnisz D, et al. Activation of proto-oncogenes by disruption of chromosome neighborhoods. *Science.* 2016; 351:1454–1458. [PubMed: 26940867]
73. Smith EM, Lajoie BR, Jain G, Dekker J. Invariant TAD Boundaries Constrain Cell-Type-Specific Looping Interactions between Promoters and Distal Elements around the CFTR Locus. *Am J Hum Genet.* 2016; 98:185–201. [PubMed: 26748519]
74. Servant N, et al. HiC-Pro: an optimized and flexible pipeline for Hi-C data processing. *Genome Biol.* 2015; 16:259. [PubMed: 26619908]
75. Geiss GK, et al. Direct multiplexed measurement of gene expression with color-coded probe pairs. *Nat Biotechnol.* 2008; 26:317–25. [PubMed: 18278033]
76. Dobin A, et al. STAR: ultrafast universal RNA-seq aligner. *Bioinformatics.* 2013; 29:15–21. [PubMed: 23104886]
77. Mudge JM, Harrow J. Creating reference gene annotation for the mouse C57BL6/J genome assembly. *Mamm Genome.* 2015; 26:366–78. [PubMed: 26187010]
78. Robinson MD, McCarthy DJ, Smyth GK. edgeR: a Bioconductor package for differential expression analysis of digital gene expression data. *Bioinformatics.* 2010; 26:139–40. [PubMed: 19910308]
79. McCarthy DJ, Chen Y, Smyth GK. Differential expression analysis of multifactor RNA-Seq experiments with respect to biological variation. *Nucleic Acids Res.* 2012; 40:4288–97. [PubMed: 22287627]
80. Borensztein M, et al. Xist-dependent imprinted X inactivation and the early developmental consequences of its failure. *Nat Struct Mol Biol.* 2017; 24:226–233. [PubMed: 28134930]

81. Chaumeil J, Augui S, Chow JC, Heard E. 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. *Methods Mol Biol.* 2008; 463:297–308. [PubMed: 18951174]
82. Engreitz JM, et al. RNA-RNA interactions enable specific targeting of noncoding RNAs to nascent Pre-mRNAs and chromatin sites. *Cell.* 2014; 159:188–199. [PubMed: 25259926]
83. Chen CK, et al. Xist recruits the X chromosome to the nuclear lamina to enable chromosome-wide silencing. *Science.* 2016; 354:468–472. [PubMed: 27492478]
84. Zeng PY, Vakoc CR, Chen ZC, Blobel GA, Berger SL. In vivo dual cross-linking for identification of indirect DNA-associated proteins by chromatin immunoprecipitation. *Biotechniques.* 2006; 41:694. [PubMed: 17191611]

