



**HAL**  
open science

# Differential Chromosome Conformations as Hallmarks of Cellular Identity Revealed by Mathematical Polymer Modeling

Imen Lassadi, Alain Kamgoué, Isabelle Goiffon, Nicolas Tanguy-Le-Gac, Kerstin Bystricky

## ► To cite this version:

Imen Lassadi, Alain Kamgoué, Isabelle Goiffon, Nicolas Tanguy-Le-Gac, Kerstin Bystricky. Differential Chromosome Conformations as Hallmarks of Cellular Identity Revealed by Mathematical Polymer Modeling. PLoS Computational Biology, 2015, 11 (6), pp.e1004306. <10.1371/journal.pcbi.1004306>. <hal-03026206>

**HAL Id: hal-03026206**

**<https://hal.science/hal-03026206v1>**

Submitted on 3 Dec 2020

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



HAL Authorization

RESEARCH ARTICLE

# Differential Chromosome Conformations as Hallmarks of Cellular Identity Revealed by Mathematical Polymer Modeling

Imen Lassadi<sup>1,2</sup>✉, Alain Kamgoué<sup>1,2</sup>✉\*, Isabelle Goiffon<sup>1,2</sup>, Nicolas Tanguy-le-Gac<sup>1,2</sup>, Kerstin Bystricky<sup>1,2</sup>\*

**1** University of Toulouse, UPS, Toulouse, France, **2** Laboratoire de Biologie Moléculaire Eucaryote, CNRS, LBME, Toulouse, France

✉ These authors contributed equally to this work.

\* [kerstin.bystricky@biotoul.fr](mailto:kerstin.bystricky@biotoul.fr)



**OPEN ACCESS**

**Citation:** Lassadi I, Kamgoué A, Goiffon I, Tanguy-le-Gac N, Bystricky K (2015) Differential Chromosome Conformations as Hallmarks of Cellular Identity Revealed by Mathematical Polymer Modeling. *PLoS Comput Biol* 11(6): e1004306. doi:10.1371/journal.pcbi.1004306

**Editor:** Alexandre V Morozov, Rutgers University, UNITED STATES

**Received:** February 5, 2015

**Accepted:** April 28, 2015

**Published:** June 1, 2015

**Copyright:** © 2015 Lassadi et al. This is an open access article distributed under the terms of the [Creative Commons Attribution License](http://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

**Data Availability Statement:** The authors confirm that all data underlying the findings are fully available without restriction. Mathematical analysis and modeling was programmed using MatLab. Data and codes are deposited at <http://dx.doi.org/10.5061/dryad.rp3g3>

**Funding:** Financial support was provided by the Institut Universitaire de France (IUF), the Agence Nationale de la Recherche (ANR) and a HFSP research grant RGP0044. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

## Abstract

Inherently dynamic, chromosomes adopt many different conformations in response to DNA metabolism. Models of chromosome organization in the yeast nucleus obtained from genome-wide chromosome conformation data or biophysical simulations provide important insights into the average behavior but fail to reveal features from dynamic or transient events that are only visible in a fraction of cells at any given moment. We developed a method to determine chromosome conformation from relative positions of three fluorescently tagged DNA in living cells imaged in 3D. Cell type specific chromosome folding properties could be assigned based on positional combinations between three loci on yeast chromosome 3. We determined that the shorter left arm of chromosome 3 is extended in *MAT $\alpha$*  cells, but can be crumpled in *MAT $a$*  cells. Furthermore, we implemented a new mathematical model that provides for the first time an estimate of the relative physical constraint of three linked loci related to cellular identity. Variations in this estimate allowed us to predict functional consequences from chromatin structural alterations in *asf1* and recombination enhancer deletion mutant cells. The computational method is applicable to identify and characterize dynamic chromosome conformations in any cell type.

## Author Summary

The spatial organization of the genome inside eukaryotic cell nuclei has been shown to play a role in transcription, replication, recombination and DNA repair. Probabilistic models have correlated structural fluctuations with these processes, but methods to detect transient features describing chromosome conformation are lacking. We developed a new fluorescent repressor-operator system (FROS) based on the  $\lambda$  repressor and combined it with the existing lac and tet FROS to measure the relative spatial positioning between three labelled DNA loci on chromosome 3 in live yeast cells. To quantitatively analyze and interpret the data we applied an original computational method that relies on the geometrical distribution of the three tags. Our results show that the conformation of the small

**Competing Interests:** The authors have declared that no competing interests exist.

yeast chromosome 3 is mating type specific in G1. Differential folding of the left arm of this chromosome can be attributed to a small DNA element which could explain why loci on this arm may be excluded from recombination with the *MAT* locus on the right arm of chromosome 3. Chromatin structural properties altered in the absence of the Asf1 histone chaperone contribute to the lineage specific chromosome organization and the relative position of the three mating type loci.

## Introduction

The three-dimensional organization of the genome was shown to dynamically adapt to nuclear function and complexity [1–3]. Chromatin fibers can be represented by a polymer random coil adopting a considerable, largely unappreciated number of states in response to DNA metabolism [4–7]. How the intrinsic folding of a chromosome within the nucleus and the relative position of loci on specific chromosomes contribute to these processes is not known. Chromosome conformation capture techniques that rely on protein-DNA cross-linking have provided precious information on the frequency of three-dimensional long range molecular contacts between genomic DNA segments [4,8–10]. Microscopy approaches are, on the other hand, necessary to further our understanding of the dynamics of DNA transactions because they allow analysis of live and single cells [11,12]. Using two-color distance measurements between two fluorescently labelled loci in 3D in fixed or living cells has yielded data to infer chromatin compaction parameters by polymer modeling [13–15]. These parameters were also used to model 4C and HiC data (for example see [4,8,12]) or to simulate genome-wide chromosome organization [2]. Because the nucleosome fiber is highly flexible, inferring fiber properties from measuring distances between two labeled loci implies making a number of assumptions to describe the actual path of the fiber. In order to obtain spatial, 3D information, at least 3 points in space are needed. Three points provide geometrical information that can be used to establish physical models. In the nucleus, these points can include reference structures, such as the nuclear envelope or the nucleolus in yeast, or three distinct DNA tags. Few previous studies used three labeled loci: in fixed, mammalian cells, distances and angles within a triangle formed by three probes in the same nucleus were measured to study changes in chromatin domain compaction [16]; in live bacteria cells, three labels were used to determine the position of two chromosomal loci relative to a third, reference one, in 2D [17]. A geometrical interpretation of changes in chromosome folding was not proposed.

The challenge for the analysis of three points in space is to develop mathematical algorithms allowing comparison of two sets of data, because commonly used tests for comparing distributions (1D, using Wilcoxon or KS tests) are no longer applicable. New computational tools are needed to identify specific features which may not be the most prominent ones, notably in cellular systems evolving over time or in space. To address this need, we developed a new system to fluorescently label three distinct genomic loci in living cells simultaneously and implemented mathematical algorithms to analyze the relative 3D positions of the labelled DNA loci.

We used *S. cerevisiae* chromosome 3 (Chr3) as a model to study its folding. Chr3 is a short chromosome of only 320 kb with a tripartite organization: two AT rich domains flank a GC-rich centromere proximal region and forms a ring-like structure mediated by frequent contacts between heterochromatin loci near the ends [18–20]. This chromosome has received particular attention because the study of the mating type loci, and their interconversion, contributed to fundamental knowledge about cell lineage control, silencing and recombination [21]. Numerous genetic data contributed to a better understanding of the directionality of the mating type

switch, yet the underlying mechanism is not known [22,23]. A role for chromosome architecture is usually invoked as a possible driving force for donor choice [24–27]. We have the possibility now to address this question by studying the nuclear position of the three mating type loci in the nucleus simultaneously.

In this study, we adapted a third system, based on the  $\lambda$  operator [28], from bacteria to yeast. In combination with the widely used Tet and Lac operators, we simultaneously tag three distinct loci in live yeast cells. We then developed a computational approach to demonstrate, that the frequency of specific positional combinations between three loci point to folding properties of a small chromosome. We find that folding of the left arm is different in *MATa* and *MAT $\alpha$*  cells. In strains lacking either a component involved in chromatin compaction, the chaperone Asf1, or the recombination enhancer (RE), the mating type specific conformation of Chr3 is altered. Our results suggest that chromosomal organization brought about by fiber folding and heterochromatic domains contribute to control of the yeast mating type loci by regulating long-range contacts.

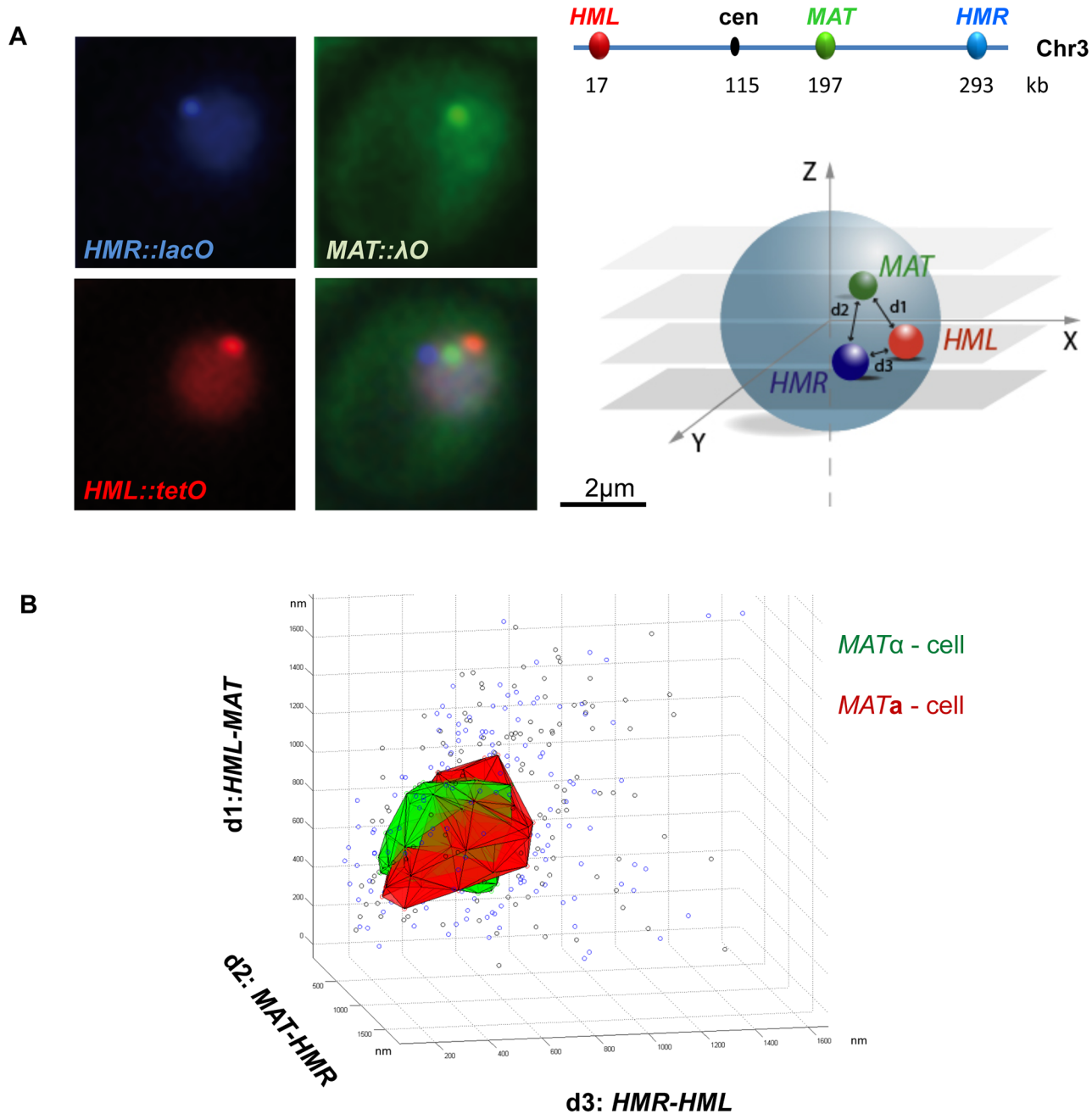
## Methods

### Lambda cloning and strain construction

To simultaneously label three distinct loci, we adapted the bacterial  $\lambda$  repressor operator system ( $\lambda$ O and  $\lambda$ Ci) for use in yeast [28,29] and combined it with the existing Lac and Tet systems [30] (Fig 1A). Strains bearing three distinct labels were created by integration of constructs encoding fluorescent repressor fusion proteins followed by the integration of the operator sequences into the yeast genome by transformation. Yeast strains are listed in Table 1.

