bioRxiv preprint doi: https://doi.org/10.1101/405969; this version posted November 25, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

Nanoscale mapping of chromatin dynamics in living cells

Haitham A. Shaban[1,2] *, Roman Barth[1,3] and Kerstin Bystricky[1] *

1: Laboratoire de Biologie Moléculaire Eucaryote (LBME), Centre de Biologie Intégrative (CBI), CNRS; University of Toulouse, UPS; 31062 Toulouse; France; 2: Spectroscopy Department, Physics Division, National Research Centre, Dokki, Cairo, Egypt; 3: Present Address: Bionanoscience Department, Faculty of Applied Sciences, Delft University of Technology; 2628 CJ Delft; The Netherlands

* Correspondence: Email haitham.shaban@ibcg.biotoul.fr , kerstin.bystricky@ibcg.biotoul.fr

ABSTRACT

Principles of genome folding and their relationship to function depend on understanding conformational changes of the chromatin fiber in real-time. However, bulk chromatin motion has not been analysed at high resolution. We developed Hi-D, a method to quantitatively map dynamics of chromatin and abundant nuclear proteins for every pixel simultaneously over the entire nucleus from fluorescence image series. Hi-D combines reconstruction of chromatin motion using computer vision and classification of local diffusion processes by Bayesian inference. We discovered that DNA dynamics in the nuclear interior are spatially partitioned into 0.3 – 3 µm domains in a mosaic-like manner which is uncoupled from chromatin compaction. This pattern was remodeled in response to transcriptional activity accompanied by global reduction of DNA dynamics. In contrast, we found that RNA polymerase II dynamics are, compared to actively transcribing cells, reduced in quiescent cells and when elongation was inhibited. Our results strengthen the hypothesis that chromatin dynamics are dictated by interchromatin contacts, rather than chromatin density. Hi-D can be applied to any dense and bulk structures opening new perspectives towards understanding motion of nuclear molecules in the context of nuclear architecture.

bioRxiv preprint doi: https://doi.org/10.1101/405969; this version posted November 25, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

INTRODUCTION

Spatial organization and dynamics of chromatin correlate with cell function and fate [1]. In mammalian cells, chromosomes occupy territories whose localization is stable over the cell cycle [2]. Each chromosome contains dense heterochromatin and open euchromatin areas whose extent is indicative of cellular activity [3]. Transitions within and between eu- and heterochromatin involve multiple levels of chromatin organization from nucleosome density, long-range looping and domain folding that adapt to and enable DNA processing [4]. Structural models derived from contact and crosslinking frequencies [5–7] are consistent with the view that the genome is partitioned into compartments and sub-compartments [8]. It is now becoming increasingly clear that such nuclear compartments are also dynamic entities whose formation and maintenance over time is of great interest in order to understand the mechanisms and function of genome folding [9]. Tracking of labelled single DNA loci [10–14] or chromatin domains [15,16] demonstrated that chromatin motion is highly heterogeneous at short time intervals. However, sparse loci are difficult to place in the context of global chromatin organization and its enduring configurational changes [17]. On the other hand, locally restricted genomic processes can also not be inferred from quantities averaged over the entire nucleus [18,19]. A first study analyzing bulk chromatin motion at nanoscale resolution revealed dynamic partitioning of chromatin into a number of nuclear sub-regions which show correlated motion of chromatin in the micrometer range [20]. However, to further elucidate underlying physical mechanisms driving chromatin dynamics and giving rise to their spatial coherence, an integrated characterization of the stochastic motion of nuclear molecules in their native, and physiological environment at the local scale is needed. Likewise, whether compaction of chromatin determines its spatial coherence or whether chromatin dynamics are distinct in open and closed chromatin is still a matter of debate.

To decipher the relationship between genome compaction, activity and functional dynamics, we developed Hi-D, a method to overcome the limitations of sparse and ensemble approaches for imaging dense structures such as chromatin and nuclear proteins. Hi-D, a combination of dense optical

bioRxiv preprint doi: https://doi.org/10.1101/405969; this version posted November 25, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

flow reconstruction to quantify the local motion of chromatin at sub-pixel accuracy from live cell imaging, and a Bayesian inference approach to precisely classify local types of diffusion. Biophysical properties such as diffusion constants and anomalous exponents were determined for each pixel to create two dimensional maps of chromatin dynamics at single pixel resolution over the entire nucleus in living single cells. We find that DNA compaction and dynamics do not necessarily correlate. Instead, we strengthen the hypothesis that chromatin dynamics are dictated by interchromatin contacts and confinement, rather than chromatin density.

RESULTS

Hi-D maps genome dynamic properties at nanoscale spatial resolution in living cells

Motion observed within a series of conventional confocal fluorescence microscopy images was quantitatively reconstructed by a dense Optical Flow method [20] and trajectories were calculated by integration of the resulting flow fields (Figure 1a, Supplementary Note 1). The type of diffusion characterizing each pixel's chromatin motion was determined in an unbiased manner using Bayesian inference to simultaneously test a set of common models to fit each trajectory's Mean Squared Displacement (MSD) curve [21] (Figure 1b, left panel, Supplementary Figure 2). The five principal models are shown as a coloured map directly on the nucleus (Figure 1b, right panel) (Methods section). In general, we found that only a small fraction of trajectories displayed directed diffusion (Figure 1b), while the bulk of chromatin exhibited sub-diffusive behaviour. Over the duration of the experiment, distinguishing MSD between the comparable cases of anomalous and confined diffusion is a challenging task. By examining a range of parameters governing different types of diffusion, our results suggest that chromatin diffusion in human U2OS cells can be adequately described as anomalous to avoid misinterpretation (Supplementary Note 2). Biophysical parameters calculated for each pixel (diffusion constant D, anomalous exponent α and drift velocity V) are presented in color-coded 2D heat-maps (Figure 1c) (Methods section) and form islands of irregular shape and dimensions

bioRxiv preprint doi: https://doi.org/10.1101/405969; this version posted November 25, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

(Figure 1c). These parameter maps clearly demonstrate that chromatin exhibits spatially heterogeneous and partitioned dynamics forming a mosaic of domains with similar values characterizing chromatin motion. This view is in line with current insights that chromatin dynamics are spatially correlated in the micrometer-range [18,20]. To further explore and characterize this heterogeneous distribution, parameter value distributions were deconvolved into contributing sub-populations using a General Mixture model (GMM) (Figure 1c, Methods, and Supplementary Figure 4). We identified three populations of chromatin mobility which we refer to as slow, intermediate and fast (Methods, exemplary in Figure 1c), irrespective of the parameter under consideration (diffusion constant or anomalous exponent) or transcription state of the cell. We found that directed motion involving a drift velocity (V) was present to a much lower extent than free and anomalous diffusion resulting in much less data for V than for the other two parameters (Figure 1c). Hence drift velocity was not retained for further analysis.

Validation of the Hi-D approach in simulation and experiment

In order to examine the accuracy of calculated trajectories and associated diffusion constants by Hi-D, we compared Hi-D results to dynamic multiple-target tracing (MTT), a Single Particle Tracking (SPT) method which is commonly used for dense molecule tracking [22] (Figure 2a,b; Supplementary Note 2; Supplementary Figure 7). While the SPT method outperforms the Hi-D approach in scenarios of sparsely labelled molecules (Figure 2a), Hi-D calculations result in considerably more accurate estimates for local diffusion constants than SPT estimates in scenarios of densely labelled molecules or structures of heterogeneous labelling density such as chromatin (Figure 2b). Hi-D therefore constitutes an approach to extract dynamic information of biomolecules with dense labeling where SPT cannot. To ensure that the calculated dynamics are not a consequence of imaging noise, we experimentally validated the sensitivity of the approach by calculating the MSD for formaldehyde fixed and living U2OS cells labelled by DNA-SiR in quiescence (- serum) or normal growth (+ serum). Diffusion constants derived from the MSD curves by Bayesian inference were about two orders of

bioRxiv preprint doi: https://doi.org/10.1101/405969; this version posted November 25, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

magnitude greater in living cells than in fixed cells (Figure 2c) confirming that Hi-D enables quantifying DNA dynamics well above the noise background. To confirm the robustness of extracted parameter values with respect to varying levels of imaging noise, Hi-D was applied to imaged nuclei to which noise was artificially added. The signal-to-noise ratio (SNR) of the original nuclei were about 26 dB and the SNR was reduced to 16 dB. The computed distribution of diffusion constants is consistent up to a lower limit of about 20 dB, below which the distribution is considerably biased towards larger values and broadens (Figure 2d, Supplementary Figure 8). Likewise, features of the spatial map of diffusion constants is equally conserved for SNR values as low as 20 dB, demonstrating the robustness of Hi-D for varying imaging noise (Supplementary Figure 8). In analogy to the robustness to varying SNR levels, Hi-D is thus robust to photobleaching effects (provided that SNR ≥ 20dB) since flow fields are only estimated between consecutive images, between which illumination changes due to photobleaching are usually negligible. Furthermore, Hi-D was also shown to be robust to slight variations in the time intervals of acquired time series as long as the expected motion between frames was in the order of the pixel size of the imaging system used (Supplementary Figure 9). We further validated Hi-D against iMSD, a well-established method to extract dynamic information of dense molecules, based on the spatial correlation function of intensity fluctuations caused by diffusing molecules, which are recorded using camera-based systems [23]. Using successive applications of iMSD to overlapping regions of interest, we computed a diffusion map similar to Hi-D derived maps (Supplementary Note 4). Quantitatively, both methods yield diffusion constants in the same order of magnitude (Hi-D $(1.6 \pm 0.8) \cdot 10^{-3}\ \mu m^2/s$, iMSD $(2.2 \pm 4.5) \cdot 10^{-3}\ \mu m^2/s$, mean $\pm$ standard deviation), which are consistent with reported values using single particle tracking and correlation spectroscopy methods in interphase chromatin [14,18] . However, the distribution of values derived by iMSD was considerably broader then the distribution revealed by Hi-D (Figure 2f). The distribution of anomalous exponents computed by iMSD showed many spurious values at the limit of the scale, while Hi-D returns reasonable values (Figure 2g). We thus conclude that Hi-D reveals dynamic parameters of the same order of magnitude as iMSD, a comparable method, but is advantageous in the estimation of

bioRxiv preprint doi: https://doi.org/10.1101/405969; this version posted November 25, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

multiple parameters simultaneously, by virtue of the featured Bayesian model selection. Hi-D is thus an accurate, robust, fast and easy to use readout of dynamics of nuclear macromolecules.

