PAR-TERRA directs homologous sex chromosome pairing

Hsueh-Ping Chu1–3, John E Froberg1–3, Barry Kesner1–3, Hyun Jung Oh1–3, Fei Ji2,4, Ruslan Sadreyev2,4, Stefan F Pinter1–3,5 & Jeannie T Lee1–3

In mammals, homologous chromosomes rarely pair outside meiosis. One exception is the X chromosome, which transiently pairs during X-chromosome inactivation (XCI). How two chromosomes find each other in 3D space is not known. Here, we reveal a required interaction between the X-inactivation center (Xic) and the telomere in mouse embryonic stem (ES) cells. The subtelomeric, pseudoautosomal regions (PARs) of the two sex chromosomes (X and Y) also undergo pairing in both female and male cells. PARs transcribe a class of telomeric RNA, dubbed PAR-TERRA, which accounts for a vast majority of all TERRA transcripts. PAR-TERRA binds throughout the genome, including to the PAR and Xic. During X-chromosome pairing, PAR-TERRA anchors the Xic to the PAR, creating a ‘tetrad’ of pairwise homologous interactions (Xic–Xic, PAR–PAR, and Xic–PAR). Xic pairing occurs within the tetrad. Depleting PAR-TERRA abrogates pairing and blocks initiation of XCI, whereas autosomal PAR-TERRA induces ectopic pairing. We propose a ‘constrained diffusion model’ in which PAR-TERRA creates an interaction hub to guide Xic homology searching during XCI.

The mammalian genome is ubiquitously transcribed, and the ends of telomeres are no exception. The telomere produces a heterogeneous population of long noncoding RNAs known as TERRA1–3. TERRA contains the telomeric repeat sequence UUUAGGG and sequences unique to subtelomeric regions of each chromosome. TERRA biology has been of major interest, as it has been linked to many human diseases and forms an integral part of telomeric architecture4,5. Although TERRA research has focused exclusively on telomeres, only a fraction of detectable TERRA transcripts reside at telomeres6 and TERRA has a propensity to localize near the inactive X chromosome (Xi) of female cells2,3. While TERRA also concentrates next to the Y chromosome7, a functional link to sex chromosome biology has remained elusive. Until now, efforts to understand functional linkages have been hampered by difficulties in depleting TERRA, an incomplete understanding of TERRA’s transcriptional origin, and limited knowledge of its genomic targets.

Here we investigate TERRA’s potential link to XCI. XCI inactivates one X chromosome in mammalian females to balance gene expression with XY males. XCI is driven by the Xic via noncoding genes Xite, Tsix, and Xist7–11. During early development, a counting mechanism determines the X-chromosome number and initiates XCI in cells with two or more X chromosomes12. An allelic-choice mechanism then randomly selects one active and one inactive X (Xa and Xi). On the Xi elect, the Xist gene initiates chromosome-wide silencing, whereas Tsix antagonizes Xist action on the Xa elect. The switch from biallelic to monoallelic Tsix expression marks the onset of XCI and establishes allelic choice13,14. Interestingly, the transition of Tsix to a monoallelic state coincides with a homologous pairing event between two Xic alleles15–19.

A required 15-kb region within Tsix and Xite, the pairing center16,17, is thought to provide a platform on which biallelically bound transcription factors could be partitioned to one allele during pairing and, in this way, ensure that only one Tsix allele continues to be expressed17,18,20. Abolishing pairing results in the loss of the mutually exclusive designation of Xa and Xi, whereas ectopic pairing prevents initiation of XCI by preventing trans interactions between two Xic alleles15,17,21. Thus, pairing appears to be critical to counting and allelic choice.

X–X pairing is one of the only known examples of homologous pairing in the mammalian soma, although the emergence of other examples22–25 suggests that somatic pairing may be more prevalent than currently suspected. X–X pairing is an excellent model system for the investigation of pairing mechanisms, as it can be recapitulated by a genetically tractable system: mouse embryonic stem (ES) cells. In female mouse ES cells, X–X pairing is observed between days 1 and 4 of differentiation17,26, and contact occurs for <30 min16,19. Recent evidence suggests that Tsix and Xite RNAs co-transcriptionally recruit the chromosomal architectural protein CTCF, which in turn serves as interchromosomal ‘glue’16,27. How the two pairing centers find each other in 3D space is currently unknown. Here we investigate this question and uncover a central role for the PAR through the RNA tether that we call PAR-TERRA.

RESULTS

Telomeric pairing of sex chromosomes in mouse ES cells

In a previous study, we noted that a significant number of female ES cells show a coalescence of X-chromosome telomeric ends early in cell differentiation3. Because Xic–Xic pairing is one of the earliest measurable
events in differentiating female cells, this prior observation led us to question whether telomeric agglomeration may be related to Xic pairing. We performed 3D DNA fluorescence in situ hybridization (FISH) assays and measured interallelic distances for the Xic (Xist and Tsix), the telomere, and various unrelated X-linked loci (Fig. 1a,b and Supplementary Fig. 1a–c). For all pairing assays, we simultaneously tracked at least two X-linked loci (labeled with different fluorophores) in order to avoid scoring sister chromatids and ensure inclusion of only distinct alleles. To track X-linked telomeres, we labeled distal chromosome X (chrX) BAC probes RP23-461E16, containing Arhgap6, and RP24-500I4, mapping within the PAR. The PAR is a subtelomeric region shared by chrX and chrY28,29. To isolate PAR-specific probes, we generated unique PCR fragments from RP24-500I4 to create a cocktail containing the fragments P3, P4, P5, P6, and P8 that specifically identifies the ends of both sex chromosomes (P34568; Supplementary Fig. 1d–f). We plotted interallelic distances as normalized distance (ND, normalized to nuclear diameter) between 0.0 and 1.0 and applied the Kolmogorov–Smirnov (KS) test to assess statistical significance (Supplementary Fig. 1g). An increase in the number of nuclei with interallelic distances <0.1 ND (<1 µm) is broadly defined as ‘paired’16–19,26.