Expression of an operon containing the phage  $\lambda$ CI repressor gene fused to the gene encoding YFP was placed under the control of the pURA3 promoter. The  $\lambda$  repressor sequence bears amino acid modifications G48S and Y210H which strengthen DNA binding and reduce tetramerization, respectively [28]. The  $\lambda$ CI sequence was amplified by PCR using pRFG116 and kindly provided by Dr D Chatteraj. During the PCR reaction a NotI recognition site was introduced. The pURA3 promoter was amplified from the pCJ97 plasmid with an introduction of a NheI site. The two fragments were then digested by NheI and ligated to generate a pURA3- $\lambda$ CI fragment. This fragment was double-digested with EcoRI and NotI, and ligated to the double-digested (EcoRI/NotI) pCJ97 plasmid bearing a YFP sequence.

Following the validation steps discussed above, we subcloned the pURA3- $\lambda$ CI-YFP in a plasmid bearing pHIS-CFP-lacI-pUra3-TetR-YFP (named pGVH30 [18]), plasmid which allows expression of two fusion proteins from a single plasmid thus requiring a unique selection marker (ADE2), finally leading to plasmid pIL01. The  $\lambda$ O were extracted from the bacterial pRFB122 plasmid by an XhoI digestion and ligated to the pSR6 plasmid [31] and digested by XhoI/SalI. Integration of the operator repeats was performed using a cloning free technique. The method entails insertion of a marker gene generated by PCR using long primers, with the optimal size of the locus-specific primer tails varying from 60 to 80 nt, near the locus of interest. The marker is then replaced by the operator repeats found in pSR plasmids. The following PCR-amplified genomic fragments (SGD coordinates) were used for insertion within 0.5 to 4 kb from the respective loci: 15160 kb to 15773 kb for *HML*, 294898 kb to 295245 kb for *HMR*, 90917 kb to 92521 kb for *LEU*, 239254 kb to 240927 kb for *ARS1413* and 16431 kb to 17993-kb for *TelVI*. The PCR-amplified sequences of *HMR* and *HML* were cloned into a pAFS52-lacO plasmid and in the pAFS59-tetO bearing plasmid, respectively [18]. The lambda operator repeats were integrated at 197197–197310 kb on Chr3 to label the *MAT* locus. The *ASF1* gene (196334 kb to 197080 kb on Chr10) or the RE (29083 kb to 29748 kb on Chr3) were replaced by the hygromycin resistance gene amplified from pAG32 (1743bp). Alternatively, the RE



**Fig 1. Cell type specific spatial organization of the three mating type loci on yeast chromosome 3.** A) Representative wide-field fluorescent images of the the mRFP-tetR, λcl-YFP and CFP-lacI foci at *HML*, *MAT* and *HMR*. Insertion sites of the TetO, λO and LacO arrays on S.c. Chr3 are shown. B) Combinations of the geometric coordinates d1, d2, d3 are plotted for each nucleus in *MATa* (black dots) and *MATα* (blue dots). The 50% most frequent combinations in *MATa* (red) and *MATα* (green) are included in a 3D volume.

doi:10.1371/journal.pcbi.1004306.g001

(28987–29852 on Chr3) was replaced by loxP flanked hygromycin resistance gene amplified from pGI10. Galactose induced expression of the recombinase from pHS47 induces deletion of the entire cassette (162bp remaining plasmid sequences). The lambda operator (λO) system comprises a relatively small number of repeated binding sites; 64 repeats compared to the usual 128–256 repeats of TetO and LacO arrays. The focus formed by binding of multiple repressor

**Table 1. Yeast strains used in this study.**

Name	Genotype
yIL02	<i>Mata ade2-1 can1-100 his3-11, 15 leu2-3, 112 trp1-1 ura3-1 ade2-1::His3p-CFP-lacI-URA3p-lacI-YFP-ADE2, TetR-mRFP:NAT1</i>
yIL03	<i>Mata ade2-1 can1-100 his3-11, 15 leu2-3, 112 trp1-1 ura3-1 ade2-1::His3p-CFP-lacI-URA3p-lacI-YFP-ADE2, TetR-mRFP:NAT1</i>
yIL30	<i>Mata ade2-1 can1-100 his3-11, 15 leu2-3, 112 trp1-1 ura3-1 ade2-1::His3p-CFP-lacI-URA3p-lacI-YFP-ADE2, TetR-mRFP:NAT1, MAT5::λO-HIS, HMR::LacO-TRP, HML::TetO-LEU</i>
yIL31	<i>Mata ade2-1 can1-100 his3-11, 15 leu2-3, 112 trp1-1 ura3-1 ade2-1::His3p-CFP-lacI-URA3p-lacI-YFP-ADE2, TetR-mRFP:NAT1, MAT5::λO-HIS, HMR::LacO-TRP, HML::TetO-LEU</i>
yIL30 <i>asf1</i>	<i>Mata ade2-1 can1-100 his3-11, 15 leu2-3, 112 trp1-1 ura3-1 ade2-1::His3p-CFP-lacI-URA3p-lacI-YFP-ADE2, TetR-mRFP:NAT1, MAT5::λO-HIS, HMR::LacO-TRP, HML::TetO-LEU, asf1::HYG</i>
yIL31 <i>asf1</i>	<i>Mata ade2-1 can1-100 his3-11, 15 leu2-3, 112 trp1-1 ura3-1 ade2-1::His3p-CFP-lacI-URA3p-lacI-YFP-ADE2, TetR-mRFP:NAT1, MAT5::λO-HIS, HMR::LacO-TRP, HML::TetO-LEU, asf1::HYG</i>
yIL30 Δ <i>re</i>	<i>Mata ade2-1 can1-100 his3-11, 15 leu2-3, 112 trp1-1 ura3-1 ade2-1::His3p-CFP-lacI-URA3p-lacI-YFP-ADE2, TetR-mRFP:NAT1, MAT5::λO-HIS, HMR::LacO-TRP, HML::TetO-LEU, re::HYG</i>
yIL31 Δ <i>re</i>	<i>Mata ade2-1 can1-100 his3-11, 15 leu2-3, 112 trp1-1 ura3-1 ade2-1::His3p-CFP-lacI-URA3p-lacI-YFP-ADE2, TetR-mRFP:NAT1, MAT5::λO-HIS, HMR::LacO-TRP, HML::TetO-LEU, re::HYG</i>
yIL32	<i>Mata ade2-1 can1-100 his3-11, 15 leu2-3, 112 trp1-1 ura3-1, ade2-1::His3p-CFP-lacI-URA3p-lacI-YFP-ADE2, MAT5::λO-HIS, TetR-mRFP-NAT, HML::LacO-TRP, Leu::TetO-LEU</i>
yIL33	<i>Mata ade2-1 can1-100 his3-11, 15 leu2-3, 112 trp1-1 ura3-1, ade2-1::His3p-CFP-lacI-URA3p-lacI-YFP-ADE2, MAT5::λO-HIS, TetR-mRFP-NAT, HML::LacO-TRP, Leu::TetO-LEU</i>
yIL34	<i>Mata ade2-1 can1-100 his3-11, 15 leu2-3, 112 trp1-1 ura3-1, ade2-1::His3p-CFP-lacI-URA3p-lacI-YFP-ADE2, MAT5::λO-HIS, TetR-mRFP-NAT, ARS1413::LacO-TRP, VI-L::TetO-LEU</i>
yIL35	<i>Mata ade2-1 can1-100 his3-11, 15 leu2-3, 112 trp1-1 ura3-1, ade2-1::His3p-CFP-lacI-URA3p-lacI-YFP-ADE2, MAT5::λO-HIS, TetR-mRFP-NAT, ARS1413::LacO-TRP, VI-L::TetO-LEU</i>

doi:10.1371/journal.pcbi.1004306.t001

fusion proteins to the arrays integrated into the genome is easily detectable using conventional fluorescence microscopy (Fig 1A). The constitutive expression of the  $\lambda$ C1 repressor fused to YFP and its binding to a short 1.5kb DNA  $\lambda$ O fragment was not toxic to the cell, which displayed growth rates identical to unmodified yeast cultures.  $\lambda$ C1-YFP fusion proteins diffuse freely in the cytoplasm, and in the nucleoplasm apart from the vacuole (Fig 1A). The focus formed at a site near the *MAT* locus (197 kb along right arm of Chr3) tagged using  $\lambda$ O was positioned in the center of the nuclear lumen with the same frequency as the same genetic locus tagged with LacO.

### Microscopy and image analysis

Live microscopy was performed using an Olympus IX-81 wide-field fluorescence microscope, equipped with a CoolSNAPHQ camera (Roper Scientific) and a Polychrome V (Till Photonics), electric piezo with accuracy of 10 nm and imaged through an Olympus oil immersion objective 100X PLANAPO NA1.4. Yeast cells were spread on a concave microscopy slide filled with SD-agarose (YNB + 2% sugar/ carbon source + 3%(w/v) agarose). Acquisition of CFP, mRFP and YFP was performed in 3D (21 focal planes at 0.2  $\mu$ m distance intervals; 500 ms, 300 ms, 500 ms acquisition times respectively). Fluorescent intensities of the acquired spots were identical in both cell types demonstrating that the inserted operator arrays maintained the same size. The x, y and z coordinates for each focus were automatically measured using the “Spot distance” plug in on image J (<http://bigwww.epfl.ch/sage/soft/spotdistance/>; D Sage, EPFL; [29]). This program uses multi-channel z-stack images to localize the center of the nucleus based on the background fluorescence of the CFP-lacI. The position of each focus is assigned to the center of gravity of the fluorescence around the brightest pixel found in this nucleus in a filtered version of the image. The signals are scored on 3D stacks using at least 200 nuclei,

monitoring nuclear integrity and cell cycle stage through bud shape and nuclear diameter. Distances between pairs of loci can be determined in 3D using this Image J plugin. Statistical analysis was performed using the Wilcoxon test.

