LETTER

Single-nucleus Hi-C reveals unique chromatin reorganization at oocyte-to-zygote transition

Ilya M. Flyamer[1,2,3]*†, Johanna Gassler[1]*, Maxim Imakaev[4,5]*, Hugo B. Brandão[6], Sergey V. Ulianov[2,3], Nezar Abdennur[7], Sergey V. Razin[2,3], Leonid A. Mirny[4,5,6]§ & Kikuë Tachibana-Konwalski[1]§

Chromatin is reprogrammed after fertilization to produce a totipotent zygote with the potential to generate a new organism[1]. The maternal genome inherited from the oocyte and the paternal genome provided by sperm coexist as separate haploid nuclei in the zygote. How these two epigenetically distinct genomes are spatially organized is poorly understood. Existing chromosome conformation capture-based methods[2–5] are not applicable to oocytes and zygotes owing to a paucity of material. To study three-dimensional chromatin organization in rare cell types, we developed a single-nucleus Hi-C (high-resolution chromosome conformation capture) protocol that provides greater than tenfold more contacts per cell than the previous method[2]. Here we show that chromatin architecture is uniquely reorganized during the oocyte-to-zygote transition in mice and is distinct in paternal and maternal nuclei within single-cell zygotes. Features of genomic organization including compartments, topologically associating domains (TADs) and loops are present in individual oocytes when averaged over the genome, but the presence of each feature at a locus varies between cells. At the sub-megabase level, we observed stochastic clusters of contacts that can occur across TAD boundaries but average into TADs. Notably, we found that TADs and loops, but not compartments, are present in zygotic maternal chromatin, suggesting that these are generated by different mechanisms. Our results demonstrate that the global chromatin organization of zygote nuclei is fundamentally different from that of other interphase cells. An understanding of this zygotic chromatin 'ground state' could potentially provide insights into reprogramming cells to a state of totipotency.

To investigate 3D genome organization in nuclei of single cells, we developed a genome-wide high-resolution *in situ* Hi-C approach. Conventional Hi-C methods include biotin incorporation and enrichment for ligated fragments[6], which might limit fragment retrieval. We simplified the protocol by omitting these steps, similarly to genome conformation capture[7] (Fig. 1a, Extended Data Fig. 1, see Supplementary Methods). To verify the protocol, we compared data from population and single-cell data from K562 (human chronic myelogenous leukaemia) cells and obtained a contact probability $P_c(s)$, dependent on genomic distance, s, that matched conventional *in situ* Hi-C on bulk K562 cells[8] (Fig. 1b). When applied to oocytes (Fig. 1c), our method was efficient at capturing chromosomal interactions: single-nucleus Hi-C (snHi-C) revealed up to 1.9×10^6 contacts per cell after filtering, yielding 1–2 orders of magnitude more contacts than published single-cell Hi-C data[2] and exceeding contact frequencies in a recent single-cell Hi-C preprint and report[9,10]. Half of the cells had $>3.39 \times 10^5$ contacts per cell and 7.1% had $>1 \times 10^6$ contacts per cell (Supplementary Table 1). These high-density snHi-C data enabled us to examine chromatin features directly in single-cell maps.

To investigate higher-level chromatin organization in oocytes, we examined how contact probability $P_c(s)$[6,11] depends on genomic distance in individual cells and pooled data. In oocytes, the shapes of $P_c(s)$ curves were consistent between individual cells (Fig. 1d) but markedly different from the characteristic shape in other mammalian interphase cells (Fig. 1e). For genomic separation $s > 1$ Mb, we observed steeper ($\sim s^{-1.5}$) decay in oocyte $P_c(s)$, closer to the random walks of yeast chromosomes[12–14]. Our simulations (see below) showed that steeper $P_c(s)$ can be attributed to the larger volume of oocyte nuclei ($\sim 25\,\mu m$ versus $\leq 10\,\mu m$ diameter in somatic cells, Extended Data Fig. 2a).

Another major feature of mammalian chromosomes is segregation into active and inactive (A–B) compartments[6]. Although assignment of compartments from snHi-C data was impossible owing to its sparsity, an enrichment of interactions between the same compartment type and depletion between different types became evident in individual cells when compartments were assigned using GC content (Fig. 2a) or population Hi-C from other cell types (Extended Data Figs 3, 9). We also examined whether loops[8] and TADs[15,16], prominent functional features of chromatin organization[17–19], are present in individual cells. Averaging over all genomic positions of loops and TADs identified in population Hi-C[8] (for CH12-LX cells, see Supplementary Methods) revealed that both are present in individual oocytes as average enrichments (Fig. 2a, Extended Data Fig. 3b) but vary between cells (Extended Data Fig. 3c), reflecting both inter-cell variability and variations in experimental conditions. We conclude that a single nucleus shows enrichments of interactions between regions of the same compartment type, within TADs and at loops.

Using snHi-C data, we asked whether TADs constitute physically isolated domains in individual cells or reflect a mere tendency of chromosomes in a cell population to interact more within and less outside of a domain. We envisioned three scenarios (Fig. 2b): (i) all population-identified TADs are present in every individual cell; (ii) only population TAD boundaries are present, but individual TADs can be missing or fused in single cells; (iii) contacts may be clustered in individual cells, but clusters do not always match population TADs, revealing them only as an average feature. To distinguish between these scenarios, we examined a region in the snHi-C maps with most contacts (Fig. 2c, d) and segmented chromosomes into domains of enriched contact frequency (contact clusters) using an exact segmentation algorithm that maximizes modularity (see Supplementary Methods; comparable results for modularity segmentation of population TADs were obtained using an algorithm from ref. 20, see Extended Data Fig. 4a). We found that single-cell contact clusters do not always match population TADs, as contact clusters are highly variable and frequently occur across TAD borders. Nevertheless, variable contact clusters averaged into TADs when pooled together (Fig. 2d). The high cell-to-cell variations
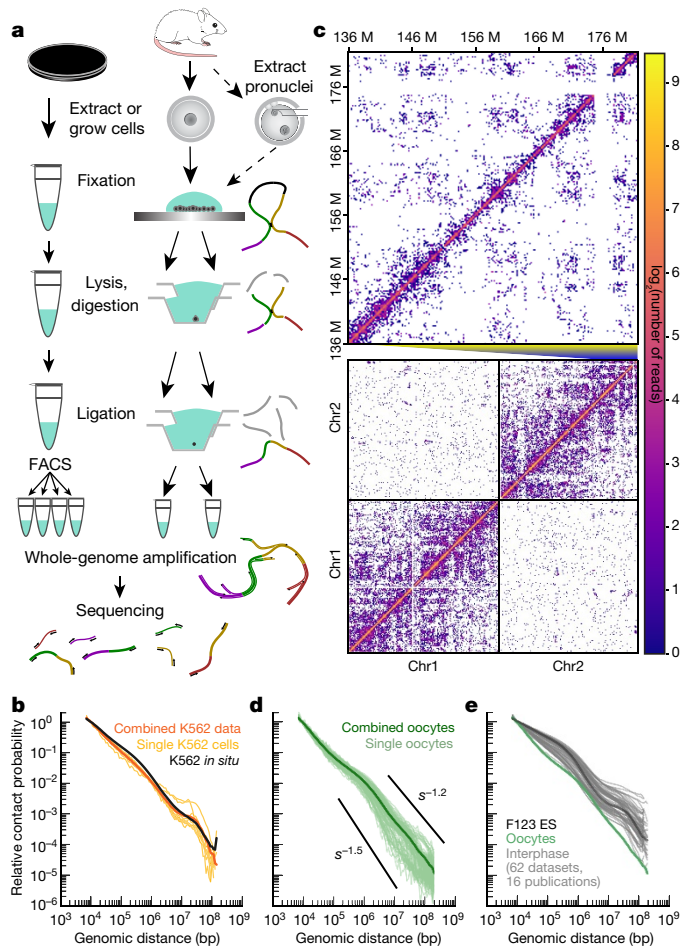
Figure 1 | Genome-wide high-resolution single-nucleus Hi-C approach. a, snHi-C workflow for cell culture and oocytes/zygotes. **b,** Dependence of contact probability on genomic separation, $P_c(s)$, for single K562 cells ($n = 9$) with >30,000 total contacts (yellow), combined single-cell K562 data from all cells (orange) and published population K562 data[8] (black). $P(s)$ here and below were normalized to be 1 at 10 kb. **c,** Example contact map from a single oocyte (cell 1). Below: chromosomes 1 and 2 at 1 Mb resolution. Above: fragment of chromosome 2 at 200 kb resolution. **d,** $P_c(s)$ in single oocytes with >30,000 total contacts ($n = 84$, from more than four biological replicates using 2–4 females each) and in combined data. Black lines show slopes for $P_c(s) = s^{-1.5}$ and $P_c(s) = s^{-1.2}$. **e,** $P_c(s)$ of oocytes (green, combined data) compared to published conventional Hi-C datasets (grey) with highlighted curve for ES cells[30] (black).

in contact patterns cannot be solely explained by experimental DNA loss because we often observe a presence of border-violating clusters rather than absence, and second, such patterns are also observed in regions of high read coverage (Extended Data Fig. 4b, c).