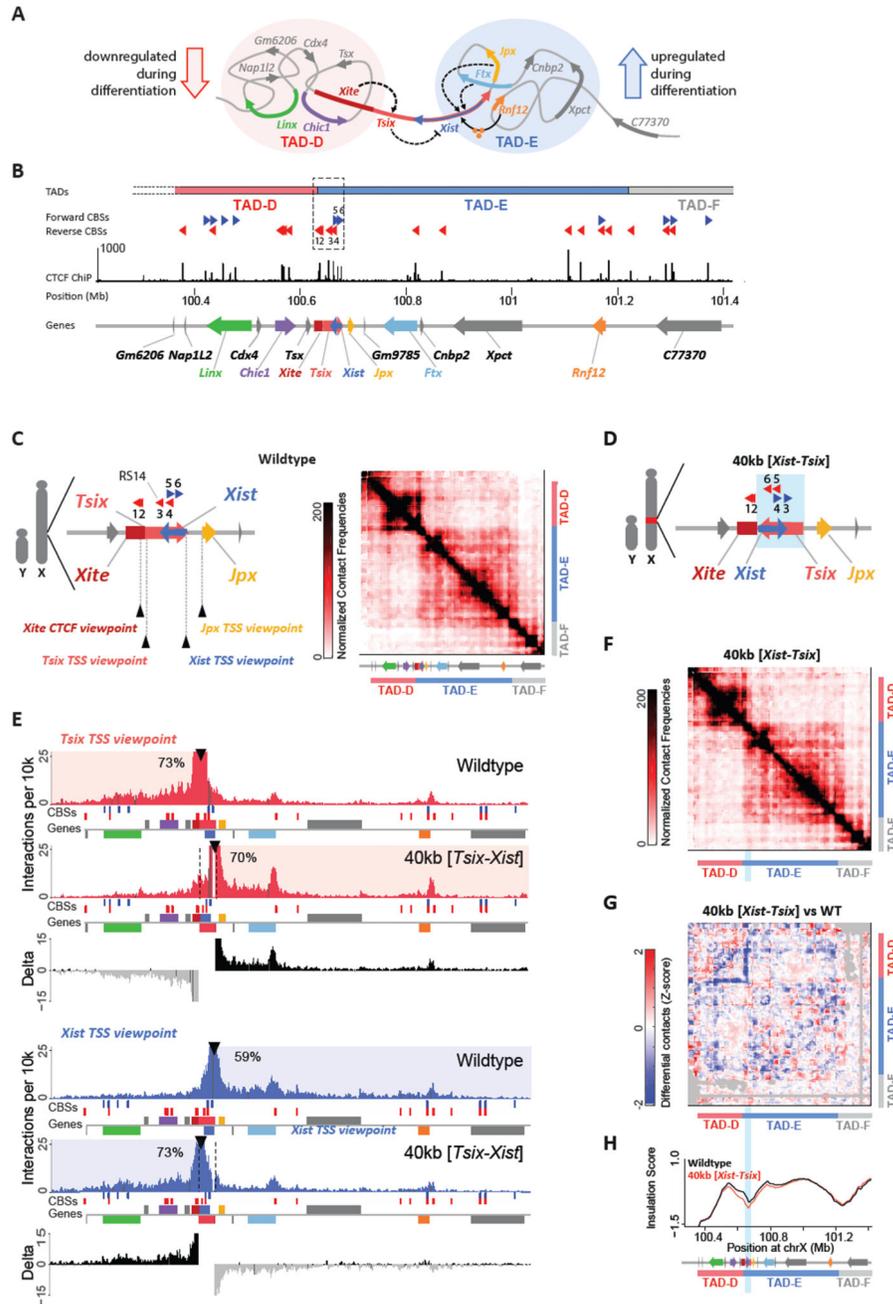


Fig. 1. Genomic inversion of the *Xist*/*Tsix* loci switches their promoters into each other's original TADs.

(a) Schematic illustration of the *Xic* organized into two TADs. Red (TAD-D) and blue circles (TAD-E) represent the interaction environments of *Tsix* and *Xist*. Dashed arrows indicate *cis*-activation. Black arrow from *Rnf12* indicates *trans* activation by the protein RNF12 (orange balls). Dashed line from *Tsix* to *Xist* indicates anti-sense repression. (b) Linear visualization of the *Xic* organized into two TADs. Dashed box represents the previously described 68-kb boundary deletion 3. TADs determined using the insulation

score, are depicted as colored bars (red, blue, grey). Dotted lines at start of TAD-D indicate undefined TAD structure due to repetitive sequences. CTCF binding sites (CBSs) in forward (blue) and reverse (red) orientation, and CTCF ChIP-seq signal in E14 mESCs 14. Gene annotation from UCSC RefSeq mm9 61, except for *Xite* (see Methods) and *Linx 3*. (c) Schematic representation of boundary region and its CBSs in male mESCs. Probes for capture enrichment (viewpoint) depicted with black arrowheads. On the right, 5C chromosome conformation contact frequencies in wild type male mESCs. Map represents normalized and binned pool of 2 samples. (d) Schematic representation of the 40-kb [*Tsix-Xist*] inversion in male mESCs. Blue box indicates inverted region. (e) Capture-C profiles for *Tsix* (red) and *Xist* (blue) viewpoints, in wild type and 40-kb [*Tsix-Xist*] inversion. Viewpoints depicted with black arrowheads. Underneath, forward CBSs in blue, reverse CBSs in red, and genes as colored boxes as in A, forward-oriented genes above and reverse-oriented genes below the grey line. Differential interaction frequencies of wild type minus inversion interaction frequencies in black (interaction gain) and grey (interaction loss). Percentage of normalized interactions on either side of the viewpoints is indicated. (f) 5C chromosome conformation contact frequencies in male mESCs harboring a 40-kb [*Tsix-Xist*] inversion; map represents inverted genome. Pool of two normalized and binned replicates, as in 1C. (g) Differential map represents subtraction of wild type from inversion Z-scores. Gray pixels correspond to filtered interactions that did not meet the quality control threshold. (h) Insulation scores, wild type in black and 40-kb [*Tsix-Xist*] inversion in red. Inverted region represented by light blue box. Chromosome positions are aligned with panels F and G. **Note:** Each genomic alteration has been generated twice (two independent cell lines). Results for the first clone are shown here, results for the second clone are shown in Supplementary Figures 2-3.

