

A Sponsored Supplement to *Science*

How deep learning is used within microscopy

Sponsored by



Produced by the *Science*/AAAS
Custom Publishing Office

Science



ScienceRobotics.org

DOESN'T YOUR RESEARCH DESERVE THE BEST READERS?

Submit your research: [cts.ScienceMag.org](https://cts.sciencemag.org)

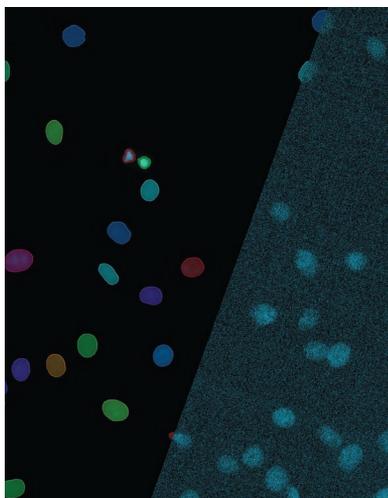
ScienceRobotics



 Twitter: @SciRobotics

 Facebook: @ScienceRobotics

Table of Contents



About the Cover:

DAPI (4', 6-diamidino-2-phenylindole, dihydrochloride) stained cell nuclei are purposefully underexposed to limit the specimen's exposure to near-UV light (right). Nikon's Enhance.ai deep learning analysis module is used to restore the signal-to-noise ratio to normally exposed DAPI staining, allowing for easy segmentation and counting (left panel shows Enhance.ai restoration).

This booklet was produced by the *Science*/AAAS Custom Publishing Office and sponsored by Nikon.

The primer and white paper that appear in this supplement have not been peer-reviewed nor have they been assessed by *Science*.

Editor: Jackie Oberst, Ph.D.
Proofreader/Copyeditor: Bob French
Designer: JD Huntsinger

STEPHANIE O'CONNOR,
SENIOR SALES MANAGER

Custom Publishing
Science | AAAS
soconnor@aaas.org
Tel: 202-336-9603

© 2020 by The American
Association for the Advancement
of Science.
All rights reserved.
16 October 2020

Introductions

2 Small images, big picture: Artificial intelligence in microscopy

Jackie Oberst
Science/AAAS

3 Image analysis: Bringing AI into focus

John Allen
Applications & Marketing Specialist
Nikon Instruments Inc.

Primer

4 Artificial intelligence for microscopy: What is deep learning and why is it making such a splash?

Florian Jug

Research Articles

5 Single-cell morphology encodes metastatic potential

Pei-Hsun Wu, Daniele M. Gilkes, Jude M. Phillip *et al.*

12 Coupling chromatin structure and dynamics by live super-resolution imaging

R. Barth, K. Bystricky, H. A. Shaban

23 Single yeast cell nanomotions correlate with cellular activity

Ronnie G. Willaert, Pieterjan Vanden Boer, Anton Malovichko *et al.*

White Paper

31 Utilizing deep learning tools to extend acquisition time and to design a label-free imaging analysis protocol

Michael A. Davis
Senior Manager, Products & Technology
Nikon Instruments Inc.



“The microscope—
a mainstay of
most research
laboratories—
has also become
self-learning.”

Small images, big picture: Artificial intelligence in microscopy

Artificial intelligence (AI) is more than just a buzzword—it is now a permanent fixture in everyday life and in scientific research. That voice you hear on the other end of your customer call, the ads you see on your computer screen after you conduct an internet search, the home security system that can tell the difference between you and an intruder—those are all a product of AI. And yet, despite its ubiquity, opinion surveys reveal that most people seem confused by the different terms used to describe AI and still do not know exactly what it is. Many confuse it with malevolent machines, a la Hollywood’s Skynet from *Terminator* or the evil computer HAL seen in *2001: A Space Odyssey*.

Part of the confusion stems from the fact that AI has many names that are often used interchangeably and incorrectly, terms such as “machine learning,” “neural networks,” and “deep learning.” AI had its beginnings in the 1950s, when Alan Turing speculated about “thinking machines” that could reason at the level of a human being. His famous “Turing Test,” or “imitation game,” specifies that in order for computers to pass as “thinking” in an autonomous manner, they would need to complete a series of reasoning puzzles. The actual term “AI” was soon coined by John McCarthy, who described it as “getting a computer to do things which, when done by people, are said to involve intelligence.” To this day, scientists continue to argue over the definition of AI, with a recent 2018 paper listing three qualities that characterize it: intentionality, intelligence, and adaptability.¹

Whether or not a consensus will one day be reached over this term, the reality is that AI is used in all forms of research, from diagnostics and drug discovery to imaging software, as is discussed in this supplement on AI in microscopy. The microscope—a mainstay of most research laboratories—has also become self-learning. While researchers do not have to become computer scientists to use these devices, a better understanding of AI’s capabilities can help them get faster and possibly better results.

Jackie Oberst, Ph.D.
Custom Publishing Office
Science/AAAS

Footnote

1. D.M. West, J.R. Allen. “How artificial intelligence is transforming the world.” <https://www.brookings.edu/research/how-artificial-intelligence-is-transforming-the-world/>



Image analysis: Bringing AI into focus

Artificial intelligence (AI) is experiencing a surge of renewed interest in recent years. AI technology, based on concepts such as machine learning and deep learning, has seemingly limitless applications in fields ranging from self-driving vehicle design, to cybersecurity, to (of course) scientific research and health care. This technology has also greatly affected how we approach optical microscopy imaging and analysis for biological research. However, given that the mainstream applications for new AI methods are still building momentum, it may not be clear to researchers the types of scenarios that are well suited to AI-based solutions. Our goal in this supplement is to help inform readers of the possibilities presented by deep learning-based imaging and analysis methods for microscopy.

This e-book includes an educational primer on deep learning in microscopy by Florian Jug, a group leader at the MPI-CBG Center for Systems Biology in Dresden, Germany, and expert in quantitative bioimaging methods utilizing deep learning. Also included are a number of recent research articles highlighting the use of machine learning or deep learning for image analysis.

Willaert et al. build upon their previous work showing that most cells exhibit nanoscale oscillations proportional in magnitude to cell activity. They take a close look at these oscillations in single yeast cells in response to various perturbations, using a deep learning-based cell detection algorithm that allows for fast and accurate identification of large numbers of cells, and explore the potential application of this method as an antifungal susceptibility test. Barth, Bystricky, and Shaban apply superresolution single-molecule localization microscopy (SMLM) to capture fine changes in chromatin conformation. Live-cell SMLM is a particularly difficult method requiring a careful balance between signal-to-noise ratio (SNR), emitter density, and other factors in order to obtain a sufficient number of high-quality localizations without sacrificing too much spatial or temporal resolution. To address these difficulties, they turned to the Deep-STORM (stochastic optical reconstruction microscopy) analysis method developed by Yoav Shechtman and colleagues, which utilizes a convolutional neural network to quickly create superresolved images from low-SNR and high-emitter density data. Finally, Wu et al. explore the use of cell morphology as an indicator of metastatic potential in a breast cancer cell line. Using an unsupervised clustering analysis of cell morphology parameters, they were able to define seven cell morphology classes in a population of over 30,000 cells and correlate them with well-defined gene expression classes, as well as tumorigenicity and metastatic potential. Their work offers an economical approach for evaluating tumor cell heterogeneity and identifying lethal cell subtypes.

To bring the power of AI-based image analysis to our customers, Nikon has developed a number of "NIS.ai" deep learning-based software modules for various applications. Several options are available for image improvement via approaches such as denoising and out-of-focus blur removal. Analysis capabilities include image segmentation, prediction of staining features in one channel from features in a different channel [e.g., prediction of DNA staining by DAPI (4', 6-diamidino-2-phenylindole, dihydrochloride) from DIC (differential interference contrast) or phase-contrast images], and more. Nikon is committed to the continuing development, testing, and support of reliable and user-friendly tools utilizing deep learning. The utility of NIS.ai modules for denoising, feature prediction, and image segmentation is explored in our included white paper.

We hope the content provided here provides insights into how AI is already being applied in various research applications, and inspires readers to think about how such methods can be used to further their own research.

John Allen

Applications & Marketing Specialist
Nikon Instruments Inc.

Artificial intelligence for microscopy: What is deep learning and why is it making such a splash?

Florian Jug, Ph.D., Max Planck Institute of Molecular Cell Biology and Genetics, Dresden, Germany

Let's start by disambiguating "artificial intelligence" (AI), "machine learning" (ML), and "deep learning" (DL). AI is a general term for the theory and development of computer systems able to perform tasks normally requiring human intelligence. ML is the subset of AI concerning algorithms and statistical models for performing a task without pre-programmed instructions. Hence, ML systems are designed to exploit training data to infer relevant correlations that are applied toward solving the learned task. DL is a sub-discipline of ML, where learning is performed by artificial neural networks (ANNs).

ANNs, described in Figure 1, draw inspiration from networks of neurons in our brain, with one artificial neuron (node) receiving inputs from many other nodes via weighted connections. An output is computed by summing the weighted inputs and feeding the result through a nonlinear activation function. Nodes are typically grouped into layers, and connections are generally made between neighboring layers. If many layers are sequentially connected, the network is called "deep." The first/last layers are the input/output layers, respectively. Training makes the output layer learn to predict the desired solution for a given input, which is achieved by adjusting all weighted connections in the ANN via "error back-propagation," a supervised approach requiring paired training data of inputs and corresponding (ground truth) outputs.

Researching DL for microscopy will inevitably introduce you to the U-Net (1), a fully convolutional autoencoder with skip connections that takes an input of a certain size and propagates it through several contracting layers of decreasing size. The "bottleneck" layer is smallest and is followed by a sequence of expansive layers that are successively upsampled back to the size of the input.

Contracting and expanding layers of equal size are connected via directed skip connections, ensuring re-localization of image features during upsampling. U-Net can be used to learn many tasks, and its fully convolutional nature enables it to be applied to variably sized input images.

Arguably the most widespread applications for U-Net are image segmentation and image restoration. While the original Content-Aware image REstoration (CARE) work needs pairs of low- and high-signal-to-noise ratio images for training (2), more recent work requires only single copies of noisy images (3). Other methods cover applications from point spread function design for improved 3D superresolution microscopy (4) to label-free prediction (5).

What will the future bring? I believe that DL and ML methods will continue to improve microscopy-based acquisition and analysis workflows. As a research community we must develop a common understanding of proper application of these powerful tools. They must not be black boxes (i.e., impenetrable), and common infrastructures are needed for model and data storage to ensure their practical benefit for researchers both with and without computational backgrounds.

1. O. Ronneberger, P. Fischer, T. Brox, *Lect. Notes Comput. Sci.* **9351**, 234–241 (2015).
2. M. Weigert *et al.*, *Nat. Methods* **15**, 1090–1097 (2018).
3. A. Krull, T.-O. Buchholz, F. Jug, *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit.*, Long Beach, CA, USA, 2124–2132 (2019).
4. E. Nehme *et al.*, *Nat. Methods* **17**, 734–740 (2020).
5. C. Ounkomol, S. Seshamani, M. M. Maleckar, F. Collman, G. R. Johnson, *Nat. Methods* **15**, 917–920 (2018).

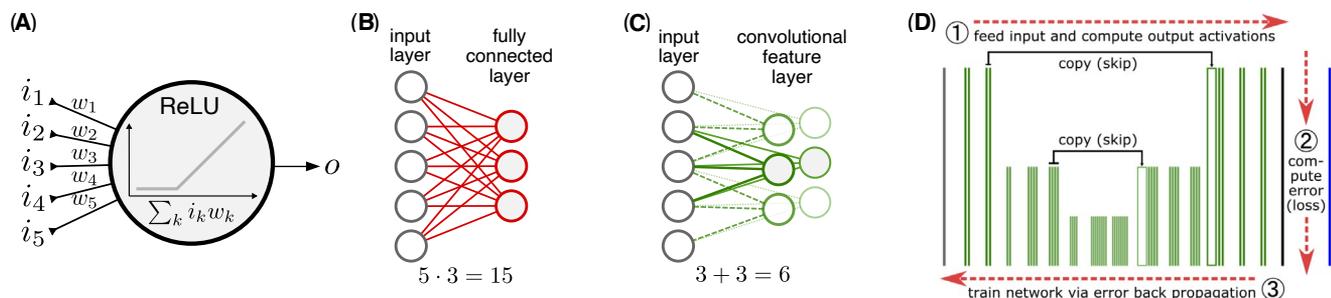


Figure 1: Artificial neural networks, summarized. (A) A node receives weighted inputs and feeds their sum through a nonlinear activation function to calculate output. While fully connected layers (B) require many connections, convolutional layers (C) only connect to nearby input nodes and reuse the same weights for each node in one feature layer. (Total number of weights computed below sketches.) (D) One training step consists of feeding input and computing network activations, computing the error (loss) compared to ground truth (blue line), then modifying all network weights via error back propagation.

CANCER

Single-cell morphology encodes metastatic potential

Pei-Hsun Wu^{1,2*}, Daniele M. Gilkes^{1,3}, Jude M. Phillip², Akshay Narkar⁴, Thomas Wen-Tao Cheng^{2,5}, Jorge Marchand², Meng-Hong Lee^{1,2}, Rong Li^{2,3,4,6}, Denis Wirtz^{1,2,3*}

A central goal of precision medicine is to predict disease outcomes and design treatments based on multidimensional information from afflicted cells and tissues. Cell morphology is an emergent readout of the molecular underpinnings of a cell's functions and, thus, can be used as a method to define the functional state of an individual cell. We measured 216 features derived from cell and nucleus morphology for more than 30,000 breast cancer cells. We find that single cell-derived clones (SCCs) established from the same parental cells exhibit distinct and heritable morphological traits associated with genomic (ploidy) and transcriptomic phenotypes. Using unsupervised clustering analysis, we find that the morphological classes of SCCs predict distinct tumorigenic and metastatic potentials in vivo using multiple mouse models of breast cancer. These findings lay the groundwork for using quantitative morpho-profiling in vitro as a potentially convenient and economical method for phenotyping function in cancer in vivo.

Copyright © 2020
The Authors, some
rights reserved;
exclusive licensee
American Association
for the Advancement
of Science. No claim to
original U.S. Government
Works. Distributed
under a Creative
Commons Attribution
NonCommercial
License 4.0 (CC BY-NC).

INTRODUCTION

Much effort is being made to explore the predictive power of genomic alterations in the detection and prognosis of diseases (1–4). However, a high degree of genomic instability in advanced cancers with metastatic disease endows these genomes with a myriad of abnormalities affecting the expression and function of tens of thousands of genes (5, 6). Recent studies show that individual (clonal) cells can display a broad landscape of properties, such as different gene expression patterns (7) and invasive behaviors (8), further increasing the challenge of deciphering the molecular basis of metastasis in cancer. A potential solution to this problem is to use a surrogate readout of a combinatorial set of genomic alterations that lead to similar outcomes. Previous studies using RNAi screens have shown that cell morphology (CM) can be an informative readout that is highly associated with molecular underpinnings (9, 10). Furthermore, recent studies indicate that the morphological status of cells can be linked to fundamental physiological properties of cells, such as cell cycle progression (11), cell-matrix adhesion properties, responsiveness to drug (8–10), aging (12), gene expression patterns (7), and invasiveness potential (8). To this end, CM in a defined environment is an emergent, yet relatively easily measurable, outcome resulting from the coupling between a cell's biochemistry and its biophysics that are ultimately encoded by the cell genome.

RESULTS

CM is a highly heritable trait at the single-cell level

We used a long-term, time-lapse recording of MDA-MB-231 human breast cancer cells growing in vitro, which readily suggested a high degree of cellular heterogeneity, including large variations in cell motility, cell size, and CM (movie S1). To determine whether the phenotypic traits presented by individual cells were stochastic or persistent, we used an ultralow-density growth assay to assess the morphology

of individual cells in colonies. Cells were morphologically similar to other cells in the same colony but distinct from cells in other colonies. The morphological traits of an individual cell persisted over extended periods of time (>1 month in culture). This observation suggested that morphological traits of individual parental cells were passed on to their progeny either by inheritance or by sharing a similar local microenvironment (Fig. 1A). A similar phenomenon was observed with six cancer cell lines derived from primary pancreatic tumors and metastases (fig. S1).

To further investigate the clonal architecture of CM, we generated single-cell clones (SCCs) obtained through the expansion of individual parental MDA-MB-231 breast cancer cells. Cells in each SCC displayed a distinct morphology (Fig. 1, B and C, and fig. S2) (7). To quantitatively describe the morphological spectrum of SCCs, we measured the morphology of cells in 14 SCCs and the parental cell line using a previously developed high-throughput microscopy and analysis system (11, 13–18). Briefly, cells and their nuclei were fluorescently stained and imaged using widefield fluorescence microscopy. For each well, a ~6 mm by 6 mm field of view was imaged and reconstructed from 81 (9 by 9) image tiles collected with a 10× objective. The morphology of cells was then automatically measured using a custom software (see details in Materials and Methods).

It has been previously shown that using a limited number of representative cell shapes is an effective strategy to explore complex CM datasets (9, 10). Here, we found that the morphology of cells in SCCs was categorized into seven CM classes (denoted A to G), which were themselves derived from a clustering analysis based on morphological features describing all >30,000 cells analyzed (Fig. 1D and Materials and Methods). This analysis provides visual and quantitative representations of CM across SCCs by assessing the distributions of these seven CM classes (e.g., Fig. 1E). These CM classes are associated with distinct properties of traditional CM parameters such as size, shape factor, and aspect ratio of cells and nuclei (fig. S3A).

On the basis of unsupervised hierarchical clustering of the CM distributions of the 14 SCCs, we classified the SCCs into six distinct morpho-types (M1 to M6) (Fig. 1F). All SCCs showed a certain degree of morphological heterogeneity as measured by Shannon's entropy of the morpho-types. The parental cells had a substantially higher level of morphological heterogeneity compared with SCCs (fig. S3, B and C). The global CM distribution obtained by ensemble-averaging the CM distributions of the 14 SCCs (denoted here <SCC>)

¹Johns Hopkins Physical Sciences–Oncology Center, The Johns Hopkins University, Baltimore, MD 21218, USA. ²Department of Chemical and Biomolecular Engineering, The Johns Hopkins University, Baltimore, MD 21218, USA. ³Department of Oncology, The Johns Hopkins University School of Medicine, Baltimore, MD 21231, USA. ⁴Department of Cell Biology, The Johns Hopkins University School of Medicine, Baltimore, MD 21231, USA. ⁵Division of Vascular and Endovascular Surgery, Boston Medical Center, Boston University School of Medicine, Boston, MA 02118, USA. ⁶Mechanobiology Institute National University of Singapore, 5A Engineering Drive 1, Singapore 117411, Singapore. *Corresponding author. Email: pww@jhu.edu (P.-H.W.); wirtz@jhu.edu (D.W.)

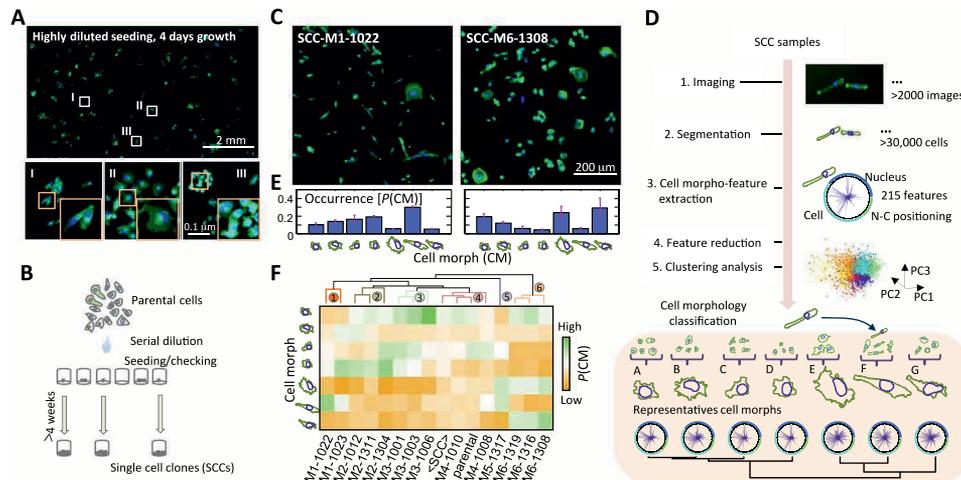


Fig. 1. Cell polymorphism—cell morphology is a highly heritable trait at the single-cell level. (A) Nuclei (blue) and F-actin organization (green) of MDA-MB-231 breast cancer cells after growth for 4 days from a sparse initial seeding density, showing how cells formed several spatially and morphologically distinct clusters. Representative high-resolution images of different clonal cells highlighted as I, II, and III are shown at the bottom. (B) Schematic plot showing the serial dilution procedure used to establish single-cell clones (SCCs) from a parental cell population. (C) Nuclei (blue) and F-actin organization (green) in cells of two established SCCs, SCC-M1-1022 and SCC-M6-1308, displaying distinct morpho-types. (D) Flow diagram illustrating the process used to quantify cell morphology (CM) through an unsupervised machine learning approach. A classifier model was built on the basis of all 14 SCCs and the parental MDA-MB-231 cells through principal component analysis and *k*-means clustering analysis. The morphology of all measured cells was classified into one of seven cell morph classes. Representative CM for each cell morph class (A to G) is shown at the bottom. (E) The fraction of cells in each cell morph class was used to quantitatively represent morpho-types of SCCs. Cell morph class fraction profiles for SCC-M1-1022 and SCC-M6-1308 are shown in the histograms. (F) Unsupervised hierarchical clustering of the SCCs based on their morpho-types (i.e., fraction of cells in cell morph classes A to G). The names of established SCCs were further marked as M1 to M6 based on six distinct cell morpho-type clusters revealed in the dendrogram.

was approximately the same as the CM distribution of the parental cells (Fig. 1F). Furthermore, SCCs displayed morphologies similar to colonies in the ultralow-density growth assay (Fig. 1, A and C).

Together, these results suggest that the parental MDA-MB-231 breast cancer cell population is composed of distinct, coexisting classes of cells with heritable morphological traits that persist over long time scales.

Morphological phenotypes in vitro and differential tumor progression in vivo

Individual MDA-MB-231 breast cancer cells can show distinct behavior in vivo, including differential ability to disseminate from the primary tumor and differential organ dissemination (8). To determine whether different cell morphologies of SCCs derived from MDA-MB-231 breast cancer cells corresponded to distinct outcomes in vivo, SCCs were injected into the mammary fat pad of mice, which were monitored for 50 days. We found that SCCs displayed a broad range of tumorigenicity (as measured by the weight and size of the primary tumors) and metastatic potential (as measured by human DNA content in the lungs) (Fig. 2A and fig. S4). For instance, the subclone SCC-M5-1317 formed tumors that were 50% larger than tumors produced by the parental MDA-MB-231 cells, but had a 99% decrease in its ability to form metastasis compared with parental cells (fig. S4). In contrast, SCC-M6-1308 and SCC-M6-1319 formed tumors of sizes similar to those produced by parental cells, but displayed 5 to 10 times more effective metastasis than parental cells. We also identified a group of SCCs (e.g., SCC-M2-1012) that formed small tumors that did not metastasize. Implanted SCCs also produced substantially different numbers of circulating tumor cells (CTCs) in the blood (Fig. 2B). Analysis of histological sections of the mouse lungs showed multiple metastatic lesions in mice bearing tumors formed by SCC-M6-1308 and SCC-M6-1319, but no metastatic

lesions for SCC-M5-1317 and SCC-M2-1012 (Fig. 2C and fig. S4). Short tandem repeat (STR) analysis showed that SCCs with distinct metastatic potential (SCC-M1-1022, SCC-M6-1308, and parental cells) had the exact same STR profiles, confirming their common ancestral origin.

We classified the SCCs into four grades of aggressiveness based on their tumorigenicity and metastatic potentials: (i) low tumorigenicity (LT), (ii) tumorigenic (T), (iii) metastatic (M), and (iv) hypermetastatic (HM) (summary of information about SCCs is given in table S1). We found only a weak correlation between tumor size and lung metastasis (Pearson’s correlation coefficient $\gamma = 0.32$) (Fig. 2A). This is consistent with the fact that SCCs that were highly tumorigenic could be either metastatic or not metastatic. In contrast, the number of CTCs per volume of blood was highly correlated with lung metastasis ($\gamma = 0.97$) (Fig. 2B), but poorly correlated with tumor size (fig. S4). The high correlation between the number of CTCs and metastatic burden in the lungs indicates that SCCs that can extravasate to blood vessels have higher potential for lung metastasis. Nevertheless, that these CTCs could come from lung metastases as well cannot be ruled out (19). We note that SCCs enriched with cells displaying a spindle-like morphology (i.e., cell shape with a high aspect ratio) are not the ones showing a high metastatic potential (fig. S4H).

Together, these results indicate that SCCs with the same morpho-types displayed similar in vivo outcomes, including tumorigenicity, tumor cells in circulation, and metastatic potential (Fig. 2A).

