nature genetics

Nuclear organization of active and inactive chromatin domains uncovered by chromosome conformation capture–on-chip (4C)

Marieke Simonis[1], Petra Klous[1], Erik Splinter[1], Yuri Moshkin[2], Rob Willemsen[3], Elzo de Wit[4], Bas van Steensel[4] & Wouter de Laat[1]

The spatial organization of DNA in the cell nucleus is an emerging key contributor to genomic function[1–12]. We developed 4C technology (chromosome conformation capture (3C)-on-chip), which allows for an unbiased genome-wide search for DNA loci that contact a given locus in the nuclear space. We demonstrate here that active and inactive genes are engaged in many long-range intrachromosomal interactions and can also form interchromosomal contacts. The active β-globin locus in fetal liver preferentially contacts transcribed, but not necessarily tissue-specific, loci elsewhere on chromosome 7, whereas the inactive locus in fetal brain contacts different transcriptionally silent loci. A housekeeping gene in a gene-dense region on chromosome 8 forms long-range contacts predominantly with other active gene clusters, both in *cis* and in *trans*, and many of these intra- and interchromosomal interactions are conserved between the tissues analyzed. Our data demonstrate that chromosomes fold into areas of active chromatin and areas of inactive chromatin and establish 4C technology as a powerful tool to study nuclear architecture.

Our understanding of genomic organization in the nuclear space is based mostly on microscopy studies that often use FISH to visualize selected parts of the genome. However, FISH, no matter how revealing, can analyze only a limited number of DNA loci simultaneously and therefore produces largely anecdotal observations. In order to get a rigorous picture of nuclear architecture, there is a need for high-throughput technology that can systematically screen the whole genome in an unbiased manner for DNA loci that contact each other in the nuclear space. To this end we have developed 4C technology, which combines chromosome conformation capture (3C) technology[13] with dedicated microarrays.

An outline of the 4C procedure is given in **Figure 1a,** and it is explained in detail in the Methods section. In short, 4C involves PCR amplification of DNA fragments cross-linked and ligated to a DNA restriction fragment of choice (here, *Hin*dIII fragments). Typically, this yields a pattern of PCR fragments specific for a given tissue and highly reproducible between independent PCR reactions (**Fig. 1b**). The amplified material, representing the fragment's genomic environment, is labeled and hybridized to a tailored microarray that contains probes each located <100 bp from a different *Hin*dIII restriction end in the genome (**Fig. 1c**). The array used for this study (from Nimblegen Systems) covered seven complete mouse chromosomes.

We applied 4C technology to characterize the genomic environment of the active and inactive mouse β-globin locus, located in a large olfactory receptor gene cluster on chromosome 7. We focused our analysis on a restriction fragment containing hypersensitive site 2 (HS2) of the β-globin locus control region (LCR). Both in embryonic day (E)14.5 liver, where the β-globin genes are highly transcribed, and in E14.5 brain, where the locus is inactive, we found that the great majority of interactions were with sequences on chromosome 7, and we detected very few LCR interactions with six unrelated chromosomes (8, 10, 11, 12, 14 and 15; **Fig. 2a**). We found the strongest signals on chromosome 7 within a 5- to 10-Mb region centered around the chromosomal position of β-globin, in agreement with the idea that interaction frequencies are inversely proportional to the distance (in bp) between physically linked DNA sequences[13]. It was not possible to interpret the interactions in this region quantitatively because corresponding probes on the array were saturated (see Methods).

Both in fetal liver and in brain, we identified clusters of 20–50 positive signals juxtaposed on chromosome 7, often at chromosomal locations tens of Mb away from β-globin (**Fig. 2b,c**). To determine the statistical significance of these clusters, we ordered data of individual experiments on chromosomal maps and analyzed them using a running mean algorithm with a window size of approximately 60 kb. We then used the running mean distribution of randomly shuffled data to set a threshold value, allowing a false discovery rate of 5%. This analysis identified 66 clusters in fetal liver and 45 in brain that were reproducibly found in duplicate experiments (**Fig. 2d–f**). Indeed, high-resolution FISH confirmed that such clusters truly represent loci that interact frequently (see below). Thus, 4C technology

[1]Department of Cell Biology and Genetics, [2]Department of Biochemistry, [3]Department of Clinical Genetics, Erasmus Medical Centre, PO Box 2040, 3000 CA Rotterdam, The Netherlands. [4]Division of Molecular Biology, Netherlands Cancer Institute, Plesmanlaan 121, 1066 CX Amsterdam, The Netherlands. Correspondence should be addressed to W.d.L. (w.delaat@erasmusmc.nl).

Received 25 May; accepted 6 September; published online 8 October 2006; doi:10.1038/ng1896

