

Research

Active chromatin and transcription play a key role in chromosome partitioning into topologically associating domains

Sergey V. Uljanov,^{1,2,12} Ekaterina E. Khrameeva,^{3,4,12} Alexey A. Gavrillov,¹ Ilya M. Flyamer,^{1,2} Pavel Kos,⁵ Elena A. Mikhaleva,⁶ Aleksey A. Penin,^{4,7,8} Maria D. Logacheva,^{4,8,9} Maxim V. Imakaev,¹⁰ Alexander Chertovich,⁵ Mikhail S. Gelfand,^{4,11} Yuri Y. Shevelyov,⁶ and Sergey V. Razin^{1,2}

¹Institute of Gene Biology, RAS, 119334 Moscow, Russia; ²Department of Molecular Biology, Lomonosov Moscow State University, 119991 Moscow, Russia; ³Skolkovo Institute of Science and Technology, 143026 Skolkovo, Russia; ⁴Institute for Information Transmission Problems (Kharkevich Institute), RAS, 127051 Moscow, Russia; ⁵Physics Department, Lomonosov Moscow State University, 119991 Moscow, Russia; ⁶Department of Molecular Genetics of Cell, Institute of Molecular Genetics, RAS, 123182 Moscow, Russia; ⁷Department of Genetics, Faculty of Biology, Lomonosov Moscow State University, 119991 Moscow, Russia; ⁸A.N. Belozersky Institute of Physico-Chemical Biology, Lomonosov Moscow State University, 119234 Moscow, Russia; ⁹Pirogov Russian National Research Medical University, 117997 Moscow, Russia; ¹⁰Department of Physics, Massachusetts Institute of Technology, Cambridge, Massachusetts 02139, USA; ¹¹Faculty of Bioengineering and Bioinformatics, Lomonosov Moscow State University, 119234 Moscow, Russia

Recent advances enabled by the Hi-C technique have unraveled many principles of chromosomal folding that were subsequently linked to disease and gene regulation. In particular, Hi-C revealed that chromosomes of animals are organized into topologically associating domains (TADs), evolutionary conserved compact chromatin domains that influence gene expression. Mechanisms that underlie partitioning of the genome into TADs remain poorly understood. To explore principles of TAD folding in *Drosophila melanogaster*, we performed Hi-C and poly(A)⁺ RNA-seq in four cell lines of various origins (S2, Kcl67, DmBG3-c2, and OSC). Contrary to previous studies, we find that regions between TADs (i.e., the inter-TADs and TAD boundaries) in *Drosophila* are only weakly enriched with the insulator protein dCTCF, while another insulator protein Su(Hw) is preferentially present within TADs. However, *Drosophila* inter-TADs harbor active chromatin and constitutively transcribed (housekeeping) genes. Accordingly, we find that binding of insulator proteins dCTCF and Su(Hw) predicts TAD boundaries much worse than active chromatin marks do. Interestingly, inter-TADs correspond to decompacted inter-bands of polytene chromosomes, whereas TADs mostly correspond to densely packed bands. Collectively, our results suggest that TADs are condensed chromatin domains depleted in active chromatin marks, separated by regions of active chromatin. We propose the mechanism of TAD self-assembly based on the ability of nucleosomes from inactive chromatin to aggregate, and lack of this ability in acetylated nucleosomal arrays. Finally, we test this hypothesis by polymer simulations and find that TAD partitioning may be explained by different modes of inter-nucleosomal interactions for active and inactive chromatin.

[Supplemental material is available for this article.]

Recently developed 3C-based methods coupled with high-throughput sequencing have enabled genome-wide investigation of chromatin organization. Studies performed in human (Lieberman-Aiden et al. 2009; Dixon et al. 2012; Rao et al. 2014), mouse (Dixon et al. 2012; Nora et al. 2012), *Drosophila* (Hou et al. 2012; Sexton et al. 2012), yeasts (Duan et al. 2011; Mizuguchi et al. 2014), *Arabidopsis* (Feng et al. 2014; Grob et al. 2014), and several other species (Le et al. 2013; Lemieux et al. 2013; Vietri Rudan et al. 2015) have unraveled general principles of genome folding. Chromosomes in mammals and *Drosophila* are organized hierarchically. At the megabase scale, mammalian chromosomes are

partitioned into active and inactive compartments. At the sub-megabase scale, these compartments are subdivided into a set of self-interacting domains called topologically associating domains (TADs); TADs themselves are often hierarchical and are split into smaller domains. Similar to mammals, *Drosophila* chromosomes are partitioned into TADs that are interspaced with short boundaries or longer inter-TAD regions (inter-TADs).

Partitioning of mammalian genomes into TADs appears to be largely cell lineage-independent and evolutionarily conserved (Dixon et al. 2012; Vietri Rudan et al. 2015). Disruption of

¹²These authors contributed equally to this work.

Corresponding author: sergey.v.razin@usa.net

Article published online before print. Article, supplemental material, and publication date are at <http://www.genome.org/cgi/doi/10.1101/gr.196006.115>.

© 2016 Uljanov et al. This article is distributed exclusively by Cold Spring Harbor Laboratory Press for the first six months after the full-issue publication date (see <http://genome.cshlp.org/site/misc/terms.xhtml>). After six months, it is available under a Creative Commons License (Attribution-NonCommercial 4.0 International), as described at <http://creativecommons.org/licenses/by-nc/4.0/>.

certain TAD boundaries leads to developmental defects in humans and mice (Lupiáñez et al. 2015). TADs correlate with units of replication timing regulation in mammals (Pope et al. 2014) and colocalize with epigenetic domains (either active or repressed) in *Drosophila* (Sexton et al. 2012). The internal structure of TADs was reported to change in response to environmental stress (Li et al. 2015), during cell differentiation (Williamson et al. 2014; Dixon et al. 2015), and embryonic development (Berlivet et al. 2013). In addition, comparative Hi-C analysis has demonstrated that genomic rearrangements between related mammalian species occur predominantly at TAD boundaries (Vietri Rudan et al. 2015). Consequently, TADs appear to evolve primarily as constant and unsplit units.

Previous studies in *Drosophila* embryonic nuclei and embryo-derived Kc167 cells (Hou et al. 2012; Sexton et al. 2012) detected TADs of various sizes roughly corresponding to epigenetic domains. Additionally, long-range genomic contacts and clustering of pericentromeric regions were revealed, and TAD boundaries were found to be enriched with active chromatin marks and insulator proteins. Both active and inactive TADs were identified, and their spatial segregation was observed.

Despite extensive studies, mechanisms underlying TAD formation remain obscure. Architectural proteins, including cohesin and CTCF, are often found at TAD boundaries; thus, they have been proposed to play a key role in the demarcation of TADs (Hou et al. 2012; Van Bortle et al. 2014). However, several studies suggest that other mechanisms may be responsible for partitioning and formation of TADs. Firstly, depletion of various insulator proteins did not affect the profile of chromosome partitioning into TADs but rather decreased intra-TAD interactions (Sofueva et al. 2013; Li et al. 2015). Secondly, CTCF may mediate loops that occur between the start and the end of the so-called “loop domains” (Rao et al. 2014). However, domains of similar sizes but without a loop were observed as well (so-called “ordinary domains”) (Rao et al. 2014). Thirdly, polymer simulations of a permanent chromatin loop yield a noticeable interaction between the loop bases on a simulated Hi-C map but without the characteristic square shape of a TAD (Doyle et al. 2014). Loops of this kind are thought to occur between insulator proteins such as Su(Hw) in the “topological insulation” model (Gohl et al. 2011). Finally, chromosomal domains similar to TADs in the bacterium *Caulobacter crescentus* are demarcated by actively transcribed genes and are not affected by the knockout of SMC, a homolog of cohesin subunits (Le et al. 2013).

Here, we present evidence that questions the role of insulators in the organization of TAD boundaries in *Drosophila*. Our results suggest that TADs are self-organized and potentially highly dynamic structures formed by numerous transient interactions between nucleosomes of inactive chromatin, while inter-TADs and TAD boundaries contain highly acetylated nucleosomes that are less prone to interactions. Finally, we develop a polymer model of TAD formation based on the two types of nucleosomes and find that a polymer composed of active and inactive chromatin blocks forms TADs on a simulated Hi-C map.

Results

Whole-genome maps of chromatin folding in a set of cultured *Drosophila* cell lines

We performed Hi-C in four cultured *Drosophila melanogaster* cell lines representing different lineages and development stages;

Schneider-2 (S2) and Kc167 cells were derived from late embryos (Echalier and Ohanessian 1969; Schneider 1972); DmBG3-c2 (BG3) cells were derived from the central nervous system of third-instar larvae (Ui et al. 1994); and ovarian somatic cells (OSCs) were derived from adult fly ovaries (Niki et al. 2006). For each cell line, Hi-C was performed in two biological replicates using the HindIII restriction enzyme, according to a previously published protocol (Belton et al. 2012) with minor modifications (see Methods). Each Hi-C library was sequenced to about 40 million paired-end reads per replicate using Illumina HiSeq 2000. As two biological replicates demonstrated high correlation (Supplemental Table S1; Supplemental Fig. S1A), we combined them and obtained 9–21 million sequenced ligation junctions per cell line after all filtration steps (Supplemental Table S2). Resulting Hi-C data were binned at 20-kb resolution for each *Drosophila* cell line. Our Hi-C interaction maps are in good agreement with previously published data (Hou et al. 2012). In particular, we observe high frequency of interactions of pericentromeric regions with each other and with the heterochromatin-rich Chromosome 4 in all cell lines and do not find significant interactions of telomeres (Supplemental Fig. S2).

One of the striking features of the *Drosophila* Hi-C maps are bright spots of intensive Hi-C signals located far off the diagonal (Supplemental Fig. S2). These spots could be interpreted as long-range contacts of genome fragments (Sexton et al. 2012) separated by hundreds of kilobases or even megabases. Alternatively, these spots could reflect differences between the genomes of the cultured cell lines and the *Drosophila* reference genome (BDGP R5/dm3) due to genomic rearrangements, as observed previously in the human HeLa cells (Naumova et al. 2013). To determine the origin of these spots, we sequenced genomes of the four studied cell lines and annotated genomic rearrangements relative to the reference genome (see Methods). For the S2 and Kc167 cells, all spots (except one in S2) originated from extended inversions, deletions, and duplications. In contrast, only four out of 15 such spots in the BG3 cells and three out of 16 spots in OSCs corresponded to genomic rearrangements (Supplemental Table S3). Other spots likely represent true long-distance contacts. Interestingly, locations of these contacts were the same in the BG3 and OSC lines and were in agreement with previously published data in *Drosophila* embryos (Sexton et al. 2012). Additionally, in contrast to previous observations in *Drosophila* (Tolhuis et al. 2011), we found long-range inter-arm and even inter-chromosomal interactions between Polycomb-occupied regions in the BG3 cells (Supplemental Fig. S3) that had been previously observed only between domains located in the same chromosomal arm. Although detailed analysis of long-distance contacts is beyond the scope of this study, we note that annotation of such contacts should take into account the possibility of genomic rearrangements.

We next annotated TADs in the four studied cell lines using the Armatus software (Filippova et al. 2014) in which the average size and the number of TADs are determined by the scaling parameter γ . We selected γ individually for each cell line (from 1.03 to 1.26) to obtain a similar number and distribution of TAD lengths in all cell lines (Supplemental Table S5; Supplemental Fig. S4). Extremely large TADs (600 kb–3 Mb) were further divided into smaller domains using higher values of γ (Fig. 1A; see Methods for details). For further analysis, a TAD was defined as a genomic region containing two or more internal bins located between boundary bins. In each cell line, we annotated ~580 TADs ranging from 80 to 640 kb and covering 83% of the analyzed genome (Supplemental Fig. S4; Supplemental Tables S4, S5).

