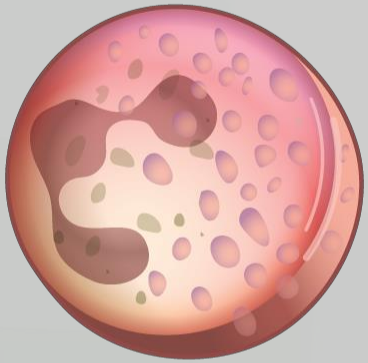


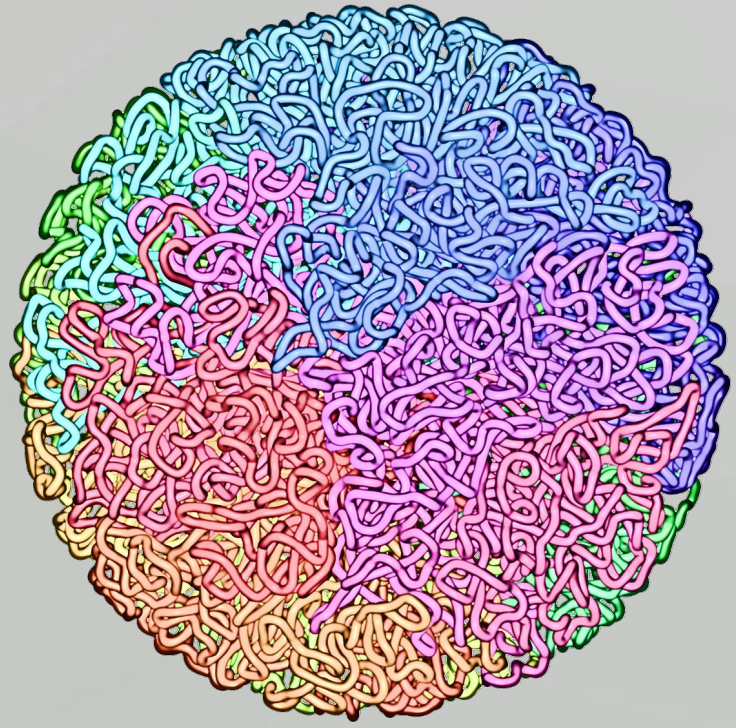
A Human Cell

Human DNA

Every cell carries 6 billion DNA base pairs

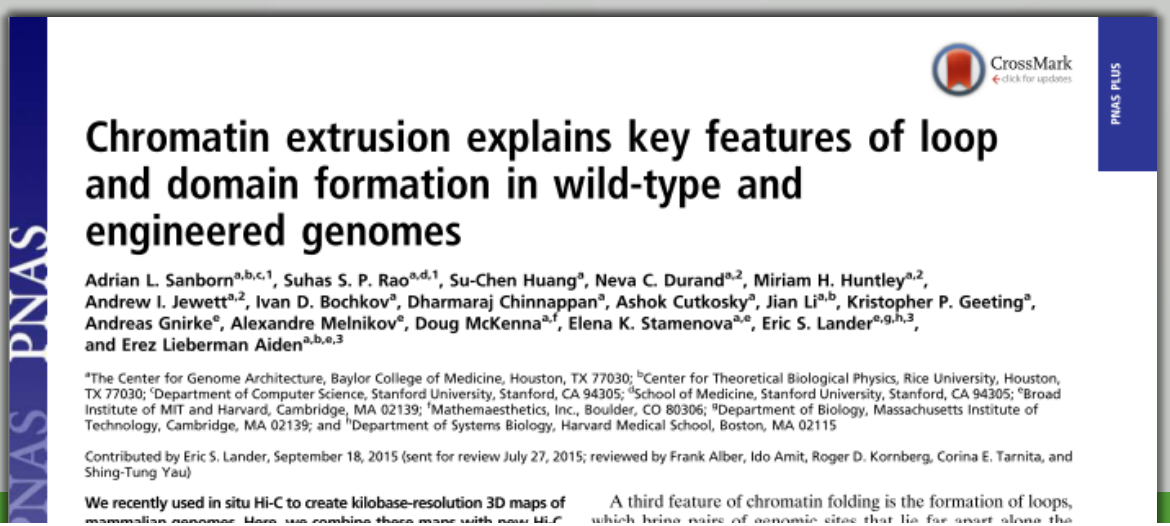
When stretched out, this would extend about **two meters**

A nucleus of just five **5 micrometers**



Chromatin extrusion explains key features of loop and domain formation in wild-type and engineered genomes

Proc. Natl. Acad. Sci 112, E6456-65 (2015)



PNAS PLUS

CrossMark click for updates

Chromatin extrusion explains key features of loop and domain formation in wild-type and engineered genomes

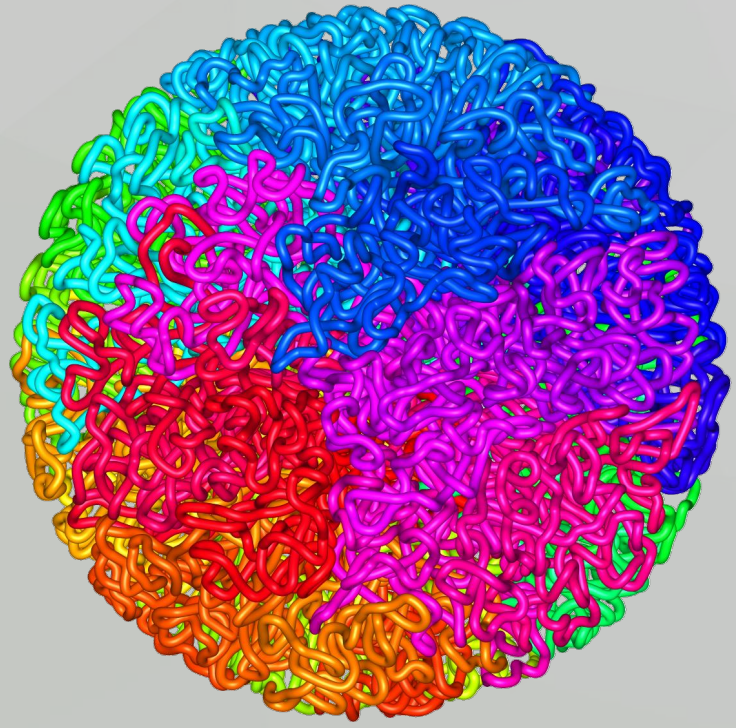
Adrian L. Sanborn[a,b,c,1], Suhas S. P. Rao[a,d,1], Su-Chen Huang[a], Neva C. Durand[a,2], Miriam H. Huntley[a,2], Andrew I. Jewett[a,2], Ivan D. Bochkov[a], Dharmaraj Chinnappan[a], Ashok Cutkosky[a], Jian Li[a,b], Kristopher P. Geeting[a], Andreas Gnirke[e], Alexandre Melnikov[e], Doug McKenna[a,f], Elena K. Stamenova[a,e], Eric S. Lander[e,g,h,3], and Erez Lieberman Aiden[a,b,e,3]

[a]The Center for Genome Architecture, Baylor College of Medicine, Houston, TX 77030; [b]Center for Theoretical Biological Physics, Rice University, Houston, TX 77030; [c]Department of Computer Science, Stanford University, Stanford, CA 94305; [d]School of Medicine, Stanford University, Stanford, CA 94305; [e]Broad Institute of MIT and Harvard, Cambridge, MA 02139; [f]Mathemaesthetics, Inc., Boulder, CO 80306; [g]Department of Biology, Massachusetts Institute of Technology, Cambridge, MA 02139; and [h]Department of Systems Biology, Harvard Medical School, Boston, MA 02115

Contributed by Eric S. Lander, September 18, 2015 (sent for review July 27, 2015; reviewed by Frank Alber, Ido Amit, Roger D. Kornberg, Corina E. Tarnita, and Shing-Tung Yau)

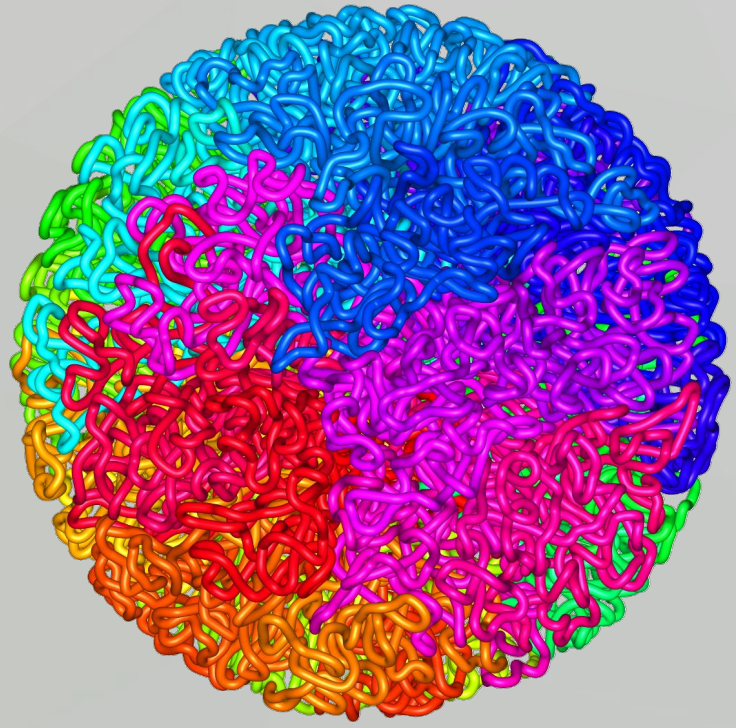
We recently used in situ Hi-C to create kilobase-resolution 3D maps of

A third feature of chromatin folding is the formation of loops,



Genome Architecture

- DNA packaging and 3D structure and how it affects Gene expression
- Mutations
- CTCF: Topologically associating domain (TAD)



Genome Architecture

It had previously been proposed that mammalian chromosomes form **fractal globules.**

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.