We observed a significant reduction in average interchromosomal distances between days 0 and 4 (d0 and d4) not only at the Xics but also at the distal ends of chrX. A leftward shift in the full distribution and an overall increase in the number of allelic pairs in the 0.0–0.1 ND bin indicated allelic clustering of Xics and telomeres. By contrast, few
Figure 2. PAR transcripts produced by the sex chromosomes. (a) The pseudoautosomal region (PAR) at the distal ends of chrX and chrY. Dotted purple lines indicate that this region is incompletely sequenced and assembled in the current genome assemblies (mm9, mm10). Mid1 and Erdr1 have repeated fragments within PAR (brown and purple triangles, respectively). The telomeric repeats (red bars) are present within PAR. PAR BAC clones: 15 kb RP24-143B12 and the ~146 kb RP24-500I4. (b) Quantitation of percent overlapping TERRA and PAR RNA signals for the experiments in c. Three biological replicates (independent cell cultures) showed similar results. (c) Two-color RNA FISH detecting TERRA (Alexa488, green) and PAR transcripts (BAC probes, Cy3, red) in ES cells. (d) Top: map of sub-BAC probes and PCR amplicons. Left: northern blot analysis of PAR-TERRA RNA using either TERRA or PAR-36k oligo probes in ES cells on different differentiation days. GAPDH, loading control. Right: primer extension using an antisense TERRA oligo probe with PCR amplification using PAR-specific primer pairs located at 33, 36, and 39 k (kb) from the end of BAC RP24-500I4. +, with RT; −, without RT. Uncropped gel images are shown in Supplementary Data Set 1. (e) RNA FISH indicating colocalization of TERRA and PAR signals at both large and small foci in ES cells. Three-color RNA FISH (top row): TERRA oligo probe (cyan), PAR-specific probe, 47k (green) and 29k (red). Two-color RNA FISH (bottom row): TERRA oligo probe (green), PAR-specific probe, 31k (cyan). DAPI (blue) for nuclear stain. Right graph, quantitation of colocalization, reported as percent. (f) IGV screenshots of TERRA-capture RNA-seq experiments show deduced PAR-TERRA transcription start sites. PAR transcripts are linked to telomeric repeat RNA. RT was conducted with TERRA-specific primers versus random hexamers, as indicated.
Hprt alleles fell in the 0.0–0.1 ND range. Because pairing is transient (t_{1/2} < 30 min) and only a few percent of asynchronously differentiating ES cells are paired in a single snapshot, we focused on nuclei in the first decile—the 10% with the shortest interallelic distances, or ‘top 10%’ (Fig. 1b). Within this fraction, the transition from d0 to d4 resulted in a decrease in allelic Xic and (Fig. 1c and Supplementary Fig. 1a–c,g). Clustering was most dramatic from the PAR probe RP24-500H (Fig. 1c,d and Supplementary Fig. 1a,f,h). At d4, ~14% of PAR–PAR measurements fell within the 0.0–0.1 ND bin, and the top 10% of nuclei showed a major reduction of interallelic distances (Supplementary Fig. 1h). Thus, like the Xics, telomeric ends of chrX undergo transient allelic pairing.

Given that chrY also has a PAR, we asked whether PAR pairing also occurs in differentiating male ES cells. Strikingly, inter-PAR distances decreased in male ES cells (Fig. 1c,d and Supplementary Fig. 1h), indicating that the X and Y PARs indeed pair. In agreement, X- and Y-painting probes also showed transient clustering of chrX and chrY, but they came together without merging (Fig. 1e). By performing RNA FISH on the same nuclei in order to detect TERRA3, we observed a merging of two TERRA signals (‘two dots’) into one (‘one dot’; dots unresolvable by light microscopy) (Fig. 1f), suggesting that X–Y pairing is confined to the PAR. Among the first decile, the number of nuclei with X–Y distances of <0.1 ND increased significantly (Fig. 1f). We conclude that sex chromosomes (X and Y) undergo transient PAR–PAR pairing during the same time window as Xic–Xic pairing.

Identification of sex-linked subtelomeric transcripts: PAR-TERRA

Our previous work had revealed a large telomeric RNA cluster next to sex chromosomes2. By performing RNA and DNA FISH using TERRA oligo probes, we confirmed two prominent TERRA foci at

