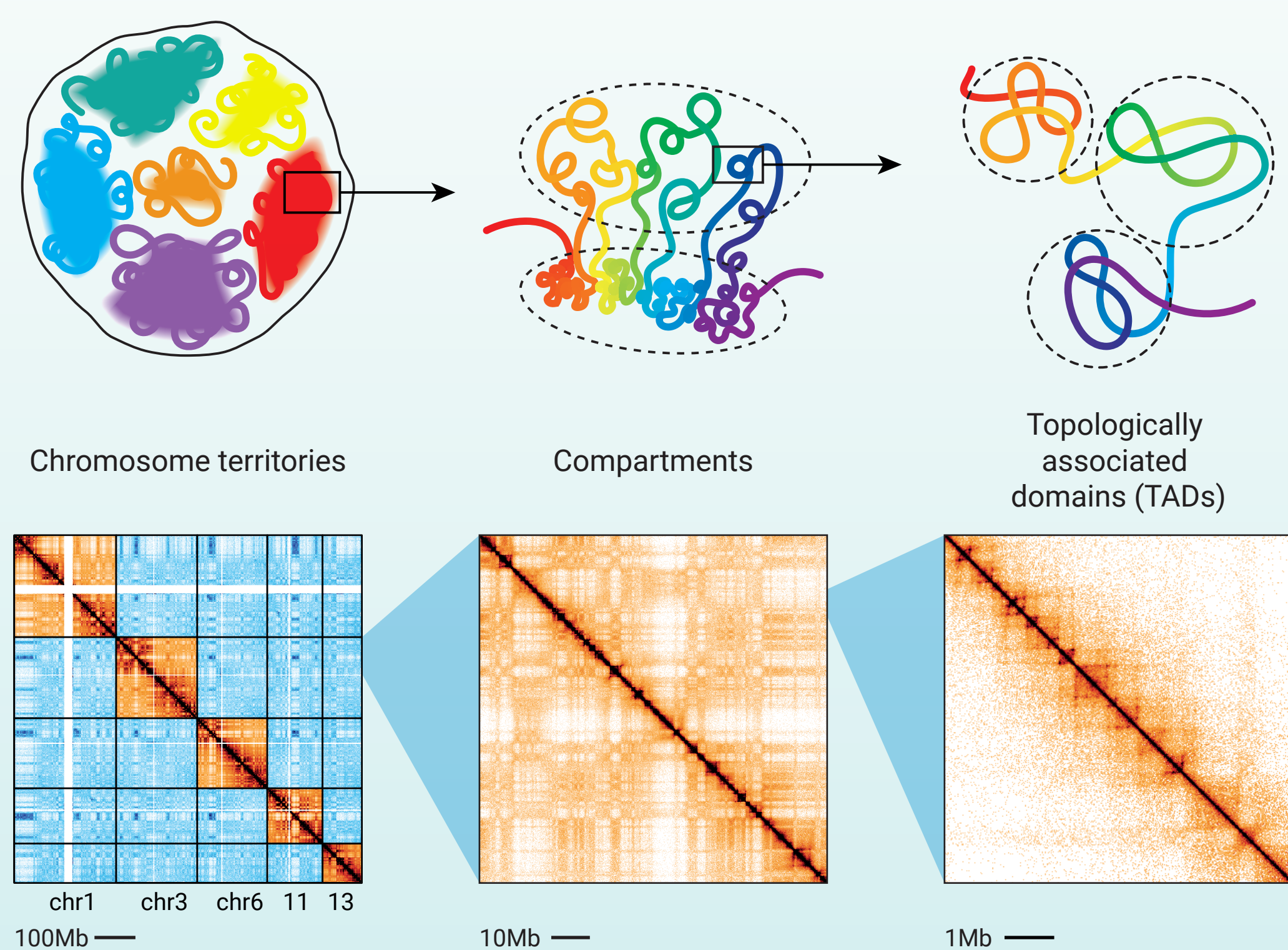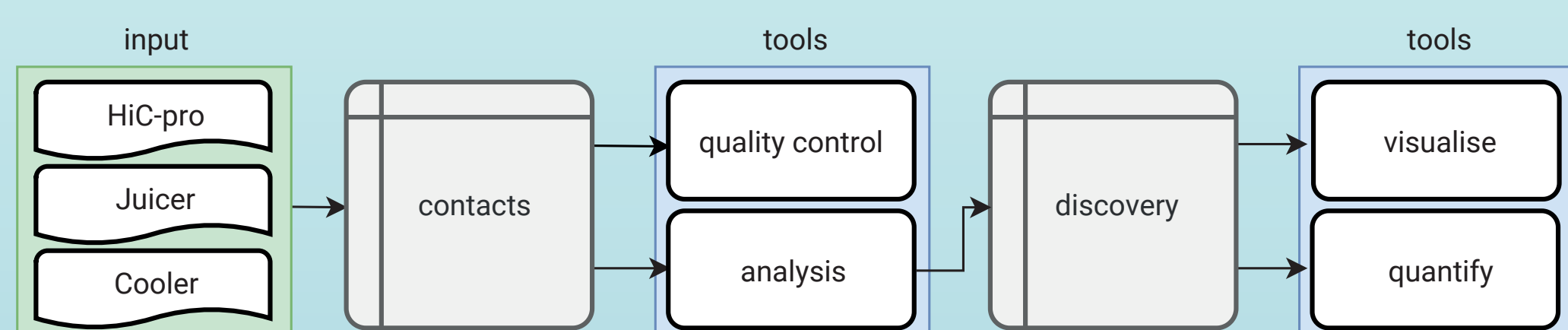# GEN🔬VA
Explore the Hi-Cs

