



# Depletion of the Chromatin Looping Proteins CTCF and Cohesin Causes Chromatin Compaction: Insight into Chromatin Folding by Polymer Modelling

Marielis Tark-Dame<sup>1</sup>, Hansjoerg Jerabek<sup>2</sup>, Erik M. M. Manders<sup>3,4</sup>, Dieter W. Heermann<sup>2\*</sup>, Roel van Driel<sup>1\*</sup>

<sup>1</sup> Swammerdam Institute for Life Sciences, University of Amsterdam, Amsterdam, The Netherlands, <sup>2</sup> Institute for Theoretical Physics, Heidelberg University, Heidelberg, Germany, <sup>3</sup> Van Leeuwenhoek Centre for Advanced Microscopy, University of Amsterdam, Amsterdam, The Netherlands, <sup>4</sup> Department for Molecular Biotechnology, University of Gent, Gent, Belgium

## Abstract

Folding of the chromosomal fibre in interphase nuclei is an important element in the regulation of gene expression. For instance, physical contacts between promoters and enhancers are a key element in cell-type-specific transcription. We know remarkably little about the principles that control chromosome folding. Here we explore the view that intrachromosomal interactions, forming a complex pattern of loops, are a key element in chromosome folding. CTCF and cohesin are two abundant looping proteins of interphase chromosomes of higher eukaryotes. To investigate the role of looping in large-scale (supra Mb) folding of human chromosomes, we knocked down the gene that codes for CTCF and the one coding for Rad21, an essential subunit of cohesin. We measured the effect on chromosome folding using systematic 3D fluorescent *in situ* hybridization (FISH). Results show that chromatin becomes more compact after reducing the concentration of these two looping proteins. The molecular basis for this counter-intuitive behaviour is explored by polymer modelling using the Dynamic Loop model (Bohn M, Heermann DW (2010) Diffusion-driven looping provides a consistent framework for chromatin organization. *PLoS ONE* 5: e12218.). We show that compaction can be explained by selectively decreasing the number of short-range loops, leaving long-range looping unchanged. In support of this model prediction it has recently been shown by others that CTCF and cohesin indeed are responsible primarily for short-range looping. Our results suggest that the local and the overall changes in of chromosome structure are controlled by a delicate balance between short-range and long-range loops, allowing easy switching between, for instance, open and more compact chromatin states.

**Citation:** Tark-Dame M, Jerabek H, Manders EMM, Heermann DW, van Driel R (2014) Depletion of the Chromatin Looping Proteins CTCF and Cohesin Causes Chromatin Compaction: Insight into Chromatin Folding by Polymer Modelling. *PLoS Comput Biol* 10(10): e1003877. doi:10.1371/journal.pcbi.1003877

**Editor:** Marc A. Marti-Renom, CNAG - Centre Nacional d'Anàlisi Genòmica and CRG - Centre de Regulació Genòmica, Spain

**Received:** March 19, 2014; **Accepted:** August 18, 2014; **Published:** October 9, 2014

**Copyright:** © 2014 Tark-Dame et al. This is an open-access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

**Data Availability:** The authors confirm that all data underlying the findings are fully available without restriction. Raw experimental data and computational data can be accessed via the repositories of University of Amsterdam (<http://dare.uva.nl>) and Heidelberg University (<https://heidata.uni-heidelberg.de>) using the manuscript author data. Direct link to the data: <http://hdl.handle.net/11245/1.430734>.

**Funding:** This work was supported by a grant of the Dutch Foundation for Fundamental Research on Matter (FOM) as part of the research programme Physics of the Genome (Grant 07DNAA11 to RvD). Hansjoerg Jerabek was funded by the Heidelberg Graduate School of Mathematical and Computational Methods for the Sciences and the Institute for Theoretical Physics. The Netherlands Organization for Scientific Research (NWO) is acknowledged for a grant for using supercomputing facilities (Grant MP-190-12 to RvD). The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

**Competing Interests:** The authors have declared that no competing interests exist.

\* Email: [heermann@tphys.uni-heidelberg.de](mailto:heermann@tphys.uni-heidelberg.de) (DWH); [r.vandriel@uva.nl](mailto:r.vandriel@uva.nl) (RvD)

## Introduction

Chromosomes are highly folded inside the interphase nucleus. Their structure has been extensively studied by light and electron microscopy and more recently by genome-wide mapping of intra- and inter-chromosomal contacts. The picture that is emerging is that the chromosomal fibre is packed in a hierarchical way on different length scales [1–3]. Complete chromosomes are confined to chromosome territories (CTs) that intermingle only to a limited extent [4–6]. Another well-defined level is that of the topologically associated domains (TADs), which are distinct sub-chromosomal structures in the Mb range [7–10]. These two levels of chromosomal organisation are evolutionary conserved in metazoans. Chromosomal folding is intimately linked with genome function. TADs are functional genomic units, each containing genes that often are transcriptionally or epigenetically co-regulated

[2,9]. Specific sequence elements between TADs, called insulators, confine long-range regulatory interactions between for instance enhancers and promoters to the individual TADs [11]. TADs coincide with DNA replication units, which show a distinct pattern of replication timing that is remarkably well conserved between mouse and man [12–15]. Another striking aspect of chromatin folding is the considerable cell-to-cell variation observed in populations of otherwise identical cells [16,17]. Overall, chromosomal architecture seems a mix of well-defined and probabilistic components.

Despite the extensive information about chromosomal architecture and its importance for the functioning of the genome, underlying principles that direct chromosome folding are still elusive. What controls the dynamic folding of chromosomes? Such fundamental insight can be captured and explored in predictive computational models [18–21]. Models help to identify critical

## Author Summary

Folding of chromosomes in interphase nuclei of higher eukaryotes is a key element in regulating gene expression. The mechanisms that control chromatin folding are largely unknown. We have shown earlier that looping is a fundamental aspect of large-scale chromatin structure. Two abundant looping proteins are known: CTCF and cohesin. Here we combine quantitative fluorescent *in situ* hybridisation experiments in human cells with polymer modelling to unravel mechanisms of chromatin folding. We show that chromatin becomes more compact after depletion of looping proteins. This is remarkable, since polymer models describing chromatin predict decompaction. We present a polymer model that shows that specific biologically relevant looping regimes give rise to this behaviour. Importantly, chromosome conformation capture studies of mammalian chromatin support such looping regimes. Our results indicate that the local and overall compaction of the chromatin is defined by a subtle balance between short and long range loops; this may explain cell cycle and genome activity dependent structural transitions of chromatin.

experiments and make sure that proposed mechanisms are physically and thermodynamically sound. Ideally, models should correctly reproduce the hierarchical architecture of interphase chromosomes, the probabilistic aspects of chromatin folding, as well as the structural transitions that chromosomes undergo during mitosis, meiosis and cell differentiation. Importantly, models should be able to make predictions that can be experimentally tested. Since chromosomal fibres are long flexible one-dimensional structures, polymer models are a good first approximation for chromosomes. Recently, a variety of such models have been proposed based on information from genome contact maps [20,22–24]. So far, many models have not yet been thoroughly tested experimentally, which makes it difficult to assess their quality.

Recently, the Dynamic Loop polymer model (DL model) has been proposed, based on systematic 3D fluorescent *in situ* hybridisation (FISH) measurements on primary human fibroblasts, in combination with basic polymer physics [16,25]. The DL model is based on first principles and its main parameter is the looping probability, describing the chance that two non-adjacent monomers of the polymer make contact, i.e. form a loop. The DL model shows that a linear polymer with randomly positioned loops with a broad length distribution correctly recapitulates three basic aspects of interphase chromatin: (i) the formation of chromosome territories, due to the entropy-driven segregation of looped chromosomes, (ii) the presence of discrete sub-chromosomal domains that differ in chromatin compaction due to local differences in looping frequency, and (iii) the considerable cell-to-cell variation in chromatin folding [25–28]. In the DL model diffusional behaviour of the chromatin fibre, which is considerable in interphase nuclei [29], drives the dynamic formation of loops. Importantly, the DL model is compatible with results of genome-wide intra-chromosomal contact mapping experiments [30]. The model indicates that the distribution of loops along the chromosome controls chromatin folding. To unveil further principles of chromosome folding we now test predictions of the DL model by manipulating the chromosome looping frequency and measuring the effect on chromosomal structure.

Two abundant nuclear proteins that are involved in chromatin looping are CTCF and cohesin [31–33]. Their binding sites on the

genome in part co-localise indicating that they functionally cooperate [34,35]. Reducing the levels of CTCF and Rad21, an essential component of the cohesin complex, in the cell has been shown to reduce the number of chromatin loops [36–38]. Here we analyse by quantitative 3D FISH the effect of CTCF and/or Rad21 knockdown on chromosome structure in human primary G1 fibroblasts. The DL model, similar to most other polymer models, predicts that upon reduction of the number of chromatin loops chromosomes become less compact. In striking contrast to this prediction our experiments show that they become considerably more compact after knocking down CTCF and Rad21. These observations put major constraints on polymer models that aim to recapitulate the behaviour of interphase chromosomes. We show that the DL model can be adapted to correctly describe the observed compaction. Systematic model simulations strongly suggest that the looping regime, i.e. loop frequency along the chromosomal fibre and loop size distribution, is the key variable that controls chromosome folding. The cell is able to manipulate chromosome folding by regulating the activity and concentration of looping proteins, such as CTCF and cohesin. The model makes specific predictions that are supported by recent experiments of others, underscoring the potential of the DL model to explore underlying principles of chromosome folding. Predictions of the adapted DL model about structural transitions that chromosomes undergo during cell differentiation, mitosis and meiosis are briefly discussed.

