Ultrastructural study of transcription factories in mouse erythroblasts

Christopher H. Eskiw*,‡ and Peter Fraser

Nuclear Dynamics Laboratory, The Babraham Institute, Babraham Research Campus, Cambridge CB22 3AT, UK

*Present address: Center for Cell and Chromosome Biology, Heinz Wolff Building, Brunel University, Uxbridge, Middlesex UB8 3PH, UK

‡Author for correspondence (Christopher.Eskiw@brunel.ac.uk)

Accepted 16 June 2011
Journal of Cell Science 124, 3676–3683
© 2011. Published by The Company of Biologists Ltd
doi: 10.1242/jcs.087981

Summary

RNA polymerase II (RNAPII) transcription has been proposed to occur at transcription factories; nuclear focal accumulations of the active, phosphorylated forms of RNAPII. The low ratio of transcription factories to active genes and transcription units suggests that genes must share factories. Our previous analyses using light microscopy have indicated that multiple genes could share the same factory. Furthermore, we found that a small number of specialized transcription factories containing high levels of the erythroid-specific transcription factor KLF1 preferentially transcribed a network of KLF1-regulated genes. Here we used correlative light microscopy in combination with energy filtering transmission electron microscopy (EFTEM) and electron microscopy in situ hybridization (EMISH) to analyse transcription factories, transcribing genes, and their nuclear environments at the ultrastructural level in ex vivo mouse foetal liver erythroblasts. We show that transcription factories in this tissue can be recognized as large nitrogen-rich structures with a mean diameter of 130 nm, which is considerably larger than that previously seen in transformed cultured cell lines. We show that KLF1-specialized factories are significantly larger, with the majority of measured factories occupying the upper 25th percentile of this distribution with an average diameter of 174 nm. In addition, we show that very highly transcribed genes associated with erythroid differentiation tend to occupy and share the largest factories with an average diameter of 198 nm. Our results suggest that individual factories are dynamically organized and able to respond to the increased transcriptional load imposed by multiple highly transcribed genes by significantly increasing in size.

Key words: Transcription factory, Electron microscopy

Introduction

The mode of transcription in which RNA polymerase II (RNAPII) scans the genome for available promoters, followed by binding, transcription initiation, and tracking along the template during elongation has been challenged by recent studies examining the location of active transcription units in the nucleus. Nascent transcript labelling studies indicate that RNAPII transcription occurs primarily, if not exclusively, at transcription factories, which are focal accumulations of phosphorylated RNAPII (Iborra et al., 1996; Pombo et al., 1999). Several RNA immunoFISH (fluorescence in situ hybridization combined with immunofluorescence) studies examining nascent transcript signals for over a dozen genes have shown that they invariably overlap with RNAPII transcription factories (Mitchell and Fraser, 2008; Osborne et al., 2004; Osborne et al., 2007; Ragoczy et al., 2006; Schoenfelder et al., 2010), which supports the proposal that RNAPII transcription occurs exclusively in factories. Estimations of the number of serine-5-phosphorylated RNAPII (RNAPII-PS5) transcription factories in several mouse tissues using immunofluorescence light microscopy have indicated that each nucleus contains only several hundreds of factories. Genome-wide gene analyses have shown that these tissues actively express greater than 6000 genes (>12,000 alleles) (Schoenfelder et al., 2010). This suggests that actively transcribed genes share these sites, and indeed nascent RNA FISH signals from genes separated by tens of millions of bases in cis, or from different chromosomes (in trans), frequently overlap at shared transcription factories (Mitchell and Fraser, 2008; Osborne et al., 2004; Osborne et al., 2007; Papantonis et al., 2010; Schoenfelder et al., 2010). In each of these cases, close spatial proximity of colocalizing genes has been confirmed by chromosome conformation capture (3C) assays.

The high frequencies of observed factory sharing coupled with the failure to observe transcription outside factories, suggests that genes either move to factories immediately after RNAPII binding and transcript initiation, or that RNAPII contact and initiation of transcription occur after relocation to a factory. Induced genes rapidly associate with factories and colocalize with constitutively transcribed genes both intra- and interchromosomally, suggesting that many alleles move to preassembled factories when they become active, rather than nucleate their own factories (Dhar and Wong-Riley, 2010; Osborne et al., 2007; Papantonis et al., 2010). 3C analyses confirmed that intra- and interchromosomal contacts are induced coincident with transcriptional activation, and 3D DNA FISH inter-gene measurements suggested that long-range associations at factories occur through movement of the induced gene toward a preassembled factory (Osborne et al., 2007). This dynamic model is supported by the finding that long-range intra- and interchromosomal interactions between genes are lost and constitutively active genes dissociate from transcription factories when transcription initiation is blocked by heat shock, whereas RNAPII-PS5 foci are still observable in the absence of transcription (Mitchell and Fraser, 2008). Shortly after return to 37℃, active genes were once again transcribed at factories,

which indicates that gene association is dynamic and transcription-dependent. Detailed analyses of the long-range contacts between two distal induced genes revealed dynamic contacts between the two genes that followed the progress of the polymerase complex through the transcription unit, consistent with a mechanism in which the two transcription units slid past each other during transcription, as if being reeled through a shared factory (Papantonis et al., 2010).

Transcription factories appear to specialize. Episomal constructs driven by various promoters resulted in the clustering of episomes at a subset of factories dependent upon the similarities of their promoters (Xu and Cook, 2008). Constructs with the same promoter driving different reporters tended to colocalize during transcription, whereas those with heterologous promoters tended to localize to different sites. Similarly, networks of endogenous genes tend to cluster. A genome-wide enhanced 4C screen (e4C; a variant of 3C) for genes that share factories with the mouse α- (*Hba*) and β-globin (*Hbb*) genes revealed intra- and interchromosomal associations with hundreds of other gene loci from nearly all chromosomes in erythroblast nuclei (Schoenfelder et al., 2010). These studies revealed spatial networks of preferred transcription partners that were overrepresented in genes regulated by the erythroid-specific Kruppel-like transcription factor, KLF1. Immunolocalization of KLF1 in erythroblast nuclei revealed 30–40 foci, the vast majority of which overlapped with RNAPII factories. Investigation of the position of active alleles of KLF1-regulated genes suggested that they had a higher probability of being active if associated with a KLF1 factory and, when there, were clustered with several other KLF1-regulated genes. Collectively, these data demonstrate the compartmentalization of the transcriptional machinery in mammalian nuclei and indicate that the genome is non-randomly organized to take advantage of specialized sites that might be optimized for efficient transcription of co-regulated gene networks.

Previous studies using energy filtering transmission electron microscopy (EFTEM) have revealed that transcription factories in HeLa nuclei are nitrogen-rich proteinacious structures with a mass of approximately 12 MDa and a mean diameter of 87 nm (Eskiw et al., 2008). Here, we used correlative microscopy on ultrathin physical sections in combination with EFTEM to show the size and mass distribution of transcription factories in mouse foetal liver erythroblasts. We showed that factories are located in regions of low chromatin density with an average of 545 factories per nucleus. EFTEM demonstrated that these foci are protein-rich domains with a mean diameter of 130 nm, containing on average 24 MDa of protein. Using this method, we also compared the size and mass distribution of factories associated with the transcription factor KLF1. We also investigated factories containing the highly transcribed *Hbb* and *Hba* gene loci using a novel electron microscopy in situ hybridization (EMISH) technique. Together, these data characterize transcription factories in primary mouse erythroblasts tissues at the ultrastructural level and begin to describe the physical diversity of the nuclear structures involved in transcribing different subsets of genes.

