LETTER

doi:10.1038/nature11049

Spatial partitioning of the regulatory landscape of the X-inactivation centre

Elphège P. Nora[1,2,3], Bryan R. Lajoie[4]*, Edda G. Schulz[1,2,3]*, Luca Giorgetti[1,2,3]*, Ikuhiro Okamoto[1,2,3], Nicolas Servant[1,5,6], Tristan Piolot[1,2,3], Nynke L. van Berkum[4], Johannes Meisig[7], John Sedat[8], Joost Gribnau[9], Emmanuel Barillot[1,5,6], Nils Blüthgen[7], Job Dekker[4] & Edith Heard[1,2,3]

In eukaryotes transcriptional regulation often involves multiple long-range elements and is influenced by the genomic environment[1]. A prime example of this concerns the mouse X-inactivation centre (*Xic*), which orchestrates the initiation of X-chromosome inactivation (XCI) by controlling the expression of the non-protein-coding *Xist* transcript. The extent of *Xic* sequences required for the proper regulation of *Xist* remains unknown. Here we use chromosome conformation capture carbon-copy (5C)[2] and super-resolution microscopy to analyse the spatial organization of a 4.5-megabases (Mb) region including *Xist*. We discover a series of discrete 200-kilobase to 1 Mb topologically associating domains (TADs), present both before and after cell differentiation and on the active and inactive X. TADs align with, but do not rely on, several domain-wide features of the epigenome, such as H3K27me3 or H3K9me2 blocks and lamina-associated domains. TADs also align with coordinately regulated gene clusters. Disruption of a TAD boundary causes ectopic chromosomal contacts and long-range transcriptional misregulation. The *Xist*/*Tsix* sense/antisense unit illustrates how TADs enable the spatial segregation of oppositely regulated chromosomal neighbourhoods, with the respective promoters of *Xist* and *Tsix* lying in adjacent TADs, each containing their known positive regulators. We identify a novel distal regulatory region of *Tsix* within its TAD, which produces a long intervening RNA, *Linx*. In addition to uncovering a new principle of *cis*-regulatory architecture of mammalian chromosomes, our study sets the stage for the full genetic dissection of the X-inactivation centre.

The X-inactivation centre was originally defined by deletions and translocations as a region spanning several megabases[3,4], and contains several elements known to affect *Xist* activity, including its repressive antisense transcript *Tsix* and its regulators *Xite*, *DXPas34* and *Tsx*[5,6]. However, additional control elements must exist, as single-copy transgenes encompassing *Xist* and up to 460 kb of flanking sequences are unable to recapitulate proper *Xist* regulation[7]. To characterize the *cis*-regulatory landscape of the *Xic* in an unbiased approach, we performed 5C[2] across a 4.5-Mb region containing *Xist*. We designed 5C-Forward and 5C-Reverse oligonucleotides following an alternating scheme[2], thereby simultaneously interrogating nearly 250,000 possible chromosomal contacts in parallel, with a mean resolution of 10–20 kb (Fig. 1a; see Supplementary Methods). Analysis of undifferentiated mouse embryonic stem cells (ESCs) revealed that long-range (>50 kb) contacts preferentially occur within a series of discrete genomic blocks, each covering 0.2–1 Mb (Fig. 1b). These blocks differ from the higher-order organization recently observed by Hi-C[8], corresponding to much larger domains of open or closed chromatin, that come together in the nucleus to form A and B types of compartments[8]. Instead, our 5C analysis shows self-associating chromosomal domains occurring at the sub-megabase scale. The size and location of these domains is identical in male and female mouse ESCs (Supplementary Fig. 1) and in different mouse ESC lines (Supplementary Fig. 2 and Supplementary Data 1).

To examine this organization with an alternative approach, we performed three-dimensional DNA fluorescent *in situ* hybridization (FISH) in male mouse ESCs. Nuclear distances were found to be significantly shorter between probes lying in the same 5C domain than in different domains (Fig. 1c, d), and a strong correlation was found between three-dimensional distances and 5C counts (Supplementary Fig. 3a, b). Furthermore, using pools of tiled bacterial artificial chromosome (BAC) probes spanning up to 1 Mb and structured illumination microscopy, we found that large DNA segments belonging to the same 5C domain colocalize to a greater extent than DNA segments located in adjacent domains (Fig. 1e), and this throughout the cell cycle (Supplementary Fig. 3c, d). Based on 5C and FISH data, we conclude that chromatin folding at the sub-megabase scale is not random, and partitions this chromosomal region into a succession of topologically associating domains (TADs).

We next investigated what might drive chromatin folding in TADs. We first noticed a striking alignment between TADs and the large blocks of H3K27me3 and H3K9me2 (ref. 9) that are known to exist throughout the mammalian genomes[10–13] (for example, TAD E, Fig. 2 and Supplementary Fig. 4). We therefore examined 5C profiles of $G9a^{-/-}$ (also known as *Ehmt2*) mouse ESCs, which lack H3K9me2, notably at the *Xic*[14], and $Eed^{-/-}$ mouse ESCs, which lack H3K27me3 (ref. 15). No obvious change in overall chromatin conformation was observed, and TADs were not affected either in size or position in these mutants (Fig. 2 and Supplementary Fig. 4b). Thus TAD formation is not due to domain-wide H3K27me3 or H3K9me2 enrichment. Instead, such segmental chromatin blocks might actually be delimited by the spatial partitioning of chromosomes into TADs.