Single-cell biophysical property maps of genome conformation and behaviour

To concomitantly monitor position and distribution of the DNA mobility populations in different biological conditions, we determined Hi-D maps for a single serum-starved and then stimulated cell (Figure 3a). Cells grown in a serum-free medium arrest in G_0 in which transcription is globally inhibited; adding serum to the medium brings cells back into the cell cycle and stimulates transcriptional activity [20,24,25]. Diffusion constants were calculated for each pixel based on the model selected by Bayesian inference. Small diffusion constants were prominently located at the nuclear envelope (dark blue). Plotting the average diffusion constants versus the distance from the nuclear periphery showed that the mobility within a rim of 1 µm from the periphery increases linearly before assuming a nearly constant value in the inner volume of the nucleus (Supplementary Figure 10). In numerous sites across the remaining nuclear volume, fast diffusive areas of irregular dimensions spanning $0.3 - 3\ \mu m$ in diameter (yellow areas in Figure 3a) are imbedded in the bulk of moderately dynamic chromatin. Areas of different parameter values seamlessly transition into one another without clearly defined boundaries, evident of spatially correlated chromatin dynamics [20]. Upon serum stimulation, the spatial distribution of high and low diffusion constants was largely conserved (compare the presence of yellow regions in the quiescent and serum stimulated cell in Figure 3a), but the diffusion constant was globally strongly reduced about one order of magnitude. Deconvolution of the distribution of diffusion constants and labelling of the pixels according to their population (Figure 3b; slow - red, intermediate – orange and fast - yellow) confirms this observation. Deconvolution constitutes a relative classification of regions with higher/lower or average values of a given parameter compared to other regions within the same nucleus and permits visualization of the redistribution of populations upon varying nuclear activity. In particular, the population diffusing quickly with respect to the bulk chromatin in quiescent cells is reduced upon serum stimulation (Figure 3c), while most pixels adjoining

bioRxiv preprint doi: https://doi.org/10.1101/405969; this version posted November 25, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

the other populations are re-classified as intermediate populations. Connected areas with high mobility appear eroded and become reduced to their central region (Figure 3b). In contrast, the slow population occupying ~ 6 % of the nuclear area, which was almost exclusively located at the nuclear periphery (Supplementary Figure 10), was invariant to transcriptional changes (Figure 3b,c). Despite considerable relative reorganization of populations, the type of diffusion governing the derived parameter maps showed only moderate changes upon inducing transcriptional activity (Figure 3d).

Anomalous diffusion dominated across the entire nucleus ($0.3 \leq \alpha \leq 0.73$) forming a mosaic-like pattern, which underwent, compared to maps of the diffusion constant, considerable remodeling upon transcriptional activation (Figure 3e, f). Within this pattern, patches of super-diffusive (red: $\alpha > 1$) motion segregated into distinct islands which became more fragmented upon serum stimulation. Random contacts or re-distribution of existing contacts of the interchromatin with itself may give rise such variations of the anomalous exponent upon serum stimulation [26]. Since, in a polymer modeling approach, the diffusion constant of chromatin fibers was unaffected for moderate degrees of cross-linking [27], we expect that association of proteins with DNA upon serum stimulation could favor global decrease of mobility in vivo.

Hi-D reveals high-resolution spatial changes in mobility and anomaly of chromatin diffusion in single cells. Further investigation may tell us if all or a subset of visible physical domains correspond to the ones determined using chromosome conformation capture (Hi-C).

Transcription status modulates chromatin diffusion processes

To further explore the relationship between global chromatin dynamics and transcriptional activity, we examined the dynamics of RNA polymerase II (RPB1-Dendra2; RNA Pol II) in live U2OS cell nuclei (Figure 4a) at different transcriptional states. Hi-D analysis resolved three mobility populations of RNA Pol II (Figure 4b), which is consistent with previous studies that revealed three kinetically different

bioRxiv preprint doi: https://doi.org/10.1101/405969; this version posted November 25, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

groups of RNA Pol II based on their chromatin-binding times [28,29]. The three dynamic populations in actively transcribing cells (grown in normal condition) exhibited significantly higher diffusion constants as in transcriptionally less active cells (serum-starved cells) (Figure 4b). In quiescent cells, the share of quickly diffusing RNA Pol II complexes was reduced compared to actively transcribing cells and, upon elongation inhibition using 5,6-Dichloro-1-β-D-ribofuranosylbenzimidazole (DRB), the slowly diffusing fraction was greater than in untreated cells, indicating tenacious immobilization of RNA Pol II on the DNA template after initiation (Figure 4c). The average diffusion constants in serum starved and DRB treated cells are similar for the three populations, suggesting that RNA Pol II not phosphorylated in absence of serum coherent with published data identifying serum starved cells in the first place [25]. We then compared the effect of transcriptional activity on chromatin dynamics in serum-starved and –stimulated cells. In contrast to RNA Pol II mobility, the average diffusion constant of DNA in serum-starved U2OS nuclei decreased by nearly one order of magnitude, for all three populations upon addition of serum, and arresting RNA Pol II before elongation did not change the observed diffusion constants, compared to undisturbed transcription (Figure 4d). These cell-population results are consistent with results from single-cell analyses (Figure 3) and strengthen our hypothesis that processes involved in cellular functions considerably reduce the diffusion of chromatin. While in a quiescent state, only a fraction of nuclear processes is active, allowing chromatin to be more mobile. Upon activation, binding of transcription complexes and other proteins acting on the DNA cause the observed reduction in chromatin dynamics. In agreement with this hypothesis, previously enhanced spatial correlation of chromatin dynamics in serum-stimulated cells is also likely accompanied by manifestation of DNA-protein interactions and possibly increased interchromatin contacts, compared to quiescent cells [20].

In serum-stimulated cells, sub-diffusive and super-diffusive regimes were intensified (Figure 4e). The presence of a population with α close to 0.5 in both conditions is indicative of a ground-state Rouse-like chromatin fiber architecture [30,31]. Upon serum stimulation, anomalous diffusion became predominant and its value ($\alpha \sim 0.33$) is indicative of entangled polymers [32], and was seen for a

bioRxiv preprint doi: https://doi.org/10.1101/405969; this version posted November 25, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

single labelled site [13]. The entanglement could effectively be established by random DNA-crosslinking by protein complexes, as became also apparent from polymer simulations inspired by chromosomal capture data [33]. Serum induction also activates the transcription by increase of RNAPII-Ser5P levels that leads to binding of the RNA Pol II complex to DNA, whose motion is therefore greatly constrained to the local DNA's motion. Elongation inhibition however, does likely not change local concentrations of RNA Pol II along the DNA, which also became apparent from chromatin coherence data [20].

Chromatin dynamics is uncoupled from compaction

We next asked if chromatin dynamics is influenced by the compaction of chromatin since heterochromatin is widely believed to be less dynamic than euchromatin [16]. Eu- and heterochromatin domains were determined in serum-starved and –stimulated cells by the fluorescence intensity as explained in [34] (Figure 5a). The average flow magnitude between all subsequent frames was independent of the compaction state of chromatin (Figure 5b) and likewise, the diffusion constant as a measure of the time averaged magnitude of dynamics did not show any difference between euchromatin and heterochromatin (Figure 5c). Heterochromatin at the nuclear periphery overlapped with the peripheral slow motion domain (Figure 5d) consistent with previous findings [16]. Inside the nucleus however, we did not observe any tendency of heterochromatin being associated with a specific mobility population. Instead, we observed that mobility populations were distributed equally among euchromatin and heterochromatin regions (Figure 5e) with the exception in serum-starved cells, that the part of heterochromatin was slightly greater in the slow diffusion population at the nuclear periphery. Furthermore, we found that regions characterized by a specific anomalous exponent did not show a tendency to overlap with either eu- or heterochromatin (Figure 5f, g). These results also hold for MCF7 cells and different fluorescent markers for chromatin (Supplementary Figure 11). These findings were confirmed in NIH3T3 cells expressing both mCherry-

bioRxiv preprint doi: https://doi.org/10.1101/405969; this version posted November 25, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

H2A and GFP-HP1α, two well-established markers for the eu- and heterochromatin (Supplementary Figure 12). We could not detect reduced dynamics of either of the two types of chromatin, but instead, Hi-D analysis hints towards previously proposed liquid phase-separation of HP1α [35]. These results thus suggest that chromatin undergoes diffusion processes which are, in general, unrelated to the compaction level of chromatin. However, compact chromatin is characterized by an increased contact frequency of the chromatin fiber with itself which seems plausible to enhance the extent of coherent chromatin motion. To test this hypothesis, we calculated Moran's Index of Spatial Autocorrelation [36] for the flow magnitude assessed at different time lags in either eu- or heterochromatin (Figure 5h). We found that heterochromatin exhibits enhanced spatial autocorrelation compared to euchromatin across all accessible time lags, which is in line with our hypothesis. Furthermore, spatial autocorrelation decreases with increasing time lag in serum-starved cells, while in serum-stimulated cells, autocorrelation is enhanced in the long-time limit (over 30 seconds). This points to active processes establishing spatial coherence in the long-term [18,20] while random processes such as thermal fluctuations decrease autocorrelation in the range of 10 seconds and over the whole time span in serum-starved cells.

DISCUSSION

Hi-D is a single cell approach that enables tracking of dense structures such as chromatin directly without losing active fluorophore density and with no need for prior experience in sophisticated labelling preparations or advanced microscopy [37]. We showed that Hi-D is an accurate and robust read-out of chromatin dynamics and that the information gain through image analysis afforded by Hi-D alleviates the incompatibility of conventional microscopy for nanoscale mapping of chromatin dynamic properties in living cells.

Hi-D analysis revealed that DNA dynamics can roughly be classified into three sub-populations. The use of three distinct populations does not reflect the real, likely fluent transitions between the

bioRxiv preprint doi: https://doi.org/10.1101/405969; this version posted November 25, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

populations, but emerged as a handy means of visualization and characterization of the highly complex and heterogeneous dynamic landscape of chromatin. The first, a slow mobility fraction prominently located at the nuclear rim, is reminiscent of lamina associated domains (LADs) [38] and the dynamic response to transcriptional stimuli supports the hypothesis that LADs play an important role in attenuating transcription activity and of gene expression patterns [39]. Although less mobile chromatin at the nuclear periphery largely overlapped with long known perinuclear heterochromatin [40,41], Hi-D analysis remarkably points to an overall absence of correlation between chromatin compaction and mobility. Instead, we could strengthen the hypothesis that heterochromatin domains are formed by a liquid phase-separation mechanism, by detection of a characteristic mobility reduction across phase boundaries.

Indeed, sub-regions displaying intermediate and highly diffusive regimes were distributed in a mosaic-like pattern throughout the nucleus, some of them also at the nuclear periphery. Heterochromatin therefore does not exhibit low mobility in nuclear space in general, but may be divided into a more viscous component with reduced mobility due to anchoring to a nuclear lamina and a more rigid elastic LAD tethered to nuclear structures [38,42–44]. The extent of the third, highly mobile fraction, which dominated in the quiescent state, decreased dramatically when cells were serum stimulated. This switch in chromatin mobility is suggestive of dramatically altered molecular crowding, as a result accrued concentration of players of all types in a highly transcribing nucleus, in particular ribosome biogenesis [20,45–47].

