nature genetics

SUPPLEMENTARY INFORMATION

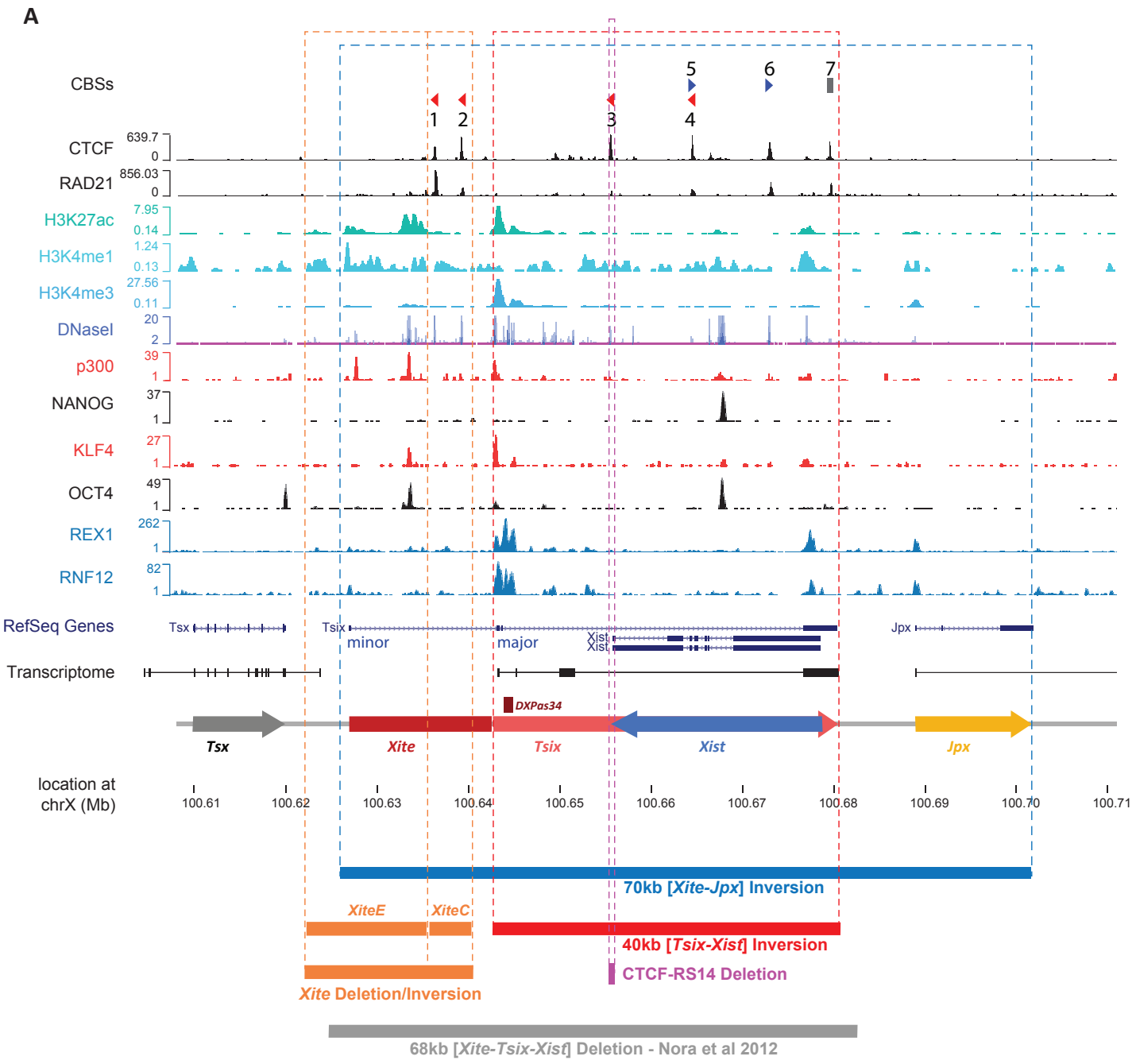
https://doi.org/10.1038/s41588-019-0412-0

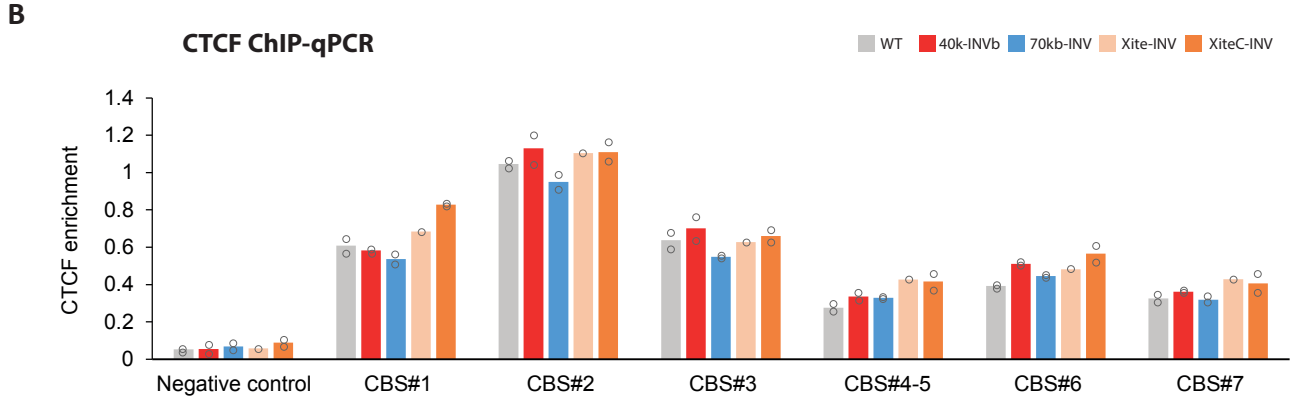
In the format provided by the authors and unedited.

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

Joke G. van Bemmel[1,2,14,16*], Rafael Galupa[1,15,16], Chris Gard[1], Nicolas Servant[3,4], Christel Picard[1], James Davies[5], Anthony James Szempruch[6], Yinxiu Zhan[7,8], Jan J. Żylicz[1,9], Elphège P. Nora[10], Sonia Lameiras[11], Elzo de Wit[12], David Gentien[13], Sylvain Baulande[11], Luca Giorgetti[7], Mitchell Guttman[6], Jim R. Hughes[5], Douglas R. Higgs[5], Joost Gribnau[2] and Edith Heard[1*]

[1]Institut Curie, CNRS UMR3215, INSERM U934, Paris, France. [2]Department of Developmental Biology, Erasmus Medical Center Rotterdam (EMC), Rotterdam, the Netherlands. [3]Institut Curie, PSL Research University, INSERM, U900, Paris, France. [4]MINES ParisTech, PSL Research University, CBIO-Centre for Computational Biology, Paris, France. [5]Medical Research Council, Molecular Haematology Unit, Weatherall Institute of Molecular Medicine, University of Oxford, John Radcliffe Hospital, Oxford, UK. [6]Division of Biology and Biological Engineering, California Institute of Technology, Pasadena, CA, USA. [7]Friedrich Miescher Institute for Biomedical Research, Basel, Switzerland. [8]University of Basel, Basel, Switzerland. [9]University of Cambridge, Department of Physiology, Development and Neuroscience, Cambridge, UK. [10]Gladstone Institute of Cardiovascular Disease and Roddenberry Center for Stem Cell Biology and Medicine at Gladstone, San Francisco, CA, USA. [11]Institut Curie Genomics of Excellence (ICGex) Platform, Institut Curie Research Center, Paris, France. [12]Oncode Institute and Division of Gene Regulation, the Netherlands Cancer Institute, Amsterdam, the Netherlands. [13]Institut Curie, PSL Research University, Translational Research Department, Genomics Platform, Paris, France. [14]Present address: Gladstone Institute of Cardiovascular Diseases, San Francisco, CA, USA. [15]Present address: European Molecular Biology Laboratory, Heidelberg, Germany. [16]These authors contributed equally: Joke G. van Bemmel, Rafael Galupa. *e-mail: jokevbemmel@gmail.com; edith.heard@curie.fr

