Title: Heterogeneous chromatin mobility derived from chromatin states is a determinant of genome organisation in *S. cerevisiae*

Authors: Sven A. Sewitz[1,2,3]†, Zahra Fahmi[1]†, Latifa Aljebali[1,3,4], Jeremy Bancroft[2], Otávio J. B. Brustolini[2,3,5], Hicham Saad[6], Isabelle Goiffon[6], Csilla Varnai[1], Steven Wingett[1,7], Hua Wong[8,9], Biola-Maria Javierre[1], Stefan Schoenfelder[1], Simon Andrews[7], Stephen G. Oliver[2,3], Peter Fraser[1], Kerstin Bystricky[6], and Karen Lipkow[1,2,3]*

Affiliations:

[1]Nuclear Dynamics Programme, Babraham Institute, Cambridge, UK

[2]Cambridge Systems Biology Centre, University of Cambridge, Cambridge, UK

[3]Department of Biochemistry, University of Cambridge, Cambridge, UK

[4]Biochemistry Department, College of Science, King Saud University, Riyadh, Saudi Arabia

[5]Agricultural Biotechnology Institute, Federal University of Viçosa-MG, Brazil

[6]Laboratoire de Biologie Moléculaire Eucaryote (LBME), Centre de Biologie Intégrative (CBI) Université de Toulouse, UPS, CNRS Toulouse, France

[7]Bioinformatics Facility, The Babraham Institute, Cambridge, UK

[8]Unité Imagerie et Modélisation, Institut Pasteur, Paris, France

[9]Université de Reims Champagne-Ardenne, France

†Contributed equally to this work.

*Correspondence to: karen.lipkow@babraham.ac.uk

Abstract: Spatial organisation of the genome is essential for the activation of correct gene expression profiles, but the mechanisms that shape three-dimensional genome organisation in eukaryotes are still far from understood. Here, we develop a new approach, combining bioinformatic determination of chromatin states, dynamic polymer modelling of genome structure, quantitative microscopy and Hi-C to demonstrate that differential mobility of yeast chromosome segments leads to self-organisation of the genome in three dimensions. More than forty percent of chromatin-associated proteins display a poised distribution and coordinated relocations. They are distributed heterogeneously along the chromosome, and by simulating the dynamics of this heteropolymer, we observe structural features that match our experimental results. Further, we show that this mechanism directly contributes to the directed relocalisation of active genes to the nuclear periphery.

One Sentence Summary:

Unequal protein occupancy and chromosome segment mobility drive 3D organisation of the genome.

Main Text: Eukaryotic genomes are highly organised in three dimensions (*1*, *2*) and this spatial organisation has to be maintained in order to achieve the correct gene expression profiles (*3–6*). The 3D organisation of the genome is thus central to many aspects of cell biology and has been intensely investigated during normal growth (*7–9*), differentiation (*10–12*), cell division (*13*), senescence (*14*), and disease (*5*, *6*, *15*). In the budding yeast *Saccharomyces cerevisiae*, the target genes of most transcription factors are enriched in specific regions along the chromosome in one dimension (*16*), or in the genome in three dimensions (*17*). A central question in the field is by which mechanisms this 3D organisation is achieved.

Any mechanism that organises genome structure has to do so in a highly dynamic nucleoplasm (*18*, *19*). The prevalent view is that 3D genome organisation comes about despite the known fluctuations of the chromatin fibre. Most studies limit their focus on stable interactions between DNA-bound proteins that connect two chromatin loci (*9*, *20–26*). Here, we propose and validate a fundamentally different mechanism: The mobility of the chromatin fibre is not uniform along its length, but heterogeneous, matching the unequal protein binding along the genome. This leads to thermodynamically driven self-organisation, which we observe experimentally, and which we show to have important functional implications.

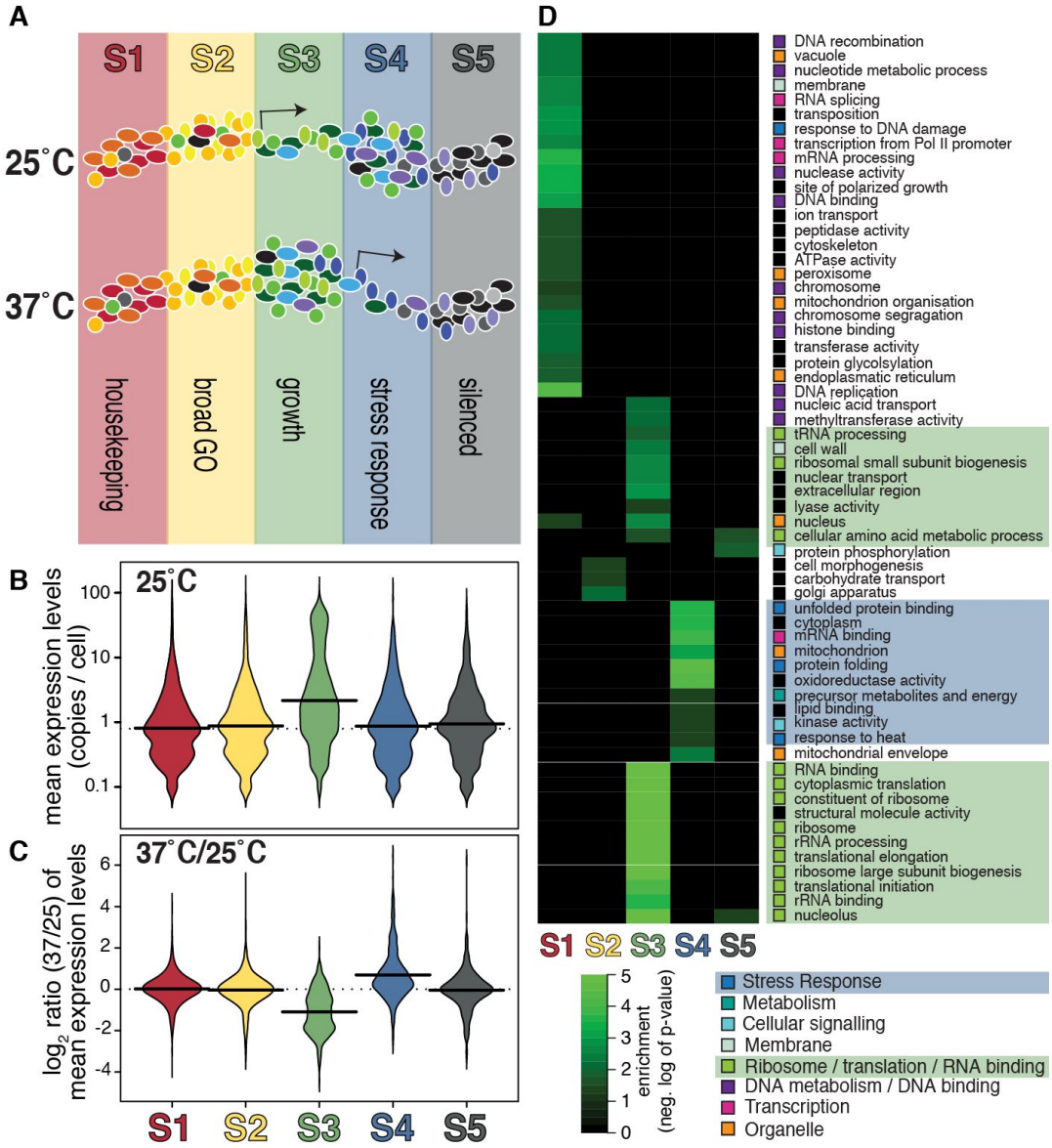
Determination and characterisation of chromatin states. In order to analyse the global effects of the protein binding on the dynamics of yeast chromosomes, and incorporate these data into a computational model, we first determined chromatin states in yeast. Chromatin states (also named 'chromatin colours') are an important conceptual advance in the field of chromatin biology. Here, chromatin modifications (*27*), chromatin-associated proteins (*28*), or a

combination thereof (*29*, *30*) are functionally categorised into groups or states, giving a chromatin-centric annotation of the genome. The resulting chromatin states were shown to correspond to differences in transcriptional activity, including the developmental regulation of genes, features that had not been used to define the states (*28–30*). Recently, chromatin states have also been linked to 3D genome organisation (*20*, *21*, *31–33*).

To determine chromatin states for the budding yeast *Saccharomyces cerevisiae*, we have modified the method of (*28*) to employ yeast chromatin immunoprecipitation (ChIP) data as input (*34*). This method determines chromatin states from quantitative protein binding data alone. As input, we used the reported genome-wide binding profile of 201 chromatin-associated proteins, measured in cells grown at 25°C, and 15 minutes after shifting the culture to 37°C (heat-shock) (*35*) (Fig. 1, Fig. S1). Five states effectively differentiated the protein binding profiles between the states (see SI for more details). The same procedure was performed independently for the 25˚C and 37˚C data. The genes in each state were counted and the states were numbered S1-S5 according to their decreasing frequency at 25˚C (Fig. S1D, Fig. S2A).

Transcriptional properties of chromatin states. At 25°C, S3 contains more highly expressed genes and significantly higher median expression than all other states (Fig. 1B). At 37°C, expression is reduced by 53% in S3, and increased by 62% in S4, while expression of genes in S1, S2 and S5 remains largely unchanged (Fig. 1C; Table S2, column 5).

S1 harbours categories typical of housekeeping genes (*36*) (Fig. 1D). Genes in S2 show very limited GO enrichment, indicating that these genes are distributed over many functional categories. Genes assigned to S3 cover the diverse processes and functions necessary for maintaining high levels of translation during rapid growth. Genes in S4 include functions such as unfolded protein binding, response to heat, and protein chaperones; all typical for heat-shock response genes. S5 is enriched for protein phosphorylation, amino acid metabolism and the nucleolus; it also harbours genes that code for proteins located at telomeres, such as the telomere-binding protein Cdc13p (*37*). These ontologies match the change in expression profile observed upon heat-shock, as genes required for translation (S3) are known to be highly expressed under favourable conditions and to be repressed when the cells are stressed (*38*, *39*)(*39*), while heat-shock genes (S4) are expected to be upregulated at 37°C (Fig. 1A-C).



