Europe PMC Funders Group
Author Manuscript
Nature. Author manuscript; available in PMC 2014 April 03.

Published in final edited form as:
Nature. 2013 October 3; 502(7469): . doi:10.1038/nature12593.

Single cell Hi-C reveals cell-to-cell variability in chromosome structure

Takashi Nagano[#1], **Yaniv Lubling**[#2], **Tim J. Stevens**[#3], **Stefan Schoenfelder**[1], **Eitan Yaffe**[2], **Wendy Dean**[4], **Ernest D. Laue**[3], **Amos Tanay**[2], and **Peter Fraser**[1]

[1]Nuclear Dynamics Programme, The Babraham Institute, Cambridge, UK

[2]Department of Computer Science and Applied Mathematics and Department of Biological Regulation, Weizmann Institute, Rehovot, Israel

[3]Department of Biochemistry, University of Cambridge, UK

[4]Epigenetics Programme, The Babraham Institute, Cambridge, UK

[#] These authors contributed equally to this work.

Abstract

Large-scale chromosome structure and spatial nuclear arrangement have been linked to control of gene expression and DNA replication and repair. Genomic techniques based on chromosome conformation capture assess contacts for millions of loci simultaneously, but do so by averaging chromosome conformations from millions of nuclei. Here we introduce single cell Hi-C, combined with genome-wide statistical analysis and structural modeling of single copy X chromosomes, to show that individual chromosomes maintain domain organisation at the megabase scale, but show variable cell-to-cell chromosome territory structures at larger scales. Despite this structural stochasticity, localisation of active gene domains to boundaries of territories is a hallmark of chromosomal conformation. Single cell Hi-C data bridge current gaps between genomics and microscopy studies of chromosomes, demonstrating how modular organisation underlies dynamic chromosome structure, and how this structure is probabilistically linked with genome activity patterns.

Chromosome conformation capture[1] (3C) and derivative methods (4C, 5C and Hi-C)[2-6] have enabled the detection of chromosome organisation in the 3D space of the nucleus. These methods assess millions of cells and are increasingly used to calculate conformations of a range of genomic regions, from individual loci to whole genomes[3,7-11]. However, fluorescence in situ hybridisation (FISH) analyses show that genotypically and phenotypically identical cells have non-random, but highly variable genome and chromosome conformations[4,12,13] probably due to the dynamic and stochastic nature of chromosomal structures[14-16]. Therefore, whilst 3C-based analyses can be used to estimate

Correspondence and requests for materials should be addressed to PF (peter.fraser@babraham.ac.uk) for the single cell Hi-C method, AT (amos.tanay@weizmann.ac.il) for the statistical analysis, or EDL (e.d.laue@bioc.cam.ac.uk) for the structural modelling..
Author Contributions TN and PF devised the single cell Hi-C method. TN performed single cell Hi-C and DNA FISH experiments. SS carried out ensemble Hi-C experiments. WD microscopically isolated single cells. YL, EY and AT processed and statistically analyzed the sequence data. TJS and EDL developed the approach to structural modelling and analysed X chromosome structures. TJS wrote the software for 3D modeling, analysis and visualisation of chromosome structures. TN, YL, TJS, EDL, AT and PF contributed to writing the manuscript, with inputs from all other authors.

Data deposited in NCBI's Gene Expression Omnibus (Nagano *et al*., 2013) and are accessible through GEO Series accession number GSE48262 (http://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE48262).

The authors declare that they have no competing financial interests.

Europe PMC Funders Author Manuscripts
Europe PMC Funders Author Manuscripts

Europe PMC Funders Author Manuscripts

Europe PMC Funders Author Manuscripts

an average conformation, it cannot be assumed to represent one simple and recurrent chromosomal structure. To move from probabilistic chromosome conformations averaged from millions of cells towards determination of chromosome and genome structure in individual cells, we developed single cell Hi-C, which has the power to detect thousands of simultaneous chromatin contacts in a single cell.

Single cell Hi-C

We modified the conventional or "ensemble" Hi-C protocol[3] to create a method to determine the contacts in an individual nucleus (Fig. 1a, Supplementary Information). We used male, mouse, spleenic $CD4^+$ T cells, differentiated *in vitro* to T helper (Th1) cells to produce a population of cells (>95% $CD4^+$), of which 69% have 2n genome content, reflecting mature cell withdrawal from the cell cycle. Chromatin cross-linking, restriction enzyme (Bgl II or Dpn II) digestion, biotin fill-in and ligation were performed in nuclei (Fig. 1a and Extended Data Fig. 1a) as opposed to ensemble Hi-C where ligation is performed after nuclear lysis and dilution of chromatin complexes[3]. We then selected individual nuclei under the microscope, placed them in individual tubes, reversed cross-links, and purified biotinylated Hi-C ligation junctions on streptavidin-coated beads. The captured ligation junctions were then digested with a second restiction enzyme (Alu I) to fragment the DNA, and ligated to customized Illumina adapters with unique 3 bp identification tags. Single cell Hi-C libraries were then PCR amplified, size selected and characterized by multiplexed, paired-end sequencing.

De-multiplexed single cell Hi-C libraries were next filtered thoroughly to systematically remove several sources of noise (Extended Data Fig. 1b-f, Supplementary Information). Hi-C in male diploid cells can theoretically give rise to at most two ligation products per autosomal restriction fragment end, and one product per fragment end from the single X chromosome. Using Bgl II, the total number of distinct mappable fragment-end pairs per single cell cannot therefore exceed 1,201,870 (Extended Data Fig. 1g, Supplementary Information). In practice, deep sequencing of the single cell Hi-C libraries demonstrated that following stringent filtering our current scheme allows recovery of up to 2.5% of this theoretical potential, and has identified at least 1000 distinct Hi-C pairings in half (37/74) of the cells. Deep sequencing confirmed saturation of the libraries' complexity, and allowed elimination of spurious flow cell read pairings and additional biases (Extended Data Tables 1-3). Based on additional quality metrics we selected ten single cell datasets, containing 11,159-30,671 distinct fragment-end pairs for subsequent in-depth analysis (Extended Data Fig. 1h-l). Visualization of the single cell maps suggested that despite their inherent sparseness, they clearly reflect hallmarks of chromosomal organization, including frequent *cis*-contacts along the matrix diagonal and notably, highly clustered *trans*-chromosomal contacts between specific chromosomes (Fig. 1b).