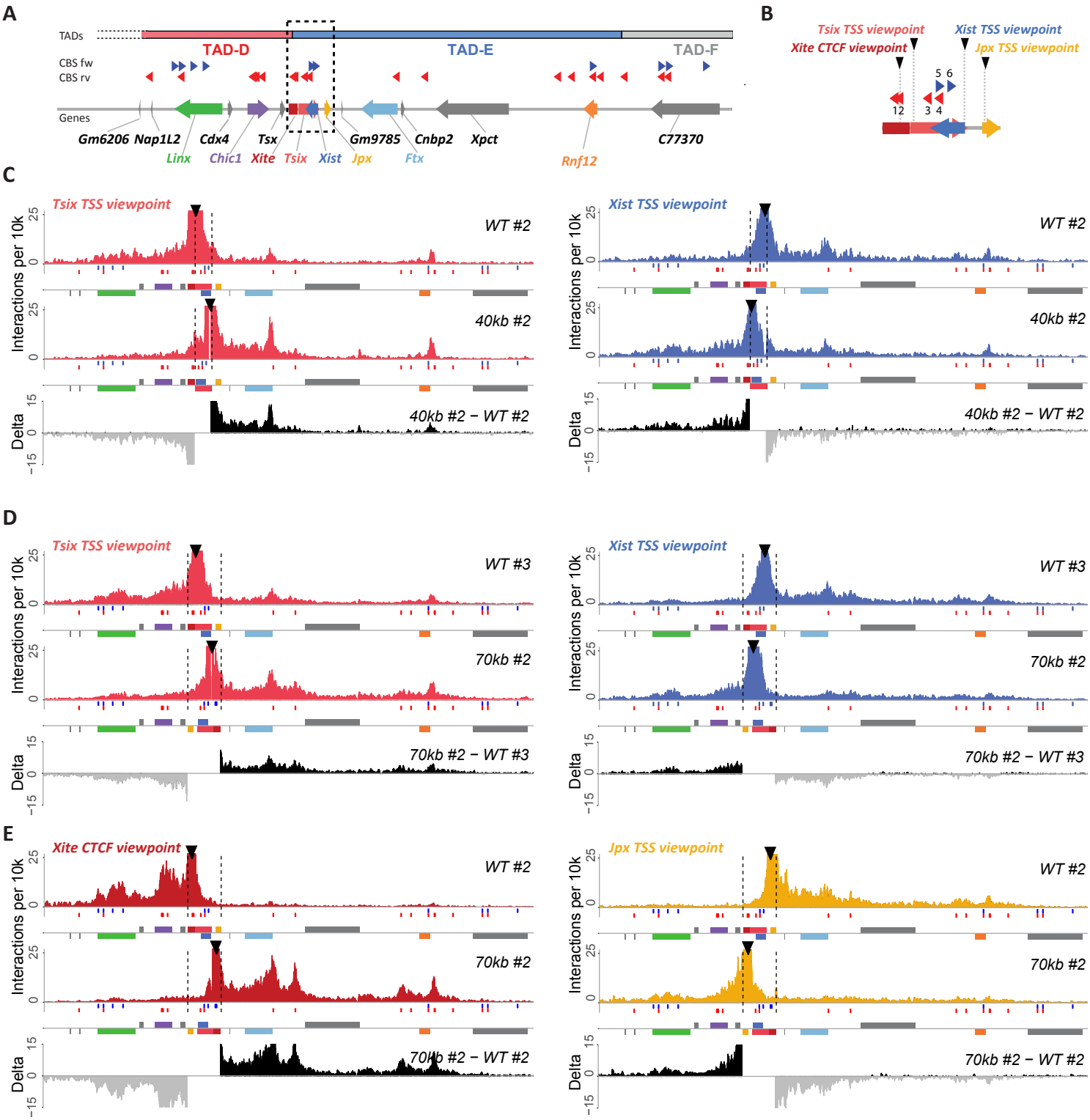

A
CBSs
1
2
3
4
5
6
7
CTCF
639.7
0
RAD21
856.03
0
H3K27ac
7.95
0.14
H3K4me1
1.24
0.13
H3K4me3
27.56
0.11
DNaseI
20
2
p300
39
1
NANOG
37
1
KLF4
27
1
OCT4
49
1
REX1
262
1
RNF12
82
1
RefSeq Genes
Tsx
Tsix
minor
major
Xist
Xist
Jpx
Transcriptome
DXPas34
Tsx
Xite
Tsix
Xist
Jpx
location at chrX (Mb)
100.61
100.62
100.63
100.64
100.65
100.66
100.67
100.68
100.69
100.70
100.71
70kb [Xite-Jpx] Inversion
XiteE
XiteC
40kb [Tsix-Xist] Inversion
CTCF-RS14 Deletion
Xite Deletion/Inversion
68kb [Xite-Tsix-Xist] Deletion - Nora et al 2012


B
CTCF ChIP-qPCR
WT
40k-INVb
70kb-INV
Xite-INV
XiteC-INV
CTCF enrichment
1.4
1.2
1
0.8
0.6
0.4
0.2
0
Negative control
CBS#1
CBS#2
CBS#3
CBS#4-5
CBS#6
CBS#7

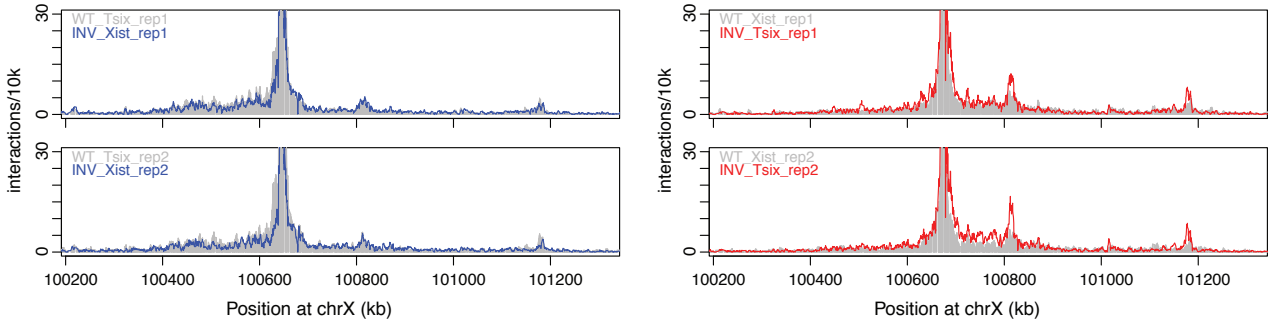
Supplementary Fig. 1:

A) Overview of chromatin features at the inverted or deleted regions in male mESCs: CRISPR-mediated mutations and inversions depicted with dashed boxes, 40kb [*Tsix-Xist*] inversion (red), 70kb [*Xite-Jpx*] inversion (blue), *Xite*-related deletions or inversions (orange) and CTCF deletion (purple). CTCF and RAD21 ChiP-seq in E14 mESCs (Nora, Goloborodko et al. 2017), H3K27ac, H3K4me1 and H3K4me3 ChiP-seq in E14 mESCs (Yue, Cheng et al. 2014)'(Consortium 2012), and DNaseI hypersensitivity in CJ7 mESCs (Yue, Cheng et al. 2014)'(Consortium 2012), p300 ChIP-seq in 129JAE-C57/B6 mESCs (Creyghton, Cheng et al. 2010), NANOG and OCT4 in V6.5 (C57BL/6-129) mESCs (Marson, Levine et al. 2008), KLF4 in E14 mESCs (Chen, Xu et al. 2008), and REX1 and RNF12 in F1 2-1 (129/Sv-Cast/Ei) mESCs (Gontan, Achame et al. 2012). Underneath, UCSC RefSeq gene annotation (Kent, Sugnet et al. 2002) and RNA-seq based transcriptome in V6.5 mESCs (Guttman, Garber et al. 2010), the annotations used in this paper (including Xite, see methods) and the genomic location (Mb) on the X chromosome in mm9. Tsix minor and major annotation from (Stavropoulos, Rowntree et al. 2005).

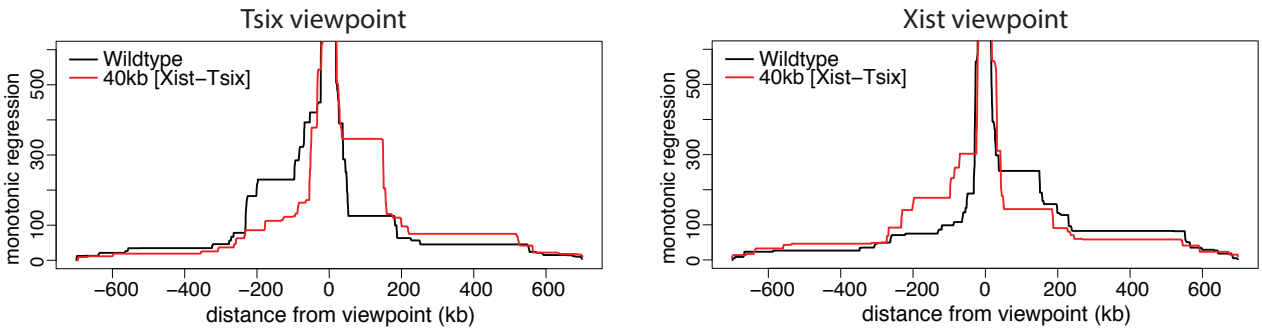
B) CTCF ChIP-qPCR results show no differences in CTCF binding within the inverted region between WT and INV cells. CTCF enrichment is relative to input and normalized to positive control, a CBS lying outside the X chromosome. Center values represent average of converted mean CT values from two independent clones for each genotype, except for Xite-INV. Mean CT values for each clone is an average from three technical replicates.


A
TADs
TAD-D
TAD-E
TAD-F
CBS fw
CBS rv
Genes
Gm6206 Nap1L2 Cdx4 Tsx Gm9785 Cnbp2 Xpct C77370
Linx Chic1 Xite Tsix Xist Jpx Ftx Rnf12
B
Tsix TSS viewpoint
Xite CTCF viewpoint
Xist TSS viewpoint
Jpx TSS viewpoint
5 6
12 3 4
C
Tsix TSS viewpoint
Xist TSS viewpoint
WT #2
40kb #2
40kb #2 − WT #2
Interactions per 10k
Delta
D
Tsix TSS viewpoint
Xist TSS viewpoint
WT #3
70kb #2
70kb #2 − WT #3
E
Xite CTCF viewpoint
Jpx TSS viewpoint
WT #2
70kb #2
70kb #2 − WT #2

F


WT_Tsix_rep1
INV_Xist_rep1
WT_Tsix_rep2
INV_Xist_rep2
interactions/10k
0
30
100200
100400
100600
100800
101000
101200
Position at chrX (kb)
WT_Xist_rep1
INV_Tsix_rep1
WT_Xist_rep2
INV_Tsix_rep2
interactions/10k
0
30
100200
100400
100600
100800
101000
101200
Position at chrX (kb)

G


Tsix viewpoint
Wildtype
40kb [Xist−Tsix]
monotonic regression
0
100
300
500
−600
−400
−200
0
200
400
600
distance from viewpoint (kb)
Xist viewpoint
Wildtype
40kb [Xist−Tsix]
monotonic regression
0
100
300
500
−600
−400
−200
0
200
400
600
distance from viewpoint (kb)

Supplementary Fig. 2:

Note: Each inversion has been generated twice in independent cell lines. Capture-C results for the second clones are shown here, results for the first clones are shown in Figure 1-2.

A) TAD-D and -E as in Fig. 1B. Dashed box represents zoom-in in B.

B) Zoom-in of dashed box in A. Probes for capture enrichment (viewpoint) depicted with arrowheads. *Xite* viewpoint at CTCF site and *Xist*, *Tsix* and *Jpx* viewpoints at transcription start sites. CBSs in blue (forward) and red (reverse).

