

Statistical analysis of the three-dimensional structure of centromeric heterochromatin in interphase nuclei

M. BEIL, F. FLEISCHER*, S. PASCHKE & V. SCHMIDT†

Department of Internal Medicine I, University Hospital Ulm, D-89070 Ulm, Germany

*Department of Applied Information Processing & Department of Stochastics, and

†Department of Stochastics, University Ulm, D-89069 Ulm, Germany

Key words. Centromeres, confocal microscopy, heterochromatin, point field characteristics, spatial statistics.

Summary

Translocation of genes into the pericentromeric heterochromatin occurs during cellular differentiation and leads to a long-term silencing of these genes. Consequently, a structural remodelling of this heterochromatin compartment is observed during differentiation but this remains to be defined from a topological point of view. In a previous study, we analysed the three-dimensional (3D) distribution patterns of centromere clusters (chromocentres) by confocal scanning laser microscopy and found that differentiation of the promyelocytic leukaemia cell line NB4 along the neutrophil lineage is associated with a progressive clustering of centromeres. This clustering was reflected by a decreased number of detectable chromocentres, i.e. groups of centromeres with a distance below the diffraction-limited resolution of optical microscopy. The purpose of this study was to perform a statistical analysis of the 3D distribution of chromocentres in NB4 cells. Several point field characteristics (Ripley's K-function, L-function, pair correlation function, nearest-neighbour distribution function) were investigated to describe the topology of chromocentres during differentiation of NB4 cells. The pair correlation function revealed a higher frequency of chromocentre distances between 350 nm and 800 nm in undifferentiated NB4 cells as compared with differentiated cells. The L-function and the nearest-neighbour distribution function confirmed these results. These data imply the existence of intranuclear heterochromatin zones formed by functionally related centromeric regions. In view of the observed decrease in the number of detectable chromocentres during differentiation, we hypothesize that these zones with a diameter of 350–800 nm in undifferentiated NB4 cells contract into zones with a diameter below 350 nm in differentiated cells.

Introduction

It is well established that the regulation of transcription involves ligand–promotor interactions as well as modulation of DNA conformation, which is referred to as chromatin structure (Cremer & Cremer, 2001). The decondensed form of chromatin, i.e. euchromatin, is characterized by a high degree of spatial accessibility of DNA to the transcriptional machinery and therefore seems to represent the functional units of transcription (Dillon & Sabbattini, 2000). In contrast, heterochromatin is the densely packed form of DNA and is, in general, transcriptionally silent (Chubb & Bickmore, 2003). A widely accepted model of nuclear architecture is based on the assumption that individual chromosomes occupy distinct territories in interphase nuclei and that the transcriptional status of genes is affected by their position inside these territories (Cremer & Cremer, 2001). Heterochromatin is principally found in centromeric and pericentromeric regions of chromosomes (Gilbert *et al.*, 2003). These heterochromatin regions can induce transcriptional repression of juxtaposed genes (Brown *et al.*, 1997). This process remains poorly understood, but the three-dimensional (3D) organization of interphase chromosomes, i.e. the higher order chromatin structure, appears to play an important role (Perrod & Gasser, 2003).

The co-ordinated activation and silencing of genes during cellular differentiation requires a large-scale remodelling of chromatin architecture (Cremer & Cremer, 2001). Once defined during differentiation, higher order chromatin structures seem to remain stable in interphase nuclei (Sadoni *et al.*, 1999). Long-term silencing of genes, which is necessary for adopting differentiated functions, is thought to be associated with a positioning of these genes into transcriptionally silent nuclear compartments, i.e. heterochromatin. Consequently, the nuclei of terminally differentiated cells show large domains of heterochromatin (Chubb & Bickmore, 2003). Although the positioning of individual genes into the pericentromeric chromatin has already been studied, the overall structural characteristics

of this heterochromatin compartment remained to be determined. Previous studies described a progressive clustering of interphase centromeres during cellular differentiation (Martou & de Boni, 2000; Alcobia *et al.*, 2003). Furthermore, a translocation of centromeres from the nuclear periphery to the centrally located nucleolus was observed during postnatal development of Purkinje neurons (Martou & de Boni, 2000).

In a recent study, we investigated centromere distribution patterns during differentiation of acute promyelocytic leukaemia (APL) cells along the neutrophil lineage (Beil *et al.*, 2002). The position of centromeres served as a surrogate marker for the localization of pericentromeric heterochromatin during the G0/G1 phase of the cell cycle. Owing to the diffraction-limited resolution of optical microscopy we had to use the notion 'chromocentre' for a group of centromeres with a distance below the limit of optical resolution. In this previous study, the number of detectable chromocentres was found to be significantly reduced during differentiation, indicating a clustering of centromeres. The 3D distribution of chromocentres was evaluated by determining the mean and variance of the edge length of the minimal spanning tree (MST) constructed by using the 3D coordinates of the chromocentres. The results obtained by this method suggested that a large-scale remodeling of higher order chromatin structure occurs during differentiation and eventually may lead to a random distribution of chromocentres in the nucleus of differentiated APL cells.

Although MST features were shown to be very effective in classifying simulated random distributions (Wallet & Dussert, 1997), the validity of this approach might be restricted by the dependence of MST feature values on the number of chromocentres, i.e. the intranuclear density of nodes and also by the fact that the spatial distribution of a point process is very complicated in general. Therefore, we have now performed a statistical analysis to investigate further the 3D chromocentre distribution with respect to the topological alterations during differentiation of APL cells with density-independent methods. These methods analyse point field characteristics such as Ripley's K-function, the L-function, the pair correlation function and the (first) nearest-neighbour distribution function. These characteristics are widely used in the statistical analysis of spatial point patterns (Stoyan & Stoyan, 1994; Stoyan *et al.*, 1995; Diggle, 2003; Ripley, 2004) and can provide useful information about structural aspects of the spatial distribution of point patterns.

Materials and methods

Experimental procedures

The NB4 cell line was established from a patient with APL and carries the *t*(15,17) translocation that is found in most cases of APL (Lanotte *et al.*, 1991). That translocation fuses the PML gene on chromosome 15 with the gene of the retinoic acid

receptor alpha on chromosome 17 (de The *et al.*, 1990). Owing to the function of the fusion protein, pharmacological doses of all-*trans* retinoic acid (ATRA) induce differentiation of promyelocytic leukaemia cells along the neutrophil pathway (Fenaux *et al.*, 1997).

