The Role of Chromosome—Nuclear Envelope Attachments in 3D Genome Organization

I. V. Sharakhov1,2*, S. M. Bondarenko2, G. N. Artemov2, and A. V. Onufriev3,4

1Virginia Polytechnic Institute and State University, Fralin Life Science Institute, Department of Entomology, 24061 Blacksburg, VA, USA; E-mail: igor@vt.edu
2Tomsk State University, Laboratory of Ecology, Genetics and Environmental Protection, Tomsk, Russia
3Virginia Polytechnic Institute and State University, Department of Physics, 24060 Blacksburg, VA, USA
4Virginia Polytechnic Institute and State University, Department of Computer Science, 24061 Blacksburg, VA, USA

Received October 20, 2017
Revision received December 11, 2017

Abstract—Chromosomes are intricately folded and packaged in the cell nucleus and interact with the nuclear envelope. This complex nuclear architecture has a profound effect on how the genome works and how the cells function. The main goal of review is to highlight recent studies on the effect of chromosome—nuclear envelope interactions on chromatin folding and function in the nucleus. The data obtained suggest that chromosome—nuclear envelope attachments are important for the organization of nuclear architecture in various organisms. A combination of experimental cell biology methods with computational modeling offers a unique opportunity to explore the fundamental relationships between different aspects of 3D genome organization in greater details. This powerful interdisciplinary approach could reveal how the organization and function of the genome in the nuclear space is affected by the chromosome—nuclear envelope attachments and will enable the development of novel approaches to regulate gene expression.

DOI: 10.1134/S0006297918040065

Keywords: chromosomes, Drosophila, nuclear architecture, nuclear envelope, mosquitoes, modeling

Non-random 3D organization of the genome plays a crucial role in the regulation of gene expression, DNA double-strand break repair, and DNA replication [1-4]. Deep understanding of cell physiology will require elucidation of the structure-function relationships in the nucleus, which is a complex, interacting, and dynamic biological system. In eukaryotes, the genomic material — chromatin fiber — is packed in a complex, hierarchical manner [5, 6]. Unlike most proteins that adopt the same unique compact 3D shape in all cells, the conformational states of chromatin fibers are not nearly as compact or ordered. It is likely that these states are governed by a set of principles most of which are still poorly understood. In contrast to the unique native (folded) state of proteins, the state of folded chromatin in the cell nucleus is expected to depend on its interaction with the nuclear envelope (NE) [7, 8]. Experimental studies of polytene and regular diploid (non-polytene) nuclei in Drosophila and other eukaryotes have revealed at least two types of interactions in the nucleus: chromosome—NE (Chr—NE) attachments and chromosome—chromosome (Chr—Chr) contacts.

Organization and functional significance of Chr—NE attachments. Statistically significant high-frequency attachments exist between certain chromosomal loci and the NE. Several lines of evidence, including microscopy and genome-wide mapping studies using DamID (DNA adenine methyltransferase identification), show that in Drosophila and human nuclei, gene-poor and transcriptionally silent regions tend to form high-frequency contacts with NE [9-12]. In D. melanogaster, lamina-associated domains (LADs) usually span from 7 to 700 kb (average size, 90 kb) and bear on average seven genes per LAD [13]. Lamin binding is linked to a combination of several features including late replication, large intergenic region size, low gene expression level, and the lack of active histone marks [9, 12]. Depletion of lamin B (Dm0) in Drosophila cells causes visible detachment of chromosomal loci from the NE [14, 15]. In contrast, removal of

Abbreviations: Chr, chromosome; DamID, DNA adenine methyltransferase identification; FISH, fluorescence in situ hybridization; Hi-C, chromosome conformation capture; LAD, lamina-associated domain; NE, nuclear envelope; TAD, topologically associated domain.

* To whom correspondence should be addressed.
essentially all lamins in mouse embryonic stem cells does not have any detectable effect on the Chr–NE attachments [16] indicating that lamins are not required for the formation of LADs in every cell type of every organism. Many Chr–NE attachments are tissue-specific and, hence, may be responsible for the tissue-specific gene expression [1, 17, 18]. Chr–NE attachments regulate gene expression by repositioning a gene closer to the NE, i.e., to the region where transcription is usually repressed [1, 9, 14]. For example, in D. melanogaster somatic cells, testis-specific genes tend to be organized in clusters that are embedded in repressive nuclear peripheral compartments. Depletion of lamin B (Dm0) results in the derepression of these genes in somatic tissues together with their detachment from the NE [14]. Another consequence of lamin B depletion in the nucleus is misexpression of key developmental genes. Thus, detachment of the hunchback gene from the NE in D. melanogaster coincides with the reduction of its silencing and extension of the time window for neuroblast competence [15]. Moreover, the loss of lamin B in Drosophila fat body during aging leads to systemic inflammation and disruption of tissue homeostasis [19].