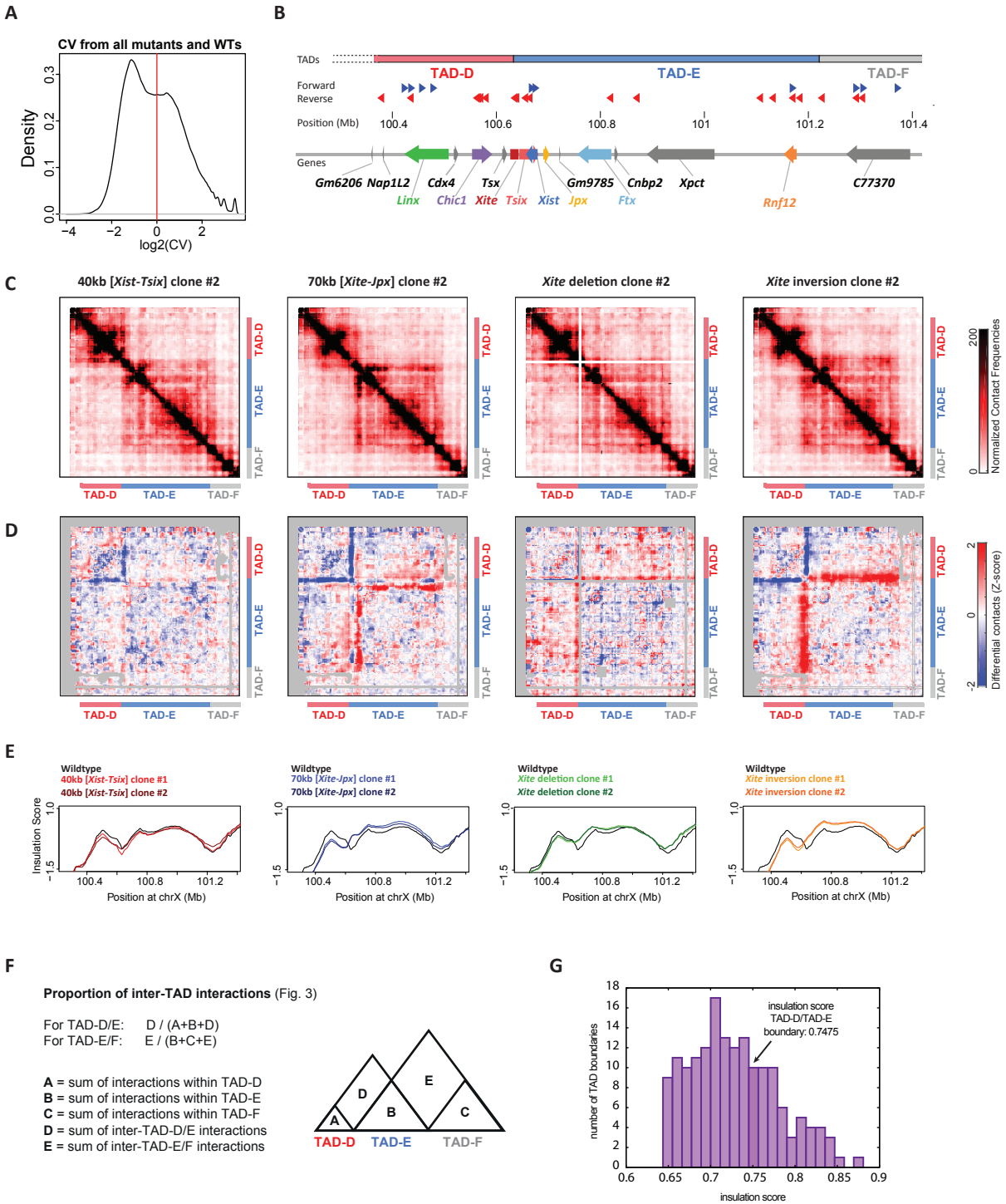
C) Capture-C profiles for *Tsix* (red) and *Xist* (blue) viewpoints in wild type and 40kb [*Tsix-Xist*] inversion. Profiles represent number of interactions for each DpnII fragment per 10,000 total interactions within the analyzed region (chrX:100214149-101420149), smoothened with a running mean of 7 DpnII fragments. Underneath the profiles, forward CBSs in blue, reverse CBSs in red and genes are annotated by 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 are depicted in black (interaction gain) and grey (interaction loss).

D) Capture-C profiles and differential interaction frequencies, as in C, for *Tsix* (red) and *Xist* (blue) viewpoints in wild type and 70kb [*Xite-Jpx*] inversion.

E) Capture-C profiles and differential interaction frequencies, as in C, for *Xite* (dark red) and *Jpx* (yellow) viewpoints in wild type and 70kb [*Xite-Jpx*] inversion.

F) Capture-C profiles of 40kb inverted *Xist* (blue) and 40kb inverted *Tsix* (red) imposed on Capture-C profiles of respectively wildtype *Tsix* (grey) and wildtype *Xist* (grey). Capture-C profiles are represented as in C.

G) Monotonic regression model of the Capture-C signal is calculated using peakC (Geeven, Teunissen et al. 2018) (de Leeuw 2009) and plotted as a function of the distance to the *Tsix* and *Xist* viewpoint, for wildtype (black) compared to 40kb[*Tsix-Xist*] inverion. (red).



F

Proportion of inter-TAD interactions (Fig. 3)

For TAD-D/E: D / (A+B+D)
For TAD-E/F: E / (B+C+E)

A = sum of interactions within TAD-D
B = sum of interactions within TAD-E
C = sum of interactions within TAD-F
D = sum of inter-TAD-D/E interactions
E = sum of inter-TAD-E/F interactions

G

Supplementary Fig. 3:

A) Distribution of the coefficient of variation (CV) of interactions (pixels in the 5C map) in a 10x10 square centered on every interaction for all 5C samples. Red line indicates applied threshold of CV=1 (See Methods), chosen based on bimodal distribution.

B-E) 5C results for the second independent cell line of each genomic alteration. Results for the first clones are shown in main Fig. 1-2.

B) Linear visualization of the *Xic* organized into two TADs as in Fig. 1. TAD boundaries are determined using the insulation score, according to (Crane, Bian et al. 2015). Dotted lines at start of TAD-D indicate undefined TAD structure due to repetitive sequences. CTCF bound sequence motifs (CBSs) in forward (blue) and reverse (red) orientation (Nora, Goloborodko et al. 2017). Gene annotation from UCSC RefSeq mm9 (Kent, Sugnet et al. 2002), except for *Xite* (see Methods) and *Linx* (Nora, Lajoie et al. 2012).

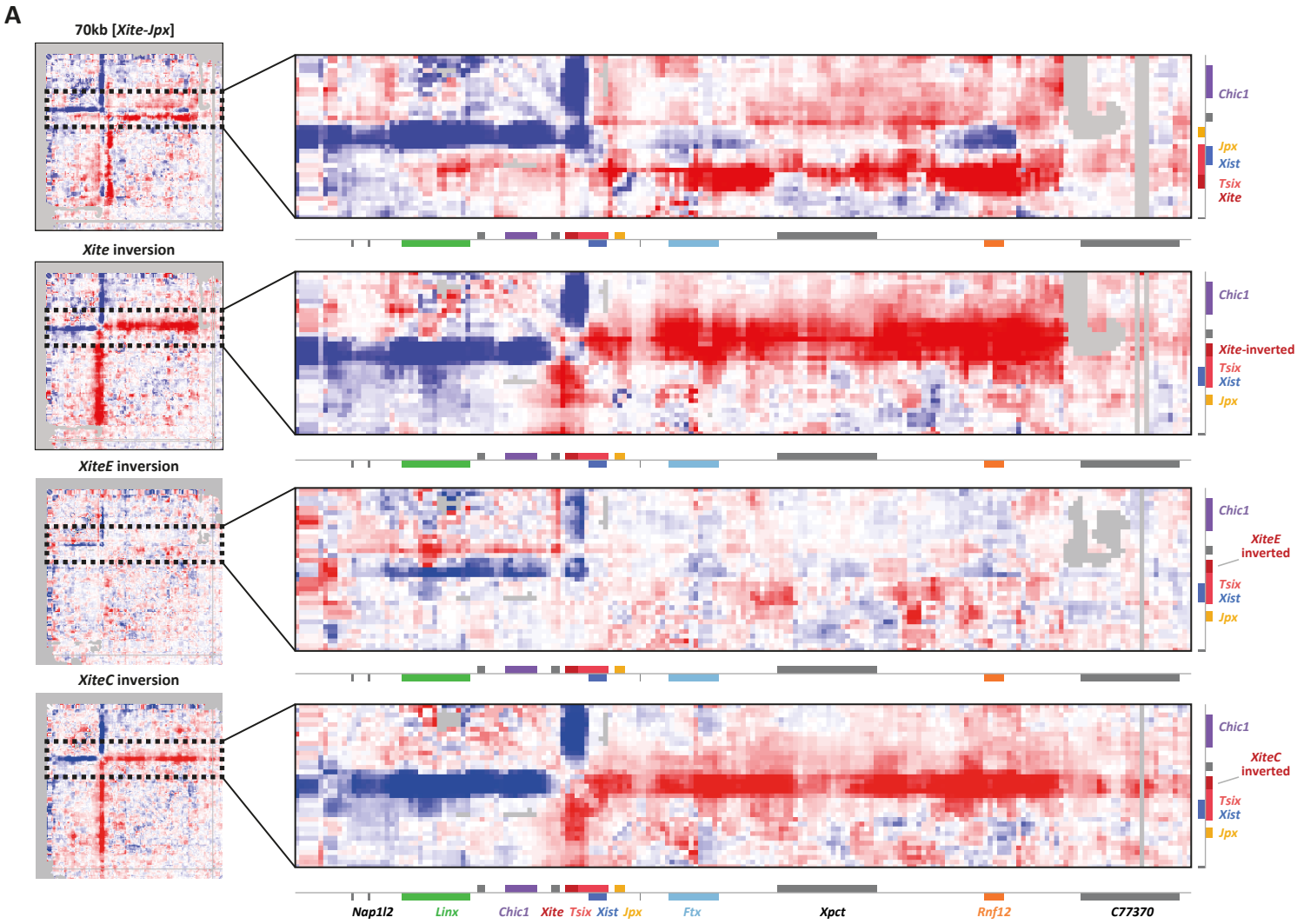
C) 5C chromosome conformation contact frequencies in male mESCs harboring a 40kb [*Tsix-Xist*] inversion, 70kb [*Xite-Jpx*] inversion, *Xite* deletion or *Xite* inversion. Pools of two replicates, data has been inverted accordingly (i.e. maps represent the inverted genome), normalized and binned as in Fig. 1-2.