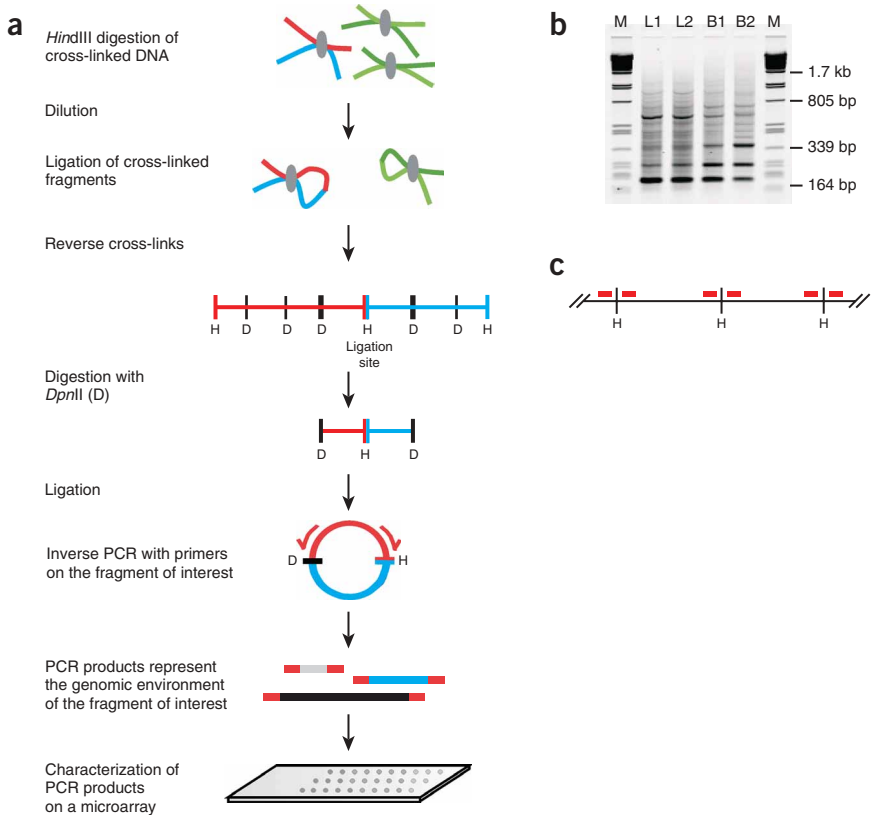

Figure 1 4C technology. (**a**) Outline of 4C procedure. Briefly, 3C analysis is performed as usual, but the PCR step is omitted. The 3C template contains bait (for example, a restriction fragment encompassing a gene) ligated to many different fragments (representing this gene's genomic environment). The ligated fragments are cleaved by a frequently cutting secondary restriction enzyme and are subsequently religated to form small DNA circles that are amplified by inverse PCR (30 cycles) using bait-specific primers facing outward. (**b**) PCR results separated by gel electrophoresis from two independent fetal liver (L1, L2) and brain (B1, B2) samples. (**c**) Schematic representation of the location of the microarray probes. Probes were designed within 100 bp of *Hin*dIII sites. Thus, each probe analyzes one possible ligation partner.

identifies long-range interacting loci by the detection of independent ligation events with multiple restriction fragments clustered at a chromosomal position. A completely independent series of 4C experiments that focused on a fragment 50 kb downstream containing the β-globin–like gene *Hbb-b1* gave almost identical results (**Supplementary Fig. 1** online).

A comparison between the two tissues showed that the actively transcribed β-globin locus in fetal liver interacts with a completely different set of loci on chromosome 7 from its transcriptionally silent counterpart in brain ($\tau = -0.03$; Spearman's rank correlation). This excluded that results were influenced by the sequence composition of the probes. In fetal liver, the interacting DNA segments were located within a 70-Mb region centered around the β-globin locus, with the majority (40/66) located toward the telomere of the chromosome. In fetal brain, we found interacting loci at similar or even greater distances from β-globin than in fetal liver, with the great majority of interactions (43/45) located toward the centromere of chromosome 7 (**Fig. 2f**).

Although the average size of interacting areas in fetal liver and brain was comparable (183 kb and 159 kb, respectively), we observed marked differences in their gene content and activity, the latter being determined by Affymetrix expression array analysis. In fetal liver, 80% of the β-globin interacting loci contained one or more actively transcribed genes, whereas in fetal brain, the great majority (87%) did not show any detectable gene activity (**Fig. 3**). Thus, the β-globin locus contacts different types of genomic regions in the two tissues (**Supplementary Table 1** online). Notably, 4C technology identified *Uros*, *Eraf* and *Kcnq1* (all ~30 Mb away from β-globin) as genes interacting with the active β-globin locus in fetal liver, in agreement with previous observations made by FISH[8] (**Supplementary Fig. 2** online). Notably, in brain, we observed contacts with two other olfactory receptor gene clusters present on chromosome 7 that were located at each side of, and 17 and 37 Mb away from, β-globin.

Uros, *Eraf* and the genes encoding β-globin are all erythroid-specific genes that may be regulated by the same set of transcription factors, and it is an attractive idea that these factors coordinate the expression of their target genes in the nuclear space. We compared Affymetrix expression array data from E14.5 liver with that of brain to identify genes expressed preferentially (that is, showing more than fivefold greater expression) in fetal liver. Of the 560 active genes on chromosome 7, 15% were 'fetal liver–specific'; this was true for 13% of the 156 active genes within the colocalizing areas. More notably, 49 out of 66 (74%) interacting regions did not contain a 'fetal liver–specific' gene. Thus, we find no evidence for the intrachromosomal clustering of tissue-specific genes.

As the β-globin genes are transcribed at exceptionally high rates, we subsequently asked whether the locus preferentially interacts with other regions of high transcriptional activity. Using Affymetrix counts as a measure for gene activity, we performed a running sum algorithm to measure overall transcriptional activity within 200-kb regions around actively transcribed genes. This analysis showed that transcriptional activity around interacting genes was not higher than around noninteracting active genes on chromosome 7 ($P = 0.9867$; Wilcoxon rank sum).

We next investigated whether a gene that is expressed similarly in both tissues also switches its genomic environment. *Rad23a* is a ubiquitously expressed DNA repair gene that resides in a gene-dense cluster of predominantly housekeeping genes on chromosome 8. Both in E14.5 liver and in brain, this gene and many of its direct neighbors are active. We performed 4C analysis and identified many long-range interactions with loci up to 70 Mb away from *Rad23a*. Notably, interactions with *Rad23a* were highly correlated between fetal liver and brain ($\tau = 0.73$; Spearman's rank correlation) (**Fig. 4a**), providing evidence for a general chromosomal folding pattern that is conserved between different cell types. Again, a shared hallmark of the interacting loci was that they contained actively transcribed genes. In both tissues, roughly 70% contained at least one active gene (**Fig. 4b**,**c**). Regions around interacting genes showed statistically significant higher levels of gene activity than for active genes elsewhere on the chromosome, as determined by a running sum algorithm ($P < 0.001$ for both tissues).

