Chromosome contact maps are bifractal

S. Pigolotti[1], M. H. Jensen[2], Y. Zhan[3,4], and G. Tiana[5]

[1]*Biological Complexity Unit, Okinawa Institute of Science and Technology Graduate University, Onna, Okinawa 904-0495, Japan.*

[2]*The Niels Bohr Institute, Blegdamsvej 16, 2100 Copenhagen Ø, Denmark.*

[3]*Friedrich Miescher Institute for Biomedical Research, Maulbeerstrasse 66, 4058 Basel, Switzerland.*

[4]*University of Basel, Petersplatz 1, 4001 Basel, Switzerland*

[5]*Department of Physics and Center for Complexity and Biosystems, Università degli Studi di Milano and INFN, via Celoria 16, 20133 Milano, Italy.*

Modern biological techniques such as Hi–C permit to measure probabilities that different chromosomal regions are close in space. These probabilities can be visualised as matrices called contact maps. Contact maps are characterized by self-similar blocks along the diagonal, corresponding to a hierarchy of tightly packed chromosomal domains. Statistical properties of these maps, and thus of chromosomal interactions, are usually summarized by the scaling of the contact probability as a function of the genomic distance. This approach is fraught with limitations, since such scaling is usually rather noisy. Here, we introduce a multifractal analysis of chromosomal contact maps. Our analysis reveals that Hi–C maps are bifractal, i.e. they are complex geometrical objects characterized by two distinct fractal dimensions. To rationalize this observation, we introduce a model that describes chromosomes

as a hierarchical set of domains nested in each other and we solve it exactly. The predicted multifractal spectrum is in excellent quantitative agreement with experimental data. Our theory yields to a more robust estimation of the scaling exponent of the contact probability than existing methods. By applying this method to experimental data, we detect subtle conformational changes among chromosomes during differentiation of human stem cells.

During cellular interphase, mammalian chromosomes assume a globular structure in the nucleus [1,2]. Their conformational properties can be studied *in vivo* with a set of techniques called chromosome conformation capture (3C) [3], most notably their genome-wide version called Hi–C [4]. Results of Hi–C experiments can be represented as matrices whose elements are proportional to the probability that two chromosome regions are in contact in space [5]. Hi–C measurements have paved the way for a mechanistic understanding of chromosome folding. In particular, they have revealed that mammalian chromosomes are characterized by a hierarchy of nested, tightly connected structures [9]. Structures below the megabase scale are usually called topological associating domains (TADs). They are particularly relevant for gene regulation [6,8], although they do not seem to be privileged over other levels in the hierarchy from a structural point of view [9]. It has been suggested that the hierarchical structure of Hi–C matrices can be studied by means of an iterative coarse-graining procedure [10].

Information contained in Hi–C matrices is often characterized by the decay of the contact probability $p(\ell)$ of pairs of chromosomal regions i and j with respect to their genomic distance

$\ell \equiv |i - j|$. Such decay appears to follow a power law

$$p(\ell) \sim \ell^{-\beta}. \tag{1}$$

The contact probability exponent β is often estimated to be slightly smaller than one [4,11,12,26]. Such low values are incompatible with simple equilibrium homopolymeric models [13]. In contrast, non-equilibrium models such as the crumpled globule [4,13] are able to account for such low exponents. Other proposed mechanisms include mediation of polymer interaction by other molecules [14], active loop-extrusion [15], and finite–size effects in heteropolymers [16]. In any case, the apparent power-law range of the contact probability is usually limited to one decade or less, see e.g. [4,11,12]. Therefore, estimates of the exponent β are rather sensitive to the choice of the fitting range.

In this paper, we robustly characterize scale-invariant properties of chromosomes using the theory of multifractals. This theory has been developed to characterize heterogeneous systems characterized by scale invariance [18–20]. This analysis reveals that chromosome contact maps are *bifractal*, i.e. multifractal objects characterized by two distinct fractal dimensions, a behavior previously reported in studies of surface roughness [21], distribution of matter in the universe [22] and turbulence [23].

We introduce our theory using a Hi–C map of chromosome 1 in mouse embryonic stem cells, see Fig. 1**a**. The contact probability seems to decay as a power law, at least for relatively short genomic distances, see Fig. 1**b**. However, the local logarithmic slope of the contact probability does not present the clear plateau characteristic of a power law behavior, see Fig. 1**c**.

To characterize scaling properties of chromosomes in a more rigorous way, we study Hi–C maps as multifractals. A multifractal is a system described in terms of a density, that in our case is the density of number of counts in the Hi–C map. We study the scaling property of this density by constructing two-dimensional histograms of the Hi–C map on grids of different linear resolution ϵ, see Fig. 1**d**. Geometrical structures of linear size smaller than ϵ are not resolved in these maps. The smallest possible value of ϵ is the resolution of the original Hi–C map, in our case $\epsilon = 4 \cdot 10^4$ bp. We define the probability $p_{ij}(\epsilon)$ in bin at coordinates i, j at resolution ϵ. We always work with normalized maps, so that $\sum_{ij} p_{ij}(\epsilon) = 1$ for all choices of ϵ.

