



09
JAN.
2018

🕒 de 11h à 12h

SÉMINAIRE

Effective normalization for copy number variation in Hi-C data

Nelle Varoquaux

Normalization is essential to ensure accurate analysis and proper interpretation of sequencing data. Chromosome conformation data, such as Hi-C, is no different. The most widely used type of normalization of Hi-C data casts estimations of unwanted effects as a matrix balancing problem, relying on the assumption that all genomic regions interact as much as any other. Here, we show that these approaches, while very effective on fully haploid or diploid genome, fail to correct for unwanted effects in the presence of copy number variations. We propose a simple extension to matrix balancing methods that properly models the copy-number variation effects. Our approach can either retain the copy-number variation effects or remove it. We show that this leads to better downstream analysis of the three-dimensional organization of rearranged genome.