The morphological diversity of SCC correlates with distinct gene expression patterns

We next determined whether distinct morpho-types and corresponding tumorigenicity and metastatic potential were associated with distinct gene expression patterns. Transcriptomic microarray analysis showed that the gene expression profiles of SCCs at approximately

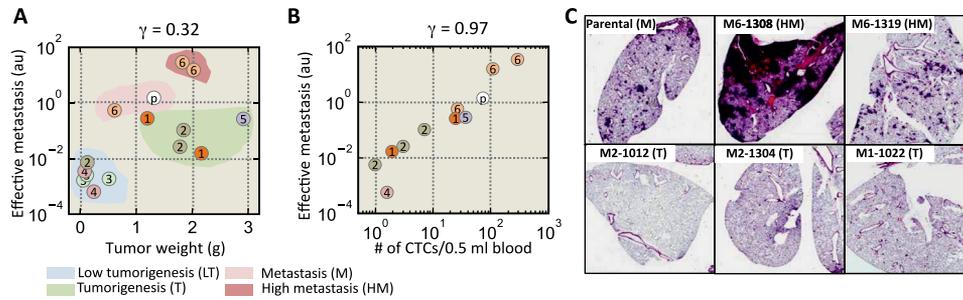


Fig. 2. Morphological phenotypes in vitro and differential tumor progression in vivo. (A and B) Scatter plot showing both tumor size and the extent of lung metastasis resulting from the injection of 14 SCCs and parental MDA-MB-231 cells into the mammary pad of SCID mice. At least four mice were tested for each SCC (A). The number within each circle represents the morpho-type class of the corresponding SCC. On the basis of tumorigenicity and metastatic burden in the lung, these SCCs were further classified into four groups: low tumorigenicity (LT), tumorigenic (T), metastatic (M), and hypermetastatic (HM). The Pearson's correlation coefficient between the effective metastasis and tumor weight among all SCCs is 0.32. The number of circulating tumor cells (CTCs) is highly correlated with lung metastasis, with a correlation coefficient of 0.96 (B). (C) Histological sections of mice lung show that clear metastatic lesions are present for SCC-M6-1308, SCC-M6-1319, and parental cells, but not in other SCCs, including SCC-M2-1012, SCC-M2-1304, and SCC-M2-1022. au, arbitrary units.

the same passage number were strongly associated with specific SCC morpho-types (Fig. 3A). SCCs belonging to the same morpho-type were located in close proximity of each other in the gene expression space, spanned by the first two principal components PC1 and PC2. Similar results were obtained using an unsupervised hierarchy clustering analysis based on expression data from genes that displayed the highest expression variations—measured by standard deviations—among SCCs (200 gene probes with 144 unique genes; see full list in table S2).

Overall, we found that there were four distinct gene expression subtypes among SCCs and the parental breast cancer cells (Fig. 3B). In the list of 200 gene candidates that were potentially responsible for the cell polymorphism, the SPANX family (*SPANXB2* and *SPANXE*)—cancer-testis antigens that are often highly expressed in tumor cells—featured the most variation, with approximately 1000-fold difference between SCCs with the lowest and highest levels of expression. A recent study has shown that the expression of members of the SPANX gene family promotes breast cancer invasion (20). Several genes in this list have been previously associated with patient survival and cancer metastasis, including *CDH11* (21), *KISS1* (22), *MAGEA3* (12, 23), *MAGEC1* (24), *TNFSF10* (25), *CXCR4* (26), and *GDF15* (18).

Mutual correlations between morpho-types, gene expression classes, and aggressiveness further confirmed that gene expression profiles (Fig. 3) and aggressiveness in vivo (Fig. 2, A to C) are reflected by the morphology of SCCs (Fig. 3C). This analysis shows that cells of LT are particularly small (enriched in cell morphs A and C); enrichment in elongated cells (cell morph F) was found only in groups of tumorigenic—but not metastatic—SCCs. The shapes of cells that were exclusively metastatic exhibited enriched cell morphs E and G: Their morphology tended to be rounder and larger (enriched in cell morph E) (Fig. 3C).

Distinct gene expression profiles of SCCs reveal prognostic genes

We further evaluated the genes that were differentially expressed among functionally distinct SCCs. A total of 218 genes (table S3) were either significantly down-regulated or significantly up-regulated [>5 -fold and P value from one-way analysis of variance (ANOVA) <0.05] when comparing SCCs of different tumorigenicity and metastatic potential (Fig. 4A). Among these 218 genes, 189 genes (87%) were associated with the comparison of LT and M' tumors, in contrast to 38 genes that were associated with the comparison of T and M' tumors (Fig. 4A).

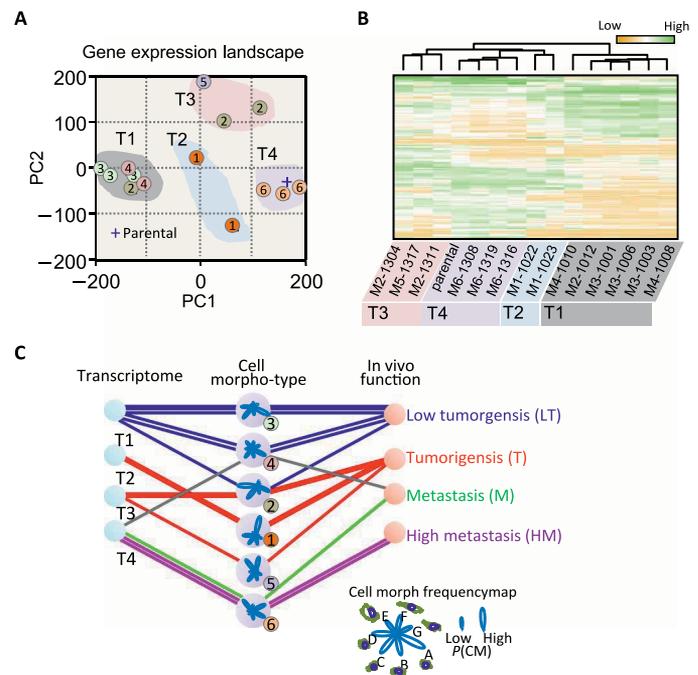


Fig. 3. Morphological diversity of SCCs is driven by distinct gene expression patterns. (A) The first and second principal components obtained from the principal component analysis of gene expression data were used to show the landscape of whole genome expression profile of the SCCs. The number within each circle represents the morpho-type class of each SCC. SCCs with the same morpho-type classes in general clustered together. (B) Unsupervised hierarchy clustering analysis using differentially expressed genes among these SCCs (see detailed list of genes in table S2) shows four distinct gene expression classes (G1 to G4). SCCs within the same morpho-type class are classified within the same gene expression class with the exception of SCC-M2-1012. SCCs within G1 and G3 gene expression classes exhibit multiple morpho-type classes. (C) Diagram showing mutual relations between morpho-type, gene expression class, and outcomes in vivo for different SCCs. Polar-petal plots were used to visualize fraction profiles of cell morph classes for the six different morpho-types. The length of a petal indicates the fraction size for the corresponding CM class.

This indicates that at the transcriptomic level, SCCs of LT were more different from metastatic SCCs (M') than tumorigenic SCCs. Of 38 genes that were differentially regulated between T and M', 28 (74%)

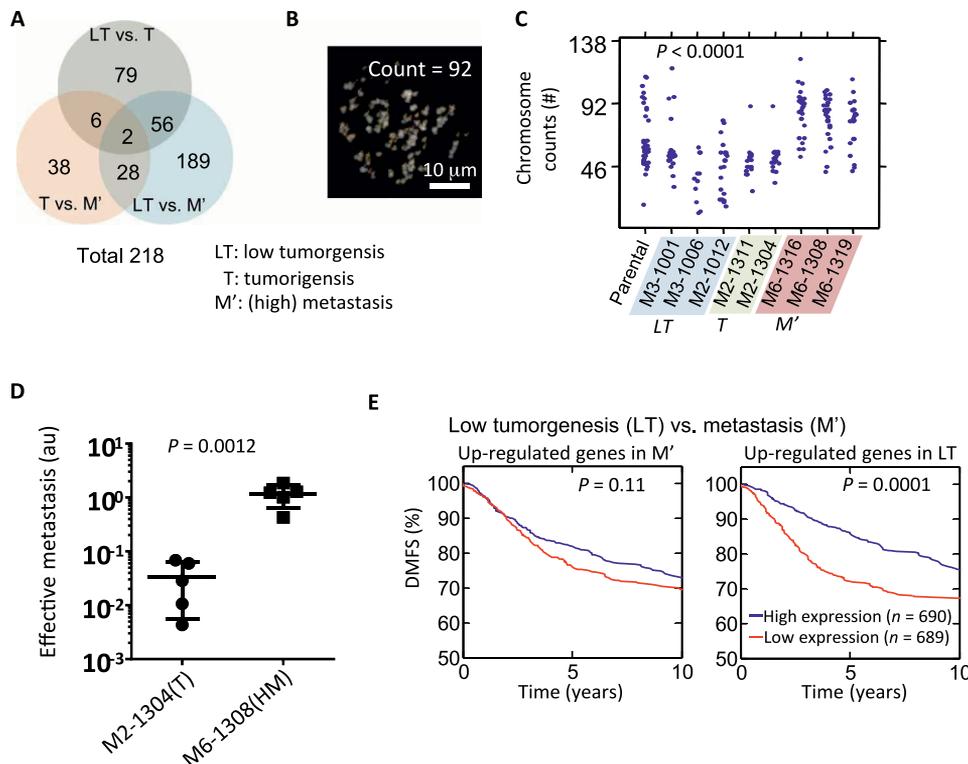


Fig. 4. Distinct gene expression profiles of SCCs reveal prognostic genes. (A) Venn diagram showing the number of genes that are found to be significantly different (>5-fold and P value from one-way ANOVA < 0.05) between three different in vivo grades of aggressiveness for SCCs (i.e., LT versus T, T versus M', and LT versus M'). M' includes both M and HM. (B and C) Representative image showing 4',6-diamidino-2-phenylindole (DAPI)-stained spreading chromosome of SCC-M6-1308 (B). Chromosome number counted using the metaphase spreading assay for parental cells ($n = 44$), and cells from SCC-M3-1001 ($n = 24$), SCC-M3-1006 ($n = 11$), SCC-M2-1012 ($n = 22$), SCC-M2-1311 ($n = 18$), SCC-M2-1304 ($n = 18$), SCC-M6-1316 ($n = 26$), SCC-M6-1308 ($n = 31$), and SCC-M6-1319 ($n = 22$). One-way ANOVA test shows there is a significant difference, with a $P < 0.0001$ (C). (D) Score for effective metastasis to the lung in the tail-vein injection mouse model ($n = 5$) shows significant difference ($P = 0.0012$) by Student t test between tumorigenic clone SCC-M2-1304 (mean lung effective metastasis score, 0.034) and metastatic clone SCC-M6-1308 (1.159). (E) Differentially expressed genes between LT SCC versus M' SCC were used to investigate their prognostic power. A cohort of 1379 tumors from patients with breast cancer was used to test the predictive potential of identified gene sets. Patients were separated into two groups based on the average expression level of these identified genes, and the Kaplan-Meier survival curves for the two groups of patients were plotted. For the genes that were up-regulated in the M' SCCs, no significant prognostic effect was found. However, the results show that patients with higher expression levels of metastasis suppressor genes (i.e., up-regulated genes in LT) have a significantly longer survival time than those with low expression ($P = 0.0001$). P value is evaluated using log-rank test.

also could differentiate LT from T tumors, suggesting that tumorigenic (T) SCCs represent an intermediate transcriptomic state between LT SCCs and M' SCCs.

To explore the relation between morpho-types and gene expression of SCCs, we cross-compared the list of 218 genes with the list of 883 genes functionally annotated as “regulation of cell shape” (GO:0000902) by gene ontology (27–29). We found that 22 genes co-occurred in both lists, which corresponds to a P value of 2.42×10^{-4} (table S4). Since our findings show that morpho-types of SCCs are highly associated with their functions in vivo, this result strongly suggests that the morphological heterogeneity of SCCs is a result of differential expressed genes, such as interleukin-6 (IL-6), IL-7R, etc., which may play a role in tumor progression (30).

We further investigated the potential mechanisms by which morpho-type M6 may encode metastatic potential in vivo. M6 is characterized by cells displaying large nuclei, which may be indicative of an increase in ploidy (31, 32). To test this hypothesis, we analyzed the degree of ploidy of SCCs. Measurement of the distribution of the number of chromosomes for each SCC, using the metaphase-spread assay, showed that SCCs with high metastatic potential (M' SCC)

displayed a substantial higher average number of chromosomes (77 to 86) than LT SCCs and T SCCs (50 to 59) and parental MDA-MB-231 cells (~59) (Fig. 4C). This is consistent with experimental and clinical evidence that suggests that tetraploidization is a frequent genomic abnormality associated with enhanced metastasis, possibly due to a high rate of aneuploidy production in subsequent cell divisions and/or better tolerance of aneuploidy due to higher basal ploidy (33–36). All SCCs exhibited wide distributions in chromosome numbers (Fig. 4B), suggesting that high chromosome instability is inherent in all cells derived from the parental breast cancer cell population.

Another morphological characteristic of the morpho-type M6 is the highly symmetric shape of the cells compared with the much more spindle-like morphology of the other morpho-types on two-dimensional (2D) glass surfaces. As cell shape is often a readout of the production of cortical cytoskeletal forces (17), we compared the motility of different SCCs in 3D matrices of controlled collagen I content (37, 38). For all motility models tested, including cells on 2D substrates of controlled stiffness and cells embedded in 3D collagen matrices (39, 40), we did not find a correlation between motility and metastatic potential (fig. S5). The proliferation rate of SCCs in culture showed that the

highly metastatic SCC clones and parental MDA-MB-231 cells had the highest population growth rates among SCCs (fig. S5B).

Direct injection of tumorigenic and metastatic SCCs into blood vessels through the tail vein showed that metastatic SCCs (SCC-M6-1308) could more effectively extravasate and colonize the lung, suggesting that the metastatic potential of the SCCs is determined by their lung-seeding capacity (Fig. 4D and fig. S6).

Last, we compared the distant metastasis-free survival (DMFS) of patients with breast cancer, which was stratified by expression levels using a cohort of 1379 tumors obtained from patients with breast cancer (GOBO database, Gene expression-based outcome for breast cancer) (41). Of the 218 genes identified by our SCC classification, we found 155 genes in this cohort (see summary in table S3). Kaplan-Meier survival analysis showed that patients with tumors showing a higher level of expression of tumor-suppressor genes—the genes that were up-regulated when comparing LT and M—had a significantly improved DMFS ($P = 0.0001$) than patients with tumors showing low expression of these genes (Fig. 4D). Consistent with this result, tumors of patients with a high expression of genes that were up-regulated in T or LT tumors in comparison to M' or T tumors had significantly longer DMFS (fig. S7), confirming the role of cancer cell polymorphism in tumor evolution and progression.

DISCUSSION

Together, our analysis of SSCs derived from a parental breast cancer cell line demonstrates clonally persistent morphological heterogeneity. These SSCs show a wide range of distinct tumorigenic and metastatic potentials *in vivo*. The progression and outcomes of SSC-derived cancers in mice are associated with distinct patterns of gene expression. The same genes that are differentially regulated when comparing metastatic to nonmetastatic SCCs are of prognostic value to assess metastasis-free patient survival. These results support our hypothesis that CM is a holistic readout (in physics, CM would be called an “emergent” property) of the complex genomic and gene expression changes in cancer cells. The morphological features that predict metastatic potential are associated with increased ploidy. High basal ploidy provides better tolerance of diverse aneuploid karyotypes, which produce the phenotypic variation driving adaption of metastatic tumors to novel microenvironments.

We anticipate that incorporating single-cell analysis of intratumoral heterogeneity could further improve diagnosis and prognosis for individual patients and that quantitative cell phenotyping analysis *in vitro* could offer an effective and economical method to decipher complex cellular heterogeneity in tumors to identify lethal cancer cell subtypes for diagnostic and therapeutic purposes.

MATERIALS AND METHODS

Cell lines and culture

The parental breast cancer cell line MDA-MB-231 (42) and derived SCCs were maintained in high-glucose (4.5 mg/ml) Dulbecco's minimum essential medium supplemented with 10% fetal bovine serum and 1% penicillin-streptomycin. Cells were maintained at 37°C in a 5% CO₂, 95% air incubator.

Establishment of SCCs

A suspension of parental MDA-MB-231 cells was diluted using culture medium to a cell density of approximately 1 cell/0.1 μ l. A droplet of

0.1 μ l of cell suspension was placed in each well of a 96-well plate by pipetting followed by microscopy inspection to examine the number of cells in the deposited droplet. For wells containing a single cell, 200 μ l of culture medium was subsequently added to allow for cell growth into SCCs. The culture medium was then replaced regularly every 3 to 4 days, and SCCs were subsequently transferred to 24-well plates, 6-well plates, and 10-cm petri dishes after they became confluent. SCCs were then frozen down and thawed for further experiments.

Orthotopic implantation and metastasis assays

Studies using 7- to 10-week-old female severe combined immunodeficient (SCID) mice [National Cancer Institute (NCI)] were performed according to protocols approved by the Johns Hopkins University Animal Care and Use Committee. Briefly, 2×10^6 cells were resuspended in 1:1 ratio of phosphate-buffered saline (PBS) to Matrigel (BD Biosciences) and injected into the second left mammary fat pad. Tumor growth was monitored by caliper measurements. Tumor volume (cubic millimeter) was calculated as length by width by depth by 0.52. After indicated times, mice were sacrificed, and the lungs were perfused with PBS. The left lung was inflated by injecting with low-melting point agarose. Uninflated lungs were used for human genomic DNA extraction. Lungs were digested with lysis buffer and proteinase K at 55°C overnight, and genomic DNA was isolated by phenol/chloroform extraction and isopropanol precipitation. Genomic DNA (200 ng) was used for quantitative polymerase chain reaction (qPCR) to quantify human HK2 and mouse 18S transcripts.

To count CTCs, 500 μ l of blood from each mouse was collected. Red blood cells were lysed using ammonium chloride solution (Stem Cell Technologies, catalog no. 07800). RNA from the remaining cells in the blood was extracted (Life Technologies, catalog no. 15596-026) and reverse transcribed to complementary DNA (cDNA; Bio-Rad iScript Reverse Transcriptase, catalog no. 170-8840). The cDNA was then used for qPCR to quantify human-only 18S rRNA and mouse and human 18S rRNA. In each sample, we measured normalized human 18S gene expression by $2^{-C_{\text{sample}}}$ where $C_{\text{sample}} = C_{\text{hu}} - \text{average } C_{(\text{hu and mu } 18\text{S})}$. The calibration curve between measured 18S gene expression and the number of MDA-MB-231 cells was obtained by spiking controlled numbers of MDA-MB-231 cells in naïve mouse blood samples.

For the tail-vein injection model, MDA-MB-231 subclones were harvested by trypsinization, resuspended at 10^7 cells/ml in PBS, and injected (1×10^6 cells) intravenously into SCID mice. After 2 weeks, lungs were perfused with PBS. One lung was inflated for formalin fixation and paraffin embedding. The other lung was used to isolate genomic DNA for qPCR analysis with human-specific HK2 primers.

Immunostaining and fluorescence microscopy

Approximately 12,000 cells were plated in each well of a 24-well glass bottom plate (MatTek, MA), corresponding to approximately 20% surface coverage to ensure single-cell dispersion. After 16 hours of incubation, cells were fixed with 3.7% paraformaldehyde for 12 min at room temperature. Cells were then permeabilized with 0.1% Triton X-100 (Sigma-Aldrich) for 10 min; nonspecific binding was blocked with PBS supplemented with 1% albumin from bovine serum for 40 min. Nuclear DNA was stained with Hoechst 33342 (Sigma-Aldrich) at 1:50 dilution; F-actin was stained with phalloidin Alexa Fluor 488 (Invitrogen) at a 1:40 dilution. Fluorescently labeled cell samples were visualized with a Nikon digital sight DS-Qi1MC camera mounted on a Nikon TE300 epifluorescence microscope (Nikon Melville, NY)

and equipped with a motorized stage and motorized excitation and emission filters (Prior Scientific, Rockland, MA) controlled by NIS-Elements (Nikon). For each sample, 81 (9-by-9 square grid) fields of view from a low-magnification lens (10× Plan Fluor lens; numerical aperture, 0.3; Nikon) were used, which covered a contiguous area of 6.03 mm by 4.73 mm (28.5 mm²). The fluorescence channels for Hoechst 33342 and Alexa Fluor 488 were recorded to obtain the necessary morphometric information about the nucleus and cellular body of each individual cell within the scanning region.

Analysis of CM

Image processing for quantification of cellular morphological features from fluorescence images was carried out using a custom program developed in MATLAB (Mathworks, MA) (13–15, 43). In brief, we first segmented individual cells and their nuclei. We used five different categories of morphological features with a total number of 215 features to characterize nucleus and cell shapes. These features correspond to classes of morphological features that include basic morphological features, boundary signature, curvature, nucleus-cell positioning, and protrusion (fig. S8). The full list of features is summarized in table S5. In general, basic morphology features are features such as area, perimeter, long axis, short axis, and aspect ratio. Boundary signature of a shape (R) is the distance profile from all boundary coordinates to the centroid points of a shape, and boundary signature features are the statistical profiles of R , such as mean, median, and SD. To obtain curvature features, we first calculated the curvature (k) along the boundary of smoothed cell shapes. The smoothing of shape is processed by convolving the x and y coordinates of the shape with a 1D Gaussian filter, which has unit SD and size of 11 pixels. Statistical descriptors for the curvature along the boundary of shape, such as mean, median, SD, were measured as curvature features. The detailed list of statistical descriptors used can be found in the table S5. The same statistical properties used for boundary signature were extracted to represent the curvature features of a shape. Nucleus-cell positioning profile (R') is represented by the distance from the nucleus edge to the cell edges in different orientations based on the centroid of the nucleus. The nucleus-cell positioning features are a set of statistical properties of R' . For the quantification of protrusion morphology, we adopted a previous approach (15, 43). In brief, we first determined the morphological skeleton of individual cell contours and identified the main body region of the cells. The protrusions were identified as the skeletal structures that were extended beyond the main body of the cell. The protrusions were further classified into two subtypes: primary and secondary protrusions. The primary protrusions were considered to be the protrusions stemming directly from the cell body, while the secondary protrusions were the ones branching from other protrusions. The length of each protrusion was measured, and the total number of protrusions for individual cells was determined as the summation of primary and secondary protrusions. The total number of protrusions, mean length of protrusions, primary protrusion number, secondary protrusion number, and the ratio of secondary to primary protrusions were used as parts of the CM features. To quantitatively classify cell morpho-types, the morphology feature space of cells was first reduced and was represented by projection scores at 36 eigenvectors that spanned 95% of variations of among all measured cells from the principal component analysis (fig. S3). K -means clustering analysis with cityblock distance function was implemented to identify the seven distinct clusters among CM data of all measured cells.

Metaphase spreading assay

Cells were grown up to 60% confluency after plating. Colcemid (Invitrogen) was added to the cultural medium at a concentration of 100 ng/ml, and cells were incubated at 37°C for 3 to 4 hours. Cells were harvested using trypsin and resuspend in 1 ml of culture medium after spinning down. Five microliters of 37°C prewarmed KCl was added slowly to the cell suspension and incubated at room temperature for 7 to 10 min followed by adding 120 μ l of freshly prepared fixative solution (methanol:acetic acid in 3:1 volume ratio). Cells were incubated in 9.5 ml of fixative solution for 10 min after being spun down at 1000 rpm for 8 min and having discarded the supernatant. Cells were then resuspended in 0.3 ml of fixative and dropped on a glass slide before being placed onto slide warmer at 65°C for 20 min followed by treatment with RNase A (1 mg/ml; 1:100 from Qiagen) and propidium iodide (1 mg/ml stock and 1:1000 final) in 2× SSC for ~45 min at 37°C. Slides were air dried before mounting using mounting medium with DAPI (Vectashield). Chromosome spreads were imaged with 63× oil objective mounted on a Ti-E microscope (Nikon) and analyzed using previously established software (44, 45).

Microarray transcriptional profiling

Total RNA was isolated from MDA-MB-231 cells and its SCCs with the RNeasy Mini kit and analyzed using the Affymetrix GeneChip PrimeView Human Gene Expression Array (Johns Hopkins Deep Sequencing and Microarray Core Facility). Partek Genomic Suite was used to normalize expression data of all extended level probe sets using the following options: GC content prebackground adjustment, robust multi-array average (RMA) background correction, and quantile normalization. Gene expression level was defined as the average expression level of all exons for that gene. One-way ANOVA test was used to obtain P value and fold change (FC) values. The differentially expressed genes were detected for $P < 0.05$ and $|FC| > 5$ (linear).

Statistics

Data were represented as averages \pm SEM unless otherwise specified. One-way ANOVA test was performed to determine significance using MATLAB (MathWorks) unless otherwise specified.

SUPPLEMENTARY MATERIALS

Supplementary material for this article is available at <http://advances.sciencemag.org/cgi/content/full/6/4/eaaw6938/DC1>

Fig. S1. Cell polymorphism in cancer cells derived from primary pancreatic tumors and metastases.

Fig. S2. Morphological spectrum of SCCs.

Fig. S3. Quantitative morphology analysis of metastasis of breast cancer cells.

Fig. S4. SCCs derived from the same parental breast cancer cells show divergent invasive behavior in vivo.

Fig. S5. Analysis of SCC with migration and proliferation assay.

Fig. S6. Lung tissue sections in tail vein metastasis model for different SCCs.

Fig. S7. Distinct gene expression profiles of SCCs reveal prognostic genes.

Fig. S8. Morphological features.

Table S1. Summary of SCC information.

Table S2. Genes that are highly differentially expressed among SCP.

Table S3. Differentially expressed genes among LT versus T versus M.

Table S4. List of genes affiliated with regulation of cell shape.

Table S5. Summary of CM features.

Table S6. Whole-genome gene expression data of MDA-MB-231 SCCs.

[View/request a protocol for this paper from Bio-protocol.](#)

REFERENCES AND NOTES

1. M. R. Stratton, P. J. Campbell, P. A. Futreal, The cancer genome. *Nature* **458**, 719–724 (2009).
2. B. Vogelstein, N. Papadopoulos, V. E. Velculescu, S. Zhou, L. A. Diaz, K. W. Kinzler, Cancer genome landscapes. *Science* **339**, 1546–1558 (2013).

