

Phase transitions in the nucleus of cells

Theories from physics and new experimental data are revealing the mechanisms that control structure and function of the human genome.

Mario Nicodemi, Dip.to di Fisica, Uni.NA "*Federico II*", INFN, IT and M. Delbruck Center, Berlin, DE

Team: Andrea Chiariello, Simona Bianco, Andrea Esposito, Mattia Conte, Alex Abraham, Francesca Vercellone, ...

The end of the beginning



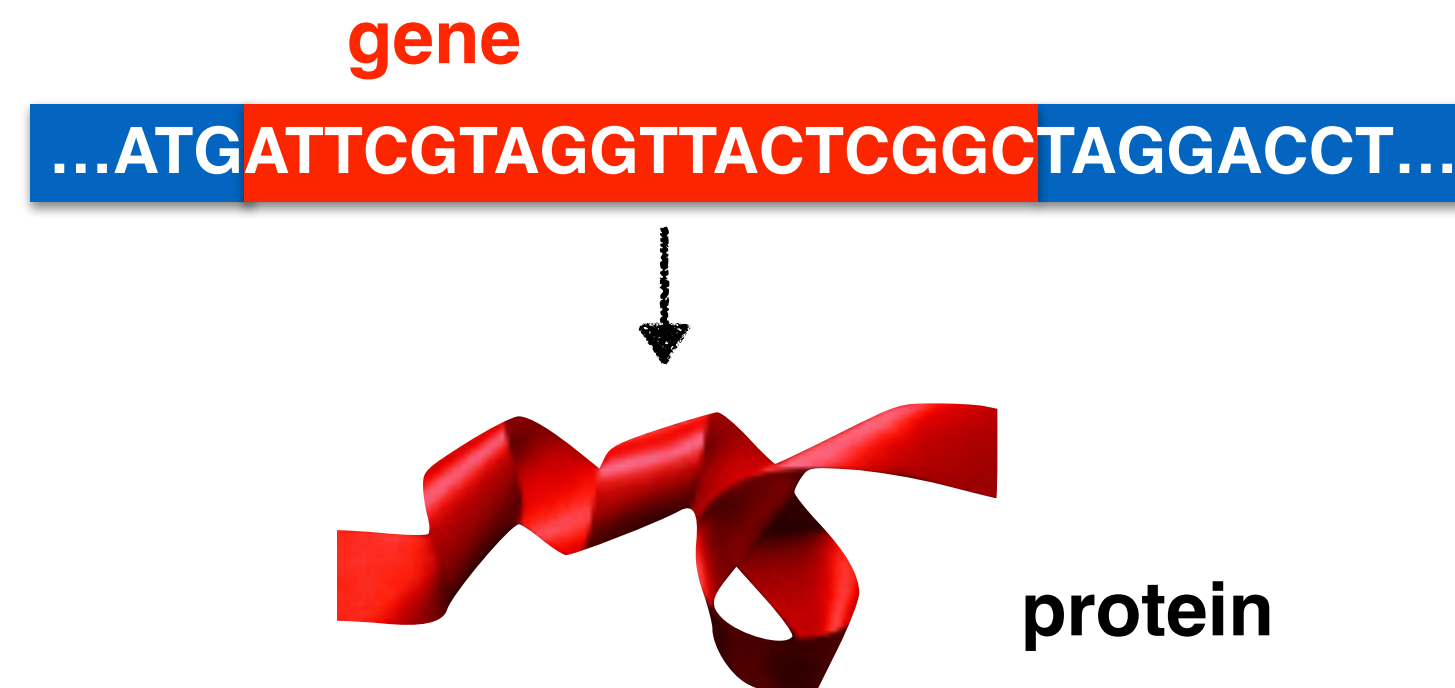
Our genome is sequenced:

... ATGTTAGACGT ...

... but how is it regulated in health and disease?

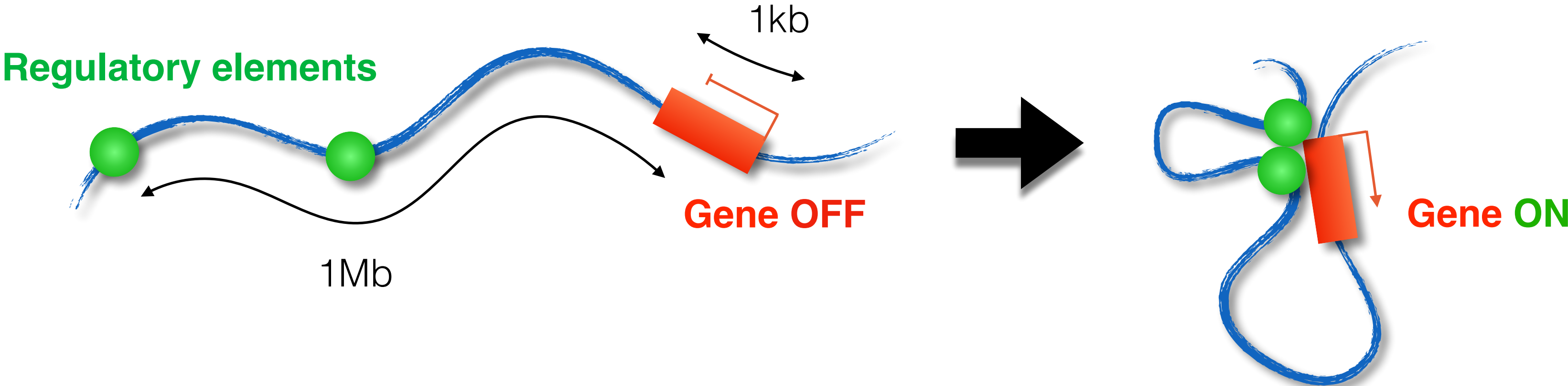
Human DNA

Our **DNA** has 3 billion bases (A,T,C,G)



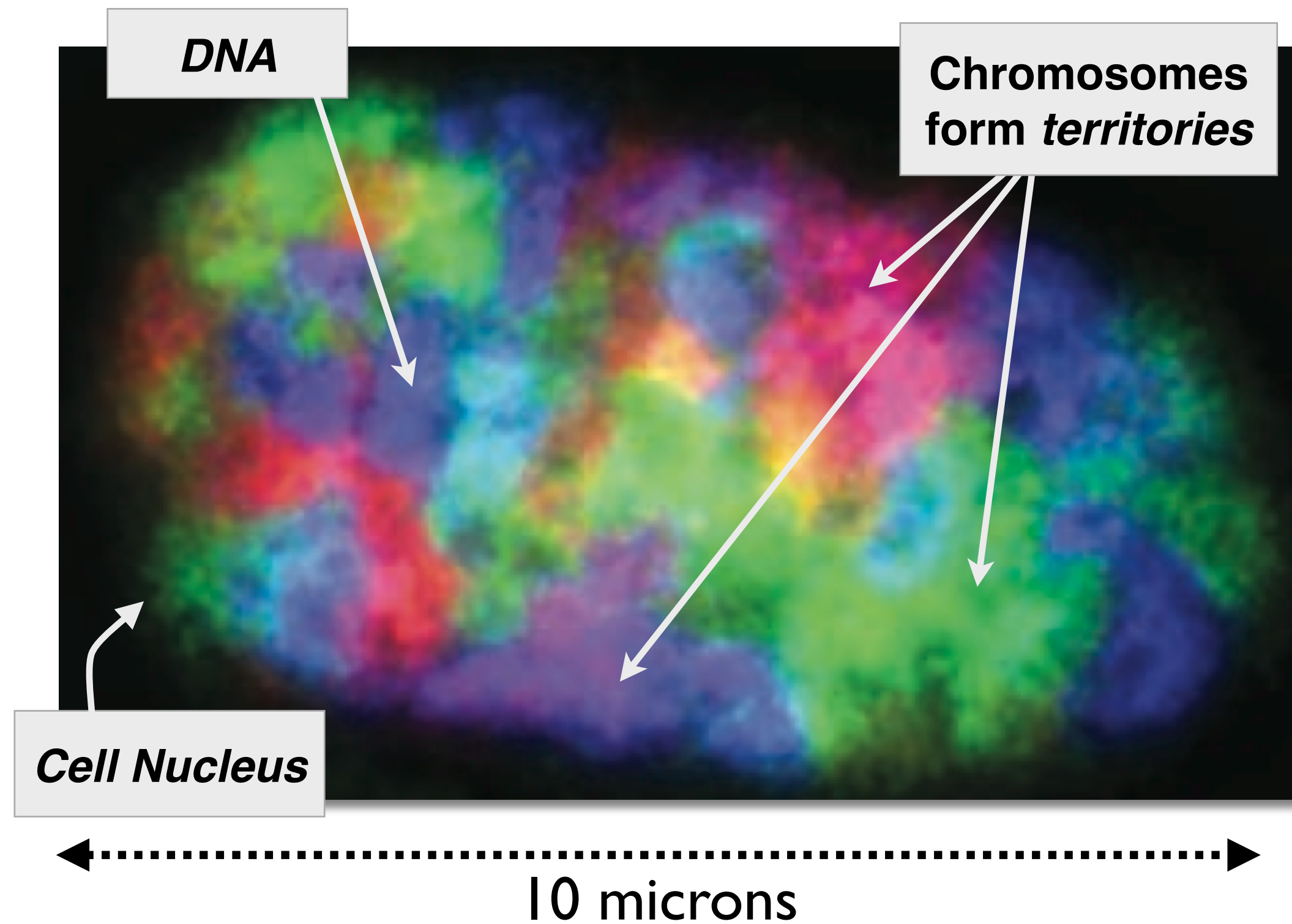
Not just a linear sequence

Non-coding DNA hides the key to the regulation of our 20000 genes



A non random organisation

Chromosomes form territories in the cell nucleus (Cremer&Cremer 1990's)

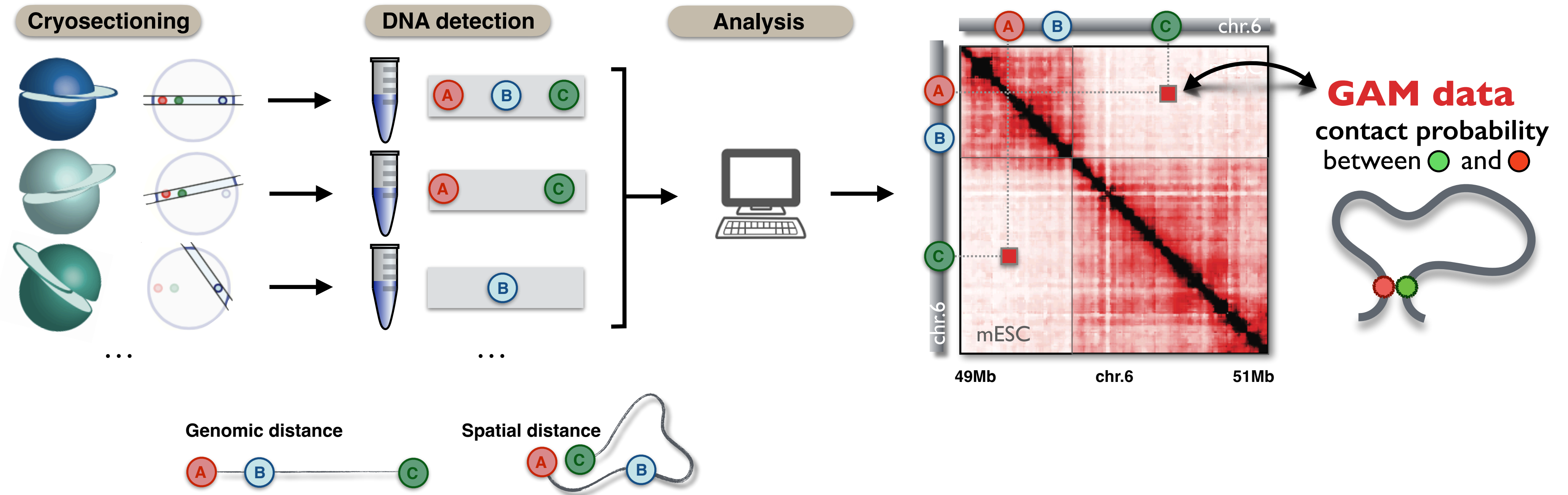


Microscopy image (FISH)