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.

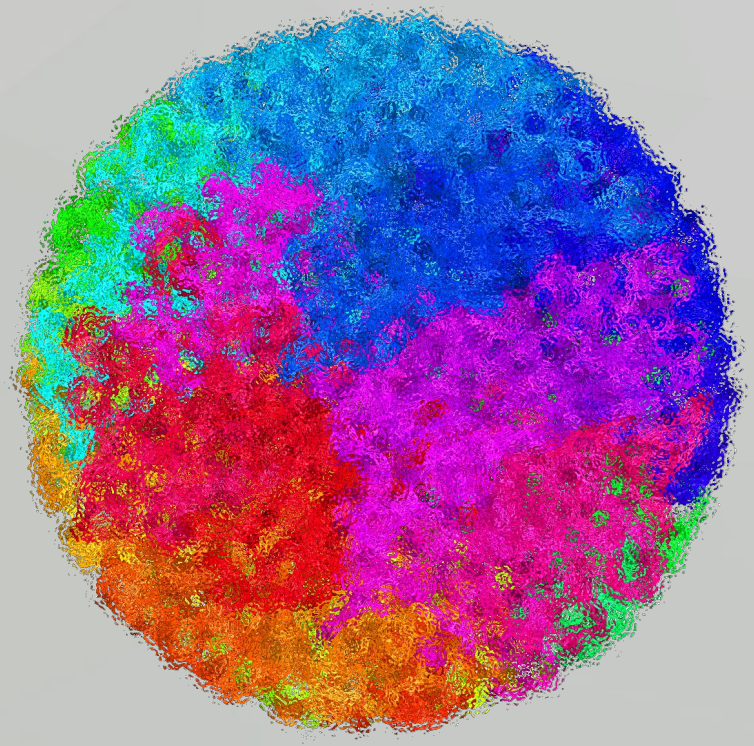
We constructed a genome-wide contact matrix M by dividing the genome into 1-Mb regions ("loci") and defining the matrix entry m_{ij} to be the number of ligation products between locus i and locus j (*10*). This matrix reflects an ensemble average of the interactions present in the original sample of 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

Fractals

- Fractals are shapes with non-integer dimension
- Fractal globules are polymers that have compact local and global scaling
- They can be modeled through a lattice walk in which every point is only visited once and no paths intersection, this prevents knot formation.
- Minkowski (box-counting) dimension
- $N(1/n) \approx Cn^d$

Is chromosomal DNA a **fractal globule**?



Low Res Hi-C

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.

[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,

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 di-

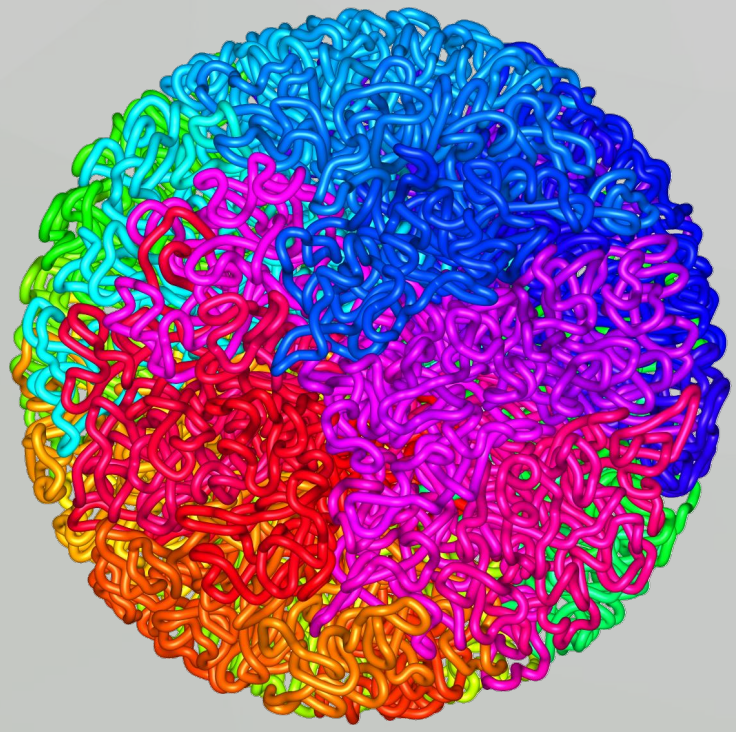
We created a Hi-C library from a karyotypically normal human lymphoblastoid cell line (GM06990) and sequenced it on two lanes of an Illumina Genome Analyzer (Illumina, San Diego, CA), generating 8.4 million read pairs that could be uniquely aligned to the human genome reference sequence; of these, 6.7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



High Res Hi-C

Article

Cell

A 3D Map of the Human Genome at Kilobase Resolution Reveals Principles of Chromatin Looping

Suhas S.P. Rao,[1,2,3,4,10] Miriam H. Huntley,[1,2,3,4,5,10] Neva C. Durand,[1,2,3,4] Elena K. Stamenova,[1,2,3,4] Ivan D. Bochkov,[1,2,3] James T. Robinson,[1,4] Adrian L. Sanborn,[1,2,3,6] Ido Machol,[1,2,3] Arina D. Omer,[1,2,3] Eric S. Lander,[4,7,8,*] and Erez Lieberman Aiden[1,2,3,4,9,*]

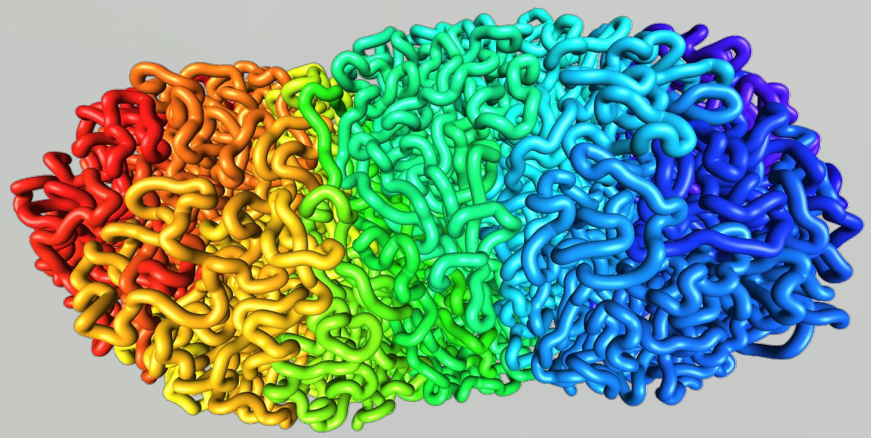
[1]The Center for Genome Architecture, Baylor College of Medicine, Houston, TX 77030, USA
[2]Department of Molecular and Human Genetics, Baylor College of Medicine, Houston, TX 77030, USA
[3]Department of Computer Science, Department of Computational and Applied Mathematics, Rice University, Houston, TX 77005, USA
[4]Broad Institute of MIT and Harvard, Cambridge, MA 02139, USA
[5]School of Engineering and Applied Sciences, Harvard University, Cambridge, MA 02138, USA
[6]Department of Computer Science, Stanford University, Stanford, CA 94305, USA
[7]Department of Biology, Massachusetts Institute of Technology (MIT), Cambridge, MA 02139, USA
[8]Department of Systems Biology, Harvard Medical School, Boston, MA 02115, USA
[9]Center for Theoretical Biological Physics, Rice University, Houston, TX 77030, USA
[10]Co-first author
*Correspondence: lander@broadinstitute.org (E.S.L.), erez@erez.com (E.L.A.)
http://dx.doi.org/10.1016/j.cell.2014.11.021

SUMMARY

We use in situ Hi-C to probe the 3D architecture of genomes, constructing haploid and diploid maps of nine cell types. The densest, in human lymphoblastoid cells, contains 4.9 billion contacts, achieving 1 kb resolution. We find that genomes are partitioned into contact domains (median length, 185 kb), which are associated with distinct patterns of histone marks and segregate into six subcompartments. We identify ~10,000 loops. These loops frequently link promoters and enhancers, correlate with gene activation, and show conservation across cell types and species. Loop anchors typically occur at domain boundaries and bind CTCF. CTCF sites at loop anchors occur predominantly (>90%) in a convergent orientation, with the asymmetric motifs "facing"

Various methods have emerged to assess the 3D architecture of the nucleus. In one seminal study, the binding of a protein to sites at opposite ends of a restriction fragment created a loop, which was detectable because it promoted the formation of DNA circles in the presence of ligase. Removal of the protein or either of its binding sites disrupted the loop, eliminating this "cyclization enhancement" (Mukherjee et al., 1988). Subsequent adaptations of cyclization enhancement made it possible to analyze chromatin folding in vivo, including nuclear ligation assay (Cullen et al., 1993) and chromosome conformation capture (Dekker et al., 2002), which analyze contacts made by a single locus, extensions such as 5C for examining several loci simultaneously (Dostie et al., 2006), and methods such as ChIA-PET for examining all loci bound by a specific protein (Fullwood et al., 2009).