[Fig1]

Fig. 1. Chromatin states in *S. cerevisiae* and their characteristics. (A) Overview: We defined chromatin states using statistical methods, from the combination of chromatin-associated proteins bound (coloured ellipses) (see Fig. S1). The states were ordered by genome coverage, named S1 to S5, and assigned a colour (red, yellow, green, blue, grey). At 25°C, the genes in S4 have a greater amount of proteins bound, while genes in S3 have higher expression (kinked arrow). At 37°C, this is reversed, with genes in S3 showing higher protein occupancy, and S4 genes higher expression. A high level gene ontology analysis per state is shown. **(B)** Expression analysis of chromatin states from yeast grown at 25˚C. Thick horizontal lines: mean; dashed line: total average. At 25°C, S3 (green) has a higher number of highly expressed genes ($p<2.2e-16$); expression array data from (*40*). **(C)** After shift to 37°C, the ratio of expression (37°C/25°C) shows that genes in S3 are repressed, while those in S4 are upregulated; expression array data from (*41*). **(D)** Gene Ontology enrichment analysis of genes in each state, using Ontologizer (*42*).

Protein occupancy and poising. We investigated the distribution of proteins across chromatin states. For each protein, we applied a binary hidden Markov Model (HMM) to the raw ChIP data, determining which loci were occupied (see (*34*)). We then quantified the fraction of loci of a given state that are occupied by that protein. We determined these 'fraction occupied' values for all states and all proteins, at both temperatures. Fig. S3 plots the values for all proteins that display at least a 20% difference in occupancy between any two states.

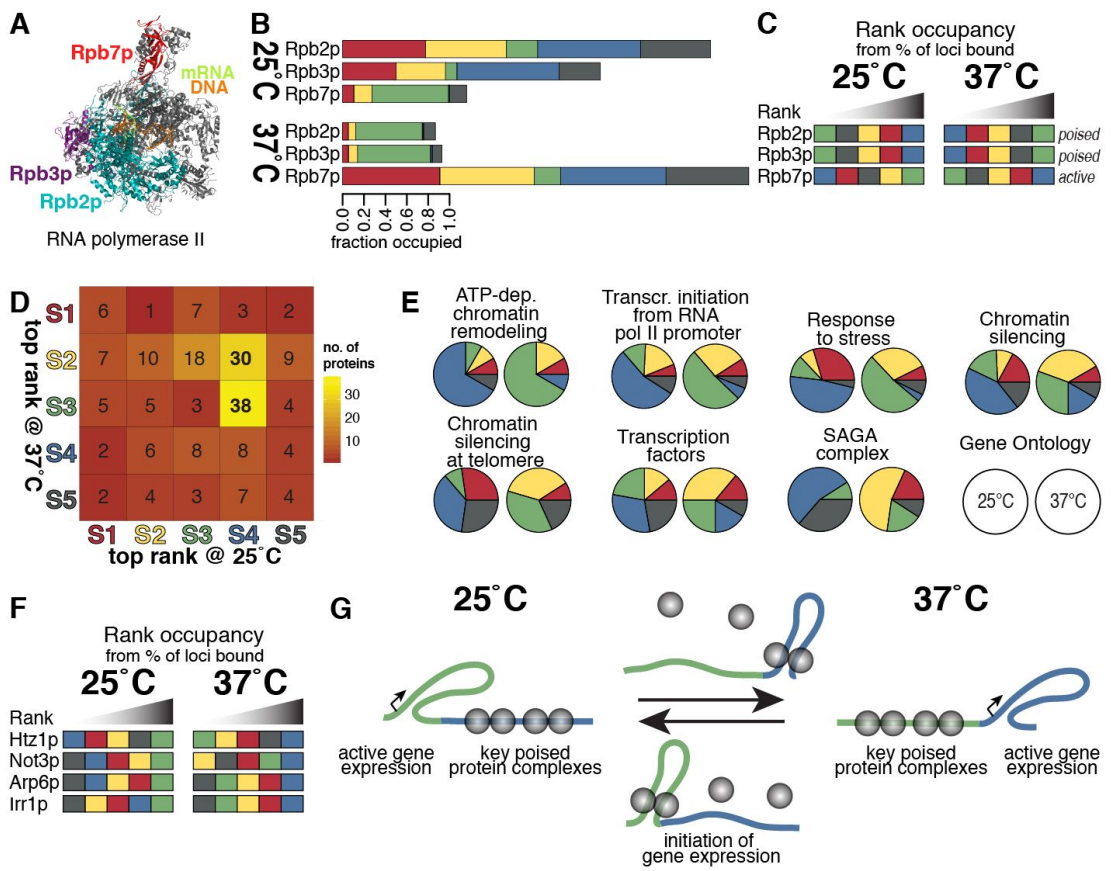
The three RNA polymerase II subunits present in the ChIP dataset, Rpb2p, Rpb3p and Rpb7p (Fig. 2A), display different binding patterns across chromatin states, and their fraction occupied values per state are shown in Fig. 2B. Interpretation is made easier when assessing the rank protein occupancy (Rank occupancy, Fig. 2C). At 25°C, Rpb2p and Rpb3p have the highest binding rank to S4 genes, and the lowest binding rank to S3 genes. The rank order for both these RNA polymerase subunits is identical, reflecting the close contacts they make within the RNA Pol II enzyme (Fig. 2A). After the shift to 37°C, both Rpb2p and Rpb3p relocate, and the ranking of S3 and S4 is reversed: now S3 is bound most (highest rank), and S4 the least (lowest rank) (Fig. 2C). The rank order of both Rpb2p and Rpb3p is again identical. This is in stark contrast to the rank occupancy for Rpb7p: At 25°C, the highest occupancy is in S3, and at 37°C in S4 (Fig. 2C). At both temperatures, the preferred location of Rpb7p coincides with the chromatin state which shows the highest level of expression (Fig. 1B). The Rpb2p and Rpb3p subunits, however, show a distribution that is termed 'poised': ready for immediate activation (*43*). This means that the highest levels of binding are to the state that becomes active under different temperature conditions. Note that here poised genes are defined by protein occupancy and not by histone modification (*44*).

The difference between subunit distributions (Fig. 2A) can be explained by the structural composition of the 12-subunit RNA Pol II enzyme. Rpb2p and Rpb3p make extensive contacts to each other and are part of the core enzyme complex, while Rpb7p binds only minimally to the core complex. Rpb7 (together with Rpb4) is in a diffusible subunit, known to be present in substoichiometric concentrations (*45*), and to participate in the stress response (*46*, *47*). RNA Pol II has been reported to be poised during stationary phase (*43*). Our analysis demonstrates that a large fraction of the main RNA Pol II complex is poised, both in exponential phase at favourable temperatures (25°C), and during heat stress (37°C), and that the location of the poised polymerase changes. Our results show that the presence of Rpb7p is linked to actively transcribed genes, coinciding with its described role during activation of transcription (*48*).

We investigated whether other proteins also showed a poised distribution. Of the 201 proteins tested, 68 proteins (42%) have their highest levels of occupancy in S4 at 25°C. At 37°C, these proteins bind strongest to S3 and S2, with 38 of those proteins binding with the highest rank to S3 (Fig. 2D). This indicates that significantly more proteins than previously described are poised, and that much of the previously described widespread movement of proteins upon heat-shock (*35*) involves a concerted migration from S4 to S3 (Fig. 1A, 2G).

Grouping the proteins according to molecular function shows that not all are poised to the same degree (Fig. 2E). ATP-dependent chromatin remodellers show the highest level of poising (66.7%, Fig. 2E), followed by the protein components responsible for transcription initiation from an RNA Pol II promoter (54.5% at 25°C, 51.2% at 37°C). Transcription factors show the lowest levels of poising (19.4%). Hence, we conclude that chromatin remodellers hold the gene in an activation-ready state, with transcriptions factors acting to trigger gene expression and binding at the time when the gene product is required.

Of the 201 proteins analysed, the rank occupancy of only five proteins that differ by more than 20% between any two states, correlated positively with transcriptional activity at both temperatures: Htz1p, Not3p, Arp6p, Irr1p and Rpb7p (Fig. 2C,F). All these proteins have well established roles at sites of active transcription (*49–55*). This validates our chromatin state analysis, and shows how remarkably few proteins consistently co-localise with the most active genes, since a significant proportion of the chromatin-associated proteins are located at other sites.



[Fig2]

Fig. 2. Wide-spread, consistent poising in the yeast genome. (A) Structure of RNA polymerase II in complex with DNA and RNA transcript. Crystal structure from PDB 3HOU (*56*). RNA Pol II is a complex of 12 proteins, Rpb1p to Rpb12p. The three proteins included in the analysed dataset (*35*) are