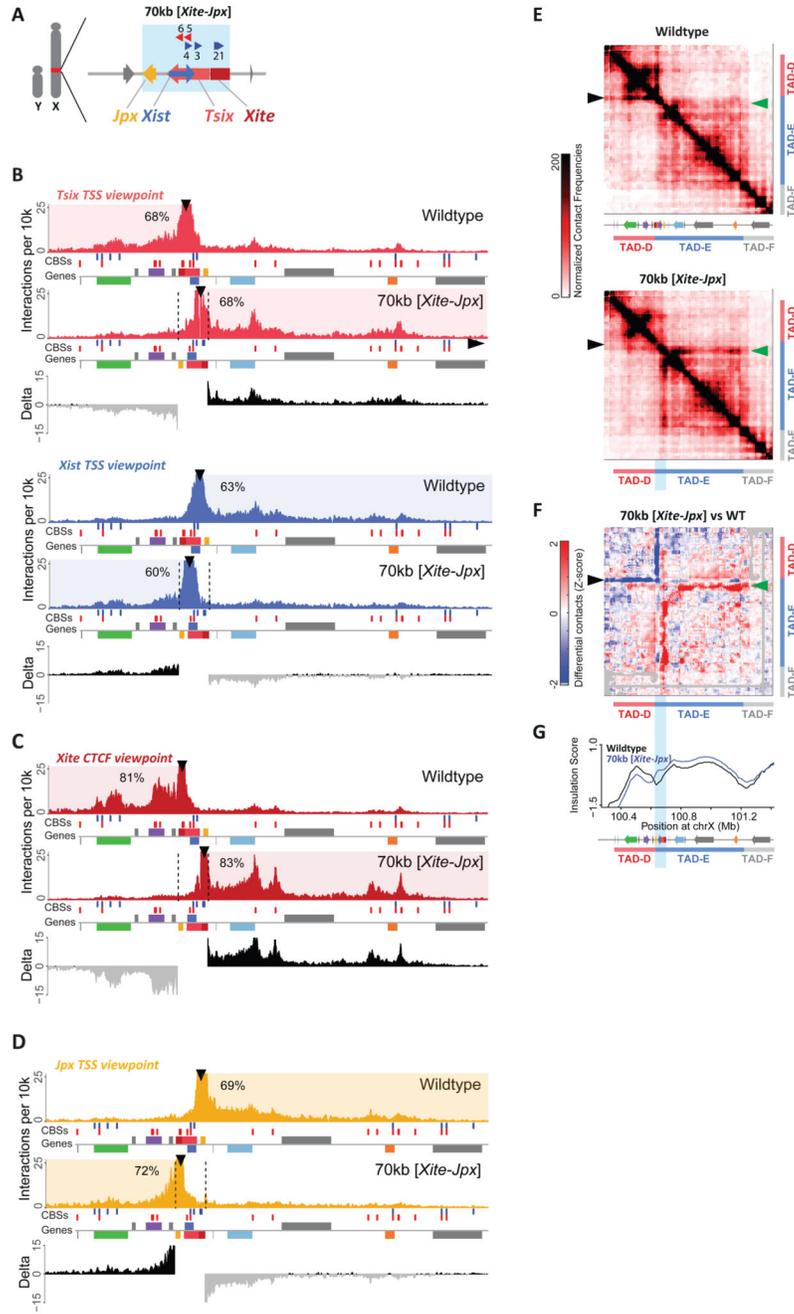


Fig. 2. Genomic inversion of the *Xist/Tsix* loci along with *Xite* and *Jpx* leads to topological changes within the *Xic*.
(a) Schematic representation of the 70-kb [*Xite-Jpx*] inversion in male mESCs. Blue box indicates inverted region. **(b)** Capture-C profiles and differential interaction frequencies, as in 1E, for *Tsix* (red) and *Xist* (blue) viewpoints in wild type and 70-kb [*Xite-Jpx*] inversion. **(c)** Capture-C profiles and differential interaction frequencies, as in 1E, for *Xite* (dark red) viewpoint in wild type and 70-kb [*Xite-Jpx*] inversion. **(d)** Capture-C profiles and differential interaction frequencies, as in 1E, for *Jpx* (yellow) viewpoint in wild type and 70-

