

Omni-C Data Analysis

Michael Landi & Andreas Gisel

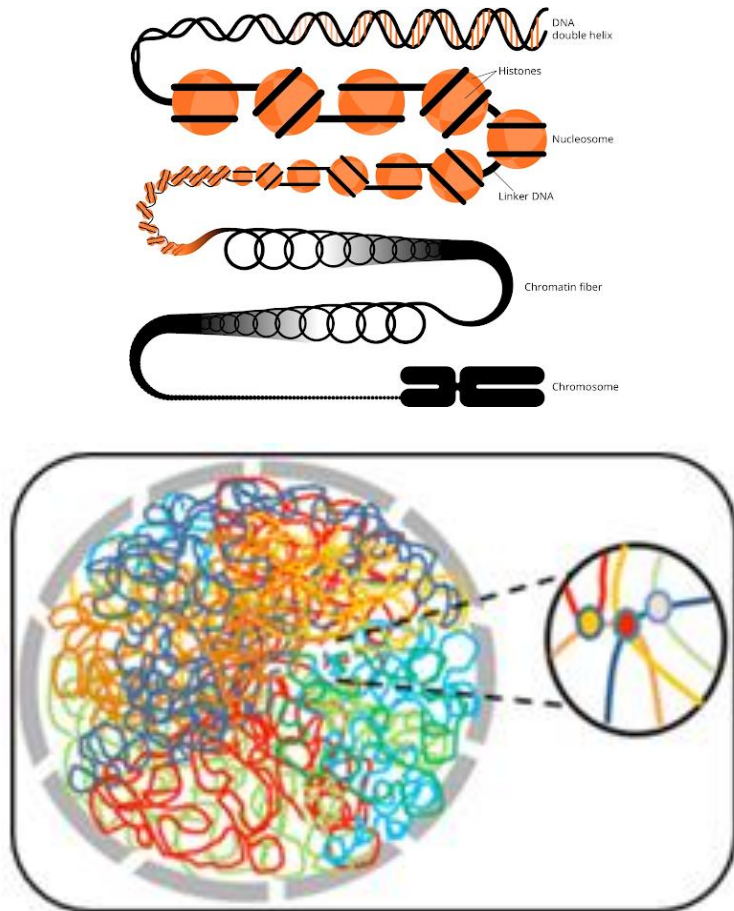
IITA- Bioinformatics
Kenya & Nigeria



African Star Apple Assembly workshop 2 – 6 June 2025

Omni-C Data Analysis

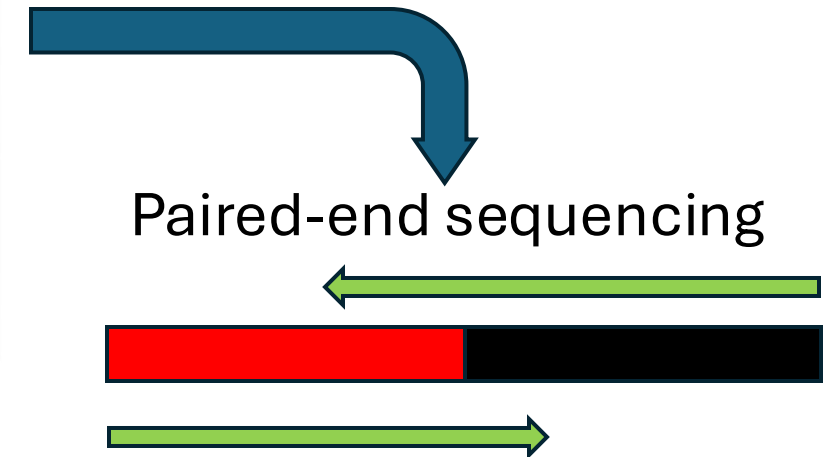
