Large-scale periodicity of nucleosome positioning signal in pericentric regions of chromosomes (Drosophila melanogaster)

V.B. Fedoseyeva* and A.A. Alexandrov

Institute of Molecular Genetics, Kurchatov Sq. 2, Moscow 123182, Russian Federation

Communicated by Ramaswamy H. Sarma

(Received 24 July 2013; final version received 10 September 2013)

Nucleosome positioning signal (NPS) in heterochromatin is not uniform. We suggest the analysis of its heterogeneity by correlation with periodic function (analog of Fourier analysis). It was established the periodical repetition of the nucleosome clusters of large size in pericentric regions in a discontinuous manner. In the 3L pericentric region, it was revealed the domination of 78–85 kbp wavelength in the correlation coefficient profile and also strong presentation of 50 kbp signal. In further to centromere position, the 69 kbp value strongly dominates as well as the 50 kbp value in the closest proximity. In addition to the long wavelength signals, there are plenty of short wavelengths signals especially in the closest vicinity to centromere. In some positions throughout pericentric region of 2L chromosome, there are two sizes of repeated intermingled correlation signals (50, and 75 kbp) with dominating value of 75 kbp in proximity and 50 kbp distantly to centromere, the situation for 2R is analogous. Some genes with long introns support these quantitative characteristics of NPSs and to some extent their dominating character in each region. The characteristic repeat periods for 3L pericentric region coincide with the distances between heterochromatin epigenetic mark clusters and their distribution throughout this region for fly embryos, larvae, and some cell lines.

Keywords: heterochromatin; nucleosome positioning; peri-centromere

Introduction

Earlier, some evidences were presented that pericentric heterochromatin is not uniform in some relations: the effect of variegation when gene is inserted in its regions (Schotta, Ebert, Dom, & Reuter, 2003), influence on the expression level by environmental sequences or their features (Howe, Dimitri, Berloco, & Wakimoto, 1995). Coexistence of active genes (light, concertina, AGO3, and others) and silencing region (Sun, Cuaycong, & Elgin, 2001), plenty of simple and different classes repeats also evidence to the complexity of heterochromatin. Characteristic feature of pericentric chromatin is the specific nucleosome modifications H3K9me2 (Noma, Allis, & Grewal, 2001; Stewart, Li, & Wong, 2005; Verschure et al., 2005) and interactions of nucleosome assembly with heterochromatin proteins (HP1 and other variants) that recognize this epigenetic signature (Jacob & Khorasanizadeh, 2002; Nielsen et al., 2002). Heterogeneous profile of HP1 binding data within some regions (De Wit, Greil, & van Steensel, 2007), nucleosome modification data (dimethylation of lysine in ninth position of H3 histone) (Riddle et al., 2011) evidence about the heterochromatin discontinuity. Transition induction to silencing by small RNA cognate to retrotransposons and multi-protein complex (members of Argonaute family, HP1, histone H3 specific methyltransferases and others) is shown for heterochromatin regions (Fagegaltier et al., 2009; Pal-Bhadra et al., 2004). However, region length of silencing spread further from the initiation point is limited by 10 kbp at least for four chromosomes (Haynes, Gracheva, & Elgin, 2006). In pericentric heterochromatin, this limitation probably is diminished due to high density of cognate small RNA (Aravin et al., 2003; Reinhart & Bratel, 2002). The intricacy persists of how dense heterochromatin causes silencing of transposable elements but permits expression of embedded genes. In heterochromatin, pericentric 10-nucleosome – size arrays were observed in micrococcus nuclease (MNase) digestion compared to 5- or 6-nucleosome size as the largest fragment observed for the genome as a whole (Sun et al., 2001). Knowledge of the precise nucleosome location and its analysis in this region is essential elucidation of the context for the various features of pericentric heterochromatin and its influence on chromosome
architecture. We possess the opportunity to map the nucleosome occupancy signal along the DNA in large scale and analyze its features. We used previously published algorithm of the nucleosome positioning calculations on the bases of two hierarchy ordered trinucleotide matrices with adapted usage of them to AT-rich or GC-rich fragments (Fedoseyeva & Alexandrov, 2007).