Single cell and ensemble Hi-C similarity

We used the same population of $CD4^+$ Th1 cells to generate an ensemble Hi-C library. Sequencing and analysis[17] of 190 million read pairs produced a contact map representing the mean contact enrichments within approximately 10 million nuclei. The probability of observing a contact between two chromosomal elements decays with linear distance following a power law regime for distances larger than 100 kb[3,18]. We found similar regimes for the ensemble, individual cells and a pool of 60 single cells (Fig. 1c). Moreover, after normalizing the matrices given this canonical trend, comparison of intra-chromosomal interaction intensities for the pool and ensemble, by global correlation analysis of contact enrichment values at 1 Mb resolution generates a highly significant correspondence (Fig. 1d). This is emphasized by the high similarity observed in comparisons of individual

Europe PMC Funders Author Manuscripts

Europe PMC Funders Author Manuscripts

chromosomes from ensemble and pooled Hi-C maps (Fig. 1e). In summary, despite different experimental procedures and sparse nature of the single cell matrices, the pooled matrix retains the most prominent properties of the ensemble map, confirming the validity of the approach and prompting us to further explore the similarities and differences among the individual cell chromosomal conformations.

Intra- and inter-domain contacts

A key architectural feature of ensemble Hi-C datasets is their topological domain structure[18-20]. As expected 1403 domains were identified in the Th1 cell ensemble Hi-C map[18] (Supplementary Information Table 1, and Supplementary Information). We used the ensemble domains to ask whether the same domain structure can be observed at the single cell level. Visual inspection of the domain structure overlaid on individual intra-chromosomal contact maps (Fig. 2a), and global statistical analysis of the ratios between intra- and inter-domain contact intensities in individual cells (approximately 2-fold enrichment on scales of 100 Kb to 1 Mb, Fig. 2b and Extended Data Fig. 2a), both supported the idea that domains are observed consistently in the single cell maps. To test whether domain structures are variable between individual cells, we estimated the distributions of intra-domain contact enrichments across cells and compared it to the distributions derived from reshuffled maps. We reasoned that cell-to-cell variation in intra-domain contact intensities would result in an increase of the variance of this distribution compared to the expected variance resulting from sampling contacts in uniformly (shuffled) intra-domain contacts. The data however (Fig. 2c), showed that the distributions for the intra-domain enrichments in real cells are not more varied than expected (Kolmogorov-Smirnov $p < 0.52$). A similar observation was derived by comparison of the correlations between intra-domain contact enrichments for pairs of real and pairs of reshuffled maps (Extended Data Fig. 2b). While this analysis cannot quantify variability in the high-resolution internal structure of domains, the data suggests that domain intactness is generally conserved at the single cell level.

Visual comparison of whole chromosome contact maps (Fig. 2d) suggested that unlike intra-domain interactions, inter-domain contacts within single cell chromosomes are structured non-uniformly. The maps showed large-scale structures as indicated, for example, by specific insulation points separating chromosomes into two or more mega-domains in a cell-specific fashion. To rule out the possibility that this can be explained by sparse sampling of contacts in each single cell map we again used reshuffled controls. In each map (real or randomized) we quantified the frequency of loci that strongly polarize the matrix into two weakly connected submatrices (using an insulation score; Supplementary Methods). We confirmed that single cell maps indeed show many more such loci than reshuffled maps (Fig. 2e and Extended Data Fig. 2c). The reshuffled controls made by mixing contacts from different single cell maps, are in fact similar to sparse versions of the ensemble map, which do not show specific structure at the intra-domain level. Along similar lines, the correlation in contact intensities between domains on the same chromosome in pairs of single cell maps is lower compared to reshuffled controls (Fig 2f). Taken together, these data show that domains form a robust and recurrent conformational basis that is evident in each of the single cells. However, inter-domain contacts are highly variable between individual cells, suggesting large-scale differences in higher-order chromosome folding that are obscured in ensemble maps, averaged over millions of such structures[21].

3D structural modeling of X chromosomes

To determine whether the single cell Hi-C data is consistent with unique chromosome conformations we developed a modeling approach to reconstruct the conformations of the

Europe PMC Funders Author Manuscripts

Europe PMC Funders Author Manuscripts

single-copy, male X chromosome. We used intra-chromosomal contacts as distance restraints and calculated structural models using a simulated annealing protocol to condense a particle-on-a-string representation of individual chromosomes from random initial conformations (Supplementary Information), to produce both fine-scale and low-resolution models, with backbone particles representing either 50 or 500 kb of the chromosome, respectively. For fine-scale calculations, each intra-chromosomal contact restrained its precise position on the chromosome, while low-resolution calculations combined contacts into larger bins. Tests of our simulation protocol demonstrated that restraint density was the most important parameter for modeling (Extended Data Fig. 3a, b). Hence, from the ten high-quality single cell datasets, we selected six with the largest numbers of intra-chromosomal X contacts, plus one with a lower number of contacts (cell-9) for contrast.