Results

Transcription factories in mouse erythroblasts

We used immunolabelling of active, RNAPII-PS5 in embryonic day 14.5 (e14.5) mouse foetal liver cells to determine the distribution of transcription factories (Fig. 1). In whole cells (Fig. 1A–C), many factories are discernable as foci; however, the numbers of these are difficult to assess due to the high amount of signal present in other focal planes and the poor *z*-resolution of light microscopy. To improve image resolution, immunostained cells were embedded in epoxy resin and 90 nm ultrathin sections prepared. Imaging of factories in ultrathin physical sections greatly improves *z*-resolution to the thickness of the section, through removal of out-of-focus light. Imaging by this method demonstrates that RNAPII-PS5 is located in foci primarily in regions of low chromatin density (Fig. 1D–F). The numbers of transcription factories were counted in 34 medial, physical sections from different nuclei (supplementary material Table S1; raw counts per section). Performing stereological calculations to account for the Holmes effect [see Materials and Methods, and (Eskiw et al., 2008)], we determined that there were an average of 545±275 factories per nucleus with a range of 120–1044 (Fig. 1G). Previous studies have shown there are a minimum of 6000 active genes (potentially 12,000 active alleles) at this stage of erythroid development (Schoenfelder et al., 2010). Even if only half of the potentially active alleles are transcribing at any given moment in a single cell, these results suggest that the average transcription factory contains approximately ten transcription units.

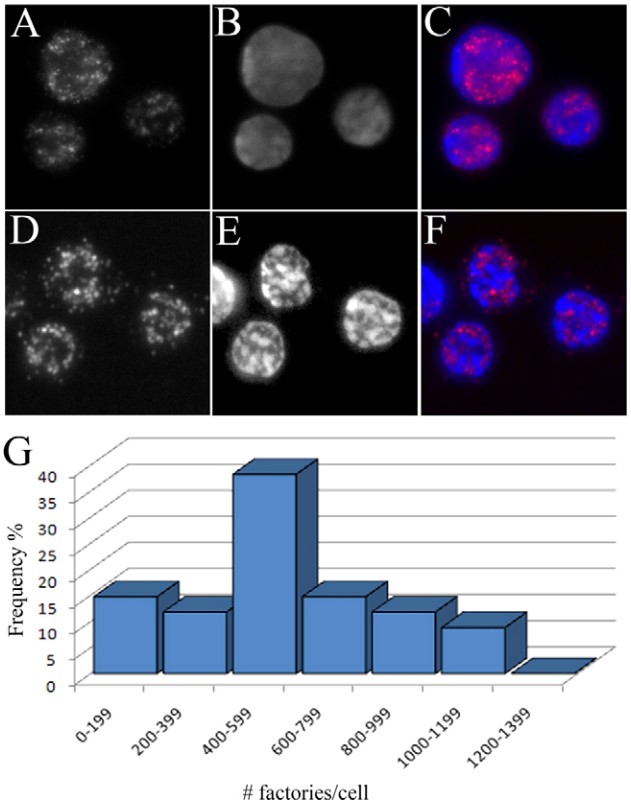

Fig. 1. RNPII-PS5 transcription factories. Ex vivo e14.5 foetal liver cells were immunolabelled for RNPII-PS5. (**A**) Distribution of transcription factories in whole cells. (**B**) Chromatin is counterstained with Hoechst 33342 (blue). **(C)** Merged image from A (red) and B (blue). (**D**) Fluorescence micrographs of 90-nm physical section of e14.5 foetal liver cells labelled for RNPII-PS5. (**E**) DNA in section counterstained with Hoechst 33342. (**F**) Merged image of D and E, which demonstrates the distribution of transcription factories (red) with regards to chromatin (blue). (**G**) Distribution of the number of transcription factories per nucleus after extrapolation from physical sections and correction for the Holmes effect.

Journal of Cell Science

We then performed correlative microscopy in combination with EFTEM to quantify the size and mass distribution of transcription factories at the ultrastructural level. Briefly, after immunolabelling of e14.5 liver cells, embedding and ultrathin sectioning, fluorescence images of transcription factories in physical sections were collected (Fig. 2A–C) and the positions of these cells marked. EFTEM was then performed on the transcription factories identified within these nuclei. EFTEM detects the energy loss of electrons that have interacted with specific atoms within the specimen and provides quantitative elemental information on the composition of structures (Eskiw et al., 2008). For example, electrons that interact with phosphorus (P) atoms within the specimen lose 132 e^-V of energy, whereas electrons that interact with nitrogen (N) atoms lose 409 e^-V of energy. Images, or 'maps', of P and N are then generated to demonstrate the position of these atoms within the specimen. First, a low magnification mass-sensitive (Fig. 2D) image was collected and overlaid with the fluorescence micrographs to indicate the location of transcription factories within the specimen (Fig. 2E). High magnification P maps (Fig. 2G) and N maps (Fig. 2H) of the regions corresponding to factories were then collected and false coloured red (P) and green (N). Merged images identify structures that are phosphorous-rich, such as chromatin or ribonucleoprotein complexes (RNP) or nitrogen-rich proteinacious structures. We identified transcription factories by correlating the RNAPII-PS5 fluorescence signal overlaid on the high magnification P and N maps (Fig. 2I,J). Previous studies (Eskiw et al., 2008) have shown that transcription factories are N-rich structures, composed primarily of protein, that make contact with the surrounding chromatin (Fig. 2J). Consistent with previous work, we found large N-rich structures underlying the RNAPII-PS5 fluorescence signals and defined the boundaries of transcription factories as the points at which the N:P ratio changed dramatically (supplementary material Fig. S1). We analysed 30 transcription factories identified by this method. Few P atoms were detected within the transcription factory structure, consistent with structural RNA components or the heavily phosphorylated C-terminal domain (CTD) of the large subunit of RNAPII. However, the level of P within the structure was not consistent with the presence of chromatin within the factory, in agreement with previous studies indicating that transcription occurs at the surface of factories (Eskiw et al., 2008). P-rich structures present near the surface of factories are consistent with RNP particles and chromatin. We defined transcription factories based on the continuity of the N-rich structures underlying the fluorescence signals. Factories had a normal distribution of diameters with a mean of 130 nm, after correction for the Holmes effect (Fig. 3). This mean diameter is significantly larger than the 87 nm mean diameter for transcription factories in HeLa cell nuclei (Eskiw et al., 2008).

