Comprehensive Mapping of Long-Range Interactions Reveals Folding Principles of the Human Genome

Erez Lieberman-Aiden, *et al.*

Science **326**, 289 (2009);
DOI: 10.1126/science.1181369

This copy is for your personal, non-commercial use only.

If you wish to distribute this article to others, you can order high-quality copies for your colleagues, clients, or customers by clicking here.

Permission to republish or repurpose articles or portions of articles can be obtained by following the guidelines here.

The following resources related to this article are available online at www.sciencemag.org (this infomation is current as of January 7, 2011):

Updated information and services, including high-resolution figures, can be found in the online version of this article at:
http://www.sciencemag.org/content/326/5950/289.full.html

Supporting Online Material can be found at:
http://www.sciencemag.org/content/suppl/2009/10/08/326.5950.289.DC1.html
http://www.sciencemag.org/content/suppl/2009/10/08/326.5950.289.DC2.html

This article **cites 34 articles**, 14 of which can be accessed free:
http://www.sciencemag.org/content/326/5950/289.full.html#ref-list-1

This article has been **cited by** 19 article(s) on the ISI Web of Science

This article has been **cited by** 32 articles hosted by HighWire Press; see:
http://www.sciencemag.org/content/326/5950/289.full.html#related-urls

This article appears in the following **subject collections:**
Genetics
http://www.sciencemag.org/cgi/collection/genetics

Science (print ISSN 0036-8075; online ISSN 1095-9203) is published weekly, except the last week in December, by the American Association for the Advancement of Science, 1200 New York Avenue NW, Washington, DC 20005. Copyright 2009 by the American Association for the Advancement of Science; all rights reserved. The title *Science* is a registered trademark of AAAS.

Bill & Melinda Gates Foundation); the U.S. Agency for International Development (USAID); and the National Institute of Allergy and Infectious Diseases, NIH, AI33292 (D.R.B.). The contents are the responsibility of the authors and do not necessarily reflect the views of USAID or the U.S. government. The authors declare competing financial interests. Protocol G Principal Investigators: G. Miiro, J. Serwanga, A. Pozniak, D. McPhee, O. Manigart, L. Mwananyanda, E. Karita, A. Inwoley, W. Jaoko, J. DeHovitz, L. G. Bekker, P. Pitisuttithum, R. Paris, and S. Allen.

Supporting Online Material
www.sciencemag.org/cgi/content/full/1178746/DC1
Materials and Methods
SOM Text
Figs. S1 to S10
Tables S1 to S6
References

7 July 2009; accepted 26 August 2009
Published online 3 September 2009;
10.1126/science.1178746
Include this information when citing this paper.

Comprehensive Mapping of Long-Range Interactions Reveals Folding Principles of the Human Genome

Erez Lieberman-Aiden,[1,2,3,4]* Nynke L. van Berkum,[5]* Louise Williams,[1] Maxim Imakaev,[2] Tobias Ragoczy,[6,7] Agnes Telling,[6,7] Ido Amit,[1] Bryan R. Lajoie,[5] Peter J. Sabo,[8] Michael O. Dorschner,[8] Richard Sandstrom,[8] Bradley Bernstein,[1,9] M. A. Bender,[10] Mark Groudine,[6,7] Andreas Gnirke,[1] John Stamatoyannopoulos,[8] Leonid A. Mirny,[2,11] Eric S. Lander,[1,12,13]† Job Dekker[5]†

We describe Hi-C, a method that probes the three-dimensional architecture of whole genomes by coupling proximity-based ligation with massively parallel sequencing. We constructed spatial proximity maps of the human genome with Hi-C at a resolution of 1 megabase. These maps confirm the presence of chromosome territories and the spatial proximity of small, gene-rich chromosomes. We identified an additional level of genome organization that is characterized by the spatial segregation of open and closed chromatin to form two genome-wide compartments. At the megabase scale, the chromatin conformation is consistent with a fractal globule, a knot-free, polymer conformation that enables maximally dense packing while preserving the ability to easily fold and unfold any genomic locus. The fractal globule is distinct from the more commonly used globular equilibrium model. Our results demonstrate the power of Hi-C to map the dynamic conformations of whole genomes.

The three-dimensional (3D) conformation of chromosomes is involved in compartmentalizing the nucleus and bringing widely separated functional elements into close spatial proximity (*1*–*5*). Understanding how chromosomes fold can provide insight into the complex relationships between chromatin structure, gene activity, and the functional state of the cell. Yet beyond the scale of nucleosomes, little is known about chromatin organization.

Long-range interactions between specific pairs of loci can be evaluated with chromosome conformation capture (3C), using spatially constrained ligation followed by locus-specific polymerase chain reaction (PCR) (*6*). Adaptations of 3C have extended the process with the use of inverse PCR (4C) (*7*, *8*) or multiplexed ligation-mediated amplification (5C) (*9*). Still, these techniques require choosing a set of target loci and do not allow unbiased genomewide analysis.

Here, we report a method called Hi-C that adapts the above approach to enable purification of ligation products followed by massively parallel sequencing. Hi-C allows unbiased identification of chromatin interactions across an entire genome.We briefly summarize the process: cells are crosslinked with formaldehyde; DNA is digested with a restriction enzyme that leaves a 5′ overhang; the 5′ overhang is filled, including a biotinylated residue; and the resulting blunt-end fragments are ligated under dilute conditions that favor ligation events between the cross-linked DNA fragments. The resulting DNA sample contains ligation products consisting of fragments that were originally in close spatial proximity in the nucleus, marked with biotin at the junction. A Hi-C library is created by shearing the DNA and selecting the biotin-containing fragments with streptavidin beads. The library is then analyzed by using massively parallel DNA sequencing, producing a catalog of interacting fragments (Fig. 1A) (*10*).