Materials and methods

Nucleotide sequences were used from GenBank sequences (Release5) for 2L chromosome beginning with 21,800 kbp up to the satellite repeats, that is 1200 kbp by length, as well as for 2R beginning from 1 up to 1200 kbp, for 3L chromosome the beginning point is 22,730 kbp and up to the centromere, that is near 1800 kbp by length, for 3R from 1st nucleotide up to 600 kbp. Calculation of nucleosome positioning signal (NPS) in conventional units was performed according to the method (Fedoseyeva & Alexandrov, 2007) by using Turbo Pascal calculating program. Initial value is the maximum value of NPS represented approximately for Turbo Pascal calculating program. Initial value is the averaging the method (Fedoseyeva & Alexandrov, 2007) by using

In Figure 2, the CCS data are presented throughout the nucleotide fragment of 2R arm starting with 1st nucleotide up to 1200 kbp. Pericentric heterochromatin part may be divided into 4 portions in accordance with our data: closest, proximal, middle, and distant. The closest to centromere region is characterized by 13, 16, and 35 kbp (peak IV) wavelength peaks (Figure 2(A)–(C)). Proximal portion lasts up to 550 kbp and is characterized by 37–45 and 80 kbp peaks (Figure 2(B),(C), III, II). Distal region includes 55 kbp peak and to a smaller extent 80 kbp one (Figure 2(A)–(C), peak I, II). Further, we compared the distant with proximal portions in the 600 kbp windows (Figure 2(E)). As a whole, in the range of long wavelengths, the distant portion is predominantly characterized by 55–60 kbp (Figure 2(E), peak I), the proximal portion by 45 kbp (Figure 2(E), peak III) as well as by 74–80 kbp (Figure 2(E), peak II) peaks. The comparison of two genes with large introns enriched by sequence repeats reveals their strong influence on the corresponding CCS profile each in its region, distant and proximal (Figure 2(D),(E)). CCSs for intermediate portion are of sparse values for long wavelength and are

