Supporting Information

Study of RNA Polymerase II Clustering inside Live-Cell Nuclei Using Bayesian Nanoscopy

Xuanze Chen,[a,b,1] Mian Wei,[a,1] M. Mocarlo Zheng,[a,c,1] Jiaxi Zhao,[d] Huiwen Hao,[a] Lei Chang,[a] Peng Xi[b] and Yujie Sun[a,2]

[a]State Key Laboratory of Membrane Biology, Biodynamic Optical Imaging Center (BIOPIC), School of Life Sciences, Peking University, Beijing 100871, China

[b]Department of Biomedical Engineering, College of Engineering, Peking University, Beijing 100871, China

[c]School of Physics, Peking University, Beijing 100871, China

[d]Department of Physics, Tsinghua University, Beijing 100084, China

[1]Contributed equally to this work

[2]To whom correspondence should be addressed. E-mail: sun_yujie@pku.edu.cn

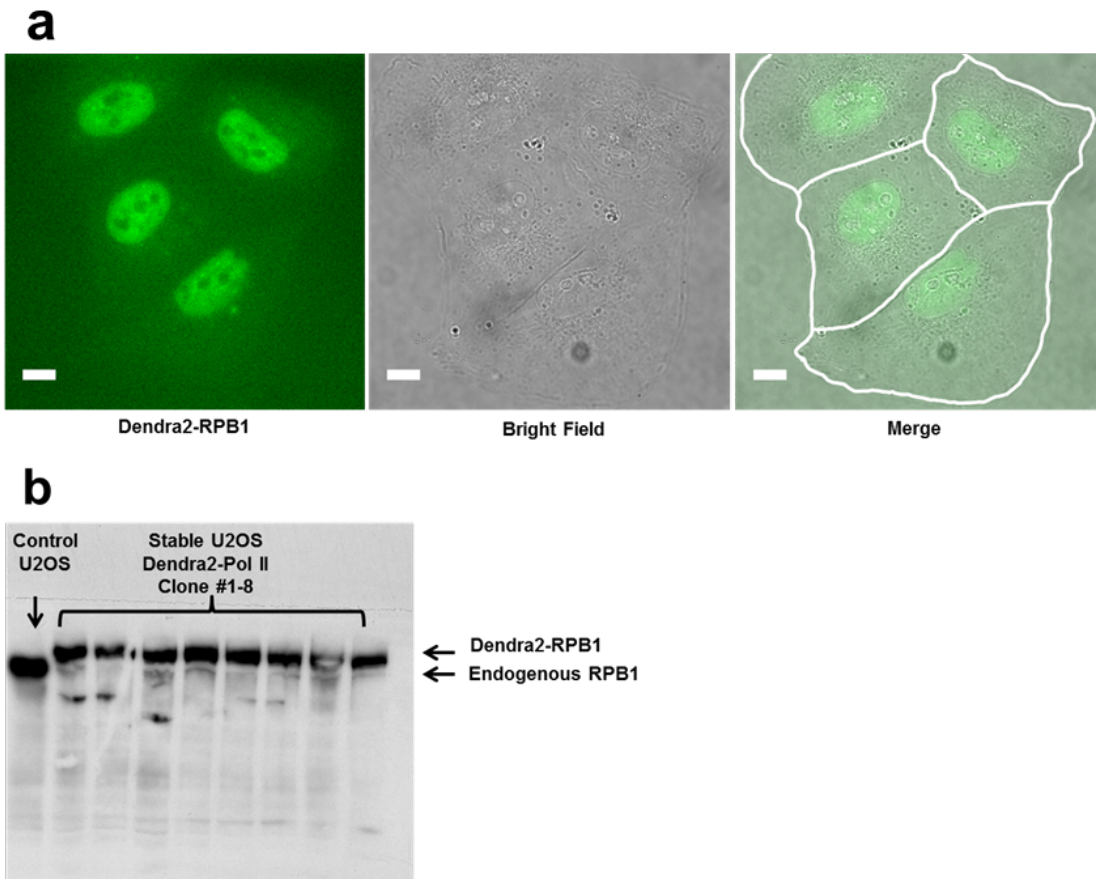

Figure S1. A U2OS cell line that stably expresses the Pol II catalytic subunit (RPB1) fused with Dendra2. (a) Live cell imaging of U2OS cell line shows nuclear localization of Dendra2-RPB1. 488 nm excitation (left). Bright field (middle). Merge (right). White lines delineate cell borders. Scale bar is 10 μm. (b) Western blot shows that endogenous RPB1 is completely replaced with Dendra2-RPB1. The first lane from the left corresponds to wild-type U2OS cells. The following eight lanes correspond to different single clones of U2OS cells stably expressing Dendra2-RPB1. The band shift between the lane of control U2OS and those of stable U2OS cells indicates that Dendra2 is fused to RPB1.