We created a Hi-C library from a karyotypically normal human lymphoblastoid cell line (GM06990) and sequenced it on two lanes of an Illumina Genome Analyzer (Illumina, San Diego, CA), generating 8.4 million read pairs that could be uniquely aligned to the human genome reference sequence; of these, 6.7 million corresponded to long-range contacts between segments >20 kb apart.

We constructed a genome-wide contact matrix M by dividing the genome into 1-Mb regions ("loci") and defining the matrix entry m_{ij} to be the number of ligation products between locus i and locus j (*10*). This matrix reflects an ensemble average of the interactions present in the original sample of cells; it can be visually represented as a heatmap, with intensity indicating contact frequency (Fig. 1B).

We tested whether Hi-C results were reproducible by repeating the experiment with the same restriction enzyme (HindIII) and with a different one (NcoI). We observed that contact matrices for these new libraries (Fig. 1, C and D) were extremely similar to the original contact matrix [Pearson's $r = 0.990$ (HindIII) and $r = 0.814$ (NcoI); P was negligible ($<10^{-300}$) in both cases]. We therefore combined the three data sets in subsequent analyses.

We first tested whether our data are consistent with known features of genome organization (*1*): specifically, chromosome territories (the tendency of distant loci on the same chromosome to be near one another in space) and patterns in subnuclear positioning (the tendency of certain chromosome pairs to be near one another).

We calculated the average intrachromosomal contact probability, $I_n(s)$, for pairs of loci separated by a genomic distance s (distance in base pairs along the nucleotide sequence) on chromosome n. $I_n(s)$ decreases monotonically on every chromosome, suggesting polymer-like behavior in which the 3D distance between loci increases with increasing genomic distance; these findings are in agreement with 3C and fluorescence in situ hybridization (FISH) (*6*, *11*). Even at distances greater than 200 Mb, $I_n(s)$ is always much greater than the average contact probability between different chromosomes (Fig. 2A). This implies the existence of chromosome territories.

Interchromosomal contact probabilities between pairs of chromosomes (Fig. 2B) show that small, gene-rich chromosomes (chromosomes 16, 17, 19, 20, 21, and 22) preferentially interact with each other. This is consistent with FISH studies showing that these chromosomes frequently colocalize in the center of the nucleus

---

[1] Broad Institute of Harvard and Massachusetts Institute of Technology (MIT), MA 02139, USA. [2] Division of Health Sciences and Technology, MIT, Cambridge, MA 02139, USA. [3] Program for Evolutionary Dynamics, Department of Organismic and Evolutionary Biology, Department of Mathematics, Harvard University, Cambridge, MA 02138, USA. [4] Department of Applied Mathematics, Harvard University, Cambridge, MA 02138, USA. [5] Program in Gene Function and Expression and Department of Biochemistry and Molecular Pharmacology, University of Massachusetts Medical School, Worcester, MA 01605, USA. [6] Fred Hutchinson Cancer Research Center, Seattle, WA 98109, USA. [7] Department of Radiation Oncology, University of Washington School of Medicine, Seattle, WA 98195, USA. [8] Department of Genome Sciences, University of Washington, Seattle, WA 98195, USA. [9] Department of Pathology, Harvard Medical School, Boston, MA 02115, USA. [10] Department of Pediatrics, University of Washington, Seattle, WA 98195, USA. [11] Department of Physics, MIT, Cambridge, MA 02139, USA. [12] Department of Biology, MIT, Cambridge, MA 02139, USA. [13] Department of Systems Biology, Harvard Medical School, Boston, MA 02115, USA.

*These authors contributed equally to this work.
†To whom correspondence should be addressed. E-mail: lander@broadinstitute.org (E.S.L.); job.dekker@umassmed.edu (J.D.)

(*12*, *13*). Interestingly, chromosome 18, which is small but gene-poor, does not interact frequently with the other small chromosomes; this agrees with FISH studies showing that chromosome 18 tends to be located near the nuclear periphery (*14*).

We then zoomed in on individual chromosomes to explore whether there are chromosomal regions that preferentially associate with each other. Because sequence proximity strongly influences contact probability, we defined a normalized contact matrix M^* by dividing each entry in the contact matrix by the genome-wide average contact probability for loci at that genomic distance (*10*). The normalized matrix shows many large blocks of enriched and depleted interactions, generating a plaid pattern (Fig. 3B). If two loci (here 1-Mb regions) are nearby in space, we reasoned that they will share neighbors and have correlated interaction profiles. We therefore defined a correlation matrix C in which c_{ij} is the Pearson correlation between the *i*th row and *j*th column of M^*. This process dramatically sharpened the plaid pattern (Fig. 3C); 71% of the resulting matrix entries represent statistically significant correlations ($P \leq 0.05$).