The procedures for cell culture, specimen preparation, immunofluorescence microscopy and 3D image analysis are described in detail in Beil *et al.* (2002). Briefly, differentiation of NB4 cells was induced by incubating cells with $5 \mu\text{mol L}^{-1}$ ATRA (Sigma, St Louis, MO, U.S.A.) for 4 days. Visualization of centromeres was based on immunofluorescence staining of centromere-associated proteins with CREST serum (Euroimmun Corp., Gross Groenau, Germany). First, suspension cultures of NB4 cells were diluted 4 : 1 with 16% formaldehyde to yield a final concentration of 4% formaldehyde. After 5 min, cells were centrifuged at 300 r.p.m. for 10 min onto Super-Frost slides (Mediate, Burgdorf, Germany) using a Shandon Cytospin 3 (Life Sciences International, Cheshire, U.K.). Thereafter, slides were incubated with human CREST serum diluted 1 : 2 in phosphate-buffered saline (PBS) overnight at 4 °C. After washing in PBS, slides were incubated with Alexa 488-conjugated anti-human IgG antibodies (Molecular Probes, Eugene, OR, U.S.A.) diluted 1 : 1000 in PBS for 1 h at room temperature. To prevent the RNA binding of the DNA stain YoPro-3 (Molecular Probes), RNA was degraded by incubating cells with 2 mg mL^{-1} RNase A (Sigma). Thereafter, nuclear DNA was stained with YoPro-3 at a concentration of $1 \mu\text{mol L}^{-1}$ for 1 h. After final washing, cells were mounted in Mowiol (Calbiochem, Bad Soden, Germany).

Acquisition of 3D images was performed by confocal scanning laser microscopy (voxel size: 98 nm in lateral and 168 nm in axial direction). Segmentation of cell nuclei and chromocentres was performed automatically with an interactive control. Cell nuclei were segmented on each confocal plane using a contour-following procedure. As the background of the images was almost completely devoid of fluorescence signals, the threshold for detecting the DNA counterstain (YoPro-3) could be set to a minimum level. Three-dimensional images of nuclei were reconstructed from the contours defined at each confocal plane. DNA content was determined by integrating the fluorescence intensity of YoPro-3. To select only diploid nuclei for further analysis, nuclei with a DNA content exceeding the G0/G1 peak of the DNA histogram were excluded. Segmentation of chromocentres was performed in two steps. First, objects at each confocal section were segmented by edge detection. An edge was defined as a fluorescence intensity difference between adjacent points greater than four times the standard deviation of the fluorescence signal as measured in control images. In a second step, 3D chromocentres were reconstructed by analysing series of 2D profiles. The centre of gravity was used to define the 3D coordinates for each chromocentre. The final analysis included 28 cell nuclei from untreated controls and 27 cell nuclei from ATRA-differentiated NB4 cells (Figs 1 and 2).

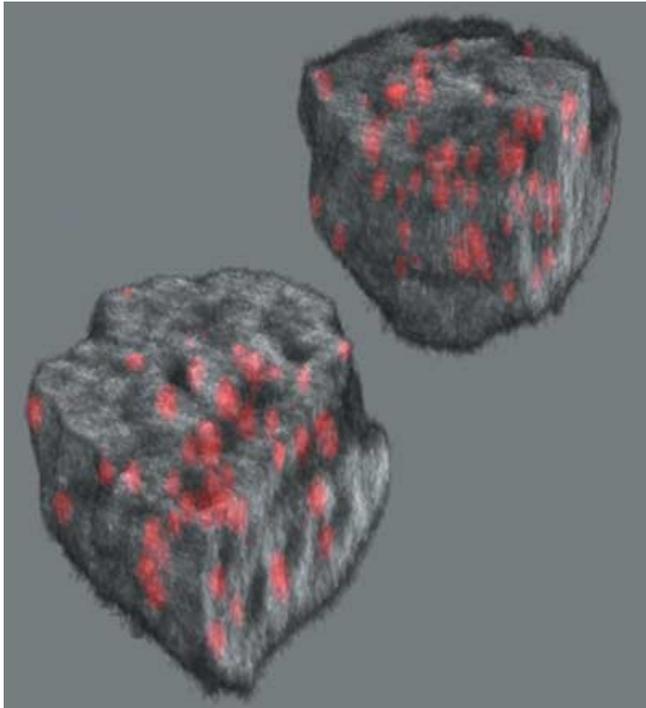


Fig. 1. Three dimensional reconstruction of NB4 cell nuclei (DNA shown in grey levels) and chromocentres including chromocentre clusters (shown in red).

Statistical analysis

Data analysis was performed using the GeoStoch library system. GeoStoch is a Java-based open-library system developed by the Department of Applied Information Processing and the Department of Stochastics of the University of Ulm, and can be used for stochastic–geometric modelling and spatial statistical analysis of image data (Mayer, 2003; Mayer *et al.*, 2004; <http://www.geostoch.de>). Statistical comparison of groups was based using the Wilcoxon–Mann–Whitney test. The real sampling regions are not known, and therefore assumed sampling regions were constructed as follows: for all three coordinates the smallest and largest values appearing in a sample were determined and denoted as x_{\min} , x_{\max} , y_{\min} , y_{\max} , z_{\min} and z_{\max} ,

respectively. Then the eight vertices of the assumed sampling cuboid were given by all possible combinations of the three coordinate pairs $\{x_{\min}, x_{\max}\}$, $\{y_{\min}, y_{\max}\}$ and $\{z_{\min}, z_{\max}\}$. In the following the positions of the chromocentres have been considered as realizations of point processes in R^3 . For the point field characteristics considered in the present paper, estimators of spatial Horvitz–Thompson type (Horvitz & Thompson, 1952; Miles, 1974) described in Section 3 have been used. An edge-effect correction was thus ensured.

Point field characteristics and their estimators

In the following let $X = \{X_n\}$ be a stationary and isotropic random point field in R^3 and let $X(B) = \#\{n: X_n \in B\}$ denote the number of points X_n of X located in a set B .