### Mating type switching assay

Expression of the HO-endonuclease gene from the pHO (URA selection[32]) was induced by addition of 2% filtered galactose (SIGMA) to a yeast culture exponentially growing in 2% raffinose. For quantitative PCR reactions, Bio-Rad SybrGreen Supermix was used in the presence of 30ng of genomic DNA and 0.26μM of each primer. PCR were run on ABI 7900. The MAT a-specific primers were 5'-GGCATTACTCCACTTCAAGT (P1) and 5'-ATGTGAACCGCATGGGCAGT (P2). The MATα-specific primers were 5'-ATGTGAACCGCATGGGCAGT (P2) and 5'-GCAGCACGGAATATGGGACT (P3). Primers specific for the ARG5, 6 locus 5'-CAAGGATCCAGCAAAGTTGGGTGAAGTATGGTA and 5'-GAAGGATCCAAATTTGTCTAGTGTGGGAACG or for the actin locus were used for normalization. All qPCR assays were accompanied by reactions using dilutions of genomic DNA from wt strains' 0h input sample to assess the linearity of the PCR signal and to create calibration curves.

### Modeling experience and problem formulation

One considers a fiber with  $N$  nodes. Each state  $E_k$  of this fiber defined in a reference frame  $R_k$  is represented by  $N$  nodes:

$$(x_1, y_1, z_1)_{R_k}, (x_2, y_2, z_2)_{R_k}, \dots, (x_N, y_N, z_N)_{R_k} \quad k = 1 \dots M$$

For all states  $E_{k_1}$  and  $E_{k_2}$ ,  $R_{k_1} \neq R_{k_2}$  since  $k_1 \neq k_2$ .

On a chromosome represented as a polymer fiber, the center of gravity of each fluorescently labelled locus represents a node.  $N$  loci can be represented by  $(x_i, y_i, z_i)_{i=1 \dots N}$ .

If we denote  $\theta_i$  the angle around the node  $i$ .

$$d_i^1 = \begin{cases} d_{i,i+1} & \text{if } i < N \\ d_{N,1} & \text{if } i = N \end{cases} \quad (1)$$

$$d_i^2 = \begin{cases} d_{i-1,i} & \text{if } i > 1 \\ d_{N,1} & \text{if } i = 1 \end{cases} \quad (2)$$

For each local analysis around node  $i$ , we make the following change of variables:

$$(x_{i-1}, y_{i-1}, z_{i-1}, x_i, y_i, z_i, x_{i+1}, y_{i+1}, z_{i+1}) \succ (d_i^1, d_i^2, \theta_i)$$

The main advantage of these new variables is their invariance during the reference frame change. To analyze the variations of positions around the node  $i$ , we introduce the normalized principal components analysis operator  $PCA_{norm}$  and consider a set of experimental data

$\{(d_i^{1,k}, d_i^{2,k}, \theta_i^k)_{i=1 \dots N}\}_{k=1 \dots M}$  acquired around all nodes  $i$  in all nuclei of state  $E_k$ . We denote by  $C_i$  and  $P_i$  the correlation matrix of the principal component analysis and the diagonal matrix of eigenvalues.

$$[C_i, P_i] = PCA_{norm} \{(d_i^{1,k}, d_i^{2,k}, \theta_i^k)\}_{k=1 \dots M} \quad (3)$$

For all nuclei, we can write the coordinates of red, green and blue loci in the referential linked to the microscope as  $(x_r, y_r, z_r)_{R_{microscope}}, (x_g, y_g, z_g)_{R_{microscope}}$ . Taking the red locus as the

origin, we can write

$$\begin{aligned} (x_r, y_r, z_r)_{R_{\text{microscope}}} &\equiv (0, 0, 0)_{R_{\text{microscopenuclei}}}, \\ (x_g, y_g, z_g)_{R_{\text{microscope}}} &\equiv (x_g - x_r, y_g - y_r, z_g - z_r)_{R_{\text{microscopenuclei}}}, \\ (x_b, y_b, z_b)_{R_{\text{microscope}}} &\equiv (x_b - x_r, y_b - y_r, z_b - z_r)_{R_{\text{microscopenuclei}}}. \end{aligned}$$

Referential  $R_{\text{microscopenuclei}}$  is called pseudo-referential because its origin is inside the nucleus and its basis vectors are linked to the microscope. To overcome the orientation problem inherent to the nuclear sphericity, we can define new variables simple enough to analyze the position of the three loci as (1 and 1) (S1 Fig). Let us take:

$$d_1 = [(x_g - x_r)^2 + (y_g - y_r)^2 + (z_g - z_r)^2]^{\frac{1}{2}} \tag{4}$$

$$d_2 = [(x_b - x_r)^2 + (y_b - y_r)^2 + (z_b - z_r)^2]^{\frac{1}{2}} \tag{5}$$

and

$$\theta = \arccos\left(\frac{(x_g - x_r, y_g - y_r, z_g - z_r) \circ (x_b - x_r, y_b - y_r, z_b - z_r)}{d_1 \cdot d_2}\right) \tag{6}$$

where  $\circ$  is the scalar product.

With these new variables each nucleus is represented by three variables instead of nine:  $(x_r, y_r, z_r, x_g, y_g, z_g, x_b, y_b, z_b) > (d_1, d_2, \theta)$

## Variables and referential

**Projection.** Each nucleus is represented by the three components  $(d_1, d_2, \theta)$ . We project all points  $(d_1^i, d_2^i, \theta^i)_{i=1..M}$  in  $(d_1, d_2)$ ,  $(\theta, d_1)$  and  $(\theta, d_2)$  planes (Fig 2A and S1 Fig),

$(X_i, X_j) \equiv (X_i, X_j)_{X_i = \theta, d_1, d_2 / X_j = \theta, d_1, d_2 / i \neq j}$ . In each plane, the probability density function is estimated by Parzen-Rozenblatt method [33].

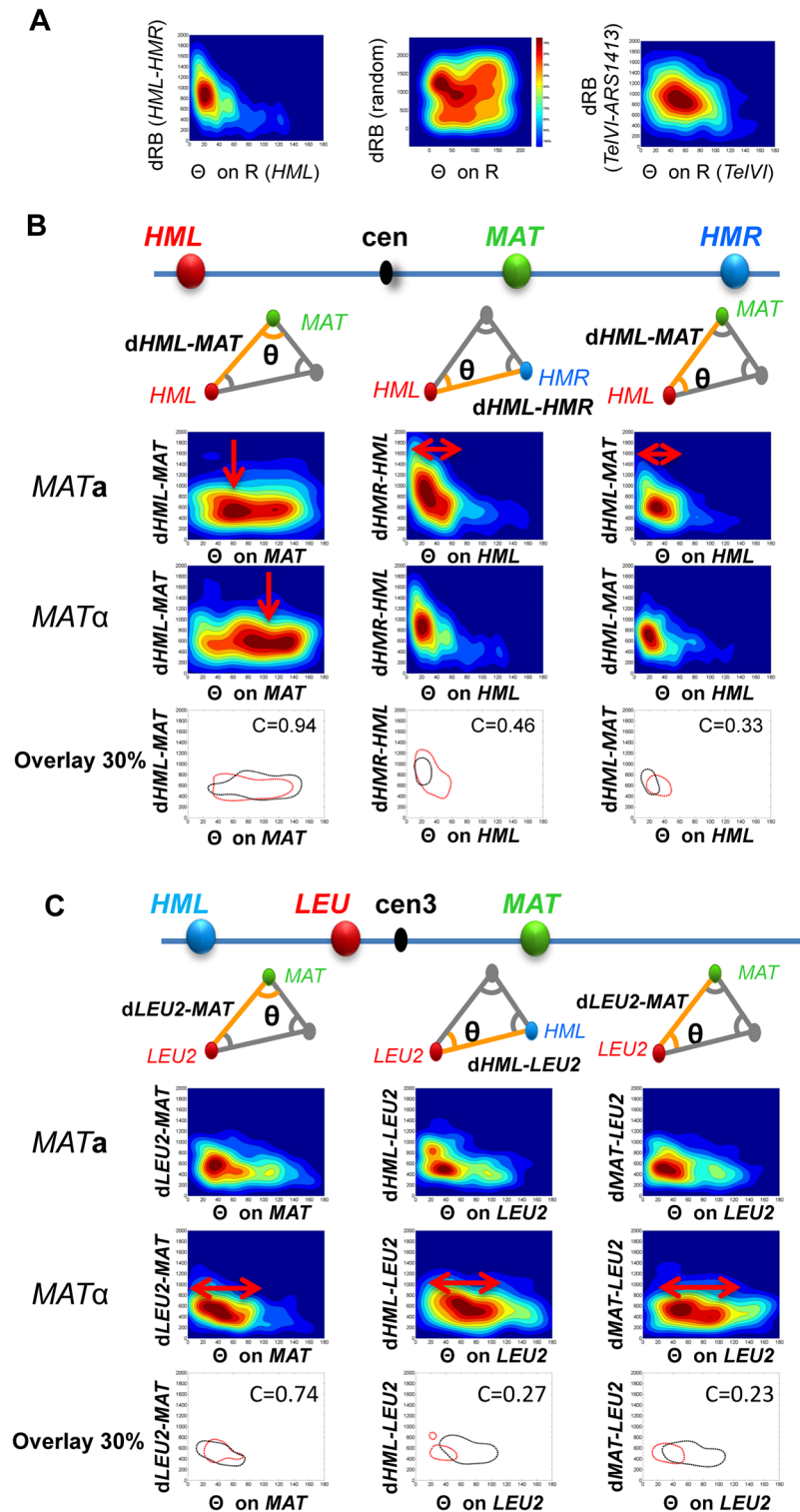
**Kernel density estimation.** In the Parzen-Rozenblatt methods, estimations lie in the choice of the kernel  $K$ . We use a Gaussian kernel[34]. For  $x, y$  in  $\mathbb{R}^2$ , the kernel  $K$  is given by:

$$K(x, y, t) = \frac{1}{2\pi t} e^{-\frac{(x-y)^2}{2t}} \tag{7}$$

This Kernel allows us to compute the probability of the density function  $f$ . The cumulative distribution function  $F(x, y)$  is plotted within ten regions defined as

$$U_k = \left\{ (x, y) \ / \ F(x, y) \leq \frac{k}{10} \right\} \tag{8}$$

. Each density map  $(X_i, X_j)$  is composed of ten levels  $U_k = (X_i^{kl}, X_j^{kl})_{k=1..10 / l=1..card(X_i^k)}$  where  $card(X_i^k)$  is the number of nuclei inside level  $k$ , giving the relation  $U_1 \subset U_2 \subset \dots \subset U_{10}$