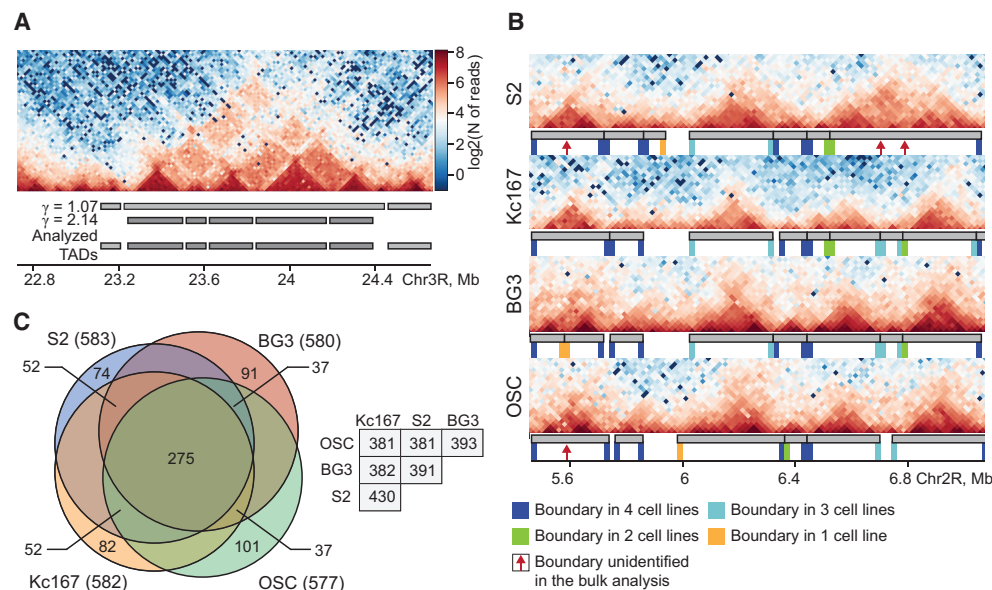


Figure 1. Genomic positions of topologically associating domains (TADs) are largely conserved among *Drosophila* cells of different origins. (A) A fragment of the Hi-C interaction map (heat map) for the BG3 cell line. Gray rectangles below the heat map represent TADs predicted with two different values of the scaling parameter γ . This illustrates a two-step TAD prediction method utilizing a higher γ value for large TADs containing internal self-interacting domains. (B) Heat maps of a 1.6-Mb region of Chromosome 2R with annotated TADs. Positions of boundary bins are indicated by colored rectangles under the TADs map. The color code for the boundary bins is shown at the bottom. Red arrows indicate unidentified weak boundaries. The resolution of all heat maps is 20 kb. (C) Venn diagram showing the numbers of TADs shared between all studied cell lines (with both boundaries located at the same genomic bin or at adjacent bins). Total numbers of TADs in each cell line are shown in parentheses. Numbers of TADs shared between different pairs of the cell lines are shown to the right of the diagram.

The comparison of TAD profiles in the four studied cell lines demonstrated that 47% of TADs are located at the same position (± 1 bin) in all cell lines, $\sim 67\%$ of TADs are shared by any two lines, and only $\sim 15\%$ of TADs occur in only one cell line (Fig. 1B, C). These estimates are in agreement with Hou et al. (2012), who have reported that 42% of TADs are shared between Kc167 cells and 16-h embryos.

Active chromatin and transcription but not architectural proteins dCTCF and Su(Hw) are typical for TAD boundaries and inter-TAD regions

Previous studies revealed that TAD boundaries in 16-h *Drosophila* embryos (Sexton et al. 2012) and in embryonic-derived Kc167 cells (Hou et al. 2012) are enriched with active chromatin marks and insulator proteins binding sites. Our analysis confirmed the preferential location of active chromatin marks at inter-TADs and TAD boundaries in *Drosophila* cells of different origin and developmental stages. For example, both RNA polymerase II and H3K27ac (active marks) were enriched within inter-TADs, whereas histone H1, which is likely involved in chromatin compaction (Hizume et al. 2005; Thomas and Stott 2012), was overrepresented in TADs (Fig. 2A). However, the segregation of active and repressive marks was not perfect, and peaks of RNA polymerase II and/or H3K27ac were found within TADs (Fig. 2A).

Analysis of the distribution of active and repressed chromatin with respect to TAD positions showed that active chromatin types represented by colors (Filion et al. 2010; Kharchenko et al. 2011) are strongly enriched at TAD boundaries and within inter-TADs in all examined cell lines (Fig. 2B; Supplemental Fig. S5A). Remarkably, in the Kc167 cells, inter-TADs and TAD boundaries were especially enriched in YELLOW chromatin (Fig. 2B), typical

for genomic regions harboring housekeeping genes (Filion et al. 2010). The proportion of certain types of active chromatin (RED in BG3 and S2, YELLOW in Kc167) changed dramatically from inter-TADs to TADs.

Next, we analyzed the distributions of several individual chromatin marks near TAD boundaries (Fig. 2C; Supplemental Fig. S6). Active chromatin marks (H3K27ac, H3K4me1, H3K4me3, H3K36me3, H4K16ac) were preferably present at TAD boundaries and within inter-TADs; the same was true for RNA polymerase II, the ISWI subunit of the chromatin remodeling complex, and the set of proteins strongly enriched at polytene chromosome interbands (Supplemental Fig. S6; Zhimulev et al. 2014). In contrast, repressive histone modification H3K27me3, linker histone H1, and core histone H3 (which reflects nucleosome occupancy) were depleted within inter-TADs and enriched within TADs in all cell lines.

Surprisingly, in contrast to active chromatin marks, the insulator (architectural) protein dCTCF and the SMC3 subunit of cohesin were only slightly enriched at TAD boundaries and within inter-TADs (Fig. 2C). Another insulator protein Su(Hw) was preferentially present within TADs. Among the insulator proteins analyzed, only BEAF-32 and CP190 demonstrated preferential binding within inter-TADs, with Z-values comparable to those of active chromatin marks. Notably, both proteins have been reported to be involved in either chromatin opening or transcription initiation (Bartkuhn et al. 2009; Jiang et al. 2009; Ahanger et al. 2014). To ensure that we did not miss any TAD boundaries enriched with dCTCF or Su(Hw), we repeated the TAD calling procedure for values of γ ranging from 0 to 3.2 with a step of 0.2. We did not observe any pronounced enrichment of dCTCF and Su(Hw) at TAD boundaries or within inter-TADs for any values of γ (Supplemental Fig. S7). Notably, a mark of active promoters H3K4me3 and other active

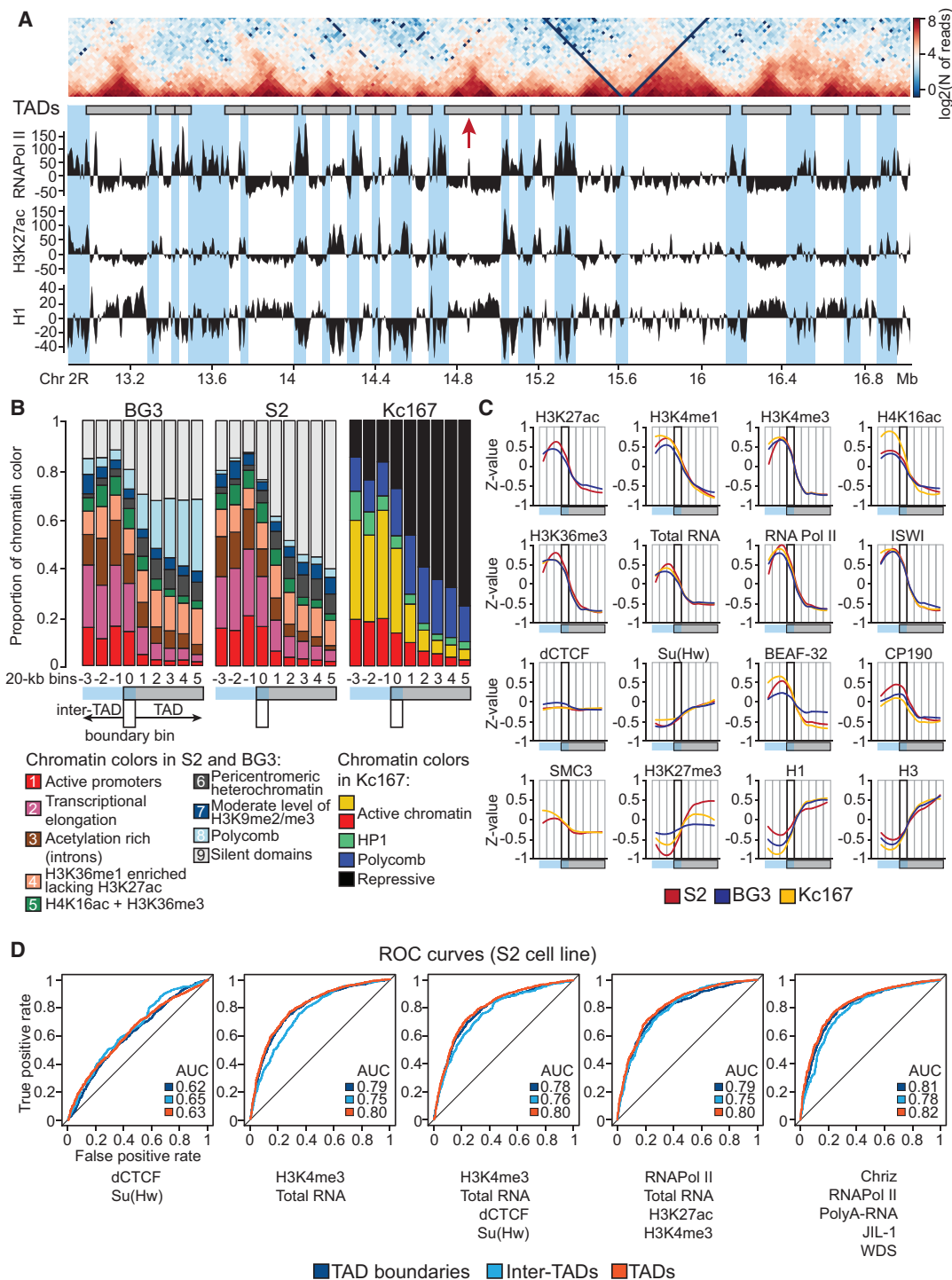


Figure 2. Partitioning of chromosomes into TADs and inter-TADs reflects distributions of active and repressed chromatin regions. (A) Heat map of a 4-Mb region of Chromosome 2R aligned with modENCODE tracks of RNA polymerase II, H3K27ac, and histone H1 for BG3 cells. TAD boundaries and inter-TAD regions are shaded in blue. Red arrow indicates an unidentified weak TAD boundary. (B) Distribution of chromatin types (colors) near TAD boundaries. Box plots show proportions of chromatin colors in bins located at the same position relative to a TAD boundary, averaged over all TADs. TAD boundaries were aligned so that inter-TAD bins were placed to the left of the boundary bin (bin 0) and TAD bins were placed to the right of the boundary bin. The region from the third bin in the inter-TAD (left) to the fifth bin in the TAD (right) is shown. Plots with longer distances from the boundary are shown in Supplemental Figure S5A. Chromatin colors are taken from Filion et al. (2010) for the Kc167 cells and from Kharchenko et al. (2011) for the S2 and BG3 cells. The OSC cells were not analyzed, as epigenetic data were not available for this cell line. The P-values are presented in Supplemental Table S6. (C) Distribution of the individual chromatin marks and proteins near TAD boundaries. Curves smoothed with LOESS show the median Z-transformed values in the groups of bins described above. Thick rectangles show TAD boundary bins. Box plots are shown in Supplemental Figure S6. P-values are presented in Supplemental Table S6. (D) Prediction of TADs, inter-TADs, and TAD boundaries in the S2 cells using logistic regression models based on the genome distribution of active chromatin marks or architectural proteins dCTCF and Su(Hw). Receiver operating characteristics (ROC curves) and AUC (area under the curve) values are shown.

chromatin marks (data not shown) were strongly enriched within inter-TADs and TAD boundaries at all values of γ (Supplemental Fig. S7).

To precisely relate gene expression to TAD profiles, we performed RNA-seq analysis of poly(A)⁺ transcriptomes in the four studied cell lines (Supplemental Table S7; Supplemental Fig. S8). As expected, genes in inter-TADs and at TAD boundaries were transcribed at significantly higher levels than genes inside TADs. Moreover, internal TAD regions contained a large number of nontranscribed bins (Fig. 3A,B; Supplemental Fig. S8C).

