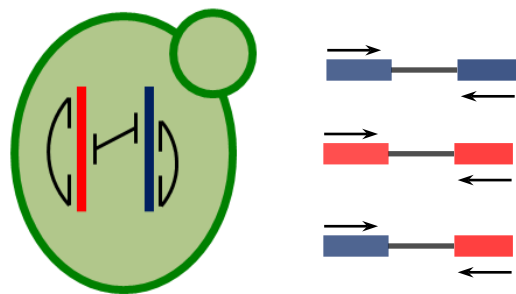
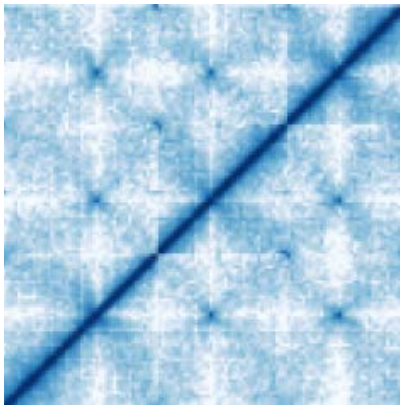


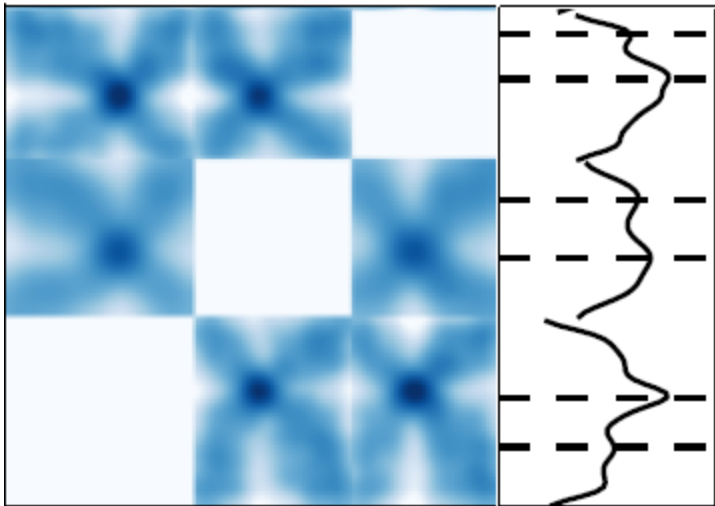
1. Mapping and filtering of Hi-C data.



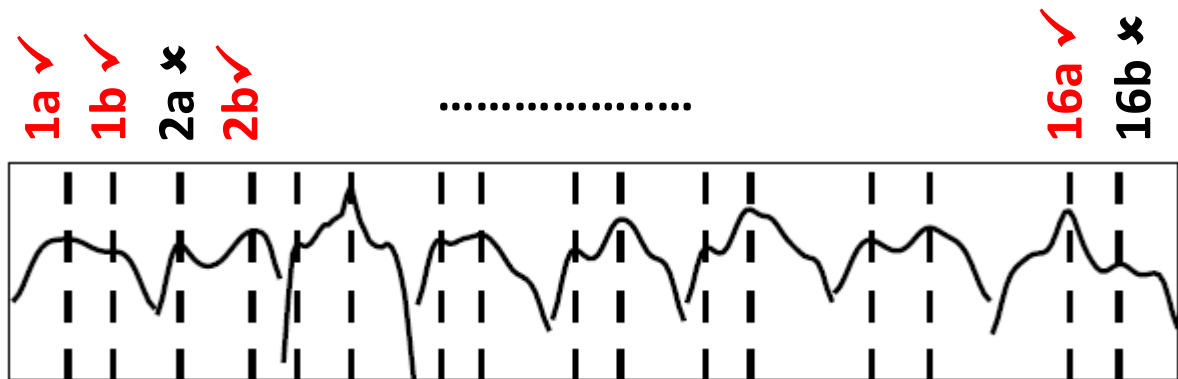
2. Normalizing and smoothing the contact maps.



3. Finding peaks in marginalized trans contact counts.

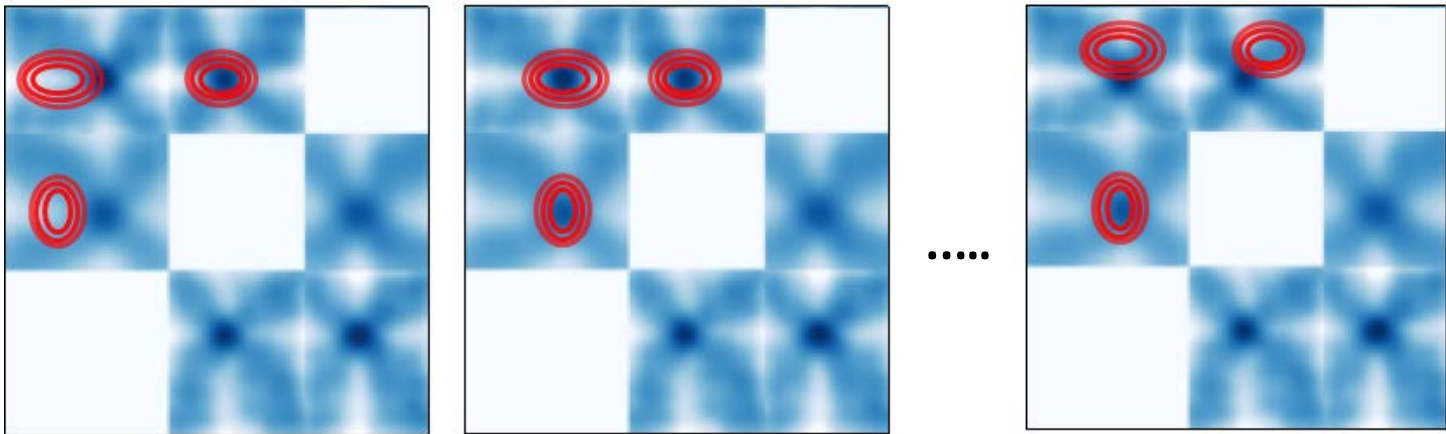


4. Selecting candidate centromeres.



candidate set = {{1a, 2b, ..., 16a}, {1b, 2b, 16a}, ...}

5. Jointly optimizing centromere coordinates over all chromosomes.

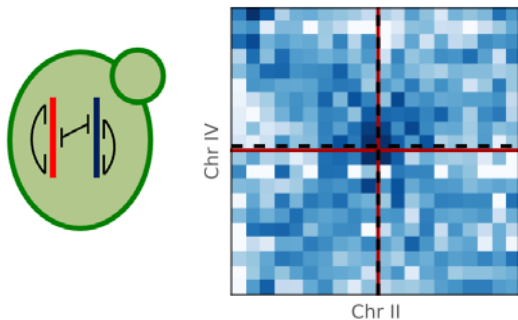


$f(c_{ij} - \text{fit}_1) = 5$

$f(c_{ij} - \text{fit}_1) = 1 \checkmark$

$f(c_{ij} - \text{fit}_n) = 3$

6a. Measuring accuracy (centromeres known)



6b. Calling de-novo centromeres (centromeres unknown)