Heterogeneous chromatin motion arises due to irregular protein binding along the chromosomes and can lead to thermodynamic or electrostatic self-organisation of nuclear compartments [48,49]. Local patches of large anomalous exponents indicate super-diffusive behaviour of chromatin which may result, among others, from thermal and/or active noise acting on the chromatin fiber [50,51], even in quiescent cells. Chromatin patches with $\alpha < 0.3$ and $\alpha > 1$ respectively correspond in size to one or a few TADs [52]. These two types of patches are present as islands within the general chromatin fraction governed by an anomalous exponent $0.3 < \alpha < 1$. This organization is in good agreement

bioRxiv preprint doi: https://doi.org/10.1101/405969; this version posted November 25, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

with the chromosome territory – interchromatin compartment model [2]. In conclusion, the combination of diffusion constant and anomalous exponent maps provides an integrated view on chromatin dynamics and yields insights into possible mechanisms driving dynamics as well as the local chromosomal organization and re-organization during genomic processes.

Our results support the hypothesis that short treatment of DRB is sufficient to halt most of a cell's transcription[53], but it is believed to have a slight impact on the chromatin configuration of promoters, as well as on upstream events and assembly dynamics of the preinitiation complex (PIC) [54]. In particular, our results show that a few minutes of DRB treatment affects only the assembly dynamics of the RNA Pol II itself. In contrast, longer treatment with DRB as well as longer serum starvation, are expected to result in massive PIC disassembly. We believe therefore that increased chromatin mobility upon serum starvation correlates with PIC breakdown. Conversely, when serum is present in the medium, activity of RNA Pol II and of PIC enzymes is restored and PIC will assemble onto DNA, both pervasively across the genome at background levels ("pervasive transcription") and more stably at genes [55,56]. Reduced mobility of chromatin in the presence of serum suggests that stable PIC binding may serve as an anchoring function for individual chromatin fibres, on top of its essential function in transcription initiation.

Hi-D can be applied to real-time imaging of any fluorescent molecule to obtain comprehensive maps of their dynamic behavior in response to stimuli, inhibitors or disruptors of nuclear functions and integrity. Here, Hi-D is developed to track bulk objectives in two dimensions, however, it could be extended to full three-dimensional analysis. It could also be combined with the tracking of specific loci to zoom in on their environment.

bioRxiv preprint doi: https://doi.org/10.1101/405969; this version posted November 25, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

METHODS

Cell Culture. A Human U2OS osterosarcoma cell line (for DNA imaging) and MCF-7 cells (ATCC) were maintained in Dulbecco's modified Eagle's medium (DMEM) containing phenol red-free and DMED-12 (Sigma-Aldrich), respectively. For RNA Pol II imaging, a U2OS cell line stably expressing RPB1 (subunit of RNA Pol II) fused with Dendra2 was a gift from Maxime Dahan (Curie Institute, Paris) and constructed as previously described in [46]. Medium was supplemented with Glutamax containing 50 µg/ml gentamicin (Sigma-Aldrich), 10% Fetal bovine serum (FBS), 1 mM sodium pyruvate (Sigma-Aldrich) and G418 0.5 mg/ml (Sigma-Aldrich) at 37°C with 5% CO2. Cells were plated for 24 h on 35 mm petri dishes with a #1.5 coverslip like bottom (µ-Dish, Ibidi, Biovalley) with a density of about 10^5 cells/dish.

Cell starvation and stimulation. For starvation mode, cells were incubated for 24 h at 37°C before imaging with serum-free medium (DMEM, Glutamax containing 50 µg/ml gentamicin, 1 mM sodium pyruvate, and G418 0.5 mg/ml). Just before imaging, cells were mounted in L-15 medium. For stimulation, 10% FBS was added to the L-15 medium for 10-15 minutes.

DNA staining. U2OS and MCF-7 cell lines were labelled by using SiR-DNA (SiR-Hoechst) kit (Spirochrome AG). For DNA was labelled as described in [57]. Briefly, we diluted 1 mM stock solution in cell culture medium to concentration of 2 µM and vortexed briefly. On the day of the imaging, the culture medium was changed to medium containing SiR-fluorophores and incubated at 37°C for 30-60 minutes. Before imaging, the medium was changed to L-15 medium (Liebovitz's, Gibco) for live imaging.

DRB treatment: The U2OS cells for both DNA and RNA Pol II images were treated with 100 µM 5,6 Dichlorobenzimidazole 1-β-D-ribofuranoside (DRB; Sigma-Aldrich) for transcription inhibition prior to live-cell image acquisition. DRB was diluted in the L-15 (Leibovitz) imaging medium that was supplemented with 10% FBS, DMEM, Glutamax containing 50 µg/ml gentamicin, 1 mM sodium

bioRxiv preprint doi: https://doi.org/10.1101/405969; this version posted November 25, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

pyruvate, and G418 0.5 mg/ml. The imaging medium was changed with fresh L-15 medium containing DRB and incubated under the microscope for 15 min before imaging.

Cell fixation: U2OS cells were washed with a pre-warmed (37 °C) phosphate buffered saline (PBS) and followed by fixation with 4% (vol/vol) Paraformaldehyde in PBS for 10-20 min at room temperature. Images were recorded at room temperature in PBS, after washing the cells with PBS (three times, 5 min each).

DNA live cell imaging: Cells were placed in a 37 °C humid incubator by controlling the temperature and CO_2 flow using H201-couple with temperature and CO_2 units. Live chromatin imaging was acquired using a DMI8 inverted automated microscope (Leica Microsystems) featuring a confocal spinning disk unit (CSU-X1-M1N, Yokogawa). An integrated laser engine (ILE 400, Andor) was used for excitation with a selected wavelength of 647 nm and 140 mW as excitation power. A 100x oil immersion objective (Leica HCX-PL-APO) with a 1.4 NA was chosen for a high resolution imaging. Fluorescence emission of the SiR–Hoechst was filtered by a single-band bandpass filter (FF01-650/13-25, Semrock, Inc.). Image series of 150 frames (5 fps), with exposure time of 150ms per frame were acquired using Metamorph software (Molecular Devices), and detected using sCMOS cameras (ORCA-Flash4.0 V2) and (1×1 binning), with sample pixel size of 65 nm. All series were recorded at 37°C.

RNA Pol II live cell imaging: Image series of 150 frames were recorded with an exposure time of 200 ms using a Nipkow-disk confocal system (Revolution, Andor) featuring a confocal spinning disk unit (CSU22, Yokogawa). A diode-pumped solid-state laser with a single wavelength of a 488 nm (Coherent) at 5-10% laser power was used for excitation of RPB1 fused to Dendra2. A 100x an oil immersion objective (Plan Apo 1.42, Nikon) was used for imaging. The fluorescent emission signal was filtered through an emission filter (ET525/30-25, Semrock, Inc.), and detected at 512/18 nm on a cooled electron multiplying charge-coupled device camera (iXon Ultra 888), with sample pixel size of 88 nm.

bioRxiv preprint doi: https://doi.org/10.1101/405969; this version posted November 25, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

Image processing

Denoising. Raw images were denoised using non-iterative bilateral filtering [58]. While Gaussian blurring only accounts for the spatial distance of a pixel and its neighbourhood, bilateral filtering additionally takes the difference in intensity values into account and is therefore an edge-preserving method. Abrupt transitions from high- to low-intensity regions (e.g. heterochromatin to euchromatin) are not over-smoothed. Images of varying noise levels were treated with a bilateral filter with half-size of the Gaussian bilateral filter window of 5 pixels, the spatial-domain standard deviation value was set to 5 pixels and the intensity-domain standard deviation was varied from 0.3 to 0.8 for decreasing levels of the signal-to-noise ratio from 26 dB to 16 dB ($SNR = 10\log_{10}(I^2/\sigma^2)$, with the signal power I^2 and the noise variance σ^2).

MSD analysis and model selection by using Bayesian inference

In order to carry out a MSD analysis locally, the spatial dependency of the Mean Squared Displacement (MSD) can be written explicitly:

$$\mathrm{MSD}(\vec{r}_0, \tau) = \langle \left|\vec{\xi}_{r_0}(t+\tau) - \vec{\xi}_{r_0}(t)\right|^2 \rangle_t,$$

where $\vec{\xi}_{r_0}(t)$ is the position at time t of a virtual particle with initial position $\vec{r}_0$, $\tau = \{\Delta t, 2\Delta t, \dots, (N-1)\Delta t\}$ are time lags where Δt is the time difference between subsequent images and the average $<\cdot>_t$ is taken over time. The resulting MSD is a function of the initial position $\vec{r}_0$ and the time lag τ.

MSD models: The MSD can be expressed analytically for anomalous diffusion (DA), confined diffusion (DR) and directed motion (V) in two dimensions as

$$MSD_{DA}(\tau) = 4D_\alpha \tau^\alpha \tag{1}$$

$$MSD_{DR}(\tau) = R_C^2\left(1 - \mathrm{e}^{-\frac{4D\tau}{R_C^2}}\right) \tag{2}$$

$$MSD_V(\tau) = v^2\tau^2 \tag{3}$$

bioRxiv preprint doi: https://doi.org/10.1101/405969; this version posted November 25, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

where D_α is the diffusion constant in units of $\mu m^2/s^\alpha$, α is its anomalous exponent, v $[\mu m/s]$ its velocity and R_C $[\mu m]$ is the radius of a sphere within the particle is confined [59]. The case $\alpha = 1$ is known as free diffusion, $0 < \alpha < 1$ corresponds to anomalous diffusion and $1 < \alpha \leq 2$ corresponds to superdiffusion. Strictly speaking, each generalized diffusion constant D_α has different units, corresponding to the specific value of α. However, we refer to it as the diffusion constant D throughout the text for simplicity. Additionally to eq. (1)-(3), different types of motion can appear overlaying, resulting in a linear combination of the equations above. For example, anomalous motion can be superimposed on an underlying drift and the resulting MSD reads $MSD_{DAV}(\tau) = MSD_{DA}(\tau) + MSD_V(\tau)$. We found that anomalous and confined diffusion appears very similar in experimental data and therefore decided in favor for anomalous diffusion to describe our data (Supplementary Note 3). The abbreviations used in this study are summarized in Table 1. As experimental data is usually subject to noise, a constant offset o is added to every model.

