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

Supplementary Methods

Mouse oocyte collection, culture and formaldehyde fixation

The care and use of the mice were carried out in agreement with the authorizing committee according to the Austrian Animal Welfare law and the guidelines of the International guiding principles for biomedical research involving animals (CIOMS, the Council for International Organizations of Medical Sciences). Fully grown denuded GV (germinal vesicle) oocytes were isolated from 2-3 month old C57BL/6J female mice by puncturing of ovaries with hypodermic needles[29]. Oocytes were incubated for at least 2 hours before the procedure to ensure viability. Before formaldehyde fixation, the *zona pellucida* was removed by treatment with Acidic Tyrode's solution (Sigma). Oocytes were fixed in with 2% formaldehyde for 15 minutes at room temperature. Oocytes were washed through drops of complete M2. For SN and NSN distinction oocytes were stained for 15 min with 0.2 µg/ml Hoechst (33342, Invitrogen). Cells were then imaged using a confocal microscope (LSM510 Axiovert 200M, Zeiss) and sorted according to their chromatin structure before proceeding with the Hi-C protocol. No blinding or randomization were used for handling of the cells during the snHi C protocol. No statistical methods were used to estimate the sample size.

Mouse zygote collection, culture and formaldehyde fixation

3-5-week-old C57BL/6J females were superovulated by intraperitoneal injection of PMSG (pregnant mare's serum gonadotropin; 5 IU, Folligon, Intervet) followed by hCG (human chorionic gonadotropin; 5 IU, Chorulon, Intervet) injection 48 hours later. Natural matings to C57BL/6J males were set up overnight. Zygotes were released from the ampullae and treated with hyaluronidase (Sigma) to remove surrounding cumulus cells. For isolation of maternal and paternal nuclei, zygotes around 19-22 hours post hCG injection (corresponding to about 7-10 hours post fertilization) were pre-incubated in culture medium supplemented with 5 µg/ml cytochalasin B (Sigma) and 1 µM nocodazole (Sigma) for 15 minutes. The *zona pellucida* of the zygotes was opened using an XYClone laser ablator (Hamilton Thorne) and nuclei were extracted one by one. The maternal and paternal origin of the nuclei was determined taking size of the nuclei and position in relation to second polar body into account. Nuclei were fixed in separate drops of M2 (not containing BSA) with 2% formaldehyde for 15 minutes. To stop the reaction, nuclei were washed through drops of complete M2. For zygote experiments without pronuclear extraction, zygotes were treated identically to oocytes. Here nuclei were separated into individual wells after SDS treatment and kept separate in the subsequent steps of the protocol. No blinding or randomization were used for handling of the cells during the snHi C protocol. No statistical methods were used to estimate the sample size.

Single-nucleus Hi-C on mouse oocytes and zygotes

After fixation, single oocytes, zygotes or isolated nuclei of zygotes were washed through 9 µl of lysis buffer (10 mM Tris-HCl pH 8.0, 10 mM NaCl, 0.5% (v/v) NP-40 substitute (Sigma), 1% (v/v) Triton X-100 (Sigma), 1× Halt™ Protease Inhibitor Cocktail (Thermo Scientific)) in a 10 µl well of a microplate (Nunc™ MicroWell™ MiniTrays) and incubated for at least 15 minutes on ice. Similarly cells or extracted nuclei were washed through 9 µl of PBS, then through 9 µl of 1× NEB3 buffer supplemented with 0.6% SDS where they were incubated for 2 hours at 37° with shaking in humidified atmosphere. When starting with whole zygotes, during this stage nuclei were released from the cytoplasm and manually separated and

placed in different wells after SDS lysis; when possible, their maternal or paternal origin was assigned based on size. Then each of the cells or extracted nuclei was washed through a well with 1× DpnII buffer (NEB) with 1× BSA (NEB) and transferred into the next well with the same solution, but also supplemented with 5 U DpnII (NEB; final concentration 0.556 U/ul). After overnight (14-16 hours) incubation in a humidified incubator at 37° the nuclei were washed through 1 well containing 9 µl PBS, then through a well with 9 µl of 1× ligation buffer (50 mM Tris-HCl, 10 mM $MgCl_2$, 1 mM ATP, 10 mM DTT, pH 7.5). After that each nucleus was transferred into a well with 9 µl of 1×ligation buffer containing 5 U T4 DNA ligase (Thermo Scientific) and they were incubated at 16°C for 4.5 hours with 50 rpm rotation, and then were kept for 30 min at room temperature. DNA from each nucleus was amplified using illustra GenomiPhi v2 DNA amplification kit (GE Healthcare) according to a published[31] protocol or using illustra Single Cell GenomiPhi DNA amplification kit according to manufacturer's recommendations, but both protocols were modified to include a decrosslinking step of an overnight (14-16 hours) incubation at 65°C. After the whole-genome amplification procedure the DNA was purified using AMPure XP magnetic beads (Beckman Coulter) and quantified using NanoDrop™ 1000 spectrophotometer (Thermo Scientific). 1 µg (or all available amount, if the yield was lower) of DNA was diluted to 500 µl with sonication buffer (50 mM Tris-HCl pH 8.0, 10 mM EDTA, 0.1% SDS) and was sonicated using Branson Sonifier 150 for 20 seconds with power setting 3. The sonicated DNA was purified either on AMPure XP magnetic beads (Beckman Coulter), or on QIAquick PCR Purification Kit (QIAGEN), and was submitted for library preparation (with 300-1,300 bp size selection) and Illumina sequencing (paired-ends 125 bp each side on HiSeq 2500 v4) to VBCF NGS Unit (csf.ac.at), between 10 and 24 cells per lane. Experiments were repeated several times independently by two researchers in the lab.

Single-nucleus Hi-C on K562 cells

K562 (a gift from Alexander Stark lab) cells were cultured in RPMI-1640 medium supplemented with 10% FBS, 2 mM L-Glutamine and Pen/Strep (100 U/ml and 100 µg/ml). 2-4 million cells were fixed for 10 minutes by resuspending the cell pellet in 5 ml full culture medium supplemented with 1% formaldehyde (Sigma). The reaction was quenched by addition of 2 M glycine to a final concentration of 125 mM and incubation for 5 minutes on ice. Cells were pelleted by centrifugation, washed by resuspension in 1 ml of PBS and pelleted again. Then the cells were resuspended in 1.5 ml of lysis buffer (50 mM Tris-HCl pH 8.0, 150 mM NaCl, 0.5% (v/v) NP-40 substitute (Sigma), 1% (v/v) Triton-X100 (Sigma), 1× Halt™ Protease Inhibitor Cocktail (Thermo Scientific)) and incubated on ice for at least 15 minutes. Cells were pelleted by centrifugation at 2,500g for 5 minutes, resuspended in 100 µl of 1×NEBuffer 3 and pelleted again. The pellet was resuspended in 100 µl of 0.3% SDS in 1×NEBuffer 3 and incubated at 37°C for 1 hour. Then the suspension was diluted with 330 µl of 1× NEBuffer 3 and 53 µl of 20% Triton X-100 (Sigma), and incubated at 37°C for 1 hour to quench SDS. The cells were pelleted by centrifugation at 3,000g for 5 minutes, resuspended in 250 µl of 1× DpnII buffer (NEB). 20 µl were taken as a chromatin integrity control, then 600 U DpnII enzyme (NEB) were added, and the chromatin was digested overnight (14-16 hours) at 37°C with rotation. In the morning 200 more units of DpnII were added and the cells were incubated for additional 2 hours. DpnII was inactivated by incubation at 65°C for 20 minutes. 20 µl was taken as a digestion control. Nuclei were pelleted by centrifugation at 3,000g for 5 minutes, resuspended in 100 µl of 1× T4 DNA ligase buffer (Thermo Scientific) and pelleted again. The pellet was resuspended in 500 µl of 1× T4 DNA ligase buffer, and 50 U T4 DNA ligase (Thermo Scientific) were added. The sample was mixed by inversion and incubated with rotation for 4 hours at 16°C and then for 30 minutes at room temperature. Then the nuclei were pelleted by centrifugation at 3,000g for 5 minutes and resuspended in 120 µl of sterile PBS. 20 µl were taken as a ligation control and 20 µl were taken for DNA purification and