3. J. G. Paez, P. A. Jänne, J. C. Lee, S. Tracy, H. Greulich, S. Gabriel, P. Herman, F. J. Kaye, N. Lindeman, T. J. Boggon, K. Naoki, H. Sasaki, Y. Fujii, M. J. Eck, W. R. Sellers, B. E. Johnson, M. Meyerson, EGFR mutations in lung cancer: Correlation with clinical response to gefitinib therapy. *Science* **304**, 1497–1500 (2004).
4. Y. Miki, J. Swensen, D. Shattuck-Eidens, P. Futreal, K. Harshman, S. Tavtigian, Q. Liu, C. Cochran, L. Bennett, W. Ding *et al.*, A strong candidate for the breast and ovarian-cancer susceptibility gene *brca1*. *Science* **266**, 66–71 (1994).
5. S. Negrini, V. G. Gorgoulis, T. D. Halazonetis, Genomic instability - an evolving hallmark of cancer. *Nat. Rev. Mol. Cell Biol.* **11**, 220–228 (2010).
6. T. A. Potapova, J. Zhu, R. Li, Aneuploidy and chromosomal instability: A vicious cycle driving cellular evolution and cancer genome chaos. *Cancer Metast. Rev.* **32**, 377–389 (2013).
7. A. Nguyen, M. Yoshida, H. Goodarzi, S. F. Tavazoie, Highly variable cancer subpopulations that exhibit enhanced transcriptome variability and metastatic fitness. *Nat. Commun.* **7**, 11246 (2016).
8. A. J. Minn, Y. Kang, I. Serganova, G. P. Gupta, D. D. Giri, M. Doubrovin, V. Ponomarev, W. L. Gerald, R. Blasberg, J. Massagué, Distinct organ-specific metastatic potential of individual breast cancer cells and primary tumors. *J. Clin. Invest.* **115**, 44–55 (2005).
9. Z. Yin, A. Sadok, H. Sailem, A. McCarthy, X. Xia, F. Li, M. A. Garcia, L. Evans, A. R. Barr, N. Perrimon, C. J. Marshall, S. T. C. Wong, C. Bakal, A screen for morphological complexity identifies regulators of switch-like transitions between discrete cell shapes. *Nat. Cell Biol.* **15**, 860–871 (2013).
10. C. Bakal, J. Aach, G. Church, N. Perrimon, Quantitative morphological signatures define local signaling networks regulating cell morphology. *Science* **316**, 1753–1756 (2007).
11. P. H. Wu, J. M. Phillip, S. B. Khatau, W. C. Chen, J. Stirman, S. Rosseel, K. Tschudi, J. Van Patten, M. Wong, S. Gupta, A. S. Baras, J. T. Leek, A. Maitra, D. Wirtz, Evolution of cellular morpho-phenotypes in cancer metastasis. *Sci. Rep.* **5**, 18437 (2015).
12. M. Ayyoub, C. M. Scarlata, A. Hamai, P. Pignoni, D. Valmori, Expression of MAGE-A3/6 in primary breast cancer is associated with hormone receptor negative status, high histologic grade, and poor survival. *J. Immunother.* **37**, 73–76 (2014).
13. W. C. Chen, P. H. Wu, J. M. Phillip, S. B. Khatau, J. M. Choi, M. R. Dallas, K. Konstantopoulos, S. X. Sun, J. S. H. Lee, D. Hodzic, D. Wirtz, Functional interplay between the cell cycle and cell phenotypes. *Integr. Biol.* **5**, 523–534 (2013).
14. P. H. Wu, S. H. Hung, T. N. Ren, I. M. Shih, Y. Tseng, Cell cycle-dependent alteration in NAC1 nuclear body dynamics and morphology. *Phys. Biol.* **8**, 015005 (2011).
15. M. H. Lee, P. H. Wu, D. Gilkes, I. Aifuwa, D. Wirtz, Normal mammary epithelial cells promote carcinoma basement membrane invasion by inducing microtubule-rich protrusions. *Oncotarget* **6**, 32634–32645 (2015).
16. A. B. Chambliss, P. H. Wu, W. C. Chen, S. X. Sun, D. Wirtz, Simultaneously defining cell phenotypes, cell cycle, and chromatin modifications at single-cell resolution. *FASEB J.* **27**, 2667–2676 (2013).
17. J. M. Phillip, P. H. Wu, D. M. Gilkes, W. Williams, S. McGovern, J. Daya, J. Chen, I. Aifuwa, J. S. H. Lee, R. Fan, J. Walston, D. Wirtz, Biophysical and biomolecular determination of cellular age in humans. *Nat. Biomed. Eng.* **1**, 0093 (2017).
18. K. M. A. Yong, Y. Zeng, D. Vindivich, J. M. Phillip, P. H. Wu, D. Wirtz, R. H. Getzenberg, Morphological effects on expression of growth differentiation factor 15 (GDF15), a marker of metastasis. *J. Cell. Physiol.* **229**, 362–373 (2014).
19. A. J. Minn, G. P. Gupta, P. M. Siegel, P. D. Bos, W. Shu, D. D. Giri, A. Viale, A. B. Olshen, W. L. Gerald, J. Massagué, Genes that mediate breast cancer metastasis to lung. *Nature* **436**, 518–524 (2005).
20. E. A. Maine, J. M. Westcott, A. M. Precht, T. T. Dang, A. W. Whitehurst, G. W. Pearson, The cancer-testis antigens SPANX-A/C/D and CTAG2 promote breast cancer invasion. *Oncotarget* **7**, 14708–14726 (2016).
21. M. J. Pishvaian, C. M. Feltes, P. Thompson, M. J. Bussemakers, J. A. Schalken, S. W. Byers, Cadherin-11 is expressed in invasive breast cancer cell lines. *Cancer Res.* **59**, 947–952 (1999).
22. T. A. Martin, G. Watkins, W. G. Jiang, KISS-1 expression in human breast cancer. *Clin. Exp. Metast.* **22**, 503–511 (2005).
23. J. Kim, H. A. Reber, O. J. Hines, K. K. Kazanjian, A. Tran, X. Ye, F. F. Amersi, S. R. Martinez, S. M. Dry, A. J. Bilchik, D. S. B. Hoon, The clinical significance of MAGEA3 expression in pancreatic cancer. *Int. J. Cancer* **118**, 2269–2275 (2006).
24. S. Kruger, V. Ola, A. C. Feller, D. Fischer, M. Friedrich, Expression of cancer-testis antigen CT7 (MAGE-C1) in breast cancer: An immunohistochemical study with emphasis on prognostic utility. *Pathol. Oncol. Res.* **13**, 91–96 (2007).
25. K. Kuribayashi, G. Krigsfeld, W. Wang, J. Xu, P. A. Mayes, D. T. Dicker, G. S. Wu, W. S. El-Deiry, TNFSF10 (TRAIL), a p53 target gene that mediates p53-dependent cell death. *Cancer Biol. Ther.* **7**, 2034–2038 (2014).
26. A. Muller, B. Homey, H. Soto, N. Ge, D. Catron, M. E. Buchanan, T. McClanahan, E. Murphy, W. Yuan, S. N. Wagner, J. L. Barrera, A. Mohar, E. Verástegui, A. Zlotnik, Involvement of chemokine receptors in breast cancer metastasis. *Nature* **410**, 50–56 (2001).
27. M. Ashburner, C. A. Ball, J. A. Blake, D. Botstein, H. Butler, J. M. Cherry, A. P. Davis, K. Dolinski, S. S. Dwight, J. T. Eppig, M. A. Harris, D. P. Hill, L. Issel-Tarver, A. Kasarskis, S. Lewis, J. C. Matese, J. E. Richardson, M. Ringwald, G. M. Rubin, G. Sherlock, Gene Ontology: Tool for the unification of biology. *Nat. Genet.* **25**, 25–29 (2000).
28. Gene Ontology Consortium, Gene Ontology Consortium: Going forward. *Nucleic Acids Res.* **43**, D1049–D1056 (2015).
29. S. Carbon, A. Ireland, C. J. Mungall, S. Q. Shu, B. Marshall, S. Lewis; the AmiGO Hub; the Web Presence Working Group, AmiGO: Online access to ontology and annotation data. *Bioinformatics* **25**, 288–289 (2009).
30. H. Jayatilaka, P. Tyle, J. J. Chen, M. Kwak, J. Ju, H. J. Kim, J. S. H. Lee, P. H. Wu, D. M. Gilkes, R. Fan, D. Wirtz, Synergistic IL-6 and IL-8 paracrine signalling pathway infers a strategy to inhibit tumour cell migration. *Nat. Commun.* **8**, 15584 (2017).
31. T. Watanabe, Y. Tanaka, Age-related alterations in the size of human hepatocytes. A study of mononuclear and binucleate cells. *Virchows Arch. B. Cell Pathol. Incl. Mol. Pathol.* **39**, 9–20 (1982).
32. I. Conlon, M. Raff, Size control in animal development. *Cell* **96**, 235–244 (1999).
33. O. S. Frankfurt, H. K. Slocum, Y. M. Rustum, S. G. Arbuck, Z. P. Pavelic, N. Petrelli, R. P. Huben, E. J. Pontes, W. R. Greco, Flow cytometric analysis of DNA aneuploidy in primary and metastatic human solid tumors. *Cytometry* **5**, 71–80 (1984).
34. D. Korenaga, T. Okamura, A. Saito, H. Baba, K. Sugimachi, DNA Ploidy Is Closely linked to tumor invasion, lymph-node metastasis, and prognosis in clinical gastric-cancer. *Cancer* **62**, 309–313 (1988).
35. N. Navin, J. Kendall, J. Troge, P. Andrews, L. Rodgers, J. McIndoo, K. Cook, A. Stepansky, D. Levy, D. Esposito, L. Muthuswamy, A. Krasnitz, W. R. McCombie, J. Hicks, M. Wigler, Tumour evolution inferred by single-cell sequencing. *Nature* **472**, 90–94 (2011).
36. N. Pavelka, G. Rancati, R. Li, Dr Jekyll and Mr Hyde: Role of aneuploidy in cellular adaptation and cancer. *Curr. Opin. Cell Biol.* **22**, 809–815 (2010).
37. S. I. Fraley, Y. Feng, R. Krishnamurthy, D. H. Kim, A. Celedon, G. D. Longmore, D. Wirtz, A distinctive role for focal adhesion proteins in three-dimensional cell motility. *Nat. Cell Biol.* **12**, 598–604 (2010).
38. S. I. Fraley, P. H. Wu, L. He, Y. Feng, R. Krishnamurthy, G. D. Longmore, D. Wirtz, Three-dimensional matrix fiber alignment modulates cell migration and MT1-MMP utility by spatially and temporally directing protrusions. *Sci. Rep.* **5**, 14580 (2015).
39. P. H. Wu, A. Giri, S. X. Sun, D. Wirtz, Three-dimensional cell migration does not follow a random walk. *Proc. Natl. Acad. Sci. U.S.A.* **111**, 3949–3954 (2014).
40. P. H. Wu, A. Giri, D. Wirtz, Statistical analysis of cell migration in 3D using the anisotropic persistent random walk model. *Nat. Protoc.* **10**, 517–527 (2015).
41. M. Ringner, E. Fredlund, J. Hakkinen, A. Borg, J. Staaf, GBO: gene expression-based outcome for breast cancer online. *PLOS ONE* **6**, e17911 (2011).
42. M. H. Lee, P. H. Wu, J. R. Staunton, R. Ros, G. D. Longmore, D. Wirtz, Mismatch in mechanical and adhesive properties induces pulsating cancer cell migration in epithelial monolayer. *Biophys. J.* **102**, 2731–2741 (2012).
43. Q. Y. Guo, J. M. Phillip, S. Majumdar, P. H. Wu, J. Chen, X. Calderón-Colón, O. Schein, B. J. Smith, M. M. Trexler, D. Wirtz, J. H. Elisseeff, Modulation of keratocyte phenotype by collagen fibril nanoarchitecture in membranes for corneal repair. *Biomaterials* **34**, 9365–9372 (2013).
44. W. Deng, S. W. Tsao, J. N. Lucas, C. S. Leung, A. L. M. Cheung, A new method for improving metaphase chromosome spreading. *Cytom. Part. A* **51A**, 46–51 (2003).
45. T. A. Potapova, J. R. Unruh, A. C. Box, W. D. Bradford, C. W. Seidel, B. D. Slaughter, S. Sivagnanam, Y. Wu, R. Li, Karyotyping human and mouse cells using probes from single-sorted chromosomes and open source software. *Biotechniques* **59**, 335–346 (2015).

Acknowledgments: We thank K. Burns and F. M. Gabhann for critically reading the manuscript. **Funding:** We acknowledge funding from the National Cancer Institute (U54CA143868 and R01CA174388), the National Institute of Neurological Disorders and Stroke (R21NS087485), the National Institute of General Medical Sciences (R35GM118172), and the American Heart Association (12POST12050638). **Author contributions:** P.-H.W. and D.W. designed the experiments. P.-H.W., D.M.G., J.M.P., A.N., J.M., T.W.-T.C., and M.-H.L. collected the data. P.-H.W. developed the analytical tools and analyzed the data. P.-H.W. composed the figures and wrote the Supplementary Materials and supplementary figures. P.-H.W. and D.W. wrote the manuscript. D.M.G. and R.L. edited the manuscript. **Competing interests:** D.W., P.-H.W., and J.M.P. are inventors on a patent related to this work (no. US8934698B2, expiring 12 October 2032). All authors declare that they have no other competing interests. **Data and materials availability:** All data needed to evaluate the conclusions in the paper are present in the paper and/or the Supplementary Materials. Additional data related to this paper may be requested from the authors.

Submitted 16 January 2019
 Accepted 13 November 2019
 Published 22 January 2020
 10.1126/sciadv.aaw6938

Citation: P.-H. Wu, D. M. Gilkes, J. M. Phillip, A. Narkar, T. W.-T. Cheng, J. Marchand, M.-H. Lee, R. Li, D. Wirtz, Single-cell morphology encodes metastatic potential. *Sci. Adv.* **6**, eaaw6938 (2020).

CELL BIOLOGY

Coupling chromatin structure and dynamics by live super-resolution imaging

R. Barth^{1,2}, K. Bystricky^{1,3}, H. A. Shaban^{1,4,*†}

Chromatin conformation regulates gene expression and thus, constant remodeling of chromatin structure is essential to guarantee proper cell function. To gain insight into the spatiotemporal organization of the genome, we use high-density photoactivated localization microscopy and deep learning to obtain temporally resolved super-resolution images of chromatin in living cells. In combination with high-resolution dense motion reconstruction, we find elongated ~45- to 90-nm-wide chromatin “blobs.” A computational chromatin model suggests that these blobs are dynamically associating chromatin fragments in close physical and genomic proximity and adopt topologically associated domain-like interactions in the time-average limit. Experimentally, we found that chromatin exhibits a spatiotemporal correlation over ~4 μm in space and tens of seconds in time, while chromatin dynamics are correlated over ~6 μm and last 40 s. Notably, chromatin structure and dynamics are closely related, which may constitute a mechanism to grant access to regions with high local chromatin concentration.

INTRODUCTION

The three-dimensional organization of the eukaryotic genome plays a central role in gene regulation (1). Its spatial organization has been prominently characterized by molecular and cellular approaches including high-throughput chromosome conformation capture (Hi-C) (2) and fluorescent in situ hybridization (3). Topologically associated domains (TADs), genomic regions that display a high degree of interaction, were revealed and found to be a key architectural feature (4). Direct three-dimensional localization microscopy of the chromatin fiber at the nanoscale (5) confirmed the presence of TADs in single cells but also, among others, revealed great structural variation of chromatin architecture (3). To comprehensively resolve the spatial heterogeneity of chromatin, super-resolution microscopy must be used. Previous work showed that nucleosomes are distributed as segregated, nanometer-sized accumulations throughout the nucleus (6–8) and that the epigenetic state of a locus has a large impact on its folding (9, 10). However, to resolve the fine structure of chromatin, high labeling densities, long acquisition times, and, often, cell fixation are required. This precludes capturing dynamic processes of chromatin in single live cells, yet chromatin moves at different spatial and temporal scales.

The first efforts to relate chromatin organization and its dynamics were made using a combination of photoactivated localization microscopy (PALM) and tracking of single nucleosomes (11). It could be shown that nucleosomes mostly move coherently with their underlying domains, in accordance with conventional microscopy data (12); however, a quantitative link between the observed dynamics and the surrounding chromatin structure could not yet be established in real time. Although it is becoming increasingly clear that chromatin motion and long-range interactions are key to genome organization and gene regulation (13), tools to detect and to define

bulk chromatin motion simultaneously at divergent spatiotemporal scales and high resolution are still missing.

Here, we apply deep learning-based PALM (Deep-PALM) for temporally resolved super-resolution imaging of chromatin in vivo. Deep-PALM acquires a single resolved image in a few hundred milliseconds with a spatial resolution of ~60 nm. We observed elongated ~45- to 90-nm-wide chromatin domain “blobs.” Using a computational chromosome model, we inferred that blobs are highly dynamic entities, which dynamically assemble and disassemble. Consisting of chromatin in close physical and genomic proximity, our chromosome model indicates that blobs, nevertheless, adopt TAD-like interaction patterns when chromatin configurations are averaged over time. Using a combination of Deep-PALM and high-resolution dense motion reconstruction (14), we simultaneously analyzed both structural and dynamic properties of chromatin. Our analysis emphasizes the presence of spatiotemporal cross-correlations between chromatin structure and dynamics, extending several micrometers in space and tens of seconds in time. Furthermore, extraction and statistical mapping of multiple parameters from the dynamic behavior of chromatin blobs show that chromatin density regulates local chromatin dynamics.

RESULTS

Deep-PALM reveals dynamic chromatin remodeling in living cells

Super-resolution imaging of complex and compact macromolecules such as chromatin requires dense labeling of the chromatin fiber to resolve fine features. We use Deep-STORM, a method that uses a deep convolutional neural network (CNN) to predict super-resolution images from stochastically blinking emitters (Fig. 1A; see Materials and Methods) (15). The CNN was trained to specific labeling densities for live-cell chromatin imaging using a photoactivated fluorophore (PATagRFP); we therefore refer to the method as Deep-PALM. We chose three labeling densities 4, 6, and 9 emitters/ μm^2 per frame in the ON-state to test on the basis of the comparison of simulated and experimental wide-field images (fig. S1A). The CNN trained with 9 emitters/ μm^2 performed significantly worse than the other CNNs and was thus excluded from further analysis (fig. S1B; see Materials

Copyright © 2020
The Authors, some
rights reserved;
exclusive licensee
American Association
for the Advancement
of Science. No claim to
original U.S. Government
Works. Distributed
under a Creative
Commons Attribution
NonCommercial
License 4.0 (CC BY-NC).

¹Laboratoire de Biologie Moléculaire Eucaryote (LBME), Centre de Biologie Intégrative (CBI), CNRS, University of Toulouse, UPS, 31062 Toulouse, France. ²Department of Bionanoscience, Delft University of Technology, 2628 CJ, Delft, Netherlands. ³Institut Universitaire de France (IUF), Paris, France. ⁴Spectroscopy Department, Physics Division, National Research Centre, Dokki, 12622 Cairo, Egypt.

*Present address: Center for Advanced Imaging, Faculty of Arts and Sciences, North-west Building, Harvard University, Cambridge, MA 02138, USA.

†Corresponding author. Email: h_shaban@g.harvard.edu

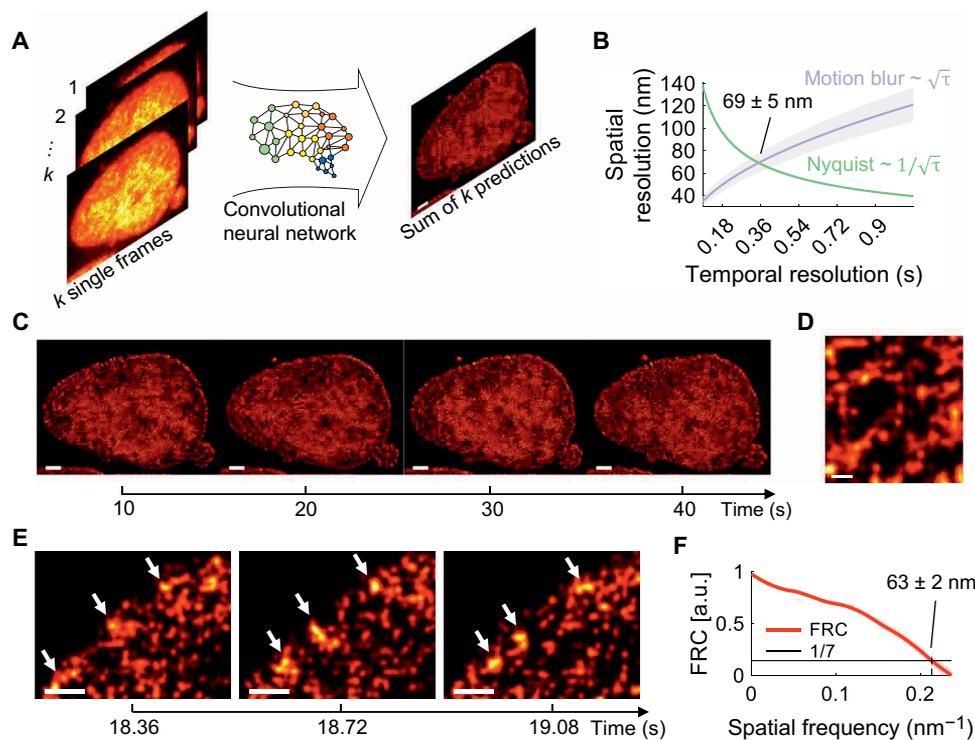


Fig. 1. Temporally resolved super-resolution images of chromatin in U2OS nuclei. (A) Wide-field images of U2OS nuclei expressing H2B-PATagRFP are input to a trained CNN, and predictions from multiple input frames are summed to construct a super-resolved image of chromatin in vivo. (B) The resolution trade-off between the prolonged acquisition of emitter localizations (green line) and motion blur due to diffusion of the underlying diffusion processes (purple line). For our experimental data, the localization density per second is $\rho = (2.4 \pm 0.1) \mu\text{m}^{-2}\text{s}^{-1}$, the diffusion constant is $D = (3.4 \pm 0.8) \cdot 10^{-3} \mu\text{m}^2\text{s}^{-1}$ (see fig. S8B), and the acquisition time per frame is $\tau = 30$ ms. The spatial resolution assumes a minimum (69 ± 5 nm) at a time resolution of 360 ms. (C) Super-resolution images of a single nucleus at time intervals of about 10 s. Scale bars, 2 μm . (D) Magnification of segregated accumulations of H2B within a chromatin-rich region. Scale bar, 200 nm. (E) Magnification of a stable but dynamic structure (arrows) over three consecutive images. Scale bars, 500 nm. (F) Fourier ring correlation (FRC) for super-resolved images resulting in a spatial resolution of 63 ± 2 nm. FRC was conducted on the basis of 332 consecutive super-resolved images from two cells. a.u. arbitrary units.

and Methods). We applied Deep-PALM to reconstruct an image set of labeled histone protein (H2B-PATagRFP) in human bone osteosarcoma (U2OS) cells using the networks trained on 4 and 6 emitters/ μm^2 per frame (see Materials and Methods). A varying number of predictions by the CNN of each frame of the input series were summed to reconstruct a temporal series of super-resolved images (fig. S1C). The predictions made by the CNN trained with 4 emitters/ μm^2 show large spaces devoid of signal intensity, especially at the nuclear periphery, making this CNN inadequate for live-cell super-resolution imaging of chromatin. While collecting photons from long acquisitions for super-resolution imaging is desirable in fixed cells, Deep-PALM is a live imaging approach. Summing over many individual predictions leads to considerable motion blur and thus loss in resolution. Quantitatively, the Nyquist criterion states that the image resolution $R = 2/\sqrt{\rho\tau}$ depends on ρ , the localization density per second, and the time resolution τ (16). In contrast, motion blur strictly depends on the diffusion constant D of the underlying structure $R = \sqrt{4D\tau}$. There is thus an optimum resolution due to the trade-off between increased emitter sampling and the avoidance of motion blur, which was at a time resolution of 360 ms for our experiments (Fig. 1B and fig. S1D).

Super-resolution imaging of H2B-PATagRFP in live cells at this temporal resolution shows a pronounced nuclear periphery, while fluorescent signals in the interior vary in intensity (Fig. 1C). This likely corresponds to chromatin-rich and chromatin-poor regions

(8). These regions rearrange over time, reflecting the dynamic behavior of bulk chromatin. Chromatin-rich and chromatin-poor regions were visible not only at the scale of the whole nucleus but also at the resolution of a few hundred nanometers (Fig. 1D). Within chromatin-rich regions, the intensity distribution was not uniform but exhibited spatially segregated accumulations of labeled histones of variable shape and size, reminiscent of nucleosome clutches (6), nanodomains (9, 11), or TADs (17). At the nuclear periphery, prominent structures arise. Certain chromatin structures could be observed for ~ 1 s, which underwent conformational changes during this period (Fig. 1E). The spatial resolution at which structural elements can be observed (see Materials and Methods) in time-resolved super-resolution data of chromatin was 63 ± 2 nm (Fig. 1E), slightly more optimistic than the theoretical prediction (Fig. 1B) (18).

We compared images of H2B reconstructed from 12 frames (super-resolved images) by Deep-PALM in living cells to super-resolution images reconstructed by 8000 frames of H2B in fixed cells (fig. S2, A and B). Overall, the contrast in the fixed sample appears higher, and the nuclear periphery appears more prominent than in images from living cells. However, in accordance with the previous super-resolution images of chromatin in fixed cells (6, 8, 9, 11, 17) and Deep-PALM images, we observe segregated accumulations of signal throughout the nucleus. Thus, Deep-PALM identifies spatially heterogeneous coverage of chromatin, as previously reported (6, 8, 9, 11, 17). We further monitor chromatin temporally at the nanometer scale in living cells.