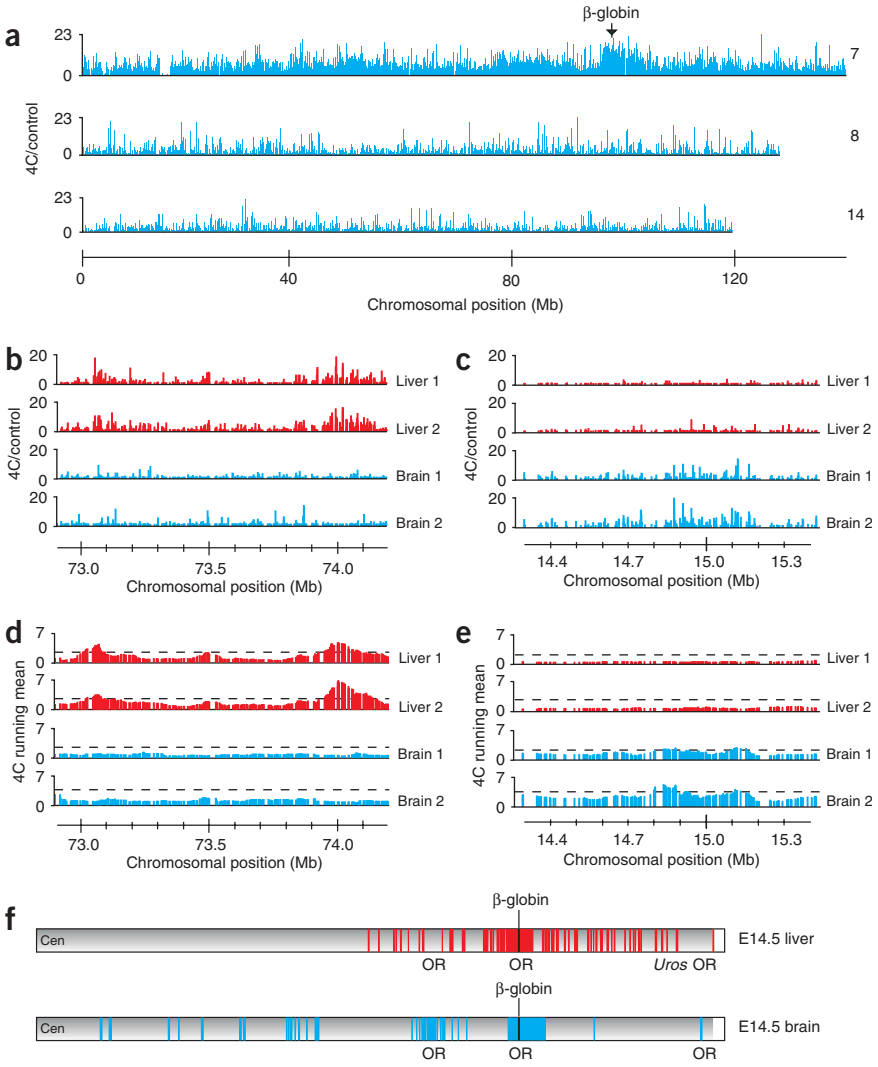

Figure 2 Long-range interactions with β-globin, as shown by 4C technology. (**a**) Unprocessed ratios of 4C-to-control hybridization signals, showing interactions of β-globin HS2 with chromosome 7 and two unrelated chromosomes (8 and 14). (**b**,**c**) Unprocessed data for two independent fetal liver and fetal brain samples plotted along two different 1- to 2-Mb regions on chromosome 7. Highly reproducible clusters of interactions are observed either in the two fetal liver samples (**b**) or the two brain samples (**c**). (**d**,**e**) Running mean data for the same regions. False discovery rate was set at 5% (dashed line). (**f**) Schematic representation of regions of interaction with active (fetal liver, top) and inactive (fetal brain, bottom) β-globin on chromosome 7. Note that interacting regions are, on average, 150–200 kb and are not drawn to scale. Chromosomal positions were based on National Center for Biotechnology (NCBI) build m34.

Thus, the *Rad23a* gene, which is located in a gene-rich region, preferentially interacts over a distance with other chromosomal regions of increased transcriptional activity.

To validate the results obtained by 4C technology, we performed cryo-FISH experiments. Cryo-FISH is a recently developed microscopy technique that has an advantage over current three-dimensional FISH in that it better preserves the nuclear ultrastructure while offering improved resolution in the *z* axis by the preparation of ultrathin cryosections[10]. Notably, 4C technology measures interaction frequencies rather than (average) distances between loci. Therefore, we verified 4C data by measuring how frequently β-globin or *Rad23a* alleles (always $n > 250$) colocalized with selected chromosomal regions in 200-nm ultrathin sections prepared from E14.5 liver and brain. We tested more than 15 loci and found that colocalization frequencies for loci positively identified by 4C technology were all significantly higher than frequencies measured for background loci ($P < 0.05$; *G* test) (**Supplementary Table 2** online). For example, distant regions that we found to interact with β-globin by 4C technology colocalized more frequently than intervening areas not detected by 4C (7.4% and 9.7% versus 3.6% and 3.5%, respectively). Also, the two distant olfactory receptor gene clusters found (by 4C) to interact with β-globin in fetal brain but not liver scored colocalization frequencies of 12.9% and 7% in brain versus 3.6% and 1.9% in liver sections (**Fig. 5**). We concluded that 4C technology faithfully identified interacting DNA loci. Next, we used cryo-FISH to demonstrate that loci identified to interact with β-globin also frequently contacted each other. This was true for two active regions separated over a large chromosomal distance in fetal liver (**Supplementary Fig. 3** online) as well as for two inactive olfactory receptor gene clusters far apart on the chromosome in brain (**Fig. 5**). Notably, we also found frequent contacts between these two distant olfactory receptor gene clusters in fetal liver, where they did not interact with the olfactory receptor gene cluster that contained the actively transcribed β-globin locus. These data provided further evidence for spatial interactions between distant olfactory receptor gene clusters[14].

FISH analysis showed that the gene-dense chromosomal region containing *Rad23a* resides mostly at the edge of (82%) or outside (14%) the territory of chromosome 8 (D. Noordermeer, M.R. Branco, A. Pombo and W.d.L., unpublished data) and we considered the possibility that *Rad23a* also interacted with regions on other chromosomes. Six unrelated chromosomes (7, 10, 11, 12, 14 and 15) were represented on our microarrays. Typically, these chromosomes showed very low 4C signals, with a few strong signals that often appeared isolated on the linear DNA template. Indeed, when we analyzed each chromosome separately by applying running mean algorithms, we identified mostly regions that contained one such very strong hybridization signal. When tested by cryo-FISH, these regions scored negative for interaction (**Supplementary Table 3** online). To better identify clusters of possibly weaker, but positive, signals on the unrelated chromosomes, we applied a running median algorithm that ignores the isolated strong signals and scores only regions containing multiple positive signals. In fetal liver and brain, 24 and 44 of these interchromosomal regions were reproducibly identified in duplicate experiments (false discovery rate of 0%). We tested two of these regions by cryo-FISH, and both showed significant colocalization with *Rad23a* ($P < 0.05$; **Supplementary Table 3**).

