Promoter type influences transcriptional topography by targeting genes to distinct nucleoplasmic sites

Joshua D. Larkin[1], Argyris Papantonis[1,2] and Peter R. Cook[1,*]

[1]Sir William Dunn School of Pathology, University of Oxford, South Parks Road, Oxford OX1 3RE, UK
[2]Center for Molecular Medicine, University of Cologne, 21 Robert-Koch Street, 50931 Cologne, Germany
*Author for correspondence (peter.cook@path.ox.ac.uk)

Accepted 31 January 2013
Journal of Cell Science 126, 2052–2059

doi: 10.1242/jcs.123653

Summary

Both the sequence of a promoter and the position of a gene in 3D nuclear space play crucial roles in gene regulation, but few studies address their inter-relationship. Using human and viral promoters on mini-chromosomes and RNA fluorescence *in situ* hybridization coupled to 'high-precision' localization, we show that promoters binding the same transcription factors and responding to the same signaling pathways tend to be co-transcribed in the same transcription factories. We go on to suggest how such spatial co-association might drive co-regulation of genes under the control of similar *cis*-elements.

Key words: Transcription factory, RNA FISH, Transcription factor, Nuclear topography, Nuclear organization

Introduction

Promoters play important roles in transcriptional regulation (Lenhard et al., 2012), acting through their differential affinities for the transcriptional machinery, as well as at other steps during the processing of the primary transcript (e.g. during cleavage, polyadenylation and RNA degradation; Wu et al., 2005; Ji et al., 2011; Trcek et al., 2011). As a result, many efforts have been made to characterize the *cis*-elements embedded within them (Koch et al., 2011; Valen and Sandelin, 2011; Rhee and Pugh, 2012) and determine which ones ensure that some genes are co-expressed (Brohée et al., 2011). As a general rule, these efforts have concentrated on the primary DNA sequence, whilst disregarding any role that the three-dimensional (3D) organization of the genome might play. However, the introduction of techniques like chromosome conformation capture (Dekker et al., 2002) and single-molecule RNA fluorescence *in situ* hybridization (FISH; Femino et al., 1998) now show that co-expressed genes are often found together in 3D nuclear space (e.g. Cai et al., 2006; Simonis et al., 2006; Dhar et al., 2009; Schoenfelder et al., 2010; Noordemeer et al., 2011; Li et al., 2012; Papantonis et al., 2012), closely associated with sub-nuclear structures known as 'transcription factories' (Cook, 2010; Edelman and Fraser, 2012). Such factories have been defined as sites where at least two different genes are transcribed at one time, and they contain many of the molecular components necessary for the production of the mature message (Cook, 2010; Melnik et al., 2011). Although it has implicitly been assumed that specific elements within promoters might drive such co-association, we now confirm that this is the case.

Our strategy is to co-transfect two plasmids bearing the same, or different, promoters (i.e. p1 and p2) into Cos-7 cells (Fig. 1A). As these plasmids encode the SV40 origin of replication (*ori*) and Cos-7 cells express the SV40 T-antigen, the plasmid DNA is assembled into mini-chromosomes which are then replicated and transcribed by the cellular machinery to give several thousand copies per cell (Mellon et al., 1981; Jackson and Cook, 1993; Dean, 1997; Xu and Cook, 2008). Previous work has shown that nascent RNAs copied from two co-transfected plasmids encoding the same (polymerase II) promoter are not randomly distributed in the nucleus (cell 1; Fig. 1A); instead, they tend to colocalize (cell 2; Fig. 1A). Replacing one polymerase II promoter with a polymerase I or III promoter abolished such colocalization (cell 3; Fig. 1A). This is consistent with transcription factories containing either polymerase I, or II, or III – but not any mixture of the active enzymes (Pombo et al., 1999; Xu and Cook, 2008). Here, we investigate the effect that different polymerase II promoters have on the topography of transcription. Plasmids p1 and p2 only differ in two respects: they carry (i) different promoters, and (ii) an insertion of intronic segments A or B (derived from different parts of intron 1 of human *SAMD4A*; Wada et al., 2009) embedded within a common intron (intron 2 of the β-globin gene; Dye et al., 2006); then, A and B transcripts are detected by RNA FISH. Both plasmids also encode EGFP (enhanced green fluorescent protein) which spontaneously concentrates in nuclei (Seibel et al., 2007), and this allows transfected cells in the population to be recognized by their fluorescence (supplementary material Fig. S1A).

Results

An example: the CMV and HIV promoters target mini-chromosomes to different sites

To establish 'proof of principle' we used plasmids encoding EGFP driven by either the cytomegalovirus (CMV) or the human immunodeficiency virus (HIV) promoter. Both are strong promoters, active in many cell types (Sabbioni et al., 1995; Matis et al., 2001; Damdindorj et al., 2012). Cos-7 cells were co-transfected either with two plasmids carrying the same CMV promoter ('p1:CMV' plus 'p2:CMV'), or with plasmids carrying different promoters ('p1:CMV' plus 'p2:HIV'). After 20 hours, the presence of nascent RNA containing intron A or B (driven by

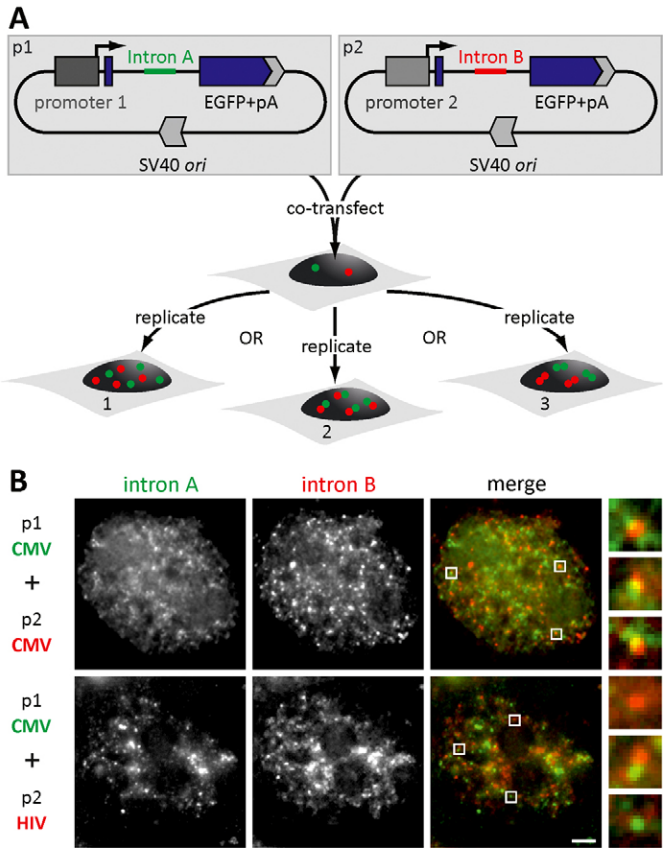


Fig. 1. Approach. (**A**) Strategy. Two constructs (p1 and p2) are co-transfected into Cos-7 cells where they replicate for 20 hours. Both vectors encode the SV40 origin of replication (*ori*; to permit replication in Cos-7 cells, which express T antigen), and *EGFP* (to permit detection of transfected cells); however, plasmids carry different promoters and intronic segments (the latter being targets for the red or green probes used for RNA FISH). Then, RNA FISH with intronic probes is used to determine whether nascent transcripts encoded by the two vectors are spread throughout nuclei (cell 1) or found in the same (cell 2) or different factories (cell 3). (**B**) Nascent transcripts copied from minichromosomes produce complex spatial patterns. Cells were transfected with constructs p1 and p2 carrying the promoters indicated, grown for 20 hours, fixed and RNA FISH performed using probes targeting introns A and B. In resulting images, signals from nascent A and B RNAs produce many diffraction-limited foci in the nucleoplasm (boxed areas are shown on the right); more colocalizing foci are seen when p1 and p2 encode the same promoter (top row). Scale bars: 3 µm (90-nm pixels in boxes).

p1 and p2, respectively) was detected by RNA FISH [Fig. 1B; for more controls, see supplementary material Fig. S1B, and Xu and Cook (Xu and Cook, 2008), who measured the number of minichromosomes and their transcripts in each cell]. Note that (i) the intron used is removed co-transcriptionally and degraded rapidly (data from HeLa cells; Dye et al., 2006) and so marks nascent RNA at the transcription site; (ii) a transcription site is recognized efficiently as a diffraction-limited spot (the nucleic acid targets span <200 nm, even if fully extended) (Wada et al., 2009; Papantonis et al., 2010; Larkin et al., 2012). As expected, nascent transcripts containing introns A and/or B are found in numerous small and discrete foci throughout the nucleoplasm; some of these appear yellow in the resulting merge (Fig. 1B). As we shall see, the transfection involving identical CMV promoters yields more yellow than the one involving different promoters (implying that they are co-transcribed in the same sites), although this is difficult to discern as patterns are so complex. There is also considerable variation from cell to cell in focal size, number, and brightness (not shown) which further complicates quantitative analysis.