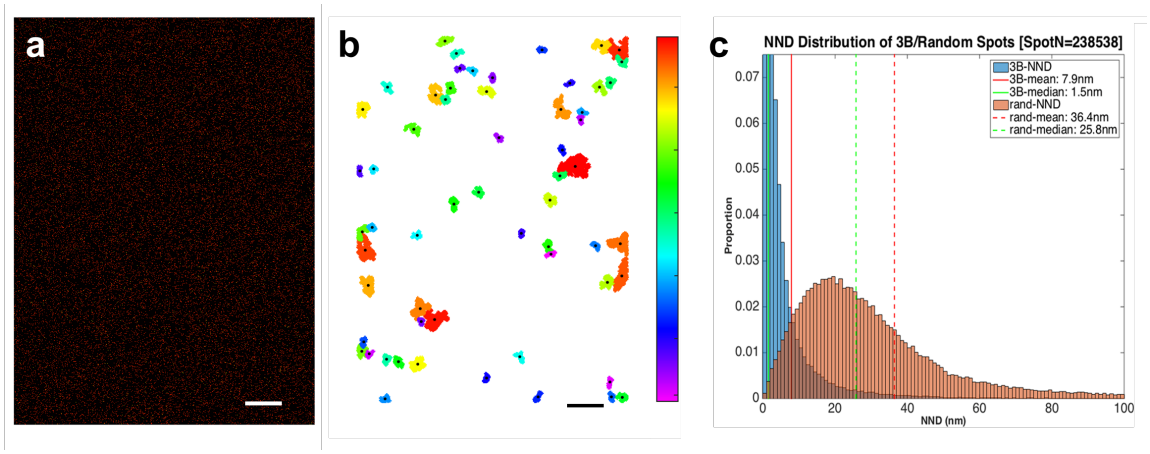

Figure S2. Comparison of cluster analysis of Pol II clusters in Figure 1b and simulated data generated by placing the same number of spots in Figure 1b randomly in the same area. (a) Randomly distributed spots with settings identical to that in Figure 1e. (b) DBSCAN-identified clusters of Figure S2a. (c) The nearest neighborhood distance (NND) distribution of 3B (Figure 1b) and random-distributed spots in Figure S2a. Scale bar is 2 μm.

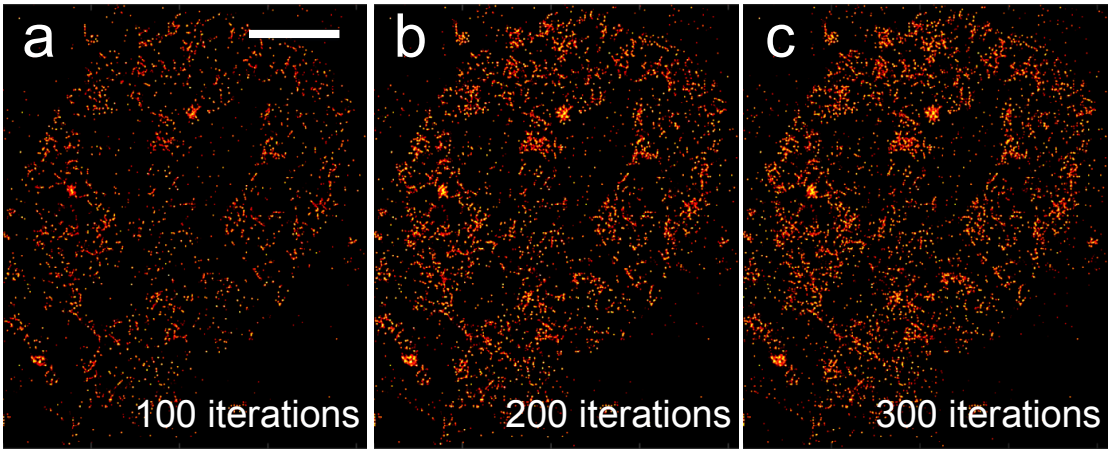

Figure S3. Live cell Bayesian super-resolution reconstruction with different iteration cycles. The results show that 200 iteration cycles are required for Bayesian reconstruction. Scale bar: 5 μm.