Intensity measure

The intensity measure Λ is defined as

$$\Lambda(B) = EX(B) \quad (1)$$

for a given set B . Hence $\Lambda(B)$ is the mean number of points in B . In the homogeneous case it suffices to regard an intensity λ since then

$$\Lambda(B) = \lambda |(B)| \quad (2)$$

where $|(B)|$ denotes the volume of B . A natural estimator for λ is given by

$$\hat{\lambda} = \frac{X(B)}{|B|}. \quad (3)$$

However, for estimation of the nearest-neighbour distance distribution a different estimator

$$\hat{\lambda}_H = \sum_{X_n \in X} \frac{1_{B-b(0,s(X_n))}(X_n)}{|B - b(0,s(X_n))|} \quad (4)$$

is applied, where $s(X_n)$ denotes the distance of X_n to its nearest neighbour, and $B - b(0,s(X_n))$ denotes the set B eroded by a ball with midpoint at the origin and radius $s(X_n)$. This intensity estimator was used in the case of the nearest-neighbour

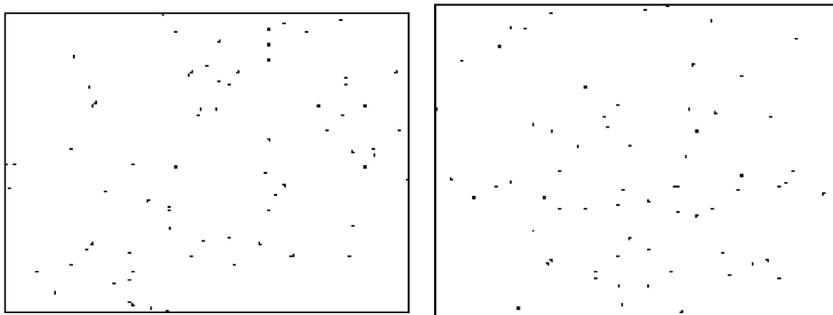


Fig. 2. Projections of the 3D chromocentre distributions of an undifferentiated NB4 cell (left) and differentiated NB4 cell (right) onto the x – y plane.

distance distribution because in Stoyan *et al.* (2001) it was shown that the estimator utilizing $\hat{\lambda}_H$ instead of $\hat{\lambda}$ usually has a smaller bias and better variance properties. Note that, following the recommendation in Stoyan & Stoyan (2000), λ^2 has been estimated by

$$\hat{\lambda}^2 = \frac{X(B)(X(B) - 1)}{|B|^2}, \quad (5)$$

because even in the Poisson case $(\hat{\lambda})^2$ is not an unbiased estimator for λ^2 .

Moment measure and product density

Let B_1 and B_2 be two sets. The second factorial moment measure $\alpha^{(2)}$ of X is defined by

$$\alpha^{(2)}(B_1 \times B_2) = E \left[\sum_{\substack{x_i, x_j \in X \\ i \neq j}} 1_{B_1}(X_i) 1_{B_2}(X_j) \right]. \quad (6)$$

Often $\alpha^{(2)}$ can be expressed using a density function $\rho^{(2)}$ as follows:

$$\alpha^{(2)}(B_1 \times B_2) = \int_{B_1} \int_{B_2} \rho^{(2)}(x_i, x_j) dx_i dx_j. \quad (7)$$

The density function $\rho^{(2)}$ is called the second product density. If one takes two balls C_1 and C_2 with infinitesimal volumes dV_1 and dV_2 and midpoints x_1 and x_2 , respectively, the probability for having in each ball at least one point of X is approximately equal to $\rho^{(2)}(x_1, x_2) dV_1 dV_2$. In the homogeneous and isotropic case, $\rho^{(2)}(x_1, x_2)$ can be replaced by $\rho^{(2)}(r)$, where $r = |x_1 - x_2|$.

As an estimator

$$\hat{\rho}^{(2)}(r) = \frac{1}{4\pi r^2} \sum_{\substack{x_i, x_j \in B \\ i \neq j}} \frac{k_h(r - |X_i - X_j|)}{|B|} \quad (8)$$

has been used (Stoyan & Stoyan, 2000), where $k_h(x)$ denotes the Epanechnikov kernel

$$k_h(x) = \frac{3}{4h} \left(1 - \frac{x^2}{h^2} \right) 1_{(-h, h)}(x) \quad (9)$$

and the sum in Eq. (8) extends over all pairs of points $X_i, X_j \in B$ with $i \neq j$. The bandwidth h has been chosen as $h = c\hat{\lambda}^{-1/3}$ with a fixed parameter $c \in \{0.06, 0.08\}$ resulting in bandwidths approximately between 120 nm and 200 nm.

Pair correlation function

The product density $\rho^{(2)}(r)$ is used to obtain the pair correlation function $g(r)$ as

$$g(x) = \frac{\rho^{(2)}(r)}{\lambda^2}. \quad (10)$$

The pair correlation function at a certain value r can be regarded as the frequency of point pairs with distance r , where $g(r) = 1$ is a base value. The pair correlation function

can be estimated by the use of estimators for $\rho^{(2)}(r)$ and λ^2 , respectively.

Note that $g(r) \geq 0$ for all distances r . In the Poisson case $g_{\text{Poi}}(r) \equiv 1$, and therefore $g(r) > 1$ indicates that there are more point pairs having distance r than in the Poisson case, whereas $g(r) < 1$ indicates that there are fewer point pairs of such a distance.

K-function

Ripley's K-function (Ripley, 1976) is defined such that $\lambda K(r)$ is the expected number of points of the stationary point field $X = \{X_n\}$ within a ball $b(X_n, r)$ centred at a randomly chosen point X_n that itself is not counted. Formally,

$$\lambda K(r) = E \sum_{X_n \in B} \frac{X(b(X_n, r)) - 1}{\lambda |B|}. \quad (11)$$

The K-function has been estimated by

$$\hat{K}(r) = \frac{\kappa(r)}{\hat{\lambda}^2} \quad (12)$$

where

$$\kappa(r) = \sum_{\substack{x_i, x_j \in B \\ i \neq j}} \frac{1_{b(0, r)} X_j - X_i}{|B_{x_j} \cap B_{x_i}|} \quad (13)$$

and $B_{x_j} = \{x + X_j : x \in B\}$ is the set B translated by the point X_j . For Poisson fields it is easy to see that $K_{\text{Poi}}(r) = 4\pi r^3/3$.