A quantitative colocalization method

As RNA FISH patterns were so complex and variable, we developed a general 'single-molecule' localization method to quantify in an unbiased manner the proximity of green and red foci in any image (Fig. 2, see also Materials and Methods). It is based upon others for localizing the peak intensity given by a diffraction-limited spot (Thompson et al., 2002; Yildiz et al., 2003; Larkin and Cook, 2012), and identifying neighboring foci (Barbini et al., 1996; McManus et al., 2006). The algorithm begins by automatically identifying diffraction-limited spots within the image, goes on to localize the peak intensity within each with ~15-nm precision, and ends with a measurement of the

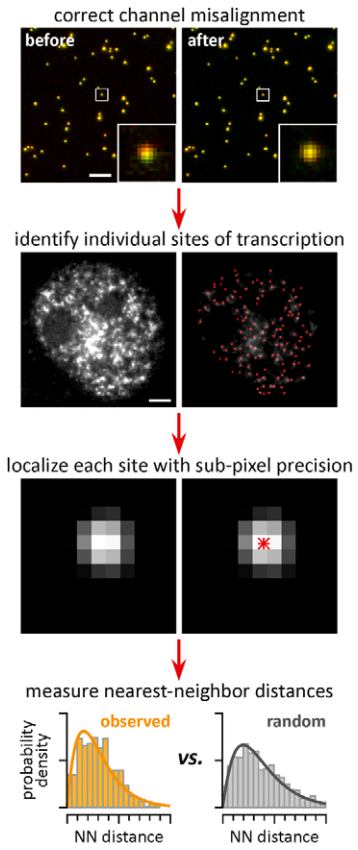


Fig. 2. Flow diagram illustrating steps in quantitative image analysis. After correcting for channel misalignment using 110-nm fluorescent beads, individual foci marking nascent transcripts were identified using shape, intensity and local contrast; focus selection proved robust despite the number of foci/nucleus varying from 1–111. Next, position of the peak intensity in a focus was localized with ~15-nm precision. Then, the distance from the peak intensity in the green channel to the nearest peak in the red channel was measured. Finally, distributions of nearest-neighbor (NN) distances between peaks in red and green channels (and vice versa) were compared with those seen in a random distribution (created for each nucleus using the same number of spots/area); significance was assessed using the two-sample Kolmogorov–Smirnov test. Scale bars: 3 µm (90-nm pixels).

distance from one peak in the green channel to its nearest neighbor (NN) in the red channel, and vice versa (with ~30-nm precision; the decrease in precision is due to imperfections in channel registration). Note that instead of assessing colocalization, which usually implies at least partial overlap of fluorescent signal from two spectrally distinct moieties, we measured proximity between two such moieties, which is represented by a definite separating distance between the two fluors. As each set of measurements is repeated for every focus in the image, the resulting distribution of distances can be compared to a random set (created for each nucleus using the same number of spots in the nucleoplasmic area); differences are then assessed statistically using the Kolmogorov–Smirnov (K-S) test (Hodges, 1958). The output of the analysis tells us whether foci in the green and red channels are distributed in a random pattern, or lie significantly closer to each other than expected by chance. In our case, we can then assess whether the nascent transcripts are in the same or different factories, and so whether the two promoters affect the spatial organization of transcription.

We first applied this method to images like those in Fig. 1B, obtained after co-transfecting 'p1:CMV' plus 'p2:CMV' carrying identical promoters. Numerous foci were identified, localized, and NN distances measured (Fig. 3A). In the experimental sample, more measurements fall in the first bin (signals thus lie 0–10 pixels apart) compared to the random sample (Fig. 3A, left). When the same data are re-plotted as a cumulative density, an excess of the shortest NN distances is again seen in the experimental sample. The two curves prove to be significantly different ($P=10^{-8}$, two-sample K-S test; Fig. 3A, right). In other words, green and red foci are found in closer proximity than expected by chance alone – consistent with nascent transcripts being produced in the same factories (as in cell 2 in Fig. 1A). In contrast, when 'p1:CMV' is paired with 'p2:HIV', the distribution of NN distances is indistinguishable from the random set ($P=0.12$, two-sample K-S test; Fig. 3B) – consistent with the two different promoters being transcribed in distinct factories (as in cell 3 in Fig. 1A). These results are summarized in rows 1 and 2 in Table 1.

Systematic analysis of human promoters reveals an effect on transcription location

We went on to test a range of different promoter combinations; these include the viral promoters mentioned above as well as several human promoters – from the constitutively expressed actin β gene (*ACTB*; Damdindorj et al., 2012), and two genes that respond to an inflammatory cytokine (*SAMD4A* and *TNFAIP2*; Wada et al., 2009; Papantonis et al., 2010). Ultimately, our goal is to assess whether a promoter is sufficient to determine where a gene is transcribed, even when used out of its genomic context.

Experiments were conducted with or without stimulation by tumour necrosis factor alpha (TNFα) for 15 minutes before fixation (as in Table 1). TNFα orchestrates the inflammatory response by signaling via nuclear factor kappa B (NFκB; Smale, 2010). The endogenous *SAMD4A* and *TNFAIP2* genes are found ~50 Mbp apart on human chromosome 14, and so rarely lie close together in 3D nuclear space. However, when activated by TNFα, 3C shows that these two chromosomal genes often come together; moreover, their nascent transcripts are found within 160 nm of each other (Papantonis et al., 2010). We therefore tested how these promoters responded to TNFα when carried on minichromosomes; after transfecting a single plasmid (bearing either the *SAMD4A* or *TNFAIP2* promoter), stimulated cells yield many more foci that unstimulated ones (sometimes up to a few orders of magnitude more; supplementary material Fig. S2). The HIV promoter – which displayed greater cell-to-cell variation (compare Fig. 1B with Fig. 3B; supplementary material Fig. S2) – proved similarly responsive; this is consistent with reports suggesting HIV responds to TNFα through binding of NFκB to the long terminal repeat (Sgarbanti et al., 2008; Zhang et al., 2011). In contrast, the number of foci obtained with the non-responsive *ACTB* promoter does not change substantially upon stimulation (supplementary material Fig. S2).