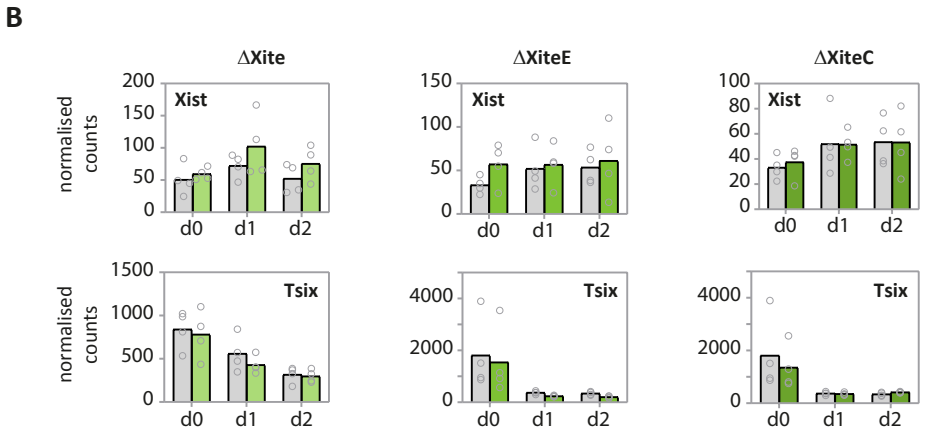
D) Differential 5C maps of mutant versus wild type mESCs. Differential maps represent the subtraction of wild type Z-scores from mutant Z-scores (see Methods). Gray pixels correspond to interactions that were filtered because they did not meet the quality control threshold (see Methods).

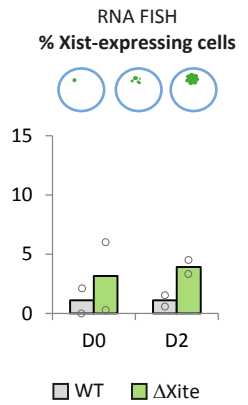
E) Insulation scores calculated according to (Crane, Bian et al. 2015). Wild type in black, 40kb [*Tsix-Xist*] inversion in red, 70kb [*Xite-Jpx*] inversion in blue, *Xite* deletion in green and *Xite* inversion in orange. Insulation scores for both clones are shown in the same graph.

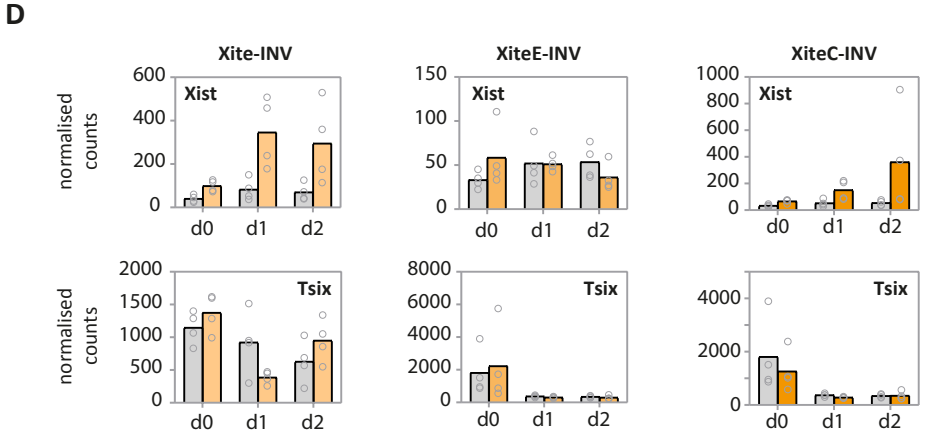
F) Supporting figure for Fig. 3C, depicting the calculations of the proportion of inter-TAD interactions.

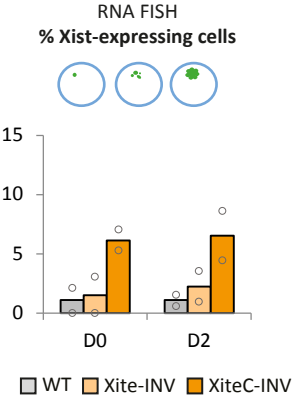
G) Distribution of insulation scores for TAD boundaries on the X chromosome. The reciprocal insulation score (Zhan, Mariani et al. 2017) was calculated for each TAD boundary using a published Hi-C dataset of mESCs (Giorgetti, Lajoie et al. 2016). The position of TADs and their boundaries on the X chromosome was determined using the CaTCH algorithm with a reciprocal insulation score of 65%.(Zhan, Mariani et al. 2017) The insulation score for the TAD-D/TAD-E boundary (indicated) is above the median.


A
70kb [Xite-Jpx]
Xite inversion
XiteE inversion
XiteC inversion
Chic1
Jpx
Xist
Tsix
Xite
Xite-inverted
XiteE inverted
XiteC inverted
Nap1l2
Linx
Chic1
Xite
Tsix
Xist
Jpx
Ftx
Xpct
Rnf12
C77370


B
ΔXite
ΔXiteE
ΔXiteC
normalised counts
Xist
Tsix
d0
d1
d2


C
RNA FISH
% Xist-expressing cells
D0
D2
WT
ΔXite


D
Xite-INV
XiteE-INV
XiteC-INV
normalised counts
Xist
Tsix
d0
d1
d2


E
RNA FISH
% Xist-expressing cells
D0
D2
WT
Xite-INV
XiteC-INV

Supplementary Fig. 4:

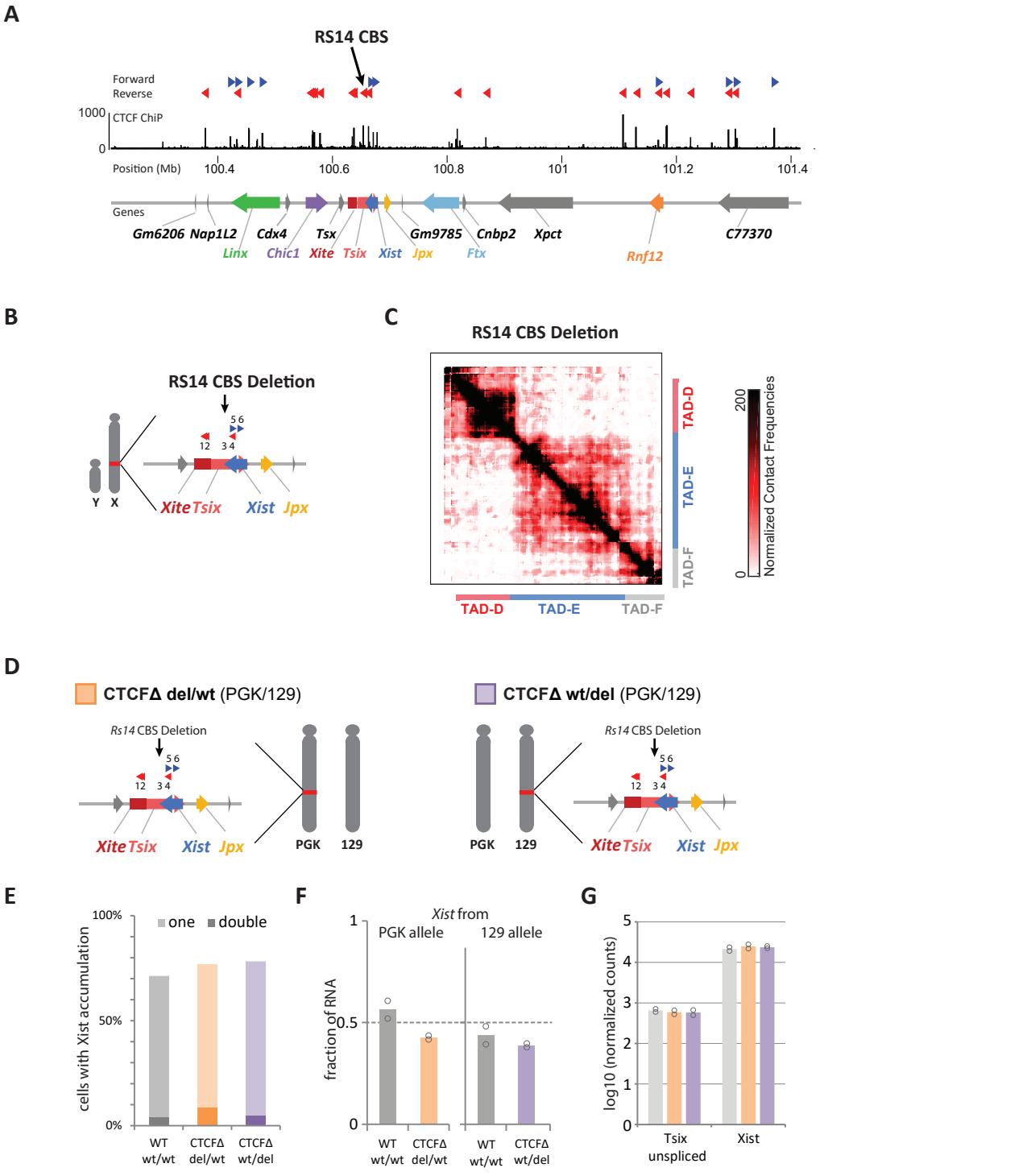
A) Zoom-ins of differential 5C maps shown in Fig. 2F, 3G and 3J. Differential maps represent the subtraction of wild type Z-scores from mutant Z-scores (see Methods). Gray pixels correspond to interactions that were filtered because they did not meet the quality control threshold (see Methods).