The plaid pattern suggests that each chromosome can be decomposed into two sets of loci (arbitrarily labeled A and B) such that contacts within each set are enriched and contacts between sets are depleted. We partitioned each chromosome

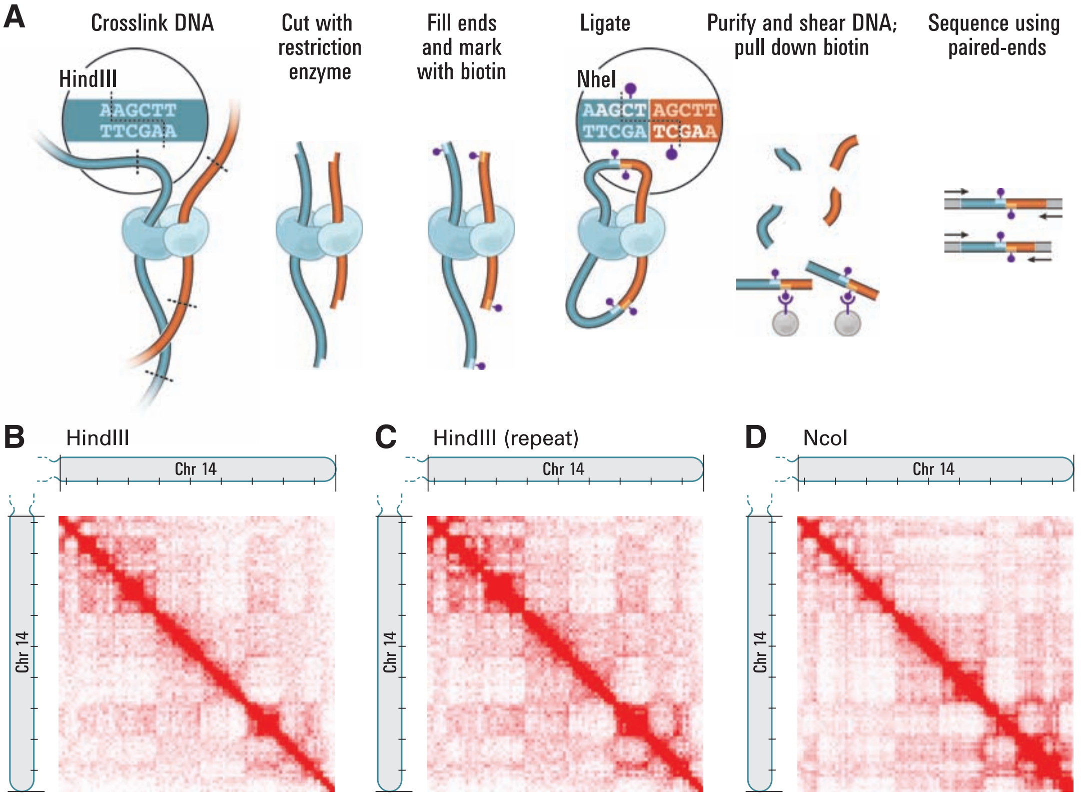

Fig. 1. Overview of Hi-C. (**A**) Cells are cross-linked with formaldehyde, resulting in covalent links between spatially adjacent chromatin segments (DNA fragments shown in dark blue, red; proteins, which can mediate such interactions, are shown in light blue and cyan). Chromatin is digested with a restriction enzyme (here, HindIII; restriction site marked by dashed line; see inset), and the resulting sticky ends are filled in with nucleotides, one of which is biotinylated (purple dot). Ligation is performed under extremely dilute conditions to create chimeric molecules; the HindIII site is lost and an NheI site is created (inset). DNA is purified and sheared. Biotinylated junctions are isolated with streptavidin beads and identified by paired-end sequencing. (**B**) Hi-C produces a genome-wide contact matrix. The submatrix shown here corresponds to intrachromosomal interactions on chromosome 14. (Chromosome 14 is acrocentric; the short arm is not shown.) Each pixel represents all interactions between a 1-Mb locus and another 1-Mb locus; intensity corresponds to the total number of reads (0 to 50). Tick marks appear every 10 Mb. (**C** and **D**) We compared the original experiment with results from a biological repeat using the same restriction enzyme [(C), range from 0 to 50 reads] and with results using a different restriction enzyme [(D), NcoI, range from 0 to 100 reads].