L-function

Often it is more convenient to scale the $K(r)$ in order to get a function equal to r for the Poisson case. Hence $L(r)$ is defined as

$$L(r) = \sqrt[3]{\frac{3K(r)}{4\pi}}. \quad (14)$$

A natural estimator for $L(r)$ is given by

$$\hat{L}(r) = \sqrt[3]{\frac{3\hat{K}(r)}{4\pi}}. \quad (15)$$

Nearest-neighbour distance distribution

The nearest-neighbour distance distribution D is the distribution function of the distance from a randomly chosen point X_n of the given stationary point field X to its nearest neighbour. Hence $D(r)$ is the probability that a randomly chosen point X_n of X has a neighbour with a distance less than or equal to r . According to Stoyan *et al.* (2001), we used the Hanisch estimator $\hat{D}_H(r) = D_H(r)/\hat{\lambda}_H$ (Hanisch, 1984) with

$$D_H(r) = \sum_{X_n \in X} \frac{1_{B-b(0, s(x_n))}(X_n) 1_{(0, r)}(s(X_n))}{|(B - b(0, s(X_n)))|}. \quad (16)$$

A useful property of the nearest-neighbour distance distribution is that in the case of stationary Poisson fields we have

$$D_{\text{Poi}}(r) = 1 - \exp\left(-\frac{4}{3}\lambda\pi r^3\right). \quad (17)$$

Therefore one can conclude that $D(r) < D_{\text{Poi}}(r)$ indicates rejection between points; by contrast, $D(r) > D_{\text{Poi}}(r)$ indicates attraction, keeping in mind that the nearest-neighbour distance distribution function is a cumulative quantity.

Results

Intensities and volumes

It is already known from Beil *et al.* (2002) that the average number of detected chromocentres was significantly decreased during differentiation (68.25 vs. 57.26). Nuclear volume ranged from $193 \mu\text{m}^3$ to $443 \mu\text{m}^3$. Regarding the volumes of the assumed sampling cuboids, the hypothesis of having the same volume before and after differentiation could not be rejected ($\alpha = 0.05$), observing mean volumes of $429.507 \mu\text{m}^3$ before differentiation and $470.929 \mu\text{m}^3$ afterwards. Hence the intensity of detected chromocentres, i.e. the average number per unit volume, is significantly decreased as well.

Averaged vs. individual values

We found that the variability in the group of differentiated as well as in the group of undifferentiated NB4 cells is greater than the difference between these two groups. Hence, in the following, means for each group are regarded instead of individual characteristics for each nucleus. For functions, the mean was taken in a point wise sense. A positive note is that by taking means of individual characteristics for each group, we do not have to be too concerned regarding regions with small numbers of points per sampling because by averaging over several sampling regions the accuracy of the estimated characteristics is improved.

Pair correlation function

Figures 3 and 4 show estimates of the pair correlation function for different parameters c . It is clear that the frequency of point-pair distances for a distance between 350 and 800 nm is higher before than after differentiation. In addition, a hardcore distance r_0 of about 350 nm can be recognized, which is determined by the limited spatial resolution of the imaging and image analysis methods. This means that all point pairs have a distance greater than r_0 . Note that the smaller hardcore values for larger values of c are due to the increased bandwidths of the Epanechnikov kernel in these cases. The results for the estimated pair correlation functions are independent of the fact that the two groups have different numbers of detectable chromocentres. For different parameters c , and therefore different bandwidths, similar results were obtained.

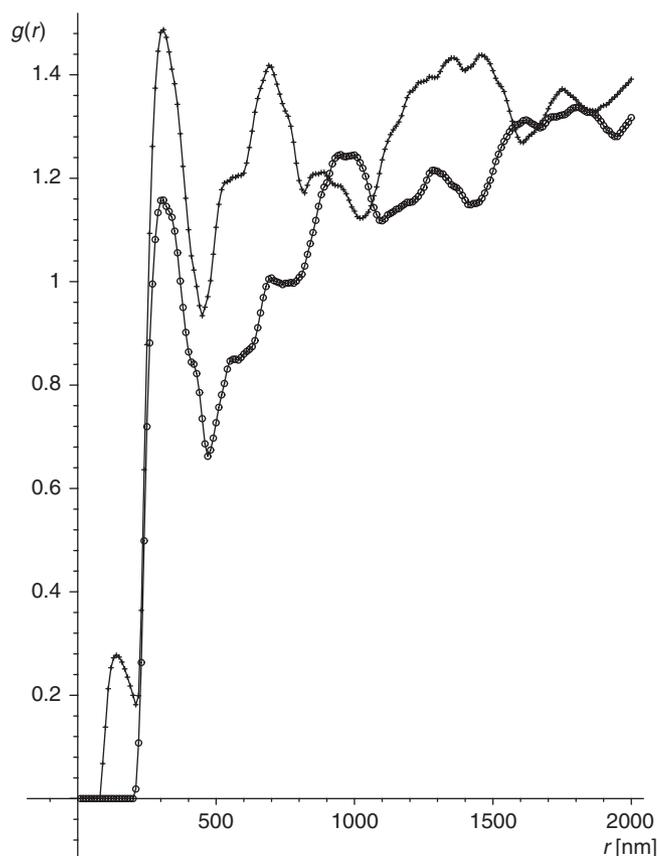


Fig. 3. Averaged estimated pair correlation function using an Epanechnikov kernel and $c = 0.06$. The group of undifferentiated NB4 cells is denoted by crosses, and the group of differentiated NB4 cells is denoted by circles.

L-function

Figure 5 shows the estimated averaged L-function and Fig. 6 shows $\hat{L}(r) - r$, where the theoretical value r for Poisson fields has been subtracted.

A similar scenario as for the pair correlation function is observed. In particular for small point-pair distances of 350–500 nm, there is a higher percentage of point pairs before than after ATRA-induced differentiation. Whereas for the group of undifferentiated cells the graph $\hat{L}(r) - r$ has a mostly positive slope in this region, which is an indicator of attraction, the group of differentiated cells shows a negative slope, indicating rejection. The same hardcore distance $r_0 \approx 350$ nm is visible. Again the results are independent of the different numbers of detectable chromocentres.

Performing a t -test for the two group samples for fixed radii shows a significant difference in the values of L-functions before and after differentiation for all radii between 350 nm and 1300 nm, especially for the region between 500 nm and 700 nm ($\alpha = 0.05$).