We then addressed whether folding in TADs is driven by discrete boundary elements at their borders. 5C was performed in a mouse ESC line carrying a 58-kb deletion (ΔXTX[16]), encompassing the boundary between the *Xist* and *Tsix* TADs (D and E; Fig. 2b). We observed ectopic contacts between sequences in TADs D and E and an altered organization of TAD E. Boundary elements can thus mediate the spatial segregation of neighbouring chromosomal segments. Within the TAD D–E boundary, a CTCF-binding site was recently implicated in insulating *Tsix* from remote regulatory influences[17]. However, alignment of CTCF- and cohesin-binding sites in mouse ESCs[18] with our 5C data showed that, although these factors are present at most TAD boundaries (Supplementary Fig. 4), they are also frequently present within TADs, excluding them as the sole determinants of TAD

[1]Institut Curie, 26 rue d'Ulm, Paris F-75248, France. [2]CNRS UMR3215, Paris F-75248, France. [3]INSERM U934, Paris F-75248, France. [4]Programs in Systems Biology and Gene Function and Expression, Department of Biochemistry and Molecular Pharmacology, University of Massachusetts Medical School, Worcester, Massachusetts 01605-0103, USA. [5]INSERM U900, Paris, F-75248 France. [6]Mines ParisTech, Fontainebleau, F-77300 France. [7]Institute of Pathology, Charité–Universitätsmedizin, 10117 Berlin, and Institute of Theoretical Biology Humboldt Universität, 10115 Berlin, Germany. [8]Department of Biochemistry and Biophysics, University of California San Francisco, San Francisco, California 94158-2517, USA. [9]Department of Reproduction and Development, Erasmus MC, University Medical Center, 3000 CA Rotterdam, The Netherlands.
*These authors contributed equally to this work.