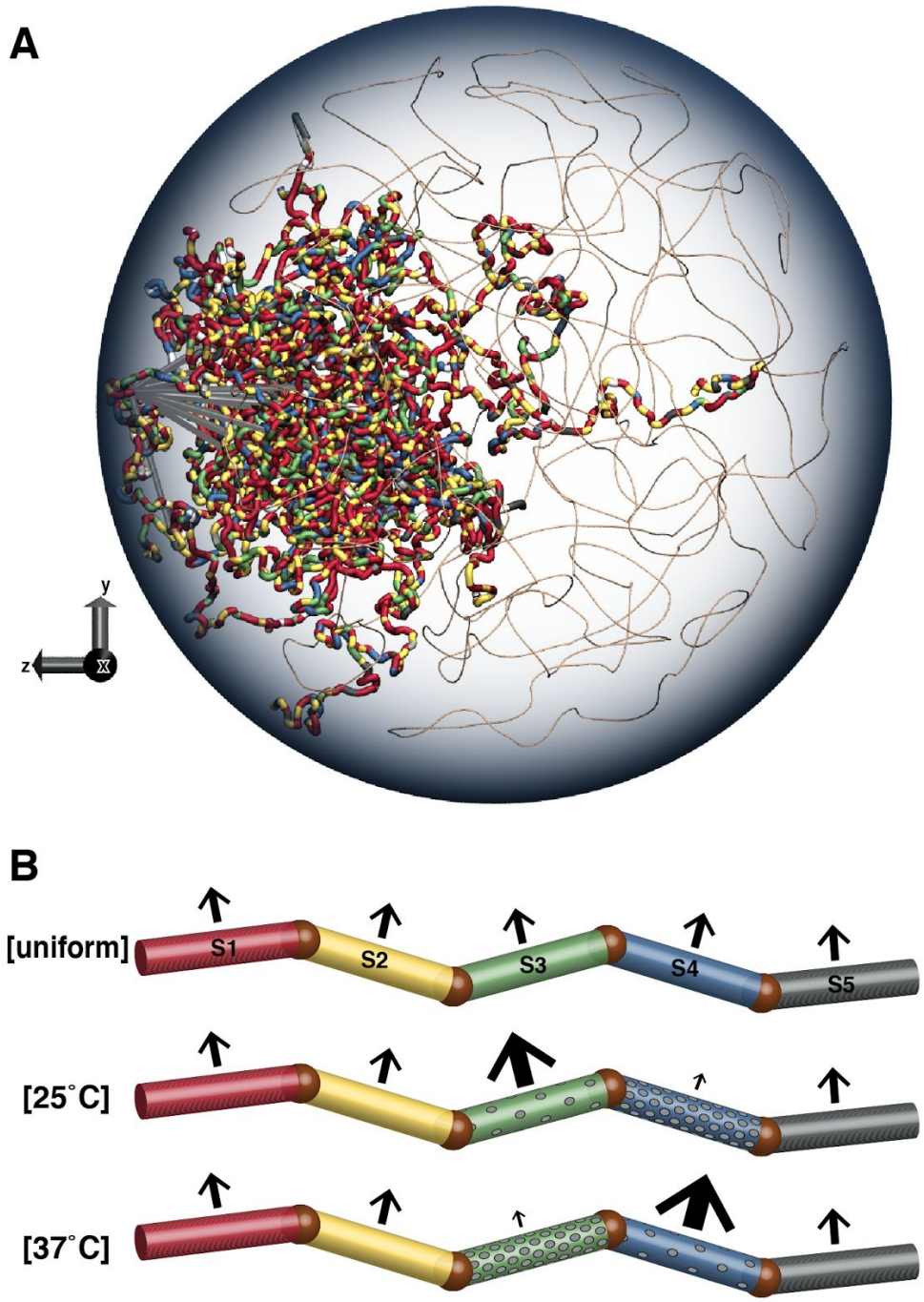
coloured: Rpb2p (teal), Rpb3p (purple) and Rpb7p (red). Note that Rpb7p is part of a small subcomplex sitting on top of the main complex. **(B)** Stacked barplots of the fraction occupied values for the three RNA Pol II subunits, for 25°C and 37°C (from Fig. S3). **(C)** Rank occupancy plot. At 25°C, both Rpb2p and Rpb3p bind most to genes in S4, indicated by the blue square to the right, while the Rpb7 subunit is bound most to S3. At 37˚C, this pattern is reversed. Rpb2p and Rpb3p are described as 'poised'. The binding location of Rpb7p coincides with actively transcribed genes ('active'). **(D)** Top-ranked binding state of all 201 proteins at 25°C and 37°C. **(E)** Poising across functional protein types. The pie charts give the number of proteins of each family that binds most to the indicated chromatin state. We here define a protein as being poised if its highest rank is S4 at 25˚C and S3 at 37˚C. **(F)** Proteins that are most bound to active genes at both temperature conditions (from those with ≥20% difference in occupancy, Fig. S3). **(G)** Schematic showing how gene expression (kinked arrow) and poising changes under different growth conditions.

Modelling Chromosome Dynamics. The significant differences in protein occupancy between chromatin states justifies modelling chromatin as a heteropolymer. We postulated that the heteropolymeric nature of chromatin affects the local mobility of each segment, leading to a distribution of segment mobilities of equivalent heterogeneity. In this context, the term mobility describes the average linear displacement of a genome segment at each time step, i.e. the instantaneous velocity of its Brownian motion. We further postulate that this heterogeneity affects the structure of the whole genome.

In order to investigate this new concept, we modified a previously developed and validated computational polymer model of the yeast genome (*57*) (Fig. 3A, Movie S1). In this model, each chromosome is a polymer of cylindrical segments, connected by ball joints, attached to the spindle pole body at the centromere, and constrained by the nuclear membrane (see further details in (*57*) and in the SI) (Fig. 3A). The mobility of genome segments is governed by Langevin dynamics, which is composed of the random force (Brownian dynamics), and the viscous drag (*58*). We call this the compound Langevin force, F_{LC}. We reduced the length of each segment to 2 kb, which is the approximate average length of a gene and its flanking sequences in *S. cerevisiae (59)*. We then assigned the corresponding chromatin state to each segment of the computational model (Fig. S4A). To simulate chromatin as a heteropolymer, we changed the forces that are applied to each segment in a state-wise manner. We chose the F_{LC} as a means to change the segments' mobility, encompassing all pertinent changes in physical properties.

We investigated the relationship between protein binding and the compound Langevin force in short test simulations (Fig. S5): Increasing the mass and radius (i.e. the protein occupancy) of a segment has the same qualitative effect on a segment's displacement per time step as decreasing the F_{LC}, and *vice versa* (compare matching columns in Fig. S5B with C). Thus, changing the F_{LC} has the desired effect of changing the mobility of genome segments in line with protein occupancy.

As shown by our state analysis of the protein binding data, the largest change in protein occupancy occurred from S4 at 25˚C to S3 at 37˚C (Fig. 2D, summarised in Fig. 2G). We therefore limited our analysis to reciprocally changing the forces applied to segments of these two states (Fig. 3B). The uniform (homopolymer) model, in which all segments have the same stochastic force applied, served as control.



[Fig3]

Fig. 3. Computational model of whole genome dynamics. (A) Snapshot of the model, with 2 kb cylindrical segments and state mapping. Centromeres are attached to short microtubules (straight grey tubes) at the spindle pole body. The segments shown as thin beige lines represent the rDNA repeats and form the nucleolus (thickness not drawn to scale). For state assignment and measurements of all segments, see Fig. S4A and Table S3. Following the convention of the literature on *S. cerevisiae* genome structures, this plot and subsequent figures are rotated such that the centromeres are shown on the left. **(B)** Each genome segment is assigned the appropriate chromatin state. To simulate different protein occupancies, the compound Langevin force (F_{LC}) that is applied to each segment is varied according to chromatin state. Representing the 25°C situation, F_{LC} is reduced for all S4 segments (small arrow) and increased for all S3 segments (large arrow). The F_{LC} is stochastic and applied to each segment in each dimension at every time step; for clarity, only one arrow per segment is drawn.

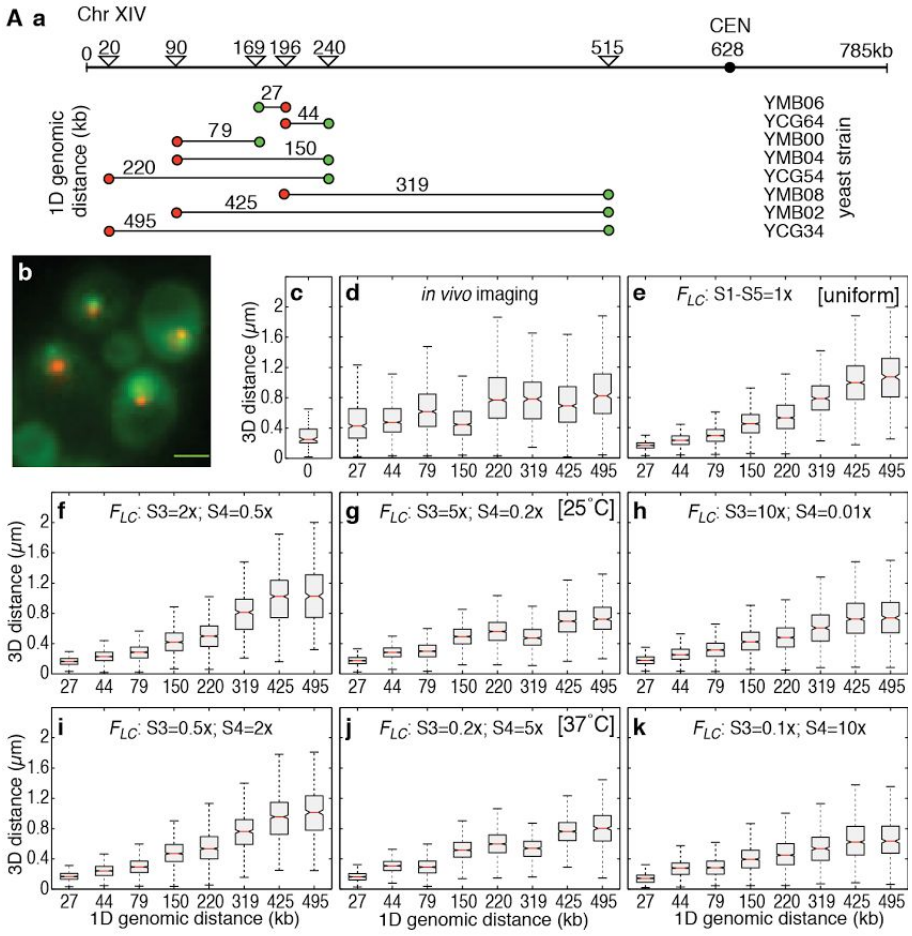
Experimental validation of the model: Microscopy. To experimentally validate our computational model, we first used live cell fluorescence imaging. We created a series of yeast strains in which two sites at genomic distances varying between 27 kb and 495 kb of the left arm of chromosome XIV are tagged by the lac and tet fluorescent operator systems (Fig. 4A(a,b), Table S4). Fluorescent signals are generated by GFP and mRFP fluorescently labelled lac and tet repressor proteins bound to their respective operator repeat sequences. Images were acquired in 3D (stacks of 21 images at z = 200 nm) and distances between the tagged sites determined using an automated ImageJ-based algorithm (*60*). Measured median distances ranged from 422 nm to 778 nm for loci separated by up to 220 kb. At greater genomic distances, the measured median did not increase further (Fig. 4A(d)). We compared these distance distributions with distributions obtained from simulations (Fig. 4A(e-k)). The uniform model predicted median distances between loci separated by 27 to 495 kb ranging from 166 nm to 1071 nm. For small genomic distances (27-79 kb), the simulated values were significantly smaller than the measured physical 3D distances due to experimental noise (Fig. 4A(c), (*61*). In contrast to the *in vivo* results, however, a uniform polymer model produces distances which increase monotonically with genomic separation without reaching a plateau (Fig. 4A(e)). Hence, we simulated distance distributions using a series of heteropolymeric models. We applied different forces F_{LC} to segments of different chromatin states. Strikingly, only the models in which the forces applied to the S3 and S4 segments are reciprocally changed 5-fold (Fig. 4A(g,j)) generate distance distributions that reach a plateau at genomic distances greater than 220 kb. Simulations based on 2-fold or 10-fold change in F_{LC} on any of the segments (Fig. 4A(f,i,h,k)) did not recapitulate the *in vivo* data.

