The transcriptional regulator CBP has defined spatial associations within interphase nuclei

Kirk J. McManus[1], David A. Stephens[2], Niall M. Adams[2], Suhail A. Islam[3], Paul S. Freemont[3]* and Michael J. Hendzel[1]*

[1]Department of Oncology, Cross Cancer Institute, 11560 University Avenue, T6G 1Z2, Edmonton, Alberta, Canada

[2]Department of Mathematics, Imperial College London, Huxley Building, 180 Queen's Gate, SW7 2BZ, London, UK

[3]Division of Molecular Biosciences, Imperial College London, South Kensington Campus, London SW7 2AY, UK

Running Title: Characterizing Nuclear Architecture

* Addresses for correspondence:

Dr. Michael J. Hendzel
Department of Oncology
Cross Cancer Institute
11560 University Avenue
Room 3332
Tel. # 780-432-8493
Fax. # 780-432-8892
email: michaelh@cancerboard.ab.ca

Professor Paul Freemont
Division of Molecular Biosciences
Imperial College London
Flowers Building
South Kensington Campus
London SW7 2AZ UK
Tel (44) 2075945327
p.freemont@imperial.ac.uk

Key Words: CBP, nucleus, spatial relationship, interphase architecture, statistical analyses

ABSTRACT

It is becoming increasingly clear that nuclear macromolecules and macromolecular complexes are compartmentalized through binding interactions into an apparent three dimensionally ordered structure. This ordering, however, does not appear to be deterministic to the extent that chromatin and non-chromatin structures maintain a strict three-dimensional arrangement. Rather, spatial ordering within the cell nucleus appears to conform to stochastic rather than deterministic spatial relationships. The stochastic nature of organization becomes particularly problematic when any attempt is made to describe the spatial relationship between proteins involved in the regulation of the genome. The CREB Binding Protein (CBP) is one such transcriptional regulator that, when visualised by confocal microscopy, reveals a highly punctate staining pattern comprising several hundred individual foci distributed within the nuclear volume. Similarly, markers for euchromatic sequences have similar patterns. Surprisingly, in most cases, the predicted one-to-one relationship between transcription factor and chromatin sequence is not observed. Consequently, to understand whether spatial relationships that are not coincident are non-random and potentially biologically important, it is necessary to develop statistical approaches. In this study, we report on the development and application of such an approach as applied to understanding the role of CBP in mediating chromatin modification and transcriptional regulation. We asked whether a statistically significant spatial relationship exists between sites enriched in CBP and components of the transcriptional machinery and sites where CBP-dependent chromatin modifications occur. Using nearest neighbor distance measurements and probability analyses, our results demonstrate that CBP has an order of spatial association with other nuclear sub-compartments. We observe closer associations between CBP and RNA polymerase II enriched foci and SC35 speckles than nascent RNA or specific acetylated histones. Furthermore, we find that CBP has a significantly higher probability of being close to its known *in vivo* substrate histone H4 lysine (Lys) 5 compared to the closely related H4 Lys 12. These data suggest, for the first time, that spatial colocalization within the nucleus may not be a complete measure of functional inter-dependence and that a measure of likelihood of association may be more revealing.

INTRODUCTION

It is now appreciated that the spatial relationships between chromatin and nonchromatin structures within the nucleoplasm are correlated with transcriptional activity. Some general rules are emerging for the organization of chromatin that are typically cited as evidence for both spatio-temporal organization of the nucleoplasm and an underlying regulated process to establish and maintain spatiotemporal organization (Croft et al., 1999; Tanabe et al., 2002). Specifically, chromosomes and regions of chromosomes segregate differently within the nucleus depending on whether or not they are rich in potentially transcribed genes. This organization has been described as a polar chromosomal organization because the individual interphase chromosome territories segregate their R-bands (gene rich) into the interior of the nucleoplasm whereas their G-bands (gene poor) are gathered against the periphery of the nucleus and against the nucleolar surface (Sadoni et al., 1999). Euchromatin sequences are further organized such that they maintain a spatial relationship with the predominant nucleoplasmic non-chromatin structure, the splicing factor compartments (Shopland et al., 2003). Smaller nonchromatin structures, such as promyelocytic leukemia (PML) and Cajal bodies, associate with specific regions of the genome (Shiels et al., 2001; Shopland et al., 2001; Smith and Lawrence, 2000). Beyond these rather general descriptors, our understanding of spatio-temporal regulation of the genome is limited. Most importantly, the most obvious predictions that arise from the molecular characterization of the RNA polymerase II transcriptional machinery, that genes represent the principal nuclear binding site for these proteins, has not been commonly observed despite obvious attention to the question (Grande et al., 1997; Verschure et al., 2002).

Careful studies have been performed with 3D deconvolution microscopy or laser scanning confocal microscopy to examine the spatial relationships between RNA polymerase II, sites of RNA polymerase II transcription, sequence-specific DNA binding proteins, and chromatin modifying machinery (Grande et al., 1997; Verschure et al., 2002). These macromolecular assemblies are consistently observed to enrich in small but abundant foci that are distinct from the larger, less abundant foci that are commonly observed when transcription factors are transiently over-expressed. The latter often do show colocalization of nuclear proteins that can interact (Fischle et al., 2002; Hendzel et al., 2001). The failure of the native small nuclear foci to colocalize has led us to propose that these may be structures that are independent of chromatin, perhaps involved in the assembly of macromolecular complexes rather than reflect sites where they function (Hendzel et al., 2001). By performing a comprehensive localization analysis for the transcriptional regulator CREB binding protein (CBP), we have revealed that, while there may be rules for the spatial organization of these foci, they are not easily linked to function by co-localizing to sites of known activities. An alternative relationship that appears to be related to function has been defined for nuclear structures involved in the expression of protein-coding genes. Structures may show clear spatial relationships that are complimentary, rather than coincident. The best characterized of these relationships include the proximity of the MHC gene cluster to PML bodies (Shiels et al., 2001) and the association of transcribed genes to splicing factor compartments (Shopland et al., 2003).

If we accept the alternative possibility for intranuclear sites enriched in proteins involved in transcriptional regulation of RNA polymerase II transcribed genes (Hendzel et al., 2001), spatial relationships rather than spatial colocalization may be an important factor in terms of function. This possibility has been introduced previously when it was determined that nuclear bodies that are related in function are often in close spatial proximity within the nucleoplasm (Schul et al., 1998). When examined by 3D deconvolution microscopy, protein-rich intranuclear sites involved in RNA polymerase II transcription typically approach the resolution limits of the microscope (~200 nm diameter). Because of their small size, we cannot be sure whether they represent single unified structures or a number of smaller sites packed close together. We can, however, make estimations about the number of macromolecules present within the structure and thus stochiometric relationships in colocalization studies. Additionally, we know with certainty, that these foci do not contain one or two copies of the protein under study, but comprise hundreds to thousands of individual macromolecules (McManus and Hendzel, 2003). In this context, the failure to observe co-enrichment of chromatin proteins or DNA in these foci may

reflect separate functions for these structures. We have previously proposed that such intranuclear foci are involved in processes like the assembly of multi-protein complexes that can then be released into the surrounding nucleoplasm where they can act directly on their chromatin target (Hendzel et al., 2001 and McManus and Hendzel 2001).

In this study, we have extended our earlier studies to determine whether CBP enriched nuclear foci have distinct non-random spatial associations within the interphase nucleus. CBP is a well-defined and important regulator of gene transcription and chromatin structure and has measurable histone acetyltransferase (HAT) activity *in vivo,* with a strong preference for specific lysines (Lys) on histones H3 and H4 (McManus and Hendzel, 2003). Furthermore, CBP is a well-established transcriptional co-activator for a large number of DNA binding proteins and is able to associate with an ever-expanding list of non-histone nuclear proteins. This leads to the attractive hypothesis that CBP may function as a molecular scaffold that associates directly or indirectly with a variety of proteins simultaneously including components of the transcriptional machinery and specific histone susbstrates (reviewed in McManus and Hendzel, 2001). In order to test this hypothesis, we have developed a distance-based statistical method that can analyse relationships between intranuclear foci in terms of inter-point distances (e.g. see Diggle, 2003). This approach is particularly appropriate since no other useful localizing information is readily available, the measurements relate to real distances and the co-ordinate system used to calculate distances is essentially arbitrary thereby allowing meaningful comparisons between individual cells and different batches of cells. Furthermore, there are a number of available methods for summarizing information from the collection of observed distances including nearest neighbor distances and mean distance to all points; for example, Noordmans et al. (1998) analyze voxel-by-voxel data across the entire 3D image in order to gain understanding of the spatial heterogeneity in signal. In contrast, our statistical approach focuses only on the point pattern corresponding to the identified objects in the image. Using this approach, we find that CBP has a hierarchy of non-random spatial relationships with a subset of nuclear compartments including RNA polymerase II transcription components and chromatin. We also find that CBP has a significantly higher probability of being spatially associated with its known *in vivo* substrate, histone H4 Lys 5, compared to the closely related H4 Lys 12. Our results have broader implications in the context of understanding nuclear organisation, where the underlying spatial mechanisms are unknown.

RESULTS

Characterization of CBP distribution relative to bulk chromatin

In order to define the properties of nuclear foci enriched in RNA polymerase II transcription factors, we determined whether or not CBP foci colocalized with reference structures of differing function. Our analysis focused on determining whether or not CBP foci colocalized with euchromatin, sites of dynamic acetylation, sites of RNA polymerase II transcription, and heterochromatin within the nucleoplasm. Figure 1 shows the relationship between CBP and the distribution of genomic DNA. The deconvolved images on the left show the relationship between the DNA distribution (red in composite images) and CBP (green in composites) in mouse 10T1/2 embryonic fibroblast cells, while those on the right depict the relationship between DNA distribution and CBP distribution in Indian muntjac fibroblast cells. The colour panels show projected 3-D images. Unlike the DNA distribution, CBP is found to concentrate in several hundred small foci. We have previously shown that eGFP-tagged CBP also enriches in small nuclear foci (McManus and Hendzel, 2001; McManus and Hendzel, 2003). Hence, this distribution is not a consequence of fixation but reflects the endogenous distribution of CBP. When line scans are used to compare the concentration of DNA with the concentration of CBP, CBP is predominantly found in chromatin depleted regions.