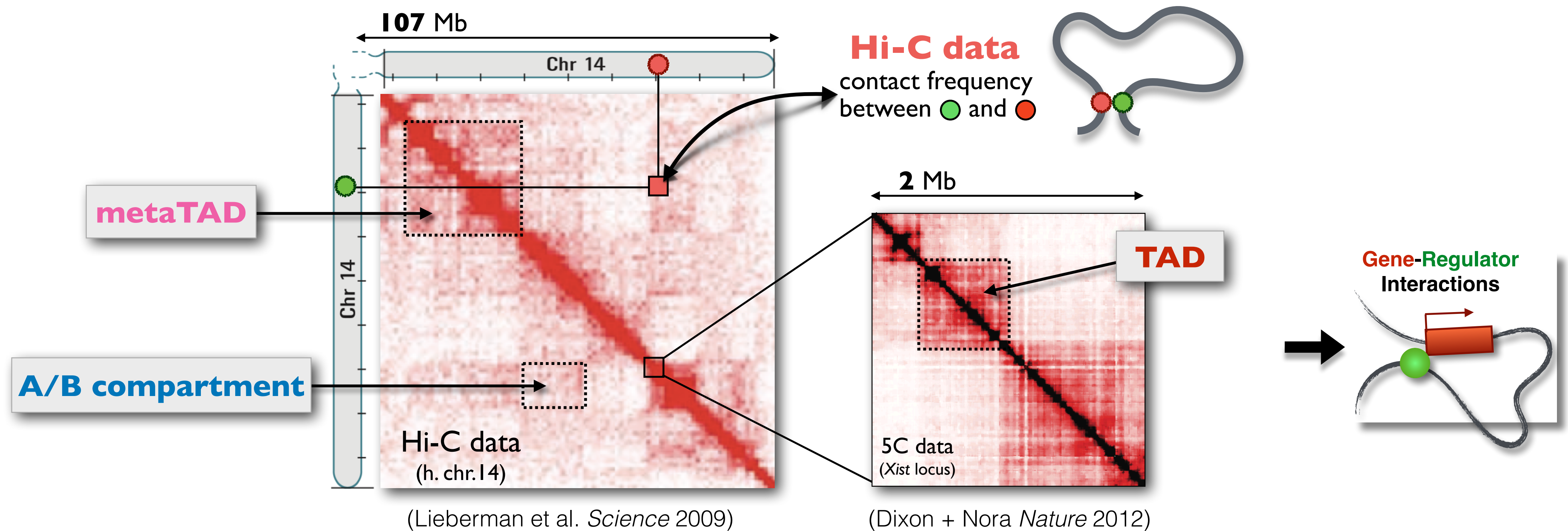
Our GAM technology

GAM probes 3D proximity by sequencing DNA from nuclear sections: spatially closer sites co-segregate more frequently.



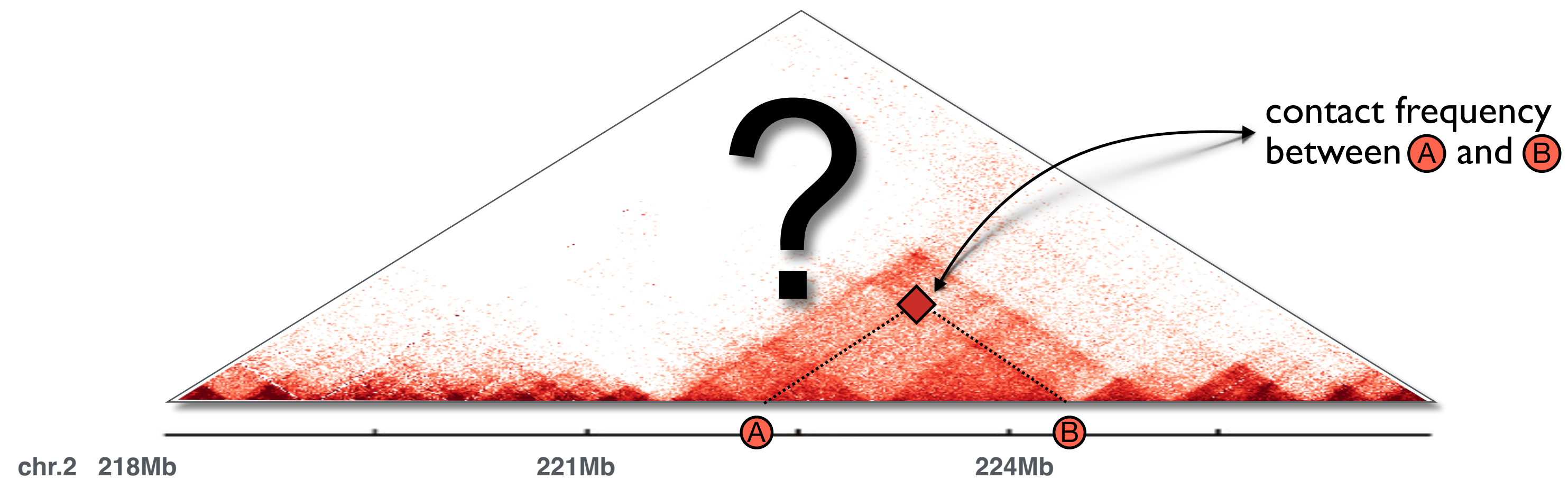
Not just a linear sequence

New quantitative technologies revealed that chromosomes have complex 3D structures.



Chromosomes are divided in 0.5-1Mb long **TADs** (Dixon; Nora 2012) and in ~10Mb **A/B compartments** (Lieberman-A. 2009). Patterns exist across chromosomal scales (Sexton 2012, Phillips-C. 2013, ...) hierarchically arranged in **metaTADs** (Fraser, Chiariello 2014, ...).

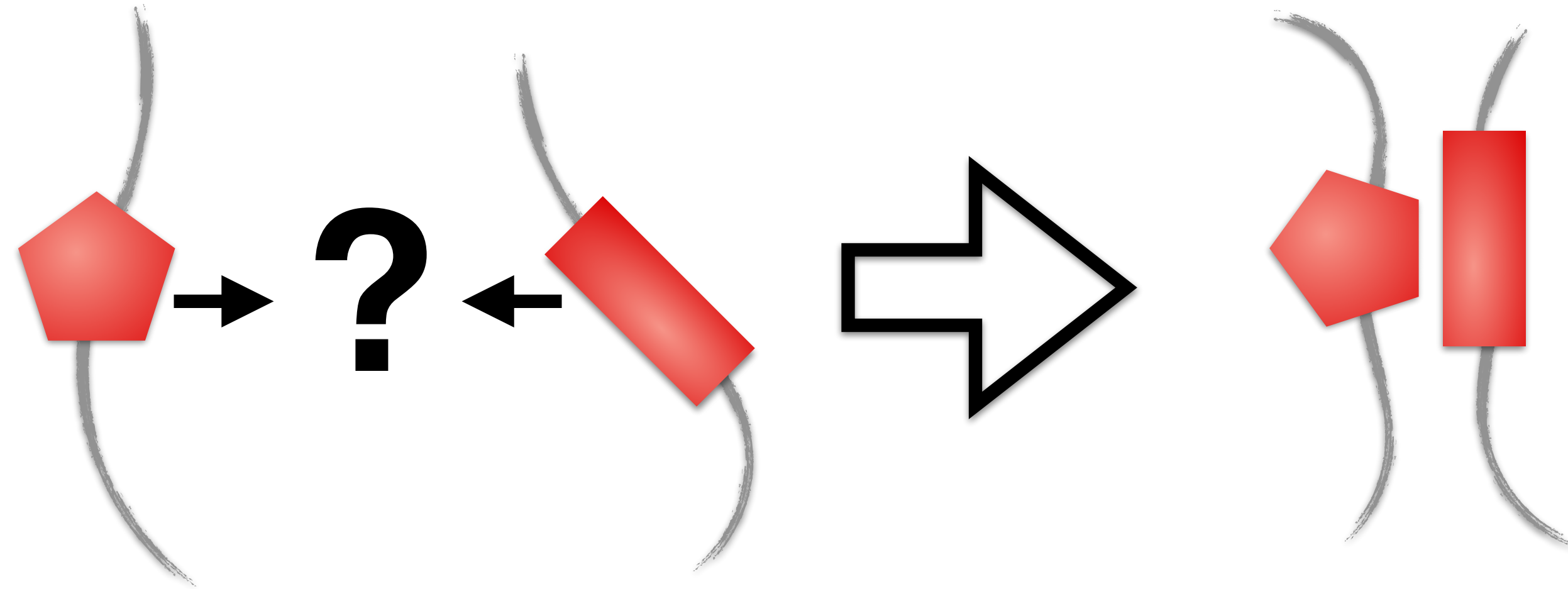
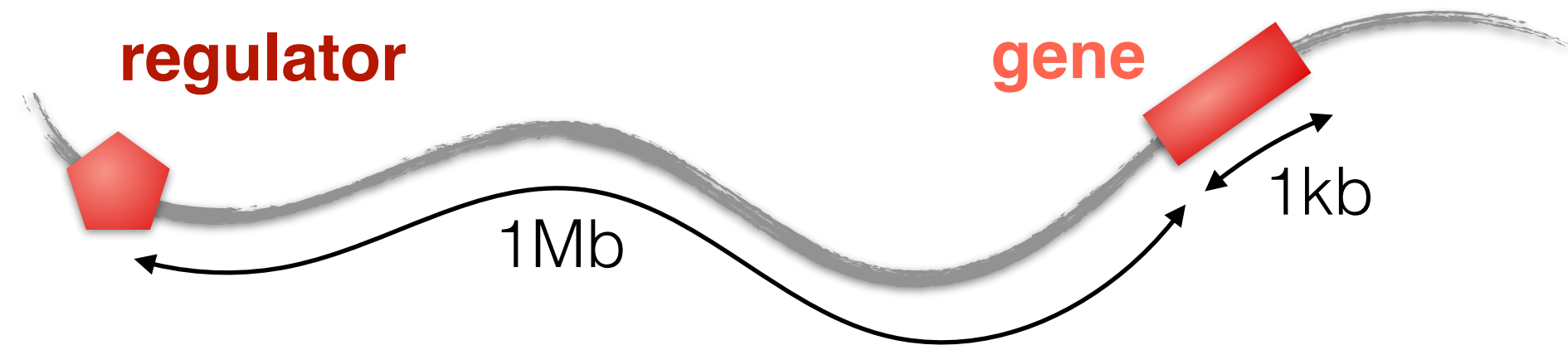
Origin of contact patterns



Principled approaches from physics can help identifying the origin of contact patterns and their molecular determinants.