## Results

Basic polymer models describe the relationship between the mean square physical distance (MSD) of two monomers of the polymer ( $\langle R^2 \rangle$ ) and their distance along the polymer, the contour distance (also referred to as contour length in our previous work), expressed in the number of monomers ( $N$ ). Experimentally, such a relationship can be measured for chromatin fibres in nuclei of fixed cells by FISH, using two probes that bind at a known genomic distance  $g$  (expressed in Mb) on the same chromosome. Confocal microscopy allows measuring their physical distance  $R$  ( $\mu\text{m}$ ) in 3D. Systematic 3D FISH permits critical comparison between model predictions and experimental observations. This approach resulted in the Dynamic Loop (DL) model, which predicts that a limited number of quasi randomly distributed loops on a chromatin fibre is sufficient to confine a chromosome to a chromosome territory [16,25]. The loop-size distribution observed in the DL model is in good agreement with what is observed in HiC experiments [30]. Moreover, the DL model shows that local differences in loop frequency along the chromatin fibre cause local differences in chromatin compaction, resulting in the formation of topological domains [27]. These and other approaches put chromatin looping on the central stage in controlling and switching large-scale chromatin structure in interphase nuclei. This implies that interfering with loop formation is likely to significantly affect chromatin structure. We tested this prediction by knocking down two proteins that play a major role in chromatin looping: CTCF and Rad21, an essential component of cohesin. The DL model predicts that a reduction of looping frequency results in chromatin expansion.

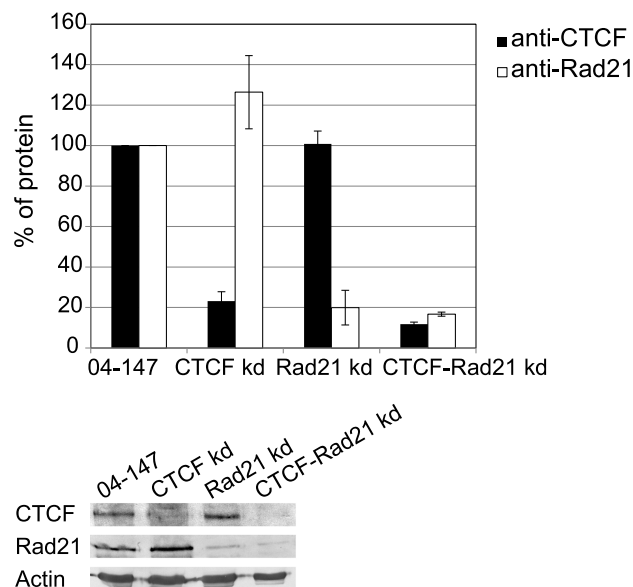
## Interfering with chromatin looping

The expression levels of CTCF and of cohesin were reduced by siRNA-mediated gene knockdown. To interfere with cohesin function we depleted its kleisin subunit Rad21. In the absence of this subunit none of the other cohesin subunits bind to chromatin [34]. The decrease in protein levels was quantified by Western

blotting of nuclear fractions. Figure 1 shows that knockdown of CTCF and cohesin individually and in combination results in a decrease of ~80% of their cellular concentration. A variety of studies have shown that such decrease in CTCF and cohesin significantly reduces chromatin looping [39,40], but has no significant effect on cell cycle progression [41]. All experiments were carried out with human primary female fibroblasts to avoid possible effects of immortalisation. Depletion of CTCF and/or cohesin had no measurable effect on the cellular and nuclear morphology and the size of the nucleus during the time course of the experiments, i.e. up to 72 hours after siRNA transfection (Figures S1 and S2). Also the viability of cells was not affected and the percentage of apoptotic cells among cells analysed did not change, as demonstrated by AnnexinV staining (Figure S3)

### Depletion of CTCF and cohesin results in chromatin compaction

The relationship between the mean square physical distance (MSD)  $\langle R^2 \rangle$  and the genomic distance  $g$  between FISH probes was determined for the same genomic regions as studied earlier in establishing the DL model, i.e. the q-arms of chromosomes 1 and 11 [16]. We concentrated on an about 3 Mb size gene-rich region gene-poor region of similar length on chromosome 1 and on a 27 Mb and a 70 Mb region on the q-arms of chromosome 1 and 11, respectively. The latter spans the complete q-arm of chromosome 11. Figures 2A and 2B show the analysed regions and the distribution of the BAC (bacterial artificial chromosome) probes used for FISH labelling. The BACs used in this study and the number of FISH probe-pairs analysed is listed in Table S1. Figure S6 shows the positions of the BACs on the genomic maps of the relevant parts of chromosomes 1 and 11 and relates them to the human transcriptome map [42], replication domains [43] and the HiC topological domains of similar cell types [7].



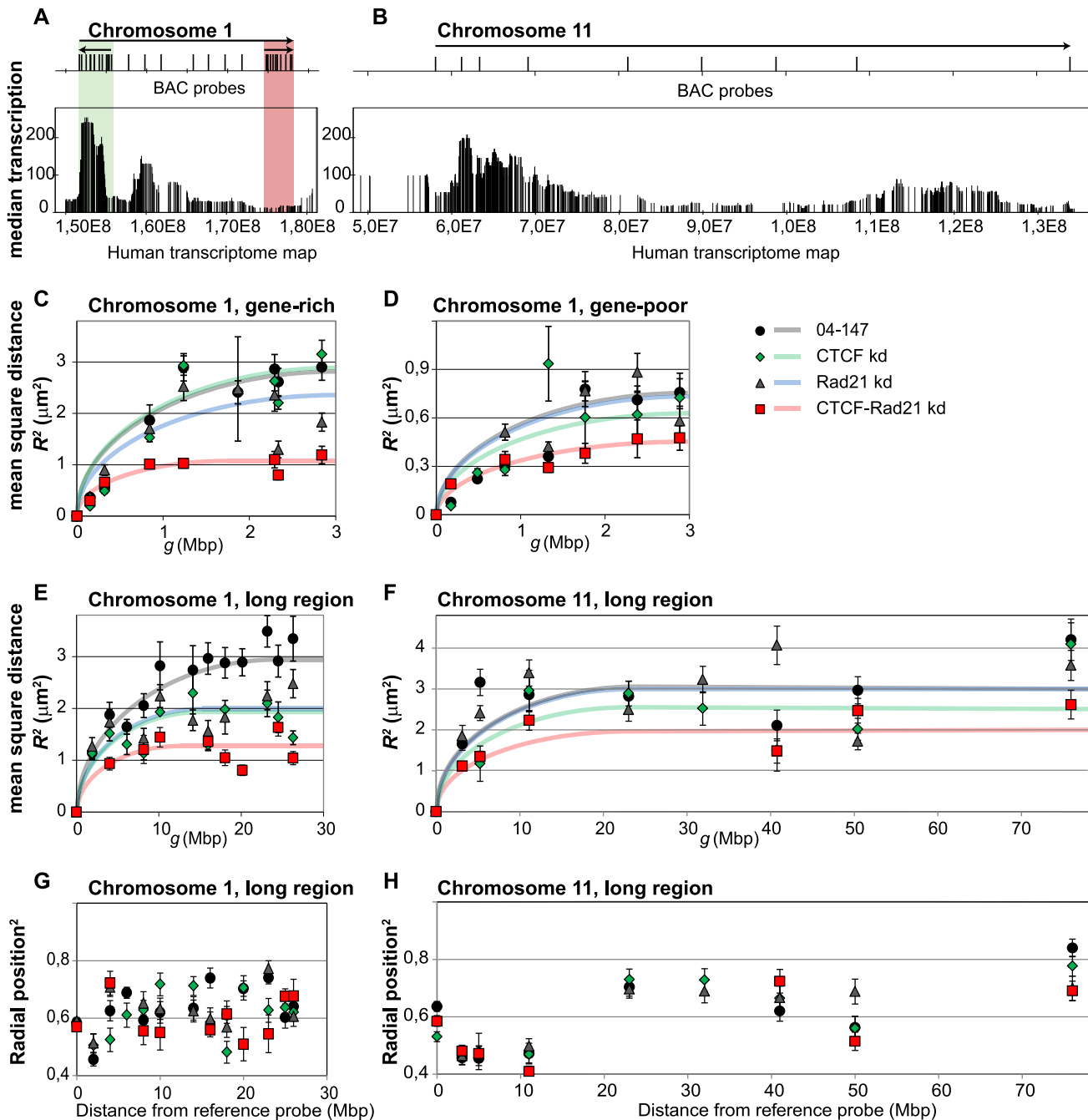
**Figure 1. Knock-down of CTCF and Rad21.** The nuclear fraction of control cells (04-147) and cells depleted of CTCF, Rad21 or both simultaneously were analysed by SDS-PAGE followed by immunoblotting.  $\beta$ -Actin was used as a control when calculating changes in CTCF and Rad21 protein levels after siRNA-mediated knock down. Average protein amounts of at least three independent experiments are shown. Error bars represent standard deviation. doi:10.1371/journal.pcbi.1003877.g001

Only cells in G1 were used for these analyses. In the randomly growing cell culture S-phase cells were identified based on their incorporation of bromodeoxy uridine (BrdU) during DNA replication and G2 cells could be recognised by their significantly larger size [16,44].