Characterization of CBP distribution relative to sites of histone modifications that demarcate euchromatin

We have previously observed that histone modifications, such as highly acetylated histone H3, are also found in small nuclear foci that enrich in chromatin-depleted regions of the nucleus. Therefore, one potential explanation for this distribution is that it reflects the distribution of transcriptionally active or potentiated regions of the genome (i.e. euchromatic). We have also shown that cells expressing CBP preferentially show increases in Lys 5 acetylation of histone H4 (McManus and Hendzel, 2003). Hence, our results implicated CBP as a histone H4 K5 acetyltransferase. Because this modification has a high turnover rate, continuous HAT activity is required to maintain K5 acetylation in these regions of the genome. Consequently, we expected that CBP would show a high degree of colocalization with acetylated K5 in these regions. To address this, we performed colocalization experiments to determine whether CBP was associated with regions that were either enriched in acetylated histone species or an unrelated control, trimethylated Lys 4 (tMeK4) of histone H3. Both histone modifications have been reported to be associated with transcriptionally active/competent regions of the genome. Figure 2A depicts deconvolved images obtained from these experiments. As expected, both antibodies show preferential staining of nuclei and, as with CBP, are enriched in several hundred small nuclear foci. Surprisingly, when CBP images (green in composites) are compared to either acetylated K5 of histone H4 (red in composites) there are very few examples where these foci colocalize – very few foci show the presence of both red and green signals (yellow in composite). Rather, most labeled regions of the nucleoplasm contain either red or green foci, but not both. This is particularly evident in the images/regions presented at higher magnification.

Although we expected CBP to colocalize with K5, it is possible that K5 sites are already fully acetylated at the time of fixation, which would negate any CBP co-localization at these sites. Nonetheless, we expected that CBP would be associated with transcriptionally active/competent regions of the genome, thereby providing the basis for its focal distribution. In order to further address this possibility, we compared the distribution of CBP with tMeK4 of histone H3. This modification has consistently been found enriched in transcriptionally active regions of the genome (Santos-Rosa et al., 2002; Schneider et al., 2004). Figure 2B compares the distribution of CBP (green in composites) with tMeK4 (red in composites). As expected, both epitopes are concentrated in small nuclear foci. When the distribution of CBP was compared to this modified histone species, however, the results once again revealed very little evidence of colocalization. Instead, these proteins appeared to be enriched in foci that were distinct from each other.

CBP distribution relative to sites of transcription

While the colocalization experiments with modified histone species indicated that CBP localization into small nuclear foci was not determined by euchromatin organization, it may be that the association of CBP with chromatin masks the histone epitopes that we used as markers

for transcriptionally active chromatin. Therefore, we again tested whether or not CBP was predominantly associated with transcriptionally active regions of the genome. In this instance, however, we used antibodies recognizing a halogenated nucleotide incorporated into nascently synthesized RNA (Figure 3A) using a brief pulse labeling with fluorouridine or an antibody recognizing RNA polymerase II (Figure 3B). While we observed examples of colocalization at the level of resolution of the fluorescence microscope (see arrows in Figures 3A and 3B), once again the majority of the CBP foci (red in composite images) existed in regions that were independent of the presence of newly synthesized RNA (green in Figure 3A composite) or RNA polymerase II (green in Figure 3B composite).

CBP distribution relative to CBP associated proteins

While the results above are not consistent with a euchromatin or transcription-dependent basis to the organization of CBP foci, proteins have been demonstrated to colocalize within these foci. For example, we have previously demonstrated that histone deacetylase-3 (HDAC3) and HDAC4 colocalize at the level of individual foci (Fischle et al., 2002). It is possible, therefore, that rather than reflect nuclear sites where these proteins carry out their activities, these nuclear foci may be involved in the assembly of multi-molecular complexes that are then available to function in the local environment. To address this, we examined the distribution of CBP relative to proteins that CBP has been previously shown to directly interact with through co-immunoprecipitation experiments. Figure 4 shows examples. Once again, although there were instances of colocalization, these results differed significantly from our previous experiments with HDAC3 and HDAC4, where their sub-nuclear distributions were very close to identical.

Statistical Approaches for Defining CBP Spatial Organization

To test the hypothesis that CBP is in spatial proximity to the chromatin that it acetylates, we characterized the spatial relationships between CBP foci and a series of other nuclear foci including sites enriched in newly synthesized RNA, RNA polymerase II, p53, CREB, and acetylated histones in the mouse 10T cell line (Table 1). We were particularly interested in defining the relationship with chromatin enriched in histone H4 acetylated at K5. Lys 5 is the least abundant acetylated species of histone H4 and is a product of CBP acetylation *in vivo* (McManus and Hendzel, 2003). In our approach, we extracted Cartesian co-ordinates for the centroid positions of individual foci and used simple inter-point distances and nearest-neighbour (NN) methods including probability measures (see Table 2 for definition of terminology). Although our approach does not take into account the irregular shape of the nucleus (for individual cells and between cells) or potential exclusion regions (e.g. nucleoli), it does have the advantage that the data can be normalized for nuclear volume across different cell images, allowing multiple observations from different cells to be used in the analysis. Furthermore, the number of observed CBP and non-CBP (NCBP) foci are sufficiently large (typically 100-300) and generally equivalent in number, to provide statistically meaningful comparisons.

The first step of our analyses involves extracting centroid coordinates (X,Y,Z) for CBP and NCBP foci from deconvoluted widefield fluorescence images. To do this we carried out a simple image reconstruction procedure of the image data, which essentially involved two steps, namely isosurface generation and foci identification via surface tracking (see Methods). We use a Marching Cubes technique (Lorensen and Cline, 1987) to construct our isosurfaces using a number of threshold values to reproduce the characteristics of the 3D image stack. Using these surface/volume reconstructions, we then identified the centers of individual 3D volumes (centroids) with a surface tracking connectivity algorithm and assigned these X,Y,Z coordinates using an arbitrary axes system. Figure 5 shows a comparison between an original image (stacked projection) and the extracted 3D-centroid points in projection showing good agreement to the original image and validating our reconstruction procedure. Each individual point represents the centroid point of the volume attributed to individual CBP enriched foci and acetylated histone K12 enriched foci shown as a scatter-plot (Fig. 5). Other approaches for confocal image reconstruction and spot detection have also been reported (Noordmans et al., 1998). The next stage of our approach is to compute NN distances between CBP and non-CBP foci. We then carried out pair-wise comparisons of these NN distance distributions. Our analysis uses the two-sample Kolmogorov-Smirnov test for point-wise equality of distribution functions (see, for example Conover (1999)). This test is preferable to the simpler Mann-Whitney two

sample test for equality of medians as it is more powerful for detecting small differences underlying distributions, although the Mann-Whitney test is a feasible alternative. In order to account for nuclei sizes, all distance measurements are on a standardized scale relative to the maximum inter-object distance within a nucleus.

Examples of these pair-wise comparisons are given in Figure 6 and supplementary Figure 1. To assess the validity of our approach in dealing with multiple observations form different cells, we compared CBP-CBP distances from different batches of 10T fibroblasts and found no significant difference between these distributions, which supports the validity of our approach (data not shown). Figure 6A graphically presents the NN distance distributions for acetylated K5 of histone H4 with acetylated K12 of histone H4. Both distributions are remarkably similar with no significant differences observed suggesting that CBP is as close to sites enriched in both acetylated lysines. In contrast a comparison between the NN distance distance distributions of 5-Fluro-uridine (FU), identifying nascent mRNA transcripts and active RNA PolII (ARNA3) shows significant differences (Figure 6C). CBP is on average closer to sites enriched in active RNA PolII than sites enriched in nascent transcripts. However for all the FU pairwise comparisons we did note a small sample of close association to CBP (Figure 6D), which could reflect sites of active transcription.

From our NN distance distributions, we can compute median distances as a way of comparing associations of sites enriched in CBP to sites enriched in other non-CBP components (Figure 7). These distances can be interpreted as estimated median NN distances with defined uncertainty intervals for each comparison; here the 95% confidence intervals for the median distance are estimated using bootstrap resampling. This general type of procedure has been much used in many fields of application including biology (see for example Manly (1997)) and involves random relabelling of objects, and recalculation of the test statistic, for a large number of bootstrap resamples, in order to gain an understanding of the variability of the statistic. For example, if an image contains N=N0+N1 objects, with N0 labelled 0 (CBP) and N1 labelled 1 (nuclear component), we may compute a summary statistic T that is informative about the positive or negative spatial association between object types. We use the median inter-object distance between the nuclear component and CBP computed over all images in the experiment. Suppose that, for the original data, the summary statistic is observed to be $T=t^*$. To obtain a standard error or uncertainty interval for the statistic, we form a pseudo data set by resampling with replacement N0 objects from the list of objects labeled 0, and then inspect the NN distances for this new set of pseudo data in order to form a new pseudo summary statistic, t_1. We then repeat this exercise B times to form a sample of pseudo statistics $t_1,\ldots, t_B$, and report the standard error (or 95% central range) derived from this sample as the estimated uncertainty measure for the summary statistic concerned. A similar method, in a similar application, was used by Knowles et al. (2000); see also the discussion below.

The results in Figure 7 are shown as box-plots with median NN distance to CBP for different nuclear components shown in ascending order. The scale of these are relative to the maximal possible distance measured across all images and do not reflect actual distances; this standardization is necessary as the images were produced at different magnifications. Interestingly, we observe three apparent groupings with one group of 'closer' CBP associations comprising active/inactive PolII and SC35 pre-mRNA splicing sites. A middle grouping comprises known and putative CBP binding proteins including CREB and phosphorylated p53 (Ser9 and Ser20). Surprisingly, the final group which shows higher median nearest neighbor distances and therefore less 'close' associations, comprises the acetylated histones H4 (Lys 5, Lys 12 and Lys 14) and sites of nascent mRNA transcripts.

Comparison of different CBP non-CBP spatial distributions

In order to compare quantitatively between different CBP and NCBP association as well as between different cells and cell lines, we used a simple probability model to determine whether the nearest neighbor to a CBP is not a CBP body taking into account the number of both CBP and NCBP (see Methods). If we observe a large value for the probability that the nearest neighbor to a CBP focus is not a CBP or vice versa, this suggests that CBP exhibits a spatial 'attraction' (association) to the other focal type. By comparing these probability values, we indirectly measure the strengths of 'attraction' between CBP and other nuclear components the

results of which are shown in Figure 8. The strengths of 'attraction' (association) are quantified as an excess odds over random between CBP and NCBP (labeled 01) or NCBP to CBP (labeled 10). Specifically, we use a binary Generalized Linear Model (GLM; see, for example, Nelder and McCullagh (1989)) with the response variable being the number of 01 associations, which is modeled as a Binomial random variable. For example, suppose there are N0 CBPs, where N01 of them have an NCBP as nearest neighbours, and N00 have a CBP as nearest neighbour. We model N01 ~ Binomial(N0,p01), where p01 is the probability that a randomly selected CBP has an NCBP as nearest neighbour. To account for the imbalance in the numbers of CBP/NCBP for different images and nuclear components, we utilize an offset model. That is, a hypothesis of "no association" might naively be thought to correspond to p01=0.5, but if, say, there is 60/40 majority of CBPS in an image series, then if there was no association between CBP and NCBP, we would obtain an estimate of p01 near to 40/100 = 0.4, and might infer negative association which is not in reality supported. The offset model uses the following formula for p01

$$\ln(p01/(1-p01)) = \alpha + \text{offset}$$

where offset = ln(N01/N00). An estimate of α reveals the nature of the association; $\alpha>0$ implies positive association, and $\alpha<0$ implies negative association.