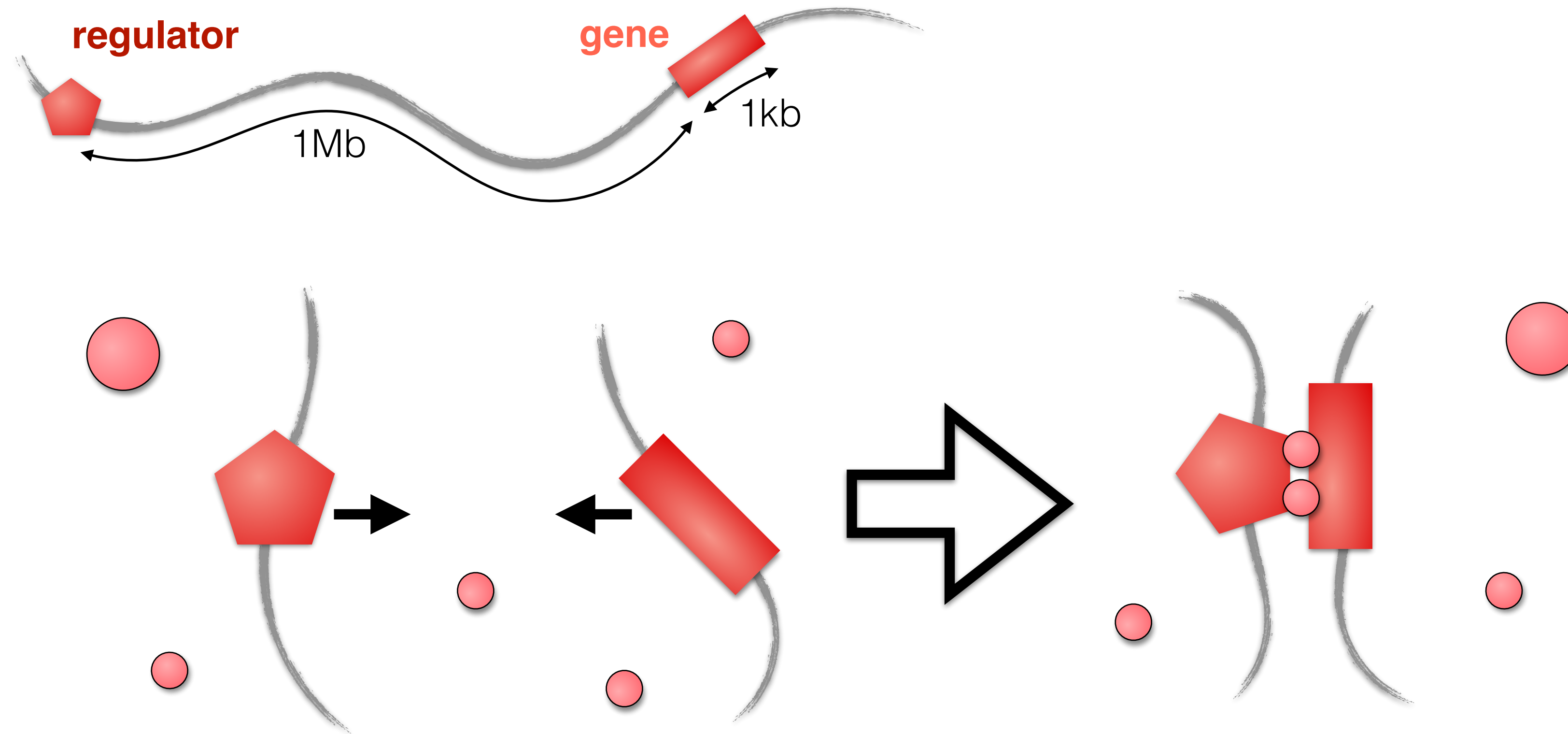
Interaction mechanism

How can genes and their distal regulators find each other?



Interaction mechanism

How can genes and their distal regulators find each other?

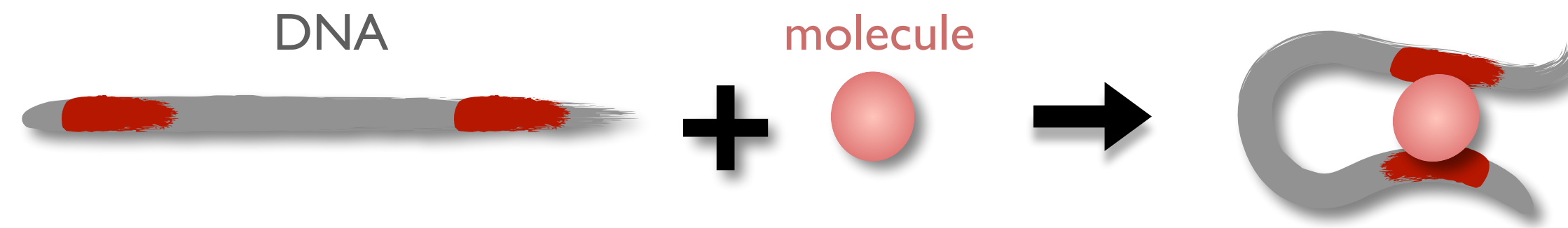


... a “**particle**” produces the interaction.

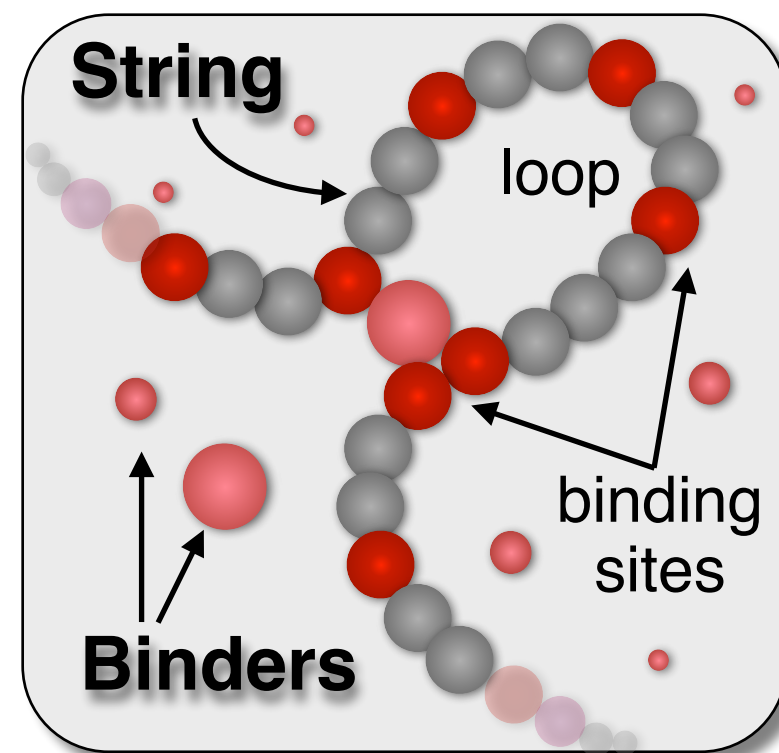
The Strings&Binders (SBS) model

Stable conformations correspond to the system thermodynamic phases.

- Scenario:






- A model of polymers & particles:

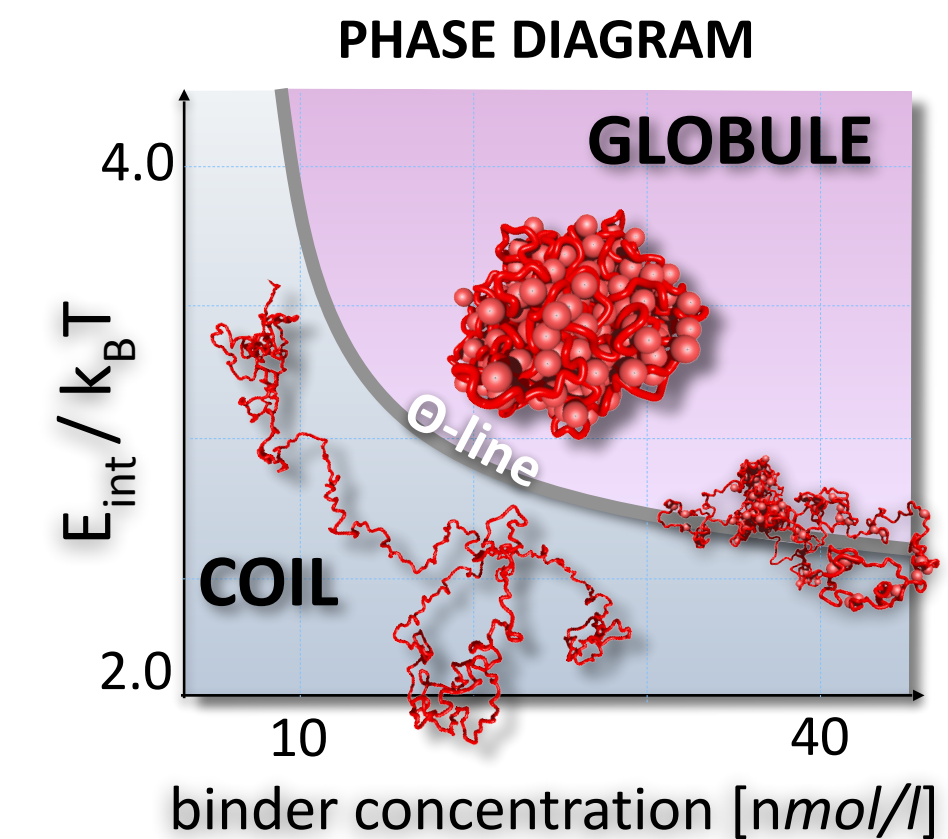


The **SBS model** is a polymer chain (SAW) with binding sites for diffusing molecules:

$$H = H^{\text{SAW}} + H^{\text{GAS}} + H^{\text{Int}}$$

SAW polymer particle gas interaction

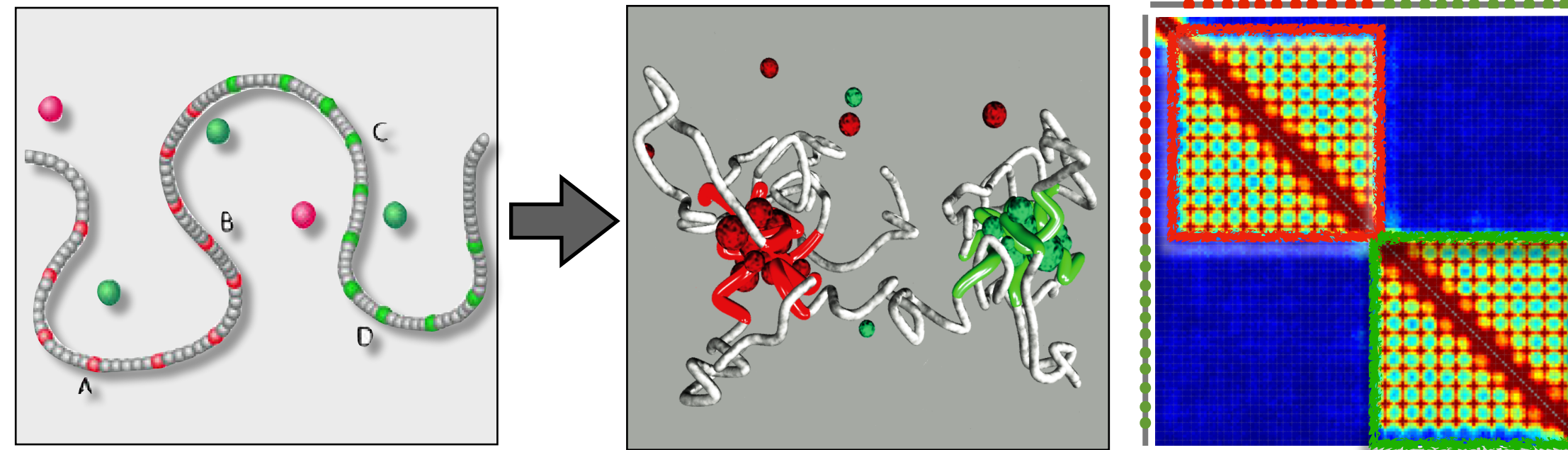


A **phase transition** controls folding *switch-like*, with no need of molecular fine tuning.