EFTEM provides quantitative information on elemental composition

We used the EFTEM images to calculate approximate protein and nucleic acid content of transcription factories through comparison with structures of known composition. Nucleosomes within the image are recognizable as N- and P-rich structures with a diameter of 11 nm and a N to P signal ratio of 1:1, consistent with previous EFTEM analyses of physical sections of human cells and purified nucleosomes (Bazett-Jones et al., 1999; Hendzel et al., 1999). Using nucleosomes as internal structures of known composition, we calculated the mass of protein within each factory. Assuming that most nucleosomes contain two of each canonical core histone (H2A, H2B, H3 and H4) (supplementary material Table S2A)

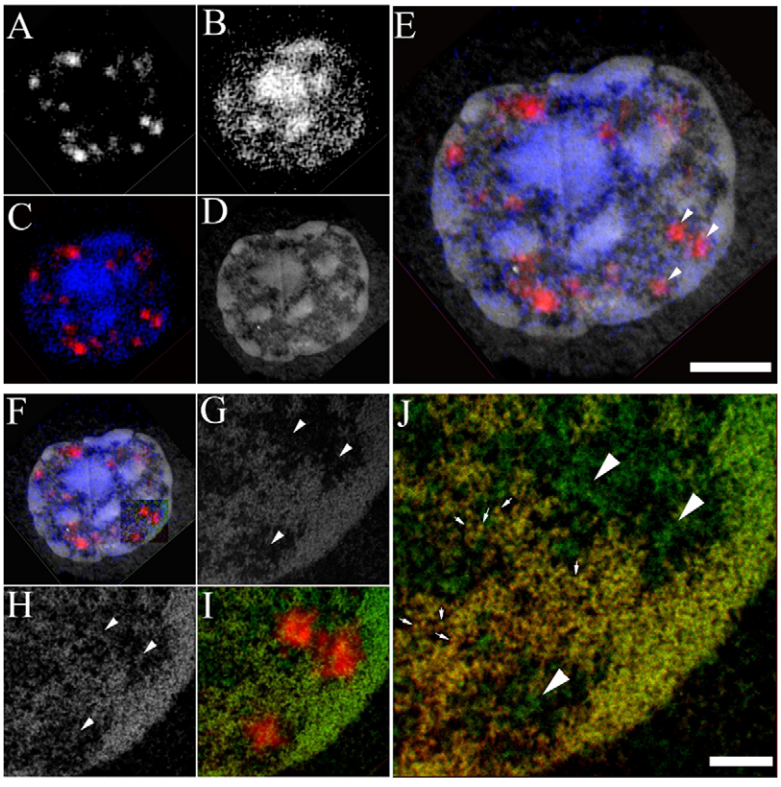

Fig. 2. Examination of transcription factories by correlative microscopy and EFTEM. Following immunolabelling for RNPII-PS5 to mark the location of transcription factories, erythroid cells from e14.5 foetal livers were embedded in resin, thin sectioned to 100 nm and imaged by fluorescence microscopy. (**A**) Distribution of transcription factories in a single physical section. (**B**) Chromatin within the nuclei is counterstained with Hoechst 33342 dye. (**C**) Merged image of transcription factories (red) and chromatin (blue). (**D**) Mass-sensitive EFTEM image of the same nucleus. (**E**) Panels C and D were merged to show the location of transcription factories at high resolution. Arrowheads indicate the structures imaged at higher magnification. (**F**) Same image as E with area of high magnification indicated. (**G**) P map of region designated in F, containing factories. (**H**) Nitrogen map of the same region as G. Arrowheads show regions identified by fluorescence. (**I**) P maps were false coloured red and N maps were false coloured green and merged, then overlaid with the fluorescence image from C. (**J**) High magnification of the region with the RNPII-PS5 fluorescence signal removed. Arrowheads indicate the structures identified by fluorescence. Small arrows indicate the locations of individual nucleosomes. Yellow structures represent chromatin or RNP-containing structures; green structures represent protein-based structures. Scale bars: 1 μm (E), 200 nm (J).

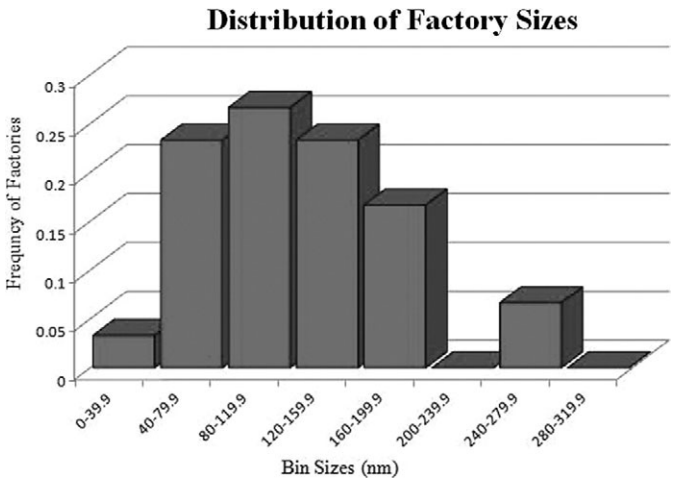

Fig. 3. Transcription factories in erythroblasts are not uniform in size. The diameters of factories identified by correlative microscopy and EFTEM were measured. Measurements were binned into groups of 40 nm increments and the frequencies of factories in each bin plotted. Data demonstrate that the distribution of factory sizes is normal, with a mean diameter of 130 nm.

wrapped by 146 bp of DNA with a GC content of 42%, we determined that there are 1980 N atoms and 292 P atoms per nucleosome (supplementary material Table S2B). Of these, 1438 N atoms would be contributed from the histone octamer and 542 from DNA. These values were used to determine the arbitrary intensity units per atom within each map (see Materials and Methods). A minimum of five nucleosomes per image were measured from random points to eliminate statistical anomalies and variations between nucleosomes. For factory masses, we determined the nucleosomal equivalents of nitrogen atoms and extrapolated the mass of protein contained within the structures. The elemental composition of transcription factories in mouse erythroblast nuclei was similar to that of transcription factories identified in HeLa cells; the structure being primarily protein-based with chromatin contacts on the surface and P-rich granules within the core. Erythroblast factories have an average protein mass of 26 MDa, considerably greater than HeLa cells.

KLF1-specialized transcription factories are larger than the population average

Many erythroid-specific genes are regulated by the erythroid kruppel-like transcription factor, KLF1. Previous studies have demonstrated that KLF1-regulated genes preferentially cluster at a subset of specialized, KLF1-enriched transcription factories (Schoenfelder et al., 2010). We immunolabelled e14.5 mouse erythroblasts for RNAPII-PS5 and KLF1, embedded them in epoxy resin and generated ultrathin sections. Fluorescence micrographs of RNAPII-PS5 factories (Fig. 4A), KLF1- (Fig. 4B) and H33342-counterstained chromatin (Fig. 4C) were merged (Fig. 4D). As before, RNAPII-PS5 foci were found in regions of low chromatin density. KFL1 was also in foci in regions of low chromatin density, with some KLF1 foci colocalized with transcription factories. We found an average of 60 ± 40 (range 8–132) KLF1-specialized transcription factories per cell after correction for the Holmes effect. Chromatin immunoprecipitation sequencing data predicts that there are between 950 and 1380 KLF1-regulated genes (Tallack et al., 2010), which provides further evidence that these genes share the limited number of KLF1-associated transcription factories.

We then used EFTEM to examine transcription factories enriched in KLF1. Fluorescence images of RNAPII and KLF1 were overlaid with low magnification mass-sensitive images (Fig. 4E) to show the areas of interest, which were then analysed at higher resolution. P (Fig. 4F) and N (Fig. 4G) maps were false coloured red and green, respectively, and merged (Fig. 4H). Line scans through KLF1-associated transcription factories demonstrated that these structures are protein-based and nucleic acid-poor (Fig. 4K), similar to other factories. However, they were significantly larger ($P=0.0021$; two-way Mann-Whitney test) with a mean diameter of 174 nm (supplementary material Fig. S2). KLF1-containing transcription factories also had significantly more mass of protein (36 MDa).