B) nCounter® mRNA expression levels in Δ*Xite,* Δ*XiteE* and Δ*XiteC.* Wildtype samples for Δ*Xite* versus Δ*XiteE/*Δ*XiteC* are different because they were processed in different batches. Bars depict the mean of four biologically independent samples.

C) *Xist* expression analyzed by RNA-FISH in male wild type (grey) and Δ*Xite* (green) mESCs at day 0 or 2 of EpiLSC differentiation. We did not consider cells in which we could not detect either *Huwe1* (X-linked) or *Xist* expression. *Xist* expression was scored based on the presence of transcript pinpoint, accumulation or clouds. Bars represent the average percentage of counted cells in two independent replicates (n>100 cells for each replicate) (Supplementary Table 6 for exact sample size details).

D) nCounter® mRNA expression levels in *Xite, XiteE* and *XiteC* inversions. Wildtype samples for *Xite*-inversion versus *XiteE/XiteC*-inversions are different because they were processed in different batches. Bars depict the mean of four biologically independent samples.

E) *Xist* expression analyzed by RNA-FISH in male wild type (grey), *Xite*-inversion (yellow) and *XiteC*-inversion (orange) mESCs at day 0 or 2 of EpiLSC differentiation. We did not consider cells in which we could not detect neither *Huwe1* (X-linked) nor *Xist* expression. *Xist* expression was scored based on the presence of transcript pinpoint, accumulation or clouds. Bars represent the average percentage of counted cells in two independent replicates (n>100 cells for each replicate) (Supplementary Table 6 for exact sample size details).


A
RS14 CBS
Forward
Reverse
CTCF ChIP
1000
0
Position (Mb)
100.4
100.6
100.8
101
101.2
101.4
Genes
Gm6206
Nap1L2
Cdx4
Tsx
Gm9785
Cnbp2
Xpct
C77370
Linx
Chic1
Xite
Tsix
Xist
Jpx
Ftx
Rnf12
B
RS14 CBS Deletion
Y
X
Xite
Tsix
Xist
Jpx
C
RS14 CBS Deletion
TAD-D
TAD-E
TAD-F
200
0
Normalized Contact Frequencies
D
CTCFΔ del/wt (PGK/129)
CTCFΔ wt/del (PGK/129)
Rs14 CBS Deletion
PGK
129
E
cells with Xist accumulation
100%
50%
0%
one
double
WT wt/wt
CTCFΔ del/wt
CTCFΔ wt/del
F
Xist from
PGK allele
129 allele
fraction of RNA
1
0.5
0
G
log10 (normalized counts)
Tsix unspliced
Xist

Supplementary Fig. 5:

A) Linear visualization of the *Xic* organized into two TADs, as in Fig. 1. CTCF binding sites (CBS) in forward (blue) and reverse (red) orientation, and CTCF ChIP-seq signal from (Nora, Goloborodko et al. 2017). Gene annotation from UCSC RefSeq mm9 (Kent, Sugnet et al. 2002), except for *Xite* (see Methods) and *Linx* (Nora, Lajoie et al. 2012). Black arrow indicates the RS14 CBS.

B) Schematic representation of male mESCs harboring an RS14 CBS deletion.

C) 5C chromosome conformation contact frequencies in male mESCs harboring a deletion of the RS14 CTCF binding site 3' of *Xist*. Map represents pool of two independent mutant cell lines, normalized and binned at 30kb (see Methods). White blocks depict non-mapped sequences. Underneath: the TAD annotation based on wild type 5C data.

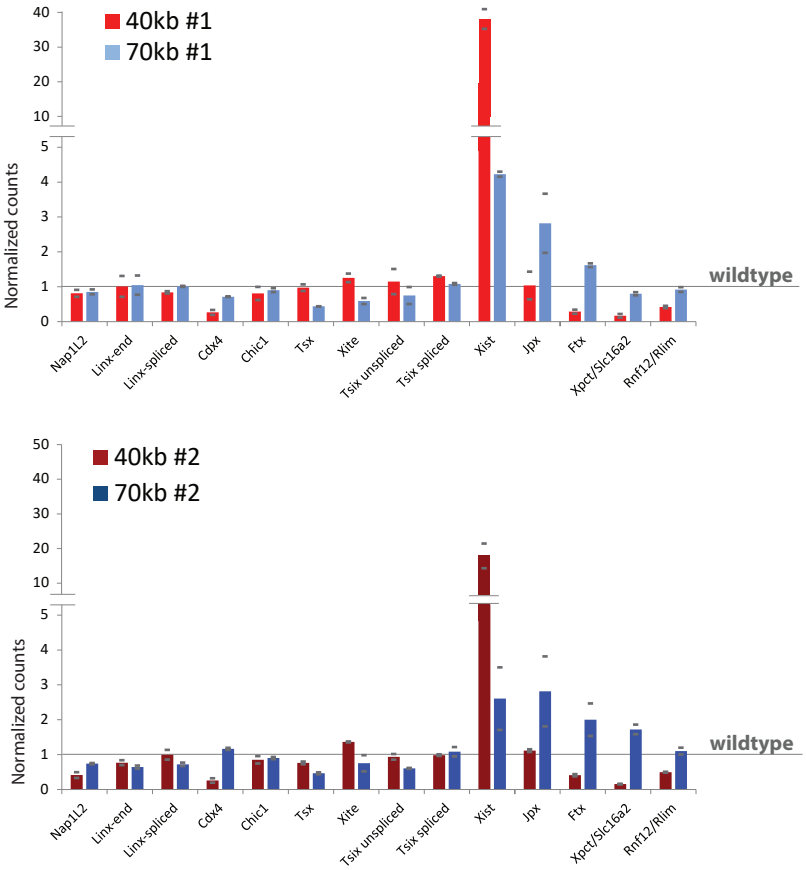
D) Schematic representation of female hybrid mESC lines harboring a heterozygous deletion of the RS14 CTCF binding site on either the PGK chromosome (orange) or the 129 chromosome (purple).

E) Efficiency of XCI analyzed by means of RNA-FISH analysis in wild type (grey), deletion/wild type (orange) and wild type/deletion (purple) female epiLSCs at day 4 of differentiation. Bars represent percentage of cells with one (light color) or two (dark color) chromosomes exhibiting *Xist* accumulation or cloud formation, for each independent cell line.

F) Allelic *Xist* expression, measured by pyrosequencing of *Xist* cDNA from female epiLSCs at day 4 of differentiation. Bars represent the fraction of *Xist* mRNA expressed from the PGK (left) or 129 (right) allele in cells with two wild type alleles (grey), in cells with the CTCF deletion on the PGK allele (orange) and cells with the CTCF deletion on the 129 allele (purple). Bars depict the mean of two independent experiments, with dots depicting theindividual experiments. CTCF deletion shows a slight decrease in *Xist* expression in *cis*.

G) nCounter® mRNA expression levels of *Xist* and *Tsix* in female epiLSCs at day 4 of differentiation harboring the CBS deletion on the PGK chromosome (orange) or the 129 chromosome (purple). Bars depict the mean of two independent experiments, with dots depicting the individual experiments.

A

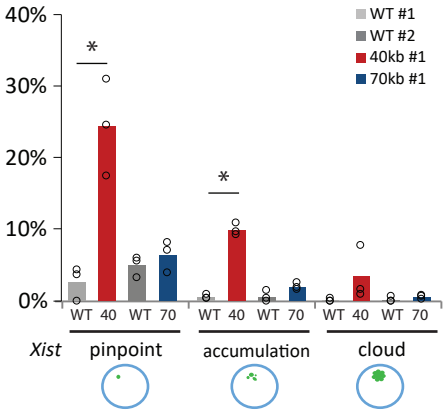


B

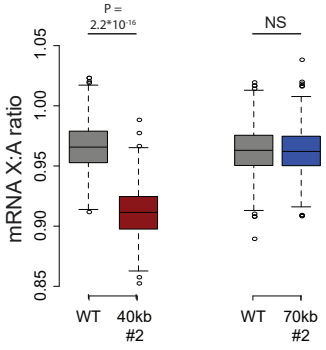
Unchanged Downregulated Upregulated according to RNAseq analysis

	Nap1l2	Cdx4	Chic1	Tsx	Xite	Tsix	Xist	Jpx	Ftx	Xpct	Rlim
40kb #1		FDR<0.05					FDR<0.05			FDR<0.05	FDR<0.05
70kb #1							FDR<0.05	FDR<0.05	FDR<0.05		
40kb #2	FDR<0.05	FDR<0.05					FDR<0.05			FDR<0.05	FDR<0.05
70kb #2							FDR<0.05	FDR<0.05	FDR<0.05		