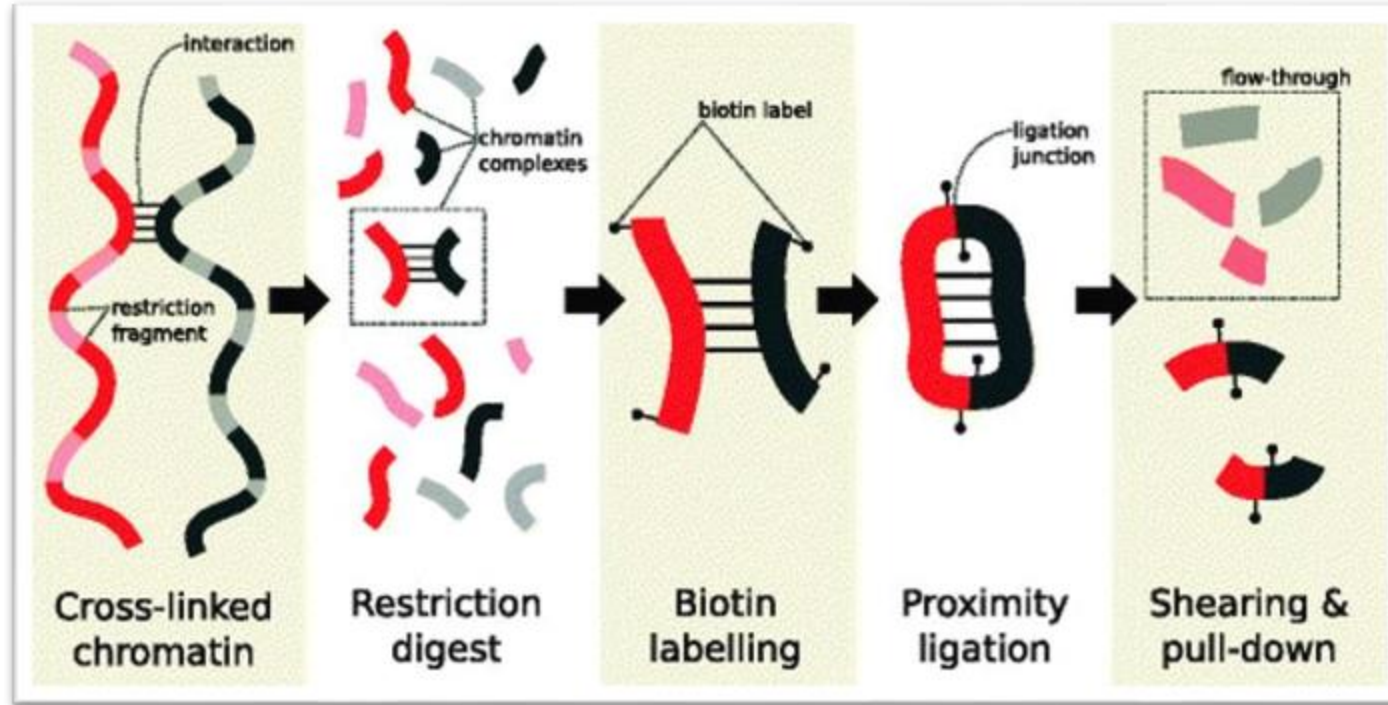
Omni-C is a high-throughput genomic and epigenomic technique to capture chromatin conformation.



Hi-C comprehensively detects genome-wide chromatin interactions in the cell nucleus.