subsequent preparation of the bulk Hi-C library. The rest of the sample was stained with Hoechst, single G1 nuclei were sorted using excitation wavelength 375 nm and forward and side scatter by FACS (FACSAriaIII machine, BD Biosciences) into wells of 12 PCR strips containing 3 µl sample buffer from illustra GenomiPhi v2 DNA amplification kit (GE Healthcare), then they were covered by 5 µl of mineral oil, incubated at 65°C overnight and amplified according to a published[31] protocol. The amplification was considered positive if the sample contained ≥1 µg DNA, and those samples were prepared for sequencing in the same way as oocyte and zygote samples. Two biological replicates of the experiment were performed.

Hi-C on MEL cells

MEL cells (ATCC) were cultured in RPMI-1640 medium supplemented with FBS and Pen/Strep (100 U/ml and 100 µg/ml). Hi-C was performed as described previously[8] with minor modifications. 7.5 million cells were fixed for 10 or 15 minutes by resuspending cell pellet in 7.5 ml full culture medium supplemented with 1% or 2% formaldehyde (Sigma), respectively. The reaction was quenched by addition of 2M glycine to a final concentration of 0.2 M and incubation for 5 minutes at room temperature on a rocker. Cells were pelleted by centrifugation at 300 g for 5 minutes at 4°C, washed by resuspension in 1 ml of PBS, pelleted and resuspended in 1 ml of PBS again. 1.9 mln cells from each sample were taken, the rest were snap-frozen for other experiments. The samples were pelleted again, then resuspended in 300 µl of cold lysis buffer (10 mM Tris-HCl pH 8.0, 10 mM NaCl, 0.2% NP-40 substitute (Sigma), 1× Halt™ Protease Inhibitor Cocktail (Thermo Scientific)) and incubated on ice for at least 15 minutes. Cells were pelleted by centrifugation at 2,500g for 5 minutes, resuspended in 500 µl of lysis buffer and pelleted again. The pellet was resuspended in 50 µl 0.3% SDS in NEBuffer 3 and incubated at 62°C for 10 minutes. Then the suspension was diluted with 145 µl of 1× NEBuffer 3, 12.5 µl mQ and 12.5 µl of 20% Triton X-100 and incubated at 37°C for 1 hour to quench SDS. 30 µl were taken as a chromatin integrity control. The cells were pelleted by centrifugation at 3,000g for 5 minutes, resuspended in 250 µl of 1× DpnII buffer (NEB) with 600 U DpnII enzyme (NEB), and the chromatin was digested overnight (14-16 hours) at 37°C with rotation. In the morning 200 more units were added and the cells were incubated for additional 2 hours. DpnII was inactivated by incubation at 65°C for 20 minutes. 40 µl was taken as a digestion control, 40 µl of mQ were added to the sample. The ends were filled-in by addition of 50 µl fill-in mix (36 ul mQ, 1.5 µl of 10 mM biotin-14-CTP, 1.5 µl of 10 mM dATP, 1.5 µl of 10 mM dGTP, 1.5 µl of 10 mM dTTP, 8 µl of 5 U/µl DNA Polymerase I, Large (Klenow) Fragment (NEB)) and incubation at 37°C for 1.5 hours. Then 900 µl of ligation mix (712 µl water, 120 µl 10× NEB T4 DNA ligase buffer (NEB), 50 µl 20% Triton X-100, 12 µl 10 mg/ml BSA (NEB), 6 µl 5 U/µl T4 DNA Ligase (Thermo Scientific)) were added to the sample. The sample was mixed by inversion and incubated with rotation for 4 hours at room temperature. Then the nuclei were pelleted by centrifugation at 3,000g for 10 minutes and resuspended in 200 µl mQ. Proteins were degraded by addition of 15 µl of 20 mg/ml proteinase K and 20 µl of 10% SDS and incubation at 55°C for 30 minutes. Then 130 µl 5M NaCl were added and the samples were incubated at 65°C overnight (14-16 hours). The DNA was ethanol precipitated and then dissolved in 500 µl 10 mM Tris-HCl pH 8.0, washed twice on Amicon 30K filter units (Millipore) with 10 mM Tris-HCl pH 8.0 and then eluted into a clean tube. The volume was adjusted to 500 µl with sonication buffer (50 mM Tris, 10 mM EDTA, 0.1% SDS) and the samples were sonicated using VirSonic 100 (Virtis) with power setting 15 for 15 seconds twice with 1 minute interval. Samples were washed twice again using Amicon 30K filter units (Millipore). All consecutive steps until PCR were performed in LoBind tubes (Eppendorf). 30 µl of 10 mg/ml Dynabeads MyOne Streptavidin T1 beads (Life technologies) were prepared for biotin pull-down by washing with 80 µl of 1× TWB (1× Tween Washing Buffer, 5 mM Tris-HCl pH 7.5, 0.5 mM EDTA, 1 M NaCl, 0.05% Tween 20). Then the beads were resuspended in 60 µl of 2× BB (2× Binding Buffer: 10 mM

Tris-HCl pH 7.5, 1 mM EDTA, 2 M NaCl) and added to the samples; the suspension was incubated for 15 minutes at room temperature with rotation. After separation on a magnet, the beads were washed with 120 µl of 1× TWB and the mixture was transferred to a new tube and heated to 55°C for 2 minutes; the wash was repeated again. The beads were resuspended in 20 µl of 1× T4 DNA Ligase buffer (NEB), transferred to a new tube, separated on a magnet and the buffer was discarded. The beads were resuspended in 100 µl of end-repair mix (85 µl of 1× T4 DNA ligase buffer with 10 mM ATP (NEB), 5 µl of 10 mM dNTP mix, 5 µl of 10 U/µl NEB T4 PNK (NEB), 4 µl of 3U/µl T4 DNA polymerase I (NEB), 1 µl of 5 U/µl NEB DNA polymerase I, Large (Klenow) Fragment (NEB)) and incubated for 30 minutes at room temperature. The beads were separated on a magnet and washed with 120 µl of 1× TWB, transferred to a new tube and heated up to 55°C for 2 minutes; the wash was repeated again. The beads were washed similarly with 100 µl of 1× NEBuffer 2 without heating. The beads were resuspended in 100 µl of dATP attachment mix: 90 µl of 1× NEBuffer 2, 5 µl of 10 mM dATP, 5 µl of 5 U/µl Klenow exo minus (NEB), incubated at 37°C for 30 minutes and then separated on a magnet. The beads were washed twice with 120 µl of 1× TWB same way as before. The beads were washed similarly with 100 µl of 1× T4 DNA Ligase buffer (NEB) and then resuspended in 50 µl of 1× T4 DNA Ligase buffer (NEB). 1 µl of T4 DNA ligase (Thermo Scientific) and 3 µl of an Illumina indexed adaptor were added to each tube, the suspension was mixed thoroughly and incubated at room temperature for 2.5 hours with rotation. The beads were separated on a magnet and washed with 50 µl of 1× TWB twice same as before. The beads were once washed in the same way with 100 µl of 1× Tris-HCl pH 8.0, then resuspended in 50 µl of the same Tris buffer. Finally, the Hi-C libraries were amplified by 13 cycles of PCR directly from 2 µl of the beads in 7 tubes using Illumina primers and protocol (Illumina, 2007) and KAPA HiFi polymerase, all reactions from each sample were combined, concentrated on Amicon filter units with two 10 mM Tris-HCl pH 8.0 washes, gel-purified with excision of fragments 300-800 bp long and submitted for Illumina sequencing (paired-end 101 bp each side on HiSeq 2000).