Results and discussion

Discontinuous periodicity of signals in pericentric heterochromatin region

In the Figure 1, the profiles are suggested for the coefficients of correlation signal (CCS) between NPS and sinusoidal function as the dependence of wavelength. The cluster size assessment of the nucleosome occupancy corresponds to the half of sinusoidal wavelength with maximal correlation value. The scan was suggested for different positions of sliding windows (200, 300, 400, and 500 kbp by length) (Figure 1(A) a–u, (B) a–l, (C) a–i, (D) a–h) bottom up with 50 or 100 kbp window displacements throughout the total nucleotide fragment (1200 kbp). Initial nucleotide corresponds to 21,800,001 bp of 2L (D.m). There are some regions corresponding to long wavelengths with significant correlation coefficient values in pericentric heterochromatin area. In this heterochromatin zone, according to our observations, it is relevant to distinguish the closest proximity, proximal, intermediate, and distant regions concerning the centromere. The nearest to the centromere portion is characterized by 38 kbp (Figure 1(A), u, peak III), and by 60 kbp (Figure 1(A), t peak IV), proximal portion by 45–50 (peak II) and with dominating 80 kbp (peak I) peaks. For distant portion, the domination is revealed approximately for 50 kbp wavelength (Figure 1(B), c, e) and there is also the fuzzy peak V (equal 60 kbp by wavelength approximately, Figure 1(A) d). Intermediate region has weakly expressed spectral characteristics. It should be noted that in dominating of 80 kbp signal, the significant role belongs to the CG40006 gene (~150 kbp Figure 1(E)) with short exons and long introns enriched of numerous sequence repeats. As is shown in the Figure 1(F), the comparison of the distant with proximal regions (each 600 kbp by length) reveals the distinct domination of 80 kbp in the proximity to the centromere. As is shown in Figure 1(D), there are some fuzzy peaks (VI peak and similar) in the long wavelengths’ region corresponding to the distances between the higher NPS regions. The histogram of the average NPS level (Figure 1(G)) also demonstrates non-uniform profile, at least two regions of higher NPSs should be distinguished.
Figure 1. Periodicity property of NPS in large-scale for pericentric region of 2L chromosome. (A) Correlation coefficient between NPS and sinusoidal function vs. wavelengths using 200 kbp window lengths, 50 kbp pace from (a) to (u). (a) Initial point is 21,800 kbp, end point 22,000 kbp. (b) from 21,850 to 22,050 kbp and so on. (B) The same as in (A), but 300 kbp window length, 100 kbp pace (a–g), 50 kbp (h–l). (C) 400 kbp window length; (D) 500 kbp window length; (E) CCS profile for CG40006, ~150 kbp window length. (F) 600 kbp window length, black marker – 1st nucleotide beginning point, light – beginning point is 600 kbp of current scale; (G) Histogram of average NPS for 200 kbp windows as in (A), N – window displacement number; (H) Scheme of region >21,800 kbp. Scale, genes, and main classes of repeats (Line, LTR, DNA, Simple, Low complexity) are presented for Simple repeats – 190, 690, 1190, 1690; AT-rich repeats – 210, 710, 1210, 1710; Line – 230, 730, 1230, 17, 030; DNA repeats – 270, 770, 1270, 1770; Pao – 290, 790, 1290, 1790; Gypsy. LTR – 310, 810, 1310, 1810; MDG – 320, 820, 1320, 1820, copia – 330,830,1330,1830; satellite –1850 at conventional units level (Y axis).
Figure 2. Periodicity property of NPSs in large-scale for pericentric region of 2R chromosome. (A) Correlation coefficient between of NPS and sinusoidal function vs. wavelength using 200 kbp sliding window, 50 kbp pace from (a) to (u). (a) Initial point is 1st bp of 2R chromosome, 200 kbp end point, (b) from 50 to 250 kbp and so on. (B) The same as in (A), but 300 kbp window length, 100 kbp pace (a–i). (C) The same as in B, but 400 kbp window length. (D) Black marker – CCS for Ir41a, light – for Gprk1; (E) 600 kbp window lengths, light marker – 1st nucleotide beginning point, black one – beginning point is 600 kbp; (F) Histogram of average NPS for 200 kbp window length as in (A), N – window displacement number; (G) Scheme of region >1 bp of 2R. Scale, genes, and main types of repeats (Sine, Line, LTR, DNA, Simple, Low complexity) are presented from genome.uscs.edu server. Simple repeats – 190, 790, 1390, 1990; AT-rich repeats – 210, 810, 1410, 2010; Line – 230, 830, 1430, 2030; DNA repeats – 270, 870, 1470, 2070; Pao family – 290, 890, 1490, 2090; Gypsy.LTR – 310, 910, 1510, 2110; MDG – 320, 920, 1520, 2120; LTR excluding gypsy – 330, 930, 1530, 2130; Pao family – 290, 890, 1490, 2090 c.u. (Y axis). Beginning point in (a) 1st nucl., (b) 300 kbp, (c) 600 kbp, (d) 900 kbp (X axis).
presented by 13, 25, and 36 kbp (Figure 2(A) j–m) peaks of intermediate wavelengths that correspond to the small or medium genes and large inter-gene spaces. Another approach is the assessment of NPS average values in sliding window (200 kbp length) for each window displacement throughout the whole DNA fragment (1200 kbp) (histogram in Figure 2(F)). It evidences to the favor of the heterochromatin heterogeneity due to wave character of NPS level. Approximately, low average values coincide with genes possessing long introns and high average values of NPS coincide with low density of small genes along with large inter-gene spaces. In the scheme (Figure 2(G) a–d), some information pool and NPS are presented. It makes possible to assess the difference between proximal, distal, and middle portions and compare these data with peculiarities elucidated above.