**Fig 2. Folding of the left arm of the chromosome 3 differs in a subset of MATa and MAT $\alpha$  cells.** A) 3D data are projected onto a unique 2D plane (eg. d1/ $\theta$ ; 9 projections can be generated from each data set. Density maps (warm colors for high density and cold colors for weak density in 10% increments) are generated for each projection in 2D. Examples shown are density maps where the origin is set to the red spot (R), the d1 distances (vector RB red to blue) are aligned and plotted relative to the angle  $\theta$  of the RGB triangle at R. left panel: R = HML, d1 = RB = HML—HMR,  $\theta$  at HML. Center panel: Example of a density map

obtained from a simulation based on a random draw of relative positions between 3 loci. Right panel:  $Te\Delta I = R$ ,  $ARS1413 = B$  and  $MAT$ . B-C)  $HML$ ,  $MAT$  and  $HMR$  (B; YIL30/YIL31;  $n = 223$  and  $276$ ) or  $LEU$ ,  $MAT$  and  $HML$  (C; YIL32/33;  $n = 409$  and  $323$ ) were labelled using  $TetO/mRFP-tetR$ ,  $\lambda O/\lambda cl-YFP$  and  $LacO/CFP-lacI$  respectively. Examples of density maps are shown to compare the distribution between the 3 loci in  $MAT\alpha$  and  $MAT\alpha$  cells. Red arrows highlight changes between  $MAT\alpha$  and  $MAT\alpha$  cells. An overlay of the contours of the density area representing 30% of the analyzed  $MAT\alpha$  (black) and  $MAT\alpha$  (red) is represented. The correlation factor  $c$  is given for 30% contour. Complete sets of density maps are shown in S2–S3 Figs.

doi:10.1371/journal.pcbi.1004306.g002

For each map  $U_k$ , we denote  $\mu_k = (\mu_k^i, \mu_k^j)$  the mean vector and  $\sigma_k = (\sigma_k^i, \sigma_k^j)$  the standard deviation vector.

$$\mu_k^i = \frac{\sum_{l=1..card(X_1^k)} X_1^{kl}}{card(X_1^k)} \tag{9}$$

$$\mu_k^j = \frac{\sum_{l=1..card(X_j^k)} X_j^{kl}}{card(X_j^k)} \tag{10}$$

$$\sigma_k^i = \sqrt{\frac{\sum_{l=card(X_1^k)} (X_1^{kl} - \mu_k^i)^2}{card(X_1^k)}} \tag{11}$$

$$\sigma_k^j = \sqrt{\frac{\sum_{l=card(X_j^k)} (X_j^{kl} - \mu_k^j)^2}{card(X_j^k)}} \tag{12}$$

Let us take two different experiments **A** and **B**, and  $(U_k)_{k=1..10}$  level intensity associated to **A** and  $(V_k)_{k=1..10}$  level intensity associated to **B**. We denote  $\mu_{kU}$  and  $\mu_{kV}$  associated mean vectors,  $\sigma_{kU}$  and  $\sigma_{kV}$  associated standard deviation vectors at the level  $k$ . For two levels,  $k_1$  in the **A**-experiment and  $k_2$  in the **B**-experiment we define

$$\rho_{k_1 k_2}^{1ij} = \frac{\min(\mu_{k_1 U}^i, \mu_{k_2 V}^i) \min(\mu_{k_1 U}^j, \mu_{k_2 V}^j)}{\max(\mu_{k_1 U}^i, \mu_{k_2 V}^i) \max(\mu_{k_1 U}^j, \mu_{k_2 V}^j)} \tag{13}$$

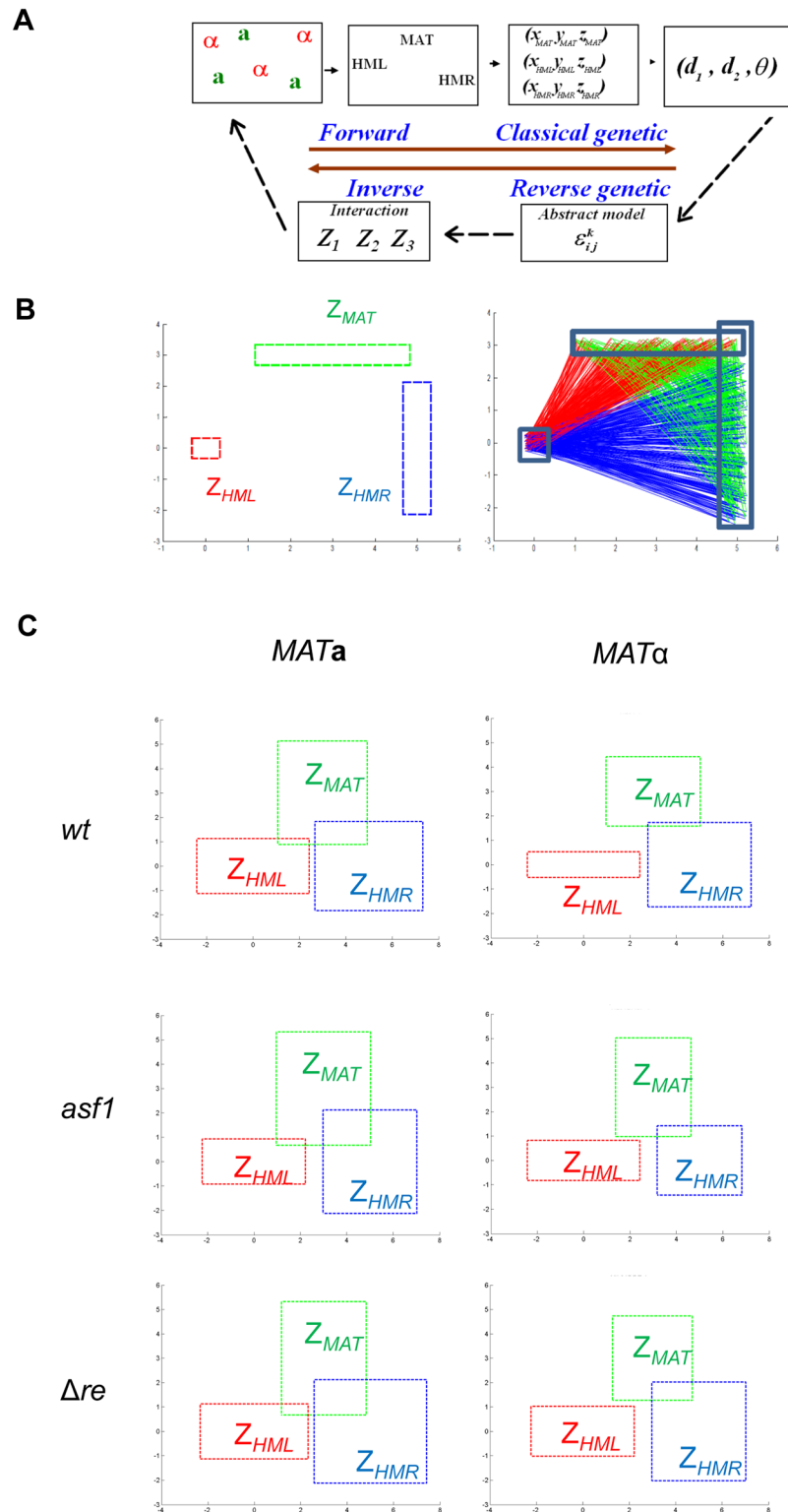
and

$$\rho_{k_1 k_2}^{2ij} = \frac{\min(\mu_{k_1 U}^i, \mu_{k_2 V}^i) \min(\mu_{k_1 U}^j, \mu_{k_2 V}^j) \min(\sigma_{k_1 U}^i, \sigma_{k_2 V}^i) \min(\sigma_{k_1 U}^j, \sigma_{k_2 V}^j)}{\max(\mu_{k_1 U}^i, \mu_{k_2 V}^i) \max(\mu_{k_1 U}^j, \mu_{k_2 V}^j) \max(\sigma_{k_1 U}^i, \sigma_{k_2 V}^i) \max(\sigma_{k_1 U}^j, \sigma_{k_2 V}^j)} \tag{14}$$

$\rho_{k_1 k_2}^{1ij}$  is a correlation in first approximation and  $\rho_{k_1 k_2}^{2ij}$  is a correlation in second approximation with  $\rho_{k_1 k_2}^{1ij} \geq \rho_{k_1 k_2}^{2ij}$ .

### Inverse problem and abstract model

We developed an original abstract model based on the assumption that each point of a fiber (or polymer) moves in a specific area (Fig 3A). If we consider a dynamic polymer in the plane in which we can define  $N$  points  $(x_i, y_i)_{i=1..N}$ . For each point  $(x_i, y_i)$ , we define the survival zone as the smallest closed set  $P^i$  where  $(x_i, y_i)$  takes its values during temporal fluctuations. Any



**Fig 3. An abstract model to determine the relative physical constraint of three linked loci in wt and mutant strains.** A) The forward mathematical approach allows determining features of a subpopulation of relative positions between 3 loci in space. The inverse computational approach consists in modeling the relative positions or survival zones of loci whose positions are spatially linked. The underlying polymer fiber imposes constraints on the positions each locus can adapt. Iterative calculation defines the survival zones of

each locus (or node) which can predict biological relevant features. B-C) Survival Zones ( $Z$ ) of *HML*, *MAT* and *HMR*. The initial position of the 3 loci is set based on the estimated conformation for Chr3 (B). Iterations were run using Eq 25 (see Methods) and the statistically most significant zones are represented for wt, *asf1* mutants and strains in which the recombination enhancer element was deleted (C).

doi:10.1371/journal.pcbi.1004306.g003

configuration  $Co_j$  of our polymer is defined as:

$$\exists (x_k^j, y_k^j)_{k=1\dots N} \in P^k / Co_j = (x_1^j, y_1^j) \rightarrow (x_2^j, y_2^j) \rightarrow \dots \rightarrow (x_N^j, y_N^j) \quad (15)$$

Let us note  $Z_{V_1}, Z_{V_2}, \dots, Z_{V_N}$  survival zones of  $(x_1, y_1), (x_2, y_2)$  and  $(x_N, y_N)$  defined as  $\mathbb{R}^2$ -closed set.

$$Z_{V_1}(\epsilon_1^1, \epsilon_2^1) = [x_1 - \epsilon_1^1; x_1 + \epsilon_1^1] \times [y_1 - \epsilon_2^1; y_1 + \epsilon_2^1] \quad (16)$$

$$Z_{V_2}(\epsilon_1^2, \epsilon_2^2) = [x_2 - \epsilon_1^2; x_2 + \epsilon_1^2] \times [y_2 - \epsilon_2^2; y_2 + \epsilon_2^2] \quad (17)$$

$$Z_{V_N}(\epsilon_1^N, \epsilon_2^N) = [x_N - \epsilon_1^N; x_N + \epsilon_1^N] \times [y_N - \epsilon_2^N; y_N + \epsilon_2^N] \quad (18)$$

We can define an abstract polymer from the knowledge of the  $N$  survival zones. Here, the abstract fiber will be close to the real fiber as soon as the survival zones will be small and the number of nodes is large enough. We can build several random configurations of our polymer knowing the survival zones.

$$(X_{1,i}, Y_{1,i})_{i=1\dots K} = \text{rand}(Z_{V_1}) \quad (19)$$

$$(X_{2,i}, Y_{2,i})_{i=1\dots K} = \text{rand}(Z_{V_2}) \quad (20)$$

$$(X_{N,i}, Y_{N,i})_{i=1\dots K} = \text{rand}(Z_{V_N}) \quad (21)$$

where  $\text{rand}(Z_{V_i})$  denotes the uniform random function in the set  $Z_{V_i}$

We can construct an abstract configuration:

$$Co_i^a = (X_{1,i}, Y_{1,i}) \rightarrow (X_{2,i}, Y_{2,i}) \rightarrow \dots \rightarrow (X_{N,i}, Y_{N,i}) \quad (22)$$

For the given abstract configuration set, we can define as (1 and 2) new system variables:

$$\{(D_{i,k}^1, D_{i,k}^2, \Theta_{i,k})_{k=1\dots M'}\}_{1\dots N}$$

Let us note  $C_i^a$  and  $P_i^a$  correlation matrix and diagonal eigenvalues matrix of the abstract system

$$[C_i^a, P_i^a] = \text{PCA}_{\text{norm}} \{(D_{i,k}^1, D_{i,k}^2, \Theta_{i,k})\}_{k=1\dots M'} \quad (23)$$

Where  $M'$  denotes the number of abstract configurations and  $N$  the number of nodes.