We assume that the system is scale invariant, at least for ϵ small enough. This means that the probability in a bin scales as a power law of the resolution:

$$p_{ij}(\epsilon) \sim \epsilon^{\alpha}, \tag{2}$$

where $\sim$ denotes the leading behavior and α is the scaling exponent associated to the density. Since the map is not homogeneous, different bins can be in principle characterized by different values of α. The number of bins $N(\alpha)$ associated to a given value of α must also scales as a power of ϵ

$$N(\alpha) \sim \rho(\alpha)\epsilon^{-D(\alpha)}. \tag{3}$$

The exponent $D(\alpha)$ characterizes the scaling of the number of bins of linear size ϵ necessary to cover the set with density exponent α and therefore can be interpreted as the fractal dimension associated to this set. The quantity $\rho(\alpha)$ is a prefactor independent of ϵ.

Computing $D(\alpha)$ directly is often unpractical. A convenient related quantity is the partition

function $Z(q,\epsilon)$, defined by [19,20]

$$Z(q,\epsilon) = \sum_{ij} \left[p_{ij}(\epsilon)\right]^q . \tag{4}$$

The name "partition function" originates from an analogy with equilibrium statistical mechanics, where the exponent q plays the role of an inverse temperature. Indeed, in an inhomogeneous system, for $q \to 0$ (high temperature), all bins give similar contributions to the sum in Eq. (4), whereas for large q (low temperature) the sum is dominated by relatively few terms characterized by largest values of the measure p_{ij}. We link the partition function with the fractal dimensions by collecting in Eq. (4) all terms with the same value of α:

$$Z(q,\epsilon) \sim \sum_{\alpha} \rho(\alpha)\epsilon^{q\alpha - D(\alpha)}. \tag{5}$$

In the limit of small ϵ, we evaluate this sum with the saddle point method. This calculation shows that, for a scale-invariant system, one also expects a power-law scaling of the partition function as a function of the resolution

$$Z(q,\epsilon) \sim \epsilon^{K(q)} \qquad \epsilon \to 0. \tag{6}$$

The function $K(q)$ is the *multifractal spectrum* and characterizes the scaling behavior with respect to the resolution of the partition function [20]. The saddle point evaluation of Eq. (5) leads to conclude that the dimensions $D(\alpha)$ and the multifractal spectrum $K(q)$ are related by a Legendre transform:

$$K(q) = \max_{\alpha}[q\alpha - D(\alpha)]. \tag{7}$$

A linear multifractal spectrum indicates that the system is homogeneous, i.e. all of its parts are characterized by the same fractal dimension $\alpha = D$ and thus by the same scaling behavior $N \sim$

ϵ^{-D} and $p_{ij} \sim \epsilon^{D}$. Conversely, a non-linear multifractal spectrum signals a diversity of scaling exponents and associated dimensions.

The partition function $Z(q, \epsilon)$ computed from Hi–C data presents a clear power law behavior as a function of ϵ, as predicted by Eq. (6), see Fig. 1**e**. In this case, for ϵ slightly larger than its minimum value, the local logarithmic slope of the partition function is essentially flat, see Fig. 1**f**. This means that the power–law behaviour of $Z(q, \epsilon)$ is much more robust than that of $p(\ell)$.

The multifractal spectrum $K(q)$ of chromosome 1 presents two different linear regimes, one at low moments with the properties of a two-dimensional object and one at high moments characterized by a fractal dimension between one and two, see Fig. 1**g**. A system with such properties is called a bifractal [21,22,22]. Since the exponent q is analogous to an inverse temperature, the sharp change of slope in the spectrum of a bifractal system is akin to a phase transition in equilibrium statistical physics [24].

To rationalize these observations, we introduce a hierarchical domain model of Hi–C maps. We define the model in terms of an iterative transformation of a measure on the unit square $[0, 1] \times [0, 1]$. At the first iteration, the measure is given by a 2×2 matrix with diagonal element a and off–diagonal elements b, with $a > b > 0$. At each following iteration n we generate a $2^n \times 2^n$ matrix. The off-diagonal element of the matrix at the previous iteration becomes a 2×2 block of identical values in the matrix at the nth iteration. The diagonal blocks are further multiplied by the original matrix (see Fig. S1 in the Supp. Mat). We impose $b = 1/2 - a$, so that the measure remains normalized at each iteration. This means that, effectively, the model is defined in terms

of a single free parameter a. Because of the normalization and the requirement that the measure should be concentrated along the diagonal, such parameter falls in the range $1/4 \leq a \leq 1/2$.