Factories transcribing the globin genes are larger than the population mean

To examine factories associated with the highly transcribed *Hbb* and *Hba* gene loci, we developed a variation of the RNA immunoFISH protocol, called electron microscopy in situ hybridization (EMISH), which allows for the detection of nascent transcripts while preserving nuclear architecture at the ultrastructural level (supplementary material Fig. S3). Erythroblasts were incubated with hapten-containing oligonucleotide probes, followed by antibody detection of nascent transcripts and transcription factories. Following detection, cells were again embedded in resin, thin sectioned, and subjected to fluorescence microscopy to identify the nascent RNA and RNAPII-PS5 signals (Fig. 5). Both *Hba* and *Hbb-b1* nascent RNA signals were detectable in physical sections, and >97% were associated with transcription factories, which was consistent with previous studies (Osborne et al., 2004; Schoenfelder et al., 2010). EFTEM was then performed on transcription factories associated with the globin nascent transcripts. P and N maps were again collected and aligned with the mass-sensitive and fluorescence images (Fig. 5F). High resolution images demonstrated a protein-rich structure, corresponding to the factory fluorescence, adjacent to RNP particles (Fig. 5G). These RNP particles corresponded to the fluorescence signal identified as *Hbb-b1* nascent transcripts. Analysis of transcription factories adjacent to *Hbb-b1* or *Hba* transcripts were indistinguishable from one another and together had a mean diameter of 194 nm, which was significantly larger ($P=0.0028$; two-way Mann-Whitney test) from the mean diameter of the population (supplementary material Fig. S2). EFTEM comparison of the ultrastructure of cells treated by EMISH and control cells demonstrated no significant differences in nuclear morphology or chromatin architecture, which indicated that this technique is suitable for analysis of transcription sites (supplementary material Fig. S3). Quantification of the mass from these factories demonstrated that they contained ~38 MDa of protein.

Genes share transcription factories at the ultrastructural level

Next, we used EMISH to investigate gene sharing at transcription factories. The results of several RNA immunoFISH studies using light microscopy led to the proposal that active genes can share factories. Although 3C-like techniques support these conclusions, others have suggested that light microscopy does not have the resolution to detect colocalization at factories (Brown et al., 2008; Lawrence and Clemson, 2008). We performed EMISH with probes for *Hba* and *Hbb* nascent transcripts (Fig. 6). We

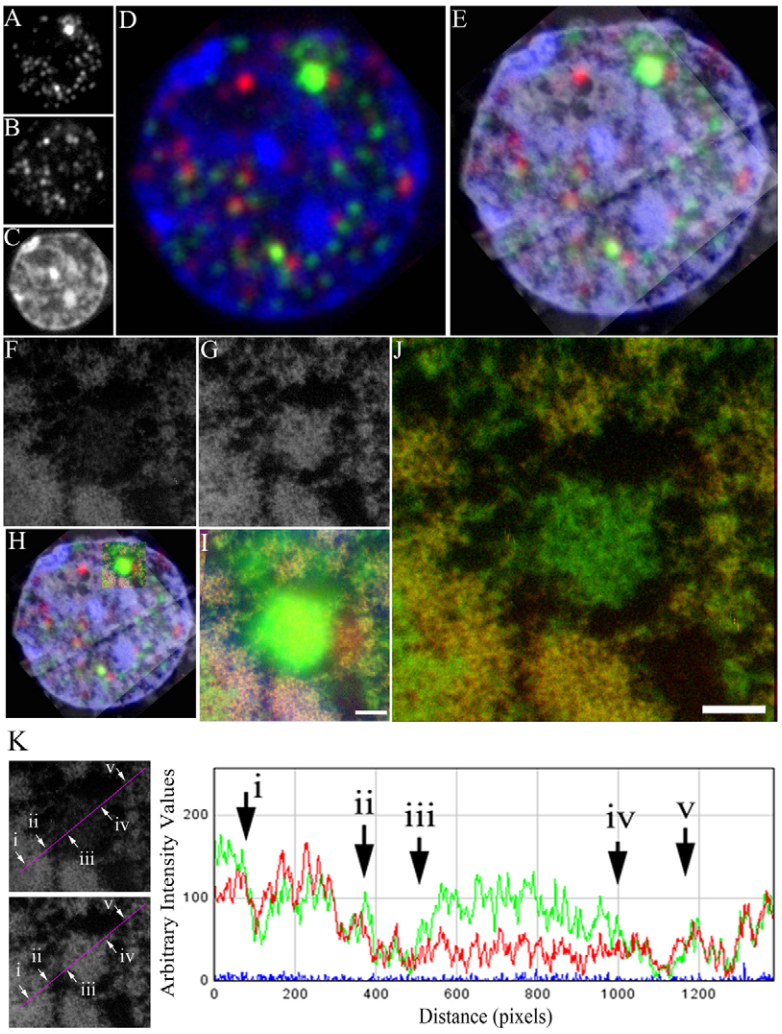

Fig. 4. Transcription factories enriched in the KLF1 are larger than the population mean. Transcription factories enriched in KLF1 were identified by correlative microscopy, and their structure analysed by EFTEM. (**A**) Immunolocalization of RNPII-PS5 in 100-nm thick physical sections. (**B**) Immunolocalization of KLF1. (**C**) Chromatin is counterstained in sections with Hoechst 33342. (**D**) Merged image indicating the localization of RNPII-PS5 (green), KLF1 (red) and chromatin (blue). (**E**) Merged image from D overlaid with the low magnification mass-sensitive electron micrograph of the same section. (**F**) P map of a region containing a KLF-enriched transcription factory. (**G**) N map of the same region imaged in F. (**H**) Overlay of the fluorescence micrographs with the low magnification mass-sensitive image combined with high magnification P and N maps. Region shown in F and G is indicated. (**I**) Merged image of false coloured P (red) and N (green) maps and fluorescence micrograph. (**J**) Merged P (red) and N (green) maps of the same region in I without fluorescence overlay showing the KLF1-enriched transcription factory. Chromatin appears yellow in colour; protein-rich structures are seen as green. (**K**) Line scans (magenta line) were performed on P (upper panel) and N (lower panel) maps through the identified factory, and the arbitrary intensity values plotted (N green, P red). Regions of interest on the maps and graph are indicated by arrows: (i) interface between the granular component of the nucleolus and heterochromatin, (ii) interface between heterochromatin and the N- and P-depleted aqueous nucleoplasm, (iv) boundary between the nucleoplasm and a Klf1-containing transcription factory, (v) boundary between the same Klf1-containing transcription factory and the nucleoplasm, (vi) interface between the nucleoplasm and chromatin. The pixels between arrows iii and iv demonstrate the relatively high N:P ratio within the factory. Scale bars: 200 nm.

located nuclei in the ultrathin sections in which the nascent RNA signals overlapped and again performed EFTEM. Fig. 6 clearly shows two overlapping nascent RNA signals for *Hbb* and *Hba* with a single underlying N-rich structure indicative of a transcription factory, which strongly supports the previous conclusion that genes share factories. Line scans indicate that this structure is N-based and low in P content (Fig. 6J), similar to factories identified previously. We confirmed that the size and composition of factories transcribing these gene pairs were indistinguishable from those identified from single-gene EMISH analyses.

Discussion

Here, we have shown that transcription factories in murine erythroblast nuclei are significantly larger than those reported in HeLa cell nuclei. To investigate the ultrastructure of sites of nascent transcript synthesis of specific genes, we developed a new and relatively simple EMISH method. We used EMSIH and correlative EM to show that specialized factories enriched in KLF1, as well as those factories transcribing the *Hba* and *Hbb* genes, are larger than the population mean. Similarly to HeLa cells (Eskiw et al., 2008), murine erythroblasts contain transcription factories that are primarily protein-based structures that contain very little phosphorus-rich material within the core of the structure. These protein-based factories make intimate contact with the surrounding chromatin environment with adjacent RNP.