We next attempted to predict inter-TADs and TAD boundaries using logistic regression, based on the combination of several active chromatin marks as well as dCTCF and Su(Hw) (Fig. 2D). Regression based on distributions of H3K4me3 and total RNA predicted inter-TADs and TAD boundaries much better than regression based on distributions of dCTCF and Su(Hw) (the AUC for TAD boundaries in these two models are 0.79 and 0.63, respectively; here and below, 50/50 cross-validation was used) (Fig. 2D). dCTCF and Su(Hw) used together with active chromatin marks did not improve predictions compared to active marks alone (the AUC for active marks + dCTCF + Su(Hw) is 0.79, same as for active marks alone) (Fig. 2D). As profiles of most chromatin marks were highly correlated, incorporating additional features associated with active chromatin did not considerably improve the prediction quality (Fig. 2D). Taken together, these results suggest that histone modifications typical for active chromatin and/or active transcription, but not binding of insulator proteins dCTCF and Su(Hw), are the main characteristic feature of TAD boundaries and inter-TAD regions in *Drosophila* cells of different origin and developmental stages.

Active chromatin marks and transcription interfere with chromatin fiber packaging into TADs

Visual inspection of heat maps revealed that TADs in *Drosophila* cultured cells are often hierarchical (Fig. 1A), similar to TADs in *Drosophila* embryos and in mammals (Sexton et al. 2012; Filippova et al. 2014). The Armatus software is capable of capturing the hierarchical structure of topological domains (Filippova et al. 2014). It identifies strong boundaries of large TADs at low values of the scaling parameter γ , and weaker boundaries (Figs. 1C, 2A, red arrow) of smaller TADs (inside large TADs) at high values (the higher the γ , the more TAD boundaries can be identified). We supposed that if active chromatin state and transcription determine the chromosome partitioning into TADs, then the propensity of a genomic bin to be identified as a boundary should depend on the proportion of active chromatin and transcription level within this bin. To test this hypothesis, we assigned each genomic bin a value of the scaling parameter γ at which this bin switches from being a part of a TAD to being a part of an inter-TAD or a TAD boundary (γ_t , or transitional γ). We observed an inverse dependence between the γ_t values and the proportion of active chromatin (Fig. 3C; Supplemental Fig. S5B), as well as between the γ_t values and the transcription level (Fig. 3D; Supplemental Fig. S8D) in the bins. At the same time, we did not observe a strong inverse dependence between the γ_t value and the occupancy of a bin by a set of architectural proteins which included dCTCF, Su(Hw), BEAF-32, and CP190 (Fig. 3E). Conversely, a strong inverse dependence was observed between the occupancy of a set of active chromatin marks (RNA polymerase II, H3K4me3, ISWI, MRG15) and the γ_t value. Both findings support our hypothesis that a high proportion of active chromatin and a high transcription level, but not

binding of architectural proteins, interfere with packaging of the chromatin fiber into TADs, thus determining genomic positions of TAD boundaries and inter-TAD regions.

Although TADs in *Drosophila* typically contain inactive chromatin, there are some exceptions (Fig. 3B; Supplemental Fig. S9). Approximately 17% of TADs contain a significant proportion of active chromatin (>30%) and harbor highly transcribed genes; it has been reported that such active TADs are generally smaller than inactive ones (Hou et al. 2012). As active chromatin state and transcription were found to interfere with chromatin packaging into TADs (see above), we evaluated whether active and inactive TADs have different chromatin density, which we estimate from the sum of contacts within the TAD (Supplemental Fig. S15A). Since larger TADs cover a larger area of the Hi-C map, we divided TADs into groups according to their length (e.g., 80, 100, 120 kb) and analyzed each group separately. The TAD density, as defined above, was negatively correlated with the level of transcription (Pearson $r = -0.34$, BG3 cells) and the proportion of RED (specific for active promoters, Pearson $r = -0.54$ in BG3 cells) and H4K16ac-rich GREEN chromatin in the BG3 and S2 cells (Pearson $r = -0.43$ in BG3 cells), as well as with YELLOW chromatin, corresponding to housekeeping genes, in Kc167 (Pearson $r = -0.31$) (Fig. 3F–H). Conversely, the repressive types of chromatin exhibited weakly positive or zero correlation with the TAD density. Collectively, these data show that TADs harboring active chromatin and transcribed genes are less compact than fully inactive TADs.

Changes of transcription level and chromatin state frequently coincide with changes of local chromatin folding

We used recently published expression data from *Drosophila* tissues and development stages (Graveley et al. 2011; Brown et al. 2014) to compile a list of housekeeping and tissue-specific genes (see Methods). We found that the coverage of housekeeping genes is much higher in inter-TADs and TAD boundaries (Fig. 4A), whereas the coverage of tissue-specific genes is higher in TADs. Additionally, we found that both housekeeping and tissue-specific enhancers, previously identified in S2 and OSC cells using STARR-seq (Zabidi et al. 2015), are enriched within inter-TADs (Fig. 4A).

To track how changes in the level of transcription correspond to changes in the chromatin folding, we performed a pairwise comparison of transcriptomes of the studied cell lines and a pairwise comparison of the TADs profiles. All transcribed bins (with RPKM > 1) (Supplemental Table S8) were divided into four groups: located within a TAD in both cell lines (group T-T), in the first or in the second cell line only (T-I and I-T), and in neither of the two (I-I). The schematic color representation of the four groups is presented in Figure 4B. Additionally, for each pair of cell lines, we defined a bin to be differentially transcribed if it had at least a four-fold difference in expression between the cell lines. As tissue-specific genes are located mainly within TADs, and housekeeping genes are located in inter-TADs or TAD boundaries (see above), the majority of differentially transcribed bins belonged to the group T-T (Fig. 4B; Supplemental Fig. S10A). Most of the bins transcribed at a similar level in both cell lines belonged to the group I-I. Notably, bins in the group I-T, where an inter-TAD changed to a TAD, frequently demonstrated a decrease in transcription level, whereas the bins in the group T-I frequently showed an increase in transcription (Fig. 4B; Supplemental Fig. S10A, yellow and orange sectors). Similar but less pronounced effects were observed when the proportion of active chromatin within a bin was

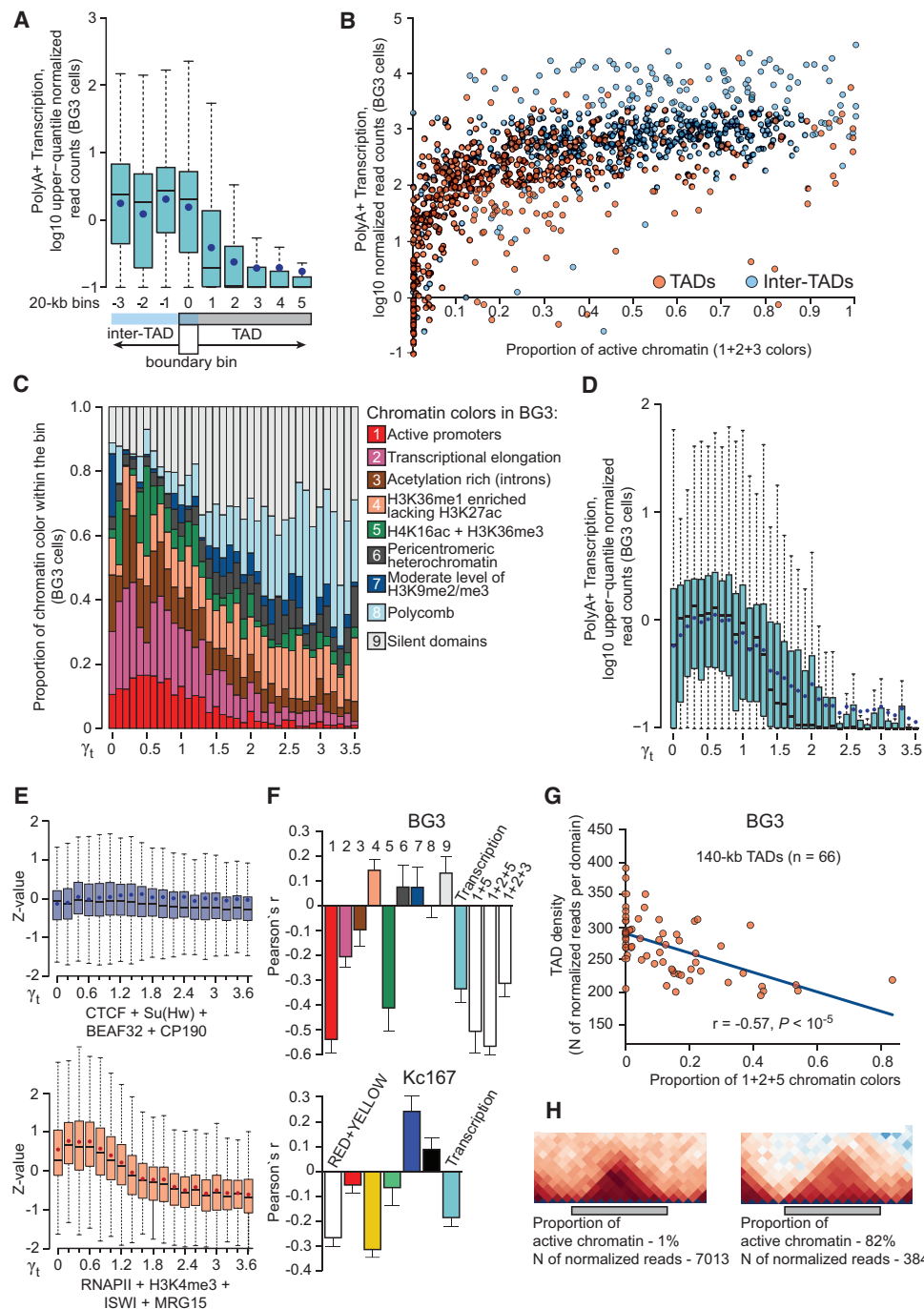


Figure 3. High transcription level and high content of active chromatin interfere with DNA packaging into TADs. (A) The level of poly(A)⁺ transcripts around TAD boundaries in the BG3 cells (data for the other cell lines are presented in Supplemental Fig. S8C). Box plots show the number of upper-quantile normalized transcriptome read counts over all bins located at the same position relative to a TAD boundary. Blue dots denote average values; medians are shown by thick black lines. (B) Scatter plot demonstrating the transcription level and fraction of active chromatin colors (sum of 1, 2, and 3 chromatin types) in individual TADs (excluding boundary bins) and inter-TADs (including TAD boundaries) in the BG3 cells. Data for the S2 and Kc167 cells are presented in Supplemental Figure S9. (C) Box plots showing inverse dependence between the proportion of active chromatin colors and γ_t (the minimal value of the scaling parameter γ required to annotate the bin as a TAD boundary or inter-TAD) in the BG3 cells. Average proportions of chromatin colors over all bins with the same γ_t are shown in each plot. Data for the S2 and Kc167 cells and similar diagrams built with γ_t ranging from 0 to 10 are shown in Supplemental Figure S5B. *P*-values are presented in Supplemental Table S6. (D) Box plots showing the inverse dependence between the transcription level within a genomic bin and γ_t in the BG3 cells. Data for the S2 and Kc167 cells are presented in Supplemental Figure S8D. (E) Box plots demonstrating the dependence between γ_t and the proportion of active chromatin marks (top) or the proportion of architectural proteins (bottom) in BG3 cells. (F) The Pearson correlation coefficients between the proportions of different chromatin colors (numbered as in C) or transcription level and the TAD density averaged over all size groups of TADs. Bar plots show the mean and variance values for the TAD size groups. (G) Scatter plot demonstrating the negative correlation between the TAD density and the proportion of active chromatin colors (sum of 1, 2, and 5 colors) (Kharchenko et al. 2011) for 140-kb-long TADs in the BG3 cell line. (H) A representative pair of TADs with different proportions of active chromatin display different TAD density.