For 3L chromosome portion beginning from 22,730 kbp up to centromere (total 1800 kbp), it should be noticed the presence of euchromatin region besides heterochromatic one. The euchromatin features are well discernible (low density of different repeats, dense presence of short and middle length genes, short inter-gene spaces) in the first 200 kbp fragments as shown in Figure 3(A) (a–d), Figure 3(B) (a–b), Figure 4(A). There are multiple insignificant peaks in the specter region up to 32 kbp wavelengths, minimal signal if any in the range of larger wavelengths according to CCS profiles. This spectral interval boundary approximately coincides with epigenetic border, e.g. abrupt strong enrichment of H3K9me2 marks as well as repeat density increase which last in further movement to centromere (Riddle et al., 2011).

There are strong dominating signals corresponding to 78–85 kbp wavelengths (Figure 3, peak I) and significant CCSs corresponding to 50 kbp approximately (Figure 3, peak II) from 200 kbp up to 1200 kbp of the current scale in such order towards centromere (Figure 3(A) f–t and Figure 3(B) c–k, informative features in Figure 4(A–D)). The dominancy of the peak I compared to the others

Figure 3. Periodicity property of (NPS) in large scale for pericentric region of 3L chromosome (1800 kbp). (A) Correlation coefficient between of NPS and sinusoidal function vs. wavelength using 300 kbp sliding window bottom up, 50 kbp pace. Initial point is 22,730 kbp, end point – 24,530 kbp. (B) The same as in A, but 400 kbp sliding window and 100 kbp pace; (C) CCS profile for the closest to centromere window (300 kbp); (D) CCS profile for fragment 1050–1350 kbp; (E) Averaged CCS profile for 300 kbp windows (50 kbp pace) throughout 200–1200 kbp fragment of current scale. (F) Profile for 1000 kbp window throughout the same fragment as in (E); (G) Histogram of average NPS values for 300 kbp windows, 50 kbp pace as in (A).
was justified by the averaging realized for spectral profiles from (f) to (t) passages (Figure 3(A), 300 kbp window length) throughout 1000 kbp fragment. This generalized profile of CCS is depicted in Figure 3(E). In Figure 3(F), CCS profile is depicted for 1000 kbp window length realized for the same boundary points. It should be concluded about the noticeable quality of NPSs phasing from the fact of peak I domination even for 1000 kbp window length. There is a distinct strong signal with 69 kbp wavelength (Figure 3(D), peak III in the 1050–1500 kbp range of current scale) in further movement towards the centromere, and also the distinct peak corresponding to 50 kbp in this wave range (Figure 3(A)–(C), peak IV, in the 1550–1750 kbp interval of current scale) is in closeness to centromere.

In the histogram (Figure 3(G)), the average values of NPS in conventional units are presented and calculated for each sliding window location in passage from (a) to (b1) (Figure 3(A)). It may be observed the crests and troughs that reflect non-uniformity of heterochromatin influenced by specific types of transposon elements as well as gene locations.

In Figure 4(A)–(F), structural features of total fragment are depicted in linear manner from 22,730 kbp up to 24,543 kbp of 3L arm. By simple viewing NPSs from the fragment beginning to the end, it is possible to verify the details of periodical properties detected above and comparative appraisal of nucleosome occurrence in genes are encountered. By viewing the scheme of genes, it should be noted that the transcription direction of almost all genes with long introns is directed to the centromere, CG40470 and jim genes are the exceptions with the opposite direction. This notion is essential as for fission yeast (Lengronne et al., 2004), where the transcriptional directions are tightly connected with site occupation of one of the cohesion proteins.

For the different repeats represented at various scheme levels, it should be noted the abundance and to