We validated frequent violations of TAD boundaries using 3D DNA fluorescence *in situ* hybridization (FISH) for equidistant pairs of probes located within a TAD and across a TAD border. Scenarios (i) and (ii), with TADs present in single cells as isolated domains, are expected to yield intra-TAD distances that are mostly shorter than inter-TAD distances. However, imaging of embryonic stem (ES) cells showed that inter-TAD distances are shorter than intra-TAD in 42% of cases, although the average inter-TAD distance is larger than the average intra-TAD (Wilcoxon test, $P = 0.007$). This indicates that although TAD borders affect the average distance and lead to preferential intra-TAD contacts, they cannot prevent contacts between TADs in single cells (Fig. 2e). Surprisingly, even the long inter-TAD pair was closer than the intra-TAD pair in 18% of cells, despite having twice the linear genome separation, having 50% larger average distance, and fourfold lower contact probability. Together, both imaging data and snHi-C

support scenario (iii), in which TADs reflect a tendency for contact enrichment arising from a diverse conformational ensemble, rather than being isolated blocks of DNA present in individual cells.

The recently proposed mechanism of TAD formation by loop extrusion and boundary insulation[21,22] provides a rationale for frequent boundary violations and, generally, TAD stochasticity. In this model, insulation prevents extruded loops from crossing boundaries, but it does not directly prevent contacts between TADs. Polymer simulations show that contact clusters naturally emerge from the 3D spatial proximity of DNA in a confined volume[22] and frequently cross TAD boundaries (Extended Data Figs 5, 6 and ref. 11). As contact clusters are also detected in K562 snHi-C (Extended Data Fig. 4d), stochastic cluster formation is probably a universal property of chromatin organization in single cells.

Next, we investigated the spatial reorganization of chromatin during the transition from transcriptionally active immature oocytes (non-surrounded nucleolus, NSN) into transcriptionally inactive mature (surrounded-nucleolus, SN) oocytes[23,24] (Fig. 3a). We observed a significant decrease in loop, TAD, and compartment strengths (Fig. 3b–d, Extended Data Fig. 7; all Mann–Whitney $P < 0.005$) during maturation, which may be related to transcriptional silencing and visual detachment of chromatin from the nuclear envelope[23,24] (Fig. 3a). Whereas combined $P_c(s)$ scalings are similar, mature oocytes display more long-range (>400 kb) contacts (Mood's equal median test $P = 0.02$) and significantly less cell-to-cell variation in $P_c(s)$ curves (Levene's test, $P = 0.007$) (Fig. 3e–g). These findings are consistent with progressive chromatin reorganization during oocyte maturation.

We next addressed the key question, namely whether and how chromatin is reorganized during the oocyte-to-zygote transition, and whether it is different between the maternal and paternal genomes that have different biological histories and epigenetic modifications[1]. Oocyte chromosomes decondense after two meiotic divisions into the maternal nucleus. However, paternal nuclei are formed from compacted sperm by replacing protamines with histones[25,26]. To determine whether chromatin architecture is inherited or established *de novo* after fertilization, we extracted maternal and paternal nuclei from zygotes predominantly in the G1 phase and performed snHi-C (Figs 1a, 4a; similar results were obtained without extracting nuclei, see Extended Data Fig. 8a, b and Supplementary Information). We detected up to 6×10^5 contacts in zygotic nuclei, which is twofold higher than in somatic cells and threefold lower than in the highest coverage oocytes. Results from averaging over TADs and loops identified previously[8] showed that these features are present at similar strengths in maternal and paternal nuclei (Fig. 4b, Extended Data Figs 3, 9). Strikingly, although A–B compartmentalization is observed in paternal nuclei, it is notably absent from maternal nuclei (Fig. 4b, c, Extended Data Fig. 9a, b). To our knowledge, this is the first example of mammalian interphase nuclei presenting essentially no A–B compartmentalization. These results further suggest that the mechanisms forming compartments are distinct from those forming TADs and loops, in agreement with a recent preprint[27].

To corroborate this novel finding, we simultaneously imaged 25 loci across chromosome 11 using 3D FISH and for each probe measured distances to the nearest probes of the same and of the different compartment type (Fig. 4d, Extended Data Fig. 10a, b). In agreement with Hi-C findings (Fig. 4c), we found that compartmentalization in ES cells is most pronounced ($P < 10^{-16}$, one-sided Mann–Whitney U-test, Extended Data Fig. 10c); compartmentalization in paternal nuclei is weak but significant ($P < 0.01$); and compartmentalization in maternal nuclei is undetectable as compared to a randomized control with shuffled probe identities ($P = 0.08$, Fig. 4d, see Supplementary Methods). The lack of compartmentalization in the maternal genome may be due to a transcriptionally inactive extended G1 phase after fertilization[1], suggesting that compartments are established *de novo* in the maternal genome[1], whilst paternal genome compartmentalization is either inherited from sperm chromatin or established faster.
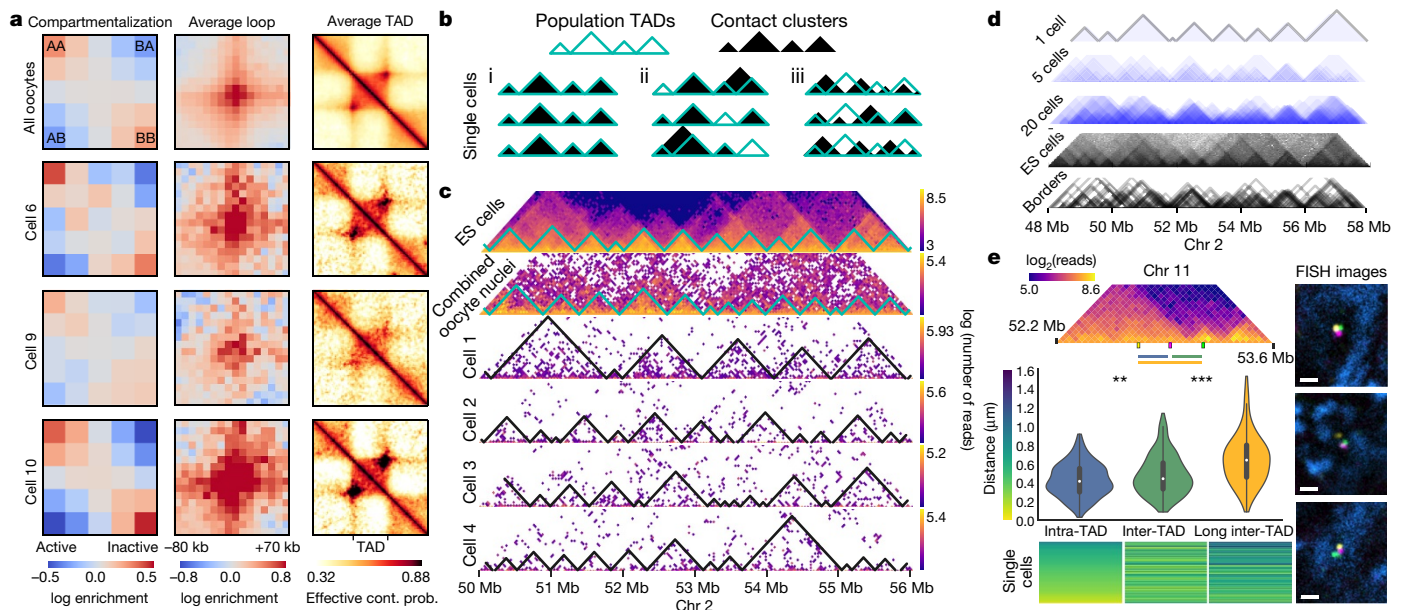
Figure 2 | snHi-C identifies compartments, TADs and chromatin loops in individual cells. a, Contact enrichment of A–B compartments averaged over genomic positions using GC content and averaging over loop and TAD positions annotated in CH12-LX cells[8] in combined and single-oocyte data. **b,** Possible scenarios that lead to TADs in combined data. **c,** Variable contact clusters of top 4 single oocytes; first row, mouse ES cells data[30]; second row, combined data from all oocytes ($n = 120$, from more than four biological replicates, using 2–4 females each). Resolution of all maps is 40 kb. Lines show boundaries of clusters identified in individual cells. **d,** Superimposing contact clusters identified in individual cells

($n = 20$) compared to population Hi-C TAD annotations. **e,** 3D FISH in mouse ES cells quantifies TAD boundary violations in single cells ($n = 211$, from two biological replicates). Top: F123 ES cell[30] Hi-C map of tested region with FISH probe locations. Middle: Distribution of measured distances, average distances (left to right) 0.428, 0.484, and 0.646 μm, relative contact probabilities from F123 cells at 20 kb resolution are 0.0095, 0.0037, 0.0024. Wilcoxon test P values: $**P = 0.007$, $***P = 2.5 \times 10^{-16}$. Bottom: Heat map of FISH measurements with colour-coded distances. Right: representative FISH images with adjusted gamma values. Scale bar, 1 μm. Probes (yellow, magenta, green) and DAPI (blue).

The compartmentalization in the paternal nuclei aligns with detection of hyperacetylated histone H4, a hallmark of active chromatin, in early G1 phase and earlier transcriptional activation[28]. These results suggest a role for transcription in genome compartmentalization. We propose that the chromatin organization of zygotic nuclei denotes a 'ground state' produced by transcriptional silencing, chromosome

condensation and an exchange in cohesin complex composition at fertilization[29].

Finally, we used polymer modelling to understand global chromosome organization in the three analysed cell types and somatic cells. A prominent feature of $P_c(s)$ curves is a steeper overall slope in oocytes and zygotes compared to somatic cells. Polymer modelling demonstrates that this steeper slope can be explained by the larger nuclear volumes of oocytes and zygotes (Extended Data Fig. 2). For maternal and paternal zygotic nuclei, experimental $P_c(s)$ curves show similar shallow slopes for genomic distances $s < 3$ Mb (Fig. 4e), probably reflecting local compaction by loop extrusion that is also observed in simulations and that underlies formation of TADs and loops[21,22]. This scaling regime in zygotes is then followed by a plateau between 3–12 Mb for the maternal genome, whereas $P_c(s)$ continues to decrease for the paternal genome. Simulations suggested that this difference may reflect previous states of differently compacted chromosomes in maternal and paternal zygotes. Simulations of decondensation subject to loop extrusion that start from a metaphase chromosome[11] (Fig. 4g, Extended Data Fig. 6) result in $P_c(s)$ that resemble those of maternal nuclei (Fig. 4h, Extended Data Fig. 2c). Analogous simulation starting from the compact fractal globule[6], as a model of protamine-compacted state (Fig. 4g, Extended Data Figs 2b, 5), can reproduce paternal $P_c(s)$ curve.