In the plots, the zero point of the vertical axis corresponds to a random spatial association. Values above this point indicate an association stronger than expected by chance, and is therefore indicative of association. The figure also includes 95% confidence intervals for each probability. Interestingly, all NCBP foci show differing strengths of CBP association above random apart from acetylated K12 of histone H4 and to some extent, fluoro-uridine. In contrast, the highest probability for CBP association is exhibited by acetylated K5 of histone H4. This is of particular significance given that the box-plots of median distances between K5 and K12 are similar (Figure 7) as are their NN distance distributions (Figure 6A).

DISCUSSION

In previous studies, we have demonstrated a strong correlation between CBP expression and the amount of histone H4 acetylated at Lys 5 (McManus and Hendzel, 2003). Lys 5 is the last acetylation site used on histone H4 and, consequently, is the least abundant of histone H4 acetylations found primarily in the tetra-acetylated form of H4 (Thorne et al., 1990; Turner et al., 1989; Zhang et al., 2002). The linear relationship between CBP expression and the amount of nuclear acetylated Lys 5 (McManus and Hendzel, 2003) and the rapid turnover of this acetylated species of histone H4 (Zhang and Nelson, 1988) prompted us to examine whether CBP was specifically enriched in the same regions of chromatin that are acetylated at Lys 5. Surprisingly, we found that this was not the case. Rather, foci enriched in CBP appear to exist independently of the chromatin that it acetylates which is consistent with the hypothesis that histone acetyltransferases exist in structures that are independent of their chromatin binding sites (Davie and Hendzel, 1994; Hendzel et al., 2001; Hendzel et al., 1994), perhaps as a mechanism to facilitate the assembly of chromatin modifying complexes.

To examine this in more detail, we have now analysed the spatial associations of foci enriched in CBP within the mouse 10T cell line with components of the transcriptional machinery as well as sites of specific chromatin modification. Using a new statistical approach, we find that CBP has a hierarchy of non-random spatial relationships with a subset of nuclear compartments maintaining a non-random spatial proximity to the chromatin that it preferentially acetylates *in vivo*. Of particular interest is the comparison of CBP proximity to foci enriched in acetylated H4 Lys 12 with acetylated H4 Lys 5. In terms of NN distance measurements, CBP appears equally close to both sites. However, what is surprising is that when a probability measure is made on the likelihood of being proximal to either site, a striking difference is observed. We find that CBP has a significantly greater chance of being close to acetylated H4 Lys 5 than Lys 12, despite acetylated Lys 12 being more abundant within the genome. We also examined the spatial relationships with a number of additional proteins, including RNA polymerase II species, sites of RNA transcription, phosphorylated species of p53, and CREB. Each of these nuclear targets are expected to colocalize with CBP if the foci enriched in CBP merely reflect sites of chromatin association, rather than nuclear structures that are independent of sites of action within chromatin. Interestingly, with the exception of K12 acetylation, all of the studied examples of nuclear proteins exhibit spatial relationships that are higher than expected from randomly arranged structures. In addition, there are varying degrees of “affinity” between the non-CBP foci and the CBP-enriched foci. From these measurements, there does appear to be a hierarchy of association for CBP foci, which is reflecting an underlying functional organisation. However, these data also suggest for the first time that spatial colocalization within the nucleus may not be a complete measure of functional inter-dependence and that a measure of likelihood of association may be more revealing. This is perhaps not surprising given the crowded nature of the inter-chromosomal space but does underlie the need for more quantitative estimates in determining functional associations. Although, current methods cannot distinguish between functional relationships that drive spatial arrangements versus underlying organisation leading to specific functional outcomes, this study does provide a first attempt at defining observed spatial arrangements within interphase nuclei. It also provides a foundation for the further study of other well-defined non-chromatin nuclear foci (e.g. PML bodies) and compartments (e.g. chromosomes, nucleoli). Furthermore our NN distance approach will significantly extend the possibilities of analysing spatial associations on smaller scales that result from improvements in optical imaging techniques.

Developing quantitative methods to understand spatial relationships within interphase nuclei has become an important area of study given that strict colocalization analysis, when applied using the highest resolutions obtainable with fluorescence microscopy, often fail to provide meaningful information. Rather, a number of studies have indicated that spatial proximity rather than spatial colocalization may be important in regulating genome function. For example, highly acetylated chromatin (Hendzel et al., 1998), some transcriptionally active genes (Moen et al., 1995; Xing et al., 1993; Xing et al., 1995) and the gene-rich R-band regions of interphase chromosomes (Shopland et al., 2003), have been reported to be juxtaposed to splicing factor compartments. Similarly, the MHC gene locus maintains a spatial relationship with PML bodies (Shiels et al.,

2001) as do active regions of the genome in general (Wang et al., 2004). Other patterns of nuclear organisation include preferences for chromosomes to occupy specific regions of the nucleus based on both chromosome size and the transcriptional capacity of the chromatin (Croft et al., 1999; Boyle et al., 2001; Cremer et al., 2001; Parada et al., 2004; Bolzer et al., 2005). In these studies, radial positioning has been established as a reliable method to demonstrate non-random distribution patterns. More recently, relative positioning and chromosome cluster analysis have been used to demonstrate tissue specificity in these patterns (Parada et al., 2004) and modelling of chromosome territory arrangement has suggested that cell type specific differences are not due solely to nuclear shape differences (Bolzer et al., 2005). Radial positioning has also been applied to nuclear compartments with more complex distribution patterns (for example centromeres, Solovei et al., 2004). However the ability to compare distribution patterns for two or more compartments simultaneously requires more sophisticated approaches and tools. Such tools are currently not available and most researchers tend to rely on either pixel overlap of fluorescent intensities (Hendzel et al., 1998; Nielsen et al., 2002; Verschure et al., 2002; Fuchsova et al., 2002) or on cross correlation analysis (Grande et al., 1996; Grande et al., 1997; van Steelsel et al., 1996; Mattern et al., 1999). In this study we have developed an alternative approach for studying nuclear organisation based on nearest neighbor distance measurements and probability estimates. Our approach is significantly different to previous studies in that it is object-based whereas most other studies rely on comparing dual labelled 3D images in terms of intensity distributions with cross-correlation as a measure of overlap compared to random. The advantages of an object-based approach is that it allows some direct spatial measure of specific associations and allows probability estimates of associations between different components given the difficulty in delineating such association in a confined nuclear volume.

In summary, we have studied the spatial associations of the CBP transcriptional regulator within interphase nuclei and have developed statistical approaches for identifying such relationships in terms of distances and probability of association. The clear spatial association of CBP-enriched foci to regions of chromatin that are selectively acetylated by CBP *in vivo* is consistent with CBP-enriched foci playing a role in targeting the enzyme to specific chromatin substrate sites. To our knowledge, these data provide the first statistical demonstration that spatial proximity rather than spatial overlap defines a functional relationship between an enzyme and chromatin substrate.

MATERIALS AND METHODS

Cell Culture

IM (male Indian Muntjac Skin Fibroblast) and 10T1/2 (C3H mouse embryo fibroblast) cells were cultured in Ham's F10 Medium plus 20% fetal bovine serum (FBS) and α-MEM plus 10% FBS, respectively, in a 37℃ incubator with 5% CO_2. Cells were plated onto sterilized glass coverslips so that they were 50 to 80% confluent the following day. Subsequent to fixation with fresh 4.0% para-formaldehyde, cells were permeablized with phosphate buffered saline (PBS) (pH 7.5) containing 0.5% Triton X-100 for 5 min.

Immunofluorescent Labeling

Cells were washed twice with PBS and a subjected to sequential series of 30 min incubations with appropriate primary and secondary antibodies. Wash steps between incubations were performed consisting of a single wash with PBS containing 0.1% Triton X-100 and two washes with PBS. The primary antibodies used fall roughly into three distinct classes of molecules, those recognizing (modified) DNA/chromatin, RNA and transcriptional regulators (including transcription factors and co-activators) and are indicated in Table 1. Primary antibodies were recognized with appropriate mouse or rabbit secondary antibodies conjugated with either Alexa-fluor 488 or Cyanin-3 (Cy-3) purchased from Molecular Probes and Jackson ImmunoResearch Laboratories, Inc., respectively. Coverslips were mounted onto slides containing approximately 10ul of a 90% glycerol-PBS-based medium containing 1mg/ml parapheylenediamine and 0.5µg/ml 4',6-diamidino-2-phenylindole (DAPI).

RNA-Labeling

To label nascent transcripts, cycling cells were incubated with 2mM fluoro-uridine (FU) for 20 min. Cells were fixed and permeablized as indicated above, and nascent transcripts with incorporated FU were identified with an anti-bromodeoxyuridine (Boehringer) primary antibody at 1:50 that is cross-reactive with FU.

Image Acquisition

Three dimensional (3D) optical series (z-series) were collected using a Zeiss Axioplan 2 digital imaging microscope equipped with a 100x (1.4 numerical aperature) plan-apochromat lens and a Coolsnap HQ cooled-CCD camera (Roper Scientific). Z-series extending above and below individual nuclei were collected at 200nm intervals with a motorized z-motor. Metamorph version 4.5r9 (Universal Imaging Corp.) was employed for computer-based acquisition of 16-bit images comprised of three (DAPI, Alexa-488 and Cy3) individual channels per image. Composite montages of collected images were assembled in Adobe Photoshop 7.0

Image Processing and Deconvolution

Individual channels from collected Z-series were imported into SoftWoRx (AppliedPrecision) and converted into DeltaVision (DV) files. Converted files were subjected to maximum-likelihood-expectation deconvolution processing using a constrained iterative algorithm and theoretical optical transfer files generated in SoftWoRx for DAPI (485nm), Alexa-488 (535nm), and Cy3 (610nm) (softWoRx). Resulting deconvolved images were used in subsequent 3D modeling. Images were then assembled in Imaris (Bitplane) and 3D projection images were generated and saved as 12-bit tiff files.

Three Dimensional Image Reconstruction

Our procedure for the 3D reconstruction of the image data essentially involves two steps:

(a) *Generation of Isosurfaces*- Treating the image stack as a 3D field of values (R,G,B components) we used and implemented the Marching Cubes algorithm (Lorensen and Cline, 1987) to construct isosurfaces. A number of threshold values (i.e. values of R,G,B components) were used to reproduce the correct number of individual bodies or main characteristics of a given image stack or experimental dataset. This procedure was carried out for each dataset manually. No statistical evaluation of the thresholding was carried out other than to compare centroid co-ordinates at different thresholds, which did not change.