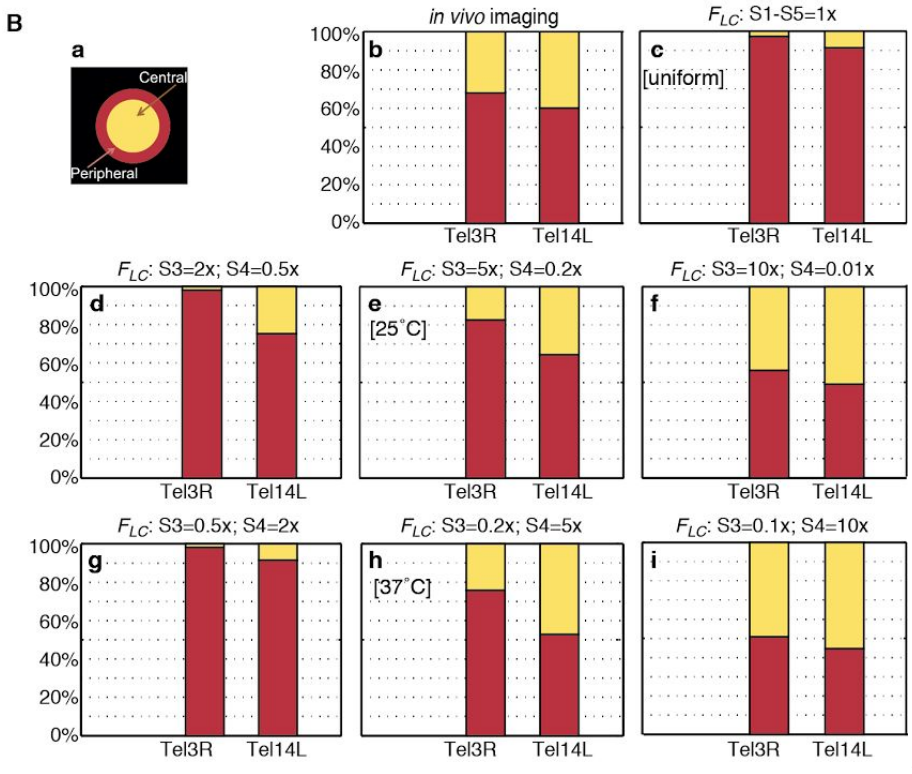
We next assessed whether our model could predict specific features of nuclear organisation. Because it was shown that budding yeast telomeres are preferentially, but not systematically located near the nuclear membrane(*18*, *62–64*), we recorded the frequency of a peripheral localisation of two different telomeres. Applying the uniform model, positions of the right telomere of Chr III (Tel3R) and the left telomere of Chr XIV (Tel14L) were assigned to the

most peripheral zone in >90% of the sampled time-steps (Fig. 4B(c), see SI for details). In contrast, simulations based on heteropolymeric models with 5x reciprocal changes in F_{LC} reduced this frequency to ~80% for Tel3R and ~65% for Tel14L (Fig. 4B(e,h)). Hence, the heteropolymer model was able to simulate the experimentally determined positions with greater accuracy than the uniform model (Fig. 4B(b))(*19*).

In both types of analyses, heteropolymers with F_{LC} changes of ≥5x resulted in a compaction of the genome. This is visible from the median 3D distances, which are reduced in comparison to the uniform simulation (Fig. 4A(g,h,j,k) vs. Fig. 4A(e)), and from the more central location of the telomeres (Fig. 4B(e,f,h,i) vs. Fig. 4B(c)). In both analyses, a reciprocal factor of 5 resulted in the best match to experimental data (Fig. 4A,B). Thus, all subsequently shown simulations used the combinations of 5x / 0.2x to represent 25˚C, or 0.2x / 5x to represent 37˚C, as F_{LC} scaling factors for S3 and S4 segments, respectively.

These results demonstrate that the fit between model and experimental data is markedly improved when simulating the chromatin fibre as a heteropolymer with differential F_{LC}. Heterogeneous mobility of chromatin segments is thus a plausible mechanism shaping chromosome conformation in yeast nuclei.


A a
Chr XIV
0 20 90 169 196 240 515 CEN 628 785kb
1D genomic distance (kb)
27 44 79 150 220 319 425 495
YMB06
YCG64
YMB00
YMB04
YCG54
YMB08
YMB02
YCG34
yeast strain
b
c
d in vivo imaging
e F_{LC}: S1-S5=1x [uniform]
f F_{LC}: S3=2x; S4=0.5x
g F_{LC}: S3=5x; S4=0.2x [25°C]
h F_{LC}: S3=10x; S4=0.01x
i F_{LC}: S3=0.5x; S4=2x
j F_{LC}: S3=0.2x; S4=5x [37°C]
k F_{LC}: S3=0.1x; S4=10x
3D distance (μm)
1D genomic distance (kb)


B
a
Central
Peripheral
in vivo imaging
b
c F_{LC}: S1-S5=1x [uniform]
d F_{LC}: S3=2x; S4=0.5x
e F_{LC}: S3=5x; S4=0.2x [25°C]
f F_{LC}: S3=10x; S4=0.01x
g F_{LC}: S3=0.5x; S4=2x
h F_{LC}: S3=0.2x; S4=5x [37°C]
i F_{LC}: S3=0.1x; S4=10x
100% 80% 60% 40% 20% 0%
Tel3R Tel14L

[Fig4]

Fig. 4. Validation of the model by microscopy. Position of labelled gene loci in confocal microscopy and in simulations. **(A)** Genomic loci at different distances from each other, and at different distances from the chromosome end, were labeled with FROS (fluorescent repressor operator system) on the left arm of Chromosome XIV. a) Positions of the insertions in kb, recognised by mRFP- and GFP-labelled fusion proteins. The distances between the insertion sites (27 kb to 495 kb) are shown above the lines connecting the red and green labels. The exact *genomic* distances between the centres of the inserted operator arrays are given in Table S4. b) Representative wide-field image, maximum projection of strain YCG54 ("220 kb"). Scale bar: 2 µm. c) Distribution of measured 3D distances in strain YPT237 (*61*), in which a single array is recognised by both fusion proteins: Control for experimental noise. d) Distribution of distances per yeast strain (n=498 to n=1233). e) Distribution of distances between equivalently mapped loci in uniform (homopolymer) simulations. f-k) Distribution of distances in heteropolymer simulations. **(B)** Telomere positions. a) The nuclei were divided into a central and peripheral zone of equal area. b) Position of labelled telomeres in live cells; n=80 (Tel3R); n=74 (Tel14L) (modified from (*19*)). c-i) Telomere positions in the simulated nuclei; n=6300 non-correlated time steps. Square brackets denote simulations with the parameter sets that were used to produce Figures 5-7.

DNA-DNA contacts captured by Hi-C and simulation. To gain detailed genome-wide insights into chromosome conformation, we performed Hi-C (*7*, *65–69*) in duplicate on yeast cultures (1) grown at 25°C, and (2) grown at 25°C and then shifted to 37°C for 15 min. This allowed us to experimentally determine the genome-wide chromatin contacts at the same growth conditions for which the chromatin states had been determined (Fig. S8). To ensure high quality of the data, we carefully optimised the Hi-C protocol for correct ligation junctions (Fig. S8A,B) and analysed the resulting data for the lack of bias between temperature conditions (Fig. S8C) and for reproducibility (Fig. S8D). We digested the DNA with *Hin*dIII, which in *S. cerevisiae* produces fragments of an average length of 2.7 kb. As there is no correlation between restriction sites and gene positions, we developed a pipeline to quantitatively map chromatin states to restriction fragments (Fig. S4B and SI text).

We mapped the Hi-C sequencing reads to the yeast genome, and filtered out experimental artefacts, PCR duplicates and physically linked segment pairs using the HiCUP pipeline (*70*). Across the four experiments, this resulted in 12 million valid, unique read pairs (Documents S4-S7). We then plotted and analysed the data at the resolution of individual restriction fragments, i.e. without binning (Fig. 5A). The full contact maps bear the hallmarks of those previously published for *S. cerevisiae*: Strong clustering of the centromeres, weaker clustering of telomeres, and only moderate enrichment of intra- versus inter-chromosomal contacts (*7*, *71*, *72*).

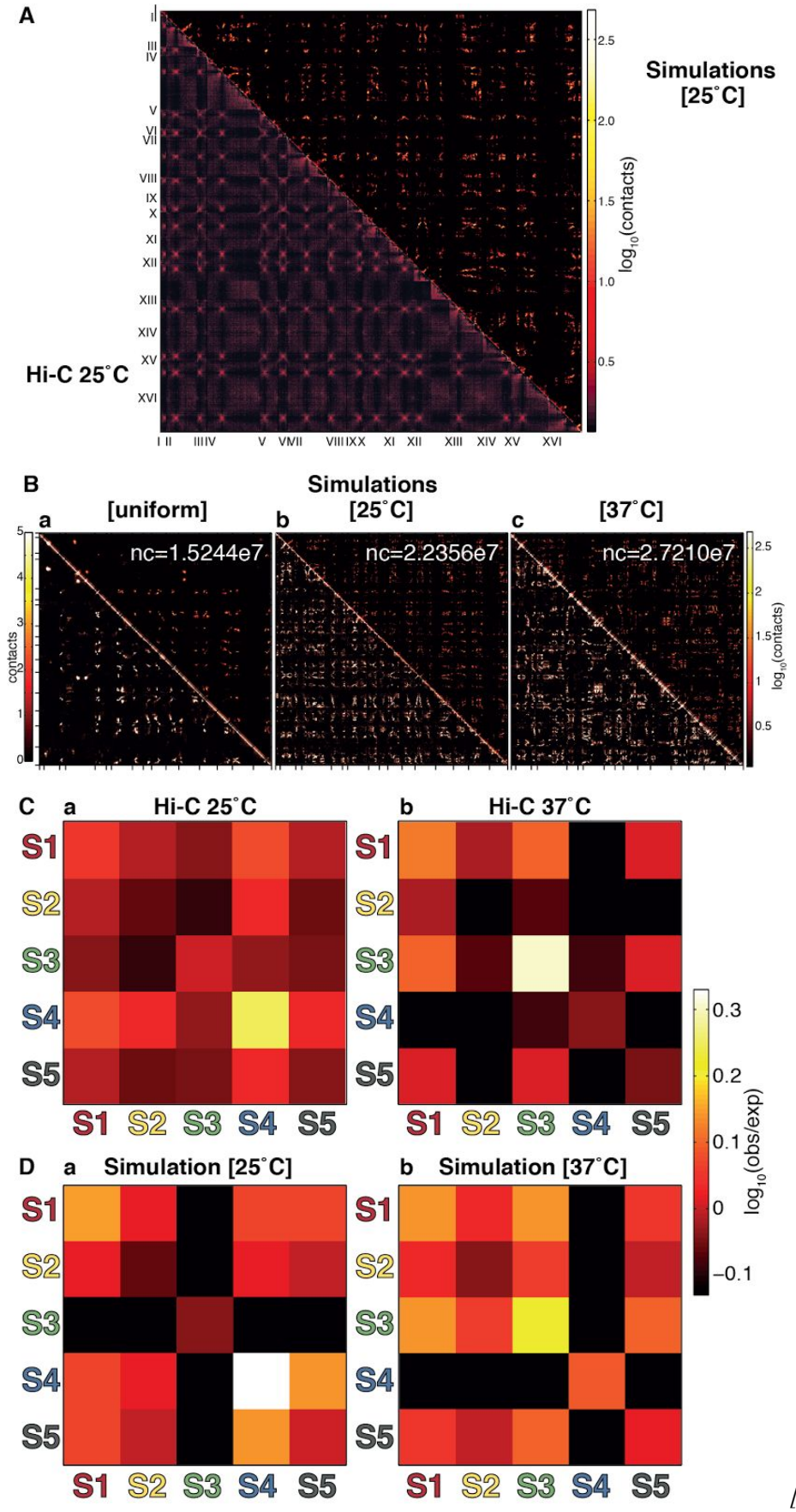
We analysed the simulations in an analogous manner, recording the incidence and position of contacts between segments throughout the simulation (see SI for details). Each of the three simulations gave rise to distinct patterns of contacts (Fig. 5B). The uniform model (Fig. 5B(a)) gave rise to markedly fewer contacts in the same number of time steps, confirming our observation that the heterogeneous mobility achieves a higher compaction of the chromosomes

(Fig. 4). As with telomere localisation (Fig. 4B(e,h)), the simulated 37˚C structure is even more compact (as indicated by the higher number of contacts) than the simulated 25˚C structure (Fig. 5B(b,c)).