kb [*Xite-Jpx*] inversion. **(e)** 5C chromosome conformation contact frequencies in wildtype male mESCs (top) or harboring a 70-kb [*Xite-Jpx*] inversion (bottom). Pool of two replicates each, data for inversion have been inverted accordingly (i.e. map represents the inverted genome), normalized and binned as in 1C. Arrowheads indicate interactions described in the text. **(f)** Differential map represents the subtraction of wild type Z-scores from inversion Z-scores (see Methods). Gray pixels correspond to interactions that were filtered because they did not meet the quality control threshold (see Methods). Arrowheads indicate interactions described in the text. **(g)** Insulation scores, wild type in black and 70-kb [*Xite-Jpx*] inversion in blue. Inverted region represented by light blue box. Chromosome positions are aligned with panels E and F. **Note:** Each genomic alteration has been generated twice (two independent cell lines). Results for the first clone are shown here, results for the second clone are shown in Supplementary Figures 2-3.

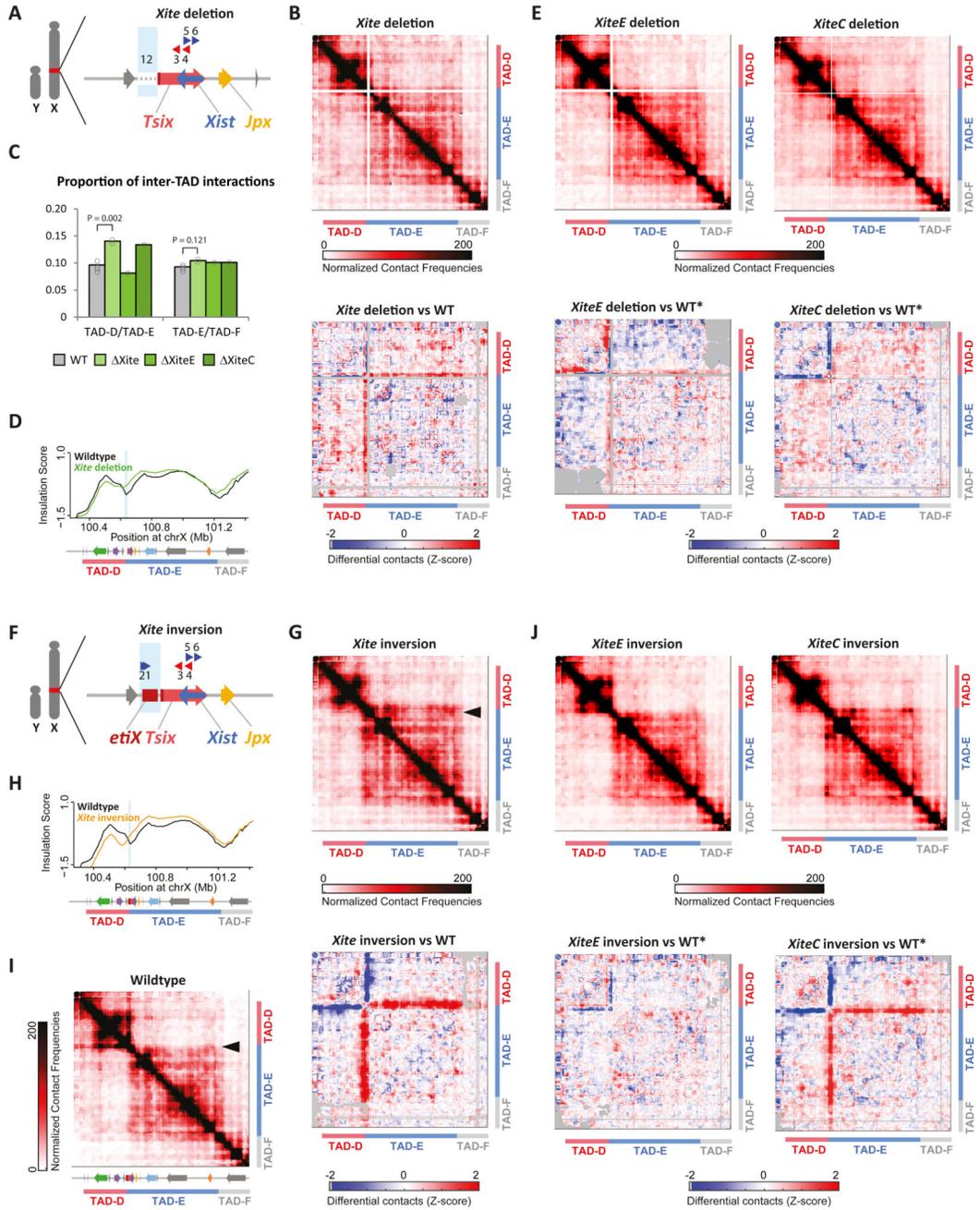


Fig. 3. *Xite* structural element is important for TAD boundary position and insulation. (a) Schematic representation of the *Xite* deletion (*Xite*) in male mESCs. Blue box indicates deleted region. (b) (e) (g) (j) Top: 5C chromosome conformation contact frequencies in *Xite* (b), *XiteE* (e, left), *XiteC* (e, right), *Xite* inversion (g), *XiteE* inversion (j, left) or *XiteC* inversion (j, right) male mESCs. Pools of two replicates normalized and binned as in 1C. Bottom: Differential 5C map of mutant versus wild type mESCs. Differential map represents the subtraction of wild type Z-scores from mutant Z-scores (see Methods). Gray pixels correspond to interactions that were filtered because they did not meet the quality

control threshold (see Methods). 5C results of *Xite* and *Xite* inversion for the first clone shown here, second clone shown in Supplementary Figure 3. For other mutants, maps represent pooled results from two independent clones. **(c)** Quantification of the increased inter-TAD interactions in *Xite* mutants. Center values represent the average of the calculated proportions: for *XiteE* and *XiteC*, two measurements each from two independent clones; for WT and *Xite*, four measurements each from two independent clones in duplicates (see Supplementary Figure 3F for details). Statistical analysis was performed using two-sample, two-tailed heteroscedastic *t* test. **(d)** Insulation scores; wild type in black and *Xite* in green. Deleted region represented by light blue box. Insulation score for one clone shown here, second clone can be found in Supplementary Figure 3E. **(f)** Schematic representation of the *Xite* inversion in male mESCs. Blue box indicates inverted region. **(h)** Insulation scores, wild type in black and *Xite* inversion in orange. Inverted region represented by light blue box. Insulation score for one clone shown here, second clone can be found in Supplementary Figure 3E. **(i)** 5C chromosome conformation contact frequencies in wildtype male mESCs, as in Figure 1C, included for clarity purposes. Arrowhead indicates interactions at the proximal end of TAD-E.

