

3D Genome Organization



Doctoral Program in

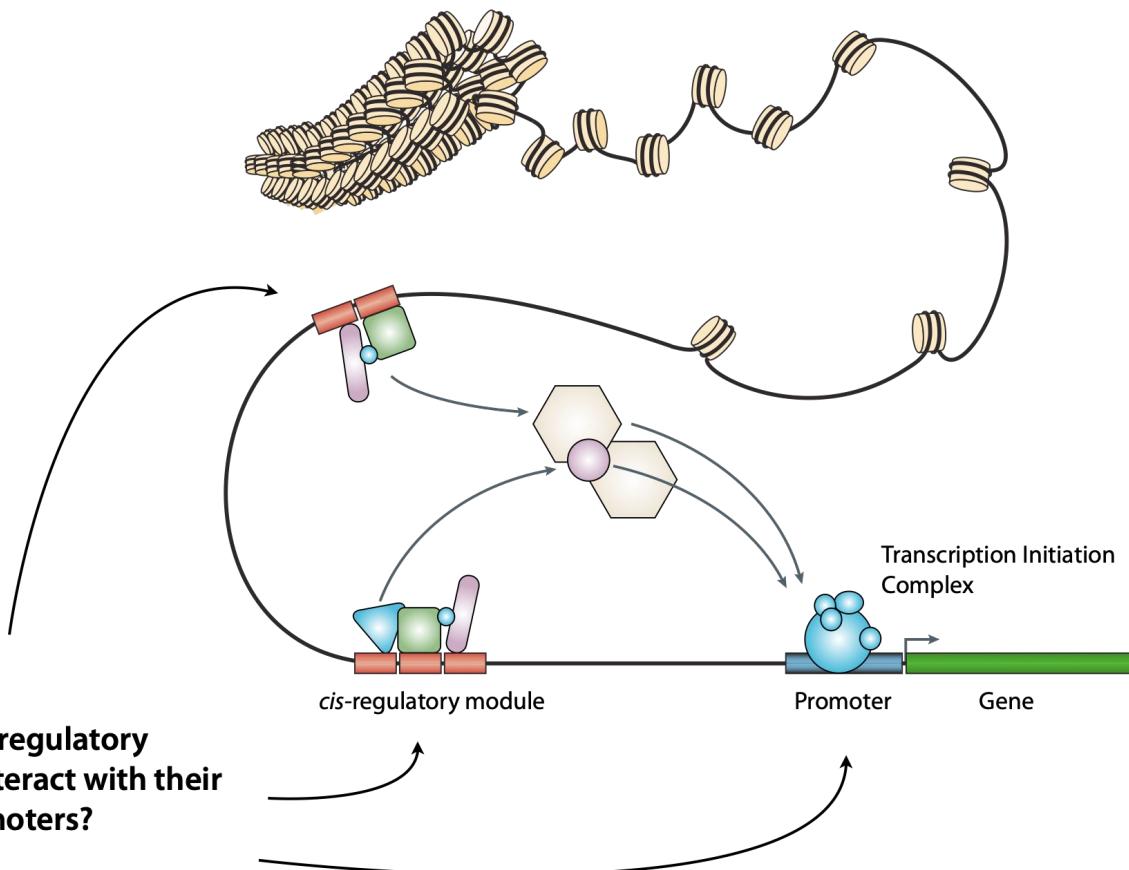
**Cell, Molecular, Developmental
Biology, and Biophysics**



JOHNS HOPKINS
UNIVERSITY

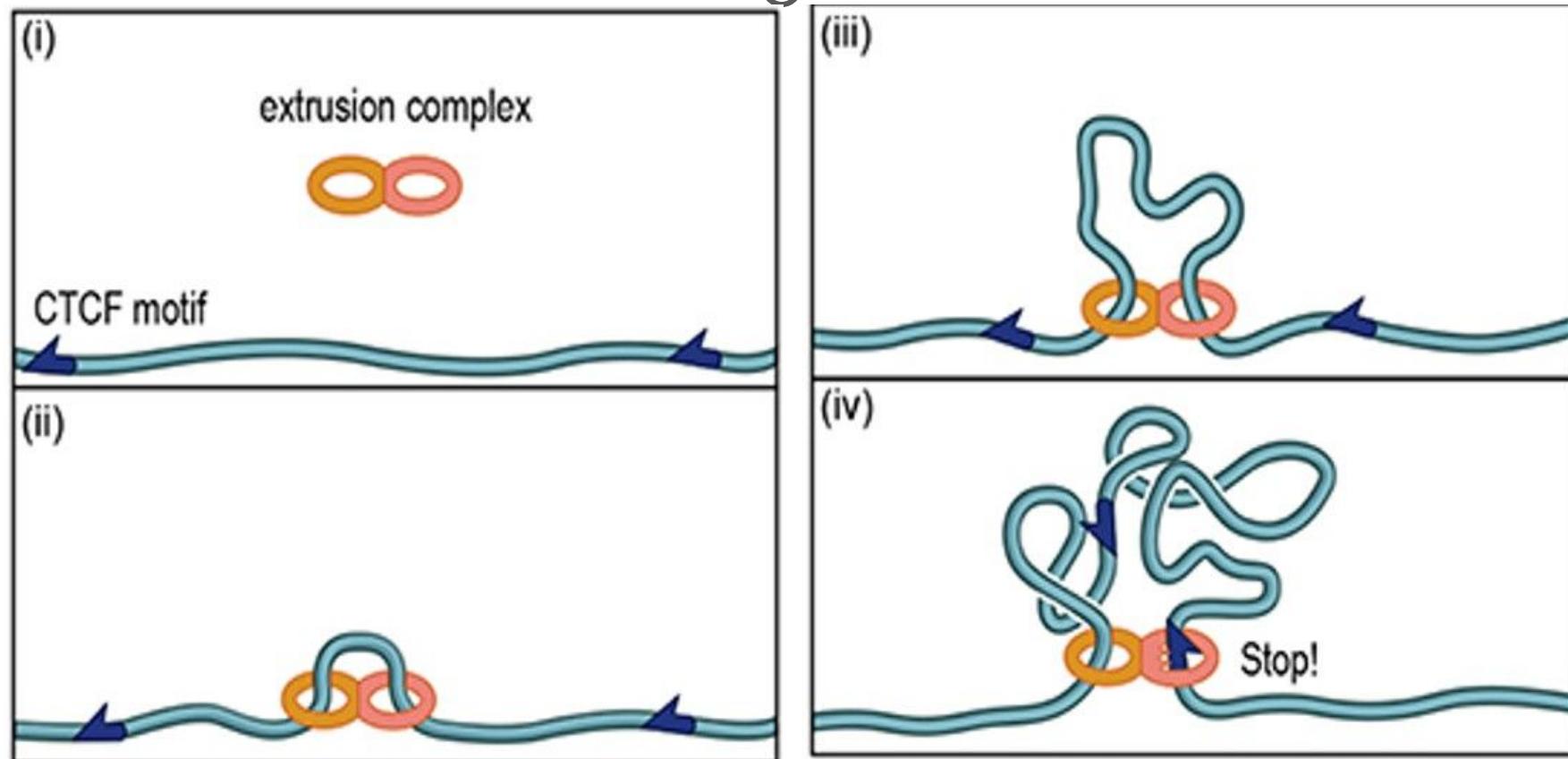
Why do we care about spatial
organization?

How do distal elements interact?



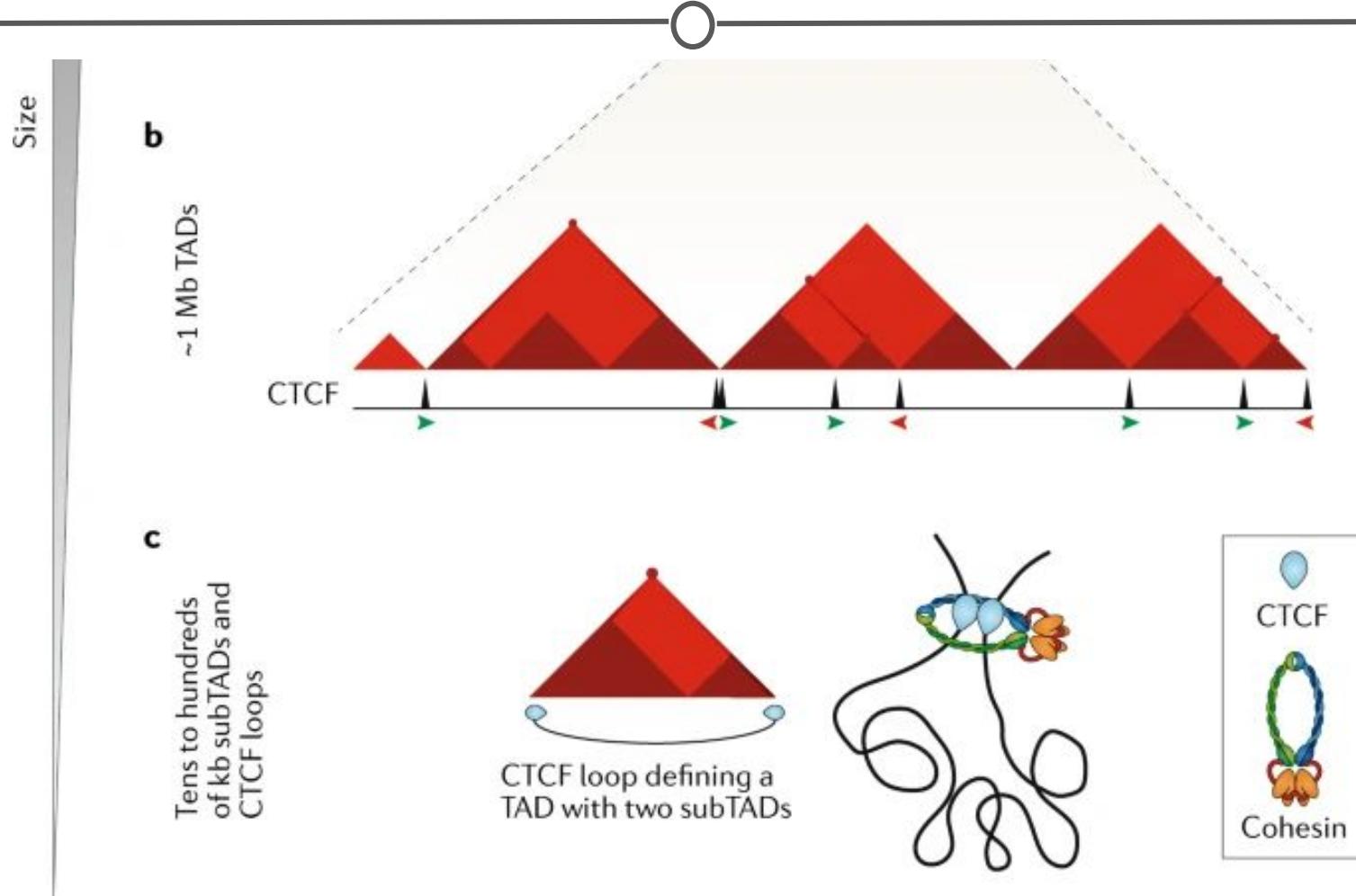
How do *cis*-regulatory modules interact with their target promoters?