To interrogate all loci at once, we developed Hi-C, which combines DNA proximity ligation with high-throughput sequencing in a genome-wide fashion (Lieberman-Aiden et al., 2009). We used



High Res Hi-C

Article

Cell

A 3D Map of the Human Genome at Kilobase Resolution Reveals Principles of Chromatin Looping

Suhas S.P. Rao,[1,2,3,4,10] Miriam H. Huntley,[1,2,3,4,5,10] Neva C. Durand,[1,2,3,4] Elena K. Stamenova,[1,2,3,4] Ivan D. Bochkov,[1,2,3] James T. Robinson,[1,4] Adrian L. Sanborn,[1,2,3,6] Ido Machol,[1,2,3] Arina D. Omer,[1,2,3] Eric S. Lander,[4,7,8,*] and Erez Lieberman Aiden[1,2,3,4,9,*]

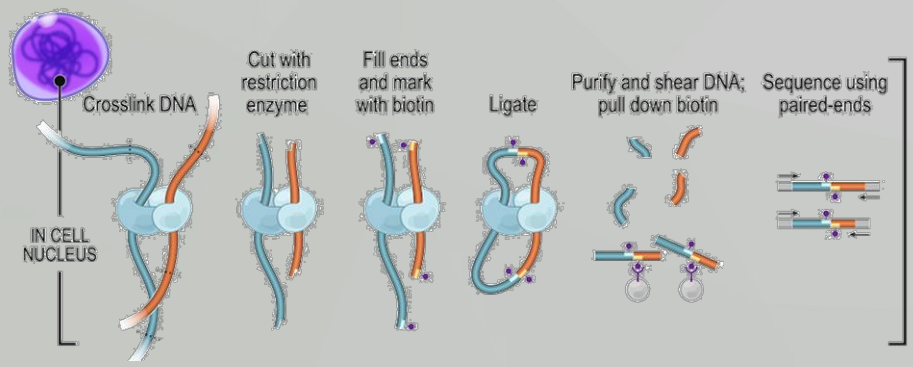
[1]The Center for Genome Architecture, Baylor College of Medicine, Houston, TX 77030, USA
[2]Department of Molecular and Human Genetics, Baylor College of Medicine, Houston, TX 77030, USA
[3]Department of Computer Science, Department of Computational and Applied Mathematics, Rice University, Houston, TX 77005, USA
[4]Broad Institute of MIT and Harvard, Cambridge, MA 02139, USA
[5]School of Engineering and Applied Sciences, Harvard University, Cambridge, MA 02138, USA
[6]Department of Computer Science, Stanford University, Stanford, CA 94305, USA
[7]Department of Biology, Massachusetts Institute of Technology (MIT), Cambridge, MA 02139, USA
[8]Department of Systems Biology, Harvard Medical School, Boston, MA 02115, USA
[9]Center for Theoretical Biological Physics, Rice University, Houston, TX 77030, USA
[10]Co-first author
*Correspondence: lander@broadinstitute.org (E.S.L.), erez@erez.com (E.L.A.)
http://dx.doi.org/10.1016/j.cell.2014.11.021

SUMMARY

We use in situ Hi-C to probe the 3D architecture of genomes, constructing haploid and diploid maps of nine cell types. The densest, in human lymphoblastoid cells, contains 4.9 billion contacts, achieving 1 kb resolution. We find that genomes are partitioned into contact domains (median length, 185 kb), which are associated with distinct patterns of histone marks and segregate into six subcompartments. We identify ~10,000 loops. These loops frequently link promoters and enhancers, correlate with gene activation, and show conservation across cell types and species. Loop anchors typically occur at domain boundaries and bind CTCF. CTCF sites at loop anchors occur predominantly (>90%) in a convergent orientation, with the asymmetric motifs "facing"

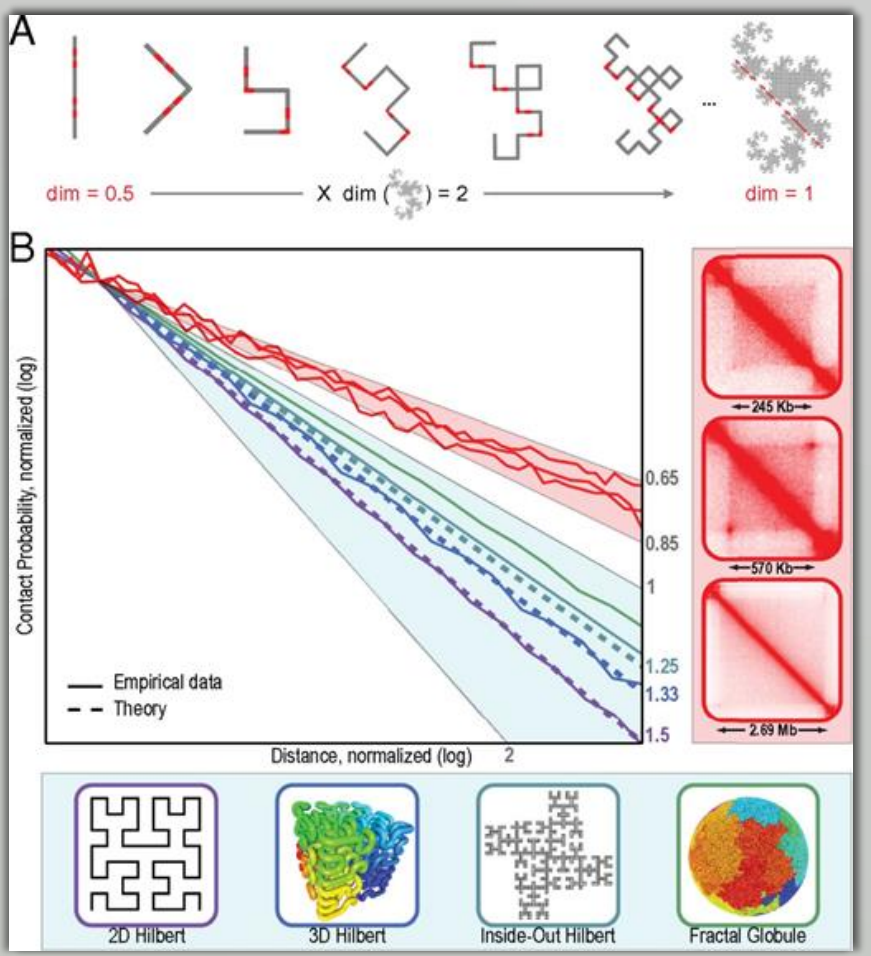
Various methods have emerged to assess the 3D architecture of the nucleus. In one seminal study, the binding of a protein to sites at opposite ends of a restriction fragment created a loop, which was detectable because it promoted the formation of DNA circles in the presence of ligase. Removal of the protein or either of its binding sites disrupted the loop, eliminating this "cyclization enhancement" (Mukherjee et al., 1988). Subsequent adaptations of cyclization enhancement made it possible to analyze chromatin folding in vivo, including nuclear ligation assay (Cullen et al., 1993) and chromosome conformation capture (Dekker et al., 2002), which analyze contacts made by a single locus, extensions such as 5C for examining several loci simultaneously (Dostie et al., 2006), and methods such as ChIA-PET for examining all loci bound by a specific protein (Fullwood et al., 2009).

To interrogate all loci at once, we developed Hi-C, which combines DNA proximity ligation with high-throughput sequencing in a genome-wide fashion (Lieberman-Aiden et al., 2009). We used



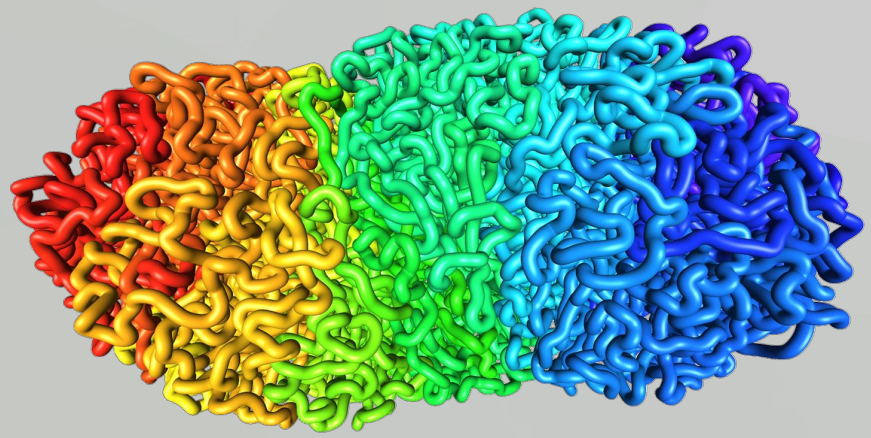
Hi-C

- A method that uses high-throughput sequencing to quantify interactions between pairs of genomic loci.
- It uses paired end sequencing where two fragments in close proximity are ligated together.
- The pair of sequences are aligned to the genome.
- Theoretically, all pairwise interactions are mapped.