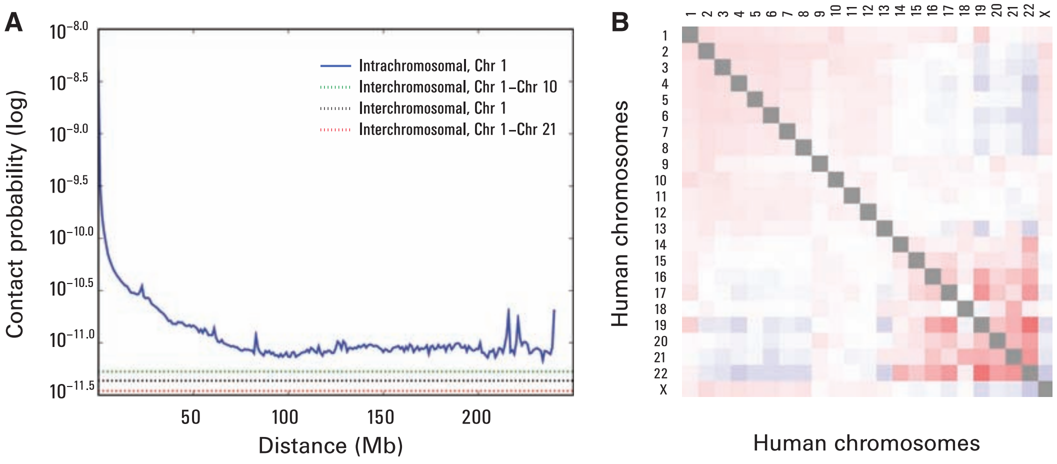

Fig. 2. The presence and organization of chromosome territories. (**A**) Probability of contact decreases as a function of genomic distance on chromosome 1, eventually reaching a plateau at ~90 Mb (blue). The level of interchromosomal contact (black dashes) differs for different pairs of chromosomes; loci on chromosome 1 are most likely to interact with loci on chromosome 10 (green dashes) and least likely to interact with loci on chromosome 21 (red dashes). Interchromosomal interactions are depleted relative to intrachromosomal interactions. (**B**) Observed/expected number of interchromosomal contacts between all pairs of chromosomes. Red indicates enrichment, and blue indicates depletion (range from 0.5 to 2). Small, gene-rich chromosomes tend to interact more with one another, suggesting that they cluster together in the nucleus.

in this way by using principal component analysis. For all but two chromosomes, the first principal component (PC) clearly corresponded to the plaid pattern (positive values defining one set, negative values the other) (fig. S1). For chromosomes 4 and 5, the first PC corresponded to the two chromosome arms, but the second PC corresponded to the plaid pattern. The entries of the PC vector reflected the sharp transitions from compartment to compartment observed within the plaid heatmaps. Moreover, the plaid patterns within each chromosome were consistent across chromosomes: the labels (A and B) could be assigned on each chromosome so that sets on different chromosomes carrying the same label had correlated contact profiles, and those carrying different labels had anticorrelated contact profiles (Fig. 3D). These results imply that the entire genome can be partitioned into two spatial compartments such that greater interaction occurs within each compartment rather than across compartments.

The Hi-C data imply that regions tend be closer in space if they belong to the same compartment (A versus B) than if they do not. We tested this by using 3D-FISH to probe four loci (L1, L2, L3, and L4) on chromosome 14 that alternate between the two compartments (L1 and L3 in compartment A; L2 and L4 in compartment B) (Fig. 3, E and F). 3D-FISH showed that L3 tends to be closer to L1 than to L2, despite the fact that L2 lies between L1 and L3 in the linear genome sequence (Fig. 3E). Similarly, we found that L2 is closer to L4 than to L3 (Fig. 3F). Comparable results were obtained for four consecutive loci on chromosome 22 (fig. S2, A and B). Taken together, these observations confirm the spatial compartmentalization