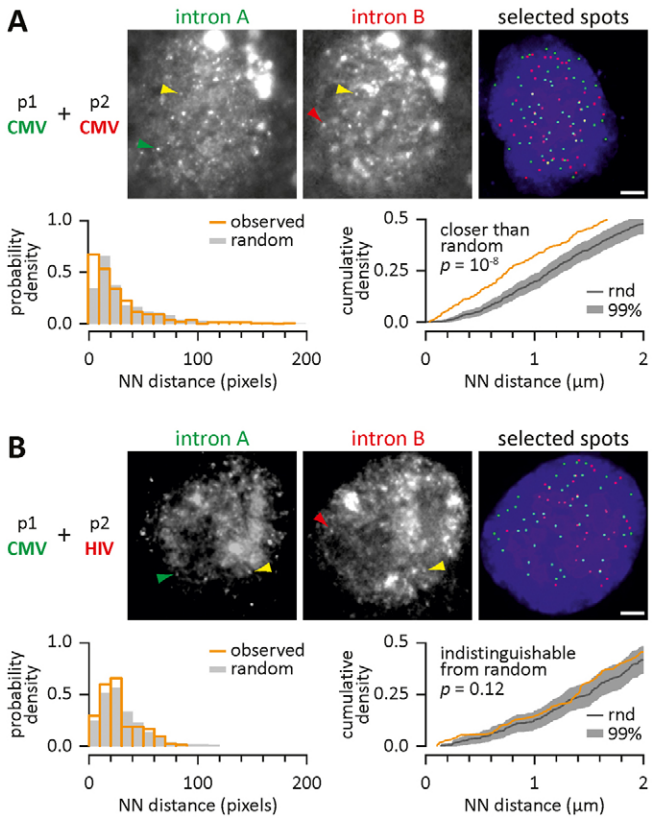


Fig. 3. An example: nascent transcripts produced from the CMV and HIV promoters are found in different locations. Cos-7 cells were co-transfected with p1 (carrying the CMV promoter plus intron A) and p2 (carrying either the CMV or HIV promoter, plus intron B). After fixation, RNA FISH was performed using green and red probes targeting introns A and B, respectively, and DNA counter-stained with DAPI. Images were collected using a wide-field microscope, individual foci marking nascent transcripts selected and nearest-neighbor (NN) distances between green and red peaks determined. (**A**) Identical promoters. A set of three views of one cell are shown; the colored image shows spots selected for NN analysis superimposed on the DAPI image. Images yield complex patterns, in which some foci do not overlap (green and red arrowheads), whereas others do (yellow arrowheads). The orange histogram shows the probability with which a NN distance is seen (10-pixel bins; corresponding probabilities for the random distribution are shown in gray); the observed probability is higher in the first bin. Histogram data are replotted as a cumulative density (orange line). The corresponding curve for the random distribution (rnd) is shown in gray with 99% confidence interval. The orange curve lies above the gray curve, indicative of significantly shorter NN distances ($P=10^{-8}$, $n=370$, two-sample K-S test). (**B**) Different promoters. A similar analysis shows that observed NN distances are indistinguishable from random ($P=0.12$, $n=189$, two-sample K-S test). Scale bars: 3 μm.

Table 1. Correlation between promoter type and nascent transcripts distribution

#	p1	p2	TNFα	Adjacent?	*P*-value	K-S statistics	Number of NNs	
							test	random
1	CMV	CMV	−	Yes	3.1×10^{-8}	0.182	406	785
2	CMV	HIV	−	No	0.12	0.105	188	360
3	*SAMD4A*	*SAMD4A*	+	Yes	1.1×10^{-5}	0.102	1223	2908
4	*SAMD4A*	*TNFAIP2*	+	Yes	1.2×10^{-3}	0.102	1021	2488
5	*SAMD4A*	*ACTB*	+	No	0.55	0.037	1218	2680
6	*SAMD4A*	CMV	+	No	0.018	0.067	1320	2771
7	*SAMD4A*	HIV	+	Yes	2.5×10^{-6}	0.133	1107	2715
8	*SAMD4A*	HIV	−	No	0.12	0.330	37	99

Cos-7 cells were co-transfected with vectors encoding promoters indicated, treated ± TNFα for 15 minutes, and distances between nascent transcripts determined. 'Yes' indicates that the result of the two-sample K-S test rejects the null hypothesis that the observed NN distribution is the same as that given by a random distribution at the 1% significance level (determined as in Fig. 2). In the absence of TNFα, promoters CMV+CMV are transcribed at the same transcription sites (row 1), while CMV+HIV are not (row 2). In the presence of TNFα, promoters *SAMD4A*+*SAMD4A* and *SAMD4A*+*TNFAIP2* are transcribed together (rows 3,4), while promoters *SAMD4A*+*ACTB* and *SAMD4A*+CMV are not (rows 5,6). Promoters *SAMD4A*+HIV are transcribed together only after stimulation with TNFα (rows 7,8).

After establishing that promoters on individual minichromosomes behave (to some degree) like their chromosomal counterparts (i.e. TNFα-responsive *SAMD4A* and *TNFAIP2* promoters respond to the cytokine, the HIV promoter responds in accordance with published data, and the *ACTB* promoter shows its expected constitutive behavior), we analyzed pairwise combinations (Table 1). The pair p1:*SAMD4A* plus p2:*SAMD4A* (in the presence of TNFα) serves as a positive control (Fig. 4, top); NN analysis shows that the resulting foci are found significantly closer together than expected by chance (Table 1, row 3). [In the absence of TNFα (when *SAMD4A* is essentially inactive), too few foci are seen to conduct a meaningful statistical analysis; nevertheless, red and green foci generally colocalize (data not shown).] With p1:*SAMD4A* plus p2:*TNFAIP2* – again in the presence of TNFα – nascent transcripts are also found in proximity (Fig. 4, middle; Table 1, row 4). When one promoter is responsive and the other non-responsive (i.e. p1:*SAMD4A* + p2:*ACTB*, or p1:*SAMD4A* + p2:CMV), nascent transcripts are no longer found together (Fig. 4, bottom; Table 1, rows 5 and 6). Finally, when the responsive *SAMD4A* promoter and HIV promoters are paired (i.e. p1:*SAMD4A* + p2:HIV), nascent transcripts again tend to be found together – but not when TNFα is omitted (Table 1, rows 7 and 8). Note that previous work has shown that *U2* genes on both host chromosomes and minichromosomes share the same factories (Xu and Cook, 2008).

Transcription factor binding correlates with transcription spatial patterns

We next examined whether promoters yielding co-associating transcripts bind the same transcription factors. Binding sites of various transcription factors in the different promoters were inferred (Fig. 5A) using the default parameters in the 'AliBaba 2' interface (Grabe, 2002). Multiple binding sites were given no more weight than single ones, and each factor was characterized as 'unique' (factors with sites only in the *SAMD4A* promoter) or 'common' (factors with sites in both the *SAMD4A* and another promoter; Fig. 5B). If a promoter has more transcription factors in common with the *SAMD4A* promoter than not, the two transcripts tended to be found together (Table 1). For example, the *SAMD4A* and *TNFAIP2* promoters share many factors and their transcripts were often in proximity, whereas those of *SAMD4A* and *ACTB* (or CMV) have less than half of their factors in common and their transcripts were apart.

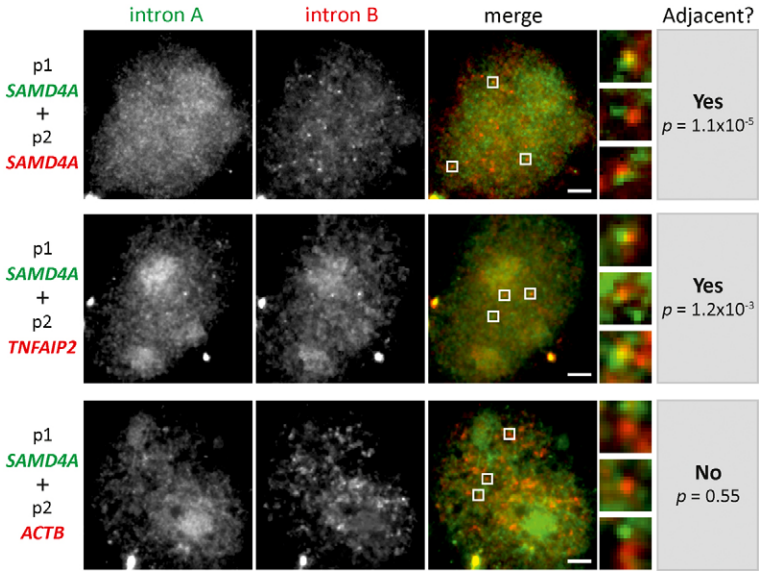


Fig. 4. Nascent transcripts produced from human TNFα-responsive promoters are found together. Cos-7 cells were co-transfected with p1 (encoding the human *SAMD4A* promoter plus intron A) and p2 (encoding either the human *SAMD4A*, *TNFAIP2* or *ACTB* promoter, plus intron B). After fixation, RNA FISH was performed using green and red probes targeting introns A and B, respectively, and DNA counter-stained with DAPI. Images were collected using a wide-field microscope and NN analysis performed (results are summarized in Table 1). Insets show examples of (boxed) foci from the merged images, and results of the NN analysis are indicated. Nascent transcripts of *SAMD4A* and *TNFAIP2* lie closer together than those of *SAMD4A* and *ACTB* (the *SAMD4A*+*SAMD4A* pair serves as a positive control). Scale bars: 3 μm.