Figures 2C and D show the relationship between the MSD of two FISH probes and their genomic distance in two about 3 Mb regions on chromosome 1q, one being gene-rich (green) and the other gene-poor (red). Results from nuclei of control cells and of cells in which CTCF and cohesin were knocked down either individually or simultaneously are shown. Simultaneous knock-down of the two proteins resulted in considerable chromatin condensation, as shown by the decrease of the MSD plateau levels. The maximum average distances between the FISH probes decreased from about 2.7  $\mu$ m to 1  $\mu$ m for the gene-rich regions and from about 0.7  $\mu$ m to close to 0.5  $\mu$ m for the gene-poor regions. Knockdown of CTCF and cohesin individually had a considerably smaller effect. Control experiments using non-target siRNA did not reveal significant changes in chromatin compaction (Figure S4). Chromatin condensation is particularly evident for gene-rich areas, which are relatively open in control cells, and is smaller for the gene-poor regions that already are relatively compact in untreated cells. Figures 2E and F show the same type of measurements for the longer genomic stretches: 27 Mb of chromosome 1q and 70 Mb of chromosome 11q. At these longer distances the decrease of CTCF or of cohesin concentration reduced the maximum probe distance from about 3  $\mu$ m to around 2  $\mu$ m for chromosome 1q, indicating chromatin condensation, as found for shorter distances. Knockdown of both proteins simultaneously reduced the average distance further to close to 1  $\mu$ m, indicating a three-fold compaction. For the 70 Mb region of chromosome 11q the effect of knocking down of CTCF and cohesin individually is less pronounced. The maximum average probe distance is about 3  $\mu$ m, similar to the value for the chromosome 1q region in control cells. Simultaneous knockdown of CTCF and cohesin reduced this value to between 2 and 2.5  $\mu$ m. What causes this quantitative difference in the behaviour of the two chromosomes is unclear. The average frequency of CTCF binding sites in both genomic regions is similar [45]. Importantly, depletion of the two looping proteins does not result in decondensation of chromatin, in contrast to what is predicted by most chromatin-inspired polymer models, including the DL model. Instead, after CTCF and cohesin depletion the opposite is observed, i.e. condensation of chromatin.

### The radial distribution of chromatin in the nucleus does not change after CTCF and cohesin knockdown

Although depletion of CTCF and cohesin causes the analysed regions to become more compact than in control cells, there are no evident changes in chromatin density when comparing DAPI staining signals of these cells (Figure S1). To find out if compaction of chromatin in CTCF and cohesin depleted cells results in large-scale chromatin rearrangements in the nucleus that remain unseen in DAPI staining, we analysed the radial position of probes used for 3D FISH measurements in the nucleus. No significant changes in radial distribution of sites on the q-arms of chromosomes 1 and 11 were observed (Figures 2G and H), indicating that depletion of CTCF and cohesin has no effect on overall nuclear organization. The position of a locus in nuclear space and its compactness are correlated: gene-rich and open chromatin regions on average are located closer to the nuclear centre, whereas compact and gene-poor regions are located closer to the nuclear periphery [44,46,47]. Figure 2G shows that the analysed regions on chromosome 1q have approximately the same



**Figure 2. *In situ* chromatin structure.** (A, B) The genomic context of analysed regions of the q-arms of human chromosome 1 (A) and chromosome 11 (B) is illustrated by the human transcriptome map. Vertical lines represent protein coding genes, the length depicting the average transcription level in a number of human tissues and cell lines [42]. The location of BAC probes used for FISH labelling are indicated above the maps. Arrows above the maps designate the regions where spatial distances between pairs of BAC probes were measured; the start of the arrow indicating the position of the reference BAC probe. On chromosome 1q three regions were analysed. Two 3 Mb regions were selected in a gene-rich and transcriptionally active region (green box) and in a region characterized by low gene density and low transcription (red box). (C–F) Graphs show the mean square physical distances  $\langle R^2 \rangle$  between the FISH probes as a function of the genomic distance  $g$  for a 3 Mb gene-rich (C) and a gene-poor (D) region on chromosome 1 q-arm. Longer regions of chromosome 1 (30 Mb) (E) and of chromosome 11 (76 Mb) (F) were also analysed. Each data point represents an average of 30–150 distance measurements in at least two independent experiments. Error bars represent standard errors. Graphs are shown for control cells (black dots) and cells after knockdown of CTCF (green diamonds), of cohesin subunit Rad21 (blue triangles) and of CTCF and Rad21 simultaneously (red squares). Results show that chromatin becomes more compact after knockdown of CTCF and Rad21 (cohesin). (G, H) Radial positions of BAC probes in nuclei for chromosome 1q (G) and chromosome 11q (H). Radial nuclear position of a BAC probe was calculated the distance between the centre of gravity of a FISH labelled genomic site and the centre of gravity of the nucleus and dividing that distance with the length of a straight line drawn from the nuclear centre to the nuclear envelope through the centre of the gravity of the labelled site. Error bars represent standard errors. Note that some chromatin regions were poorly labelled in CTCF and/or cohesin knock down conditions and therefore omitted from the graph.

doi:10.1371/journal.pcbi.1003877.g002

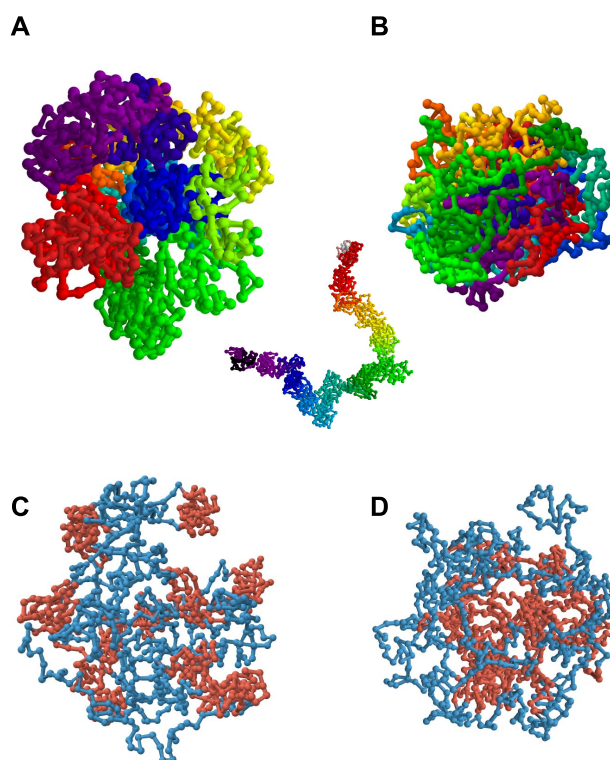
radial position. On chromosome 11 two well-defined regions can be identified (Figure 2H): the probes that label the centromere-proximal 0–11 Mb of the analysed region are closer to the nuclear centre, while the probes marking the 23–76 Mb area are located more towards nuclear periphery. This 23–76 Mb area of chromosome 11 is closer to the nuclear periphery compared to the probes on chromosome 1. In combination with measurements of relative volume of these regions carried out previously [44], these results indicate that in control cells the 23–76 Mb domain of chromosome 11 has a more compact structure than the 0–11 Mb region on chromosome 11 and the analysed regions on chromosome 1.

### Polymer modelling of interphase chromatin

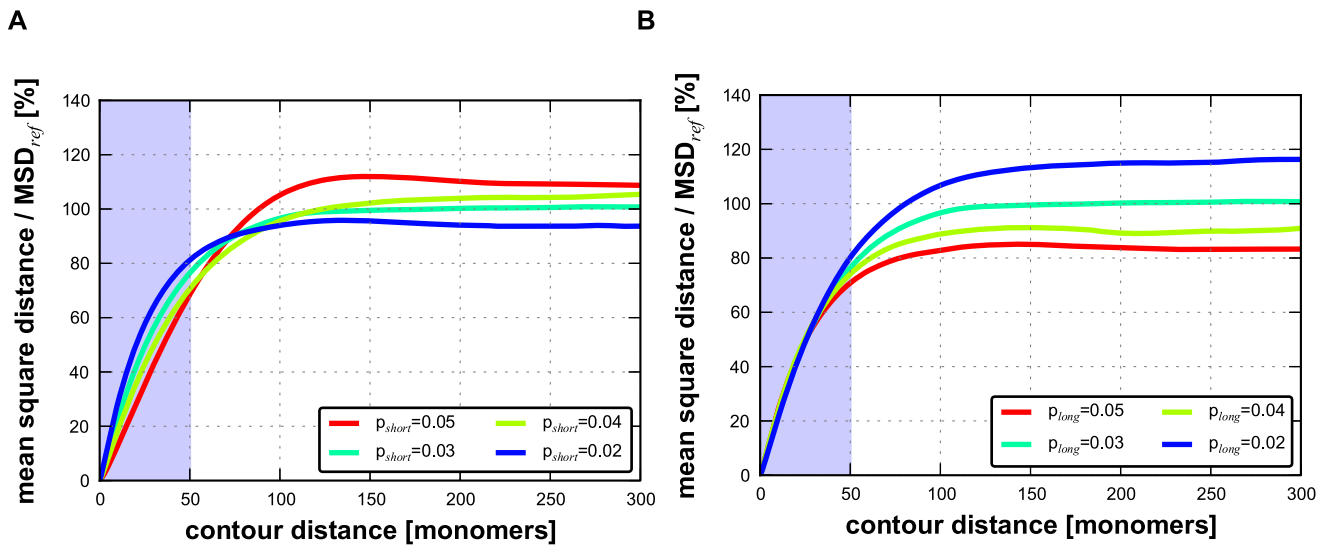
Recently, we proposed a simple polymer model that explains the confinement of an interphase chromosome to its chromosome territory. In this DL model the chromosomal fibre is represented as a self-avoiding random walk polymer forming probabilistic intra-polymer contacts between non-adjacent monomers [25]. As a consequence, loops with a broad size distribution are formed. The main model parameter is the looping probability ( $p$ ), which is a measure for the probability that a bond (loop) is formed between two non-adjacent monomers. Due to the proximity of monomers, short-range loops are formed more frequently than long-range ones, yielding a loop size distribution comparable to that observed in HiC experiments [30]. In the DL model a decrease in looping probability results in an increase of the volume occupied by the polymer, i.e. decondensation. Interestingly, our experimental results show that chromatin in interphase nuclei behaves in an opposite way, i.e. a decrease in chromatin looping due to depletion of CTCF and cohesin results in compaction. In the DL model decreasing the looping probability affect loops of all lengths. However, in cells CTCF seems involved preferentially in mediating short-range loops, i.e. below 1 Mb length [48].