We were successful in identifying sites of specific nascent transcript synthesis, through the development of EMISH, to examine factories responsible for transcribing specific genes. From these analyses, we demonstrated that transcription factories specializing in the transcription of genes involved in erythroid differentiation form a distinct subset of larger transcription factories.

To demonstrate sites of nascent transcript synthesis of specific genes we developed a new and relatively simple method that allows for ultrastructural analysis. This method is based on chemical fixation in formaldehyde of ex vivo foetal liver cells to preserve structure and the addition of single-stranded DNA (ssDNA) probes against intronic sequences to localize specific nascent transcripts. This method, in combination with EFTEM, provides the first high-resolution images of the environment in which specific genes are transcribed and provides quantitative information on the transcription factories responsible for their synthesis.

Previous studies had identified between 30–40 KLF1-enriched transcription factories in e14.5 erythroid cells (Schoenfelder et al., 2010). This was derived using confocal imaging of whole cells and performing deconvolution of the images. Here, we used physical sections to estimate the number of KLF1-enriched factories and

Journal of Cell Science

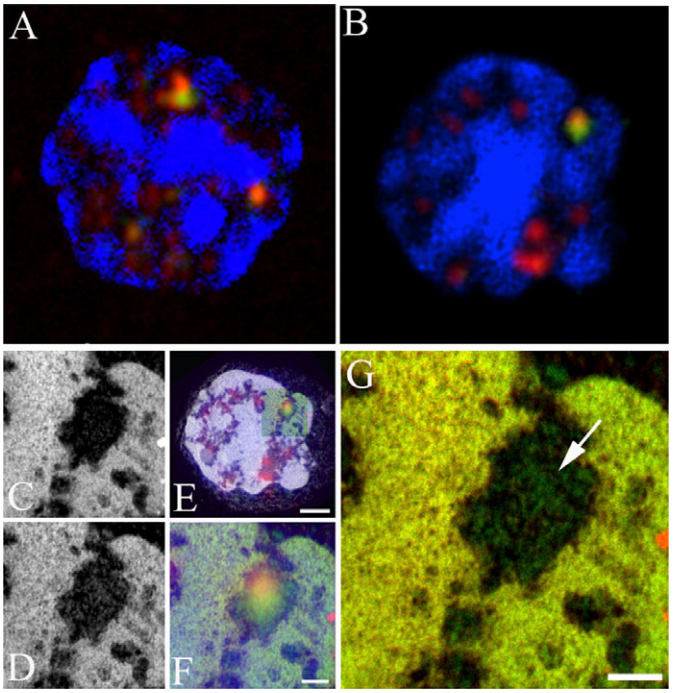

Fig. 5. Nascent transcripts are associated with protein rich transcription factories. (**A**) Day 14.5 foetal liver cells were immunolabelled for RNPII-PS5 (red) and incubated with probes against *Hba* transcripts (green), embedded, thin sectioned and imaged by fluorescence microscopy. Chromatin in the section is counterstained with Hoechst 33342 (blue). (**B**) Day 14.5 foetal liver cell labelled for RNPII-PS5 (red) and incubated with probes against *Hbb-b1* (green) transcripts followed by embedding, thin sectioning and imaging by fluorescence microscopy. Chromatin is counterstained with Hoechst 33342 (blue). (**C**) P map of a subregion of the cell in B containing the *Hbb-b1* signal imaged by EFTEM. (**D**) Corresponding N map of the same region in C. (**E**) Overlay of the low magnification mass-sensitive image with the fluorescence image demonstrating the location of the *Hbb-b1* and RNPII-PS5 signals. (**F**) P map and N map were false coloured red and green, respectively, and then merged with the fluorescence image to demonstrate the position of the transcripts and factory within the nucleus. (**G**) High resolution EFTEM image of the region containing *Hbb-b1* transcripts with the associated transcription factory identified by fluorescence signal. Scale bars: 1 μm (E), 200 nm (F,G).

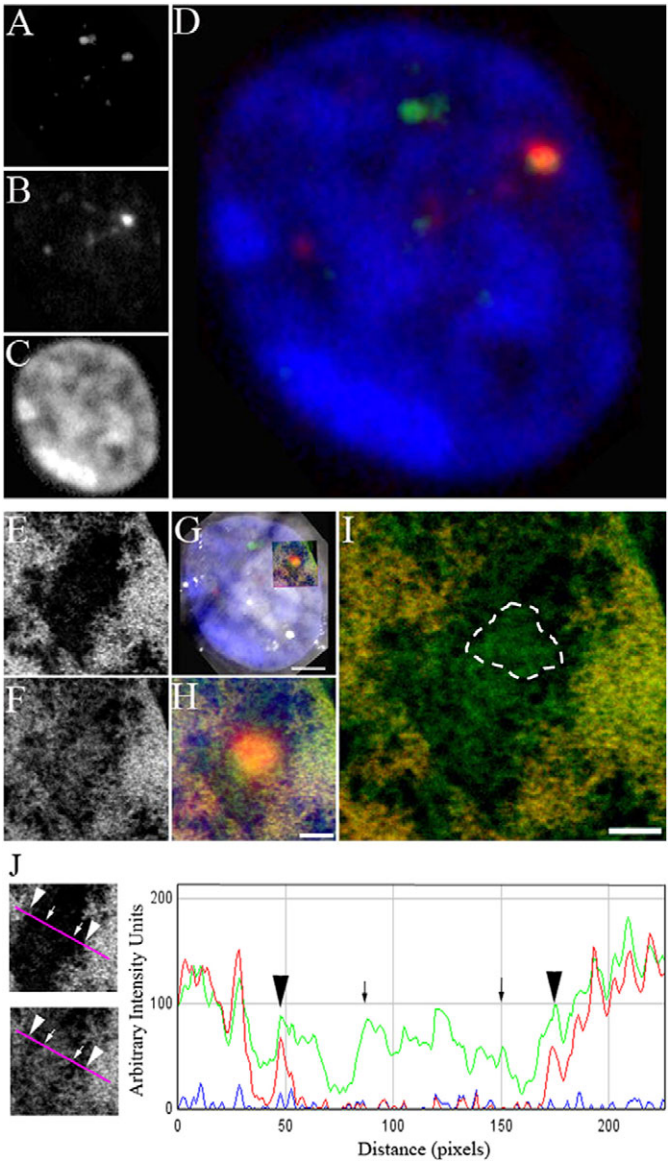

Fig. 6. Genes sharing a single transcription factory. Day 14.5 foetal liver cells were probed for *Hba* and *Hbb-b1* transcripts. (**A**) Localization of *Hba* transcripts within the 100-nm physical section. (**B**) Localization of *Hbb-b1* transcripts within the same physical section. (**C**) Image of Hoechst 33342 counterstained chromatin. (**D**) Merged image showing the localization of *Hba* (green) and *Hbb-b1* transcripts (red) within the nucleus (blue). (**E**) High resolution P map of region containing the co-associating signals from the same section. (**F**) N map of the same region as in E. (**G**) P and N maps false coloured red and green, respectively, and overlaid with the fluorescence micrographs. (**H**) High resolution image of the P and N maps overlaid with fluorescence micrograph. (**I**) High resolution overlay of the P and N maps demonstrating the environment in which the transcript signals for *Hba* and *Hbb-b1* were identified. Structure indicative of an N-rich transcription factory is indicated by the dashed line. (**J**) Line scans (magenta line) were performed on P (upper panel) and N (lower panel) maps through the identified factory and the arbitrary intensity values plotted (N green, P red). Arrows indicate the boundaries of the factory; arrowheads represent the edges of chromatin domains. Scale bars: 200 nm.

determined that there are ~60 per nucleus. Given the number of predicted KFL1-dependent genes (Tallack et al., 2010), we estimate that 15–25 genes are transcribing in a single factory, assuming that not all KFL1-dependent alleles are transcribing simultaneously. This, in combination with the loss and regaining of interactions seen from heat shock studies (Mitchell and Fraser, 2008), is compelling evidence for transcription factories mediating gene co-associations.