Taken together, these results suggest that the factors influencing $P_c(s)$ are nuclear density, memory of the previous chromosome state, and cell cycle phase. Zygotic maternal nuclei and somatic cells are both predominantly in G1 phase and recently experienced chromosome decondensation from metaphase, which makes their global genome organization most similar. Paternal nuclei have a different biological history due to chromatin compaction by protamines in sperm and are thus different from somatic cells and maternal nuclei. Oocytes experienced the last mitosis weeks or months ago and are arrested in prophase I; they therefore differ the most from somatic cells.

In summary, our work provides insights into general principles of chromosome organization and specific biological aspects of oocyte
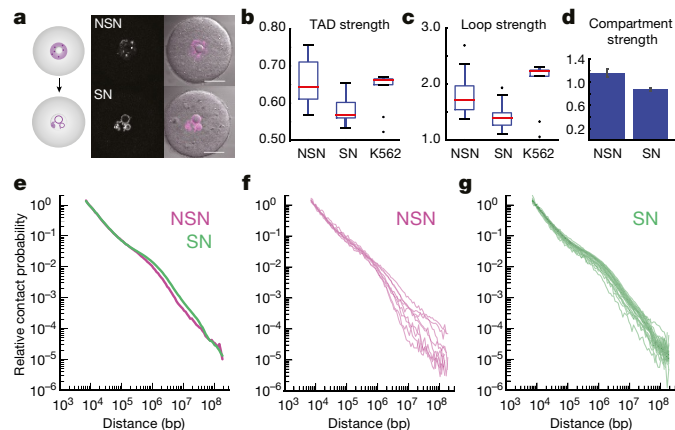

Figure 3 | Chromatin reorganization during oocyte maturation.
a, Immature (non-surrounded nucleolus, NSN) and mature (surrounded nucleolus, SN) oocytes stained with Hoechst (magenta). Scale bar, 25 μm. Images were adjusted in brightness/contrast settings in the individual channels using ImageJ. **b–d,** Comparison of average TAD strength (**b**), loop strength (**c**) and compartment strength (**d**) in NSN and SN oocytes with Hoechst staining (n (NSN) = 15, n (SN) = 30, from more than two biological replicates using 2–4 females each). Scores used described in Supplemental Methods. Error bars in **d** show s.d., obtained by bootstrapping. **e,** $P_c(s)$ in data from combined NSN and SN oocytes, scored by Hoechst staining. **f, g,** $P_c(s)$ in single Hoechst-stained NSN (**f**, $n = 9$) and SN oocytes (**g**, $n = 27$) with >30,000 total contacts.
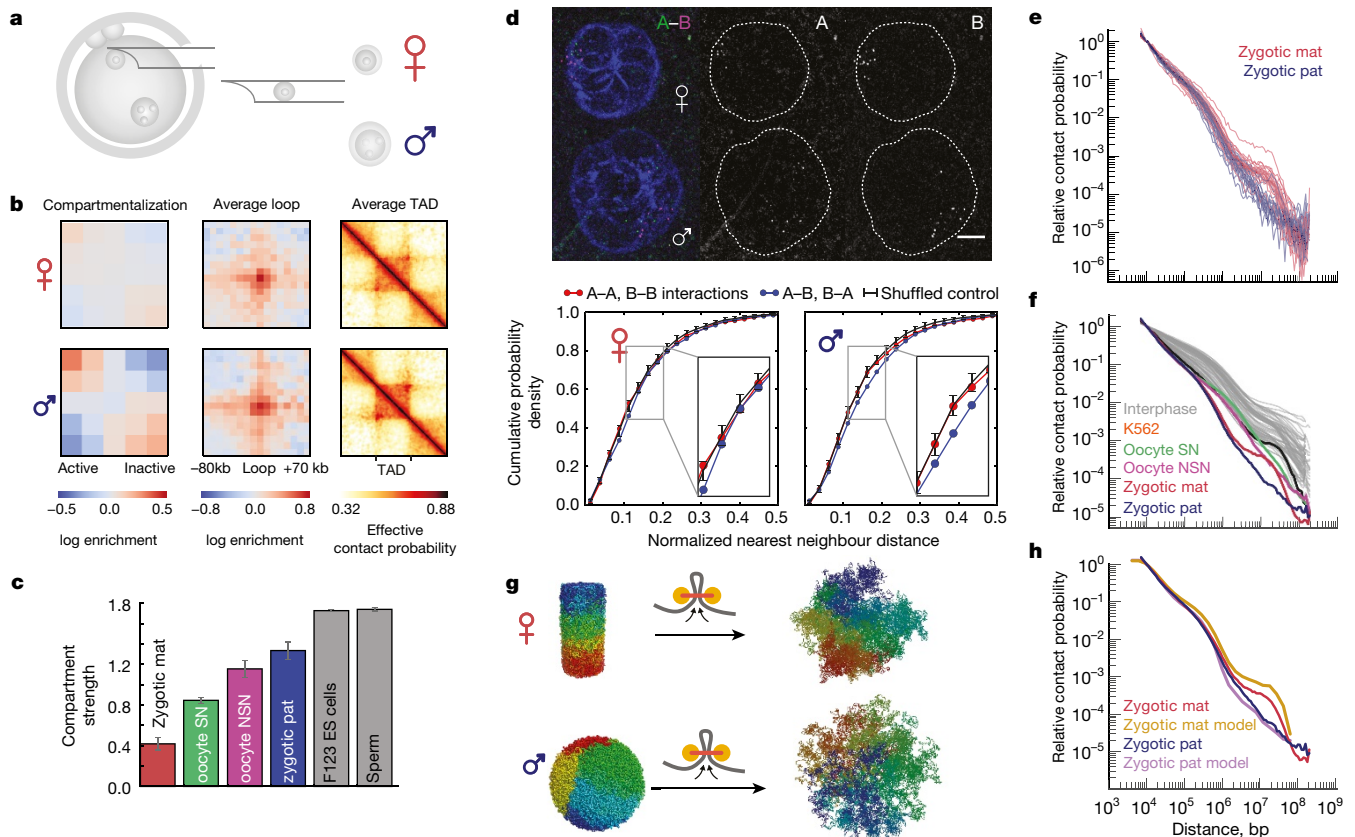
Figure 4 | Distinct chromatin architecture in haploid nuclei of totipotent zygotes. a, Extraction of nuclei from zygotes. **b**, Same as Fig. 2a but for combined data from zygote nuclei. Top, maternal ($n = 31$); bottom, paternal ($n = 24$). Data from more than three biological replicates using 3–6 females each. **c**, Comparison of compartment signal strength in combined maternal (mat) and paternal (pat) zygote nuclei, oocytes, ES cell[30] and sperm[25] data. Error bars and units as in Fig. 3d. **d**, 3D FISH compartment quantification in zygote nuclei. Top: Deconvolved zygote FISH image, adjusted in brightness/contrast settings, background subtracted using ImageJ. Scale bar, 5 μm. Bottom: Cumulative probability density plots for nearest-neighbour distances between same compartment and different compartment probe pairs for maternal ($n = 33$) and paternal

($n = 37$) nuclei. Data from one biological replicate using four females. Black denotes average reshuffled control with random compartment assignment; error bars show 5% and 95% percentiles. **e**, $P_c(s)$ in single maternal and paternal nuclei with >30,000 total contacts (maternal = 24, paternal = 20; from more than three biological replicates using 3–6 females each). **f**, $P_c(s)$ from all combined data sets (colour) contrasted with all previously published conventional Hi-C data from mammalian interphase nuclei (grey). **g**, In models of maternal and paternal nuclei, loop extrusion is initialized with a model of metaphase chromosome[11] or compact fractal globule. Colour denotes position along the chromosome. **h**, Model $P_c(s)$ compared to experimental $P_c(s)$.

and zygote genomes. We find that all known levels of chromosomal organization appear as mere tendencies that become visible when averaged over a population of cells or over loci; in single cells and at a single locus they may become overshadowed by the stochasticity of chromosome conformations. The unanticipated finding that zygotic maternal chromatin contains TADs and loops but not compartments suggests that it represents a transition towards building the embryonic chromatin organization of a totipotent cell. The difference in higher-order chromatin organization between maternal and paternal chromatin also raises the question of how paternal chromatin maintains or establishes compartments faster after fertilization. Our results, together with cohesin loader depletion experiments[27], suggest that loops and compartments are formed by distinct mechanisms. Lastly, snHi-C could enable the study of chromatin organization during development and in rare cell types, such as stem cells and distinct cells within highly heterogeneous tumours. By combining snHi-C with other single-cell approaches, including single-cell transcriptome and methylome analyses, it will be possible to build a comprehensive picture of the interplay between genome folding and transcription in generating identities of individual cells.

Online Content Methods, along with any additional Extended Data display items and Source Data, are available in the online version of the paper; references unique to these sections appear only in the online paper.