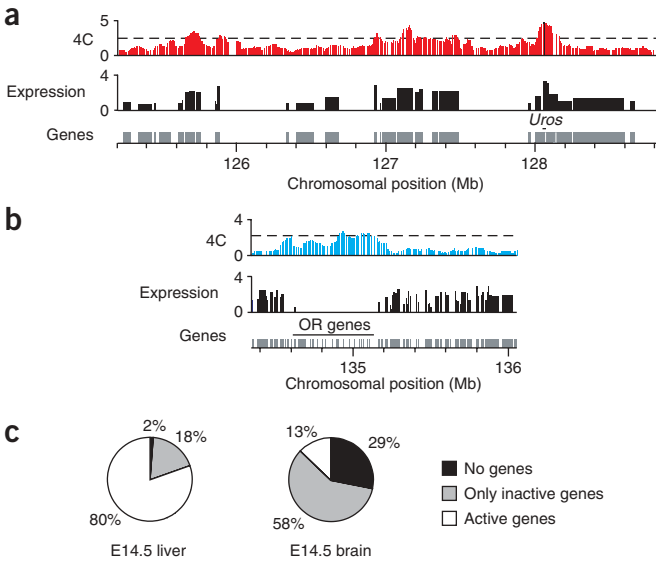

Figure 3 Active and inactive β-globin interact with active and inactive chromosomal regions, respectively. (**a**) Comparison between β-globin long-range interactions in fetal liver (4C running mean, top), microarray expression analysis in fetal liver (log scale, middle) and the location of genes (bottom) plotted along a 4-Mb region that contains the gene *Uros* (~30 Mb away from β-globin), showing that active β-globin preferentially interacts with other actively transcribed genes. (**b**) The same comparison in fetal brain around an olfactory receptor (OR) gene cluster located ~38 Mb away from β-globin, showing that inactive β-globin preferentially interacts with inactive regions. Chromosomal positions were based on NCBI build m34. (**c**) Characterization of regions interacting with β-globin in fetal liver and brain in terms of gene content and activity.

We identified *trans*-interacting regions on all six chromosomes analyzed by 4C (**Supplementary Table 4** online). A comparison between fetal liver and brain showed that the *trans*-interacting regions were often located at the same parts of the chromosomes, suggesting that the interchromosomal environment of the gene-dense region containing *Rad23a* is conserved to some extent between the tissues (**Fig. 6a**). Similar to what we had observed for the intrachromosomal contacts, interchromosomal interactions with *Rad23a* seemed biased toward regions with increased gene density (**Fig. 6b–d** and **Supplementary Fig. 4** online). These data suggested that active gene-dense regions preferentially contact each other in *cis* and in *trans*.

The β-globin locus has been shown to preferentially locate inside its own chromosome territory[15]. In agreement with this, we found only four interacting regions in *trans* in each tissue (data not shown). In fetal liver, all four regions contained actively transcribed genes, whereas in fetal brain, where the locus is inactive, three out of the four regions did not contain an active gene. Cryo-FISH on two loci in fetal liver, however, did not show colocalization frequencies significantly above background (data not shown). Thus, either the method detects some false positive inter-chromosomal interactions or is so sensitive that it identifies interactions too infrequent to be discerned from background under the microscope.

Our observations demonstrate that not only active, but also inactive, genomic regions can transiently interact over large distances with many loci in the nuclear space. The data strongly suggest that each DNA segment has its own preferred set of interactions. This implies that it is impossible to predict the long-range interaction partners of a given DNA locus without knowing the characteristics of its neighboring segments and, by extrapolation, the whole chromosome. The fact that a housekeeping gene shows interchromosomal contacts provides further evidence for extensive chromosome intermingling[10] and shows that individual loci have preferred neighbors in the nuclear space. We propose that the extensive network of long-range interactions that we have identified both between inactive and between active genomic loci reflects cell-to-cell differences in chromosome conformations in combination with dynamic movements during interphase.

METHODS

4C technology. The initial steps of the 3C technology procedure were performed as described previously[16], yielding ligation products between *Hin*dIII fragments. This *Hin*dIII-ligated 3C template (~50 μg) was digested

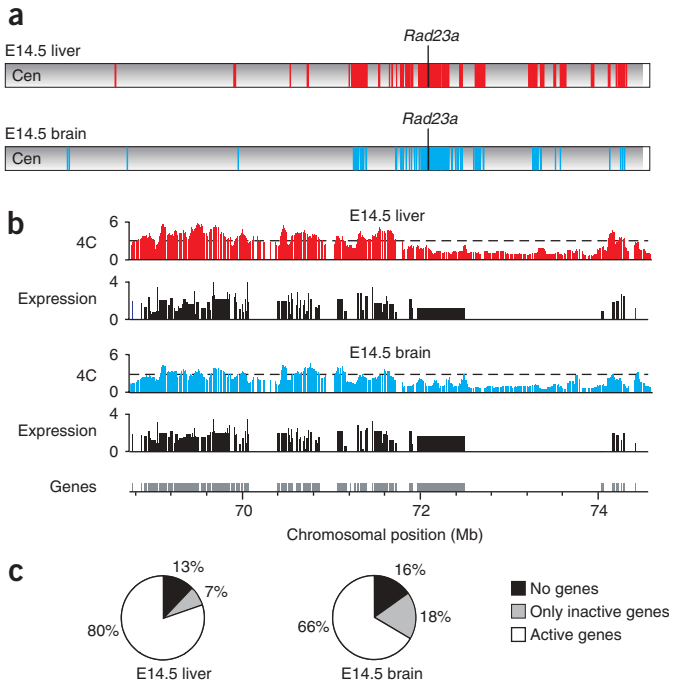

Figure 4 Ubiquitously expressed *Rad23a* interacts with very similar active regions in fetal liver and brain. (**a**) Schematic representation of regions on chromosome 8 interacting with active *Rad23a* in fetal liver and brain. Note that interacting regions are on average 150–200 kb and are not drawn to scale. (**b**) Comparison between *Rad23a* long-range interactions (4C running mean) and microarray expression analysis (log scale) in fetal liver and fetal brain. Location of genes is plotted (bottom) along a 3 Mb region of chromosome 8. Chromosomal positions were based on NCBI build m34. (**c**) Characterization of regions interacting with *Rad23a* in fetal liver and brain in terms of gene content and activity.