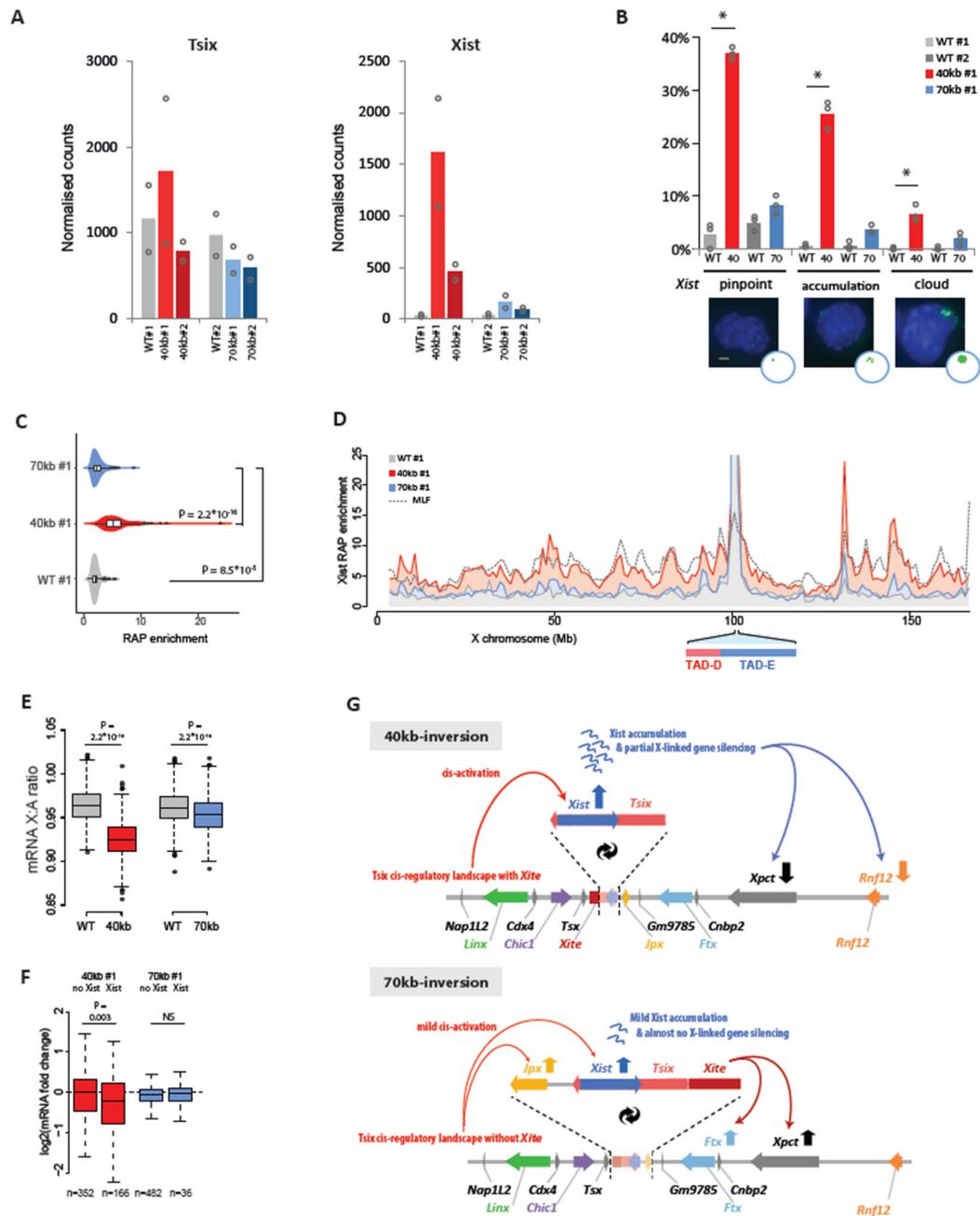


Fig. 4. *Xist/Tsix* inversions lead to ectopic *Xist* expression, *Xist* RNA coating and X-linked gene silencing in male mESCs.

(a) nCounter[®] mRNA expression levels in 40-kb inversion (red), 70-kb inversion (blue) and wild type (gray). Each inversion has its corresponding wild type comparison because they were processed in different batches. Bars depict the mean of two independent experiments, with dots depicting the independent experiments. (b) *Xist* accumulation analyzed by RNA-FISH in male wild type (gray), 40-kb inversion (red) and 70-kb inversion (blue) mESCs. Each inversion has its corresponding wild type comparison because they were processed in