Chromatin appears in elongated nanometer-sized blobs with a nonrandom spatial distribution

To quantitatively assess the spatial distribution of H2B, we developed an image segmentation scheme (see Materials and Methods; fig. S3), which allowed us to segment spatially separated accumulations of H2B signal with high fidelity (note S1 and figs. S4 and S5). Applying our segmentation scheme, ~10,000 separable elements, blob-like structures were observed for each super-resolved image (166 resolved images per movie; Fig. 2A). The experimental resolution does not enable us to elucidate their origin and formation because tracking of blobs in three dimensions would be necessary to do so (see Discussion). We therefore turned to a transferable computational model introduced by Qi and Zhang (19), which is based on one-dimensional genomics and epigenomics data, including histone modification profiles and binding sites of CTCF (CCCTC-binding factor). To compare our data to the simulations, super-resolution images were generated

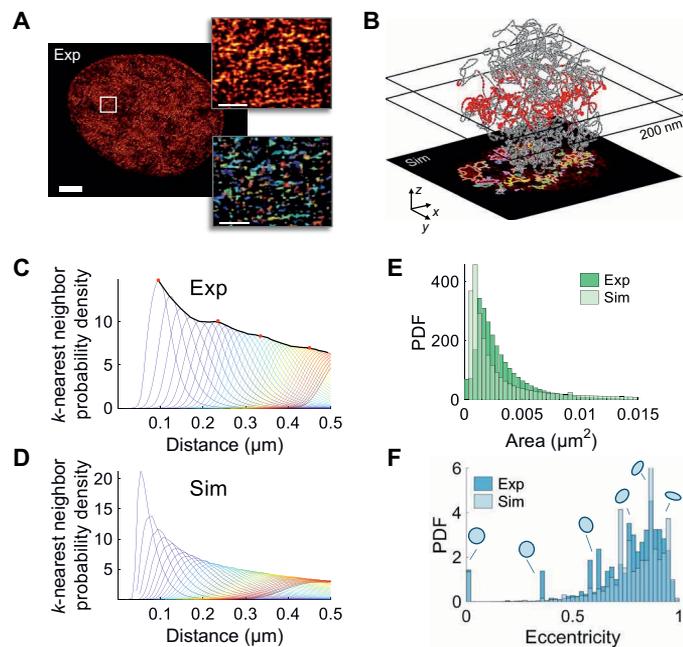


Fig. 2. Chromatin blob identification and characterization of imaged and modeled chromatin. (A) Super-resolved images show blobs of chromatin (left). These blobs are segmented (see Materials and Methods and note S1) and individually labeled by random color (right). Magnifications of the boxed regions are shown. Scale bars, 2 μm (whole nucleus); magnifications, 200 nm. (B) Generation of super-resolution images and blob identification and characterization for a 25-million base pair (Mbp) segment of chromosome 1 from GM12878 cells, as simulated in Qi and Zhang (19). Beads (5-kb genomic length) of a simulated polymer configuration within a 200-nm-thick slab are projected to the imaging plane, resembling experimental super-resolved images of live chromatin. Blobs are identified as on experimental data. (C) From the centroid positions, the NND distributions are computed for up to 40 nearest neighbors (blue to red). The envelope of the k -NND distributions (black line) shows peaks at approximately 95, 235, 335, and 450 nm (red dots). (D) k -NND distributions as in (B) for simulated data. (E) Area distribution of experimental and simulated blobs. The distribution is, in both cases, well described by a lognormal distribution with parameters $(3.3 \pm 2.8) \times 10^{-3} \mu\text{m}^2$ for experimental blobs and $(3.1 \pm 3.2) \times 10^{-3} \mu\text{m}^2$ for simulated blobs (means \pm SD). PDF, probability density function. (F) Eccentricity distribution for experimental and simulated chromatin blobs. Selected eccentricity values are illustrated by ellipses with the corresponding eccentricity. Eccentricity values range from 0, describing a circle, to 1, describing a line. Prominent peaks arise because of the discretization of chromatin blobs in pixels. The data are based on 332 consecutive super-resolved images from two cells, in each of with ~10,000 blobs were identified.

from the modeled chromosomes. Within these images, we could identify and characterize “chromatin blobs” analogously to those derived from experimental data (see Materials and Methods; Fig. 2B).

For imaged (in living and fixed cells) and modeled chromatin, we first computed the k th nearest-neighbor distance (NND; centroid-to-centroid) distributions, taking into account the nearest 1st to 40th neighbors (Fig. 2C and fig. S2, C and D, blue to red). Centroids of the nearest neighbors are (95 ± 30) nm (means \pm SD) apart, consistent with previous and our own super-resolution images of chromatin in fixed cells (9) and slightly further than what was found for clutches of nucleosomes (6). The envelope of all NND distributions (Fig. 2C, black line) shows several weak maxima at ~95, 235, 335, and 450 nm, which roughly coincide with the peaks of the 1st, 7th, 14th, and 25th nearest neighbors, respectively (Fig. 2C, red dots). In contrast, simulated data exhibit a prominent first nearest-neighbor peak at a slightly smaller distance, and higher-order NND distribution decay quickly and appear washed out (Fig. 2D). This hints toward greater levels of spatial organization of chromatin *in vivo*, which is not readily recapitulated in the used state-of-the-art chromosome model.

Next, we were interested in the typical size of chromatin blobs. Their area distribution (Fig. 2E) fit a log-normal distribution with parameters $(3.3 \pm 2.8) \times 10^{-3} \mu\text{m}^2$ (means \pm SD), which is in line with the area distribution derived from fixed samples (fig. S2E) and modeled chromosomes. Notably, blob areas vary considerably, as indicated by the high SD and the prominent tail of the area distribution toward large values. Following this, we calculated the eccentricity of each blob to resolve their shape (Fig. 2F and fig. S2F). The eccentricity is a measure of the elongation of a region reflecting the ratio of the longest chord of the shape and the shortest chord perpendicular to it (Fig. 2F; illustrated shapes at selected eccentricity values). The distribution of eccentricity values shows an accumulation of values close to 1, with a peak value of ~0.9, which shows that most blobs have an elongated, fiber-like shape and are not circular. In particular, the eccentricity value of 0.9 corresponds to a ratio between the short and long axes of the ellipse of 1:2 (see Materials and Methods), which results, considering the typical area of blobs in experimental and simulated data, in roughly 92-nm-long and 46-nm-wide blobs on average. A highly similar value was found in fixed cells (fig. S2F). The length coincides with the value found for the typical NND [Fig. 2C; (95 ± 30) nm]. However, because of the segregation of chromatin into blobs, their elongated shape, and their random orientation (Fig. 2A), the blobs cannot be closely packed throughout the nucleus. We find that chromatin has a spatially heterogeneous density, occupying 5 to 60% of the nuclear area (fig. S6, A and B), which is supported by a previous electron microscopy study (20).

Blob dimensions derived from live-cell super-resolution imaging using Deep-PALM are consistent with those found in fixed cells, thereby further validating our method, and in agreement with previously determined size ranges (6, 9). A previously published chromosome model based on Hi-C data (and thus not tuned to display blob-like structures *per se*) also displays blobs with dimensions comparable to those found here, in living cells. Together, these data strongly suggest the existence of spatially segregated chromatin structures in the sub-100-nm range.

Chromatin blobs identified by Deep-PALM are coherent with sub-TADs

The simulations offer to track each monomer (chromatin locus) unambiguously, which is currently not possible to do from experimental

data. Since the simulations show blobs comparable to those found in experiment (Fig. 2), simulations help to indicate possible mechanisms leading to the observation of chromatin blobs. For instance, because of the projection of the nuclear volume onto the imaging plane, the observed blobs could simply be overlays of distant, along the one-dimensional genome, noninteracting genomic loci. To examine this possibility, we analyzed the gap length between beads belonging to the same blob along the simulated chromosome. Beads constitute the monomers of the simulated chromosome, and each bead represents roughly 5 kb (19).

The analysis showed that the blobs are mostly made of consecutive beads along the genome, thus implying an underlying domain-like structure, similar to TADs (Fig. 3A). Using the affiliation of each

bead to an intrinsic chromatin state of the model (Fig. 3B), it became apparent that blobs along the simulated chromosome consisting mostly of active chromatin are significantly larger than those formed by inactive and repressive chromatin (Fig. 3C). These findings are in line with experimental results (10) and results from the simulations directly (19), thereby validating the projection and segmentation process.

Since chromatin is dynamic in vivo and in computer simulations, each bead can diffuse in and out of the imaging volume from frame to frame. We estimated that, on average, each bead spent approximately 1.5 s continuously within a slab of 200-nm thickness (Fig. 3D). Furthermore, a bead is, on average, found only 0.55 ± 0.33 s continuously within a blob, which corresponds to one to two experimental super-resolved images (Fig. 3D). These results suggest that chromatin blobs are highly dynamic entities, which usually form and dissemble within less than 1 s. We thus constructed a time-averaged association map for the modeled chromosomes, quantifying the frequency at which each locus is found with any other locus within one blob. The association map is comparable to interaction maps derived from Hi-C (Fig. 3E). Notably, interlocus association and Hi-C maps are strongly correlated, and the association map shows similar patterns as those identified as TADs in Hi-C maps, even for relatively distant genomic loci [>1 million base pairs (Mbp)]. A similar TAD-like organization is also apparent when the average inverse distance between loci is considered (Fig. 3F, top), suggesting that blobs could be identified in super-resolved images because of the proximity of loci within blobs in physical space. The computational chromosome model indicates that chromatin blobs identified by Deep-PALM are mostly made of continuous regions along the genome and cannot be attributed to artifacts originating from the projection of the three-dimensional genome structure to the imaging plane. The simulations further indicate that the blobs associate and dissociate within less than 1 s, but loci within blobs are likely to belong to the same TAD. Their average genomic content is 75 kb, only a fraction of typical TAD lengths in mammalian cells (average size, 880 kb) (4), suggesting that blobs likely correspond to sub-TADs or TAD nanocompartments (17).

Quantitative chromatin dynamics at nanoscale resolution

To quantify the experimentally observed chromatin dynamics at the nanoscale, down to the size of one pixel (13.5 nm), we used a dense reconstruction of flow fields, optical flow (Fig. 4A; see Materials and Methods), which was previously used to analyze images taken on confocal (12, 14), and structured illumination microscopes (8). We examined the suitability of optical flow for super-resolution on the basis of single-molecule localization images using simulations. We find that the accuracy of optical flow is slightly enhanced on super-resolved images compared to conventional fluorescence microscopy images (note S2 and fig. S7, A to C). Experimental super-resolution flow fields are illustrated on the basis of two subsequent images, between which the dynamics of structural features are apparent to the eye (fig. S7, D and E). On the nuclear periphery, connected regions spanning up to ~ 500 nm can be observed [fig. S7D (i and ii), marked by arrows]. These structures are stable for at least 360 ms but move from frame to frame. The flow field is shown on top of an overlay of the two super-resolved images and color-coded [fig. S7D (iii); the intensity in frame 1 is shown in green, the intensity in frame 2 is shown in purple, and colocalization of both is white]. Displacement vectors closely follow the redistribution of

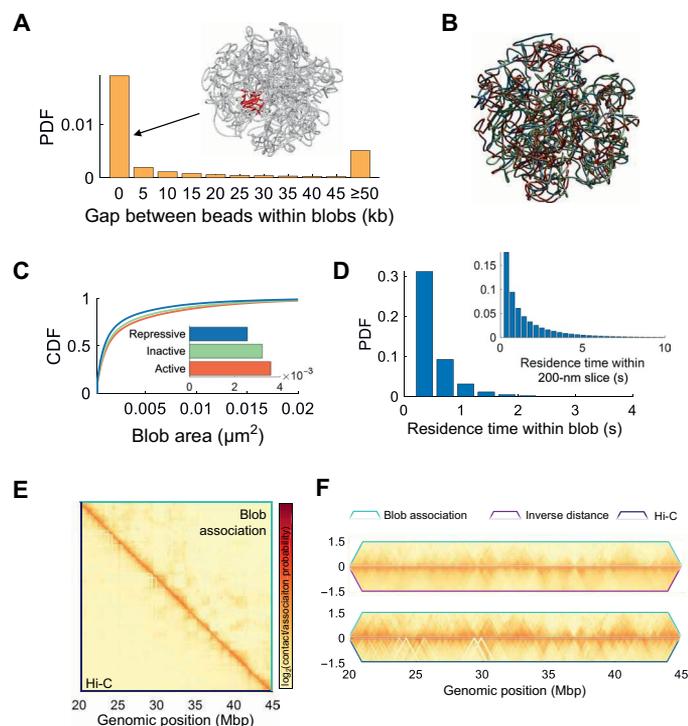


Fig. 3. Chromatin blobs on modeled chromosomes consist of continuous loci along the genome and exhibit a TAD-like time-averaged conformation. (A) Gap length between beads belonging to the same blob. An exemplary blob with small gap length is shown. The blob is mostly made of consecutive beads being in close spatial proximity. (B) A representative polymer configuration is colored according to chromatin states (red, active; green, inactive; and blue, repressive). (C) The cumulative distribution function (CDF) of clusters within active, inactive, and repressive chromatin. Inset: Mean area of clusters within the three types of chromatin. The distributions are all significantly different from each other, as determined by a two-sample Kolmogorov-Smirnov test ($P < 10^{-50}$). (D) Distribution of the continuous residence time of any monomer within a cluster (0.5 ± 0.3 s; means \pm SD). Inset: Continuous residence time of any monomer within a slab of 200-nm thickness (1.5 ± 1.6 s; means \pm SD). (E) The blob association strength between any two beads is measured as the frequency at which any two beads are found in one blob. The association map is averaged over all simulated configurations (upper triangular matrix; from simulations), and experimental Hi-C counts are shown for the same chromosome segment [lower triangular matrix; from Rao *et al.* (40)]. The association and Hi-C maps are strongly correlated [Pearson's correlation coefficient (PCC) = 0.76]. (F) Close-up views around the diagonal of Hi-C-like matrices. The association strength is shown together with the inverse distance between beads (top; PCC = 0.85) and with experimental Hi-C counts [bottom; as in (E)]. The data are based on 20,000 polymer configurations.

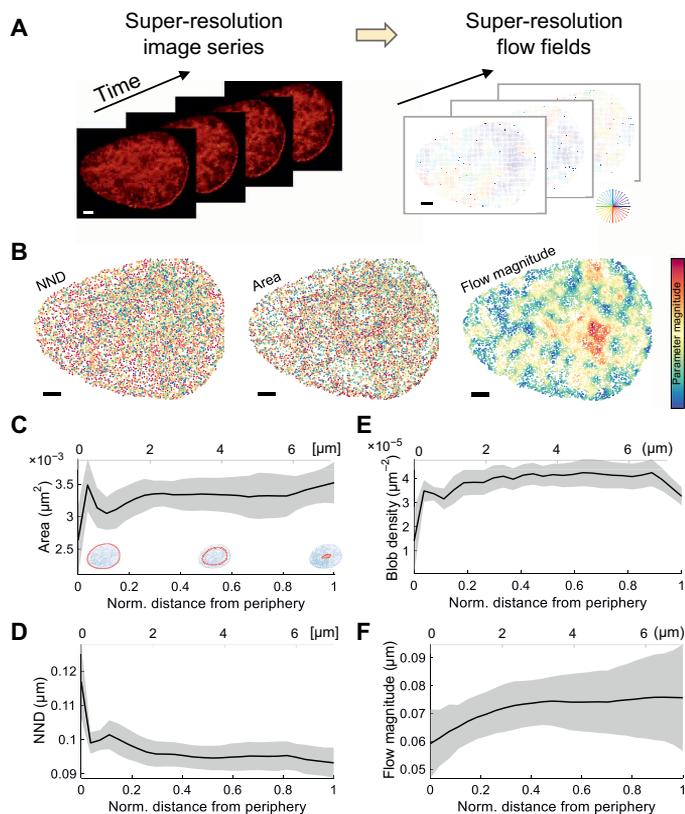


Fig. 4. Structural and dynamics blob characteristics dependent on the proximity to the nuclear periphery. (A) A time series of super-resolution images (left) is subject to optical flow (right). (B) Blobs of a representative nucleus (see movie S1) are labeled by their NND (left), area (middle), and flow magnitude (right). Colors denote the corresponding parameter magnitude. (C) The average blob area, (D) NND, (E) density, and (F) flow magnitude are shown versus the normalized distance from the nuclear periphery (lower x axis; 0 is on the periphery and 1 is at the center of the nucleus) and versus the absolute distance (upper x axis). Line and shaded area denote the means \pm SE from 322 super-resolved images of two cells. Scale bar, (A) and (B): $3 \mu\text{m}$.

intensity from frame to frame (roughly from green to purple). Similarly, structures within the nuclear interior (fig. S7E) can be followed by eye, thus further validating and justifying the use of a dense motion reconstruction as a quantification tool of super-resolved chromatin motion.

Using optical flow fields, we linked the spatial appearance of chromatin to their dynamics. Effectively, the blobs were characterized with two structural parameters (NND and area) and their flow magnitude (Fig. 4B). Movie S1 shows the time evolution of those parameters for an exemplary nucleus. Blobs at the nuclear periphery showed a distinct behavior from those in the nuclear interior. In particular, the periphery exhibits a lower density of blobs, but those appear slightly larger and are less mobile than in the nuclear interior (Fig. 4, C to F), in line with previous findings using conventional microscopy (14). The peripheral blobs are reminiscent of dense and relatively immobile heterochromatin and lamina-associated domains (21), which extend only up to $0.5 \mu\text{m}$ inside the nuclear interior. In contrast, blob dynamics increase gradually within 1 to $2 \mu\text{m}$ from the nuclear rim.

Chromatin structure and dynamics are linked

To further elucidate the relationship between chromatin structure and dynamics, we analyzed the correlation between each pair of parameters in space and time. Therefore, we computed the auto- and

cross-correlation of parameter maps with a given time lag across the entire nucleus (in space) (Fig. 5A). In general, a positive correlation denotes a low-low or a high-high relationship (a variable de-/increases when another variable de-/increases), while, analogously, a negative correlation denotes a high-low relationship. The autocorrelation of NND maps [Fig. 5A (i)] shows a positive correlation; thus, regions exist spanning 2 to $4 \mu\text{m}$, in which chromatin is either closely packed (low-low) or widely dispersed (high-high). Likewise, blobs of similar size tend to be in spatial proximity [Fig. 5A (iii)]. These regions are not stable over time but rearrange continuously, an observation bolstered by the fact that the autocorrelation diminishes with increasing time lag. The cross-correlation between NND and area [Fig. 5A (ii)] shows a negative correlation for short time lags, suggesting that large blobs appear with a high local density while small ones are more isolated. The correlation becomes slightly positive for time lags ≥ 20 s, indicating that big blobs are present in regions that were sparsely populated before and small blobs tend to accumulate in previously densely populated regions. This is in line with dynamic reorganization and reshaping of chromatin domains on a global scale, as observed in snapshots of the Deep-PALM image series (Fig. 1A).

The flow magnitude is positively correlated for all time lags, while the correlation displays a slight increase for time lags ≤ 20 s [Fig. 5A (vi)], which has also been observed previously (8, 12, 22). The spatial autocorrelation of dynamic and structural properties of chromatin are in stark contrast. While structural parameters are highly correlated at short but not at long time scales, chromatin motion is still correlated at a time scale exceeding 30 s. At very short time scales (< 100 ms), stochastic fluctuations determine the local motion of the chromatin fiber, while coherent motion becomes apparent at longer times (22). However, there exists a strong cross-correlation between structural and dynamic parameters: The cross-correlation between the NND and flow magnitude shows notable negative correlation at all time lags [Fig. 5A (iv)], strongly suggesting that sparsely distributed blobs appear less mobile than densely packed ones. The area seems to play a negligible role for short time lags, but there is a modest tendency that regions with large blobs tend to exhibit increased dynamics at later time points [≥ 10 s; Fig. 5A (v)], likely due to the strong relationship between area and NND.

In general, parameter pairs involving chromatin dynamics exhibit an extended spatial auto- or cross-correlation (up to $\sim 6 \mu\text{m}$; the lower row of Fig. 5A) compared to correlation curves including solely structural parameters (up to 3 to $4 \mu\text{m}$). Furthermore, the cross-correlation of flow magnitude and NND does not considerably change for increasing time lag, suggesting that the coupling between those parameters is characterized by an unexpectedly resilient memory, lasting for at least tens of seconds (23). Concomitantly, the spatial correlation of time-averaged NND maps and maps of the local diffusion constant of chromatin for the entire acquisition time enforces their negative correlation at the time scale of ~ 1 min (fig. S8). Such resilient memory was also proposed by a computational study that observed that interphase nuclei behave similar to concentrated solutions of unentangled ring polymers (24). Our data support the view that chromatin is mostly unentangled since entanglement would influence the anomalous exponent of genomic loci in regions of varying chromatin density (24). However, our data do not reveal a correlation between the anomalous exponent and the time-averaged chromatin density (fig. S8), in line with our previous results using conventional microscopy (14).

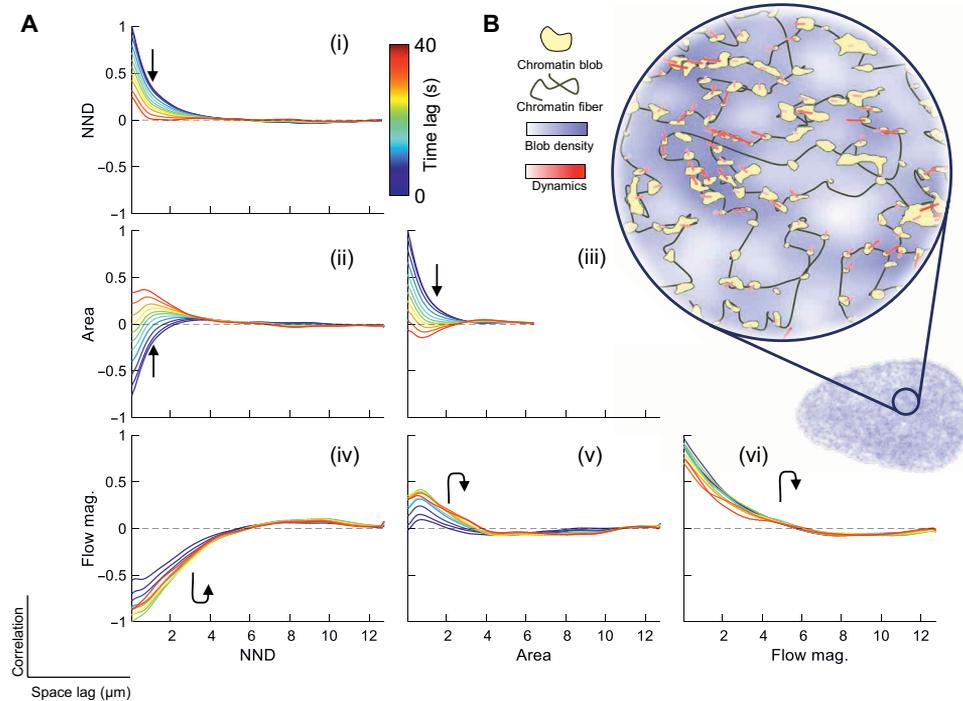


Fig. 5. Spatiotemporal correlations between structural and dynamic parameters. (A) The spatial auto- and cross-correlation between parameters were computed for different time lags. The graphs depict the correlation over space lag for each parameter pair, and different colors denote the time lag (increasing from blue to red). (B) Illustration of the instantaneous relationship between local chromatin density and dynamics. The blob density is shown in blue; the magnitude of chromatin dynamics is shown by red arrows. The consistent negative correlation between NND and flow magnitude is expressed by increased dynamics in regions of high local blob density. Data represent the average over two cells. The cells behave similarly such that error bars are omitted for the sake of clarity.

Overall, the spatial cross-correlation between chromatin structure and dynamics indicates that the NND between blobs and their mobility stand in a strong mutual, negative relationship. This relationship, however, concerns chromatin density variations at the nanoscale, but not global spatial density variations such as in euchromatin or heterochromatin (14). These results support a model in which regions with high local chromatin density, i.e., larger blobs are more prevalent and are mobile, while small blobs are sparsely distributed and less mobile (Fig. 5B). Blob density and dynamics in the long-time limit are, to an unexpectedly large extent, influenced by preceding chromatin conformations.

The local chromatin density is a key regulator of instantaneous chromatin dynamics

The spatial correlations above were only evaluated pairwise, while the behavior of every blob is likely determined by a multitude of factors in the complex energy landscape of chromatin (19, 22). Here, we aim to take a wider range of available information into account to reveal the principle parameters, driving the observed chromatin structure and dynamics. Using a microscopy-based approach, we have access to a total of six relevant structural, dynamic, and global parameters, which potentially shape the chromatin landscape in space and time (Fig. 6A). In addition to the parameters used above, we included the confinement level as a relative measure, allowing the quantification of transient confinement (see Materials and Methods). We further included the bare signal intensity of super-resolved images and, as the only static parameter, the distance from the periphery since it was shown that dynamic and structural parameters

show some dependence on this parameter (Fig. 4). We then used *t*-distributed stochastic neighbor embedding (*t*-SNE) (25), a state-of-the-art dimensionality reduction technique, to map the six-dimensional chromatin “features” (the six input parameters) into two dimensions (Fig. 6A and see note S3). The *t*-SNE algorithm projects data points such that neighbors in high-dimensional space likely stay neighbors in two-dimensional space (25). Visually apparent grouping of points (Fig. 6B) implies that grouped points exhibit great similarity with respect to all input features, and it is of interest to reveal which subset of the input features can explain the similarity among chromatin blobs best. It is likely that points appear grouped because their value of a certain input feature is considerably higher or lower than the corresponding value of other data points. We hence labeled points in *t*-SNE maps which are smaller than the first quartile point or larger than the third quartile point. Data points falling in either of the low/high partition of one input feature are colored accordingly for visualization (Fig. 6D; blue/red points, respectively). We then assigned a rank to each of the input features according to their nearest-neighbor fraction (n-n fraction): Since the *t*-SNE algorithm conserves nearest neighbors, we described the extent of grouping in *t*-SNE maps by the fraction of nearest neighbors, which fall in either one of the subpopulations of low or high points (illustrated in fig. S9). A high n-n fraction (Fig. 6C) therefore indicates that many points marked as low/high are indeed grouped by *t*-SNE and are therefore similar. The ranking (from low to high n-n fraction) reflects the potency of a given parameter to induce similar behavior between chromatin blobs with respect to all input features.