To understand the influence of protein occupancy on genome organisation, we asked whether chromatin contacts are equally distributed across all chromatin states. After mapping chromatin states to *Hin*dIII fragments, we calculated state-wise contact maps for each growth and simulation condition (Fig. 5C,D and SI). The state-wise contact maps of the experimental and simulated data are remarkably similar in a number of important aspects. Of all state-wise contacts, the highest contact frequencies at 25˚C are between *Hin*dIII fragments (Fig. 5C(a)) or model segments (Fig. 5D(a)) of the S4 state. At 37˚C, the highest contact frequencies in each case are between fragments or segments in the S3 state (Fig. 5C(b), Fig. 5D(b)). At both temperatures and in both simulations, the highest number of contacts are intra-state in the state with the highest protein occupancy (as determined experimentally) and the lowest mobility (implemented as the lowest F_{LC} in the model). At the same time, the segments with low protein occupancy and high mobility (S3 at 25°C, S4 at 37˚C) have a moderate intra-state contact frequency but show clearly reduced contacts to all other chromatin states.

Interestingly, it is not the fast-moving segments that interact most frequently (Fig. 5D). Instead, it is the slow moving segments that have the highest number of contact, and our simulations indicate that this is a result of close spatial proximity (Fig. 6). In addition, the state-wise contact maps demonstrate that there is a clear spatial separation of the states with high protein occupancy and low mobility from those with low protein occupancy and high mobility, indicated by the 'black cross' of low contact frequencies, most clearly seen in Fig. 5C(b) and 5D(a,b).

To determine how well the simulated contacts matched the experimental results, we calculated Pearson correlation coefficients between the state-wise contact maps (Table 1): There is no significant correlation between the Hi-C or simulation data at different temperature conditions, which is an indication of significant rearrangements of genome structure that occur after heat-shock. However, there is high correlation between Hi-C and simulated data at equivalent conditions. These results lead us to conclude that the overall protein occupancy of a genome segment alters its mobility and thereby is a significant determinant of 3D genome organisation in yeast.


A
Simulations
[25˚C]
Hi-C 25˚C
log10(contacts)
B
Simulations
[25˚C]
a
[uniform]
b
c
[37˚C]
nc=1.5244e7
nc=2.2356e7
nc=2.7210e7
contacts
log10(contacts)
C
a
Hi-C 25˚C
b
Hi-C 37˚C
S1
S2
S3
S4
S5
log10(obs/exp)
D
a
Simulation [25˚C]
b
Simulation [37˚C]

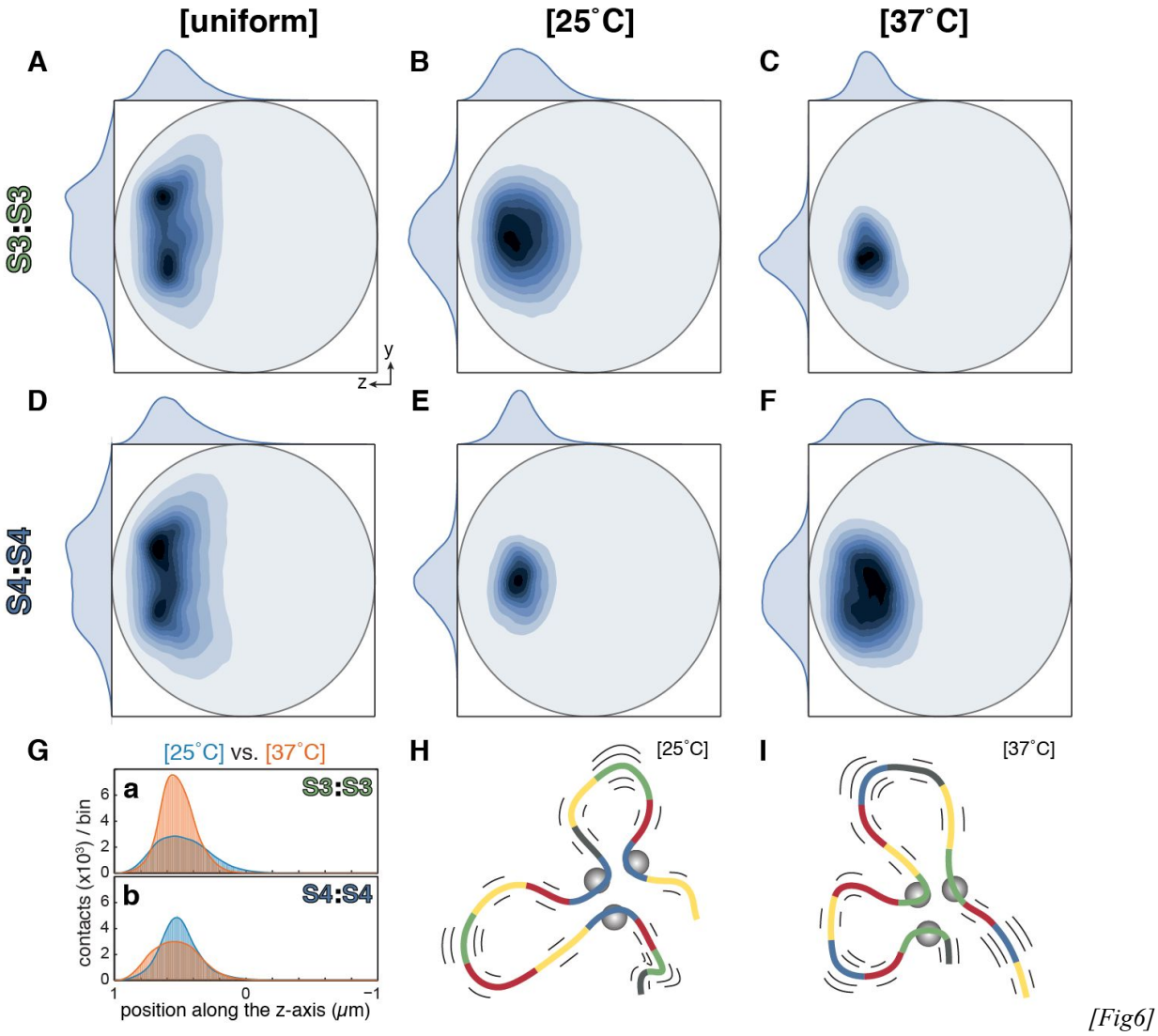
[Fig5]

Fig. 5. DNA-DNA contact maps of *S. cerevisiae* by Hi-C and simulation. (A) Full contact maps of Hi-C at 25˚C (below diagonal) and simulations representing [25˚C], sampled every 100 time steps (above diagonal). Both are drawn with single-fragment resolution. **(B)** Full contact maps of simulations at representing different setups, sampled every 100 time steps to exclude autocorrelation. nc, number of all contacts recorded in 590,000 simulated time steps. **(C)** State-wise contact maps of the Hi-C data, from both growth temperatures (25°C and 37°C), normalised by the number of expected contacts as calculated from the proportion of segments in each state. **(D)** State-wise, normalised contact maps of the simulated data, sampled every 100 time steps.

State-wise contact maps		Correlation coefficient	p-value	Significance
Hi-C 25˚C	Simulation [25˚C]	0.8328	0.000116	***
Hi-C 37˚C	Simulation [37˚C]	0.7181	0.0026	**
Simulation [25˚C]	Simulation [37˚C]	0.0654	0.8168	ns
Hi-C 25˚C	Hi-C 37˚C	- 0.0140	0.9606	ns
Hi-C 25˚C	Simulation [37˚C]	- 0.0861	0.7602	ns
Hi-C 37˚C	Simulation [25˚C]	- 0.1079	0.7020	ns

Table 1. Correlation of state-wise contact maps. Pearson correlation coefficients calculated in MATLAB between two vectors of size 15 (the numerical values of the upper-right triangular parts of matrices depicted in Fig. 5C,D). The data are sorted by correlation coefficient; ns, not significant.

Poised genes cluster in 3D. To determine where in the nucleus the simulated contacts occurred, we visualised the position of the intra-state-contacts in 2D projections (Fig. 6, Fig. S9). In all cases, the majority of the contacts are restricted to one half of the nucleus, due to the tethering of the centromeres to the spindle pole body, with the nucleolus occupying the opposite side (*57*). We quantified the level of clustering for all intra-state contacts (see Table S7, Table S8 and SI for details). In the uniform simulations, contacts are the most disperse, confirming the lack of compaction in the homopolymer (as in Fig. 4, Fig. 5B). In addition to this compaction, which affects the chromosomes in their entirety, we observe state-specific clustering of genes in 3D, where genes of one state, dispersed between all 16 chromosomes, preferentially co-localise. Between [25˚C] and [37˚C], antagonistic clustering occurs: Intra-S3 contacts cluster strongly at [37˚C], and intra-S4 contacts cluster strongly at [25˚C] (Fig. 6C,E,G), with a change of 47-fold and 18-fold respectively (Table S8, columns 5&6). For comparison, contacts in the entire genome change by only 1.3-fold (Table S8, last three rows). Preferential clustering is visualised in Fig. 6H,I: In both temperature conditions, the slowly moving, poised segments come together in dynamic clusters, located at a distance from the nuclear envelope, towards the centre of the nucleus.