The parameter a represents the weight of domains compared to the rest of the Hi–C map. Matrices with larger a have a measure more concentrated along the diagonal, whereas matrices with lower a are characterized by a more uniform measure, see Fig. 2**a**. In the limiting case $a = 1/4$ the matrix is uniform at all iterations, whereas for $a = 1/2$ the matrix is characterized by a uniform measure on the diagonal, whereas all the off-diagonal elements are zero.

The partition function of the hierarchical domain model can be expressed as

$$Z(q,n) = \sum_{k=0}^{n} \exp[\xi(k)], \tag{8}$$

where we defined the exponent

$$\begin{aligned} \xi(k) &= [n + \max(n-k-1,0)] \ln 2 + kq \ln a + \\ &+ (1-\delta_{kn}) q \ln b - \max(n-k-1,0) q \ln 4. \end{aligned} \tag{9}$$

A saddle–point estimation of the partition function gives

$$\frac{d\xi}{dk} = -\ln 2 + q \ln 4a, \tag{10}$$

which is positive for $q > q_c = (\ln 2)/(\ln 4a)$. This means that, for $q \geq q_c$, the leading term is either $\xi(n)$ or $\xi(n-1)$. Since $\xi(n) - \xi(n-1) = q \ln(a/b)$, the maximum of the exponent is attained at $k = n$. Thus, for large n, the partition function scales as $Z(q,n) \sim \epsilon^{K(q)}$ with the length scale $\epsilon = 2^{-n}$ and the multifractal spectrum

$$K(q) = -q \frac{\ln a}{\ln 2} - 1. \tag{11}$$

For $a = 1/2$, the matrix $p_{ij}(\epsilon)$ is diagonal at each iteration. In this case, Eq. (11) predicts a linear spectrum with slope $D = 1$, consistent with the fact that the geometric set is equivalent to a one dimensional line. For $a = 1/4$ the distribution is uniform on the square, and Eq. (11) correctly returns $D = 2$. Between these two limiting cases, Eq. (11) predicts a fractal distribution with a dimension $D = -\ln a/\ln 2$ between 1 and 2.

We now focus on the "high temperature phase" where $-\ln 2 + q \ln 4a < 0$. In this case, the term dominating the scaling is $k = 0$, so that

$$K(q) = 2(q - 1). \tag{12}$$

Since in the high temperature phase the scaling is determined by the terms far from the diagonal, the spectrum is that of a regular two-dimensional set.

Summarizing, our calculation predicts a multifractal spectrum characterized by two linear regimes: one with slope $-\log(a)/\log(2)$ for $q > q_c = (\ln 2)/(\ln 4a)$, Eq. (11), and one with slope 2 for $q < q_c$, Eq. (12). Such predictions are in excellent agreement with numerical simulations, as shown in Fig. 2**a**, with very small discrepancies for high values of q and low values of a arising from finite size effects. These results confirm the validity of our saddle point approximation.

Strikingly, our theory predicted extremely well also the multifractal spectra of real chromosomes , with a value of $a \approx 0.425$ and very little variability among the first three chromosomes of mouse embryonic stem cells (mean squared displacement MSD ≈ 0.0015 in all three cases), see Fig. 2**c**. To verify whether the multifractal spectrum in Fig.2 is truly a signature of a hierarchical

mechanism, we computed numerically the spectra of two null models. In both null models, the contact probability decays as a power law of the distance, to which we add domains at a single length scale M in the first null model and random noise in the second. Additional details on the null models are presented as Supplementary Information. The solution of Eqs. (11) and (12) provides a poor fit to this first null model (MSD in the range $0.01 \sim 0.02$ depending on parameters, see Supplementary Fig. S2). The second null model provides a better fit, with MSD ≈ 0.0015 for small noise and and $\beta \approx 1$, see Supplementary Fig. S2. This suggests that a pure power–law behavior with the addition of noise provides an alternative explanation for the observed scaling behavior.

Our theory accurately describes also Hi–C maps of Drosophila chromosomes (Supplementary Fig. S3, average MSD ≈ 0.001 for the first four chromosomes) and of human chromosomes (Supplementary Fig. S4, average MSD ≈ 0.005 for all chromosomes except chromosome Y). These observations support that the bifractal spectrum predicted by the hierarchical domain model is compatible with that observed in a broad class of higher organisms

From the model, we also predict the scaling of the contact probability with the genomic distance. In our framework, the contact probability $P(\ell; n)$ at the n-th iteration can be expressed by summing the probabilities of blocks at a genomic distance ℓ from the diagonal:

$$P(\ell; n) = \sum_{ij} p_{ij}(n) \delta_{|i-j|,\ell}. \tag{13}$$

From this expression, we derive that the decay exponent of the contact probability with the genomic

distance is

$$\beta = \frac{\ln(4a)}{\ln(2)}. \tag{14}$$

Derivation of Eq. (14) is presented in SI. For $1/2 \leq a \leq 1$, Eq. (14) predicts contact probability exponents in the range $\beta \in [0, 1]$. This prediction is supported by numerical simulations, see Fig. S3 in the Supp. Mat. This means that the hierarchical structure of contact matrices automatically leads to exponents $\beta < 1$, as observed in chromosomes.