(b) *Foci Identification via Surface Tracking*- The results from (a) result in the creation of a series of triangles from the 3D volume data. A surface tracking connectivity program was developed and implemented to isolate and identify individual bodies within the volume data. This program takes a triangle as a starting 'seed point' and marks all other triangles that touch it within neighbouring voxels (using a small distance threshold). The resulting group of triangles then defines an individual 'body'. This procedure is repeated until all triangles have been grouped into individual bodies. Geometric parameters (e.g. centroid points) for individual bodies are computed from the coordinates of the constituting group of triangles.

Nearest Neighbor Assignment and distances

See Table 2 for details of the mathematical terminology. Inter-point distances between centroids were computed as Euclidean distances. Nearest neighbor distances (NN distances) are on a standardized scale as measured relative to the maximal possible inert-object distance taking into account variations in nuclei size. The algorithm for computing median NN distances from inter-centroids is as follows:

For all objects in each image/in a subgroup

- compute the collection of inter-centroid distances (ICD) for all objects
- for each object i compute the smallest ICD Mi (NN distance)
- compute the median Mi across the collection of objects

The label of the nearest neighbor to each CBP centroid focus was then obtained. In situations where ties were observed for the nearest neighbor, the tie was broken randomly. Nearest neighbor labels were aggregated across all cell images, with every CBP observed retaining a label identifying the body type of it nearest neighbor. For each non-CBP type, the required probability was estimated as a constant term in a logistic regression generalized linear model (GLM) (McCullagh and Nelder, 1986), where the binomial response datum for each cell image is the number of NCBPs that had a CBP as its nearest neighbour. An offset term was included in the GLM for each cell image that accounted for the total number of CBPs and NCBPs observed in the image. Analytic confidence intervals for the log-odds ratio association parameters were verified using bootstrap re-sampling. Excess log-odds values were computed and compared against the same quantities computed under random re-labelling of objects within an image. All computations were conducted in the R and SPLUS statistics systems; R code is available from the second author.

Distance-based Assessment of CBP Spatial Organization

In our previous studies (Shiels et al., 2001; Wang et al., 2004), we described the assessment of PML nuclear body spatial organization in relation to specific genomic loci via statistical hypothesis tests (specifically, using parametric *t*-tests). Here we use non-parametric alternatives to these tests, and Monte Carlo exact methods (see Manly 1997) to assess statistical significance. The extracted 3D coordinates for CBP and NCBP foci were used to provide empirical distribution functions (edfs) of nearest neighbor distances from each non-CBP (NCBP) focus to the nearest CBP location. Differences between the distributions for different foci were then tested using a two-sample *Kolmogorov-Smirnov* statistic (see, for example Conover 1999), utilizing randomization procedures (calibration against randomly re-labeled datasets) where necessary, under the assumption that the distribution of NCBP-CBP distances is not different for different foci. Computation of p-values performed using standard asymptotic reasoning, verified using randomization procedures. See the extended discussion on page 8 of this manuscript.

ACKNOWLEDGEMENTS

We would like to thank the financial support of the Canadian Institute for Health Research and the Alberta Heritage Foundation for Medical Research for doctoral studentships to KJM and a scholarships to MJH. We also thank the *Journal of Cell Science* for providing a Travel Fellowship to KJM to conduct this collaborative effort and also Dr Carol Shiels for help in the initial stages of the statistical analyses. PF and SI would also like to thank the Wellcome Trust for funding.

REFERENCES

Bolzer, A., G. Kreth, I. Solovei, D. Koehler, K. Saracoglu, C. Fauth, S. Muller, R. Eils, C. Cremer, M.R. Speicher, and Cremer, T. (2005). Three-dimensional maps of all chromosomes in human male fibroblast nuclei and prometaphase rosettes. *PLoS Biol*. **3**:e157.

Boyle, S. S. Gilchrist, J.M. Bridger, N.L. Mahy, J.A. Ellis, and Bickmore, W.A. (2001). The spatial organization of human chromosomes within the nuclei of normal and emerin-mutant cells, *Hum Mol Genet*. **10**, 211-219.

Cremer, M., J. von Hase, T. Volm, A. Brero, G. Kreth, J. Walter, C. Fischer, I. Solovei, C. Cremer, and Cremer, T. (2001). Non-random radial higher-order chromatin arrangements in nuclei of diploid human cells. *Chromosome Res*. **9**, 541-567.

Croft, J. A., Bridger, J. M., Boyle, S., Perry, P., Teague, P. and Bickmore, W. A. (1999). Differences in the localization and morphology of chromosomes in the human nucleus, *J Cell Biol* **145**, 1119-31.

Conover, W.J. (1999). Non-Parametric Statistics, John Wiley, New York.

Davie, J. R. and Hendzel, M. J. (1994). Multiple functions of dynamic histone acetylation. *J Cell Biochem* **55**, 98-105.

Diggle, P.J. (2003) Statistical analysis of spatial point patterns. Oxford University Press.

Fischle, W., Dequiedt, F., Hendzel, M. J., Guenther, M. G., Lazar, M. A., Voelter, W. and Verdin, E. (2002) Enzymatic activity associated with class II HDACs is dependent on a multiprotein complex containing HDAC3 and SMRT/N-CoR. *Mol Cell* **9**, 45-57.

Fuchsova, B., P. Novak, J. Kafkova, and Hozak, P. (2002) Nuclear DNA helicase II is recruited to IFN-alpha-activated transcription sites at PML nuclear bodies. *J Cell Biol*. **158**, 463-73.

Grande, M.A., I. van der Kraan, B. van Steensel, W. Schul, H. de The, H.T. van der Voort, L. de Jong, and R. van Driel, R. (1996) PML-containing nuclear bodies: their spatial distribution in relation to other nuclear components. *J Cell Biochem*. **63**, 280-91.

Grande, M. A., van der Kraan, I., de Jong, L. and van Driel, R. (1997). Nuclear distribution of transcription factors in relation to sites of transcription and RNA polymerase II. *J Cell Sci* **110 (Pt 15)**, 1781-91.

Hendzel, M. J., Kruhlak, M. J. and Bazett-Jones, D. P. (1998). Organization of highly acetylated chromatin around sites of heterogeneous nuclear RNA accumulation. *Mol Biol Cell* **9**, 2491-507.

Hendzel, M. J., Kruhlak, M. J., MacLean, N. A., Boisvert, F., Lever, M. A. and Bazett-Jones, D. P. (2001). Compartmentalization of regulatory proteins in the cell nucleus. *J Steroid Biochem Mol Biol* **76**, 9-21.

Hendzel, M. J., Sun, J. M., Chen, H. Y., Rattner, J. B. and Davie, J. R. (1994). Histone acetyltransferase is associated with the nuclear matrix. *J Biol Chem* **269**, 22894-901.

Knowles, D.W., Ortiz de Solorzano, C., Jones, A., S.J. (2000) Analysis of the 3D spatial organization of cells and subcellular structures in tissue. Proc. SPIE Vol. 3921 p. 66-73, Optical Diagnostics of Living Cells III (Daniel L. Farkas, Robert C. Leif; Eds.)

Lorensen, W.E., and Cline, H.E. (1987) Marching Cubes: a high resolution 3D surface reconstruction algorithm. *Computer Graphics*. **21**, 163-169.

Manly B.F.J. (1997). Randomisation, Bootstrap and Monte Carlo Methods in Biology. Chapman & Hall, 2nd edition.

Mattern, K.A. I. van der Kraan, W. Schul, L. de Jong, and van Driel, R. (1999) Spatial organization of four hnRNP proteins in relation to sites of transcription, to nuclear speckles, and to each other in interphase nuclei and nuclear matrices of HeLa cells. *Exp Cell Res*. **246**, 461-70.

McCullagh, P. and Nelder, J. A. (1989). Generalised Linear Models, second edition. Chapman & Hall, 2nd edition.

McManus, K. J. and Hendzel, M. J. (2001). CBP, a transcriptional coactivator and acetyltransferase. *Biochem Cell Biol* **79**, 253-66.

McManus, K. J. and Hendzel, M. J. (2003). Quantitative analysis of CBP- and P300-induced histone acetylations in vivo using native chromatin. *Mol Cell Biol* **23**, 7611-27.

Moen, P. T., Jr., Smith, K. P. and Lawrence, J. B. (1995). Compartmentalization of specific pre-mRNA metabolism: an emerging view. *Hum Mol Genet* **4 Spec No**, 1779-89.

Nielsen, J.A., L.D. Hudson, and Armstrong, R.C. (2002) Nuclear organization in differentiating oligodendrocytes. *J Cell Sci.* **151**, 4071-9

Noordmans, H.J., van der Kraa K., van Driel, R. and Smeulders, A.W.M. (1998) Randomness of spatial distributions of two ptoteins in the cell nucleus involved in mRNA synthesis and their relationship *Cytometry* **33,** 297–309

Parada, L.A. S. Sotiriou, and Misteli, T (2004) Spatial genome organization. *Exp Cell Res*. **296**, 64-70.

Sadoni, N. Langer, S., Fauth, C., Bernardi, G., Cremer, T., Turner, B. M. and Zink, D. (1999). Nuclear organization of mammalian genomes. Polar chromosome territories build up functionally distinct higher order compartments. *J Cell Biol* **146**, 1211-26.

Santos-Rosa, H., Schneider, R., Bannister, A. J., Sherriff, J., Bernstein, B. E., Emre, N. C., Schreiber, S. L., Mellor, J. and Kouzarides, T. (2002). Active genes are tri-methylated at K4 of histone H3. *Nature* **419**, 407-11.

Schneider, R., Bannister, A. J., Myers, F. A., Thorne, A. W., Crane-Robinson, C. and Kouzarides, T. (2004). Histone H3 lysine 4 methylation patterns in higher eukaryotic genes. *Nat Cell Biol* **6**, 73-7.

Schul, W., de Jong, L. and van Driel, R. (1998). Nuclear neighbours: the spatial and functional organization of genes and nuclear domains. *J Cell Biochem* **70**, 159-71.

Shiels, C., Islam, S. A., Vatcheva, R., Sasieni, P., Sternberg, M. J., Freemont, P. S. and Sheer, D. (2001). PML bodies associate specifically with the MHC gene cluster in interphase nuclei. *J Cell Sci* **114**, 3705-16.

Shopland, L. S., Byron, M., Stein, J. L., Lian, J. B., Stein, G. S. and Lawrence, J. B. (2001). Replication-dependent histone gene expression is related to Cajal body (CB) association but does not require sustained CB contact. *Mol Biol Cell* **12**, 565-76.

Shopland, L. S., Johnson, C. V., Byron, M., McNeil, J. and 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**, 981-90.

Smith, K. P. and Lawrence, J. B. (2000). Interactions of U2 gene loci and their nuclear transcripts with Cajal (coiled) bodies: evidence for PreU2 within Cajal bodies. *Mol Biol Cell* **11**, 2987-98.

Solovei, I., L. Schermelleh, K. During, A. Engelhardt, S. Stein, C. Cremer, and Cremer, T. (2004) Differences in centromere positioning of cycling and postmitotic human cell types. *Chromosoma*. **112**, 410-23.