Repeat calculations starting from random positions generated 200 X chromosome models for each cell at both scales. The fine-scale models displayed very low numbers of restraint violations (Extended Data Fig. 3c). We introduced an estimated average unit DNA distance length[22] to approximate packaging of chromatin fibers (~0.15 μm/50 kb) (Supplementary Information). This resulted in models with a mean X chromosome territory diameter of 4.3 μm (range 3.3-5.9 μm), in good agreement with X chromosome paint FISH in Th1 cells (Fig. 3a; mean diameter 3.7 μm) and chromosome territory sizes in live cells[23]. We confirmed that the restrained points in a single cell are indeed close in the structures calculated from them (Extended Data Fig. 3c, d). Interestingly, the single cell distance matrix demonstrates how the network of contacts in a model imparts further structural information beyond the directly observed contacts (Extended Data Fig. 3d).

Comparison of the low-resolution models demonstrated convergence toward a single conformation for each single cell dataset (Fig. 3b and Extended Data Fig. 3e). For fine-scale models, hierarchical clustering revealed 4-5 that were most representative of the data (Fig. 3c). In all cases models from a single cell were significantly more similar to each other than to models from different cells. (Extended Data Fig. 4a, b).

Highlighting four regions of the X chromosome showed large-scale conformational differences between cells (Fig. 3d), supporting the finding of highly variable inter-domain contacts. Models created by shuffling Hi-C contacts, or combining contacts from two cells resulted in structures smaller and more compact than observed chromosome territories (Extended Data Fig. 4c, d) with many restraints stretched toward or exceeding their upper bounds (Extended Data Fig. 3c). These results reaffirm that the variation in single cell contacts is not the result of partial sampling of a single underlying structure.

We next asked whether despite their cell-to-cell variability, X chromosome structures share common folding properties that could be tested in real cells. One such important property, which is often consistent within a cell population, and with multiple potential functional implications, is localisation within the chromosomal territory relative to its surface. To predict loci with consistent positions within their chromosome territory we calculated the structural density along the X chromosome (Supplementary Information) and identified regions with consistently high or low structural density (Fig. 3e). We chose five such regions (P1-P5) with predicted positions near the surface (P1, P2, P5; low structural density) or inside (P3, P4; high structural density) the model X chromosome territories using the 1200 models from the six cells (Extended Data Fig. 4e). We then performed double label DNA FISH with X chromosome paints and P1-P5 BAC probes (Fig. 3f) to directly test these predictions. The distances between DNA FISH signals and edge of the chromosome territory in over one hundred Th1 cells showed that probes P1, P2 and P5 were indeed found predominantly outside or toward the edge of the chromosome territory, whereas signals for probes P3 and P4 were found at internal positions (Fig. 3g). These data show that despite

highly variable inter-domain structure of the X chromosomal territory, some of its key organisational properties are robustly observed across the cell population.

Domains at the interface

Overlaying data from *trans*-chromosomal contacts on the X chromosome models demonstrates that *trans*-chromosomal contacting regions are strongly enriched toward the inferred surface of the models (Fig. 3h), providing further validation. These observations prompted us to further explore the structural characteristics of interfaces between chromosomal territories, and the relationships between such interfaces and the domain structure of the territory itself. We found that *trans*-chromosomal contact enrichments of domains vary across cells (Fig. 4a), showing a significant difference between the mean contact enrichment per domain in the real and reshuffled maps ($p < 1.2e-9$, Kolmogorov-Smirnov test). The higher variance of the distribution for the real data suggests some domains are more likely to contact elements on other chromosomes. Previous work has suggested that active genomic regions on the sub-domain scale often loop out of their chromosome territories[24], which might imply less defined local domain structures and disassociation from their chromosome territory. However, our analysis shows that *trans*-contacting domains retain domain organisation, as demonstrated by the intra-domain contact probabilities within them (Fig. 4b and Extended Data Fig. 5a, b). On the other hand, *trans*-contacting domains show slightly reduced contact intensity to other domains on the same chromosome (Fig. 4c and Extended Data Fig. 5c, d), consistent with localisation on the interfaces of their territories rather than dissociation from them.

Analyses of ensemble Hi-C data have previously shown that active marks correlate with enrichment of *trans*-chromosomal contacts[3,17]. Using the single cell maps combined with annotation of domains based on their enrichment for histone H3 lysine 4 tri-methylation (H3K4me3) hotspots[25] (Fig. 4d), we tested whether this correlation is the result of low frequency re-localisation of active domains to other chromosome territories (looping out), or from frequent localisation of active domains on territory interfaces. As shown in Fig. 4e, domains with high *trans*- to *cis*-chromosomal contact ratios (excluding intra-domain) are highly correlated with H3K4me3 enrichment in all cells. However, the data show that domains (including active ones) retain their association with the territory in almost all cases. Very few domains with strong *trans*-contacts were found to lack association with their own territory (i.e., upper left points in graphs in Fig. 4e). Some of this lack of perfect territory re-localisation can be explained by having two copies of each autosomal domain, but the overall reduction with territory association for *trans*-contacting domains is much smaller than the 50% expected by this explanation (reduction estimated at 15-20% and 10% for contacts across 1-5 Mb and 10 Mb, respectively, Fig. 4c). Comparison of active domain localization shows that different active domains are highly *trans*-contacting in each cell (Extended Data Fig. 5e). Together, these data show that preferential localisation of active domains to territory interfaces is a hallmark of chromosome organisation in all cells. Active domains maintain their intra-domain organisation, and only partially lose intra-chromosomal contacts with other domains. Our data are consistent with the concept that chromosomal territories are maintained robustly despite the *trans*-chromosomal contacts between active domains.