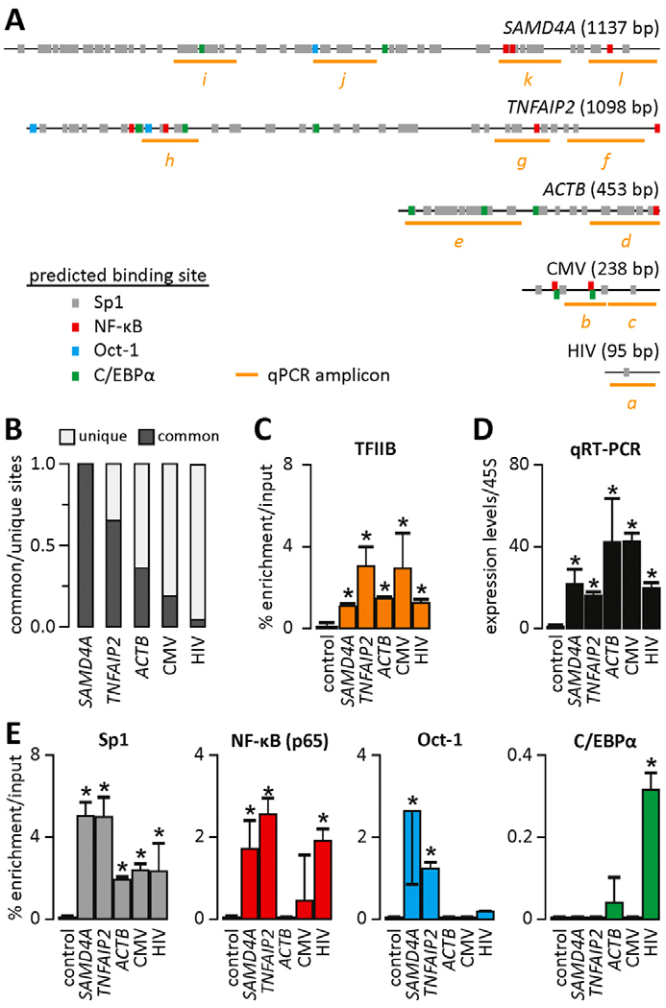


Fig. 5. Transcription factor occupancy on mini-chromosomal promoters. (**A**) Promoter maps showing binding sites for Sp1 (gray), NFκB (red), Oct-1 (blue) and C/EBPα (green) predicted using Alibaba 2. Orange lines indicate regions amplified by qPCR; letters refer to primer pairs in supplementary material Table S1. (**B**) Proportions of 'unique' and 'common' binding sites (i.e. found only in the *SAMD4A* promoter or in both the promoter indicated and the *SAMD4A* promoter). (**C**) Binding of TFIIB, a component of the basal transcriptional machinery, as assessed by ChIP-qPCR (enrichment expressed as percentage of input±s.d.). (**D**) Levels of nascent (intronic) RNA assessed by qRT-PCR. Cos-7 cells were transfected with plasmids carrying the same intronic fragments, total RNA extracted and qRT-PCR applied to assess expression levels (normalized to 45S rRNA levels±s.d.). (**E**) Transcription factor binding assessed by ChIP-qPCR (enrichment expressed as percentage of input±s.d.). In each case, all amplimers shown in A were tested (see supplementary material Table S1), but only results for the one giving the highest enrichment are shown. Sp1 binds to all promoters; NFκB binds to promoters from *SAMD4A*, *TNFAIP2* and HIV, but not to those of *ACTB* and CMV (although they encode potential NFκB-binding sites); Oct-1 bound only to the *SAMD4A* and *TNFAIP2* promoters; and C/EBPα only to the HIV promoter. *$P<0.05$ compared with control levels; two-tailed unpaired Student's *t*-test ($n=4$; two technical replicates of two biological replicates).

Finally, we used chromatin immunoprecipitation coupled to quantitative PCR (ChIP-qPCR) to verify that the mini-chromosomal promoters did indeed associate with the relevant transcriptional machinery and transcription factors (Fig. 5C,E; supplementary material Table S1). TFIIB is a general transcription factor that can be used as a marker for an active transcription complex; as expected, it bound to all promoters tested (Fig. 5C; supplementary material Table S1). RT-qPCR applied with probes targeting an intronic region common to all plasmids also confirmed that all were efficiently transcribed (Fig. 5D). All promoters contained computationally-predicted binding sites for the general transcription factor, Sp1 (Fig. 5C; gray boxes), and many of these sites were occupied (Fig. 5E). NFκB – and in particular, its p65 subunit – is a specific transcription factor that drives the inflammatory response, and all promoters but the HIV one contained predicted p65 binding sites; however, p65 failed to occupy sites in the *ACTB* and CMV promoters, whilst doing so in the HIV promoter. As a result, binding of NFκB (rather than the presence of a predicted site) serves as a good predictor of the RNA FISH results; if the transcription factor is bound to the promoters of both plasmids, their transcripts are found in proximity. Oct-1 is a specific transcription factor with potential targets only on *SAMD4A* and *TNFAIP2*, and only these promoters are occupied. C/EBPα is another specific factor that again has potential targets on all but the HIV promoter, but only this was occupied (also shown by Liu et al., 2010). These results highlight the known difficulties in using computer algorithms to predict potential binding sites, and point to NFκB binding as the (expected) driver of the transcriptional organization after stimulation with TNFα.

Discussion

Early experiments showed that the active forms of RNA polymerases II and III cluster in distinct nucleoplasmic foci – or 'factories', whilst RNA polymerase I is found in nucleoli (Pombo et al., 1999). More recently, transfection experiments showed that two identical constructs encoding polymerase II promoters tend to be co-transcribed in the same factories; inserting into one a polymerase III promoter abolished this (Xu and Cook, 2008; see also Binnie et al., 2006). This suggests that nucleoplasmic factories specialize in transcribing genes transcribed by polymerases II and III. Here, we address the question: do polymerase II factories further specialize in transcribing different gene sub-sets, and are transcription factors bound to the promoters sufficient to drive the targeting to the different factories? Our strategy is to co-transfect two (replicating) plasmids bearing the same, or different, promoters (i.e. p1 and p2) into cells, and use RNA FISH with intronic probes to see if the nascent transcripts produced from these promoters co-associate (Fig. 1A). We applied a 'single-molecule' localization method to quantify in an unbiased manner the proximity of the resulting green and red foci (Fig. 2). This allowed us to assess whether the nascent RNAs are produced in the same factories – and if promoter sequence determined the targeting (Table 1). In parallel, ChIP was used to see if selected transcription factors bound to the promoters, and so whether binding correlated with proximity (Fig. 5).

Our results indicate that when co-transfected plasmids encode the same promoter (whether CMV + CMV, or *SAMD4A* + *SAMD4A*), the two nascent transcripts are found close together; similarly, pairing promoters responding to TNFα ensures that their transcripts are found together (exemplified by *SAMD4A* + *TNFAIP2*, and *SAMD4A* + HIV; Table 1). Conversely, co-transfecting two different viral (CMV + HIV) or human promoters (*SAMD4A* + *ACTB*), or a human and a viral promoter (*SAMD4A* + *CMV*) yields random distributions of transcripts throughout the nucleoplasm (Fig. 1B; Fig. 3B; Fig. 4;

Table 1). Even though it is difficult to predict how a promoter will behave when removed from its normal genomic context (e.g. Noordemeer et al., 2008), results are consistent with plasmids being transcribed preferentially in certain (polymerase II) factories but not in others, with a promoter of <1.3 kbp (the size of the longest used here) determining the site of transcription. The latter conclusion is supported by ChIP results (supplementary material Table S1). For example, *SAMD4A* + *TNFAIP2* and *SAMD4A* + HIV pairs both yield co-localizing transcripts, and bind the critical transcription factor that drives the inflammatory response – NFκB (Smale, 2010; Fig. 5E). The latter pair (i.e. *SAMD4A* + HIV) is of particular interest. Co-association is only seen in the presence of TNFα (Table 1), which induces NFκB to flood into nuclei (Smale, 2010; Ashall et al., 2009), where it binds to both promoters (Fig. 5E). This occurs even though the HIV promoter lacks a predicted (canonical) binding site. Note also that HIV-1 has the capacity to 'hijack' the inflammatory cascade (Brenchley et al., 2006; Grossman et al., 2006) – again presumably through this mechanism.