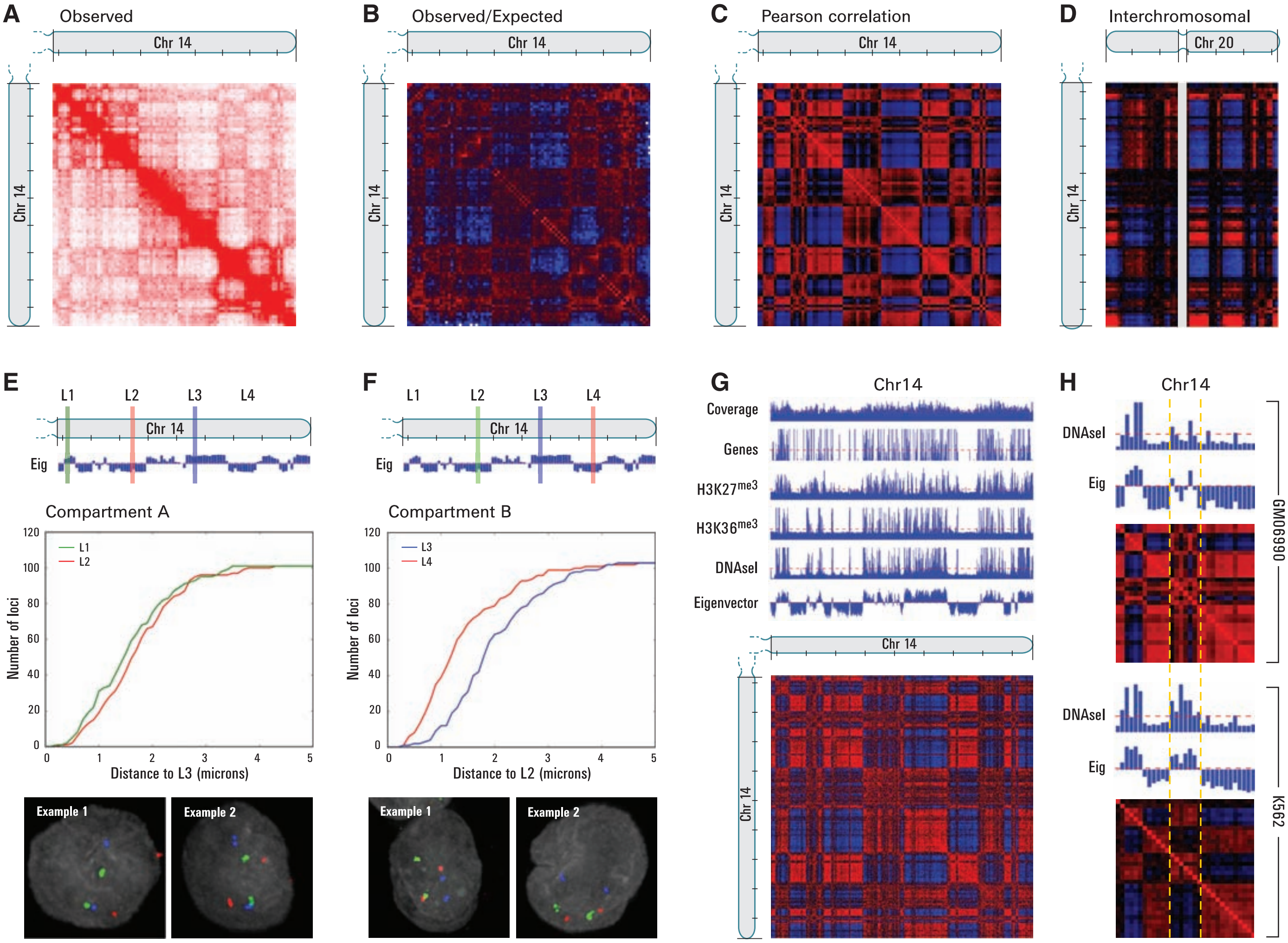

Fig. 3. The nucleus is segregated into two compartments corresponding to open and closed chromatin. (**A**) Map of chromosome 14 at a resolution of 1 Mb exhibits substructure in the form of an intense diagonal and a constellation of large blocks (three experiments combined; range from 0 to 200 reads). Tick marks appear every 10 Mb. (**B**) The observed/expected matrix shows loci with either more (red) or less (blue) interactions than would be expected, given their genomic distance (range from 0.2 to 5). (**C**) Correlation matrix illustrates the correlation [range from – (blue) to +1 (red)] between the intrachromosomal interaction profiles of every pair of 1-Mb loci along chromosome 14. The plaid pattern indicates the presence of two compartments within the chromosome. (**D**) Interchromosomal correlation map for chromosome 14 and chromosome 20 [range from –0.25 (blue) to 0.25 (red)]. The unalignable region around the centromere of chromosome 20 is indicated in gray. Each compartment on chromosome 14 has a counterpart on chromosome 20 with a very similar genome-wide interaction pattern. (**E** and **F**) We designed probes for four loci (L1, L2, L3, and L4) that lie consecutively along chromosome 14 but alternate between the two compartments [L1 and L3 in (compartment A); L2 and L4 in (compartment B)]. (E) L3 (blue) was consistently closer to L1 (green) than to L2 (red), despite the fact that L2 lies between L1 and L3 in the primary sequence of the genome. This was confirmed visually and by plotting the cumulative distribution. (F) L2 (green) was consistently closer to L4 (red) than to L3 (blue). (**G**) Correlation map of chromosome 14 at a resolution of 100 kb. The PC (eigenvector) correlates with the distribution of genes and with features of open chromatin. (**H**) A 31-Mb window from chromosome 14 is shown; the indicated region (yellow dashes) alternates between the open and the closed compartments in GM06990 (top, eigenvector and heatmap) but is predominantly open in K562 (bottom, eigenvector and heatmap). The change in compartmentalization corresponds to a shift in chromatin state (DNAseI).

of the genome inferred from Hi-C. More generally, a strong correlation was observed between the number of Hi-C reads m_{ij} and the 3D distance between locus i and locus j as measured by FISH [Spearman's $\rho = -0.916$, $P = 0.00003$ (fig. S3)], suggesting that Hi-C read count may serve as a proxy for distance.

Upon close examination of the Hi-C data, we noted that pairs of loci in compartment B showed a consistently higher interaction frequency at a given genomic distance than pairs of loci in compartment A (fig. S4). This suggests that compartment B is more densely packed (*15*). The FISH data are consistent with this observation; loci in compartment B exhibited a stronger tendency for close spatial localization.

To explore whether the two spatial compartments correspond to known features of the genome, we compared the compartments identified in our 1-Mb correlation maps with known genetic and epigenetic features. Compartment A correlates strongly with the presence of genes (Spearman's $\rho = 0.431$, $P < 10^{-137}$), higher expression [via genome-wide mRNA expression, Spearman's $\rho = 0.476$, $P < 10^{-145}$ (fig. S5)], and accessible chromatin [as measured by deoxyribonuclease I (DNAseI) sensitivity, Spearman's $\rho = 0.651$, P negligible] (*16*, *17*). Compartment A also shows enrichment for both activating (H3K36 trimethylation, Spearman's $\rho = 0.601$, $P < 10^{-296}$) and repressive (H3K27 trimethylation, Spearman's $\rho = 0.282$, $P < 10^{-56}$) chromatin marks (*18*). We repeated the above analysis at a resolution of 100 kb (Fig. 3G) and saw that, although the correlation of compartment A with all other genomic and epigenetic features remained strong (Spearman's $\rho > 0.4$, P negligible), the correlation with the sole repressive mark, H3K27 trimethylation, was dramatically attenuated (Spearman's $\rho = 0.046$, $P < 10^{-15}$). On the basis of these results we concluded that compartment A is more closely associated with open, accessible, actively transcribed chromatin.