For a given set of points, the system of survival zones that best correlates with our experimental data has to be determined. Let us note

$$M_{\Delta}^i = C_i - C_i^a \quad (24)$$

as the difference between the experimental correlation matrix and abstract correlation matrix.

The optimal abstract model will be defined by:

$$\{(x_1, y_1, \epsilon_1^1, \epsilon_2^1), (x_2, y_2, \epsilon_1^2, \epsilon_2^2), \dots, (x_N, y_N, \epsilon_1^N, \epsilon_2^N)\} = \operatorname{argmin}_{\epsilon_j^i \in D} \sum_{i=1 \dots N} \|M_{\Delta}^i\|_{\text{frob}} \quad (25)$$

$\|\cdot\|_{\text{frob}}$  denotes the Frobenian norm of a square matrix.  $D$  is the base interval.

### Inverse problem solving

Eq (25) is solved iteratively. The center of each survival zone is fixed by conservation of proportions between the three distances.

On chromosome 3, the center of *HML* positions is taken at (0,0), the center of *HMR* is fixed at (5,0). If  $d_{\text{HML-MAT}}$ ,  $d_{\text{MAT-HMR}}$  and  $d_{\text{HML-HMR}}$  are experimental mean distances, the lengths  $r_0$  and  $r_1$  are given by:

$$\frac{5}{d_{\text{HML-HMR}}} = \frac{r_0}{d_{\text{HML-MAT}}} = \frac{r_1}{d_{\text{MAT-HMR}}} \quad (26)$$

The center of the *MAT* zone is defined as the intersection of two circles. The circle with (0,0) center and  $r_0$  as radius. The second circle has (5,0) as a center and  $r_1$  as radius.

Iterations are made using the variables  $\epsilon_1^1, \epsilon_2^1, \epsilon_1^2, \epsilon_2^2, \epsilon_1^3$  and  $\epsilon_2^3$ , in the range [0;5]. Our code has been parallelized on one hundred CPU cards using MPI (Message Passing Interface). The step of subdivision for the iterative resolution is taken at 0.2 to yield  $5^{12}$  abstract configurations. Each experiment represents ~100 hours of calculation; computing of survival zones for all conditions tested took 1500 hours of calculation.

## Results and Discussion

### Differentiating geometric distribution of three distinct chromosome loci

Combinations of the widely used Tet and Lac repressors, each fused to a fluorophore with a distinct emission spectrum, has proven to be a valuable tool for simultaneous visualization of two loci [18,19,30,35–37]. In order to simultaneously label three distinct loci to track their relative position, we adapted a third system based on the the bacterial  $\lambda$  repressor operator system ( $\lambda O$  and  $\lambda Ci$ ) [19] for use in yeast (Fig 1A). Operator sequences were inserted by homologous recombination near the *HML*, *MAT* and *HMR* or near *HML*, *LEU* and *MAT* loci on Chr3. Acquisition of fluorescent RFP, YFP and CFP proteins fused to the Lac, Tet and  $\lambda$  repressors, respectively, was performed in 3D.

We introduce a normalized principal components analysis (PCA) operator to convert the automatically measured x, y and z coordinates for each focus into geometrical variables (see Methods). The resulting coordinates of the formed triangle, here  $d_1$  (side of the triangle delimited by *HML*- *MAT*),  $d_2$  (*MAT*-*HMR*) and  $d_3$  (*HMR*-*HML*), were plotted on the same graph for all analyzed cells (Fig 1B). The large variation in positional combinations of the three loci reflects the highly dynamic nature of chromosome folding in yeast [38]. The *MAT* locus is mobile [18,39,40] but the two silent mating type loci, despite their frequent juxtaposition, also change position relative to each other within 10–30 seconds [19]. 50% of the most frequent positions observed for the three loci were included in a volume whose 3D surface is colored in red for *MAT $\alpha$*  and in green for *MAT $\alpha$*  cells (Fig 1B). A large fraction of the  $d_1\_d_2\_d_3$  combinations are correlated in both mating types (intersection between the red and green iso-volumes). They represent the most probable conformations of Chr3 and are as such readily detected by other methods. Strikingly, subsets of relative triangular positions of the three mating type loci

are specific to *MATa* or *MAT $\alpha$*  cells. Our goal was to characterize the folding features leading to this variant part of the distribution and to correlate them with donor preference.

### Folding of the left arm of chromosome 3 is mating type specific

The distribution of angles within the triangle formed by *MAT*, *HML* and *HMR* in *MATa* and *MAT $\alpha$*  cells was statistically significant (S1 and S2A Figs), again suggesting that certain conformations of Chr3 could be mating type specific. To extract folding features from the distribution of the three loci, we generated 2D projections of geometrical coordinates recorded for all nuclei. In the resulting density maps (Fig 2 and S2–S7 Figs), data are grouped within 10% color-coded increments of occurrence. For each dataset, nine distinct maps are generated and compared to the equivalent maps of another dataset using PCA (see Methods). Correlation coefficients ( $c$ ) between duplicate experiments using the same strain were greater than 0.8 (S2C Fig; four independent experiments  $223 < n < 559$ ). As a control, density maps resulting from a simulation of data obtained using random positioning of three loci within a sphere of 2  $\mu\text{m}$  diameter (one example is shown in Fig 2A) were significantly different from experimental data ( $c < 0.1$ ). Furthermore, the distribution of three independent loci (*MAT* on Chr3, the right telomere of Chr5 and ARS1412 on Chr14) did not correlate with the ones on Chr3 ( $c < 0.6$  ( $n = 405$ )). Thus, density maps obtained for the distribution of the three mating type loci on Chr3 are non-random. Strikingly, the maps obtained in *a*-cells were distinct from those in  $\alpha$ -cells. This was surprising because the nuclear positions of individually labelled mating type loci were previously shown not to be statistically different in *a*- and  $\alpha$ -cells [26]. Density maps differed at the 10–50% contour levels around *HML* and *MAT* (S2C Fig). The angles formed at *HML* were significantly smaller in a fraction (~30%) of  $\alpha$ -cells compared to *a*-cells (Fig 2B) suggesting that, in  $\alpha$ -cells, loci on the left arm of Chr3 are more confined with respect to those on the right arm. Also, in about 20% of cells, the angle at *MAT* was much smaller in *a*-cells than in  $\alpha$ -cells (80°–140°) for a similar distribution of *HML*-*MAT* distances. These data expose, for the first time, that positioning of *HML* relative to *MAT* and *HMR* differs between *MATa* and *MAT $\alpha$*  cells.

Geometric analysis of another combination of three loci provides additional detail (Fig 2C and S3 Fig). We labeled two loci, *HML* and *LEU2*, on the left arm and one, *MAT*, on the right. Density maps confirm that a portion of the left arm of Chr3 is largely compressed in *MATa* cells. For example, the angle at *LEU2* formed with the vector pointing towards *HML* was, in the 30% most representative cells, significantly smaller in *MATa* than in *MAT $\alpha$*  cells ( $c = 0.27$ ). This suggests that *HML* and *LEU2* roam in a similar volume with respect to *MAT* in *MATa* but not in *MAT $\alpha$*  cells. Increased constraint of the *LEU2* locus supports the view that the left arm, or at least a large region of it, can crumple and shorten in *MATa* cells, dynamically changing between an extended conformation and a transiently more compressed one.

### Relative physical constraint characterizes differential chromosome folding

Our goal was to determine whether the linkage to a chromosome fiber contributes to the relative positions of three distinct loci. We developed an abstract polymer model to define physical constraint imposed by the fiber (Eq (25) in Methods). Each nucleus can be taken as a state of the chromosome fiber. Hence, we have up to five hundred states of our system to describe our data. In an optimization approach, we assume a known configuration of Chr3 as the initial configuration (Fig 3A) to reduce the number of parameters to be estimated. The abstract model is a model based on the assumption that each point of a fiber (or polymer) moves uniformly in a specific area or survival zone (Z). To solve Eq (25) we used an iterative method based on the variables  $\epsilon_1^1, \epsilon_2^1, \epsilon_1^2, \epsilon_2^2, \epsilon_1^3$  and  $\epsilon_2^3$ , in the range [0;5]. Our code has been parallelized on one hundred CPU