ES cell culture and 3D DNA FISH

E14 ES cells were grown feeder-free on gelatin coated coverslips in GMEM (Life Technologies) supplemented with 10% FCS, 2 mM L-Glutamine, 1 mM Sodium Pyruvate (Sigma), 1× Non-Essential Amino Acids (Sigma), 50 µM β-mercaptoethanol (Gibco), Pen/Strep (100 U/ml and 100 µg/ml) and LIF. Paraformaldehyde (pFA)-fixed cells were permeabilized in 0.5% Triton X-100, washed in PBS, and stored at -80°C. 3D-FISH was performed as in ref. 32. When using oligo probes no Cot1 or ssDNA was used, and instead of pre-annealing after denaturation probes were snap-cooled in ice-cold water. Fosmid probes used for TAD boundary violation analysis were WI1-1516E02, WI1-2538D05 and WI1-1060D24 (BACPAC). They were labelled with Biotin-16-dUTP (Roche), Alexa 594-5-dUTP (Life Technologies) and 5-Fluorescein dUTP (Enzo), respectively, same way as in ref. 32. 20 Kb long oligos probe libraries for compartment segregation analysis (MYtags®) labelled with atto-488 and atto-594 dyes to selected A- and B-compartments (see below), respectively, were designed and synthesized by MYcroarray® (Ann Arbor, MI). For imaging, slides were counterstained with DAPI and imaged on Zeiss Axioskop II MOT fluorescence microscope using Plan-neofluar or Plan apochromat 100×, 1.4 NA objectives with 64×64×200 nm voxel size; hardware was controlled using Volocity software (PerkinElmer Inc, Waltham, MA). Images were then deconvolved using calculated PSFs in Volocity (PerkinElmer Inc, Waltham, MA) with fast algorithm before positions of the probes were determined using spot detection in Imaris (8.1.2). 2 biological replicates were performed and combined data are reported with 2 outlier measurements manually removed.

3D DNA FISH on zygotes

3D DNA FISH on zygotes was performed as described previously[33] with some modifications. In brief zygotes generated by natural mating as described above were fixed in 4% PFA supplemented with 0.1% Triton X-100 for 15 min, the zona pellucida was removed using acidic Tyrode's before permeabilization with 0.2% Triton X-100 in PBS for 30 min at room temperature. Embryos were embedded in fibrin clots, washed briefly with 2% Triton X-100 in PBS before post-fixation in 2% PFA in PBS for 30 min. Permeablization in 0.1M HCl, 0.7% Triton X-100 was followed by two washes in 2x SSC. Denaturation was performed at 80°C for 45 min in 70% formamide, 2xSSC. Hybridization of probes was done overnight in 50% formamide, 2x SSC, 10% dextran sulfate, 1% Tween 20. The same probes were used as in ES cells. Washing steps were performed in 2x SSC at 45°C followed by 0.1x SSC at 60°C and 4x SSC, 0.1% Triton X-100. Slides were counterstained with Hoechst and imaged on a confocal microscope (LSM780 Axio Observer, Zeiss) using a 63×, 1.4 NA objective lens. Images were acquired with 1024 x 1024 pixels frame size, 65.9 x 65.9 x 200 nm voxel size, 2x zoom, 4 line averages and 0.79 µsec pixel dwell. Images were deconvolved using Huygens Remote Manager (v3.2.2), before positions of the probes were determined using spot detection in Imaris (8.0.2). Cells from 4 females were subjected to the protocol together and analysed separately.

Compartment FISH probe design and data analysis

FISH probes were designed to equalize the likelihood of interaction between similar and different compartment types in the absence of a segregating force (i.e. compartmentalization). A sequence of probes was generated in two colours such that number of A/A, A/B, B/A, and B/B nearest neighbours in the linear genomic sequence were equal. In addition to equal numbers of "like" and "dislike" compartment type nearest neighbours, we constrained our selection of candidate probe locations such that: 1) it covered an entire chromosome, 2) avoided poorly mapping regions, centromeres, telomeres, etc. 3) ensured equal spacing between successive probes, and 4) ensured the probes were as centered as possible in the middle of a compartment to avoid chances of ambiguity in compartment assignment. Compartment profiles were identified from population Hi-C data from mouse F123 ES cells (Selvaraj et al., 2013) mapped to the mm9 genome assembly and binned at 200 kb resolution. Compartments were designated as A or B as described previously, using eigenvector decomposition of within-chromosomal observed-over-expected maps[34]. We searched for an optimal probe design that would conform to the constraints outlined above by iterating through hundreds of possible spacings between probes, and starting positions on each chromosome. Specifying a minimum number of 10 probes for each of A and B compartments, constraining the number of A/A, A/B, B/A and B/B contacts to differ by at most 2, and specifying a threshold deviation from the eigenvector "zero" for A compartments (minimum set to 0.015) and B compartments (maximum set to -0.010), we identified a list of 20 possible candidate probe designs. The candidate probe sets were manually scored based on their mappability, proximity to centromeres, microsatellites and telomeres using the UCSC genome browser[35] and visual inspection of their positioning/centeredness in the compartments as judged by plotting the probe locations on the mESC Hi-C maps. The final probe design is seen in Extended Data Fig. 10a-b; probes were each separated by 4.6 Mbp with the initial probe centered at the 5.39 Mbp locus on chromosome 11; each probe was designed to span 20 kb from the probe center. The chosen design comprised a total of 13 A compartments and 12 B compartments creating a total of 6 A/A, 6 A/B, 6 B/A, 6 B/B nearest neighbours in the linear genome sequence.