We test more extensively our method using datasets from different Hi–C experiments using mouse embryonic stem cells (mESC). We fit the multifractal spectrum of each dataset and obtain the corresponding exponent β via Eq. (14). We first test the robustness of the multifractal analysis across two Hi–C experiments in mESC from two different labs [17,30], both following the Hi–C protocol *in solution*, see Fig. 3**a**. Both the multifractal and the power law methods predict that the variability of the exponent β across chromosomes is significantly larger than the variability of the exponent across replicates. To figure out which method performs best, we use bootstrapping to test against the null hypothesis that the values of β of different chromosomes are paired at random. The multifractal method permits to exclude random pairing of chromosomes (p–value $5.2 \cdot 10^{-4}$) much better than the direct fit (p–value 0.11, see Supplementary Table S1). Moreover, the mean squared error of β between the two replicates is smaller in the multi–fractal case compared to direct fit (0.009 vs 0.04). The χ^2 associated with the direct fit is affected by a strong systematic error, although remaining quite correlated (see Supplementary Fig. S6). This effect is much milder in the multifractal approach. The multifractal and direct fit methods are similarly robust with respect to varying the resolution of Hi–C maps (see Supplementary Fig. S7).

We apply the two methods to attempt to distinguish among experiments on the same cell lines, but following different experimental protocols. These different protocols mainly differ in the ligation step of digested DNA fragments. In the original "in solution" protocol, the ligation is performed in a diluted solution [36]. In other protocols, the ligation step is either carried out in intact nuclei ("in situ" protocol [36]) or on solid substrates ("TCC" protocol [32]). The two latter protocols are able to produce Hi-C maps with better signal-to-noise ratio compared to the in solution protocol [32,36]. Both the multifractal and the direct fit methods show that the values of β obtained from in situ and TCC experiments are markedly different from those obtained in solution (see Supplementary Fig. S6). The multifractal method estimates compatible scaling exponents for in situ and TCC protocols (the p–value associated with the null hypothesis that chromosomes are paired at random is 0.021, see Supplementary Table S2), while a direct fit does not permit to draw this conclusion (p–value 0.14). This result suggests that the in situ and TCC protocols are able to produce statistically compatible Hi–C maps due to their high signal-to-noise ratio.

We compare the two methods in detecting differences between wild–type and mutant cell lines, in which either the Nipbl gene [31] or the CTCF gene [30] is knocked down (see Fig. 3**b**). Knock–down of these genes has been shown to disrupt chromosome folding. In particular, Nipbl knock–down leads to loss of TAD structures and global changes in scaling properties[31]. We quantify average differences in exponents between the wild–type and the mutant in terms of χ^2 of pairs of chromosomes, weighted by their mean squared errors (see Supplementary Table S2 and Supplementary Fig. S7). Multifractal analysis detects more marked differences ($\chi^2 = 214$) compared to the direct fit ($\chi^2 = 4.7$), although both methods highlight a statistically significant difference

between the two sets (p–values $1.3 \cdot 10^{-5}$ for the direct fit and $< 10^{-6}$ for the multifractal). Knock–down of CTCF also causes a loss of TAD structure, but without a clear effect on genome–wide scaling [30]. Nonetheless, the values of β of different chromosomes obtained by multifractal analysis reveal a large ($\chi^2 = 86$) and significant (p–value $< 10^{-6}$) difference between the scaling properties of wild–type and CTCF–deficient cells. Also in this case, the direct fit detects less marked differences ($\chi^2 = 25$, see Supplementary Table S2).

We apply the multifractal analysis to elucidate how chromosome structure changes during cellular differentiation. To this aim, we analyze Hi–C data obtained from different human cell lines at different stages of early differentiation [35]. In temporal order, the cell lines are: trophectoderm, embyonic stem cells (ESC), mesoderm, neuronal precursor cells (NPC) and mesenchymal cells. Mesoderm and NPC can be regarded as in similar development stages. The value of β obtained by multifractal analysis (see Fig. 3**c**) tends to increase upon differentiation(p–value of Kendall's tau test applied to all chromosome pairs $< 10^{-11}$). A similarly clear conclusion cannot be obtained from the values of β estimated by direct fit (see Fig. 3**d**, p–value 0.07). The increase of β points to the idea that more differentiated cell types require more insulated domains to achieve more specialised function.