Folding of the SBS model

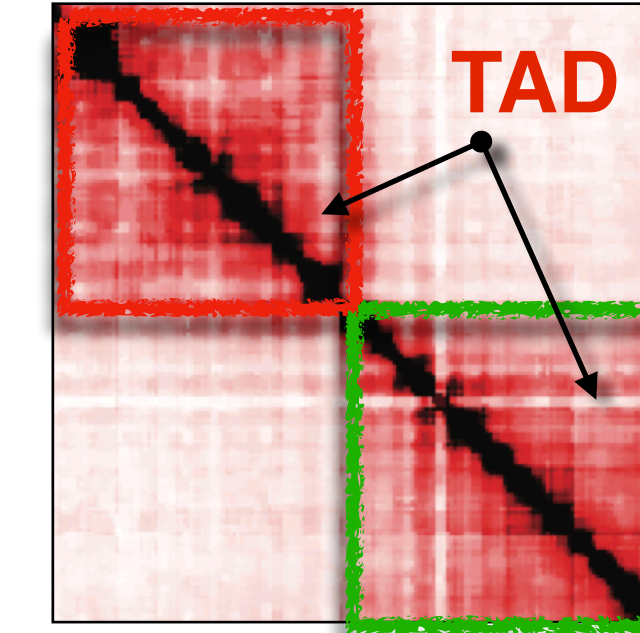
Contact patterns result from polymer micro phase-separation.

SBS TOY MODEL



(Barbieri et al. *PNAS* 2012, *Nature SMB* 2017)

5C DATA

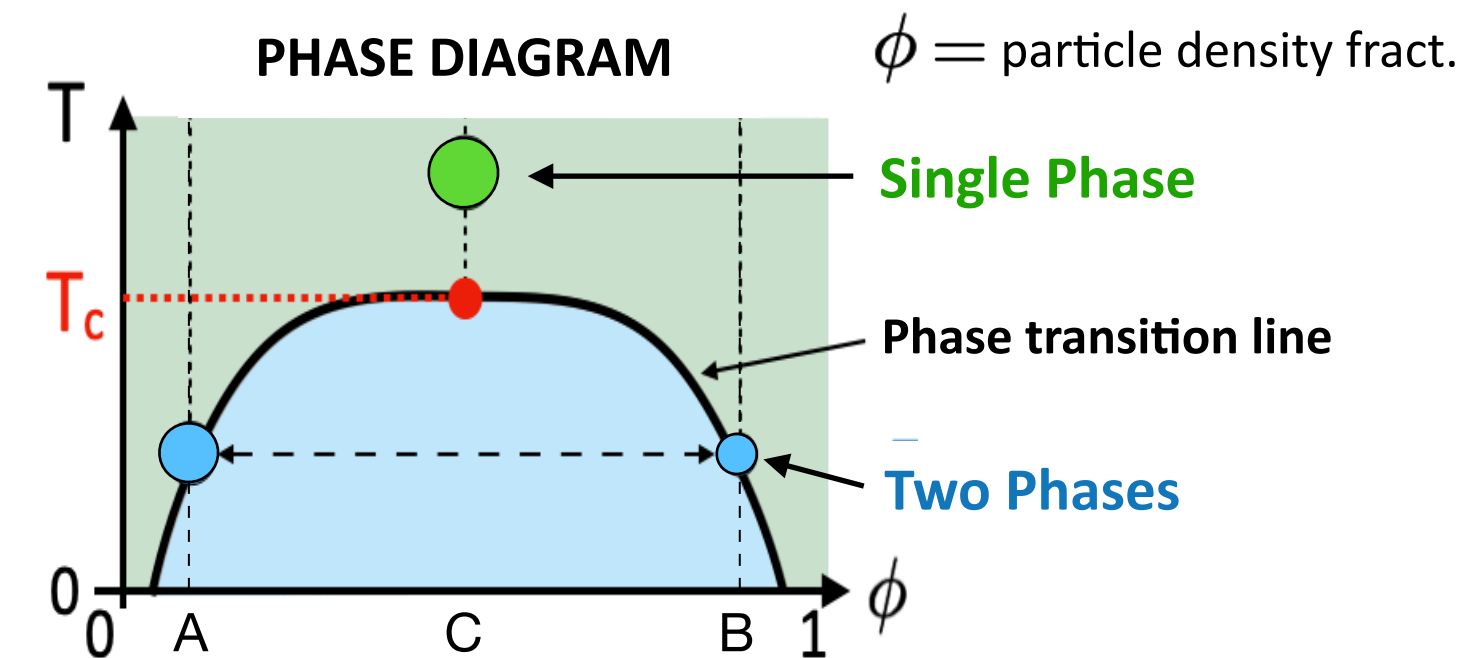
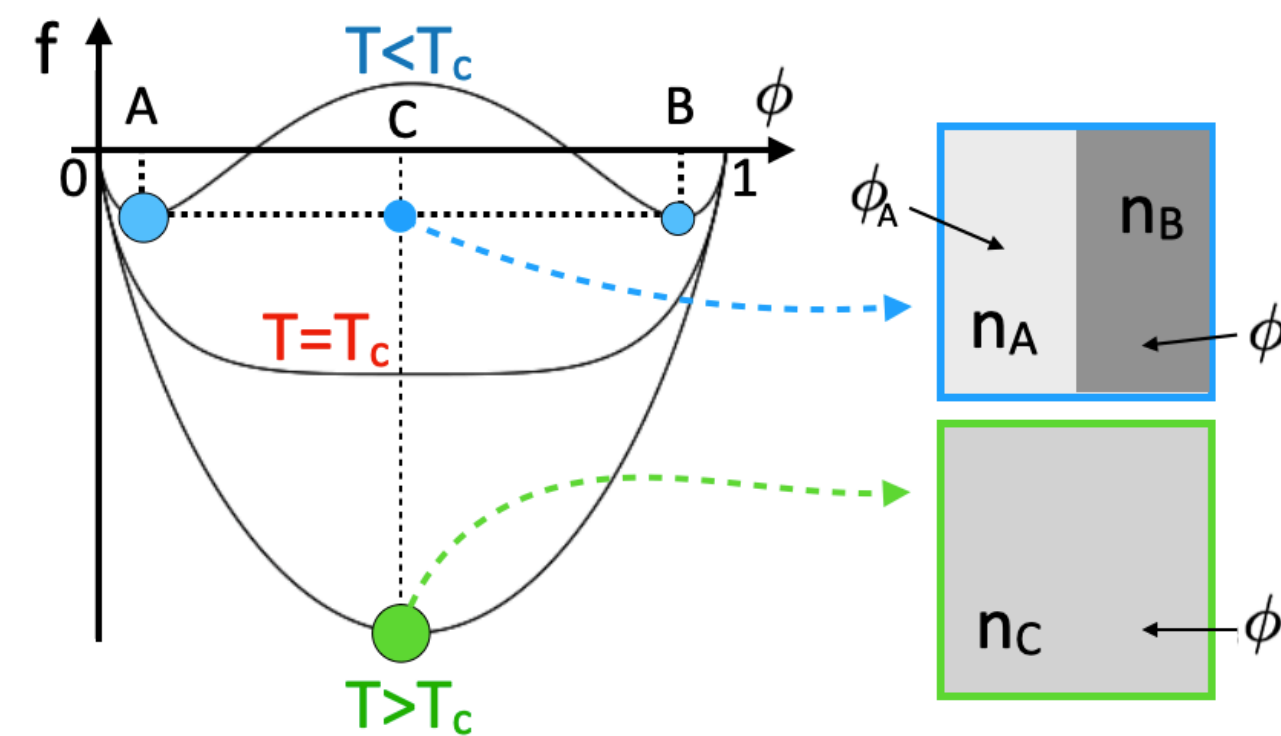


(Nora et al.; Dixon et al. *Nature* 2012)

mean-field free energy: $f(\phi) = k_B T [\phi \ln \phi + (1 - \phi) \ln(1 - \phi) + 2J\phi(1 - \phi)]$