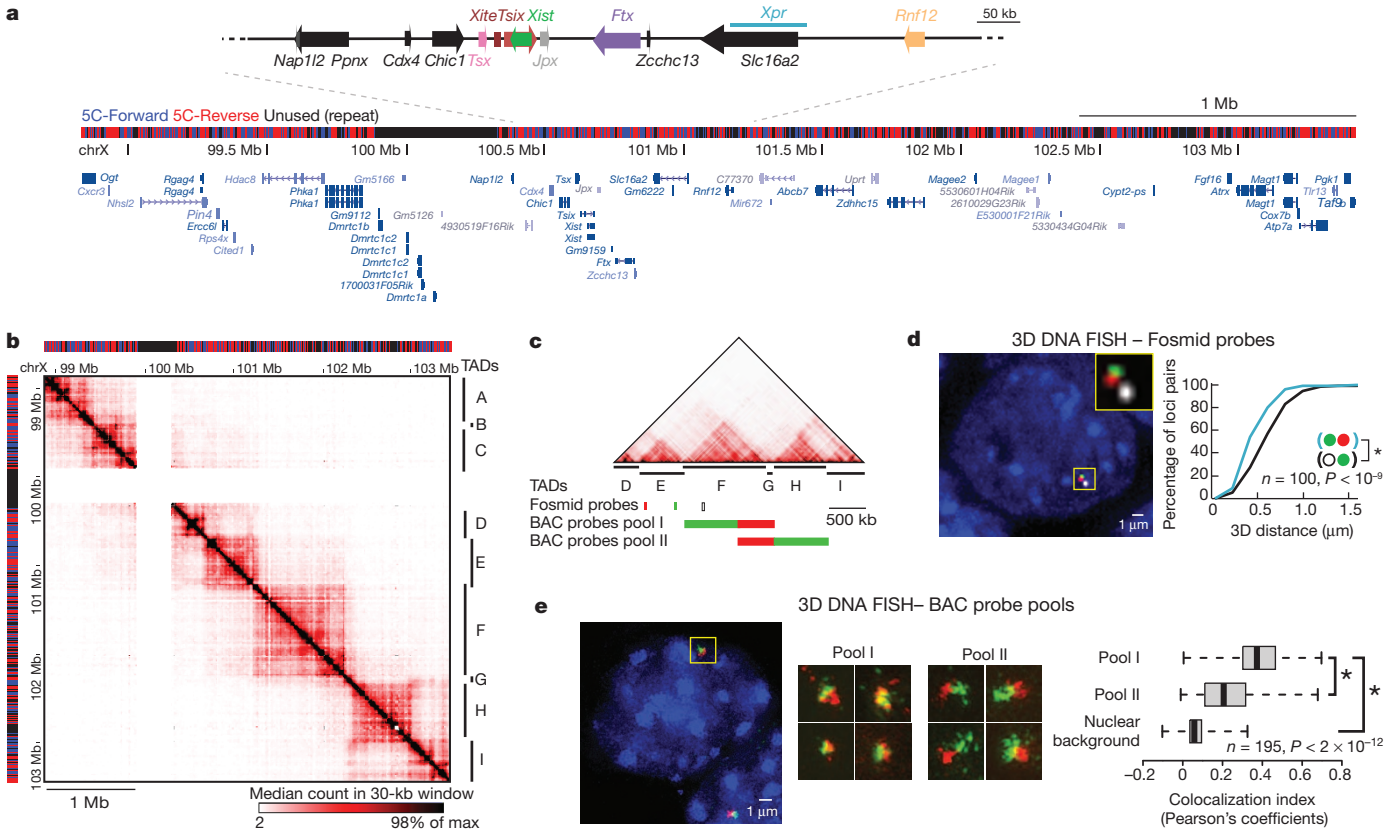

Figure 1 | Chromosome partitioning into topologically associating domains (TADs). **a**, Distribution of 5C-Forward and 5C-Reverse HindIII restriction fragments across the 4.5 Mb analysed showing positions of RefSeq genes and known XCI regulatory loci. **b**, 5C data sets from XY undifferentiated mouse ESCs (E14), displaying median counts in 30-kb windows every 6 kb. Chromosomal contacts are organized into discrete genomic blocks (TADs A–I). A region containing segmental duplications excluded from the 5C analysis is masked (white). **c**, Positions of DNA FISH probes. **d**, Interphase nuclear distances are smaller for probes in the same 5C domain. **e**, Structured illumination microscopy reveals that colocalization of neighbouring sequences is greater when they belong to the same 5C domain. Boxplots show the distribution of Pearson's correlation coefficient between red and green channels, with whiskers and boxes encompassing all and 50% of values, respectively; central bars denote the median correlation coefficient. Statistical significance was assessed using Wilcoxon's rank sum test.

positioning. Furthermore, the fact that the two neighbouring domains do not merge completely in ΔXTX cells (Fig. 2b) implies that additional elements, within TADs, can act as relays when a main boundary is removed. The factors underlying an element's capacity to act as a canonical or shadow boundary remain to be investigated.

Next we asked whether TAD organization changes during differentiation or XCI. Both male neuronal progenitors cells (NPCs) and male primary mouse embryonic fibroblasts (MEFs) show similar organization to mouse ESCs, with no obvious change in TAD positioning. However, consistent differences in the internal contacts within TADs were observed (Fig. 3a, Supplementary Figs 2 and 5). Noticeably, some TADs were found to become lamina-associated domains[19] (LADs) at certain developmental stages (Fig. 3b). Thus chromosome segmentation into TADs reveals a modular framework where changes in chromatin structure or nuclear positioning can occur in a domain-wide fashion during development.

We then assessed TAD organization on the inactive X, by combining *Xist* RNA FISH, to identify the inactive X, and super-resolution DNA FISH using BAC probe pools on female MEFs. We found that colocalization indices on the inactive X were still higher for sequences belonging to the same TAD than for neighbouring TADs (Supplementary Fig. 6a). However, the difference was significantly lower for the inactive X than for the active X. Deconvolution of the respective contributions of the active X and inactive X in 5C data from female MEFs (see Supplementary Methods and Supplementary Fig. 6) similarly revealed that global organization in TADs remains on the inactive X, albeit in a much attenuated form, but that specific long-range contacts within TADs are lost. This, together with a recent report focused on longer-range interactions[20], suggests that the inactive X has a more random chromosomal organization than its active homologue, even below the megabase scale.

We next investigated how TAD organization relates to gene expression dynamics during early differentiation. A transcriptome analysis, consisting of microarray measurements at 17 time points over the first 84 h of female mouse ESC differentiation was performed (Fig. 4a). During this time window, most genes in the 5C region were either up- or downregulated. Statistical analysis demonstrated that expression profiles of genes with promoters located within the same TAD are correlated (Fig. 4b). This correlation (median correlation coefficient cc of 0.40) is significantly higher than for genes in different domains (cc of 0.03, $P < 10^{-9}$) or for genes across the X chromosome in randomly selected, TAD-size regions (cc of 0.09, $P < 10^{-7}$). The observed correlations within TADs seem not to depend on distance between genes, and are thus distinct from previously described correlations between neighbouring genes[21] that decay on a length scale of approximately 100 kb (Supplementary Fig. 7). Our findings indicate that physical clustering within TADs may be used to coordinate gene expression patterns during development. Furthermore, deletion of the boundary between *Xist* and *Tsix* in ΔXTX cells was accompanied by long-range transcriptional misregulation (Supplementary Fig. 8), underlining the role that chromosome partitioning into TADs can play in long-range transcriptional control.

A more detailed analysis of each domain (Supplementary Fig. 7) revealed that co-expression is particularly pronounced in TADs D, E