C



D



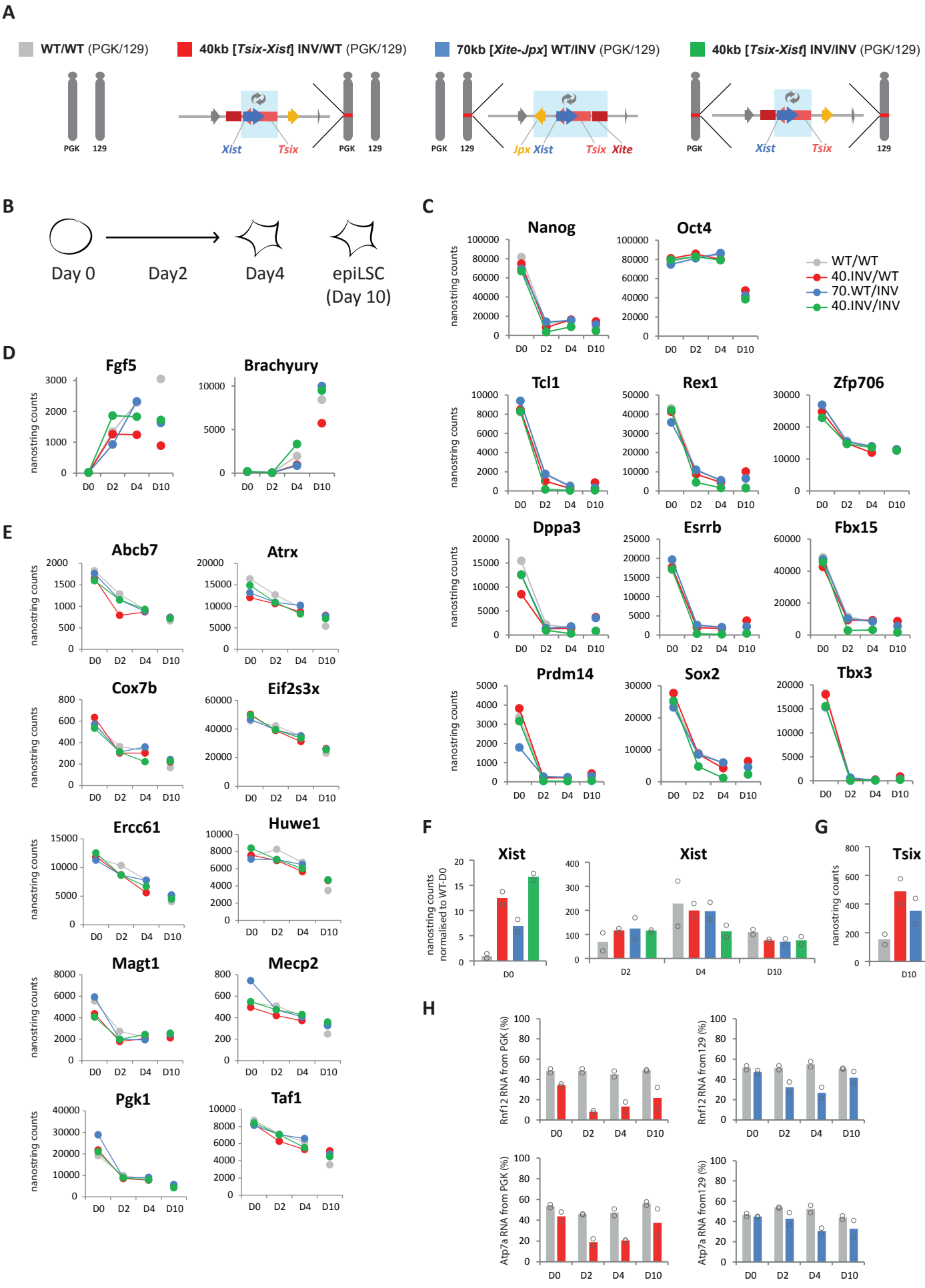
Supplementary Fig. 6:

A) mRNA expression levels from nCounter® analysis in 40kb inversion (red and dark red) and 70kb inversion (blue and dark blue) clones. mRNA levels are normalized to wild type (dashed grey line). Bars depict the mean of two independent experiments, with grey rectangles depicting the individual experiments.

B) Summary of results from differential expression analysis of mRNA sequencing data for genes within the *Xic*. Two independent experiments. Genes significantly upregulated are depicted in red, and genes significantly downregulated in green. Statistical analysis was performed by computing genewise exact tests for differences in the means between two groups of negative binomially distributed counts, and adjusted for multiple testing by Benjamini-Hochberg correction.

C) As in main Fig. 4B. *Xist* accumulation analyzed by RNA-FISH in male wild type (grey & dark grey), 40kb inversion #2 (dark red) and 70kb inversion #2 (dark blue) mESCs. Bars represent percentage of counted cells with either a *Xist* pinpoint, *Xist* accumulation or full *Xist* cloud. For each cell line three independent experiments were performed, each counting from 100 to 400 cells (Supplementary Table 6 for exact sample size details). Dots represent individual experiments. 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.

D) 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) and the minimum and maximum (whiskers) of the bootstrap ratios. Statistical analysis was performed using two-sided Wilcoxon Rank Sum test.


A
WT/WT (PGK/129)
40kb [Tsix-Xist] INV/WT (PGK/129)
70kb [Xite-Jpx] WT/INV (PGK/129)
40kb [Tsix-Xist] INV/INV (PGK/129)
PGK
129
Xist
Tsix
Jpx
Xite
B
Day 0
Day2
Day4
epiLSC
(Day 10)
C
Nanog
Oct4
Tcl1
Rex1
Zfp706
Dppa3
Esrrb
Fbx15
Prdm14
Sox2
Tbx3
nanostring counts
WT/WT
40.INV/WT
70.WT/INV
40.INV/INV
D0
D2
D4
D10
D
Fgf5
Brachyury
E
Abcb7
Atrx
Cox7b
Eif2s3x
Ercc61
Huwe1
Magt1
Mecp2
Pgk1
Taf1
F
Xist
nanostring counts
normalised to WT-D0
G
Tsix
H
Rnf12 RNA from PGK (%)
Rnf12 RNA from129 (%)
Atp7a RNA from PGK (%)
Atp7a RNA from129 (%)

Supplementary Fig. 7:

A) Schematic representation of female hybrid mESC lines harboring a heterozygous 40kb [*Tsix-Xist*] inversion on the PGK chromosome (red), a 70kb [*Xite-Jpx*] inversion on the 129 chromosome (blue) and a homozygous 40kb [Tsix-Xist] inversion (green). Light blue boxes indicate inverted region. The inverted gene localization is shown, compare to Fig. 1B for wild type localizations.

B) Schematic representation of mESC to epiLSC induction, with cell collections at day 0, day 2 and day 4 of early epiLSC induction and an additional collection in epiLSCs that were induced independently for 10 days.

C) nCounter® mRNA expression levels of pluripotency factors and mESC-specific genes in the female cell lines described in A, at 0, 2 and 4 days of mESC into epiLSC induction and in day 10 epiLSCs. Dots depict the mean of two replicates.

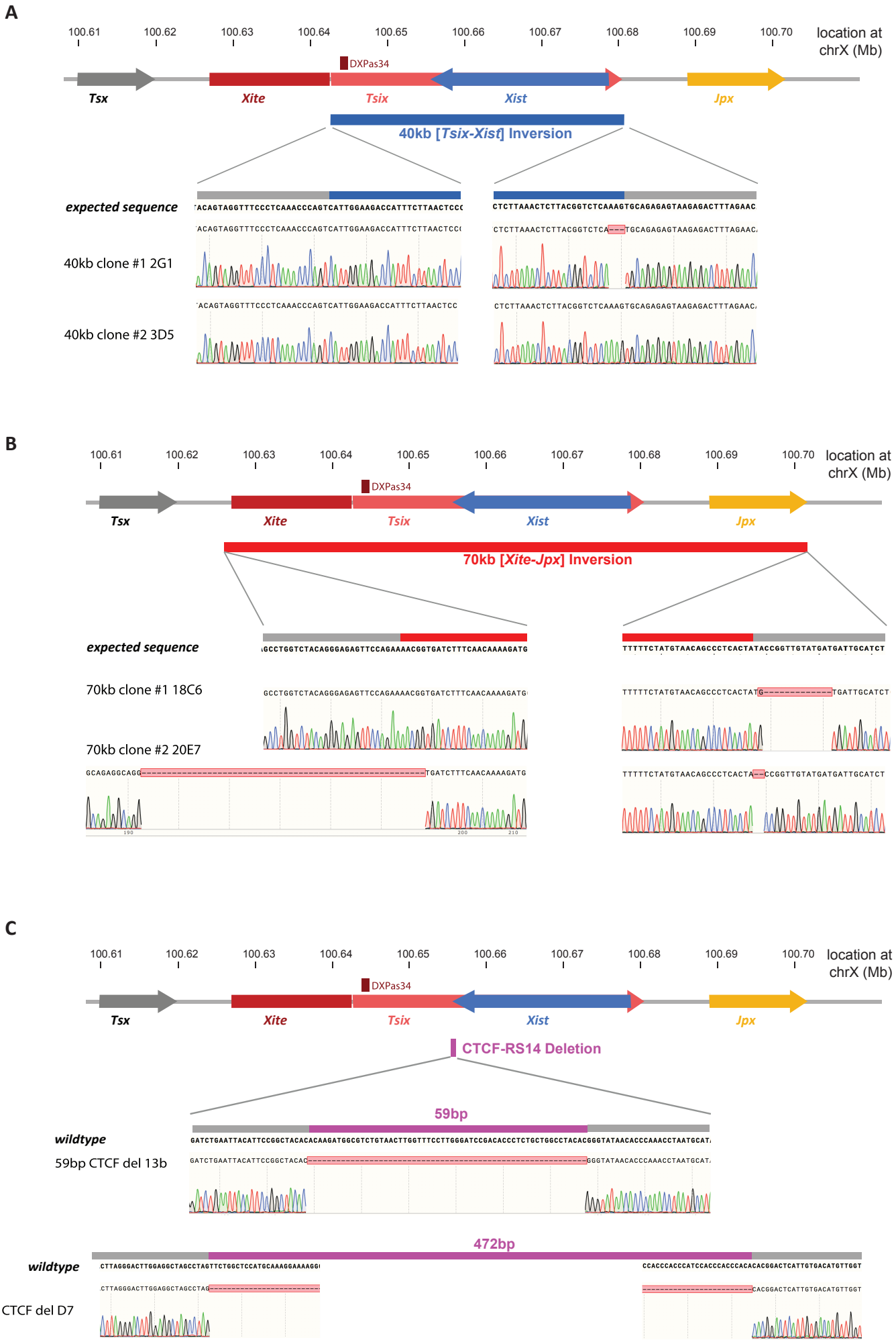
D) As in C, for two epiLSC-specific genes.