Organization of Chr–Chr interactions. Chromosomes fold in a characteristic way within a limited volume of the interphase nucleus thereby creating chromosome territories [11, 20–22]. Spatial chromosome organization is manifested in short-range and long-range intra-chromosomal interactions, as well as inter-chromosomal interactions. Long-range interactions between polytene chromosomes have been microscopically observed in Drosophila [23]. More recently, Hi-C (chromosome conformation capture) techniques [24–27] have been used to identify intra- and inter-chromosomal interactions in nonpolytenized nuclei. The Hi-C studies have led to the discovery of 3D structures that favor short-range, rather than long-range, chromosome interactions. These chromatin structures were named topologically associating domains (TADs) [28]. TADs are delimited by sharp boundaries defined by housekeeping genes and insulator-binding sites [29]; they correspond to chromosomal domains [30] and bands in polytene chromosomes [26, 31].

Functional significance of short-range Chr–Chr contacts. The functions of TADs depend on their structural organization that can be classified based on the chromatin type [2, 32]. Accordingly, several types of TADs can be recognized in Drosophila.

1) Active TADs belong to the most heterogeneous RED and YELLOW chromatin types. RED and YELLOW chromatin types represent housekeeping and tissue-specific genes, respectively, of the euchromatin. Active TADs have binding sites for many chromatin factors and are marked by multiple histone modifications, such as H3K4me, H3K36ac, H3K79ac [33]. They represent a highly accessible environment and participate in long-range interactions that favor coregulation of distant genes by increasing the local concentration of factors required for their transcription [34]. Active TADs are located predominately in the inner part of the nucleus and preferentially contact other active TADs, rather than inactive ones [2].

2) Polycomb-repressed TADs correspond to the BLUE chromatin, which is marked by the Polycomb group (PeG) proteins and H3K27me3. Polycomb-repressed TADs form long-range intra- and inter-chromosome contacts that occur in regions of high Polycomb concentration, the so-called Polycomb foci visible under a microscope [35]. Depletion of CTCF in D. melanogaster affects long-range interactions between homeotic complexes located in different Polycomb-repressed TADs, without affecting their local clustering [36].

3) Lamin-repressed TADs correspond to the BLACK chromatin (intercalary heterochromatin), marked by lamin and SUUR (suppressor of under-replication). These TADs might colocalize with lamin-associated domains (LADs). More than 5000 genes are located inside LADs and interact with lamin in Drosophila [13]. Interestingly, these genes have a very low transcriptional activity and are not associated with any active histone or heterochromatin marks. Repression at the nuclear lamina appears to be weaker compared to more specialized silencing mechanisms, such as in the Polycomb–repressed TADs [14, 37, 38]. This correlates well with the fact that H3K27me3 histone-associated Drosophila LADs contain less genes than Polycomb-repressed domains [32].

4) GREEN chromatin includes pericentromeric heterochromatin and can contain heterochromatic TADs. GREEN chromatin is marked by histone lysine methyltransferase SU(VAR)3-9, H3K9me2/3, and heterochromatin protein 1 (HP1) [32]. Average transcriptional activity is lower in heterochromatic domains than in the active euchromatic TADs but higher than in Polycomb- and lamin-repressed TADs. Long-range interactions between heterochromatic foci and euchromatic genes can drive tissue-specific gene silencing due to the proximity effect [39].

Regions between TADs (inter-TADs) correspond to decompacted interbands of polytene chromosomes in Drosophila; they consist of active chromatin and constitutively transcribed (housekeeping) genes [26].