Interestingly, domains associated with Lamin-B1[26], which are thought to be primarily inactive regions, are also found toward the surface of the models (Fig. 3h). However, these domains are highly anti-correlated with H3K4me3 domains (Spearman's correlation = −0.73) and typically depleted of *trans*-chromosomal contacts (Extended Data Fig. 5f-i). Superposition of H3K4me3, Lamin-B1 enriched domains and *trans*-chromosomal contacts on the X chromosome models illustrates spatial partitioning of the active, *trans*-contacting

Europe PMC Funders Author Manuscripts
Europe PMC Funders Author Manuscripts

Europe PMC Funders Author Manuscripts

Europe PMC Funders Author Manuscripts

regions from those that are Lamin-associated, although both types of domains tend toward the surface of the chromosome territory, supporting the above descriptions of differential positioning of domains (Extended Data Fig. 5j and Supplementary Videos 1 and 2).

Ensemble Hi-C maps generate a highly complex view of chromosomal contacts, including low intensity contacts between all possible chromosomal pairings[3,8,17,19]. In contrast, years of single cell analyses by microscopy have suggested that individual cells have much simpler and discrete chromosome structures involving a limited number of interfaces between spatially constrained chromosomal territories[27,28]. Our single cell maps bridge the gap between the genomic and imaging techniques, showing cell-specific clusters of *trans*-chromosomal contacts associating some pairs of chromosomes, and a lack of contacts between other chromosome pairs (Fig. 5a, blue). Such organisation is completely lacking in reshuffled maps (Fig. 5a, red) confirming it is not a consequence of sparse contact sampling (Extended Data Fig. 6a, b). *Trans*-chromosomal contact clusters bring pairs of domains together, as shown by comparing the enrichment in *trans*-contacts between pairs of elements connecting the same two domains and pairs connecting one domain with two different domains (Fig. 5b). Such synergistic contacting preferentially brings together pairs of active domains, with interaction between active and inactive domains being underrepresented (Fig. 5c and Extended Data Fig. 6c). Although inactive domains are depleted as a group from *trans*-chromosomal interactions (Fig. 4e), inactive domains that engage in *trans*-contacts are more likely to interact with other inactive domains. Interestingly, analysis of interacting pairs of domains suggests that the number of chromosomes contacted by each chromosome is relatively constant (less than 30% difference) despite the >3-fold change in chromosome size, the total number of *trans*-chromosomal contacts in the map, or a number of other factors (Fig. 5d and Extended Data Fig. 7a-e). We note that even though the total number of chromosome-chromosome interfaces per single cell is bounded, the detailed interface between chromosome pairs can involve multiple domain-domain contacts reflecting higher order organization (Extended Data Fig. 7f).

Overall, these results indicate that each chromosome contacts a discrete and fairly constant number of other chromosomes in a single cell, with little dependency on the chromosome size. At the single cell level both the microscopic and genomic observations therefore indicate highly defined territory structures, which can be hypothesised to harbor much of the chromosome within the territory, and expose a limited, relatively constant surface area engaged in chromosome-to-chromosome interfaces. Since these interfaces are highly variable among different cells, their averaging by ensemble Hi-C contributes toward the relatively uniform *trans*-chromosomal contact matrices previously reported.

We have presented a new experimental strategy to create Hi-C contact maps from single cells. The approach allows for characterization of thousands of simultaneous contacts occurring in individual cells, and provides unique insights into Hi-C technology and 3D chromosomal architecture (Supplementary Videos 3 and 4). Single cell contact maps reflect conservation of domain structure that was recently characterized[18-20], but show that inter-domain and *trans*-chromosomal contact structure is highly variable between individual cells. Genome-wide statistical analysis and reconstruction of the single copy X chromosome models gave us the opportunity to quantify key features of chromosomal architecture. For example, active domains tend to locate on the boundaries of their chromosomal territories in the majority of nuclei, while maintaining associations with other domains on the same chromosome. Our results do not exclude chromosome territory intermingling[29], but argue against domains becoming completely immersed in other territories. Coupled with previous observations of small and large-scale chromatin mobility[30-32] a highly dynamic view of chromosomal organization ermerges, where territories are continuously being remodeled,

Europe PMC Funders Author Manuscripts

Europe PMC Funders Author Manuscripts

while maintaining some key local (domain) and global (depth from surface) organisational features.

Online Methods

Male Th1 cells were fixed and subjected to modified Hi-C, in which nuclei were maintained through restriction enzyme digestion, biotin fill-in labelling and ligation. Single nuclei were isolated and processed to prepare single cell Hi-C libraries for paired-end sequencing.

Sequences were mapped to the mouse genome, and abnormal read pairs were discarded. Read pairs that occurred only once (without duplication) in the library sequencing were removed. We chose 10 single cell datasets for further in-depth analyses based on several quality criteria (see Supplementary Information). To validate the single cell Hi-C procedure, we pooled the single cell Hi-C datasets and compared them to ensemble Hi-C dataset prepared from approximately 10 million cells essentially as described[3]. We created reshuffled datasets by randomly redistributing contacts of the analyzed single cells to create the same number of cells with the same number of contacts in each cell as a control to statistically analyse the variation among single cell datasets.

We reconstructed three-dimensional X chromosome structure models using restrained molecular dynamics calculations employing a simulated annealing protocol. A combination of unambiguous distance restraints from the X intra-chromosomal contacts in the single cell Hi-C dataset and anti-distance restraints between regions that were found not to contact each other in the ensemble Hi-C dataset was used. To assess the precision and accuracy of the structure generation process we used the protocol to generate synthetic Hilbert curve structures, and explored the impact of varying the number of restraints. For pair-wise comparison of the structures, we calculated the root-mean-square deviation (RMSD). To compare the X chromosome models to X chromosome structure *in vivo*, we selected five loci with consistently high or low structural density in the models, and compared distances between the loci and the X chromosome territory surface in cells (DNA FISH).