Hprt

Par

Terra

Hprt

Par

Terra

Hprt

Par

Terra

Hprt

Par

Terra

Hprt

Par

Terra

Hprt

Par

Terra

Hprt

Par

Terra

Hprt

Par

Terra

Hprt

Par

Terra

Hprt

Par

Terra

Hprt

Par

Terra

Hprt

Par

Terra

Hprt

Par

Terra

Hprt

Par

Terra

Hprt

Par

Terra

Hprt

Par

Terra

Hprt

Par

Terra

Hprt

Par

Terra

Hprt

Par

Terra

Hprt

Par

Terra

Hprt

Par

Terra

Hprt

Par

Terra

Hprt

Par

Terra

Hprt

Par

Terra

Hprt

Par

Terra

Hprt

Par

Terra

Hprt

Par

Terra

Hprt

Par

Terra

Hprt

Par

Terra

Hprt

Par

Terra

Hprt

Par

Terra

Hprt

Par

Terra

Hprt

Par

Terra

Hprt

Par

Terra

Hprt

Par

Terra

Hprt

Par

Terra

Hprt

Par

Terra

Hprt

Par

Terra

Hprt

Par

Terra

Hprt

Par

Terra

Hprt

Par

Terra

Hprt

Par

Terra

Hprt

Par

Terra

Hprt

Par

Terra

Hprt

Par

Terra

Hprt

Par

Terra

Hprt

Par

Terra

Hprt

Par

Terra

Hprt

Par

Terra

Hprt

Par

Terra

Hprt

Par

Terra

Hprt

Par

Terra

Hprt

Par

Terra

Hprt

Par

Terra

Hprt

Par

Terra

Hprt

Par

Terra

Hprt

Par

Terra

Hprt

Par

Terra

Hprt

Par

Terra

Hprt

Par

Terra

Hprt

Par

Terra

Hprt

Par

Terra

Hprt

Par

Terra

Hprt

Par

Terra

Hprt

Par

Terra

Hprt

Par

Terra

Hprt

Par

Terra

Hprt

Par

Terra

Hprt

Par

Terra

Hprt

Par

Terra

Hprt

Par

Terra

Hprt

Par

Terra

Hprt

Par

Terra

Hprt

Par

Terra

Hprt

Par

Terra

Hprt

Par

Terra

Hprt

Par

Terra

Hprt

Par

Terra

Hprt

Par

Terra

Hprt

Par

Terra

Hprt

Par

Terra

Hprt

Par

Terra

Hprt

Par

Terra

Hprt

Par

Terra

Hprt

Par

Terra

Hprt

Par

Terra

Hprt

Par

Terra

Hprt

Par

Terra

Hprt

Par

Terra

Hprt

Par

Terra

Hprt

Par

Terra

Hprt

Par

Terra

Hprt

Par

Terra

Hprt

Par

Terra

Hprt

Par

Terra

Hprt

Par

Terra

Hprt

Par

Terra

Hprt

Par

Terra

Hprt

Par

Terra

Hprt

Par

Terra

Hprt

Par

Terra

Hprt

Par

Terra

Hprt

Par

Terra

Hprt

Par

Terra

Hprt

Par

Terra

Hprt

Par

Terra

Hprt

Par

Terra

Hprt

Par

Terra

Hprt

Par

Terra

Hprt

Par

Terra

Hprt

Par

Terra

Hprt

Par

Terra

Hprt

Par

Terra

Hprt

Par

Terra

Hprt

Par

Terra

Hprt

Par

Terra

Hprt

Par

Terra

Hprt

Par

Terra

Hprt

Par

Terra

Hprt

Par

Terra

Hprt

Par

Terra

Hprt

Par

Terra

Hprt

Par

Terra

Hprt

Par

Terra

Hprt

Par

Terra

Hprt

Par

Terra

Hprt

Par

Terra

Hprt

Par

Terra

Hprt

Par

Terra

Hprt

Par

Terra

Hprt

Par

Terra

Hprt

Par

Terra

Hprt

Par

Terra

Hprt

Par

Terra

Hprt

Par

Terra

Hprt

Par

Terra

Hprt

Par

Terra

Hprt

Par

Terra

Hprt

Par

Terra

Hprt

Par

Terra

Hprt

Par

Terra

Hprt

Par

Terra

Hprt

Par

Terra

Hprt

Par

Terra

Hprt

Par

Terra

Hprt

Par

Terra

Hprt

Par

Terra

Hprt

Par

Terra

Hprt

Par

Terra

Hprt

Par

Terra

Hprt

Par

Terra

Hprt

Par

Terra

Hprt

Par

Terra

Hprt

Par

Terra

Hprt

Par

Terra

Hprt

Par

Terra

Hprt

Par

Terra

Hprt

Par

Terra

Hprt

Par

Terra

Hprt

Par

Terra

Hprt

Par

Terra

Hprt

Par

Terra

Hprt

Par

Terra

Hprt

Par

Terra

Hprt

Par

Terra

Hprt

Par

Terra

Hprt

Par

Terra

Hprit
distal chrX telomeres, colocalizing with PAR DNA, in female ES cells (Supplementary Fig. 1f). Additionally, we noticed that RNA signals generated by PAR hybridization greatly resembled TERRA signals (Fig. 2a–e), suggesting that the PAR is transcribed and that its transcript accumulates. This intriguing relationship led us to explore a possible structural and functional linkage between the PAR and TERRA. Although previous cytological analysis showed TERRA localization to the ends of most if not all chromosomes1,2, it is unclear where TERRA originates. TERRA may be transcribed by all telomeres and retained in cis or transcribed by select loci and disseminated in trans to other sites. Results from one study suggest TERRA originates predominantly in chr18 (ref. 30). Because subtelomeric sequences of most mouse chromosomes have not been fully assembled, additional origins could have escaped detection. Because the PAR is contiguous to the telomere and resembles TERRA signals, we asked whether the PAR and TERRA may be one long sex-linked telomeric transcript.

The PAR comprises the coding genes Mid1, Erdr1, and Asnmt and a number of highly repetitive sequences, including internal telomeric (TTAGGG)n repeats and various intronic and exonic repeats of Mid1 and Erdr1 (Fig. 2a). Owing to the region’s repetitive nature, the current genome assembly remains incomplete here. RNA FISH using PAR BAC probes, RP24-143B12 and RP24-500I4, showed large foci that nearly perfectly overlapped TERRA signals in both male and female ES cells (Fig. 2a–c). Northern blot analyses using various PAR probes (Fig. 2d, left, and Supplementary Fig. 2a) showed that each detected a heterogeneous population dominated by a high-molecular-weight band of >9 kb, in a pattern that resembled that of TERRA1,30. PAR patterns revealed by probes 27k, 29k, 34k, 36k, and 47k (all unique to the PAR; “k” represents kb from the end of the BAC RP24-500I4) were similar and remained consistent during cell differentiation (Fig. 2d, left, and Supplementary Fig. 2a). RT primer extension using an antisense TERRA oligo yielded positive PCR amplification with PAR-specific primer pairs at 33, 36, and 39 kb (Fig. 2d, right). Notably, RNA FISH showed that PAR and TERRA signals colocalized with each other not only at sex chromosomes but also within finer speckles throughout the nucleus (>90%, Fig. 2e). Thus, PAR and TERRA show striking resemblance as viewed by means of orthogonal assays.

To investigate the possibility that PAR is continuous with TERRA, we conducted TERRA-capture RNA-seq. We first captured TERRA transcripts from total ES cell RNA using biotinylated oligos complementary to (UUAGGG)n repeats and various intronic and exonic repeats of Mid1 and Erdr1 (Fig. 2a). To map PAR-TERRA chromatin-binding sites, we performed CHIRT6, a hybrid ChiRIP3 and CHART4 method optimized for PAR-TERRA. Because capture probes could potentially interact with DNA rather than RNA, we included an RNase H elution step (Supplementary Fig. 3a). Several capture-probe sets were designed: (i) TERRA antisense (TERRA-AS), to capture transcripts containing UUAGGG, (ii) PAR, to capture PAR-containing transcripts (Supplementary Fig. 3a), and (iii) TERRA sense (TERRA-S, the reverse complement), to control for strand specificity and further rule out DNA capture. Quantitative RT-PCR indicated that PAR-TERRA transcripts were specifically enriched relative to other nuclear RNAs using TERRA-AS probes (Supplementary Fig. 3b). PAR DNA was also enriched relative to Hprt in both PAR and TERRA-S capture (Fig. 3b). Enrichment was dependent on RNase H and abolished by RNase A (Fig. 3c), indicating that the pulldown was mediated by interaction between DNA capture probes and RNA targets. To exclude artifacts due to probe hybridization to genomic DNA rather than the intended RNA target, we sequenced two critical controls: an RNase H control in which RNase H was omitted in the elution step, which should preclude elution of RNA-dependent interactions, and a TERRA-S control, which would not hybridize to TERRA RNA but could pull down genomic DNA.

We performed CHIR T paired-end sequencing of d0, d3, and d7 female mouse ES samples and mouse embryonic fibroblasts (MEF). PAR-specific binding sites were called using the peak caller MACS and compared to TERRA sites6 with normalization to input library, TERRA-S library, or no-RNase-H library, with each normalization method yielding similar results in two biological replicates (Fig. 3d and Supplementary Fig. 3c,d). In total, 2,000–5,000 significant peaks were called using PAR and TERRA probes in ES cells (d0, d3, d7), whereas very few were called for the control TERRA-S probe (Table 1). In MEFs, the number of PAR and TERRA peaks was reduced to 400–500, correlating with a six-fold reduction in MEFs relative to ES cells (Supplementary Fig. 3e). There was considerable overlap of PAR CHIR T profiles with those of TERRA, with high Pearson’s r values in correlation plots for d0, d3, and d7 ES cells and MEFs but not for comparisons to sense-CHIR T controls (Fig. 3d and Table 1). There was also high correlation between cells of various differentiation states (Supplementary Fig. 4). In MEFs, consistent with a dramatic reduction in PAR-TERRA expression, the number of peaks shared between PAR and TERRA decreased. Interestingly, in all samples, there existed a set of binding sites unique to TERRA, possibly reflecting binding sites for autosomal TERRA (Table 1). These data supported two types
of telomeric transcripts—PAR-TERRA and TERRA—with PAR-TERRA being the dominant species in ES cells.

* cis*-regulatory element annotation system (CEAS) analysis indicated that, like TERRA, PAR was enriched for binding in noncoding space (Fig. 3e–g). The strongest enrichment occurred at subtelomeric and telomeric regions (Fig. 4a,b). At chr2, chr9, chr13, and chr18, for example, strong PAR and TERRA peaks dominated the subtelomeric landscape (Fig. 4b). Moreover, PAR and TERRA binding occurred together at non-telomeric-repeat regions (Figs. 3f and 4c). Binding was also observed at internal (TTAGGG)₅ repeats present throughout the genome. The highly similar binding profiles of PAR and TERRA further support the idea of a continuous transcript. Together, these data indicate that, like TERRA, X-linked telomeric transcript PAR-TERRA targets chromatin sites both in cis and in trans.

PAR-TERRA directs telomeric sex chromosome pairing

We asked whether subtelomeric transcription may aid in pairing. Interestingly, PAR-TERRA showed greatest density at discrete positions within *Mid1, Erdr1*, and *Asmt* (Fig. 4a,d). We depleted PAR-TERRA using locked nucleic acid gapmers (LNA) directed against either PAR or TERRA sequences (Fig. 5a). At 1–6 h post treatment, only 20–30% of PAR-TERRA remained by northern analysis (Fig. 5a). RNA FISH confirmed the disappearance of PAR-TERRA foci (Fig. 5b). Moreover, depletion of PAR-TERRA targets resulted in a comitant depletion of TERRA, consistent with the idea that the two transcripts are linked.