Multifractal analysis of Hi–C data is a powerful statistical tool to characterize scaling properties of chromosomes. We have shown that contact maps of chromosomes are bifractal, i.e. they are characterized by two distinct fractal dimensions. To understand this observation, we proposed a hierarchical domain model, whose analytical solution is in striking quantitative agreement with

observed multifractal spectra. Our theory implicates a power-law scaling of the contact probability with the genomic distance with exponents β lower than one, also in agreement with experimental observations. The multifractal method is sensitive enough to discard a null model with power-law decay of the contact probability, but domains on a single length scale only. The predicted form of the multifractal spectrum provides a more stringent prediction than the contact probability exponent alone. The analysis proposed here constitutes an excellent benchmark to select among different polymer models that provide similar values of β [13–16].

Our results indicate that scaling properties of chromosomes are a direct consequence of the hierarchical structure of chromosome domains [9,27]. Recent work has suggested that such domain structure is generated by a ”hierarchical folding” mechanism, mediated by different proteins such as cohesin and CTCF [28]. The precise mechanism driving the folding of chromosomes has been subject of debate [29]. It will be interesting to test whether the activity of these factors can produce self-similar structures compatible with our observations. The shape of the multifractal spectrum is controlled by a single parameter, that also controls the contact probability exponent β via Eq. (14). The determination of β from the fit of the multifractal spectrum is a much more robust way of characterizing the Hi–C map than the direct fit of β. Multifractal analysis is importantly also more sensitive in highlighting subtle structural differences as shown in the case of Nipbl and CTCF knock–down. The analysis of the multifractal spectrum is simple and robust enough to become a routinely tool to analyze contact maps, both from polymer models and experiments, and capture subtle differences among them.

1. L. Giorgetti, R. Galupa, E. P. Nora, T. Piolot, F. Lam *et al.*, *Predictive polymer modeling reveals coupled fluctuations in chromosome conformation and transcription.* Cell **157**, 950 (2014).

2. G. Tiana, A. Amitai, T. Pollex, T. Piolot, D. Holcman *et al. Structural Fluctuations of the Chromatin Fiber within Topologically Associating Domains.* Biophys J. **110**, 1234 (2016)

3. J. Dekker, K. Rippe, M. Dekker and N. Kleckner, *Capturing chromosome conformation.* Science, **295**, 1306 (2002).

4. E. Lieberman-Aiden, N. L. van Berkum, L. Williams, M. Imakaev, T. Ragoczy, A. Telling *et al.*, *Comprehensive mapping of long-range interactions reveals folding principles of the human genome.* Science **326**, 289 (2009).

5. J. Redolfi, Y. Zhan, C. Valdes, M. Kryzhanovska, I. Misteli Guerreiro *et al.*, *DamC reveals principles of chromatin folding in vivo without crosslinking and ligation.* Nature Str. Mol. Biol. **26**, 471 (2019)

6. Q. Szabo, et al *Principles of genome folding into topologically associating domains*, Science Advances, **5**, (2019)

7. J. Nuebler, G. Fudenberg, M. Imakaev, N Abdennur and L. A. Mirny, *Chromatin Organization by an Interplay of Loop Extrusion and Compartmental Segregation.* Proc. Natl. Acad. Sci. USA **115**, E6697 (2018).

8. D. G. Lupiáñez, K. Kraft, V. Heinrich, P. Krawitz, F. Brancati, E. Klopocki *et al. Disruptions of topological chromatin domains cause pathogenic rewiring of gene-enhancer interactions.* Cell **161**, 1012 (2015).

9. Y. Zhan, L. Mariani, I Barozzi, E. G. Schulz, N. Blüthgen *et al.*, *Reciprocal insulation analysis of Hi-C data shows that TADs represent a functionally but not structurally privileged scale in the hierarchical folding of chromosomes.* Genome Research **27**, 479 (2017).

10. A. M. Chiariello, S. Bianco, C. Annunziatella, A. Esposito and M. Nicodemi, *The scaling features of the 3D organization of chromosomes are highlighted by a transformation la Kadanoff of Hi-C data.* Europhys. Lett. **120** 40004 (2017).

11. A. L. Sanborn, S. S. P. Rao, S.-C. Huang, N. C. Durand *et al. Chromatin extrusion explains key features of loop and domain formation in wild-type and engineered genomes.* Proc. Natl. Acad. Sci. USA **112**, E6456 (2017).

12. B. Bonev, N. M. Cohen Q. Szabo, L. Fritsch, G. L. Papadopoulos *et al.*, *Multiscale 3D Genome Rewiring during Mouse Neural Development.* Cell **171**, 557.e1 (2017).