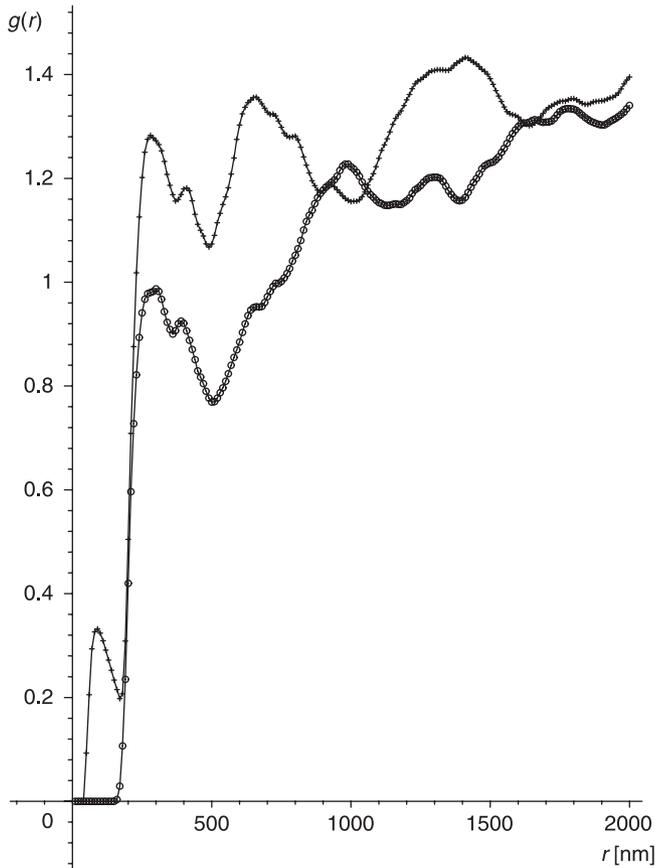


Fig. 4. Averaged estimated pair correlation function using an Epanechnikov kernel and $c = 0.08$. The group of undifferentiated NB4 cells is denoted by crosses, and the group of differentiated NB4 cells is denoted by circles.

Nearest-neighbour distance distribution

Here an interpretation is complicated by the fact that the function depends on the estimated intensities, which differ significantly for the two groups of NB4 cells. Therefore, in Fig. 7 averages of the theoretical function values for Poisson fields with same intensities as the estimated intensities of the samples are also shown. The hardcore distance r_0 is approximately equal to 350 nm; in the range between 350 nm and 500 nm the group of undifferentiated cells has a much higher frequency of nearest-neighbour point pairs than the group of differentiated cells. At approximately 700 nm the group of undifferentiated cells crosses its associated theoretical value, whereas the group of differentiated cells is crossing it at approximately 1000 nm.

Discussion

The compartmentalization of the genome in the interphase nucleus appears to be a prerequisite for a strictly regulated transcription of genes that eventually allow differentiated cells

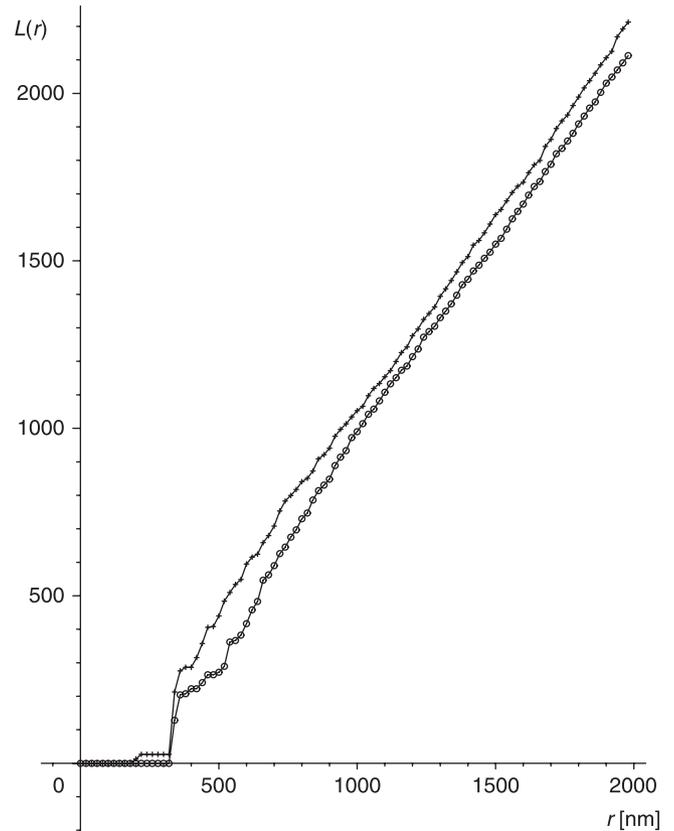


Fig. 5. Averaged estimated L-function, where crosses denote the group of undifferentiated NB4 cells and circles denote the group of NB4 differentiated cells.

to function in a programmed way (Cremer & Cremer, 2001). Although association with heterochromatin is not observed for all silent gene loci and is not necessarily incompatible with gene expression (Chubb & Bickmore, 2003), the heterochromatin compartment seems to play a pivotal role in the long-term silencing of genes during differentiation. Consequently, a translocation into heterochromatin domains during silencing was demonstrated for individual genes (Schubeler *et al.*, 2000). However, the rules that determine the dynamic architecture of heterochromatin at a global, i.e. nuclear, level of observation remained to be elucidated. Previous approaches to the analysis of chromatin architecture (Martou & de Boni, 2000; Chubb *et al.*, 2002; Kozubek *et al.*, 2002) were based on the definition of nuclear landmarks, whose distribution could depend on yet unknown processes (Spector & Gasser, 2003).