Tanabe, H., Muller, S., Neusser, M., von Hase, J., Calcagno, E., Cremer, M., Solovei, I., Cremer, C. and Cremer, T. (2002). Evolutionary conservation of chromosome territory arrangements in cell nuclei from higher primates. *Proc Natl Acad Sci U S A* **99**, 4424-9.

Thorne, A. W., Kmiciek, D., Mitchelson, K., Sautiere, P. and Crane-Robinson, C. (1990). Patterns of histone acetylation. *Eur J Biochem* **193**, 701-13.

Turner, B. M., O'Neill, L. P. and Allan, I. M. (1989). Histone H4 acetylation in human cells. Frequency of acetylation at different sites defined by immunolabeling with site-specific antibodies. *FEBS Lett* **253**, 141-5.

van Steensel, B., E.P. van Binnendijk, C.D. Hornsby, H.T. van der Voort, Z.S. Krozowski, E.R. de Kloet, and van Driel, R. (1996) Partial colocalization of glucocorticoid and mineralocorticoid receptors in discrete compartments in nuclei of rat hippocampus neurons. *J Cell Sci.* **109**, 787-92.

Verschure, P. J., Van Der Kraan, I., Enserink, J. M., Mone, M. J., Manders, E. M. and Van Driel, R. (2002). Large-scale chromatin organization and the localization of proteins involved in gene expression in human cells. *J Histochem Cytochem* **50**, 1303-12.

Wang, J., Shiels, C., Sasieni, P., Wu, P. J., Islam, S. A., Freemont, P. S. and Sheer, D. (2004). Promyelocytic leukemia nuclear bodies associate with transcriptionally active genomic regions. *J Cell Biol* **164**, 515-26.

Xing, Y., Johnson, C. V., Dobner, P. R. and Lawrence, J. B. (1993). Higher level organization of individual gene transcription and RNA splicing. *Science* **259**, 1326-30.

Xing, Y., Johnson, C. V., Moen, P. T., Jr., McNeil, J. A. and Lawrence, J. (1995). Nonrandom gene organization: structural arrangements of specific pre-mRNA transcription and splicing with SC-35 domains. *J Cell Biol* **131**, 1635-47.

Zhang, D. E. and Nelson, D. A. (1988). Histone acetylation in chicken erythrocytes. Rates of deacetylation in immature and mature red blood cells. *Biochem J* **250**, 241-5.

Zhang, K., Williams, K. E., Huang, L., Yau, P., Siino, J. S., Bradbury, E. M., Jones, P. R., Minch, M. J. and Burlingame, A. L. (2002). Histone acetylation and deacetylation: identification of acetylation and methylation sites of HeLa histone H4 by mass spectrometry. *Mol Cell Proteomics* **1**, 500-8.

TABLE I Description of antibodies and epitopes

Antibody	Epitope	Dilution	Other
8WG16	RNA pol II (unphosphorylated)	1:500	recognizes an inactive form of RNA pol II
ARNA3	RNA pol IIA (phosphorylated)	1:200	recognizes an active form of RNA pol II
FU	Fluoro-uridine	1:50	recognizes FU incorporated into nacently synthesized RNA
CREB	cAMP Responsive Element Binding Protein	1:200	known protein-protein interactions with CBP
SC35	Splicing factor compartment	1:1	delineates splicing factor compartments
AIM1	Aurora B Kinase	1:200	protein kinase that does not directly interact with CBP (control)
K5	Acetylated K5 (H4)	1:200	strong *in vivo* product of CBP HAT activity (McManus and Hendzel, 2003), enriched in euchromatin
K12	Acetylated K12 (H4)	1:2,000	weaker *in vivo* product of CBP HAT activity (McManus and Hendzel, 2003)
K14	Acetylated K14 (H3)	1:200	strong *in vivo* product of CBP HAT activity (McManus and Hendzel, 2003)
tMeK4	trimethylated K4 (H3)	1:500	unrelated histone modification, enriched in euchromatin
S9	p53, phosphorylated serine 9	1:200	transcription factor(occurs in response to DNA damage)
S20	p53, phosphorylated serine 20	1:200	transcription factor (occurs in response to DNA damage)

TABLE II Description of statistical terminology

Term	Description
Object	*Contiguous block of voxels classified via the microscopic technique as the 'specific epitope' or 'locus'*
Centroid	*Centre of mass*
Inter-Centroid Distance	*Distance (in the standardized image) between centres of mass of two selected objects*
Nearest-Neighbor Distance	*Smallest inter-centroid distance between a selected object and all other objects (NN distance)*
Co-localization	*Two object types are regarded as 'colocalizing' if the NN distance between objects is (on average) (a) below some pre-set threshold or (b) statistically significantly smaller than a population average NN distance. Two objects colocalize if they exhibit the (stochastic) tendency to lie in proximate spatial regions.*
Label	*Descriptor given to a specific types of object*
Monte Carlo Test	*A way of assessing the computed statistical significance by random relabelling of the object in order to recalculate distances (or other summary statistics) and thus to provide an empirical null distribution to calibrate the actual observed statistic.*
Bootstrap resampling	*A simulation method for estimating the uncertainty (for example, standard error) of an estimator derived in some statistical procedure. Items are resampled with replacement from the original collection, and the estimator recomputed.*
Empirical Distribution Function	*The cumulative distribution of a sample, that is, for sample of size n, the function defined by EDF(x)=(Number of data points less than or equal to x)/n. It provides an estimate of the underlying probability distribution from which the data are generated.*
CBP	*Object name corresponding to CBP*
NCBP	*Object name not corresponding to CBP but corresponding to any other object*