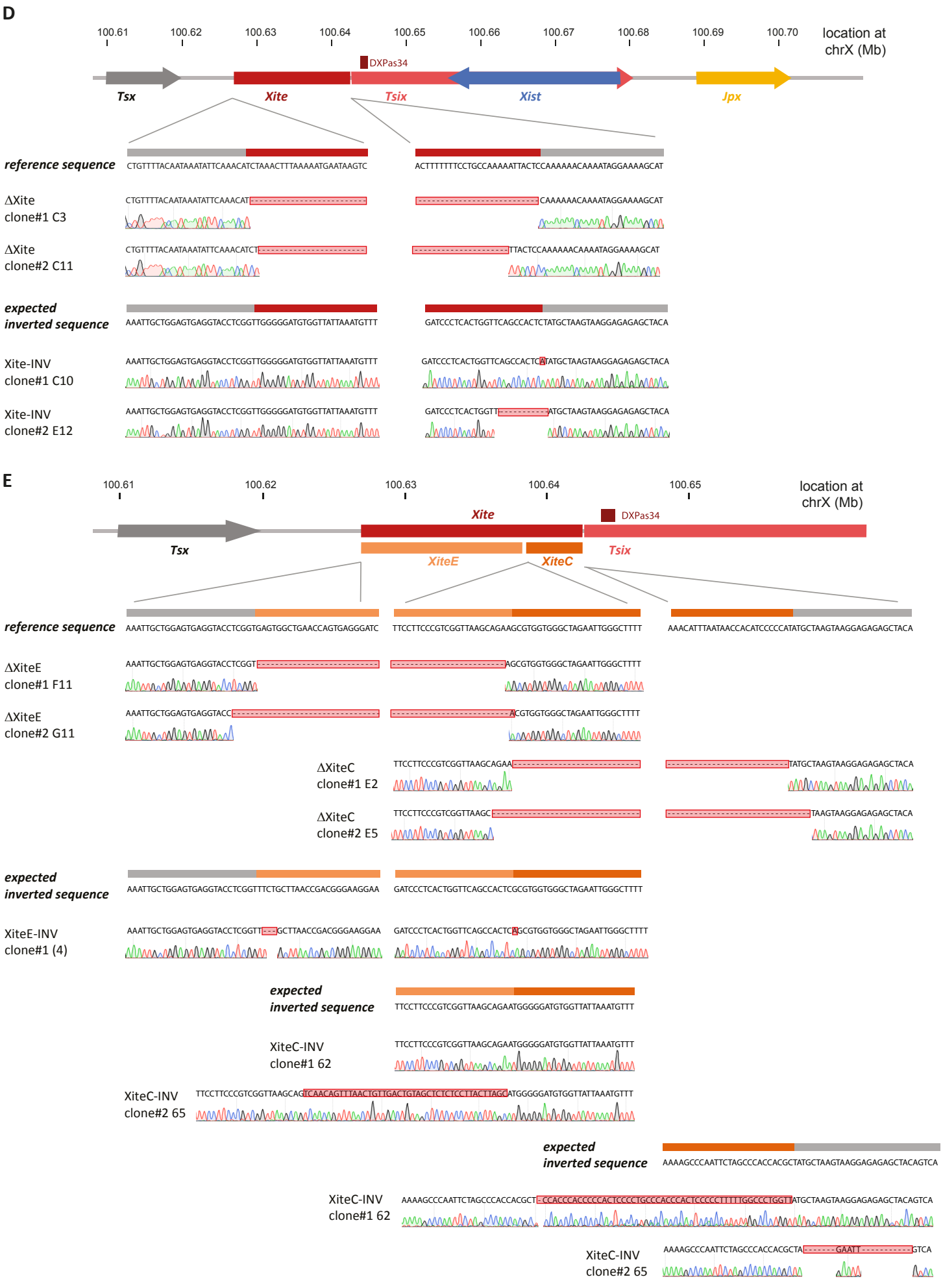
E) As in C, for X-linked genes normally silenced during epiLSC induction.

F) As in C, global *Xist* expression.

G) As in C, global *Tsix* expression.

H) Allelic *Rnf12* and *Atp7a* expression, measured by a quantification pyrosequencing-based assay of cDNA from female cells at day 0, 2, 4 and 10 of epiLSCs differentiation. Bars represent the percentage of mRNA expressed from the PGK (left) or 129 (right) allele in cells described in A. Bars depict the mean of two replicates.


A
100.61
100.62
100.63
100.64
100.65
100.66
100.67
100.68
100.69
100.70
location at chrX (Mb)
DXPas34
Tsx
Xite
Tsix
Xist
Jpx
40kb [Tsix-Xist] Inversion
expected sequence
40kb clone #1 2G1
40kb clone #2 3D5
B
location at chrX (Mb)
DXPas34
Tsx
Xite
Tsix
Xist
Jpx
70kb [Xite-Jpx] Inversion
expected sequence
70kb clone #1 18C6
70kb clone #2 20E7
C
location at chrX (Mb)
DXPas34
Tsx
Xite
Tsix
Xist
Jpx
CTCF-RS14 Deletion
59bp
wildtype
59bp CTCF del 13b
472bp
wildtype
472bp CTCF del D7


D
100.61
100.62
100.63
100.64
100.65
100.66
100.67
100.68
100.69
100.70
location at chrX (Mb)
DXPas34
Tsx
Xite
Tsix
Xist
Jpx
reference sequence
CTGTTTTACAATAAATATTCAAACATCTAAACTTTAAAAATGAATAAGTC
ACTTTTTTTCCTGCCAAAAATTACTCCAAAAAACAAAATAGGAAAAGCAT
ΔXite clone#1 C3
CTGTTTTACAATAAATATTCAAACAT----------
----------CAAAAAACAAAATAGGAAAAGCAT
ΔXite clone#2 C11
CTGTTTTACAATAAATATTCAAACATCT----------
----------TTACTCCAAAAAACAAAATAGGAAAAGCAT
expected inverted sequence
AAATTGCTGGAGTGAGGTACCTCGGTTGGGGGATGTGGTTATTAAATGTTT
GATCCCTCACTGGTTCAGCCACTCTATGCTAAGTAAGGAGAGAGCTACA
Xite-INV clone#1 C10
AAATTGCTGGAGTGAGGTACCTCGGTTGGGGGATGTGGTTATTAAATGTTT
GATCCCTCACTGGTTCAGCCACTCATATGCTAAGTAAGGAGAGAGCTACA
Xite-INV clone#2 E12
AAATTGCTGGAGTGAGGTACCTCGGTTGGGGGATGTGGTTATTAAATGTTT
GATCCCTCACTGGTT----------ATGCTAAGTAAGGAGAGAGCTACA
E
100.61
100.62
100.63
100.64
100.65
location at chrX (Mb)
Xite
DXPas34
Tsx
XiteE
XiteC
Tsix
reference sequence
AAATTGCTGGAGTGAGGTACCTCGGTGAGTGGCTGAACCAGTGAGGGATC
TTCCTTCCCGTCGGTTAAGCAGAAGCGTGGTGGGCTAGAATTGGGCTTTT
AAACATTTAATAACCACATCCCCCATATGCTAAGTAAGGAGAGAGCTACA
ΔXiteE clone#1 F11
AAATTGCTGGAGTGAGGTACCTCGGT----------
----------AGCGTGGTGGGCTAGAATTGGGCTTTT
ΔXiteE clone#2 G11
AAATTGCTGGAGTGAGGTACC----------
----------ACGTGGTGGGCTAGAATTGGGCTTTT
ΔXiteC clone#1 E2
TTCCTTCCCGTCGGTTAAGCAGAA----------
----------TATGCTAAGTAAGGAGAGAGCTACA
ΔXiteC clone#2 E5
TTCCTTCCCGTCGGTTAAGC----------
----------TAAGTAAGGAGAGAGCTACA
expected inverted sequence
AAATTGCTGGAGTGAGGTACCTCGGTTTCTGCTTAACCGACGGGAAGGAA
GATCCCTCACTGGTTCAGCCACTCGCGTGGTGGGCTAGAATTGGGCTTTT
XiteE-INV clone#1 (4)
AAATTGCTGGAGTGAGGTACCTCGGTT---GCTTAACCGACGGGAAGGAA
GATCCCTCACTGGTTCAGCCACTCAGCGTGGTGGGCTAGAATTGGGCTTTT
expected inverted sequence
TTCCTTCCCGTCGGTTAAGCAGAATGGGGGATGTGGTTATTAAATGTTT
XiteC-INV clone#1 62
TTCCTTCCCGTCGGTTAAGCAGAATGGGGGATGTGGTTATTAAATGTTT
XiteC-INV clone#2 65
TTCCTTCCCGTCGGTTAAGCAGTCAACAGTTTAACTGTTGACTGTAGCTCTCTCCTTACTTAGCATGGGGGATGTGGTTATTAAATGTTT
expected inverted sequence
AAAAGCCCAATTCTAGCCCACCACGCTATGCTAAGTAAGGAGAGAGCTACAGTCA
XiteC-INV clone#1 62
AAAAGCCCAATTCTAGCCCACCACGCT-CCACCCACCCCCACTCCCCTGCCCCACCCACTCCCCCTTTTTGGCCCTGGTTATGCTAAGTAAGGAGAGAGCTACAGTCA
XiteC-INV clone#2 65
AAAAGCCCAATTCTAGCCCACCACGCTA------GAATT----------GTCA

Supplementary Fig. 8:
Sanger sequencing results of the clones generated with following genotypes: **(A)** 40kb [*Tsix-Xist*] inversion, **(B)** 70kb [*Xite-Jpx*] inversion, **(C)** RS14 deletions, **(D)** *Xite* deletions and inversions, **(E)** *XiteE* and *XiteC* deletions and inversions.