The centromeric and pericentromeric regions of chromosomes represent an important part of the heterochromatin compartment in interphase nuclei. A previous study of our group was focused on the quantitative description of 3D distribution patterns of centromeric chromatin (chromocentres) using features of the MST (Beil *et al.*, 2002). From a mathematical point of view, this approach has several disadvantages. Quantities such as the MST edge lengths or their variance are

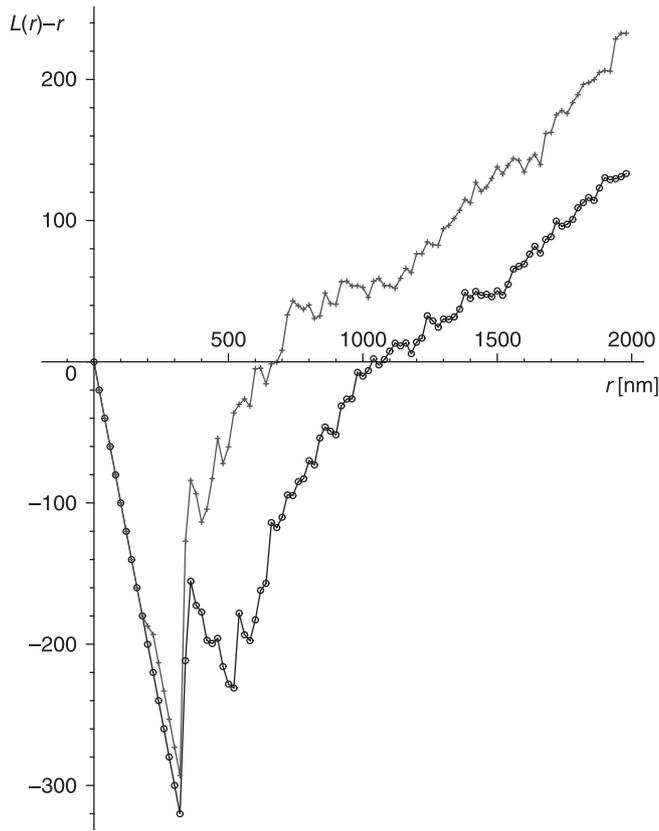


Fig. 6. Averaged estimated function $L(r) - r$, where crosses denote the group of NB4 undifferentiated cells and circles denote the group of NB4 differentiated cells.

strongly dependent on the mean number of points per volume unit. In addition to this, the methods applied in the present study allow us to gain inference regarding different specific regions of point pair distances. Thus, this novel approach provides the opportunity for a more detailed analysis of 3D chromocentre distributions.

Throughout this paper we assumed stationarity and isotropy for the point fields studied. Although formal tests have been performed that favoured our working assumption, we regard stationarity and isotropy as prior assumptions that are not in question. Because the numbers of points per sampling region are not sufficiently large, the results of our formal tests for stationarity and isotropy should not be considered as results of the statistical analysis but as hints that our working assumptions might not be poorly chosen. A similar argument can be performed with regard to tests for complete spatial randomness. Again we tested the Poisson hypothesis during preliminary investigations, which led to significant rejections but also in this case we regard the rejection of complete spatial randomness more as a prior assumption than as a result of our analysis, due to the given data basis.

Note that, although the observed point patterns are finite and bounded, we assume that they are realizations of stationary

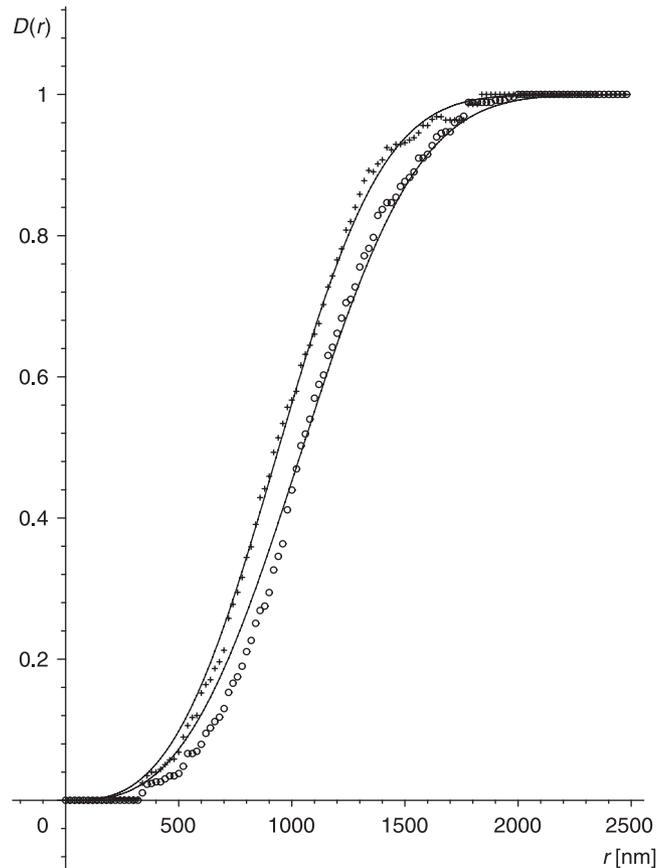


Fig. 7. Averaged estimated nearest-neighbour distance distribution, where crosses denote the group of undifferentiated NB4 cells and circles denote the group of NB4 differentiated cells.

point processes restricted to a bounded sampling region. This method of assuming unbounded stationary point processes as sources for observed realizations restricted to a bounded sampling region is a common practice because often data are given in finite sampling regions and behave in a rather different non-stationary way outside these regions (Diggle *et al.*, 2000; Schladitz *et al.*, 2003). Owing to the assumption that the observed point samples in the bounded sampling regions are extracts of unbounded realizations of stationary point processes, it is necessary, although having only bounded sampling regions, to perform edge-corrections in order to ensure compatibility with the methods applied.

We are aware that the procedure of constructing assumed sampling regions introduces small effects on the results of the estimations of point process characteristics. However, these effects do not seem to be too serious given that our main focus was on the comparison of the two different cell groups, which have similar total volumes for the assumed sampling region and therefore suffer similarly from such effects. In addition, during our preliminary investigations, tests for stationarity did not show significant rejections. At first glance the average numbers of chromocentres per cell nuclei (68.25 vs. 57.26)

do not seem to be large enough for the data analysis performed in three dimensions. But note that, in particular for the estimation of point field characteristics, an analysis is performed for mean values of several sampling regions only, which we assumed to be independent and identically distributed. As has been stated in Stoyan & Stoyan (1994) such a technique can help to improve the accuracy of the estimators.

In addition to spatial Horvitz–Thompson style estimators, other types of estimators, e.g. of Kaplan–Meier type (Baddeley & Gill, 1997), and other techniques of edge correction might also be applicable.