different batches. Bars represent mean percentage of counted cells with either a *Xist* pinpoint, *Xist* accumulation or full *Xist* cloud. Dots represent independent experiments. $n = 3$ independent experiments, each experiment counting 100 to 400 cells (Supplementary Table 6 for exact sample size details). Statistical analysis was performed on independent experiments, using two-sided Fisher's exact test, mutant versus wild type. * $P < 0.05$ in all three experiments, see Supplementary Table 8 for exact P values. Scale bar: 2 μm , images at same scale. (c) Violin plots indicating *Xist* RAP enrichment at the X chromosome. Substantial and significant enrichment in the 40-kb inversion (red), and slight but significant enrichment in the 70-kb inversion (blue) compared to wild type (gray). X-axis depicts mean RAP signal normalized to input. $n = 2$ independent experiments). Box corresponds to 25th and 75th percentiles, black line to the median, whiskers to 1.5 \times the interquartile range (IQR), and dots to individual datapoints beyond the 1.5 \times IQR. Statistical analysis was performed using two-sided Wilcoxon Rank Sum test. (d) *Xist* RAP signal (normalized to input, mean of two replicates) along the entire X chromosome in wild type (gray), 40-kb [*Tsix-Xist*] inversion (red) and 70-kb [*Xite-Jpx*] inversion (blue) male mESCs, in comparison to *Xist* RAP signal in female mouse lymphatic fibroblasts (MLF) (dotted line). (e) X-linked gene silencing represented by the mRNA expression ratio between X-linked genes and bootstrapped autosomal genes ($n = 1000$). Boxplots represent the median (black line), 25-75% (box), 1.5 \times interquartile range (whiskers), and outliers (dots) of the bootstrap ratios. Statistical analysis was performed using two-sided Wilcoxon Rank Sum test. (f) X-linked gene silencing specific for *Xist*-bound genes, as shown by the difference in \log_2 fold-change in mRNA expression levels between inversion and wild type cells, for genes not bound by *Xist* RNA (labeled as no *Xist*) versus genes that are bound by *Xist* RNA (labeled as *Xist*). Boxplots represent the median (black line), 25-75% (box) and 1.5 \times interquartile range (whiskers). Genes were classified as bound by *Xist* when RAP signal at the TSS > 7.5 , $n =$ number of genes in each group. Statistical analysis was performed using two-sided Wilcoxon Rank Sum test. (g) Schematic illustration of the effects of inverting *Tsix-Xist* (40 kb) and *Xite-Jpx* (70 kb) on transcriptional activity in male mESCs. **Note:** Each inversion has been generated in two independent cell lines. Results for the second clones are shown in Supplementary Figure 6.