← Entropy → ← Coupling →

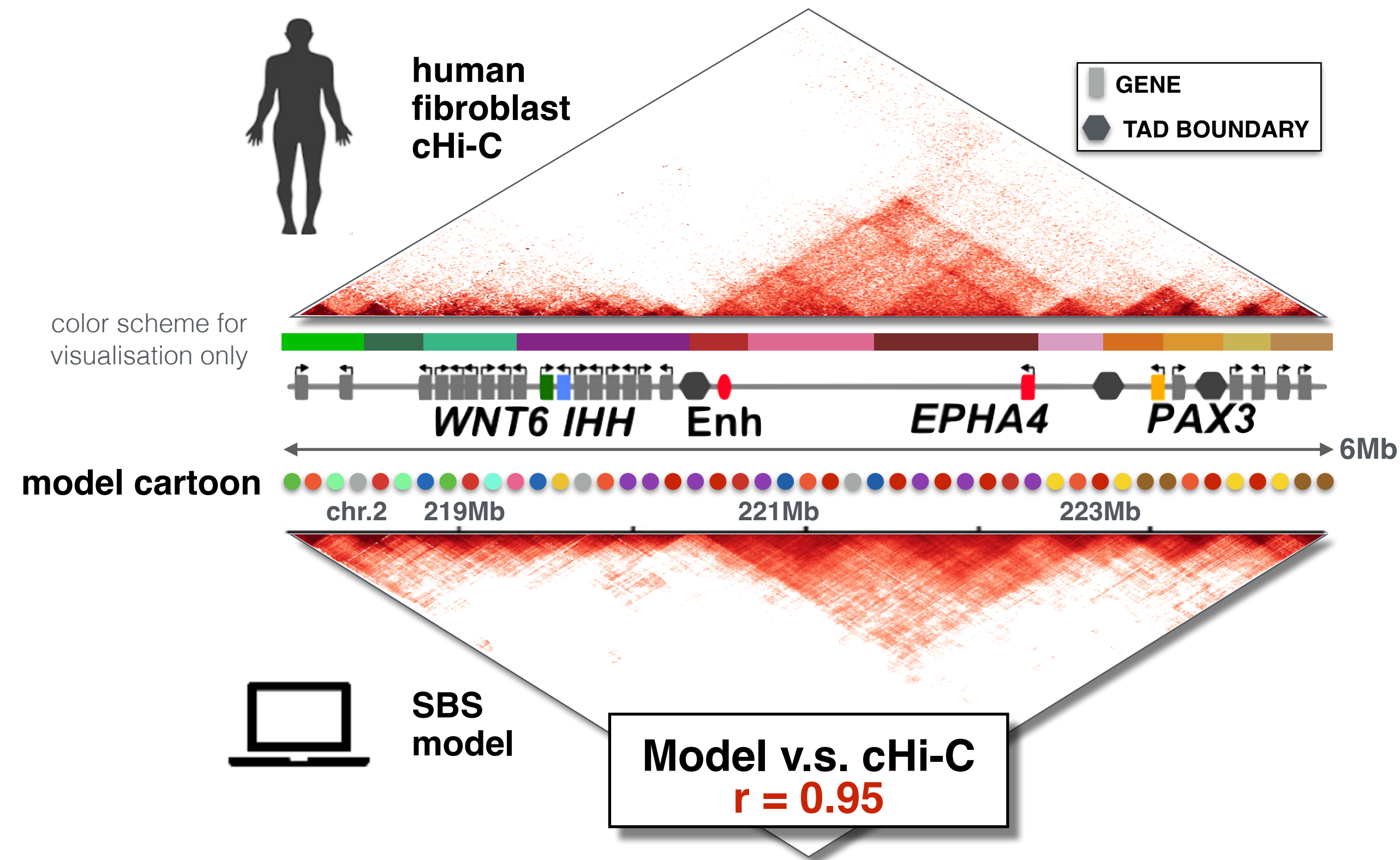
Alike water-oil phase separation



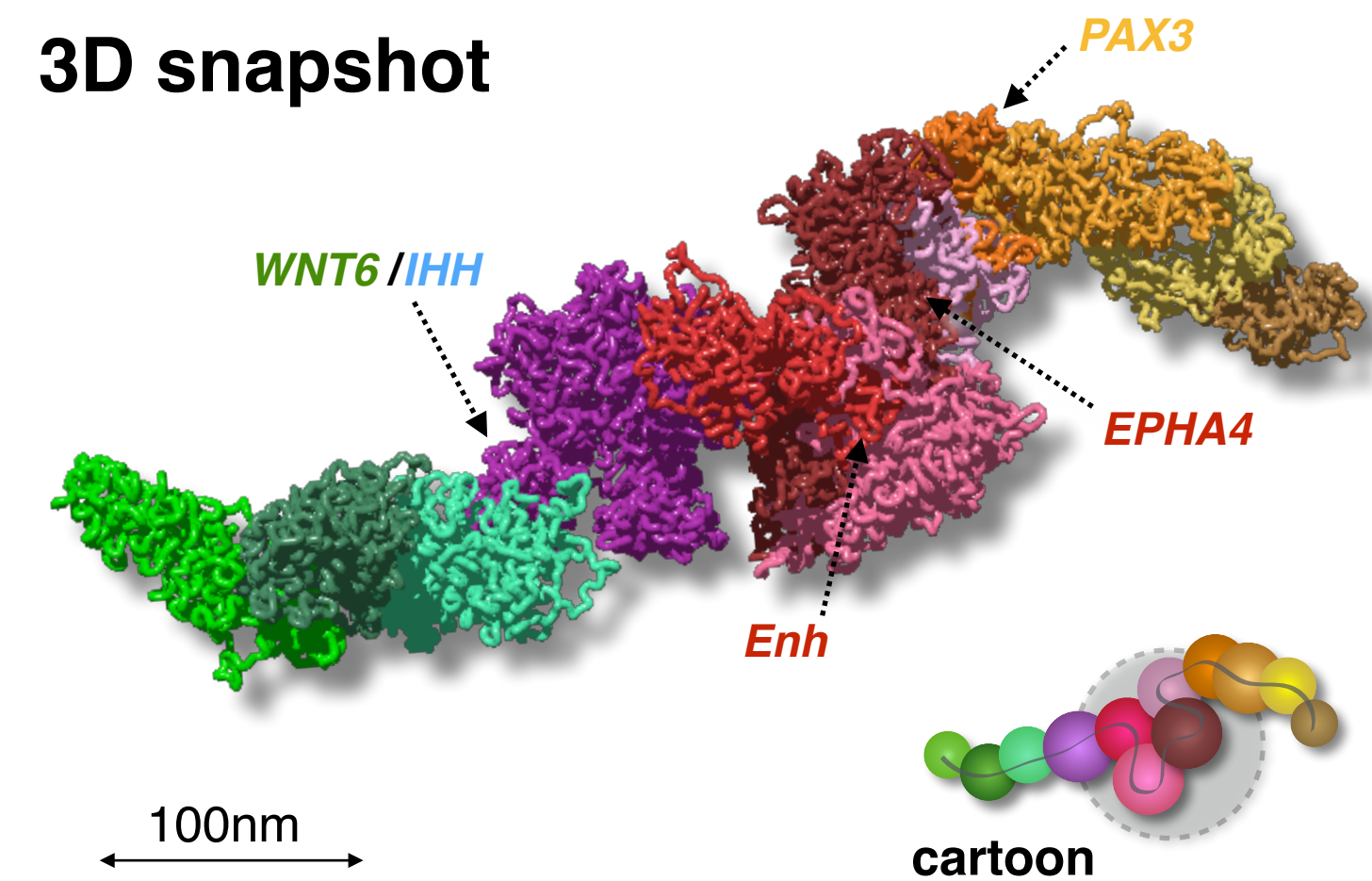
$\phi =$ particle density fract.

The *EPHA4* gene region

The SBS model explains contact maps with good accuracy.



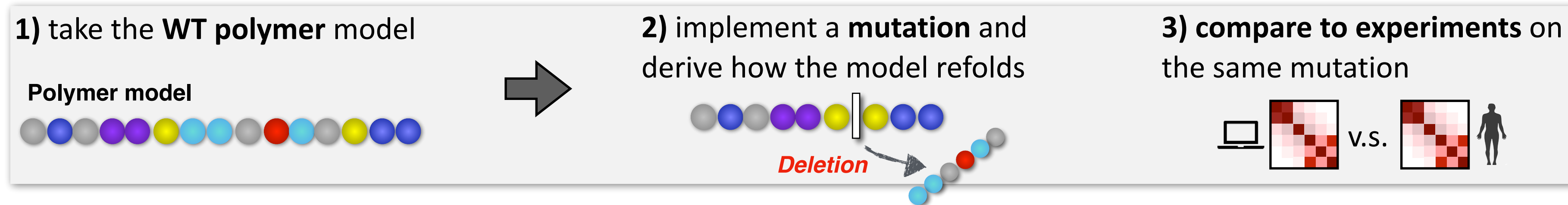
3D snapshot



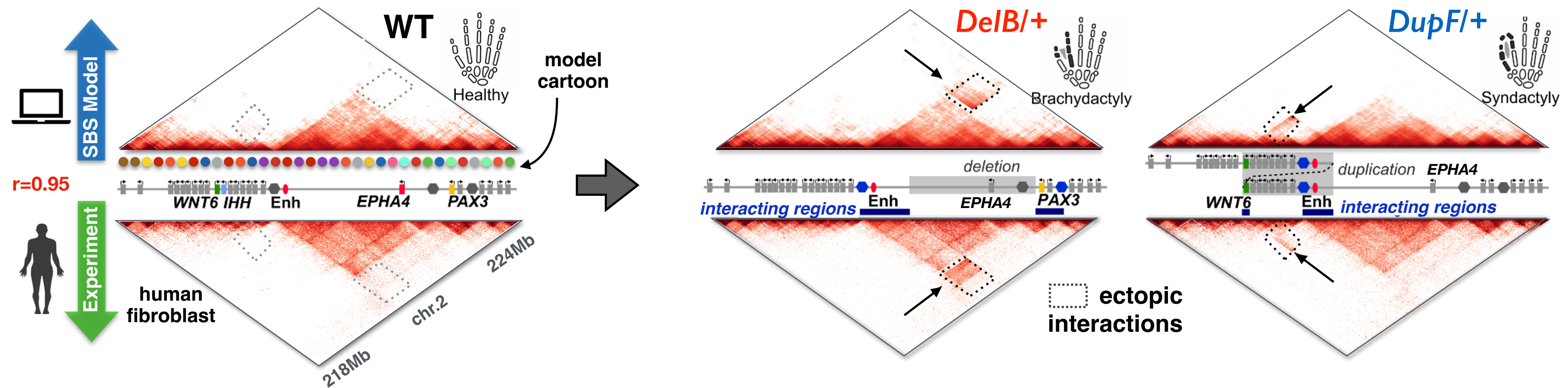
Contacts have a **complex 3D structure**, not just loops.

A stringent test of the theory

Use physics to predict the effects of mutations on 3D architecture and function of the genome.



No available fitting parameters!



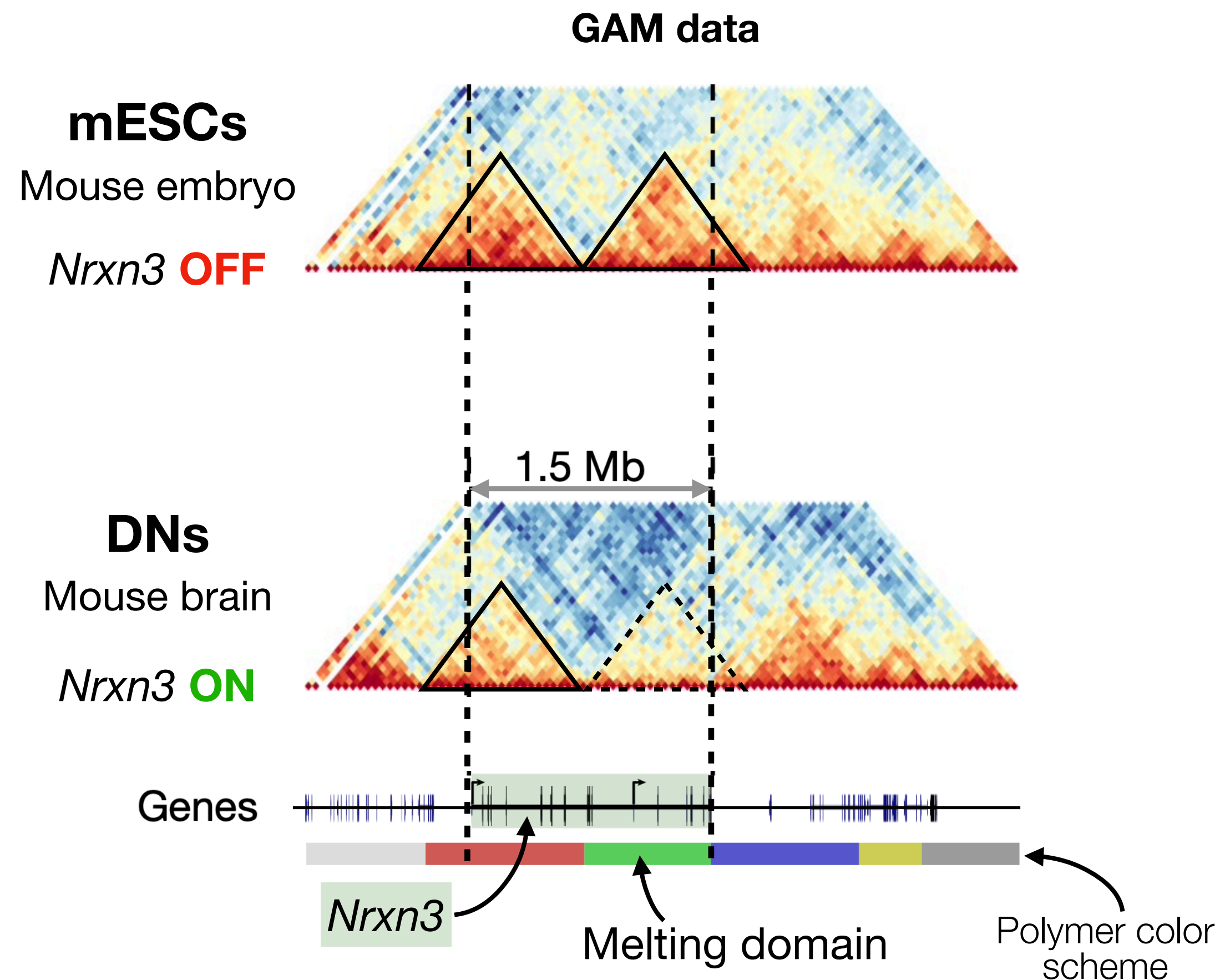
Pax3, *Wnt6* are upregulated in resp. *DelB/+*, *DupF/+* mice.