1. Zhou, L. Q. & Dean, J. Reprogramming the genome to totipotency in mouse embryos. *Trends Cell Biol.* **25,** 82–91 (2015).
2. Nagano, T. *et al.* Single-cell Hi-C reveals cell-to-cell variability in chromosome structure. *Nature* **502,** 59–64 (2013).
3. Ma, W. *et al.* Fine-scale chromatin interaction maps reveal the cis-regulatory landscape of human lincRNA genes. *Nat. Methods* **12,** 71–78 (2015).
4. Hsieh, T.-H. S. *et al.* Mapping nucleosome resolution chromosome folding in yeast by micro-C. *Cell* **162,** 108–119 (2015).
5. Barutcu, A. R. *et al.* C-ing the genome: a compendium of chromosome conformation capture methods to study higher-order chromatin organization. *J. Cell. Physiol.* **231,** 31–35 (2016).
6. Lieberman-Aiden, E. *et al.* Comprehensive mapping of long-range interactions reveals folding principles of the human genome. *Science* **326,** 289–293 (2009).
7. Rodley, C. D. M., Bertels, F., Jones, B. & O'Sullivan, J. M. Global identification of yeast chromosome interactions using Genome conformation capture. *Fungal Genet. Biol.* **46,** 879–886 (2009).
8. Rao, S. S. P. *et al.* A 3D map of the human genome at kilobase resolution reveals principles of chromatin looping. *Cell* **159,** 1665–1680 (2014).
9. Nagano, T. *et al.* Cell cycle dynamics of chromosomal organisation at single-cell resolution. Preprint at *bioRxiv* http://biorxiv.org/content/early/2016/12/15/094466 (2016).
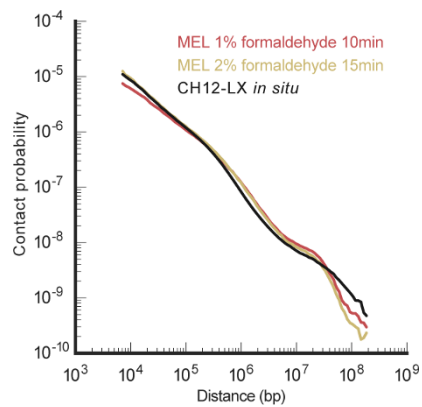
10. Ramani, V. *et al.* Massively multiplex single-cell Hi-C. *Nat. Methods* **14,** 263–266 (2017).
11. Naumova, N. *et al.* Organization of the mitotic chromosome. *Science* **342,** 948–953 (2013).
12. Mizuguchi, T. *et al.* Cohesin-dependent globules and heterochromatin shape 3D genome architecture in *S. pombe*. *Nature* **516,** 432–435 (2014).
13. Tjong, H., Gong, K., Chen, L. & Alber, F. Physical tethering and volume exclusion determine higher-order genome organization in budding yeast. *Genome Res.* **22,** 1295–1305 (2012).
14. Halverson, J. D., Smrek, J., Kremer, K. & Grosberg, A. Y. From a melt of rings to chromosome territories: the role of topological constraints in genome folding. *Rep. Prog. Phys.* **77,** 022601 (2014).
15. Dixon, J. R. *et al.* Topological domains in mammalian genomes identified by analysis of chromatin interactions. *Nature* **485,** 376–380 (2012).
16. Nora, E. P. *et al.* Spatial partitioning of the regulatory landscape of the X-inactivation centre. *Nature* **485,** 381–385 (2012).
17. Flavahan, W. A. *et al.* Insulator dysfunction and oncogene activation in IDH mutant gliomas. *Nature* **529,** 110–114 (2016).
18. Franke, M. *et al.* Formation of new chromatin domains determines pathogenicity of genomic duplications. *Nature* **538,** 265–269 (2016).
19. Lupiáñez, D. G. *et al.* Disruptions of topological chromatin domains cause pathogenic rewiring of gene-enhancer interactions. *Cell* **161,** 1012–1025 (2015).
20. Crane, E. *et al.* Condensin-driven remodelling of X chromosome topology during dosage compensation. *Nature* **523,** 240–244 (2015).
21. Sanborn, A. L. *et al.* Chromatin extrusion explains key features of loop and domain formation in wild-type and engineered genomes. *Proc. Natl Acad. Sci. USA* **112,** E6456–E6465 (2015).
22. Fudenberg, G. *et al.* Formation of chromosomal domains by loop extrusion. *Cell Reports* **15,** 2038–2049 (2016).
23. Miyara, F. *et al.* Chromatin configuration and transcriptional control in human and mouse oocytes. *Mol. Reprod. Dev.* **64,** 458–470 (2003).
24. Bouniol-Baly, C. *et al.* Differential transcriptional activity associated with chromatin configuration in fully grown mouse germinal vesicle oocytes. *Biol. Reprod.* **60,** 580–587 (1999).
25. Battulin, N. *et al.* Comparison of the three-dimensional organization of sperm and fibroblast genomes using the Hi-C approach. *Genome Biol.* **16,** 77 (2015).
26. Ward, W. S. & Coffey, D. S. DNA packaging and organization in mammalian spermatozoa: comparison with somatic cells. *Biol. Reprod.* **44,** 569–574 (1991).
27. Schwarzer, W. *et al.* Two independent modes of chromosome organization are revealed by cohesin removal. Preprint at *bioRxiv* http://biorxiv.org/content/early/2016/12/15/094185 (2016).
28. Adenot, P. G., Mercier, Y., Renard, J. P. & Thompson, E. M. Differential H4 acetylation of paternal and maternal chromatin precedes DNA replication and differential transcriptional activity in pronuclei of 1-cell mouse embryos. *Development* **124,** 4615–4625 (1997).
29. Tachibana-Konwalski, K. *et al.* Rec8-containing cohesin maintains bivalents without turnover during the growing phase of mouse oocytes. *Genes Dev.* **24,** 2505–2516 (2010).
30. Selvaraj, S., R Dixon, J., Bansal, V. & Ren, B. Whole-genome haplotype reconstruction using proximity-ligation and shotgun sequencing. *Nat. Biotechnol.* **31,** 1111–1118 (2013).

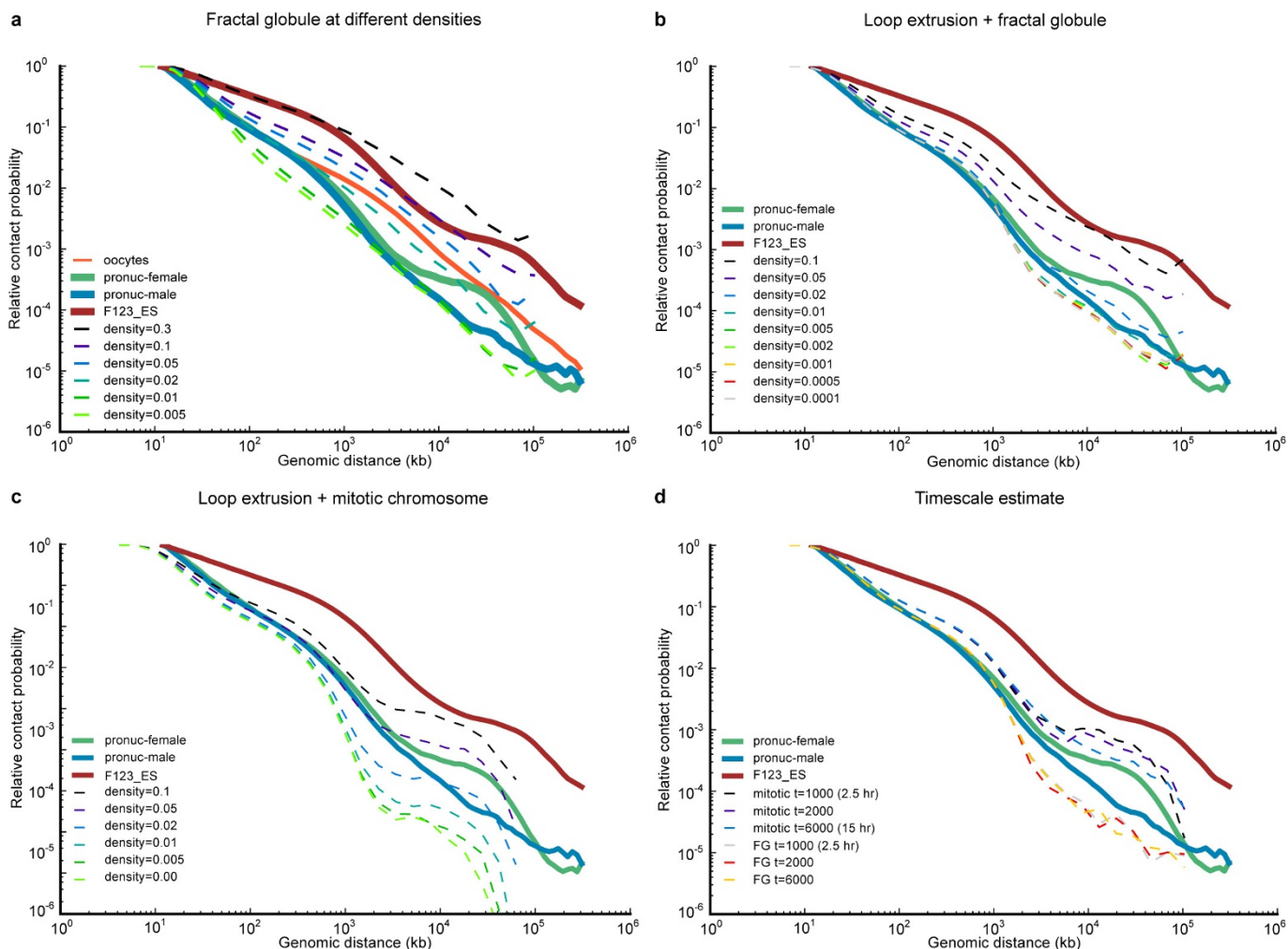**Supplementary Information** is available in the online version of the paper.