For the data analysis, pairwise distances were computed between all identified probe loci for a particular nucleus and the closest neighbour distance for each probe was stored. To control for the variation in nuclear

size between cell types and to allow easier comparison, the maximum separation between probes of any colour was stored and used to normalize the nearest neighbour probe distances. A cumulative density histogram of normalized nearest neighbour distances was computed for each of A/A, A/B, B/A, and B/B interactions for each cell type (ES cells, paternal, and maternal nuclei). To compute the null distribution and expected variance of nearest neighbour contacts, we used a bootstrapping approach: we reshuffled the A/B compartment assignments by randomly choosing new probe designations (A* and B*), such that the absolute numbers of loci were preserved (i.e. #A*=#A and #B*=#B). For each nucleus, the reshuffling procedure was repeated 1000 times, and median and the 95% confidence intervals for the expected variance of the distribution were obtained using robust statistics by ordering the set of 1000 nearest neighbour distances and selecting the values at 50% down the list for the median and 5% and 95% for the lower and upper confidence intervals of minimum distances.

Read processing

Reads were mapped to the genome using *hiclib* software; processing was done using the *hiclib* library, but using the custom pipeline tailored to single cell data described below. Mapping was done using iterative mapping as in ref. 11. Because of the whole-genome amplification in the experimental protocol (rather than PCR in conventional Hi-C), molecular products containing the same ligation junction can start in the vicinity of the ligation junction at different locations. To eliminate counting the same ligation junction many times, we applied the following filter: if two reads map to the same strands, and each side of the read is within 500 bp of any side of the other read, we retained only one copy of the read. This replaced the "PCR Duplicate filter" of the classical Hi-C data processing.

To filter the data further, we separated the genome into 500 bp bins. If any two bins were interacting more than once, only one interaction was counted. Since each region is only present in the single cell up to 4 times (in oocytes), we eliminated all 500 bp bins that form more than 8 unique interactions with other bins. We note that even in the best oocyte (1.6 million reads), each 500bp bin has on average about 0.5 reads. Filtered reads were then binned at different resolutions using read start as a read position. No normalization or iterative correction was applied to the data due to its sparsity.

Notes about previously published Hi-C datasets

In the current manuscript, we performed extensive comparison of our snHi-C data to previously available Hi-C datasets. In particular, we averaged snHi-C data over the positions of loops, TADs and compartments measured in other cell types. As such, detecting TADs and compartments in other cell types was essential for our analyses.

For the human genome, there are datasets from four publications[8,36–38] that have more *cis* (within-chromosomal) reads than any existing mouse dataset as of mid-2015. For each of these human datasets, positions of CTCF-mediated loops and TADs can be easily determined and are publicly available. Unfortunately, there is only one mouse Hi-C dataset with read depth sufficient to reliably call TADs and loops, CH12 cells from ref. 8. If loop and TAD positions were available for more cell types, it would have allowed us to more precisely compare oocytes and zygotes to other cell types. It also suggests that performing similar analyses with human cells opens possibilities to more detailed comparisons of TADs and loops, as for human cells comparisons to many cell types are possible.

Due to the lower coverage of mouse datasets, for some cell types, we had to perform pooling of all replicates of the wildtype and mutant datasets together. The rationale behind this step is that in existing datasets, there is little difference between wild type and mutant datasets in terms of the features of the Hi-C map (compartments, TADs and loops). Loops and TADs are largely stable across cell types[8,15], and thus should be similarly unchanged in different conditions of the same cell types. Even in extreme mutants, such as depletion of core architectural proteins, cohesin or CTCF, the features are present at the same locations, but are different in intensity. We note that the two top datasets, F123 ES and CH12 LX, were from wild type cells only. Pooling of the wildtype and mutant datasets was performed to lift two other Hi-C experiment, NSC and AST datasets, to the read count comparable to the top two datasets.

For comparison with the Hi-C data, we used the datasets shown in the Supplementary Table 1. For human cells, we only used *in situ* datasets from ref. 8. For mouse cells, ref. 8 only provides one dataset. Thus, we used datasets from previous publications.

For $P_c(s)$ calculation, we used most of the currently published mammalian Hi-C datasets as of late 2015. For each publication, replicates for each experiment were pooled together and processed using *hiclib*. The datasets are: ref. 6,8,11,15,30,36–45. All replicates were mapped and filtered independently and pooled together. $P_c(s)$ was computed using *hiclib* with 1kb bins.

Hi-C data analyses

$P_c(s)$ were computed using *hiclib* from data binned at 1 kb. Specifically, we used the function fragmentHiC.HiCDataset.plotScaling(); dataset was initialized with enzyme=1000 which denotes 1kb bins used instead of restriction fragments.

For single-cell data, to calculate the expected # of interaction pairs, we used a list of 1-kb bins from combined (all-hg19 or all-mm9) single-cell datasets, as not all genomic regions may be present in a low-coverage single-cell dataset. $P_c(s)$ was computed by chromosomal arm.

A/B compartment profiles from existing population-average Hi-C were calculated using between-chromosomal data, as described in ref. 34. Enrichment of A/B interactions in Hi-C data were calculated at 1-Mb resolution. For each chromosome, we first focused only on bins that have non-zero sum in the data that is being averaged. An A/B compartment profile or a GC content profile for the current chromosome was then separated into 5 bins: (min to 20th percentile), (20th percentile to 40th percentile), etc. For each pair of bins (25 pairs total), "observed over expected" values were then calculated for loci belonging to each pair of bins. Because the single-cell data was not iteratively corrected, the resulting profile can have a small bias towards more A-A or more B-B interactions; this bias was removed by iteratively correcting the resulting 5×5 (percentile-binned) matrix.

To calculate the error in evaluating the compartment strength, we created 100 5x5 compartment enrichment matrices obtained by bootstrapping. For each pixel of the 5x5 compartment enrichment map, we took all the observed-over-expected values that contributed to this pixel and took a random sample with replacement of the same size that the contributing values. We then proceeded with downstream for each of the 100 reshuffled maps.

To calculate the strength of compartment signal, we took the natural logarithm of the AA * BB / AB^2. For boxplots, we chose the top 20% of eigenvector for A, and the top 20% for B. Instead of the first eigenvector, GC content was used for the main figures, since it is known to correlate with the compartment profile with r~0.8 (see ref 34). To show that the analyses are robust to the way we define compartments, we repeated analyses using the entire A or B compartment, defined as positive or negative values of the eigenvector. The results were comparable (Extended Data Fig. 9b). Error bars for Figures 3d and 4c were calculating by taking a standard deviation of the compartment strength across the 100 maps obtained by bootstrapping.

TADs were averaged over published positions of domains from ref. 8 using data at 10-kb resolution. For mm9, domain positions were taken from the only mouse cell line, CH12-LX. For hg19, the highest quality dataset, GM12878, was used. Domains of sizes between 100 kb and 1 Mb were used. For each domain of length *length*, a map for the region ((*start* - *length*) to (*end* + *length*)) was obtained; this produced a contact that is three times bigger than a given domain. This contact map was then rescaled to a (90×90) pixel map using linear interpolation and block-averaging. In the resulting map, the mid-region pixels 30 to 60 correspond to the TAD body.

TAD strength for boxplots was quantified as the ratio of two numbers. The first number is the within-TAD intensity: the sum of the central square of the enrichment map, rows 30 to 59 and columns 30 to 59 ([30:60, 30:60] in Python slicing notation). The second number is the between-TAD intensity, ½ of the sums of the regions [0:30, 30:60] and [30:60, 60:90].

