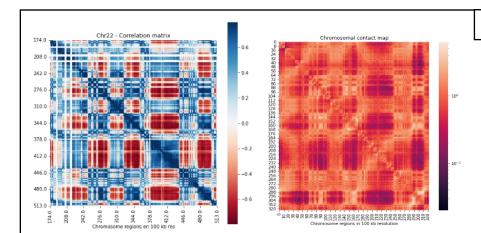
Hi-C matrix-based pipeline for multi-scale chromatin compartmental analysis

Mikal Daou, Oktavia Ścibior, Maxime Gueudré

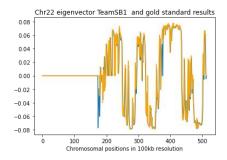
Based on Rao et al. (2014)

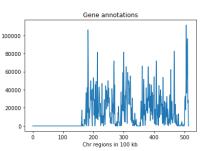
WORKFLOW Hi-C Sequencing & **RAW** data acquisition **Pre-processing of contact matrix** (bining, filtering, normalization, correlation...) **Detection of 2 main compartments** active: euchromatin inactive: heterochromatin **Comparison to GoldStandard** Sign of correlation between <0 gene density & eigenvectors **SWAP** >0 **Gene Density Analysis Detection of potential** subcompartments **EPIGENETICS MARKS HMM Model** Visual Approach: PyMol 3D **Vizualisation** p-value based **Unsupervised Clustering** approach: methods: KMeans & Welch test **Hierarchical Clustering**



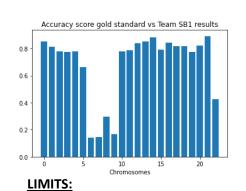
Correlation and Contact matrix

- Correlation matrix shows that compartments are NOT conserved across chromosomes
- Enhanced contacts on the diagonal





- Results matching with gold standard results
- Positive eigenvector values: enriched regions in intrachromosomal contact and open chromatin (compartment A)
- Enriched gene density in compartment A

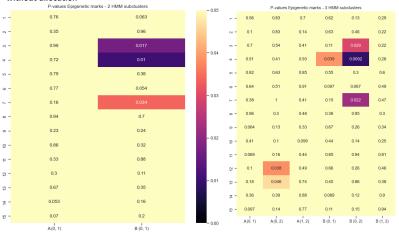


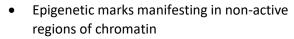
Detection of 2 main compartments

 Low accuracy: partial inversion, chromosome branches size

Detection of more than 2 compartments

p-values for 2 subcompartments in A and 3 subcompartments in B, correlation for 5 subcompartments without allocation





- Hierarchical organization of compartments
- Several approaches to detect optimal number of subcompartments

LIMITS:

- only based on intrachromosomal contacts
- tested in **ONE** cell
- only in 100kb resolution