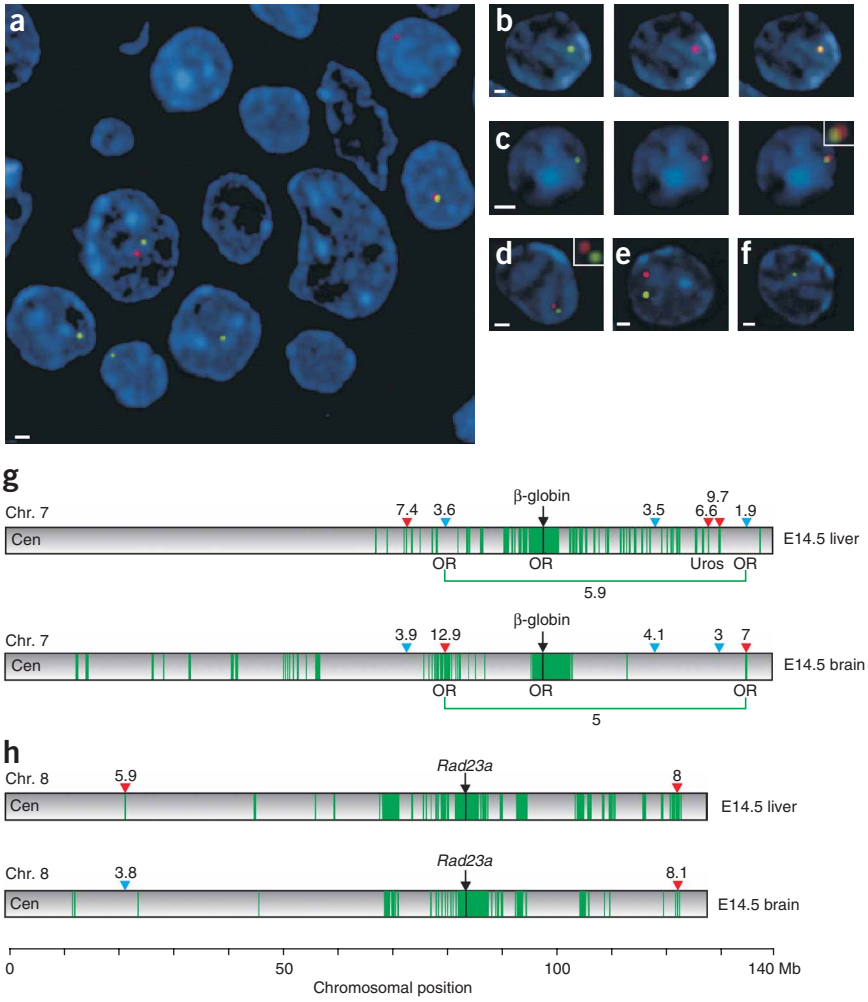

Figure 5 Cryo-FISH confirms that 4C technology truly identifies interacting regions. (**a**) Example of part of a 200-nm cryosection showing more than ten nuclei, some of which contain the β-globin locus (green) and/or *Uros* (red). Owing to sectioning, many nuclei do not show signals for these two loci. (**b**,**c**) Examples of completely (**b**) and partially (**c**) overlapping signals, which were all scored as positive for interaction. (**d**–**f**) Examples of nuclei containing non-contacting alleles (**d**,**e**) and a nucleus containing only β-globin (**f**), which were all scored as negative for interaction. Scale bars in **a**–**f**: 1 μm. (**g**–**h**) Schematic representation of cryo-FISH results. Percentages of interaction with β-globin (**g**) and *Rad23a* (**h**) are indicated above the chromosomes for regions positively identified (red arrowhead) and negatively identified (blue arrowhead) by 4C technology. The same BACs were used for the two tissues. Interaction frequencies measured by cryo-FISH between two distant olfactory receptor gene clusters in fetal liver and brain are indicated below the chromosomes. Interacting regions are on average 150–200 kb and are not drawn to scale.

overnight at a concentration of 100 ng/μl with 50 units (U) of a secondary, frequently cutting restriction enzyme (either *Dpn*II (used for HS2 and *Rad23a*) or *Nla*III (used for *Hbb-b1*)). Other combinations of restriction enzymes will also work, provided that the secondary, but not the primary, restriction enzyme is a frequent DNA cutter. To avoid constraints in DNA circle formation[17], we chose a secondary restriction enzyme that did not cut within 350–400 bp of the *Hin*dIII restriction site that demarcates the restriction fragment of interest (that is, the 'bait'). After secondary restriction enzyme digestion, DNA was phenol extracted, ethanol precipitated and subsequently ligated at low concentration (50 μg sample in 14 ml using 200 U ligase (Roche); incubated for 4 h at 16 °C) to promote *Dpn*II circle formation. Ligation products were phenol extracted and ethanol precipitated using glycogen (Roche) as a carrier (20 μg/ml). We linearized the circles of interest by digesting overnight with 50 U of a tertiary restriction enzyme that cuts the bait between the primary and secondary restriction enzyme recognition sites, using the following restriction enzymes: *Spe*I (for HS2), *Pst*I (for *Rad23a*) and *Pfl*mI (for *Hbb-b1*). This linearization step was performed to facilitate subsequent primer hybridization during the first rounds of PCR amplification. Digested products were purified using a QIAquick nucleotide removal (250) column (Qiagen).