All results obtained here – and related ones (Xu and Cook, 2008; Papantonis et al., 2012) – are simply explained as follows. We know that some factories in the nucleoplasm contain RNA polymerase II, and others RNA polymerase III (Cook, 2010; Edelman and Fraser, 2012). We imagine that each polymerase II factory will be richer in some transcription factors, and poorer in others (Papantonis et al., 2012). Then, as the promoter on a minichromosome diffuses through the nucleoplasm, it will occasionally collide with a polymerase in a factory, and – if the appropriate polymerase and transcription factor are present – binding to the factory will be stabilized. This, in turn, will increase the chances of productive initiation. As the nucleoplasm contains several thousand minichromosomes (Xu and Cook, 2008), and a typical factory ~8 active polymerases (Cook, 2010), minichromosomes sharing similar promoters will then tend to be transcribed in the same factories (Xu and Cook, 2008). In the particular case involving the *SAMD4A* + HIV promoters in unstimulated cells, both are poorly transcribed in different factories, so transcripts are not likely to be found together. But when TNFα increases the levels of active NFκB (Smale, 2010; Ashall et al., 2009), binding of the transcription factor to both promoters targets them to the same factories where they initiate more frequently. Then, these factories are the transcriptional 'hot-spots' that form the nodes in a complex network, and differential regulation in response to physiological stimuli (in our case) or disease (de la Fuente, 2010) would involve redefining the factors associated with these hot-spots.

Materials and Methods

Plasmid construction

Plasmids were based on II^{CMV}, *EGFP*, pA (Xu and Cook, 2008). This plasmid contains an intron of 925 bp with the splice donor and acceptor sites from the second intron of the human HBB gene (position 5,203,482–5,204,406 on chromosome 11). It was modified by inserting either intron A (1,386 bp from the first intron of *SAMD4A*, beginning 33,984 bp from the TSS) or intron B (1663 bp from the first intron of *SAMD4A*, beginning 114,068 bp from the TSS) between donor and acceptor sites; inserted intron A and B were prepared by PCR amplification of a BAC template (CTD-2046P20, ImaGenes GmbH). The resulting modified plasmid has a CMV promoter; this was replaced with promoters from HIV, *SAMD4A*, *TNFAIP2* or *ACTB* using *Ase*I and *Eco*RI (New England Biolabs). The HIV-1 promoter was obtained from a plasmid provided by Mick Dye (Sir William Dunn School of Pathology, University of Oxford) by PCR-amplification of an 82-bp fragment of the long terminal repeat (LTR) region (−96 to −15 bp relative to the TSS). The *SAMD4A* promoter was amplified from a BAC template (CTD-2589I5, Invitrogen) and comprises 1117 bp (−1322 to −205 bp relative to the TSS). The *TNFAIP2* promoter was amplified from human genomic DNA prepared from human umbilical vein endothelial cells (Lonza), and comprises 1086 bp (−1174 to −88 bp relative to the TSS). The *ACTB* promoter was obtained from a plasmid provided by Martin Dienstbier (Sir William Dunn School of Pathology, University of Oxford), by PCR amplification of a 441-bp fragment (−395 to +46 bp relative to the TSS).

Cell culture and transfection

Cos-7 cells were grown at 37℃ in DMEM+10% fetal calf serum (PAA) on coverslips until 70–80% confluent. FuGENE HD (Promega) was used to transfect cells with 2 μg of DNA (200 ng of each plasmid, and the remainder with sheared salmon sperm DNA), and cells grown for 20 hours. Where applicable, cells were treated with 10 ng/ml tumor necrosis factor alpha (TNFα; Peprotech) for 15 minutes prior to fixation. Fixation involved a pre-incubation in 300 mM sucrose, 100 mM NaCl, 10 mM PIPES (pH 6.8), 6 mM VRC (vanadyl ribonucleoside complex, Sigma), 3 mM $MgCl_2$, and 1.2 mM PMSF for 10 minutes on ice, followed by fixation in 4% paraformaldehyde in 1× PBS (20 minutes; room temperature). Fixed cells were then stored in 70% ethanol (4℃; overnight).

RNA fluorescence *in situ* hybridization, microscopy and image analysis

All solutions were prepared using DEPC-treated water and RNase-free vessels. Probes were sets of five HPLC-purified 50-mers targeting two intronic segments taken from the human *SAMD4A* gene (Wada et al., 2009) spanning 570 nucleotides and 272 nucleotides for intron A and B, respectively. In each 50-mer, roughly every tenth thymine was tagged with Alexa Fluor 647 or 594. Probes were purified, labeled, and labeling efficiencies calculated to be 3.5–4.5 fluors per 50-mer as described (Papantonis et al., 2010). 15 ng of each probe was mixed with 25% deionized formamide, 2× SSC, 50 ng/μl sheared salmon sperm DNA, 5× Denhardt's solution, 50 mM phosphate, and 1 mM EDTA, denatured (90℃; 10 minutes), and placed immediately on ice. Fixed cells on coverslips were removed from storage, rehydrated in PBS, and serially dehydrated in 70%, 90%, and 100% ethanol (3 minutes each). Denatured probe mixture was spotted on clean parafilm and coverslips were inverted onto the spot, then incubated in an humidified chamber (37℃; overnight). Next day, coverslips were washed in 2× SSC/50% formamide (37℃; 15 minutes), 2× SSC (37℃; 15 minutes), and 1× SSC (25℃; 15 minutes). Finally, coverslips were dehydrated in absolute ethanol, allowed to air-dry, mounted onto a slide using Vectashield (Vector Laboratories) plus 200 ng/ml DAPI (4′,6′-diamidino-2-phenylindole), and sealed with nail varnish.

Images were acquired on a wide-field Zeiss Axiovert200 microscope. The aforementioned microscope was equipped with a 63× /1.40 NA objective and a CoolSnap HQ (Photometrix) CCD camera. Fluorescence filter-sets (Semrock), comprising an exciter/dichroic/emitter, were as follows: FF01-650/13, FF660-Di01, FF01-684/24 for Alexa Fluor 647 and FF01-580/23, FF593-Di02, FF01-615/20 for Alexa Fluor 594. Measured transmission spectra were used to compute cross-channel bleed-through, which was negligible. Images were analyzed using a custom software routine written in Matlab (available upon request; summarized in Fig. 2; see also Larkin and Cook, 2012). Briefly, the routine created a mask of the nucleus before identifying candidate foci as diffraction-limited spots using normalized cross-correlation with a kernel representative of the microscope's point-spread function. Features that produced a normalized-cross-correlation value greater than the mean plus 2.5 standard deviations were selected as candidates. Features smaller than 4 pixels were removed. Candidate foci were screened for high local contrast; they were retained only if the brightest pixel was greater than the sum of all pixels at the perimeter of a 9×9 pixel region surrounding the focus (i.e. rows and columns 1, 2, 8 and 9). Finally, candidates with intensity less than one half of a standard deviation above the mean nuclear intensity were eliminated. These selection criteria consistently identified foci that agreed with manual selection. Foci were identified in each channel independently, so as not to bias co-localization results, and identical selection criteria were used for each sample. The center location of each focus was estimated using the Joint Distribution localization algorithm (Larkin and Cook, 2012). Inter-channel image misalignment was corrected by imaging 110-nm TetraSpeck beads (Invitrogen) in each channel at the same time that cell images were acquired. A region within the image of a field of beads that corresponds to the same region within the image of a field of cells was isolated, and the locations of beads in each channel determined with sub-pixel precision. A 2D spatial transform was computed to register the Alexa Fluor-647 channel to the Alexa Fluor-594 channel using bi-linear interpolation between ≥12 bead locations per nucleus. The spatial transform was applied to the cell image before nearest-neighbor (NN) distances were computed. The distance between each spot in one channel and the nearest spot in the other channel was measured for every spot identified. Distance measurements had a precision <30 nm, which includes localization uncertainty and residual misalignment error. The distribution of NN distances observed was compared to that of a randomly-distributed image of spots of the same spatial density as in each nucleus. If the observed distribution contains distances that are significantly smaller than the random distribution, we conclude the mini-chromosomes are transcribed at the same location. The statistical test used for this comparison is the

two-sample Kolmogorov–Smirnov test (Hodges, 1958) with a 99% confidence interval.