Cohesin-mediated looping

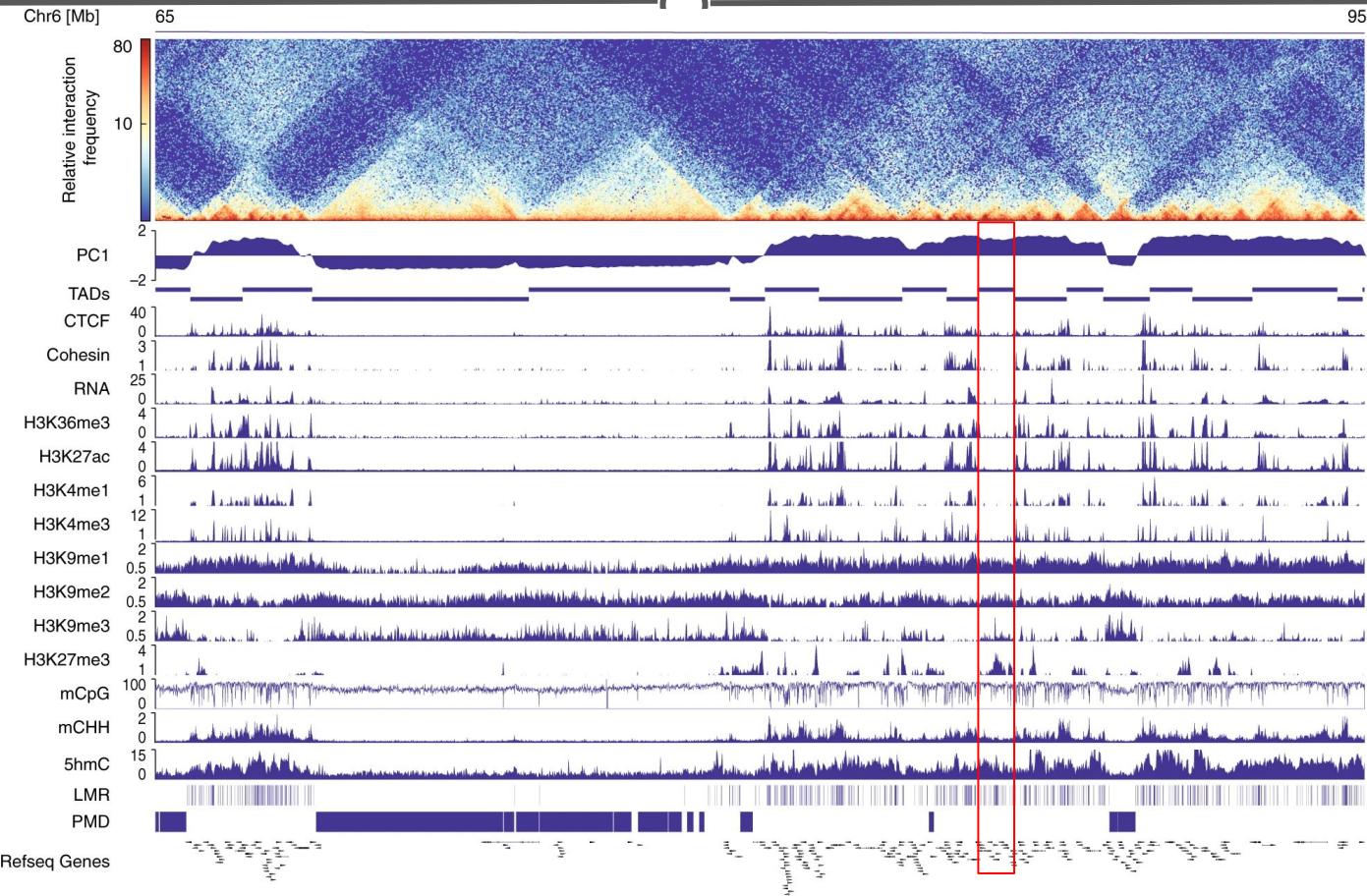


Sanborn AL, Rao SS, Huang SC, Durand NC, Huntley MH, Jewett AI, Bochkov ID, Chinnappan D, Cutkosky A, Li J, Geeting KP, Gnrke A, Melnikov A, McKenna D, Stamenova EK, Lander ES, Aiden EL. Chromatin extrusion explains key features of loop and domain formation in wild-type and engineered genomes. Proc Natl Acad Sci U S A. 2015 Nov 24;112(47):E6456-65.

Topologically associating domains (TADs)

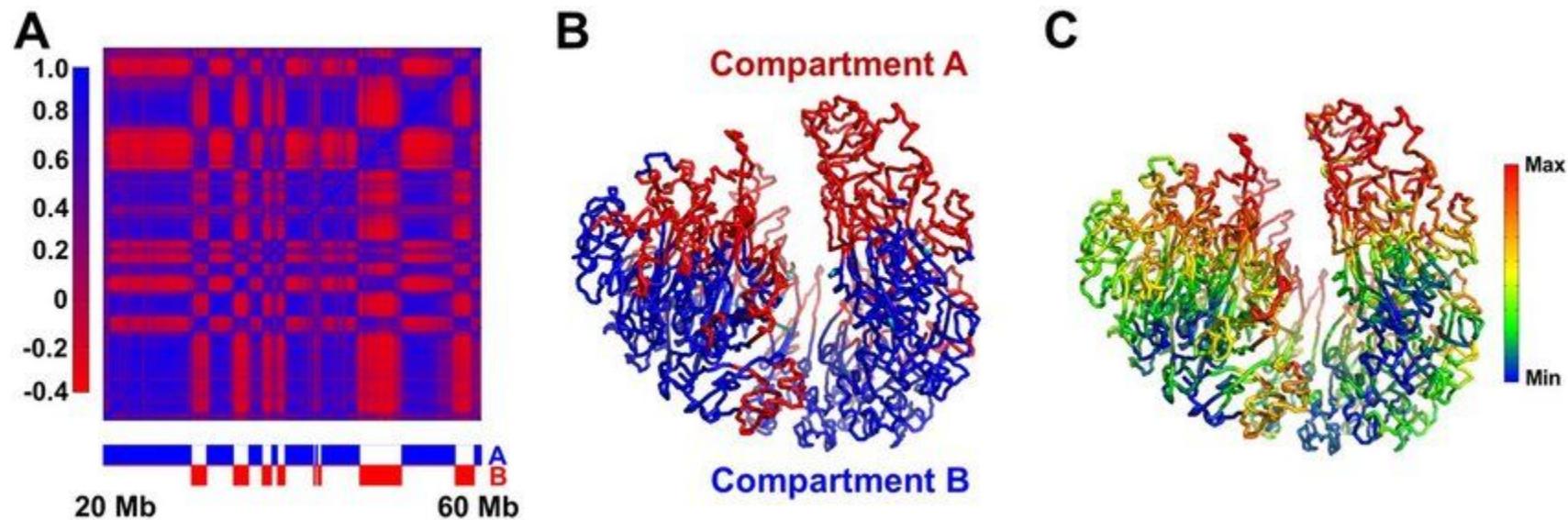


Sub-TADs delineate epigenetic marks

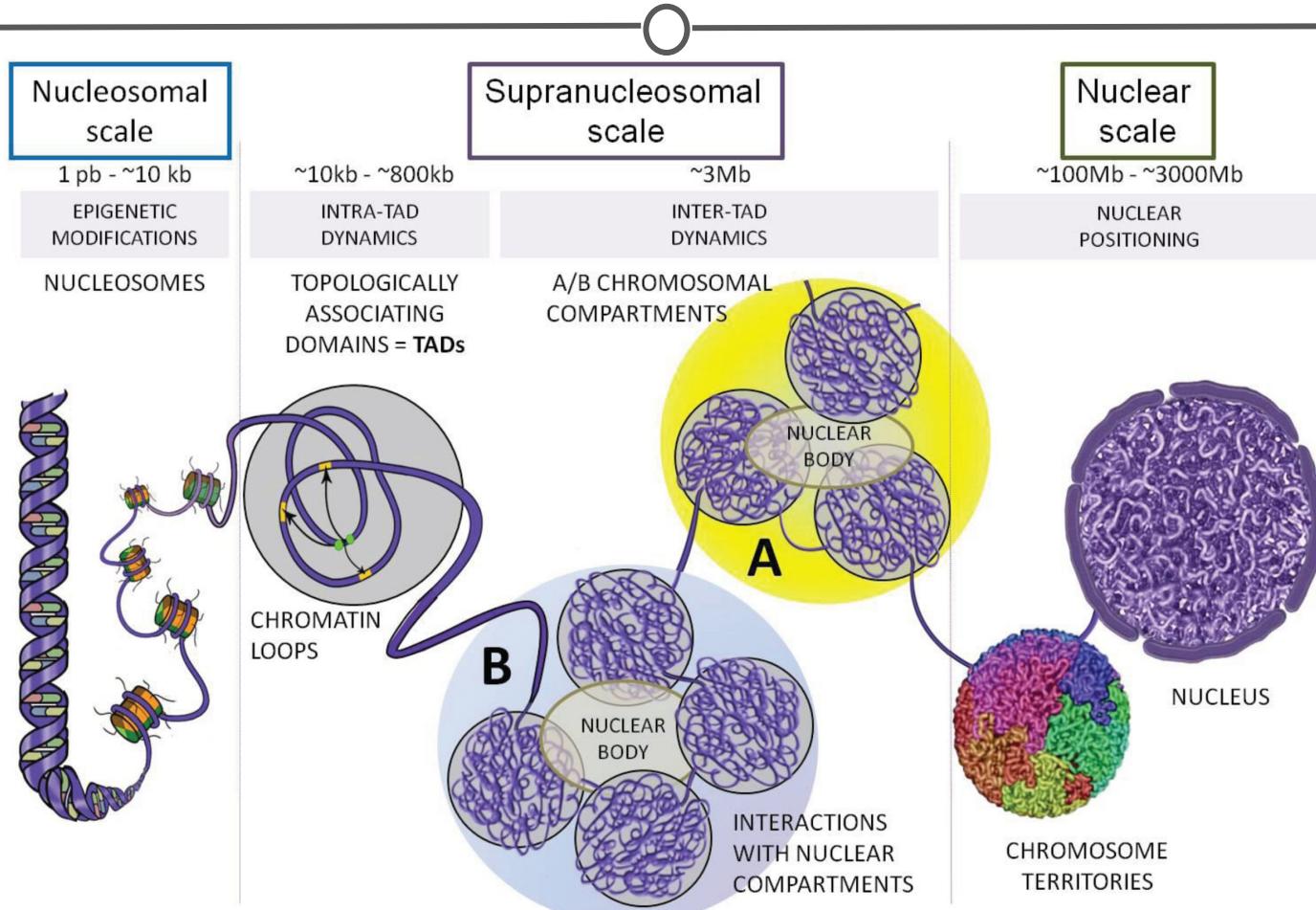


Nothjunge S, Nührenberg TG, Grüning BA, Doppler SA, Preissl S, Schwaderer M, Rommel C, Krane M, Hein L, Gilzbach R. DNA methylation signatures follow preformed chromatin compartments in cardiac myocytes. Nat Commun. 2017 Nov 21;8(1):1667.

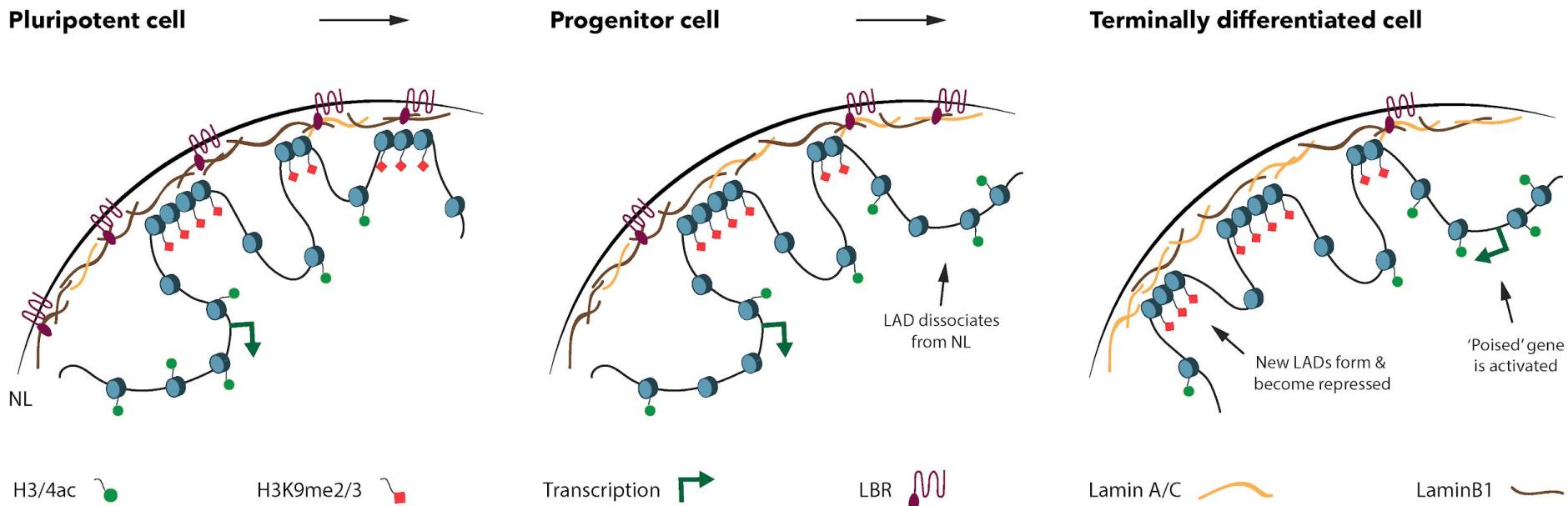
Chromatin compartmentalization



Chromosome territories



Lamina associated domains (LADs)