Figure 4. Scheme of region >22,730 kbp of 3L chromosome, scale, NPS (c. u. interval), genes and main classes of repeats (Line, LTR, DNA, Simple, Short, and Long of low complexity) are presented from ncbi.nlm.nih.gov databases. (A) From 22,730 to 23,030 kbp; (B) From 23,030 to 23,330 kbp; (C) From 23,330 to 23,630 kbp; (D) From 23,630 to 23,930 kbp; (E) From 23,930 to 24,230 kbp; (F) From 24,230 kbp to 24,543,550 bp. Main classes of repeats (Line, LTR, DNA, Simple, Low complexity) are presented at different levels: (A)–(F) Simple repeats – 190, 690, 1190; AT-rich repeats – 210, 710, 1210; Line – 230, 730 in (A,B); Line (Penelope family) – 230, 730,1230 (C)–(F); Line (Jockey, CR1, R1families) – 360, 860, 1360 (C)–(F); DNA repeats – 250, 750, 1250 (A)–(F); LTR (Gypsy, Copia, Pao families) – 295–315, 795–815, 1295–1315 (A)–(F); MDG – 320, 820, 1320 (A)–(F) at c. u. level for (A–F) sections.
some extent interchange of two main types: retrotransposon elements (Gypsy family framed by Long Terminal Repeats, Copia, Pao and others) and Line Class (Jockey, Penelope, CR1, R1 families) as well as the presence of DNA Class, short or long simple and AT-rich of low complexity repeats. Regions of high NPS value are in accordance with high-repeats density. However in detail, it is not possible to detect a strong correlation between each repeat type density and the level of NPSs with the exception of AT-rich short and long repeats that has a positive and negative impact on NPS level accordingly.

If to transfer this information especially on the geometrical proportions of chromosome architecture, then in the case of 3L, these dominating wavelengths (78–85 kbp) in distant heterochromatin area most probably correspond to circumference of cylinder as a chromosome model in simple case of one-peak local domination. Probably this representation concerns G2 and pro-metaphase stages in cell cycle. Half values of these wavelengths approximately correspond to the nucleosome occupancy cluster sizes and the dominating wavelengths by themselves correspond to the repetition period of clusters. In movement towards the centromere, the circumference length gradually diminishes from 78–85 kbp (Figure 3(B) j–k, peak I throughout 500 kbp fragment, we chose the peak of maximum wavelength in this range), further 69 kbp (Figure 3(B) m, Figure 3(D), peak III, throughout 300 kbp fragment) and at last to 50 kbp wavelength (Figure 3(B) o, Figure 3(C), peak IV, throughout 200 kbp fragment). This gradual diminishing of dominating wavelengths may be indicative of a tendency to change the chromosome cylinder circumference towards the centromere. In juxtaposed to centromere position there is CCS peak corresponding by wavelength to 8 kbp approximately, e.g. clusters of nucleosome occupancy of 4 kbp by size periodically occurring in the 8 kbp interval (for 1740–1775 kbp region in the current linear scale). It may be that these nucleosome occupancy repeats corresponding to short wavelength signals are connected with kinetochore functioning.

**Influence of genes with long introns**

In Figure 5(A), the CCS profiles are presented for separate genes with long introns located in different heterochromatin regions of 3L chromosome arm discussed above. They have strong influence definite in each location on the whole profile, especially due to their long introns full of various repeats. It is noticeable that upon the elimination of two genes AGO3 and CG70470 from 1000 kbp fragment (the central portion of pericentric heterochromatin) the peak I (82 kbp) loses to some extent the dominant property in specter profile as shown in Figure 5(B), b compared to Figure 5(B) a. The same diminishing of main signals domination are also observed for 2L pericentric chromosome region upon the deletion of CG40006 gene with long introns (compare of peak I in CCS specters in Figure 1(E),(B),(C) and Figure 1 Suppl). This observation favors to the significance of the genes with long introns filled with appreciable NPS density for periodical features of NPS distribution.

**Comparison with epigenetic marker and HP1 localization data**

It is noticeable to compare the NPS profile of our data with H3K9me2 modifications’ enrichment in constitutive heterochromatin and euchromatin-heterochromatin junction in high-resolution manner (Yasuhara & Wakimoto, 2008). Due to detailed comparison, we could notify the coincidence of the profiles of two types (NPS and H3K9me2 signature) to a significant extent for distinct genes such as CG40006, light, CG17018, Nipped-A and their environments as well as for euchromatin-heterochromatin junctions for 2L and 3L chromosomes. Another suggested opportunity is the comparison NPS data with detailed H3K9me2 mark profile of 3L pericentric region for different fly growth stages and cell lines (Riddle et al., 2011). Especially for embryos and larvae and partially for fly heads, these marks profiles are characterized by clusters of signal enrichment with inter-distances reminding the periodicity features in NPSs. Cell lines, the distinct cluster repetitions are in some cases less discernible. Profile of the repeats’ density also displays some wavy character (Riddle et al., 2011) but without strong coherence with epigenetic marks or NPS/CCS periodicity according to the data presented here.