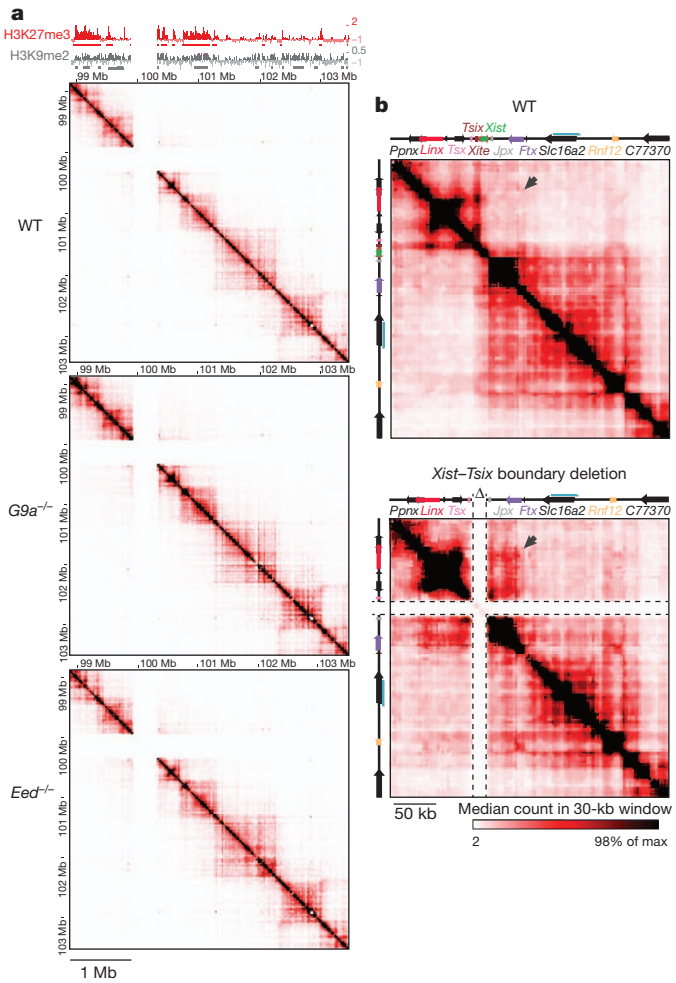

Figure 2 | Determinants of topologically associating domains. **a**, Blocks of contiguous enrichment in H3K27me3 or H3K9me2 (ref. 11) align with the position of TADs (chromatin immunoprecipitation on chip from ref. 9) in wild-type cells (TT2), but TADs are largely unaffected in the absence of H3K9me2 in male *G9a*$^{-/-}$ cells or H3K27me3 in male *Eed*$^{-/-}$ cells. **b**, Deletion of a boundary at *Xist/Tsix* disrupts folding pattern of the two neighbouring TADs.

and F (Fig. 4b, c). Although correlations are strongest within TADs, there is some correlation between TADs showing the same trend, such as TADs D and F, which are both downregulated during differentiation. Only TAD E, which contains *Xist* and all of its known positive regulators *Jpx*, *Ftx*, *Xpr/Xpct* and *Rnf12*[5] (*Jpx*, *Ftx*, *Xpct* and *Rnf12* are also known as *Enox*, *B230206F22Rik*, *Slc16a2* and *Rlim*, respectively) is anti-correlated with most other genes in the 4.5 Mb region, being upregulated during differentiation (Supplementary Fig. 7). The fact that these coordinately upregulated loci are located in the same TAD suggests that they are integrated into a similar *cis*-regulatory network, potentially sharing common *cis*-regulatory elements. We therefore predict that TAD E (~550 kb) represents the minimum 5′ regulatory region required for accurate *Xist* expression, explaining why even the largest transgenes tested so far (covering 150 kb 5′ to *Xist*, Fig. 5a) cannot recapitulate normal *Xist* expression[7].

The respective promoters of *Xist* and *Tsix* lie in two neighbouring TADs with transcription crossing the intervening boundary (Fig. 2b), consistent with previous 3C experiments[22]. Whereas the *Xist* promoter and its positive regulators are located in TAD E, the promoter of its antisense repressor, *Tsix*, lies in TAD D, which extends up to *Ppnx* (also known as *4930519F16Rik*)/*Nap1l2*, more than 200 kb away (Fig. 2b). Thus, in addition to the *Xite* enhancer, more distant elements within TAD D may participate in *Tsix* regulation. To test this we used two different single-copy transgenic mouse lines, Tg53 and Tg80 (ref. 23). Both transgenes contain *Xist*, *Tsix* and *Xite* (Fig. 5a). Tg53 encompasses the whole of TAD D, whereas Tg80 is truncated just 5′ to *Xite* (Fig. 5a and Supplementary Fig. 9). In the inner cell mass of male mouse embryos at embryonic day 4.0 (E4.0), *Tsix* transcripts could be readily detected from Tg53, as well as from the endogenous X (Fig. 5b). However, no *Tsix* expression could be detected from Tg80, which lacks the distal portion of TAD D (Fig. 5b). Thus, sequences within TAD D must contain essential elements for the correct developmental regulation of *Tsix*.