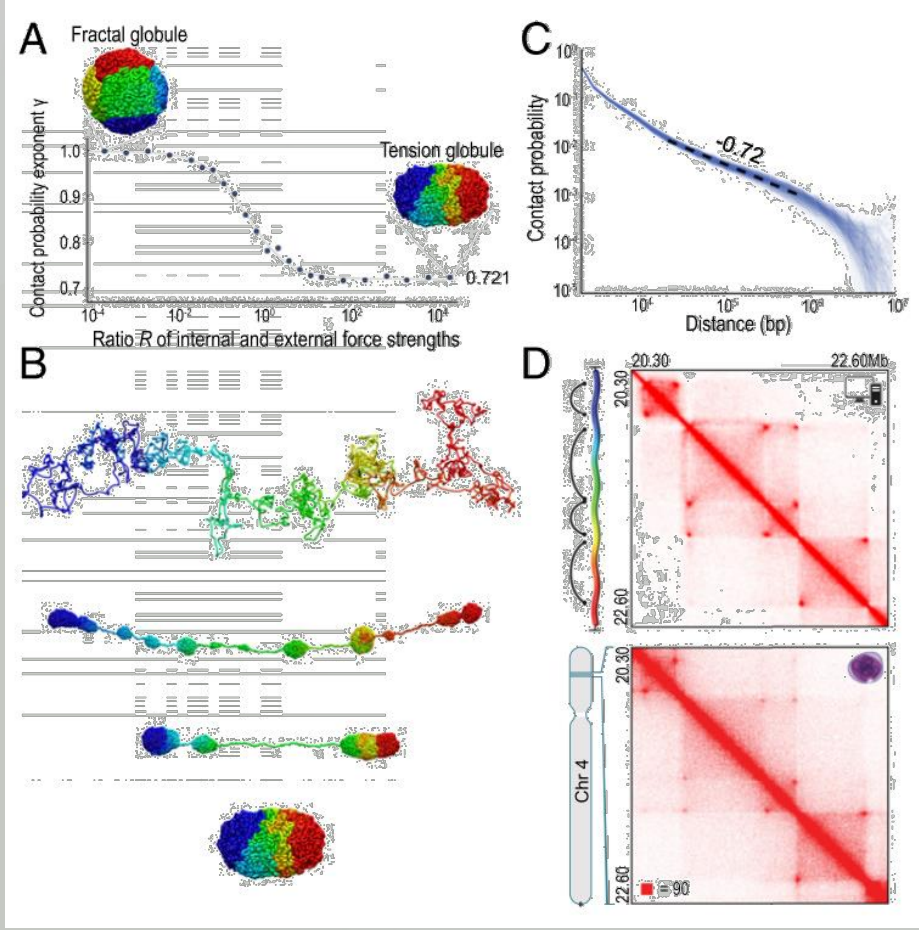
Contact Domains

- Contact probability scaling exponent γ for any fractal curve:
- $\gamma = 2 - (d_{surf}/d)$
- where d_{surf} is the dimension of the curve's surface and d is the dimension of its interior.
- **Contact domains are <u>not</u> strictly fractal!**



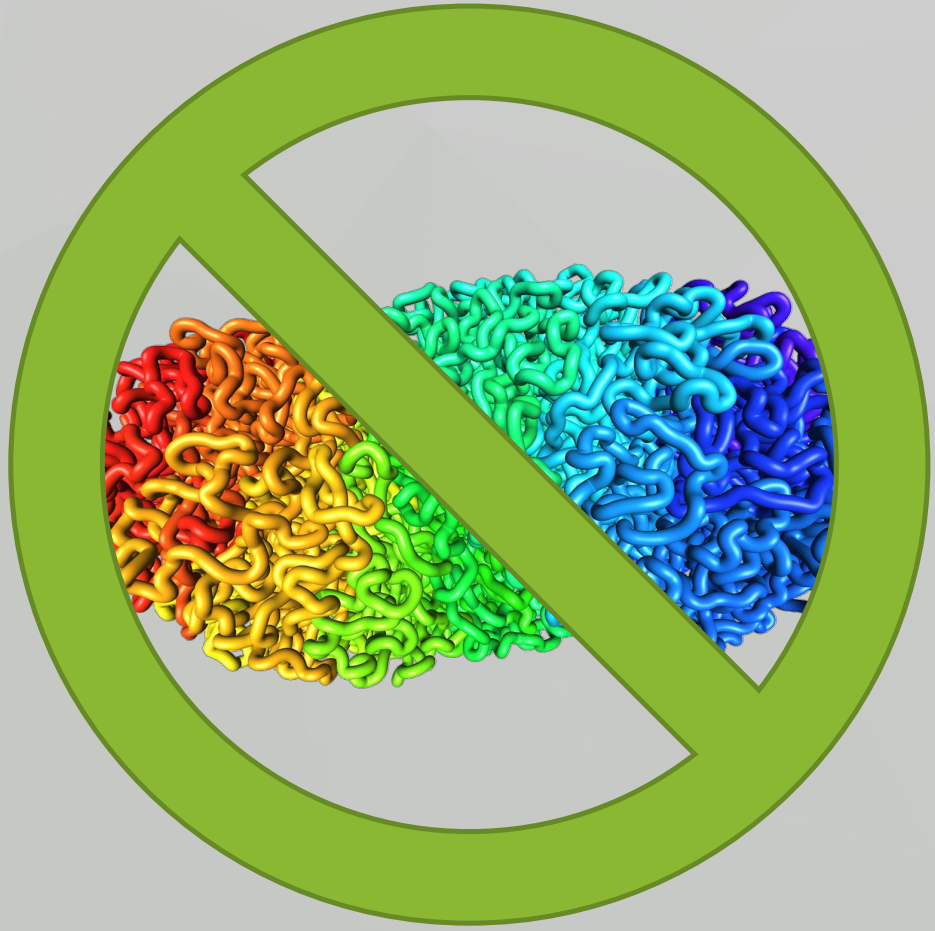
Tension Globule

- Data from Hi-C is consistent with a tension globule in which loops form by diffusion.
- Tiny globules form along an extended chain causing them to concatenate in a linear fashion.
- Contact domains form spontaneously



Tension Globule

- Data from Hi-C is consistent with a tension globule in which loops form by diffusion.
- Tiny globules form along an extended chain causing them to concatenate in a linear fashion.
- Contact domains form spontaneously

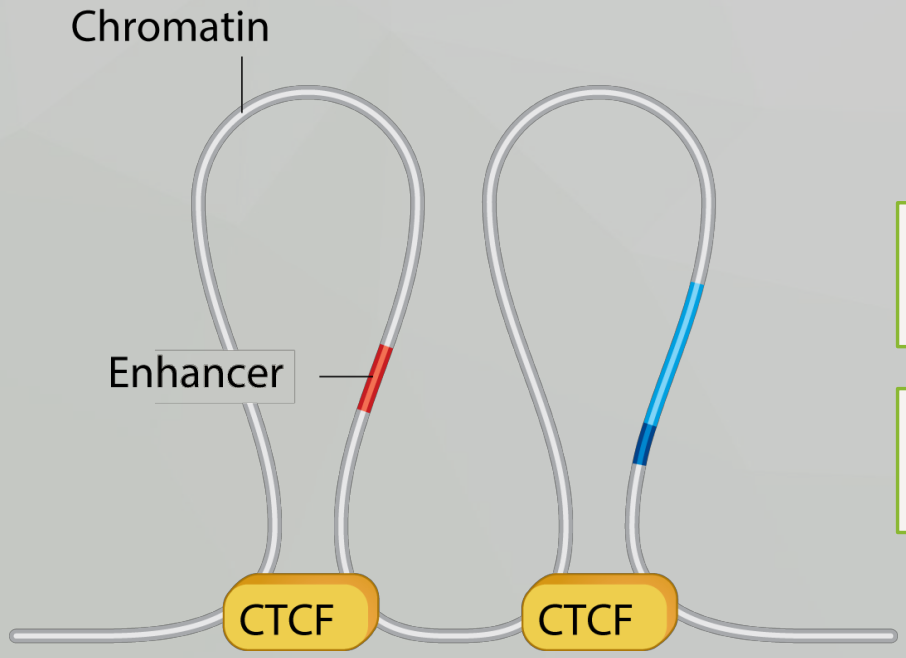


Inadequacies

- It doesn't explain the realities of loop formation
- In the tension globule model, loop formation would occur via diffusion but this would invariantly lead to loops that overlap and cause entanglements.
- However, their experimental data shows that this almost never happens.
- Further, the data doesn't explain why CTCF motifs must lie in the convergent orientation.