To explore whether the observed chromatin compaction may be caused by the predominant depletion of short-range loops, we modified the DL model by imposing different looping probabilities for short-range and long-range looping ( $p_{\text{short}}$  and  $p_{\text{long}}$ , respectively). In simulations with this adapted DL model we used polymers with a total length of 1050 monomers (for further details see Materials and Methods section). Short-range loops were defined as those spanning 50 monomers or less and long-range loops as those spanning 51 monomers or more. For comparison, for human chromosomes in sizes 50–250 Mb the cut-off between short- and long-range loops would be approximately 2.5–12.5 Mb. Subsequently, we calculated the MSD *versus* contour distance relationship for a matrix of combinations of short- and long-range looping probabilities. Figure 3A illustrates the typical compact conformation of an adapted DL polymer with high  $p_{\text{short}}$  (0.12) and low  $p_{\text{long}}$  (0.04) probabilities. The colour code labels the monomers along the polymer according to the visible spectrum along the polymer. The inset in panel A displays the conformation for the same  $p_{\text{short}}$  (0.12) after  $p_{\text{long}}$  has been set to zero, showing that a high probability of the short-range interactions ( $p_{\text{short}}$ ) in the absence of long-range interactions results in a uniform thick fibre. Figure 3B shows a configuration of the DL polymer with low  $p_{\text{short}}$  (0.04) and high  $p_{\text{long}}$  (0.12) probabilities, showing a more chaotic folding of the polymer compared to Figure 3A. Figures 4A and B exhibit examples of the calculated relationships between the MSD and the contour distance, covering a range of  $p_{\text{short}}$  values (0.02–0.05), while keeping  $p_{\text{long}}$  constant (0.03) (Figure 4A), and of  $p_{\text{long}}$  (0.02–0.05) values at constant  $p_{\text{short}}$  (0.03) (Figure 4B). The plateau level MSD value is used as a measure for the overall compaction of the polymer: a lower MSD values indicates a stronger compaction.

Remarkably, these simulations show that decreasing the long-range looping frequency  $p_{\text{long}}$  results in expansion of the polymer, while decreasing  $p_{\text{short}}$  leads to compaction. The heat map of Figure 5A displays the MSD plateau level of the polymer as a function of the  $p_{\text{short}} - p_{\text{long}}$  parameter space. The arrows depict the effect of decreasing the  $p_{\text{short}}$  value at constant  $p_{\text{long}}$  and of decreasing the  $p_{\text{long}}$  value at constant  $p_{\text{short}}$ , underscoring that compaction occurs only if the short-range looping probability is decreased, whereas a reduction of  $p_{\text{long}}$  results in de-compaction. The qualitative explanation for this behaviour is explained in the Discussion. Evidently, the DL-model model analysis does not intend to reflect the precise behaviour of interphase chromosomes. It shows how different looping regimes may dramatically affect large scale chromosome folding. As can be seen in Fig. 5 the behaviour is qualitatively the same over a large range of the  $p_{\text{long}}$  and  $p_{\text{short}}$  parameter landscape. Polymer modelling shows that the chromatin compaction we observe after knocking down CTCF and cohesin may be attributed to a specific reduction of short-range loops.



**Figure 3. Example of polymer conformations of the adapted DL model at different looping regimes.** (A) Conformation of the adapted-DL polymer with high short-range ( $p_{\text{short}} = 0.12$ ) and low long-range ( $p_{\text{long}} = 0.04$ ) looping probabilities and (B) the same polymer with low short-range ( $p_{\text{short}} = 0.04$ ) and high long-range ( $p_{\text{long}} = 0.12$ ) looping probabilities. The colour code labels the monomers of the polymer according to the visible spectrum along the length of the polymer. The inset in panel A displays the conformation for the same  $p_{\text{short}}$  (0.12) after  $p_{\text{long}}$  has been set to zero, showing that a high probability of the short-range interactions ( $p_{\text{short}}$ ) in the absence of long-range interactions results in a uniform thick fibre. Figure 3B shows a configuration of the DL polymer with low  $p_{\text{short}}$  (0.04) and high  $p_{\text{long}}$  (0.12) probabilities, showing a more chaotic folding of the polymer compared to Figure 3A. Figures 4A and B exhibit examples of the calculated relationships between the MSD and the contour distance, covering a range of  $p_{\text{short}}$  values (0.02–0.05), while keeping  $p_{\text{long}}$  constant (0.03) (Figure 4A), and of  $p_{\text{long}}$  (0.02–0.05) values at constant  $p_{\text{short}}$  (0.03) (Figure 4B). The plateau level MSD value is used as a measure for the overall compaction of the polymer: a lower MSD values indicates a stronger compaction. More adapted DL polymer conformations with the same looping probabilities as shown in (A) and (B) are shown in Figure S5. (C) Conformation of the domain-adapted DL polymer with high short-range ( $p_{\text{short}} = 0.16$ ) and low long-range ( $p_{\text{long}} = 0.02$ ) looping probabilities and (D) the same polymer with low short-range ( $p_{\text{short}} = 0.02$ ) and high long-range ( $p_{\text{long}} = 0.16$ ) looping probabilities. Here, topological domains are labelled red and non-looping linker regions blue. doi:10.1371/journal.pcbi.1003877.g003

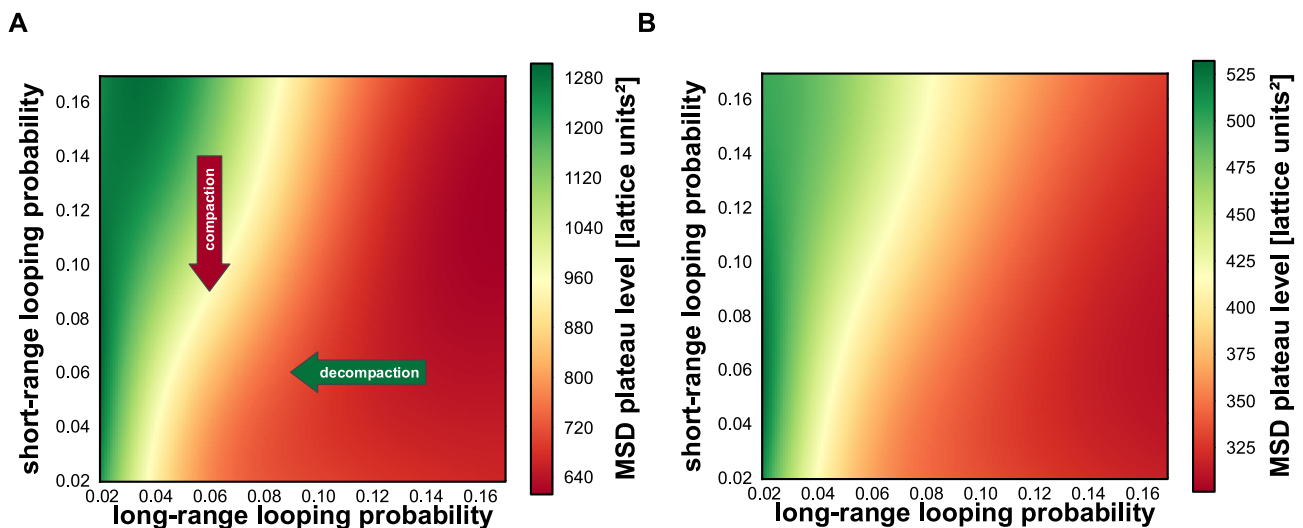


**Figure 4. Effect of short-range and long-range looping probabilities on the relationships between the MSD and the contour distance in the adapted DL polymer model.** (A) Varying the short-range looping probability at constant long-range looping probability ( $p_{long}$  0.03). (B) Varying long-range looping probability at constant short-range looping ( $p_{short}$  0.03). The MSD plateau levels is a measure of overall polymer compaction. The simulations show that decreasing the long-range looping probability results in expansion of the polymer, while lowering short-range looping probability leads to compaction. To facilitate comparison between the two effects, the MSD was normalized by the plateau level of the MSD for  $p_{long}=0.03$  and  $p_{short}=0.03$  ( $MSD_{ref}$ ) because this parameter combination is present in both figures. doi:10.1371/journal.pcbi.1003877.g004

#### Domain substructure of chromosomes

The adapted DL model does not incorporate what seems to be a key aspect of chromatin architecture, namely the presence of topological sub-chromosomal domains. Several independent lines of evidence, including replication timing, HiC analysis, clustering of epigenetic marks and electron microscopy, indicate that chromosomes of higher eukaryotes consist of arrays of topological domains in the size 0.5 to 5 Mb range [7–10]. In terms of the adapted DL model parameters this means that short-range looping interactions are locally clustered at many positions along the

polymer, rather than evenly distributed as assumed in the adapted DL model. We wondered whether a polymer consisting of an array of topological domains would recapitulate the chromatin compaction we observe in our experiments when reducing the looping probability. To force the adapted DL model polymer to mimic topological domains we assumed that the polymer consists of regions that alternately contain a 50 monomer domain with high short-range looping probability ( $p_{short}$ ) and one that is devoid of short-range loops. For simplicity, in this domain-adapted DL model the two types of regions were made equal in length (50



**Figure 5. Heat plots showing how polymer compaction varies in the  $p_{short}$  -  $p_{long}$  parameter space.** Polymer compaction is represented as MSD plateau level as shown in Figure 4. (A) adapted DL polymer, (B) domain-adapted DL polymer. The arrows in (A) show that decreasing the short-range looping probability (red arrow) results in compaction, while reducing the long-range looping probability (green arrow) leads to decompaction. In the heat plot red marks compaction, green decompaction. doi:10.1371/journal.pcbi.1003877.g005

monomers) resulting in 10 highly and 11 lowly looped domains in the model polymer of again in total 1050 monomers. Long-range loops were allowed between the compact domains only with a probability  $p_{\text{long}}$ . Figure 3C shows the characteristic domain-like structure of the domain-adapted DL polymer under conditions of high short-range looping probability ( $p_{\text{short}} = 0.16$ ) in combination with a low long-range probability ( $p_{\text{long}} = 0.02$ ). As expected, compact domains (red in Figure 3C) are formed and connected by non-structured linker stretches (blue). When lowering the looping probability in the domains to  $p_{\text{short}} = 0.02$  and increasing the looping between the domains ( $p_{\text{long}} = 0.16$ ), the less compact individual domains (red) cluster due to long-range looping (Figure 3D). The heat map of the variation of polymer compaction in the  $p_{\text{short}} - p_{\text{long}}$  parameter space is shown in Figure 5B. Comparison (Figures 5A and B) shows that the behaviour of polymers with and without domain-structure is very similar after decreasing short-range or the long-range looping probabilities. The former induces compaction and the latter results in expansion of the polymer. Evidently, introducing discrete topological domains in the adapted DL polymer model does not affect its capacity to show compaction after decreasing the short-range looping probability  $p_{\text{short}}$ .