FIGURES

Figure 1

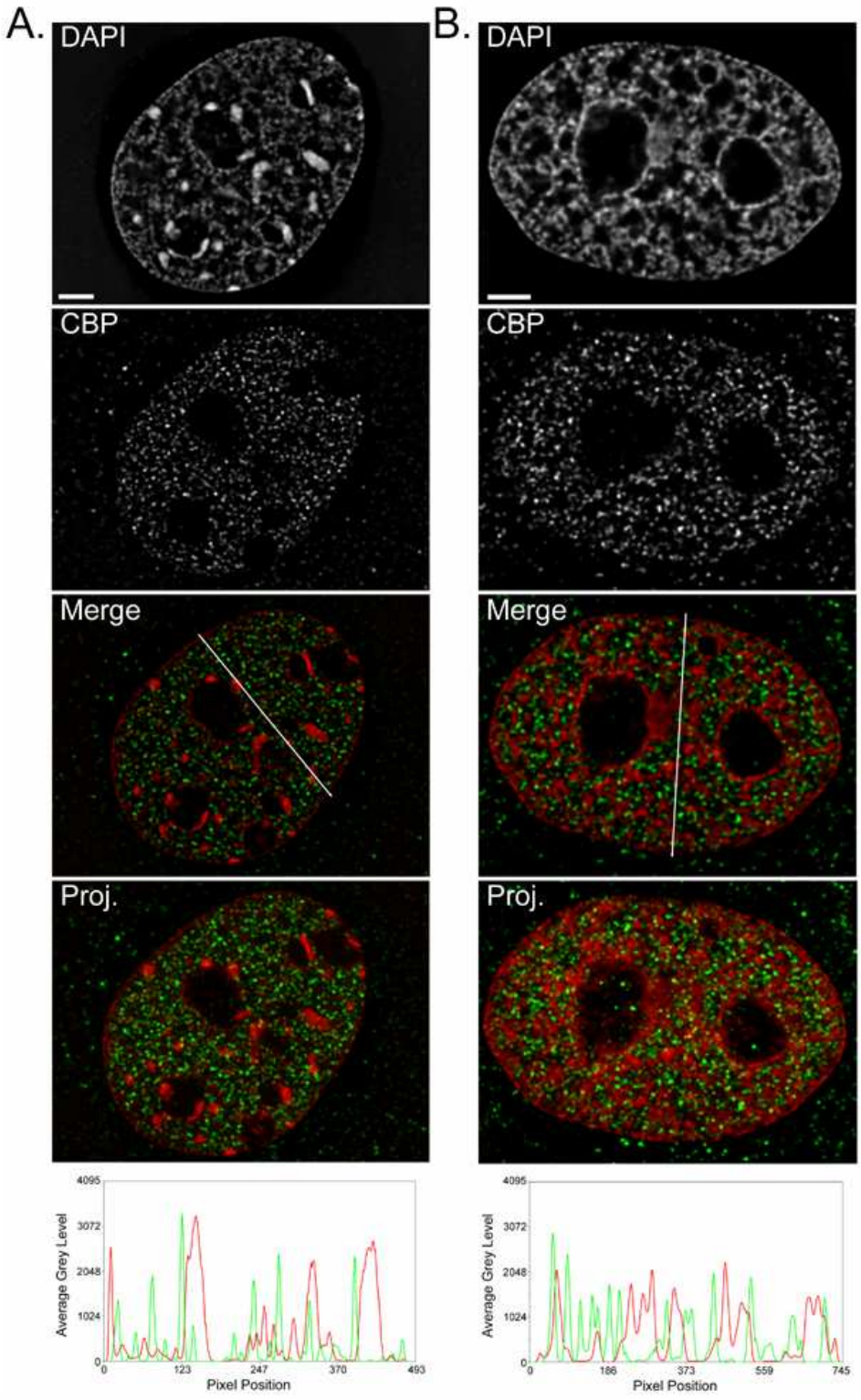


Figure 2

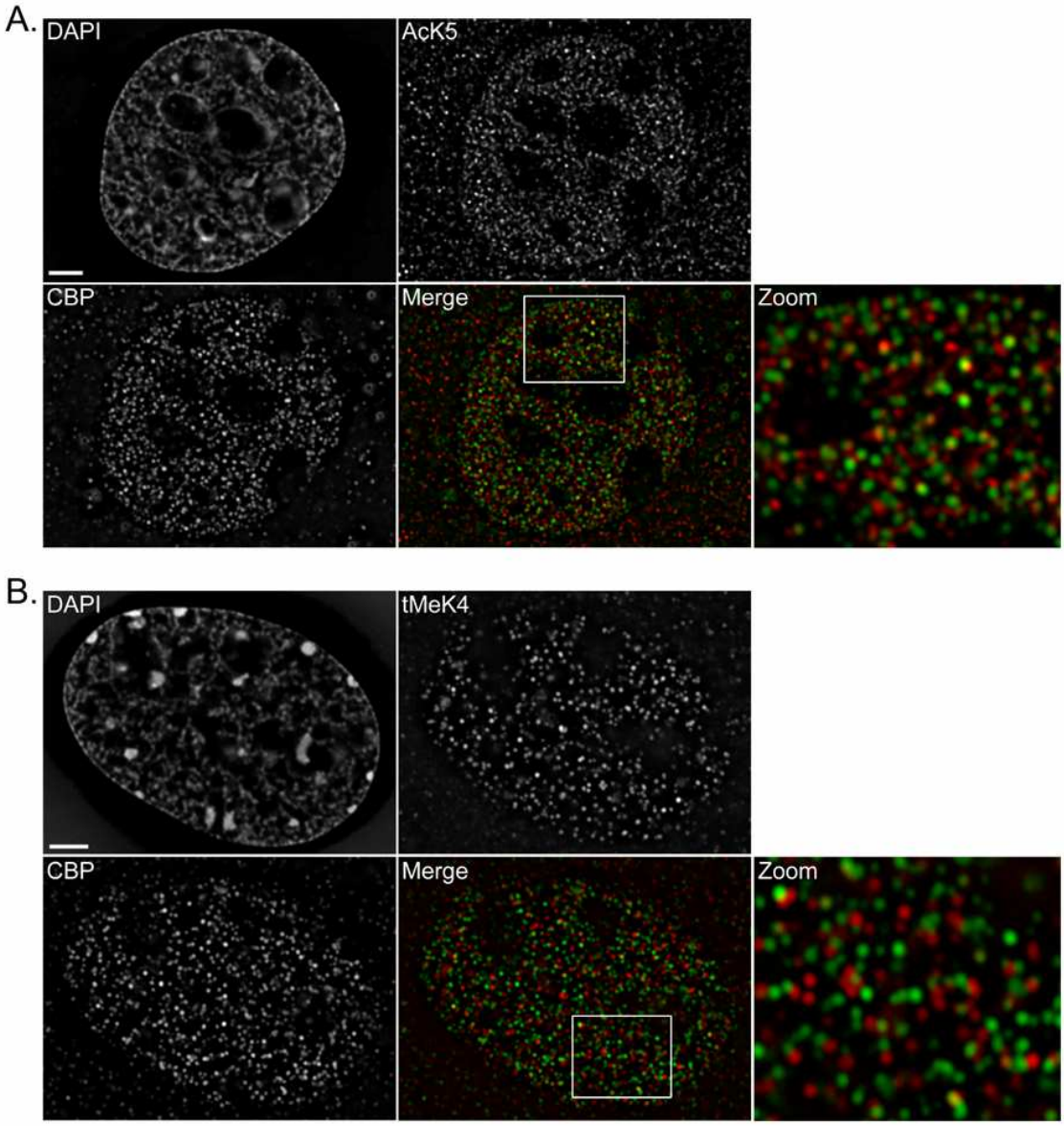


Figure 3

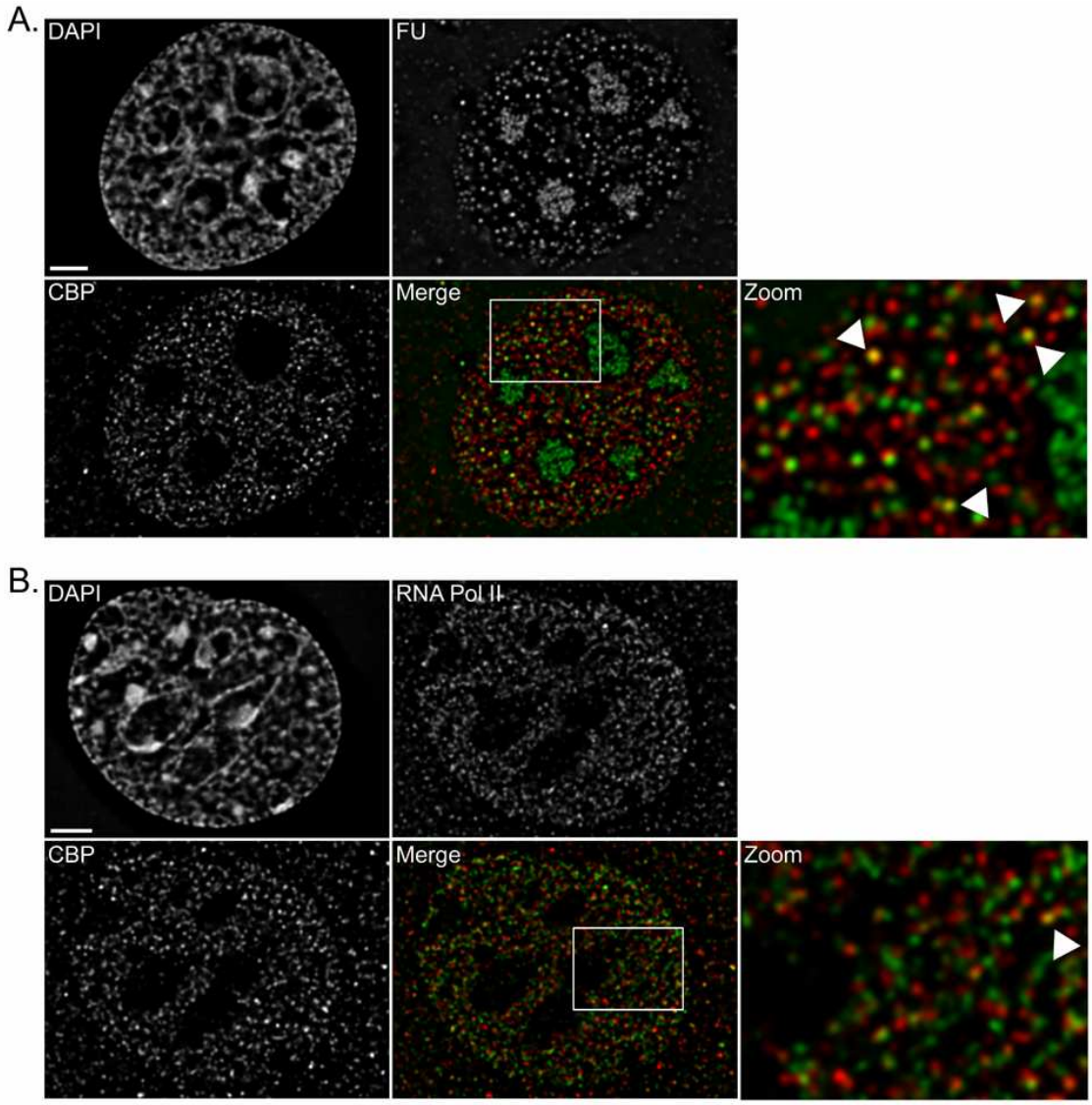


Figure 4

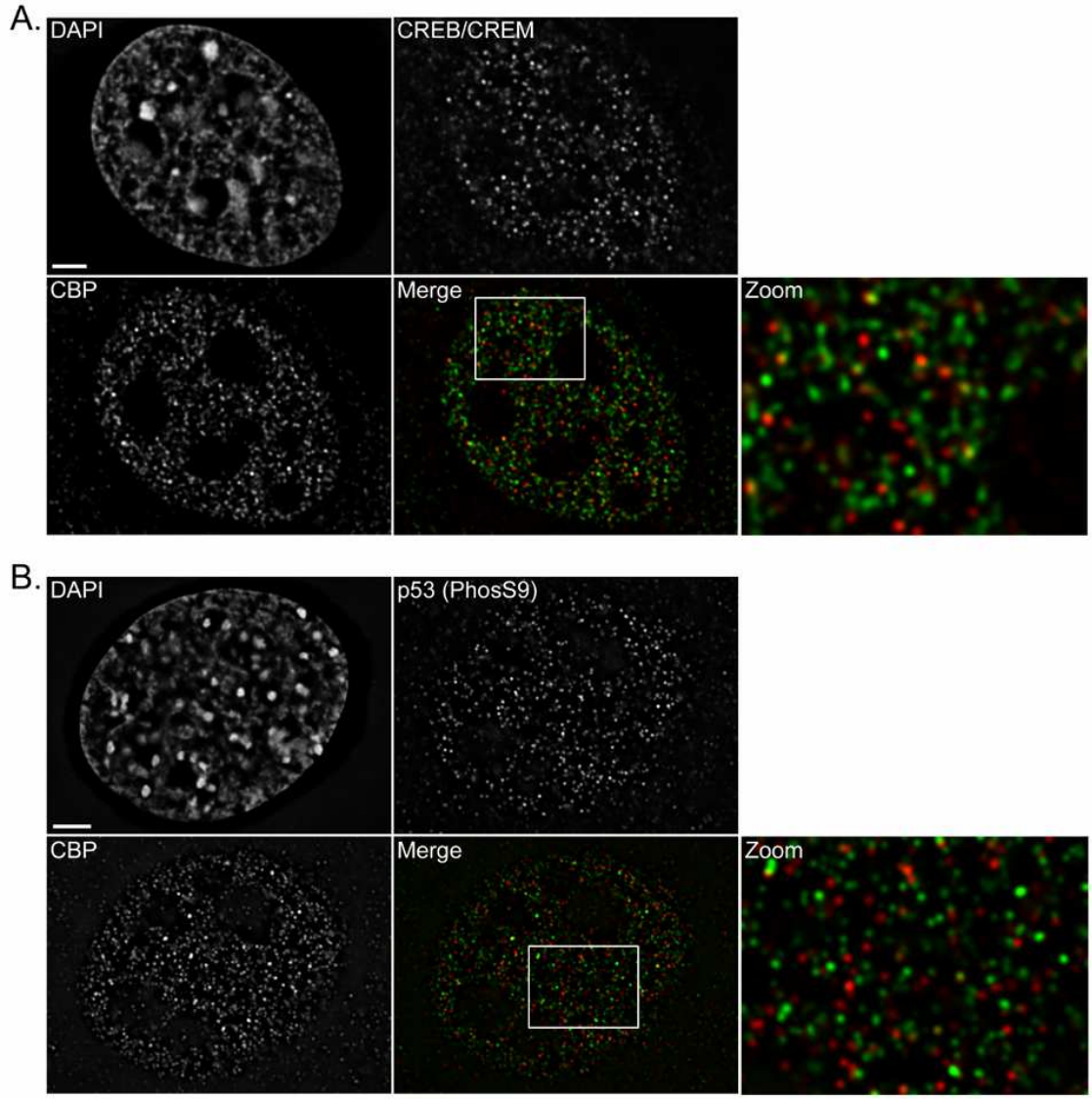

A.
DAPI
CREB/CREM
CBP
Merge
Zoom
B.
DAPI
p53 (PhosS9)
CBP
Merge
Zoom

Figure 5

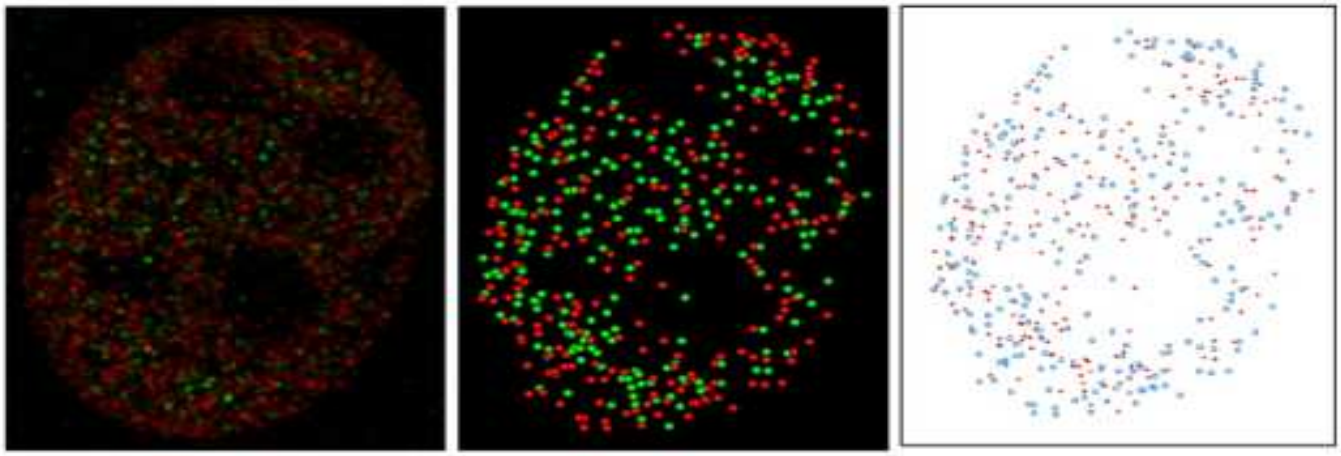
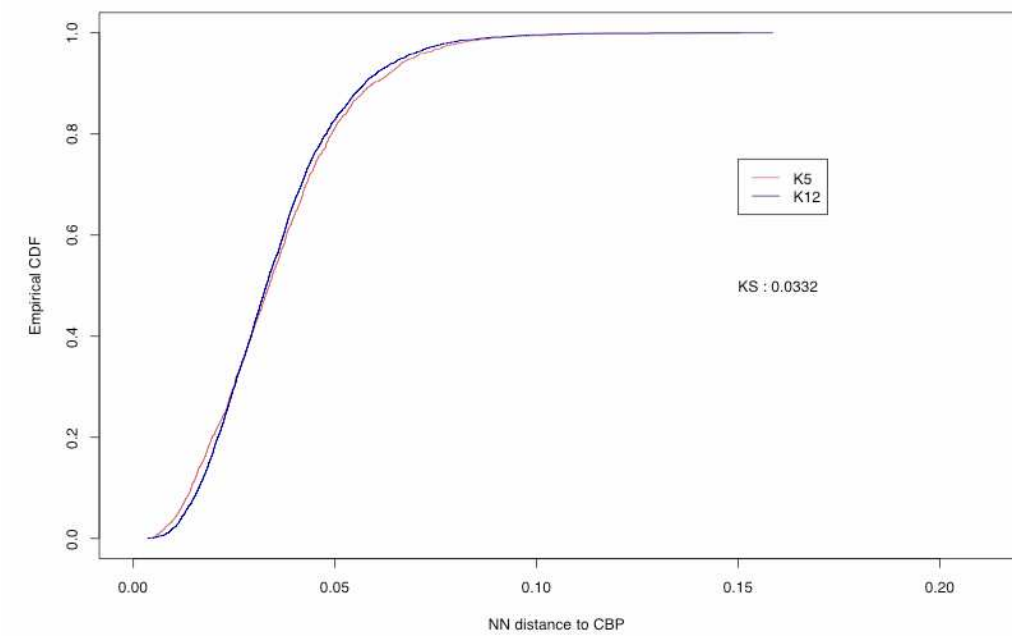
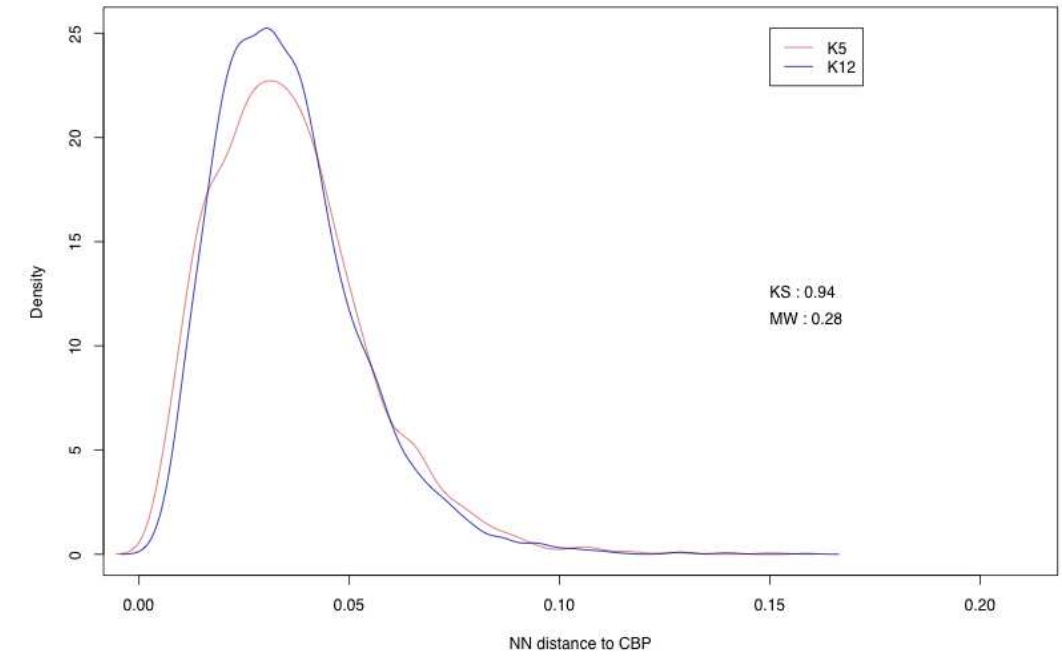


Figure 6

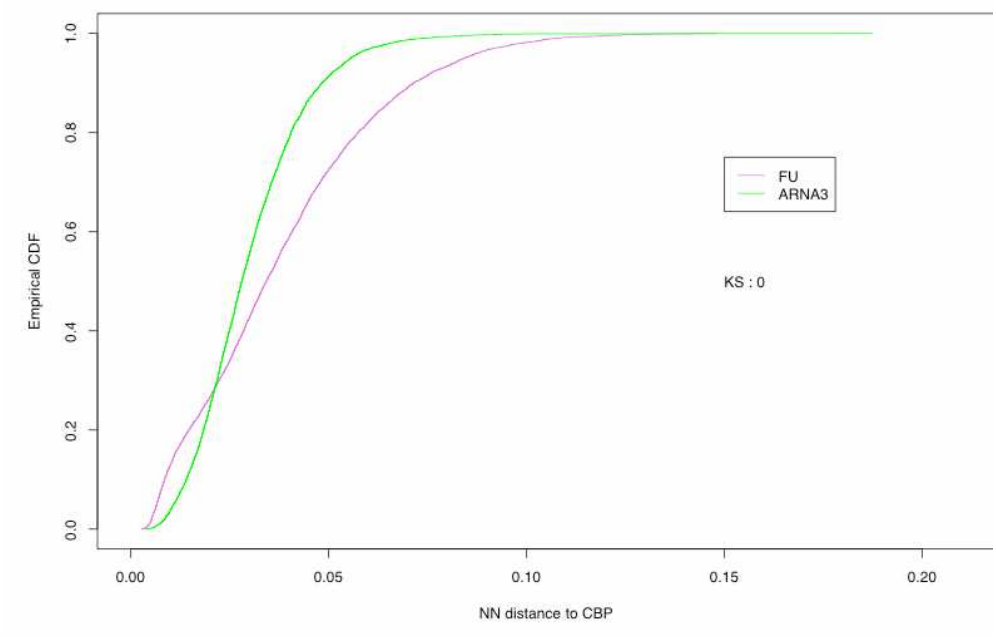