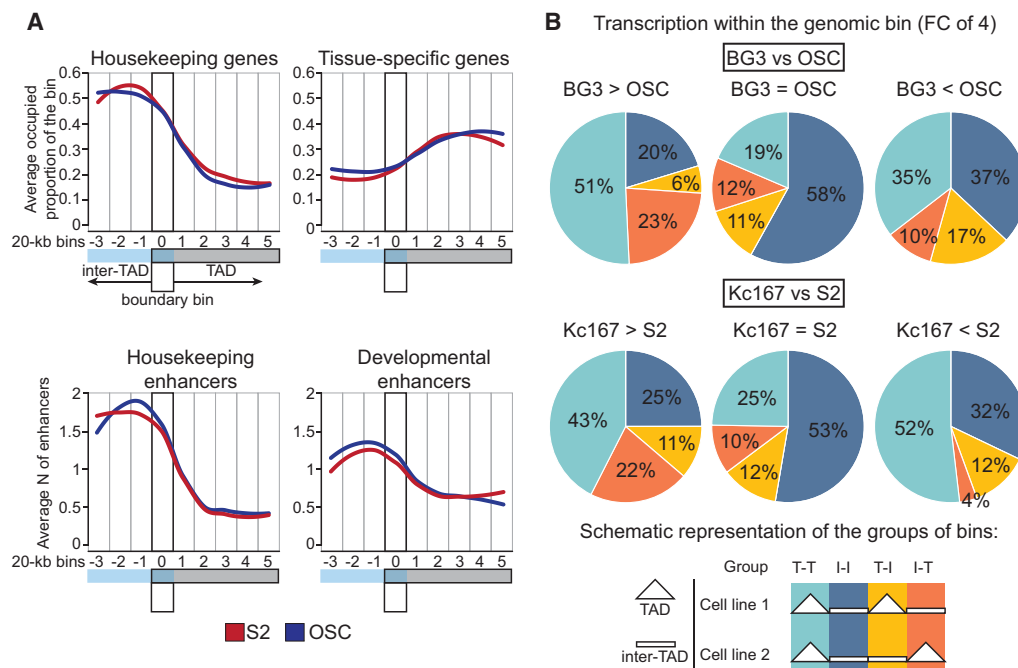


Figure 4. Distribution of housekeeping genes, intensity of transcription, and the presence of active chromatin, as related to TAD profiles. (A) Distribution of housekeeping and tissue-specific genes, housekeeping, and developmental STARR-seq-identified enhancers around TAD boundaries from the S2 cells (red) and OSCs (blue). The proportion of a bin occupied by gene bodies (upper graphs) and the average number of enhancers in a bin (lower graphs) were calculated for all bins located at the same position relative to a TAD boundary. (B) Pie charts showing the distribution of differentially or uniformly transcribed bins in the four groups of bins defined by pairwise comparison of the cell lines (see the description in the text).

used instead of the transcription level (Supplemental Fig. S10B; Supplemental Table S9). Thus, increasing transcription and/or proportion of active chromatin within a genomic bin is frequently accompanied by the transition of this bin from a TAD to an inter-TAD.

In the pairwise comparisons of the BG3 and S2 cell lines, we observed that the proportion of active and repressed chromatin varied significantly within TADs located at the exactly coinciding genomic position in the two cell lines. The same was true for the boundaries of these TADs. However, there was no correlation between the differences of proportion of active chromatin within TADs and at boundary regions [Pearson $r = (-0.07) - 0.14$, $P > 0.05$] (Supplemental Fig. S11). Thus, internal regions and boundaries of TADs are functionally distinct chromatin compartments.

Inter-TADs correspond to inter-bands in polytene chromosomes

Polytene chromosomes in *Drosophila melanogaster* are partitioned into cytologically identified bands and inter-bands. Inter-bands display a high level of transcription, decondensed chromatin, and have a lower DNA compaction ratio than bands (Vatolina et al. 2011). A set of proteins typically overrepresented in inter-bands was recently identified (Zhimulev et al. 2014). Notably, the same proteins, including histone acetyltransferase MOF, chromatin remodeling factors E(bx) (also known as NURF301), WDS and ISWI, H3S10-specific kinase JIL-1, and the chromodomain proteins Chriz (Chromator) and MRG15, demonstrated strong enrichment at the inter-TAD regions and TAD boundaries in our analysis (see Supplemental Fig. S6). It has been reported that inter-band positions predicted by profiles of proteins listed above correspond to TAD boundaries in embryonic nuclei (Zhimulev et al. 2014). In our study, we performed an extended analysis of

the relationship between the inter-bands and the chromatin topology.

We compared genomic positions of 32 cytologically and molecularly identified (Zhimulev et al. 2014) inter-bands with TADs profiles. As shown in Figure 5A, 26 out of 32 of these inter-bands are located in inter-TADs or TAD boundaries in at least one cell line (13 in all cell lines). To compare locations of inter-bands and inter-TADs genome-wide, we analyzed the distribution of inter-band-specific CYAN and BLUE chromatin types described by Zhimulev et al. (2014) in the BG3 and S2 cells (these inter-band colors should not be confused with the chromatin types in Filion et al. [2010] and Kharchenko et al. [2011]). We found that predicted inter-bands are strongly enriched within inter-TADs and TAD boundaries (Fig. 5B). In contrast, the MAGENTA chromatin type (Zhimulev et al. 2014), which is depleted of inter-band-specific protein marks, demonstrated a preferential location within TADs. Additionally, we analyzed the distribution of integration sites of the P-element, known to insert predominantly into inter-bands due to the open chromatin state of these genome regions (Zhimulev et al. 2014). As expected, P-element integration events are enriched in inter-TADs (Supplemental Fig. S6).

Analysis of TAD annotations at various values of the scaling parameter γ (see above) demonstrated that the transitional γ value γ_t depends inversely on the proportion of inter-band-specific chromatin (similarly to the proportion of active chromatin) (Fig. 5C).

Taken together, these results show that inter-TADs largely correspond to polytene chromosome inter-bands, whereas TADs correspond to bands. The visual inspection of a phase-contrast image of polytene chromosomes aligned to a Hi-C heat map confirms this conclusion (Fig. 5D). However, this correspondence is not exact. Although most of the predicted inter-bands colocalize with inter-TADs, some TADs contain a large fraction of CYAN

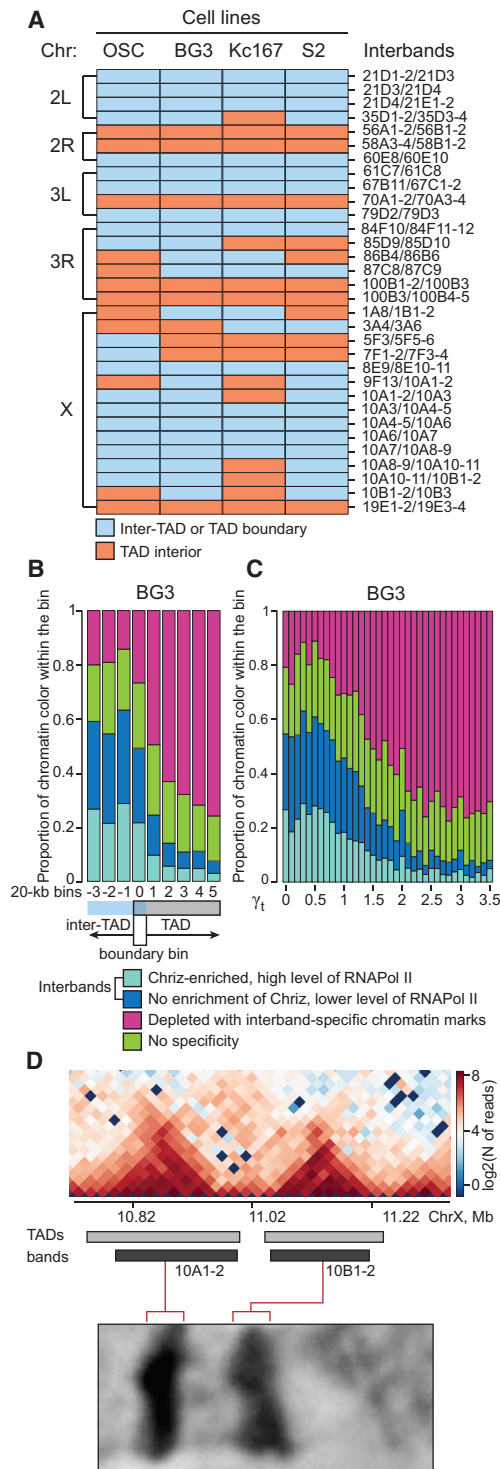


Figure 5. Inter-TADs correspond to polytene chromosome inter-bands. (A) Our annotation of TADs and inter-TADs at sites of 32 cytologically identified inter-bands (Zhimulev et al. 2014). (B) Distribution of inter-band-related chromatin colors (Zhimulev et al. 2014) near TAD boundaries. The alignment of bins relative to the TAD boundary bin is as in Figure 2B. The P -values are presented in Supplemental Table S6. (C) Bar plot showing the inverse dependence between the proportions of inter-band-specific chromatin colors within a genomic bin and γ_t . The P -values are presented in Supplemental Table S6. (D) Phase-contrast image of two bands clearly corresponding to TADs in the BG3 cells. Reprinted from Figure 4A from Vatulina et al. (2011) with permission from the publisher.

and BLUE inter-band-specific chromatin fragments (Supplemental Fig. S12).

Computer simulation of a linear polymer that folds into a set of TADs supports a key role for active chromatin in the separation of TADs

A strong correspondence between TAD partitioning and active/inactive genomic regions suggests that gene-rich segments of the genome that bear active chromatin marks are incapable or less capable of forming compact structures. Such inability to form contacts may be explained by the acetylation of histone tails, which is typical for active chromatin (Shahbazian and Grunstein 2007) and is known to interfere with inter-nucleosomal associations (Shogren-Knaak et al. 2006; Allahverdi et al. 2011). In this case, nucleosomes of inactive chromatin would form various supra-nucleosomal structures due to interactions between positively charged histone tails of one nucleosome and the acidic patch of another nucleosome (Luger et al. 1997; Schalch et al. 2005; Sinha and Shogren-Knaak 2010). These interactions are thought to contribute to folding of nucleosomal arrays into secondary chromatin structures and self-association of arrays into higher-order tertiary structures (Kalashnikova et al. 2013; Peppenella et al. 2014), which collectively may result in formation of TADs. We propose that those differences in physical properties of nucleosomal arrays in active and inactive chromatin are sufficient to direct self-assembly of TADs separated by inter-TADs.

To validate this conjecture, we performed polymer simulations of a 2-Mb region of chromatin consisting of 19 blocks of inactive chromatin (100 kb or 500 nucleosomes each) interspaced by shorter blocks of active chromatin (10 kb, or 50 nucleosomes each) (Fig. 6). We simulated polymer fiber as a sequence of nonoverlapping beads, with each bead representing one nucleosome. The nucleosomes of inactive chromatin (below referred to as inactive nucleosomes) were able to establish short-lived interactions with each other, while the nucleosomes of active chromatin (active nucleosomes) did not form these associations. Importantly, each inactive nucleosome was able to establish only one interaction with another nucleosome at a time (so-called bond saturated interactions) (see Lifshitz et al. 1976; Mukhopadhyay et al. 2011). Saturated interactions were important for TAD segregation and prevented aggregation of all inactive nucleosomes together (see Methods; Supplemental Fig. S16). Simulations were performed using dissipative particle dynamics (DPD) (Hoogerbrugge and Koelman 1992; Español and Warren 1995), using the code developed previously in Gavrilov et al. (2013) and Chertovich and Kos (2014).

Twelve independent simulations of the model were performed. Upon equilibration, reversibility in the association/dissociation of inactive beads was reached, and the corresponding chain conformation was considered as a representative example of a typical chromatin 3D structure (Fig. 6A; Supplemental Fig. S13). Finally, we averaged the structures over 12 independent runs to obtain the ensemble distance map, simulated Hi-C interactions map, and the decay of contact probability with the distance (Fig. 6D; Supplemental Fig. S15B). The Hi-C map for a group of three consecutive TADs was obtained using a contact distance of 40 nm (four nucleosome diameters) and was averaged over 17 starting positions in each of the 12 realizations.