We then assessed pairing in *d4* female ES cells at 6 h after knockdown (KD). Indeed, PAR pairing was significantly disrupted relative to scrambled (Scr) KD controls (Fig. 5c,d). This was also the case in male ES cells (Supplementary Fig. 5). Thus, PAR-TERRA is required for PAR–PAR pairing. To examine whether PAR-TERRA is sufficient for pairing, we created ES cells carrying an autosomal PAR-TERRA BAC transgene (RP24-500I4; Fig. 1a). RNA FISH showed that the transgene copy number did not exceed that of the endogenous PAR (Supplementary Fig. 6a,b). RNA FISH demonstrated transgene expression of PAR-TERRA and linked P1 vector sequences (Fig. 5e). To assess whether the transgene could pair with PAR, we performed DNA FISH using a P1 and an *Arghap6* probe. Significantly, transgenic PAR-TERRA induced ectopic pairing with the distal end of chrX (Fig. 5e,f). No pairing was observed with the *Hprt* control (Fig. 5e, bottom, and f). By performing 3-color DNA FISH for PAR, *Arghap6–Mid1* (chrX_end), and *Sry* (Y linked) (Supplementary Fig. 6c,d), we observed a simultaneous disruption of X–Y PAR pairing (Supplementary Fig. 6e), suggesting a competitive interaction between PAR sequences in vivo. Thus, PAR is both necessary and sufficient to direct interchromosomal pairing between sex chromosome telomeres (Fig. 5g).

PAR-TERRA tethers the Xic to the PAR

We next investigated the relationship between PAR and Xic pairing. Interestingly, time-course CHIRT analysis revealed hotspots of PAR-TERRA binding in the 15-kb X-pairing center of *Tsx1* and *Xite* (Fig. 6a). Binding was specific to ES cells (which are pairing competent) and not seen in MEFs (which are not pairing competent). To examine whether PAR-TERRA affects Xic–Xic pairing, we depleted PAR-TERRA, confirmed loss of PAR pairing, and performed Xic-pairing assays (Fig. 6b). Significantly, depleting PAR-TERRA impacted inter-Xic pairing in *d4* female ES cells, with a substantial decrease in the number of nuclei exhibiting interallelic distances of <0.1 ND (Fig. 6b, *P* < 0.001, two-tailed Student’s *t* test). Therefore, both PAR–PAR and Xic–Xic pairing require PAR-TERRA.

The relationship between PAR and Xic intrigued us, as meiotic chromosome pairing appears to be driven by a telomeric bouquet. This conglomeration of telomeres has been proposed to facilitate initial synopsis at one chromosomal end and enable synaptic extension through a ‘zippering’ mechanism to the other end. However, because somatic X–X pairing does not involve the full chromosome, a telomere-initiated zippering mechanism seems unlikely to drive Xic–Xic pairing. In principle, a long-range looping mechanism could bring the Xic to the PAR which, when it paired, would constrain the random walk for homology searching between two Xic alleles. To test this idea, we measured PAR–Xic distances and observed a significant decrease on *d4*, but not on *d8* (Fig. 6c). To confirm using an unbiased sampling method, we employed 4C technology (Fig. 6d). This correlative methodology allowed for unbiased sampling and confirmed PAR–Xic interactions. The relationship between PAR and Xic was validated using 4C technology (Fig. 6d). The relationship between PAR and Xic further validated our initial hypothesis.

The tetrad

A scenario involving coincident pairwise interactions would predict the existence of a ‘tetrad’ formed by homotypic (Xic–Xic, PAR–PAR)
and heterotypic (Xic–PAR) interactions. In d4 female ES cells, 5.7% of nuclei harbored a tetrad, as defined by distances of all pairwise interactions occurring within <1.5 μm, and almost all Xic–Xic pairs were found within such tetrads (Fig. 7a). These tetrads almost always occurred at the nuclear edge, consistent with chromosomal ends being anchored to the nuclear envelope. By d8, the number of tetrads (TTAGGG) increased.

**Figure 4** CHIRT-seq: PAR-TERRA RNA binds in cis and in trans throughout the genome. (a) CHIRT-seq tracks representing PAR-TERRA enrichment at chromosomal ends in female ES cells and MEFs. PAR data are compared to TERRA data and are normalized to input (TERRA/input, PAR/Input), no-RNase-H control (TERRA/no RNase H), or the sense control (TERRA/sense). (b) PAR-TERRA enrichment in subtelomeric regions of multiple autosomes in female ES cells. Red bars, TTAGGG repeats. Pink bars, sequence gaps. (c) PAR-TERRA binds to internal chromosomal regions as well. (d) PAR-TERRA binds to pseudoautosomal regions of chrX and chrY.
Can PAR-PAR pairing occur without Xic–Xic pairing and, conversely, can Xic–Xic pairing occur without PAR–PAR pairing? In d4 ES cells, tetrads accounted for approximately half of all nuclei with PAR–PAR interactions. We conclude that X–X pairing occurs in a stepwise fashion, with PAR-TERRA tethering PAR to Xic in all pairwise interactions and thereby creating an interaction ‘hub’ (Fig. 7d).

Finally, we asked whether PAR-TERRA’s effects had consequences for XCI. Indeed, RNA FISH showed that depleting PAR-TERRA precluded formation of Xist clouds in differentiating female ES cells (Fig. 8a, b). Fewer TERRA KD cells demonstrated large foci of Xist RNA than Scr KD cells on d8 (P = 0.01, 17.8% (n = 219, TERRA KD) versus 29.1% (Scr KD, n = 168); Fig. 8b). The failure of XCI was supported by a larger fraction of d8 nuclei retaining a biallelic pattern of X-linked gene expression (Fig. 8c–e). Depleting either PAR or TERRA RNA resulted in the same effect on Xist upregulation (Fig. 8d) and Atrx silencing (Fig. 8e), with the KD of PAR having an especially strong effect. By contrast, treatment with either Scr or TERRA-S control LNAs did not adversely affect XCI. PAR KD also did not affect expression of autosomal TERRA (Supplementary Fig. 7).

We conclude that PAR-TERRA RNA is required for homologous X–X pairing and for proper XCI16,17,26.
Figure 6  Intrachromosomal interactions between PAR and Xic occur in a PAR-TERRA-dependent manner. (a) Time course analysis (d0, d3, d7 ES cells; MEFs) using PAR-TERRA CHIRT-seq reveals binding sites at the Xic-pairing center. (b) PAR-TERRA KD disrupts inter-Xic pairing in d4 female ES cells at 6 h post-transfection. Left: DNA FISH using probes to the Xic, PAR, and Arhgap6 in d4 female ES cells. Arhgap6 signals served as hybridization control and ensured scoring only nuclei with two discernible signals for each probe. Right: dotplot of interallelic distances shown for the top decile of nuclei. $P$ values were determined using two-tailed Student’s $t$ test. (c) Frequency of PAR–Xic association in differentiating female ES cells (d0, d4, d8). Full distributions of PAR–Xic distances are shown. $n$, cell number. Triangles indicate mean values. $P$ values determined by the KS test. (d) Heatmap of 4C analysis with a viewpoint at PAR $Erdr1$ in d0 and d4 female mESCs and MEFs. The heatmap represents the log mean coverage for a given window size (100 K–5 Mb sliding windows). Two biological replicates show the similar pattern of interactions. Significance of interaction between Xic and PAR was determined by fitting the observed empirical distribution of normalized contact data with Weibull distribution and calculating $P$ values. The asterisk indicates the PAR–Xic interaction in d4 ES cells. (e) Strong correlation between contact frequencies (blue track, 4C) and PAR-TERRA RNA binding (green tracks, CHIRT) in female ES cells undergoing XCI (d4). (f) Scatterplot: 4C analysis of differentiating d4 female ES cells showed strong correlation between PAR-TERRA binding (CHIRT) and PAR interaction frequency (4C). Pearson’s $r = 0.69$. Each dot represents the mean coverage of 100-kb bin size for 4C (x axis) and CHIRT-seq (y axis) in the scatterplot. (g) Frequency of PAR–Xic and Hprt–Xic associations in differentiating d4 female ES cells after Scr or PAR-TERRA knockdown for 6 h. Full distributions for each pairwise distance measurements are shown. $n$, cell number. $P$ values determined by the KS test.
facilitates Xic–Xic homology searching and pairing. The two pathways are not mutually exclusive. PAR-TERRA RNA is required for all pairwise interactions.