Robin van der Weide, _Teun van den Brand_, **Elzo de Wit**
Division of Gene Regulation, Netherlands Cancer Institute / Oncode Institute

GENome Organisation Visualisation and Analysis

Hi-C data captures all-versus-all 3D DNA-DNA contact information at various levels of scale.

Chromosome territories — Compartments — Topologically associated domains (TADs)
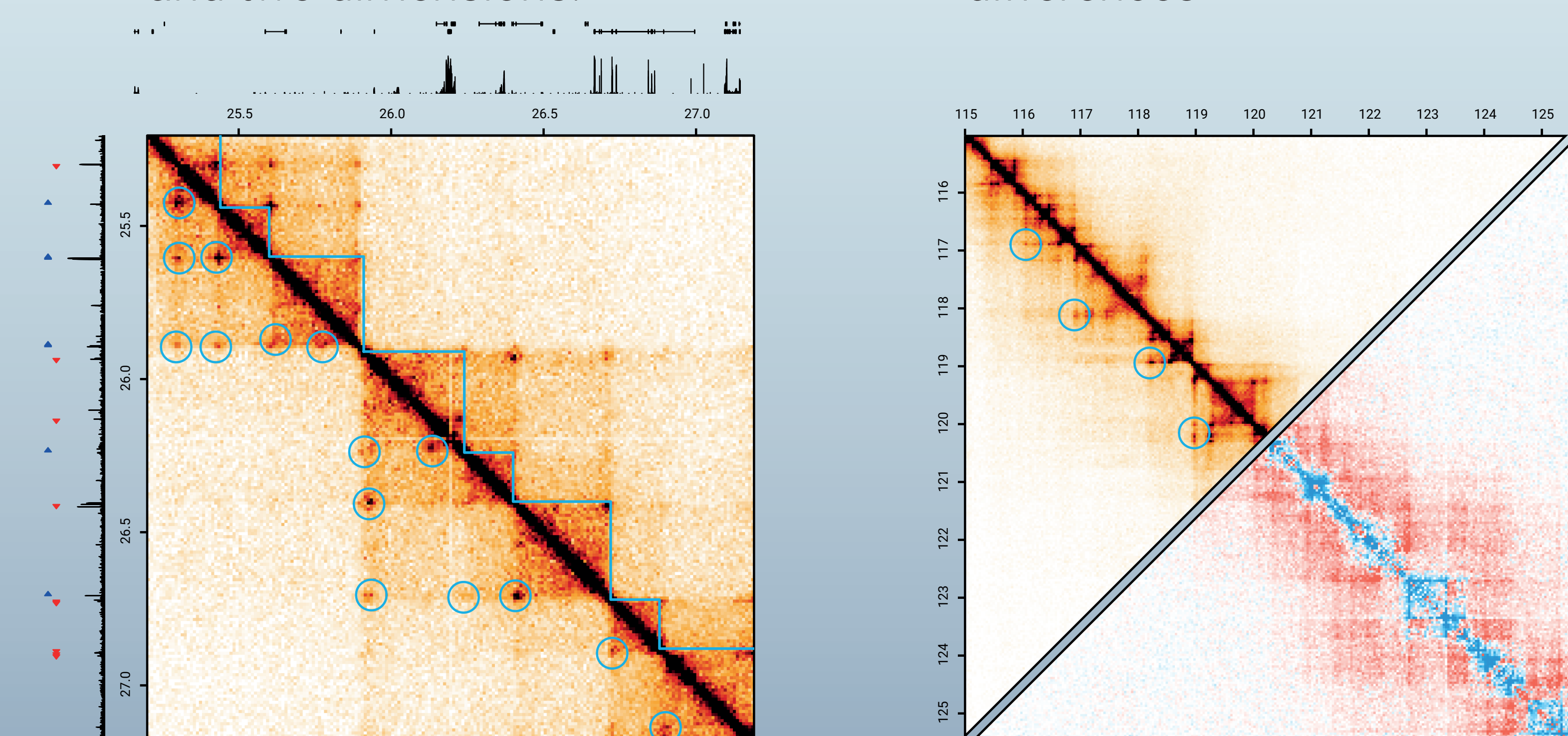
chr1 chr3 chr6 11 13
100Mb — 10Mb — 1Mb —