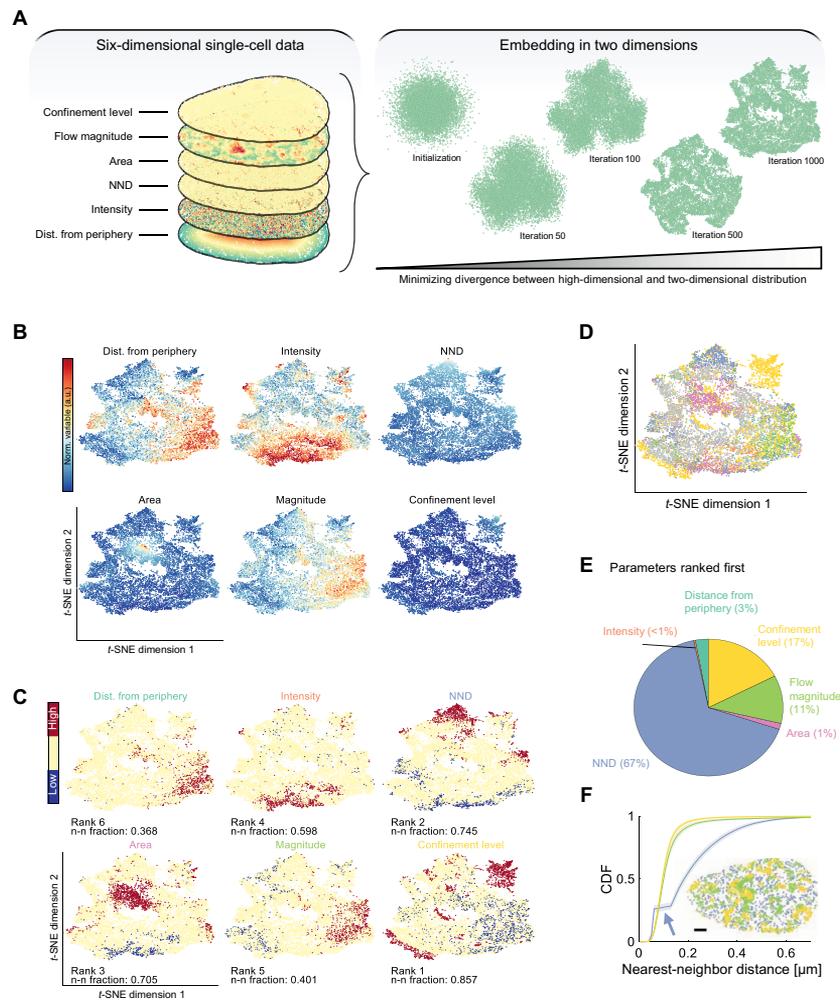


Fig. 6. Chromatin feature extraction. (A) The six-dimensional parameter space is input to the t -SNE algorithm and projected to two dimensions. (B) The two-dimensional embedding of an exemplary dataset is shown and colored according to the magnitude of each input feature (blue to red; the parameter average is shown in beige). (C) Points below the first (blue) and above the third (red) quartile points of the corresponding parameter are marked, and the parameters are ranked according to the fraction of nearest neighbors that fall in one of the marked regions. (D) Data points marked below the first or above the third quartile points are labeled according to the feature in which they were marked. Priority is given to the feature with the higher n-n fraction if necessary. (E) t -SNE analysis is carried out for each nucleus over the whole time series, and it is counted how often a parameter ranked first. The results are visualized as a pie chart. The NND predominantly ranks first in about two-thirds of all cases. (F) Marked points in (C) and (D) are mapped back onto the corresponding nuclei, and the CDF over space is shown (means \pm SE). Pie chart and CDF computations are based on 322 super-resolved images from two cells.

The relative frequency at which each parameter ranked first provides an intuitive feeling for the most “influential” parameters in the dataset (Fig. 6E). The signal intensity plays a negligible role, suggesting that our data are free of potential artifacts related to the bare signal intensity. Furthermore, the blob area and the distance from the periphery likewise do not considerably shape chromatin blobs. In contrast, the NND between blobs was found to be the main factor inducing the observed characteristics in 67% of all-time frames across all nuclei. The flow magnitude and confinement level together rank first in 26% of all cases (11 and 17%, respectively). These numbers suggest that the local chromatin density is a universal key regulator of instantaneous chromatin dynamics. Note that no temporal dependency is included in the t -SNE analysis and, thus, the feature extraction concerns only short-term (≤ 360 ms) relationships. The characteristics of roughly one-fourth of all blobs at each time point are mainly determined by similar dynamical features. Mapping chromatin blobs as marked in Fig. 6 (C and D) back to their respective positions inside the nucleus (Fig. 6F) shows that blobs with low/high flow magnitude or confinement level

markedly also grouped in physical space, which is highly reminiscent of coherent motion of chromatin (12). In contrast, blobs with extraordinary low or high NND were found interspersed throughout the nucleus, in line with spatial correlation analysis between structural and dynamic features (Fig. 5). Our results point toward a large influence of the local chromatin density on the dynamics of chromatin at the scale of a few hundred nanometers and within a few hundred milliseconds. At longer time and length scales, however, previous results suggest that this relationship is lost (14).

DISCUSSION

With Deep-PALM, we present temporally resolved super-resolution images of chromatin in living cells. Our technique identified chromatin nanodomains, named blobs, which mostly have an elongated shape, consistent with the curvilinear arrangement of chromatin, as revealed by structured illumination microscopy (8) with typical axes lengths of 45 to 90 nm. A previous study reported ~ 30 -nm-wide “clutches of nucleosomes” in fixed mammalian cells using STORM nanoscopy (6),

while the larger value obtained using Deep-PALM may be attributed to the motion blurring effect in live-cell imaging. However, histone acetylation and methylation marks were shown to form nanodomains of diameter 60 to 140 nm, respectively (9), which includes the computed dimensions for histone H2B using Deep-PALM.

To elucidate the origin of chromatin blobs, we turned to a simulated chromosome model, which displays chromatin blobs similar to our experimental data when seen in a super-resolution reconstruction. The simulations suggest that chromatin blobs consist of continuous genomic regions with an average length of 75 kb, assembling and disassembling dynamically within less than 1 s. Monomers within blobs display a distinct TAD-like association pattern in the long-time limit, suggesting that the identified blobs represent sub-TADs. Transient formation is consistent with recent findings that TADs are not stable structural elements but exhibit extensive heterogeneity and dynamics (3, 5). To experimentally probe the transient assembly of chromatin blobs, it would be interesting to track individual blobs over time. However, this is a nontrivial task. While the size (area/volume) or shape of blobs could be used to establish correspondences between blobs in subsequent frames, the framework needs to be flexible enough to allow for blob deformations since blobs likely arise stochastically and are not rigid bodies. Achieving an even shorter acquisition time per frame in the future could help minimize the influence of blob deformations and make tracking feasible. The second challenge is to distinguish between disassembly and out-of-focus diffusion of a blob. The three-dimensional imaging at sufficient spatial and temporal resolution will be helpful in the future to overcome this hurdle.

Using an optical flow approach to determine the blob dynamics instead, we found that structural and dynamic parameters exhibit extended spatial and temporal (cross-) correlations. Structural parameters such as the local chromatin density (expressed as the NND between blobs) and area lose their correlation after 3 to 4 μm and roughly 40 s in the spatial and temporal dimension, respectively. In contrast, chromatin mobility correlations extend over $\sim 6 \mu\text{m}$ and persist during the whole acquisition period (≥ 40 s). Extensive spatiotemporal correlation of chromatin dynamics has been presented previously, both experimentally (12) and in simulations (22), but was not linked to the spatiotemporal behavior of the underlying chromatin structure until now. We found that the chromatin dynamics are closely linked to the instantaneous but also to past local structural characterization of chromatin. In other words, the instantaneous local chromatin density influences chromatin dynamics in the future and vice versa. On the basis of these findings, we suggest that chromatin dynamics exhibit an extraordinary long memory. This strong temporal relationship might be established by the fact that stress propagation is affected by the folded chromosome organization (26). Fiber displacements cause structural reconfiguration, ultimately leading to a local amplification of chromatin motion in local high-density environments. This observation is also supported by the fact that increased nucleosome mobility grants chromatin accessibility even within regions of high nucleosome density (27).

Given the persistence at which correlations of chromatin structure and, foremost, dynamics occur in a spatiotemporal manner, we speculate that the interplay of chromatin structure and dynamics could involve a functional relationship (28): Transcriptional activity is closely linked to chromatin accessibility and the epigenomic state (29). Because chromatin structure and dynamics are related, dynamics could also correlate with transcriptional activity (14, 30, 31). However, it is currently unknown whether the structure-dynamics relationship revealed here is strictly mutual or whether it may be causal. Simulations hint that chromatin dynamics follows from structure (22, 23); this question

will be exciting to answer experimentally and in the light of active chromatin remodelers to elucidate a potential functional relationship to transcription. Chromatin regions that are switched from inactive to actively transcribing, for instance, undergo a structural reorganization accompanied by epigenetic modifications (32). The mechanisms driving recruitment of enzymes inducing histone modifications such as histone acetyltransferases, deacetylases, or methyltransferases are largely unknown but often involve the association to proteins (33). Their accessibility to the chromatin fiber is inter alia determined by local dynamics (27). Such a structure-dynamics feedback loop would constitute a quick and flexible way to transiently alter gene expression patterns upon reaction to external stimuli or to coregulate distant genes (1). Future work will study how structure-dynamics correlations differ in regions of different transcriptional activity and/or epigenomic states. Furthermore, probing the interactions between key transcriptional machines such as RNA polymerases with the local chromatin structure and recording their (possibly collective) dynamics could shed light into the target search and binding mechanisms of RNA polymerases with respect to the local chromatin structure. Deep-PALM in combination with optical flow paves the way to answer these questions by enabling the analysis of time-resolved super-resolution images of chromatin in living cells.

MATERIALS AND METHODS

Cell culture

Human osteosarcoma U2OS expressing H2B-PATagRFP cells were a gift from S. Huet (CNRS, UMR 6290, Institut Génétique et Développement de Rennes, Rennes, France); the histone H2B was cloned, as described previously (34). U2OS cells were cultured in Dulbecco's modified Eagle's medium [with glucose (4.5 g/liter)] supplemented with 10% fetal bovine serum (FBS), 2 mM glutamine, penicillin (100 $\mu\text{g}/\text{ml}$), and streptomycin (100 U/ml) in 5% CO_2 at 37°C. Cells were plated 24 hours before imaging on 35-mm petri dishes with a no. 1.5 coverslip-like bottom (ibidi, Biovalley) with a density of 2×10^5 cells per dish. Just before imaging, the growth medium was replaced by Leibovitz's L-15 medium (Life Technologies) supplemented with 20% FBS, 2 mM glutamine, penicillin (100 $\mu\text{g}/\text{ml}$), and streptomycin (100 U/ml).

PALM imaging in living cells

Imaging of H2B-PATagRFP in living U2OS cells was carried out on a fully automated Nikon Ti-E/B PALM (Nikon Instruments) microscope. The microscope is equipped with a full incubator enclosure with gas regulation to maintain a temperature of $\sim 37^\circ\text{C}$ for normal cell growth during live-cell imaging. Image sequences of 2000 frames were recorded with an exposure time of 30 ms per frame (33.3 frames/s). For Deep-PALM imaging, a relatively low power ($\sim 50 \text{ W}/\text{cm}^2$ at the sample) was applied for H2B-PATagRFP excitation at 561 nm and then combined with the 405 nm ($\sim 2 \text{ W}/\text{cm}^2$ at the sample) to photoactivate the molecules between the states. Note that for Deep-PALM imaging, switched fluorophores are not required to stay as long in the dark state as for conventional PALM imaging. We used oblique illumination microscopy (11) combined with total internal reflection fluorescence (TIRF) mode to illuminate a thin layer of 200 nm (axial resolution) across the nucleus. The reconstruction of super-resolved images improves the axial resolution only marginally (fig. S1, E and F). Laser beam powers were controlled by acoustic optic-modulators (AA Opto-Electronic). Both wavelengths were united into an oil immersion 1.49-NA (numerical aperture) TIRF objective (100 \times ; Nikon). An oblique illumination was applied to acquire image series with a high signal-to-noise ratio. The fluorescence emission signal was collected by using the

same objective and spectrally filtered by a Quad-Band beam splitter (ZT405/488/561/647rpc-UF2, Chroma Technology) with a Quad-Band emission filter (ZET405/488/561/647m-TRF, Chroma Technology). The signal was recorded on an electron-multiplying charge-coupled device camera (Andor iXon X3 DU-897, Andor Technology) with a pixel size of 108 nm. For axial correction, Perfect Focus System was applied to correct for defocusing. NIS-Elements software was used for acquiring the images.

PALM imaging and PALM data analysis in fixed cells

The same cell line (U2OS expressing H2B-PATagRFP), as in live-cell imaging, was used for conventional PALM imaging. Before fixation, cells were washed with phosphate-buffered saline (PBS) (three times for 5 min each) and then fixed with 4% paraformaldehyde (Sigma-Aldrich) diluted in PBS for 15 min at room temperature. A movie of 8000 frames was acquired with an exposure time of 30 ms per frame (33.3 frames/s). In comparison to Deep-PALM imaging, a relatively higher excitation laser of 561 nm (~ 60 W/cm² at the sample) was applied to photobleach H2B-PATagRFP and then combined with the 405 nm (~ 2.5 W/cm² at the sample) for photoactivating the molecules. We used the same oblique illumination microscopy combined with TIRF system, as applied in live-cell imaging.

PALM images from fixed cells were analyzed using ThunderSTORM (35). Super-resolution images were constructed by binning emitter localizations into 13.5×13.5 nm pixels and blurred by a Gaussian to match Deep-PALM images. The image segmentation was carried out as on images from living cells (see below).

Deep-PALM analysis

The CNN was trained using simulated data following Nehme *et al.* (15) for three labeling densities (4, 6, and 9 emitters/ μm^2 per frame). Raw imaging data were checked for drift, as previously described (12). The detected drift in raw images is in the range of <10 nm and therefore negligible. The accuracy of the trained net was evaluated by constructing ground truth images from the simulated emitter positions. The structural similarity index is computed to assess the similarity between reconstructed and ground truth images (36)

$$\text{SSIM} = \sum_{x,y} \frac{(2\mu_x\mu_y + C_1)(2\sigma_{xy} + C_2)}{(\mu_x^2 + \mu_y^2 + C_1)(\sigma_x^2 + \sigma_y^2 + C_2)} \quad (1)$$

where x and y are windows of the predicted and ground truth images, respectively, μ and σ denote their local means and SD, respectively, and σ_{xy} denotes their cross-variance. $C_1 = (0.01L)^2$ and $C_2 = (0.03L)^2$ are regularization constants, where L is the dynamic range of the input images. The second quantity to assess CNN accuracy is the root mean square error between the ground truth G and reconstructed image R

$$\text{RMSE} = \sqrt{\frac{1}{N} \sum (R - G)^2} \quad (2)$$

where N is the number of pixels in the images. After training, sequences of all experimental images were compared to the trained network, and predictions of single Deep-PALM images were summed to obtain a final super-resolved image. An up-sampling factor of 8 was used, resulting in an effective pixel size of $108 \text{ nm}/8 = 13.5 \text{ nm}$. A blind/referenceless image spatial quality evaluator (37) was used to determine the optimal number of predictions to be summed. For visualization,

super-resolved images were convolved with a Gaussian kernel ($\sigma = 1$ pixel) and represented using a false red, green, and blue colormap. The parameters of the three trained networks are available at <https://github.com/romanbarth/DeepPALM-trained-models>.

Fourier ring correlation analysis

Fourier ring correlation (FRC) is an unbiased method to estimate the spatial resolution in microscopy images. We follow an approach similar to the one described by Nieuwenhuizen *et al.* (38). For localization-based super-resolution techniques, the set of localizations is divided into two statistically independent subsets, and two images from these subsets are generated. The FRC is computed as the statistical correlation of the Fourier transforms of both subimages over the perimeter of circles of constant frequency in the frequency domain. Deep-PALM, however, does not result in a list of localizations, but in predicted images directly. The set of 12 predictions from Deep-PALM were thus split into two statistically independent subsets, and the method described by Nieuwenhuizen *et al.* (38) was applied.

Chromatin blob identification

The super-resolved images displayed isolated regions of accumulated emitter density. To quantitatively assess the structural information implied by this accumulation of emitters in the focal plane, we developed a segmentation scheme that aims to identify individual blobs (fig. S3). A marker-assisted watershed segmentation was adapted to accurately determine blob boundaries. For this purpose, we use the raw predictions from the deep CNN without convolution (fig. S3A). The foreground in this image is marked by regional maxima and pixels with very high density (i.e., those with $I > 0.99 I_{\text{max}}$; fig. S3B). Since blobs are characterized by surrounding pixels of considerably less density, the Euclidian distance transform is computed on the binary foreground markers. Background pixels (i.e., those pixels not belonging to any blobs) are expected to lie far away from any blob center, and thus, a good estimate for background markers are those pixels being furthest from any foreground pixel. We hence compute the watershed transform on the distance transform of foreground markers, and the resulting watershed lines depict background pixels (fig. S3C). Equipped with fore- and background markers (fig. S3D), we apply a marker-controlled watershed transform on the gradient of the input image (fig. S3E). The marker-controlled watershed imposes minima on marker pixels, preventing the formation of watershed lines across marker pixels. Therefore, the marker-controlled watershed accurately detects boundaries and blobs that might not have been previously marked as foreground (fig. S3F). Last, spurious blobs whose median- or mean intensity is below 10% of the maximum intensity are discarded, and each blob is assigned a unique label for further correspondence (fig. S3G). The area and centroid position are computed for each identified blob for further analysis. This automated segmentation scheme performs considerably better than other state-of-the-art algorithms for image segmentation because of the reliable identification of fore- and background markers accompanied by the watershed transform (note S1).

Chromatin blob properties

Centroid position, area, and eccentricity were computed. The eccentricity is computed by describing the blobs as an ellipse

$$E = \sqrt{1 - a^2/b^2} \quad (3)$$

where a and b are the short and long axes of the ellipse, respectively.

Chromatin blob identification from a computational chromatin model

We chose to use a computational chromatin model, recently introduced by Qi and Zhang (19), to elucidate the origin of experimentally determined chromatin blobs. Each bead of the model covers a sequence length of 5 kb and is assigned 1 of 15 chromatin states to distinguish promoters, enhancers, quiescent chromatin, etc. Starting from the simulated polymer configurations, we consider monomers within a 200-nm-thick slab through the center of the simulated chromosome. To generate super-resolved images as those from Deep-PALM analysis, fluorescence intensity is ascribed to each monomer. Monomer positions are subsequently discretized on a grid with 13.5-nm spacing and convolved with a narrow pointspread function, which results in images closely resembling experimental Deep-PALM images of chromatin. Chromatin blobs were then identified and characterized as on experimental data (Fig. 2, A and B). Mapping back the association of each bead to a blob (if any) allows us to analyze principles of blob formation and maintenance using the distance and the association strength between each pair of monomers, averaged over all 20,000 simulated polymer configurations.

Radial distribution function

The radial distribution function $g(r)$ (also pair correlation function) is calculated (in two dimensions) by counting the number of blobs in an annulus of radius r and thickness dr . The result is normalized by the bulk density $\rho = n/A$, with the total number of blobs n and A , the area of the nucleus, and the area of the annulus, $2\pi r dr$

$$dn(r) = \rho \cdot g(r) \cdot 2\pi r dr \quad (4)$$

Quantification of chromatin dynamics

Super-resolved images of chromatin showed spatially distributed blobs of varying size, but the resolved structure is too dense for state-of-the-art single-particle tracking methods to track. Furthermore are highly dynamic structures, assembling and disassembling within one to two super-resolved frames (Fig. 3D), which makes a single-particle tracking approach unsuitable. Instead, we used a method for dynamics reconstruction of bulk macromolecules with dense labeling, optical flow. Optical flow builds on the computation of flow fields between two successive frames of an image series. The integration of these flow fields from super-resolution images results in trajectories displaying the local motion of bulk chromatin with temporal and high spatial resolution. Further, the trajectories are classified into various diffusion models, and parameters describing the underlying motion are computed (14). Here, we use the effective diffusion coefficient D (in units of m^2/s), which reflects the magnitude of displacements between successive frames (the velocity of particles or monomers in the continuous limit) and the anomalous exponent α (14). The anomalous exponent reflects whether the diffusion is free ($\alpha = 1$, e.g., for noninteracting particles in solution), directed ($\alpha > 1$, e.g., as the result from active processes), or hindered ($\alpha < 1$, e.g., because of obstacles or an effective back-driving force). Furthermore, we compute the length of constraint L_c , which is defined as the SD of the trajectory positions with respect to its time-averaged position. Denoting $\mathbf{R}(t; \mathbf{R}_0)$, the trajectory at time t originating from \mathbf{R}_0 , the expression reads $L_c(\mathbf{R}_0) = \text{var}(\mathbf{R}(t; \mathbf{R}_0))^{1/2}$, where var denotes the variance. The length of constraint is a measure of the length scale explored of the monomer during the observation period. A complementary measure is the confinement level (39), which computes the inverse of the variance of displacements within a sliding window of

length ω : $C \propto \omega / \text{var}(\mathbf{R}(t; \mathbf{R}_0))$, where the sliding window length is set to four frames (1.44 s). Larger values of C denote a more confined state than small ones.

Spatial correlation for temporally varying parameters

The NND and the area, as well as the flow magnitude, were calculated and assigned to the blobs' centroid position. To calculate the spatial correlation between parameters, the parameters were interpolated from the scattered centroid positions onto a regular grid spanning the entire nucleus. Because not every pixel in the original super-resolved images is assigned a parameter value, we chose an effective grid spacing of five pixels (67.5 nm) for the interpolated parameter maps. After interpolation, the spatial correlation was computed between parameter pairs: Let $\mathbf{r} = (x, y)^T$ denote a position on a regular two-dimensional grid and $f(\mathbf{r}, t)$ and $g(\mathbf{r}, t)$ two scalar fields with mean zero and variance one, at time t on that grid. The time series of parameter fields consist of N time points. The spatial cross-correlation between the fields f and g , which lie a lag time τ apart, is then calculated as

$$C(\rho, \tau) = \frac{1}{N} \sum_t \frac{\sum_{x,y} f(\mathbf{r}, t) g(\mathbf{r} + \rho, t + \tau)}{\sum_{x,y} f(\mathbf{r}, t) g(\mathbf{r}, t + \tau)} \quad (5)$$

where the space lag ρ is a two-dimensional vector $\rho = (\Delta x, \Delta y)^T$. The sums in the numerator and denominator are taken over the spatial dimensions; the first sum is taken over time. The average is thus taken over all time points that are compliant with time lag τ . Subsequently, the radial average in space is taken over the correlation, thus effectively calculating the spatial correlation $C(\rho, \tau)$ over the space lag $\rho = \sqrt{\Delta x^2 + \Delta y^2}$. If $f = g$, then the spatial autocorrelation is computed.

Spatial correlation for static parameters

We denote as global parameters those that reflect the structural and dynamic behavior of chromatin spatially resolved in a time-averaged manner. Examples involve the diffusion constant, the anomalous exponent, the length of constraint, but also time-averaged NND maps, etc. (fig. S8). Those parameters are useful to determine time-universal characteristics. The spatial correlation between those parameters is equivalent to the expression given for temporally varying parameters when the temporal dimension is omitted, effectively resulting in a correlation curve $C(\rho)$.

t-distributed stochastic neighbor embedding

The distance from the periphery, intensity, their NND, area, flow magnitude, and confinement level of each identified blob form the six-dimensional-input feature space for t -SNE analysis. The parameters for each blob ($n = 3,260,232$; divided into subsets of approximately 10,000) were z -transformed before the t -SNE analysis. The t -SNE analysis was performed using MATLAB and the Statistics and Machine Learning Toolbox (Release 2017b; The MathWorks Inc., Natick, MA, USA) with the Barnes-Hut approximation. The algorithm was tested using different distance metrics and perplexity values and showed robust results within the examined ranges (note S3 and fig. S10).