Extended Data Figure 1 | Comparison of conventional and strong fixation conditions for Hi-C. $P_c(s)$ of contact probability over genomic separation has similar shape under conventional (1% of formaldehyde for 10 min) and strong (2% of formaldehyde for 15 min) fixation conditions (one replicate). $P_c(s)$ plot for the CH12-LX cell line is constructed using previously published *in situ* Hi-C data[8] and is normalized to integrate to 1.

a Fractal globule at different densities

b Loop extrusion + fractal globule

c Loop extrusion + mitotic chromosome

d Timescale estimate

Extended Data Figure 2 | Simulations of $P_c(s)$ of oocytes, maternal and paternal nuclei. a–c, $P_c(s)$ for various polymer models. All simulated $P_c(s)$ curves were calculated using contact radius of 10 monomer diameters (100 nm). Decondensed fractal globule (**a**), loop extrusion model starting with fractal globule (**b**), loop extrusion model starting with mitotic chromosome (**c**). **d**, Simulations in **a–c** and in Fig. 4h were run for 2,000 loop extrusion steps, which represents around 5 h of real time (see Supplementary Methods). In reality, zygotes spent 7–10 h

after fertilization. To ensure that $P_c(s)$ does not change greatly over this timescale, we simulated one run for three times longer (6,000 loop extrusion steps). Note that as this figure was obtained from only two simulations, and not an average of many, and therefore the $P_c(s)$ does not exactly match Fig. 4h. Even after 6,000 loop extrusion steps, the $P_c(s)$ curves starting with the fractal globule and with mitotic chromosome model are very distinct, and different by almost two orders of magnitude at 10 Mb.
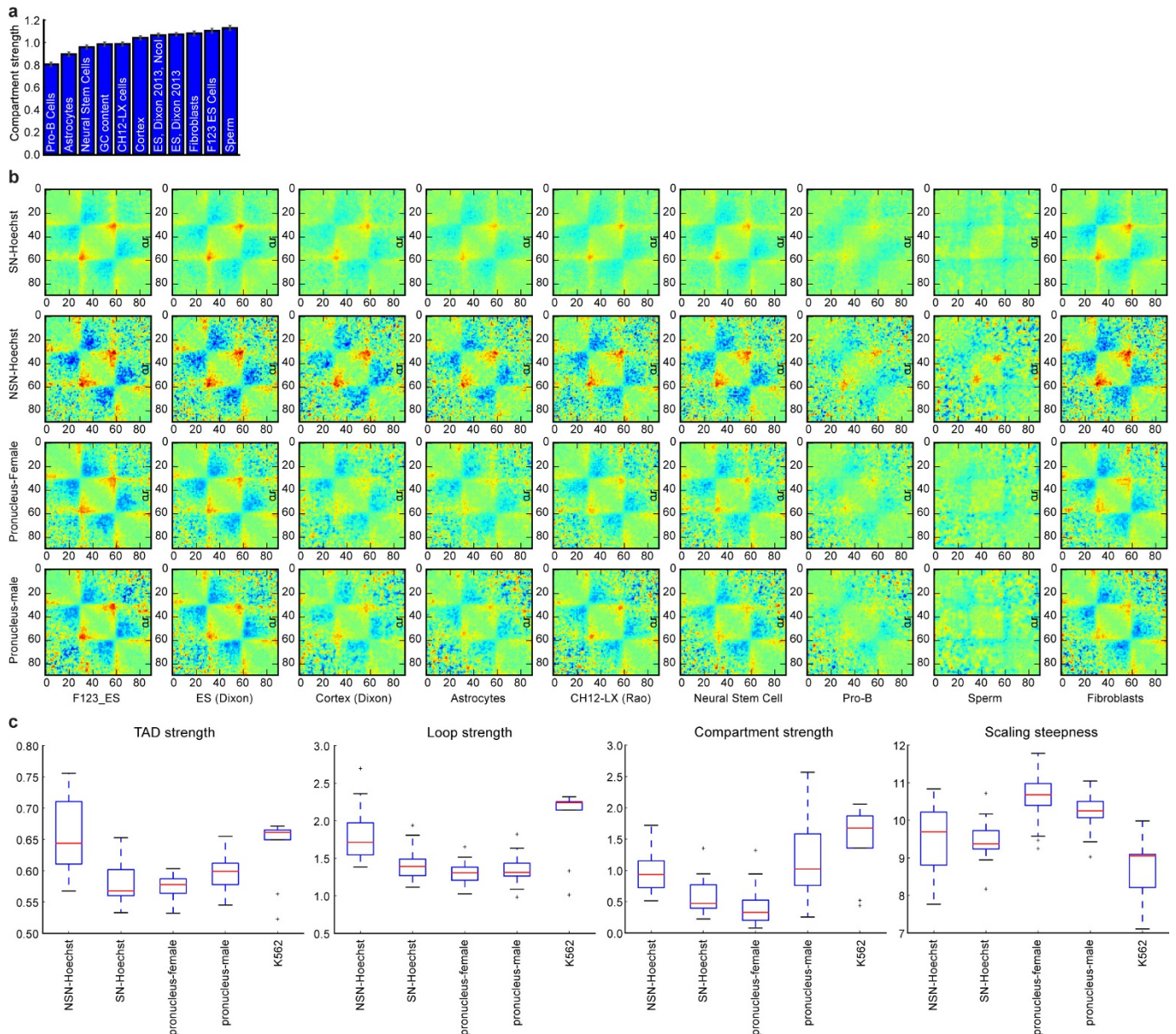
Extended Data Figure 3 | Quantification of average features in Hi-C maps. a, We quantified compartment strengths in oocytes using compartment annotation from different published data sets and quantification of compartment strength indicates that oocyte compartments are most similar to sperm, mouse ES cells and fibroblasts chromatin. Error bars as in Fig. 3d. **b**, Average TADs calculated over TADs computed from various cell types. Note that high-resolution TAD calling is only available in CH12-LX cells[8]. For this figure, all TADs were all computed using the 'lavaburst' algorithm described in the Supplementary Methods. The value plotted here is the natural log of observed-over-expected of the TAD enrichment. Unlike plots in the main figures, these are true observed-over-expected probabilities, not 'effective contact probability'. The colour map is jet, ranging from −0.5 to 0.5. **c**, TAD, loop and compartment strength as well as scaling steepness (definitions are in Supplementary Methods) in different classes of cells. Boxplots were generated using 'matplotlib' (version 1.5.1) library for Python with default parameters.
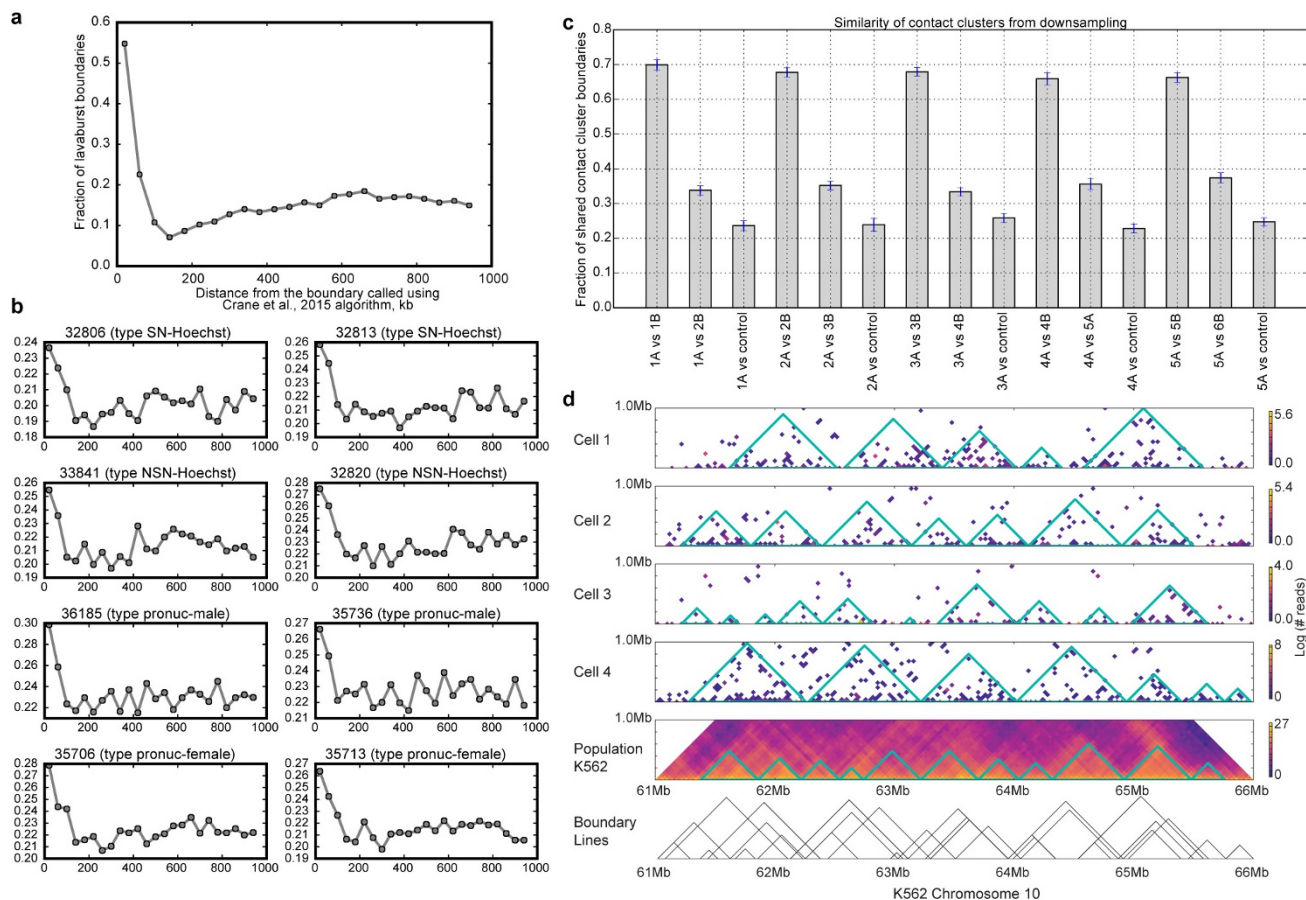
**Extended Data Figure 4 | Stochasticity of contact clusters and validation of contact cluster annotation algorithm. a**, For this figure, boundaries were called on population data from CH12-LX cells[8] at 20 kb resolution using two different methods: lavaburst with modularity score, with an average domain size of 25 bins (500 kb), and a method from ref. 20, downloaded from https://github.com/dekkerlab/crane-nature-2015 (most recent commit, August 28, 2016). The latter method was used with default parameters, on whole-chromosome heat maps. The plot shows fraction of lavaburst boundaries that are located within a certain distance of the boundaries defined in ref. 20; step is 40 kb. Modularity score boundaries align very well with boundaries called using an algorithm from ref. 20. For example, 77% of boundaries called using modularity score were within an 80 kb of algorithm boundaries from ref. 20 (32% expected if boundaries were randomized by offsetting them by 1 Mb). **b**, Same as **a**,