As bipolar properties are ascribed now to the functioning of heterochromatic protein 1 (HP1) by its property to participate in silencing and simultaneously
have expression effects on genes depending on their lengths (De Wit et al., 2007; Lundberg, Stenberg & Larsson, 2013), it was of interest to compare NPS and HP1 occupancy were presented only for restricted regions in 2L pericentromere, preferentially coincide with NPS data (It, cta genes, 21,800–22,200 kbp portion of 2L, 0–1200 kbp of 2R arms).

The higher nucleosome occupancy may lead to realization of conditions or preliminaries to transfer to silencing status in support of the initiation stages of nucleosome modification connected with the complexes of small RNA and protein modifiers with following recruitment of HP1 or HP type proteins. One of the stages is the recruitment of cohesion protective proteins to HP1 in pericentric heterochromatin. This view relies on the data concerning the cohesion protective proteins recruitment by HP1 (Kiburz et al., 2005; Sokuno & Watanabe, 2009), when it is essential that NPS clusters and perhaps HP1 have to a some extent the preference to localize approximately at one half-pipe side (half-pipe model).

In summarizing these data it should be underlined that these specters correspond to the variegation character of heterochromatin in pericentromere area. These features correlate with the variability of small and long genes locations as well as repeats density. Constantly, reminiscent values in the range of 50–85 kbp wavelengths with some interruptions for each chromosome argue in favor to the inconstant diameter of chromosome. It is possible to interpret half-pipe preference in nucleosome position in the frame of cylinder model in connection with the functions of protective proteins those in their location, have to protect cohesion complex of sister chromatids from irrelevant destruction (Kiburz et al., 2005; Sokuno & Watanabe, 2009).

It should be emphasized that in comparison between 2nd and 3rd chromosomes that constitutive heterochromatin area is distributed for 2nd chromosome on the right and left arms in juxtaposition to centromere, however for the 3rd chromosome analogous area is located predominately on the left arm. The right arm area juxtaposed to the centromere is enriched of high-density short genes, short inter-gene distances and heterochromatin properties are presented in the area more distantly >700 kbp, where large inter-gene regions are presented. So that the constitutive heterochromatin properties are more concentrated on the left arm of 3rd chromosome when compared with 2nd chromosome where they are distributed between two arms.

Acknowledgments
We are grateful to I.A. Zharinova and M.V. Suzdaleva for their technical assistance and useful advice.

Supplementary material
The supplementary material for this paper is available online at http://dx.doi.10.1080/07391102.2013.844081.