13. L. Mirny, *The fractal globule as a model of chromatin architecture in the cell.* Chrom. Res. **19**, 37 (2011).

14. M. Barbieri, M. Chotalia, J. Fraser, L. M. Lavitas, J. Dostie, A. Pombo and M. Nicodemi, *Complexity of chromatin folding is captured by the strings and binders switch model.* Proc. natl. Acad. Sci. USA **109**, 16173 (2012).

15. G. Fudenberg, M. Imakaev, C. Lu, A. Goloborodko, N. Abdennur and L. Mirny, *Formation of Chromosomal Domains by Loop Extrusion.* Cell Rep. **15**, 2038 (2016).

16. Y. Zhan, L. Giorgetti and G. Tiana, *Looping probability of random heteropolymers helps to understand the scaling properties of biopolymers.* Phys. Rev. E **94**, 032402 (2016).

17. L. Giorgetti, B. R. Lajoie, A. C. Carter, M. Attia, *et al.*, *Structural organization of the inactive X chromosome in the mouse.* Nature **535**, 575 (2016).

18. G. Parisi and U. Frisch in ”Turbulence and Predictability in Geophysical Fluid Dynamics”, p. 8487, eds. M. Ghil et al (1985).

19. R. Benzi, G. Paladin, G. Parisi, and A. Vulpiani, *On the multifractal nature of fully developed turbulence and chaotic systems.* Journal of Physics A **17**, 3521 (1984).

20. T.C. Halsey, M.H. Jensen, L.P. Kadanoff, I. Procaccia, and B. Shraiman, *Fractal measures and their singularities: The characterization of strange sets.* Phys. Rev. A **33**, 1141 (1986).

21. Bhushan B, Majumdar A. *Elastic-plastic contact model for bifractal surfaces.* Wear, **153**(1), 53-64. (1992).

22. Balian, R., and R. Schaeffer. *Scale-invariant matter distribution in the universe. ii-bifractal behaviour.* Astronomy and Astrophysics **226**, 373-414 (1989).

23. Iyer, K.P., Sreenivasan, K.R. and Yeung, P.K., *Circulation in high Reynolds number isotropic turbulence is a bifractal.* arXiv preprint arXiv:1902.07326 (2019).

24. T. Bohr and M.H. Jensen, *Order parameter, symmetry breaking, and phase transitions in the description of multifractal sets.* Phys. Rev. A **36**, 4904 (1987).

25. G. Tiana and L. Giorgetti, *Integrating experiment, theory and simulation to determine the structure and dynamics of mammalian chromosomes.* Curr. Opin. Struct. Biol. **49**, 11 (2018).

26. Y. Zhan, L. Giorgetti and G. Tiana, *Modelling genome-wide topological associating domains in mouse embryonic stem cells.* Chromosome Res. **25** 5 (2017).

27. S. S. P. Rao, M. H. Huntley, N. C. Durand, *et al*, *A 3D map of the human genome at kilobase resolution reveals principles of chromatin looping.* Cell, **159** 1665 (2014).

28. B. Bonev and G. Cavalli, *Organization and function of the 3D genome.* Nat. Rev. Gen. **17**, 661 (2016).

29. M. J. Rowley and V. G. Corces, *Organizational principles of 3D genome architecture.* Nat. Rev. Gen. **19**, 789 (2018).

30. E. P. Nora, A. Goloborodko, A. L. Valton, J. H. Gibcus, A. Uebersohn, N. Abdennur *et al. Targeted Degradation of CTCF Decouples Local Insulation of Chromosome Domains from Genomic Compartmentalization.* Cell, **169**, 930 (2017).

31. W. Schwarzer, N. Abdennur, A. Goloborodko, A., Pekowska, G. Fudenberg, Y. Loe-Mie *et al Two independent modes of chromatin organization revealed by cohesin removal.* Nature, **152**, 1270 (2017).

32. R. Kalhor, H. Tjong, N. Jayathilaka, F. Alber and L. Chen, *Genome architectures revealed by tethered chromosome conformation capture and population-based modeling.* Nature Biotechnology **30**, 90 (2012).

33. J. Redolfi, Y. Zhan, C. Valdes-Quezada, *et al DamC reveals principles of chromatin folding in vivo without crosslinking and ligation.* Nat. Struct. Mol. Biol, **26**, 471-480, (2019).

34. E. Lieberman-Aiden , *et al Comprehensive mapping of long-range interactions reveals folding principles of the human genome.* Science, **326**, 5950, (2019).

35. J. R. Dixon, *et al Chromatin architecture reorganization during stem cell differentiation.* Nature, **518**, 331336, (2015).

36. T. Nagano, *et al Comparison of Hi-C results using in-solution versus in-nucleus ligation.* Genome Biology, **16**, (2015).

Acknowledgements We thank Luca Giorgetti for help with the data.

Competing Interests The authors declare that they have no competing financial interests.