The resulting polymer conformations displayed aggregation of simulated nucleosomes into condensed blobs (analogous to

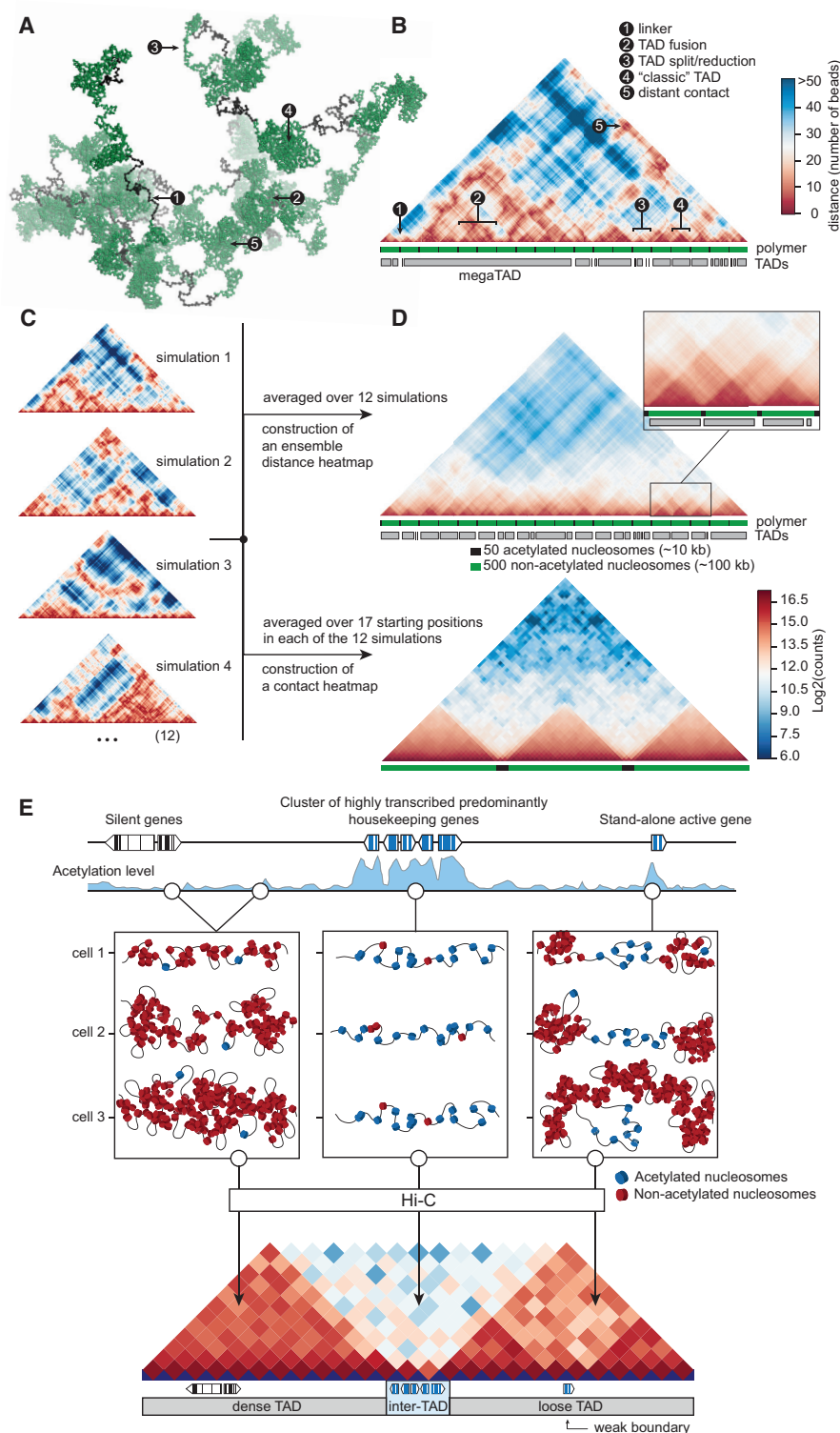


Figure 6. Computer simulation of a linear polymer that folds into a set of TADs supports a key role of hyperacetylated chromatin in separation of TADs. (A) One of the predicted spatial configurations of a polymer composed of 19 blocks of inactive (interacting) nucleosomes (500 nucleosomes each, green) interspaced by shorter blocks of active noninteracting nucleosomes (50 nucleosomes each, black). (B) Spatial proximity map (distance heat map) of the polymer configuration presented in A. Distances are measured in numbers of nucleosomes. A scheme of the model polymer and positions of TADs predicted by the Armatus algorithm are shown below the map. (C) Distance heat maps of the four individual configurations of the model polymer. Configuration 1 is used in A and B. (D) Distance heatmap (upper) and contact heatmap (simulated Hi-C map constructed at resolution of 4 kb [20 beads] for three consecutive TADs, lower) of the model polymer obtained by the averaging of the heat maps of 12 individual configurations. Notation as in B. (E) A schematic illustrating the proposed model of chromatin folding into TADs/inter-TADs, as directed by the self-association of nucleosomes. A high acetylation level of the chromatin within the genomic regions harboring actively transcribed genes interferes with chromatin packaging into TADs due to decreased inter-nucleosomal interactions.

TADs) consisting mostly of inactive nucleosomes (green color in Fig. 6; Supplemental Fig. S13) separated by relatively stretched spacers (respectively, inter-TADs) of active nucleosomes (black color in Fig. 6; Supplemental Fig. S13). Although the general rule is that inactive nucleosomes tend to occur in TADs and active nucleosomes are in inter-TADs, there are numerous exceptions. Our simulations demonstrated that TADs in individual cells can fuse and assemble into large loose aggregates (mega-TADs), separate into smaller blobs, and form long-range contacts (Fig. 6B,C; see also Supplemental Fig. S13; Supplemental Video S1). Partitioning of the polymer into TADs of different sizes was verified by running the Armatus algorithm on the maps obtained in the computer simulations (see track “TADs” below the map in Fig. 6B). Although the positions of the predicted TADs in individual simulations were not the same, the regular pattern of TADs determined by the alternation of blocks of interacting (green) and noninteracting (black) nucleosomes could be traced in each individual simulation (Fig. 6B,C; Supplemental Fig. S13). This basic pattern became much clearer after averaging all 12 simulations (Fig. 6D). In this case, Armatus produced an almost uniform TAD distribution, with TADs generally corresponding to single blocks of interacting particles (of note, we obtained the same results in modeling of a polymer composed of equal blocks of 500 acetylated and 500 nonacetylated nucleosomes) (Fig. 6D; Supplemental Fig. S14).

Discussion

Here, we perform Hi-C to analyze chromatin folding in four unrelated *Drosophila* cell lines. We find that active chromatin demarcates TADs in cells of various origins and developmental stages, while the repressed chromatin is localized within TADs. Based on this finding, we propose a mechanistic model of chromatin self-organization into TADs and inter-TADs (Fig. 6E). We suggest that chromatin is intrinsically partitioned into self-organized and potentially highly dynamic globular structures that are primarily built from nonacetylated nucleosomes and are separated by unstructured linkers composed of acetylated nucleosomes. This partitioning is seen as a TAD/inter-TAD profile after Hi-C analysis.

This study (Figs. 2–4) and others (Hou et al. 2012; Sexton et al. 2012) revealed that boundaries and inter-TADs in *Drosophila*, as opposed to TADs, are strongly enriched with active chromatin and its individual marks, as well as with active transcription and with constitutively transcribed housekeeping genes. Consequently, active chromatin marks, in the simplest case only total transcription and H3K4me3 (a mark of active promoters), can relatively well predict a TAD/inter-TAD profile (Fig. 2D). The existence of long inter-TADs (Supplemental Fig. S4) composed of active chromatin is, per se, an argument for the ability of this type of chromatin to separate TADs. Furthermore, our observations demonstrate that the presence of active chromatin and transcribed regions within TADs undermines the TAD integrity, making the TAD less compact and generating weak boundaries inside the TAD. Consequently, a bona fide TAD is inactive; TADs containing active chromatin become less dense, acquire weak internal boundaries, and eventually split into smaller TADs that are composed of inactive chromatin. The observation that the majority of housekeeping genes are located within inter-TADs and TAD boundaries suggests that evolutionary conservation and cell type independence of TAD/inter-TAD profiles may be explained by the conservation of positions of housekeeping genes along the chromosomes.

We note that chromosomal interaction domains similar to TADs have been observed in the bacterium *Caulobacter crescentus*, where they are demarcated by sites of active transcription (Le et al. 2013). Although the basic level of chromosomal folding is different in bacteria and eukaryotes, the model proposed in Le et al. (2013) and our model stem from common principles. In *Caulobacter*, active transcription is thought to disrupt the fiber of supercoils (plectonemes) by creating a stretch of nonpackaged DNA, free of plectonemes, which spatially separates chromosomal regions flanking it. In our model, transcription disrupts chromatin organization by introducing a “nonsticky” region of chromatin, which is less compact and more unfolded in space, and thus spatially separates two flanking regions.

Our computer modeling shows that stickiness of nonacetylated (inactive) nucleosomes and the absence of stickiness for acetylated (active) nucleosomes are sufficient for chromatin partitioning into TADs and inter-TADs (Fig. 6). Self-association of nucleosomes may be explained by the interaction of positively charged histone tails (in particular, the tail of histone H4) of one nucleosome with the acidic patch of histones H2A/H2B at an adjacent nucleosome (Luger et al. 1997; Schalch et al. 2005; Sinha and Shogren-Knaak 2010). Acetylation of histone tails, which is typical of active chromatin (Shahbazian and Grunstein 2007), may interfere with inter-nucleosomal associations (Shogren-Knaak et al. 2006; Allahverdi et al. 2011). In addition to a high level of histone acetylation, other features of active chromatin, including lower nucleosome density in inter-TADs, manifested as the decreased histone H3 occupancy (Fig. 2C), might contribute to the generation of TAD profiles.

It should be mentioned that a significant difference between our polymer simulations and models previously suggested by the Cavalli and Vaillant groups (Jost et al. 2014) is the use of saturating interactions between inactive nucleosomes. In the case of volume interactions, all nucleosomes of the same type adjacent in 3D space will attract each other; in the case of saturating interactions, each nucleosome may attract only one neighbor. Using volume interactions leads to the formation of a single dense blob (Supplemental Fig. S16) and does not produce TADs in a simulated Hi-C map. We note that the saturating nature of interactions between nucleosomes is based on the known properties of nucleosomal particles (Schalch et al. 2005).

Previous studies considered a variety of mechanisms that may lead to the formation of TADs. In particular, Barbieri et al. (2012) studied segregation of two TADs using cubic lattice simulations of a short 152-monomer chain consisting of two TADs, assuming that inter-monomer interactions could only form between monomers belonging to the same TAD. In our model, we show that TADs emerge without requiring such specific interactions; any two regions of sticky monomers separated by a nonsticky linker would form TADs. Another study proposed that transcription-induced supercoiling may be responsible for the formation of TADs (Benedetti et al. 2014). Although this model is consistent with our observation that sites of active transcription demarcate TAD boundaries, there is limited evidence that supercoiling of chromatinized DNA exists in *Drosophila* and other organisms. On the contrary, our model is based on known biochemical properties of nucleosomes.

The fact that a minor fraction of TADs is built mostly from active chromatin (Fig. 3B) apparently contradicts our model, suggesting that additional ways of chromatin self-organization could exist. One possibility is the establishment of long-range contacts between enhancers and their cognate promoters, as well as loops

between pairs of insulators (Doyle et al. 2014). Such loops formed inside active unstructured chromatin linkers (i.e., inter-TADs) could probably be sufficient to compact them and thus to fold into TADs.

TAD profiles of X Chromosomes are almost identical in the male and female cell lines (Supplemental Fig. S17), in agreement with recently published observations (Ramírez et al. 2015). Thus, it seems that hyperacetylation of male X Chromosomes due to dosage compensation (Akhtar and Becker 2000; Kind et al. 2008) does not generate new TAD boundaries. However, it should be noted that MOF histone acetyltransferase of the MSL complex introduces only the H4K16ac mark. Although this modification is important to prevent inter-nucleosomal interactions (Shogren-Knaak et al. 2006), acetylation at other histone positions (Pepenna et al. 2014) and H2B ubiquitylation (Fierz et al. 2011) contribute as well. Additionally, H4K16 acetylation generated by the dosage compensation system occurs preferentially at regions enriched with transcribed genes (Gilfillan et al. 2006), and hence within inter-TADs.

Our analysis does not support the previously reported (Hou et al. 2012; Sexton et al. 2012) strong enrichments of insulator proteins Su(Hw) and dCTCF at TAD boundaries in *Drosophila*. To assess the possible reasons of this divergence, we re-analyzed the dCTCF distribution with respect to TAD positions in our data set using the raw ChIP-seq data (Wood et al. 2011) used by Hou et al. (2012). No strong difference was observed in the dCTCF coverage in TADs and inter-TADs (Supplemental Fig. S18A). Interestingly, we obtained the same result while analyzing dCTCF and Su(Hw) binding within TAD boundaries identified by Hou et al. (2012) (Supplemental Fig. S18E). However, we observed a strong enrichment of dCTCF at TAD boundaries when the peak distribution was analyzed instead of read coverage (Supplemental Fig. S18B–D). Additionally, the effect was much weaker when modENCODE peaks were used. Hence, the discrepancy may be caused by a different peak calling procedure in modENCODE and in Hou et al. (2012). The biological significance of these observations remains to be determined. We note that disruption of the cohesin/CTCF complex in mammals, as well as depletion of the Vtd (also known as Rad21) cohesin subunit in *Drosophila*, did not lead to disappearance of TAD boundaries, but rather only slightly decreased interactions inside TADs (in mammals) and reduced TAD boundary strength in the *Drosophila* genome (Sofueva et al. 2013; Li et al. 2015). These observations favor a role for the cohesin/CTCF complex, which is known to form loops (Splinter et al. 2006; Holwerda and de Laat 2013), in chromatin compaction inside the TADs.