Full description of the methods can be found in the Supplementary Information.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

The authors thank I. Clay, S. Wingett, K. Tabbada, D. Bolland, S. Walker, S. Andrews, M. Spivakov, N. Cope, L. Harewood and W. Boucher for assistance. This work was supported by the Medical Research Council, the Biotechnology and Biological Sciences Research Council (to PF), the MODHEP project, the Israel Science Foundation (to AT) and the Wellcome Trust (to EDL).

References

1. Dekker J, Rippe K, Dekker M, Kleckner N. Capturing chromosome conformation. Science. 2002; 295:1306–11. [PubMed: 11847345]
2. Dostie J, et al. Chromosome Conformation Capture Carbon Copy (5C): a massively parallel solution for mapping interactions between genomic elements. Genome Res. 2006; 16:1299–309. [PubMed: 16954542]
3. Lieberman-Aiden E, et al. Comprehensive mapping of long-range interactions reveals folding principles of the human genome. Science. 2009; 326:289–93. [PubMed: 19815776]
4. Schoenfelder S, et al. Preferential associations between co-regulated genes reveal a transcriptional interactome in erythroid cells. Nat Genet. 2010; 42:53–61. [PubMed: 20010836]

Europe PMC Funders Author Manuscripts

Europe PMC Funders Author Manuscripts

5. Simonis M, et al. Nuclear organization of active and inactive chromatin domains uncovered by chromosome conformation capture-on-chip (4C). Nat Genet. 2006; 38:1348–54. [PubMed: 17033623]
6. Zhao Z, et al. Circular chromosome conformation capture (4C) uncovers extensive networks of epigenetically regulated intra- and interchromosomal interactions. Nat Genet. 2006; 38:1341–7. [PubMed: 17033624]
7. Duan Z, et al. A three-dimensional model of the yeast genome. Nature. 2010; 465:363–7. [PubMed: 20436457]
8. Kalhor R, Tjong H, Jayathilaka N, Alber F, Chen L. Genome architectures revealed by tethered chromosome conformation capture and population-based modeling. Nat Biotechnol. 2012; 30:90–8. [PubMed: 22198700]
9. Marti-Renom MA, Mirny LA. Bridging the resolution gap in structural modeling of 3D genome organization. PLoS Comput Biol. 2011; 7:e1002125. [PubMed: 21779160]
10. Tanizawa H, et al. Mapping of long-range associations throughout the fission yeast genome reveals global genome organization linked to transcriptional regulation. Nucleic Acids Res. 2010; 38:8164–77. [PubMed: 21030438]
11. van de Werken HJ, et al. Robust 4C-seq data analysis to screen for regulatory DNA interactions. Nat Methods. 2012; 9:969–72. [PubMed: 22961246]
12. Osborne CS, et al. Active genes dynamically colocalize to shared sites of ongoing transcription. Nat Genet. 2004; 36:1065–71. [PubMed: 15361872]
13. Rapkin LM, Anchel DR, Li R, Bazett-Jones DP. A view of the chromatin landscape. Micron. 2012; 43:150–8. [PubMed: 22172345]
14. Fraser P, Bickmore W. Nuclear organization of the genome and the potential for gene regulation. Nature. 2007; 447:413–7. [PubMed: 17522674]
15. Lanctot C, Cheutin T, Cremer M, Cavalli G, Cremer T. Dynamic genome architecture in the nuclear space: regulation of gene expression in three dimensions. Nat Rev Genet. 2007; 8:104–15. [PubMed: 17230197]
16. Osborne CS, et al. Myc dynamically and preferentially relocates to a transcription factory occupied by Igh. PLoS Biol. 2007; 5:e192. [PubMed: 17622196]
17. Yaffe E, Tanay A. Probabilistic modeling of Hi-C contact maps eliminates systematic biases to characterize global chromosomal architecture. Nat Genet. 2011; 43:1059–65. [PubMed: 22001755]
18. Sexton T, et al. Three-dimensional folding and functional organization principles of the Drosophila genome. Cell. 2012; 148:458–72. [PubMed: 22265598]
19. Dixon JR, et al. Topological domains in mammalian genomes identified by analysis of chromatin interactions. Nature. 2012; 485:376–80. [PubMed: 22495300]
20. Nora EP, et al. Spatial partitioning of the regulatory landscape of the X-inactivation centre. Nature. 2012; 485:381–5. [PubMed: 22495304]
21. Gibcus JH, Dekker J. The hierarchy of the 3D genome. Mol Cell. 2013; 49:773–82. [PubMed: 23473598]
22. Jhunjhunwala S, et al. The 3D structure of the immunoglobulin heavy-chain locus: implications for long-range genomic interactions. Cell. 2008; 133:265–79. [PubMed: 18423198]
23. Muller I, Boyle S, Singer RH, Bickmore WA, Chubb JR. Stable morphology, but dynamic internal reorganisation, of interphase human chromosomes in living cells. PLoS One. 2010; 5:e11560. [PubMed: 20644634]
24. Heard E, Bickmore W. The ins and outs of gene regulation and chromosome territory organisation. Curr Opin Cell Biol. 2007; 19:311–6. [PubMed: 17467967]
25. Deaton AM, et al. Cell type-specific DNA methylation at intragenic CpG islands in the immune system. Genome Res. 2011; 21:1074–86. [PubMed: 21628449]
26. Peric-Hupkes D, et al. Molecular maps of the reorganization of genome-nuclear lamina interactions during differentiation. Mol Cell. 2010; 38:603–13. [PubMed: 20513434]
27. Cremer T, Cremer C. Chromosome territories, nuclear architecture and gene regulation in mammalian cells. Nat Rev Genet. 2001; 2:292–301. [PubMed: 11283701]