## Discussion

The 46 human chromosomes, in total consisting of 5 cm nucleosomal fibre, are packed in an interphase nucleus with a diameter in the order of 10  $\mu\text{m}$ . Intra-chromosomal interactions forming chromatin loops play an important role in compacting interphase chromosomes. At the same time loops are a key component in gene regulation, for instance because they allow physical contacts between distant regulatory elements, such as promoters and enhancers [11]. Our understanding of basic principles of chromosome folding and how these affect gene regulation and other genomic functions is still limited. Polymer modelling based on experimental data sets begins to unveil general aspects of large scale (supra Mb) chromosome folding. The Dynamic Loop (DL) model shows that intra-chromosomal loops can explain the relatively compact nature of chromosome territories in interphase nuclei and why they intermingle only to a limited extent. Also, variations in looping frequency along the chromosomal fibre result in sub-chromosomal domains with different compaction [25–28].

All polymer models that are based on chromosome looping [16,49] predict that a decrease in the number of loops results in chromosome expansion. Here we test this prediction in primary human fibroblasts by reducing the concentration of two abundant proteins that are involved in loop formation, i.e. CTCF and cohesin. Knocking down individually or simultaneously the CTCF gene and a gene coding for Rad21, an essential component of cohesin, reduces their concentration to about 20% of the initial levels (Figure 1). Others have shown that such decrease of CTCF and cohesin concentrations indeed results in a decrease in the number of chromosomal loops [36–38].

Remarkably, instead of the predicted expansion we observe a significant condensation of chromosome, in full contrast to the model predictions (Figure 2). These results confirm the importance of looping in controlling how chromosomes are folded. At the same time they put severe constraints on looping-based polymer models that should recapitulate the configuration of interphase chromosomes. Because the part of the binding sites on the genome of CTCF and cohesin co-localise [34,35], it can be expected that knocking down of each of these proteins individually

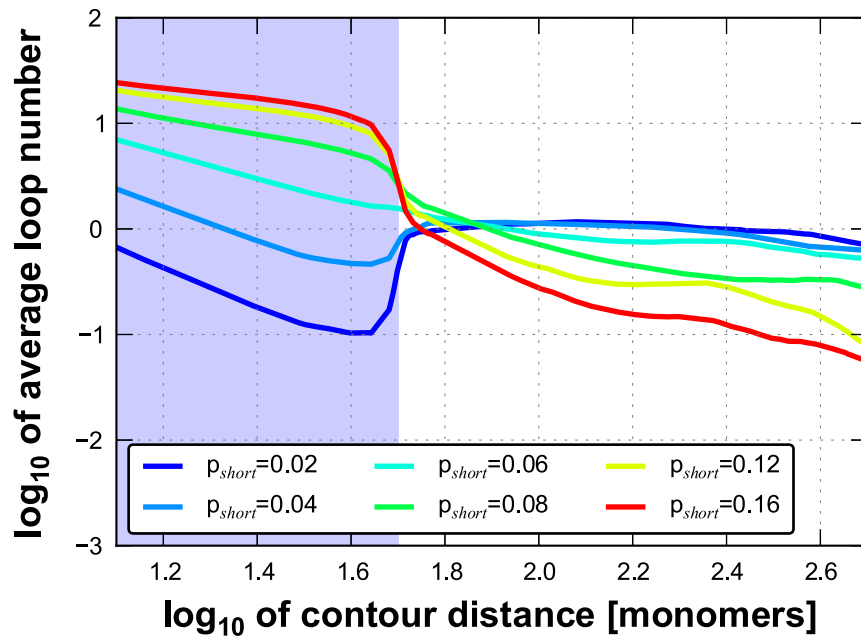
will affect similar effects, whereas their simultaneous knock down has a larger effect. This is exactly what is observed (Figure 2C–F).

To explain our observations we started from the DL model, which assumes that loops with a broad length distribution are randomly positioned along the chromosomal fibre. Systematic model simulations showed that selectively decreasing the frequency of short loops, without affecting longer loops, recapitulates the chromatin compaction observed in our experiments after decreasing the looping frequency by reducing the cellular concentration of CTCF and cohesin (Figures 4A, B and 5A). To do so we introduced separate parameters for the probability of short loops and of long loops (shorter than 5% of the polymer length and 5% or longer, respectively). Systematic simulations exploring variations of polymer compaction in the short-range vs. long-range looping probabilities show that only in a specific part of the parameter space a decrease of short-range looping frequencies results in significant condensation of the polymer. In contrast, decrease of long-range looping frequencies always results in decondensation.

The observed chromatin compaction after decreasing the number of loops is counter-intuitive. The rationale is that reducing the number of short loops leads to expansion of the polymer on the short scale, resulting in an increase in volume occupied by the polymer. However, at the same time the decrease of the number of short loops leads to reduced entropic intra-polymer repulsion [27]. Consequently, parts of the polymer that are in close spatial proximity, e.g. due to the long-range interactions, intermingle stronger, which in turn makes the formation of a long-range loops more probable. This means that for constant  $p_{\text{long}}$  the number of long-range loops ( $n_{\text{long}}$ ) indirectly increases when reducing the number of short-range loops (Figure 6). As the entropy-driven volume reduction, due to this intra-polymer intermingling, exceeds the increase of volume due to the short-range polymer swelling, the volume occupied by the complete polymer decreases, i.e. it condenses. It may be argued that after de-compaction of the polymer at low length scales the probability of long range interaction increases because more such interaction sites become available. As can be seen in Fig. 5 such increase of  $p_{\text{long}}$  would further enhance the compaction process.

The above considerations predict that CTCF and cohesin are mainly involved in short-range looping. In support of this, Handoko et al. [48] showed that CTCF is predominantly engaged in the Mb-range organisation of chromosomes. Cohesin and CTCF are known to cooperate in the spatial organisation of chromosomes [34,35]. In line with this, decreasing the activity of CTCF and cohesin preferentially affects looping at short-range, below Mb [36–38].

There is ample experimental evidence that, in contrast to what the adapted DL model assumes, the distribution of short-range loops is not random. Rather, short-range looping sites seem to cluster at multiple sites along the chromosome, resulting in topological domains in the 0.5–5 Mb size range [7–10]. Examples are distinct sub-chromosomal domains that differ in replication timing and in epigenetic state [2,9,12–15]. To mimic these topological domains in the adapted DL model, the short-range loops were clustered at multiple domains along the polymer, intervened by areas devoid of looping. Model simulations show that this redistribution of loops has little effect on the condensation-decondensation behaviour of the polymer (Figure 5B). Evidently, the domain-adapted DL polymer model reproduces the condensation of the polymer after reducing the number of short loops similar to the adapted DL model without domains (Figure 5A). Interestingly, after knockdown of CTCF and cohesin the Mb-size topological domains do not dissolve, indicating that at



**Figure 6. Relationship between looping probability and number of established loops.** When keeping the long-range looping probability ( $p_{\text{long}}$ ) constant while varying short-range looping probability ( $p_{\text{short}}$ ), the number of long-range loops ( $n_{\text{long}}$ ) increases when reducing  $p_{\text{short}}$ . This is because lowering  $p_{\text{short}}$  reduces the loop repulsion on the short scale and hence increases intermingling on the long scale, which in turn increases the frequency of long-range loop formation. Here,  $p_{\text{long}}=0.06$  and  $p_{\text{short}}$  is varied from 0.02 to 0.16. The blue area marks the short-range looping regime ( $<50$  monomers) while the rest describes the long-range looping. doi:10.1371/journal.pcbi.1003877.g006

least part if the intra-domain loops remain intact. Zuin et al [36] showed that the number of interactions between chromatin from different topological domains, i.e. long-range loops, increases. As can be seen in Figure 5, the adapted DL model predicts that in addition to the entropic effects the increase of the frequency of large loops further contributes to polymer compaction.

In recent years, several polymer models have been proposed for the folding of interphase chromatin. The ‘Strings and Binders Switch’ (SBS) model of Barbieri et al. [49]), which assumes a diffusible component responsible for loop formation by linking two monomers of the polymer, is a special case of the DL model [16,25]. In the DL model the properties of such ‘binders’ are incorporated in the looping probability parameter. As expected, the SBS model correctly mimics the formation of chromosome territories in the same way as the DL model. Another class of polymer models describes chromatin as a fractal polymer [22,23]. This type of model was proposed to explain the chromatin folding based on HiC contact maps in the 0.5–10 Mb range [22,50]. Since these models do not involve polymer looping-related parameters in their present form, they do not allow making predictions about how changes in the looping regime relate to changes in compaction. Systematic quantitative FISH measurements on chemically fixed cells, as used in this and other papers [16,49,51], do not support fractal models.