PCR reactions were performed using the Expand Long Template PCR system (Roche) using conditions carefully optimized to assure linear amplification of fragments sized up to 1.2 kb (80% of 4C PCR fragments are <600 bp). PCR conditions were as follows: 94 °C for 2 min; 30 cycles of 94 °C for 15 s, 55 °C for 1 min and 68 °C for 3 min; followed by a final step of 68 °C for 7 min. We also carefully determined the maximum amount of template that still showed a linear range of amplification. For this, serial dilutions of template were added to PCR reactions; amplified DNA material was separated on an agarose gel and PCR products were quantified using ImageQuant software. Typically, 100–200 ng of template per 50 μl PCR reaction gave products in the linear range of amplification. We pooled 16 to 32 PCR reactions and purified this 4C template using the QIAquick nucleotide removal (250) system (Qiagen). Primer sequences are listed in **Supplementary Table 5** online.

Purified 4C template was labeled and hybridized to arrays according to standard chromatin immunoprecipitation (ChIP)-chip protocols (Nimblegen Systems). Differentially labeled genomic DNA, digested with the same primary and secondary enzyme, served as a control template to correct for differences in hybridization efficiencies. For each experiment, two independently processed samples were labeled with alternate dye orientations. Data were highly reproducible between independent experiments ($t > 0.99$; Spearman's rank correlation). Under the conditions described, sequences nearby on the chromosome template were together with the 'bait' so frequently that their large overrepresentation in our hybridization samples saturated the corresponding probes. This was confirmed when we performed hybridizations with samples diluted 1:10 and 1:100 and found that signal intensity was reduced at probes outside and at the edge of, but not inside, this region (data not shown).

4C arrays. An important aspect of our strategy is the use of dedicated microarrays. Recently, others have used techniques based on the 3C method to screen for colocalizing sequences[18,19]. However, these approaches were based on the sequencing of a limited number of captured fragments and, hence, did not provide a comprehensive overview of long-range interacting DNA segments. For our arrays, we designed probes (60-mers) that each represent a different *Hin*dIII restriction end in the genome. The advantages of such a design are that (i) each probe is informative, as each analyzes an independent ligation event, greatly facilitating the interpretation of the results, and (ii) a large representation of the genome can be spotted on a single array (which is cost effective). The array used for this study (Nimblegen Systems) covered seven complete mouse chromosomes. Arrays and analyses were based on NCBI build m34. Probes (60-mers) were selected from the sequences 100 bp up- and downstream of *Hin*dIII sites. To prevent cross-hybridization, probes that had any similarity with highly abundant repeats (as judged by RepBase 10.09)[20] were removed from the probe set. In addition, probes that gave more than two BLAST hits in the genome were also removed from the probe set. Sequence

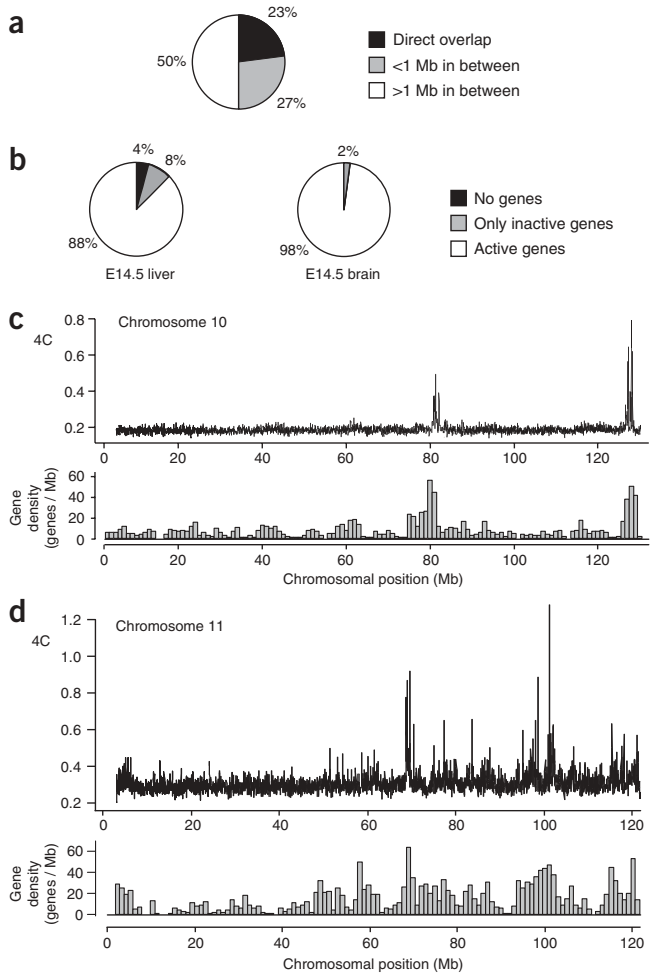

Figure 6 Interchromosomal interactions with *Rad23a*. (**a**) Interchromosomal interactions with *Rad23a* are conserved between tissues. Indicated are the percentages of *trans*-interacting regions observed in fetal liver that are identical, close to (that is, <1 Mb) or unrelated to the interchromosomal interactions detected in fetal brain. (**b**) The great majority of regions contacting *Rad23a* in *trans* contain active genes, both in fetal liver (left) and brain (right). (**c**,**d**) High-running median values of the 4C data (top) are found at chromosomal locations with a high gene density (bottom). Chromosomes 10 (**c**) and 11 (**d**) from a fetal brain sample are shown as examples. Data from fetal liver have a similar profile. Chromosomal positions were based on NCBI build m34.

alignments were performed using MegaBLAST[21] using the standard settings. A hit was defined as an alignment of 30 nt or longer.