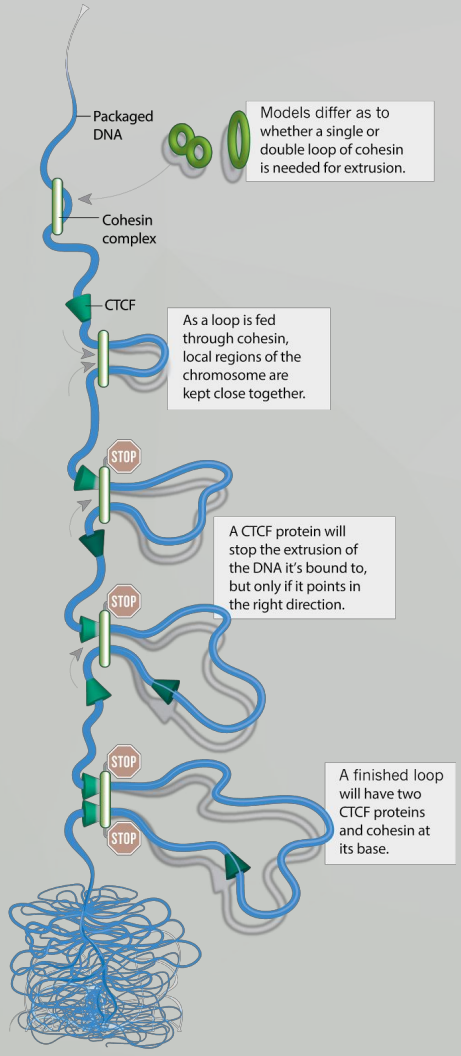
Froot Loops

- A highly nutritious, fruit-inspired cereal
- Goes best with a glass of Hi-C
- Ring-shaped topology
- Low “standard deviation” in flavor from loop to loop



Loops

- Topologically associated domains (TADs)
- CTCF

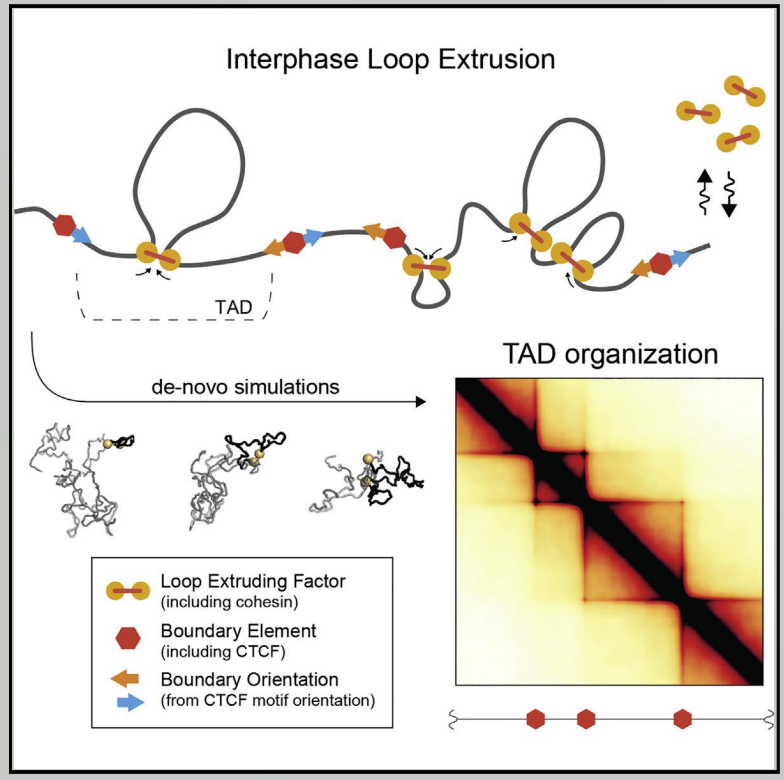


Extrusion Model

- To understand how chromatin folded in on themselves, we need to understand how loops form:
- **The extrusion model helps determine this.**

- CTCF was known to interact with cohesion at the base of each loop of uncondensed chromosomes
- CTCF had inherent directionality

- Knowing how CTCF interacts will help us understand how chromatin folds into globules and if we can use the extrusion model to predict chromatin formation and gene expression



Extrusion Model

Another group concurrently published a similar model of loop extrusion that mirrors many of the same conclusions.

OPEN ACCESS
CellPress

Cell Reports
Article

Formation of Chromosomal Domains by Loop Extrusion

Geoffrey Fudenberg,[1,2,6] Maxim Imakaev,[3,6] Carolyn Lu,[4] Anton Goloborodko,[3] Nezar Abdennur,[5] and Leonid A. Mirny[1,2,3,*]
[1]Graduate Program in Biophysics, Harvard University, Cambridge, MA 01238, USA
[2]Institute for Medical Engineering and Science, Massachusetts Institute of Technology (MIT), Cambridge, MA 02139, USA
[3]Department of Physics, Massachusetts Institute of Technology (MIT), Cambridge, MA 02139, USA
[4]Program for Research in Mathematics, Engineering and Science for High School Students (PRIMES) and Undergraduate Research Opportunities Program (UROP), MIT, Cambridge, MA 02139, USA
[5]PhD Program in Computational and Systems Biology, MIT, Cambridge, MA 02139, USA
[6]Co-first author, listed alphabetically
*Correspondence: leonid@mit.edu
http://dx.doi.org/10.1016/j.celrep.2016.04.085

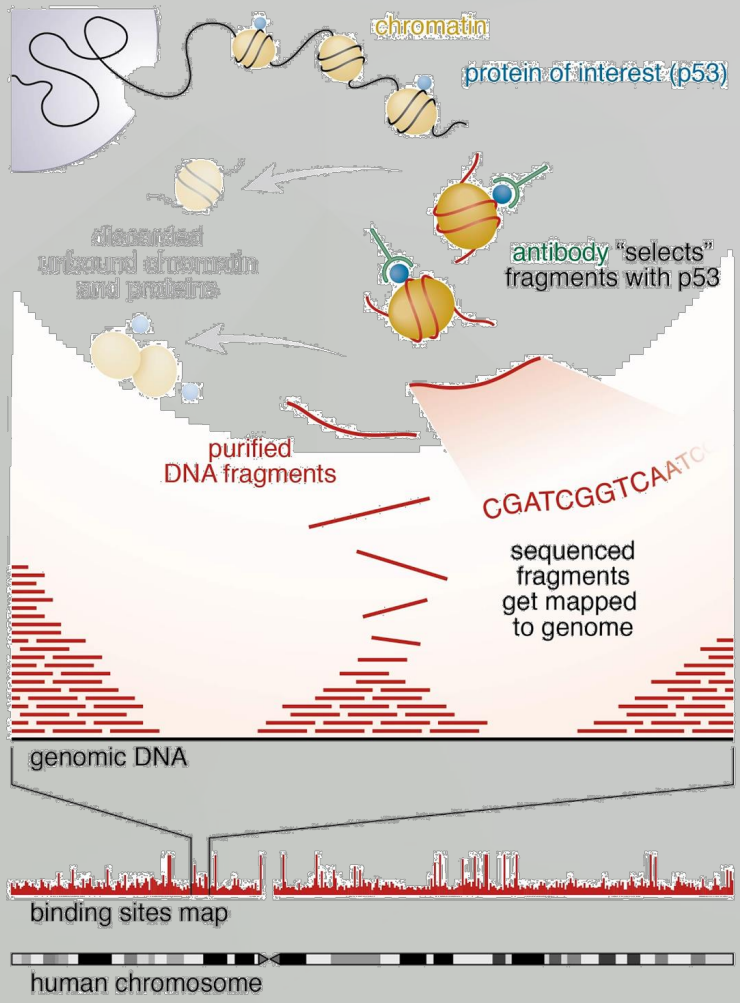
SUMMARY

Topologically associating domains (TADs) are fundamental structural and functional building blocks of human interphase chromosomes, yet the mechanisms of TAD formation remain unclear. Here, we propose that loop extrusion underlies TAD formation. In this process, *cis*-acting loop-extruding factors, likely cohesins, form progressively larger loops but stall at TAD boundaries due to interactions with boundary proteins, including CTCF. Using polymer simulations, we show that this model produces TADs and finer-scale features of Hi-C data. Each TAD emerges from multiple loops dynamically formed through extrusion, contrary to typical illustrations of single static loops. Loop extrusion both explains diverse experimental observations—including

ment (Andrey et al., 2013; Lupiáñez et al., 2015; Symmons et al., 2014).

TADs are contiguous regions of enriched contact frequency that appear as squares in a Hi-C map (Figure 1A), which are relatively insulated from neighboring regions. Many TADs have homogeneous interiors, while others have particularly enriched boundaries, or even more complex features. More recently, high-resolution maps revealed peaks of interactions between loci at the boundaries of TADs ("peak loci"; Rao et al., 2014). TADs differ from larger scale A/B compartments in that they do not necessarily form an alternating "checkerboard" pattern of enriched contact frequencies (Lajoie et al., 2015), and several TADs often reside within a single contiguous compartment (Gibcus and Dekker, 2013; Gorkin et al., 2014) (Supplemental Notes).

Although often illustrated as such, several lines of evidence indicate that TADs are not simply stable loops formed between pairs of boundary loci. First, only 50% of TADs have corner-



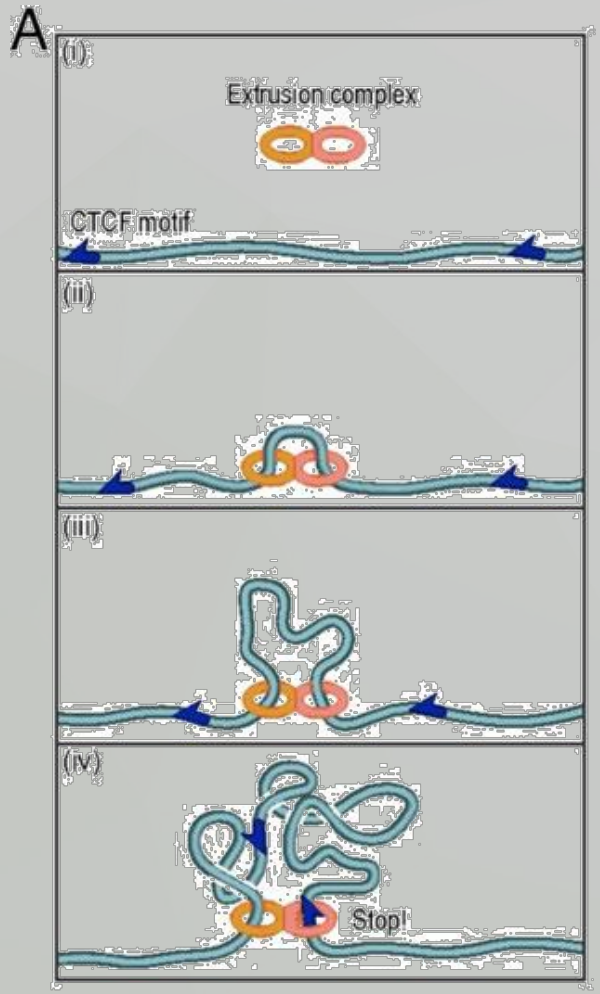
CHIP-Seq:

Protein Interactions

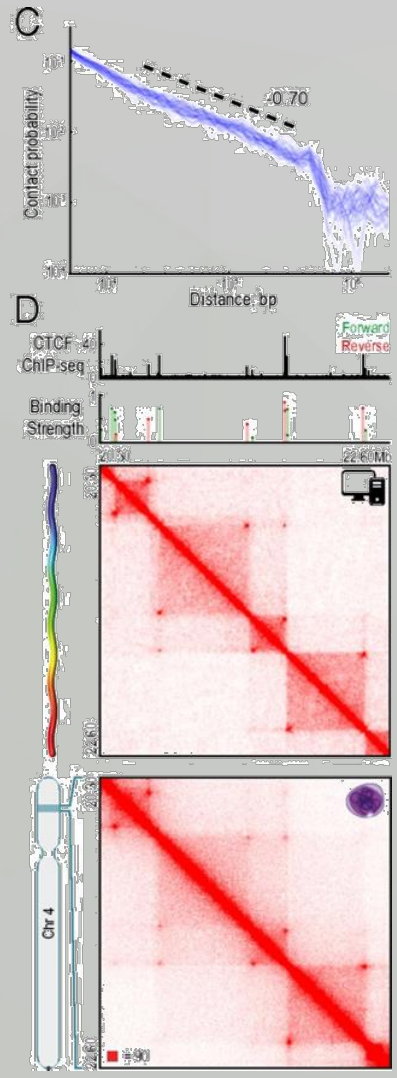
Maps global binding sites for a given protein!

How does it work?

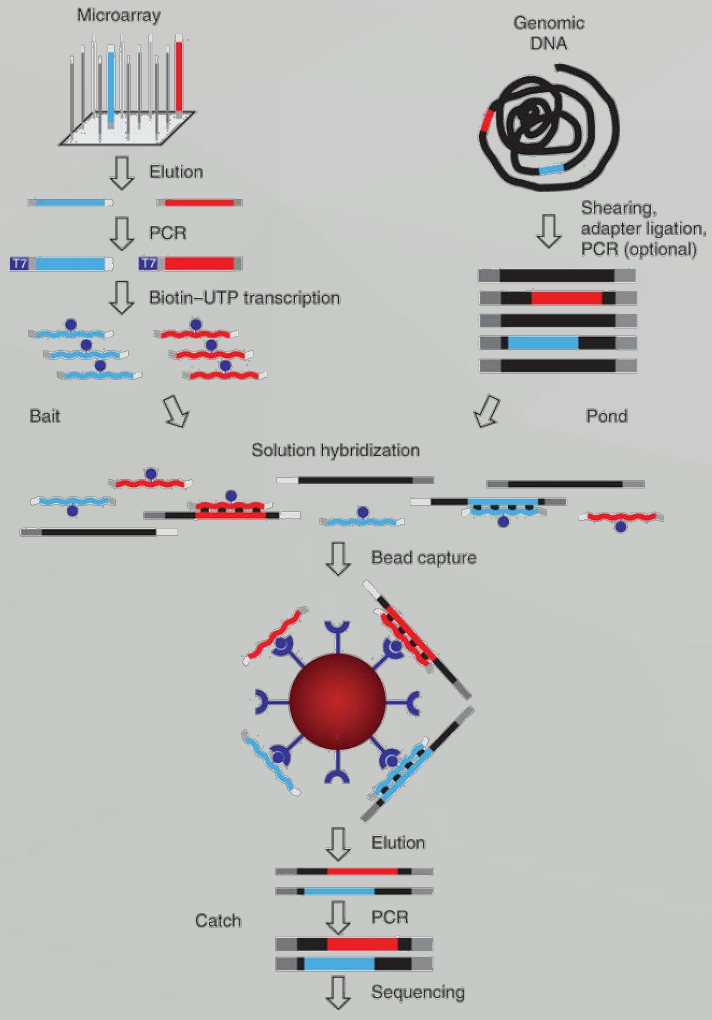
1. Crosslink DNA-protein complexes
2. Immunoprecipitate target protein
3. Purify the DNA and sequence!



Extrusion Model can recapitulate Hi-C data with CHIP-Seq alone!

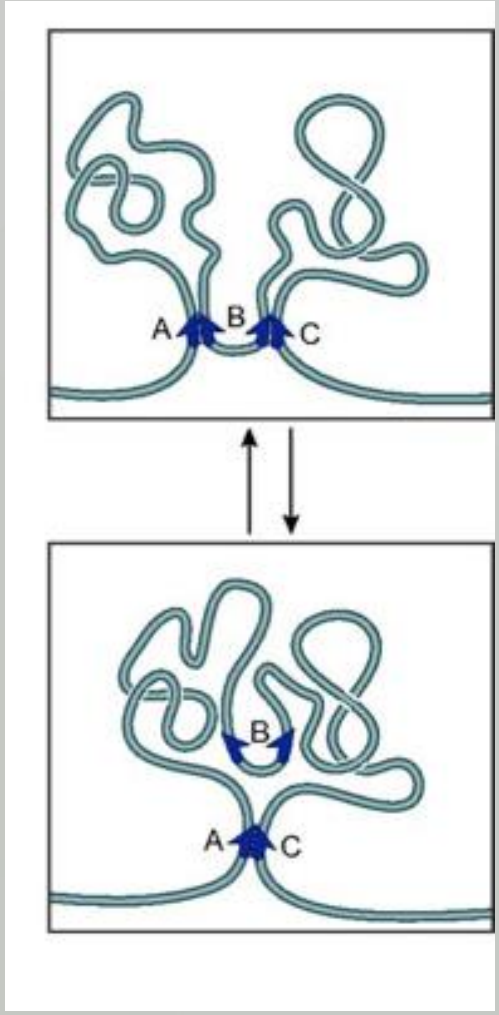


Extrusion Model can recapitulate Hi-C data with CHIP-Seq alone!



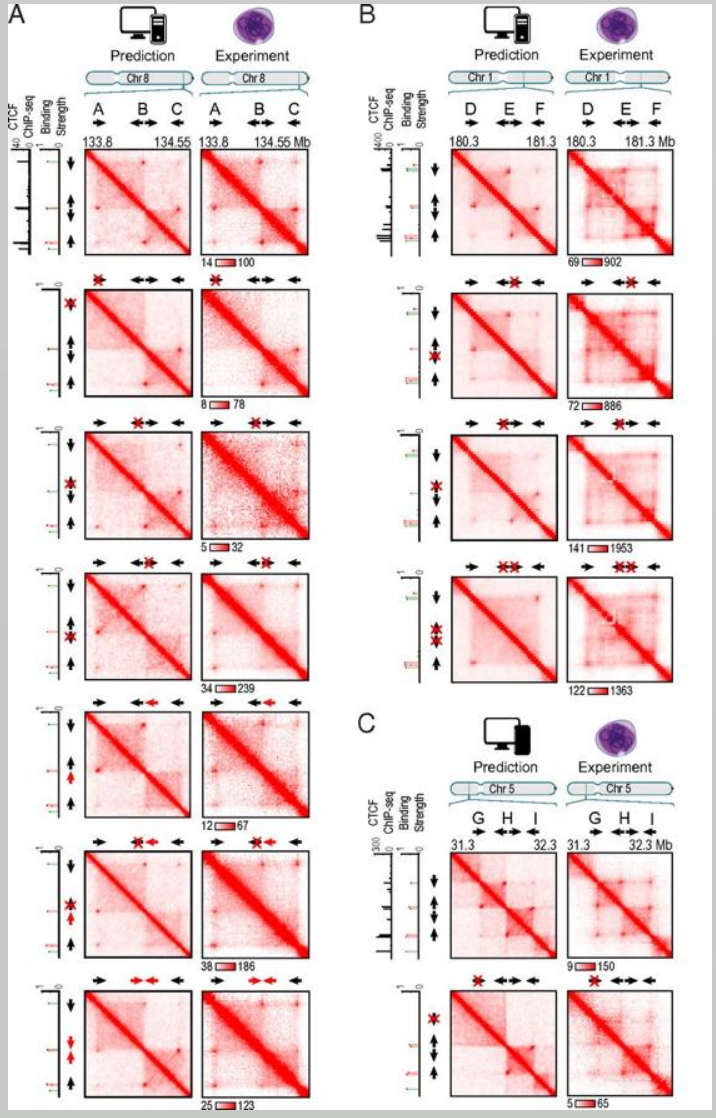
Hybrid Capture on the *in situ* library (Hi-C²)

Combines **targeted genomic capture** and **existing situ Hi-C libraries** to observe conformation changes in selected genomic regions