(Bianco et al. *Nature Gen.* 2018; Ringel et al. *Cell* 2022; Kragestein et al. *Nature Gen.* 2018)

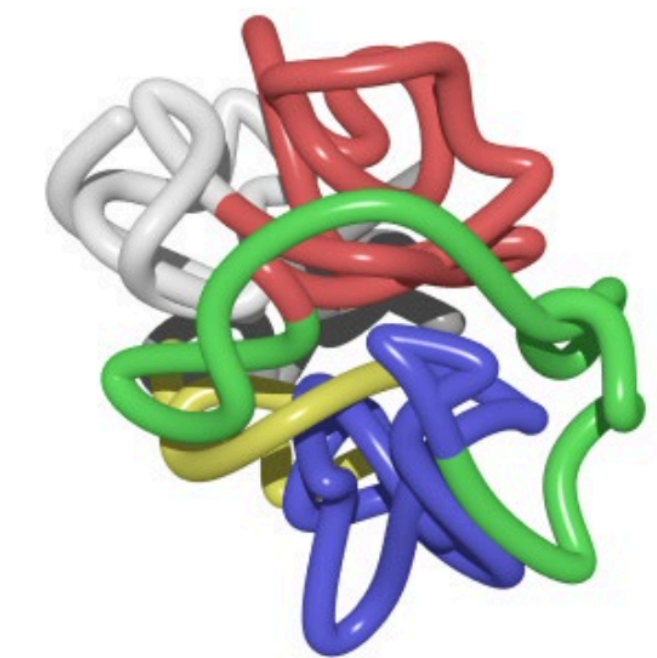
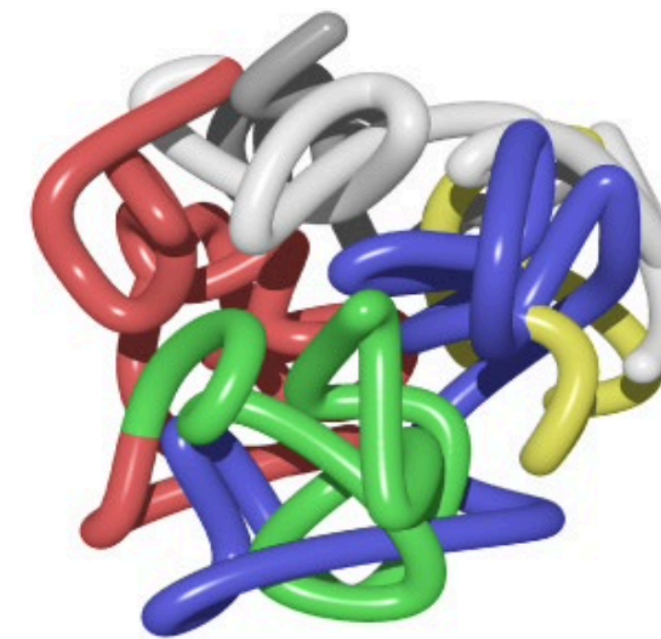
Mutations rewire regulatory contacts and induce diseases.

3D structure and gene regulation in neuronal development

Extensive cell-type specific 3D chromosome structures relate to patterns of gene expression in mouse brain.



Polymer physics

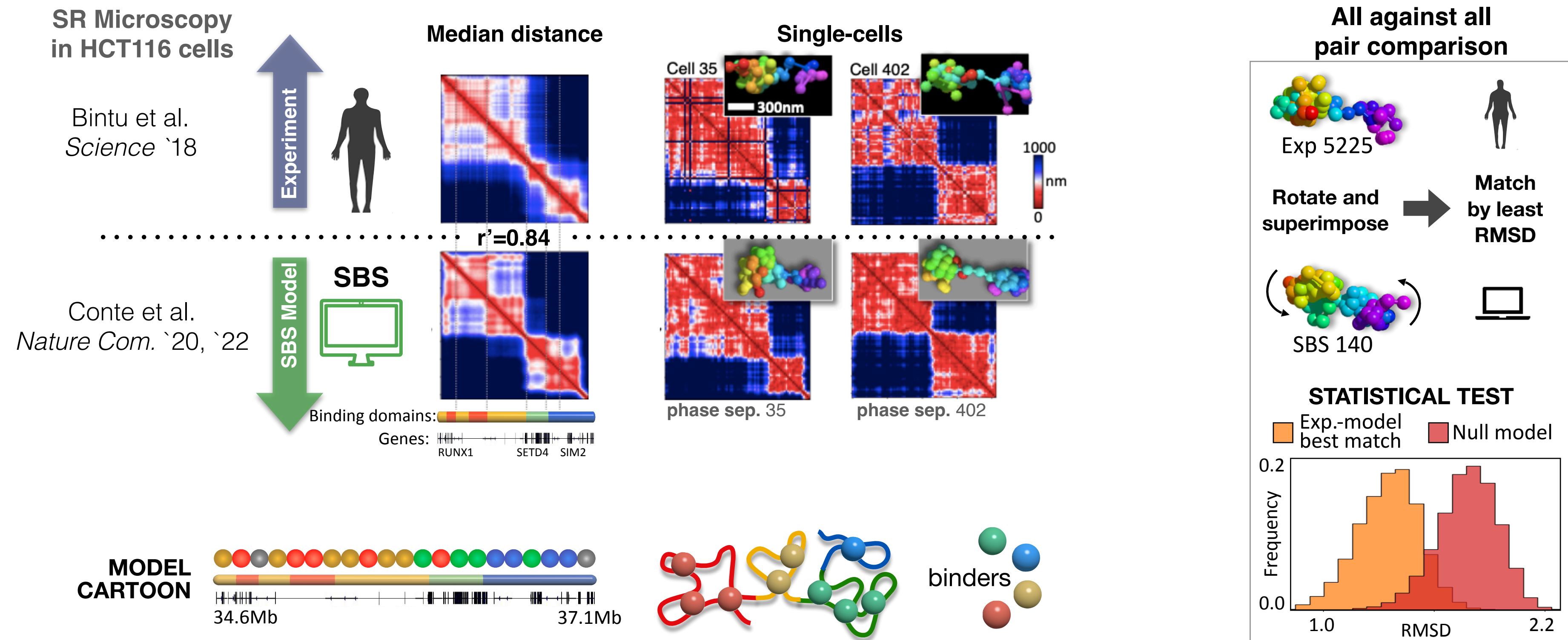


Melting domain

Nrxn3 occupies two compact TADs in mESCs but **melts** in dopaminergic neurons (DNs) where it is accessible and **expressed**.

Ensemble Distribution

Polymer physics explains the distribution of DNA 3D structure across single-cells, not just average contacts.

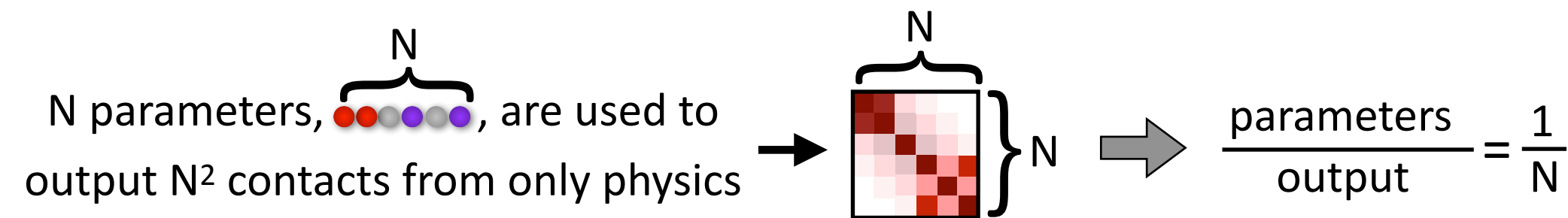
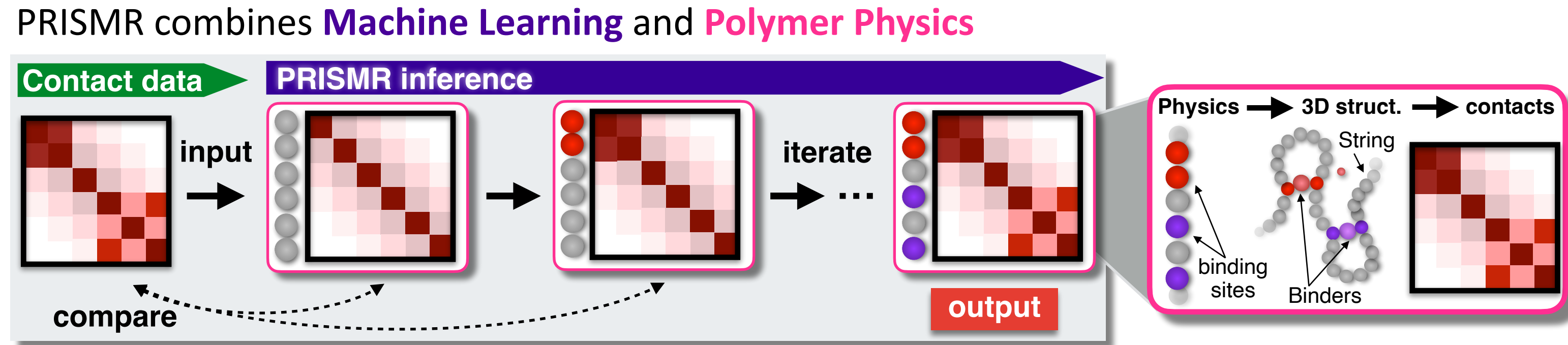


Cell-to-cell variability results from thermodynamic degeneracy of folding.

The PRISMR method

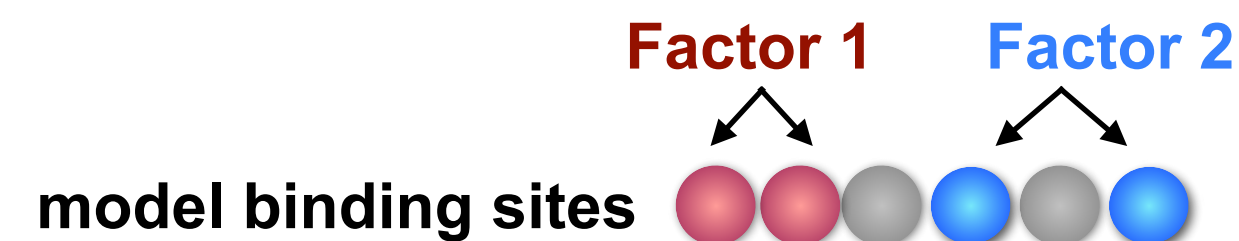
PRISMR infers the minimal polymer model that best explains a given contact matrix.

- It infers the model colours from experimental data:



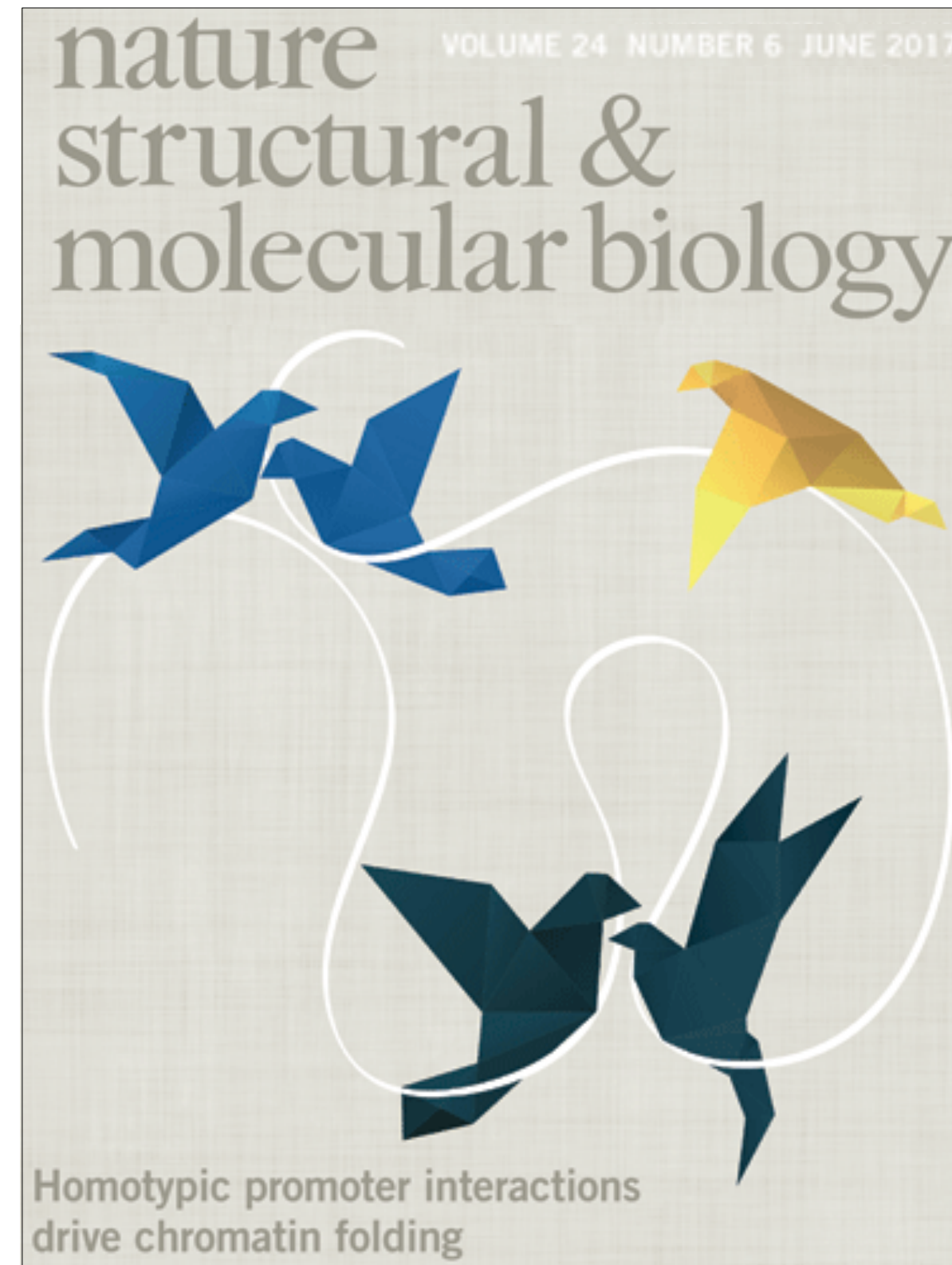
(Bianco et al. *Nature Gen.* 2018)