4C data analysis. The ratio of sample-to–genomic DNA 4C signal was calculated for each probe, and the data were visualized with SignalMap software provided by Nimblegen Systems. Data were analyzed using the R package, Spotfire and Excel. Unprocessed hybridization ratios showed clusters of 20–50 strong signals along the chromosome template. It is important to realize that each probe on the array analyzes an independent ligation event. Moreover, only two copies of a given restriction fragment are present per cell, and each can ligate only to one other restriction fragment. Therefore, the detection of independent ligation events with 20 or more neighboring restriction fragments strongly indicates that the corresponding locus contacts the 4C bait in multiple cells. To define the clusters in *cis*, we applied a running mean or running median. We used various window sizes, ranging from 9–39 probes, which all identified the same clusters. Results shown are based on a window size of 29 probes (on average 60 kb) and were compared with the running mean performed across randomized data. This was done for each array separately. Consequently, all measurements were judged relative to the amplitude and noise of that specific array. The false discovery rate (FDR), defined as (number of false positives) / (number of false positives + number of true positives), was determined as follows: (number of positives in the randomized set) / (number of positives in the data). The threshold level was determined using a top-down approach to establish the minimal value for which FDR < 0.05.

Next, biological duplicate experiments were compared. Windows that met the threshold in both duplicates were considered positive. When comparing randomized data, no windows were above threshold in both duplicates. Positive windows directly adjacent on the chromosome template were joined (no gaps allowed), creating positive areas.

In defining interacting regions in *trans*, we took a similar approach. However, there we applied a running median, again with a window size of 29 probes. The threshold was set at an FDR of 0%. Thus, a region was called interacting when, in both duplicates, the median signal was higher than any signal found in the respective randomized data sets.

Expression analysis. For each tissue, three independent microarrays were performed according to Affymetrix protocol (mouse 430_2 arrays). Data were normalized using Bioconductor RMA ca-tools, and for each probe set, the measurements of the three microarrays were averaged. In addition, when multiple probe sets represented the same gene, they were also averaged. Mas5calls was used to establish 'present', 'absent' and 'marginal' calls. Genes with a 'present' call in all three arrays and an expression value >50 were called expressed. 'Fetal liver–specific genes' were classified as genes that met our criteria of being expressed in fetal liver and having over fivefold higher expression than in fetal brain. To provide a measure of overall transcriptional activity around each gene, a running sum was applied. For this, we used log-transformed expression values. For each gene, we calculated the sum of the expression of all genes found in a window 100 kb upstream of the start and 100 kb downstream of the end of the gene, including the gene itself. We compared the resulting values for active genes found inside 4C-positive regions ($n = 124$, 123 and 208, respectively, for HS2 in liver, *Rad23a* in brain and *Rad23a* in liver) with the values obtained for active genes outside 4C-positive areas ($n = 153$, 301 and 186, respectively, where $n = 153$ corresponds to the number of active, noninteracting genes present between the most centromeric interacting region and the telomere of chromosome 7; we compared the two groups using a one-tailed Wilcoxon rank sum test.

FISH probes. The following BAC clones (BACPAC Resources Centre) were used: RP23-370E12 for β-globin, RP23-317H16 for chromosome 7 at 80.1 Mb (olfactory receptor gene cluster), RP23-334E9 for *Uros*, RP23-32C19 for chromosome 7 at 118.3 Mb, RP23-143F10 for chromosome 7 at 130.1 Mb, RP23-470N5 for chromosome 7 at 73.1 Mb, RP23-247L11 for chromosome 7 at 135.0 Mb (olfactory receptor gene cluster), RP24-136A15 for *Rad23a*, RP23-307P24 for chromosome 8 at 21.8 Mb, RP23-460F21 for chromosome 8 at 122.4 Mb, RP24-130O14 for chromosome 10 at 74.3 Mb, RP23-153N12 for chromosome 11 at 68.7 Mb, RP23-311P1 for chromosome 11 at 102.2 Mb, RP24-331N11 for chromosome 14 at 65.1 Mb and RP23-236O12 for chromosome 14 at 73.7 Mb.

Random primer–labeled probes were prepared using BioPrime Array CGH Genomic Labeling System (Invitrogen). Before labeling, DNA was digested with *Dpn*II and purified with a DNA Clean-up and Concentrator 5 Kit (Zymo Research). Digested DNA (300 ng) was labeled with SpectrumGreen dUTP (Vysis) or Alexa Fluor 594 dUTP (Molecular Probes) and purified through a GFX PCR DNA and Gel Band Purification kit (Amersham Biosciences) to remove unincorporated nucleotides. The specificity of labeled probes was tested on metaphase spreads prepared from mouse embryonic stem (ES) cells.

Cryo-FISH. Cryo-FISH was performed as described before[10]. Briefly, E14.5 liver and brain were fixed for 20 min in 4% paraformaldehyde/250 mM HEPES (pH 7.5) and were cut into small tissue blocks, followed by another fixation step of 2 h in 8% paraformaldehyde at 4 °C. Fixed tissue blocks were immersed in 2.3 M sucrose for 20 min at room temperature (18–24 °C), mounted on a specimen holder and snap-frozen in liquid nitrogen. Tissue blocks were stored in liquid nitrogen until sectioning. Ultrathin cryosections of approximately

200 nm were cut using a Reichert Ultramicrotome E equipped with a cryo-attachment (Leica). Using a loop filled with sucrose, sections were transferred to coverslips and stored at –20 °C. For hybridization, sections were washed with PBS to remove sucrose, treated with 250 ng/ml RNase in 2× SSC for 1 h at 37 °C, incubated for 10 min in 0.1 M HCl, dehydrated in a series of ethanol washes and denatured for 8 min at 80 °C in 70% formamide/2× SSC, pH 7.5. Sections were again dehydrated directly before probe hybridization. We coprecipitated 500 ng labeled probe with 5 μg of mouse Cot1 DNA (Invitrogen) and dissolved it in hybridization mix (50% formamide, 10% dextran sulfate, 2× SSC, 50 mM phosphate buffer, pH 7.5). Probes were denatured for 5 min at 95 °C, reannealed for 30 min at 37 °C and hybridized for at least 40 h at 37 °C. After posthybridization washes, nuclei were counterstained with 20 ng/ml DAPI (Sigma) in PBS/0.05% Tween-20 and mounted in Prolong Gold antifade reagent (Molecular Probes).