Figure 7 The tetrad as a hub for pairing interactions. (a) DNA FISH using probes to the Xic, PAR, and Arhgap6 in d4 female ES cells reveals a high frequency of tetrads. Mean percentages are plotted in graph, with sample sizes \( n = \text{cell number} \) and statistical analysis. \( P \) value determined using Fisher’s exact test. (b) DNA FISH shows that PAR-TERRA knockdown disrupts tetrad formation at 6 h post-transfection in d4 female ES cells. Three biological replicates (independent cell cultures) are averaged. \( n = \text{cell number} \). \( P \) value determined using Fisher’s exact test. (c) Cartoon shows possible pairing species and the prevalence (relative ratio) of each in d4 female ES cells. DNA-FISH images for two representative examples of each species are shown above the cartoon. Xic–Xic pairs without PAR interaction are not observed. (d) Cartoon shows possible pairing species and the prevalence (relative ratio) of each in d4 female ES cells. DNA-FISH images for two representative examples of each species are shown above the cartoon. Xic–Xic pairs without PAR interaction are not observed. (e) DNA-FISH using probes to the Xic, PAR, and Arhgap6 in d4 female ES cells reveals a high frequency of tetrads. Mean percentages are plotted in graph, with sample sizes \( n = \text{cell number} \) and statistical analysis. \( P \) value determined using Fisher’s exact test. (f) DNA FISH shows that PAR-TERRA knockdown disrupts tetrad formation at 6 h post-transfection in d4 female ES cells. Three biological replicates (independent cell cultures) are averaged. \( n = \text{cell number} \). \( P \) value determined using Fisher’s exact test. (g) Cartoon shows possible pairing species and the prevalence (relative ratio) of each in d4 female ES cells. DNA-FISH images for two representative examples of each species are shown above the cartoon. Xic–Xic pairs without PAR interaction are not observed. (h) Cartoon shows possible pairing species and the prevalence (relative ratio) of each in d4 female ES cells. DNA-FISH images for two representative examples of each species are shown above the cartoon. Xic–Xic pairs without PAR interaction are not observed.

**DISCUSSION**

Here we have (i) identified the subtelomeric region (PAR) of sex chromosomes as a major source of TERRA transcripts in ES cells, (ii) demonstrated that PAR-TERRA binds in cis and in trans throughout the genome, (iii) uncovered PAR-TERRA as a mediator of homologous X-chromosome pairing, and (iv) defined PAR–PAR interactions as a key to nucleating Xic–Xic pairing for subsequent initiation of XCI.

We propose a constrained diffusion model (Fig. 7d) in which telomeric pairing guides Xic–Xic pairing, and PAR-TERRA RNA serves as a tether to bring the PAR and Xic into contact within a tetrad. This interaction hub would serve to constrain the space for Xic–Xic interactions, thereby reducing the effective volume within which Xic alleles must engage in a random walk during homology searching. Our data show that all pairing occurs within the tetrad (Fig. 7). PAR transcription is key in establishing all pairwise interactions, both homotypic (PAR–PAR) and heterotypic (PAR–Xic). Two pathways can be envisioned. In pathway 1, PAR-TERRA directs PAR–PAR interactions. Independent binding of PAR-TERRA to the Xic enables PAR to interact with Xic, resulting in formation of PAR–Xic interactions, which in turn facilitate Xic–Xic homology searching and in trans pairing. Alternatively (pathway 2), PAR-TERRA directs PAR–Xic interactions to occur first in cis. PAR–PAR pairing would occur independently, and this in turn would facilitate Xic–Xic homology searching and in trans pairing. The models are not mutually exclusive. Nevertheless, additional factors must work with PAR-TERRA to control the timing of pairing, given that PAR-TERRA expression is constitutive. Our study suggests that a key difference between meiotic and somatic pairing is that, while both appear to be telomere-mediated, the latter does not apparently involve a whole-chromosome zipper mechanism.

The purpose of somatic X–X pairing has been actively debated. Our present data provide clear evidence for a role of PAR-TERRA and PAR and Xic pairing during the initiation of XCI and lend further support to the idea that allelic pairing mediates counting and allelic choice of one Xa and one X16,17. A recent study suggested that XCI can also occur in the absence of X–X pairing38, as heterokaryons made by fusing XX and XY cells occasionally resulted in XCI in the XY nucleus (where pairing was presumptively not possible). However, heterokaryons can continue to divide after fusion, in which case breakdown of the nuclear envelope could enable physical contact between X chromosomes as well as mixing of chromosomes. Alternatively, Xist dysregulation could occur in heterokaryons, akin to “chaotic XCI”21,39.

Our study shows that male cells also undergo PAR–PAR pairing. If pairing occurs for the express purpose of initiating XCI correctly, why does PAR pairing also occur in male cells? Although male cells do not initiate XCI, they nevertheless must count X chromosomes. In male cells, PAR–PAR pairing would not result in a productive Xic homology search due to the absence of a second X. Accordingly, XCI does not initiate.

An important sidebar to our discovery is that TERRA is predominantly expressed from sex chromosomes, with PAR-TERRA providing >99% of all (UUAGGG)\(_2\)-containing transcripts in mouse ES cells. There may be differences between species and cell types30,40,41. Still, one wonders why sex chromosomes produce so much TERRA. Facilitation
of homologous pairing might be only one of its sex-linked functions. Notably, somatic pairing also occurs at a number of autosomal loci, including Oct4, cytokine genes, and imprinted loci. It would be of interest to determine whether subtelomeric pairing and transcription also play a role in those contexts. Finally, our study shows that PAR-TERRA targets loci on autosomes as well. Indeed, it is now known that TERRA in general can operate in trans at nontelomeric loci. Henceforth, a major goal will be to understand whether sex-linked PAR-TERRA can influence genome-wide activity and what those activities might be.

METHODS

Methods, including statements of data availability and any associated accession codes and references, are available in the online version of the paper.

Note: Any Supplementary Information and Source Data files are available in the online version of the paper.

ACKNOWLEDGMENTS

We are grateful to Y. Jeon and H. Sunwoo for technical advice. This work was funded by a grant from the NIH (R01-GM58839) and Howard Hughes Medical Institute to J.T.L.

AUTHOR CONTRIBUTIONS

H.-P.C. and J.T.L. designed the experiments and analyzed data. H.-P.C. performed experiments, including FISH, CHIRT-seq, TERRA-capture RNA-seq and LNA knockdown. J.E.F. performed the 4C experiment. S.E.P. established the 4C protocol. H.-P.C. and H.J.O. optimized CHIRT protocols. F.J., R.S., B.K. and H.P.C. performed bioinformatics analyses. H.P.C. and J.T.L. wrote the manuscript.