We repeated our experiment with K562 cells, an erythroleukemia cell line with an aberrant karyotype (*19*). We again observed two compartments; these were similar in composition to those observed in GM06990 cells [Pearson's $r = 0.732$,

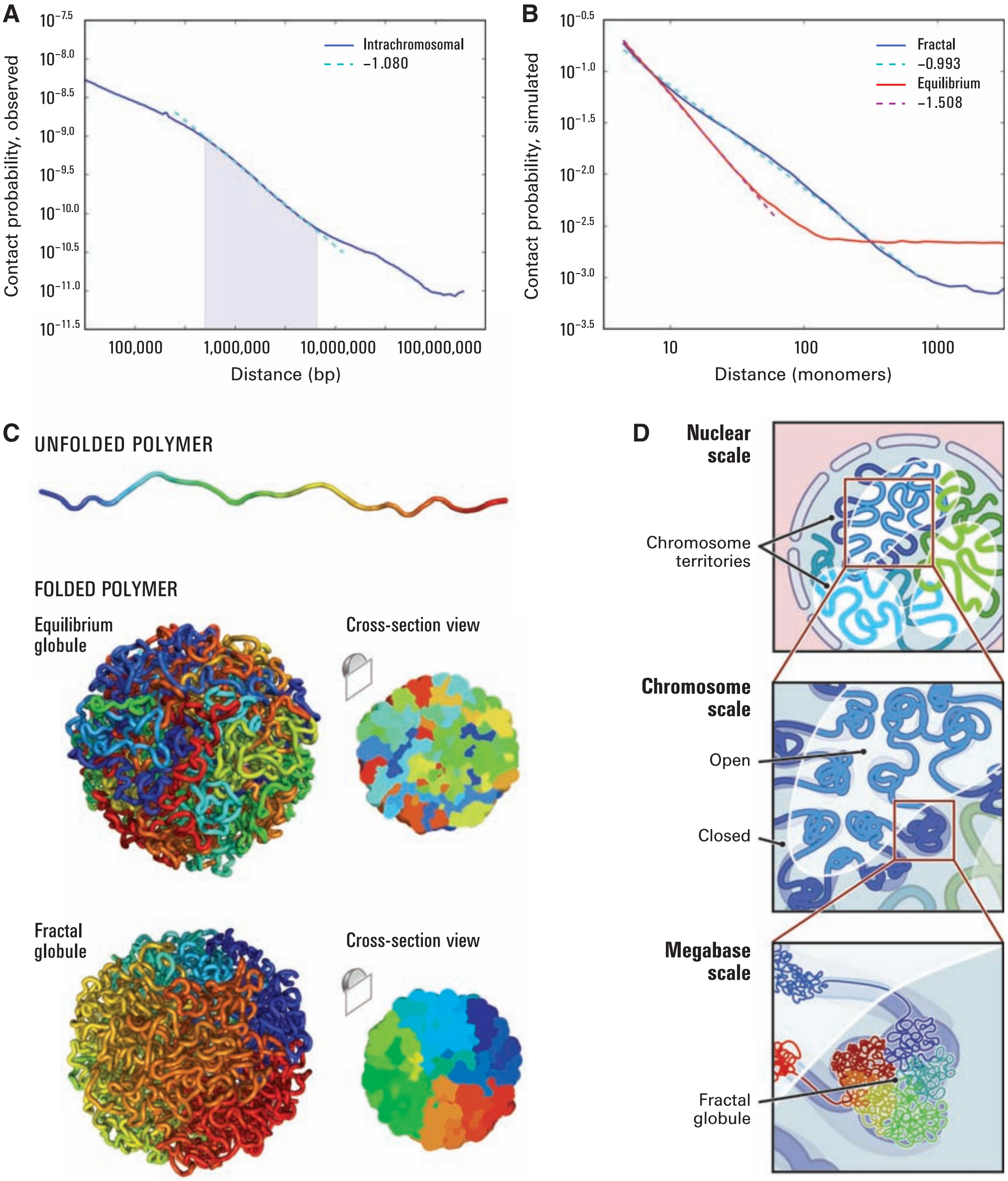

Fig. 4. The local packing of chromatin is consistent with the behavior of a fractal globule. (**A**) Contact probability as a function of genomic distance averaged across the genome (blue) shows a power law scaling between 500 kb and 7 Mb (shaded region) with a slope of –1.08 (fit shown in cyan). (**B**) Simulation results for contact probability as a function of distance (1 monomer ~ 6 nucleosomes ~ 1200 base pairs) (*10*) for equilibrium (red) and fractal (blue) globules. The slope for a fractal globule is very nearly –1 (cyan), confirming our prediction (*10*). The slope for an equilibrium globule is –3/2, matching prior theoretical expectations. The slope for the fractal globule closely resembles the slope we observed in the genome. (**C**) (Top) An unfolded polymer chain, 4000 monomers (4.8 Mb) long. Coloration corresponds to distance from one endpoint, ranging from blue to cyan, green, yellow, orange, and red. (Middle) An equilibrium globule. The structure is highly entangled; loci that are nearby along the contour (similar color) need not be nearby in 3D. (Bottom) A fractal globule. Nearby loci along the contour tend to be nearby in 3D, leading to monochromatic blocks both on the surface and in cross section. The structure lacks knots. (**D**) Genome architecture at three scales. (Top) Two compartments, corresponding to open and closed chromatin, spatially partition the genome. Chromosomes (blue, cyan, green) occupy distinct territories. (Middle) Individual chromosomes weave back and forth between the open and closed chromatin compartments. (Bottom) At the scale of single megabases, the chromosome consists of a series of fractal globules.

P negligible (fig. S6)] and showed strong correlation with open and closed chromatin states as indicated by DNAseI sensitivity (Spearman's $\rho = 0.455$, $P < 10^{-154}$).

