

The three-dimensional structure of the genome plays a key role in regulatory control of the cell. Chromosomes are organized nonrandomly inside the nucleus and form a network of interactions. Recent experimental methods such as Hi-C have been used to probe the 3D architecture of the genome, giving average pairwise contact frequencies between chromosomes. However, deducing the spatial organization of chromosomes from this data remains a challenge due to high levels of noise and technical bias. Here, we propose a novel framework that leverages 1D features of the genome (e.g. gene expression) in combination with Hi-C data to identify interacting regions in the genome. First, we find domains of high average interaction in Hi-C maps using a large average sub-matrix algorithm. Then we construct a weighted network with genetic regions as nodes and interactions as edges, where the edge weights are given by the correlation between genomic features. Individual interacting clusters are determined using weighted correlation clustering on the network. In addition to recapitulating known organizational patterns of chromosome interactions, our method provides a quantitative framework that allows to couple features of the 1D genome with 3D interactions to uncover the guiding principles of genome spatial organization and regulatory control.