(A)


Empirical CDF
0.0
0.2
0.4
0.6
0.8
1.0
0.00
0.05
0.10
0.15
0.20
NN distance to CBP
K5
K12
KS : 0.0332

(B)


Density
0
5
10
15
20
25
0.00
0.05
0.10
0.15
0.20
NN distance to CBP
K5
K12
KS : 0.94
MW : 0.28

(C)


Empirical CDF
0.0
0.2
0.4
0.6
0.8
1.0
0.00
0.05
0.10
0.15
0.20
NN distance to CBP
FU
ARNA3
KS : 0

(D)

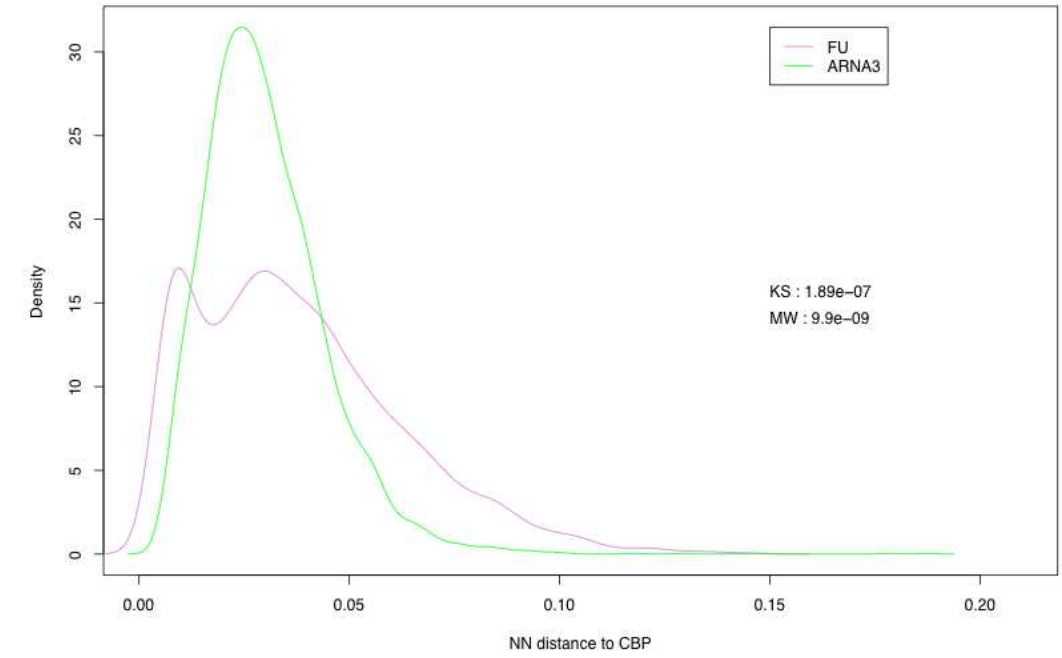

Density
0
5
10
15
20
25
30
0.00
0.05
0.10
0.15
0.20
NN distance to CBP
FU
ARNA3
KS : 1.89e−07
MW : 9.9e−09