To visualize TADs in a way that resembles Hi-C contact maps, 90×90 TAD enrichment matrices were then converted to the "effective contact probability". To this end, we multiplied the observed-over-expected value by $(\mathrm{abs}(i - j) + 10\mathrm{kb})^{-0.25}$. This created higher values near the diagonal and smaller values away from the diagonal, thus resembling the visual appearance of a TAD in ordinary Hi-C maps. We did this because it is easier to perceive TAD visually in the ordinary Hi-C map rather than in the observed-over-expected map as done in Extended Data Fig. 3. However, we could not simply perform averaging of the observed Hi-C data because of the differences in $P_c(s)$ across cell types. Thus, to allow for comparisons between cell types, we chose to normalize $P_c(s)$ out by obtaining observed-over-expected values first, and then re-introducing a particular scaling $P_c(s)$. It was deliberately chosen to be very shallow (-0.25 scaling) to increase contrast. We note that observed-over-expected average TADs shown in Extended Data Fig. 3 were not affected by this normalization.

Loop positions were taken from ref. 8, for CH12-LX or GM12878 datasets. For all loops, we summed up a 16×16 pixel region surrounding the loop (loop position was at pixel 9). 16 was chosen over an odd number of bins because it is divisible by 2 and 4, making it possible to coarse grain the map by a factor of 2 or 4 for low-coverage cells.

To control for the effects of coverage, when averaging loops in the Hi-C data, we simultaneously averaged coverage of the 16 bins of the y axis, and 16 bins of the x axis. Specifically, when adding a [x:+16, y:y+16] sub-square of the Hi-C map to the 16x16 aggregate map, we also added coverage from bins [x:x+16] to the average x-axis coverage, and coverage from bins [y:y+16] to the average y-axis coverage. After averaging, we divided the average 16x16 Hi-C map by an outer product of the two average coverages. This normalizes for the potential effect of loop anchors having different coverage in the Hi-C map.

To create the expected distribution for each loop (to control for the decay of contact probability), loop positions were offset by a uniformly chosen distance between 100 kb and 1,100 kb. To maximize statistics, each loop was offset independently 5 times, and the same averaging was performed. The observed average loop was then divided by the average expected value (offset loop).

To quantify loop intensity as a single number, we took an average Hi-C counts at the loop base, averaging over 70×70kb square centered at the bin containing the loop base. We then divided it by the same average, but in the offset loops (see previous paragraph).

To quantify $P_c(s)$ steepness, we calculated the ratio of the average value of $P_c(s)$ in the 11 - 20 kb region, and the average value in the (45 Mb - rest of the chromosome) region. Natural log of this ratio was used as a measure of scaling steepness.

Contact cluster identification algorithm

To identify contact clusters (CCs), we used a segmentation algorithm very similar to ref. 46, which divides the genome into domains in such a way as to maximize a global domain scoring function. The global optimization is carried out on a chromosome by chromosome basis. The scoring function we chose was based on network modularity[47], which is a metric widely used to detect communities in networks. The modularity score for a domain spanning genomic bins a to b inclusively is given by $S(a,b)=\sum_{i=a}^{b}\sum_{j=i}^{b}(A_{i,j}-\gamma N_{i,j})$, where A is the contact matrix and N is the corresponding matrix of a penalizing background model (a uniform matrix in our case). By restricting the solution space to contiguous segmentations, both calculating domain scores and finding the highest scoring segmentation can be reduced to $O(n^2)$ dynamic programming algorithms. Optimal segmentation, in particular, becomes the well-known max-sum algorithm on a weighted directed acyclic graph[48]. The resolution parameter γ controls the strength of the penalty and therefore the characteristic size of the domains identified.

The implementation of these and related algorithms is provided in the *lavaburst* package (https://github.com/nezar-compbio/lavaburst). We swept through different values of gamma, 6-to-24 * (length of chromosome in bins) / 3,250 using a background model of flat contact probability. Forty values uniformly spaced from 6 to 24 were chosen; then the average CC length was calculated; taking into account only CCs more than 200 kb. Out of these 40 values, we chose the one for which the average CC length was closest to 500 kb. These TADs were used for further analyses.

Contact cluster/TAD validation

To validate the contact-cluster identification protocol, we take two orthogonal approaches:

1. We compared TAD boundaries identified using our method to those identified using published algorithm from ref. 20. When boundaries were called at 20kb resolution, 77% of boundaries called using lavaburst were within 80kb of boundaries called using (Crane et al., 2015) algorithm with default parameters (32% expected if boundaries were randomized by offsetting them by 1 Mb).
2. We show that CC calling is robust to downsampling (Extended Data Fig. 4c). We downsample single cell maps for the top 5 oocytes by 50% of the reads, and we show that we can still identify 65-70% of domain boundaries allowing for an error margin of 80kb between boundaries. The fraction of CCs overlapping are ~1.5 to 2-fold enriched when comparing two independent down samplings of the same single cell than when comparing the fraction of CC overlap with the next top cell (~35-40%). We also

compare the fraction of overlap with a randomized control (~25%), which is actually somewhat close to that observed overlap between the top cell and the next best cell, indicating that single-cell contact clusters are reproducibly called, and that CCs between single cells are quite variable.

Polymer simulations

Polymer simulations were performed using *openmm-polymer* package (see for example ref. 49 or 11) that is based on OpenMM molecular dynamics engine[50,51]. 128,000 monomer long fractal globules were simulated as follows. We started with a 128,000 monomer self-avoiding walk, generated by pivot algorithm using hard spheres of radius 0.5 connected by hard bonds of length 1 with 30*128000 attempted pivot moves. Molecular dynamics (*openmm-polymer*) was then used to perform the collapse of a self-avoiding walk. Neighbouring monomers were held together by harmonic bonds with energy U=50*kT*(x-1)^2. Collapse was simulated as a constant force of 1kT/monomer pointing to the center of mass of the polymer. The force was acting on any particles that were no more than 1 monomer inside the confining sphere (energy of the collapsing potential at the confining sphere was 1kT). Confining sphere was chosen to have volume 256,000, which roughly corresponds to density of 0.5. Variable Langevin integrator with error tolerance of 0.001 was used in *openm-polymer*. Langevin thermostat was set to the value that did not let kinetic energy rise above 3kT during collapse. After 99 percent of the particles crossed the confining sphere, simulations were performed for 3,000 timesteps, and then stopped.

Fractal globules were then expanded to the given density. Simulations were performed similarly, but with much weaker thermostat (0.0005 in *openmm-polymer*). Spherical confinement was increased to allow larger density-by-volume, and the compact fractal globule was allowed to relax into the volume. As soon as 1% of the particles crossed the 0.9 * r, where r is the radius of the confined volume, simulations were stopped.
Simulations of loop extrusion were performed as described previously[22]. For convenience, size of the large TAD was reduced to 1,100 from 1,200, so that TAD sizes (300 + 600 + 1,100) sum to 2,000. A chain of 128,000 monomers was used, corresponding to 77 Mb of DNA, as in ref. 11. We used lifetime=300, since it better reproduces properties of *in vivo* Hi-C maps[22]. We performed two steps of loop extrusion per 10000-timestep block of molecular dynamics, 4 times more MD/LEF steps than in ref. 22, to allow for better mixing of different TADs at low densities. A non-expanded fractal globule was taken as a starting configuration. Simulations were performed in periodic boundary conditions, just as in ref. 22. Loop extrusion was simulated for 2,000 blocks, and done in 40 replicates.