- The biological nature of colours is next searched through experiments:



(see, e.g., Baribieri et al. *Nature SMB* 2017; Esposito *Cell Rep.* 2022)

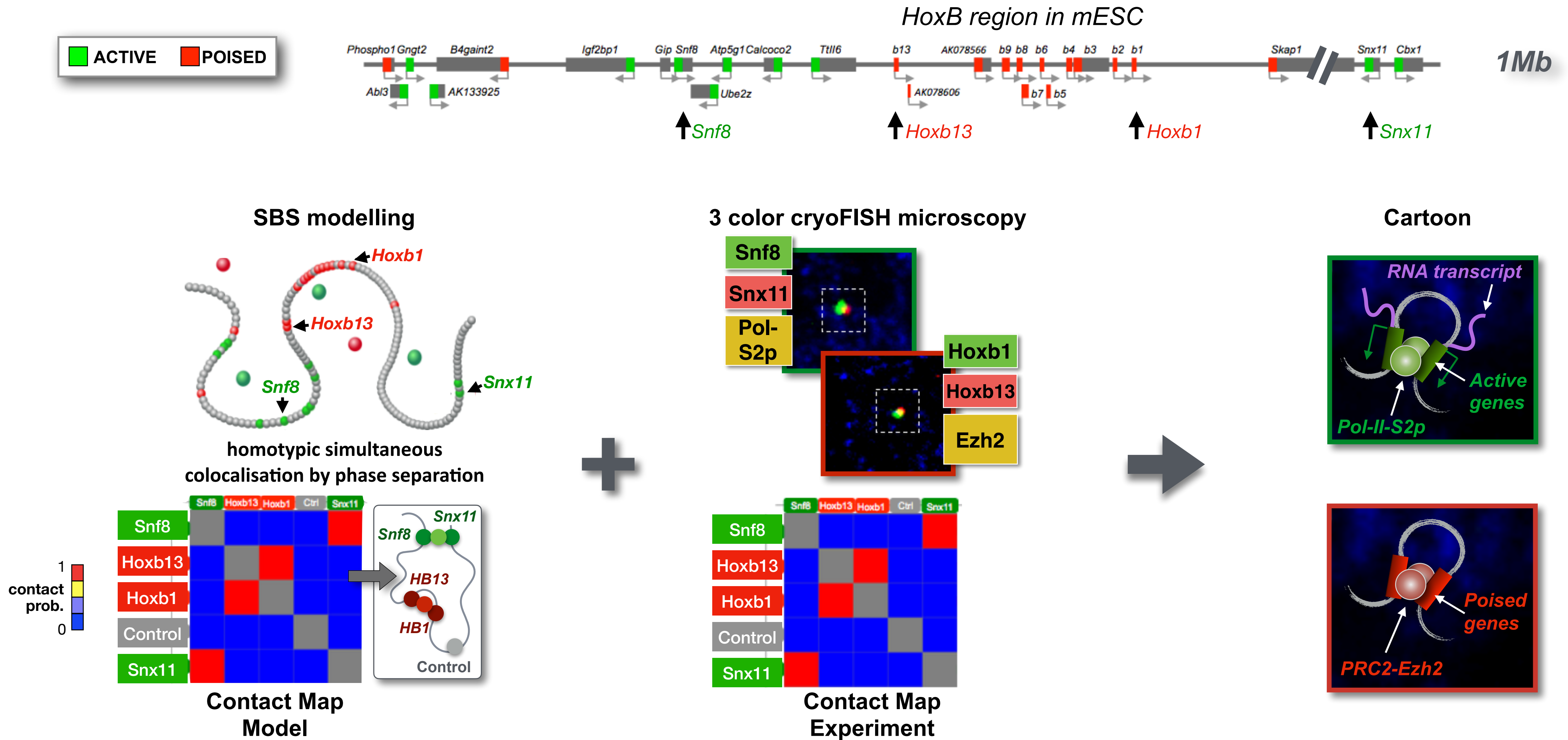
Binding particles



The combination of **microscopy** and **SBS model** helps identifying molecular factors shaping folding.

Binding factors in the *HoxB* region in mESCs

Active / poised gene promoters colocalize homotypically resp. with Pol-II-S2p / PRC2-Ezh2.



Multiple factors contribute to folding

Different binding site types of the model correlate with distinct combinations of DNA marks and binding factors.

- SBS model** of the locus in HCT116



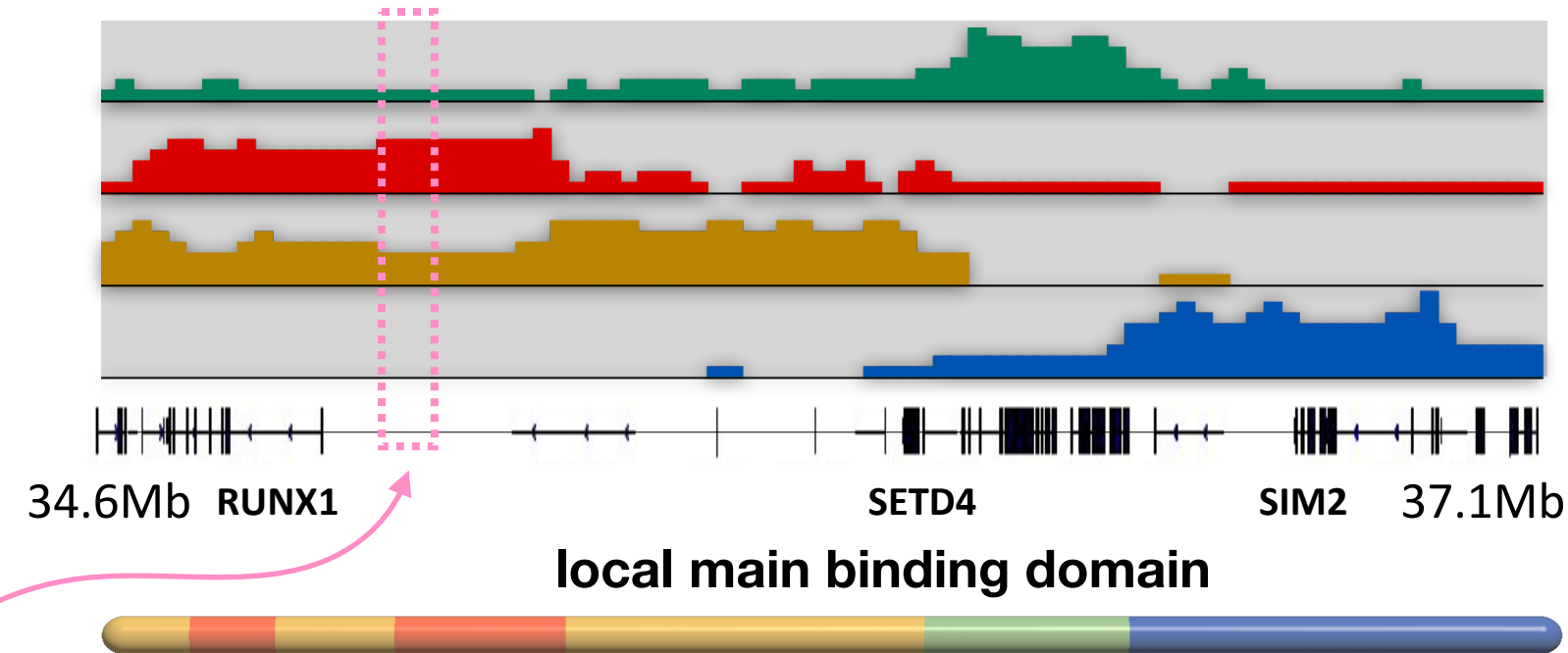
CARTOON:

4 types of binding sites

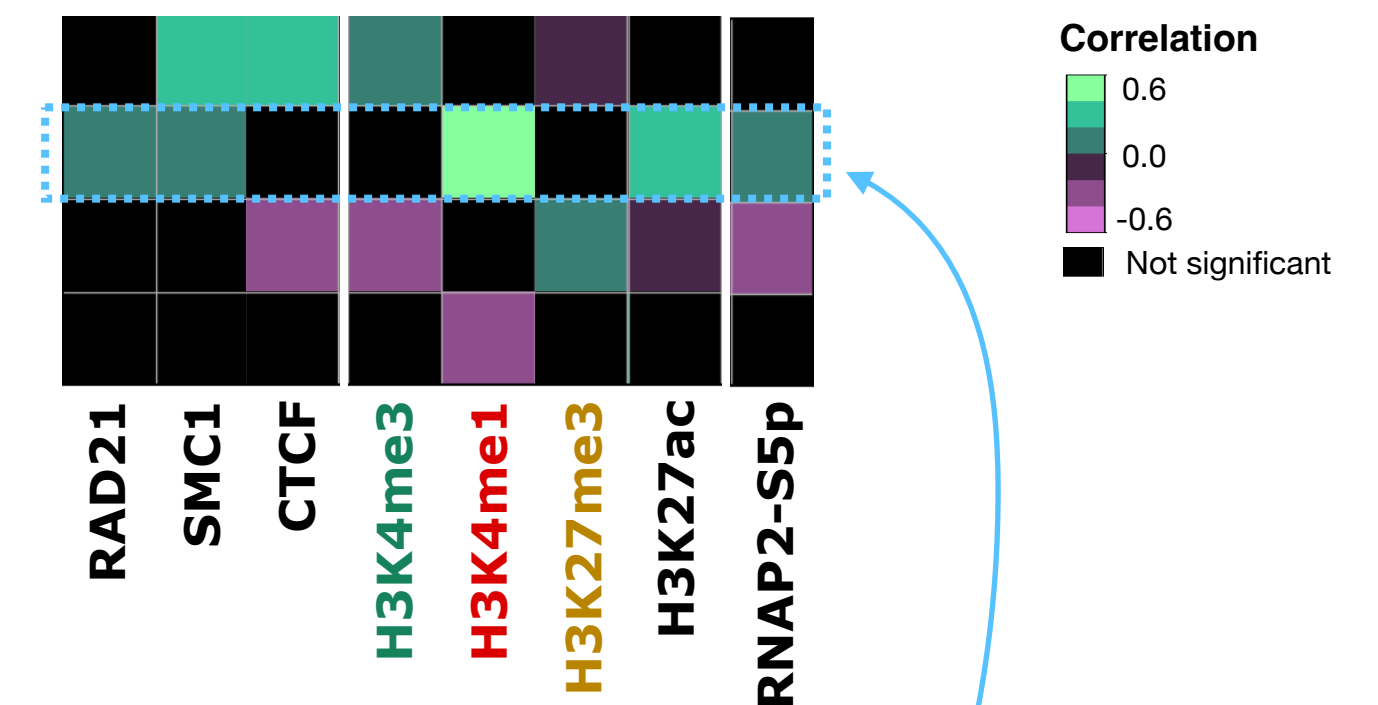


Binding domains **combinatorially overlap** along DNA

Abundance of **model binding sites**:



Correlation with 1D DNA marks:

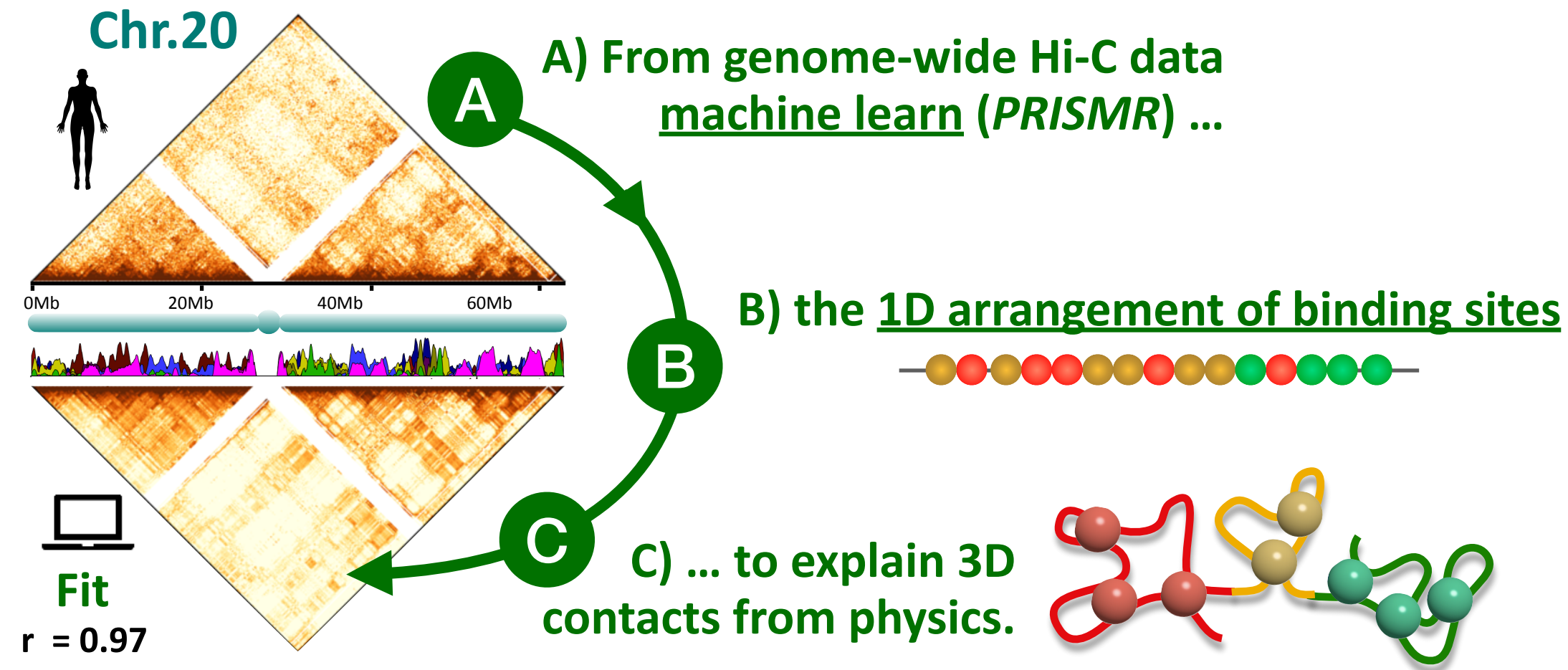


Binding domains **correlate** with **combinations** of factors

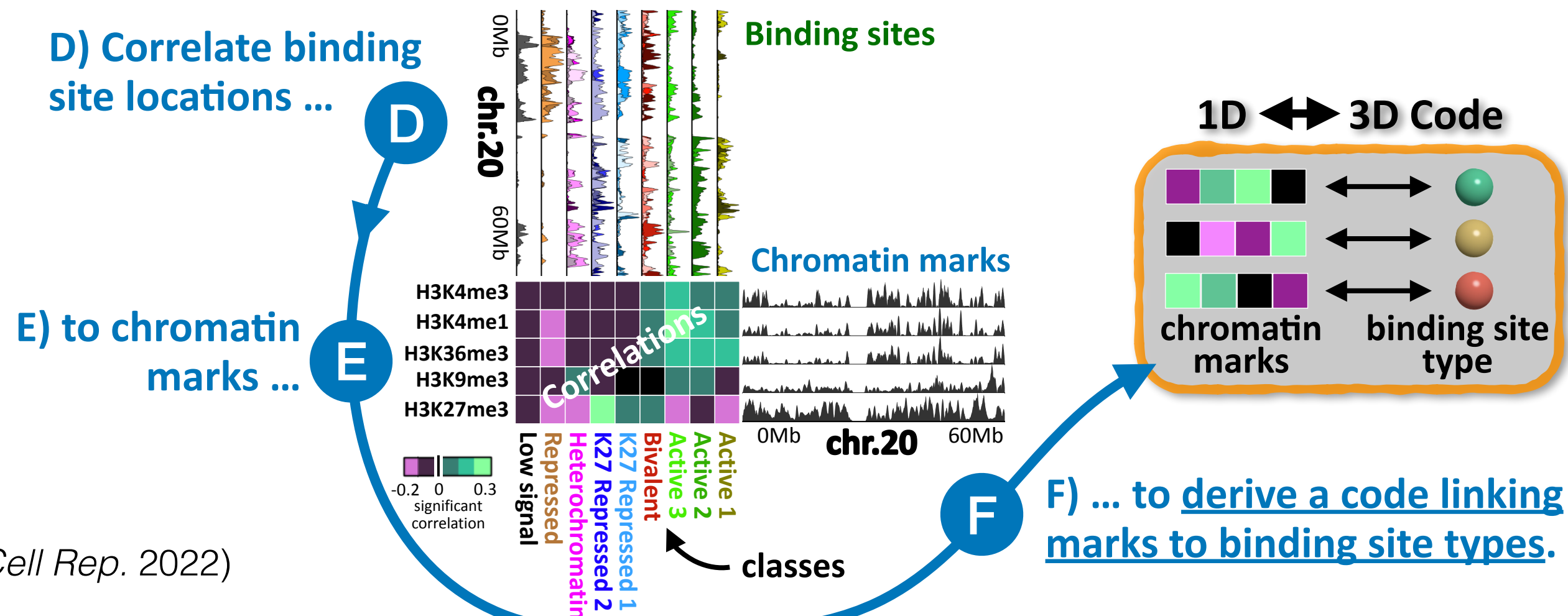
From 1D sequence to model binders to 3D structure

A combinatorial code of 1D chromatin marks predicts 3D structure genome wide (1D ↔ 3D code).

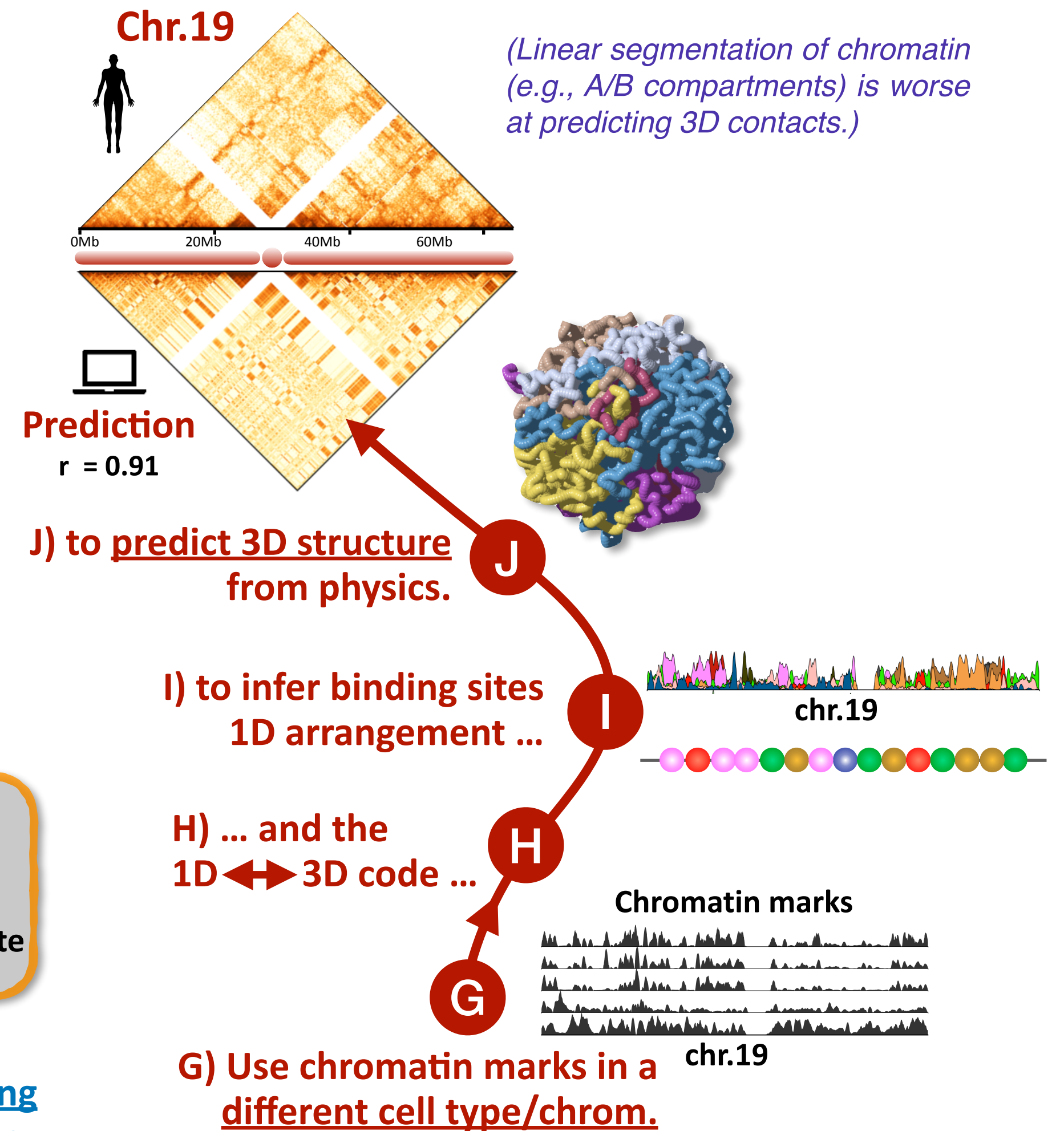
1) Infer model binding sites:



2) Correlate bind. sites to histone marks:



3) Predict 3D structure:



Conclusions

- **Chromosome structure** is shaped by **phase transitions**, which control genome functions.
- **Physics predicts** 3D impact of **mutations** and origin of **associated diseases**.
- **Machine Learning** & **physics** combined discover **new molecular factors** defining DNA 3D organisation (e.g., Pol-II).

“No field is making more progress than biology and [...] the most powerful assumption to understand life is that everything that living things do can be derived in terms of physics and atoms.”

Richard P. Feynman