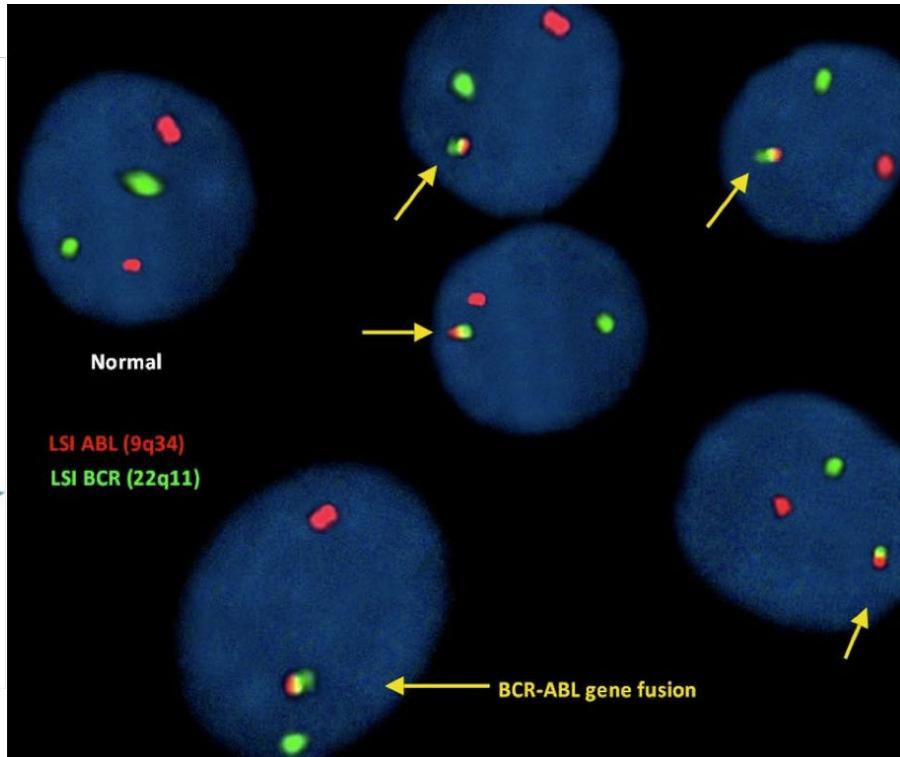
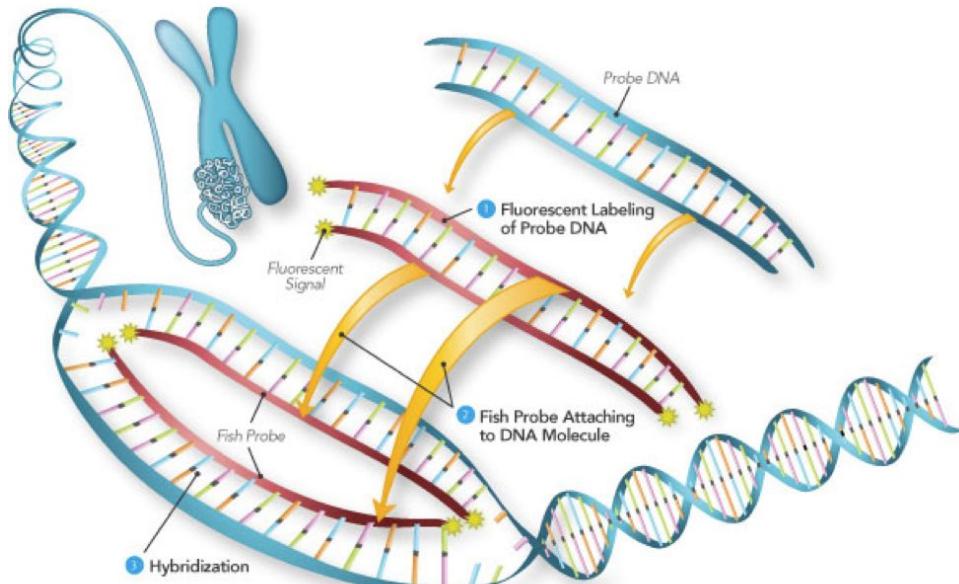


Lochs SJA, Kefalopoulou S, Kind J. Lamina Associated Domains and Gene Regulation in Development and Cancer. Cells. 2019 Mar 21;8(3):271.

Assaying spatial organization



Fluorescence in-situ hybridization (FISH)



<https://www.semrock.com/fish.aspx>

https://www.researchgate.net/figure/Fluorescence-in-situ-hybridization-FISH-for-the-detection-of-9-22q34-q11-BCR_fig1_301779737

Limitations of FISH



- Low throughput
- Limited resolution
- Limited number of unique channels

Can we turn this into a sequencing problem?

Enter 3C (and all it's little baby Cs)



REPORTS

antibody to chicken C μ (M1) (Southern Biotechnology Associates, Birmingham, AL) and then with polyclonal fluorescein isothiocyanate-conjugated goat antibodies to mouse IgG (Fab) $_2$ (Sigma). Predominantly slgM(+) subclones were excluded from the analysis, because they most likely originated from cell that were already slgM(+) at the time of subcloning.

23. For Ig light chain sequencing, PCR amplification and sequencing of the rearranged light chain V segments were performed as previously described

(19), except that high-fidelity PfuTurbo polymerase (Stratagene) was used with primer pair V λ 1/V λ 2 for PCR, and primer V λ 3 was used for sequencing (17). Only one nucleotide change, which most likely reflects a PCR-introduced artifact, was noticed in the V-j-3' intron region in a total of 80 0.5-kb-long sequences from AID $^{-/-}$ cells.

24. We thank M. Reth and T. Brummer for kindly providing the MerCreMer plasmid vector; P. Carninci and Y. Hayashizaki for construction of the riken1 bursal

cDNA library; A. Peters and K. Jablonski for excellent technical help; and C. Stocking and J. Löhrer for carefully reading the manuscript. Supported by grant Bu 631/2-1 from the Deutsche Forschungsgemeinschaft, by the European Union Framework V programs "Chicken Image" and "Genetics in a Cell Line," and by Japan Society for the Promotion of Science Postdoctoral Fellowships for Research Abroad.

22 October 2001; accepted 18 December 2001

Capturing Chromosome Conformation

Job Dekker,^{1*} Karsten Rippe,² Martijn Dekker,³ Nancy Kleckner¹

We describe an approach to detect the frequency of interaction between any two genomic loci. Generation of a matrix of interaction frequencies between sites on the same or different chromosomes reveals their relative spatial disposition and provides information about the physical properties of the chromatin fiber. This methodology can be applied to the spatial organization of entire genomes in organisms from bacteria to human. Using the yeast *Saccharomyces cerevisiae*, we could confirm known qualitative features of chromosome organization within the nucleus and dynamic changes in that organization during meiosis. We also analyzed yeast chromosome III at the G₁ stage of the cell cycle. We found that chromatin is highly flexible throughout. Furthermore, functionally distinct AT- and GC-rich domains were found to exhibit different conformations, and a population-average 3D model of chromosome III could be determined. Chromosome III emerges as a contorted ring.

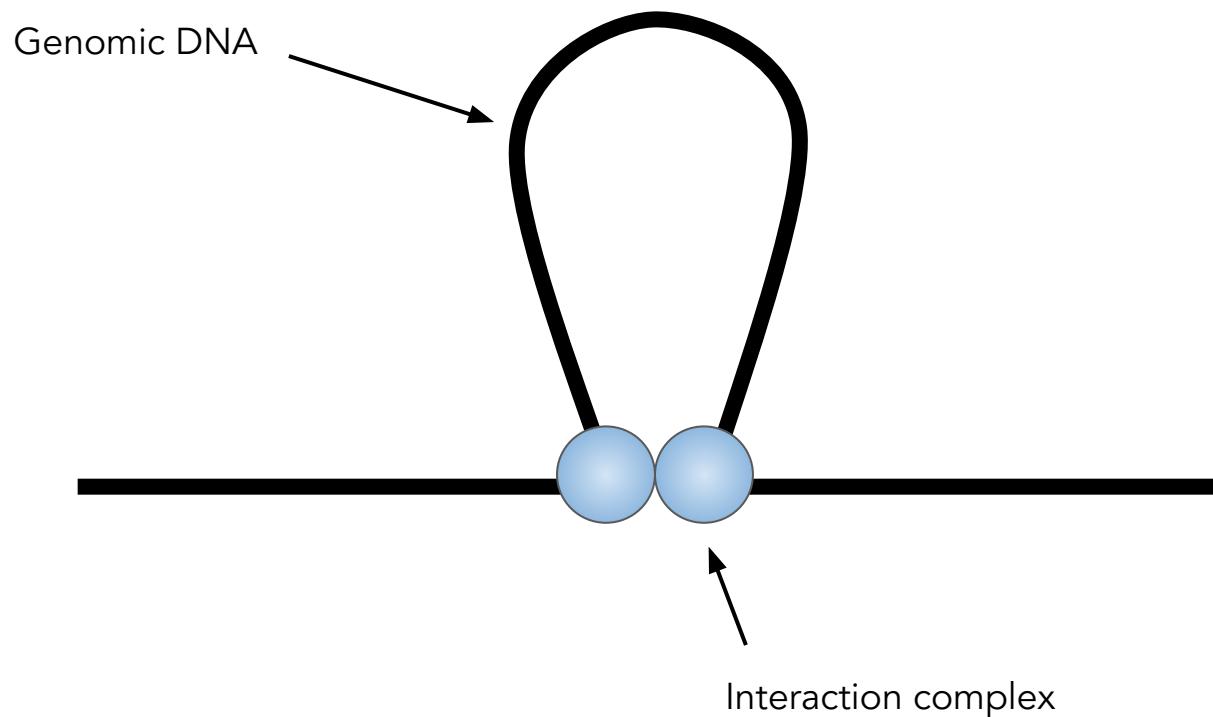
Important chromosomal activities have been linked with both structural properties and spatial conformations of chromosomes. Local properties of the chromatin fiber influence gene expression, origin firing, and DNA re-

affords a resolution of 100 to 200 nm at best, which is insufficient to define chromosome conformation. DNA binding proteins fused to green fluorescent protein permit visualization of individual loci, but only a few positions

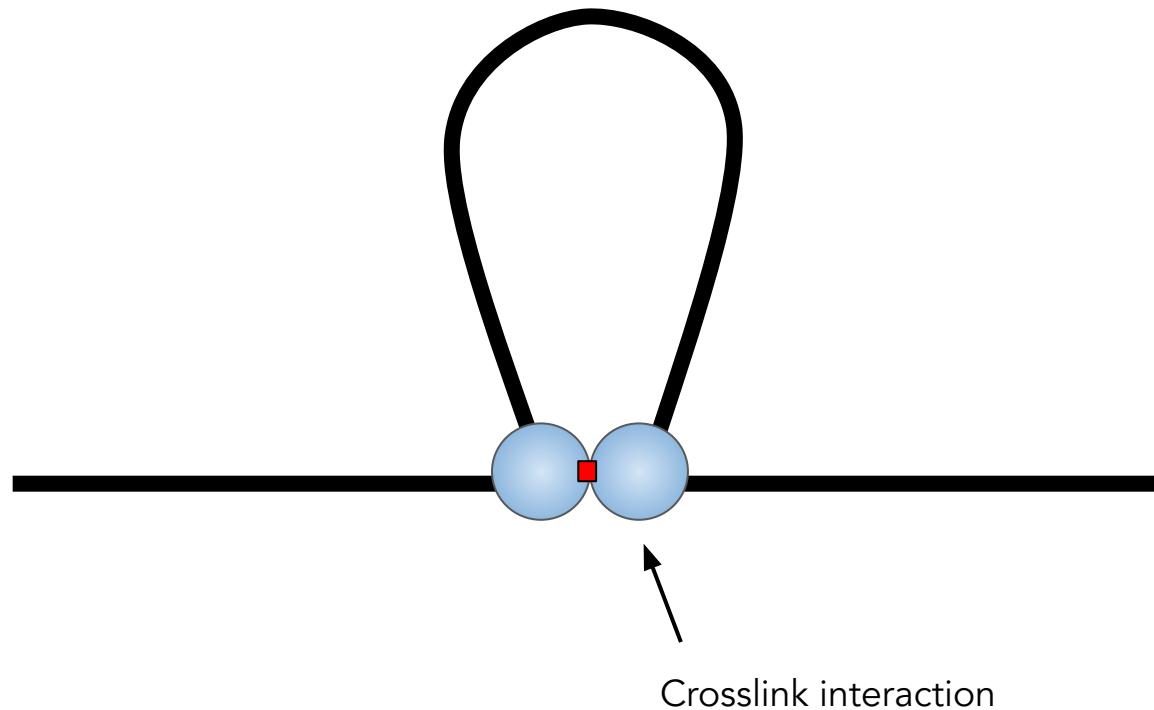
of purified nuclei is largely intact, as shown below.