Simulations initialized with the mitotic chromosome were initialized with best-fitting loops-on-the-scaffold model[11], loop size 120 kb. For each run, a random run out of 5 runs made for ref. 11 was selected, and a random conformation was selected from the second half of this run.

Our model is non-equilibrium by design (loop extruding factors constantly pump energy into the system); however, the system should likely eventually achieve a state of dynamic equilibrium. It would mean that if we waited long enough to achieve dynamic equilibrium, then any observables, such as $P_c(s)$, would no longer change with time and would not depend on the initial conformation. We started our model with two different conformations, fractal globule and mitotic chromosome model, and obtained different $P_c(s)$ curves, which indicates that our model did not reach the dynamic equilibrium, and it takes longer for a system of this size to achieve it.

To make sure our simulation time is long enough , we mapped the timescale of our simulations to real time. In our simulations, 2,000 blocks would correspond to each loop extruding factor exchanging on average 13.3 times (lifetime=300, and two LEF steps per block, yield one exchange per 150 blocks). Assuming cohesins are indeed LEFs[22], the residence time of cohesin on chromatin was estimated to be on the order of 25 minutes in rat kidney cells in interphase (G1)[52]; a recent preprint also reported residence time of 22 in mouse ES cells[53]. With that residence time, 13.3 exchanges would correspond to the time scale of around 5-6 hours, slightly less than the age of zygotes in our analysis (roughly 7-10 hours since fertilization). To ensure that 7-10 hours do not lead to a strong shift in $P_c(s)$, we performed two simulation runs of the paternal and two of the maternal zygote model for 3 times the original simulation time: 6,000 blocks, or 15-18 hours of real time. Resulting Pc(s) curves were mostly unchanged, and had the same shape at 15 hours (6,000 blocks) as at 2.5 hours (1000 blocks). This comparison implies that, at the time scale of around a day, chromatin in nuclei does not achieve equilibration, and can indeed retain the memory of the previous state. We note that this analysis should be taken as a rough estimate, since it assumes cohesin turnaround time in zygotes is the same as in somatic cells. Equilibration time may also be strongly affected by chromatin properties such as Topo II activity, anchoring to the lamina or nuclear bodies, and other known and unknown unknowns.

We also note that since oocytes have loops and TADs, they could also undergo the process of loop extrusion. However, oocytes have spent much more time than nuclei since the last cell division: in our experimental setup 2-3 months (unlike sperm, that is generated throughout the lifetime, oocytes are arrested in prophase I before birth and only a subset matures and divides upon ovulation every estrus cycle throughout the lifetime of the mouse). If loop extrusion continues for the entire life of the oocyte, this time may be sufficient for the loop extrusion to achieve dynamic equilibrium. Dynamic equilibrium would be consistent with the $s^{-1.5}$ scaling (random walk) of contact probability at the scale larger than the scale set by loop extrusion, i.e. >1Mb. However, simulations at a timescale of months are currently inaccessible with the molecular dynamics engine used here, as it would take many months of wall-clock time to simulate one trajectory.

Code availability

Polymer simulation code, including the loop extrusion code, is available in the "examples" directory of the *openmm-polymer* library https://bitbucket.org/mirnylab/openmm-polymer. Data processing code will be released as an example for the *hiclib* package https://bitbucket.org/mirnylab/hiclib. Modularity score domain segmentation algorithm is available as a part of *lavaburst* package https://github.com/nezar-compbio/lavaburst.

Data availability

All sequencing data in support of the findings of this study have been deposited in the Gene Expression Omnibus (GEO) under accession number GSE80006. Source data for figures (Fig. 1b, d, e, Fig. 2c-e, Fig. 3b-g, Fig. 4c-f, Extended Data: Fig. 1, Fig. 2, Fig. 3a, c, Fig. 4a-d, Fig. 7b, Fig. 8a-c, Fig. 10a-c) are provided with the paper.

Supplementary References

31. Kumar, G., Garnova, E., Reagin, M. & Vidali, A. Improved multiple displacement amplification with ϕ29 DNA polymerase for genotyping of single human cells. *BioTechniques* **44,** 879–890 (2008).
32. Eskeland, R. *et al.* Ring1B Compacts Chromatin Structure and Represses Gene Expression Independent of Histone Ubiquitination. *Mol. Cell* **38,** 452–464 (2010).
33. Probst, A. V., Santos, F., Reik, W., Almouzni, G. & Dean, W. Structural differences in centromeric heterochromatin are spatially reconciled on fertilisation in the mouse zygote. *Chromosoma* **116,** 403–415 (2007).
34. Imakaev, M. *et al.* Iterative correction of Hi-C data reveals hallmarks of chromosome organization. *Nat. Methods* **9,** 999–1003 (2012).
35. Kent, W. J. *et al.* The human genome browser at UCSC. *Genome Res.* **12,** 996–1006 (2002).
36. Dixon, J. R. *et al.* Chromatin architecture reorganization during stem cell differentiation. *Nature* **518,** 331–336 (2015).
37. Zuin, J. *et al.* Cohesin and CTCF differentially affect chromatin architecture and gene expression in human cells. *Proc. Natl. Acad. Sci.* **111,** 996–1001 (2013).
38. Jin, F. *et al.* A high-resolution map of the three-dimensional chromatin interactome in human cells. *Nature* **503,** 290–294 (2013).
39. 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).
40. Sofueva, S. *et al.* Cohesin-mediated interactions organize chromosomal domain architecture. *EMBO J.* **32,** 3119–3129 (2013).
41. Zhang, Y. *et al.* Spatial Organization of the Mouse Genome and Its Role in Recurrent Chromosomal Translocations. *Cell* **148,** 908–921 (2012).
42. 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).
43. 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).
44. 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).
45. Chandra, T. *et al.* Global Reorganization of the Nuclear Landscape in Senescent Cells. *Cell Rep.* **10,** 471–483 (2015).

46. Filippova, D., Patro, R., Duggal, G. & Kingsford, C. in *Algorithms in Bioinformatics* **8126,** 300–312 (Springer, 2013).
47. Newman, M. E. J. & Girvan, M. Finding and evaluating community structure in networks. *Phys. Rev. E* **69,** 026113 (2004).
48. Bishop, C. Pattern Recognition and Machine Learning. *Springer* (2006).
49. Le, T. B. K., Imakaev, M. V., Mirny, L. A. & Laub, M. T. High-Resolution Mapping of the Spatial Organization of a Bacterial Chromosome. *Science* **342,** 731–734 (2013).
50. Eastman, P. *et al.* OpenMM 4: A Reusable, Extensible, Hardware Independent Library for High Performance Molecular Simulation. *J. Chem. Theory Comput.* **9,** 461–469 (2013).
51. Eastman, P. *et al.* OpenMM 7: Rapid Development of High Performance Algorithms for Molecular Dynamics. Preprint at http://biorxiv.org/content/early/2016/12/06/091801 (2016). doi:10.1101/091801
52. Gerlich, D., Koch, B., Dupeux, F., Peters, J.-M. & Ellenberg, J. Live-cell imaging reveals a stable cohesin-chromatin interaction after but not before DNA replication. *Curr. Biol. CB* **16,** 1571–1578 (2006).
53. Hansen, A. S., Pustova, I., Cattoglio, C., Tjian, R. & Darzacq, X. CTCF and Cohesin Regulate Chromatin Loop Stability with Distinct Dynamics. Preprint at http://biorxiv.org/content/early/2016/12/13/093476 (2016). doi:10.1101/093476

Supplementary Table 1

Information about datasets generated in this study.