but for top two single cells in each set. **c**, Contact cluster calling is robust to downsampling. From each of the top five single-cell oocytes, we obtained two maps down-sampled by 50% (1A and 1B from cell 1, etc.). We then compared contact clusters called in the two same-cell downsampled maps to each other (1A versus 1B), two maps from different cells (1A versus 2B), and each map to its randomly-shuffled control (1A versus control). Two maps from the same cell overlap by 65–70% of domain boundaries with 80 kb error margin. Overlap between different cells is about 1.5 times less (30–40%), and overlap with the reshuffled control is about 20–30%. Displayed are the average over all chromosomes and 95% confidence intervals of the fraction of overlap. **d**, The Hi-C contact cluster annotation of the top four single cell K562 cells is compared with the published population Hi-C map[8].
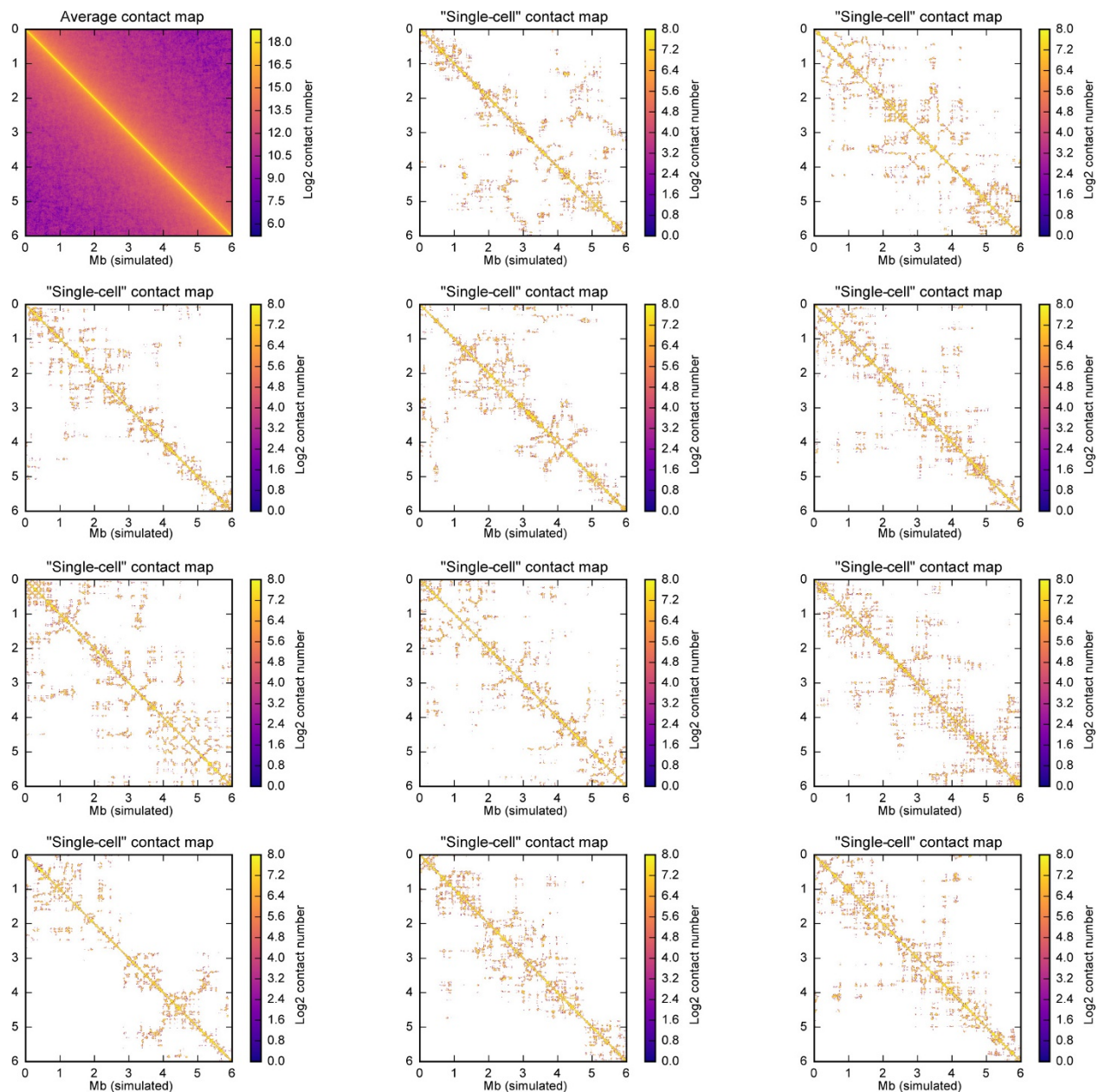
Extended Data Figure 5 | Contact clusters can form in polymer models with no average structure. This figure shows contact maps of a 10,000 monomer region in fractal globules at density 0.05 (see Supplementary Methods for the fractal globule creation descriptions). Each contact map was calculated with a contact radius of 10, at bin size of 16 monomers (approximately 10 kb, if we assume 600-bp monomers as in the other models or in refs 11, 22). First map (top left) shows a population average contact map calculated from 2,000 independent realizations. Fractal globule is a model in which monomers are all treated equally and have no specific organization; therefore, a population average contact map of the fractal globule would be completely uniform (for example, contact probability only depends on the distance between the two regions). Each of the remaining 11 maps shows a 'single-cell' contact map from