SUPPLEMENTARY MATERIALS

Supplementary material for this article is available at <http://advances.sciencemag.org/cgi/content/full/6/27/eaaz2196/DC1>

[View/request a protocol for this paper from Bio-protocol.](#)

REFERENCES AND NOTES

1. B. van Steensel, E. E. M. Furlong, The role of transcription in shaping the spatial organization of the genome. *Nat. Rev. Mol. Cell Biol.* **20**, 327–337 (2019).
2. E. Lieberman-Aiden, N. L. van Berkum, L. Williams, M. Imakaev, T. Ragoczy, A. Telling, I. Amit, B. R. Lajoie, P. J. Sabo, M. O. Dorschner, R. Sandstrom, B. Bernstein, M. A. Bender, M. Groudine, A. Gnirke, J. Stamatoyannopoulos, L. A. Mirny, E. S. Lander, J. Dekker, Comprehensive mapping of long-range interactions reveals folding principles of the human genome. *Science* **326**, 289–293 (2009).
3. E. H. Finn, G. Pegoraro, H. B. Brandão, A.-L. Valton, M. E. Oomen, J. Dekker, L. Mirny, T. Misteli, Extensive heterogeneity and intrinsic variation in spatial genome organization. *Cell* **176**, 1502–1515.e10 (2019).
4. J. R. Dixon, S. Selvaraj, F. Yue, A. Kim, Y. Li, Y. Shen, M. Hu, J. S. Liu, B. Ren, Topological domains in mammalian genomes identified by analysis of chromatin interactions. *Nature* **485**, 376–380 (2012).
5. B. Bintu, L. J. Mateo, J.-H. Su, N. A. Sinnott-Armstrong, M. Parker, S. Kinrot, K. Yamaya, A. N. Boettiger, X. Zhuang, Super-resolution chromatin tracing reveals domains and cooperative interactions in single cells. *Science* **362**, eaau1783 (2018).
6. M. A. Ricci, C. Manzo, M. F. García-Parajo, M. Lakadamyali, M. P. Cosma, Chromatin fibers are formed by heterogeneous groups of nucleosomes in vivo. *Cell* **160**, 1145–1158 (2015).
7. K. Fang, X. Chen, X. Li, Y. Shen, J. Sun, D. M. Czajkowsky, Z. Shao, Super-resolution imaging of individual human subchromosomal regions in situ reveals nanoscopic building blocks of higher-order structure. *ACS Nano* **12**, 4909–4918 (2018).
8. E. Miron, R. Oldenkamp, D. M. S. Pinto, J. M. Brown, A. R. Carvalho Faria, H. A. Shaban, J. D. P. Rhodes, C. Innocent, S. de Ornellas, V. Buckle, L. Schermelleh, Chromatin arranges in filaments of blobs with nanoscale functional zonation. *bioRxiv* doi.org/10.1101/566638 (2019).
9. J. Xu, H. Ma, J. Jin, S. Uttam, R. Fu, Y. Huang, Y. Liu, Super-resolution imaging of higher-order chromatin structures at different epigenomic states in single mammalian cells. *Cell Rep.* **24**, 873–882 (2018).
10. A. N. Boettiger, B. Bintu, J. R. Moffitt, S. Wang, B. J. Believeau, G. Fudenberg, M. Imakaev, L. A. Mirny, C.-T. Wu, X. Zhuang, Super-resolution imaging reveals distinct chromatin folding for different epigenetic states. *Nature* **529**, 418–422 (2016).
11. T. Nozaki, R. Imai, M. Tanbo, R. Nagashima, S. Tamura, T. Tani, Dynamic organization of chromatin domains revealed by super-resolution live-cell imaging. *Mol. Cell* **10**, 1–12 (2017).
12. H. A. Shaban, R. Barth, K. Bystricky, Formation of correlated chromatin domains at nanoscale dynamic resolution during transcription. *Nucleic Acids Res.* **46**, e77 (2018).
13. W. Xiang, M. J. Roberti, J.-K. Hériché, S. Huet, S. Alexander, J. Ellenberg, Correlative live and super-resolution imaging reveals the dynamic structure of replication domains. *J. Cell Biol.* **217**, 1973–1984 (2018).
14. H. A. Shaban, R. Barth, L. Recoules, K. Bystricky, Hi-D: nanoscale mapping of nuclear dynamics in single living cells. *Genome Biol.* **21**, 95 (2020).
15. E. Nehme, L. E. Weiss, T. Michaeli, Y. Shechtman, Deep-STORM: Super-resolution single-molecule microscopy by deep learning. *Optica* **5**, 458–464 (2018).
16. H. Shroff, C. G. Galbraith, J. A. Galbraith, E. Betzig, Live-cell photoactivated localization microscopy of nanoscale adhesion dynamics. *Nat. Methods* **5**, 417–423 (2008).
17. Q. Szabo, D. Jost, J.-M. M. Chang, D. I. Cattoni, G. L. Papadopoulos, B. Bonev, T. Sexton, J. Gurgó, C. Jacquier, M. Nollmann, F. Bantignies, G. Cavalli, TADs are 3D structural units of higher-order chromosome organization in *Drosophila*. *Sci. Adv.* **4**, eaar8082 (2018).
18. R. J. Marsh, K. Pfisterer, P. Bennett, L. M. Hirvonen, M. Gautel, G. E. Jones, S. Cox, Artifact-free high-density localization microscopy analysis. *Nat. Methods* **15**, 689–692 (2018).
19. Y. Qi, B. Zhang, Predicting three-dimensional genome organization with chromatin states. *PLOS Comput. Biol.* **15**, e1007024 (2019).
20. H. D. Ou, S. Phan, T. J. Deerinck, A. Thor, M. H. Ellisman, C. C. O’Shea, ChromEMT: Visualizing 3D chromatin structure and compaction in interphase and mitotic cells. *Science* **357**, eaag0025 (2017).
21. J. Kind, L. Pagie, H. Ortobozkoyun, S. Boyle, S. S. de Vries, H. Janssen, M. Amendola, L. D. Nolen, W. A. Bickmore, B. van Steensel, Single-cell dynamics of genome-nuclear lamina interactions. *Cell* **153**, 178–192 (2013).
22. M. Di Pierro, D. A. Potoyan, P. G. Wolynes, J. N. Onuchic, Anomalous diffusion, spatial coherence, and viscoelasticity from the energy landscape of human chromosomes. *Proc. Natl. Acad. Sci. U.S.A.* **115**, 7753–7758 (2018).
23. L. Liu, G. Shi, D. Thirumalai, C. Hyeon, Chain organization of human interphase chromosome determines the spatiotemporal dynamics of chromatin loci. *PLOS Comput. Biol.* **14**, e1006617 (2018).
24. A. Rosa, R. Everaers, Structure and dynamics of interphase chromosomes. *PLOS Comput. Biol.* **4**, e1000153 (2008).
25. L. Van der Maaten, G. Hinton, Visualizing data using t-SNE. *J. Machine Learning Res.* **9**, 2579–2605 (2008).
26. K. E. Polovnikov, M. Gherardi, M. Cosentino-Lagomarsino, M. V. Tamm, Fractal folding and medium viscoelasticity contribute jointly to chromosome dynamics. *Phys. Rev. Lett.* **120**, 088101 (2018).
27. S. Hihara, C.-G. Pack, K. Kaizu, T. Tani, T. Hanafusa, T. Nozaki, S. Takemoto, T. Yoshimi, H. Yokota, N. Imamoto, Y. Sako, M. Kinjo, K. Takahashi, T. Nagai, K. Maeshima, Local nucleosome dynamics facilitate chromatin accessibility in living mammalian cells. *Cell Rep.* **2**, 1645–1656 (2012).
28. L. Giorgetti, R. Galupa, E. P. Nora, T. Piolot, F. Lam, J. Dekker, G. Tiana, E. Heard, Predictive polymer modeling reveals coupled fluctuations in chromosome conformation and transcription. *Cell* **157**, 950–963 (2014).
29. S. M. Görisch, M. Wachsmuth, K. F. Tóth, P. Lichter, K. Rippe, Histone acetylation increases chromatin accessibility. *J. Cell Sci.* **118**, 5825–5834 (2005).
30. T. Germier, S. Kocanova, N. Walther, A. Bancaud, H. A. Shaban, H. Sellou, A. Z. Politi, J. Ellenberg, F. Gallardo, K. Bystricky, Real-time imaging of a single gene reveals transcription-initiated local confinement. *Biophys. J.* **113**, 1383–1394 (2017).
31. R. Nagashima, K. Hibino, S. S. Ashwin, M. Babokhov, S. Fujishiro, R. Imai, T. Nozaki, S. Tamura, T. Tani, H. Kimura, M. Shribak, M. T. Kanemaki, M. Sasai, K. Maeshima, Single nucleosome imaging reveals loose genome chromatin networks via active RNA polymerase II. *J. Cell Biol.* **218**, 1511–1530 (2019).
32. N. Soshnikova, D. Duboule, Epigenetic temporal control of mouse Hox genes in vivo. *Science* **324**, 1320–1323 (2009).
33. M. Laisné, N. Gupta, O. Kirsh, S. Pradhan, P.-A. Defossez, Mechanisms of DNA methyltransferase recruitment in mammals. *Genes* **9**, 617 (2018).
34. H. Sellou, T. Lebeaupin, C. Chapuis, R. Smith, A. Hegele, H. R. Singh, M. Kozlowski, S. Bultmann, A. G. Ladurner, G. Timinsky, S. Huet, The poly(ADP-ribose)-dependent chromatin remodeler Alc1 induces local chromatin relaxation upon DNA damage. *Mol. Biol. Cell* **27**, 3791–3799 (2016).
35. M. Ovesný, P. Křížek, J. Borkovec, Z. Švindrych, G. M. Hagen, ThunderSTORM: A comprehensive ImageJ plug-in for PALM and STORM data analysis and super-resolution imaging. *Bioinformatics* **30**, 2389–2390 (2014).
36. Z. Wang, A. C. Bovik, H. R. Sheikh, E. P. Simoncelli, Image quality assessment: From error visibility to structural similarity. *IEEE Trans. Image Process.* **13**, 600–612 (2004).
37. A. Mittal, A. K. Moorthy, A. C. Bovik, No-reference image quality assessment in the spatial domain. *IEEE Trans. Image Process.* **21**, 4695–4708 (2012).
38. R. P. J. Nieuwenhuizen, K. A. Lidke, M. Bates, D. L. Puig, D. Grünwald, S. Stallinga, B. Rieger, Measuring image resolution in optical nanoscopy. *Nat. Methods* **10**, 557–562 (2013).
39. A. Sergé, N. Bertaux, H. Rigneault, D. Marguet, Dynamic multiple-target tracing to probe spatiotemporal cartography of cell membranes. *Nat. Methods* **5**, 687–694 (2008).
40. S. S. P. Rao, M. H. Huntley, N. C. Durand, E. K. Stamenova, I. D. Bochkov, J. T. Robinson, A. L. Sanborn, I. Machol, A. D. Omer, E. S. Lander, E. L. Aiden, A 3D map of the human genome at kilobase resolution reveals principles of chromatin looping. *Cell* **159**, 1665–1680 (2014).
41. W. M. Rand, Objective criteria for the evaluation of clustering methods. *J. Am. Stat. Assoc.* **66**, 846–850 (1971).
42. S. Beucher, B. Marcotegui, B. P. Marcotegui, P algorithm, a dramatic enhancement of the waterfall transformation (2009); <https://hal-mines-paristech.archives-ouvertes.fr/hal-00835019>.

Acknowledgments: We acknowledge support from the Pôle Scientifique de Modélisation Numérique, ENS de Lyon for providing computational resources. We thank B. Zhang (Massachusetts Institute of Technology) for providing data of simulated chromosomes and S. Kocanova (LBME, CBI-CNRS; University of Toulouse) for providing PALM videos for fixed cells. We thank H. Babcock (Harvard University), A. Seeber (Harvard University), and M. Tamm (Moscow State University) for valuable feedback on the manuscript. **Funding:** This publication is based upon work from COST Action CA18127, supported by COST (European Cooperation in Science and Technology). This work is supported by Agence Nationale de la Recherche (ANR) ANDY and Sinfonie grants. **Author contributions:** H.A.S. designed and supervised the project. R.B. designed the data analysis and wrote the code. H.A.S. carried out experimental work. R.B. carried out the data analysis. H.A.S. and R.B. interpreted results. H.A.S., R.B., and K.B. wrote the manuscript. **Competing interests:** The authors declare that they have no competing interests. **Data and materials availability:** All data needed to evaluate the conclusions in the paper are present in the paper and/or the Supplementary Materials. Additional data related to this paper may be requested from the authors.

Submitted 20 September 2019

Accepted 16 April 2020

Published 1 July 2020

10.1126/sciadv.aaz2196

Citation: R. Barth, K. Bystricky, H. A. Shaban, Coupling chromatin structure and dynamics by live super-resolution imaging. *Sci. Adv.* **6**, eaaz2196 (2020).

CELL BIOLOGY

Single yeast cell nanomotions correlate with cellular activity

Ronnie G. Willaert^{1,2,3,4,*}, Pieterjan Vanden Boer^{1,2,3†}, Anton Malovichko^{1,5†}, Mitchel Alioscha-Perez^{1,6}, Ksenija Radotić⁷, Dragana Bartolić⁷, Aleksandar Kalauzi⁷, Maria Ines Villalba⁸, Dominique Sanglard⁹, Giovanni Dietler^{1,5}, Hichem Sahli^{1,6,10,11}, Sandor Kasas^{1,5,12*}

Living single yeast cells show a specific cellular motion at the nanometer scale with a magnitude that is proportional to the cellular activity of the cell. We characterized this cellular nanomotion pattern of nonattached single yeast cells using classical optical microscopy. The distribution of the cellular displacements over a short time period is distinct from random motion. The range and shape of such nanomotion displacement distributions change substantially according to the metabolic state of the cell. The analysis of the nanomotion frequency pattern demonstrated that single living yeast cells oscillate at relatively low frequencies of around 2 hertz. The simplicity of the technique should open the way to numerous applications among which antifungal susceptibility tests seem the most straightforward.

INTRODUCTION

Our team developed previously an atomic force microscopy (AFM)–based assay to assess the effects of chemicals on the viability of bacteria (1). The detection is based on the observation that living organisms oscillate at a nanometric scale and transfer these oscillations to the AFM cantilever onto which they are attached. These oscillations stop as soon as the viability of the cells is compromised. We demonstrated that these oscillations are present in living bacteria, yeasts, plant, and mammalian cells (2). The nanomotions of living bacterial cells that are attached to a surface have been confirmed by using other detection methods such as plasmonic imaging of the *z* motion of bacteria (3), tracking the submicron scale *x-y* motion of attached bacteria (4), sensing of attached bacterial vibrations with phase noise of a resonant crystal (5), and subcellular fluctuation imaging (6).

We recently noticed that an optical microscope equipped with a video camera can detect the motions of single yeast cells that are sedimented on a glass surface. In this work, we explored the nanomotion of nonattached yeast cells using this optical nanomotion detection (ONMD) method. The cellular *x-y* motions were monitored by recording 12-s-long movies (1000 frames) taken at a magnification of $\times 400$ (Fig. 1). By periodically recording these movies, tem-

poral behavior of the cells was characterized as a function of different chemical and physical stimuli (Fig. 1, A and B). To track the cellular motions of single cells, we used a cross-correlation image registration algorithm (7). The algorithm is based on the initial estimation of the cross-correlation peak between the first and every subsequent frame. It provides a numerical value for the image translation with a subpixel (submicrometer) resolution. The cell displacement for each frame, the trajectories of tracked cells (Fig. 1C), and the root mean square of the total displacement (Fig. 1F) were calculated. Single-cell nanomotions were characterized by plotting the distribution of the displacements per frame as a violin plot (Fig. 1D). The motions of the set of 20 cells were characterized by plotting grouped cellular displacements per frame as violin plots and the total displacements of 20 cells over 1000 frames as box-and-whisker plots (Fig. 1E).

RESULTS

First, we compared single-cell nanomotions of *Saccharomyces cerevisiae* cells that were grown in the presence of nutrients [by growing them in yeast extract, peptone, and dextrose (YPD) growth medium] to cells that were in a nutrient-free physiological phosphate-buffered saline (PBS) buffer. Single-cell displacements were recorded every hour during 4 hours (Fig. 2, A and B, and fig. S1, A and B). Actively growing single cells showed a large distribution of displacements. The distribution of the displacements is not symmetric, and this reflects the nonrandom behavior of the cells (as could also be observed from the *x-y* displacements graphs in Fig. 3), i.e., cells can make jumps from time to time. This motion behavior is also reflected in the shape of the violin plots that represents the displacements distribution. In this set, a few cells (one to three) display a very small displacement distribution and can be classified as inactive. In contrast, significantly more inactive cells were present in the absence of nutrients, especially after 3 to 4 hours of incubation (Fig. 2B and fig. S1B). This behavior is also reflected in the grouped displacements violin plots (Fig. 2, A and B, bottom) and the total displacements boxplots (fig. S1, C and D). In these last plots, the adaptation of the cells to the new growth condition can clearly be observed, i.e., a significant increase

Copyright © 2020
The Authors, some
rights reserved;
exclusive licensee
American Association
for the Advancement
of Science. No claim to
original U.S. Government
Works. Distributed
under a Creative
Commons Attribution
NonCommercial
License 4.0 (CC BY-NC).

¹International Joint Research Group BioNanotechnology & NanoMedicine (NANO), Vrije Universiteit Brussel—Ecole Polytechnique de Lausanne (EPFL), B-1050 Brussels, Belgium—B-1015 Lausanne, Switzerland. ²Structural Biology Brussels (SBB), Department of Bioengineering Sciences, Vrije Universiteit Brussel, B-1050 Brussels, Belgium. ³Alliance Research Group NanoMicrobiology (NAMI), Vrije Universiteit Brussel, Brussels B-1050, Belgium—Ghent University, B-9000 Ghent, Belgium. ⁴Visiting professor, Department of Bioscience Engineering, University Antwerp, B-2020 Antwerp, Belgium. ⁵Laboratoire de Physique de la Matière Vivante, Ecole Polytechnique Fédérale de Lausanne (EPFL), CH-1015 Lausanne, Switzerland. ⁶Electronics and Informatics Dept (ETRO), AVSP Lab, Vrije Universiteit Brussel, B-1050 Brussels, Belgium. ⁷Institute for Multidisciplinary Research, University of Belgrade, 11000 Beograd, Serbia. ⁸Centro de Investigación y Desarrollo en Fermentaciones Industriales, Universidad Nacional de La Plata, 1900, La Plata, Argentina. ⁹Institute of Microbiology, Lausanne University Hospital and University of Lausanne, CH-1011 Lausanne, Switzerland. ¹⁰Interuniversity Microelectronics Centre (IMEC), B-3001 Heverlee, Belgium. ¹¹Visiting professor, Shaanxi Provincial Key Lab on Speech and Image Information Processing, Northwestern Polytechnical University (NPU), Xi'an, China. ¹²Unité Facultaire d'Anatomie et de Morphologie (UFAM), CUMRL, University of Lausanne, CH-1005 Lausanne, Switzerland. *Corresponding author. Email: ronnie.willaert@vub.be (R.G.W.); sandor.kasas@epfl.ch (S.K.) †These authors contributed equally to this work.

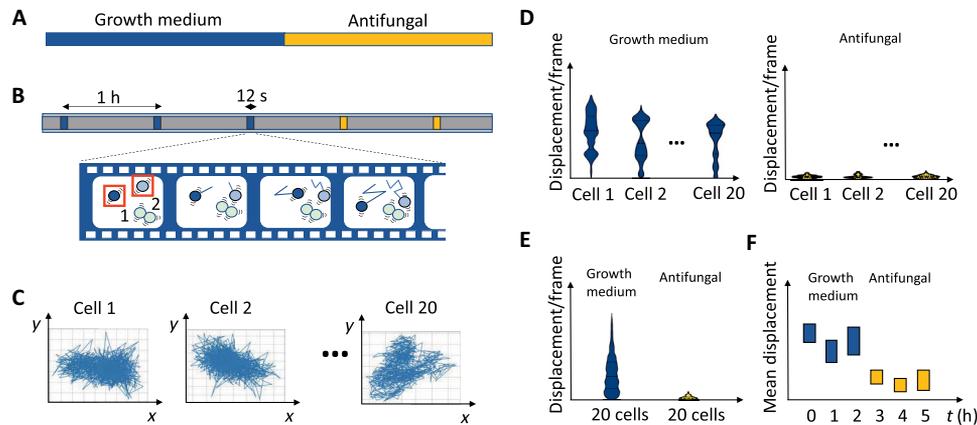


Fig. 1. Overview of the optical nanomotion method to detect cellular nanomotions. (A) Time period of cells in growth medium followed by, e.g., an antifungal treatment. (B) At different time points, movies are recorded of 1000 frames. Cell movements within the box are detected and analyzed. (C) The x - y displacements of individual cells (typically 20 cells) are calculated using the cross-correlation algorithm. (D) For each cell, the displacement per frame is calculated, and this distribution is represented by a combined violin and box plot. (E) The displacement per frame for all cells for a condition/sampling point is represented as a combined violin plot and box plot. (F) The mean of the total displacements of 20 cells is calculated for each condition/sampling point and represented in a box plot.

of the total displacement after 1 hour, in contrast with the measurements obtained in PBS.

The cellular nanomotions were also compared to the motions of silica beads recorded in the same conditions (fig. S2, A to D). The distributions of the displacements were symmetric. The magnitudes of the motions were much reduced compared to living cells and were of the same order of dead cells (fig. S2E).

To assess the effect of the temperature on the nanomotion pattern of yeast, we monitored the cellular oscillations at different temperatures in the range of 13° to 35°C (Fig. 2, C and D, and fig. S3). Each strain is characterized by a different distribution of grouped displacement distributions. For both yeast strains, a maximal activity was detected at around 30°C. This value is in concordance with the documented optimal growth temperature of 30° to 35°C for *S. cerevisiae* BY4742 (8) and 33° to 38°C for *Candida albicans* (9). These experiments demonstrate that the magnitude of the cellular activity is proportional to the magnitude of the distribution of the displacements and the total displacement over 1000 frames.

Next, we characterized the cellular motions of *Candida* and *S. cerevisiae* cells when they are exposed to a killing agent. The *Candida* species can be involved in candidiasis, which is a human fungal infection that can be hard to treat due to the acquired resistance (10). Some of the evaluated yeast strains were hypersusceptible or resistant to the applied antifungal drugs to challenge their viability. First, we explored the effect of a high ethanol concentration (70%) on *C. albicans*, *Candida glabrata*, *Candida lusitanae*, and *S. cerevisiae* cells. The x - y displacements are quickly and markedly reduced after adding ethanol (Fig. 3 and movies S1 and S2). The displacement distributions of the single cells (Fig. 4), the grouped displacements of 20 cells [Fig. 4, E and F (top)], and the total displacements [Fig. 4, E and F (bottom)] were reduced. Ten minutes after the addition of ethanol, a reduction of the displacements in the set of 20 cells for all strains is observed. *C. glabrata* and—in a lesser extent—*S. cerevisiae* are somewhat more ethanol tolerant than *C. glabrata* and *C. lusitanae*, since a significant decrease of the total displacement was only observed after 60 min. Ethanol tolerance is strain dependent (11, 12), affects the growth rate, and will impair the cell membrane integrity (13), which results in ionic species permeability and leakage of metabo-

lites (14); and freely diffuses inside the cell, where it directly perturbs and denatures intracellular proteins (15).

Second, we assessed the effect of killing of cells on the change in cellular nanomotions by exposing them to different concentrations of various antifungals. The x - y displacements were significantly reduced after treatment of the cells with antifungals (Fig. 3B). The effect of amphotericin B, caspofungin, and fluconazole on the cellular motions of *C. albicans* DSY294 wild-type strain is shown in Fig. 5 (A to C) and fig. S4 (A to C). The polyene amphotericin B selectively binds to ergosterol in the cell membrane and causes the formation of pores (which results in a quicker death), whereas the azole fluconazole selectively inhibits cytochrome P450-dependent lanosterol 14- α -demethylase, and the echinocandin caspofungin inhibits fungal β -1,3-glucan synthase (16). For all three antifungals, a significant decrease in the cellular displacements and total displacement was detected after 1-hour treatment. Amphotericin B at a high concentration is very effective, since after already 1 hour, the cellular nanomotion decreased to values close to those recorded on dead cells in the whole set of 20 cells (fig. S4A). Caspofungin did not affect the nanomotion of the candidin-resistant *C. albicans* DSY4614 clinical strain (Fig. 5E) but did kill the hypersusceptible (mutant for efflux systems) *C. albicans* DSY1024 strain (Fig. 5D and table S1). The reported caspofungin minimal inhibitory concentration (MIC) values for *C. albicans* are in the range of 0.03 to 8 $\mu\text{g/ml}$ (17–19). Fluconazole decreased the cellular nanomotions of *C. albicans* DSY294 significantly after 1 hour (Fig. 5C).

To observe the life-death transition, we exposed *C. albicans* DSY294 clinical wild-type strain to lower amphotericin concentrations, including the MIC (which has been reported for *C. albicans* DSY strains as 0.5 $\mu\text{g/ml}$) (20, 21). The total displacements curves show that there is an increase after 1 hour for a concentration of 10 $\mu\text{g/ml}$ (Fig. 5F); the increase is reduced after 2-hour treatment (Fig. 5G). A similar amphotericin B response was recorded for the *C. albicans* CAF2-1 wild-type strain (fig. S6). On the single-cell level, a larger number of cells show a significantly reduced motion for amphotericin B concentration in the range of 0.1 to 0.5 $\mu\text{g/ml}$ (figs. S5 and S6), which corresponds to the reported MIC value of 0.3 $\mu\text{g/ml}$ for this strain (17).

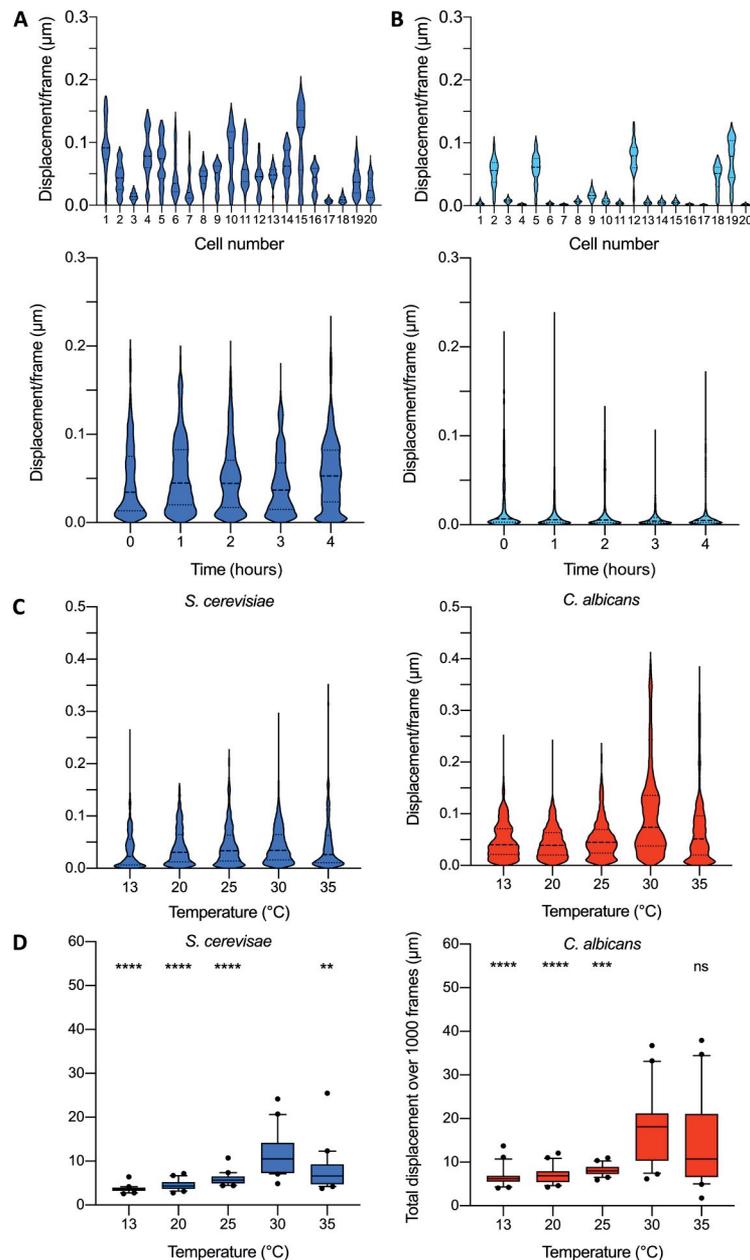


Fig. 2. Effect of the nutritional environment and the temperature on the cellular nanomotions of yeast cells. (A) The distribution of the displacements per frame of 20 *S. cerevisiae* BY4742 cells growing in YPD growth medium after 2 hours (top). Time evolution of the merged distributions of the displacements for 20 cells (bottom). (B) The distribution of the displacements per frame of 20 *S. cerevisiae* BY4742 cells present in PBS after 2 hours (top). Time evolution of the merged distributions of the displacements for 20 cells (bottom). (C) Time evolution of grouped displacements in growth medium. Effect of the temperature on the displacement distribution (20 cells) for *S. cerevisiae* BY4742 and *C. albicans* DSY294. (D) Effect of the temperature on the total displacement of *S. cerevisiae* BY4742 and *C. albicans* DSY294. Wilcoxon test: **** $P < 0.0001$; *** $P < 0.001$; ** $P < 0.01$; ns, not significant.