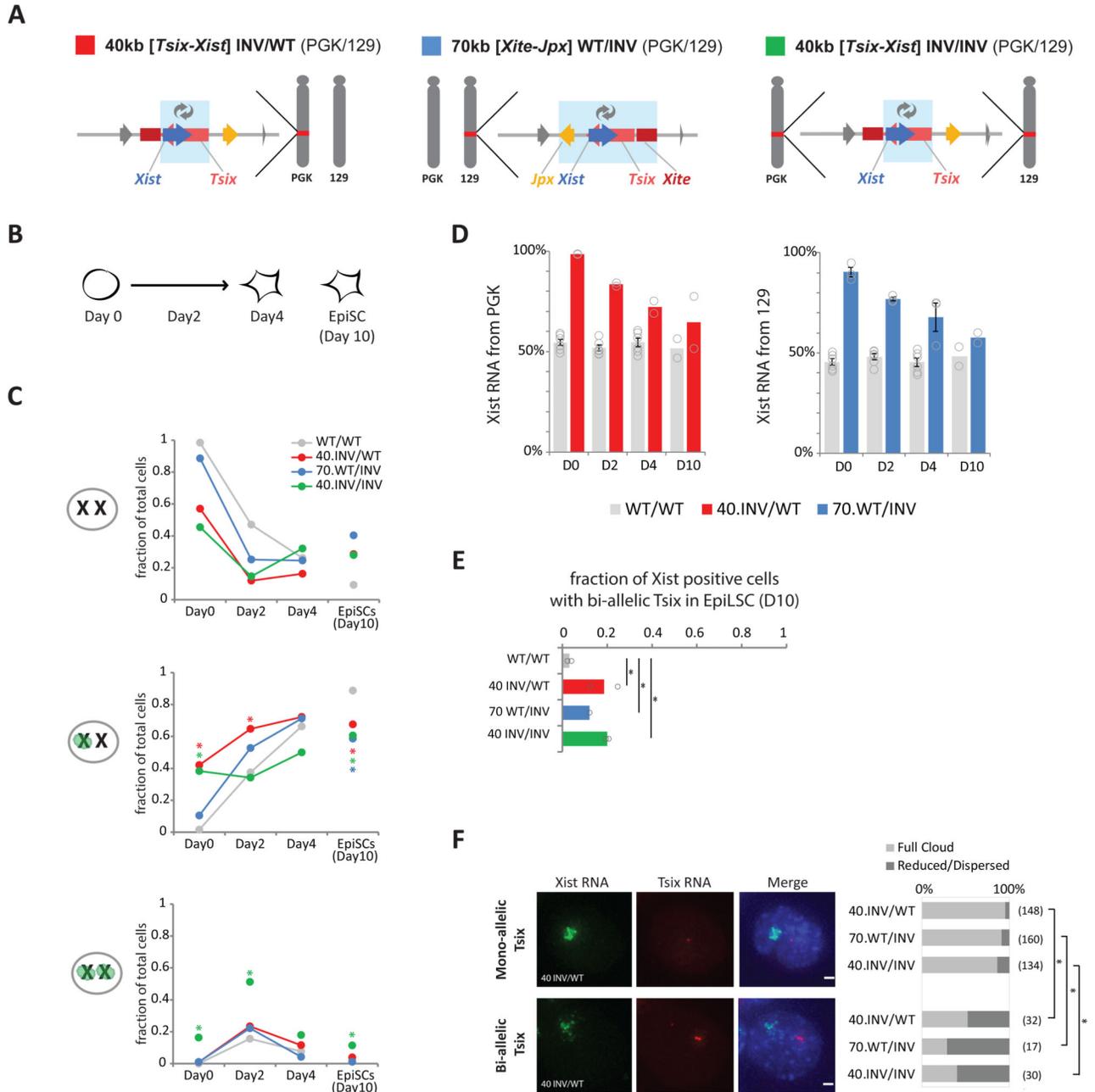


Fig. 5. *Xist/Tsix* inversions affect the initiation of XCI and the expression dynamics of *Xist* and *Tsix* during differentiation of female mESCs.

(a) Schematic representation of female hybrid mESC lines harboring a heterozygous 40-kb [*Tsix-Xist*] inversion on the PGK chromosome (red), a 70-kb [*Xite-Jpx*] inversion on the 129 chromosome (blue) and a homozygous 40-kb [*Tsix-Xist*] inversion (green). Light blue boxes indicate inverted region. (b) Schematic representation of mESC to EpiLSC differentiation and time points analyzed. (c) *Xist* accumulation analyzed by RNA-FISH in the female cell lines described in A. Bars represent mean fraction of counted cells with an

Xist accumulation or cloud at none, one and both X chromosomes. For each cell line, at least two independent experiments were performed, each counting >100 cells (Supplementary Table 6 for number of independent experiments per cell line and exact sample size details). Statistical analysis was performed on independent experiments, using two-sided Fisher's exact test, mutant versus wild type per time point. * $P < 0.05$ in all experiments, see Supplementary Table 8 for exact P values. **(d)** Allelic *Xist* expression, measured by pyrosequencing of *Xist* cDNA. Bars represent the percentage of *Xist* mRNA expressed from the PGK (left) or 129 (right) allele in cells described in A. Bars depict the mean of two or more independent experiments, with dots depicting each experiment, and error bars representing SEM in case of $n > 3$ (number of independent experiments per line as in