Functional significance of long-range Chr–Chr contacts. Most active or potentially active euchromatin TADs are located in the inner part of the nucleus where they form long-range contacts with other active or potentially active TADs, resulting in the formation of larger active transcriptional factories [2]. Transcriptional factories represent a highly accessible environment and are formed by the interactions that favor coregulation of distant genes by increasing the local amounts of factors required for their transcription [34]. In contrast, PeG-associated TADs participate in long-range intra- and inter-chromo-
omal contacts and form strongly repressed larger hubs with high PcG concentration [40]. These large PcG hubs are visible under a light microscope [35]. Such long-range Chr–Chr interactions are usually tissue-specific because they occur between genes sharing either PcG-mediated repression [41, 42] or activation by transcription factors [43, 44] specifically in cell types where the regulation is mediated.

The interplay between Chr–NE attachments and Chr–Chr contacts. The driving forces that bring together long-range chromosomal contacts are not well-understood. It is possible that several active TADs form transcription factories, because transcriptional complexes cooperatively bring chromosomes toward a more internal position in the nucleus. In contrast, chromosome loci depleted in active TADs would be pushed to a more peripheral location. However, recent data indicate that chromosomal loci have the peripheral location because of specific interactions with the NE [1, 9, 13, 45]. Therefore, the internal positioning of active chromatin could be due to a passive force, resulting in preferential positioning of specific chromosomal loci at the nuclear periphery. Indeed, recent modeling results indicate that the number of Chr–NE attachments may affect the frequency of Chr–Chr interactions in D. melanogaster. Introduction of Chr–NE attachments decreases the overall probability of Chr–Chr contacts in trans configuration but increases the frequency of Chr–Chr interactions in cis configuration [46]. Attachment to the NE could constrain chromatin condensation and prevent chromosome territories from freely diffusing and mixing in the entire nuclear volume, which would be critical for maintaining the delicate 3D genome architecture. Experimental studies demonstrated the role of Chr–NE attachments in the maintenance of nuclear architecture and separation of chromosome territories [47, 48], as well as in the reduction of recombination between distant genomic loci and lowering the chances of non-allelic recombination and double-strand break repair [1, 49]. It is possible that Chr–NE attachments regulate expression of genes located not only at the nuclear periphery but also genes involved in long-range Chr–Chr interactions. Moreover, spatial networks of inter-connected genes may be a result of chromosome folding influenced by Chr–NE attachments. The functional significance of such configurations is that the long-range contact between two loci increases concentrations of common regulatory factors. When a factor dissociates, it is more likely to be retrapped by the cluster formed by long-range contacts than to diffuse away to another single TAD [50]. Also, interacting TADs may ensure robust gene regulation by constraining genes to the areas inaccessible for long-range interactions [51].

Studies of Chr–NE attachments using polytene chromosomes. The link between Chr–NE attachments and 3D genome organization can be uncovered with the help of giant (polytene) chromosomes in Drosophila and mosquitoes. Each of these chromosomes contain 1024 banded replicas of the corresponding “regular” interphase chromosome and, thus, the entire chromosome set of a single nucleus becomes readily visible under a light microscope [22, 52, 53]. Drosophila polytene chromosomes have the same banding patterns as nonpolytene chromosomes, as it was found by analyzing the data on localization of interband- and band-specific proteins (modENCODE) and genomic positions of interbands and bands [54–57]. Recent Hi-C studies have revealed that TADs, and the regions between them (which are conserved between polytene and diploid cells) correspond to the bands and interbands of polytene chromosomes in Drosophila [26, 31]. Moreover, polytene chromosomes and nonpolytene chromosomes have similar sets of Chr–NE attachments, as seen by comparing DamID data obtained in embryonic cells [9] with light microscopy data obtained in salivary glands [11, 17, 58]. Belyaeva et al. [55] compared distribution of LADs [13] and localization of 60 regions of intercalary heterochromatin in polytene chromosomal arms and found complete overlap for four regions of intercalary heterochromatin. Most of the intercalary heterochromatin regions have shown partial overlap with LADs, six regions have shown no overlap with any of the LADs, and one region of the intercalary heterochromatin encompassed five separate LADs. A microscopic study of polytene nuclei offers important methodological advantages. The large size of a polytene nucleus (30–35 µm) facilitates discrimination between the peripheral position of a chromosomal locus and its actual attachment to the NE. Characteristic chromatin fibers that physically penetrate the NE, as well as tight bends and strong twists, can often be seen where a chromosomal locus attaches to the periphery [11, 23, 59, 60]. Spatial positions of multiple regions in a nucleus can be determined by reading their banding pattern with the help of oligopaint probes [61]. Microscopy data can be converted into R(x, y, z) curves corresponding to each chromosome, and the corresponding nuclear architecture can be computationally simulated. This level of structural completeness is currently not possible in the studies of regular nonpolytene chromosomes.