ID	Cis contacts	Total number of contacts	Cis/total	Genome assembly	Cell type
K562 B bulk	1712155	2141382	0.8	hg19	K562
K562 A bulk	1634607	2036405	0.8	hg19	K562
MEL-1	979087	1411002	0.69	mm9	MEL 1% 10 min
MEL-2	1508371	1850370	0.82	mm9	MEL 2% 15 min
1	1654572	1906436	0.87	mm9	NSN
2	1028451	1216953	0.85	mm9	SN-Hoechst
3	1005906	1046441	0.96	mm9	SN
4	953597	1065855	0.89	mm9	NSN
5	946028	1049207	0.9	mm9	NSN
6	917752	1099449	0.83	mm9	SN-Hoechst
7	903545	987696	0.91	mm9	Intermediate
8	877764	971972	0.9	mm9	SN
9	855559	1014984	0.84	mm9	SN-Hoechst
10	792569	887821	0.89	mm9	NSN
11	778100	902876	0.86	mm9	SN-Hoechst
12	769912	821470	0.94	mm9	NSN
13	769548	909180	0.85	mm9	SN
14	735118	778341	0.94	mm9	SN
15	706447	804586	0.88	mm9	NSN
16	691097	776203	0.89	mm9	SN-Hoechst
17	663795	722963	0.92	mm9	NSN
18	655405	725650	0.9	mm9	SN-Hoechst
19	603615	694711	0.87	mm9	SN-Hoechst
20	596623	647722	0.92	mm9	SN
21	594051	690512	0.86	mm9	SN-Hoechst
22	570747	625399	0.91	mm9	SN-Hoechst
23	557556	646846	0.86	mm9	SN-Hoechst
24	524460	621074	0.84	mm9	zygotic-undef (no extr.)
25	520881	563257	0.92	mm9	NSN
26	491389	579188	0.85	mm9	NSN-Hoechst
27	486062	509435	0.95	mm9	SN
28	477672	511384	0.93	mm9	SN-Hoechst
29	469079	508844	0.92	mm9	SN
30	460853	490535	0.94	mm9	NSN

31	452916	492881	0.92	mm9	SN
32	449533	479894	0.94	mm9	NSN
33	443935	489659	0.91	mm9	SN-Hoechst
34	422711	456903	0.93	mm9	SN-Hoechst
35	419636	469021	0.89	mm9	SN
36	400372	447953	0.89	mm9	zygotic-both (no extr.)
37	400110	427149	0.94	mm9	SN
38	399408	459218	0.87	mm9	zygotic-undef (no extr.)
39	391416	428725	0.91	mm9	SN
40	372588	390217	0.95	mm9	NSN
41	367433	391242	0.94	mm9	NSN
42	361786	399426	0.91	mm9	SN-Hoechst
43	358549	392837	0.91	mm9	Intermediate
44	354119	378687	0.94	mm9	NSN
45	347083	416018	0.83	mm9	zygotic-undef (no extr.)
46	345064	363080	0.95	mm9	SN
47	344244	373838	0.92	mm9	NSN
48	343263	360477	0.95	mm9	NSN
49	336908	395663	0.85	mm9	SN-Hoechst
50	334359	391879	0.85	mm9	Intermediate-Hoechst
51	333105	363143	0.92	mm9	SN
52	313038	347617	0.9	mm9	SN-Hoechst
53	312800	343709	0.91	mm9	zygotic-both (no extr.)
54	306169	356255	0.86	hg19	K562
55	304903	330434	0.92	mm9	NSN
56	300355	337020	0.89	mm9	SN-Hoechst
57	296848	320661	0.93	mm9	Intermediate-Hoechst
58	295863	354583	0.83	hg19	K562
59	294495	322679	0.91	mm9	NSN
60	293952	317901	0.92	mm9	zygotic-mat
61	293139	333587	0.88	mm9	SN-Hoechst
62	290996	351271	0.83	mm9	NSN
63	288412	332125	0.87	mm9	zygotic-pat
64	281766	302849	0.93	mm9	zygotic-mat
65	270460	289585	0.93	mm9	NSN
66	268825	299345	0.9	mm9	zygotic-mat
67	263708	302507	0.87	mm9	zygotic-mat
68	261481	277926	0.94	mm9	SN

69	255175	278726	0.92	mm9	zygotic-mat (no extr.)
70	252210	302366	0.83	mm9	zygotic-undef (no extr.)
71	242668	261620	0.93	mm9	zygotic-mat
72	241311	285306	0.85	hg19	K562
73	236970	270266	0.88	mm9	zygotic-mat
74	234198	267166	0.88	mm9	zygotic-pat
75	232580	268704	0.87	mm9	zygotic-pat (no extr.)
76	230829	246753	0.94	mm9	zygotic-mat
77	229153	280728	0.82	mm9	zygotic-pat (no extr.)
78	225399	246815	0.91	mm9	zygotic-mat (no extr.)
79	223809	263643	0.85	hg19	K562
80	223283	256935	0.87	mm9	zygotic-pat (no extr.)
81	214224	230927	0.93	mm9	zygotic-mat
82	209103	238291	0.88	mm9	SN
83	205606	220979	0.93	mm9	SN
84	203853	225864	0.9	mm9	Intermediate-Hoechst
85	199366	225577	0.88	mm9	zygotic-pat
86	197463	209818	0.94	mm9	NSN-Hoechst
87	193017	214367	0.9	mm9	zygotic-pat (no extr.)
88	192458	218039	0.88	mm9	SN-Hoechst
89	188618	205165	0.92	mm9	zygotic-pat
90	187669	207337	0.91	mm9	zygotic-pat (no extr.)
91	186658	222418	0.84	mm9	zygotic-mat
92	181463	190918	0.95	mm9	zygotic-mat
93	181244	217247	0.83	mm9	NSN-Hoechst
94	179259	197391	0.91	mm9	zygotic-mat
95	177505	201220	0.88	mm9	zygotic-mat
96	175104	197126	0.89	mm9	zygotic-pat
97	171311	211220	0.81	mm9	zygotic-pat
98	171288	194490	0.88	mm9	zygotic-pat (no extr.)
99	169583	187830	0.9	mm9	SN
100	168394	185570	0.91	mm9	zygotic-pat
101	161008	176370	0.91	mm9	SN
102	157045	182949	0.86	hg19	K562
103	152970	172050	0.89	mm9	zygotic-undef (no extr.)
104	150923	161318	0.94	mm9	zygotic-mat
105	147475	162889	0.91	mm9	zygotic-pat
106	137709	155382	0.89	mm9	SN-Hoechst