Abbreviation	**Model**	**Formula**
D	Free diffusion	$MSD_D(\tau) = 4D\tau + o$
DA	Anomalous diffusion	$MSD_{DA}(\tau) = 4D_\alpha\tau^\alpha + o$
V	Drift	$MSD_V(\tau) = v^2\tau^2 + o$
DV	Free diffusion + drift	$MSD_{DV}(\tau) = 4D\tau + v^2\tau^2 + o$
DAV	Anomalous diffusion + drift	$MSD_{DAV}(\tau) = 4D_\alpha\tau^\alpha + v^2\tau^2 + o$

Table 1: Overview over possible Mean Squared Displacement models

MSD model selection: The MSD is calculated for every pixel independently, resulting in a space- and time lag-dependent MSD. It is known that living cells can behave largely heterogeneous[3,60]. Ad-hoc, it is not known which particle undergoes which kind of diffusion. Fitting a MSD curve with a wrong model might result in poor fits and highly inaccurate determination of the mentioned parameters. For this reason, we use a Bayesian inference approach to test different models for any given MSD curve as proposed by Monnier *et al.* [21]. Given the data $Y = \{Y_1, \dots, Y_n\}$ and K model candidates $M =$

bioRxiv preprint doi: https://doi.org/10.1101/405969; this version posted November 25, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

$\{M_1, \dots, M_K\}$, each with its own (multidimensional) parameter set $\theta = \{\theta_1, \dots, \theta_K\}$, we want to find the model $M_k(Y, \theta_k)$ such that the probability that $M_k(Y, \theta_k)$ describes the data, given the set of models to test, is maximal. By Bayes' theorem, the probability for each model is given by

$$P(M_k|Y) = \frac{P(Y|M_k)P(M_k)}{P(Y)}$$

If there is no reason to prefer one model over the other, the prior probability of each model $P(M_k)$ is equal. The parameter set which is used to describe the data, given a fixed model, strongly influences the probability. Therefore, it is crucial to estimate the optimal parameters for every model in order to calculate the model probabilities. The probability that the data Y is observed, given the model M_k described by the model function $M_k(x; \theta_k)$ and any parameter set θ_k is approximated by a general multivariate Gaussian function [61]

$$P(Y|\theta_k, M_k) = \frac{1}{\sqrt{(2\pi)^n \det(C)}} \cdot \exp\left\{-\frac{1}{2}[Y - M_k(x; \theta_k)]^T \cdot C^{-1} \cdot [Y - M_k(x; \theta_k)]\right\}$$

where C is the empirical covariance matrix of the data and the prefactor is a normalizing factor. This equation has an intuitive meaning. Assume we test a model M_k parametrized by θ_k to find out if it describes the data Y. The exponential function consists of the term $[Y - M_k(x; \theta_k)]$, i.e. the residuals of the data and the describing model. If the residuals are small, i.e. the model describes the data well, the exponent is small and the probability $P(Y|\theta_k, M_k)$ seeks 1. On the other hand, the worse the fit, the greater the resulting residuals and the probability seeks asymptotically to 0. The factor C^{-1} accounts for the covariance in the data. The covariance matrix for a set of MSD curves normally shows large values for large time lags as the uncertainty increases and MSD curves diverge. The covariance matrix implicitly introduces a weight to the data, which is small for large variances and large where the data spreads little. This fact avoids cutting of the MSD curve after a specific number of time lags, but instead includes all available time lags weighted by the covariance matrix. The approach is illustrated in (Supplementary Figure 6b) with the covariance matrix exemplary shown in the inset. In

bioRxiv preprint doi: https://doi.org/10.1101/405969; this version posted November 25, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

case of uncorrelated errors, non-diagonal elements are zero, but the approach keeps its validity [62] and follows an ordinary least-squares regression.

Given the best estimate of the parameter set for a model, the model and its corresponding parameters are chosen so that their probability to describe the data is maximal: $\hat{\theta}_{k,MLE} = \arg\max_{\theta_k} P(Y|\theta_k, M_k)$.

It has to be stressed that values of the anomalous exponent scatter around 1, but do not assume the value 1 (e.g. Figure 1c, middle panel). This is due to the model selection procedure, selecting the simplest model consisting with the data. In case that the underlying motion is well described by free diffusion, α is inherently set to 1 and classified as free diffusion rather than anomalous diffusion. The descriptions of free diffusion or anomalous diffusion with $\alpha = 1$ are equivalent, but the free diffusion model contains one parameter less and is therefore preferred leading to abundance of α values close to 1 in the parameter maps and histograms. To carry out the MSD analysis locally, we choose to take the 3x3 neighborhood of a pixel, detect possible outliers therein by the interquartile range criterion [63] and calculate the error covariance matrix of the data within the pixel's neighborhood. The restriction to a single pixel and its neighborhood allows us to carry out the MSD analysis of trajectories locally, in contrast to an ensemble MSD in previous studies [18], revealing only average information over many trajectories. The choice of a 3x3 window is reasonable with regard to the equivalently chosen filter size in the Optical Flow estimation. The flow field in this region is therefore assumed to be sufficiently smooth. All calculations, except for the General Mixture Model analysis, were carried out using MATLAB (MATLAB Release 2017a, The MathWorks, Inc., Natick, Massachusetts, United States) on a 64-bit Intel Xeon CPUE5-2609 1.90 GHz workstation with 64 GB RAM and running Microsoft Windows 10 Professional.

Deconvolution of sub-populations

Regarding especially the distribution of the diffusion constant, an analytical expression can be found assuming that the diffusion constant was calculated from a freely diffusing particle ($\alpha = 1$)[64]. However, we find anomalous diffusion to a large extent in our data (e.g. Figure 2b, d and Figure 4d)

bioRxiv preprint doi: https://doi.org/10.1101/405969; this version posted November 25, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

and, to our knowledge, an analytical expression cannot be found for distributions of anomalous exponent, radius of confinement and drift velocity. We therefore aim to deconvolve the parameter sets in a rather general manner, for which we use a General Mixture model (GMM), a probabilistic model composed of multiple distributions and corresponding weights. We describe each data point as a portion of a normal or log-normal distribution described by

$$f_N(Y|\mu,\sigma^2) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(Y-\mu)^2}{2\sigma^2}} \text{ and}$$

$$f_L(Y|\mu,\sigma^2) = \frac{1}{Y\sqrt{2\pi}s} e^{-\frac{(\ln(Y)-m)^2}{2s^2}},$$

respectively. The logarithmic mean m and standard deviation s are related to the mean and standard deviation of the normal distribution via [65].

$$\mu = \exp\left(m + \frac{s^2}{2}\right)$$

$$\sigma^2 = \exp(2m + s^2)\left(e^{\sigma^2} - 1\right)$$

We consider up to three subpopulations to be found in our data and model the total density estimate as a superposition of one, two or three subpopulations, i.e. the Mixture Model reads

$$f_{GMM}(Y) = \sum_k w_k\, f_k\left(Y\middle|\mu_k,\sigma_k^2\right)$$

for both normal and log-normal distributions, where to sum goes to 1, 2 or 3. The variable w_k describes the weights for each population (or component), which satisfy $0 \leq w_k \leq 1$ and sum up to unity. The weights of each component are directly proportional to the area of the histogram covered by this component and therefore its existence in the data set.

General Mixture Model analysis

Let $Y = \{Y_1, \dots, Y_n\}$ denote n data points. For the scope of this description, assume Y to be a one-dimensional variable. Further assume that the data cannot be described by a single distribution, but by a mixture of distributions. A deconvolution of the data into sub-populations faces the following problem: Given a label for each data point, denoting the affiliation to a population, one could group

bioRxiv preprint doi: https://doi.org/10.1101/405969; this version posted November 25, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

corresponding data points and find the parameters of each population separately using a maximum likelihood estimation or other methods. On the other hand, if we had given the model parameters for each population, labels could in principle be inferred from the likelihood of a data point being described by a population or another. The problem can be formulated by Bayes' rule (M indicates model, D indicates data)

$$P(M|D)P(D) = P(D|M)P(M).$$

Here, $P(M|D)$ is the posterior probability of the model given the data, which is the aim to calculate. We assign a data point to the component, which maximizes $P(M|D)$. The probability to observe the data given a model is described by $P(D|M)$, i.e. the likelihood function. $P(M)$ is the prior for the models to be chosen from. In our case, we have no prior beliefs on the models (all models are equally likely) such that $P(M)$ is uniform. Lastly, the probability $P(D)$ does not depend on the models and can therefore be dropped.

Unfortunately, neither labels, that is $P(M|D)$, nor model parameters and weights are known a priori. The problem can be approached by an Expectation-Maximization (EM) scheme: Without any prior beliefs about the data distribution, one starts with a simple estimate of model parameters, e.g. a k-means clustering estimate and iterates subsequently between the two following steps until convergence:

Expectation step: Calculation of the probability that the component with the current parameter estimate generated the sample, i.e. $P(D|M)$.

Maximization step: Update the current parameter estimate for each component by means of a weighted maximum likelihood estimate, where the weight is the probability that the component generated the sample.

We illustrate the results of the EM algorithm exemplary in Supplementary Figure 8. From the input data (Supplementary Figure 8a), represented as histogram, both the likelihood $P(D|M)$ (Supplementary Figure 8b) and the posterior (Supplementary Figure 8c) is obtained. The sum of sub-

bioRxiv preprint doi: https://doi.org/10.1101/405969; this version posted November 25, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

populations corresponds to the overall probability distribution (shown in black) with different model parameters and weights found by maximizing the likelihood function. The posterior describes the probability of data points to fall under each population, i.e. $\sum_k P(M_k|D) = 1$. The data points are assigned to those population, for which $P(M_k|D)$ is maximum, resulting in labeled data. The labels are subsequently mapped in two dimensions, visualizing spatial correspondence of slow, intermediate and fast sub-populations (Supplementary Figure 8d). The GMM analysis is carried out using the pomegranate machine learning package for probabilistic modeling in Python [66].

Selection of subpopulations by the Bayesian Information Criterion (BIC)

A priori, it is not unambiguously clear from how many populations the data is sampled and which form the subpopulations take. We therefore assess the suitability of each model by means of the Bayesian Information Criterion (BIC), which is calculated by[67]

$$BIC = -2\ln(\hat{L}) + p\, ln(n), \tag{3}$$

where $\hat{L}$ is the maximum likelihood of the maximum likelihood estimation (MLE) estimate, p denotes the number of parameters in the model and n is the number of data points used for the fit. Among a family of models, the one with the lowest BIC is considered to describe the data best, taking into account competing complexity of models. A large likelihood of a model favors it to describe the data well, while on the other hand the model is penalized if many parameters are involved in the model by the second term in (3). Therefore, the BIC prevents overfitting. In order to judge which model is appropriate for our data, we tested all considered models for each histogram and assessed the optimal model by means of the BIC. The fraction of all histograms which described best by one of the six models considered is given in Table 2. Based on the objective judgement of the fit using the BIC, we chose for each parameter the model which best describes the largest fraction of histograms (Table 2, bold cells).

bioRxiv preprint doi: https://doi.org/10.1101/405969; this version posted November 25, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

	Normal distribution			Log-normal distribution		
#populations	1	2	3	1	2	3
D	0.01	0.01	0.06	0.01	0.40	**0.51**
α	0	0.05	**0.68**	0.02	0.03	0.22

Table 2: Fraction of histograms over all parameters best described by one of the six models considered. The highest fraction is shown in bold.