Correspondence Correspondence and requests for materials should be addressed to S.P. (email: simone.pigolotti@oist.jp).

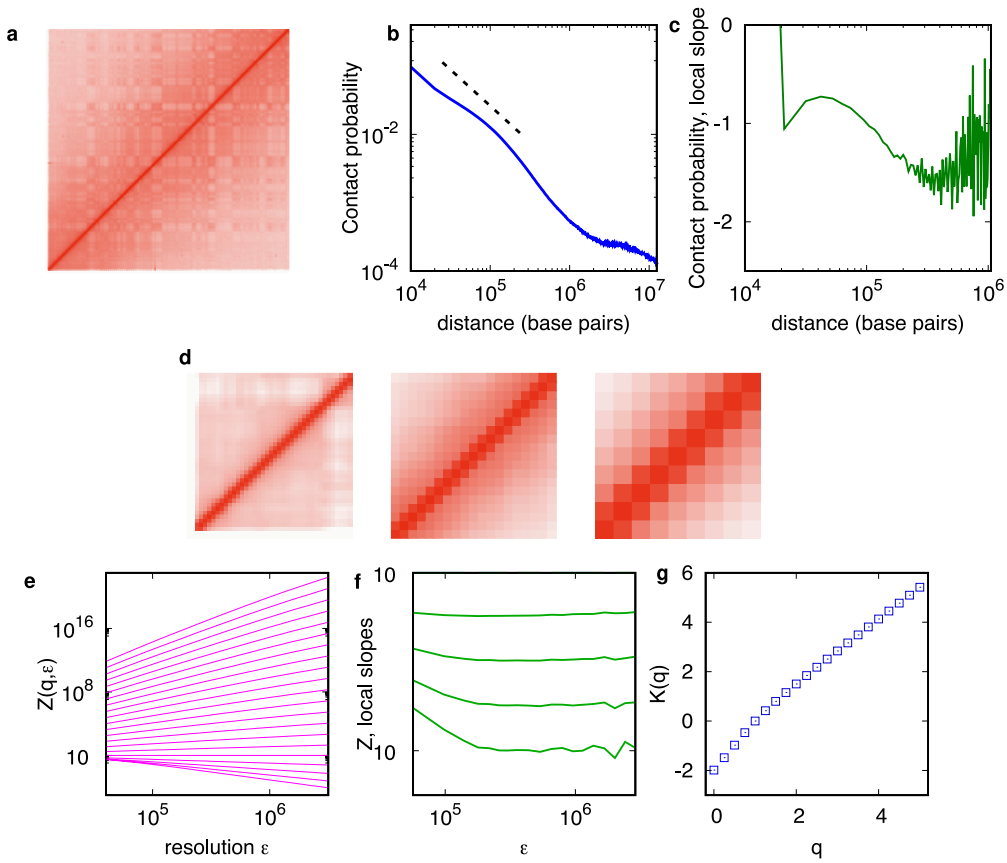

Figure 1: **a**, Hi–C map of chromosome 1 of mouse embryonic stem cells (mESC) [17]. Darker shades of red correspond to higher values of the contact probability p_{ij} **b**, Scaling of the average contact probability $p(\ell)$ as a function of the distance $\ell = |i-j|$. The continuous line is a best-fit of a power law, Eq. (1), in the range of distances $[10^5, 10^6]$, yielding $\beta \approx 1.06$. **c**, Local logarithmic slope $\mathrm{d}\ln P(\ell)/\mathrm{d}(\ln \ell)$ of the contact probability. **d**, Coarse-graining of the Hi–C map at increasing values of the resolution $\epsilon = 10^6, 10^7, 2\ 10^7$ bp. **e**, Scaling of $Z(q)$ as a function of resolution, see Eqs. (4) and (6). Different lines correspond to different values of q, increasing from top to bottom from $q = 0$ to $q = 5$ at intervals of $\Delta q = 0.25$. **f**, Local logarithmic slopes of the first momenta of $Z(q)$. **g**, Corresponding multifractal spectrum $K(q)$. Here and throughout the paper, spectra are obtained by a fit in log-log scale of $Z(q, \epsilon)$ versus ϵ in the range $\epsilon \in [10^5, 10^7]$ unless specified otherwise.