Europe PMC Funders Author Manuscripts

Europe PMC Funders Author Manuscripts

28. Misteli T. Beyond the sequence: cellular organization of genome function. Cell. 2007; 128:787–800. [PubMed: 17320514]

29. Branco MR, Pombo A. Intermingling of Chromosome Territories in Interphase Suggests Role in Translocations and Transcription-Dependent Associations. PLoS Biol. 2006; 4:e138. [PubMed: 16623600]

30. Chuang CH, et al. Long-range directional movement of an interphase chromosome site. Curr Biol. 2006; 16:825–31. [PubMed: 16631592]

31. Chubb JR, Boyle S, Perry P, Bickmore WA. Chromatin motion is constrained by association with nuclear compartments in human cells. Curr Biol. 2002; 12:439–45. [PubMed: 11909528]

32. Dundr M, et al. Actin-dependent intranuclear repositioning of an active gene locus in vivo. J Cell Biol. 2007; 179:1095–103. [PubMed: 18070915]

Europe PMC Funders Author Manuscripts

Europe PMC Funders Author Manuscripts

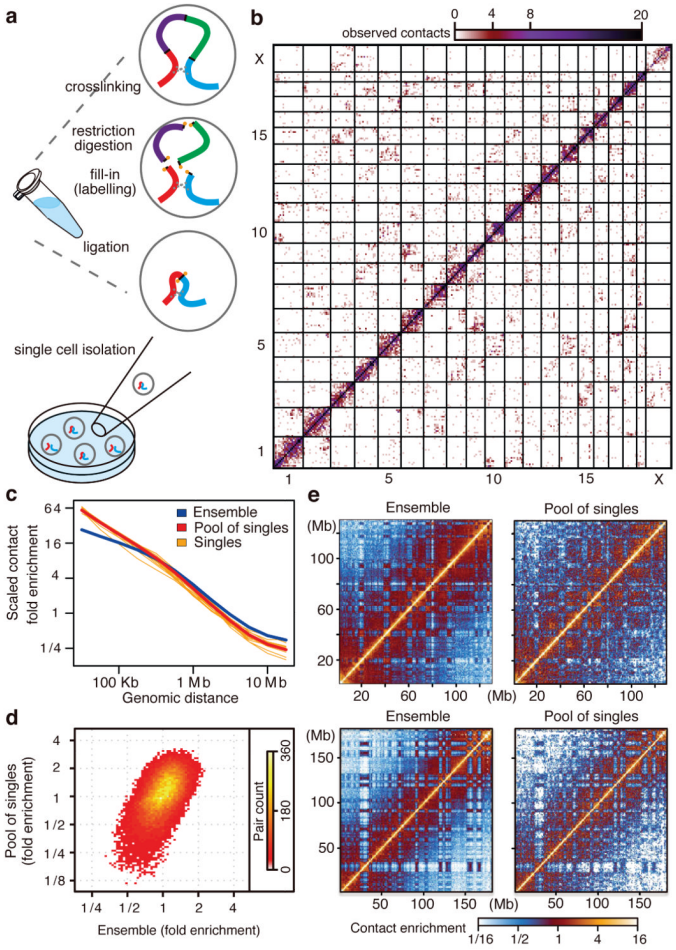

Figure 1. Single cell and ensemble Hi-C

a, Single cell Hi-C method. **b**, Single cell Hi-C heatmap (cell-5), coverage for 10 Mb bins. **c**, Contact enrichment versus genomic distance, from ensemble Hi-C, pool of 60 single cells and 10 individual cells, scaled to normalise sequencing depths. **d**, Normalising by the trends in **c**, intra-chromosomal contact enrichments for 1 Mb square bins, comparing ensemble and pooled single cell Hi-C (Spearman correlation = 0.56). **e**, Intra-chromosomal contact enrichment maps of ensemble and pooled single cell Hi-C, for chromosome 10 (top) and chromosome 2 (bottom), using variable bin sizes.