The GENOVA package accepts data from three most-used mapping pipelines to build 'contacts' objects. Analysis functions process these objects into 'discovery' objects which are easily visualised and quantified.

input: HiC-pro, Juicer, Cooler → contacts → tools: quality control, analysis → discovery → tools: visualise, quantify

Compute genome-wide scores and aggregations

Compartment scores discriminate active (A) from inactive (B) chromatin.

Insulation scores can be used to call TADs based on local minima.

Global Aggregates: saddle, RCP — genome, data
Local Aggregates: ARA, APA, ATA, PE-SCAn — genome, data

Compartment / Insulation / Hi-C
Location Chr1 (Mb)

Robustness and performance

Aggregation functions reveal similar structures across different sequencing depths.

1.6M reads — 25.6M reads
3' border, 5' border
5' border, 3' border
mean contacts ≥80 ... 0

Variability in analysis outcomes reflects variability in biology.

technical replicates
biological replicates
biological different
Spearman's ρ (insulation score)
0   0.25   0.50   0.75   1

The majority of analysis functions in GENOVA run in under a minute for 1-2 samples and typical input.

10kb resolution
GENOVA
Juicer
seconds elapsed
number of loops (APA)
4   32   256   2048

Annotation and visualisation of Hi-C data

Rich annotations of a diverse set of features in both one and two dimensions.