[Fig6]

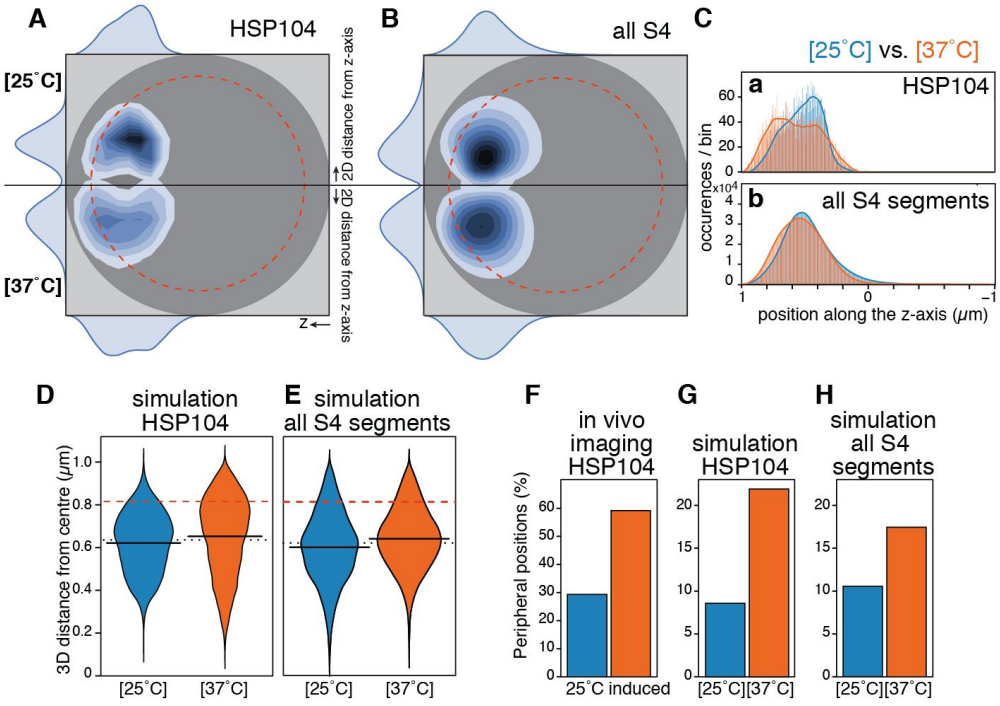
Fig. 6. Positions of intra-state contacts in the model. (A-F) 2D projection density plots of spherical yeast nuclei show the positions of simulated contacts between segments of the same state, in the same 590,000 time steps plotted in Fig. 5. Nuclei are oriented as in Fig. 3A, with the spindle pole body to the centre left. The distributions along the projected axes are given outside the box. The plots are normalised to the highest density in each dataset. (A-C) S3 intra-state contacts; (D-F) S4 intra-state contacts. (A,D) uniform model; (B,E) with [25°C] conditions; (C,F) with [37°C] conditions. **(G)** Quantitative histograms of the contacts along the z dimension, in 500 bins of 4 nm width. Profiles at [25˚C] (cyan) and [37˚C] (orange) are overlaid. They demonstrate the strong clustering of S3 contacts at 37˚C **(a)** and of S4 contacts at 25˚C **(b)**. **(H,I)** Schematic of how differences in mobility caused by increased binding of poised proteins at S4 segments at 25˚C (H), or S3 segments at 37˚C (I), lead to clustering in space. Grey spheres represent poised protein complexes; the number and length of the flanking black lines symbolise the degree of segment mobility.

Relocalisation of activated genes towards the nuclear periphery. In yeast, several genes have been described to relocate to the nuclear periphery upon activation and to interact with nuclear

pore complexes, aiding the export of mRNA into the cytoplasm (*73–78*). An example is the gene coding for the stress-induced disaggregase HSP104 (*79*, *80*), which moves to the nuclear periphery upon induction (*81*). Several molecular mechanisms have been proposed (*52*, *73*, *77*, *81–83*), but the phenomenon is not yet fully understood (*84*). Seeing that our models resulted in changes in location of S3 and S4 genes, we set out to test whether the changes in physical properties of the heteropolymer would suffice to deliver individual genes to the periphery.

We took the simulated positions of the segment corresponding to the *HSP104* gene at both temperatures and plotted the distribution of positions in radial density plots (Fig. 7A). These density plots show that, at 25°C, the gene is located within a small area and at distance from the nuclear periphery (Fig. 7A, [25°C]). At 37°C, the position of the gene has changed, with a broader distribution and moving closer to the periphery (Fig. 7A, [37°C]; Fig. 7C(a)). The number of times the gene was found within the peripheral area (outside of the dashed circle) more than doubled (Fig. 7D,G) ($p<2.2e-16$, Wilcoxon rank sum test). Confocal microscopy imaging had shown a very similar increase in peripheral location of the *HSP104* gene upon induction (*81*) (Fig. 7F).

This difference in location between the two temperature conditions can be seen in the collective shift of distribution of all S4 genes (Fig. 7B,C(b),E,H). The difference in peripheral location is highly significant (Fig. 7E,H; $p<2.2e-16$). Simulating chromosomes as heteropolymers with differential and changing mobility thus allows us to propose this as a general mechanism facilitating the relocation of activated genes to the periphery without the requirement for additional energy sources or specific factors.



[Fig7]

Fig. 7. Peripherality of heat-shock gene *HSP104* and all genes in S4 before and after induction. (A, B) Plots of the radial density around the z-axis of the simulated positions of the segment representing the *HSP104* gene, which belongs to S4, at 25°C (top) and at 37°C (bottom) throughout the recorded 590,000 time steps. The SPB is located at the center to the left. The peripheral zone, defined as covering the outermost 33% of the nuclear area, is indicated (dashed red circle). **(B)** The location of all S4 gene segments, plotted as in (A). The non-normalised distributions of *HSP104* **(a)** or all S4 segment positions **(b)** along the z-axis are overlaid for both simulated temperatures. Histograms calculated for 4 nm bins (vertical bars), and smoothed with Loess smoothing (thick lines), demonstrating the shift towards the periphery at 37˚C. **(D,E)** Distribution of the linear 3D distances of *HSP104* or all S4 segments from the nuclear centre. **(F)** Corresponding experimental values of *HSP104* (modified from (*81*)) of non-induced and induced cells. **(G,H)** Simulated data, quantified from (D,E).

Discussion

A growing body of evidence has shown that genome architecture is closely linked to gene expression and, in higher organisms, cellular differentiation (*11*, *85*, *86*). We now propose a novel mechanism in which chromatin states, which have different expression profiles, show distinct spatial organisation. In this mechanism, the organisation results from the heterogeneous dynamic behaviour of the chromatin fibre, which we derive from the known heterogeneous protein occupancy. Of note, it is not the activity of a small number of specific proteins that shape the overall genome organisation, but the cumulative effect of all chromatin-associated proteins on the mobility of each genome segment. Crucially, functional organisation is achieved even in the absence of explicit protein-protein interactions.

The definition of chromatin states allowed us to determine that a large proportion of chromatin-associated proteins, not just RNA polymerase subunits, are poised. This poising is visible at both growth temperatures, and involves rapid and coordinated movement of 42% of chromatin-associated proteins during the transition. At each temperature, the poised proteins are attached to those genes, which are currently less active, but which will need to be transcribed quickly if conditions change. It is easy to envision that this widespread poising greatly increases the cell's ability to react quickly, something that is of great benefit for a free-living, single celled organism encountering unbuffered and rapidly changing external conditions.

Our results demonstrate that, given the known architectural constraints of the *S. cerevisiae* genome, the differential mobility of chromatin segments is a significant determinant of overall 3D genome organisation. This requires a highly dynamic system, in which the chromosomes are in constant motion. The self-organisation we observe does not rely on protein-protein interactions, such as those mediated by cohesin (*87*) or CTCF in higher eukaryotes (*88*). Intriguingly, it has recently been shown that the major structural compartments in the genomes of mouse embryonic stem cells remain intact even when CTCF is depleted (*89*), indicating that

other, CTCF-independent, mechanisms are involved. These mechanisms remain to be elucidated and could include self-organisation by differential mobility.

The changes in dynamics we implemented were based on the observed changes in protein occupancy. We mapped the chromatin states onto a computational model that had been built and validated using experimental measurements (*57*). We then changed only one variable, the compound Langevin force F_{LC}. This approach is very different from using iterative fitting to change a large number of parameters, which can lead to impressive correlations to experimental data (*90*), but makes the interpretation of the biophysical basis of these parameters opaque. In contrast to this, heterogeneous segment mobility is more directly tangible.

The heterogeneous mobility that we infer has several important consequences. Firstly, it leads to an overall compaction of the genome, as seen by the position of the telomeres and 3D distances of chromosomal loci (Fig. 4) and the clustering analysis of contacts (Fig. 6, Table S7-S8) (the uniform conformation is always less compact than the two heteropolymeric ones). This is in addition to the compaction of the chromatin fibre achieved by nucleosomes, which silences and protects inactive stretches. The general compaction we observe keeps most of the genome at a distance from the nuclear envelope, leaving more of the peripheral region free for induced genes.

Secondly, our simulations reveal that changes in the distribution of segment mobilities lead to re-organisation of the genome. Specifically, they show that at 37°C, the S4 genes are located closer to the periphery. This represents an astonishingly simple mechanism to re-locate active genes to the nuclear periphery. We are aware that other specific interactions are also involved (e.g. Mex67p and nuclear pore complex (NPC) proteins (*82*, *91*)). Interestingly, the relocation of osmotic shock response genes to the nuclear periphery has been shown to occur even in the absence of nuclear pore complex proteins (*75*). This supports the contention that our mobility-based mechanism can act independently of known specific mechanisms. Together, the dynamics of self-organisation that stem from physics and the specific activities of biomolecules can achieve a directed and specific response.

Finally, the spatial clustering of the poised state segments is relevant in the context of the observed clustering of proteins that lead to the formation of transcription factories (*75*). While still speculative, this clustering offers the opportunity to help all proteins that are constituents of transcription factories to find each other prior to activation of transcription. We suggest that understanding changes in chromosome dynamics will shed new light on the mechanisms responsible for regulating gene expression.

Taken together, the results of our study demonstrate that by simply assigning different mobilities to chromosomal segments, biologically meaningful 3D organisation is achieved. This is more than a proof of concept: In our interdisciplinary approach, the assignment of states has been extracted from comprehensive experimental data sets, and the quality of the mobility

changes was derived from their further analysis. After careful mapping of the chromatin states to both a heteropolymer model and the genome, key predictions of the model were supported by a variety of experimental findings. Especially higher eukaryotes have evolved additional means to organise the genome, such as bridging of CTCF sites and attachment to nuclear pores. It remains to be quantified to which extent these mechanisms are redundant, refine and complement each other.

