DeepHiC: Greatly Enhancing Chromatin Interaction Information Using Deep Learning

Chun Li

Case Western Reserve University
E-mail: lichun1668@gmail.com

Abstract: Hi-C is a high-throughput sequencing technique for studying chromatin conformations. Due to relatively high cost, a typical Hi-C dataset for a human sample has 150-300 million read pairs, which is often used for detecting chromatin conformations at the 40kb resolution. Results at finer resolutions require a lot more data (e.g., 10kb would require 3 billion read pairs). To enhance the quality of Hi-C data, we develop DeepHiC, a deep learning algorithm, to enhance the information of a typical dataset for analysis at the 10kb resolution. We evaluate the method from several perspectives: Spearman's correlation, effective depth, and chromatin loop calling. We also demonstrate the performance of the algorithm for its potential use across tissue types and across species.