The compartment patterns in K562 and GM06990 are similar, but there are many loci in the open compartment in one cell type and the closed compartment in the other (Fig. 3H). Examining these discordant loci on karyotypically normal chromosomes in K562 (*19*), we observed a strong correlation between the compartment pattern in a cell type and chromatin accessibility in that same cell type (GM06990, Spearman's $\rho = 0.384$, $P = 0.012$; K562, Spearman's $\rho = 0.366$, $P = 0.017$). Thus, even in a highly rearranged genome, spatial compartmentalization correlates strongly with chromatin state.

Our results demonstrate that open and closed chromatin domains throughout the genome occupy different spatial compartments in the nucleus. These findings expand on studies of individual loci that have observed particular instances of such interactions, both between distantly located active genes and between distantly located inactive genes (*8*, *20*–*24*).

Lastly, we sought to explore chromatin structure within compartments. We closely examined the average behavior of intrachromosomal contact probability as a function of genomic distance, calculating the genome-wide distribution $I(s)$. When plotted on log-log axes, $I(s)$ exhibits a prominent power law scaling between ~500 kb and ~7 Mb, where contact probability scales as s^{-1} (Fig. 4A). This range corresponds to the known size of open and closed chromatin domains.

Power-law dependencies can arise from polymer-like behavior (*25*). Various authors have proposed that chromosomal regions can be modeled as an "equilibrium globule": a compact, densely knotted configuration originally used to describe a polymer in a poor solvent at equilibrium (*26*, *27*). [Historically, this specific model has often been referred to simply as a "globule"; some authors have used the term "equilibrium globule" to distinguish it from other globular states (see below).] Grosberg *et al*. proposed an alternative model, theorizing that polymers, including interphase DNA, can self-organize into a long-lived, nonequilibrium conformation that they described as a "fractal globule" (*28*, *29*). This highly compact state is formed by an unentangled polymer when it crumples into a series of small globules in a "beads-on-a-string" configuration. These beads serve as monomers in subsequent rounds of spontaneous crumpling until only a single globule-of-globules-of-globules remains. The resulting structure resembles a Peano curve, a continuous fractal trajectory that densely fills 3D space without crossing itself (*30*). Fractal globules are an attractive structure for chromatin segments because they lack knots (*31*) and would facilitate unfolding and refolding, for example, during gene activation, gene repression, or the cell cycle. In a fractal globule, contiguous regions of the genome tend to form spatial sectors whose size corresponds to the length of the original region (Fig. 4C). In contrast, an equilibrium globule is highly knotted and lacks such sectors; instead, linear and spatial positions are largely decorrelated after, at most, a few megabases (Fig. 4C). The fractal globule has not previously been observed (*29*, *31*).

The equilibrium globule and fractal globule models make very different predictions concerning the scaling of contact probability with genomic distance s. The equilibrium globule model predicts that contact probability will scale as $s^{-3/2}$, which we do not observe in our data. We analytically derived the contact probability for a fractal globule and found that it decays as s^{-1} (*10*); this corresponds closely with the prominent scaling we observed ($s^{-1.08}$).

The equilibrium and fractal globule models also make differing predictions about the 3D distance between pairs of loci ($s^{1/2}$ for an equilibrium globule, $s^{1/3}$ for a fractal globule). Although 3D distance is not directly measured by Hi-C, we note that a recent paper using 3D-FISH reported an $s^{1/3}$ scaling for genomic distances between 500 kb and 2 Mb (*27*).

We used Monte Carlo simulations to construct ensembles of fractal globules and equilibrium globules (500 each). The properties of the ensembles matched the theoretically derived scalings for contact probability (for fractal globules, s^{-1}, and for equilibrium globules, $s^{-3/2}$) and 3D distance (for fractal globules $s^{1/3}$, for equilibrium globules $s^{1/2}$). These simulations also illustrated the lack of entanglements [measured by using the knot-theoretic Alexander polynomial (*10*, *32*)] and the formation of spatial sectors within a fractal globule (Fig. 4B).

We conclude that, at the scale of several megabases, the data are consistent with a fractal globule model for chromatin organization. Of course, we cannot rule out the possibility that other forms of regular organization might lead to similar findings.

We focused here on interactions at relatively large scales. Hi-C can also be used to construct comprehensive, genome-wide interaction maps at finer scales by increasing the number of reads. This should enable the mapping of specific long-range interactions between enhancers, silencers, and insulators (*33*–*35*). To increase the resolution by a factor of n, one must increase the number of reads by a factor of n^2. As the cost of sequencing falls, detecting finer interactions should become increasingly feasible. In addition, one can focus on subsets of the genome by using chromatin immunoprecipitation or hybrid capture (*36*, *37*).