COMPETING FINANCIAL INTERESTS

The authors declare no competing financial interests.
1. Azzalin, C.M., Reichenbach, P., Khoriauli, L., Giulotto, E. & Lingner, J. Telomeric repeat containing RNA and chromatin factors at mammalian chromosome ends. Science 318, 798–801 (2007).

2. Schoeftner, S. & Blasco, M.A. Developmentally regulated transcription of mammalian telomeres by DNA-dependent RNA polymerase II. Nat. Cell Biol. 10, 228–236 (2008).

3. Zhang, L.-F. Telomeric RNAs mark sex chromosomes in stem cells. Genetics 182, 685–698 (2009).

4. Chu, H.-P. Telomeric functions grounding on TERRA firma. Trends Cell Biol. 25, 29–36 (2015).

5. Maicher, A., Kastner, L., Dees, M. & Luke, B. Deregulated telomere transcription causes replication-dependent telomere shortening and promotes cellular senescence. Nucleic Acids Res. 40, 6649–6659 (2012).

6. Chu, H.-P. et al. TERRA RNA antagonizes ATRX and protects telomeres. Cell (in the press).

7. Starmer, J. & Magnuson, T. A new model for random X chromosome inactivation. Development 136, 1–10 (2009).

8. Distech, C.M. Dosage compensation of the sex chromosomes. Annu. Rev. Genet. 46, 537–560 (2012).

9. Lee, J.T. Gracefully ageing at 50, X-chromosome inactivation becomes a

10. Wutz, A. Gene silencing in X-chromosome inactivation: advances in understanding facultative heterochromatin formation. Nat. Rev. Genet. 12, 542–553 (2011).

11. van Bemmel, J.G., Mira-Bontenbal, H. & Gribnau, J. Cis- and trans-regulation in X inactivation. Chromosome 125, 41–50 (2016).

12. Sun, S. et al. Jpx RNA activates Xist by evicting CTCF. Cell 153, 1537–1551 (2013).

13. Lee, J.T., Davudov, L.S. & Warszawsky, D. Tsix, a gene antisense to Xist at the X-inactivation centre. Nat. Genet. 21, 400–404 (1999).

14. Brown, C.J. et al. The human XIST gene: analysis of a 17 kb inactive X-specific RNA that contains conserved repeats and is highly localized within the nucleus. Cell 71, 527–542 (1992).

15. Carrel, L. Molecular biology. “X”-rated chromosomal rendezvous. Science 311, 1107–1109 (2006).

16. Xu, N., Donohoe, M.E., Silva, S.S. & Lee, J.T. Evidence that homologous X-chromosome pairing requires transcription and Ctcf protein. Nat. Genet. 39, 1390–1396 (2007).

17. Xu, N., Tsai, C.L. & Lee, J.T. Transient homologous chromosome pairing marks the onset of X inactivation. Science 311, 1149–1152 (2006).

18. Donohoe, M.E., Silva, S.S., Prliter, S.F., Xu, N. & Lee, J.T. The pluripotency factor Oct4 interacts with Ctcf and also controls X-chromosome pairing and counting. Nature 460, 128–132 (2009).

19. Masui, O. et al. Live-cell chromosome dynamics and outcome of X chromosome pairing events during ES cell differentiation. Cell 145, 447–459 (2011).

20. Scialdone, A. & Nicolodi, M. Mechanics and dynamics of X-chromosome pairing at X inactivation. PLOS Comput. Biol. 4, e1000244 (2008).

21. Lee, J.T. Regulation of X-chromosome counting by Tsix and Xite sequences. Science 309, 768–771 (2005).

22. LaSalle, J.M. & Lalonde, M. Homologous association of oppositely imprinted chromosomal domains. Science 272, 725–728 (1996).

23. Brandt, V.L., Hewitt, S.L. & Skok, J.A. It takes two: communication between homologous alleles preserves genomic stability during VDJ recombination. Nucleic Acids 1, 23–29 (2010).

24. Hogan, M.S., Parfitt, D.E., Zepeda-Mendoza, C.J., Shen, M.M. & Spector, D.L. Transient pairing of homologous Oct4 alleles accompanies the onset of embryonic stem cell differentiation. Cell Stem Cell 16, 275–288 (2015).

25. Zorza, C.E. et al. Mysos VI regulates gene pairing and transcriptional pause release in T cells. Proc. Natl. Acad. Sci. USA 112, E1587–E1593 (2015).

26. Bacher, C.P. et al. Transient colocalization of X-inactivation centres accompanies the initiation of X inactivation. Nat. Cell Biol. 8, 293–299 (2006).

27. Kung, J.T. et al. Locus-specific targeting to the X chromosome revealed by the RNA interactome of CTCF. Mol. Cell 57, 361–375 (2015).

28. Berleth, J.B., Yang, F. & Distech, C.M. Escape from X inactivation in mice and humans. Genome Biol. 11, 213 (2010).

29. Carrel, L. & Willard, H.F. X-inactivation profile reveals extensive variability in X-linked gene expression in females. Nature 434, 400–404 (2005).

30. López de Silanes, I. et al. Identification of TERRA locus unveils a telomere protection role through association to nearly all chromosomes. Nat. Commun. 5, 4723 (2014).

31. Chu, C., Qu, K., Zhong, F.L., Artandi, S.E. & Chang, H.Y. Genomic maps of long noncoding RNA occupancy reveal principles of RNA-chromatin interactions. Mol. Cell 44, 667–678 (2011).

32. Simon, M.D. et al. The genomic binding sites of a noncoding RNA. Proc. Natl. Acad. Sci. USA 108, 20497–20502 (2011).

33. Xiang, Y., Miller, D.E., Ross, E.J., Sanchez Alvarado, A. & Hawley, R.S. Synaptosomal complex extension from clustered telomeres mediates full-length chromosome pairing in Schmidtea mediterranea. Proc. Natl. Acad. Sci. USA 111, E1519–E1568 (2014).

34. Maguire, M.P. The mechanism of meiotic homologue pairing. J. Theor. Biol. 106, 605–616 (1984).

35. Rockmill, B. & Roeder, G.S. Telomere-mediated chromosome pairing during meiosis in budding yeast. Genes Dev. 12, 2574–2586 (1998).

36. Reig-Viader, R. et al. Telomeric repeat-containing RNA and telomerase in human fetal oocytes. Hum. Reprod. 28, 414–422 (2013).

37. Zicker, D. & Kleckner, N. Recombination, pairing, and synopsis of homologs during meiosis. Cold Spring Harb. Perspect. Biol. 7, a016626 (2015).

38. Barakat, T.S. et al. The trans-activator RNF12 and cis-acting elements effectuate X chromosome inactivation independent of X-pairing. Mol. Cell 53, 965–978 (2014).

39. Lee, J.T. Homozygous Tsix mutant mice reveal a sex-ratio distortion and revert to random X-inactivation. Nat. Genet. 32, 195–200 (2002).

40. Feretaki, M. & Lingner, J. A practical qPCR approach to detect TERRA, the elusive telomeric repeat-containing RNA. Methods 114, 39–45 (2017).

41. Farnung, B.D., Giulotto, E. & Azzalin, C.M. Promoting transcription of chromosome ends. Transcription 1, 140–143 (2010).