With regard to the choice of the parameter c for estimations of pair correlation functions, we consider that the resulting bandwidths with lengths of 120–200 nm represent a good balance for this particular data between necessary smoothing and good characterization of the features of the averaged estimated pair correlation function. Determination of bandwidths that depend on estimated intensities and a fixed parameter is a quite common tool (e.g. see Stoyan & Stoyan, 1994).

Clustering of nuclear structures and loci is considered to be an important mechanism regulating the functional organization of chromatin. Telomeres were shown to exhibit a dynamic clustering *in vivo* (Molenaar *et al.*, 2003). Intranuclear SC-35 domains appear to constitute centres of clustered euchromatin (Shopland *et al.*, 2003). Clustering patterns of centromeres were related to the stage of cellular differentiation (Park & de Boni, 1992). In our previous study, we observed a progressive clustering of centromeres after differentiation of NB4 cells with ATRA (Beil *et al.*, 2002). A similar observation has recently been reported by Alcobia *et al.* (2003) for lymphopoiesis. Owing to the diffraction-limited resolution of optical microscopy, which does not permit the detection of centromere distances below 350 nm, clusters (chromocentres) represent groups of centromeres with a distance below that threshold. In the present study, we analysed the distance of chromocentres and found a higher frequency of chromocentre distances between 350 and 800 nm for undifferentiated cells in comparison with ATRA-differentiated NB4 cells (Figs 3, 4 and 6). These new data imply the existence of heterochromatin zones with a range of 350–800 nm containing functionally related centromeric regions. The centromeres in these zones agglomerate during ATRA-induced differentiation of NB4 cells. The resulting clusters, i.e. chromocentres, then contain centromeres located within a sphere having a diameter of less than 350 nm. This model indicates that the chromatin remodelling during differentiation of NB4 cells exceeds a simple expansion of already existing heterochromatin loci and involves an intranuclear movement of centromeric regions. Such a movement of chromosomal regions would also require a remodelling of chromosome territories providing access of centromeric regions to the periphery of chromosome territories to allow clustering with centromeric regions of neighbouring chromosomes. The existence of heterochromatin zones containing centromeres of specific chromosomes would imply that the restructuring of

these chromosome territories has to proceed in a co-ordinated non-random way during the differentiation-induced ‘collapse’ of these heterochromatin zones. This model is in accordance with a topological model for gene regulation based on the structural remodelling of chromosome territories during modulation of transcription (Park & de Boni, 1999; Cremer & Cremer, 2001).

With regard to the estimated point field characteristics and some preliminary formal tests it seems that the 3D distribution of chromocentres is not completely random in undifferentiated as well as in ATRA-differentiated NB4 cells. However, caution needs to be exercised in making any statements here because the average number per sampling region is not very large and therefore the significance of formal tests is weakened. The main focus of this study was therefore not to test the two groups for complete spatial randomness, but to detect differences between them. A rejection of complete spatial randomness would be in opposition to a previous hypothesis that was based on the comparison of chromocentre distributions in NB4 cells with simulated completely random patterns (Beil *et al.*, 2002). By contrast, it would be in accordance with other studies, which suggested that interphase centromeres are not arranged in a completely random way (Manuelidis, 1984; Haaf & Schmid, 1991; Janevski *et al.*, 1995).

The molecular mechanisms governing higher order chromatin architecture are currently under investigation. HP1 and MENT are potential regulators of heterochromatin architecture (Kellum & Alberts, 1995; Grigoryev *et al.*, 1999). Transcription factors were also shown to induce changes of higher order chromatin structure (Lundgren *et al.*, 2000). The methods used in the present study could become a valuable tool in analysing structural changes of the heterochromatin compartment induced by targeting specific molecular pathways.

References

- Alcobia, I., Quina, A.S., Neves, H., Clode, N. & Parreira, L. (2003) The spatial organization of centromeric heterochromatin during normal human lymphopoiesis: evidence for ontogenically determined spatial patterns. *Exp. Cell Res.* **290**, 358–369.
- Baddeley, A.J. & Gill, R.D. (1997) Kaplan–Meier estimators of distance distribution for spatial point processes. *Ann. Statistics*, **25**, 263–292.
- Beil, M., Durschmied, D., Paschke, S., Schreiner, B., Nolte, U., Bruel, A. & Irinopoulou, T. (2002) Spatial distribution patterns of interphase centromeres during retinoic acid-induced differentiation of promyelocytic leukemia cells. *Cytometry*, **47**, 217–225.
- Brown, K.E., Guest, S.S., Smale, S.T., Hahm, K., Merckenschlager, M. & Fisher, A.G. (1997) Association of transcriptionally silent genes with Ikaros complexes at centromeric heterochromatin. *Cell*, **91**, 845–854.
- Chubb, J.R. & Bickmore, W.A. (2003) Considering nuclear compartmentalization in the light of nuclear dynamics. *Cell*, **112**, 403–406.
- Chubb, J.R., Boyle, S., Perry, P. & Bickmore, W.A. (2002) Chromatin motion is constrained by association with nuclear compartments in human cells. *Curr. Biol.* **12**, 439–445.