For quantification of cross-linking frequencies, cross-linked DNA is digested with a restriction enzyme and then subjected to ligation at very low DNA concentration. Under such conditions, ligation of cross-linked fragments, which is intramolecular, is strongly favored over ligation of random fragments, which is intermolecular. Cross-linking is then reversed and individual ligation products are detected and quantified by the polymerase chain reaction (PCR) using locus-specific primers. Control template is generated in which all possible ligation products are present in equal abundance (7). The cross-linking frequency (X) of two specific loci is determined by quantitative PCR reactions using control and cross-linked templates, and X is expressed as the ratio of the amount of product obtained using the cross-linked template to the amount of product obtained with the control template (Fig. 1B). X should be directly proportional to the frequency with which the two corresponding genomic sites interact (10).

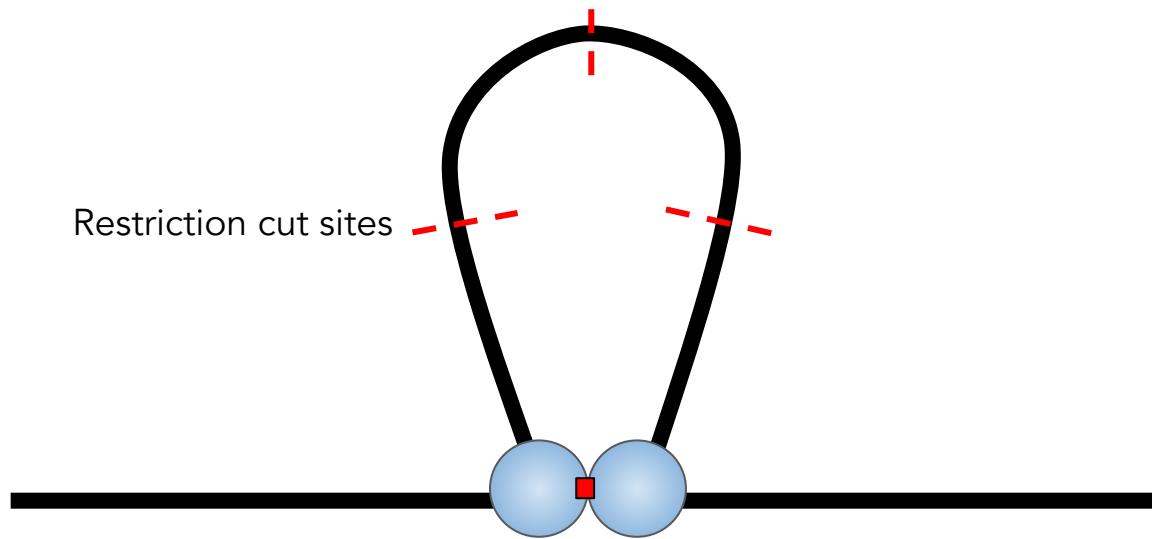
Chromosome conformation capture (3C)



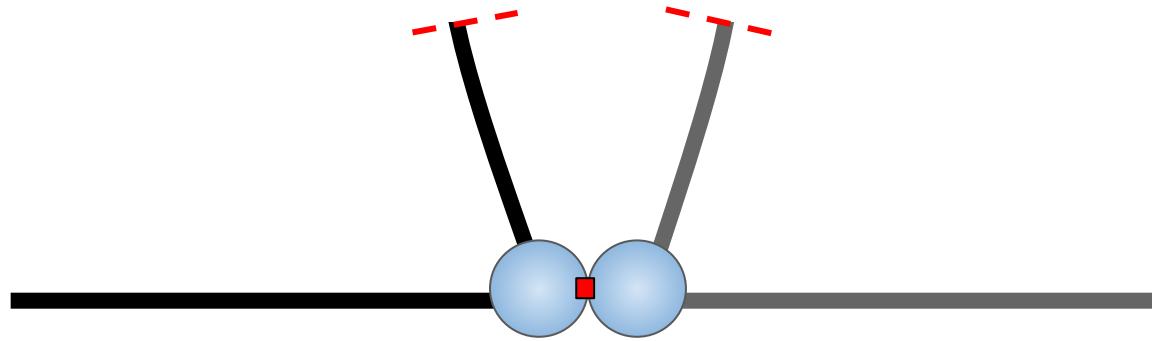
Chromosome conformation capture (3C)



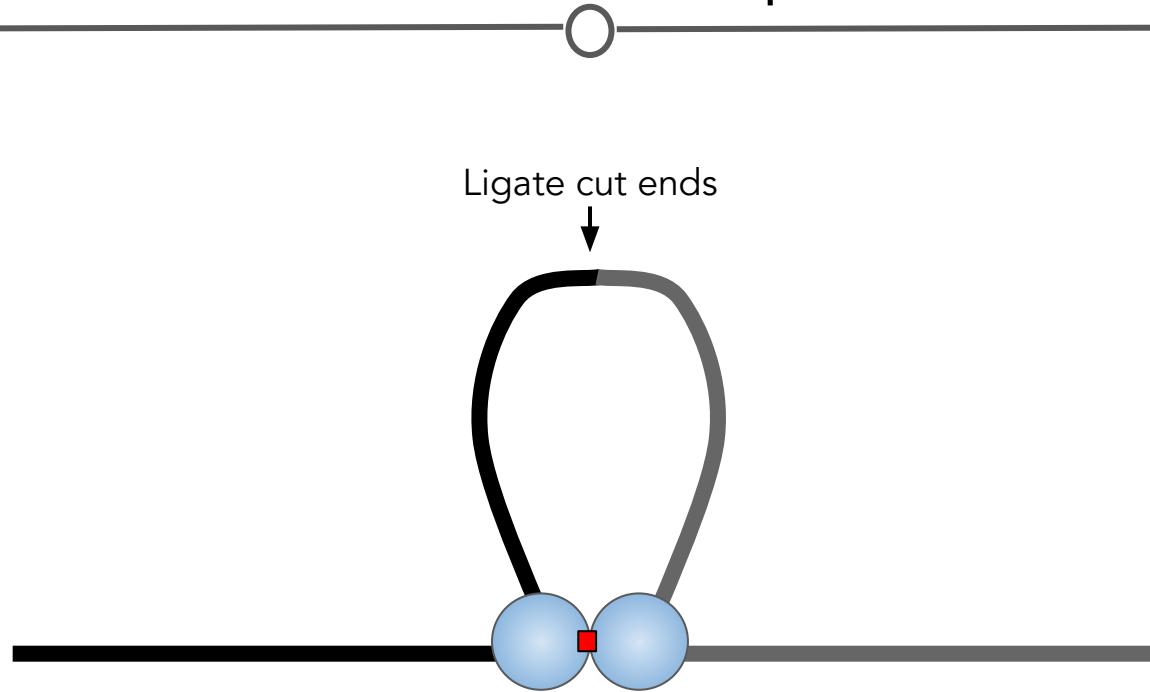
Chromosome conformation capture (3C)



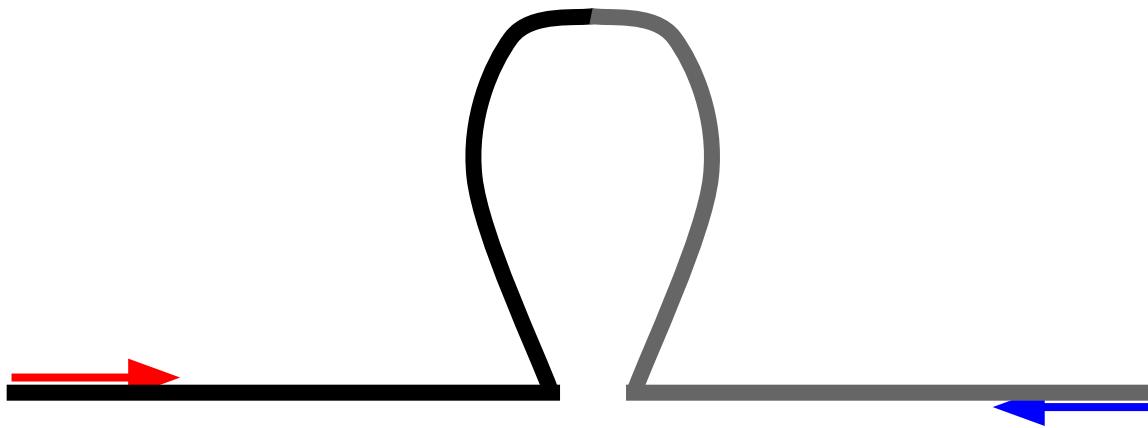
Chromosome conformation capture (3C)



Chromosome conformation capture (3C)



Chromosome conformation capture (3C)



Use specific primers to amplify hybrid fragments of interest.

Survey numerous 3C products to establish interaction profile across distance

Hi-C, mmm tasty

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

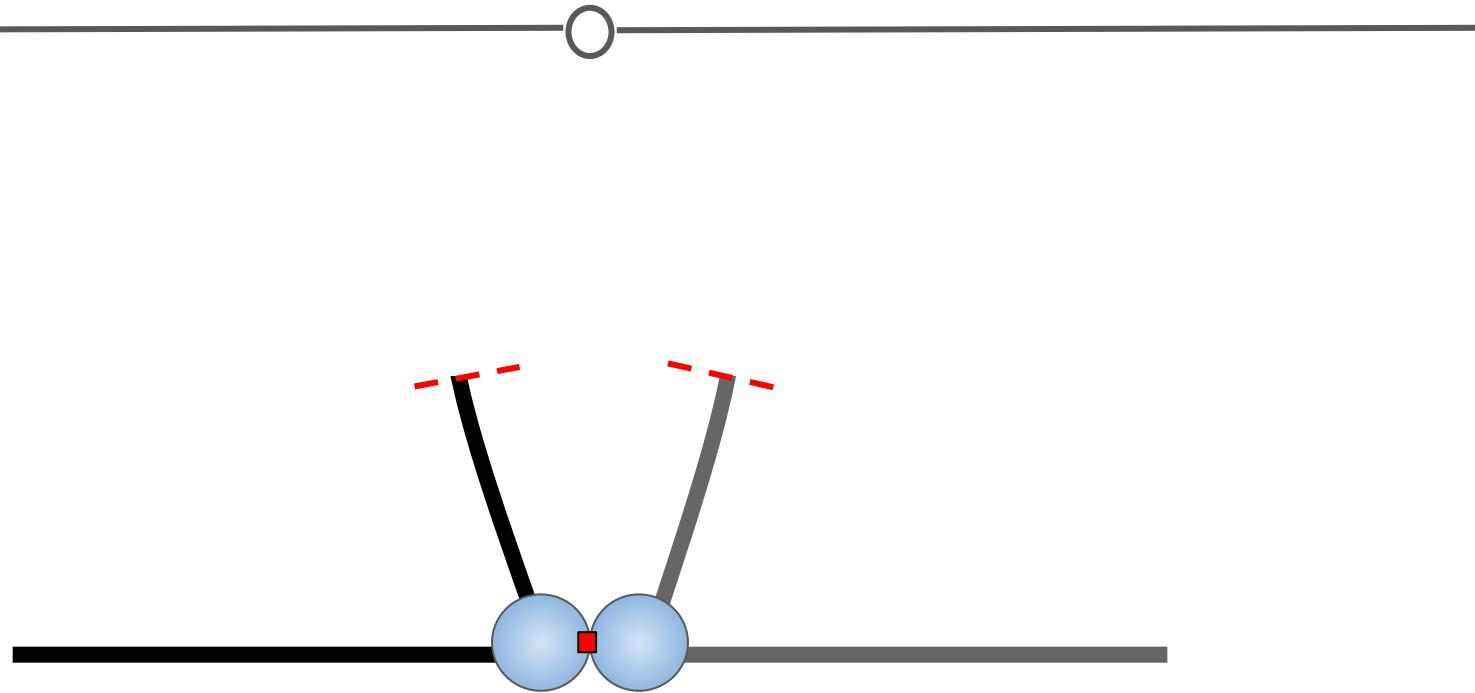
We created a Hi-C library from a karyotypically normal human lymphoblastoid cell line (GM06990) and sequenced it on two lanes of an Illumina Genome Analyzer (Illumina, San Diego, CA), generating 8.4 million read pairs that could be uniquely aligned to the human genome reference sequence; of these, 6.7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