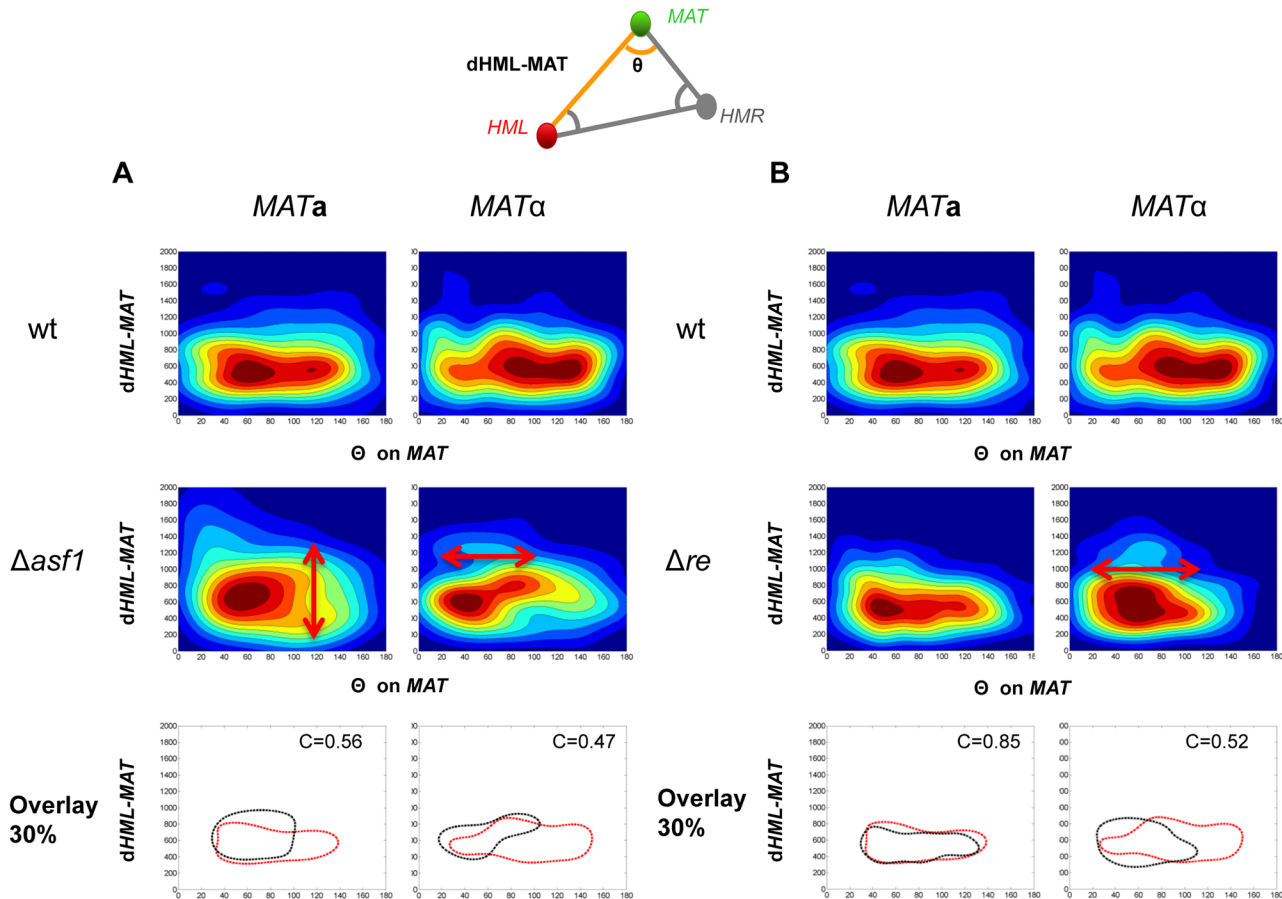
cards. The step of subdivision for the iterative resolution is taken at 0.2 to yield  $5.10^{12}$  abstract configurations. We can thus determine interactions between the survival zones  $Z_{MAT}$ ,  $Z_{HML}$  and  $Z_{HMR}$  as the extent by which two points roam in the same space while being under the physical constraint of the fiber.  $Z_{MAT}$  and  $Z_{HMR}$  partially overlap in **a**-cells (Fig 3) in agreement with the fact that chromatin in yeast is highly flexible [38], notably between these two loci 100kb apart on the same chromosome arm. In contrast,  $Z_{HML}$  is excluded from the two other survival zones. In **a**-cells,  $Z_{HML}$  is significantly greater than in  $\alpha$ -cells consistent with our finding that the left arm is more crumpled and flexible in a subset of nuclei (Fig 2).

In addition, we asked whether our abstract mathematical model could inform on physical properties of the chromatin fiber in yeast. First simulations (S4 Fig) suggest that the determined survival zones  $Z$  of the three mating type loci correspond to rather flexible, moderately constrained polymer fibers. For example, the survival zones of three linked loci corresponding to  $Z_1 = [-1;1] \times [-1;1]$ ,  $Z_2 = [-1;5] \times [-1;5]$  and  $Z_3 = [3.5;6.5] \times [-1.5;1.5]$  are closer to simulation B than to A or C (compare experimental data in Fig 3 to S4 Fig). The simulated survival zones  $Z_2$  and  $Z_3$  in our example are separated by  $<0.5$  or less. If we assign a value of  $Z = 5$  to the  $\sim 100$ kb contour length which corresponds to the separation between *MAT* ( $Z_2$ ) and *HMR* ( $Z_3$ ) the intervening fibers's flexibility is  $<10$ kb. In future work, using different sets of multiple labels at varying distances along chromosomes, the abstract mathematical model presented in this study and polymer modeling will allow to better define chromatin fiber properties.

### Chromatin structural properties control mating type specific folding of chromosome 3

We then tested whether our abstract model could be applied to predict functional consequences in mutant cells. We deleted the *asf1* gene, coding for a histone chaperone. *Asf1* was previously shown to regulate juxtaposition of the HM loci [19]. An increase in the random kinetics of the loci is expected due to general chromatin decompaction in the absence of *Asf1* [41]. All distances measured between the three loci on Chr3 in *asf1* strains increased by nearly 20% (S5 and S6 Figs;  $p < 0.007$ ). Interestingly, the distribution of angles formed around each locus was the same in wt and *asf1* mutant *MATa* but varied in *MAT $\alpha$* . This suggests that decompaction of chromatin in the absence of *Asf1* leads to an overall extension of the chromatin fiber without compromising the folding properties of Chr3 in **a**-cells. In  $\alpha$ -cells, however, density maps change to resemble those of wt *MATa* cells (Fig 4 and S5 Fig). Computing the survival zones of *HML* relative to *MAT* and *HMR* in  $\alpha$ -cells showed that *HML* is less constrained in *asf1* mutants than in wild type (Fig 3).  $Z_{MAT}$  and  $Z_{HMR}$  partially overlap and the zone of *HML* expands, again, similar to the situation determined in wt **a**-cells. Thus, the extended, stiffer conformation of the left arm of Chr3 in  $\alpha$ -cells seems to be dependent on chromatin structural features brought by *Asf1*. The lesser impact of chromatin structure in **a**-cells suggests that the folding of the left arm of Chr3 in *MATa* is intrinsic and that in  $\alpha$ -cells certain conformations are excluded due to chromatin structural properties. To test whether a greater contact probability between *HML* and the right arm of Chr3 would favor recombination, we determined donor preference in *asf1* mutant cells. We found that in  $\alpha$ -cells, usage of *HML* increased nearly four-fold in the absence of *Asf1*. Hence, reduced physical constraint of the left arm of Chr3 in a subset of cells correlates with improved recombination competence.

Finally, we asked whether the Recombination enhancer (RE), a  $<1$ kb DNA element located 17 kb to the right of *HML $\alpha$* , shown to be required for recombinational competence of a large region (40 kb) near the left end of Chr3 in *MATa* [22], is involved in folding of Chr3. Deletion of the RE region reduces the use of *HML* to repair *MATa* from  $>80\%$  to  $<10\%$  [22,23]. In *MAT $\alpha$* , the RE is in a heterochromatin configuration and non-functional, leaving the left arm

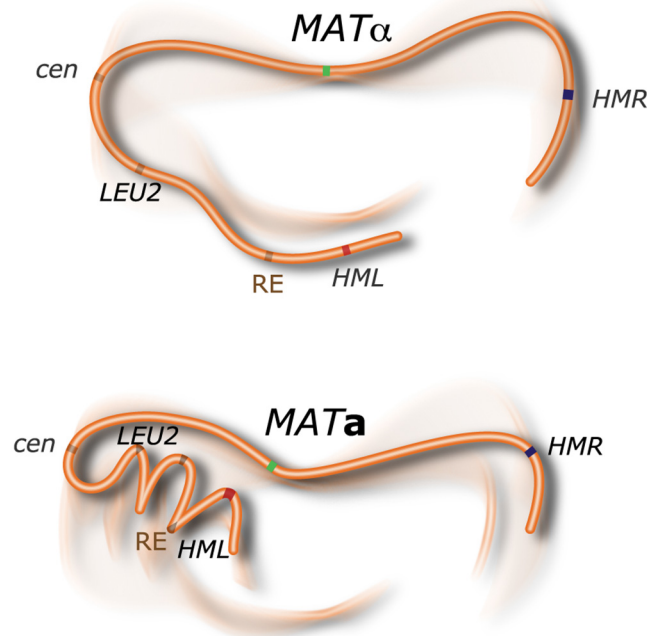


**Fig 4. Conformation of chromosome 3 is altered in *asf1* mutant cells and in the absence of the recombination enhancer element.** As in Fig 2B, *HML*, *MAT* and *HMR* were labelled. Examples of density maps ( $d = HML-MAT/\theta$  at *MAT*) are shown to compare the distribution between the 3 loci in wt and mutant cells. Red arrows highlight changes. An overlay of the contours of the density area representing 30% of the analyzed wt (black) and mutant (red) is represented. The correlation factor  $c$  is given for 30% contour. Complete sets of density maps are shown in S5 and S7 Figs. A) wt versus *asf1*. B) wt versus a strain in which the recombination enhancer element ( $\Delta re$ ) was deleted.

doi:10.1371/journal.pcbi.1004306.g004

incompetent for recombination [42]. We found that the distribution of the three loci was altered in both mating types in the absence of the RE, although the differences in density maps are more pronounced in  $\alpha$ - than in  $a$ - cells (Fig 4, S6 and S7 Figs). In  $a$ -cells, only the most frequently observed wt conformations of the three loci remained. The relative positioning of *HML* and *MAT* varied and this was more noticeable in  $\alpha$ -cells (S7 and S8 Figs). To verify that this effect was due to the RE element rather than the insertion of the hygromycin resistance gene we deleted the RE region using the *cre/lox* method. The generated density maps of the relative positions of the three loci, although in a different manner, also show differences to the wt maps. Notwithstanding that we cannot formally rule out that replacing the RE with heterologous sequences (the deletion of 803bp or the addition of 1078bp within the left arm, 29kb from the telomere) might affect chromosome folding, RE specific sequences appear to mediate some of the detected cell type dependent differences.

It was previously debated whether the RE may be involved in changing the localization or the higher-order organization of the entire left arm of Chr3 to make it more flexible in pairing with the recipient site in *MATa* cells [24,25]. Our data suggest that folding of the left arm is such that at a given moment it can be in the proximity of the *MAT* locus without being pre-



**Fig 5. Schematic representation of the folding of chromosome 3 in *MATa* and *MATα* cells—only prominent features determined in this study are included.** Labeled loci are indicated. Not to scale.

doi:10.1371/journal.pcbi.1004306.g005

folded or permanently in a recombination favorable position (see model Fig 5). Thus, in *MATa* cells, donor preference seems to only be imposed at commitment to recombination following cleavage of the *MAT* locus through recruitment of repair and recombination factors to RE elements or even synthesis of non-coding transcripts [23,43]. Strikingly, in  $\alpha$ -cells, the repressed RE element appears, at least in part, to be responsible for the extended conformation of the left arm of Chr3. Hence, the heterochromatin complex at the RE may sequester part of the chromosome, an organization that could counteract looping and may limit recombination aptitude of this chromosome arm.