Within TAD D, several significant looping events involving the *Tsix* promoter or its enhancer *Xite* were detected (Figs 2b and 5a, Supplementary Fig. 10). Alignment of 5C maps with chromatin signatures of enhancers in mouse ESCs (Supplementary Fig. 11) suggested the existence of multiple regulatory elements within this region. We also identified a transcript initiating approximately 50 kb upstream of the *Ppnx* promoter (Fig. 5a), from a region bound by pluripotency factors and corresponding to a predicted promoter for a large (80 kb) intervening non-coding RNA (lincRNA[24], Supplementary Fig. 12) which we termed *Linx* (large intervening transcript in the *Xic*). *Linx* RNA shares several features with non-coding RNAs, such as accumulation around its transcription site[25] (Fig. 5c), nuclear enrichment and abundance of the unspliced form[26] (Supplementary Fig. 12 and 13). *Linx* and *Tsix* are co-expressed in the inner cell mass of blastocysts from E3.5–4.0 onwards, as well as in male and female mouse ESCs (Fig. 5c). *Linx* RNA is not detected earlier in embryogenesis, nor in extra-embryonic lineages, implying an epiblast-specific function

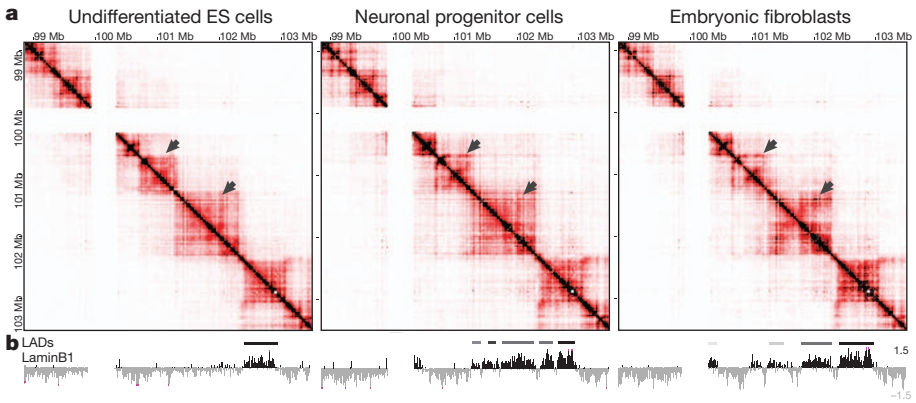

Figure 3 | Dynamics of topologically associating domains during cell differentiation. **a**, Comparison of 5C data from male mouse ESCs (E14), NPCs (E14) and primary MEFs reveals general conservation of TAD positions during differentiation, but differences in their internal organization (arrows highlight examples of tissue-specific patterns). **b**, Lamina-associated domains (LADs, from ref. 19) align with TADs. Chromosomal positions of tissue-specific LADs reflect gain of lamina association by TADs, as well as internal reorganization of lamina-associated TADs during differentiation.

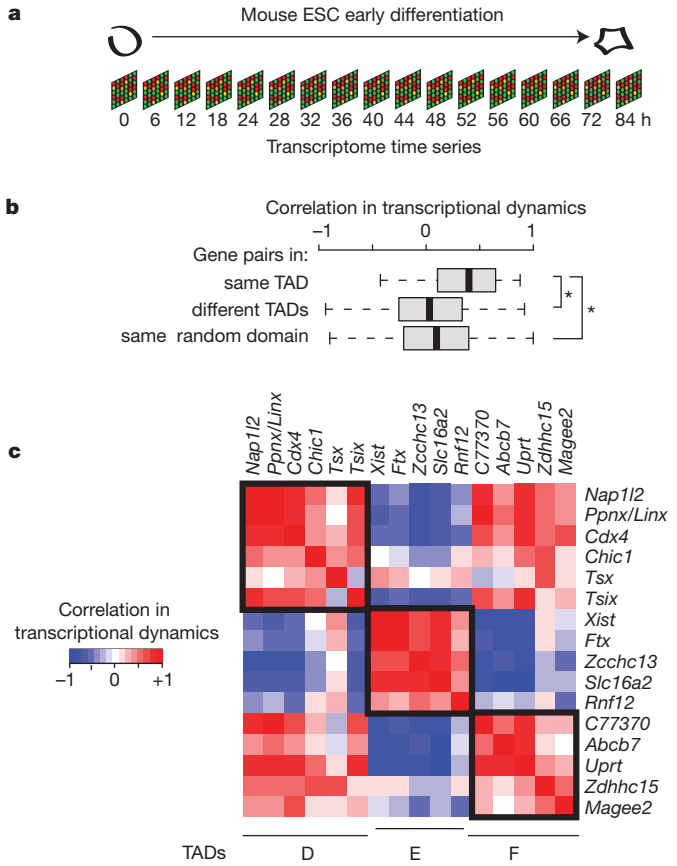

Figure 4 | Transcriptional co-regulation within topologically associating domains. **a**, Female mouse ESCs were differentiated towards the epiblast stem cell lineage for 84 h. Transcript levels were measured every 4–6 h at 17 different time points by microarray analysis. **b**, Pearson's correlation coefficients over all time points were calculated for gene pairs lying in the same TAD, pairs in different TADs and for pairs in randomly defined domains on the X chromosome that contain a similar number of genes and are of comparable size. Boxplots show the distribution of Pearson's correlation coefficients, with whiskers and boxes encompassing all and 50% of values, respectively, and central bars denoting the median correlation coefficient. * represents significant difference with $P < 10^{-7}$ using Wilcoxon's rank sum test. **c**, Pearson's correlation coefficients for gene pairs in TADs D, E and F with red denoting positive and blue negative correlation. Boxes indicate the TAD boundaries.