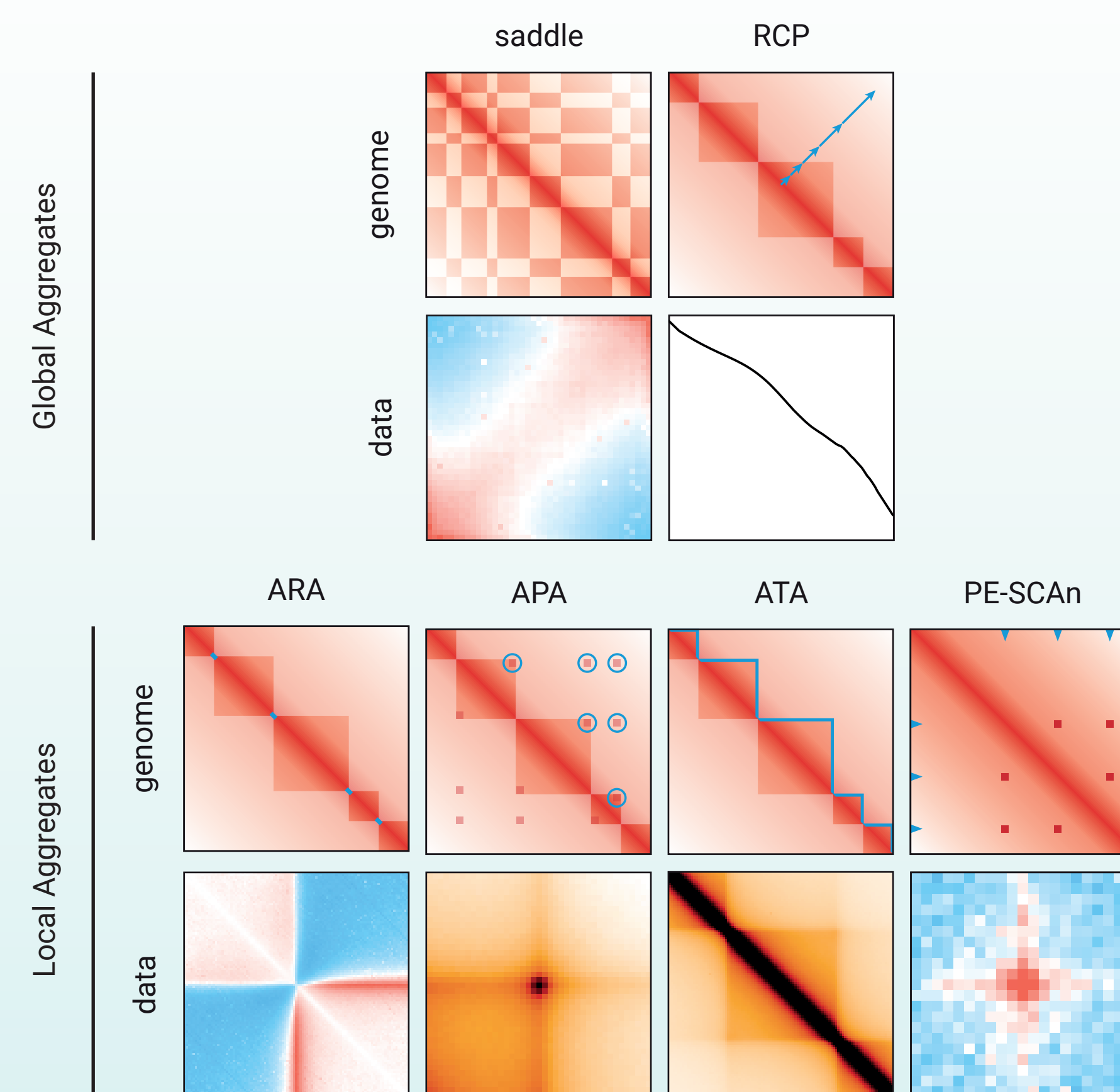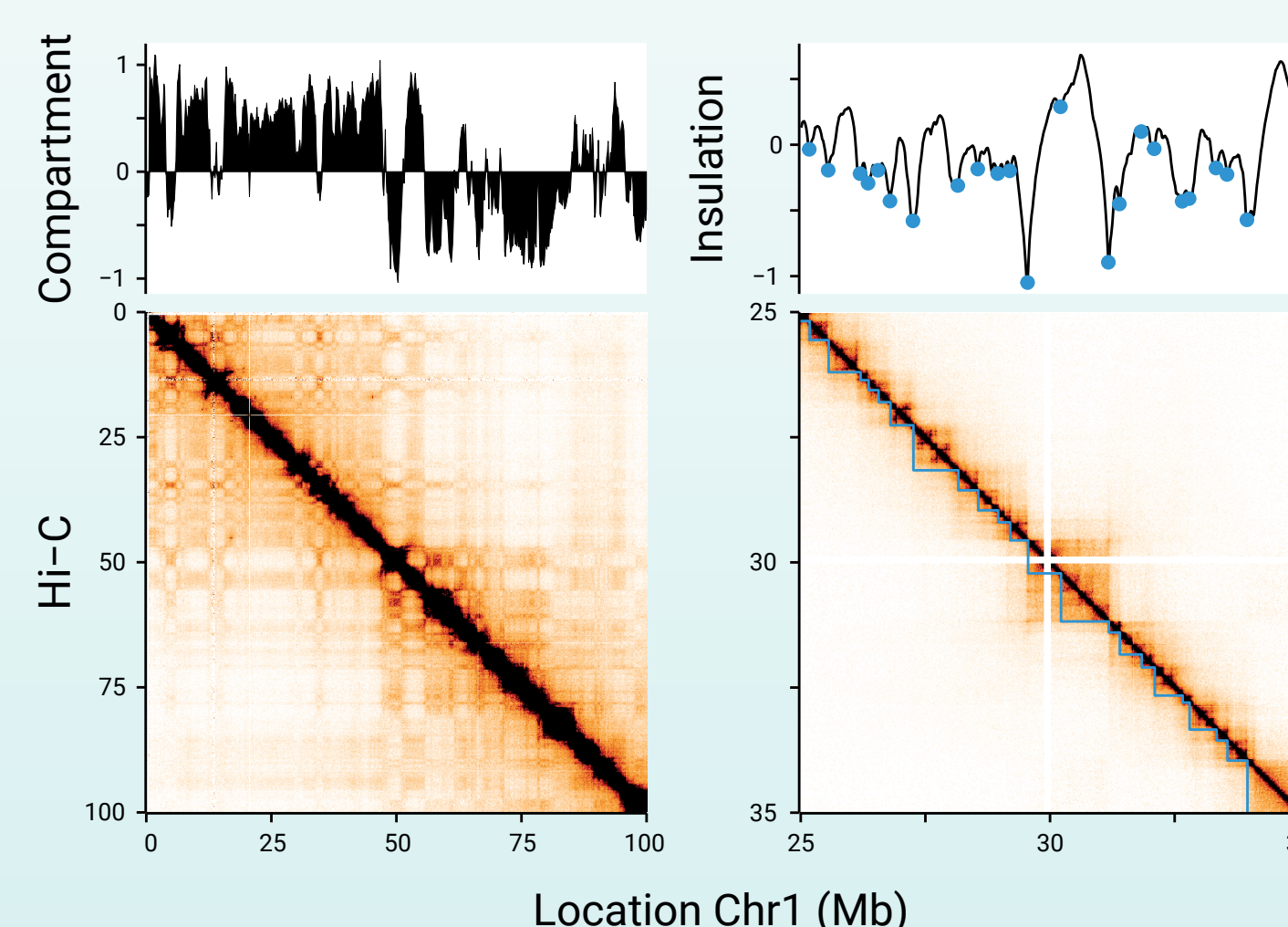