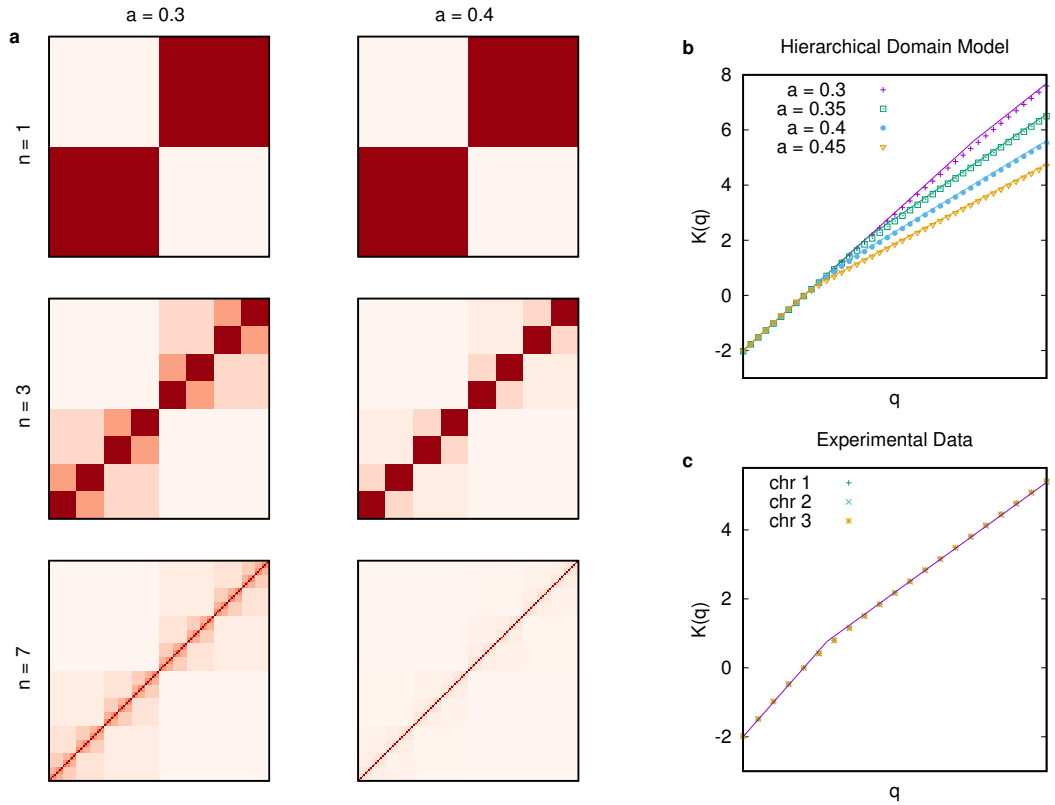

Figure 2: **a**, Construction of the hierarchical domain model for two different values of the parameter a. **b**, Colored symbols are the multifractal spectrum calculated by the hierarchical domain mode at different values of a, the solid curve is the prediction of Eqs. (11) and (12). **c**, The multifractal spectra of the first three chromosomes of mESC (colored symbols) are barely distinguishable from each other. The solid curve is a fit of Eqs. (11) and (12), giving $a \approx 0.413$.

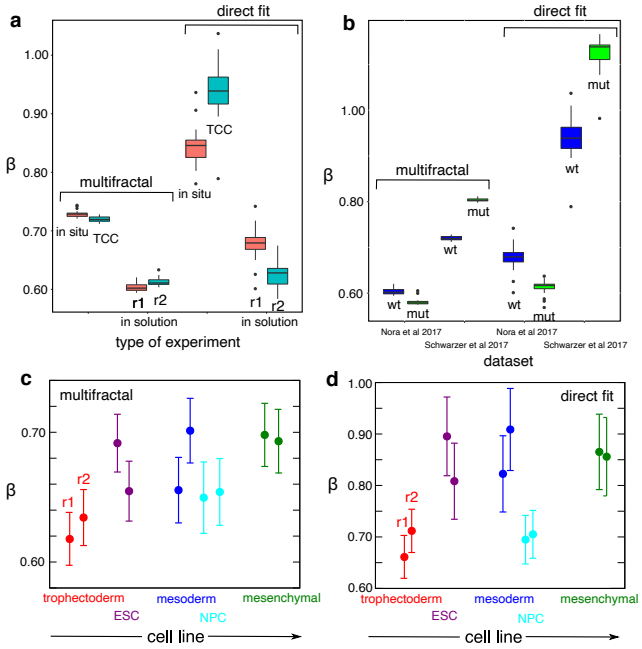

Figure 3: **a** Exponents β calculated from the multifractal spectrum using Eq. (14) and from a direct fit of the data (as in Fig. 1B) with the alternative Hi–C methods. One replicate of in situ [33], TCC [31] and two independent replicates of in solution [17,30] Hi–C are shown in situ [32,33] and in solution [17,30]. The black bar indicates the median over the chromosomes, the colored box indicates the first and third quartiles and black dots are outliers. **b** Value of β obtained for the wild–type (wt) system and for a mutant (mut) in which either CTCF [30] or Nipbl [31] is depleted in experiments. **c** Exponents β obtained with the multifractal method for cell lines at different stages of human stem cell development [35]. **d** Exponents obtained by direct fit for the same data.