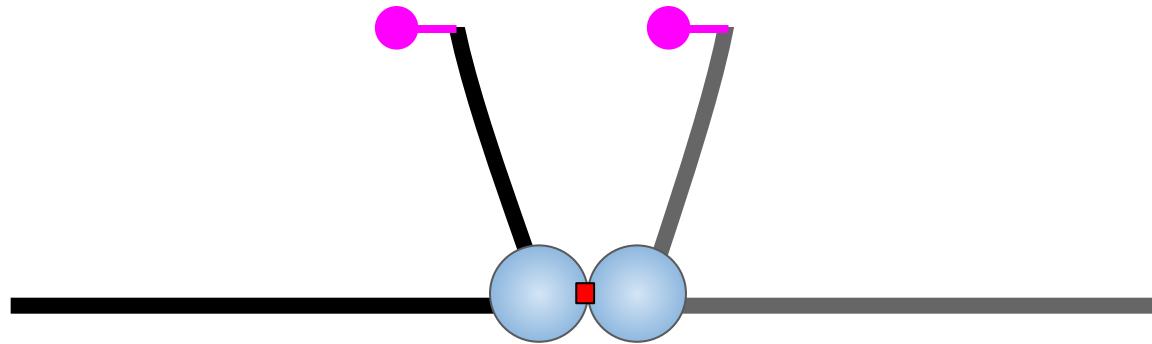
Hi-C



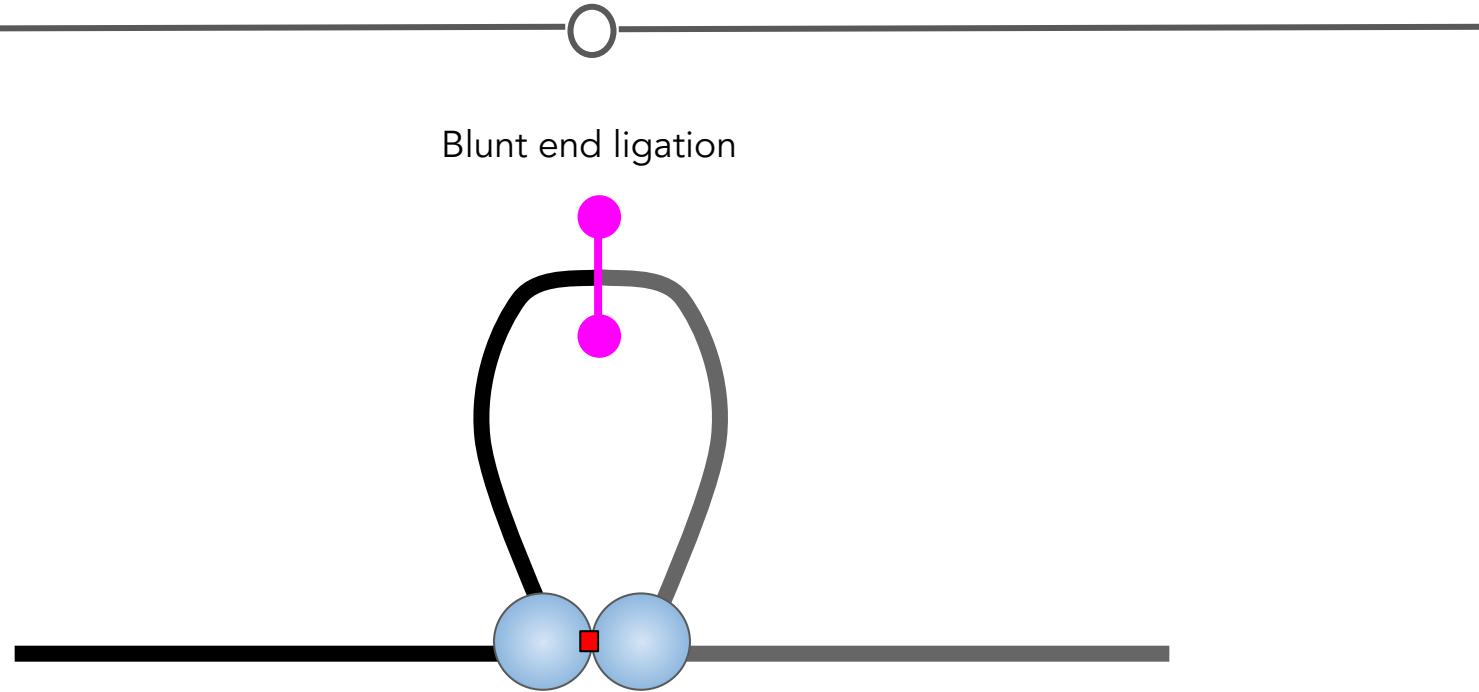
Hi-C



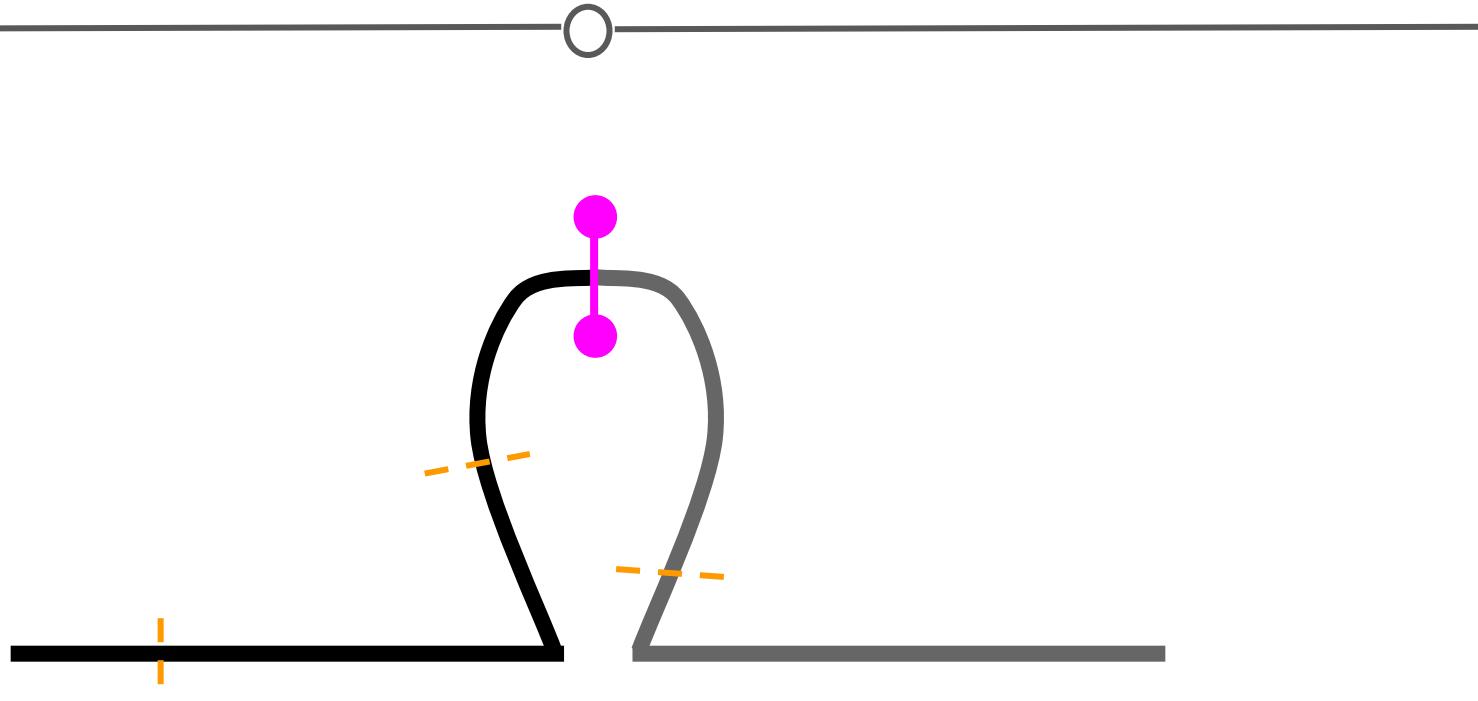
Fill in with biotinylated nucleotides



Hi-C

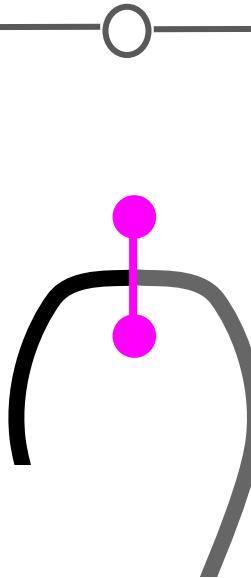


Hi-C



Purify DNA and shear randomly

Hi-C

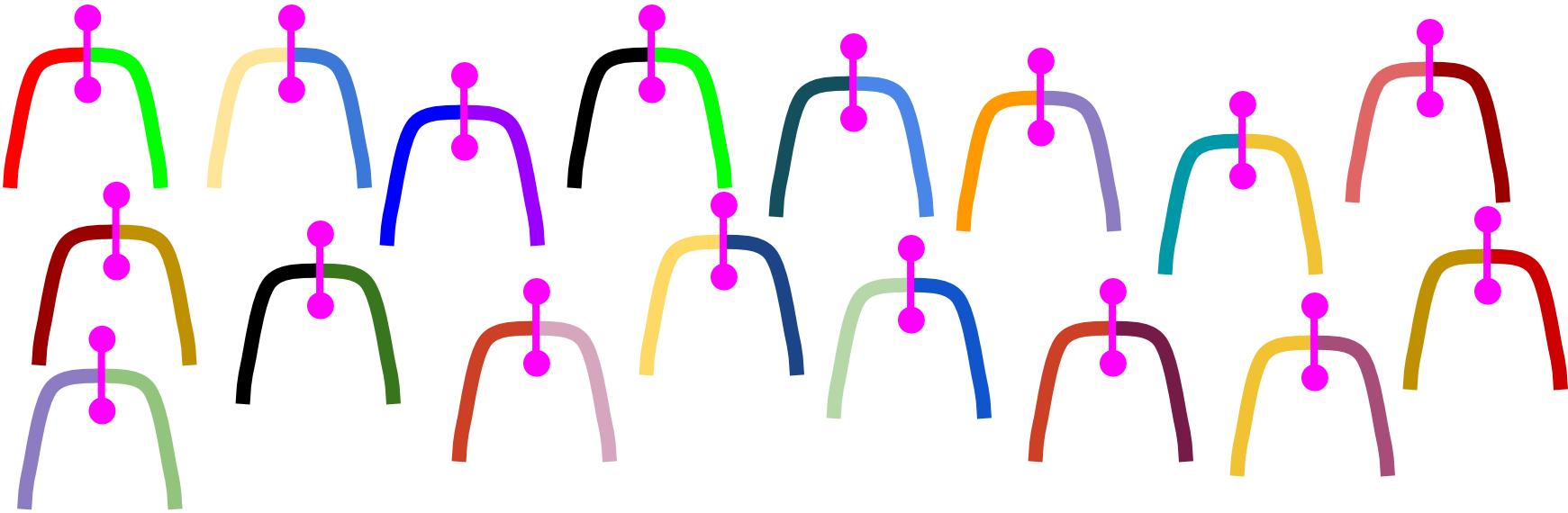