Binding of insulator proteins might contribute to establishing TAD boundaries through introducing active chromatin marks. Indeed, when inserted into an ectopic position, a classical insulator triggers hyperacetylation of the local chromatin domain (Mutskov et al. 2002) and recruits chromatin-remodeling complexes (Li et al. 2011). However, absence of strong enrichment of dCTCF at TAD boundaries and preferential location of Su(Hw) inside TADs mean that, at least, dCTCF- and Su(Hw)-dependent insulators are not the major determinants of TAD boundaries and inter-TADs.

TADs are predicted based on the analysis of averaged data from a cell population. Although they are usually represented as large chromatin globules (for review, see Mirny 2011; Fudenberg and Mirny 2012; Bickmore 2013; de Graaf and van Steensel 2013; Dekker et al. 2013; Gibcus and Dekker 2013), direct experimental evidence for the existence of such globules in individual cells is controversial. Using confocal and 3D-SIM microscopy,

Cremer and colleagues observed ~1-Mb globular domains within chromosomal territories (Cremer and Cremer 2001; Kolbl et al. 2012; Smeets et al. 2014). However, using STORM microscopy, Ricci et al. found that chromatin in individual mammalian cells is organized into “clutches” composed of several nucleosomes, and that increased histone acetylation dramatically reduces the size of these clutches (Ricci et al. 2015). It is thus possible that sub-megabase TADs revealed by Hi-C represent a set of nucleosome clutches separated by relatively short spacers of various sizes. These short clutches may occupy various positions within TADs in different cells and stochastically assemble to form short-living aggregates. The stochastic nature of TADs is supported by our computer simulations.

Methods

Cell cultures

Drosophila melanogaster Schneider-2 (S2) and Kc167 cell cultures obtained from the collection of IMG RAS were grown at 25°C in Schneider's *Drosophila* Medium (Gibco) supplemented with 10% heat-inactivated fetal bovine serum (FBS, Gibco), 50 units/mL penicillin, and 50 µg/mL streptomycin. An ML-DmBG3-c2 (BG3) cell culture obtained from the *Drosophila* Genomics Resource Center was grown under the same conditions, except that half of Schneider's *Drosophila* Medium was substituted with Shields and Sang M3 insect medium (Sigma). Ovarian somatic cells (OSCs), kindly provided by M. Siomi, were cultured in Shields and Sang M3 insect medium (Sigma-Aldrich) supplemented with 10% heat-inactivated FBS (Gibco), 10% fly extract (<http://biology.st-andrews.ac.uk/sites/flycell/flyextract.html>), 10 µg/mL insulin (Sigma-Aldrich), 0.6 mg/mL glutathione (Sigma-Aldrich), 50 units/mL penicillin, and 50 µg/mL streptomycin.

Generation and analysis of Hi-C libraries

Hi-C of each cell line was performed in two biological replicates using the HindIII-HF restriction endonuclease according to a previously published protocol (Belton et al. 2012), with minor modifications (see Supplemental Methods). Hi-C libraries were sequenced using Illumina HiSeq 2000, and reads were mapped to the *D. melanogaster* reference genome (version dm3) using Bowtie 2 v2.2.1 (with the –very-sensitive option) (Langmead and Salzberg 2012). The Hi-C data were processed using the ICE pipeline v0.9 (20 iterations of iterative correction) as described (Imakaev et al. 2012).

TADs were predicted using the Armatus software v1.0 (Filippova et al. 2014), in which the average size and the number of TADs are determined by the scaling parameter γ . Initially, we manually selected parameter γ to achieve good partitioning of TADs. However, such annotation of TADs contained several extremely large TADs (up to 3 Mb) that could be visually subdivided into smaller domains with relatively weak inter-domain interactions (Fig. 1A). To annotate these subdomains as separate TADs, it was necessary to increase the value of the scaling parameter γ . However, doing this genome-wide would lead to an excessive splitting of many bona fide TADs into clusters of smaller TADs that did not look like independent TADs. To solve the problem, TAD annotation was performed in two steps. At the first step, the scaling parameter was set to 1.03–1.26 for the whole genome. At the second step, TADs larger than 600 kb were further split into smaller TADs with the scaling parameter set to 2.06–2.52. After that, the smallest of the detected TADs (60-kb TADs composed of one internal and two boundary 20-kb bins) were annotated as inter-TADs due to their poorly resolved internal structure.

Poly(A)⁺ RNA-seq: experimental procedures and analysis

RNA extraction was carried out using an RNeasy Mini kit (Qiagen) following the manufacturer's instructions. RNA quality was assessed using capillary electrophoresis with a Bioanalyzer 2100 (Agilent). For library preparation, a TruSeq RNA Sample Prep kit v2 (Illumina) was used following the manufacturer's instructions. After preparation, libraries were quantified using a Qubit fluorometer and quantitative PCR and sequenced with HiSeq 2000 with read lengths of 51 nt. Reads were mapped to the *D. melanogaster* reference genome (version dm3) and FlyBase gene annotation v5.48 using TopHat2 v2.0.12 (with options -g 1 -x 1 -M -no-novel-indels) (Supplemental Table S9; Kim et al. 2013). Replicates demonstrated very high similarity according to PCA (Supplemental Fig. S8A) and cluster analysis (Supplemental Fig. S8B). Replicates were not merged to correctly estimate dispersion in further differential expression analyses. BEDTools v2.16.2 (Quinlan and Hall 2010) was used to calculate the read coverage in 20-kb genomic bins and genes, according to FlyBase annotation v5.48. Differentially expressed bins and genes were identified with the edgeR package using the RLE normalization method (McCarthy et al. 2012).

De novo genome sequencing for the studied cell cultures

High-molecular-weight genomic DNA isolated from the S2, Kc167, BG3, and OSC cells according to a standard method (Maniatis et al. 1982) was fragmented into 8–10 kb by “Evrogen” (<http://www.evrogen.com/>) with the use of the Nextera Mate Pair library Sample Prep kit (Illumina). The libraries were prepared with the use of the TruSeq DNA Sample Prep kit and quantitated by qPCR; the libraries were pooled and sequenced from both fragment ends in one lane for 111 cycles using HiSeq 2500 in the high output mode with TruSeq SBS sequencing chemistry version 3. FASTQ files were generated and de-multiplexed with the bcl2fastq v1.8.4 Conversion Software (Illumina). Only read pairs that contained the circularization adaptor CTGTCTCTTATACA CATCT were used in the further analysis because they were guaranteed to derive from real mate-pair fragments. The circularization adaptor was trimmed using a custom script. The BreakDancer (with options -t -q 10 -l) (Chen et al. 2009) and Delly (with option -t TRA) (Rausch et al. 2012) programs were used to predict genomic rearrangements (Supplemental Table S11).

modENCODE data processing

Normalized and smoothed WIG files corresponding to chromatin marks (Celniker et al. 2009) were downloaded from the modENCODE database. Similarly to Khrameeva et al. (2012), to measure the chromatin mark density $S(i)$ in each 20-kb bin i , we multiplied the height of each chromatin mark peak (H_k) by the fraction of the bin i intersecting with the peak (W_k) and then summed the results for all peaks $1...n$ in the bin i :

$$S(i) = \sum_{k=1}^n (H_k \cdot W_k).$$

$S(i)$ values were further Z-transformed to generate distributions of different comparable chromatin marks. If there were several data sets available for one chromatin mark, we averaged their Z-transformed density values.

STARR-seq (Zabidi et al. 2015), inter-bands (Zhimulev et al. 2014), P-elements (FlyBase), CP190, and dCTCF (Wood et al. 2011) data sets were processed similarly.

Prediction of TADs by chromatin marks

The logistic regression model implemented via the glm function in R (R Core Team 2015) was used to model the outcome variable as a linear combination of predictor variables. The outcome variable was binary (0/1), where 0 denoted bins within TADs and 1 denoted bins within inter-TADs. Z-transformed values of chromatin mark densities in bins were used as predictor variables. ROC curves and area under the curve (AUC) values were calculated with the pROC package in R. As a control, we shuffled the outcome variable and repeated the model calculation to estimate the quality of the model prediction.

Housekeeping and tissue-specific genes

A previously published data set (Brown et al. 2014) was used to compile a list of tissue-specific genes. We divided samples from this study into eight groups, with each group corresponding to a particular developmental stage (Supplemental Table S10). A gene was assumed to be tissue-specific within a developmental stage group if its maximum RPKM value was fourfold higher than its minimum RPKM value in the group. The final set of tissue-specific genes was composed of 3634 genes that were tissue-specific in five or more developmental stage groups. Another data set (Graveley et al. 2011) was used to compose a list of 5906 housekeeping genes (genes with RPKM > 1 in all samples).

Computer modeling and simulation technique

Computer simulation of a single copolymer chain in a periodic boundary box was performed using the dissipative particle dynamics algorithm with a simulation code, which was previously applied to an analysis of the microphase segregation of copolymer melts and the collapse of long single chains (Gavrilov et al. 2013; Chertovich and Kos 2014). DPD is a version of coarse-grained molecular dynamics adapted to polymers and mapped onto the classical lattice Flory-Huggins theory (Groot and Warren 1997). In short, macromolecules are represented in terms of the bead-and-spring model, with particles interacting by a conservative force (repulsion), a dissipative force (friction), and a random force (heat generator). A soft repulsive potential enhances the stability of the numerical scheme for integrating the equations of motion, making it possible to increase the time step and thus access large time scales when complex polymeric structures are analyzed. It has been shown previously that this model describes the dynamics of polymer melts well at all timescales.

Consider an ensemble of particles (beads) obeying the Newton equations of motion

$$\begin{aligned} \frac{dr_i}{dt} &= v_i; \quad m_i \frac{dv_i}{dt} = f_i, \\ f_i &= \sum_{j \neq i} (F_{ij}^b + F_{ij}^c + F_{ij}^d + F_{ij}^r), \end{aligned} \quad (1)$$

where r_i , m_i , v_i are the coordinate, mass, and velocity of an i -th bead, respectively, and f_i is the force acting on it. The summation is performed over all other beads within the cut-off radius r_c . Below, we assume that all quantities entered into Eq. (1) are dimensionless and for simplicity set r_c and m_i for any i to unity. The first two terms in the sum are the conservative forces.

Macromolecules are represented in terms of the bead-and-spring model. F_{ij}^b is a spring force describing the chain connectivity of beads:

$$F_{ij}^b = -K(r_{ij} - l) \frac{r_{ij}}{r_{ij}}, \quad (2)$$

where K is the bond stiffness. If beads i and j are not connected, then $\mathbf{F}_{ij}^b = 0$. \mathbf{F}_{ij}^c is a soft core repulsion between the i - and j -th beads:

$$\mathbf{F}_{ij}^c = \begin{cases} a_{ij}(1 - r_{ij})\mathbf{r}_{ij}/r_{ij}, & r_{ij} \leq 1 \\ 0, & r_{ij} > 1 \end{cases}, \quad (3)$$

where a_{ij} is the maximum repulsion between beads i and j attained at $r_i = r_j$. Because \mathbf{F}_{ij}^c has no singularity at zero distance, a much larger time step than in a standard molecular dynamics could be used.

Other constituents of \mathbf{f}_i are a random force, \mathbf{F}_{ij}^r , and a dissipative force, \mathbf{F}_{ij}^d , acting as the heat source and medium friction, respectively, and are taken as dictated by the Groot-Warren thermostat. A more detailed description of the simulation methodology may be found in Groot and Warren (1997).

It has been shown that the DPD method is consistent with both the scaling theory of polymers (e.g., it provides correct relationships between the average radius of gyration of a coil and the number of units in the coil) and the Rouse dynamics. In our study, we set $a_{ii} = a_{ij} = 150$, $l = 0.5$, and $K = 150$. The other parameters were the DPD number density $\rho = 3$, the noise parameter $\sigma = 3$, and the integration time step $\Delta t = 0.04$.