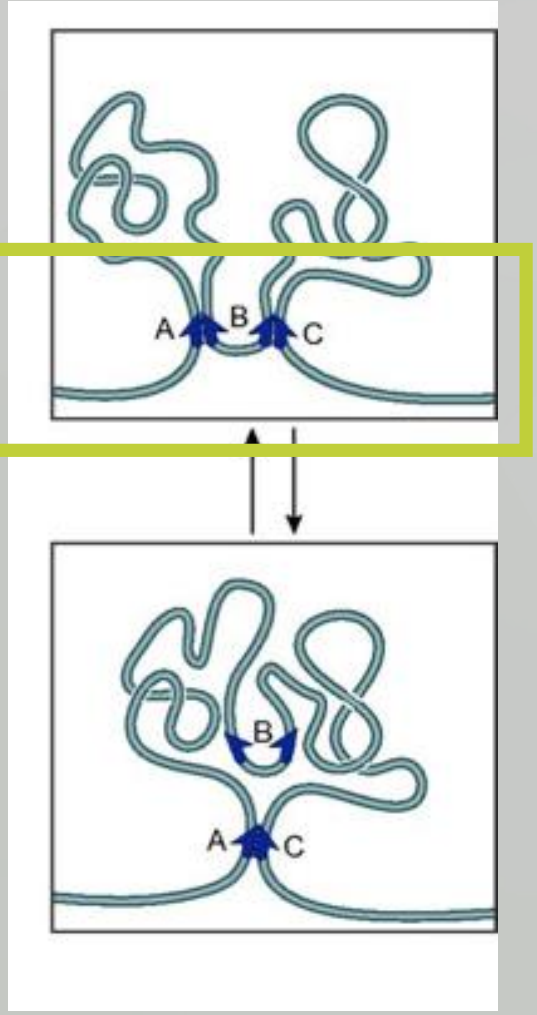


CRISPR Engineering Genomes Hi-C

Locus A: forward-oriented CTCF motif (AF)
Locus B: reverse-oriented CTCF motif (BR)
followed by forward-oriented motif (BF)
Locus C: reverse-oriented CTCF motif (CR)

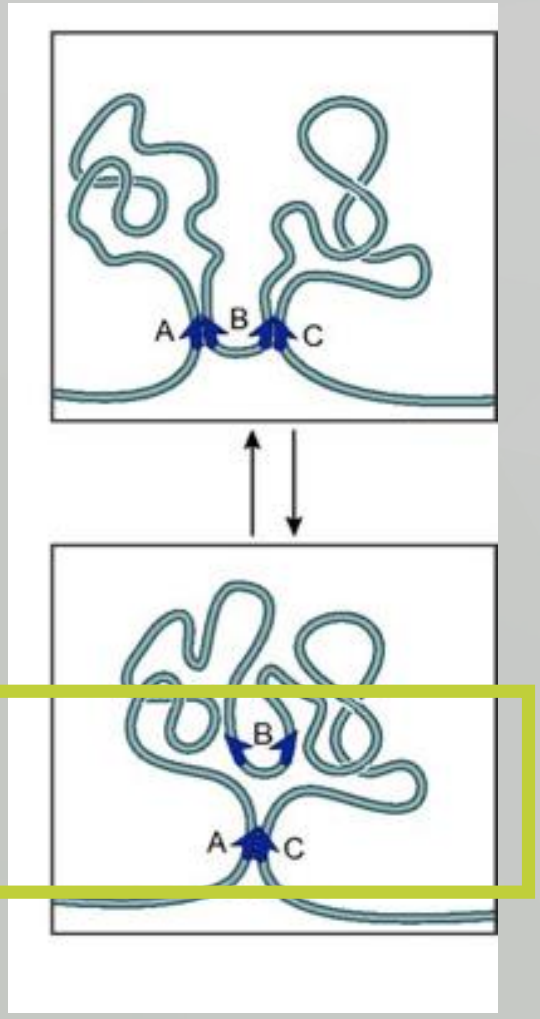


Disrupting CTCF motifs impairs loop formation!



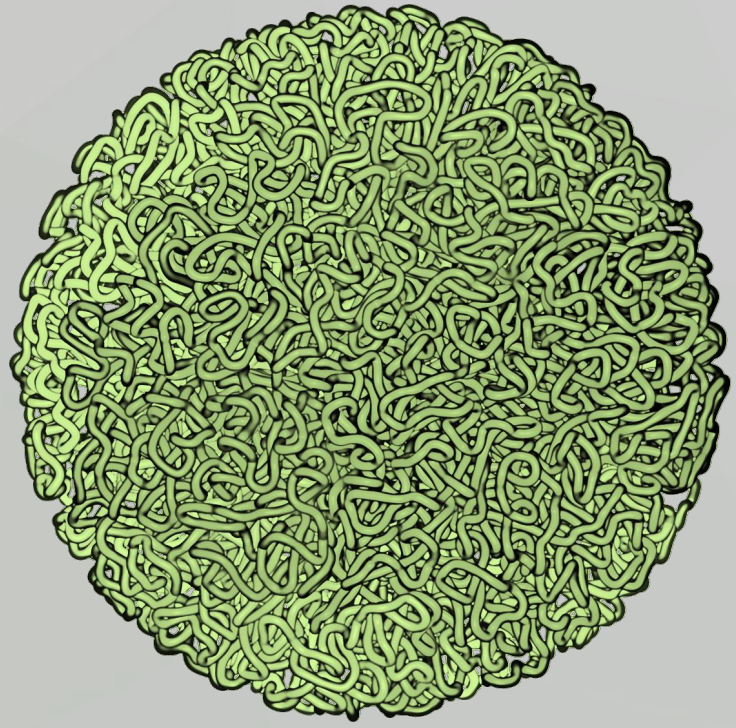
CRISPR Engineering Genomes Hi-C

1. **Consecutive loops tend to occur together**
2. Larger loops tend to occur in a cell-type dependent manner



CRISPR Engineering Genomes Hi-C

1. Consecutive loops tend to occur together
2. **Larger loops tend to occur in a cell-type dependent manner**



Take-aways

- DNA in the nucleus is organized into well-defined compartments and domains
- A loop extrusion model explains many of the key characteristics of experimental data
- The major role of CTCF and cohesin may be to form and maintain these loop structures


What's on the Horizon?

Experimentally proving the loop extrusion model

REPORT

CHROMOSOMES

Bacillus subtilis SMC complexes juxtapose chromosome arms as they travel from origin to terminus

Xindan Wang,[1]*† Hugo B. Brandão,[2] Tung B. K. Le,[3]‡ Michael T. Laub,[3,4] David Z. Rudner[1]†

Structural maintenance of chromosomes (SMC) complexes play critical roles in chromosome dynamics in virtually all organisms, but how they function remains poorly understood. In the bacterium *Bacillus subtilis*, SMC-condensin complexes are topologically loaded at centromeric sites adjacent to the replication origin. Here we provide evidence that these ring-shaped assemblies tether the left and right chromosome arms together while traveling from the origin to the terminus (>2 megabases) at rates >50 kilobases per minute. Condensin movement scales linearly with time, providing evidence for an active transport mechanism. These data support a model in which SMC complexes function by processively enlarging DNA loops. Loop formation followed by processive enlargement provides a mechanism by which condensin complexes compact and resolve sister chromatids in mitosis and by which cohesin generates topologically associating domains during interphase.

Recent chromosome conformation capture (Hi-C) studies (*1*–*4*) and polymer simulations (*2*, *5*–*7*) have reinvigorated a model proposed over a decade ago (*8*) in which structural maintenance of chromosomes (SMC) complexes generate DNA loops through processive loop enlargement (also referred to as loop extrusion) (*5*, *6*). In this model, these ring-shaped complexes encircle the DNA flanking their loading site, tethering the DNA duplexes together. As these tethers move away from their loading site, they generate loops. Moreover, if SMC rings are continuously loaded at the same site, then the DNA duplexes within the loop segment become juxtaposed (Fig. 1E). De novo loop generation along chromosome arms provides a simple solution to explain how condensin complexes could compact and resolve replicated chromosomes into rod-shaped structures during mitosis and provides a mechanism for the formation of topologically associating domains (TADs) by the SMC-cohesin complex during interphase (*3*, *6*, *9*). Although compelling in its simplicity, this model remains largely untested.

In the bacterium *Bacillus subtilis*, the SMC-condensin complex is required for the resolution and segregation of newly replicated sister origins (*10*, *11*). Condensin is recruited to the origin by the broadly conserved partitioning protein ParB bound to centromeric *parS* sites located adjacent to the origin (*12*–*14*). Like its eukaryotic counterparts, *B. subtilis* condensin encircles chromosomal DNA in vivo, and topological entrapment is strongly reduced in the absence of ParB, suggesting that most condensin is loaded onto the chromosome at *parS* sites (*15*). Recent Hi-C studies in *B. subtilis* revealed that recruitment of condensin to origin-proximal *parS* sites is required for the alignment of the left and right chromosome arms (*1*, *4*). However, the mechanism by which condensin promotes the juxtaposition of DNA flanking its loading site remains unknown.