## Conclusion

We present a new system for labelling and visualizing a specific chromosomal site by fluorescence microscopy in living cells. We show that a small sequence element can influence the folding of an entire chromosome. Our new methodology allows three individual chromosomal sites to be imaged simultaneously in living yeast. Quantitative data can be obtained because all cells are labeled identically and permanently. The computational strategy used to evaluate the relative distribution of three objects simultaneously in 3D represents a powerful tool for studying chromosome biology and is applicable to the analysis of any three simultaneously labelled sites in any cell type. It allows identifying transient and unstable conformations of chromosomes which are statistically not the most frequently detected ones, yet may be relevant for regulating DNA processes. This view is also supported by recent studies using polymer modeling of chromatin which revealed that fluctuations in transcriptional activity correlated with probabilistic organization of the *Tsix* gene domain [44] and with enhancer-promoter communication via modulatory chromatin looping[45]. Our method is highly complementary to genome-wide chromosome conformation capture approaches and necessary to validate models from

simulations. It is also amenable to investigation of chromosomal rearrangements governing changes in DNA-related processes in higher eukaryotes.

## Supporting Information

**S1 Fig. Variables and referential determination.** A) To change the 3D referential, geometric variables are calculated from the microscope coordinates. A single data point is represented for each nucleus by the three components  $d_1, d_2, d_3$  and their associated angles. 3D data are projected onto a unique 2D plane (eg.  $d_1/q$ ); 9 projections can be generated from each data set. Density maps (warm colors for high density and cold colors for weak density in 10% increments) are generated for each projection in 2D. 3 Examples are shown. B) Kernel density estimation to compute the probability of the density function  $f$ .

(EPS)

**S2 Fig. Geometric organization of the mating type loci in *MATa* and *MATα* cells.** Density maps of HM loci in *MATa* (YIL30,  $n = 223$ ) and *MATα* (YIL31,  $n = 276$ ) wild-type cells. A) Boxplots of pairwise distances between TetR-mRFP (*HML*), CFP-LacI (*HMR*), and YFP- $\lambda$ CI (*MAT*), foci (left panel) and the angles formed at each locus (right panel) in G1. Boxes represent the interquartile range (IQR) between lower and upper 25% quartiles, red bands correspond to the median distance or angle, and outliers are indicated by crosses. A median distribution Wilcoxon test was used to obtain p values. B) All generated maps are represented, and correspond to the triangulation between the position of TetR-mRFP (*HML*), CFP-LacI (*HMR*) and YFP- $\lambda$ CI (*MAT*) foci relative to one distance or angle at each locus. Correlation coefficients ( $c$ ) obtained from the comparison between two strains at each of the ten incremental combination frequency levels is shown in [Table 1](#). Red, dark pink and pink shades correspond to  $c < 0.5$ ,  $c < 0.6$  and  $c < 0.7$  respectively.

(EPS)

**S3 Fig. Folding of the left arm of chromosome 3 is different in *MATa* compared to *MATα* cells.** Geometric organization of *LEU*, *HML* and *MAT* in *MATa* (YIL32;  $n = 409$ ) and *MATα* (YIL33;  $n = 323$ ) wild-type cells. A) boxplots of pairwise distances between TetR-mRFP (*LEU2*), CFP-LacI (*HML*), and YFP- $\lambda$ CI (*MAT*), foci (left panel) and the angles formed at each locus (right panel) in G1. Boxes represent the interquartile range (IQR) between lower and upper 25% quartiles, red bands correspond to the median distance or angle, and outliers are indicated by crosses. A median distribution Wilcoxon test was used to obtain p values. B) All generated density maps are represented, and correspond to the triangulation between the position of the three foci relative to one distance or angle at each locus. Correlation coefficients ( $c$ ) obtained from the comparison between two strains at each of the ten incremental combination frequency levels is shown in [Table 1](#). Red, dark pink and pink shades correspond to  $c < 0.5$ ,  $c < 0.6$  and  $c < 0.7$  respectively.

(EPS)

**S4 Fig. Simulation of the relative physical constraint of three linked loci using our abstract mathematical model to assess the properties of the chromatin fiber.** Survival zones  $Z$  of three linked loci whose parameters were flexible (A), moderately constrained (B) or constrained (C) were simulated using the same iterative algorithm as for [Fig 3](#).

(EPS)

**S5 Fig. *Asf1* deletion increases the distance between the three mating type loci and affects the organization of the left arm of the chromosome 3 in *MATα* and not in *MATa* cells.** Density maps represent the geometric 3D organization of the mating type loci in wild-type and *asf1*

deleted *MATa* and *MATα*-cells.  
(EPS)

**S6 Fig. *Asf1* deletion increases the distance between the three mating type loci and affects the organization of the left arm of the chromosome 3 in *MATα* and not in *MATa* cells.** A) Boxplots represent the distances measured between the three tagged loci in wild-type and *asf1* *MATa* (n = 242) and *MATα* (n = 116) cells. Boxplots of pairwise distances between TetR-mRFP (*HML*), CFP-LacI (*HMR*) and YFP-λcI (*MAT*) foci (left panel) and the angles formed at each locus (right panel) in G1. Boxes represent the interquartile range (IQR) between lower and upper 25% quartiles, red bands correspond to the median distance or angle, and outliers are indicated by crosses. A median distribution Wilcoxon test was used to obtain p values. B) Correlation coefficients (c) obtained from the comparison between two strains at each of the ten incremental combination frequency levels is shown in [Table 1](#). Red, dark pink and pink shades correspond to  $c < 0.5$ ,  $c < 0.6$  and  $c < 0.7$  respectively.  
(EPS)

**S7 Fig. *RE* deletion alters the organization of chromosome 3 in both mating type cells differentially.** Density maps represent the geometric organization of the mating type loci in wild-type and *re* deleted *MATa* and *MATα*.  
(EPS)

**S8 Fig. *RE* deletion alters the organization of chromosome 3 in both mating type cells differentially.** A) Boxplots representing the distances between the three tagged loci and its comparison between wild-type and *re* deleted in *MATa* (n = 500) and *MATα* (n = 324). Boxplots of pairwise distances between TetR-mRFP (*HML*), CFP-LacI (*HMR*) and YFP-λcI foci (*MAT*) (left panel) and the angles formed at each locus (right panel) in G1. Boxes represent the interquartile range (IQR) between lower and upper 25% quartiles, red bands correspond to the median distance or angle, and outliers are indicated by crosses. A median distribution Wilcoxon test was used to obtain p values. B) Correlation coefficients (c) obtained from the comparison between two strains at each of the ten incremental combination frequency levels is shown in [Table 1](#). Red, dark pink and pink shades correspond to  $c < 0.5$ ,  $c < 0.6$  and  $c < 0.7$ , respectively.  
(EPS)

## Acknowledgments

We are grateful to Dr D Chatteraj for kindly providing plasmids pRFG116 and pRFB122. We thank Daniel Sage for sharing the Image J '3 spot distance' plugin, and David Villa, Photographic service of the FRBT, Toulouse, for the schematic representation of chromosome 3. We acknowledge the TRI-Genotoul Imaging platform. We benefitted from stimulating discussions with Job Dekker and his group.

## Author Contributions

Conceived and designed the experiments: KB. Performed the experiments: IL IG NTIG. Analyzed the data: IL AK IG KB. Contributed reagents/materials/analysis tools: IG IL AK. Wrote the paper: IL AK KB. Designed and programmed the algorithms used in the analysis: AK.

## References

1. Duan Z, Blau CA. The genome in space and time: Does form always follow function?: How does the spatial and temporal organization of a eukaryotic genome reflect and influence its functions? *BioEssays*. 2012; 34:800–10. doi: [10.1002/bies.201200034](https://doi.org/10.1002/bies.201200034) PMID: [22777837](https://pubmed.ncbi.nlm.nih.gov/22777837/)