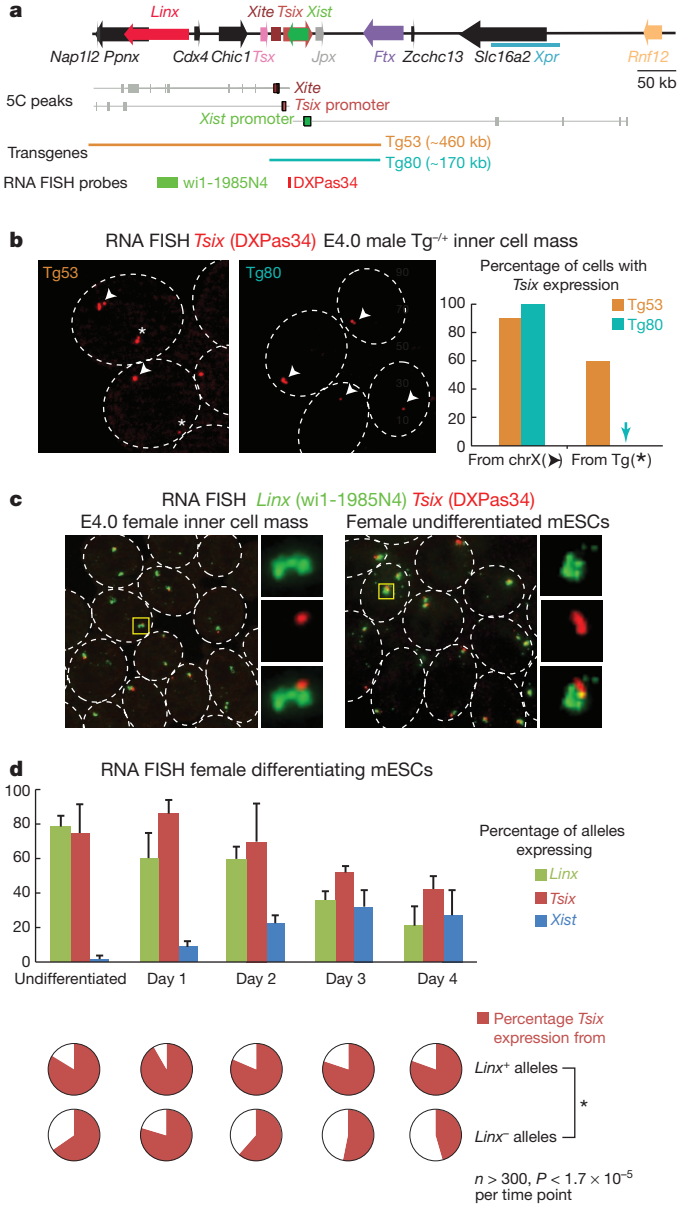

Figure 5 | 5C maps reveal new regulatory regions in the *Xic*. **a**, Statistically significant looping events (5C peaks) for restriction fragments within *Xite*, *Tsix* promoter or *Xist* promoter within their respective TAD, in male (E14) mouse ESCs. The Tg80 YAC transgene lacks genomic elements found to interact physically with *Xite/Tsix* that are present in Tg53. **b**, RNA FISH analysis of *Tsix* expression is detected in the inner cell masses of heterozygous transgenic male E4.0 embryos by RNA FISH from single-copy paternally inherited Tg53 but not Tg80 transgenes. Transgenic (star) and endogenous *Tsix* alleles (arrowhead) were discriminated by subsequent DNA FISH as in Supplementary Fig. 5. $n = 20$ inner cell mass cells (two embryos each). **c**, *Linx* transcripts (green, wi1-1985N4 probe) are expressed in both E4.0 inner cell mass cells and mouse ESCs, together with *Tsix* (red, DXPas34 probe), and unspliced transcripts accumulate locally in a characteristic cloud-like shape. **d**, RNA FISH in differentiating female mouse ESCs revealing synchronous downregulation of *Linx* and *Tsix* with concomitant upregulation of *Xist* (detected with a strand-specific probe). Bars are the standard deviation around the mean of three experiments. Triple-colour RNA FISH allows simultaneous detection of *Linx*, *Tsix* and *Xist* RNAs. Scoring of *Xist*-negative alleles demonstrates that before *Xist* upregulation *Tsix* expression is more frequent from *Linx*-expressing alleles than from *Linx* non-expressing alleles, at all time points tested. Presented is the mean of three experiments. Statistical differences were assessed using Fisher's exact test. Cells were differentiated in monolayers by withdrawal of leukaemia inhibitory factor (LIF).

(Supplementary Fig. 9). Triple RNA FISH for *Linx*, *Tsix* and *Xist* in differentiating female mouse ESCs (Supplementary Fig. 14) revealed that before *Xist* upregulation, the probability of *Tsix* expression from alleles co-expressing *Linx* is significantly higher than from alleles that do not express *Linx* (Fig. 5d). Furthermore, *Linx* expression is frequently monoallelic, even before *Xist* upregulation (Supplementary Fig. 14), revealing a transcriptional asymmetry of the two *Xic* alleles before XCI. Taken together, our experiments based on 5C, transgenesis and RNA FISH, point towards a role for *Linx* in the long-range transcriptional regulation of *Tsix* — either through its chromosomal association with *Xite* and/or via the RNA it produces. This analysis of the *Xist/Tsix* region illustrates how spatial compartmentalization of chromosomal neighbourhoods in TADs partitions the *Xic* into two large regulatory domains, with opposite transcriptional fates (Supplementary Fig. 15).

