De Novo Chromosome Structure Prediction

MICHELE DI PIERRO, RYAN R. CHENG, Rice University, EREZ LIEBERMAN-AIDEN, Baylor College of Medicine, PETER G. WOLYNES, JOSE’ N. ONUCHIC, Rice University — Chromatin consists of DNA and hundreds of proteins that interact with the genetic material. In vivo, chromatin folds into nonrandom structures. The physical mechanism leading to these characteristic conformations, however, remains poorly understood.

We recently introduced MiChroM [1], a model that generates chromosome conformations by using the idea that chromatin can be subdivided into types based on its biochemical interactions. Here we extend and complete our previous finding by showing that structural chromatin types can be inferred from ChIP-Seq data. Chromatin types, which are distinct from DNA sequence, are partially epigenetically controlled and change during cell differentiation, thus constituting a link between epigenetics, chromosomal organization, and cell development. We show that, for GM12878 lymphoblastoid cells we are able to predict accurate chromosome structures with the only input of genomic data. The degree of accuracy achieved by our prediction supports the viability of the proposed physical mechanism of chromatin folding and makes the computational model a powerful tool for future investigations.