2. Rosa A, Zimmer C. Computational models of large-scale genome architecture. [Internet]. 1st ed. International review of cell and molecular biology. Elsevier Inc.; 2014 [cited 2014 Dec 8]. 275–349 p. Available from: <http://www.ncbi.nlm.nih.gov/pubmed/24380598> doi: [10.1016/B978-0-12-800046-5.00009-6](https://doi.org/10.1016/B978-0-12-800046-5.00009-6) PMID: [24380598](https://pubmed.ncbi.nlm.nih.gov/24380598/)
3. Bickmore W a, van Steensel B. Genome architecture: domain organization of interphase chromosomes. Cell [Internet]. Elsevier Inc.; 2013 Mar 14 [cited 2014 Jul 15]; 152(6):1270–84. Available from: <http://www.ncbi.nlm.nih.gov/pubmed/23498936> doi: [10.1016/j.cell.2013.02.001](https://doi.org/10.1016/j.cell.2013.02.001) PMID: [23498936](https://pubmed.ncbi.nlm.nih.gov/23498936/)
4. Lieberman-Aiden E, van Berkum NL, Williams L, Imakaev M, Ragoczy T, Telling A, et al. Comprehensive mapping of long-range interactions reveals folding principles of the human genome. Science. 2009; 326:289–93. doi: [10.1126/science.1181369](https://doi.org/10.1126/science.1181369) PMID: [19815776](https://pubmed.ncbi.nlm.nih.gov/19815776/)
5. Alipour E, Marko JF. Self-organization of domain structures by DNA-loop-extruding enzymes. Nucleic Acids Res. 2012; 40(22):11202–12. doi: [10.1093/nar/gks925](https://doi.org/10.1093/nar/gks925) PMID: [23074191](https://pubmed.ncbi.nlm.nih.gov/23074191/)
6. Bohn M, Heermann DW. Repulsive forces between looping chromosomes induce entropy-driven segregation. PLoS One. 2011; 6(1):1–8.
7. Münkler C, Eils R, Dietzel S, Zink D, Mehring C, Wedemann G, et al. Compartmentalization of interphase chromosomes observed in simulation and experiment. J Mol Biol. 1999; 285:1053–65. PMID: [9887267](https://pubmed.ncbi.nlm.nih.gov/9887267/)
8. Duan Z, Andronescu M, Schutz K, Mcllwain S, Kim YJ, Lee C, et al. A three-dimensional model of the yeast genome. Nature. 2010; 465:363–7. doi: [10.1038/nature08973](https://doi.org/10.1038/nature08973) PMID: [20436457](https://pubmed.ncbi.nlm.nih.gov/20436457/)
9. Noordermeer D, Leleu M, Schorderet P, Joye E, Chabaud F, Duboule D. Temporal dynamics and developmental memory of 3D chromatin architecture at Hox gene loci. Elife [Internet]. 2014 Jan [cited 2014 Oct 17];3:e02557. Available from: <http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=4017647&tool=pmcentrez&rendertype=abstract>
10. Rodley CDM, Bertels F, Jones B, O’Sullivan JM. Global identification of yeast chromosome interactions using Genome conformation capture. Fungal Genet Biol [Internet]. Elsevier Inc.; 2009; 46(11):879–86. Available from: [10.1016/j.fgb.2009.07.006](https://doi.org/10.1016/j.fgb.2009.07.006) doi: [10.1016/j.fgb.2009.07.006](https://doi.org/10.1016/j.fgb.2009.07.006) PMID: [19628047](https://pubmed.ncbi.nlm.nih.gov/19628047/)
11. Crosetto N, Bienko M, Oudenaarden A Van. and beyond. Nat Publ Gr [Internet]. Nature Publishing Group; 2014;(December):1–10. Available from: <http://dx.doi.org/10.1038/nrg3832>
12. Nagano T, Lubling Y, Stevens TJ, Schoenfelder S, Yaffe E, Dean W, et al. Single-cell Hi-C reveals cell-to-cell variability in chromosome structure. Nature [Internet]. Nature Publishing Group; 2013 Oct 3 [cited 2014 Jul 10]; 502(7469):59–64. Available from: <http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=3869051&tool=pmcentrez&rendertype=abstract> doi: [10.1038/nature12593](https://doi.org/10.1038/nature12593) PMID: [24067610](https://pubmed.ncbi.nlm.nih.gov/24067610/)
13. Van den Engh G, Sachs R, Trask BJ. Estimating genomic distance from DNA sequence location in cell nuclei by a random walk model. Science. 1992; 257:1410–2. PMID: [1388286](https://pubmed.ncbi.nlm.nih.gov/1388286/)
14. Dekker J. Mapping in vivo chromatin interactions in yeast suggests an extended chromatin fiber with regional variation in compaction. J Biol Chem. 2008; 283:34532–40. doi: [10.1074/jbc.M806479200](https://doi.org/10.1074/jbc.M806479200) PMID: [18930918](https://pubmed.ncbi.nlm.nih.gov/18930918/)
15. Bystricky K, Heun P, Gehlen L, Langowski J, Gasser SM. Long-range compaction and flexibility of interphase chromatin in budding yeast analyzed by high-resolution imaging techniques. Proc Natl Acad Sci U S A. 2004; 101(47):16495–500. PMID: [15545610](https://pubmed.ncbi.nlm.nih.gov/15545610/)
16. Kawamura R, Tanabe H, Wada T, Saitoh S, Fukushima Y, Wakui K. Visualization of the spatial positioning of the SNRPN, UBE3A, and GABRB3 genes in the normal human nucleus by three-color 3D fluorescence in situ hybridization. Chromosom Res. 2012; 20:659–72. doi: [10.1007/s10577-012-9300-5](https://doi.org/10.1007/s10577-012-9300-5) PMID: [22801776](https://pubmed.ncbi.nlm.nih.gov/22801776/)
17. Wiggins PA, Cheveralls KC, Martin JS, Lintner R, Kondev J. Strong intranucleoid interactions organize the Escherichia coli chromosome into a nucleoid filament. Proc Natl Acad Sci U S A. 2010; 107:4991–5. doi: [10.1073/pnas.0912062107](https://doi.org/10.1073/pnas.0912062107) PMID: [20194778](https://pubmed.ncbi.nlm.nih.gov/20194778/)
18. Bystricky K, Laroche T, van Houwe G, Blaszczyk M, Gasser SM. Chromosome looping in yeast: telomere pairing and coordinated movement reflect anchoring efficiency and territorial organization. J Cell Biol. 2005; 168(3):375–87. PMID: [15684028](https://pubmed.ncbi.nlm.nih.gov/15684028/)
19. Miele A, Bystricky K, Dekker J. Yeast silent mating type loci form heterochromatic clusters through silencer protein-dependent long-range interactions. PLoS Genet. 2009; 5(5).
20. Dekker J, Rippe K, Dekker M, Kleckner N. Capturing chromosome conformation. Science. 2002; 295 (February):1306–11.
21. Haber JE. Mating-type genes and MAT switching in Saccharomyces cerevisiae. Genetics. 2012; 191:33–64. doi: [10.1534/genetics.111.134577](https://doi.org/10.1534/genetics.111.134577) PMID: [22555442](https://pubmed.ncbi.nlm.nih.gov/22555442/)
22. Wu X, Haber JE. A 700 bp cis-Acting Region Controls Mating-Type Dependent Recombination Along the Entire Left Arm of Yeast Chromosome III. 1996; 87:277–85. PMID: [8861911](https://pubmed.ncbi.nlm.nih.gov/8861911/)

23. Szeto L, Fafalios MK, Zhong H, Vershon AK, Broach JR.?? 2p controls donor preference during mating type interconversion in yeast by inactivating a recombinational enhancer of chromosome III. *Genes Dev.* 1997; 11:1899–911. PMID: [9271114](#)
24. Simon P, Houston P, Broach J. Directional bias during mating type switching in *Saccharomyces* is independent of chromosomal architecture. 2002; 21(9). PMID: [11980725](#)
25. Bressan D a, Vazquez J, Haber JE. Mating type-dependent constraints on the mobility of the left arm of yeast chromosome III. *J Cell Biol* [Internet]. 2004 Feb 2 [cited 2014 Sep 9]; 164(3):361–71. Available from: <http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=2172233&tool=pmcentrez&rendertype=abstract> PMID: [14745000](#)
26. Bystricky K, Van Attikum H, Montiel M-D, Dion V, Gehlen L, Gasser SM. Regulation of nuclear positioning and dynamics of the silent mating type loci by the yeast Ku70/Ku80 complex. *Mol Cell Biol.* 2009; 29(3):835–48. doi: [10.1128/MCB.01009-08](#) PMID: [19047366](#)
27. Coïc E, Richard G-F, Haber JE. *Saccharomyces cerevisiae* donor preference during mating-type switching is dependent on chromosome architecture and organization. *Genetics* [Internet]. 2006 Jul [cited 2014 Oct 18]; 173(3):1197–206. Available from: <http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=1526691&tool=pmcentrez&rendertype=abstract> PMID: [16624909](#)
28. Fekete R a, Chattoraj DK. A cis-acting sequence involved in chromosome segregation in *Escherichia coli*. *Mol Microbiol* [Internet]. 2005 Jan [cited 2014 Oct 18]; 55(1):175–83. Available from: <http://www.ncbi.nlm.nih.gov/pubmed/15612926> PMID: [15612926](#)
29. Lassadi I, Bystricky K. Tracking of single and multiple genomic loci in living yeast cells. *Methods Mol Biol.* 2011; 745:499–522. doi: [10.1007/978-1-61779-129-1\\_29](#) PMID: [21660713](#)
30. Straight AF, Belmont AS, Robinett CC, Murray AW. GFP tagging of budding yeast chromosomes reveals that protein–protein interactions can mediate sister chromatid cohesion. *Curr Biol* [Internet]. 1996 Dec; 6(12):1599–608. Available from: <http://linkinghub.elsevier.com/retrieve/pii/S0960982202707835> PMID: [8994824](#)
31. Longtine MS, Iii AMK, Demarini DJ, Shah NG. Additional Modules for Versatile and Economical PCR-based Gene Deletion and Modification in *Saccharomyces cerevisiae*. 1998; 961(February):953–61. PMID: [9717241](#)
32. Connolly B, White CI, Haber JE. Physical monitoring of mating type switching in *Saccharomyces cerevisiae*. *Mol Cell Biol.* 1988; 8(6):2342–9. PMID: [2841579](#)
33. Parzen E. On Estimation of a Probability Density Function and Mode Emanuel Parzen. *Ann Math Stat.* 1962; 33:1065–76.
34. Botev ZI, Grotowski JF, Kroese DP. Kernel density estimation via diffusion. *The Annals of Statistics.* 2010. p. 2916–57.
35. Robinett CC, Straight A, Li G, Wilhelm C, Sudlow G, Murray A, et al. In vivo localization of DNA sequences and visualization of large-scale chromatin organization using lac operator/repressor recognition. *J Cell Biol.* 1996; 135:1685–700. PMID: [8991083](#)
36. Roukos V, Voss TC, Schmidt CK, Lee S, Wangsa D, Misteli T. Spatial dynamics of chromosome translocations in living cells. *Science* [Internet]. 2013; 341:660–4. Available from: <http://www.ncbi.nlm.nih.gov/pubmed/23929981> doi: [10.1126/science.1237150](#) PMID: [23929981](#)
37. Soutoglou E, Dorn JF, Sengupta K, Jasin M, Nussenzweig A, Ried T, et al. Positional stability of single double-strand breaks in mammalian cells. *Nat Cell Biol* [Internet]. 2007 Jun [cited 2014 Oct 5]; 9(6):675–82. Available from: <http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=2442898&tool=pmcentrez&rendertype=abstract> PMID: [17486118](#)
38. Hajjoul H, Mathon J, Ranchon H, Goiffon I, Mozziconacci J, Albert B, et al. High-throughput chromatin motion tracking in living yeast reveals the flexibility of the fiber throughout the genome. *Genome Res.* 2013; 23:1829–38. doi: [10.1101/gr.157008.113](#) PMID: [24077391](#)
39. Hajjoul H, Kocanova S, Lassadi I, Bystricky K, Bancaud A. Lab-on-Chip for fast 3D particle tracking in living cells. *Lab Chip.* 2009; 9(21):3054–8. doi: [10.1039/b909016a](#) PMID: [19823719](#)
40. Nagai S, Dubrana K, Tsai-Pflugfelder M, Davidson MB, Roberts TM, Brown GW, et al. Functional targeting of DNA damage to a nuclear pore-associated SUMO-dependent ubiquitin ligase. *Science* [Internet]. 2008 Oct 24 [cited 2014 Oct 1]; 322(5901):597–602. Available from: <http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=3518492&tool=pmcentrez&rendertype=abstract> doi: [10.1126/science.1162790](#) PMID: [18948542](#)
41. Corpet A, Almouzni G. Making copies of chromatin: the challenge of nucleosomal organization and epigenetic information. *Trends Cell Biol* [Internet]. 2009 Jan [cited 2014 Aug 27]; 19(1):29–41. Available from: <http://www.ncbi.nlm.nih.gov/pubmed/19027300> doi: [10.1016/j.tcb.2008.10.002](#) PMID: [19027300](#)
42. Weiss K, Simpson RT. Cell type-specific chromatin organization of the region that governs directionality of yeast mating type switching. 1997; 16(14):4352–60. PMID: [9250679](#)

43. Coïc E, Sun K, Wu C, Haber JE. Cell cycle-dependent regulation of *Saccharomyces cerevisiae* donor preference during mating-type switching by SBF (Swi4/Swi6) and Fkh1. *Mol Cell Biol* [Internet]. 2006 Jul [cited 2014 Oct 18]; 26(14):5470–80. Available from: <http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=1592702&tool=pmcentrez&rendertype=abstract> PMID: 16809780
44. Giorgetti L, Galupa R, Nora EP, Piolot T, Lam F, Dekker J, et al. Predictive polymer modeling reveals coupled fluctuations in chromosome conformation and transcription. *Cell*. 2014; 157(4):950–63. doi: [10.1016/j.cell.2014.03.025](https://doi.org/10.1016/j.cell.2014.03.025) PMID: 24813616
45. Doyle B, Fudenberg G, Imakaev M, Mirny L a. Chromatin Loops as Allosteric Modulators of Enhancer-Promoter Interactions. *PLoS Comput Biol* [Internet]. 2014; 10(10):e1003867. Available from: <http://dx.plos.org/10.1371/journal.pcbi.1003867> doi: [10.1371/journal.pcbi.1003867](https://doi.org/10.1371/journal.pcbi.1003867) PMID: 25340767