Note that we use RNA FISH with intronic probes and wide-field microscopy to localize the two nascent transcripts relative to each other. As the two transcripts lie so close together, we then infer that the two are found in the same factory. Obviously, it would be an advantage if the factory could also be labelled directly, and if resolution were improved. However, after immunolabeling the polymerase and imaging in a wide-field microscope, many foci fill much of the nucleus (so signal from the FISH probes inevitably overlaps the immunolabeling signal, and this signal also contributes to background in the FISH channel). [For a discussion of a related point, see supplementary information in Melnik et al. (Melnik et al., 2011), where different detection approaches are compared (i.e. co-immunoprecipitation, *in situ* proximity ligation, immunofluorescence, immunofluorescence coupled to antibody blocking).] We prefer wide-field microscopy because of the low signal-to-noise ratio in our images (for an example see Fig. 3); in most cases, a confocal microscope is not sensitive enough to allow us to distinguish individual sites of transcription (an exception is shown in supplementary material Fig. S1C, where the bright signal presumably results from multiple transcription events occurring at a single location).

Chromatin immunoprecipitation

Approximately 2×10^7 Cos-7 cells, transfected with the appropriate constructs, were cross-linked (10 minutes; 20℃) in 1% paraformaldehyde (Electron Microscopy Sciences), before chromatin was prepared using the ChIP-IT Express kit (Active motif) as per the manufacturer's instructions. In brief, nuclei were isolated by an NP-40 treatment, chromatin was sheared using a Bioruptor (Diagenode) to ~250 bp, and chromatin was isolated. Immunoprecipitations were performed (~25 µg of chromatin in each) using polyclonal antibodies against the p65 subunit of NFκB (4 µg/reaction; sc-372X, Santa Cruz Biotechnology), Sp1 (5 µg/reaction; 39058, Active motif), Oct-1 (2 µg/reaction; sc-232X, Santa Cruz Biotechnology), C/EBPα (5 µg/reaction; 39306, Active motif), TFIIB (2 µg/reaction; sc-274X, Santa Cruz Biotechnology), and anti-FLAG (2 µg/reaction; F-3165, Sigma-Aldrich) as a non-specific control. DNA was purified using an EZNA MicroElute Cycle-Pure kit (Omega Bio-tek) prior to quantitative real-time PCR (Platinum SYBR Green qPCR Mix-UDG, Invitrogen). All PCR oligonucleotide primers were designed in Primer 3.0 Plus (http://www.bioinformatics.nl/cgi-bin/primer3plus/primer3plus.cgi to have a length of 20–22 nucleotides, a Tm of 62℃, and to yield amplimers of 75–225 bp by using default 'qPCR' settings; primer sequences provided in supplementary material Table S1). Reactions were 50℃/2 minutes, 95℃/5 minutes, and 40 cycles at 95℃/15 seconds, and 60℃/50 seconds using a Rotor-Gene 3000 cycler (Corbett). The presence of single amplimers was confirmed by melting curve analysis, and data was analyzed as described (Nelson et al., 2006) to obtain enrichments relative to input. Two-tailed *P* values for unpaired Student's *t*-tests were calculated using GraphPad (http://www.graphpad.com); they were considered significant when <0.05.

Quantitative reverse-transcriptase PCR

Total RNA was isolated using TRIzol (Invitrogen) from cells stimulated with TNFα for 15 minutes, treated with RQ1 DNase (1 unit of D*N*ase/µg of total RNA; 37℃ for 45 minutes; Promega), and nascent RNA amplified (54℃/10 minutes followed by 1 cycle at 95℃/5 minutes, 40 cycles of 95℃/15 seconds, 60℃/50 seconds, and a single cycle at 40℃/2 minutes; Rotor-Gene 3000 cycler, Corbett) using the One-Step qRT-PCR kit (Invitrogen) with primers targeting intronic region A. The presence of single amplimers was confirmed by melting curve analysis. Reactions in which Platinum Taq polymerase (Invitrogen) replaced the RTase/Taq polymerase mix were performed to ensure amplimers did not result from residual genomic DNA.

Acknowledgements

We thank Jon Bartlett for help, Charlotte S. Mas Cabré for constructs, Martin Dienstbier for the *ACTB* promoter, Mick Dye for the HIV promoter and Tatsuhiko Kodama for RNA FISH probes. The authors declare no competing interests.

Author contributions

J.D.L., A.P. and P.R.C. conceived experiments; J.D.L. produced constructs and analysis algorithms, and performed RNA FISH experiments; A.P. performed ChIP and qRT-PCR; J.D.L., A.P. and P.R.C. analyzed data and wrote the manuscript.

Funding

This work was supported by the Biotechnology and Biological Sciences Research Council via the ERASysBio+/FP7 initiative (to A.P.); and the Wellcome Trust [grant number 086017 to J.D.L.]. Deposited in PMC for release after 6 months.

Supplementary material available online at
http://jcs.biologists.org/lookup/suppl/doi:10.1242/jcs.123653/-/DC1

References

Ashall, L., Horton, C. A., Nelson, D. E., Paszek, P., Harper, C. V., Sillitoe, K., Ryan, S., Spiller, D. G., Unitt, J. F., Broomhead, D. S. et al. (2009). Pulsatile stimulation determines timing and specificity of NF-kappaB-dependent transcription. *Science* **324**, 242-246.

Barbini, P., Cevenini, G. and Massai, M. R. (1996). Nearest-neighbor analysis of spatial point patterns: application to biomedical image interpretation. *Comput. Biomed. Res.* **29**, 482-493.

Binnie, A., Castelo-Branco, P., Monks, J. and Proudfoot, N. J. (2006). Homologous gene sequences mediate transcription-domain formation. *J. Cell Sci.* **119**, 3876-3887.

Brenchley, J. M., Price, D. A. and Douek, D. C. (2006). HIV disease: fallout from a mucosal catastrophe? *Nat. Immunol.* **7**, 235-239.

Brohée, S., Janky, R., Abdel-Sater, F., Vanderstocken, G., André, B. and van Helden, J. (2011). Unraveling networks of co-regulated genes on the sole basis of genome sequences. *Nucleic Acids Res.* **39**, 6340-6358.

Cai, S., Lee, C. C. and Kohwi-Shigematsu, T. (2006). SATB1 packages densely looped, transcriptionally active chromatin for coordinated expression of cytokine genes. *Nat. Genet.* **38**, 1278-1288.

Cook, P. R. (2010). A model for all genomes: the role of transcription factories. *J. Mol. Biol.* **395**, 1-10.

Damdindorj, L., Karnan, S., Ota, A., Takahashi, M., Konishi, Y., Hossain, E., Hosokawa, Y. and Konishi, H. (2012). Assessment of the long-term transcriptional activity of a 550-bp-long human β-actin promoter region. *Plasmid* **68**, 195-200.

de la Fuente, A. (2010). From 'differential expression' to 'differential networking' – identification of dysfunctional regulatory networks in diseases. *Trends Genet.* **26**, 326-333.

Dean, D. A. (1997). Import of plasmid DNA into the nucleus is sequence specific. *Exp. Cell Res.* **230**, 293-302.

Dekker, J., Rippe, K., Dekker, M. and Kleckner, N. (2002). Capturing chromosome conformation. *Science* **295**, 1306-1311.

Dhar, S. S., Ongwijitwat, S. and Wong-Riley, M. T. (2009). Chromosome conformation capture of all 13 genomic Loci in the transcriptional regulation of the multisubunit bigenomic cytochrome C oxidase in neurons. *J. Biol. Chem.* **284**, 18644-18650.

Dye, M. J., Gromak, N. and Proudfoot, N. J. (2006). Exon tethering in transcription by RNA polymerase II. *Mol. Cell* **21**, 849-859.

Edelman, L. B. and Fraser, P. (2012). Transcription factories: genetic programming in three dimensions. *Curr. Opin. Genet. Dev.* **22**, 110-114.

Femino, A. M., Fay, F. S., Fogarty, K. and Singer, R. H. (1998). Visualization of single RNA transcripts in situ. *Science* **280**, 585-590.

Grabe, N. (2002). AliBaba2: context specific identification of transcription factor binding sites. *In Silico Biol.* **2**, S1-S15.

Grossman, Z., Meier-Schellersheim, M., Paul, W. E. and Picker, L. J. (2006). Pathogenesis of HIV infection: what the virus spares is as important as what it destroys. *Nat. Med.* **12**, 289-295.