Analysis of the volume indicates that transcription factory structures might differ from cell type to cell type. Comparisons of the mean factory volume demonstrated that transcription factories in cells from the erythroid lineage (1.05×10^{-3} μm^3) are three times that of HeLa cell factories (0.35×10^{-3} μm^3), and that KLF1-associated factories are eight times larger (2.76×10^{-3} μm^3). The spatial distribution of transcription factories within the nucleoplasm was approximately 7.4 factories/μm^3, consistent with the values predicted by Faro-Trindade and Cook, (2006) (between 6.8 and 14 factories/μm^3). This average value, however, is slightly misleading because the distribution of factories is not uniform throughout the nucleoplasm, tending to be primarily in regions of low chromatin density. The amount of nucleoplasm with low

Journal of Cell Science

chromatin density does vary dramatically between cell types and therefore the likelihood of observing transcription factories is highly dependent on the location examined. Regardless, this density does allow rough prediction of the total number of transcription factories present within a nucleus of given volume.

Our transcription factory density calculation is reassuring. The value demonstrates that we detected a significant proportion of total transcription sites and that our calculations are accurate. We were unable to label structures with gold particles, as it is notoriously difficult to label nuclear proteins with gold-conjugated antibodies. Attempts were made to post-section label or to label thick-cut cryo-sections; however, this was unsuccessful probably due to insignificant epitope presentation. Due to these technical limitations, we used correlative microscopy to successfully identify transcription factories. Correlative microscopy allowed us to identify the structures of interest without the use of gold particles. This is fortuitous because gold particles, and other electron-dense materials, mask the underlying ultrastructural and elemental information and so we were able gather unobstructed images of the entire physical structure of transcription factories.

One caveat of our method is the use of antibody labelling to identify structures of interest because antibodies contain N atoms that add mass to our predicted values. This is, however, a requirement and the minimal amount of antibodies were used to perform the labelling. For example, we performed direct labelling of primary antibodies to increase signal. We chose to directly label ssDNA probes for the purpose of EMISH, which proved to be adequate for the detection of nascent transcripts in whole cells; however, upon physical sectioning not enough fluorescently labelled material was present to be detected. We are therefore confident that we have limited the amount of material added to provide reliable detection of the structures of interest.

The quantitative elemental information collected using EFTEM allowed us to predict the mass of transcription factories. This was based on using nucleosomes as internal standards for elemental composition. Nucleosomes are the only unit of chromatin with known dimensions and compositions: 11 nm nucleosomal core consisting of eight histones and ~146 bp of DNA. Although we have made assumptions that the histones are canonical and that the DNA is made up of 42% GC content, nucleosomes containing a histone variant or alterations in GC compositions would not significantly change the predicted 1980 N atoms and 292 P atoms. Using these values, we predicted that the mean mass of protein within a transcription factory from e14.5 erythroid cells was 26 MDa. The protein mass is the equivalent to 52 RNAPII holoenzyme complexes, based on the amino acid sequences of proteins in the RNAPII holoenzyme complex. However, it is likely that other accessory proteins required for initiation, elongation and termination are present, as well as a host of proteins ranging from histone-modifying enzymes to specific transcription factors and adaptor proteins. This mass and the co-association of genes at a single factory strongly support the view that these large complexes contain more than a single polymerase or single gene. This becomes more apparent when examining the average mass of protein associated with transcription factories enriched in KLF1 (36 MDa) or those factories transcribing the *Hbb* and *Hba* genes (38 MDa).

Studies have shown that transcriptional machinery can accumulate at artificially inserted sites within the nucleus (Janicki et al., 2004; Kumaran and Spector, 2008). These artificially induced arrays appear to accumulate large amounts of RNAPII, building 'super factories' that might reflect the high rate of transcription and the demand of high amounts of transcriptional machinery. This might be parallel to the puffs in *Drosophila* polytene chromosomes. These puffs are several hundred copies of transcribing heat shock genes with a large accumulation of transcriptional machinery; essentially a large super factory (Yao et al., 2007). In both the artificially induced systems and polytene chromosomes it appears that factories grow in size; however, it is unclear whether there was a smaller factory at these sites prior to transcriptional activation, which would indicate that factories grow in size as the transcriptional load increases. This could account for the increased size in KLF1-enriched factories as well as in those factories that are transcribing highly expressed genes of the erythroid lineage. These genes have a higher transcriptional load and, therefore, require more polymerase and hence larger factories. This increase in transcription factory size could represent both the accumulation of RNAPII components as well as accessory proteins (such as splicing machinery) needed for efficient transcription to occur.

Materials and Methods

Cell preparation

Plugged blb/c or blk6 female mice were terminated at e14.5 by cervical dislocation. Foetuses were removed and placed in ice-cold PBS. Foetal livers were removed, placed in 1 ml ice-cold PBS and gently homogenized. 60 µl of cell suspension was placed on polylysine-coated slides (Sigma) and the cells allowed to attach for 2 minutes. Slides were then placed in PBS containing 4% formaldehyde for 5 minutes at room temperature. Slides were then washed for 5 minutes in PBS at room temperature and permeabilized for 10 minutes in PBS containing 0.5% Triton-X 100. Slides were rinsed in PBS for 5 minutes at room temperature and then placed in 70% ethanol at room temperature for 10 minutes. Slides were stored in 70% ethanol at −20℃. The ultrastructure appeared consistent with previous reports, showing regions of both heterochromatin and euchromatin, as well as other nuclear structures such as nuclear pore complexes that are indicative of good nuclear preservation. All animal experiments were performed according to the relevant regulatory standards.

Stereological calculation of factories per nucleus

The number of foci was determined by acquiring fluorescence images of single sections, containing RNAPII-PS5-labelled e14.5 foetal liver cells, at constant exposure times. Images were processed identically and thresholds set. Masks of images were then created and the number of the factories within each physical slice of nucleus determined. We used the diameter of factories as determined by electron microscopy, i.e. 130 nm, to calculate the number of factories per cell. Given the thickness of the section and the diameter of nucleus, we extrapolated the total number of factories for the entire nuclear volume. This number was then corrected for the Holmes effect, which states that because the structure is larger than the section thickness, objects will be represented multiple times. We therefore derived the formula $S/(D+S)$ to calculate the number of factories per nucleus, where D is the diameter of factories and S is the section thickness (90 nm in these experiments). This formula demonstrates that only 41% of the factories calculated for the entire nuclear volume can be considered because they will be represented in at least two of the physical sections.

Probes

Probes for *Hba* and *Hbb-b1* transcripts (Eurogentec, Southampton, UK) contained three modifications of either dinitrophenyl (DNP) or digoxigenin (DIG) at the 5′-end of the oligonucleotide.