References
Aravin, A. A., Lagos-Quintant, M., Yalcin, A., Zavolan, M., Marks, D., Snyder, B., ... Gaasterland, T. (2003). The small RNA profile during Drosophila melanogaster development. Developmental Cell, 5, 337–350.
De Wit, E., Greil, F., & van Steensel, B. (2007). High-resolution mapping reveals links of HP1 with active and inactive chromatin components. PLoS Genetics, 3, e38.
Fagegaltier, D., Bouge, A.-L., Berry, B., Poisot, E., Sismeiro, O., Coppee, J.-Y., ... Theodore, L. (2009). The endogenous siRNA pathway is involved in heterochromatin formation in Drosophila. Proceedings of the National Academy of Sciences of the United States of America, 106, 21258–21263.
Fedoseyeva, V. B., & Alexandrov, A. A. (2007). Analysis and development of the computer methods of nucleosome localization on DNA fragments with different AT-content. Journal of Biomolecular Structure and Dynamics, 24, 481–488.
Haynes, K. A., Gracheva, E., & Elgin, S. C. (2006). A distinct type of heterochromatin within Drosophila melanogaster chromosome four. Genetics, 175, 1539–1542.
Hower, M., Dimitri, P., Berloco, M., & Wakimoto, B. (1995). Cis-effects of heterochromatin on heterochromatic and euchromatic gene activity in Drosophila melanogaster. Genetics, 140, 1033–1045.
Jacob, S. A., & Khorasanizadeh, S. (2002). Structure of HP1 chromodomain bound to lysine 9-methylated histone H3 tail. Science, 295, 2080–2083.
Kiburz, B. M., Reynolds, D. B., Meqee, P. C., Marston, A. L., Lee, B. A., Lee, T. I., ... Levin, S. S. (2005). The core centromere and Sgo1 establish a 50-kb cohesion-protected domain around centromeres during meiosis I. Genes & Development, 19, 3017–3030.
Lengronne, A., Kato, Y., Mori, S., Yokobayashi, S., Kelly, G. P., Itoh, T., ... Watanabe, Y. (2004). Cohesin relocation from sites of chromosomal loading to places of convergent transcription. Nature, 430, 573–578.
Lundberg, L. E., Stenberg, P., & Larsson, J. (2013). HP1α, su(var)-3-9, SETDB1 and POF stimulate or repress gene expression depending on genomic position, gene length and expression pattern in Drosophila melanogaster. Nucleic Acids Research, 41, 4481–4494.
Nielsen, P. R., Nielstespach, D., Mott, H. R., Callaghan, J., Bannister, A., Kouzarides, T., ... Murzin, A. G. (2002). Structure of the HP1 chromodomain bound to histone h3 methylated at lysine 9. Nature, 416, 103–107.
Noma, K., Allis, C. D., & Grewal, S. I. (2001). Transitions in distinct histone H3 methylation patterns at the heterochromatin domain boundaries. Science, 293, 1150–1155.
Pal-Bhadra, M., Leibovitch, B. A., Gandhi, S. G., Chikka, M. R., Bhadra, U., Birchler, J. A., & Elgin, S. C. (2004). Heterochromatic silencing and HP1 localization in Drosophila are dependent on the RNAi machinery. Science, 303, 669–672.
Reihart, B. J., & Bratel, D. P. (2002). Small RNAs correspond to centromere heterochromatin repeats. Science, 297, 1831–1838.
Riddle, N. C., Monoda, A., Kharchenko, P. V., Alekseyenko, A. A., Schwartz, Y. B., Tolstorukov, M. Y., ... Gorchakov, A. A. (2011). Plasticity in patterns of histone modifications and chromosomal proteins in Drosophila heterochromatin. Genome Research, 21, 147–163.
Schotta, G., Ebert, A., Dorn, R., & Reuter, G. (2003). Position-effect variegation and the genetic dissection of chromatin regulation in Drosophila. Seminars in Cell & Developmental Biology, 14, 67–75.

Sokuno, T., & Watanabe, Y. (2009). Studies of meiosis disclose distinct roles of cohesion in the core centromere and peri-centromeric regions. Chromosome Research, 17, 239–249.

Stewart, M. D., Li, J., & Wong, J. (2005). Relationship between histone H3 lysine 9 methylation, transcription repression and heterochromatin protein1 recruitment. Molecular and Cellular Biology, 25, 2525–2538.

Sun, F.-L., Cuaycong, M. H., & Elgin, S. C. R. (2001). Long-range nucleosome ordering is associated with Gene silencing in Drosophila melanogaster pericentric heterochromatin. Molecular and Cellular Biology, 21, 2867–2879.

Verschure, P., Van der Kraan, I., De Leeuw, W., Van der Vlag, J., Carpenter, A. E., Belmont, A. S., & Van Drier, R. (2005). In vivo HP1 targeting causes large-scale chromatin condensation and enhanced histone lysine methylation. Molecular and Cellular Biology, 25, 4552–4564.

Yasuhara, J. C., & Wakimoto, B. T. (2008). Molecular landscape of modified histones in Drosophila heterochromatic genes and euchromatic transition zones. PLoS Genetics, 4, e16.