42. Spilianakis, C.G., Lalioti, M.D., Town, T., Lee, G.R. & Flavell, R.A. Interchromosomal associations between alternatively expressed loci. Nature 435, 637–645 (2005).
**ONLINE METHODS**

**FISH & pairing assays.** Male (J1) and female (16.7) ES cells were cytospun onto glass slides and permeabilized with CSK buffer containing 0.5% Triton X-100 and fixed in 4% paraformaldehyde. DNA oligos probes for RNA FISH were ordered from Integrated DNA Technologies. For TERRA: (TAACCC)3, -Alexa488-3’ and 5’-Cy5 (TAACCC). For 14 oligos: 14-47k 5’-Alexa488-TGG-ACT GAC CTC TGG CCA CTG GGT GGC CTC CGT GTG CAG ACC TGA GGT T; I4 31k: 5’-Cy5-CTC TGT GTC TCT CTC CTG CTC TGG CAG ACC TGA GGT T; I4 31k: 31K-CTC TGT GTC TCT CTC CTG CTC TGG CAG ACC TGA GGT T; I4 31K: TAT CTG GGC CTC CGT GTG CAG ACC TGA GGT T; I4 31k: 5’-Cy5-CTC TGT GTC TCT CTC CTG CTC TGG CAG ACC TGA GGT T; I4 31k: TAT CTG GGC CTC CGT GTG CAG ACC TGA GGT T; I4 31k: 5’-Cy5-CTC TGT GTC TCT CTC CTG CTC TGG CAG ACC TGA GGT T; I4 31k: TAT CTG GGC CTC CGT GTG CAG ACC TGA GGT T; I4 31k: 5’-Cy5-CTC TGT GTC TCT CTC CTG CTC TGG CAG ACC TGA GGT T; I4 31k: TAT CTG GGC CTC CGT GTG CAG ACC TGA GGT T; I4 31k: 5’-Cy5-CTC TGT GTC TCT CTC CTG CTC TGG CAG ACC TGA GGT T; I4 31k: TAT CTG GGC CTC CGT GTG CAG ACC TGA GGT T; I4 31k: 5’-Cy5-CTC TGT GTC TCT CTC CTG CTC TGG CAG ACC TGA GGT T; I4 31k: TAT CTG GGC CTC CGT GTG CAG ACC TGA GGT T; I4 31k: 5’-Cy5-CTC TGT GTC TCT CTC CTG CTC TGG CAG ACC TGA GGT T; I4 31k: TAT CTG GGC CTC CGT GTG CAG ACC TGA GGT T; I4 31k: 5’-Cy5-CTC TGT GTC TCT CTC CTG CTC TGG CAG ACC TGA GGT T; I4 31k: TAT CTG GGC CTC CGT GTG CAG ACC TGA GGT T; I4 31k: 5’-Cy5-CTC TGT GTC TCT CTC CTG CTC TGG CAG ACC TGA GGT T; I4 31k: TAT CTG GGC CTC CGT GTG CAG ACC TGA GGT T; I4 31k: 5’-Cy5-CTC TGT GTC TCT CTC CTG CTC TGG CAG ACC TGA GGT T; I4 31k: TAT CTG GGC CTC CGT GTG CAG ACC TGA GGT T; I4 31k: 5’-Cy5-CTC TGT GTC TCT CTC CTG CTC TGG CAG ACC TGA GGT T; I4 31k: TAT CTG GGC CTC CGT GTG CAG ACC A

**CHIRT-seq analysis.** CHIRT6 combines ChiRP and CHART protocols31,32 as follows. We used a minimum number of capture probes to reduce off-target effects and increased the shearing size to 0.5–3 kb to preserve IncRNA integrity. Because we observed that RNase H is not active in SDS, we used NP40 instead of SDS or N-lauroyl sarcosine during DNA elution to preserve RNase H activity. In brief, 15 million exponentially PCR-amplified ES cells were washed with PBS and resuspended in 10 ml of PBS. Cells were then fixed by adding 10 ml of 2% glutaraldehyde at RT for 10 min, and crosslinking was quenched with 0.125 M glycine for 5 min. Cells were spun at 2000 g for 5 min at 4 °C, washed with cold PBS, resuspn, snap frozen in liquid nitrogen, and stored at –80 °C. For differentiating ES cells (d3, d7), embryoid bodies were trypsinized and filtered with cell striainers (40 µm) before the above steps. For nuclei isolation during CHIRT, cells were thawed, resuspended in 1 ml of swelling buffer (0.1 M Tris pH 7.0 10 mM KOAc, 15 mM MgOAc, 1% NP40, 1 mM DTT, 1 mM PMSF, 100 µM Superase-In (Ambion)) for 10 min on ice, dounced and pelleted at 2500 g for 5 min. Nuclei were further lyzed in 50 µM Tris pH 7.0, 10 mM EDTA, 1% SDS, 1 mM DTT, 1 mM PMSF, protease inhibitor, 100 µM Superase-In on ice for 10 min, and sonicated (Bioruptor) for a 0.5–3 kb size range. Lysates were spun down at 13,000 r.p.m. for 5 min to remove debris, snap frozen in liquid nitrogen, and stored in –80 °C. Streptavidin–magnetic C1 (Life Technologies) beads were blocked with 300 µg/ml yeast total RNA, and 1 mg/ml BSA for 1 h at 37 °C, and resuspended in 1× hybridization buffer (1 vol of lysate bufferplus 2 volume of 2× hybridization buffer). Lysates were diluted in two volumes of 2× hybridization buffer (lyse: 2× hyc: 1:2) (75 mM NaCl, 1% SDS, 50 µM Tris pH 7.0, 1 mM EDTA, 15% formamide, 1 mM DTT, PMSF, protease inhibitor, 100 U/µl Superase-In), precleaned with Streptavidin magnetic C1 beads at 37 °C for 1 h (100 µl of beads for 1 ml lysates), and incubated with pooled probes (100 pmol for 3 ml of diluted cell lysates) at 37 °C for 3 h. 300 µl washed and blocked C1 beads were added per 100 pmol of pools, and the reaction was mixed for another 1 h at 37 °C. DNA probes for CHIRT were ordered from Integrated DNA Technologies and labeled with 3′ biotin-TEG. PAR DNA probes are as follows: 36K: GAGCGCGCTAGTCTGACAGGAT, 29K: CTCCGTCGGCAGCTGAGGTTT, 34K: CCGCTACTACCATCAGAGA, 31K: TCTCTGCTCTGTCGCTAC. TERRA-AS probe, TAACCCAGACTAACCCTA. TERRA-sense probe, CTCTGCTCTGTCGCTAC. TERRA-AS probe, TAACCTAACCCTAACTA. TERRA-sense probe, CTCTGCTCTGTCGCTAC.