EMISH

Slides with attached cells were removed from −20℃ 70% ethanol, and incubated at room temperature in 70% ethanol. Cells were rehydrated in PBS at room temperature for 5 minutes. Slides were equilibrated in 50% formamide, 2× saline–sodium citrate buffer for 15 minutes at room temperature. Then, 100 µl of hybridization mix (Schoenfelder et al., 2010) containing the probe was placed on the slide and incubated at 37℃ for 1 hour. Probes for *Hbbb1* and *Hba* were used at concentrations described previously (Schoenfelder et al., 2010). Following hybridization, slides were washed for 5 minutes in Tris–saline solution (TS) at room temperature and then in TS containing 0.1% Triton-X 100 for 5 minutes. Slides were blocked in blocking agent (Roche) and the probes detected by antibody labelling. DNP-containing probes were detected using rat anti-DNP antibody conjugates to A488 dye (1:50, Roche) and then donkey anti-rat FITC (1:200,

Journal of Cell Science

Jackson Scientific). DIG-labelled probes were detected using sheep anti-DIG (1:1000, Roche) conjugated to A555 dye and donkey anti-goat Cy3 conjugates (1:200, Jackson Scientific). RNAPII was detected using a 1:200 dilution of a rabbit antibody against the serine 5 phosphorylation of the C-terminal domain of Rpb1 (Abcam, AB5131) directly conjugated to A555 dye (Invitrogen).

EFTEM

Cells were prepared for EFTEM after EMISH or immunofluorescence by first post-fixing slides overnight at 4℃ in 8% formaldehyde in PBS. After quenching in 155 mM glycine for 30 minutes at room temperature, cells were dehydrated in a graded series of ethanol washes (50, 70, 90 and 100%) at room temperature. Cells were then infiltrated with Quetol 651 (Agar Scientific) epoxy resin mix (with nonenyl succinic anhydride, NSA, and methyl-5-norbornene-2,3-dicarboxylic anhydride, NMA) for 2 hours with two changes, and then infiltrated with Quetol resin with hardener (2% final concentration) for 2 hours before polymerization overnight at 65℃. Physical sections were cut on a Lieca ultracut microtome, and fluorescence images collected on an Olympus X61 upright epi-fluorescence microscope. EFTEM images were collected on a Tecnia 20 (FEI) TEM with a 200 $Ke^{-}V$ field emission gun fitted with a post-column Gatan imaging filter.

Diameter calculation

To determine the diameters of the factories, the structures identified by correlative microscopy were measured by EFTEM and the area occupied by the N-rich factory was calculated. Since the aspect ratio (length of longest axis divided by the length of the perpendicular axis) of these structures was <1.5, we could treat these structures as circles while introducing $<3\%$ error in the calculation (Weibel, 1979). Using $r=\sqrt{(A/((2/3)\pi))}$, where r is radius and A is area, we determined the average radius of a factory. Because the average diameter of the factories is larger than the section thickness (~90–100 nm sections) we needed to correct for the Holmes effect using Fullman's formula (Weibel, 1979). The corrected value for the diameter of transcription factories was calculated to be ~130 nm (129.6 nm).

Mass calculation

Mass calculation for the protein and nucleic acid content was done using the same formula previously described (Eskiw et al., 2008). Mass estimates were calculated as nucleosome equivalents and converted to RNAPII core enzyme (murine Rpb1-12) equivalents. The area of a nucleosome (94.99 nm or 78 pixels) was multiplied by the average intensity to give the total intensity units for N or P. This number was then divided by the total number of atoms within the nucleosome to give the average intensity units per atom. Each nucleosome on average will contain 292 P atoms, which is equal to 146 bp of DNA. Given an average GC content, 542 N atoms per nucleosome will be present from DNA, and using the canonical histone sequences 1438 N atoms will be present from the nucleosomal core proteins for a total of 1980 N atoms per nucleosome. The total integrated intensities (area multiplied by average intensity units per pixel) for N and P were divided by the average intensity per atom derived from nucleosomal calculations, and the total number of N and P atoms determined per factory by dividing the total intensity of N or P within the identified area by the intensity of units per atom. Correction for the Holmes effect was performed as previously described (Pombo and Cook, 1996).

Funding

This work was supported by the Medical Research Council and the Biotechnology and Biological Sciences Research Council, UK [grant number BB/E017460/1]. P.J.F. is a Senior Fellow of the MRC [grant number G117/530]. Deposited in PMC for release after 6 months.

Supplementary material available online at
http://jcs.biologists.org/lookup/suppl/doi:10.1242/jcs.087981/-/DC1

References

Bazett-Jones, D. P., Hendzel, M. J. and Kruhlak, M. J. (1999). Stoichiometric analysis of protein- and nucleic acid-based structures in the cell nucleus. *Micron* **30**, 151-157.

Brown, J. M., Green, J., das Neves, R. P., Wallace, H. A., Smith, A. J., Hughes, J., Gray, N., Taylor, S., Wood, W. G., Higgs, D. R. et al. (2008). Association between active genes occurs at nuclear speckles and is modulated by chromatin environment. *J. Cell Biol.* **182**, 1083-1097.

Dhar, S. S. and Wong-Riley, M. T. (2010). Chromosome conformation capture of transcriptional interactions between cytochrome c oxidase genes and genes of glutamatergic synaptic transmission in neurons. *J. Neurochem.* **115**, 676-683.

Eskiw, C. H., Rapp, A., Carter, D. R. and Cook, P. R. (2008). RNA polymerase II activity is located on the surface of protein-rich transcription factories. *J. Cell Sci.* **121**, 1999-2007.

Faro-Trindade, I. and Cook, P. R. (2006). A conserved organization of transcription during embryonic stem cell differentiation and in cells with high C value. *Mol. Biol. Cell* **17**, 2910-2920.

Hendzel, M. J., Boisvert, F. and Bazett-Jones, D. P. (1999). Direct visualization of a protein nuclear architecture. *Mol. Biol. Cell* **10**, 2051-2062.

Iborra, F. J., Pombo, A., Jackson, D. A. and Cook, P. R. (1996). Active RNA polymerases are localized within discrete transcription ''factories'' in human nuclei. *J. Cell Sci.* **109**, 1427-1436.

Janicki, S. M., Tsukamoto, T., Salghetti, S. E., Tansey, W. P., Sachidanandam, R., Prasanth, K. V., Ried, T., Shav-Tal, Y., Bertrand, E., Singer, R. H. et al. (2004). From silencing to gene expression: real-time analysis in single cells. *Cell* **116**, 683-698.

Kumaran, R. I. and Spector, D. L. (2008). A genetic locus targeted to the nuclear periphery in living cells maintains its transcriptional competence. *J. Cell Biol.* **180**, 51-65.

Lawrence, J. B. and Clemson, C. M. (2008). Gene associations: true romance or chance meeting in a nuclear neighborhood? *J. Cell Biol.* **182**, 1035-1038.

Mitchell, J. A. and Fraser, P. (2008). Transcription factories are nuclear subcompartments that remain in the absence of transcription. *Genes Dev.* **22**, 20-25.

Osborne, C. S., Chakalova, L., Brown, K. E., Carter, D., Horton, A., Debrand, E., Goyenechea, B., Mitchell, J. A., Lopes, S., Reik, W. et al. (2004). Active genes dynamically colocalize to shared sites of ongoing transcription. *Nat. Genet.* **36**, 1065-1071.