We can only speculate about the biological relevance of chromatin compaction after decreasing the concentration of the looping proteins CTCF and cohesin. Conceivably, this may be a step in the transition to metaphase chromosomes at the onset of cell division. A recent genome-wide interactome study shows that there are marked changes in intra-chromosomal contacts in metaphase chromosomes compared to the situation in interphase [52]. Whatever the mechanism of metaphase chromosome formation is, it most likely involves the disappearance of long-range loops. The inset in Figure 3A suggestively shows that

complete abolishment of long-range loops results in metaphase-like structures in a state before axial condensation.

Taken together, the adapted DL model is successful in predicting various key properties of interphase chromosomes: (i) distinct and poorly intermingling chromosome territories, (ii) local differences in chromatin compaction along the chromosomal fibre, and (iii) chromosome compaction after reducing the looping frequency. The model shows that chromosomal looping constitutes the basis for this behaviour. Relatively subtle changes in the balance between short-range and long-range looping probabilities lead to changes in local and overall chromosome folding and compaction, probably including the interphase-metaphase transitions.

## Materials and Methods

### Cell culture

Human primary female fibroblasts (04–147) were cultured in DMEM containing 10% FCS, 20 mM glutamine, 60  $\mu\text{g}/\text{mL}$  penicillin and 100  $\mu\text{g}/\text{mL}$  streptomycin (Gibco, Life Technologies Corporation, Carlsbad, CA, USA). Cells were used up to passage 25 to avoid effects related to senescence.

### RNA interference

siRNA transfections were performed using lipofectamine RNAiMAX (Invitrogen, Life Technologies Corporation, Carlsbad, CA, USA) according to the reverse transfection protocol of the manufacturer. siRNAs used are listed in Table S2. The transfection efficiency of siRNA was estimated with the BLOCK-iT Alexa FluorRed Fluorescent Control (Invitrogen, Life Technologies Corporation, Carlsbad, CA, USA) in three independent experiments. In each experiment 100 nuclei were scored based on DAPI signal, 99% of which showed red fluorescence and therefore had been transfected successfully. Knockdowns were verified by



Western blot. Cells were resuspended in lysis buffer (10 mM HEPES, 10 mM KCl, 1.5 mM MgCl<sub>2</sub>, 0.34M sucrose, 10% glycerol, 1 mM DTT, complete protease inhibitor cocktail (Roche, F. Hoffmann-La Roche Ltd, Basle, Switzerland)) containing 0.1% Triton X-100, incubated 10 min on ice and centrifuged 5 min at 1300×g. Pellets were resuspended and washed once. We used anti-CTCF polyclonal antibody (07-729) from Millipore (Millipore, Temecula, CA, USA) and anti-Rad21 polyclonal antibody (abb992) from Abcam (Abcam plc, Cambridge, UK), both in 1:1000 dilutions to detect CTCF and Rad21 protein levels. As a control antibody against β-Actin was used - anti-β-Actin (A1978) from Sigma-Aldrich (Sigma-Aldrich, Saint Louis, MO, USA) in 1:5000 dilution. Secondary antibodies: AP-conjugated Anti-Rabbit IgG (111-055-003) and AP-conjugated Anti-Mouse IgG (115-055-003) both from Jackson (Jackson ImmunoResearch Laboratories Inc, West Grove, PA, USA) were used in dilution 1:5000. Signals were quantified using ImageJ software (<http://rsb.info.nih.gov/ij/>).

### Measuring apoptotic cells

The fraction of apoptotic cells on slides was estimated by FITC Annexin V (BD Pharmingen, Becton, Dickinson and Company, Franklin Lakes, NJ, USA) staining. 100 cells of control, CTCF knockdown, Rad21 knockdown and CTCF-Rad21 knockdown populations were selected by DAPI signal, the fraction of Annexin V positive cells was measured in this population. The same procedure was repeated for three independent siRNA knockdown experiments. Results obtained by AnnexinV staining were confirmed by the Tunel apoptosis assay (Promega, Madison, WI, USA).

### Labelling of BAC probes and fluorescence in situ hybridization

BACs were selected from the BAC clones available in the RP11-collection at the Sanger Institute. Genomic distances were defined as the distance between centres of the BACs. BAC DNA was isolated using the Qiagen REAL prep 96 kit (Qiagen, Qiagen Benelux BV, Venlo, Netherlands). Nick-translation was used to label the probes, either with digoxigenin or biotin (Roche, F. Hoffmann-La Roche Ltd, Basle, Switzerland). FISH was carried out as described in [44]. FITC-conjugated antibodies (Roche, F. Hoffmann-La Roche Ltd, Basle, Switzerland) and Cy3-conjugated streptavidin (Jackson ImmunoResearch Laboratories, Inc., West Grove, PA) were used to visualize the hybridization signals. DAPI (4',6'-diamidino-2-phenylindole) (Sigma-Aldrich, Saint Louis, MO, USA) was used to outline the cell nucleus.

To exclude S-phase cells from the analysis the culture was incubated with bromodeoxyuridine (BrdU) to label replicating cells 30 min before fixation, followed by immunolabelling in combination with FISH labelling as described in [44].

### Imaging

Twelve-bit 3D images were recorded using Nikon A1R confocal laser-scanning microscope (Nikon, Tokyo, Japan) equipped with a 100×/1.49 NA Apo TIRF DIC objective (Nikon), using a diode laser at 405 nm, an Ar laser at 488 nm and a diode-pumped solid-state laser at 561 nm to excite DAPI, FITC and Cy3, respectively. Fluorescence was detected with the following bandpass filters: 425–475 nm (DAPI), 500–550 nm (FITC) and 570–620 nm (Cy3). Images were scanned with a voxel size of 60×60×100 nm.

### Image processing and data analysis

All confocal images were subject to deconvolution using Huygens Professional 3.7 software (Scientific Volume Imaging, Hilversum, The Netherlands) using the measured point spread function (PSF) for each channel and the classical maximum likelihood estimation algorithm. The PSF was obtained by imaging Tetraspeck Fluorescent Microsphere Standards with a diameter of 200 nm (Invitrogen). The signal to noise ratios and background intensities were estimated for each channel and averaged from several images. These values were used as a standard for batch processing deconvolution of all the image stacks.

Automated image analysis was carried out on deconvolved datasets with the ARGOS software (<http://homepages.cwi.nl/~wimc/argos>) to identify nuclear sites labelled by BACs and to compute their 3D position in the nucleus as described in [44]. Chromatic aberration was measured via Tetraspeck Fluorescent Microsphere Standards with a diameter of 200 nm (Invitrogen) and corrected for in the analysis. After background subtraction, images were treated with a bandpass filter to remove noise. Subsequently, images were segmented and ensembles of interconnected voxels were regarded as the site labelled by a BAC. The centre of mass was calculated for each labelled site at a sub-voxel resolution and 3D distances between BACs were measured. To estimate the systematic measuring error we hybridized cells with a mixture of the same BAC marked with two different fluorophores and measured the distances between the two signals. Accuracy of measurements was better than 50 nm in three dimensions.

The radial nuclear position  $p_n$  ( $p_n = r_o/r_n$ ) of BAC probe indicates positioning of the centre of gravity of a FISH labelled site on the line drawn from the centre of gravity of the nucleus to nuclear envelope.  $p_n$  was calculated as previously described [44]: the distance from the centre of gravity of a BAC probe and the centre of the nucleus ( $r_o$ ) was divided by the length of a line from the nuclear centre to the nuclear envelope through the centre of gravity of the BAC probe ( $r_n$ ).  $p_n$  value close to 1 indicates positioning of the BAC probe at nuclear periphery while  $p_n$  value close to 0 indicates its positioning close to the centre of gravity of the nucleus.

### Dynamic Loop polymer model

Chromatin fibres are in general modelled as un-branched polymer chains. The modelled polymers are coarse-grained versions of the real chromosomes, where each monomer represents a stretch of DNA. If the length of this stretch is much larger than the persistence length of the chromatin fibre, we can expect the polymer to be totally flexible and are allowed to neglect the influence of the local bending rigidity. With a persistence length below 250 nm [53] and a fibre packing, where a length of 10 nm = 1 kb, the monomers should represent DNA stretches with a length of at least 50 kb.

In this study, we performed Monte-Carlo simulations [54] with an implementation of the Dynamic Loop model to generate chromosomal conformations. For the polymer chains that represent the interphase chromosomes we used the well-tested bond fluctuation method [55,56] to investigate their structure, as well as their dynamics. In the simulations a monomer of the polymer chain is randomly selected and, if possible, randomly moved to one of its nearest neighbours on the lattice. Excluded volume interactions are taken into account by preventing a lattice site to be occupied by more than one monomer. When simulating  $N$  monomers we define one Monte-Carlo step (MCS) to correspond to  $N$  moves, i.e. on average each monomer is translated once during a MCS. Due to the fact that the polymer conformations only exhibit slight changes from one MCS to the

next, the time span when two conformations can be considered to be independent must be determined. Therefore, we calculate the autocorrelation function of the polymer's squared radius of gyration  $Rg^2(t)$  which is a measure for structural correlation. We obtain the estimated autocorrelation time  $\tau_{ac}$  by applying an exponential fit to the autocorrelation function. Finally, we set  $5\tau_{ac}$  MCS as the time span above which two conformations are expected to be independent.