11 single conformations. Note the high degree of variability between single-conformation contact map, despite the complete homogeneity of the average contact map. See supplementary figure 17 in ref. 11 for similar maps from our model of mitotic chromosomes. Note that, unlike in Hi-C, where each fragment end can form only one contact, in our simulations we record all contacts happening within the contact radius 10, and each monomer can form many contacts. Thus, this map shows more contacts than a single-cell Hi-C map would, even if Hi-C had the same capture radius. The map thus shows all potential contacts that could be extracted from a single conformation if sn-HiC was 'performed' on the same conformation many times, each time choosing one neighbour within the contact radius of 10.
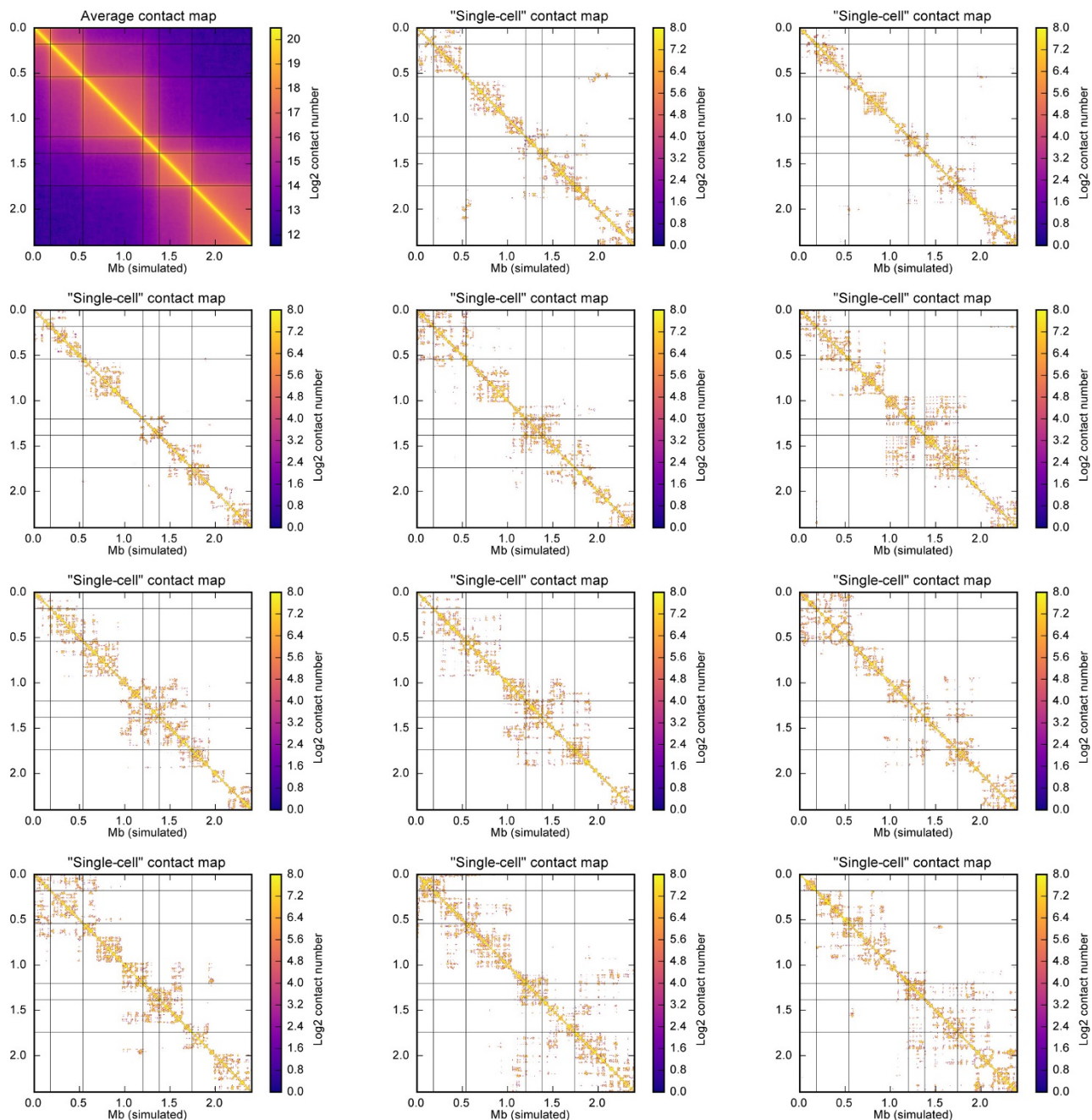
Extended Data Figure 6 | TADs are not visible in single polymers undergoing loop extrusion. Similar to Extended Data Fig. 5, but for our model of loop extrusion starting with mitotic-like conformation (maternal nuclei). In this model, a 77-Mb chromosome (600-bp monomers; 128,000 monomers) is divided into 64 blocks of 3 TADs each. TAD sizes are 300, 600, and 1,100 monomers (180 kb, 360 kb and 660 kb). See ref. 22 and Supplementary Methods for the description of the model. Thin grey lines denote TAD boundaries on all heat maps. Each panel shows a block of 6 consecutive TADs, 4,000 monomers, or 2.4 Mb. Contact map is calculated at contact radius 10, and for bin size of 6 kb (10 monomers). For a population average map, 15,000 conformations were used. From each of 50 independent runs, we sampled 10 conformation at block numbers 1,100, 1,200,..., 2,000. From each conformation, we sampled 30 non-overlapping blocks of 6 TADs (4,000 monomers) starting at monomers 4,000, 8,000, 12,000, …, 120,000) totalling 15,000 blocks. Single-cell map was calculated from a single randomly chosen block.
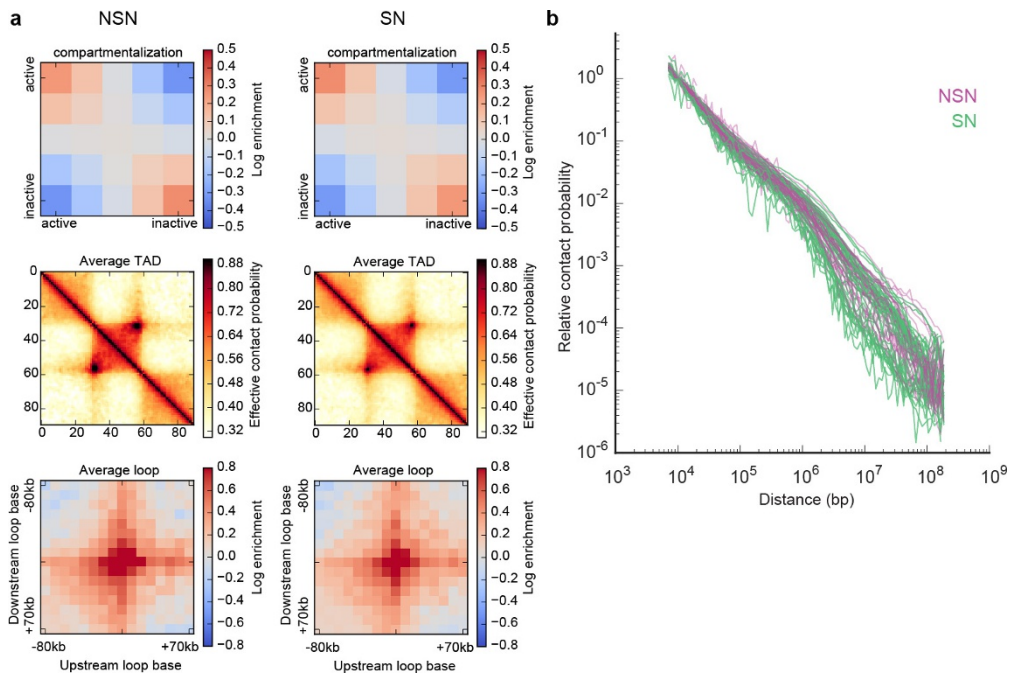
Extended Data Figure 7 | sn-HiC results for NSN and SN oocytes sorted by DIC scoring. Note that Hoechst staining (see Fig. 3) is necessary for proper sorting of NSN and SN oocyte populations. **a**, Compartment signal, average TAD, average loop in oocytes staged by DIC with no DNA staining (n (NSN) = 29, n (SN) = 40, from more than three biological replicates using 2–4 females). **b**, $P_c(s)$ (for cells with >30,000 contacts, n (NSN) = 25, n (SN) = 30, from more than three biological replicates using 2–4 females) for oocytes staged by DIC with no DNA staining.
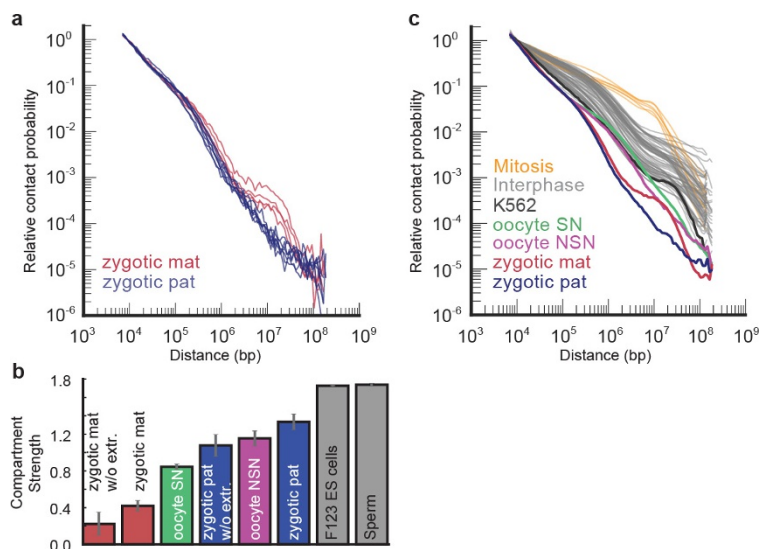
Extended Data Figure 8 | $P_c(s)$ and compartment strength in zygote nuclei in comparison to other cell types. a, $P_c(s)$ for maternal (mat) and paternal (pat) zygotic nuclei with >30,000 total contacts analysed without nuclear extraction (n (maternal) = 4, n (paternal) = 7, from more than two biological repli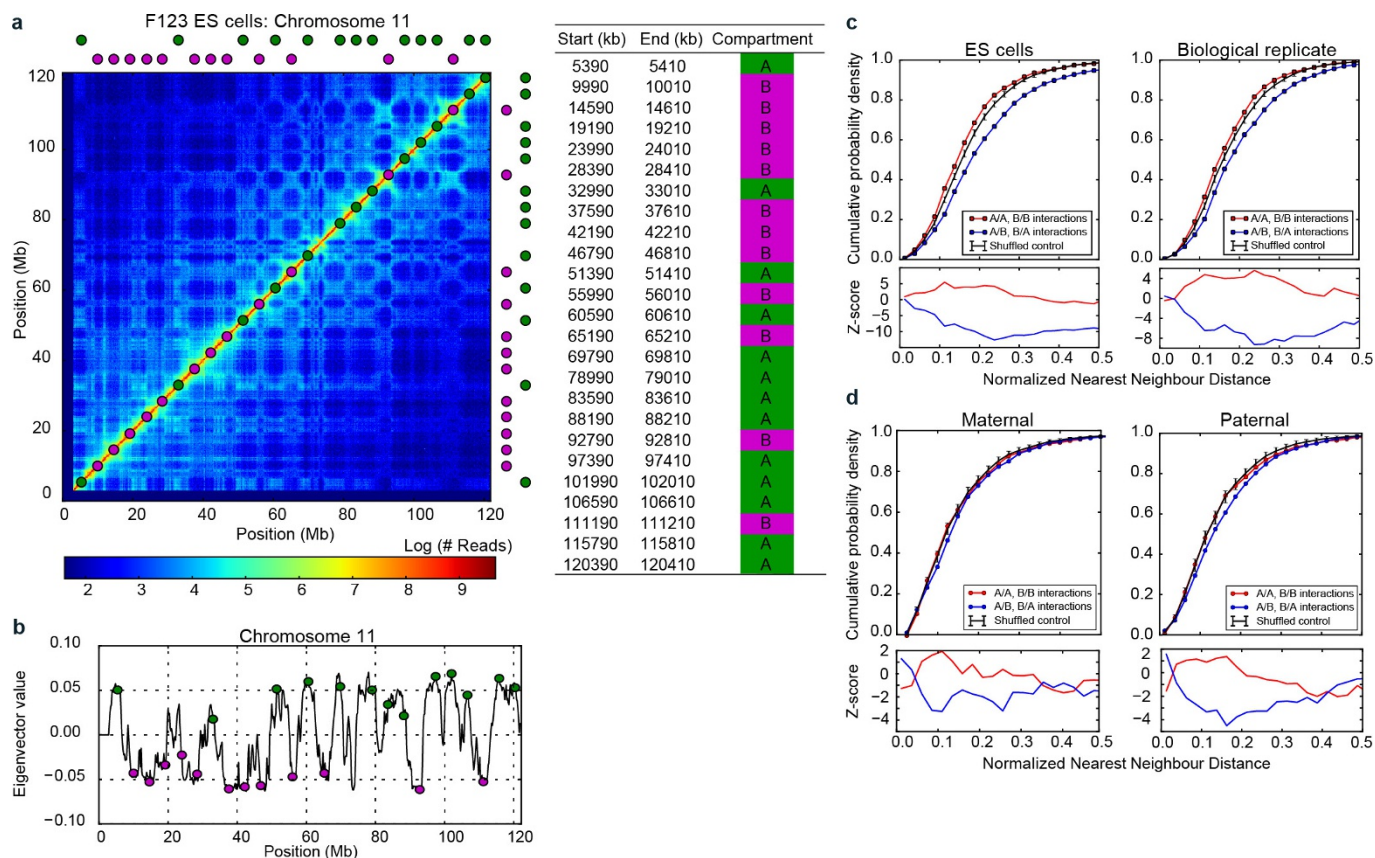cates using 4–6 females). **b,** Comparison of compartment signal strength in combined maternal and paternal zygote nuclei with or without using nuclear extraction, with NSN and SN oocytes (staged with Hoechst staining), and published ES cell[30] and sperm[25] data. **c,** $P_c(s)$ for K562 cells, paternal and maternal nuclei, NSN and SN oocytes (this work), interphase cells[6,8,11,15,30–40] and mitotic chromosomes[11].