In conclusion, our study reveals that sub-megabase folding of mammalian chromosomes results in the self-association of large chromosomal neighbourhoods in the three-dimensional space of the nucleus. The stability of such partitioning throughout differentiation, X inactivation and in cell lines with impaired histone-modifying machineries, indicates that this level of chromosomal organization may provide a basic framework onto which other domain-wide

features, such as lamina association and blocks of histone modification, can be dynamically overlaid. Our data also point to a role for TADs in shaping regulatory landscapes, by defining the extent of sequences that belong to the same regulatory neighbourhood. We anticipate that TADs may underlie regulatory domains previously proposed on the basis of functional and synteny conservation studies[27,28]. We believe that the principles we have revealed here will not be restricted to the *Xic*, as spatial partitioning of chromosomal neighbourhoods occurs throughout the genome of mouse and human[29], as well as *Drosophila*[30] and *E. coli*[31]. We have shown that TAD boundaries can have a critical role in high-order chromatin folding and proper long-range transcriptional control. Future work will clarify the mechanisms driving this level of chromosomal organization, and to what extent it generally contributes to transcriptional regulation. In summary, our study provides new insights into the *cis*-regulatory architecture of chromosomes that orchestrates transcriptional dynamics during development, and paves the way to dissecting the constellation of control elements of *Xist* and its regulators within the *Xic*.

METHODS SUMMARY

5C was performed on mouse ESCs, mouse NPCs and primary MEFs following a previously described protocol[2] with modifications, and sequenced on one lane of an Illumina GAIIx. RNA and DNA FISH were performed on mouse ESCs and inner cell masses extracted from pre-implantation embryos as previously described[7], with modifications. Full experimental and bioinformatic methods are detailed in Supplementary Information.

Received 3 October 2011; accepted 22 March 2012.
Published online 11 April 2012.

1. Kleinjan, D. A. & Lettice, L. A. Long-range gene control and genetic disease. *Adv. Genet.* **61,** 339–388 (2008).
2. Dostie, J. *et al.* Chromosome conformation capture carbon copy (5C): a massively parallel solution for mapping interactions between genomic elements. *Genome Res.* **16,** 1299–1309 (2006).
3. Rastan, S. Non-random X-chromosome inactivation in mouse X-autosome translocation embryos–location of the inactivation centre. *J. Embryol. Exp. Morphol.* **78,** 1–22 (1983).
4. Rastan, S. & Robertson, E. J. X-chromosome deletions in embryo-derived (EK) cell lines associated with lack of X-chromosome inactivation. *J. Embryol. Exp. Morphol.* **90,** 379–388 (1985).
5. Augui, S., Nora, E. P. & Heard, E. Regulation of X-chromosome inactivation by the X-inactivation centre. *Nature Rev. Genet.* **12,** 429–442 (2011).
6. Anguera, M. C. *et al. Tsx* produces a long noncoding RNA and has general functions in the germline, stem cells, and brain. *PLoS Genet.* **7,** e1002248 (2011).
7. Heard, E., Mongelard, F., Arnaud, D. & Avner, P. *Xist* yeast artificial chromosome transgenes function as X-inactivation centers only in multicopy arrays and not as single copies. *Mol. Cell. Biol.* **19,** 3156–3156 (1999).
8. Lieberman-Aiden, E. *et al.* Comprehensive mapping of long-range interactions reveals folding principles of the human genome. *Science* **326,** 289–293 (2009).
9. Marks, H. *et al.* High-resolution analysis of epigenetic changes associated with X inactivation. *Genome Res.* **19,** 1361–1373 (2009).
10. Pauler, F. M. *et al.* H3K27me3 forms BLOCs over silent genes and intergenic regions and specifies a histone banding pattern on a mouse autosomal chromosome. *Genome Res.* **19,** 221–233 (2009).
11. Wen, B., Wu, H., Shinkai, Y., Irizarry, R. A. & Feinberg, A. P. Large histone H3 lysine 9 dimethylated chromatin blocks distinguish differentiated from embryonic stem cells. *Nature Genet.* **41,** 246–250 (2009).
12. Lienert, F. *et al.* Genomic prevalence of heterochromatic H3K9me2 and transcription do not discriminate pluripotent from terminally differentiated cells. *PLoS Genet.* **7,** e1002090 (2011).
13. Hawkins, R. D. *et al.* Distinct epigenomic landscapes of pluripotent and lineage-committed human cells. *Cell Stem Cell* **6,** 479–491 (2010).
14. Rougeulle, C. *et al.* Differential histone H3 Lys-9 and Lys-27 methylation profiles on the X chromosome. *Mol. Cell. Biol.* **24,** 5475–5484 (2004).
15. Montgomery, N. D. *et al.* The murine polycomb group protein Eed is required for global histone H3 lysine-27 methylation. *Curr. Biol.* **15,** 942–947 (2005).
16. Monkhorst, K., Jonkers, I., Rentmeester, E., Grosveld, F. & Gribnau, J. X Inactivation counting and choice is a stochastic process: evidence for involvement of an X-linked activator. *Cell* **132,** 410–421 (2008).
17. Spencer, R. J. *et al.* A boundary element between *Tsix* and *Xist* binds the chromatin insulator Ctcf and contributes to initiation of X chromosome inactivation. *Genetics* CrossRef (2011).
18. Kagey, M. H. *et al.* Mediator and cohesin connect gene expression and chromatin architecture. *Nature* **467,** 430–435 (2010).
19. Peric-Hupkes, D. *et al.* Molecular maps of the reorganization of genome-nuclear lamina interactions during differentiation. *Mol. Cell* **38,** 603–613 (2010).
20. Splinter, E. *et al.* The inactive X chromosome adopts a unique three-dimensional conformation that is dependent on Xist RNA. *Genes Dev.* **25,** 1371–1383 (2011).
21. Caron, H. *et al.* The human transcriptome map: clustering of highly expressed genes in chromosomal domains. *Science* **291,** 1289–1292 (2001).
22. Tsai, C.-L., Rowntree, R. K., Cohen, D. E. & Lee, J. T. Higher order chromatin structure at the X-inactivation center via looping DNA. *Dev. Biol.* **319,** 416–425 (2008).
23. Heard, E. *et al.* Transgenic mice carrying an *Xist*-containing YAC. *Hum. Mol. Genet.* **5,** 441–450 (1996).
24. Guttman, M. *et al.* Chromatin signature reveals over a thousand highly conserved large non-coding RNAs in mammals. *Nature* **458,** 223–227 (2009).
25. Khalil, A. M. *et al.* Many human large intergenic noncoding RNAs associate with chromatin-modifying complexes and affect gene expression. *Proc. Natl Acad. Sci. USA* **106,** 11667–11672 (2009).
26. Seidl, C. I. M., Stricker, S. H. & Barlow, D. P. The imprinted *Air* ncRNA is an atypical RNAPII transcript that evades splicing and escapes nuclear export. *EMBO J.* **25,** 3565–3575 (2006).
27. Ruf, S. *et al.* Large-scale analysis of the regulatory architecture of the mouse genome with a transposon-associated sensor. *Nature Genet.* **43,** 379–386 (2011).
28. Kikuta, H. *et al.* Genomic regulatory blocks encompass multiple neighboring genes and maintain conserved synteny in vertebrates. *Genome Res.* **17,** 545–555 (2007).
29. Dixon, J. R. *et al.* Topological domains in mammalian genomes identified by analysis of chromatin interactions. *Nature* doi:10.1038/nature11082 (this issue).
30. Sexton, T. *et al.* Three-dimensional folding and functional organization principles of the *Drosophila* genome. *Cell* **148,** 458–472 (2012).
31. Mercier, R. *et al.* The MatP/*matS* site-specific system organizes the terminus region of the *E. coli* chromosome into a macrodomain. *Cell* **135,** 475–485 (2008).