### Chromatin looping

For the interactions between the chromatin fibres we used the Dynamic Loop (DL) model [25]. In this model two monomers that are non-adjacent along the fibre are only allowed to interact if they are in spatial proximity, i.e. if the distance between them is below a certain cutoff. When two monomers  $i$  and  $j$  approach each other due to diffusional motion and the cutoff condition is fulfilled, a bond can be established between them with a certain probability  $p_{bond,ij}$ : the binding or looping probability. In case the bond is formed, a lifetime  $t_{bond}$  chosen from a Poisson distribution with mean  $T_{bond}$  is assigned to it, determining when the bond dissociates. Hence, the bonds can frequently change during the course of the simulation, which mimics the effect of the highly dynamic DNA-DNA interactions. In contrast to this, the bonds along the backbone of the polymer are fixed and cannot break open.

### Polymer model simulations

For the simulations with the domain-adapted DL and with the adapted DL model, the chromosome-representing polymers have a length of 1050 monomers. We have to simulate thousands of independent chromosome conformations for each parameter set to get statistically reliable values for the different chromosome properties. Hence, the polymer length is limited to  $\sim 1000$  monomers in order to finish the simulations in a manageable timeframe. We set the length to exactly 1050 monomers because in the domain-adapted DL model, the highly active as well as the lowly active domains have a size of 50 monomers and the polymer consists of 21 domains (11 lowly active interspersed by 10 highly active regions). We also use 1050 monomers for the adapted DL model to ease comparison with the results from the domain-adapted DL model. The size of the Monte-Carlo lattice is set to a large value ( $L=500$ ) to avoid interactions caused by periodic boundary conditions. The mean of the Poisson distributions that is used to determine the bond lifetime is set to 8000 MCS. In the adapted DL model,  $p_{bond,ij}$  can take 2 different values, namely  $p_{short}$  if  $|i-j| \leq 50$  and  $p_{long}$  otherwise. In the domain-adapted DL model, the bond formation probabilities  $p_{bond,ij}$  between the monomers are set to  $p_{short}$  if monomers  $i$  and  $j$  are inside the same domain, or to  $p_{long}$  if  $i$  and  $j$  belong to different domains.

We start each simulation by creating a random self-avoiding walk (polymer where bonds are only established between adjacent monomers). We continue with a first equilibration run of  $10^8$  MCS where only homogeneous bond formation is allowed. As  $p_{bond}$  is set to be constant along the fibre, we obtain a homogenous starting conformation, meaning a polymer with a uniform structure. In a second equilibration run of  $10^8$  MCS we allow heterogeneous looping as defined by the domain-adapted DL and the adapted DL model. When all equilibration runs are finished, the main simulation starts and the chromosomal conformation is saved every  $10^7$  MCS. This is done until at least 1000 independent conformations are generated.

## Supporting Information

**Figure S1 Cellular and nuclear morphology of cells depleted in CTCF and/or cohesin.** Phase contrast images (A) and images of DAPI stained cells with visible FISH probes (B) showing 04-147 primary fibroblasts before and after siRNA mediated depletion of CTCF, Rad21 and simultaneous depletion of CTCF and Rad21. For (B) deconvolved 3D confocal images were flattened using ImageJ. DAPI signal is a standard deviation projection over all stacks, FISH signals are maximum intensity projections.  
(TIF)

**Figure S2 Nuclear size of cells depleted in CTCF and/or cohesin.** 3D imaged nuclei were automatically detected and their volume measured in ARGOS software (<http://homepages.cwi.nl/~wimc/argos>). Box plot shows volumes of at least 250 nuclei per condition. Outliers are indicated in red. There is no statistical difference between datasets obtained from control cells and cells depleted in CTCF, Rad21 or CTCF and Rad21 (t-test).  
(EPS)

**Figure S3 Fraction of apoptotic cells in experiments.** Fraction of apoptotic cells on slides was estimated by FITC Annexin V staining. 100 cells of control, CTCF knockdown, Rad21 knockdown and CTCF-Rad21 knockdown populations were selected by DAPI signal, the fraction of Annexin V positive cells was measured in this population. The same procedure was repeated for three independent siRNA knockdown experiments. Error bars represent standard deviation.  
(EPS)

**Figure S4 FISH measurements of cells transfected with non-target siRNA.** To test whether the siRNA transfection process itself affects chromatin compaction we carried out FISH measurements on cells transfected with 150 nM non-target siRNA (Negative Control Medium GC Duplex, Invitrogen). Transfections were carried out as in the case of CTCF Rad21 double depletion. Error bars represent standard error of mean.  
(EPS)

**Figure S5 Example of polymer conformations of the adapted DL model at different looping regimes.** (A) Conformation of a adapted DL polymer with high short-range ( $p_{short} = 0.12$ ) and low long-range ( $p_{long} = 0.04$ ) looping probabilities and (B) the same polymer with low short-range ( $p_{short} = 0.04$ ) and high long-range ( $p_{long} = 0.12$ ) looping probabilities. The colour code labels the monomers of the polymer according to the visible spectrum.  
(EPS)

**Figure S6 Genomic alignment of transcriptome maps, replication domains, topological domains data, and BAC probes.** Alignment of the transcriptome map (04-147 primary fibroblasts [42]), replication timing domains (IMR90 foetal lung fibroblasts [43]), coordinates of HiC topological domains (IMR90 [7]) and the positions of BAC probes used in this study (A) for regions analysed on Chromosome 1 and (B) for the region on Chromosome 11. On the human transcriptome map vertical lines represent protein coding genes, the length depicting the average transcription level in a number of human tissues and cell lines. Replication-timing domains are illustrated by peaks and valleys corresponding to early and late-replicating regions, respectively. Coordinates of topological domains identified using HiC are shown as blocks (HiC data processing described in reference [7]). For better visibility of domain borders alternating blocks have different height; regions devoid of blocks are linker areas between

domains. Triangles below the plot indicate positions of BAC probes. Reference probes for each dataset are indicated in red. The same probes are indicated above the graph of the transcriptome map by bars (as on figure 2). Gene-rich and gene-poor genomic regions analyzed are indicated by green and red bar, respectively. (EPS)

**Table S1 BACs used and their positions on human genome assembly 18 (Hg18).**

(DOCX)

**Table S2 siRNAs used.**

(DOCX)