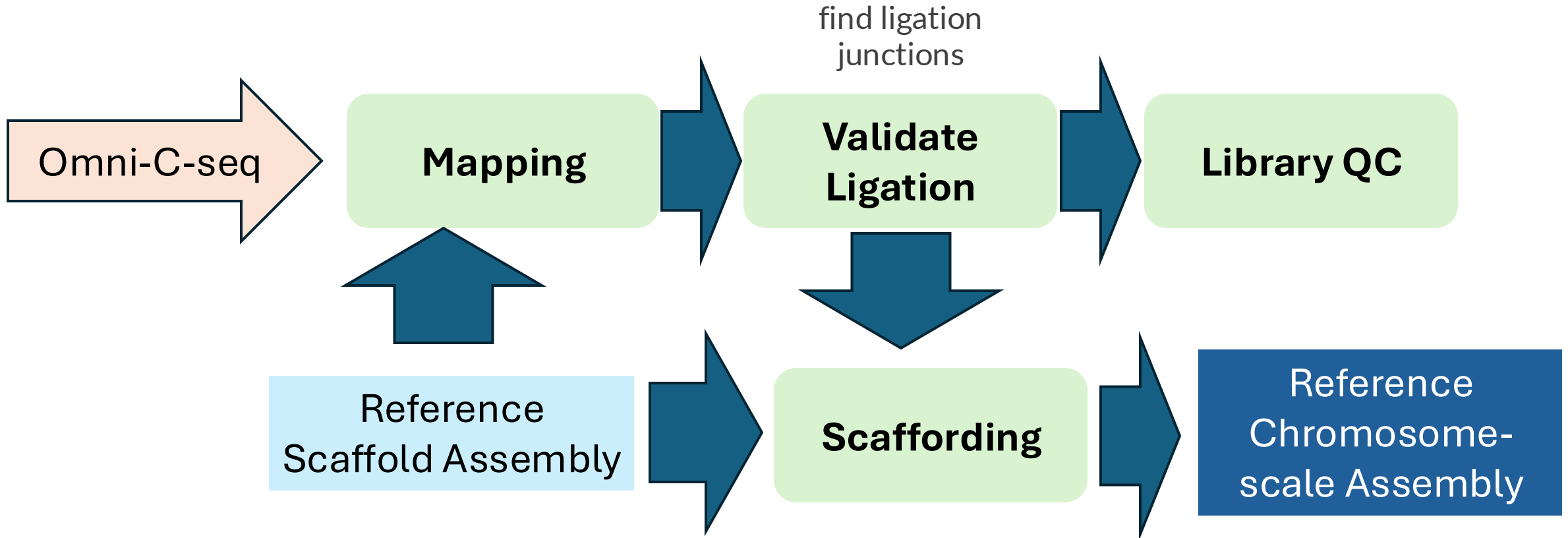
Hi-C measures the frequency (as an average over a cell population) at which two DNA fragments physically associate in 3D space, linking chromosomal structure directly to the genomic sequence.

Omni-C Data Analysis

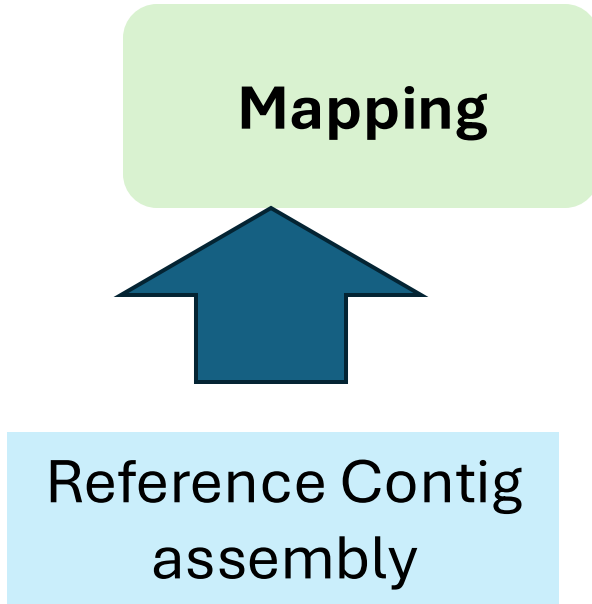


The Dovetail™ Omni-C™ library uses a sequence-independent endonuclease (compared to Hi-C, using restriction enzymes) for chromatin digestion prior to proximity ligation and library generation.