Osborne, C. S., Chakalova, L., Mitchell, J. A., Horton, A., Wood, A. L., Bolland, D. J., Corcoran, A. E. and Fraser, P. (2007). Myc dynamically and preferentially relocates to a transcription factory occupied by igh. *PLoS Biol.* **5**, e192.

Papantonis, A., Larkin, J. D., Wada, Y., Ohta, Y., Ihara, S., Kodama, T. and Cook, P. R. (2010). Active RNA polymerases: mobile or immobile molecular machines? *PLoS Biol.* **8**, e1000419.

Pombo, A. and Cook, P. R. (1996). The localization of sites containing nascent RNA and splicing factors. *Exp. Cell Res.* **229**, 201-203.

Pombo, A., Hollinshead, M. and Cook, P. R. (1999). Bridging the resolution gap: Imaging the same transcription factories in cryosections by light and electron microscopy. *J. Histochem. Cytochem.* **47**, 471-480.

Ragoczy, T., Bender, M. A., Telling, A., Byron, R. and Groudine, M. (2006). The locus control region is required for association of the murine beta-globin locus with engaged transcription factories during erythroid maturation. *Genes Dev.* **20**, 1447-1457.

Schoenfelder, S., Sexton, T., Chakalova, L., Cope, N. F., Horton, A., Andrews, S., Kurukuti, S., Mitchell, J. A., Umlauf, D., Dimitrova, D. S. et al. (2010). Preferential associations between co-regulated genes reveal a transcriptional interactome in erythroid cells. *Nat. Genet* **42**, 53-61.

Tallack, M. R., Whitington, T., Yuen, W. S., Wainwright, E. N., Keys, J. R., Gardiner, B. B., Nourbakhsh, E., Cloonan, N., Grimmond, S. M., Bailey, T. L. et al. (2010). A global role for KLF1 in erythropoiesis revealed by ChIP-seq in primary erythroid cells. *Genome Res.* **20**, 1052-1063.

Weibel, E. R. (1979). *Stereological Methods: Practical Methods for Biological Morphometry*. London: Academic Press.

Xu, M. and Cook, P. R. (2008). Similar active genes cluster in specialized transcription factories. *J. Cell Biol.* **181**, 615-623.

Yao, J., Ardehali, M. B., Fecko, C. J., Webb, W. W. and Lis, J. T. (2007). Intranuclear distribution and local dynamics of RNA polymerase II during transcription activation. *Mol. Cell* **28**, 978-990.

Table S1. Values used for determining the true number of factories per cell

Area of cell (nm^2)	# Factories	Area (μm^2)	Radius (μm)	Diameter (μm)	Volume (μm^3)	Volume of slice (μm^3)	Factories/ μm^3	Total # of factories / nucleus	True # of factories (# *0.41)
31447330	29	31.44	3.16	6.33	132.69	2.83	10.24	1326.90	544.03
24890776	33	24.89	2.82	5.63	93.44	2.24	14.73	1376.46	564.35
23913117	30	23.91	2.76	5.52	87.99	2.15	13.94	1226.51	502.87
22873055	54	22.87	2.70	5.40	82.31	2.06	26.23	2159.17	885.26
25169513	26	25.17	2.83	5.66	95.01	2.27	11.48	1090.54	447.12
13179672	16	13.18	2.05	4.10	36.00	1.17	13.49	485.63	199.11
30810812	45	30.81	3.13	6.27	128.69	2.78	16.23	2088.31	856.21
29766589	50	29.77	3.08	6.16	122.20	2.68	18.66	2280.69	935.08
23963040	37	23.96	2.76	5.53	88.27	2.16	17.16	1514.27	620.85
16441308	30	16.44	2.29	4.58	50.16	1.48	20.27	1017.00	416.97
12643000	17	12.64	2.01	4.01	33.83	1.14	14.94	505.37	207.20
13591537	13	13.59	2.08	4.16	37.70	1.22	10.63	400.69	164.28
15426207	16	15.43	2.22	4.43	45.59	1.39	11.52	525.39	215.41
16557795	10	16.56	2.30	4.59	50.70	1.49	6.71	340.20	139.48
27769669	24	27.77	2.97	5.95	110.11	2.50	9.60	1057.37	433.52
17423127	20	17.42	2.36	4.71	54.72	1.57	12.75	697.95	286.16
14477670	11	14.4	2.15	4.30	41.45	1.30	8.44	349.92	143.47
26646401	44	26.65	2.91	5.83	103.50	2.40	18.35	1898.91	778.55
19286919	27	19.29	2.48	4.96	63.73	1.74	15.55	991.35	406.45
29604339	56	29.60	3.07	6.14	121.20	2.66	21.018	2547.40	1044.43
21084147	37	21.08	2.59	5.18	72.85	1.90	19.50	1420.40	582.36
27399407	32	27.40	2.95	5.91	107.92	2.47	12.981	1400.40	574.16
27391086	57	27.39	2.95	5.91	107.87	2.47	23.12	2494.08	1022.57
22831452	36	22.83	2.90	5.39	82.09	2.05	17.52	1438.14	589.64
23742547	45	23.74	2.75	5.50	87.05	2.14	21.06	1833.19	751.61
17693543	19	17.69	2.37	4.75	56.00	1.59	11.93	668.18	273.95
12210334	10	12.21	1.97	3.94	32.10	1.10	9.10	292.14	119.78
33348564	41	33.35	3.26	6.52	144.91	3.00	13.66	1979.49	811.59
26554876	31	26.56	2.91	5.82	102.97	2.39	12.97	1335.57	547.58
29096789	56	29.10	3.04	6.09	118.10	2.62	21.38	2525.47	1035.44
23696784	30	23.70	2.75	5.49	86.80	2.13	14.07	1220.95	500.59
21783069	39	21.78	2.63	5.27	76.50	1.96	19.89	1521.79	623.93
24770129	32	24.77	2.81	5.62	92.77	2.23	14.35	1331.51	545.91
21129910	46	21.13	2.59	5.19	73.08	1.90	24.19	1767.82	724.80

((D+S)-(D))/D+S, where D = diameter of factory and S = section thickness
((129+90)-(129))/(129+90) = 0.41 therefore 41% of factories should be considered
0.41×1326 = 545
Since a factory will have to be in a minimum of two sections based on the diameter, not all factories counted can be considered in the calculation.

Table S2A. Nitrogen atoms contributed to a nucleosome by a conical histone octamer

AA	N	Total # AAs	2×	Total N
Ala	1	49	98	98
Cys	1	3	6	6
Asp	2	14	28	56
Glu	1	20	40	40
Phe	1	11	22	22
Gly	1	46	92	92
His	3	14	28	84
Ile	1	25	50	50
Lys	2	54	108	216
Leu	1	41	82	82
Met	1	9	18	18
Asn	2	12	24	48
Pro	1	22	44	44
Gln	2	19	38	76
Arg	3	51	102	306
Ser	1	24	48	48
Thr	1	28	56	56
Val	1	32	64	64
Trp	2	1	2	4
Tyr	1	14	28	28
		Total N		1438

Table S2B. Nitrogen atoms contributed to a nucleosome by DNA

42% GC content	**N**	**% of Genome**	**Residues/ nucleosome**	**N/residue nucleosome**	
A	5	29	42.34	211.7	
T	2	29	42.34	84.68	
G	5	21	30.66	153.3	
C	3	21	30.66	91.98	
	Total N		541.66	542	