SUPPLEMENTARY INFORMATION

The Supplementary Information contains details on the conversion from flow fields to trajectories, detailed discussion on the distinction between confined and anomalous diffusion and a comparison of Single Particle Tracking and Optical Flow for reconstruction of motion in intermediate to high labelling density scenarios.

DATA AVAILABILITY

The code of the HiD method can be downloaded from https://github.com/romanbarth/Hi-D , and raw data are available on request.

ACKNOWLEDGMENTS

We thank Amy Strom (Gary Karpen Lab, UC Berkeley, USA) for providing images of HP1 in NIH3T3 cells, Silvia Kocanova (CBI Toulouse) for assistance with acquiring some movies, Raphael Mourad (CBI Toulouse) and Genevieve Fourel (ENS Lyon) for fruitful discussions, and Alain Kamgoué for assistance with cluster computer setting up. We acknowledge support from the LITC imaging platform, CBI Toulouse. This work was supported by grants (to K. B.) from the ANR, IDEX Toulouse strategic actions, the Foundation ARC and INSERM Cancer and Epigenetics.

bioRxiv preprint doi: https://doi.org/10.1101/405969; this version posted November 25, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

AUTHOR CONTRIBUTIONS

H.A.S. conceived the method, performed the experimental work, developed the data processing algorithm and analyzed the data; R.B, developed the data processing algorithm, wrote the code, carried out simulations and analyzed the data; H.A.S. and K.B. interpreted the results; H.A.S., R.B. and K.B. wrote the manuscript.

CONFLICT OF INTEREST

The authors declare no competing financial interests.

REFERENCES

1. Serizay J, Ahringer J. ScienceDirect Genome organization at different scales : nature , formation and function. Curr Opin Cell Biol. Elsevier Ltd; 2018;52:145–53.

2. Cremer T, Cremer M, Hübner B, Strickfaden H, Smeets D, Popken J, et al. The 4D nucleome: Evidence for a dynamic nuclear landscape based on co-aligned active and inactive nuclear compartments. FEBS Lett. 2015;589:2931–43.

3. Banerjee B, Bhattacharya D, Shivashankar G V. Chromatin structure exhibits spatio-temporal heterogeneity within the cell nucleus. Biophys J. 2006;91:2297–303.

4. Fraser P, Bickmore W. Nuclear organization of the genome and the potential for gene regulation. Nature. 2007. p. 413–7.

5. Lieberman-aiden E, Berkum NL Van, Williams L, Imakaev M, Ragoczy T, Telling A, et al. Comprehensive Mapping of Long-Range Interactions Reveals Folding Principles of the Human Genome. Science (80-). 2009;326:289–93.

6. Dixon JR, Selvaraj S, Yue F, Kim A, Li Y, Shen Y, et al. Topological domains in mammalian genomes identified by analysis of chromatin interactions. Nature. 2012;485:376–80.

7. Beagrie RA, Scialdone A, Schueler M, Kraemer DCA, Chotalia M, Xie SQ, et al. Complex multi-enhancer contacts captured by genome architecture mapping. Nature. 2017;543:519–24.

8. Rao SSP, Huntley MH, Durand NC, Stamenova EK, Bochkov ID, Robinson JT, et al. A 3D map of the human genome at kilobase resolution reveals principles of chromatin looping. Cell. 2014;159:1665–80.

9. Hansen AS, Pustova I, Cattoglio C, Tjian R, Darzacq X. CTCF and cohesin regulate chromatin loop stability with distinct dynamics. Elife. 2017;

10. Chubb JR, Boyle S, Perry P, Bickmore WA. Chromatin motion is constrained by association with

bioRxiv preprint doi: https://doi.org/10.1101/405969; this version posted November 25, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

nuclear compartments in human cells. Curr Biol. 2002;12:439–45.

11. Chuang CH, Carpenter AE, Fuchsova B, Johnson T, de Lanerolle P, Belmont AS. Long-Range Directional Movement of an Interphase Chromosome Site. Curr Biol. 2006;16:825–31.

12. Chen B, Gilbert LA, Cimini BA, Schnitzbauer J, Zhang W, Li GW, et al. Dynamic imaging of genomic loci in living human cells by an optimized CRISPR/Cas system. Cell. 2013;155:1479–91.

13. Germier T, Kocanova S, Walther N, Elien Bancaud A, Shaban HAHA, Sellou H, et al. Real-Time Imaging of a Single Gene Reveals Transcription-Initiated Local Confinement. Biophys J. 2017;113:1383–94.

14. Levi V, Ruan Q, Plutz M, Belmont AS, Gratton E. Chromatin dynamics in interphase cells revealed by tracking in a two-photon excitation microscope. Biophys J [Internet]. 2005;89:4275–85. Available from: http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=1366992&tool=pmcentrez&rendertype=abstract

15. Bornfleth H, Edelmann P, Zink D, Cremer T, Cremer C. Quantitative motion analysis of subchromosomal foci in living cells using four-dimensional microscopy. Biophys J. 1999;77:2871–86.

16. Nozaki T, Imai R, Tanbo M, Nagashima R, Tamura S, Tani T. Dynamic Organization of Chromatin Domains Revealed by Super-Resolution Live-Cell Imaging. Mol Cell [Internet]. 2017;10:1–12. Available from: http://dx.doi.org/10.1016/j.molcel.2017.06.018

17. Di Pierro M, Potoyan DA, Wolynes PG, Onuchic JN. Anomalous diffusion, spatial coherence, and viscoelasticity from the energy landscape of human chromosomes. Proc Natl Acad Sci [Internet]. 2018;115:7753–8. Available from: http://www.pnas.org/lookup/doi/10.1073/pnas.1806297115

18. Zidovska A, Weitz D a, Mitchison TJ. Micron-scale coherence in interphase chromatin dynamics. Proc Natl Acad Sci U S A. 2013;110:15555–60.

19. Shinkai S, Nozaki T, Maeshima K, Togashi Y. Dynamic Nucleosome Movement Provides Structural Information of Topological Chromatin Domains in Living Human Cells. PLoS Comput Biol. 2016;12.

20. Shaban HA, Barth R, Bystricky K. Formation of correlated chromatin domains at nanoscale dynamic resolution during transcription. Nucleic Acids Res [Internet]. 2018;46:e77. Available from: http://dx.doi.org/10.1093/nar/gky269

21. Monnier N, Guo S-M, Mori M, He J, Lénárt P, Bathe M. Bayesian Approach to MSD-Based Analysis of Particle Motion in Live Cells. Biophys J [Internet]. 2012;103:616–26. Available from: http://linkinghub.elsevier.com/retrieve/pii/S0006349512007187

22. Sergé A, Bertaux N, Rigneault H, Marguet D. Dynamic multiple-target tracing to probe spatiotemporal cartography of cell membranes. Nat Methods. 2008;5:687–94.

23. Di Rienzo C, Gratton E, Beltram F, Cardarelli F. Fast spatiotemporal correlation spectroscopy to

bioRxiv preprint doi: https://doi.org/10.1101/405969; this version posted November 25, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

determine protein lateral diffusion laws in live cell membranes. Proc Natl Acad Sci. 2013;

24. Cho WK, Jayanth N, English BP, Inoue T, Andrews JO, Conway W, et al. RNA Polymerase II cluster dynamics predict mRNA output in living cells. Elife. 2016;5.

25. Nagashima R, Hibino K, Ashwin SS, Babokhov M, Fujishiro S, Imai R, et al. Single nucleosome imaging reveals loose genome chromatin networks via active RNA polymerase II. J Cell Biol. 2019;

26. Shukron O, Seeber A, Amitai A, Holcman D. Single particle trajectory statistic to reconstruct chromatin organization and dynamics. bioRxiv. 2019;559369.

27. Amitai A, Seeber A, Gasser SM, Holcman D. Visualization of Chromatin Decompaction and Break Site Extrusion as Predicted by Statistical Polymer Modeling of Single-Locus Trajectories. Cell Rep. 2017;18:1200–14.

28. Steurer B, Janssens RC, Geverts B, Geijer ME, Wienholz F, Theil AF, et al. Live-cell analysis of endogenous GFP-RPB1 uncovers rapid turnover of initiating and promoter-paused RNA Polymerase II. Proc Natl Acad Sci. 2018;

29. Darzacq X, Shav-Tal Y, de Turris V, Brody Y, Shenoy SM, Phair RD, et al. In vivo dynamics of RNA polymerase II transcription. Nat Struct Mol Biol. 2007;14:796–806.

30. Hajjoul H, Mathon J, Ranchon H, Goiffon I, Mozziconacci J, Albert B, et al. High-throughput chromatin motion tracking in living yeast reveals the flexibility of the fiber throughout the genome. Genome Res. 2013;23:1829–38.

31. Ghosh SK, Jost D. How epigenome drives chromatin folding and dynamics, insights from efficient coarse-grained models of chromosomes. Zhong S, editor. PLOS Comput Biol. Public Library of Science; 2018;14:e1006159.

32. Doi M, Edwards SF. The Theory of Polymer Dynamics. Clarendon Press; 1988.

33. Shukron O, Holcman D. Transient chromatin properties revealed by polymer models and stochastic simulations constructed from Chromosomal Capture data. PLoS Comput Biol. 2017;

34. Wachsmuth M, Knoch TA, Rippe K. Dynamic properties of independent chromatin domains measured by correlation spectroscopy in living cells. Epigenetics and Chromatin. 2016;9:1–20.

35. Strom AR, Emelyanov A V., Mir M, Fyodorov D V., Darzacq X, Karpen GH. Phase separation drives heterochromatin domain formation. Nature. 2017;547:241–5.

36. Chen Y. New Approaches for Calculating Moran's Index of Spatial Autocorrelation. PLoS One. 2013;

37. Manley S, Gillette JM, Patterson GH, Shroff H, Hess HF, Betzig E, et al. High-density mapping of single-molecule trajectories with photoactivated localization microscopy. Nat Methods. 2008;5:155–7.

38. Kind J, Pagie L, Ortabozkoyun H, Boyle S, De Vries SS, Janssen H, et al. Single-cell dynamics of

bioRxiv preprint doi: https://doi.org/10.1101/405969; this version posted November 25, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

genome-nuclear lamina interactions. Cell. 2013;153:178–92.

39. Akhtar W, De Jong J, Pindyurin A V., Pagie L, Meuleman W, De Ridder J, et al. Chromatin position effects assayed by thousands of reporters integrated in parallel. Cell. 2013;154:914–27.

40. Heitz E. Das Heterochromatin der Moose. Jahrbücher für wissenschaftliche Bot. 1928;69:762–818.

41. Solovei I, Wang AS, Thanisch K, Schmidt CS, Krebs S, Zwerger M, et al. LBR and lamin A/C sequentially tether peripheral heterochromatin and inversely regulate differentiation. Cell. 2013;152:584–98.