We used rather long chains with $N \sim 10^4$ monomer units. The starting conformation of a polymer chain was a random walk trajectory in the $48 \times 48 \times 48$ simulation box. Conformational changes occur during the equilibration process and are caused by the reversible association of inactive beads (nonacetylated nucleosomes). We set probabilities of $P1 = 0.1$ for forming associations between nonacetylated beads neighboring in space (green beads on Fig. 6A) and $P2 = 0.01$ for destroying such associations. Thus in equilibrium, $\sim 90\%$ of all possible inactive beads form associations. To run simulations, we used an in-house domain-decomposition parallelized DPD code and performed simulations using the MSU supercomputer facilities (<http://hpc.msu.ru/>). For averaging over different realizations, we performed 12 independent runs, starting with different random walk trajectories. Each conformation was evaluated during the time of 2×10^5 DPD steps on 256 processors, which takes ~ 1 d per independent run.

The Armatus software v1.0 (Filippova et al. 2014) was applied to predict TADs in the structures averaged over 12 independent runs. As an input, we used 5-kb-binned distance matrices that were transformed into contact frequency matrices using inverse function, normalized to the upper quartile, and \log_2 -transformed.

Data access

Raw sequencing reads for the Hi-C libraries, RNA-seq, and de novo sequencing of the genomes of the studied cell lines, as well as iteratively corrected Hi-C heat maps from this study, have been submitted to NCBI Gene Expression Omnibus (GEO; <http://www.ncbi.nlm.nih.gov/geo/>) under accession number GSE69013.

Acknowledgments

This study was supported by the Russian Foundation for Basic Research grants (13-00-40086-K, 14-04-91331, 14-04-00010, 14-03-00825), and by the Presidium of the Russian Academy of Sciences grants (MCB). A.A.G. and I.M.F. were supported by the Russian Science Foundation (project 14-14-01088), and M.S.G. was supported by the Russian Science Foundation (grant 14-24-00155). We thank the MSU supercomputer center for providing the computational facilities.

References

- Ahanger SH, Gunther K, Weth O, Bartkuhn M, Bionde RR, Shouche YS, Renkawitz R. 2014. Ectopically tethered CP190 induces large-scale chromatin decondensation. *Sci Rep* **4**: 3917.
- Akhtar A, Becker PB. 2000. Activation of transcription through histone H4 acetylation by MOF, an acetyltransferase essential for dosage compensation in *Drosophila*. *Mol Cell* **5**: 367–375.
- Allahverdi A, Yang R, Korolev N, Fan Y, Davey CA, Liu CF, Nordenskiöld L. 2011. The effects of histone H4 tail acetylations on cation-induced chromatin folding and self-association. *Nucleic Acids Res* **39**: 1680–1691.
- Barbieri M, Chotalia M, Fraser J, Lavitas LM, Dostie J, Pombo A, Nicodemi M. 2012. Complexity of chromatin folding is captured by the strings and binders switch model. *Proc Natl Acad Sci* **109**: 16173–16178.
- Bartkuhn M, Straub T, Herold M, Herrmann M, Rathke C, Saumweber H, Gilfillan GD, Becker PB, Renkawitz R. 2009. Active promoters and insulators are marked by the centrosomal protein 190. *EMBO J* **28**: 877–888.
- Belton JM, McCord RP, Gibcus JH, Naumova N, Zhan Y, Dekker J. 2012. Hi-C: a comprehensive technique to capture the conformation of genomes. *Methods* **58**: 268–276.
- Benedetti F, Dorier J, Burnier Y, Stasiak A. 2014. Models that include supercoiling of topological domains reproduce several known features of interphase chromosomes. *Nucleic Acids Res* **42**: 2848–2855.
- Berlivet S, Paquette D, Dumouchel A, Langlais D, Dostie J, Kmita M. 2013. Clustering of tissue-specific sub-TADs accompanies the regulation of *HoxA* genes in developing limbs. *PLoS Genet* **9**: e1004018.
- Bickmore WA. 2013. The spatial organization of the human genome. *Annu Rev Genomics Hum Genet* **14**: 67–84.
- Brown JB, Boley N, Eisman R, May GE, Stoiber MH, Duff MO, Booth BW, Wen J, Park S, Suzuki AM, et al. 2014. Diversity and dynamics of the *Drosophila* transcriptome. *Nature* **512**: 393–399.
- Celniker SE, Dillon LA, Gerstein MB, Gunsalus KC, Henikoff S, Karpen GH, Kellis M, Lai EC, Lieb JD, MacAlpine DM, et al. 2009. Unlocking the secrets of the genome. *Nature* **459**: 927–930.
- Chen K, Wallis JW, McLellan MD, Larson DE, Kalicki JM, Pohl CS, McGrath SD, Wendl MC, Zhang Q, Locke DP, et al. 2009. BreakDancer: an algorithm for high-resolution mapping of genomic structural variation. *Nat Methods* **6**: 677–681.
- Chertovich A, Kos P. 2014. Crumpled globule formation during collapse of a long flexible and semiflexible polymer in poor solvent. *J Chem Phys* **141**: 134903.
- Cremer T, Cremer C. 2001. Chromosome territories, nuclear architecture and gene regulation in mammalian cells. *Nat Rev Genet* **2**: 292–301.
- de Graaf CA, van Steensel B. 2013. Chromatin organization: form to function. *Curr Opin Genet Dev* **23**: 185–190.
- Dekker J, Marti-Renom MA, Mirny LA. 2013. Exploring the three-dimensional organization of genomes: interpreting chromatin interaction data. *Nat Rev Genet* **14**: 390–403.
- Dixon JR, Selvaraj S, Yue F, Kim A, Li Y, Shen Y, Hu M, Liu JS, Ren B. 2012. Topological domains in mammalian genomes identified by analysis of chromatin interactions. *Nature* **485**: 376–380.
- Dixon JR, Jung I, Selvaraj S, Shen Y, Antosiewicz-Bourget JE, Lee AY, Ye Z, Kim A, Rajagopal N, Xie W, et al. 2015. Chromatin architecture reorganization during stem cell differentiation. *Nature* **518**: 331–336.
- Doyle B, Fudenberg G, Imakaev M, Mirny LA. 2014. Chromatin loops as allosteric modulators of enhancer-promoter interactions. *PLoS Comput Biol* **10**: e1003867.
- Duan Z, Andronescu M, Schutz K, McIlwain S, Kim YJ, Lee C, Shendure J, Fields S, Blau CA, Noble WS. 2011. A three-dimensional model of the yeast genome. *Nature* **465**: 363–367.
- Echalier G, Ohanessian A. 1969. [Isolation, in tissue culture, of *Drosophila melanogaster* cell lines]. *C R Acad Sci Hebd Seances Acad Sci D* **268**: 1771–1773.
- Español P, Warren P. 1995. Statistical mechanics of dissipative particle dynamics. *Europhys Lett* **30**: 191.
- Feng S, Cokus SJ, Schubert V, Zhai J, Pellegrini M, Jacobsen SE. 2014. Genome-wide Hi-C analyses in wild-type and mutants reveal high-resolution chromatin interactions in *Arabidopsis*. *Mol Cell* **55**: 694–707.
- Fierz B, Chatterjee C, McGinty RK, Bar-Dagan M, Raleigh DP, Muir TW. 2011. Histone H2B ubiquitylation disrupts local and higher-order chromatin compaction. *Nat Chem Biol* **7**: 113–119.
- Filion GJ, van Bommel JG, Braunschweig U, Talhout W, Kind J, Ward LD, Brugman W, de Castro IJ, Kerkhoven RM, Bussemaker HJ, et al. 2010. Systematic protein location mapping reveals five principal chromatin types in *Drosophila* cells. *Cell* **143**: 212–224.
- Filippova D, Patro R, Duggal G, Kingsford C. 2014. Identification of alternative topological domains in chromatin. *Algorithms Mol Biol* **9**: 14.
- Fudenberg G, Mirny LA. 2012. Higher-order chromatin structure: bridging physics and biology. *Curr Opin Genet Dev* **22**: 115–124.