Select out biotinylated fragments and sequence from both ends

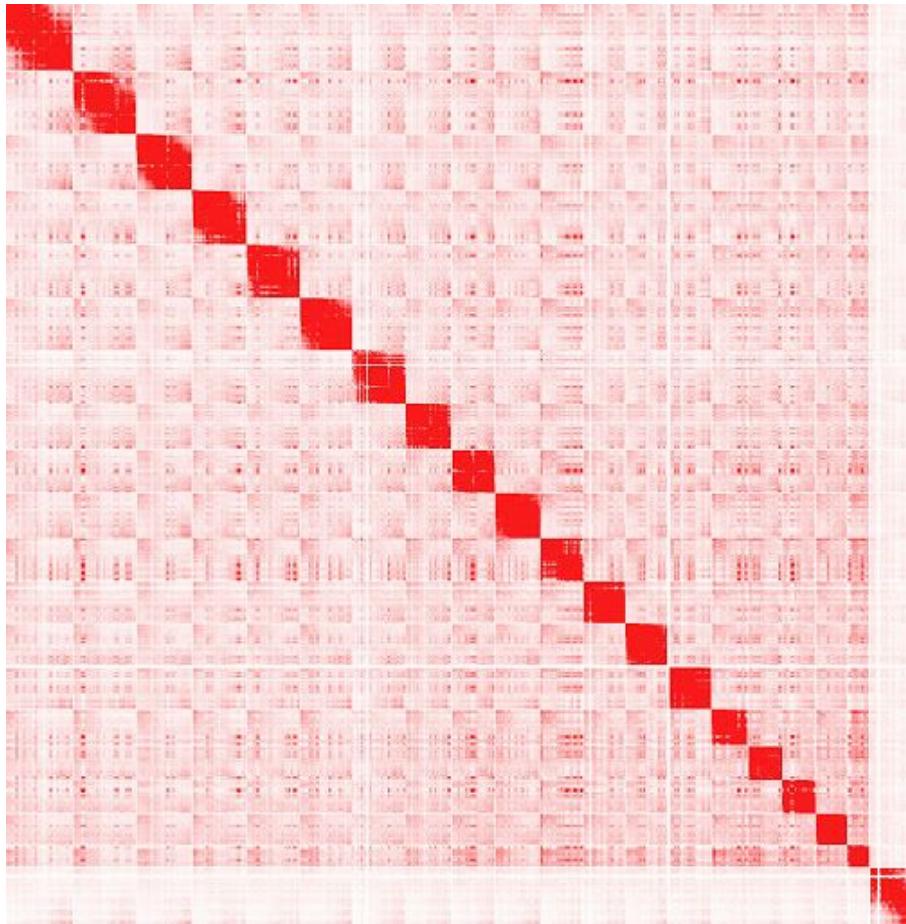
Hi-C



Sequence millions of fragments simultaneously

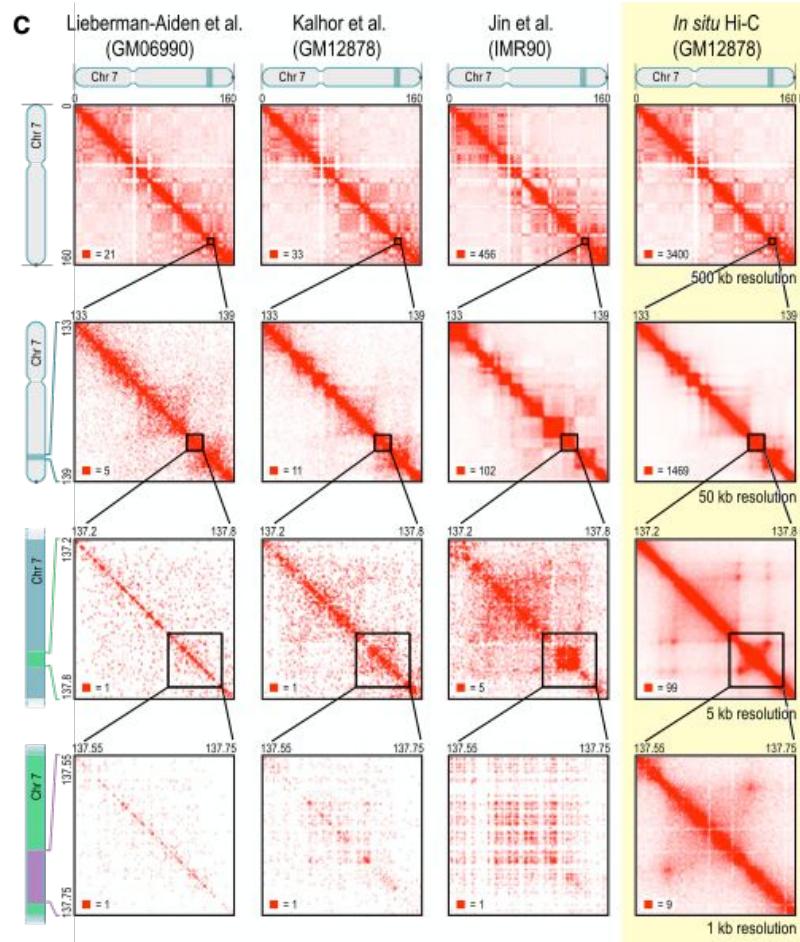


Hi-C Data

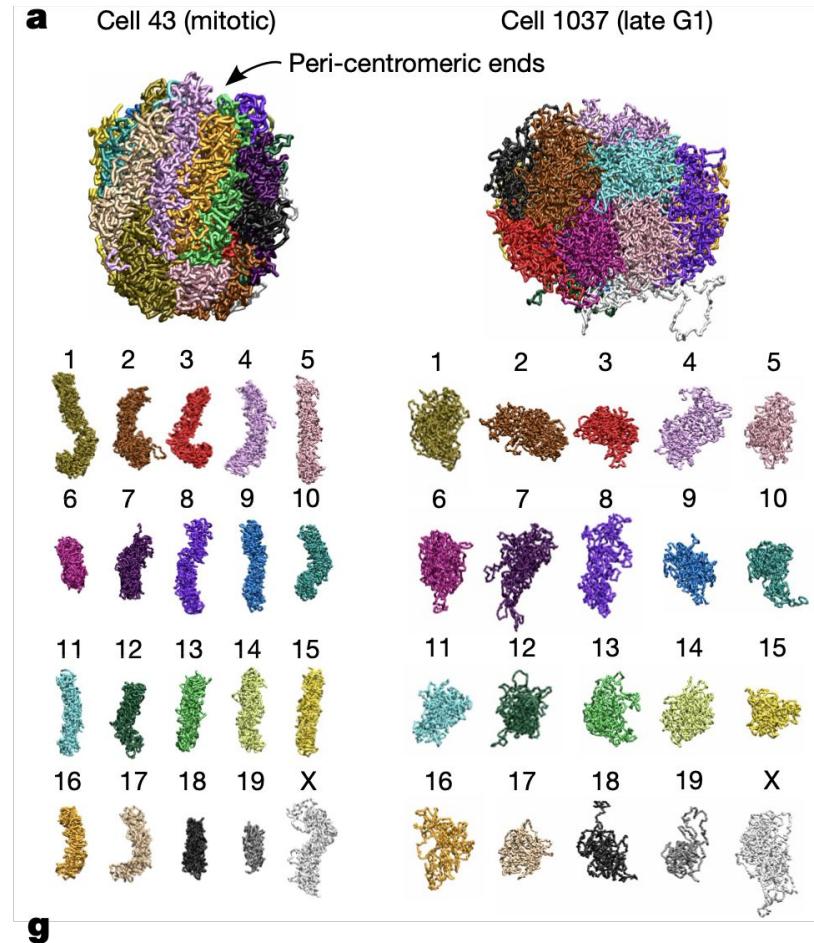
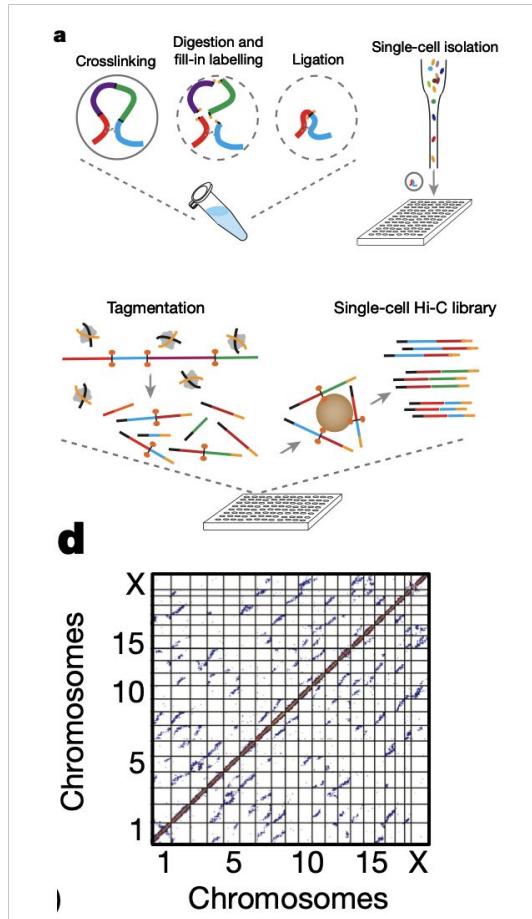