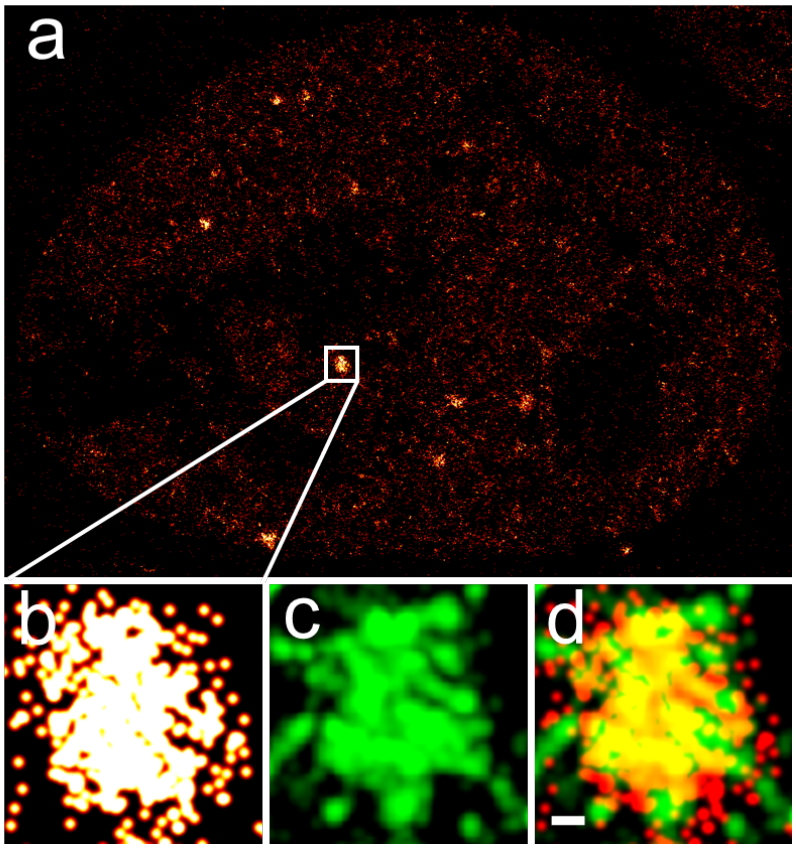

Figure S4. Live cell Bayesian nanoscopy and tcPALM of Pol II clustering. (a) tcPALM super-resolution image reconstructed using 10000 raw frames. (b-d) Zoom-in area of (a), Bayesian nanoscopy result (1500 raw frames) and merged image, respectively. The result shown in (d) shows evident colocalization between tcPALM result and Bayesian nanoscopy result. Scale bar: 100 nm.

Supplementary Note 1. Cluster analysis using DBSCAN

DBSCAN is a density-based algorithm for identifying clusters, which is of good efficiency for large spatial databases with noise and arbitrary shaped clusters.[1] To be specific: any spot shall be included into a cluster so long as the number of one spot's neighborhoods within one cluster exceeds some threshold, which ensures one cluster any arbitrary shape rather than some other methods that only take the gravity center of a cluster into consideration; any spot shall be eliminated so long as the number of the spot's neighborhoods within any cluster around couldn't exceed some threshold, which ensures an efficient de-noise by excluding those outliers.

Bayesian analysis of Blinking and Bleaching data (3B) method utilizes Bayesian model to reconstruct a probability-density map *via* throwing numerous single spots onto the map iteration by iteration,[2] where spots tend to be localized at those locations with more fluctuations. Yet the precision of localizing a 3B spot is limited due to the noisy background, whereby the final visually distinguishable resolution is ~50nm.

Based on those facts above, it is convenient to apply DBSCAN to the clustering analysis of Pol II with 3B database, considering its large databases, possible arbitrary shaped clusters and existence of outlying 3B spots. Besides (x,y)-coordinates of all the 3B spots derived from 3B analysis, another two inputs should be specified for DBSCAN here: neighborhood radius (Eps) and minimal number of objects considered as a cluster (k). We set Eps as 50nm, which is inspired by 3B's visual resolution. We set k as 1.5, *i.e.* any lonely spot (no neighbor 3B spot within Eps) shall be regarded as outliers (noise).

Supplementary Note 2. 3B Colormap generation

We linearly map each pixel's intensity (I) into the interval of [0,1] with the algorithm:

$$\mathrm{I}' = \frac{\mathrm{I} - \{\mathrm{I}\}_{\min}}{\{\mathrm{I}\}_{\max} - \{\mathrm{I}\}_{\min}}$$

and get the normalized density image (denoted as $\{\mathrm{I}'\}$). Then the normalized density image is given a false RGB colormap using the following formula:

$$\begin{aligned} r &= \min\left(1,\ 3\mathrm{I}'\right) \\ g &= \min\left(1,\ \max\left(0, 3\mathrm{I}' - 1\right)\right) \\ b &= \min\left(1,\ \max\left(0,\ 3\mathrm{I}' - 2\right)\right) \end{aligned}$$

where "r", "g" and "b" refer to red, green and blue, respectively.

Movie S1: the raw time-lapse movie that was used to construct Bayesian nanoscopy results of normal RNA Pol II shown in Figure 1b.

Supplementary References:

1. Ester, M.; Kriegel, H.-P.; Sander, J.; Xu, X. A Density-Based Algorithm for Discovering Clusters in Large Spatial Databases with Noise. In Proceedings of the Second International Conference on Knowledge Discovery and Data Mining (KDD-96); Simoudis, E., Han, J., Fayyad, U., Eds.; AAAI Press: Palo Alto, CA, 1996; pp 226−231.
2. Cox, S.; Rosten, E.; Monypenny, J.; Jovanovic-Talisman, T.; Burnette, D. T.; Lippincott-Schwartz, J.; Jones, G. E.; Heintzmann, R. Bayesian Localization Microscopy Reveals Nanoscale Podosome Dynamics. *Nat. Methods* **2012**, 9, 195-200.