Images were collected with a Zeiss Axio Imager Z1 epifluorescence microscope (100× plan apochromat, 1.4× oil objective), equipped with a charge-coupled device (CCD) camera and Isis FISH Imaging System software (Metasystems). A minimum of 250 β-globin or *Rad23a* alleles were analyzed and scored (by a person not knowing the probe combination applied to the sections) as overlapping or nonoverlapping with BACs located elsewhere in the genome. Replicated goodness-of-fit tests (*G* statistic)[22] were performed to assess significance of differences between values measured for 4C-positive versus 4C-negative regions. An overview of the results is provided in **Supplementary Tables 2** and **3**. The frequencies measured by cryo-FISH were considerably lower than those reported by others based on two-dimensional and three-dimensional FISH[8,9]. Although cryo-FISH may slightly underestimate actual interaction frequencies owing to sectioning, we expect that its increased resolution will provide more accurate measurements.

URLs. The R package can be downloaded from http://www.r-project.org. RMA ca-tools and Mas5calls for the analysis of Affymetrix microarray expression data can be found at http://www.bioconductor.org.

Accession codes. Gene Expression Omnibus (GEO): GSE5891.

Note: Supplementary information is available on the Nature Genetics website.

ACKNOWLEDGMENTS

We thank F. Grosveld for support and discussion and S. van Baal, M. Branco, A. Pombo, P. Verrijzer, J. Hou, B. Eussen, A. de Klein, T. de Vries Lentsch, D. Noordermeer and R.-J. Palstra for assistance. This work was supported by grants from the Dutch Scientific Organization (NWO) (912-04-082) and the Netherlands Genomics Initiative (050-71-324) to W.L.

AUTHOR CONTRIBUTIONS

M.S. set up, performed and analyzed 4C experiments and wrote the paper; P.K. performed cryo-FISH experiments; E.S. helped set up 4C; Y.M. helped with statistical analysis; R.W. prepared cryosections; E.d.W. and B.v.S. provided probe sequences and helped with statistical analysis and W.d.L. designed experiments and wrote the paper.

COMPETING INTERESTS STATEMENT

The authors declare that they have no competing financial interests.

Published online at http://www.nature.com/naturegenetics
Reprints and permissions information is available online at http://npg.nature.com/reprintsandpermissions/

1. Misteli, T. Concepts in nuclear architecture. *Bioessays* **27**, 477–487 (2005).
2. Sproul, D., Gilbert, N. & Bickmore, W.A. The role of chromatin structure in regulating the expression of clustered genes. *Nat. Rev. Genet.* **6**, 775–781 (2005).
3. Chakalova, L., Debrand, E., Mitchell, J.A., Osborne, C.S. & Fraser, P. Replication and transcription: shaping the landscape of the genome. *Nat. Rev. Genet.* **6**, 669–677 (2005).
4. Volpi, E.V. *et al.* Large-scale chromatin organization of the major histocompatibility complex and other regions of human chromosome 6 and its response to interferon in interphase nuclei. *J. Cell Sci.* **113**, 1565–1576 (2000).
5. Chambeyron, S. & Bickmore, W.A. Chromatin decondensation and nuclear reorganization of the HoxB locus upon induction of transcription. *Genes Dev.* **18**, 1119–1130 (2004).
6. Brown, K.E. *et al.* Association of transcriptionally silent genes with Ikaros complexes at centromeric heterochromatin. *Cell* **91**, 845–854 (1997).
7. Grogan, J.L. *et al.* Early transcription and silencing of cytokine genes underlie polarization of T helper cell subsets. *Immunity* **14**, 205–215 (2001).
8. Osborne, C.S. *et al.* Active genes dynamically colocalize to shared sites of ongoing transcription. *Nat. Genet.* **36**, 1065–1071 (2004).
9. Spilianakis, C.G., Lalioti, M.D., Town, T., Lee, G.R. & Flavell, R.A. Interchromosomal associations between alternatively expressed loci. *Nature* **435**, 637–645 (2005).
10. Branco, M.R. & Pombo, A. Intermingling of chromosome territories in interphase suggests role in translocations and transcription-dependent associations. *PLoS Biol.* **4**, e138 (2006).
11. Roix, J.J., McQueen, P.G., Munson, P.J., Parada, L.A. & Misteli, T. Spatial proximity of translocation-prone gene loci in human lymphomas. *Nat. Genet.* **34**, 287–291 (2003).
12. Lemaitre, J.M., Danis, E., Pasero, P., Vassetzky, Y. & Mechali, M. Mitotic remodeling of the replicon and chromosome structure. *Cell* **123**, 787–801 (2005).
13. Dekker, J., Rippe, K., Dekker, M. & Kleckner, N. Capturing chromosome conformation. *Science* **295**, 1306–1311 (2002).
14. Lomvardas, S. *et al.* Interchromosomal interactions and olfactory receptor choice. *Cell* **126**, 403–413 (2006).
15. Brown, J.M. *et al.* Coregulated human globin genes are frequently in spatial proximity when active. *J. Cell Biol.* **172**, 177–187 (2006).
16. Splinter, E., Grosveld, F. & de Laat, W. 3C technology: analyzing the spatial organization of genomic loci in vivo. *Methods Enzymol.* **375**, 493–507 (2004).
17. Rippe, K., von Hippel, P.H. & Langowski, J. Action at a distance: DNA-looping and initiation of transcription. *Trends Biochem. Sci.* **20**, 500–506 (1995).
18. Ling, J.Q. *et al.* CTCF mediates interchromosomal colocalization between Igf2/H19 and Wsb1/Nf1. *Science* **312**, 207–208 (2006).
19. Wurtele, H. & Chartrand, P. Genome-wide scanning of HoxB1-associated loci in mouse ES cells using an open-ended chromosome conformation capture methodology. *Chromosome Res.* **14**, 477–495 (2006).
20. Jurka, J. *et al.* Repbase Update, a database of eukaryotic repetitive elements. *Cytogenet. Genome Res.* **110**, 462–467 (2005).
21. Zhang, Z., Schwartz, S., Wagner, L. & Miller, W. A greedy algorithm for aligning DNA sequences. *J. Comput. Biol.* **7**, 203–214 (2000).
22. Sokal, R.R. & Rohlf, F.J. *Biometry: the Principles and Practice of Statistics in Biological Research* 3rd edn. (W.H. Freeman, New York, 1995).