Hodges, J. L. (1958). The significance probability of the Smirnov two-sample test. *Arkiv för Matematik* **3**, 469-486.

Jackson, D. A. and Cook, P. R. (1993). Transcriptionally active minichromosomes are attached transiently in nuclei through transcription units. *J. Cell Sci.* **105**, 1143-1150.

Ji, Z., Luo, W., Li, W., Hoque, M., Pan, Z., Zhao, Y. and Tian, B. (2011). Transcriptional activity regulates alternative cleavage and polyadenylation. *Mol. Syst. Biol.* **7**, 534.

Koch, F., Fenouil, R., Gut, M., Cauchy, P., Albert, T. K., Zacarias-Cabeza, J., Spicuglia, S., de la Chapelle, A. L., Heidemann, M., Hintermair, C. et al. (2011). Transcription initiation platforms and GTF recruitment at tissue-specific enhancers and promoters. *Nat. Struct. Mol. Biol.* **18**, 956-963.

Larkin, J. D. and Cook, P. R. (2012). Maximum precision closed-form solution for localizing diffraction-limited spots in noisy images. *Opt. Express* **20**, 18478-18493.

Larkin, J. D., Cook, P. R. and Papantonis, A. (2012). Dynamic reconfiguration of long human genes during one transcription cycle. *Mol. Cell. Biol.* **32**, 2738-2747.

Lenhard, B., Sandelin, A. and Carninci, P. (2012). Metazoan promoters: emerging characteristics and insights into transcriptional regulation. *Nat. Rev. Genet.* **13**, 233-245.

Li, G., Ruan, X., Auerbach, R. K., Sandhu, K. S., Zheng, M., Wang, P., Poh, H. M., Goh, Y., Lim, J., Zhang, J. et al. (2012). Extensive promoter-centered chromatin interactions provide a topological basis for transcription regulation. *Cell* **148**, 84-98.

Liu, Y., Nonnemacher, M. R., Stauff, D. L., Li, L., Banerjee, A., Irish, B., Kilareski, E., Rajagopalan, N., Suchitra, J. B., Khan, Z. K. et al. (2010). Structural and functional studies of CCAAT/enhancer binding sites within the human immunodeficiency virus type 1 subtype C LTR. *Biomed. Pharmacother.* **64**, 672-680.

Matis, C., Chomez, P., Picard, J. and Rezsohazy, R. (2001). Differential and opposed transcriptional effects of protein fusions containing the VP16 activation domain. *FEBS Lett.* **499**, 92-96.

McManus, K. J., Stephens, D. A., Adams, N. M., Islam, S. A., Freemont, P. S. and Hendzel, M. J. (2006). The transcriptional regulator CBP has defined spatial associations within interphase nuclei. *PLoS Comput. Biol.* **2**, e139.

Mellon, P., Parker, V., Gluzman, Y. and Maniatis, T. (1981). Identification of DNA sequences required for transcription of the human alpha 1-globin gene in a new SV40 host-vector system. *Cell* **27**, 279-288.

Melnik, S., Deng, B., Papantonis, A., Baboo, S., Carr, I. M. and Cook, P. R. (2011). The proteomes of transcription factories containing RNA polymerases I, II or III. *Nat. Methods* **8**, 963-968.

Nelson, J. D., Denisenko, O. and Bomsztyk, K. (2006). Protocol for the fast chromatin immunoprecipitation (ChIP) method. *Nat. Protoc.* **1**, 179-185.

Noordermeer, D., Branco, M. R., Splinter, E., Klous, P., van Ijcken, W., Swagemakers, S., Koutsourakis, M., van der Spek, P., Pombo, A. and de Laat, W. (2008). Transcription and chromatin organization of a housekeeping gene cluster containing an integrated beta-globin locus control region. *PLoS Genet.* **4**, e1000016.

Noordermeer, D., Leleu, M., Splinter, E., Rougemont, J., De Laat, W. and Duboule, D. (2011). The dynamic architecture of Hox gene clusters. *Science* **334**, 222-225.

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.

Papantonis, A., Kohro, T., Baboo, S., Larkin, J. D., Deng, B., Short, P., Tsutsumi, S., Taylor, S., Kanki, Y., Kobayashi, M. et al. (2012). TNFα signals through specialized factories where responsive coding and miRNA genes are transcribed. *EMBO J.* **31**, 4404-4414.

Pombo, A., Jackson, D. A., Hollinshead, M., Wang, Z., Roeder, R. G. and Cook, P. R. (1999). Regional specialization in human nuclei: visualization of discrete sites of transcription by RNA polymerase III. *EMBO J.* **18**, 2241-2253.

Rhee, H. S. and Pugh, B. F. (2012). Genome-wide structure and organization of eukaryotic pre-initiation complexes. *Nature* **483**, 295-301.

Sabbioni, S., Negrini, M., Rimessi, P., Manservigi, R. and Barbanti-Brodano, G. (1995). A BK virus episomal vector for constitutive high expression of exogenous cDNAs in human cells. *Arch. Virol.* **140**, 335-339.

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.

Seibel, N. M., Eljouni, J., Nalaskowski, M. M. and Hampe, W. (2007). Nuclear localization of enhanced green fluorescent protein homomultimers. *Anal. Biochem.* **368**, 95-99.

Sgarbanti, M., Remoli, A. L., Marsili, G., Ridolfi, B., Borsetti, A., Perrotti, E., Orsatti, R., Ilari, R., Sernicola, L., Stellacci, E. et al. (2008). IRF-1 is required for full NF-kappaB transcriptional activity at the human immunodeficiency virus type 1 long terminal repeat enhancer. *J. Virol.* **82**, 3632-3641.

Simonis, M., Klous, P., Splinter, E., Moshkin, Y., Willemsen, R., de Wit, E., van Steensel, B. and de Laat, W. (2006). Nuclear organization of active and inactive chromatin domains uncovered by chromosome conformation capture-on-chip (4C). *Nat. Genet.* **38**, 1348-1354.

Smale, S. T. (2010). Selective transcription in response to an inflammatory stimulus. *Cell* **140**, 833-844.

Thompson, R. E., Larson, D. R. and Webb, W. W. (2002). Precise nanometer localization analysis for individual fluorescent probes. *Biophys. J.* **82**, 2775-2783.

Trcek, T., Larson, D. R., Moldón, A., Query, C. C. and Singer, R. H. (2011). Single-molecule mRNA decay measurements reveal promoter- regulated mRNA stability in yeast. *Cell* **147**, 1484-1497.

Valen, E. and Sandelin, A. (2011). Genomic and chromatin signals underlying transcription start-site selection. *Trends Genet.* **27**, 475-485.

Wada, Y., Ohta, Y., Xu, M., Tsutsumi, S., Minami, T., Inoue, K., Komura, D., Kitakami, J., Oshida, N., Papantonis, A. et al. (2009). A wave of nascent transcription on activated human genes. *Proc. Natl. Acad. Sci. USA* **106**, 18357-18361.

Wu, Q., Hwang, C. K., Yao, S., Law, P. Y., Loh, H. H. and Wei, L. N. (2005). A major species of mouse mu-opioid receptor mRNA and its promoter-dependent functional polyadenylation signal. *Mol. Pharmacol.* **68**, 279-285.

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

Yildiz, A., Forkey, J. N., McKinney, S. A., Ha, T., Goldman, Y. E. and Selvin, P. R. (2003). Myosin V walks hand-over-hand: single fluorophore imaging with 1.5-nm localization. *Science* **300**, 2061-2065.

Zhang, H. S., Sang, W. W., Ruan, Z. and Wang, Y. O. (2011). Akt/Nox2/NF-κB signaling pathway is involved in Tat-induced HIV-1 long terminal repeat (LTR) transactivation. *Arch. Biochem. Biophys.* **505**, 266-272.