C). **(e)** Bi-allelic *Tsix* expression measured by RNA FISH in the female cell lines described in A. Bars represent mean fraction of *Xist* positive cells with bi-allelic *Tsix* expression. For each cell line, two independent experiments were performed, each counting >100 cells (Supplementary Table 6 for exact sample size details). Dots represent the independent experiments. Statistical analysis was performed on independent experiments, using Pearson's Chi-squared test with Yates' continuity correction: * $P < 0.01$ in both experiments. See Supplementary Table 8 for exact P values. **(f)** Quantification of *Xist* RNA cloud formation in the female cell lines described in A, at day 10 of EpiLSC induction. Typical example of a proper *Xist* RNA cloud in a 40 INV/WT cell that expresses *Tsix* RNA from one X chromosome (upper panels). Typical example of a reduced and dispersed *Xist* RNA cloud formation in a 40 INV/WT cell that expresses *Tsix* from both X chromosomes (lower panels). Bar plot shows mean percentages of cells with full *Xist* RNA cloud formation (light gray) and reduced/dispersed *Xist* RNA cloud formation (dark gray) in *Xist*-positive cells expressing *Tsix* RNA from one or neither of the X chromosomes (top bars) or cells expressing *Tsix* RNA from both X chromosomes (bottom bars). $n =$ two independent experiments, total number of cells in the two experiments per group noted between brackets. Statistical analysis was performed on each independent experiment using two-sided Fisher's exact test: * $P < 0.01$ in both experiments. See Supplementary Table 8 for exact P values.