42. Falk M, Feodorova Y, Naumova N, Imakaev M, Lajoie RB, Leonhardt H, et al. Heterochromatin drives organization of conventional and inverted nuclei Martin. bioRxiv [Internet]. 2018;1–19. Available from: http://dx.doi.org/10.1101/244038

43. van Steensel B, Belmont AS. Lamina-Associated Domains: Links with Chromosome Architecture, Heterochromatin, and Gene Repression. Cell. 2017;

44. Kind J, Pagie L, De Vries SS, Nahidiazar L, Dey SS, Bienko M, et al. Genome-wide Maps of Nuclear Lamina Interactions in Single Human Cells. Cell. 2015;

45. Ghamari A, van de Corput MPC, Thongjuea S, van Cappellen WA, van Ijcken W, van Haren J, et al. In vivo live imaging of RNA polymerase II transcription factories in primary cells. Genes Dev [Internet]. 2013;27:767–77. Available from: http://genesdev.cshlp.org/cgi/doi/10.1101/gad.216200.113%5Cnhttp://genesdev.cshlp.org/content/27/7/767.abstract%0Ahttp://www.ncbi.nlm.nih.gov/pubmed/23592796%0Ahttp://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=PMC3639417

46. Cisse II, Izeddin I, Causse SZ, Boudarene L, Senecal A, Muresan L, et al. Real-time dynamics of RNA polymerase II clustering in live human cells. Science (80-). 2013;341:664–7.

47. Norred SE, Caveney PM, Chauhan G, Collier LK, Collier CP, Abel SM, et al. Macromolecular Crowding Induces Spatial Correlations That Control Gene Expression Bursting Patterns. ACS Synth Biol [Internet]. American Chemical Society; 2018;7:1251–8. Available from: https://doi.org/10.1021/acssynbio.8b00139

48. Haddad N, Jost D, Vaillant C. Perspectives: using polymer modeling to understand the formation and function of nuclear compartments. Chromosom. Res. 2017. p. 35–50.

49. Sewitz SA, Fahmi Z, Aljebali L, Bancroft J, Brustolini OJB, Saad H, et al. Heterogeneous chromatin mobility derived from chromatin states is a determinant of genome organisation in S. cerevisiae. bioRxiv [Internet]. 2017;106344. Available from: https://www.biorxiv.org/content/early/2017/03/02/106344.full.pdf+html

50. Sakaue T, Saito T. Active diffusion of model chromosomal loci driven by athermal noise. Soft

bioRxiv preprint doi: https://doi.org/10.1101/405969; this version posted November 25, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

Matter. 2017;

51. Osmanović D, Rabin Y. Dynamics of active Rouse chains. Soft Matter. 2017;

52. Giorgetti L, Galupa R, Nora EP, Piolot T, Lam F, Dekker J, et al. Predictive polymer modeling reveals coupled fluctuations in chromosome conformation and transcription. Cell. 2014;157:950–63.

53. Bensaude O. Inhibiting eukaryotic transcription: Which compound to choose? How to evaluate its activity? Transcription. 2011;

54. Sainsbury S, Bernecky C, Cramer P. Structural basis of transcription initiation by RNA polymerase II. Nat. Rev. Mol. Cell Biol. 2015.

55. Kapranov P, Cheng J, Dike S, Nix DA, Duttagupta R, Willingham AT, et al. RNA maps reveal new RNA classes and a possible function for pervasive transcription. Science (80-). 2007;

56. Lu Z, Lin Z. Pervasive and dynamic transcription initiation in Saccharomyces cerevisiae . Genome Res. 2019;

57. Lukinavičius G, Blaukopf C, Pershagen E, Schena A, Derivery E, Gonzalez-gaitan M, et al. A far-red DNA stain for live-cell nanoscopy. Nat Commun [Internet]. 2015;61:3–5. Available from: http://dx.doi.org/10.1038/ncomms9497

58. Tomasi C, Manduchi R. Bilateral Filtering for Gray and Color Images. Proc 1998 IEEE Int Conf Comput Vis. 1998;

59. Saxton MJ, Jacobson K. Single-particle tracking: applications to membrane dynamics. Annu Rev Biophys Biomol Struct. 1997;26:373–99.

60. Dickerson D, Gierliński M, Singh V, Kitamura E, Ball G, Tanaka TU, et al. High resolution imaging reveals heterogeneity in chromatin states between cells that is not inherited through cell division. BMC Cell Biol. 2016;17.

61. Seber GAF, Wild CJ. Nonlinear Regression. John Wiley Sons. 2003.

62. He J, Guo S-M, Bathe M. Bayesian approach to the analysis of fluorescence correlation spectroscopy data I: theory. Anal Chem [Internet]. 2012;84:3871–9. Available from: http://pubs.acs.org/doi/abs/10.1021/ac2034369%5Cnpapers2://publication/doi/10.1021/ac2034369%5Cnhttp://www.ncbi.nlm.nih.gov/pubmed/22423978

63. Rousseeuw PJ, Croux C. Alternatives to the Median Absolute Deviation. J Am Stat Assoc. 1993;88:1273–83.

64. Vrljic M, Nishimura SY, Brasselet S, Moerner WE, McConnell HM. Translational diffusion of individual class II MHC membrane proteins in cells. Biophys J. 2002;83:2681–92.

65. Mood AM, Graybill FA, Boes DC. Introduction to the Theory of Statistics. Book. 1974;3:540–1.

66. Schreiber J. Pomegranate: fast and flexible probabilistic modeling in python. 2017;

67. Schwarz G. Estimating the Dimension of a Model. Ann Stat. Institute of Mathematical Statistics;

bioRxiv preprint doi: https://doi.org/10.1101/405969; this version posted November 25, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

1978;6:461–4.

Figure Legends

Figure 1: Hi-D enables spatially resolved mapping of genome dynamic properties at nanoscale resolution in living cells. Workflow: **a)** A series of N=150 confocal microscopy images acquired at 5 fps (left) (here SiR DNA stained living U2OS cells). Dense optical flow was applied to define (N-1) flow fields of the images (center, color coded) based on fluorescence intensity of each pixel (size = 65 nm). Individual trajectories are reconstructed over the duration of acquisition (right). **b)** MSD model selection (left): Trajectories of a 3x3 neighborhood of every pixel are used to calculate a mean MSD curve and its corresponding covariance matrix. By a Bayesian inference approach, the type of diffusion fitting each individual curve is chosen (free diffusion (D), anomalous diffusion (DA), directed motion (V) or a combination (DV) or (DAV). The spatial distribution of the selected models for each pixel is shown as a color map. **c)** Maps of biophysical parameters (D, α and V) extracted from the best describing model per pixel reveal local dynamic behavior of DNA in large domains. The distribution is deconvolved using a General Mixture Model.

Figure 2: Experimental validation of the Hi-D approach. a) Exemplary frame of a simulated time series with low density ($0.001/px^3$) of emitters undergoing Brownian motion convolved by a typical point spread function (left). The time series is subject to Hi-D and Single Particle Tracking estimating the trajectories of emitters. From the estimated trajectories, the MSD is computed and compared to the ground truth diffusion constant. The relative error in the determined diffusion constant is shown. **b)** High density ($0.02/px^3$) of emitters with patches of super-high density ($0.035/px^3$) encircled for visualization, imitating regions of densely packed chromatin. Dashed lines show the optimal value, i.e. perfect agreement between estimation and ground truth. Red lines indicate the median value. Data from 10 independent simulations. Statistical significance assessed by a two-sample Kolmogorov-Smirnov test (***: $p < 0.001$). **c)** MSD curves computed in fixed and living quiescent (- Serum) and serum stimulated (+ Serum) U2OS cells. Diffusion constants for the three average curves were derived by regression yielding $D = (0.87 \pm 0.1) \cdot 10^{-3}\ \mu m^2/s$ for quiescent, $D = (2.6 \pm 0.1) \cdot 10^{-4}\ \mu m^2/s$ for stimulated and $D = (6.1 \pm 0.1) \cdot 10^{-6}\ \mu m^2/s$ for fixed cells. MSD curves show considerably higher MSD values for living cells and diffusion constants are two orders of magnitude higher for living cells thus confirming the detection of motion well above noise background. **d)** Diffusion constants derived from a nucleus corrupted with varying levels of signal-to-noise ratio. Results are consistent up to a lower bound of $\sim$ 20 dB. **e)** Map of diffusion constants

bioRxiv preprint doi: https://doi.org/10.1101/405969; this version posted November 25, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

computed by Hi-D (left) and iMSD (right). Diffusion constants are color-coded from their minimum to their maximum value (blue to yellow; for absolute values see f)). Red arrows indicate regions of high mobility detected by both methods. **f)** Diffusion constants shown in e) and **g)** corresponding values of the anomalous exponent computed by Hi-D (blue) and iMSD (red).