31. Dixon, J. R. *et al.* Chromatin architecture reorganization during stem cell differentiation. *Nature* **518,** 331–336 (2015).
32. Zuin, J. *et al.* Cohesin and CTCF differentially affect chromatin architecture and gene expression in human cells. *Proc. Natl Acad. Sci. USA* **111,** 996–1001 (2014).
33. Jin, F. *et al.* A high-resolution map of the three-dimensional chromatin interactome in human cells. *Nature* **503,** 290–294 (2013).
34. Wang, Z. *et al.* The properties of genome conformation and spatial gene interaction and regulation networks of normal and malignant human cell types. *PLoS One* **8,** e58793 (2013).
35. Sofueva, S. *et al.* Cohesin-mediated interactions organize chromosomal domain architecture. *EMBO J.* **32,** 3119–3129 (2013).
36. Zhang, Y. *et al.* Spatial organization of the mouse genome and its role in recurrent chromosomal translocations. *Cell* **148,** 908–921 (2012).
37. McCord, R. P. *et al.* Correlated alterations in genome organization, histone methylation, and DNA-lamin A/C interactions in Hutchinson–Gilford progeria syndrome. *Genome Res.* **23,** 260–269 (2013).
38. Lin, Y. C. *et al.* Global changes in the nuclear positioning of genes and intra- and interdomain genomic interactions that orchestrate B cell fate. *Nat. Immunol.* **13,** 1196–1204 (2012).
39. Kalhor, R., Tjong, H., Jayathilaka, N., Alber, F. & Chen, L. Solid-phase chromosome conformation capture for structural characterization of genome architectures. *Nat. Biotechnol.* **30,** 90–98 (2011).
40. Chandra, T. *et al.* Global reorganization of the nuclear landscape in senescent cells. *Cell Reports* **10,** 471–483 (2015).
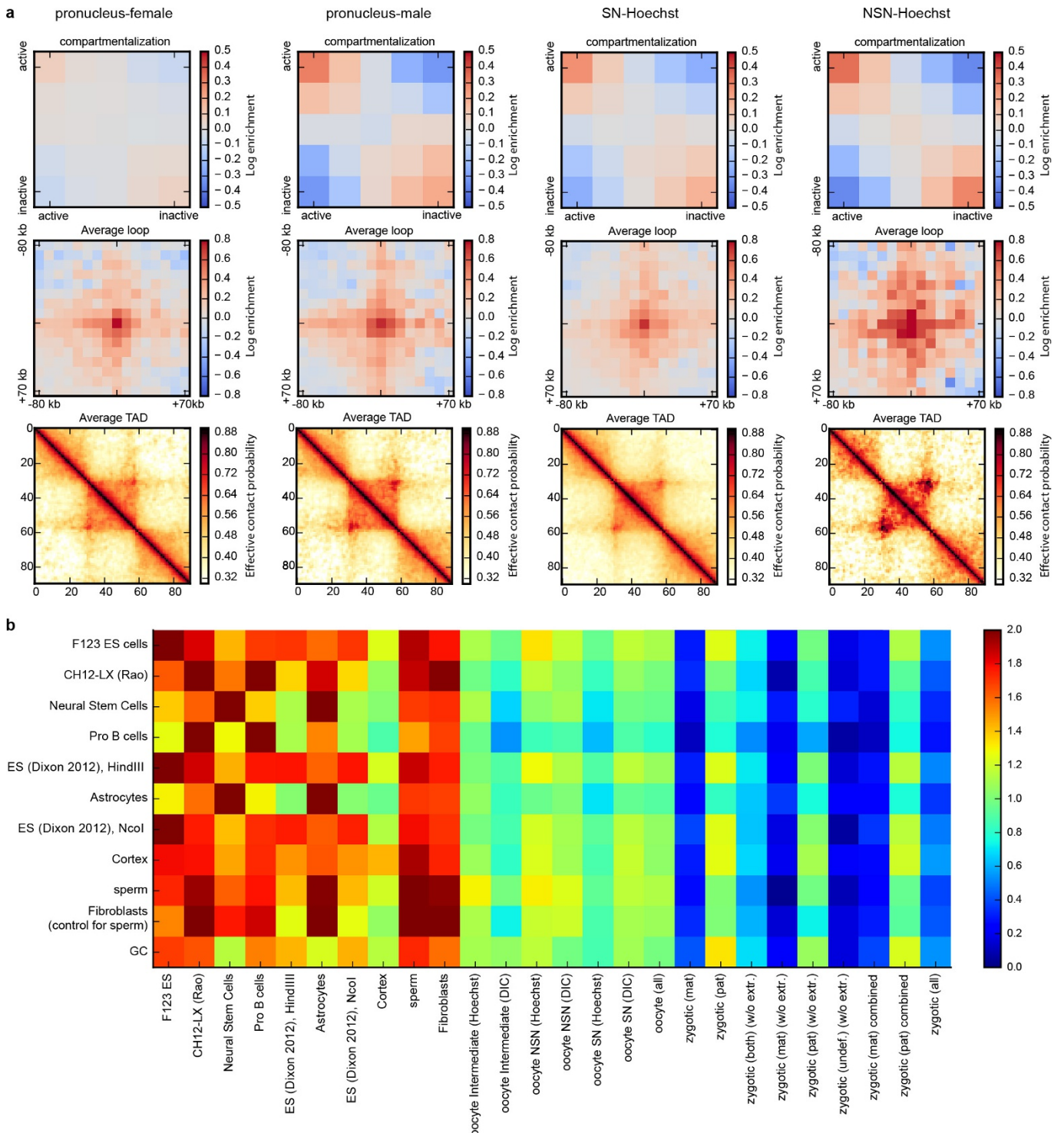
Extended Data Figure 9 | Comparison of all mm9 data sets. a, Same as Fig. 4b, but for oocytes and zygotic nuclei together. **b,** Compartment strength quantified in different data sets (columns, both published and from this study) on the basis of compartment annotation from published data sets (rows). The highest values in each column represent cell types that are most similar to the data of interest. Note that the first nine columns have the highest value on the main diagonal, which correspond to compartment strength evaluated using eigenvector (compartment profile) from the same data set. Also note that cortex cells have similar compartment strength to oocytes and paternal zygotic nuclei.

a F123 ES cells: Chromosome 11

Start (kb)	End (kb)	Compartment
5390	5410	A
9990	10010	B
14590	14610	B
19190	19210	B
23990	24010	B
28390	28410	B
32990	33010	A
37590	37610	B
42190	42210	B
46790	46810	B
51390	51410	A
55990	56010	B
60590	60610	A
65190	65210	B
69790	69810	A
78990	79010	A
83590	83610	A
88190	88210	A
92790	92810	B
97390	97410	A
101990	102010	A
106590	106610	A
111190	111210	B
115790	115810	A
120390	120410	A

Extended Data Figure 10 | Design and validation of FISH probes for quantification of compartments. a, FISH probe design for quantifying compartment segregation; left: probe locations are shown superimposed on the Hi-C data (200 kb resolution from F123 ES cells[30], right: exact locations of designed probes. **b**, Probe locations shown relative to the profile of compartment strength (200 kb resolution) as measured by the first eigenvector of the Hi-C map eigenvector decomposition. **c, d**, Top: nearest neighbour FISH distances—the same as curves in Fig. 4d—but shown for ES cells (n (replicate 1) = 87, n (replicate 2) = 78) (**c**) and maternal (n = 33) and paternal (n = 37) zygotic nuclei. Data from one biological replicate using four females. **d**, Bottom: z-scores showing the number of standard deviations from the expected minimum distance distribution of the control data; the control distribution was obtained from randomly reshuffling probe colours as described in the Supplementary Methods.