<http://homer.ucsd.edu/homer/interactions/HiCmatrices.html>

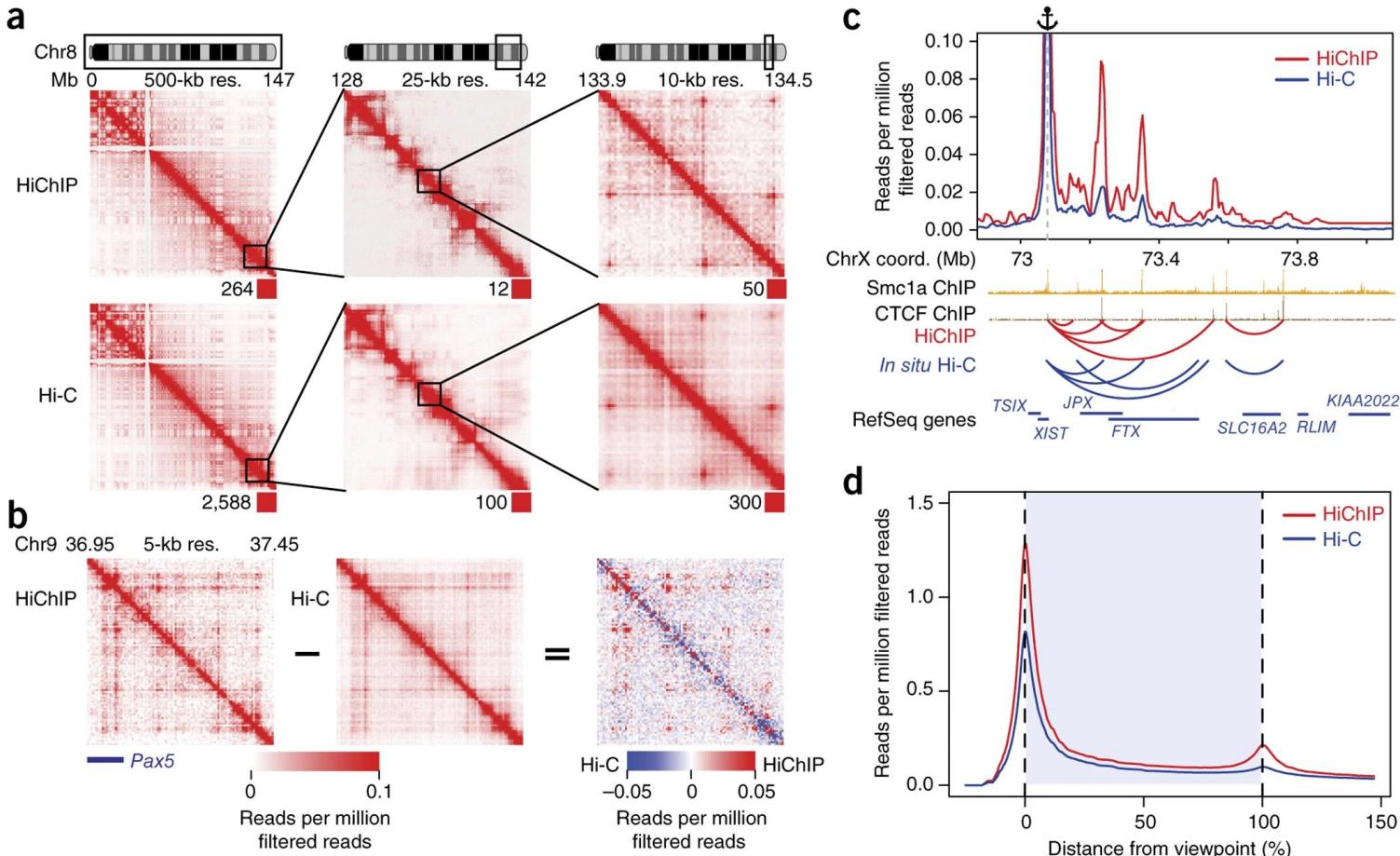
Improving resolution



Single cell Hi-C



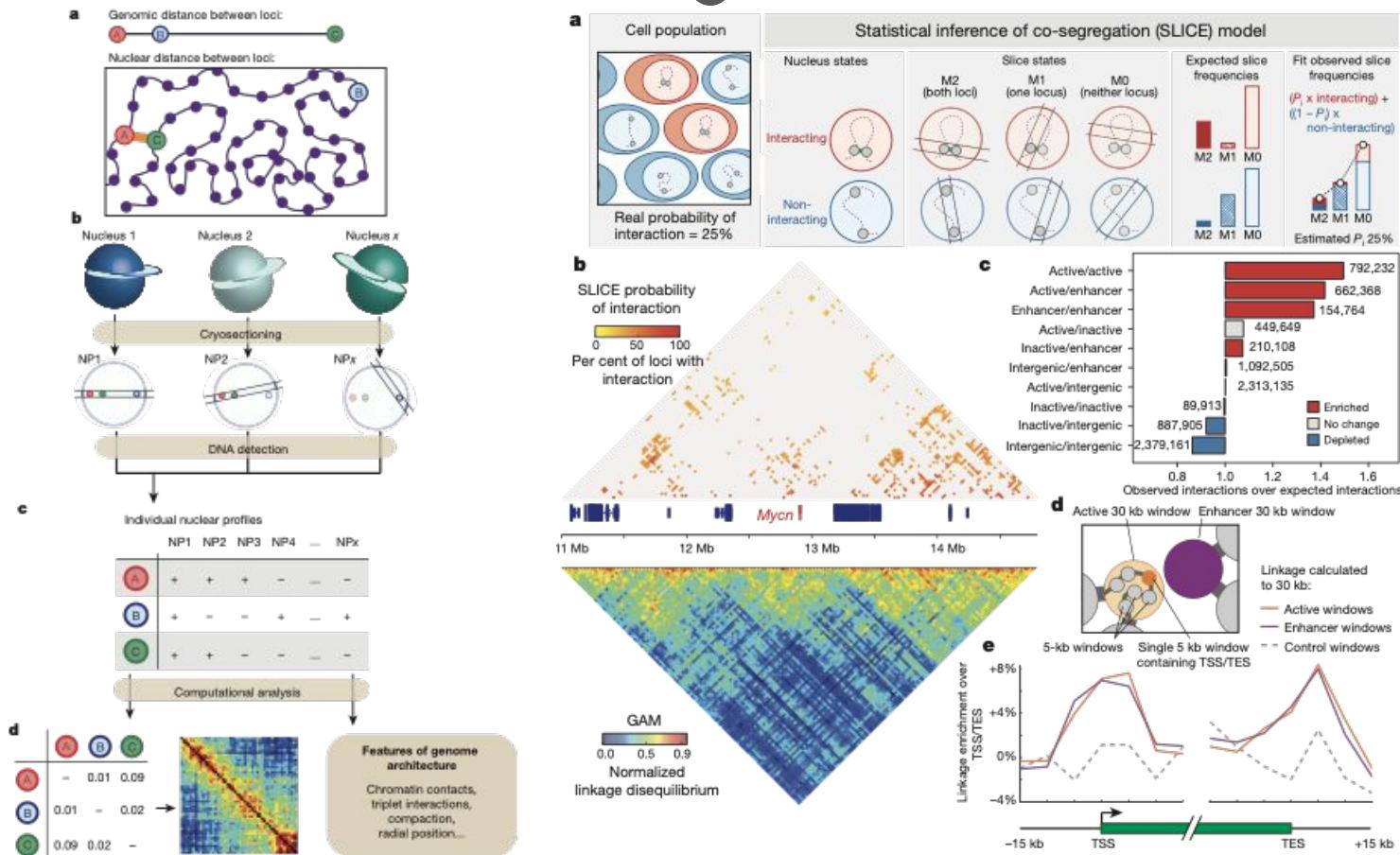
HiChIP



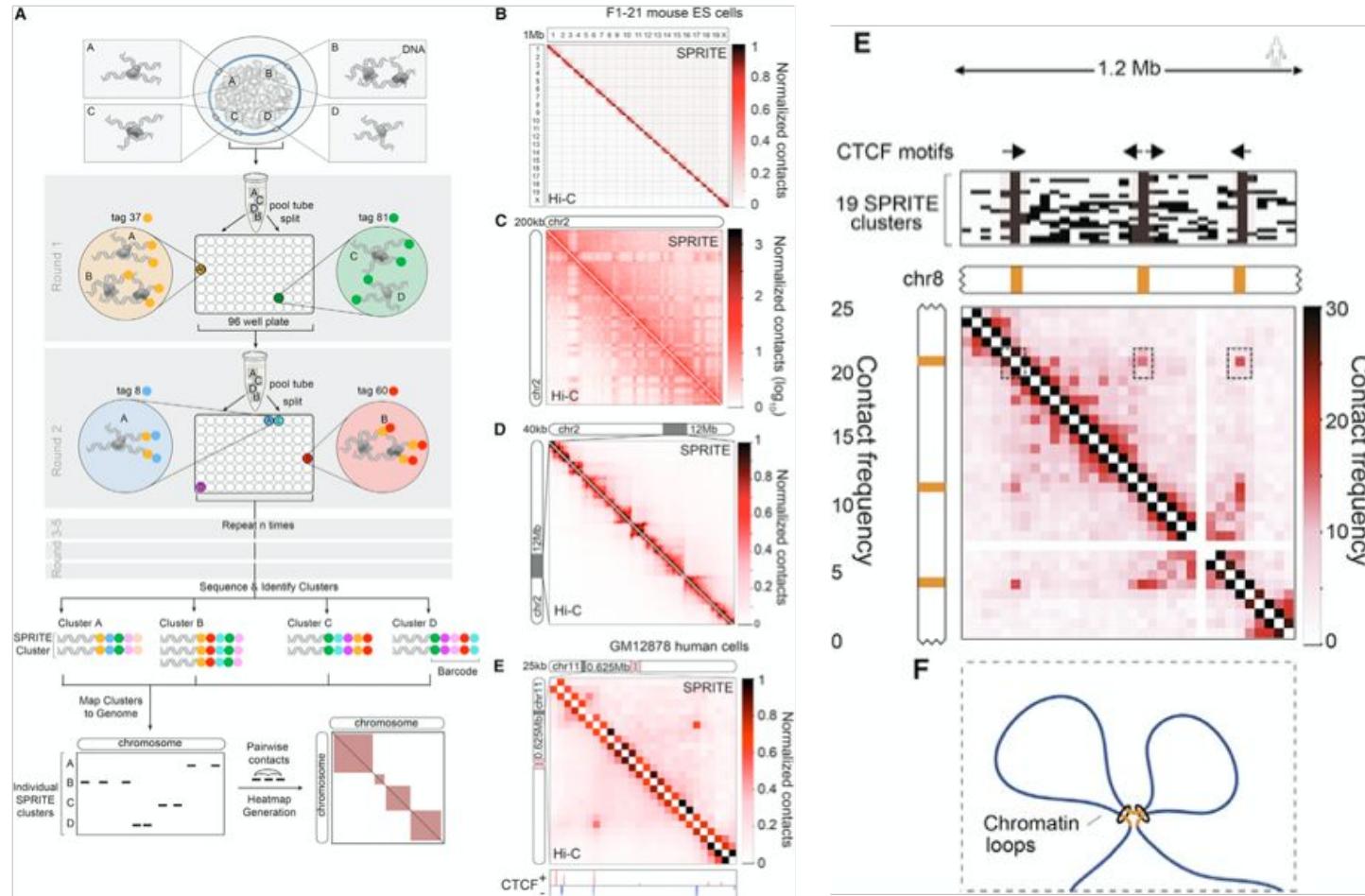
Non-Hi-C approaches



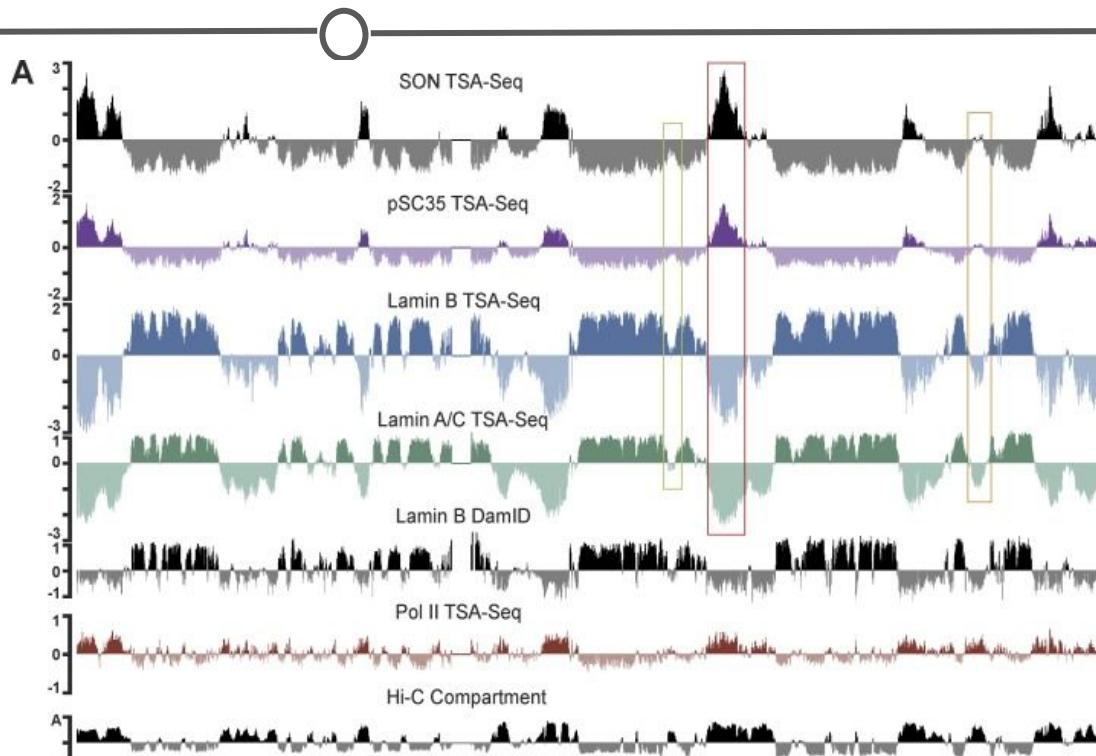
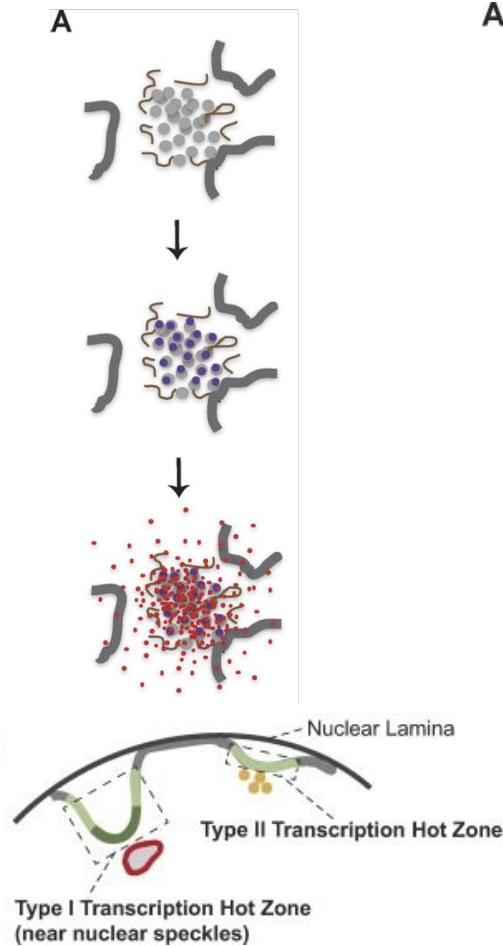
Genome architecture mapping (GAM)