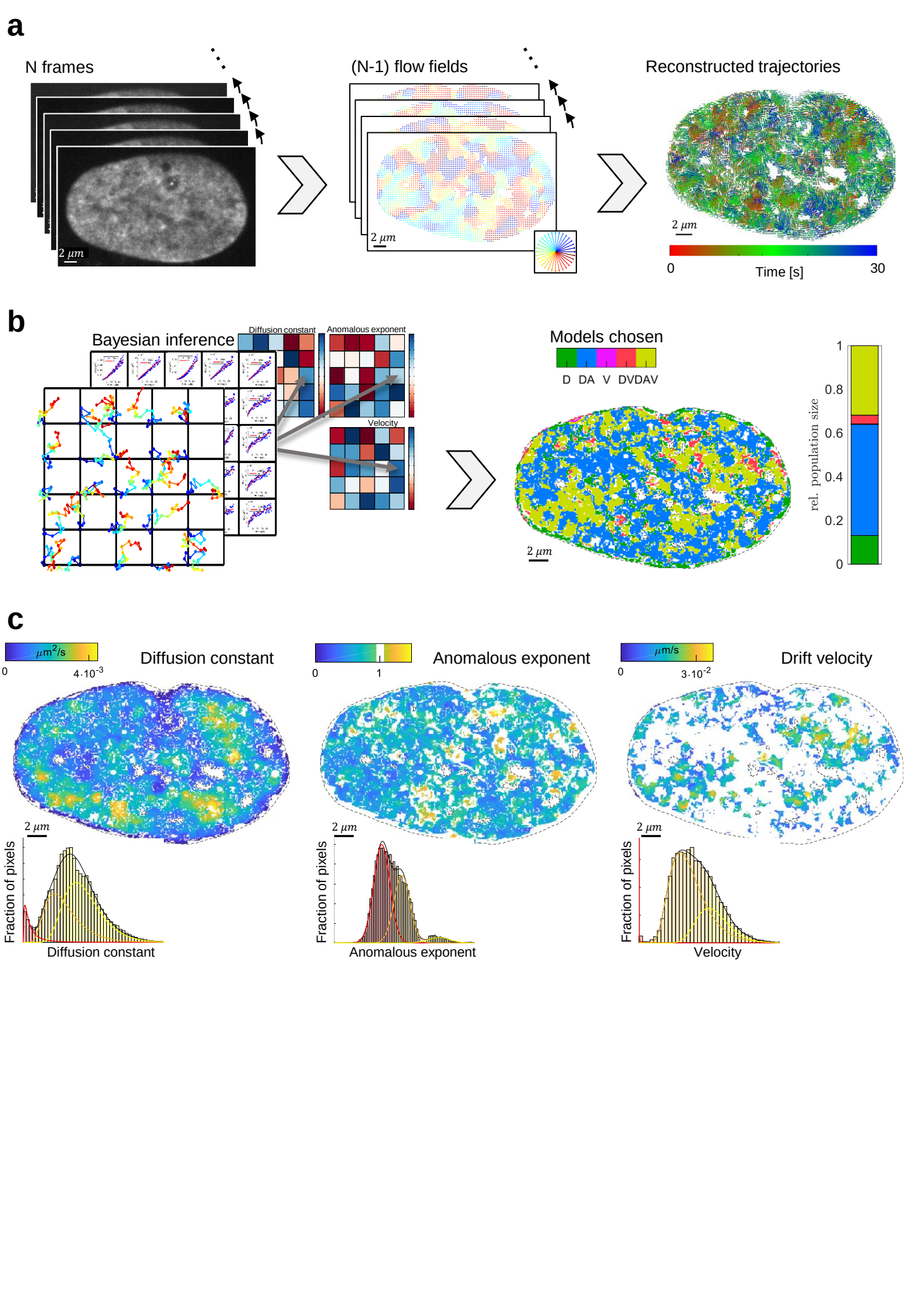
Figure 3: Hi-D maps single-cell biophysical properties of genome conformation and behaviour. a) Diffusion constant spatially mapped onto the nucleus of a quiescent cell (left) and the same nucleus upon serum stimulation **b)** Spatial distribution of populations found by the GMM for the diffusion constant for quiescent (left) and actively transcribing cells (right). **c)** Relative share of populations on the cell volume. Numbers in percent. **d)** For each population in b), the relative share of chosen MSD models is represented as a stacked histogram in the quiescent state (left bars) and actively transcribing state (right bars). **e)** Spatial distribution of populations for the anomalous exponent for quiescent (left) and actively transcribing cells (right). **f)** Detailed insight into the spatial patterning of the low population of the anomalous exponent. The intermediate and high population are shown in light and dark red respectively.

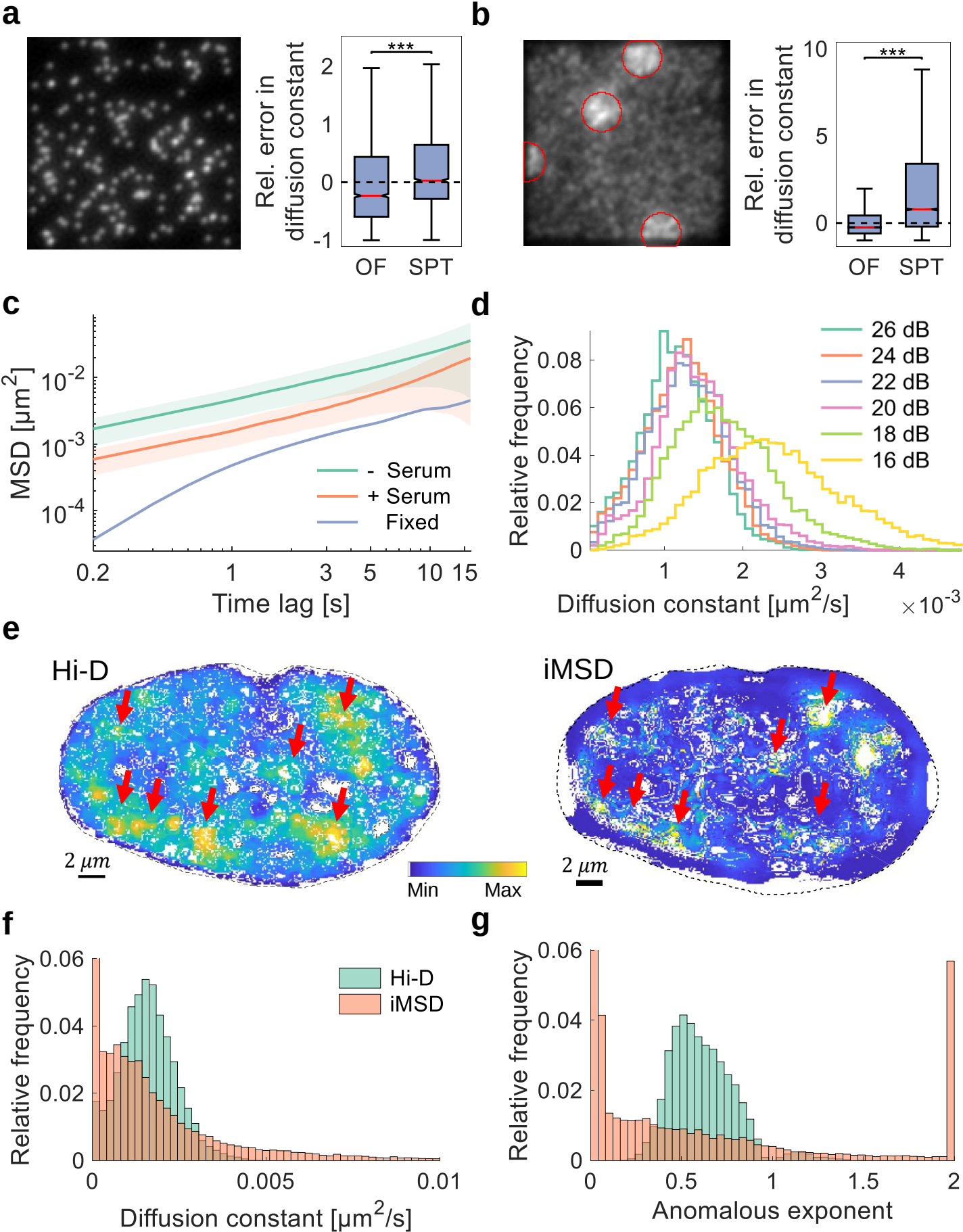
Figure 4: Biophysical properties of chromatin are sensitive to transcriptional activity. a) Fluorescence image of RPB1, a RNA polymerase II subunit, fused to Dendra2 (left), the spatial mapping of the diffusion constant (middle) and population deconvolution (right). **b)** Violin plots of the mean diffusion constant of RNA Pol II (n = 19 cells) for all three mobility groups in quiescent (-Serum), actively transcribing (+ Serum) and elongation inhibited (+ DRB) cells are shown; dark blue, light blue and green denote the slow, intermediate and fast population, respectively. **c)** Relative share of the populations on the cell volume for starved, stimulated and DRB-treated cells. Numbers in percent. **d)** As b) for the diffusion constant and **d)** the anomalous exponent of DNA dynamics (n=13 cells); red, gold and yellow denote the slow, intermediate and fast population, respectively. Statistical significance for assessed by a Friedman test (*: $p < 0.05$, **: $p < 0.01$, ***: $p < 0.001$).

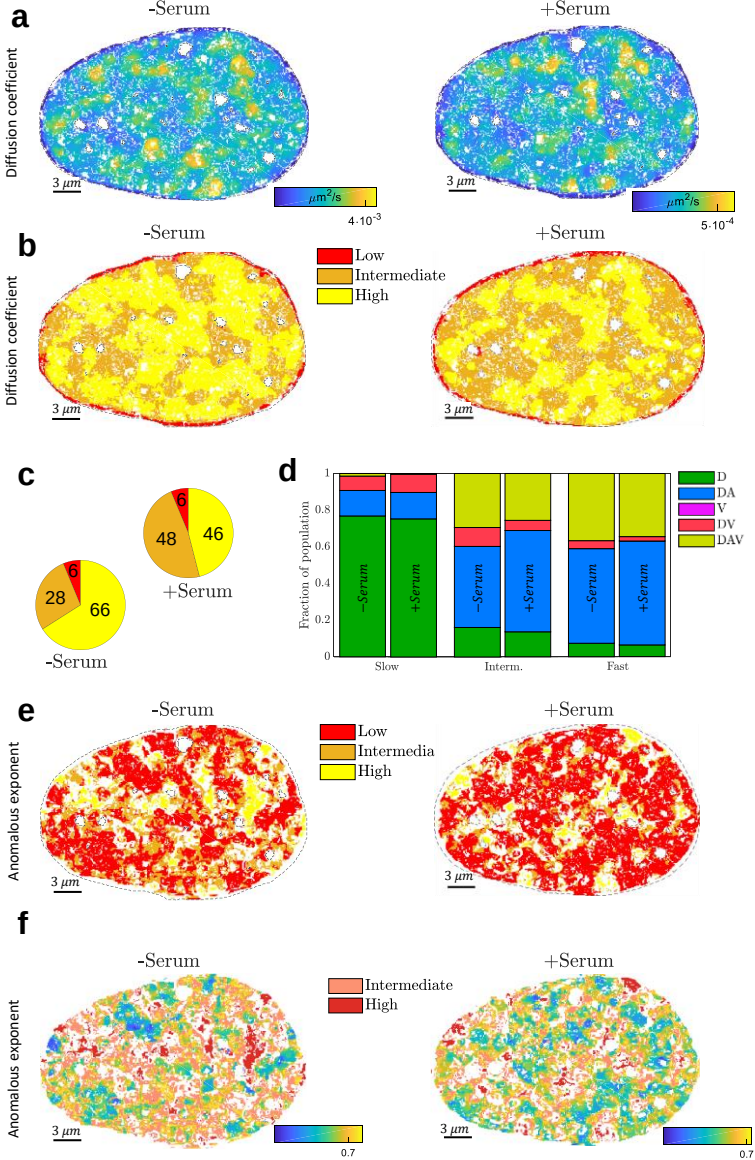
Figure 5: Chromatin compaction and dynamics do no spatially correlate. a) Spatial classification of signal intensity into euchromatin and heterochromatin [34] overlaid on an exemplary fluorescence