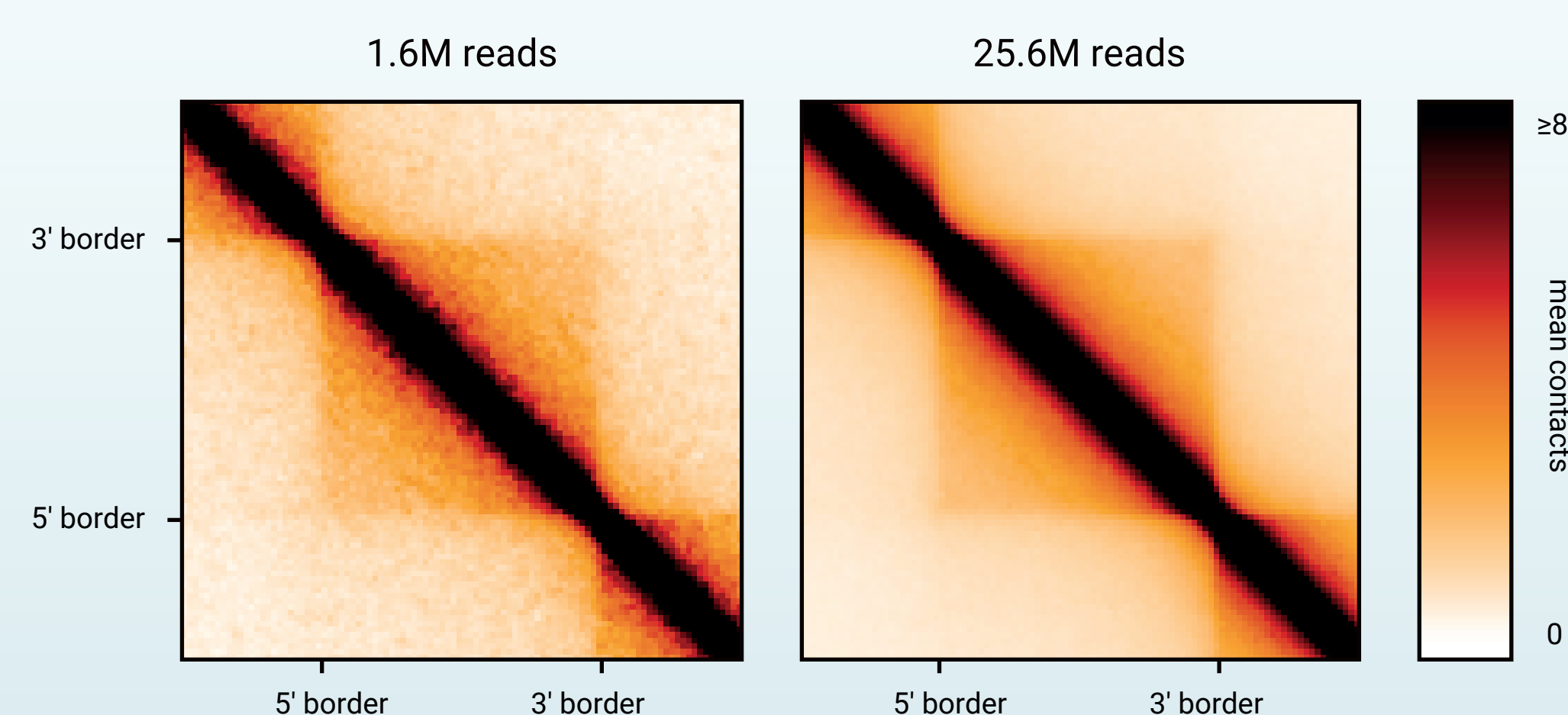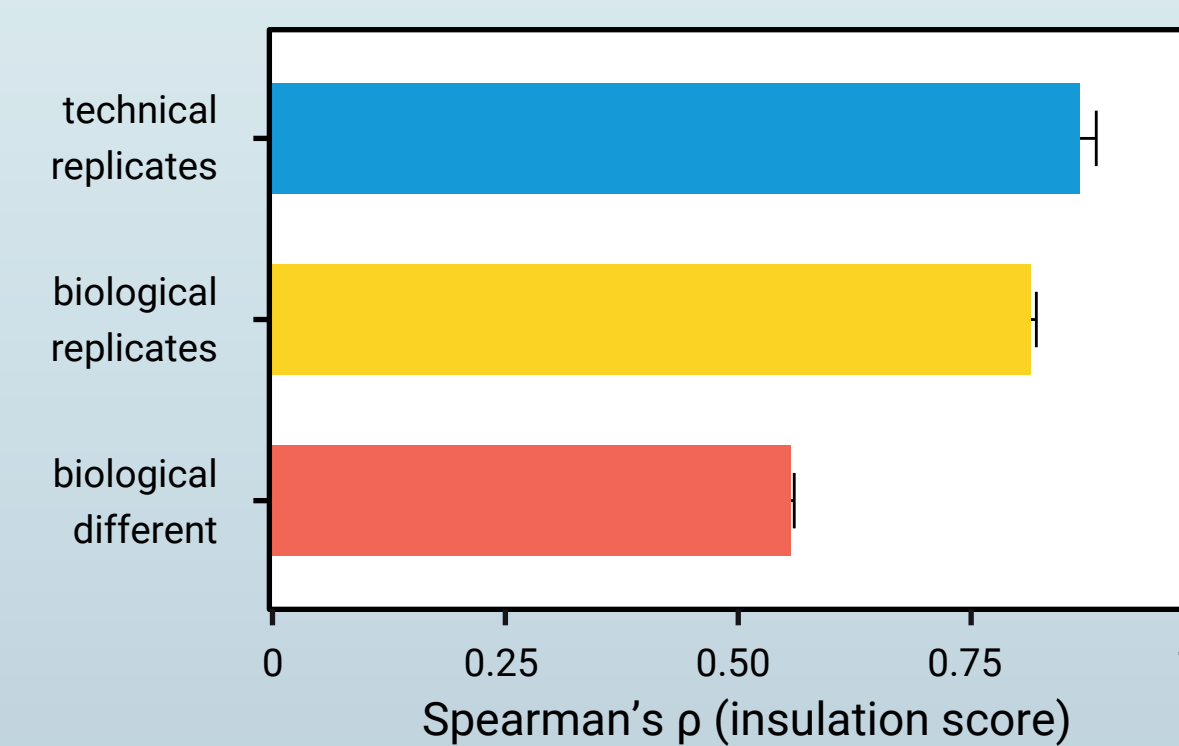