Supplementary Information is linked to the online version of the paper at www.nature.com/nature.

Acknowledgements We thank T. Pollex and T. Forné for experimental help; the imaging facility PICTIBiSA@BDD for technical assistance, D. Gentien and C. Hego for microarray hybridizations. We thank K. Bernhard, F. Stewart and A. Smith for protocols and material for 2i culture and EpiSC differentiation. We are grateful to members of the E.H. laboratory for critical input. This work was funded by grants from the Ministère de la Recherche et de l'Enseignement Supérieur and the ARC (to E.P.N.); a HFSP Long term fellowship (LT000597/2010-L) (to E.G.S.). EU EpiGeneSys FP7 Network of Excellence no. 257082, the Fondation pour la Recherche Medicale, ANR, ERC Advanced Investigator award no. 250367 and EU FP7 SYBOSS grant no. 242129 (to E.H.). N.B. was supported by BMBF (FORSYS) and EMBO (fellowship ASTF 307-2011). J.D., B.R.L. and N.L.v.B. were supported by NIH (R01 HG003143) and a W. M. Keck Foundation Distinguished Young Scholar Award.

Author Contributions E.P.N. performed and analysed 3C, 5C, (RT–)qPCR, immunofluorescence, RNA and DNA FISH. B.R.L. and N.L.v.B. helped in the design and/or the analysis of 3C and 5C. L.G. performed 3C, FISH and 5C analysis. E.G.S. generated the time-course transcriptomic data, which was analysed by J.M. and N.B.; I.O. performed FISH on pre-implantation embryos. J.G. donated the XTX mouse ESC line. N.S. and E.B. helped in the epigenomic and 5C analyses. J.S. and T.P. set up OMX microscopy and analysis and T.P. performed structured illumination microscopy and image analysis. The manuscript was written by E.P.N., J.D. and E.H. with contribution from E.G.S. and input from all authors.

Author Information High-throughput data are deposited in Gene Expression Omnibus under accession number GSE35721 for all 5C experiments and GSE34243 for expression microarrays. Reprints and permissions information is available at www.nature.com/reprints. The authors declare no competing financial interests. Readers are welcome to comment on the online version of this article at www.nature.com/nature. Correspondence and requests for materials should be addressed to E.H. (edith.heard@curie.fr) or J.D. (job.dekker@umassmed.edu).