Studies of 3D organization of chromosomes in mosquitoes. Previous studies have shown tissue-specific features of chromosome attachments to the NE in various organisms including malaria mosquitoes [62–65]. Early works on mosquitoes were among the first to demonstrate substantial differences in the spatial organization of polytene chromosomes in the nuclei of salivary gland cells, malpighian tubules, and ovarian nurse cells in malaria mosquitoes of the Anopheles maculipennis complex [64, 66]. The chromosomes had typical Rabl organization in salivary glands and malpighian tubules. In contrast, the ovarian nurse cells had no single chromosome (and Rabl organization), and some of their chromosomes attached to the NE with the pericentromeric beta-heterochro-
matin or specific loci of intercalary heterochromatin [67]. Another study focused on tissue-specific features of chromosome architecture in closely related malaria mosquitoes [63]. This study showed that the X chromosome–NE attachment is stronger in nurse cells than in salivary glands of An. messeae. Microdissected DNA probes corresponding to the regions of the X chromosome (2BC) and 3R chromosome (32D) attachment to the NE were hybridized with intact nuclei of nurse cells, salivary gland cells, follicle epithelium cells, and imaginal disc cells in 3D-FISH (3D fluorescence in situ hybridization) experiments. Only salivary gland cells and follicle epithelium cells had no statistical differences in the interposition of 2BC and 32D. On average, the X chromosome and chromosome 3R were located closer to each other in somatic tissue cell than in nurse cells. In the nuclei of imaginal disc cells, the distance between the X and 3R chromosomes was intermediate between those observed in the nurse cells and somatic cells. A more recent study conducted a thorough analysis of tissue-specific dynamics of the X chromosome in follicular epithelium and nurse cells of An. atroparvus ovaries [65]. A fluorescently labeled X chromosome painting probe was hybridized with formaldehyde-fixed ovaries of mosquitoes using the 3D-FISH method, and the results were processed with the TANGO software for chromosome spatial organization analysis [68]. It was shown that the volume and position of the X chromosome are tissue-specific. Unlike nurse cell nuclei, the growth of follicular epithelium nuclei was not accompanied with the proportional growth of the X chromosome. The dynamics of the X chromosome–NE attachment regions was also tissue-specific (Fig. 1). During the nucleus growth in nurse cells, the polytene X chromosome moved toward the periphery, whereas in follicular epithelium, the X chromosome telomere moved from the nuclear envelope to the intranuclear space [65]. It is possible that Chr–NE attachments facilitate or prevent tissue-specific Chr–Chr contacts, where actively transcribed genes colocalize and exchange transcription sites.

In addition to tissue-specific differences, essential differences in the nuclear architecture among seven species of the An. maculipennis complex have been found [69]. These closely related species of malaria mosquitoes, some of which are homosequential (e.g., An. maculipennis and An. subalpinus), had different patterns of Chr–NE attachments. Spatial organization of chromosomes in ovarian nurse cell of experimental F1 hybrids (An. mac-
ulipennis × An. subalpinus and An. sacharovi × An. matriotinus) confirmed species-specific chromosome organization in the parental species. However, despite species-specific Chr–NE attachments, no differences in the 3D organization of nurse cell chromosomes between An. messelae and An. atroparvus have been observed [63].

More detailed studies are needed to address species-specific aspects of the nuclear architecture. Such studies will be essential for assessing the possibility of using 3D genome organization as a phylogenetic marker for understanding the process of mosquito speciation.