References and Notes

1. T. Cremer, C. Cremer, *Nat. Rev. Genet.* **2**, 292 (2001).
2. T. Sexton, H. Schober, P. Fraser, S. M. Gasser, *Nat. Struct. Mol. Biol.* **14**, 1049 (2007).
3. J. Dekker, *Science* **319**, 1793 (2008).
4. T. Misteli, *Cell* **128**, 787 (2007).
5. S. T. Kosak, M. Groudine, *Genes Dev.* **18**, 1371 (2004).
6. J. Dekker, K. Rippe, M. Dekker, N. Kleckner, *Science* **295**, 1306 (2002).
7. Z. Zhao *et al.*, *Nat. Genet.* **38**, 1341 (2006).
8. M. Simonis *et al.*, *Nat. Genet.* **38**, 1348 (2006).
9. J. Dostie *et al.*, *Genome Res.* **16**, 1299 (2006).
10. Materials and methods are available as supporting material on *Science* Online.
11. H. Yokota, G. van den Engh, J. E. Hearst, R. K. Sachs, B. J. Trask, *J. Cell Biol.* **130**, 1239 (1995).
12. S. Boyle *et al.*, *Hum. Mol. Genet.* **10**, 211 (2001).
13. H. Tanabe, F. A. Habermann, I. Solovei, M. Cremer, T. Cremer, *Mutat. Res.* **504**, 37 (2002).
14. J. A. Croft *et al.*, *J. Cell Biol.* **145**, 1119 (1999).
15. J. Dekker, *J. Biol. Chem.* **283**, 34532 (2008).
16. P. J. Sabo *et al.*, *Nat. Methods* **3**, 511 (2006).
17. J. R. Hesselberth *et al.*, *Nat. Methods* **6**, 283 (2009).
18. T. S. Mikkelsen *et al.*, *Nature* **448**, 553 (2007).
19. S. Naumann, D. Reutzel, M. Speicher, H. J. Decker, *Leuk. Res.* **25**, 313 (2001).
20. C. S. Osborne *et al.*, *Nat. Genet.* **36**, 1065 (2004).
21. J. M. Brown *et al.*, *J. Cell Biol.* **182**, 1083 (2008).
22. A. F. Dernburg *et al.*, *Cell* **85**, 745 (1996).
23. L. S. Shopland *et al.*, *J. Cell Biol.* **174**, 27 (2006).
24. P. Fraser, W. Bickmore, *Nature* **447**, 413 (2007).
25. P. G. d. Gennes, *Scaling Concepts in Polymer Physics* (Cornell Univ. Press, Ithaca, NY, 1979).
26. C. Munkel, J. Langowski, *Phys. Rev. E* **57**, 5888 (1998).
27. J. Mateos-Langerak *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* **106**, 3812 (2009).
28. A. Y. Grosberg, S. K. Nechaev, E. I. Shakhnovich, *J. Phys. France* **49**, 2095 (1988).
29. A. Grosberg, Y. Rabin, S. Havlin, A. Neer, *Europhys. Lett.* **23**, 373 (1993).
30. B. B. Mandelbrot, *The Fractal Geometry of Nature* (Freeman, New York, ed. 2, 1983).
31. O. A. Vasilyev, S. K. Nechaev, *Theor. Math. Phys.* **134**, 142 (2003).
32. G. Kolesov, P. Virnau, M. Kardar, L. A. Mirny, *Nucleic Acids Res.* **35**, W425 (2007).
33. E. M. Blackwood, J. T. Kadonaga, *Science* **281**, 60 (1998).
34. A. C. Bell, A. G. West, G. Felsenfeld, *Science* **291**, 447 (2001).
35. J. E. Phillips, V. G. Corces, *Cell* **137**, 1194 (2009).
36. M. H. Kuo, C. D. Allis, *Methods* **19**, 425 (1999).
37. A. Gnirke *et al.*, *Nat. Biotechnol.* **27**, 182 (2009).

38. Supported by a Fannie and John Hertz Foundation graduate fellowship, a National Defense Science and Engineering graduate fellowship, an NSF graduate fellowship, the National Space Biomedical Research Institute, and grant no. T32 HG002295 from the National Human Genome Research Institute (NHGRI) (E.L.); a fellowship from the American Society of Hematology (T.R.); award no. R01HL06544 from the National Heart, Lung, and Blood Institute and R37DK44746 from the National Institute of Diabetes and Digestive and Kidney Diseases (M.G.); NIH grant U54HG004592 (J.S.); i2b2 (Informatics for Integrating Biology and the Bedside), the NIH-supported Center for Biomedical Computing at Brigham and Women's Hospital (L.A.M.), grant no. HG003143 from the NHGRI, and a Keck Foundation distinguished young scholar award (J.D.). We thank J. Goldy, K. Lee, S. Vong, and M. Weaver for assistance with DNaseI experiments; A. Kosmrlj for discussions and code; A. P. Aiden, X. R. Bao, M. Brenner, D. Galas, W. Gosper, A. Jaffer, A. Melnikov, A. Miele, G. Giannoukos, C. Nusbaum, A. J. M. Walhout, L. Wood, and K. Zeldovich for discussions; and L. Gaffney and B. Wong for help with visualization. We also acknowledge the ENCODE chromatin group at Broad Institute and Massachusetts General Hospital. Hi-C sequence data has been deposited at the GEO database (www.ncbi.nlm.nih.gov/geo/), accession no. GSE18199. Expression data are also available at GEO, accession no. GSE18350. Chromatin immunoprecipitation sequence (ChIP-Seq) data and DNaseI sensitivity data are available at the University of California Santa Cruz (UCSC) browser (http://genome.ucsc.edu/). Additional visualizations are available at http://hic.umassmed.edu. A provisional patent on the Hi-C method (no. 61/100,151) is under review.

Supporting Online Material

www.sciencemag.org/cgi/content/full/326/5950/289/DC1
Materials and Methods
Figs. S1 to S32

1 September 2009; accepted 18 September 2009
10.1126/science.1181369