Split-Pool Recognition of Interactions by Tag Extension (SPRITE)



TSA-seq



- Horseradish peroxidase generates tyramide free-radicals
- Using biotin-tyramide, chromatin can be labeled and enriched

Hi-C use cases

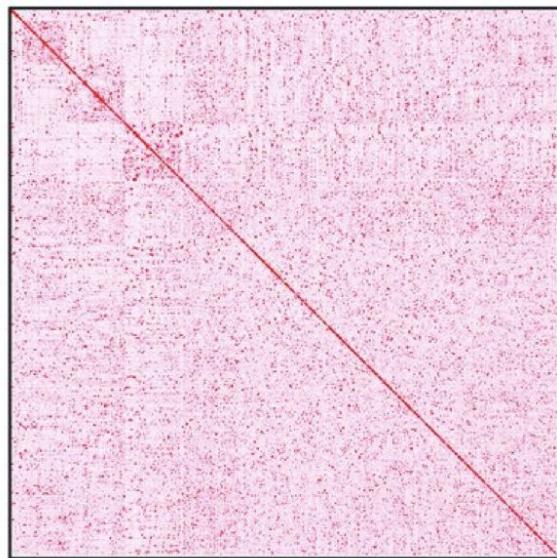


Genome assembly

Draft assembly

Lnk 1 Lnk 2 Lnk 3 U

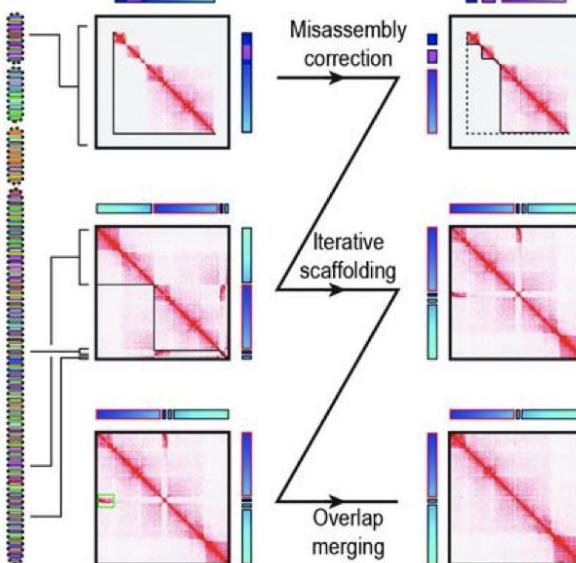
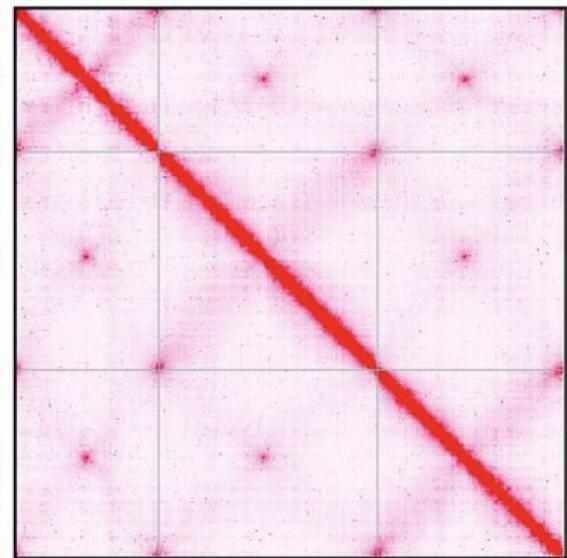
Hi-C contact matrix



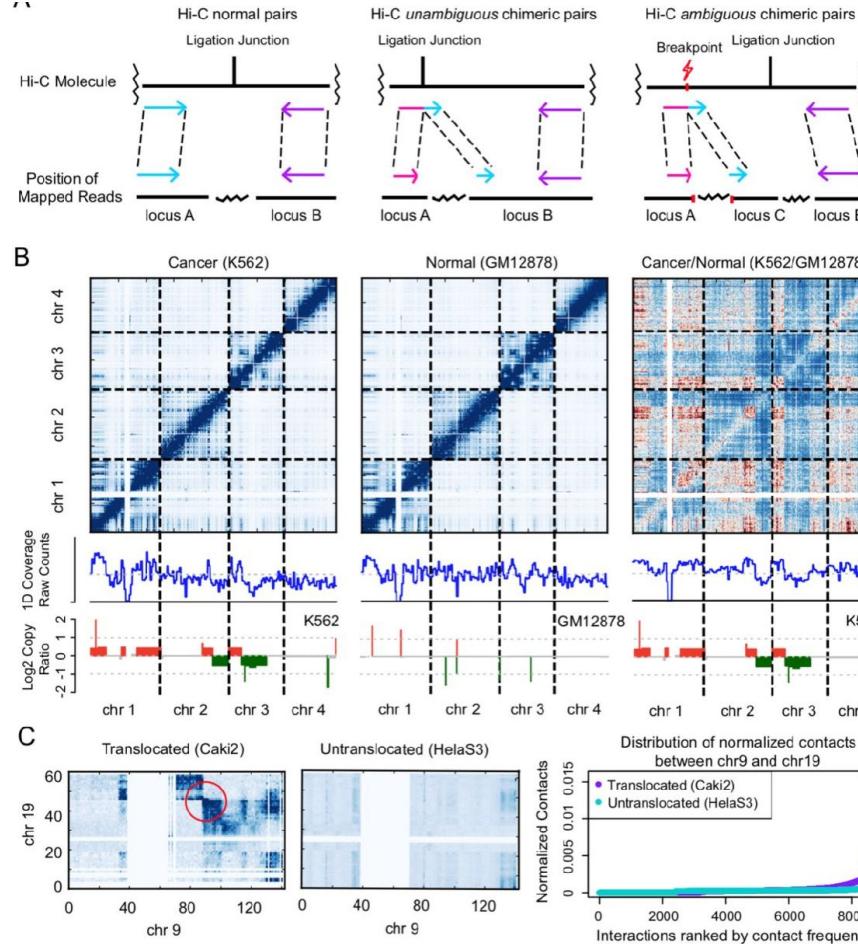
End-to-end assembly

chr 1 chr 2 chr 3

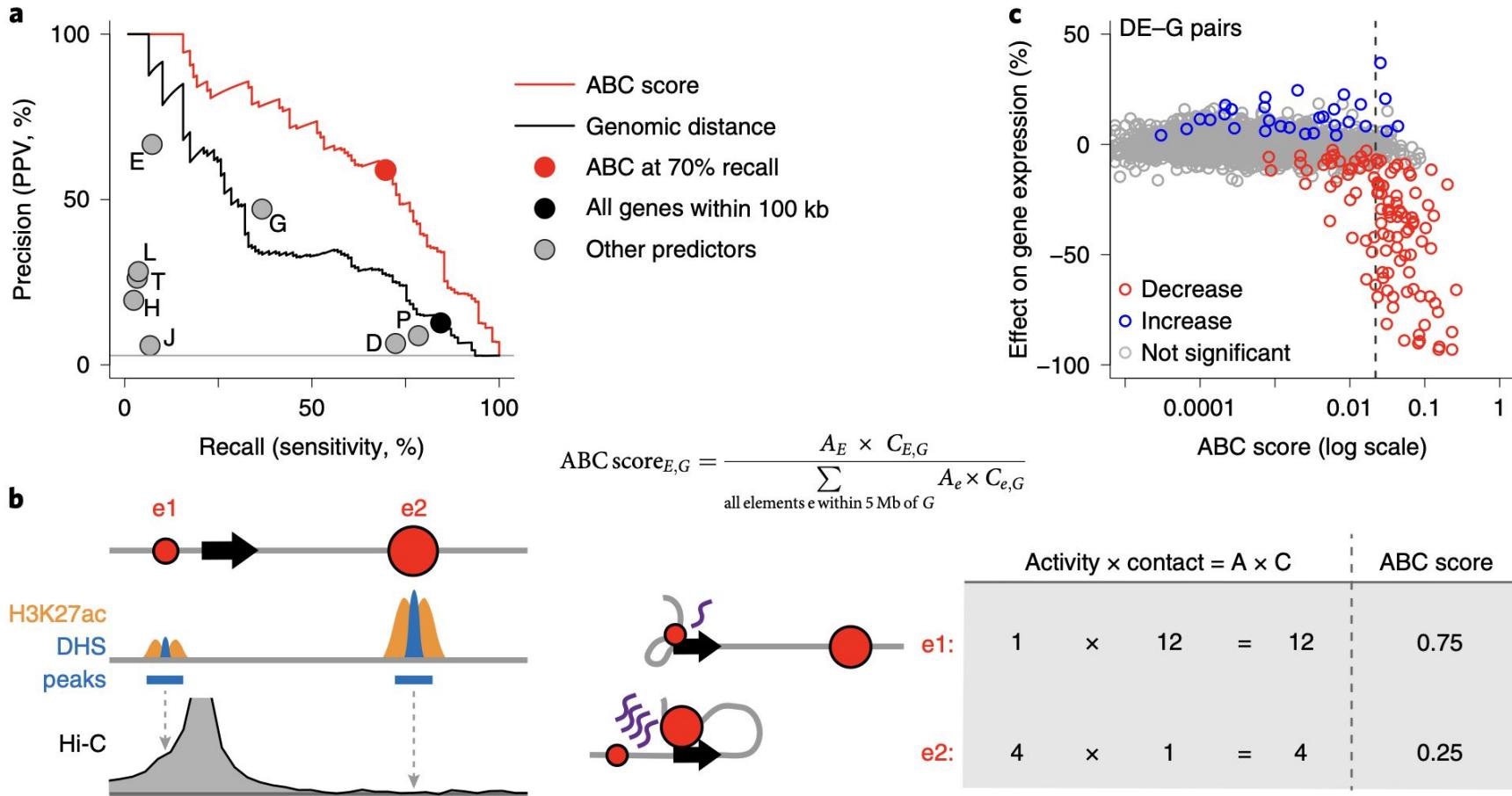
1p 1q 2p 2q 3p 3q



Structural variation detection



Gene regulation



Hi-C Resources



- Genome browsers
 - HiGlass (<https://higlass.io>)
 - HiC-Browser
(<http://promoter.bx.psu.edu/hi-c/index.html>)
- Analysis tools
 - HiFive (<https://github.com/bxlab/hifive>)
 - HiCExplorer
(<https://galaxyproject.github.io/training-material/topics/epigenetics/tutorials/hicexplorer/tutorial.html>)