- Cremer, T. & Cremer, C. (2001) Chromosome territories. nuclear architecture and gene regulation in mammalian cells. *Nat. Rev. Genet.* **2**, 292–301.
- Diggle, P.J. (2003) *Statistical Analysis of Spatial Point Patterns*. Oxford University Press, Oxford.
- Diggle, P.J., Mateu, J. & Clough, H.E. (2000) A comparison between parametric and non-parametric approaches to the analysis of replicated spatial point patterns. *Adv. Appl. Probability*, **32**, 331–343.
- Dillon, N. & Sabbattini, P. (2000) Functional gene expression domains: defining the functional unit of eukaryotic gene regulation. *Bioessays*, **22**, 657–665.
- Fenaux, P., Chomienne, C. & Degos, L. (1997) Acute promyelocytic leukemia: biology and treatment. *Semin. Oncol.* **24**, 92–102.
- Gilbert, N., Boyle, S., Sutherland, H., de Las Heras, J., Allan, J., Jenuwein, T. & Bickmore, W.A. (2003) Formation of facultative heterochromatin in the absence of HP1. *EMBO J.* **22**, 5540–5550.
- Grigoryev, S.A., Bednar, J. & Woodcock, C.L. (1999) MENT, a heterochromatin protein that mediates higher order chromatin folding, is a new serpin family member. *J. Biol. Chem.* **274**, 5626–5636.
- Haaf, T. & Schmid, M. (1991) Chromosome topology in mammalian interphase nuclei. *Exp. Cell Res.* **192**, 325–332.
- Hanisch, K.-H. (1984) Some remarks on estimators of the distribution function of nearest-neighbor distance in stationary spatial point patterns. *Statistics*, **15**, 409–412.
- Horvitz, D.G. & Thompson, D.J. (1952) A generalization of sampling without replacement from a finite universe. *J. Am. Statist. Assoc.* **47**, 663–685.
- Janevski, J., Park, P.C. & De Boni, U. (1995) Organization of centromeric domains in hepatocyte nuclei: rearrangement associated with de novo activation of the vitellogenin gene family in *Xenopus laevis*. *Exp. Cell Res.* **217**, 227–239.
- Kellum, R. & Alberts, B.M. (1995) Heterochromatin protein 1 is required for correct chromosome segregation in *Drosophila* embryos. *J. Cell Sci.* **108**, 1419–1431.
- Kozubek, S., Lukasova, E., Jirsova, P., Koutna, I., Kozubek, M., Ganova, A., Bartova, E., Falk, M. & Pasekova, R. (2002) 3D structure of the human genome: order in randomness. *Chromosoma*, **111**, 321–331.
- Lanotte, M., Martin-Thouvenin, V., Najman, S., Balerini, P., Valensi, F. & Berger, R. (1991) NB4, a maturation inducible cell line with t(15;17) marker isolated from a human acute promyelocytic leukemia (M3). *Blood*, **77**, 1080–1086.
- Lundgren, M., Chow, C.M., Sabbattini, P., Georgiou, A., Minaee, S. & Dillon, N. (2000) Transcription factor dosage affects changes in higher order chromatin structure associated with activation of a heterochromatic gene. *Cell*, **103**, 733–743.
- Manuelidis, L. (1984) Different central nervous system cell types display distinct and nonrandom arrangements of satellite DNA sequences. *Proc. Natl Acad. Sci. U.S.A.*, **81**, 3123–3127.
- Martou, G. & De Boni, U. (2000) Nuclear topology of murine cerebellar Purkinje neurons: changes as a function of development. *Exp. Cell Res.* **256**, 131–139.
- Mayer, J. (2003) *On quality improvement of scientific software: theory, methods, and application in the GeoStoch Development*. Doctoral dissertation, University of Ulm.
- Mayer, J., Schmidt, V. & Schweiggert, F. (2004) A unified simulation framework for spatial stochastic models. *Simulation Modelling Practice Theory*, **12**, 307–326.
- Miles, R.E. (1974) On the elimination of edge-effects in planar sampling. *Stochastic Geometry: a Tribute to the Memory of Rollo Davidson* (ed. by E. F. Harding and D. G. Kendall), pp. 228–247. J. Wiley & Sons, Chichester.
- Molenaar, C., Wiesmeijer, K., Verwoerd, N.P., Khazen, S., Eils, R., Tanke, H.J. & Dirks, R.W. (2003) Visualizing telomere dynamics in living mammalian cells using PNA probes. *EMBO J.* **22**, 6631–6641.
- Park, P.C. & De Boni, U. (1992) Spatial rearrangement and enhanced clustering of kinetochores in interphase nuclei of dorsal root ganglion neurons in vitro: association with nucleolar fusion. *Exp. Cell Res.* **203**, 222–229.
- Park, P.C. & De Boni, U. (1999) Dynamics of structure–function relationships in interphase nuclei. *Life Sci.* **64**, 1703–1718.
- Perrod, S. & Gasser, S.M. (2003) Long-range silencing and position effects at telomeres and centromeres: parallels and differences. *Cell Mol. Life Sci.* **60**, 2303–2318.
- Ripley, B.D. (1976) The second-order analysis of stationary point processes. *J. Appl. Probability*, **13**, 255–266.
- Ripley, B.D. (2004) *Spatial Statistics*. Wiley-VCH, Weinheim.
- Sadoni, N., Langer, S., Fauth, C., Bernardi, G., Cremer, T., Turner, B.M. & Zink, D. (1999) Nuclear organization of mammalian genomes. Polar chromosome territories build up functionally distinct higher order compartments. *J. Cell Biol.* **146**, 1211–1226.
- Schladitz, K., Särkkä, A., Pavenstädt, I., Haferkamp, O. & Mattfeldt, T. (2003) Statistical analysis of intramembranous particles using fracture specimens. *J. Microsc.* **211**, 137–153.
- Schubeler, D., Francastel, C., Cimbara, D.M., Reik, A., Martin, D.I. & Groudine, M. (2000) Nuclear localization and histone acetylation: a pathway for chromatin opening and transcriptional activation of the human beta-globin locus. *Genes Dev.* **14**, 940–950.
- Shopland, L.S., Johnson, C.V., Byron, M., McNeil, J. & Lawrence, J.B. (2003) Clustering of multiple specific genes and gene-rich R-bands around SC-35 domains: evidence for local euchromatic neighborhoods. *J. Cell Biol.* **162**, 981–990.
- Spector, D.L. & Gasser, S.M. (2003) A molecular dissection of nuclear function. *EMBO Report*, **4**, 18–23.
- Stoyan, D., Kendall, W.S. & Mecke, J. (1995) *Stochastic Geometry and its Applications*. J. Wiley & Sons, Chichester.
- Stoyan, D. & Stoyan, H. (1994) *Fractals, Random Shapes and Point Fields. Methods of Geometrical Statistics*. J. Wiley & Sons, Chichester.
- Stoyan, D. & Stoyan, H. (2000) Improving ratio estimators of second order point process characteristics. *Scand. J. Statistics*, **27**, 641–656.
- Stoyan, D., Stoyan, H., Tscheschel, A. & Mattfeldt, T. (2001) On the estimation of distance distribution functions for point processes and random sets. *Image Anal. Stereol.* **20**, 65–69.
- de The, H., Chomienne, C., Lanotte, M., Degos, L. & Dejean, A. (1990) The t(15;17) translocation of acute promyelocytic leukaemia fuses the retinoic acid receptor alpha gene to a novel transcribed locus. *Nature*, **347**, 558–561.
- Wallet, F. & Dussert, C. (1997) Multifactorial comparative study of spatial point pattern analysis methods. *J. Theoret. Biol.* **187**, 437–447.