To evaluate whether the cells interact significantly with the glass surface, which could influence the measured displacements, we compared the cellular nanomotions on a cell-repellent surface [poly-L-lysine (PLL)-polyethylene glycol (PEG)-coated glass surface] to the ones on a nontreated glass surface. Comparable results were obtained for *C. albicans* DSY294 and DSY1024 treated with caspofungin (fig. S7). The motion of silica beads on a PLL-PEG-coated surface showed that there was a small effect on the displacement distribution and the total displacement (fig. S2, A to D). When

the *Candida* cells were adhered to the glass surface by concanavalin A, the displacements were reduced strongly and no significant difference between caspofungin treated and nontreated cells could be detected (fig. S2, F to K).

To highlight differences in the oscillation pattern occurring during the life-death transition, we performed numerical analysis in the frequency domain (fig. S8). Every frequency range is characterized by its low- and high-frequency limit. With the present method of recording, the low-frequency limit is determined as the reciprocal value

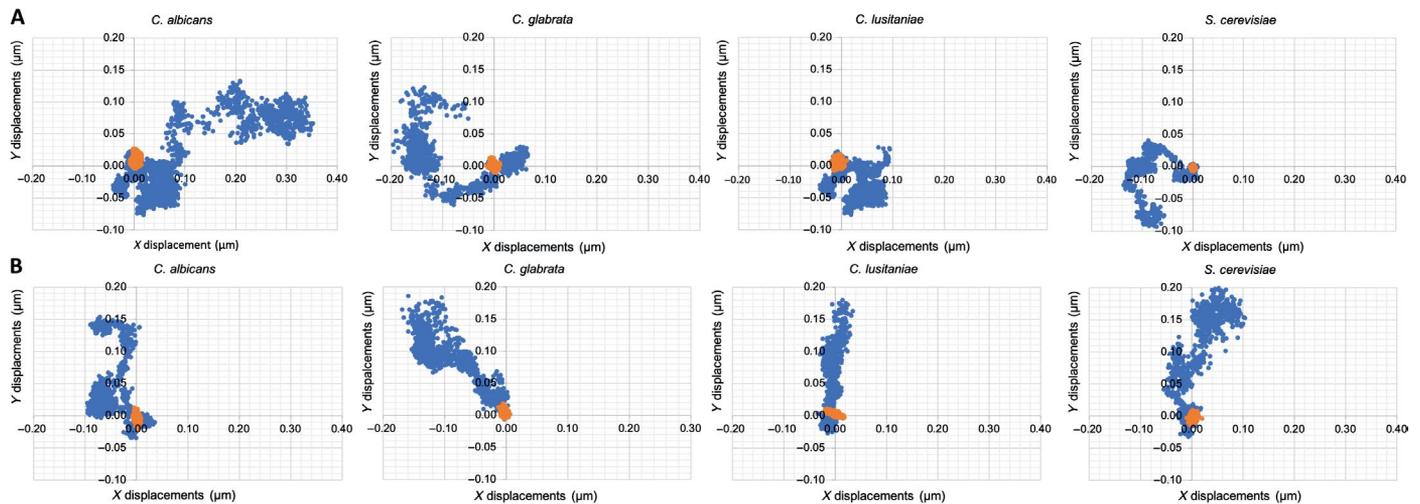


Fig. 3. Monitoring of life-death transition by observing cellular nanomotions of dying yeast cells in the presence of ethanol or antifungal. Effect of (A) ethanol (70%, v/v) and (B) caspofungin (100 $\mu\text{g/ml}$) on the x - y displacements of *C. albicans* DSY294, *C. glabrata* DSY562, *C. lusitanae* DSY4606, and *S. cerevisiae* BY4742 cells during 12 s (1000 frames, 83 fps). Blue dots represent the position of the cell without ethanol or caspofungin treatment; the orange dots represent the positions after the treatment.

of the signal duration, which equals to 0.083 Hz for 12 s. Calculating the high limit is more complex and varies between different cells. We named the high-frequency limit as critical frequency (f_{crit}) and devised a procedure for its calculation. These analyses revealed that untreated resistant *C. albicans* cells show a maximum activity in a frequency range (0.083 - f_{crit}) where f_{crit} varied, varying from 0.7 to 1.5 Hz. *C. albicans* DSY294 had also an f_{crit} at a higher frequency (2.5 to 3.0 Hz) (fig. S8C), whereas caspofungin-treated cells presented a broadening of their f_{crit} , which was most extended for the hypersusceptible DSY1024 cells (fig. S8D). In that case, we also observed an additional f_{crit} around 2 Hz. These results show that dying cells exposed to low antifungal concentrations are characterized by an increased cellular oscillation frequency, and the largest difference of the oscillation pattern in the frequency domain between living and death yeast cells is located in the very low-frequency range.

The data processing was further accelerated by using a deep learning algorithm that detects individual yeast cells (fig. S9A and movies S3 and S4). Deep learning models have been widely successful in automated objects (22, 23) and cells detection (24, 25) tasks. Compared to a manual selection of cells, this approach could permit to analyze automatically a significantly larger number of cells (100 to 1000). The developed algorithm is based on a medium-sized shallow You Only Look Once (YOLO) (24) architecture. As a proof of principle, we reanalyzed the video data demonstrating the influence of the temperature (fig. S9, B and C). The increase in the samples size for the analysis resulted in a smaller distribution of the total displacements.

DISCUSSION

The newly developed ONMD method is based on basic laboratory material, i.e., an optical microscope, a camera, and a computer. The technique is label free and does not require the attachment of the living sample onto a substrate. The method can address the nanomotion pattern of single cells and cellular populations. ONMD has the potential to detect a single resistant cell in a large population. A limitation of the current ONMD method is that only the x - y mo-

tions are recorded. The measurement envelope could be extended by considering also the z axis motion (3), which will further increase the sensitivity of the method.

We could link the cellular nanomotions of single yeast cells to its metabolic activity by comparing the nanomotions of the cells in the presence and absence of nutrients, as well as by detecting a maximum cellular motion at the optimal growth temperature. Living single-cell nanomotions show a nonrandom behavior as was clear from the x - y displacements graphs and the distribution of the displacements during 1000 frames.

The nanomotion analysis of increasing amphotericin B concentration on *C. albicans* DSY294 and *C. albicans* CAF2-1 showed that the analysis based on the distribution of the displacements per frame of single cells seems to be more sensitive than those based on the total displacement of the whole cellular population. The effect of concentrations as low as the MIC of this antifungal became noticeable in a population of 20 cells. In addition, the results showed that at amphotericin B concentrations of around 10 times the MIC, the cellular nanomotion is increased. This indicates that the mechanism of action of amphotericin B (which binds selectively to ergosterol in the cell membrane and causes the formation of pores) (26) increases the motion of the cells, i.e., the antifungal action increases the metabolic activity of the cells, probably due to an increased activity of the efflux pumps. An increase of the nanomotion (measured by the AFM cantilever method) of the bacterium *Bordetella pertussis* for the antibiotic was also previously observed (27).

The analysis of the nanomotion frequency pattern demonstrated that single living yeast cells oscillate at relatively low frequencies of around 2 Hz. These results complement those published by Gimzewski and coworkers (28) who highlighted a periodical motion of the *S. cerevisiae* cell wall in the range of 0.8 to 1.6 kHz. These measurements were accomplished by AFM on yeast cells that were mechanically trapped into a filter pore. An ultrasonic excitation and interferometric motion detection permitted to detect resonance frequencies of single *S. cerevisiae* cells in the range of 330 kHz, which correspond to rigid body oscillations of the cell (29). These high-frequency ranges are too high to be measured with our optical microscopy setup. Therefore, the

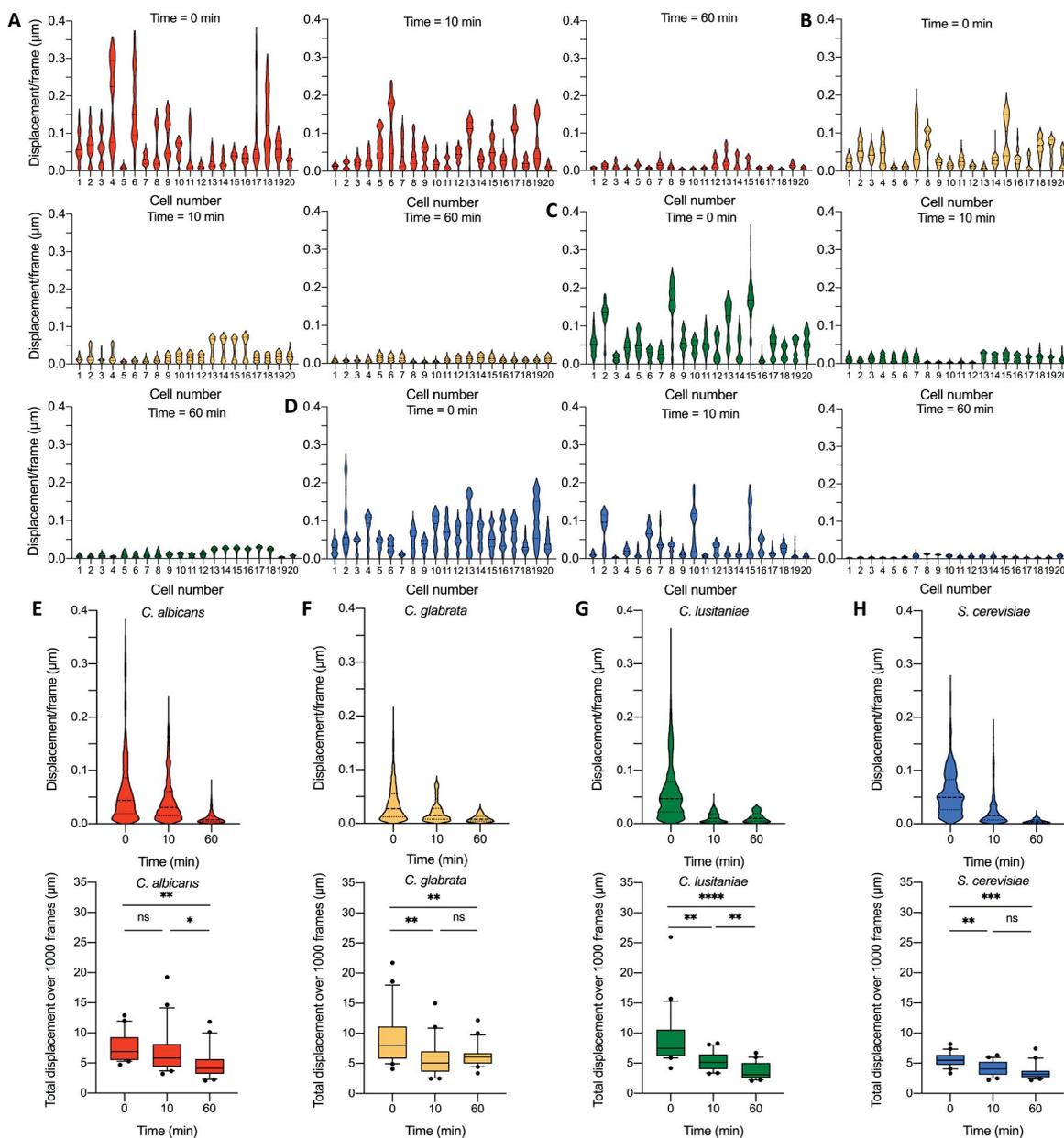


Fig. 4. Life-death transition by observing cellular nanomotions of yeast cells in the presence of ethanol. The distribution of the displacements of 20 cells at time 0 min (left) and 60 min (right) for (A) *C. albicans* DSY294, (B) *C. glabrata* DSY562, (C) *C. lusitaniae* DSY4606, and (D) *S. cerevisiae* BY4742. Time evolution of displacements/frame per frame of 20 cells (top) for (E) *C. albicans* DSY294, (F) *C. glabrata* DSY562, (G) *C. lusitaniae* DSY4606, and (H) *S. cerevisiae* BY4742 and the corresponding graphs of the total displacement during 12 s measurement as a function of time (bottom). Wilcoxon test: **** $P < 0.0001$; *** $P < 0.001$; ** $P < 0.01$; * $P < 0.1$.

low-frequency oscillations that we observed probably correspond to the whole-body displacements of single yeast cells. Future experiments involving high-speed optical microscopy and AFM-based measurements should highlight the full spectrum of cellular oscillations and a possible contribution of low-frequency cell wall oscillations.

The molecular processes that could cause the observed oscillations have not been investigated yet in detail. Additional experiments consisting in blocking or activating molecular actors (processes) would permit to better understand the observed phenomena. The nanomotion signal is made of vibrations arising from many metabolically related sources that combine energy consumption with local movement

or molecule redistributions (30). Cellular nanomotion could arise from processes such as DNA replication, DNA transcription, protein assembly, cytoskeleton rearrangement, ionic pumps activity, organelle transport, etc. The involvement of the cytoskeleton has already been demonstrated by depolymerizing the actin cytoskeleton of osteoblasts by cytochalasin, which resulted in a reduced cellular motion (as measured by the AFM cantilever method) (2). In addition, conformational changes of proteins [as was demonstrated for human topoisomerase II (31)] could contribute to nanomotion.

By automatizing the cell recognition using a deep learning algorithm, we could avoid manual cell detection, extend the number of

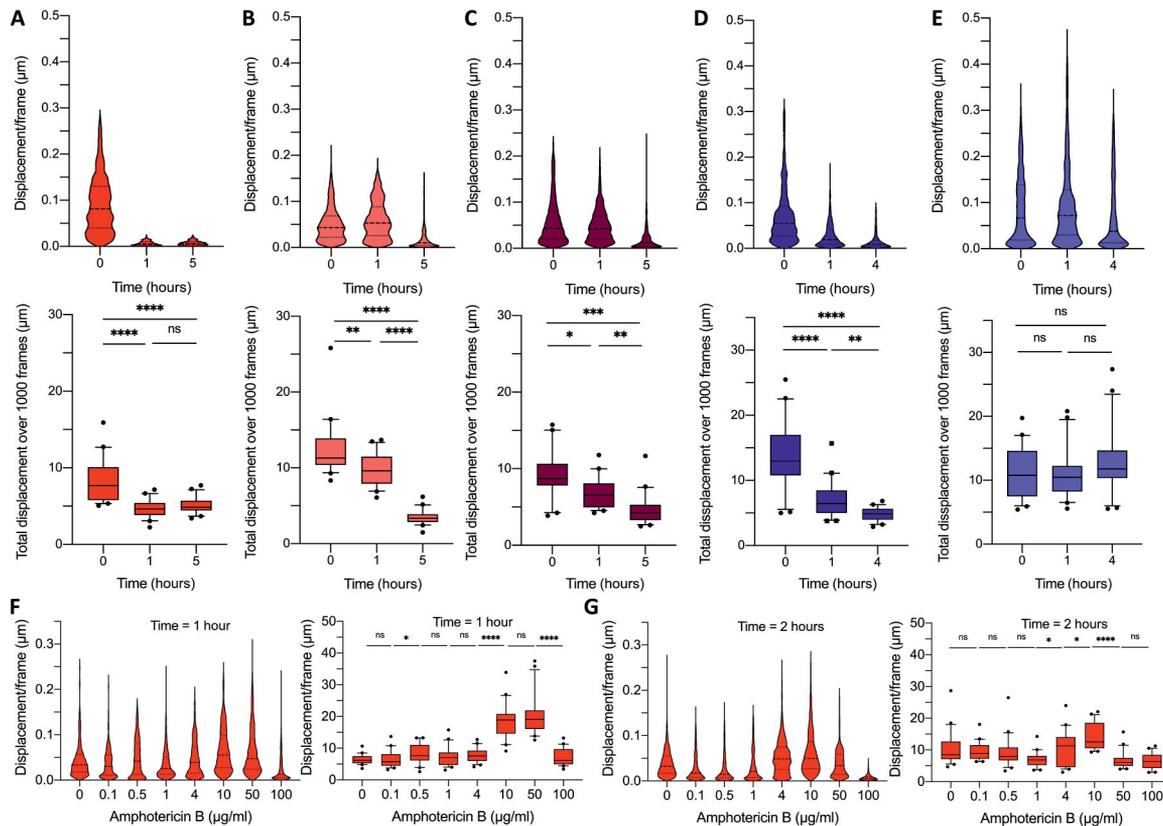


Fig. 5. Effect of antifungals on the cellular nanomotion of *C. albicans*. (A) Effect of amphotericin B (500 µg/ml), (B) caspofungin (100 µg/ml), and (C) fluconazole (400 µg/ml) on *C. albicans* DSY294. Time evolution of displacement distributions of 20 cells of (top) and corresponding graphs of the total displacement during 12 s (bottom). Effect of caspofungin (10 µg/ml) on (D) the hypersusceptible *C. albicans* DSY1024 and (E) the candid-resistant *C. albicans* DSY4614. Time evolution of displacement distributions of 20 cells of (top) and corresponding graphs of the total displacement (bottom). (F) Effect of increasing amphotericin B concentrations on *C. albicans* DSY294: cellular displacements (left) and the total displacement for 20 cells (right) after 1-hour treatment and (G) after 2-hour treatment. Wilcoxon test: **** $P < 0.0001$; *** $P < 0.001$; ** $P < 0.01$; * $P < 0.1$.

analyzed cells, and reduce the processing time. In addition, this opens the way to analyze a larger population of cells. Future developments will include dedicated microfluidic chip development and software optimization to run the acquisition and/or the data processing steps onto a low-end computer. These developments could eventually lead to an easy operational mobile device that can be directly implemented in hospitals or even in remote doctor's practices in developing world countries where it will allow to perform antifungal susceptibility testing in the earliest possible treatment stage and make the appropriate decision for a personalized effective antifungal therapy.

MATERIALS AND METHODS

Strains and cell growth

All yeast strains (table S1) were cultured by inoculating 10 ml of YPD [yeast extract (10 g/liter), peptone (20 g/liter), and dextrose (20 g/liter)] medium with a colony from a YPD agar (20 g/liter) plate. The cultures were grown overnight in Erlenmeyer flasks [30°C and 200 revolutions per minute (rpm)]. The overnight cultures were 10- to 20-fold diluted in 5-ml YPD medium to obtain an optical density at 600 nm (OD_{600nm}) of 0.5 and were then allowed to grow in Erlenmeyer flasks for 1 hour at 30°C and 200 rpm. The cultures were further diluted afterward, depending on the cell concentration (OD_{600nm} value) to obtain an optimal number of cells for visualization.

Optical nanomotion experiment

Ten microliters of each yeast cell culture was dispensed in one of the microwells, using a 4 Well FulTrac micro-Insert (Ibidi, Germany) in an imaging micro-dish (Ibidi, Germany). The yeast cells were allowed to sediment for a period of 10 min before starting the measurement. The motion of cells was observed by taking movies of 1000 frames with a framerate of 84 frames per second (fps) using an electron multiplying charge-coupled device camera (Andor iXon, Oxford Instruments) using a Nikon TE-2000 microscope with a 40× objective. The petri dish was kept at 30°C using a microscope stage top incubator (Ibidi, Germany).

For the experiments where the cell activity in PBS [NaCl (8 mg/ml), KCl (0.20 mg/ml), Na_2HPO_4 (1.44 mg/ml), and KH_2PO_4 (0.24 mg/ml)] was compared to YPD growth medium, the overnight cultured cells were 1000-fold diluted in either YPD medium or PBS and immediately dispensed in the microwells. The nanomotions of yeast cells were measured every hour during 4 hours. For the experiments where the effect of the temperature on the metabolism was evaluated, the temperature inside the microwells was controlled by adapting the temperature of water (from a recirculating water bath) circulating around the microwells. The temperature was successively adapted from 13° to 20°, 25°, 30°, and lastly, 35°C. The yeast cells were allowed to adapt during 20 min to each temperature before measuring the nanomotion of the cells. For the experiments with glass surface treatment, the glass surface was coated with concanavalin A (2 mg/ml; Sigma)

or PLL-g-PEG (0.1 mg/ml; SuSoS AG, Switzerland) by incubating the glass surface for 30 min with the coating solutions.

The effect of ethanol on the viability of the cells was compared with cells grown in YPD medium. Therefore, first, cell nanomotion videos were recorded, and next (after approximately 1 min), 200 μ l of ethanol (70%, v/v) was added to the top chamber of the four micro-Insert wells. Videos were recorded every hour during a period of 5 hours. Caspofungin, amphotericin B, and fluconazole were used as antifungals to assess their effect on cell viability. Before starting the treatment with the antifungal, reference (no treatment) videos were recorded of yeast cells in YPD medium (time = 0 hours). Then (at time = 1 min), 200 μ l of a certain concentration of antifungal was added to the chamber (200 μ l) above the four microwells.

As abiotic reference particles, silica microbeads (monodisperse silica standard, Whitehouse Scientific) with a diameter of 3 μ m were used. The beads were dissolved in YPD medium, and the measurements were performed at 30°C.

Nanomotion detection software

The ONMD algorithm calculates the cell displacement for each frame and saves the trajectories of tracked cells as well as the root mean square of the displacement to a MS Excel file. The main part of the program is based on the algorithm of Guizar-Sicairos *et al.* (7), ported from MATLAB to Python in the open-source image processing library sci-kit image (32). The Python package nd2 reader (Verweij R, online: www.lighthacking.nl/nd2reader/) was used to import the videos in the ND2 Nikon format.

Deep learning cell detection

The described nanomotion analysis previously described starts from the position of each individual cell, which is currently provided through a manual selection of the bounding box of a cell in the first video frame. However, the number of videos and cells present in the videos can be considerably large, especially when determining the MIC or the impact of different temperature conditions. Providing cell detection at every few frames instead of only in the first video frame allows rectifying the position of those cells that drifts away from their initial positions. In these cases, manual detection efforts could span over several hours for multiframe annotations and thus must be replaced by an automatic detection process of the cells. Therefore, we decided to use a deep learning algorithm, i.e., a medium-sized YOLO (24), to automatize cell detection. The training process is performed using a set of 50,000 synthetic cell images randomly generated. These synthetic images, obtained using a phase-contrast imaging model we previously proposed (33), look very similar compared to the cell images obtained with the microscope in terms of cells distribution, illumination, and imaging artefacts. Once the YOLO model has been trained, it is then used to automatically detect cells in real microscopic images, and each detection is then used as initial position to calculate the optical nanomotion with the cross-correlation algorithm previously described. The overall processing pipeline starts from a bulk of video sequences and performs a per-frame automated single-cell detection. Next, the results are refined by averaging the cells position and detection confidence across frames to correct abrupt changes between consecutive frames. Last, the position of cells detected with high confidence (i.e., >0.6) is provided as input to the nanomotion analysis algorithm (fig. S11A and movies S1 to S5).

Automation of the process

The overall process consists of reading a sequence of frames as images containing the cells and applying the above described cell de-

tection process every 10 frames to obtain a set of bounding boxes indicating the location of each individual cell automatically detected. On the basis of their positions, the cells are tracked across time as the video analysis process unfolds. At the end, in a subsequence refinement stage, we find the contours of cells (34) indicated by the bounding boxes to confirm that their center roughly corresponds to that of the bounding boxes, and any mismatch is then used to penalize the cells detection confidence proportionally. In addition, the bounding boxes position, and detection confidence is averaged across frames to avoid sharp changes the cells detection positions.

Calculation of the frequency region of optically recorded cell movements and the critical frequency

The approach based on calculating fast Fourier transform (FFT) amplitude spectrum, computed for two signals of cell movements—one in the horizontal (“x”) and the other in the vertical (“y”) direction—was developed to calculate the frequency region and critical frequency of optically recorded yeast cell movements. We named this method Double Region Interpolation Method. The idea is to perform two linear interpolations on these averaged amplitudes, in two different frequency regions. One should be performed on the wideband noise part of the spectrum, sufficiently distant from the cell activity region (e.g., 5 to 10 Hz). The other is performed on the low-frequency region, where the cell is active. Crossing point between these two straight lines determines the upper critical frequency of the cell activity region.

Our approach to determine the frequency range of the cell movements is based on the direct application of the FFT algorithm on two detrended signals: one obtained for movements in the horizontal (*x*) and the other in the vertical (*y*) direction. Both *x* and *y* spectra have similar profiles; however, some amplitudes were slightly different. We therefore derived their frequency-by-frequency average to equally incorporate frequencies of cell movements in both directions. Since signals were recorded with a sampling frequency of 83 fps, we visually inspected frequencies up to 41 Hz and confirmed that cell movements are confined to a few Hz only (fig. S8A, inset). As each signal contained 1000 samples (duration of 12.0482 s), each FFT spectrum was taken from the first 12 s, resulting in a frequency resolution of 0.083 Hz. To maintain this resolution, FFT was applied to each 12-s signal as a whole, thus avoiding any shorter moving windows. The task of determining the frequency range of cell movements implies assessing both its upper and lower critical frequencies.

Statistical analysis

Violin and box-and-whisker (10th to 90th percentile) plots were created with Prism8 (GraphPad). A Wilcoxon matched-pairs signed-rank test was performed to determine the significant differences between conditions (boxplots of mean total displacements over 1000 frames).

SUPPLEMENTARY MATERIALS

Supplementary material for this article is available at <http://advances.sciencemag.org/cgi/content/full/6/26/eaba3139/DC1>

REFERENCES AND NOTES

1. G. Longo, L. Alonso-Sarduy, L. Marques Rio, A. Bizzini, A. Trampuz, J. Notz, G. Dietler, S. Kasas, Rapid detection of bacterial resistance to antibiotics using AFM cantilevers as nanomechanical sensors. *Nat. Nanotechnol.* **8**, 522–526 (2013).
2. S. Kasas, F. S. Ruggeri, C. Benadiba, C. Maillard, P. Stupar, H. Tournu, G. Dietler, G. Longo, Detecting nanoscale vibrations as signature of life. *Proc. Natl. Acad. Sci. U.S.A.* **112**, 378–381 (2015).