Figure 7

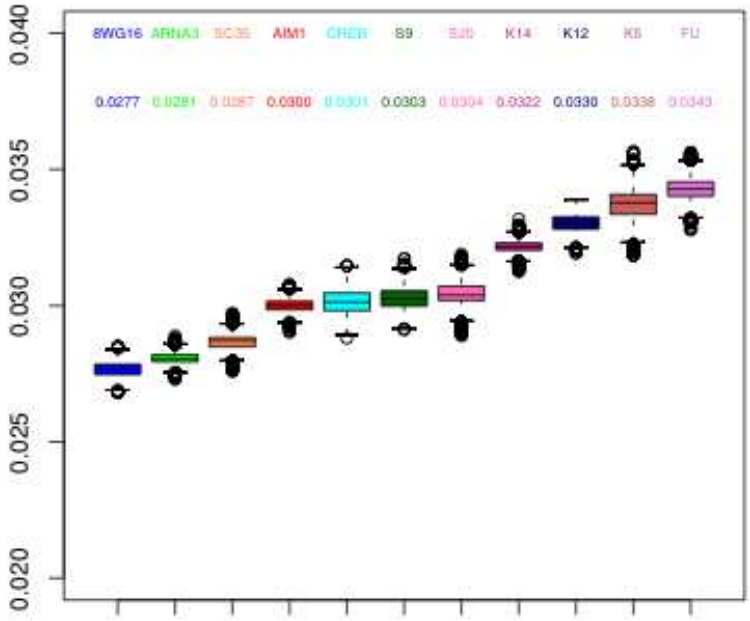

0.040
0.035
0.030
0.025
0.020
8WG16 ARNA3 SC35 AIM1 CREB S9 SJ5 K14 K12 K5 FU
0.0277 0.0281 0.0287 0.0300 0.0301 0.0303 0.0304 0.0322 0.0330 0.0338 0.0343

Figure 8

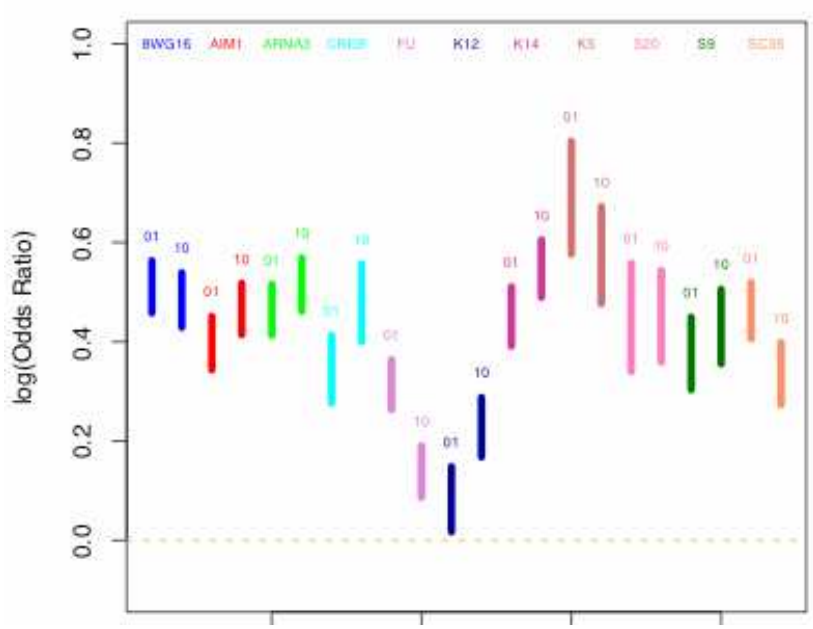


FIGURE LEGENDS

Figure 1: Spatial Relationship Between Endogenous CBP and DNA

Shown here are representative high-resolution (100×) deconvolved images of an interphase 10T1/2 (A) and IM (B) cells immunofluorescently labeled with anti-CBP and counterstained with DAPI. A combined image of the single planes is presented as the Merge with the CBP and DNA shown in green and red respectively. A 3D projection of the entire nucleus is presented (Proj). The linescan (bottom row) demonstrates that endogenous CBP preferentially localizes within euchromatin (less intense DAPI signal) or in regions immediately adjacent to intensely staining heterochromatic regions (intense DAPI signal). Scale bars represent 3µm.

Figure 2: Spatial Relationship Between CBP and Specific Histone H3 Posttranslational Modifications

The spatial relationship between CBP and a known product of its HAT activity, acetylated K5 (H4) (A), or an unrelated histone H3 modification, trimethylated K4 (H3) (B) was investigated. Representative and deconvolved high-resolution images (100×) of interphase 10T1/2 cells immunofluorescently labeled with anti-CBP and either anti-AcK9 or anti-tMeK4 counterstained with DAPI are presented. The merged image presents a single plane from the 3D projection where CBP and AcK9 or tMeK4 are shown in green and red respectively. A region, identified by the white box, is further magnified to show the spatial relationship between CBP and either posttranslational histone modification. Scale bars represent 3µm.

Figure 3: Spatial Relationship Between CBP and Transcription.

The spatial relationship between CBP and transcription as identified by FU incorporation in nascent RNA transcripts (A) or immunofluorescent labeling of RNA polymerase II (8WG16) (B) was investigated. Representative and deconvolved high-resolution images (100×) of interphase 10T1/2 cells immunofluorescently labeled with anti-CBP and either anti-FU or anti-8WG16 counterstained with DAPI are presented. The merged image presents a single plane from the 3D projection where CBP and FU or 8WG16 are shown in green and red respectively. A region, identified by the white box, is further magnified to show the spatial relationship between CBP and either posttranslational histone modification. Scale bars represent 3µm.

Figure 4: Spatial Relationship Between CBP and Regulators of Transcription

The spatial relationship between CBP and CREB/CREM (A) and phosphorylated serine 9 of p53 (B) was investigated. Representative and deconvolved high-resolution images (100×) of interphase 10T1/2 cells immunofluorescently labeled with anti-CBP and either anti-CREB/CREM or anti-p53(phosS9) counterstained with DAPI are presented. The merged image presents a single plane from the 3D projection where CBP and CREB/CREM or p53(phosS9) are shown in green and red respectively. A region, identified by the white box, is further magnified to show the spatial relationship between CBP and either posttranslational histone modification. Scale bars represent 3µm.

Figure 5: Spatial analysis of CBP and acetylated histones in three-dimensionally preserved mouse 10T fibroblast nuclei.

Projected images of CBP (green) and acetylated histone K12 (red) were collected by widefield fluorescence microscopy in 200 nm z-steps followed by deconvolution. (A) The deconvoluted images were processed using the program Image3dV to obtain centroid positions of each fluorescent foci and displayed using the graphics program PREPI. (B) A scatterplot of the extracted centroid positions showing good agreements with the original projected confocal image (C).

Figure 6: Distribution plots of Nearest Neighbor Distances for non-CBP foci to CBP.

Empirical Distribution Function (EDF) and density plots are shown for pairwise comparison of NN distances (see text and Table II for details). The density plot represents density estimates of the NN distance distribution (NCBP to CBP) for nuclear components taken pairwise. KS is the p-value in the Kolmogorov-Smirnov test while MW refers to the Mann-Whitney statistic. (A) EDF plot comparing acetylated Lys5 (K5) and Lys12 (K12) of Histone H4. No statistical difference is observed suggesting that CBP is as close to both K5 and K12 foci (B) Density distribution plot for

the same pairwise comparisons showing a similar trend as in (A). (C) EDF plot comparing NN distances for CBP to Fluoro-uridine (FU) and active PolII (ARNA3). The FU foci distribution is significantly different to the active PolII foci suggesting that CBP is closer to PolII than nascent mRNA transcripts. Note the shape of the FU distribution at smaller CBP nearest neighbor distances indicating that a sub-population of FU foci lie close to CBP foci. (D) Density distribution plot for the same comparison. Note the bomodal distribution for FU indicating a population of FU foci that lie closer to CBP than active PolII and could represent sites of active transcription.

Figure 7: Box-plot of median nearest neighbour distances for different nuclear components to CBP foci.

The vertical axis represents relative distances to the maximal possible distance in microns with the median distances for each foci highlighted. The box represents the interquartile range of these distances. The "whiskers" indicate the 5% and 95% quantiles, and distances observed outside this range are indicated with circles. The nuclear foci analysed for nearest neighbor distances to CBP are: 8WG16: RNA PolII (hypo-phosphorylated); ARNA3: PolII (hyper-phosphorylated); SC35: pre mRNA splicing speckles; AIM1: Aurora B kinase; CREB: Creb/Crem transcriptional activator; S9: p53 phosphorylated Ser9; S20: p53 phosphorylated Ser20; K14: acetylated Lys14 Histone H4; K12: acetylated Lys12 Histone H4; K5 acetylated Lys5 Histone H4; FU: Fluoro-uridine.

Figure 8: Plot of 'attraction' or association of different nuclear components with CBP foci.

The vertical axis represents the Log of the excess odds ratio over randomness with a value of zero equivalent to random association. The different nuclear body components are labelled as in Fig. 3. Non-CBP 'association' to CBP is labelled 01 with CBP 'association' to NCBP labelled 10. The bars represent a 95% confidence interval for the mean association (see text for further text). Note the significant difference in 'probability of association' between K5 and K12.

SUPPLEMENTARY INFORMATION

Supplementary Figure 1: Empirical Distribution Function plots of Nearest Neighbor Distances for non-CBP foci to CBP. Kolmogorov-Smirnov statistics were used to compare empirical distribution functions (edfs) for CBP to NCBP nearest-neighbour distances, with computation of p-values performed using standard asymptotic reasoning, verified using randomization procedures. Each pair-wise plot compares the named components in terms of their nearest neighbor distances to CBP foci. The labels for each non-CBP componenst is as follows: WG16: RNA PolII (hypo-phosphorylated); ARNA3: PolII (hyper-phosphorylated); SC35: pre mRNA splicing speckles; AIM1: Aurora B kinase; CREB: Creb/Crem transcriptional activator; S9: p53 phosphorylated Ser9; S20: p53 phosphorylated Ser20; K14: acetylated Lys14 Histone H4; K12: acetylated Lys12 Histone H4; K5 acetylated Lys5 Histone H4; FU: Fluoro-uridine. Distance have been normalised and are therefore relative distances. The Ks statistic is given in the box and refers to how significant the observed differences.

Movie 1: Centroid Positions of Acetylated K14 and CBP in Indian muntjac fibroblasts. IM cells were paraformaldehyde fixed, permeabilized, immunofluorescently labeled with anti-AcK14 and anti-CBP and 3D image series were acquired at 200nm intervals. Deconvolved images were processed in Image3D and centroid positions were assigned to both AcK14 (red) and CBP (green) foci and viewed in PREPI. Shown here is a 360° rotation of the centroid positions that presents the 3D spatial relationships between AcK14 and CBP foci.