Supplementary Note 1

The "flame" observed in the 70kb inversion likely arises due to the inversion of the CTCF sites within *Xite*, as we show that the structural element within *Xite*, which harbours CTCF-bound sites, is associated with the structural phenotypes observed when deleting or inverting *Xite* (Figure 3). This is also consistent with the role of CTCF in mediating the formation of loops and TADs (Nora et al. 2017). Moreover, orientation of CTCF sites has been shown to be critical for the formation of loops (reviewed in Ghirlando and Felsenfeld, 2016), and the new orientation of the *Xite* CTCF sites in the 70kb inversion (compare Fig. 2A with Fig. 1C) is consistent with increased interactions along TAD-E. "Flames" have been described in the literature, from a theoretical point of view, as "straight lines often emerging from the borders of TADs" (Fudenberg et al. 2016, 2017). We would like to underline that the "flame" we observe in the 70kb inversion might not exactly match the description of the theoretical "flame", as ours is not a homogeneous gain of interactions along a TAD, as some regions gain more interactions than others (especially the ones containing CTCF sites, see *Xite* Capture-C profile in Fig. 2C). Nevertheless, we believe that our "flame" arises via similar mechanisms as described for the theoretical flames. Simulations of the "loop extrusion model" predict the formation of flames (Fudenberg et al. 2016, 2017), which arise "when the loop extrusion factors become halted on one side at a barrier while continuing to extrude from the other side" (Fudenberg et al. 2017). CTCF has been proposed to act as a barrier for loop extrusion factors, which also supports our idea that the inversion of the *Xite* CTCF sites underlies the formation of our "flame".

Supplementary Note 2

Besides the differences in inter-TAD interactions, we also note a few other important changes in the differential maps (Fig. 3B, 3E). (i) For the *Xite*-deletion (Fig. 3B), the differential map shows a strong gain of interactions between sequences right upstream of the deletion and sequences along the *Xist*-TAD. This reveals a change in the boundary position, as determined in Fig. 3D and apparent in the 5C map (Fig. 3B, bottom), in the absence of *Xite*. These results reinforce the idea that *Xite* plays an important role in defining the boundary between the two TADs. (ii) For the *XiteE* deletion (Fig. 3E), the differential map shows a strong gain of interactions between sequences right upstream of the deletion and sequences along the *Tsix*-TAD. We think this is because the sequences right upstream of the deletion are now juxtaposed to the *XiteC* element, which "drags" them along in its interactions. Please note that we have corrected our Z-score calculations (underlying the differential maps) to account for linear distance effects due to the deletions; the fact that we nevertheless observe this localized gain of interactions means that these new interactions are strongly above what would be expected just due to a change in linear distance. The fact that this happens in such a localized manner is intriguing. It is known that the probability of contacts decays quickly with linear genomic distance, which might at least partially explain the localized effect. (iii) For the *XiteC* deletion (Fig. 3E), the differential map shows a strong loss of interactions between sequences right upstream of the deletion and sequences along the *Tsix*-TAD. This suggests that *XiteC* normally "drags" those sequences along while interacting with sequences in *Tsix*-TAD - also compare the individual wild type and *Δ*XiteC 5C maps (Fig. 3), in which

the normal interactions between *Xite* and *Linx* and *Chic1* within the *Tsix*-TAD are much weakened. This likely explains the localized loss of interactions seen in the differential map.

Supplementary Note 3

One of the CBSs at the TAD-D/TAD-E boundary, named RS14 in Spencer et al (CBS #3 in Fig. 1B and Supplementary Fig. 1), has been proposed as a primary determinant of the TAD boundary (Jegu et al. 2017), and a critical factor for *Xist* regulation (Spencer et al. 2011). This was based on a 2.3kb deletion that included not only RS14 but also overlapping sequences of *Xist* and *Tsix* loci (Spencer et al. 2011). To explore in further detail the role of RS14 in the topological organization of the *Xic*, we generated 59bp and 500bp CRISPR/Cas9-mediated deletions (see Methods) in male and female mESCs (Supplementary Fig. 1). 5C analysis of mutant male ESCs revealed no striking changes in the organization of the TADs or the TAD boundary (Supplementary Fig. 5), and we did not detect significant effects on *Xist* regulation in either mutant male mESCs (data not shown), or heterozygous female mESCs and EpiLSCs (Supplementary Fig. 5). Thus, RS14 alone cannot account for the physical and transcriptional insulation of TADs D and E within the *Xic*.

Supplementary Note 4

Using nCounter® technology (NanoString; see Methods), we assessed the expression status of *Xist* and *Tsix* in wild type and *Xite* mutant mESCs (Supplementary Fig. 4B,D), as well as of other genes included in our probe codeset (data not shown). *Xist* expression was slightly increased in ΔXite mutants at each time point (Supplementary Fig. 4B), which we confirmed by Xist RNA FISH (Supplementary Fig. 4C). The same trend seems to be present in ΔXiteE mutants, but not ΔXiteC (Supplementary Fig. 4B), consistent with the *Xite* enhancer element having a repressive effect on *Xist* expression (Ogawa and Lee, 2003). These expression changes were not, however, statistically significant. Upon inversion of the *Xite* element, *Xist* expression was significantly upregulated during differentiation when compared to control (Supplementary Fig. 4D; p=0.0005, two-factor ANOVA with replication). The same effect was observed upon inversion of *XiteC* (Supplementary Fig. 4D; p=0.0403, two-factor ANOVA with replication), but not of *XiteE*. These results suggest that *Xist* expression is sensitive to the orientation of *Xite* structural element (*XiteC*) and that the topological changes upon *XiteC* inversion affect *Xist* expression during differentiation (Fig. 3, Supplementary Fig. 4D). We further confirmed these results by Xist RNA FISH (Supplementary Fig. 4E; p=0.018 for Δ*XiteC*, two-factor ANOVA with replication). In all our *Xite* mutants, *Tsix* expression was not significantly affected during differentiation.

Supplementary Note 5

The presence of ectopic expression of *Xist* in male mESCs harboring an inversion suggests that it does not induce systematic or complete XCI; otherwise cells would be eliminated due to functional nullisomy after the silencing of their single X chromosome. Nevertheless, a certain degree of counter-selection could explain why we observed higher percentages of cells exhibiting ectopic *Xist* expression in female mESCs compared to male mESCs. This

could also be due to female cells harboring a double dose of the X-linked activators, which could facilitate the ectopic *Xist* expression in female mESCs.

Supplementary Note 6

Our results show that moving *Xist* to the TAD-D landscape leads to ectopic upregulation of *Xist* in female mESCs but hampers the expression and/or accumulation of Xist RNA throughout EpiLSC differentiation. Hampered Xist accumulation in both 40kb and 70kb inversions could be due to: (i) less efficient activation of *Xist*, due to the loss of its endogenous cis-regulatory landscape; (ii) less stably maintained ectopic *Xist* upregulation from the inverted allele, as the *cis*-regulatory environment of *Tsix* becomes less active during differentiation; or (iii) *Xist*-repression mediated by *Tsix*, if *Tsix* expression is somehow prolonged within TAD-E, which becomes more active during early differentiation. These possibilities are not mutually exclusive; we show that *Tsix* expression is prolonged in our female inversion cell lines (Fig. 5E-F), suggesting that this could be indeed one of the reasons for hampered *Xist* expression.

Supplementary Note 7

Xist allelic ratios and Xist RNA cloud quantification in our mutant female cell lines (Fig. 5D, 5F) suggest that *Xist* expression is impaired, not properly sustained or lost from the inverted allele during long-term differentiation. This is also supported by total *Xist* RNA quantification in control and mutant female EpiLSCs at day 10 of differentiation – all the inversion genotypes show reduced levels of *Xist* compared to control (Supplementary Fig. 7F). We next investigated how XCI is affected in the female inversion cell lines. Total expression levels of X-linked genes suggest that gene silencing is occurring during differentiation as expected (Supplementary Fig. 7E); however, allelic expression analysis reveals that reduced *Xist* expression from the inverted allele was accompanied by increased expression of X-linked genes from the same allele (Supplementary Fig. 7H). At early time points of differentiation (days 2 and 4), X-linked expression is biased: *Rnf12* and *Atp7a* expression levels from the inverted allele are substantially decreased (Supplementary Fig. 7H). This is consistent with ectopic *Xist* expression from the inverted allele leading to gene silencing in cis. At day 10 of differentiation, the allelic ratios of *Rnf12* and *Atp7a* expression are still biased; however, expression levels from the inverted allele are increased compared to day 4, and show decreased difference compared to wild type levels at the same time point. This suggests that the female inversion cell lines not only show impaired expression of *Xist*, but also show signs of impaired X-inactivation. These results suggest that the inactivate state of the X chromosome at the earlier time points might still be reversed; this probably results from the loss of *Xist* RNA coating and/or expression from the X chromosome harbouring the inverted allele.

Supplementary Note 8

Both *Tsix* and *Xist* are controlled by the pluripotency network, and this regulation appears to be more direct at *Tsix* than it is at *Xist*. *Xist* is bound by pluripotency factors mainly within its intron-1 (Navarro et al. 2008) and deletion of *Xist* intron-1 does not affect XCI regulation (Minkovsky et al. 2013). Although *Xist* is regulated by the pluripotency network, this

regulation could be indirect. *Tsix* on the other hand, is bound at its promoter by pluripotency factors (Navarro et al. 2010), and might therefore be less sensitive to the loss of its long-range cis-regulatory environment in the ES state, when the pluripotency network is active. The *Xist* promoter, which is lowly expressed in the ES state, may simply be inactive by the lack of activating cues and therefore more susceptible to being influenced by long-range regulatory elements in undifferentiated ESCs. Thus, switching the cis-regulation landscape of *Xist* and *Tsix* in the ES state might have more of an impact on the "lowly expressed" status of the *Xist* promoter than on the activated status of the *Tsix* promoter, which is probably sustained more directly by the pluripotency factors.