**TERRA-capture RNA-seq.** 20 µg of Trizol-purified total RNA was treated with 4 U of TURBO DNase at 37 °C for 10 min 100 µl with RNase inhibitor, 10 nM of ribonuclease vanadyl complex (VRC). 5 mM EDTA was added to disrupt TAT CTG GGC CTC CGT GTG CAG ACC TGA GGT T; I4 31k: 5’-Cy5-CTC TGT GTC TCT CTC CTG CTC TGG CAG ACC TGA GGT T; I4 31k: TAT CTG GGC CTC CGT GTG CAG ACC TGA GGT T; I4 31k: 5’-Cy5-CTC TGT GTC TCT CTC CTG CTC TGG CAG ACC TGA GGT T; I4 31k: TAT CTG GGC CTC CGT GTG CAG ACC A

© 2017 Nature America, Inc., part of Springer Nature. All rights reserved.
enrichment and maximum likelihood estimates, resulted in similar distribution patterns. Scatter plots for correlation analysis used input-normalized coverage produced by SPP, windowed by 3 kb bins and filtered for enriched bins with an averaged density of at least 4. Peaks were called by MACS (1.4.2)41 using normalization to indicated controls (input, sense, no RNase H or pre-RNase A), and filtered by peak length greater than 1 kb and >10-fold enrichment. Metagene profiles were produced by software CEAS (0.9.9.7)42 using two-fold enriched over input wig files and bed files produced by MACS peak calling. At least two biological replicates of each CHIRT experiment were examined and they showed similar results.

PAR-TERRA knockdown analysis. 2 × 106 ES cells (16.7, cas/mus hybrid) were grown in antibiotic-free media and differentiated into embryonic bodies for 4 d. Cells then were trypsinized and transfected with LNA gapmers (Exiqon) at 2 µM in 100 µl media solution using A30 program (Lonza) for pairing assay. For Xist RNA FISH, cells were transfected with 10 nM of LNA using Lipofectamine 2000 or RNAiMax, and collected at 6, 24, 48, 96 h after transfection. Gapmer LNAs with a phosphothiolated backbone were: Scr, 5′-CAG GTC TAT ACA CCA C-3′; PAR, 5′-TCT CGT TCT TCG C-3′; TERRA, 5′-TAA CCC TAA CCC TAA C-3′; sense, 5′-TTA GGG TTA GGG TTA G-3′. Multiple biological replicates were examined for each knockdown experiment. Degree of knockdown was determined by quantitative northern analysis.

Northern blotting analysis. Probes used were as follows: TERRA: TAAACCTTAAACC TAAACCTTAAACC; GAPDH: GTAGACGCAAGCCTTGACCAC CGGCGTACCCCAT; I4-29k: TAATCTGAATATCTGGGCCTCCGTGCA; I4-31k: GTCTCTGTGTCTGTCTCTCTGTCTCTGTCGCTAA CCGGCCTCACCCCATT; PAR-TERRA knockdown analysis. Two biological replicates of each CHIRT experiment were examined and they verified duplicates. If multiple reads at the same position arise from duplicates, they will result in multiple reads that align to exactly the same coordinates and have the same sequence. These reads are indistinguishable from PCR duplicates, which is an extension of our pipeline for aligning Hi-C reads allele-specifically48, which is an extension of our pipeline for aligning ChIP-seq reads allele-specifically27,49,50. Next, we employed our paired-end sequencing strategy allows us to sequence the viewpoint primer in read 1 and then identify the potential interaction sequence that has ligation into the viewpoint in read 2. This allows us to identify library molecules that actually contain the viewpoint, as opposed to library molecules that arise from off-target priming or spurious amplification of from the universal adaptor sequences, thus allowing us to eliminate several sources of noise from our data. To achieve this, we first filtered read pairs so that we include only read pairs with the viewpoint sequence present at the beginning of read 1. We included a 6-nucleotide random barcode sequence at the beginning of read 2; we stripped this sequence from the beginning of read 2 and appended it to the name of the read in the fastq file. We also removed the Fat1 sequence next to the barcode. Finally, we trimmed read 2 for sequence past DpnII sites, as we are interested in the molecules that ligated onto the DpnII site near the viewpoint primer. All filtering steps were conducted using cutadapt 1.7.1 using its built-in paired-end filtering options.

After these filtering steps, read 2 was aligned to the mm9 reference genome using the same single-end alignment strategy we have previously used to align Hi-C reads allele-specifically48, which is an extension of our pipeline for aligning ChIP-seq reads allele-specifically27,48,50. Next, we employed our random-barcode strategy to eliminate PCR duplicates. Properly identifying and removing PCR duplicates in 4C experiments is usually not possible using standard methods, because the positions of 4C reads are ‘anchored’ by the positions of the first and second restriction sites. Thus, genuine interactions will result in multiple reads that align to exactly the same coordinates and have the same sequence. These reads are indistinguishable from PCR duplicates, but because we have included a random barcode sequence, we can properly identify duplicates. If multiple reads at the same position arise from duplication during PCR, then all reads will have the same barcode, whereas if they...
arise from a ligation event present multiple times in the population, then the reads will have distinct barcodes. Thus, we collapsed the counts of all reads aligning with the same barcode to the same position to 1, eliminating PCR duplicates from the data. Read counts were normalized at each position by read depth and represent read counts as reads/million reads (RPM). Only uniquely mapped reads were used in the downstream analysis. Reads counts were normalized in 10 Kb genomic bins by 4cseq_pipeline. Statistical significance of detected interaction domains was estimated as $P$ values based on fitting the observed empirical distribution of normalized contact data with Weibull distribution.

Growth conditions and ES cell differentiation. Cells were all tested and confirmed that they were free from mycoplasma contamination. Mouse ES cells were cultured in regular ES medium (500 ml DMEM with the addition of 1 ml of β-mercaptoethanol, 6 ml of MEM NEAA, 25 ml of 7.5% NaHCO$_3$, 6 ml of GlutaMAX-1, 15 ml of 1M HEPES, 90 ml of FBS, 300 µl of LIF, 6 ml of PEN/STREP) with feeders. ES cell differentiation was initiated by removing LIF and cultured in low adherent plates to form embryonic bodies in suspension (day 0 to day 4). At differentiation day 5–8, embryonic bodies were plated onto gelatinized plates to allow them to attach and outgrow.

Data availability. CHIRT-seq, TERRA-captured RNA seq, and 4C-seq data have been deposited in GEO (GSE69887).

Kharchenko, P.V., Tolstorukov, M.Y. & Park, P.J. Design and analysis of ChIP-seq experiments for DNA-binding proteins. Nat. Biotechnol. 26, 1351–1359 (2008).

Zhang, Y. et al. Model-based analysis of ChIP-Seq (MACS). Genome Biol. 9, R137 (2008).

Shin, H., Liu, T., Manrai, A.K. & Liu, X.S. CEAS: cis-regulatory element annotation system. Bioinformatics 25, 2605–2606 (2009).

Nagano, T. et al. Comparison of Hi-C results using in-solution versus in-nucleus ligation. Genome Biol. 16, 175 (2015).

Apostolou, E. et al. Genome-wide chromatin interactions of the Nanog locus in pluripotency, differentiation, and reprogramming. Cell Stem Cell 12, 699–712 (2013).

Minaigi, A. et al. Chromosomes. A comprehensive Xist interactome reveals cohesin repulsion and an RNA-directed chromosome conformation. Science 349, aab2276 (2015).

Pinter, S.F. et al. Spreading of X chromosome inactivation via a hierarchy of defined Polycomb stations. Genome Res. 22, 1864–1876 (2012).

Yildirim, E., Sadreyev, R.I., Pinter, S.F. & Lee, J.T. X-chromosome hyperactivation in mammals via nonlinear relationships between chromatin states and transcription. Nat. Struct. Mol. Biol. 19, 56–61 (2011).

van de Werken, H.J. et al. Robust 4C-seq data analysis to screen for regulatory DNA interactions. Nat. Methods 9, 969–972 (2012).