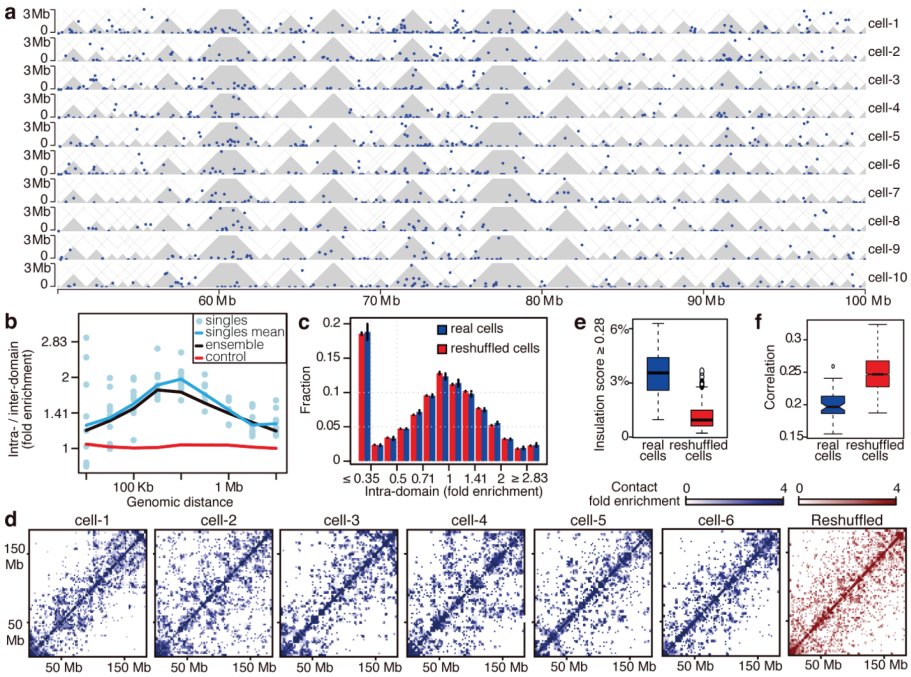

Figure 2. Conserved intra-domain, but not inter-domain structure in single cells

a, Individual intra-chromosomal contacts of 50 Mb region of chromosome 2 up to 3 Mb distance (blue dots), domains (grey). **b**, Ratios between intra-domain and inter-domain contact enrichments over genomic distance. Control is combined trend of 10 single cells calculated by repeatedly shifting the domains randomly. **c**, Distribution of intra-domain contact enrichments per domain from 9 cells (where Bgl II was used) and reshuffled datasets (black bars, standard errors). **d**, Maps of inter-domain contacts intensities for chromosome 2 from individual cells and reshuffled controls using variable bin sizes. **e**, Distribution of percentage of loci with high insulation scores in single vs. reshuffled cells. **f**, For all pairs of single cells, the correlations between inter-domain contact numbers of all pairs of domains within the same chromosome were computed. Shown are the distributions of these correlations in the real and reshuffled cells.

Europe PMC Funders Author Manuscripts
Europe PMC Funders Author Manuscripts

Europe PMC Funders Author Manuscripts

Europe PMC Funders Author Manuscripts

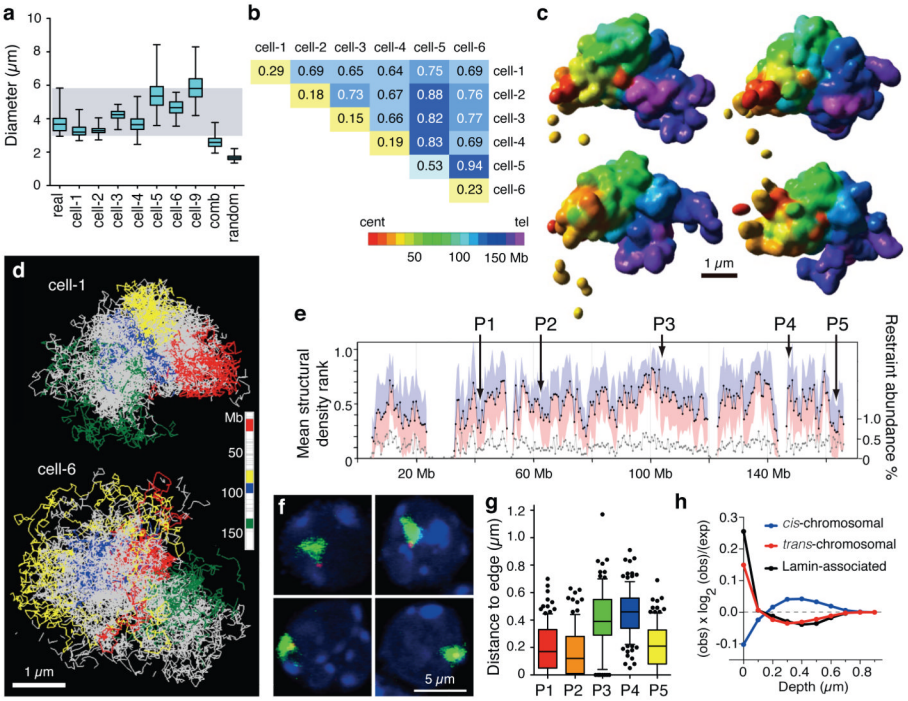

Figure 3. Structural modeling of X chromosomes

a, Distribution of longest diameter of X chromosome paint DNA FISH signals in 62 male Th1 cells (real), 200 structural models calculated for each single cell (cell-1 to -9), 200 structures from combined dataset (cell-1 and -2; comb) and 200 structures from 20 randomised cell-1 datasets (random; 10 calculations per dataset). Whiskers denote minimum and maximum. **b**, Average coordinate root-mean-square deviation (RMSD) values in microns comparing 200 low-resolution structural models for each cell and between cells. **c**, Four surface-rendered models of the X chromosome from cell-1, which are most representative of the data based on hierarchical clustering of pair-wise RMSD values (Supplementary Information). Scale bar, 1 μm. **d**, Structural ensembles of the four most representative fine-scale models for cell-1 and cell-6, with four large regions coloured. Scale bar, 1 μm. **e**, Mean structural density rank for 500 kb regions (black) from 6 × 200 fine-scale models from cell-1 to -6. Standard deviation (blue/pink). Abundance of intra-chromosomal restraints (grey, right axis). DNA FISH probes (P1-P5) are indicated. **f**, DNA FISH on Th1 cells. X chromosome paint (green) and specific locus signals (red). **g**, Distribution of DNA FISH distance measurements between signal centres for probes P1 - P5 and edge of the X chromosome territory in Th1 cells (n = 114, 113, 105, 115, 108 for P1-P5). Whiskers denote 10 and 90 percentiles. **h**, Enrichment of *cis*-, *trans*-contacts and Lamin-B1 associated domains at various depths on the chromosome models relative to null hypothesis of random positions.