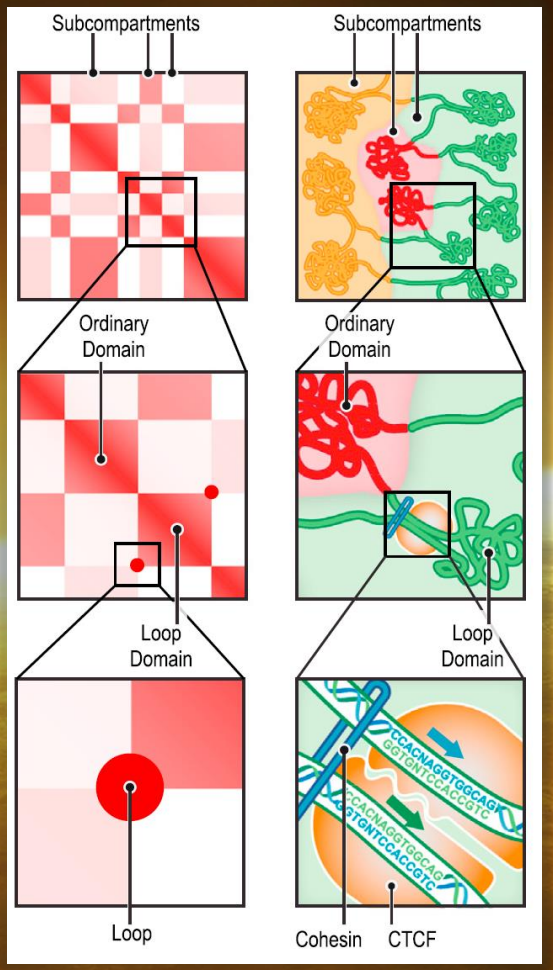
To investigate whether condensin-dependent DNA juxtaposition (or "zip-up") initiates at *parS* and progresses down the flanking DNA, we used a strain with a single origin-proximal *parS* site (at −1°, where a degree is ~11.2 kb of DNA and "−" is counterclockwise from the origin of replication) that harbors an isopropyl-β-D-thiogalactopyranoside (IPTG)–inducible allele of *parB* as the sole source of the loader, enabling us to follow zip-up dynamics. Interarm interactions were monitored by Hi-C at 5-min intervals after ParB induction (Fig. 1A to C). left and right arms. Analysis of all seven Hi-C time points indicated that the two arms zipped-up at a nearly constant rate of 52 ± 5 kb/min (figs. S1 and S2). Chromatin immunoprecipitation–sequencing (ChIP-seq) using antibodies against SMC (anti-SMC) before and after ParB induction revealed modest SMC enrichment that correlated with the extent of interarm interaction observed by Hi-C (Fig. 1D and fig. S1). These results and results of previous studies (*1*, *4*, *15*, *16*) suggest that condensin is loaded at *parS* and then progressively accumulates along the flanking DNA (Fig. 1E and fig. S3, A to C).

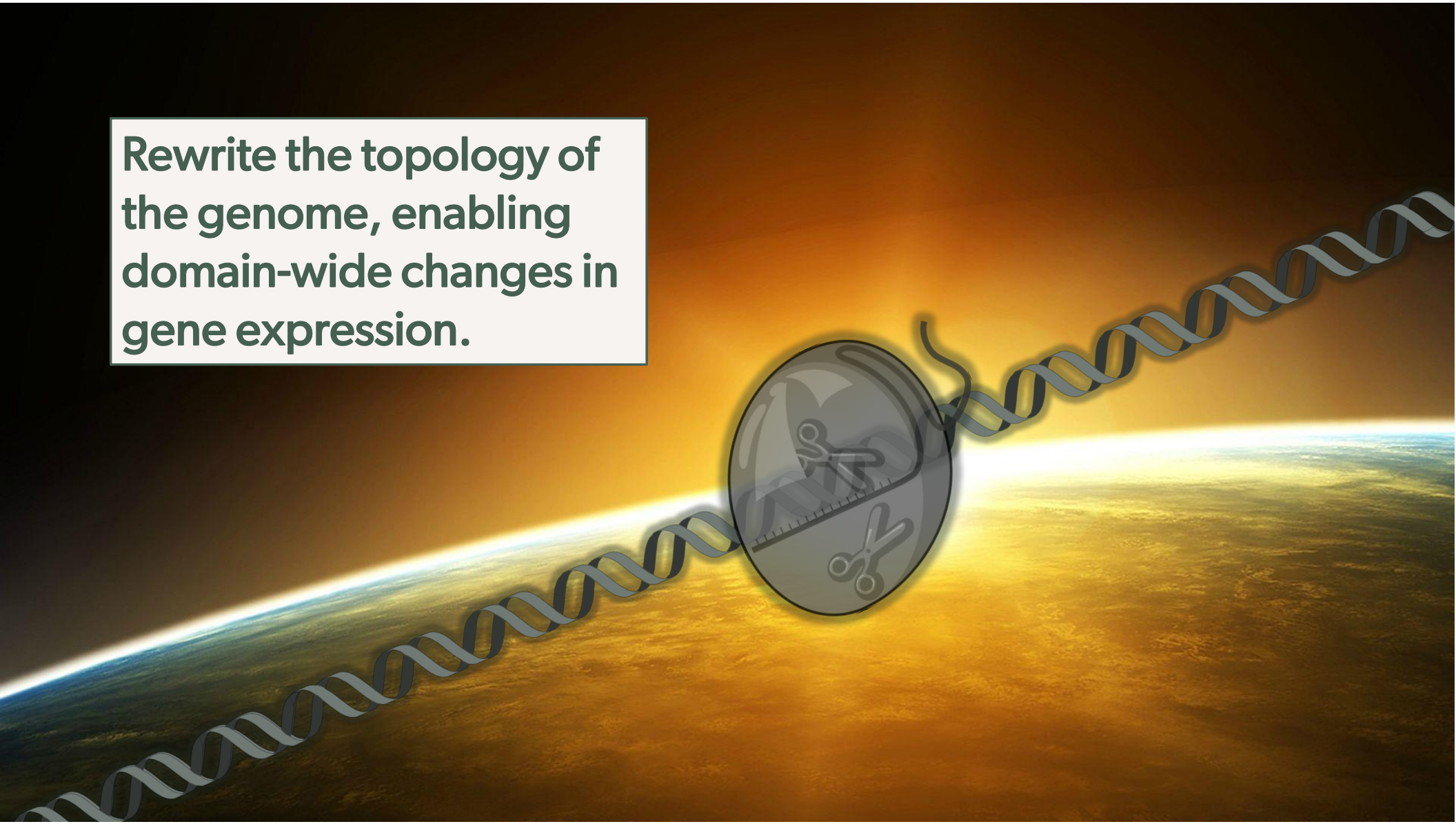
Condensin promotes the juxtaposition of large tracts of DNA flanking *parS* sites inserted at ectopic chromosomal positions (*4*). To determine whether SMC was specifically enriched along these juxtaposed regions, we performed Hi-C and ChIP-seq on four *B. subtilis* strains, each with an ectopic *parS* site at a different position along the left chromosome arm (Fig. 2 and fig. S4). The Hi-C contact maps indicated that DNA flanking the ectopic *parS* sites interacted, giving characteristic zip-up patterns (Fig. 2A and fig. S4A). As observed previously (*4*), the zip-ups were asymmetric, containing more terminus-proximal than origin-proximal DNA (discussed below). Importantly, the ChIP-seq profiles revealed strong SMC enrichment along the DNA flanking the *parS* sites that correlated with the extent of juxtaposition (Fig. 2B and figs. S4B, S5, and S6, A and C). Results of ChIP-seq with antibodies against the other two subunits of the condensin complex (subunits ScpA and ScpB) showed similar enrichment profiles (fig. S4D), whereas ParB enrichment was limited to small chromosomal regions (12 to 23 kb) centered on *parS* (*12*) (fig. S4C).

For unknown reasons, regardless of where the ectopic *parS* site was inserted along the left or right arm, the zip-up did not extend beyond a ~170-kb region surrounding the replication terminus (Fig. 2A and figs. S4A and S7). Examination of SMC enrichment in the ChIP-seq profiles indicated that condensin did not appreciably accumulate in this region (Fig. 2B and figs. S4B and S6A), suggesting that the complexes were actively dissociated. Similarly, SMC did not accumulate where enrichment ended on the origin-proximal side of *parS* (Fig. 2B and figs. S4B and S6A). The mechanism by which condensin is released from the DNA is currently unknown. A comparison of the Hi-C contact maps from the four strains revealed that the greater the distance between the ectopic *parS* site and the terminus region, the greater the extent of juxtaposition from *parS* toward the origin (Fig. 2 and figs. S4, A and B, and S5, A and C). Similarly, the distance between the ectopic *parS* and the terminus region correlated with the extent of SMC enrichment from the *parS* toward the origin. These data indicate that

[1]Department of Microbiology and Immunobiology, Harvard Medical School, 77 Avenue Louis Pasteur, Boston, MA 02115, USA. [2]Graduate Program in Biophysics, Harvard University, Cambridge, MA 02138, USA. [3]Department of Biology

What mechanisms are responsible for compartments and subcompartments?




Rewrite the topology of the genome, enabling domain-wide changes in gene expression.

References:

- Wang, X., Brandão, H. B., Le, T. B. K., Laub, M. T. & Rudner, D. Z. Bacillus subtilis SMC complexes juxtapose chromosome arms as they travel from origin to terminus. Science (80-.). 527, 524–527 (2017).
- Rao, S. S. P. et al. A 3D map of the human genome at kilobase resolution reveals principles of chromatin looping. Cell 159, 1665–1680 (2014).
- Lieberman-aiden, E. et al. Comprehensive Mapping of Long-Range Interactions Reveals Folding Principles of the Human Genome. Science (80-.). 326, 289–293 (2009).
- Fudenberg, G. et al. Formation of Chromosomal Domains by Loop Extrusion. Cell Rep. 15, 2038–2049 (2016).
- Sanborn, A. L. et al. Chromatin extrusion explains key features of loop and domain formation in wild-type and engineered genomes. Proc. Natl. Acad. Sci. 112, 201518552 (2015).