## References

- Cavalli G, Misteli T (2013) Functional implications of genome topology. *Nature structural & molecular biology* 20: 290–299. doi:10.1038/nsmb.2474.
- Gibcus JH, Dekker J (2013) The Hierarchy of the 3D Genome. *Molecular Cell* 49: 773–782. doi:10.1016/j.molcel.2013.02.011.
- Bickmore WA (2013) The spatial organization of the human genome. *Annual review of genomics and human genetics* 14: 67–84. doi:10.1146/annurev-genom-091212-153515.
- Cremer T, Cremer M (2001) Chromosome territories, nuclear architecture and gene regulation in mammalian cells. *Nature reviews Genetics* 2: 292–301. doi:10.1038/35066075.
- Cremer T, Cremer M (2010) Chromosome territories. *Cold Spring Harbor perspectives in biology* 2: a003889. doi:10.1101/cshperspect.a003889.
- Pombo A, Branco MR (2007) Functional organisation of the genome during interphase. *Current opinion in genetics & development* 17: 451–455. doi:10.1016/j.gde.2007.08.008.
- Dixon JR, Selvaraj S, Yue F, Kim A, Li Y, et al. (2012) Topological domains in mammalian genomes identified by analysis of chromatin interactions. *Nature* 485: 376–380. doi:10.1038/nature11082.
- Sexton T, Yaffe E, Kenigsberg E, Bantignies F, Leblanc B, et al. (2012) Three-dimensional folding and functional organization principles of the *Drosophila* genome. *Cell* 148: 458–472. doi:10.1016/j.cell.2012.01.010.
- Nora EP, Lajoie BR, Schulz EG, Giorgetti L, Okamoto I, et al. (2012) Spatial partitioning of the regulatory landscape of the X-inactivation centre. *Nature* 485: 381–385. doi:10.1038/nature11049.
- Hou C, Li L, Qin ZS, Corces VG (2012) Gene density, transcription, and insulators contribute to the partition of the *Drosophila* genome into physical domains. *Molecular cell* 48: 471–484. doi:10.1016/j.molcel.2012.08.031.
- Phillips-Cremins JE, Sauria MEG, Sanyal A, Gerasimova TI, Lajoie BR, et al. (2013) Architectural Protein Subclasses Shape 3D Organization of Genomes during Lineage Commitment. *Cell* 153: 1281–1295. doi:10.1016/j.cell.2013.04.053.
- Moindrot B, Audit B, Klous P, Baker A, Thermes C, et al. (2012) 3D chromatin conformation correlates with replication timing and is conserved in resting cells. *Nucleic Acids Research* 40: 9470–9481. doi:10.1093/nar/gks736.
- Ryba T, Hiratani I, Lu J, Itoh M, Kulik M, et al. (2010) Evolutionarily conserved replication timing profiles predict long-range chromatin interactions and distinguish closely related cell types. *Genome research* 20: 761–770. doi:10.1101/gr.099655.109.
- Pope BD, Hiratani I, Gilbert DM (2010) Domain-wide regulation of DNA replication timing during mammalian development. *Chromosome research: an international journal on the molecular, supramolecular and evolutionary aspects of chromosome biology* 18: 127–136. doi:10.1007/s10577-009-9100-8.
- Pope BD, Gilbert DM (2013) The replication domain model: regulating replicon firing in the context of large-scale chromosome architecture. *Journal of molecular biology* 425: 4690–4695. doi:10.1016/j.jmb.2013.04.014.
- Mateos-Langerak J, Bohn M, De Leeuw W, Giromus O, Manders EMM, et al. (2009) Spatially confined folding of chromatin in the interphase nucleus. *Proceedings of the National Academy of Sciences of the United States of America* 106: 3812–3817. doi:10.1073/pnas.0809501106.
- Nagano T, Lubling Y, Stevens TJ, Schoenfelder S, Yaffe E, et al. (2013) Single-cell Hi-C reveals cell-to-cell variability in chromosome structure. *Nature* 502: 59–64. doi:10.1038/nature12593.
- Tark-Dame M, Van Driel R, Heermann DW (2011) Chromatin folding—from biology to polymer models and back. *Journal of cell science* 124: 839–845. doi:10.1242/jcs.077628.
- Fudenberg G, Mirny LA (2012) Higher-order chromatin structure: bridging physics and biology. *Current opinion in genetics & development* 22: 115–124. doi:10.1016/j.gde.2012.01.006.
- Huet S, Lavelle C, Ranchon H, Carrivain P, Victor J-M, et al. (2014) Relevance and limitations of crowding, fractal, and polymer models to describe nuclear architecture. *International review of cell and molecular biology* 307: 443–479. doi:10.1016/B978-0-12-800046-5.00013-8.
- Rosa A, Zimmer C (2014) Computational models of large-scale genome architecture. *International review of cell and molecular biology* 307: 275–349. doi:10.1016/B978-0-12-800046-5.00009-6.
- Lieberman-Aiden E, Van Berkum NL, Williams L, Imakaev M, Ragozcy T, et al. (2009) Comprehensive mapping of long-range interactions reveals folding principles of the human genome. *Science (New York, NY)* 326: 289–293. doi:10.1126/science.1181369.
- Mirny LA (2011) The fractal globule as a model of chromatin architecture in the cell. *Chromosome research: an international journal on the molecular, supramolecular and evolutionary aspects of chromosome biology* 19: 37–51. doi:10.1007/s10577-010-9177-0.
- Dekker J, Marti-Renom MA, Mirny LA (2013) Exploring the three-dimensional organization of genomes: interpreting chromatin interaction data. *Nature reviews Genetics* 14: 390–403. doi:10.1038/nrg3454.
- Bohn M, Heermann DW (2010) Diffusion-driven looping provides a consistent framework for chromatin organization. *PLoS one* 5: e12218. doi:10.1371/journal.pone.0012218.
- Heermann DW, Jerabek H, Liu L, Li Y (2012) A model for the 3D chromatin architecture of pro and eukaryotes. *Methods (San Diego, Calif)* 58: 307–314. doi:10.1016/j.ymeth.2012.04.010.
- Bohn M, Heermann DW (2011) Repulsive forces between looping chromosomes induce entropy-driven segregation. *PLoS one* 6: e14428. doi:10.1371/journal.pone.0014428.
- Heermann DW (2011) Physical nuclear organization: loops and entropy. *Current opinion in cell biology* 23: 332–337. doi:10.1016/j.ceb.2011.03.010.
- Dion V, Gasser SM (2013) Chromatin movement in the maintenance of genome stability. *Cell* 152: 1355–1364. doi:10.1016/j.cell.2013.02.010.
- Jerabek H, Heermann DW (2012) Expression-dependent folding of interphase chromatin. *PLoS one* 7: e37525. doi:10.1371/journal.pone.0037525.
- Holwerda SJB, De Laat W (2013) CTCF: the protein, the binding partners, the binding sites and their chromatin loops. *Philosophical transactions of the Royal Society of London Series B, Biological sciences* 368: 20120369. doi:10.1098/rstb.2012.0369.
- Dorsett D, Merkenschlager M (2013) Cohesin at active genes: a unifying theme for cohesin and gene expression from model organisms to humans. *Current opinion in cell biology* 25: 327–333. doi:10.1016/j.ceb.2013.02.003.
- Merkenschlager M, Odom DT (2013) CTCF and cohesin: linking gene regulatory elements with their targets. *Cell* 152: 1285–1297. doi:10.1016/j.cell.2013.02.029.
- Wendt KS, Yoshida K, Itoh T, Bando M, Koch B, et al. (2008) Cohesin mediates transcriptional insulation by CCCTC-binding factor. *Nature* 451: 796–801. doi:10.1038/nature06634.
- Parelho F, Hadjir S, Spivakov M, Leleu M, Sauer S, et al. (2008) Cohesins functionally associate with CTCF on mammalian chromosome arms. *Cell* 132: 422–433. doi:10.1016/j.cell.2008.01.011.
- Zuin J, Dixon JR, Van der Reijden MJA, Ye Z, Kolovos P, et al. (2014) Cohesin and CTCF Differentially Affect Chromatin Architecture and Gene Expression in Human Cells. *PNAS* 111: 996–1001.
- Sofueva S, Yaffe E, Chan W-C, Georgopoulou D, Vietri Rudan M, et al. (2013) Cohesin-mediated interactions organize chromosomal domain architecture. *The EMBO journal* 32: 3119–3129. doi:10.1038/emboj.2013.237.
- Seitan VC, Faure AJ, Zhan Y, McCord RP, Lajoie BR, et al. (2013) Cohesin-based chromatin interactions enable regulated gene expression within preexisting architectural compartments. *Genome research* 23: 2066–2077. doi:10.1101/gr.161620.113.
- Nativio R, Wendt KS, Ito Y, Huddleston JE, Uribe-Lewis S, et al. (2009) Cohesin is required for higher-order chromatin conformation at the imprinted IGF2-H19 locus. *PLoS genetics* 5: e1000739. doi:10.1371/journal.pgen.1000739.
- Hou C, Dale R, Dean A (2010) Cell type specificity of chromatin organization mediated by CTCF and cohesin. *Proceedings of the National Academy of Sciences of the United States of America* 107: 3651–3656. doi:10.1073/pnas.0912087107.
- Schmidt D, Schwalie PC, Ross-Innes CS, Hurtado A, Brown GD, et al. (2010) A CTCF-independent role for cohesin in tissue-specific transcription. *Genome research* 20: 578–588. doi:10.1101/gr.100479.109.
- Caron H, Van Schaik B, Van der Mee M, Baas F, Riggins G, et al. (2001) The human transcriptome map: clustering of highly expressed genes in chromosomal

## Acknowledgments

The authors thank Drs. Paul Fransz, Paul Verbruggen, Remus Dame and Till Bey for their comments on the manuscript.

## Author Contributions

Conceived and designed the experiments: RvD DWH MTD HJ. Performed the experiments: MTD HJ. Analyzed the data: MTD HJ. Contributed reagents/materials/analysis tools: EMMM. Wrote the paper: MTD RvD HJ DWH.

- domains. *Science* (New York, NY) 291: 1289–1292. doi:10.1126/science.1056794.
43. Pope BD, Chandra T, Buckley Q, Hoare M, Ryba T, et al. (2012) Replication-timing boundaries facilitate cell-type and species-specific regulation of a rearranged human chromosome in mouse. *Human molecular genetics*. doi:10.1093/hmg/dds232.
  44. Goetze S, Mateos-Langerak J, Gierman HJ, De Leeuw W, Giromus O, et al. (2007) The three-dimensional structure of human interphase chromosomes is related to the transcriptome map. *Molecular and cellular biology* 27: 4475–4487. doi:10.1128/MCB.00208-07.
  45. Rhee HS, Pugh BF (2011) Comprehensive Genome-wide Protein-DNA Interactions Detected at Single-Nucleotide Resolution. *Cell* 147: 1408–1419. doi:10.1016/j.cell.2011.11.013.
  46. Takizawa T, Gudla PR, Guo L, Lockett S, Misteli T (2008) Allele-specific nuclear positioning of the monoallelically expressed astrocyte marker GFAP. *Genes & development* 22: 489–498. doi:10.1101/gad.1634608.
  47. Ferrai C, De Castro IJ, Lavitas L, Chotalia M, Pombo A (2010) Gene positioning. *Cold Spring Harbor perspectives in biology* 2: a000588. doi:10.1101/cshperspect.a000588.
  48. Handoko L, Xu H, Li G, Ngan CY, Chew E, et al. (2011) CTCF-mediated functional chromatin interactome in pluripotent cells. *Nature genetics* 43: 630–638. doi:10.1038/ng.857.
  49. Barbieri M, Chotalia M, Fraser J, Lavitas L-M, Dostie J, et al. (2012) Complexity of chromatin folding is captured by the strings and binders switch model. *Proceedings of the National Academy of Sciences of the United States of America* 109: 16173–16178. doi:10.1073/pnas.1204799109.
  50. Bancaud A, Lavelle C, Huet S, Ellenberg J (2012) A fractal model for nuclear organization: current evidence and biological implications. *Nucleic acids research* 40: 8783–8792. doi:10.1093/nar/gks586.
  51. Yokota H, Van den Engh G, Hearst JE, Sachs RK, Trask BJ (1995) Evidence for the organization of chromatin in megabase pair-sized loops arranged along a random walk path in the human G0/G1 interphase nucleus. *The Journal of cell biology* 130: 1239–1249.
  52. Naumova N, Imakaev M, Fudenberg G, Zhan Y, Lajoie BR, et al. (2013) Organization of the mitotic chromosome. *Science* (New York, NY) 342: 948–953. doi:10.1126/science.1236083.
  53. Langowski J (2006) Polymer chain models of DNA and chromatin. *The European physical journal E, Soft matter* 19: 241–249. doi:10.1140/epje/i2005-10067-9.
  54. Heermann DW, Binder K (2010) *Monte Carlo Simulations in Statistical Physics*. fifth. Heidelberg: Springer. doi:10.1007/978-3-642-03163-2.
  55. Carmesin I, Kremer K (1988) The Bond Fluctuation Method: A New Effective Algorithm for the Dynamics of Polymers in All Spatial Dimensions. *Macromolecules* 21: 2819–2823.
  56. Deutsch HP, Binder K (1991) Interdiffusion and self-diffusion in polymer mixtures: A Monte Carlo study. *J Chem Phys* 94: 2294–2304.