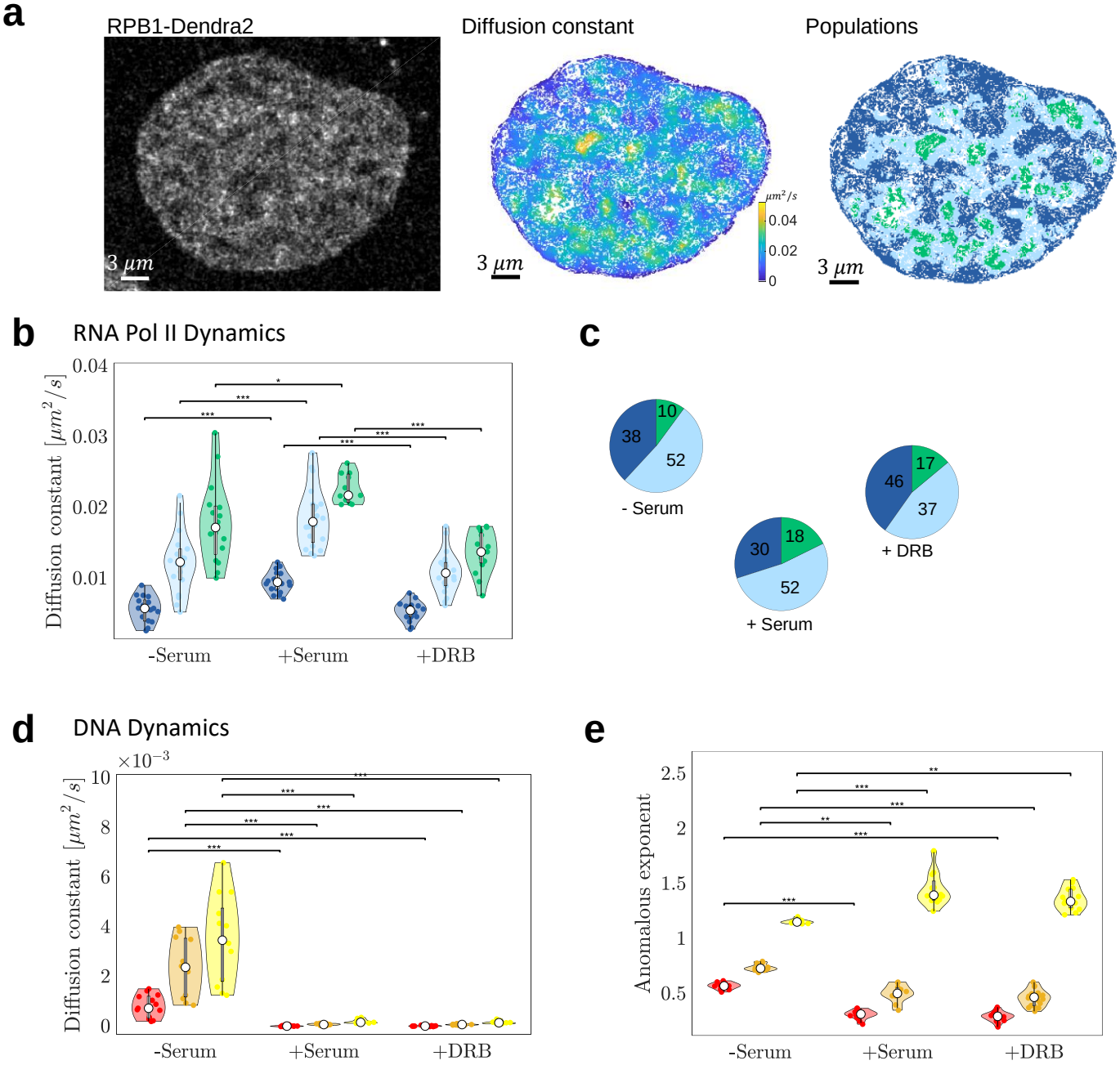
bioRxiv preprint doi: https://doi.org/10.1101/405969; this version posted November 25, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

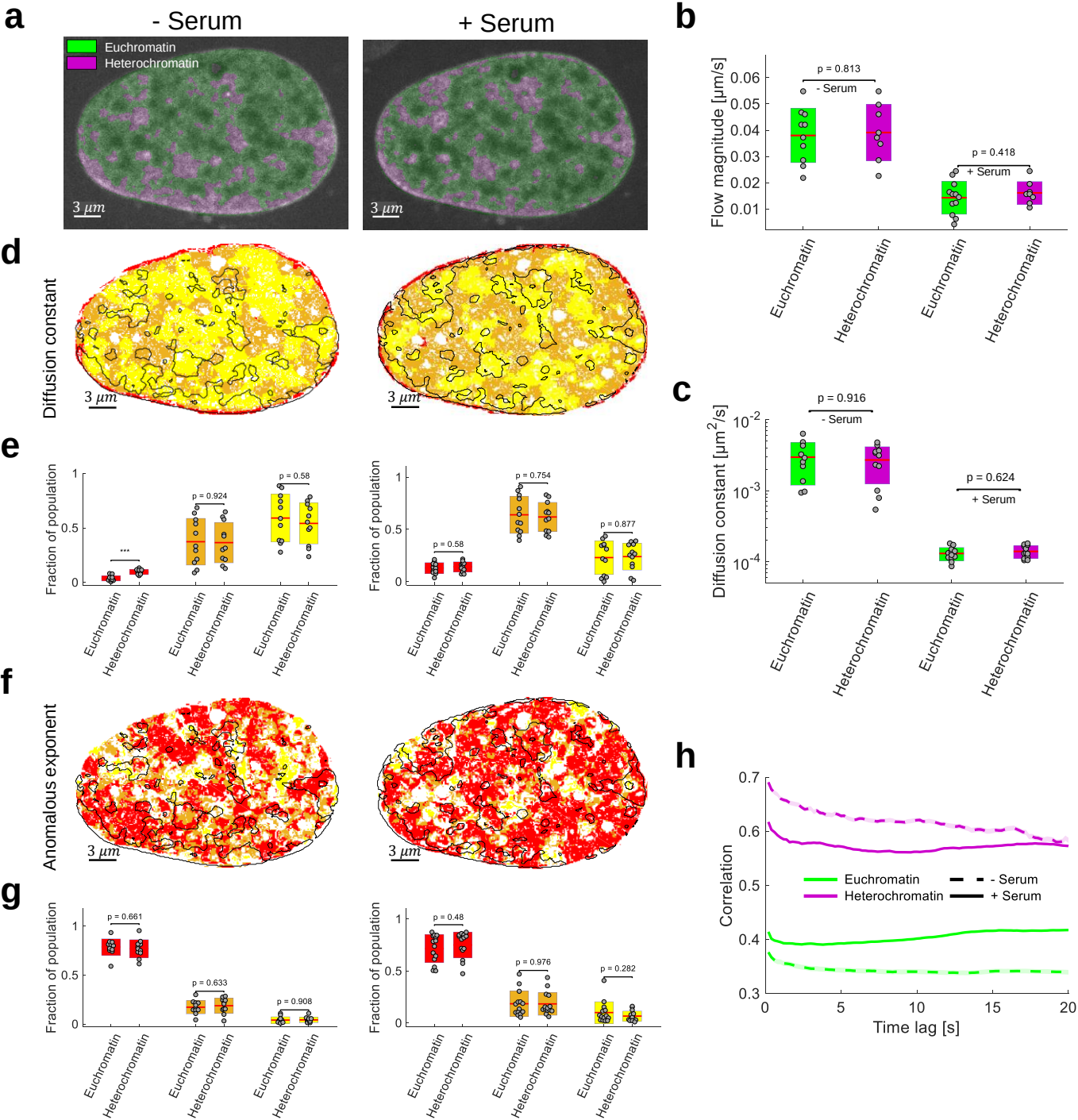
image for quiescent (left) and stimulated (right) cells. **b)** Average flow magnitude and **c)** diffusion constant in euchromatin and heterochromatin for starved (left) and serum stimulated cells (right). Statistical significance assessed by a two-sample t-test. **d)** Overlay with the diffusion populations found by Hi-D. Black solid line corresponds to eu- / heterochromatin region boundaries. **e)** Diffusion populations show a similar distribution over hetero- and euchromatin. The colors refer to the slow, intermediate and high population respectively and each point corresponds to one nucleus. Statistical significance assessed by a two-sample t-test (*: $p < 0.05$, **: $p < 0.01$, ***: $p < 0.001$). **f-g)** Anomalous exponent as d-e). **h)** Spatial autocorrelation at euchromatin (green) and heterochromatin (purple) of the flow magnitude between all accessible time lags in quiescent and serum-stimulated cells.


a
N frames
2 μm
(N-1) flow fields
2 μm
Reconstructed trajectories
2 μm
0
Time [s]
30
b
Bayesian inference
Diffusion constant
Anomalous exponent
Velocity
Models chosen
D DA V DVDAV
2 μm
rel. population size
1
0.8
0.6
0.4
0.2
0
c
μm²/s
0
4·10⁻³
Diffusion constant
2 μm
Fraction of pixels
Diffusion constant
0
1
Anomalous exponent
2 μm
Fraction of pixels
Anomalous exponent
μm/s
0
3·10⁻²
Drift velocity
2 μm
Fraction of pixels
Velocity


a
Rel. error in diffusion constant
***
OF
SPT
b
Rel. error in diffusion constant
***
OF
SPT
c
MSD [μm²]
Time lag [s]
- Serum
+ Serum
Fixed
d
Relative frequency
Diffusion constant [μm²/s]
×10⁻³
26 dB
24 dB
22 dB
20 dB
18 dB
16 dB
e
Hi-D
iMSD
2 μm
Min
Max
f
Relative frequency
Diffusion constant [μm²/s]
Hi-D
iMSD
g
Relative frequency
Anomalous exponent


a
-Serum
+Serum
Diffusion coefficient
3 μm
μm²/s
4·10⁻³
5·10⁻⁴
b
-Serum
+Serum
Low
Intermediate
High
Diffusion coefficient
3 μm
c
6
28
66
-Serum
6
48
46
+Serum
d
Fraction of population
1
0.8
0.6
0.4
0.2
0
D
DA
V
DV
DAV
−Serum
+Serum
Slow
Interm.
Fast
e
-Serum
+Serum
Low
Intermedia
High
Anomalous exponent
3 μm
f
-Serum
+Serum
Intermediate
High
Anomalous exponent
3 μm
0.7


a
RPB1-Dendra2
Diffusion constant
Populations
3 μm
3 μm
3 μm
μm²/s
0.04
0.02
0
b
RNA Pol II Dynamics
Diffusion constant [μm²/s]
0.04
0.03
0.02
0.01
-Serum
+Serum
+DRB
c
10
38
52
- Serum
30
18
52
+ Serum
17
46
37
+ DRB
d
DNA Dynamics
×10⁻³
Diffusion constant [μm²/s]
10
8
6
4
2
0
-Serum
+Serum
+DRB
e
Anomalous exponent
2.5
2
1.5
1
0.5
-Serum
+Serum
+DRB


a
- Serum
+ Serum
Euchromatin
Heterochromatin
3 μm
b
Flow magnitude [μm/s]
p = 0.813
- Serum
p = 0.418
+ Serum
Euchromatin
Heterochromatin
c
Diffusion constant [μm²/s]
p = 0.916
- Serum
p = 0.624
+ Serum
d
Diffusion constant
e
Fraction of population
p = 0.924
p = 0.58
p = 0.754
p = 0.877
f
Anomalous exponent
g
p = 0.661
p = 0.633
p = 0.908
p = 0.48
p = 0.976
p = 0.282
h
Correlation
Euchromatin
Heterochromatin
- Serum
+ Serum
Time lag [s]