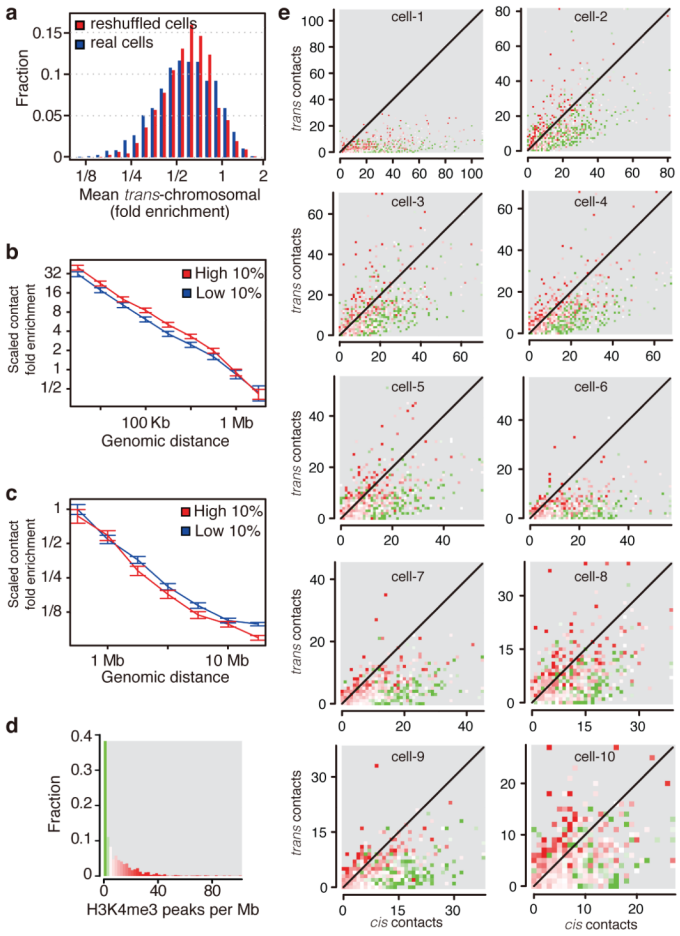

Figure 4. Active domains localise to territory interfaces

a, Distribution of *trans*-chromosomal contact enrichments of each domain averaged across real and reshuffled cells. Reshuffling maintains the number of *cis* and *trans* contacts within each cell and chromosome. **b**, Intra-domain contact enrichment over genomic distance for high vs. low *trans*-chromosomal contacting domains selected independently in each cell, with 95% confidence intervals. **c**, Same sets as in **b** but plotting the enrichment of inter-domain contacts. **d**, Distribution of H3K4me3 peak density in domains (number of peaks divided by size), color-coded according to density. **e**, Domains plotted according to number of *trans*- and *cis*-chromosomal (excluding intra-domain) contacts, color coded for H3K4me3 density as in **d**.

Europe PMC Funders Author Manuscripts
Europe PMC Funders Author Manuscripts

Europe PMC Funders Author Manuscripts

Europe PMC Funders Author Manuscripts

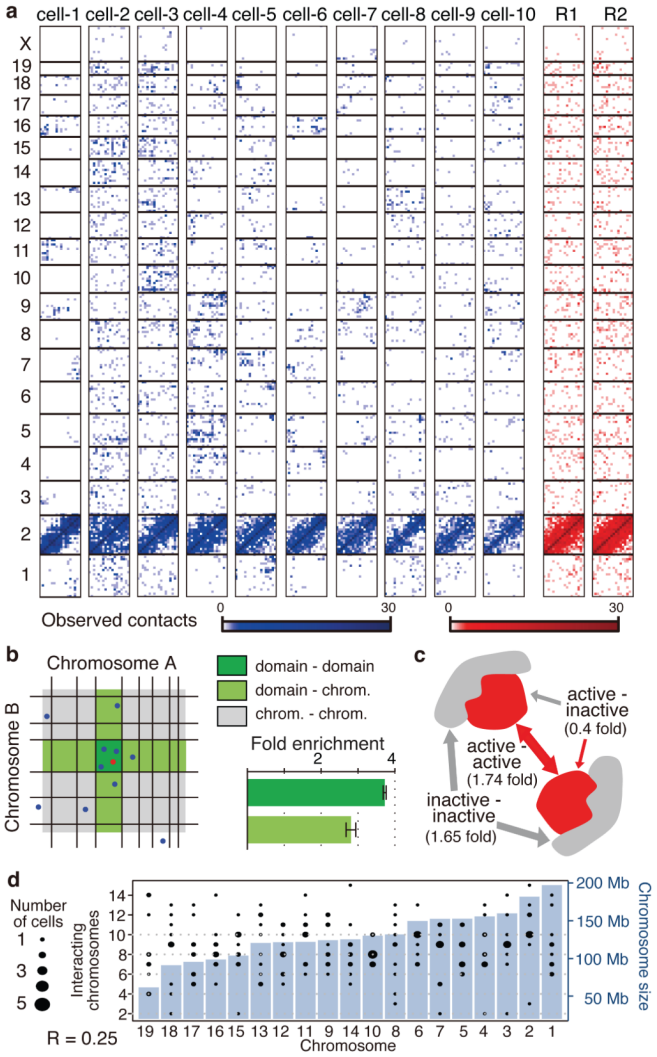

Figure 5. Chromosomal interfaces

a, All *trans*-chromosomal contacts formed by chromosome 2 in real cells (blue) and reshuffled (red). **b**, Schematic diagram of a chromosomal interface between linearly adjacent domains, their borders marked in black on two chromosomes, A and B. We considered each of the two contacting fragments of every *trans*-chromosomal contact and classified every nearby *trans*-chromosomal contact as domain-domain, domain-chromosome and chromosome-chromosome, the latter being used as background for normalisation (Supplementary Information). Contact under consideration (red), nearby contacts (blue). Fold enrichments shown for each group type (error bars, standard deviation). **c**, *Trans*-chromosomal contacts are highly significantly enriched between active domains (H3K4me3 enriched) or between inactive domains, but not mixed interaction (chi-square test; p = 5.8e-18; even after taking account of the generally higher connectivity of active domains). **d**, Bar graph depicting mouse autosomes ordered by size with number of interacting chromosomes per single cell (black circles depict the distribution over individual cells). Mean number of interacting chromosomes changes modestly (30%) with chromosome size, suggesting a highly organized territory structure with surface that is not scaling with chromosome length.