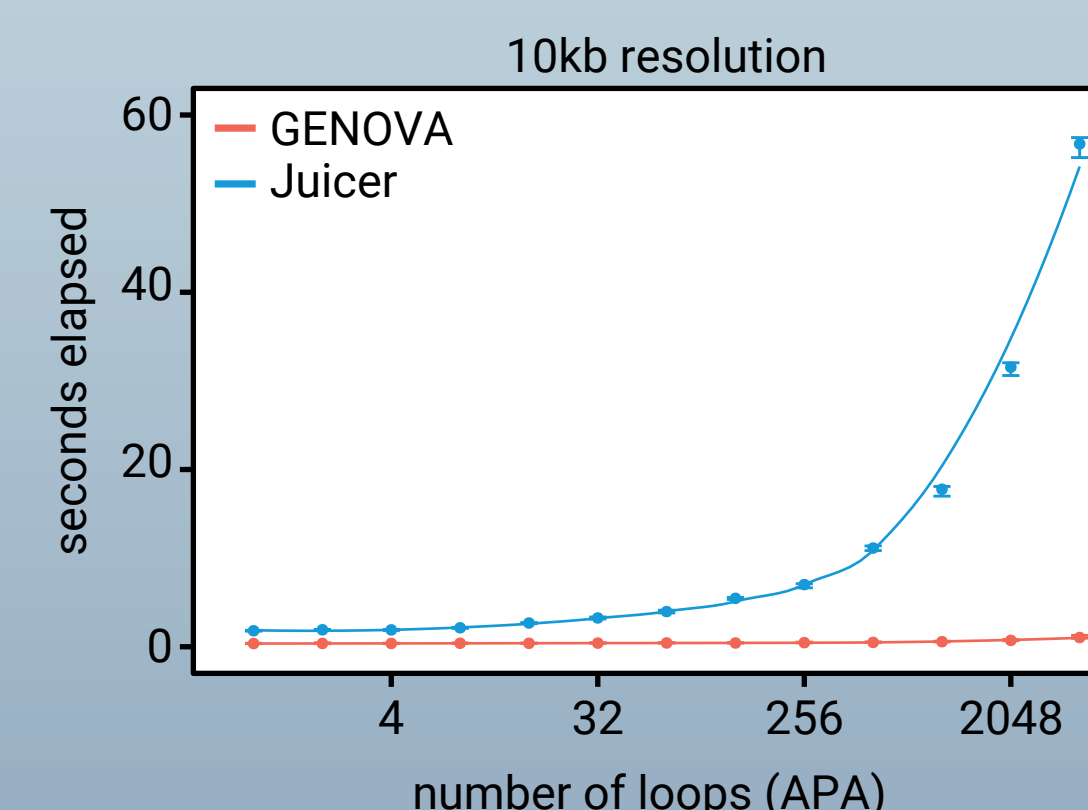Compare experiments by dual triangles or by differences

NETHERLANDS CANCER INSTITUTE
ANTONI VAN LEEUWENHOEK

Oncode Institute