Possible role of NE–chromosome attachments in genome rearrangements. Cytogenetic and genomic studies in various organisms have provided evidence for nonuniform distribution of chromosomal rearrangements and breakpoints [70–72]. Chromosomal rearrangements, which are believed to be affected by nuclear architecture, are linked to cancer [73]. Unequal chances of chromosome arms and lineages in generating and fixing inversions have serious consequences for differential adaptive and evolutionary plasticity of organisms. Therefore, it is important to understand the forces that govern the structural malleability of genomes. Analysis of nuclear architecture in diverse systems revealed a general principle: because of nonrandom nuclear organization, certain loci are closer together than other loci and, therefore, are more likely to interact and generate rearrangements [74, 75]. Computer modeling has been instrumental in helping to understand the link between chromosome folding and chromosome rearrangements. Such models have demonstrated that specific loci are much closer to each other in the nuclear space than it could be expected based on their genomic locations, which explains their frequent involvement in chromosomal rearrangements [76, 77]. A study of the 3D organization of the yeast genome revealed that location of inter-chromosomal contacts correlates with chromosomal fragile sites where chromosomal breakpoints occur in evolution [78]. Another study detected rearrangement events in Saccharomyces species and observed increased frequency of the rearrangements between regions with higher frequency of contacts [79]. These results show how 3D chromosomal organization can influence evolutionary events. Chr–NE attachments might indirectly affect the frequency of evolutionary rearrangements by changing the probability of long-range Chr–Chr interactions.

Computational modeling of the 3D organization of chromosomes: effects of Chr–NE attachments. 3D organization of eukaryotic genomes is extremely intricate and much more complex than organization of proteins or protein complexes. At the same time, all experimental methods available to study various aspects of 3D organization of eukaryotic genomes are inherently limited in their resolution, both spatial and temporal. Compared to structural biology, there is an additional complexity: both Chr–NE attachments and Chr–Chr contacts appear to be stochastic to some degree, as evidenced by Dam1D and Hi-C experiments [80–82]. All these factors combined make computational modeling an essential tool to complement the experimental studies. The use of computational modeling in close integration with experiments can generate verifiable hypotheses that otherwise would be very difficult or impossible to arrive at. Equally importantly, models can rationalize experimental observations and suggest explanations of the basic mechanisms of nuclear architecture. While computation is becoming commonplace in this field [24, 26, 83–85], Chr–NE attachments have received relatively little attention in nuclear architecture modeling studies compared to the Chr–Chr interactions [26, 86, 87] and other aspects of genomic architecture. Although a few published studies included Chr–NE attachments in their models [86, 88, 89], they provided no comprehensive analysis of the role of these attachments in the regulation of Chr–Chr contacts. Also, results of computational modeling of yeast genome architecture, which may lack chromosome territories and typical TADs [90], cannot be readily extrapolated to the 3D genome organization of higher eukaryotes. Recent simulation study [91] that generated thousands of models of Drosophila polytene nuclei (Fig. 2), allowed researchers to determine an objective threshold of statistical significance for Chr–NE attachments originally based on only 24 experimental nuclei [11]. This simulation work identified 33 additional positions of statistically significant Chr–NE attachments in the experimental nuclei [11]. Of the total 48 Chr–NE attachments, 34 were intercalary chromatin, 3 were euchromatin, and 11 Chr–NE attachments displayed some properties of intercalary chromatin by being late replicating regions. These results suggest that affinity for the NE can change gradually, with the highest affinity almost exclusively possessed by intercalary chromatin, and the lowest affinity for the NE being mostly, but not exclusively, a property of euchromatin. The newly identified Chr–NE attachments

![Fig. 2. Representation of a simulated model nucleus (right) compared to experiment (left) for the D. melanogaster salivary glands [91].](image)
are real because they were confirmed in an independent DamID study [9]. The model of Drosophila polytene nucleus has been used to investigate if and how the presence and the number of Chr–NE attachments affect chromosome territories and Chr–Chr contacts. Model nuclei with more numerous Chr–NE attachments formed more distinct chromosome territories, and their chromosomes intertwined less frequently. They also displayed higher probabilities of intra-chromosome and inter-arm contacts. In contrast, inter-chromosome and inter-arm contacts become less common in nuclei with more numerous Chr–NE attachments [46]. Statistical models that generate snapshots of configurational ensembles have succeeded in reconstructing the folding of D. melanogaster chromosomes through the integration of experimental data on Chr–Chr and Chr–NE interactions [85]. However, dynamic models are needed to predict changes in Chr–NE and Chr–Chr interactions in the interphase nucleus.

Acknowledgments

The review of experimental and computational studies of Drosophila, yeast, and mammalian genomes was supported by the National Science Foundation grant MCB-1715207 to IVS and AVO. Our recent research on mosquitoes was supported by the grant 15-14-20011 to IVS from the Russian Science Foundation.

REFERENCES