Acknowledgments: We are indebted to Jonathan Baxter, Stephanie Schalbetter, Job Dekker and Jon-Matthew Belton for sharing their unpublished yeast Hi-C protocol and advice. We thank Kristina Tabbada for sequencing, Charlotte Kaplan Galimow and Matthias Benoit for help with generating yeast strains, and Jonathan Houseley for advice. We thank Christophe Zimmer for sharing the code of their published yeast model, Pascal Carrivain for advice on Langevin equations and ODE, and Guillaume Filion for providing the HMMt package and advice. The Hi-C data is archived at the Gene Expression Omnibus (http://www.ncbi.nlm.nih.gov/geo/), accession number XXX.

We acknowledge funding by a Royal Society University Research Fellowship (KL), Microsoft Research Faculty Fellowship (KL), the Babraham Institute (KL, SAS), Islamic Development Bank (ZF), King Saud University (LA), CNPq (OJBB), EC FP7 contract 201142 (SGO), Biotechnology and Biological Sciences Research Council UK grant BBS/E/B/000C0405 (PF), MRC MR/L007150/1 (PF), ERC Advanced Grant DEVOCHROMO (PF), Agence Nationale pour la Recherche ANR *ANDY* (KB), Institut Universitaire de France (KB).

References:

1. F. W. Albert, L. Kruglyak, The role of regulatory variation in complex traits and disease. *Nat. Rev. Genet.* **16**, 197–212 (2015).
2. S. Schoenfelder *et al.*, Polycomb repressive complex PRC1 spatially constrains the mouse embryonic stem cell genome. *Nat. Genet.* **47**, 1179–1186 (2015).
3. D. G. Lupiáñez *et al.*, Disruptions of topological chromatin domains cause pathogenic rewiring of gene-enhancer interactions. *Cell*. **161**, 1012–1025 (2015).
4. M. Franke *et al.*, Formation of new chromatin domains determines pathogenicity of genomic duplications. *Nature*. **538**, 265–269 (2016).
5. W. A. Flavahan *et al.*, Insulator dysfunction and oncogene activation in IDH mutant gliomas. *Nature*. **529**, 110–114 (2016).
6. P. C. Taberlay *et al.*, Three-dimensional disorganization of the cancer genome occurs coincident with long-range genetic and epigenetic alterations. *Genome Res.* **26**, 719–731 (2016).
7. Z. Duan *et al.*, A three-dimensional model of the yeast genome. *Nature*. **465**, 363–367 (2010).
8. F. Jin *et al.*, A high-resolution map of the three-dimensional chromatin interactome in human cells. *Nature*. **503**, 290–294 (2013).
9. L. Giorgetti *et al.*, Predictive Polymer Modeling Reveals Coupled Fluctuations in Chromosome Conformation and Transcription. *Cell*. **157**, 950–963 (2014).

10. S. Schoenfelder *et al.*, The pluripotent regulatory circuitry connecting promoters to their long-range interacting elements. *Genome Res.* **25**, 582–597 (2015).
11. B. M. Javierre *et al.*, Lineage-Specific Genome Architecture Links Enhancers and Non-coding Disease Variants to Target Gene Promoters. *Cell*. **167**, 1369–1384 (2016).
12. J. R. Dixon *et al.*, Chromatin architecture reorganization during stem cell differentiation. *Nature*. **518**, 331–336 (2015).
13. N. Naumova *et al.*, Organization of the mitotic chromosome. *Science*. **342**, 948–953 (2013).
14. T. Chandra *et al.*, Global Reorganization of the Nuclear Landscape in Senescent Cells. *CellReports*. **10**, 471–483 (2015).
15. R. P. McCord *et al.*, Correlated alterations in genome organization, histone methylation, and DNA-lamin A/C interactions in Hutchinson-Gilford progeria syndrome. *Genome Res.* **23**, 260–269 (2013).
16. S. C. Janga, J. Collado-Vides, M. M. Babu, Transcriptional regulation constrains the organization of genes on eukaryotic chromosomes. *Proc. Natl. Acad. Sci. U. S. A*. **105**, 15761–15766 (2008).
17. S. Ben-Elazar, Z. Yakhini, I. Yanai, Spatial localization of co-regulated genes exceeds genomic gene clustering in the Saccharomyces cerevisiae genome. *Nucleic Acids Res.* **41**, 2191–2201 (2013).
18. P. Heun, T. Laroche, K. Shimada, P. Furrer, S. M. Gasser, Chromosome dynamics in the yeast interphase nucleus. *Science*. **294**, 2181–2186 (2001).
19. H. Hajjoul *et al.*, High-throughput chromatin motion tracking in living yeast reveals the flexibility of the fiber throughout the genome. *Genome Res.* **23**, 1829–1838 (2013).
20. A. M. Chiariello, C. Annunziatella, S. Bianco, A. Esposito, M. Nicodemi, Polymer physics of chromosome large-scale 3D organisation. *Sci. Rep.* **6**, 29775 (2016).
21. D. Jost, P. Carrivain, G. Cavalli, C. Vaillant, Modeling epigenome folding: formation and dynamics of topologically associated chromatin domains. *Nucleic Acids Res.* **42**, 9553–9561 (2014).
22. C. A. Brackley *et al.*, Predicting the three-dimensional folding of cis-regulatory regions in mammalian genomes using bioinformatic data and polymer models. *Genome Biol.* **17**, 1–16 (2016).
23. M. Barbieri *et al.*, A polymer model explains the complexity of large-scale chromatin folding. *Nucleus*. **4**, 267–273 (2013).
24. O. Kurkcuoglu, P. A. Bates, Mechanism of Cohesin Loading onto Chromosomes: A Conformational Dynamics Study. *Biophys J*. **99**, 1212–1220 (2010).
25. T. M. K. Cheng *et al.*, A simple biophysical model emulates budding yeast chromosome condensation. *eLife*. **4**, e05565 (2015).
26. M. Tark-Dame, H. Jerabek, E. M. M. Manders, D. W. Heermann, R. van Driel, Depletion of the Chromatin Looping Proteins CTCF and Cohesin Causes Chromatin Compaction: Insight into Chromatin Folding by Polymer Modelling. *PLoS Comput. Biol.* **10**, e1003877–12 (2014).
27. P. V. Kharchenko *et al.*, Comprehensive analysis of the chromatin landscape in Drosophila melanogaster. *Nature*. **471**, 480–485 (2011).
28. G. J. Filion *et al.*, Systematic Protein Location Mapping Reveals Five Principal Chromatin Types in Drosophila Cells. *Cell*. **143**, 212–224 (2010).
29. J. Ernst, M. Kellis, Discovery and characterization of chromatin states for systematic annotation of the human genome. *Nat. Biotechnol.* **28**, 817–825 (2010).
30. M. Kasowski *et al.*, Extensive Variation in Chromatin States Across Humans. *Science*. **342**, 750–752 (2013).
31. A. N. Boettiger *et al.*, Super-resolution imaging reveals distinct chromatin folding for different epigenetic states. *Nature*. **529**, 418–422 (2016).
32. C. A. de Graaf, B. van Steensel, Chromatin organization: form to function. *Curr. Opin. Genet. Dev.* **23**, 185–190 (2013).
33. J. Fraser *et al.*, Hierarchical folding and reorganization of chromosomes are linked to transcriptional

changes in cellular differentiation. *Mol. Syst. Biol.* **11**, 852 (2015).
34. S. Sewitz, K. Lipkow, *Systems Biology Approaches for Understanding Genome Architecture, in Chromosome Architecture*, M. C. Leake, Ed. (Springer New York, 2016), *Methods in Molecular Biology*, pp. 109–126.
35. B. J. Venters *et al.*, A Comprehensive Genomic Binding Map of Gene and Chromatin Regulatory Proteins in Saccharomyces. *Mol. Cell*. **41**, 480–492 (2011).
36. J. Parenteau *et al.*, Deletion of Many Yeast Introns Reveals a Minority of Genes that Require Splicing for Function. *Mol. Biol. Cell*. **19**, 1932–1941 (2008).
37. T. R. Hughes, R. G. Weilbaecher, M. Walterscheid, V. Lundblad, Identification of the single-strand telomeric DNA binding domain of the Saccharomyces cerevisiae Cdc13 protein. *Proc. Natl. Acad. Sci. U. S. A.* **97**, 6457–6462 (2000).
38. D. A. Lashkari *et al.*, Yeast microarrays for genome wide parallel genetic and gene expression analysis. *Proc. Natl. Acad. Sci. U. S. A.* **94**, 13057–13062 (1997).
39. M. H. Herruer *et al.*, Mild temperature shock affects transcription of yeast ribosomal protein genes as well as the stability of their mRNAs. *Nucleic Acids Res.* **16**, 7917–7929 (1988).
40. F. C. Holstege *et al.*, Dissecting the regulatory circuitry of a eukaryotic genome. *Cell*. **95**, 717–728 (1998).
41. S. J. Zanton, B. F. Pugh, Changes in genomewide occupancy of core transcriptional regulators during heat stress. *Proc. Natl. Acad. Sci. U. S. A.* **101**, 16843–16848 (2004).
42. S. Bauer, S. Grossmann, M. Vingron, P. N. Robinson, Ontologizer 2.0--a multifunctional tool for GO term enrichment analysis and data exploration. *Bioinformatics*. **24**, 1650–1651 (2008).
43. M. Radonjic *et al.*, Genome-wide analyses reveal RNA polymerase II located upstream of genes poised for rapid response upon S. cerevisiae stationary phase exit. *Mol. Cell*. **18**, 171–183 (2005).
44. B. E. Bernstein *et al.*, A bivalent chromatin structure marks key developmental genes in embryonic stem cells. *Cell*. **125**, 315–326 (2006).
45. G. J. Jensen, G. Meredith, D. A. Bushnell, R. D. Kornberg, Structure of wild-type yeast RNA polymerase II and location of Rpb4 and Rpb7. *EMBO J.* **17**, 2353–2358 (1998).
46. S. R. Singh, Domainal organization of the lower eukaryotic homologs of the yeast RNA polymerase II core subunit Rpb7 reflects functional conservation. *Nucleic Acids Res.* **32**, 201–210 (2004).
47. M. Choder, R. A. Young, A portion of RNA polymerase II molecules has a component essential for stress responses and stress survival. *Mol. Cell. Biol.* **13**, 6984–6991 (1993).
48. S. M. Orlicky, P. T. Tran, M. H. Sayre, A. M. Edwards, Dissociable Rpb4-Rpb7 Subassembly of RNA Polymerase II Binds to Single-strand Nucleic Acid and Mediates a Post-recruitment Step in Transcription Initiation. *J. Biol. Chem.* **276**, 10097–10102 (2001).
49. K. Yen, V. Vinayachandran, B. F. Pugh, SWR-C and INO80 chromatin remodelers recognize nucleosome-free regions near +1 nucleosomes. *Cell*. **154**, 1246–1256 (2013).
50. I. Albert *et al.*, Translational and rotational settings of H2A.Z nucleosomes across the Saccharomyces cerevisiae genome. *Nature*. **446**, 572–576 (2007).
51. M. A. Collart, O. O. Panasenko, S. I. Nikolaev, The Not3/5 subunit of the Ccr4-Not complex: A central regulator of gene expression that integrates signals between the cytoplasm and the nucleus in eukaryotic cells. *Cell. Signal.* **25**, 743–751 (2013).
52. B. Steglich, S. Sazer, K. Ekwall, Transcriptional regulation at the yeast nuclear envelope. *Nucleus* (2013) (available at http://www.tandfonline.com/doi/abs/10.4161/nucl.26394).
53. D. Dorsett, M. Merkenschlager, Cohesin at active genes: a unifying theme for cohesin and gene expression from model organisms to humans. *Curr. Opin. Cell Biol.* **25**, 327–333 (2013).
54. S. Holwerda, W. de Laat, Chromatin loops, gene positioning, and gene expression. *Front. Genet.* **3**, 217 (2012).
55. J. Huang, J.-M. Hsu, B. C. Laurent, The RSC nucleosome-remodeling complex is required for