Journal of Cell Science

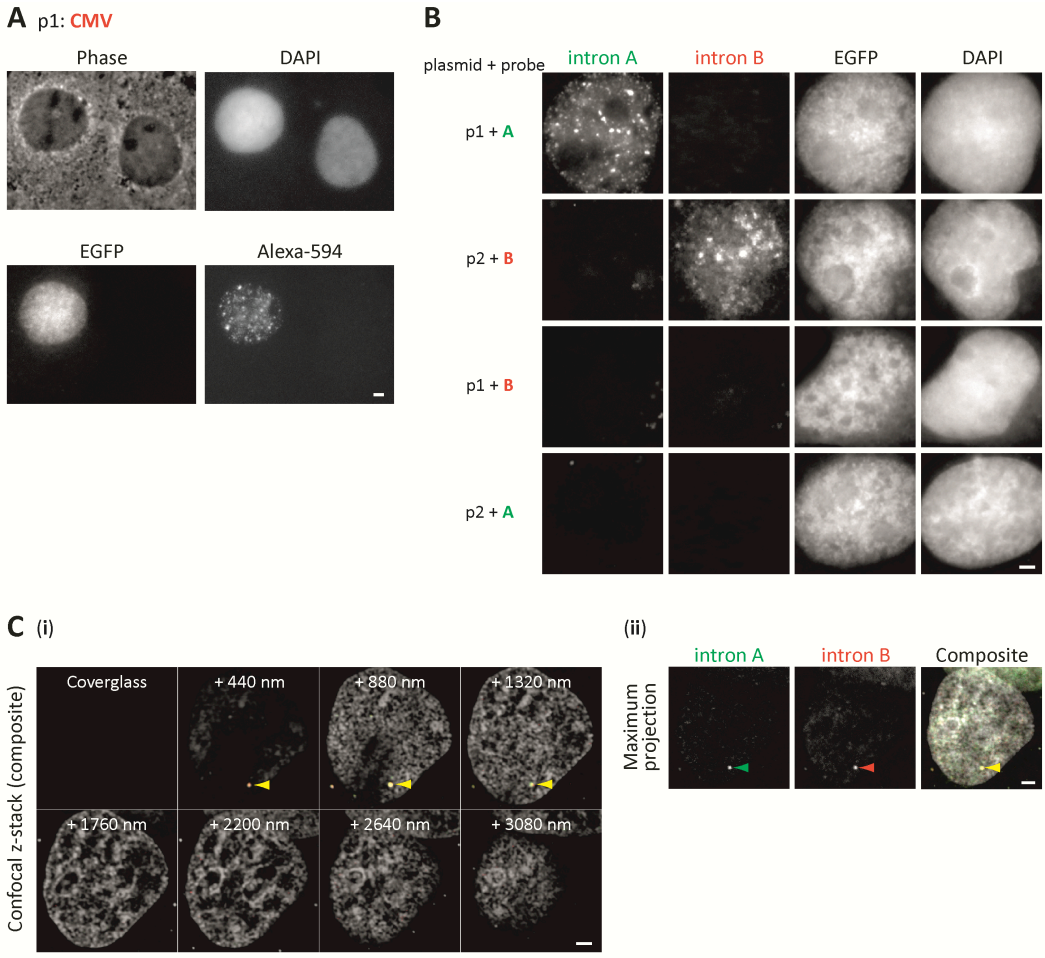
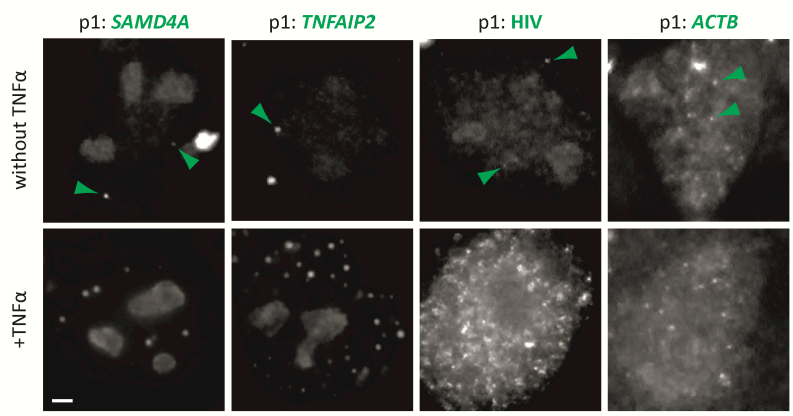


Fig. S1. RNA FISH probes specifically detect single sites of transcription. Bars: 2 μm. (**A**) A control. Cos-7 cells were transfected with one plasmid (p1:CMV), fixed, RNA FISH performed using probes tagged with Alexa-594 targeting intron A, DNA counter-stained with DAPI, and cells imaged using a wide-field microscope. Four views of two cells are shown; only the cell on the left is transfected (identified by EGFP expression), and only it contains intron A RNA in discrete foci (Alexa-594 fluorescence). (**B**) Probe specificity. Cos-7 cells were transfected with either p1 or p2 (carrying the CMV promoter plus intron A or B, respectively) and fixed, RNA FISH performed using probes targeting one or other intron, DNA counter-stained with DAPI, and cells imaged using a wide-field microscope. Sets of four views of 4 cells are shown; all cells express EGFP (and so were transfected). Green foci marking intron A RNA are only seen in cells transfected with p1 using probe A (green), and red foci marking intron B RNA are only seen with p2 using probe B (red). As probes target intronic RNA, strong signal is restricted to nuclei (although some non-specific hybridization and auto-fluorescence contribute background). (**C**) Sites of minichromosome transcription are confined to the nucleus. Cos-7 cells were transfected with plasmids containing the *SAMD4A*

and *TNFAIP2* promoters, and transcripts detected using probes targeting introns, DNA counter-stained with DAPI, and cells imaged using a confocal microscope. (**i**) A *z*-stack of 8 sections (each 440 nm thick; channels merged). Yellow arrowheads: a yellow focus seen in the nuclear interior in three sections. (**ii**) Maximum projection of the stack showing colocalization of the two intronic signals.



Fig, S2. Some minichromosomal genes respond to TNFα. Cos-7 cells were transfected with p1 carrying the promoter from either *SAMD4A*, *TNFAIP2*, HIV, or *ACTB* (plus intron A), treated for 15 min ± TNFα, fixed, RNA FISH performed using probes targeting intron A, and cells imaged using a wide-field microscope (arrowheads mark typical foci). When mini-chromosomes encode one of the three TNFα-responsive promoters (but not the *ACTB* promoter), stimulation increases the number of foci. Bar: 2 μm.

Table S1. Primers for ChIP-qPCR and transcription factor binding detected. Primers are shown sequentially and in respect to the amplicons (*a*-*l*) depicted in Figure 5A. Y: transcription factor binding detected; N: no binding detected.

			transcription factor ChIP				
amplicon	**primer name**	**sequence (5'—> 3')**	**Sp1**	**NFκB**	**Oct1**	**C/EBPα**	**TFIIB**
a	HIVpF	tctctctggttagaccagatctgag	**Y**	**Y**	N	**Y**	**Y**
	HIVpR	gctttattgaggcttaagcagtg					
b	CMVp65F	ccaagtctccaccccattg	N	N	N	N	N
	CMVp65R	acattttggaaagtcccgttg					
c	CMVSp1F	taacaactccgccccattg	**Y**	N	N	N	**Y**
	CMVSp1R	ggttcactaaaccagctctgc					
d	ACTBproxF	gttccgaaagttgccttttatg	**Y**	N	N	N	**Y**
	ACTBproxR	caaaggcgaggctctgtg					
e	ACTBdistF	agtgcccaagagatgtccac	**Y**	N	N	N	N
	ACTBdistR	aagaggaggagggagagttttg					
f	TNFAIPSp1F	tctggggctgtgaagtattg	**Y**	**Y**	N	N	**Y**
	TNFAIPSp1R	gtctgtttcacccattcagc					
g	TNFAIPp65F	gatccagggaggtgtccatt	N	**Y**	N	N	N
	TNFAIPp65R	ccacacagaggggacttt					
h	TNFAIPdistF	cttgcctctctggatgtagctt	N	N	**Y**	N	N
	TNFAIPdistR	agcttgggcttggaactca					
i	SAMcebpF	aggcccagggagacctc	N	N	N	N	N
	SAMcebpR	aaactctctccacccaaagc					
j	SAMoctF	taagaggaggaggaaggaactc	N	N	**Y**	N	N
	SAMoctR	gcgcagaaaacttgcaatg					
k	SAMprox1F	cgtgaagggtccgagttg	**Y**	N	N	N	**Y**
	SAMprox1R	ccacaaggcacacaactgaa					
l	SAMprox2F	gatttcggcgaacccaa	N	**Y**	N	N	**Y**
	SAMprox2R	agtgagacctcggctagcg					