107	132746	157988	0.84	mm9	zygotic-both (no extr.)
108	124809	138511	0.9	mm9	zygotic-pat
109	124723	133595	0.93	mm9	zygotic-mat
110	121183	141322	0.86	hg19	K562
111	119212	144671	0.82	mm9	zygotic-mat
112	114852	127135	0.9	mm9	zygotic-mat
113	109722	120604	0.91	mm9	zygotic-pat
114	107583	118808	0.91	mm9	SN
115	106804	114422	0.93	mm9	zygotic-mat
116	105728	129172	0.82	mm9	zygotic-pat
117	102388	108576	0.94	mm9	zygotic-mat
118	100428	119991	0.84	mm9	zygotic-pat
119	99698	109258	0.91	mm9	zygotic-mat (no extr.)
120	96689	107627	0.9	mm9	Intermediate-Hoechst
121	93872	106859	0.88	mm9	zygotic-pat
122	91077	100297	0.91	mm9	SN
123	89768	106925	0.84	hg19	K562
124	89071	95147	0.94	mm9	SN-Hoechst
125	89036	95598	0.93	mm9	SN
126	88262	96196	0.92	mm9	zygotic-pat
127	84969	97841	0.87	mm9	zygotic-pat
128	79256	84403	0.94	mm9	zygotic-mat
129	79020	82495	0.96	mm9	zygotic-mat
130	75229	84338	0.89	mm9	SN
131	74492	84208	0.88	mm9	NSN-Hoechst
132	72793	79219	0.92	mm9	SN
133	71669	79753	0.9	mm9	zygotic-pat (no extr.)
134	68999	75900	0.91	mm9	NSN
135	68169	76682	0.89	mm9	SN-Hoechst
136	66023	75417	0.88	mm9	zygotic-mat (no extr.)
137	64535	76052	0.85	mm9	zygotic-pat
138	63151	70973	0.89	mm9	NSN-Hoechst
139	62521	68294	0.92	mm9	zygotic-pat
140	61430	71542	0.86	mm9	NSN-Hoechst
141	61376	66783	0.92	hg19	K562
142	58458	61834	0.95	mm9	zygotic-mat
143	58122	61824	0.94	mm9	SN-Hoechst
144	57930	61915	0.94	mm9	SN

145	56497	60703	0.93	mm9	NSN-Hoechst
146	55995	65564	0.85	mm9	NSN-Hoechst
147	54881	59651	0.92	mm9	zygotic-pat
148	53442	58859	0.91	mm9	zygotic-mat
149	50679	57053	0.89	mm9	zygotic-both (no extr.)
150	50148	55926	0.9	mm9	SN
151	49920	52448	0.95	mm9	zygotic-mat
152	47267	51180	0.92	mm9	NSN
153	47260	51483	0.92	mm9	NSN
154	43662	47079	0.93	mm9	zygotic-pat
155	43290	50762	0.85	mm9	NSN
156	42929	49745	0.86	hg19	K562
157	41626	46438	0.9	mm9	SN-Hoechst
158	39929	44367	0.9	mm9	zygotic-pat
159	38496	44683	0.86	mm9	NSN
160	38311	42595	0.9	mm9	NSN-Hoechst
161	36982	41711	0.89	mm9	SN
162	35545	40290	0.88	mm9	SN-Hoechst
163	34152	39314	0.87	mm9	SN
164	31490	35685	0.88	mm9	SN-Hoechst
165	30942	35149	0.88	mm9	SN
166	30656	35826	0.86	mm9	SN-Hoechst
167	29690	32022	0.93	mm9	SN
168	29407	33863	0.87	mm9	NSN
169	28939	30939	0.94	mm9	zygotic-mat
170	28306	30283	0.93	mm9	SN
171	24789	28057	0.88	mm9	NSN-Hoechst
172	24496	25867	0.95	mm9	SN
173	22933	26596	0.86	mm9	SN
174	22203	25263	0.88	mm9	SN
175	22059	24629	0.9	mm9	SN
176	21920	25125	0.87	mm9	NSN-Hoechst
177	21668	23350	0.93	mm9	SN
178	20481	21934	0.93	mm9	zygotic-undef (no extr.)
179	20076	22891	0.88	mm9	NSN-Hoechst
180	19150	21304	0.9	mm9	NSN
181	19049	21161	0.9	mm9	SN
182	18260	20813	0.88	mm9	NSN

183	18193	21820	0.83	mm9	zygotic-undef (no extr.)
184	17556	20480	0.86	mm9	zygotic-pat (no extr.)
185	17505	20913	0.84	mm9	SN
186	16864	19048	0.89	mm9	NSN-Hoechst
187	16741	18808	0.89	mm9	zygotic-mat
188	16517	18621	0.89	mm9	NSN-Hoechst
189	16048	18039	0.89	mm9	NSN
190	15135	17565	0.86	mm9	zygotic-mat
191	14671	16145	0.91	mm9	SN
192	13584	15794	0.86	mm9	SN-Hoechst
193	13109	14684	0.89	mm9	NSN-Hoechst
194	12855	14544	0.88	mm9	SN
195	12687	14722	0.86	mm9	zygotic-pat
196	12498	14035	0.89	mm9	zygotic-undef (no extr.)
197	11159	12590	0.89	mm9	SN-Hoechst
198	11056	11950	0.93	mm9	zygotic-mat
199	10460	12975	0.81	mm9	zygotic-pat (no extr.)
200	10360	11976	0.87	mm9	zygotic-undef (no extr.)
201	10026	11593	0.86	mm9	SN-Hoechst
202	8959	9987	0.9	mm9	zygotic-undef (no extr.)
203	7749	9570	0.81	mm9	zygotic-pat
204	6875	7883	0.87	mm9	zygotic-pat (no extr.)
205	6454	9172	0.7	mm9	zygotic-mat
206	6217	7621	0.82	mm9	zygotic-pat
207	5879	7769	0.76	mm9	zygotic-mat (no extr.)
208	5708	6953	0.82	mm9	zygotic-undef (no extr.)
209	4815	6139	0.78	mm9	zygotic-pat
210	4727	5928	0.8	mm9	zygotic-mat
211	4682	5241	0.89	mm9	zygotic-undef (no extr.)
212	4191	5479	0.76	mm9	zygotic-pat (no extr.)
213	4152	4717	0.88	mm9	zygotic-mat
214	4010	5377	0.75	mm9	zygotic-mat (no extr.)
215	3407	5616	0.61	hg19	K562
216	2922	4112	0.71	mm9	zygotic-mat
217	1878	3269	0.57	hg19	K562
218	1372	1508	0.91	mm9	NSN
219	1187	1428	0.83	mm9	SN
220	966	2009	0.48	hg19	K562

221	802	1890	0.42	hg19	K562
222	716	1790	0.4	hg19	K562
223	579	1430	0.4	hg19	K562
224	510	1284	0.4	hg19	K562
225	495	1560	0.32	hg19	K562
226	354	1258	0.28	hg19	K562
227	338	1467	0.23	hg19	K562
228	310	1386	0.22	hg19	K562
229	257	1155	0.22	hg19	K562
230	249	1403	0.18	hg19	K562
231	230	1297	0.18	hg19	K562
232	219	1017	0.22	hg19	K562
233	217	1154	0.19	hg19	K562
234	203	1171	0.17	hg19	K562
235	203	1130	0.18	hg19	K562
236	200	1181	0.17	hg19	K562
237	188	1302	0.14	hg19	K562
238	169	1029	0.16	hg19	K562
239	169	1352	0.13	hg19	K562
240	135	1023	0.13	hg19	K562
241	90	1057	0.09	hg19	K562
242	83	1015	0.08	hg19	K562