Cohesin's association with chromosome arms. *Mol. Cell*. **13**, 739–750 (2004).
56. J. F. Sydow *et al.*, Structural basis of transcription: mismatch-specific fidelity mechanisms and paused RNA polymerase II with frayed RNA. *Mol. Cell*. **34**, 710–721 (2009).
57. H. Wong *et al.*, A predictive computational model of the dynamic 3D interphase yeast nucleus. *Curr. Biol.* **22**, 1881–1890 (2012).
58. D. S. Lemons, A. Gythiel, Paul Langevin's 1908 paper "on the theory of Brownian motion"["Sur la théorie du mouvement brownien," CR Acad. Sci.(Paris) 146, 530–533 (1908)]. *Am. J. Phys.* (1997) (available at http://scitation.aip.org/content/aapt/journal/ajp/65/11/10.1119/1.18725).
59. A. Goffeau *et al.*, Life with 6000 genes. *Science*. **274**, 546, 563–7 (1996).
60. I. Lassadi, A. Kamgoué, I. Goiffon, N. Tanguy-le-Gac, K. Bystricky, Differential Chromosome Conformations as Hallmarks of Cellular Identity Revealed by Mathematical Polymer Modeling. *PLoS Comput. Biol.* **11**, e1004306–21 (2015).
61. P. Therizols, T. Duong, B. Dujon, C. Zimmer, E. Fabre, Chromosome arm length and nuclear constraints determine the dynamic relationship of yeast subtelomeres. *Proc. Natl. Acad. Sci. U. S. A.* **107**, 2025–2030 (2010).
62. M. Gotta *et al.*, The clustering of telomeres and colocalization with Rap1, Sir3, and Sir4 proteins in wild-type Saccharomyces cerevisiae. *J. Cell Biol.* **134**, 1349–1363 (1996).
63. F. Hediger, F. R. Neumann, G. Van Houwe, K. Dubrana, S. M. Gasser, Live imaging of telomeres: yKu and Sir proteins define redundant telomere-anchoring pathways in yeast. *Curr. Biol.* **12**, 2076–2089 (2002).
64. K. Bystricky, T. Laroche, G. Van Houwe, M. Blaszczyk, S. M. Gasser, Chromosome looping in yeast: telomere pairing and coordinated movement reflect anchoring efficiency and territorial organization. *J. Cell Biol.* **168**, 375–387 (2005).
65. E. Lieberman-Aiden *et al.*, Comprehensive Mapping of Long-Range Interactions Reveals Folding Principles of the Human Genome. *Science*. **326**, 289–293 (2009).
66. N. L. van Berkum *et al.*, Hi-C: a method to study the three-dimensional architecture of genomes. *J. Vis. Exp.* (2010) (available at http://www.jove.com/index/Details.stp?ID=1869).
67. J.-M. Belton *et al.*, Hi-C: a comprehensive technique to capture the conformation of genomes. *Methods*. **58**, 268–276 (2012).
68. J.-M. Belton, J. Dekker, Hi-C in Budding Yeast. *Cold Spring Harb. Protoc.* **2015**, 649–661 (2015).
69. T. Nagano *et al.*, Comparison of Hi-C results using in-solution versus in-nucleus ligation. *Genome Biol.* **16**, 1068 (2015).
70. S. Wingett *et al.*, HiCUP: pipeline for mapping and processing Hi-C data. *F1000Res.* **4**, 1310 (2015).
71. M. Guidi *et al.*, Spatial reorganization of telomeres in long-lived quiescent cells. *Genome Biol.* **16**, 206 (2015).
72. T.-H. S. Hsieh *et al.*, Mapping Nucleosome Resolution Chromosome Folding in Yeast by Micro-C. *Cell*. **162**, 108–119 (2015).
73. D. Egecioglu, J. H. Brickner, Gene positioning and expression. *Curr. Opin. Cell Biol.* **23**, 338–345 (2011).
74. J. H. Brickner, P. Walter, Gene recruitment of the activated INO1 locus to the nuclear membrane. **2**, e342 (2004).
75. D. Guet *et al.*, Combining Spinach-tagged RNA and gene localization to image gene expression in live yeast. *Nat. Commun.* **6**, 8882 (2015).
76. J. M. Casolari, C. R. Brown, D. A. Drubin, O. J. Rando, P. A. Silver, Developmentally induced changes in transcriptional program alter spatial organization across chromosomes. *Genes Dev.* **19**, 1188–1198 (2005).
77. D. A. Drubin, A. M. Garakani, P. A. Silver, Motion as a phenotype: the use of live-cell imaging and machine visual screening to characterize transcription-dependent chromosome dynamics. *BMC Cell*

Biol. **7**, 19 (2006).
78. G. G. Cabal *et al.*, SAGA interacting factors confine sub-diffusion of transcribed genes to the nuclear envelope. *Nature*. **441**, 770–773 (2006).
79. Y. Sanchez, J. Taulien, K. A. Borkovich, S. Lindquist, Hsp104 is required for tolerance to many forms of stress. *EMBO J.* **11**, 2357–2364 (1992).
80. D. A. Parsell, A. S. Kowal, M. A. Singer, S. Lindquist, Protein disaggregation mediated by heat-shock protein Hsp104. *Nature*. **372**, 475–478 (1994).
81. G. Dieppois, N. Iglesias, F. Stutz, Cotranscriptional recruitment to the mRNA export receptor Mex67p contributes to nuclear pore anchoring of activated genes. *Mol. Cell. Biol.* **26**, 7858–7870 (2006).
82. C. Randise-Hinchliff, J. H. Brickner, Transcription factors dynamically control the spatial organization of the yeast genome. *Nucleus*. **7**, 369–374 (2016).
83. S. Ahmed *et al.*, DNA zip codes control an ancient mechanism for gene targeting to the nuclear periphery. *Nat. Cell Biol.* **12**, 111–118 (2010).
84. C. Randise-Hinchliff *et al.*, Strategies to regulate transcription factor–mediated gene positioning and interchromosomal clustering at the nuclear periphery. *J. Cell Biol.* **212**, 633–646 (2016).
85. X. Ji *et al.*, 3D Chromosome Regulatory Landscape of Human Pluripotent Cells. *Stem Cells*. **18**, 262–275 (2016).
86. A. D. Schmitt *et al.*, A Compendium of Chromatin Contact Maps Reveals Spatially Active Regions in the Human Genome. *Cell Rep.* **17**, 2042–2059 (2016).
87. K. Nasmyth, C. H. Haering, Cohesin: its roles and mechanisms. *Annu. Rev. Genet.* **43**, 525–558 (2009).
88. B. Bonev, G. Cavalli, Organization and function of the 3D genome. *Nat. Rev. Genet.* **17**, 661–678 (2016).
89. E. P. Nora *et al.*, Targeted degradation of CTCF decouples local insulation of chromosome domains from higher‐order genomic compartmentalization. *bioRxiv*, , doi:10.1101/095802.
90. M. Di Pierro, B. Zhang, E. L. Aiden, P. G. Wolynes, J. N. Onuchic, Transferable model for chromosome architecture. *Proc. Natl. Acad. Sci. U. S. A.* **113**, 12168–12173 (2016).
91. P. Vinciguerra, F. Stutz, mRNA export: an assembly line from genes to nuclear pores. *Curr. Opin. Cell Biol.* **16**, 285–292 (2004).
92. J. Dekker, Mapping in vivo chromatin interactions in yeast suggests an extended chromatin fiber with regional variation in compaction. *J. Biol. Chem.* **283**, 34532–34540 (2008).
93. T. Schalch, S. Duda, D. F. Sargent, T. J. Richmond, X-ray structure of a tetranucleosome and its implications for the chromatin fibre. *Nature*. **436**, 138–141 (2005).
94. Y. Cui, C. L. Denis, In vivo evidence that defects in the transcriptional elongation factors RPB2, TFIIS, and SPT5 enhance upstream poly(A) site utilization. *Mol. Cell. Biol.* **23**, 7887–7901 (2003).
95. H. C. Berg, *Random Walks in Biology* (Princeton University Press, 1993), *Princeton paperbacks*.
96. P. Carrivain, M. Barbi, J.-M. Victor, In silico single-molecule manipulation of DNA with rigid body dynamics. *PLoS Comput. Biol.* **10**, e1003456 (2014).
97. A. D. A. K. S. William Humphrey, VMD: visual molecular dynamics. *J. Mol. Graph.* **14**, 33–38 (1996).
98. K. Bystricky, P. Heun, L. Gehlen, J. Langowski, S. M. Gasser, Long-range compaction and flexibility of interphase chromatin in budding yeast analyzed by high-resolution imaging techniques. *PNAS*. **101**, 16495–16500 (2004).
99. A. Miele, K. Bystricky, J. Dekker, Yeast silent mating type loci form heterochromatic clusters through silencer protein-dependent long-range interactions. *PLoS Genet.* **5**, e1000478 (2009).
100. S. R. Engel *et al.*, The reference genome sequence of Saccharomyces cerevisiae: then and now. *G3* . **4**, 389–398 (2014).