- Gavrilov AA, Kudryavtsev YV, Chertovich AV. 2013. Phase diagrams of block copolymer melts by dissipative particle dynamics simulations. *J Chem Phys* **139**: 224901.
- Gibcus JH, Dekker J. 2013. The hierarchy of the 3D genome. *Mol Cell* **49**: 773–782.
- Gilfillan GD, Straub T, de Wit E, Greil F, Lamm R, van Steensel B, Becker PB. 2006. Chromosome-wide gene-specific targeting of the *Drosophila* dosage compensation complex. *Genes Dev* **20**: 858–870.
- Gohl D, Aoki T, Blanton J, Shanower G, Kappes G, Schedl P. 2011. Mechanism of chromosomal boundary action: roadblock, sink, or loop? *Genetics* **187**: 731–748.
- Graveley BR, Brooks AN, Carlson JW, Duff MO, Landolin JM, Yang L, Artieri CG, van Baren MJ, Boley N, Booth BW, et al. 2011. The developmental transcriptome of *Drosophila melanogaster*. *Nature* **471**: 473–479.
- Grob S, Schmid MW, Grossniklaus U. 2014. Hi-C analysis in *Arabidopsis* identifies the KNOT, a structure with similarities to the flamenco locus of *Drosophila*. *Mol Cell* **55**: 678–693.
- Groot RD, Warren PB. 1997. Dissipative particle dynamics: bridging the gap between atomistic and mesoscopic simulation. *J Chem Phys* **107**: 4423.
- Hizume K, Yoshimura SH, Takeyasu K. 2005. Linker histone H1 per se can induce three-dimensional folding of chromatin fiber. *Biochemistry* **44**: 12978–12989.
- Holwerda SJ, de Laat W. 2013. CTCF: the protein, the binding partners, the binding sites and their chromatin loops. *Philos Trans R Soc Lond B Biol Sci* **368**: 20120369.
- Hoogerbrugge PJ, Koelman JMVA. 1992. Simulating microscopic hydrodynamic phenomena with dissipative particle dynamics. *Europhys Lett* **19**: 155.
- 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. *Mol Cell* **48**: 471–484.
- Imakaev M, Fudenberg G, McCord RP, Naumova N, Goloborodko A, Lajoie BR, Dekker J, Mirny LA. 2012. Iterative correction of Hi-C data reveals hallmarks of chromosome organization. *Nat Methods* **9**: 999–1003.
- Jiang N, Emberly E, Cuvier O, Hart CM. 2009. Genome-wide mapping of boundary element-associated factor (BEAF) binding sites in *Drosophila melanogaster* links BEAF to transcription. *Mol Cell Biol* **29**: 3556–3568.
- Jost D, Carrivain P, Cavalli G, Vaillant C. 2014. Modeling epigenome folding: formation and dynamics of topologically associated chromatin domains. *Nucleic Acids Res* **42**: 9553–9561.
- Kalashnikova AA, Porter-Goff ME, Muthurajan UM, Luger K, Hansen JC. 2013. The role of the nucleosome acidic patch in modulating higher order chromatin structure. *J R Soc Interface* **10**: 20121022.
- Kharchenko PV, Alekseyenko AA, Schwartz YB, Minoda A, Riddle NC, Ernst J, Sabo PJ, Larschan E, Gorchakov AA, Gu T, et al. 2011. Comprehensive analysis of the chromatin landscape in *Drosophila melanogaster*. *Nature* **471**: 480–485.
- Khrameeva EE, Mironov AA, Fedonin GG, Khaitovich P, Gelfand MS. 2012. Spatial proximity and similarity of the epigenetic state of genome domains. *PLoS One* **7**: e33947.
- Kim D, Pertea G, Trapnell C, Pimentel H, Kelley R, Salzberg SL. 2013. TopHat2: accurate alignment of transcriptomes in the presence of insertions, deletions and gene fusions. *Genome Biol* **14**: R36.
- Kind J, Vaquerizas JM, Gebhardt P, Gentzel M, Luscombe NM, Bertone P, Akhtar A. 2008. Genome-wide analysis reveals MOF as a key regulator of dosage compensation and gene expression in *Drosophila*. *Cell* **133**: 813–828.
- Kolbl AC, Weigl D, Mulaw M, Thormeyer T, Bohlander SK, Cremer T, Dietzel S. 2012. The radial nuclear positioning of genes correlates with features of megabase-sized chromatin domains. *Chromosome Res* **20**: 735–752.
- Langmead B, Salzberg SL. 2012. Fast gapped-read alignment with Bowtie 2. *Nat Methods* **9**: 357–359.
- Le TB, Imakaev MV, Mirny LA, Laub MT. 2013. High-resolution mapping of the spatial organization of a bacterial chromosome. *Science* **342**: 731–734.
- Lemieux JE, Kyes SA, Otto TD, Feller AI, Eastman RT, Pinches RA, Berriman M, Su XZ, Newbold CI. 2013. Genome-wide profiling of chromosome interactions in *Plasmodium falciparum* characterizes nuclear architecture and reconfigurations associated with antigenic variation. *Mol Microbiol* **90**: 519–537.
- Li X, Wang S, Li Y, Deng C, Steiner LA, Xiao H, Wu C, Bungert J, Gallagher PG, Felsenfeld G, et al. 2011. Chromatin boundaries require functional collaboration between the hSET1 and NURF complexes. *Blood* **118**: 1386–1394.
- Li L, Lyu X, Hou C, Takenaka N, Nguyen HQ, Ong CT, Cubenas-Potts C, Hu M, Lei EP, Bosco G, et al. 2015. Widespread rearrangement of 3D chromatin organization underlies polycomb-mediated stress-induced silencing. *Mol Cell* **58**: 216–231.
- Lieberman-Aiden E, van Berkum NL, Williams L, Imakaev M, Ragoczy T, Telling A, Amit I, Lajoie BR, Sabo PJ, Dorschner MO, et al. 2009. Comprehensive mapping of long-range interactions reveals folding principles of the human genome. *Science* **326**: 289–293.
- Lifshitz IM, Grosberg AY, Khokhlov AR. 1976. Structure of a polymer globule formed by saturating bonds. *J Exp Theor Phys* **44**: 855–860.
- Luger K, Mader AW, Richmond RK, Sargent DF, Richmond TJ. 1997. Crystal structure of the nucleosome core particle at 2.8 Å resolution. *Nature* **389**: 251–260.
- Lupiáñez DG, Kraft K, Heinrich V, Krawitz P, Brancati F, Klopocki E, Horn D, Kayserili H, Opitz JM, Laxova R, et al. 2015. Disruptions of topological chromatin domains cause pathogenic rewiring of gene-enhancer interactions. *Cell* **161**: 1012–1025.
- Maniatis T, Fritsch EF, Sambrook J. 1982. *Molecular cloning: a laboratory manual*. Cold Spring Harbor Laboratory, Cold Spring Harbor, NY.
- McCarthy DJ, Chen Y, Smyth GK. 2012. Differential expression analysis of multifactor RNA-Seq experiments with respect to biological variation. *Nucleic Acids Res* **40**: e288–e297.
- Mirny LA. 2011. The fractal globule as a model of chromatin architecture in the cell. *Chromosome Res* **19**: 37–51.
- Mizuguchi T, Fudenberg G, Mehta S, Belton JM, Taneja N, Folco HD, Fitzgerald P, Dekker J, Mirny L, Barrowman J, et al. 2014. Cohesin-dependent globules and heterochromatin shape 3D genome architecture in *S. pombe*. *Nature* **516**: 432–435.
- Mukhopadhyay S, Schedl P, Studitsky VM, Sengupta AM. 2011. Theoretical analysis of the role of chromatin interactions in long-range action of enhancers and insulators. *Proc Natl Acad Sci* **108**: 19919–19924.
- Mutskov VJ, Farrell CM, Wade PA, Wolffe AP, Felsenfeld G. 2002. The barrier function of an insulator couples high histone acetylation levels with specific protection of promoter DNA from methylation. *Genes Dev* **16**: 1540–1554.
- Naumova N, Imakaev M, Fudenberg G, Zhan Y, Lajoie BR, Mirny LA, Dekker J. 2013. Organization of the mitotic chromosome. *Science* **342**: 948–953.
- Niki Y, Yamaguchi T, Mahowald AP. 2006. Establishment of stable cell lines of *Drosophila* germ-line stem cells. *Proc Natl Acad Sci* **103**: 16325–16330.
- Nora EP, Lajoie BR, Schulz EG, Giorgetti L, Okamoto I, Servant N, Piolot T, van Berkum NL, Meisig J, Sedat J, et al. 2012. Spatial partitioning of the regulatory landscape of the X-inactivation centre. *Nature* **485**: 381–385.
- Pepinella S, Murphy KJ, Hayes JJ. 2014. Intra- and inter-nucleosome interactions of the core histone tail domains in higher-order chromatin structure. *Chromosoma* **123**: 3–13.
- Pope BD, Ryba T, Dileep V, Yue F, Wu W, Denas O, Vera DL, Wang Y, Hansen RS, Canfield TK, et al. 2014. Topologically associating domains are stable units of replication-timing regulation. *Nature* **515**: 402–405.
- Quinlan AR, Hall IM. 2010. BEDTools: a flexible suite of utilities for comparing genomic features. *Bioinformatics* **26**: 841–842.
- R Core Team. 2015. *R: a language and environment for statistical computing*. R Foundation for Statistical Computing, Vienna, Austria. <http://www.R-project.org/>.
- Ramírez F, Lingg T, Toscano S, Lam KC, Georgiev P, Chung HR, Lajoie BR, de Wit E, Zhan Y, de Laat W, et al. 2015. High-affinity sites form an interaction network to facilitate spreading of the MSL complex across the X chromosome in *Drosophila*. *Mol Cell* **60**: 146–162.
- Rao SS, Huntley MH, Durand NC, Stamenova EK, Bochkov ID, Robinson JT, Sanborn AL, Machol I, Omer AD, Lander ES, et al. 2014. A 3D map of the human genome at kilobase resolution reveals principles of chromatin looping. *Cell* **159**: 1665–1680.
- Rausch T, Zichner T, Schlattl A, Stutz AM, Benes V, Korbel JO. 2012. DELLY: structural variant discovery by integrated paired-end and split-read analysis. *Bioinformatics* **28**: i333–i339.
- Ricci MA, Manzo C, Garcia-Parajo MF, Lakadamyali M, Cosma MP. 2015. Chromatin fibers are formed by heterogeneous groups of nucleosomes in vivo. *Cell* **160**: 1145–1158.
- Schalch T, Duda S, Sargent DF, Richmond TJ. 2005. X-ray structure of a tetranucleosome and its implications for the chromatin fibre. *Nature* **436**: 138–141.
- Schneider I. 1972. Cell lines derived from late embryonic stages of *Drosophila melanogaster*. *J Embryol Exp Morphol* **27**: 353–365.
- Sexton T, Yaffe E, Kenigsberg E, Bantignova F, Leblanc B, Hoichman M, Parrinello H, Tanay A, Cavalli G. 2012. Three-dimensional folding and functional organization principles of the *Drosophila* genome. *Cell* **148**: 458–472.
- Shahbazian MD, Grunstein M. 2007. Functions of site-specific histone acetylation and deacetylation. *Annu Rev Biochem* **76**: 75–100.
- Shogren-Knaak M, Ishii H, Sun JM, Pazin MJ, Davie JR, Peterson CL. 2006. Histone H4-K16 acetylation controls chromatin structure and protein interactions. *Science* **311**: 844–847.
- Sinha D, Shogren-Knaak MA. 2010. Role of direct interactions between the histone H4 Tail and the H2A core in long range nucleosome contacts. *J Biol Chem* **285**: 16572–16581.
- Smeets D, Markaki Y, Schmid VJ, Kraus F, Tattermusch A, Cerase A, Sterr M, Fiedler J, Demmerle J, Popken J, et al. 2014. Three-dimensional super-resolution microscopy of the inactive X chromosome territory reveals

- a collapse of its active nuclear compartment harboring distinct Xist RNA foci. *Epigenetics Chromatin* **7**: 8.
- Sofueva S, Yaffe E, Chan WC, Georgopoulou D, Vietri Rudan M, Mira-Bontenbal H, Pollard SM, Schroth GP, Tanay A, Hadjur S. 2013. Cohesin-mediated interactions organize chromosomal domain architecture. *EMBO J* **32**: 3119–3129.
- Splinter E, Heath H, Kooren J, Palstra RJ, Klous P, Grosveld F, Galjart N, de Laat W. 2006. CTCF mediates long-range chromatin looping and local histone modification in the β -globin locus. *Genes Dev* **20**: 2349–2354.
- Thomas JO, Stott K. 2012. H1 and HMGB1: modulators of chromatin structure. *Biochem Soc Trans* **40**: 341–346.
- Tolhuis B, Blom M, Kerkhoven RM, Pagie L, Teunissen H, Nieuwland M, Simonis M, de Laat W, van Lohuizen M, van Steensel B. 2011. Interactions among Polycomb domains are guided by chromosome architecture. *PLoS Genet* **7**: e1001343.
- Ui K, Nishihara S, Sakuma M, Togashi S, Ueda R, Miyata Y, Miyake T. 1994. Newly established cell lines from *Drosophila* larval CNS express neural specific characteristics. *In Vitro Cell Dev Biol Anim* **30A**: 209–216.
- Van Bortle K, Nichols MH, Li L, Ong CT, Takenaka N, Qin ZS, Corces VG. 2014. Insulator function and topological domain border strength scale with architectural protein occupancy. *Genome Biol* **15**: R82.
- Vatolina TY, Boldyreva LV, Demakova OV, Demakov SA, Kokoza EB, Semeshin VF, Babenko VN, Goncharov FP, Belyaeva ES, Zhimulev IF. 2011. Identical functional organization of nonpolytene and polytene chromosomes in *Drosophila melanogaster*. *PLoS One* **6**: e25960.
- Vietri Rudan M, Barrington C, Henderson S, Ernst C, Odom DT, Tanay A, Hadjur S. 2015. Comparative Hi-C reveals that CTCF underlies evolution of chromosomal domain architecture. *Cell Rep* **10**: 1297–1309.
- Williamson I, Berlivet S, Eskeland R, Boyle S, Illingworth RS, Paquette D, Dostie J, Bickmore WA. 2014. Spatial genome organization: contrasting views from chromosome conformation capture and fluorescence in situ hybridization. *Genes Dev* **28**: 2778–2791.
- Wood AM, Van Bortle K, Ramos E, Takenaka N, Rohrbaugh M, Jones BC, Jones KC, Corces VG. 2011. Regulation of chromatin organization and inducible gene expression by a *Drosophila* insulator. *Mol Cell* **44**: 29–38.
- Zabidi MA, Arnold CD, Scherhuber K, Pagani M, Rath M, Frank O, Stark A. 2015. Enhancer-core-promoter specificity separates developmental and housekeeping gene regulation. *Nature* **518**: 556–559.
- Zhimulev IF, Zykova TY, Goncharov FP, Khoroshko VA, Demakova OV, Semeshin VF, Pokholkova GV, Boldyreva LV, Demidova DS, Babenko VN, et al. 2014. Genetic organization of interphase chromosome bands and interbands in *Drosophila melanogaster*. *PLoS One* **9**: e101631.

Received June 17, 2015; accepted in revised form October 26, 2015.



Active chromatin and transcription play a key role in chromosome partitioning into topologically associating domains

Sergey V. Ulianov, Ekaterina E. Khrameeva, Alexey A. Gavrilov, et al.

Genome Res. 2016 26: 70-84 originally published online October 30, 2015

Access the most recent version at doi:[10.1101/gr.196006.115](https://doi.org/10.1101/gr.196006.115)

Supplemental Material <http://genome.cshlp.org/content/suppl/2015/10/30/gr.196006.115.DC1>

References This article cites 90 articles, 18 of which can be accessed free at:
<http://genome.cshlp.org/content/26/1/70.full.html#ref-list-1>

Creative Commons License This article is distributed exclusively by Cold Spring Harbor Laboratory Press for the first six months after the full-issue publication date (see <http://genome.cshlp.org/site/misc/terms.xhtml>). After six months, it is available under a Creative Commons License (Attribution-NonCommercial 4.0 International), as described at <http://creativecommons.org/licenses/by-nc/4.0/>.

Email Alerting Service Receive free email alerts when new articles cite this article - sign up in the box at the top right corner of the article or [click here](#).

Affordable, Accurate
Sequencing.



To subscribe to *Genome Research* go to:
<https://genome.cshlp.org/subscriptions>