3. K. Syal, R. Iriya, Y. Yang, H. Yu, S. Wang, S. E. Haydel, H.-Y. Chen, N. Tao, Antimicrobial susceptibility test with plasmonic imaging and tracking of single bacterial motions on nanometer scale. *ACS Nano* **10**, 845–852 (2016).
4. K. Syal, S. Shen, Y. Yang, S. Wang, S. E. Haydel, N. Tao, Rapid antibiotic susceptibility testing of uropathogenic *E. coli* by tracking submicron scale motion of single bacterial cells. *ACS Sens.* **2**, 1231–1239 (2017).
5. W. L. Johnson, D. C. France, N. S. Rentz, W. T. Cordell, F. L. Walls, Sensing bacterial vibrations and early response to antibiotics with phase noise of a resonant crystal. *Sci. Rep.* **7**, 12138 (2017).
6. C. R. Bermingham, I. Murillo, A. D. J. Payot, K. C. Balram, M. B. Kloucek, S. Hanna, N. M. Redmond, H. Baxter, R. Oulton, M. B. Avison, M. Antognozzi, Imaging of sub-cellular fluctuations provides a rapid way to observe bacterial viability and response to antibiotics. *bioRxiv*, 460139 (2018).
7. M. Guizar-Sicairos, S. T. Thurman, J. R. Fienup, Efficient subpixel image registration algorithms. *Opt. Lett.* **33**, 156–158 (2008).
8. J.-S. Lee, E.-H. Park, J.-W. Kim, S.-H. Yeo, M.-D. Kim, Growth and fermentation characteristics of *Saccharomyces cerevisiae* NK28 isolated from kiwi fruit. *J. Microbiol. Biotechnol.* **23**, 1253–1259 (2013).
9. M. Lemos-Carolino, A. Madeira-Lopes, N. Van Uden, The temperature profile of the pathogenic yeast *Candida albicans*. *Z. Allg. Mikrobiol.* **22**, 705–709 (1982).
10. D. Sanglard, Emerging threats in antifungal-resistant fungal pathogens. *Front. Med.* **3**, 11 (2016).
11. M. L. Zeuthen, N. Dabrowa, C. M. Aniebo, D. H. Howard, Ethanol tolerance and the induction of stress proteins by ethanol in *Candida albicans*. *J. Gen. Microbiol.* **134**, 1375–1384 (1988).
12. A. K. Pandey, M. Kumar, S. Kumari, P. Kumari, F. Yusuf, S. Jakeer, S. Naz, P. Chandna, I. Bhatnagar, N. A. Gaur, Evaluation of divergent yeast genera for fermentation-associated stresses and identification of a robust sugarcane distillery waste isolate *Saccharomyces cerevisiae* NGY10 for lignocellulosic ethanol production in SHF and SSF. *Biotechnol. Biofuels* **12**, 40 (2019).
13. M. Schiavone, C. Formosa-Dague, C. Elsztein, M.-A. Teste, H. Martin-Yken, M. A. De Morais Jr., E. Dague, J. M. François, Evidence for a role for the plasma membrane in the nanomechanical properties of the cell wall as revealed by an atomic force microscopy study of the response of *Saccharomyces cerevisiae* to ethanol stress. *Appl. Environ. Microbiol.* **82**, 4789–4801 (2016).
14. D. Stanley, A. Bandara, S. Fraser, P. J. Chambers, G. A. Stanley, The ethanol stress response and ethanol tolerance of *Saccharomyces cerevisiae*. *J. Appl. Microbiol.* **109**, 13–24 (2010).
15. C. M. Henderson, D. E. Block, Examining the role of membrane lipid composition in determining the ethanol tolerance of *Saccharomyces cerevisiae*. *Appl. Environ. Microbiol.* **80**, 2966–2972 (2014).
16. R. G. Willaert, Micro- and nanoscale approaches in antifungal drug discovery. *Fermentation* **4**, 43 (2018).
17. D. Sanglard, F. Ischer, O. Marchetti, J. Entenza, J. Bille, Calcineurin A of *Candida albicans*: Involvement in antifungal tolerance, cell morphogenesis and virulence. *Mol. Microbiol.* **48**, 959–976 (2003).
18. N. P. Wiederhold, J. L. Grabinski, G. Garcia-Effron, D. S. Perlin, S. A. Lee, Pyrosequencing to detect mutations in *FKS1* that confer reduced echinocandin susceptibility in *Candida albicans*. *Antimicrob. Agents Chemother.* **52**, 4145–4148 (2008).
19. R. A. Cordeiro, C. E. C. Teixeira, R. S. N. Brillhante, D. S. C. M. Castelo-Branco, M. A. N. Paiva, J. J. Giffoni Leite, D. T. Lima, A. J. Monteiro, J. J. C. Sidrim, M. F. G. Rocha, Minimum inhibitory concentrations of amphotericin B, azoles and caspofungin against *Candida* species are reduced by farnesol. *Med. Mycol.* **51**, 53–59 (2013).
20. D. Sanglard, F. Ischer, D. Calabrese, P. A. Majcherczyk, J. Bille, The ATP binding cassette transporter gene *CgCDR1* from *Candida glabrata* is involved in the resistance of clinical isolates to azole antifungal agents. *Antimicrob. Agents Chemother.* **43**, 2753–2765 (1999).
21. L. Vale-Silva, E. Beaudoin, V. D. T. Tran, D. Sanglard, Comparative genomics of two sequential *Candida glabrata* clinical isolates. *G3* **7**, 2413–2426 (2017).
22. L. Castrejón, K. Kundu, R. Urtasun, S. Fidler, Annotating object instances with a Polygon-RNN. *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2017, pp. 4485–4493.
23. O. Russakovsky, J. Deng, H. Su, J. Krause, S. Satheesh, S. Ma, Z. Huang, A. Karpathy, A. Khosla, M. Bernstein, A. C. Berg, L. Fei-Fei, ImageNet large scale visual recognition challenge. *Int. J. Comput. Vis.* **115**, 211–252 (2015).
24. J. Redmon, S. Divvala, R. Girshick, A. Farhadi, You only look once: Unified, real-time object detection. *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2016, pp. 779–788.
25. O. Ronneberger, P. Fischer, T. Brox, in *Lecture Notes in Computer Science* (Springer Verlag, 2015), vol. 9351, pp. 234–241.
26. J. Brajtburg, W. G. Powderly, G. S. Kobayashi, G. Medoff, Amphotericin B: Current understanding of mechanisms of action. *Antimicrob. Agents Chemother.* **34**, 183–188 (1990).
27. M. I. Villalba, P. Stupar, W. Chomicki, M. Bertacchi, G. Dietler, L. Arnal, M. E. Vela, O. Yantorno, S. Kasas, Nanomotion detection method for testing antibiotic resistance and susceptibility of slow-growing bacteria. *Small* **14**, 1702671 (2018).
28. A. E. Pelling, S. Sehati, E. B. Gralla, J. S. Valentine, J. K. Gimzewski, Local nanomechanical motion of the cell wall of *Saccharomyces cerevisiae*. *Science* **305**, 1147–1150 (2004).
29. B. Farzi, C. Cetinkaya, Micromechanical and surface adhesive properties of single *Saccharomyces cerevisiae* cells. *J. Phys. D Appl. Phys.* **50**, 375401 (2017).
30. A. C. Kohler, L. Venturelli, G. Longo, G. Dietler, S. Kasas, Nanomotion detection based on atomic force microscopy cantilevers. *Cell Surf.* **5**, 100021 (2019).
31. L. Alonso-Sarduy, P. De Los Rios, F. Benedetti, D. Vobornik, G. Dietler, S. Kasas, G. Longo, Real-time monitoring of protein conformational changes using a nano-mechanical sensor. *PLOS One* **9**, e103674 (2014).
32. S. van der Walt, J. L. Schönberger, J. Nunez-Iglesias, F. Boulogne, J. D. Warner, N. Yager, E. Goullart, T. Yu; scikit-image contributors, Scikit-image: Image processing in python. *PeerJ* **2**, e453 (2014).
33. M. Alioscha-Perez, R. Willaert, H. Tournu, P. Van Dijck, H. Sahli, in *Lecture Notes in Computer Science* (Springer, Berlin, Heidelberg, 2013), vol. 8259, pp. 25–32.
34. A. Trujillo-Pino, K. Krissian, M. Alemán-Flores, D. Santana-Cedrés, Accurate subpixel edge location based on partial area effect. *Image Vis. Comput.* **31**, 72–90 (2013).
35. W. A. Fonzi, M. Y. Irwin, Isogenic strain construction and gene mapping in *Candida albicans*. *Genetics* **134**, 717–728 (1993).
36. O. Marchetti, P. A. Majcherczyk, M. P. Glauser, J. Bille, P. Moreillon, D. Sanglard, Sensitive bioassay for determination of fluconazole concentrations in plasma using a *Candida albicans* mutant hypersusceptible to azoles. *Antimicrob. Agents Chemother.* **45**, 696–700 (2001).
37. A. T. Coste, M. Karababa, F. Ischer, J. Bille, D. Sanglard, *TAC1*, transcriptional activator of *CDR* genes, is a new transcription factor involved in the regulation of *Candida albicans* ABC transporters *CDR1* and *CDR2*. *Eukaryot. Cell* **3**, 1639–1652 (2004).
38. C. B. Brachmann, A. Davies, G. J. Cost, E. Caputo, J. Li, P. Hieter, J. D. Boeke, Designer deletion strains derived from *Saccharomyces cerevisiae* S288C: A useful set of strains and plasmids for PCR-mediated gene disruption and other applications. *Yeast* **14**, 115–132 (1998).

Acknowledgments: The Research Council of the Vrije Universiteit Brussel (Belgium) and the University of Ghent (Belgium) are acknowledged to support the Alliance Research Group VUB-UGhent NanoMicrobiology (NAMI). We also would like to thank R. Behrend from Geneva Observatory for help in developing the cells tracking software. **Funding:** This work was supported by the Belgian Federal Science Policy Office (Belspo) and the European Space Agency (ESA) PRODEX program (Yeast Bioreactor project). The Research Council of the Vrije Universiteit Brussel (Belgium) support the International Joint Research Group (IJRG) VUB-EPFL BioNanotechnology & NanoMedicine (NANO) and the Strategic Research Program (project SRP11-M3D2). Swiss members of the team were funded by the Swiss National Grants 200021-144321, 407240-167137, and CRSI15_173863; the Gebert Ruff Stiftung GRS-024/14; and NASA NNH16ZDA001N-CLDTCH. Serbian members of the team were funded by the grant 173017 from the Ministry of Education, Science and Technological Development of the Republic of Serbia. **Author contributions:** R.G.W. and S.K. conceived the idea for the research, designed the experiments, and supervised the writing of the manuscript. P.V.B. performed the optical nanomotion experiments. R.G.W., S.K., P.V.B., A.M., M.A.-P., D.S., and G.D. interpreted the nanomotion results and/or contributed to the optimization of the method. A.M. and S.K. introduced the cross-correlation method to analyze the data. A.M. developed the cross-correlation analysis program. R.G.W. and S.K. introduced the displacement/frame distribution analysis (violin plot analysis). R.G.W., S.K., and M.I.V. contributed to the violin plot analyses. A.M., P.V.B., K.R., A.K., and M.A.-P. contributed to the different steps of the data analyses. H.S. and M.A.-P. introduced the AI cell recognition/detection method. M.A.-P. wrote the cell detection program. K.R., D.B., and A.K. contributed via the analysis of the motion signal and related frequency analyses. All authors discussed the results and contributed to and commented on the manuscript. R.G.W. and S.K. conceived the overall project and supervised the research. **Competing interests:** R.G.W., S.K., A.M., P.V.B., M.A.-P., and H.S. are authors on a patent application filed by Vrije Universiteit Brussel and Ecole Polytechnique Fédérale de Lausanne (no. PCT/EP 2020/062143. Submitted April 2019). The authors declare that they have no other competing interests. **Data and materials availability:** All data needed to evaluate the conclusions in the paper are present in the paper and/or the Supplementary Materials. Additional data related to this paper may be requested from the authors.

Submitted 22 November 2019

Accepted 12 May 2020

Published 24 June 2020

10.1126/sciadv.aba3139

Citation: R. G. Willaert, P. Vanden Boer, A. Malovichko, M. Alioscha-Perez, K. Radotić, D. Bartolić, A. Kalauzi, M. I. Villalba, D. Sanglard, G. Dietler, H. Sahli, S. Kasas, Single yeast cell nanomotions correlate with cellular activity. *Sci. Adv.* **6**, eaba3139 (2020).

Utilizing deep learning tools to extend acquisition time and to design a label-free imaging analysis protocol

Traditional point-scanning confocal instruments are widely used to achieve high-resolution fluorescence images. Unlike camera-based confocal systems, the rastering nature of point scanners, which build images point by point rather than all at once, allows for digitization with scan zooming. The primary benefit of this method is that it takes advantage of the microscope objective lens's full numerical aperture, even at low magnification. Point-scanning confocals have the key abilities of oversampling Nyquist resolution even at low magnifications and have variable emission pinholes, whereas camera-based confocals generally cannot come close to Nyquist due to pixel size on the sensor, and usually have fixed-diameter emission pinholes.

With the advent of resonant scanning mirrors, point scanners have become more relevant for live imaging. Not only can they scan zoom and achieve high spatial resolution, but they can also image with ultrashort dwell times (typically 100 nanoseconds/pixel, or at least 20x shorter time than classic point scanners), minimizing photodamage and maximizing frame rates. Such short dwell times also mean fewer emission photons being collected in that small amount of time. This means new challenges: Not only do detectors have to be very sensitive, but other efforts such as increasing excitation power, line averaging, or increasing the pinhole size are usually employed to increase signal.

Spheroid imaging is an ideal target for low-magnification confocal due to the size and volume of these samples, and live spheroids require resonant scanning to achieve long timelapses. To further extend the number of frames that can be captured and to minimize the laser exposure to the sample, deep learning (DL) denoising can be employed.

In this example, spheroids were acquired as Z stacks with a Nikon A1R point-scanning confocal using a resonant scanner with a 0.1-micro-second (μ s) dwell time and a line average of 2x (frame rate = 15 fps at full frame). Laser power was minimized to <1mW and the confocal pinhole set to 1 Airy unit (AU) at 20x 0.95 numerical aperture (NA). These settings were determined by measuring the signal-to-noise ratio (SNR) of the sample, and adjusting the laser power, gain, and line averaging

until the sample had a SNR that was not limited by read noise (shot-noise limited). Though this SNR was statistically shot-noise limited, the output image quality was not sufficient for downstream analysis.

NIS-Elements **Denoise.ai** was then applied postacquisition to the shot-noise limited data sets to remove the Poisson shot-noise component of the images. This DL classifier is based on a pretrained convolutional neural network (CNN) that can identify and remove the noise component of the images. Since the noise from the confocal's gallium arsenide phosphide (GaAsP) detectors is primarily shot noise, the result images are shot-noise free (**Figure 1**).

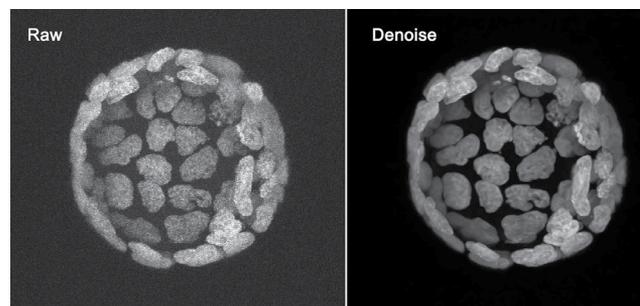


Figure 1: Maximum intensity projection of one-time point Z stack from a spheroid timelapse, demonstrating the effectiveness of applying DL denoising to resonant confocal images.

Applying denoising to recover SNR allows for significantly more acquisition frames than would normally be required to achieve adequate SNR. Normally, higher laser power and/or a trade-off in confocal Z resolution by opening the emission pinhole would be required, and as a result, fewer total frames could be acquired before the sample was bleached. With Denoise.ai, reduced excitation exposure allowed for over 10,000 frames to be acquired of each spheroid using a point-scanning confocal without significant bleaching.

Other DL tools available in NIS-Elements provided the opportunity to prepare a pilot study for spheroid imaging, which would allow the omission of acquiring the fluorescence channel entirely. NIS-Elements **Convert.ai** is a user-customizable trained CNN that allows users to provide data sets for DL and easily generate output classifiers to apply to subsequent images.

Michael A. Davis

Sr. Manager; Products & Technology

Nikon Instruments Inc., Melville, NY, USA
mike.davis@nikon.com

In this preliminary study, several spheroids were imaged over 10,000 frames in time and Z series, and Denoise.ai was applied to remove the Poisson shot noise. Two channels were collected: a differential interference contrast (DIC) channel and a fluorescence channel (which received the Denoise.ai processing). The goal was to train the Convert.ai CNN to recognize on DIC images where the corresponding fluorescence signal was present, so that only DIC images would subsequently be needed to be acquired in later experiments. Essentially, this was a test of using DL for a label-free imaging application.

Omitting fluorescence imaging would significantly impact the effective imaging lifetime and number of frames that could be acquired. This would allow for much more frequent time intervals and/or longer acquisition duration.

Nikon's NIS-Elements AI tools allow microscopists to apply DL to images in a simple interface such that anyone can apply DL to image data. In this case, several pairs of images from multiple spheroid samples were provided to the CNN: a DIC image and a fluorescent image (ground truth). Training was performed via NIS-Elements on a graphics processing unit-enabled graphics board for 1,000 iterations/epoch. Convert.ai then output a trained classifier that could be applied to any new data sets acquired under the same conditions, but requiring only DIC images.

Though the final goal of this study was to acquire only DIC images, the Convert.ai classifier was applied to newly acquired data sets in which the fluorescence label channel was also still captured, for purposes of validation of the DL result versus actual ground truth. The validation data reveals that DL was able to reconstruct the fluorescence channel adequately enough for the intended analysis purposes: to enable counting of cell number and to identify mitotic events over the timelapse. (Figure 2) DL did not reconstruct every nuance of the fluorescence

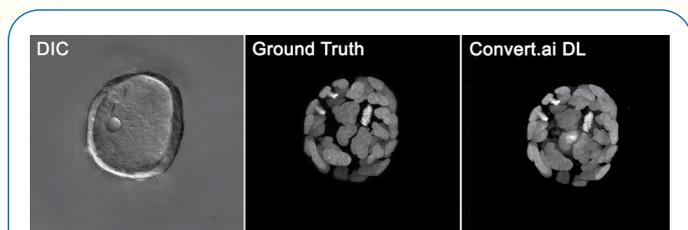


Figure 2: Spheroid DIC, ground truth fluorescence, and Convert.ai DL reconstructed channels shown as 2D projections. DL was able to adequately reconstruct the position and shape of the fluorescence label.

shape and texture, but for this study's analytical requirements these were not necessary. Only the number of objects and basic morphology were needed for purposes of counting and tracking.

One final analysis feature that was desired from the DIC data was the measurement of the spheroid dimension and shape using the center-plane DIC image. Further analysis interests were to compute areas occupied by proliferating, quiescent, or dead cells over time. Classic segmentation techniques are challenging with DIC images, due to embossing and shadowing. Intensity-based thresholding fails to effectively segment the spheroids, especially differentiating the outer proliferating cell layer from the rest of the spheroid volume.

DL was again applied to these images. A small number of spheroid images were hand-traced to define areas of interest, then an NIS-Elements CNN called **Segment.ai** was trained using both the DIC and reconstructed fluorescence channels, creating an output classifier generating a binary mask. When applied to other spheroid data where only DIC was acquired and the fluorescence channel reconstructed, these two channels together were used to reconstruct a segmentation mask over the DIC channel (Figure 3). This binary overlay was an accurate tracing that could further be easily analyzed for size and shape over many timelapses of multiple spheroids.

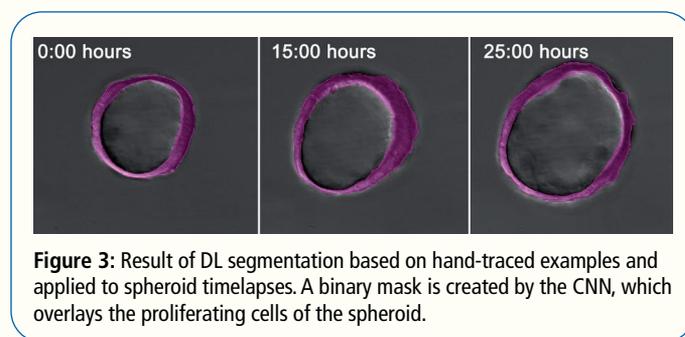


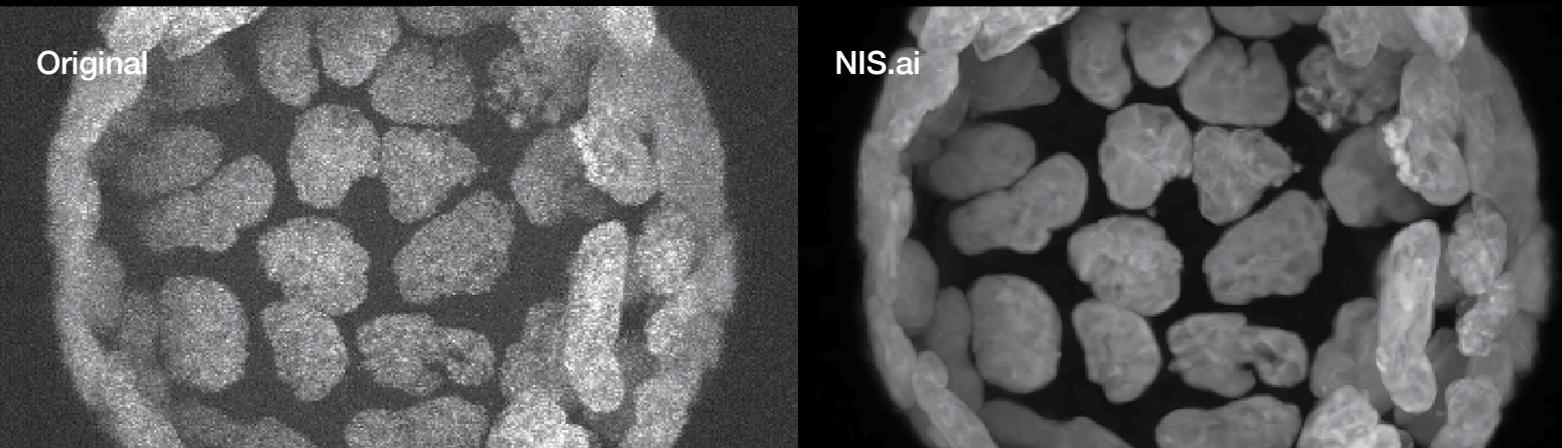
Figure 3: Result of DL segmentation based on hand-traced examples and applied to spheroid timelapses. A binary mask is created by the CNN, which overlays the proliferating cells of the spheroid.

This application illustrates three very powerful DL tools that can be used to completely change the capabilities of a microscope system: first, using more gentle acquisition settings while still achieving good SNR, allowing significantly more acquisition time points; second, using DL to reconstruct fluorescence such that only less-harmful DIC imaging is necessary for the experiment, further extending possible acquisition duration and/or frequency; and third, using DL to aid in the segmentation of DIC images for shape analysis, which is very difficult to perform using traditional intensity-based thresholding.



Introducing Deep Learning for **Image Acquisition**

NIS.ai expands Nikon's NIS-Elements microscope imaging platform by building in automated acquisition parameters powered by deep learning methods.



Microscopy images are never perfect, but suffer from noise, low signal, out-of-focus blur, and other imperfections. Nikon offers multiple NIS.ai modules, each approaching these challenges from a different direction in order to improve data quality and utility. This allows users to choose the approach that works best for their data, whether it be denoising, blur prediction and removal, or even trained feature prediction.

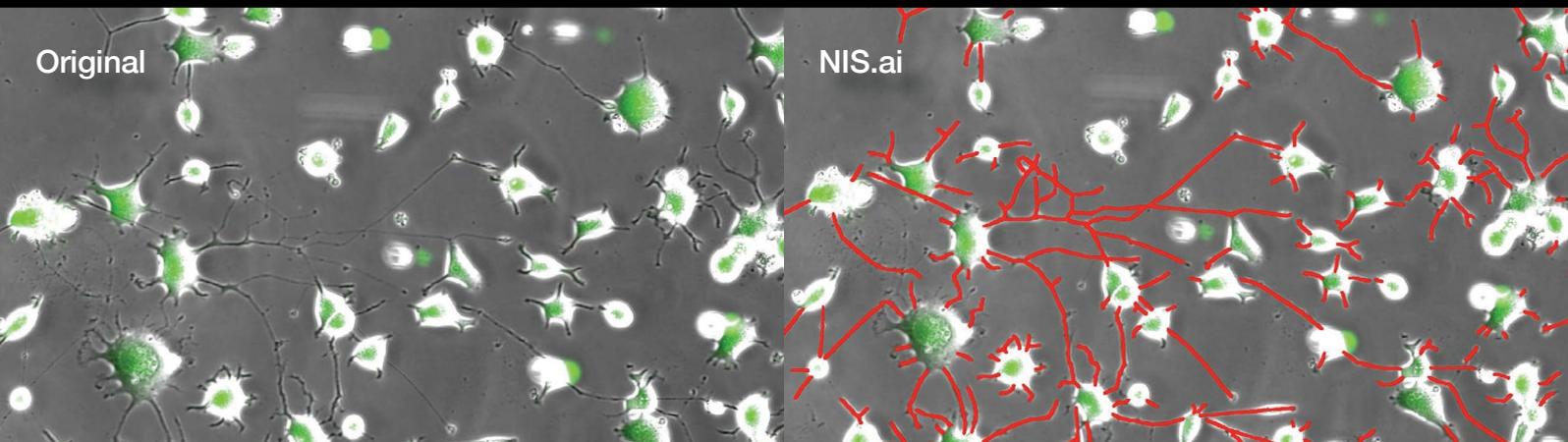
For more information, visit www.microscope.healthcare.nikon.com/nis-ai





Introducing Deep Learning for **Image Analysis**

NIS.ai expands Nikon's NIS-Elements microscope imaging platform with tools dedicated to simplifying previously complex or difficult analysis routines.



Today's data analysis tasks require intelligence. Classic numerical approaches such as intensity thresholding aren't always robust when considering complex image data. Deep learning allows for the characterization and consideration of numerous (and often subtle) image features in concert, rather than relying upon a single feature such as intensity. Tasks once best accomplished by hand, such as image segmentation, can be performed automatically.

For more information, visit www.microscope.healthcare.nikon.com/nis-ai