Omni-C Data Analysis



Omni-C Data Analysis



For mapping, we need a specific reference index

We use bwa (<https://bio-bwa.sourceforge.net/>)

We need to clean the OmniC data

We need to run the mapper

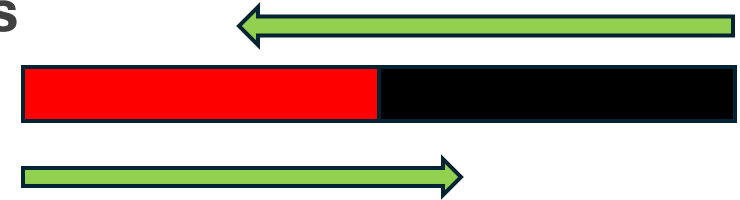
Omni-C Data Analysis

find ligation
junctions

Validate Ligation

Recording valid ligation events

- pairtools parse



Sorting the pairsam file

- pairtools sort

Removig PCR duplicates

- pairtools dedup

Generating .pairs and bam files

- pairtools split

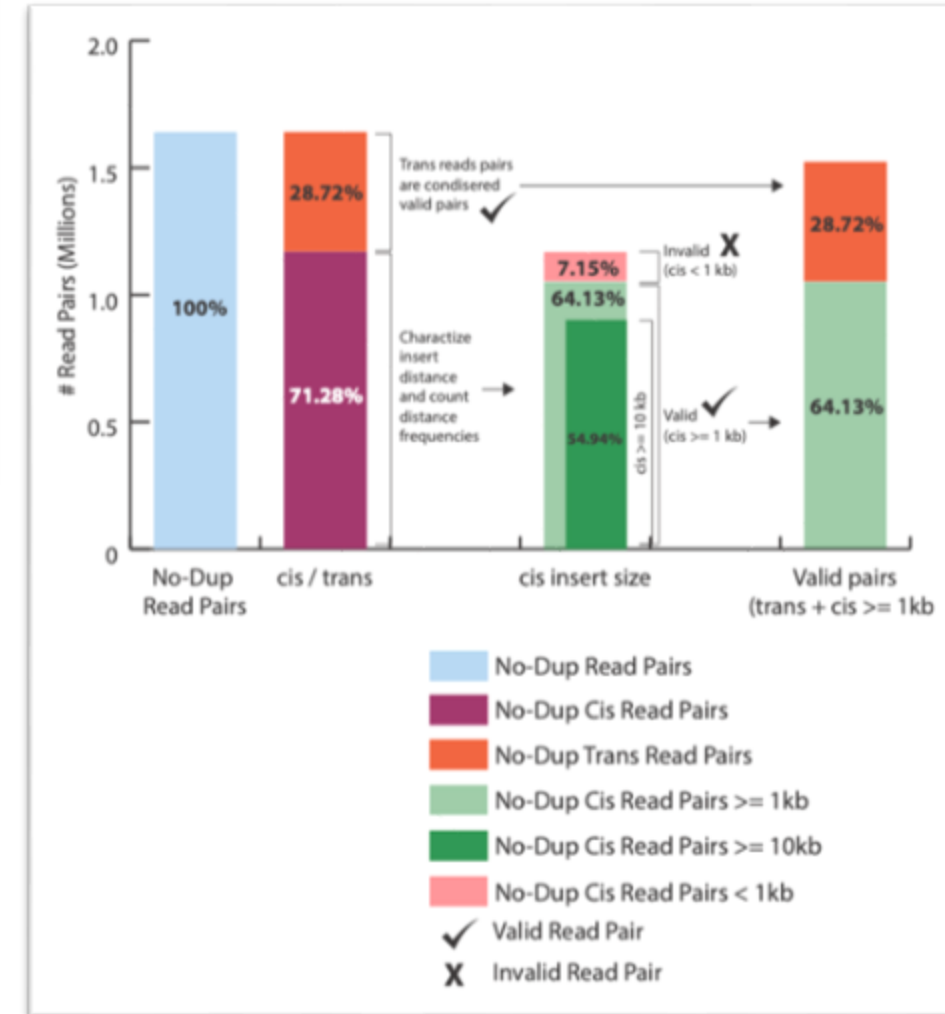
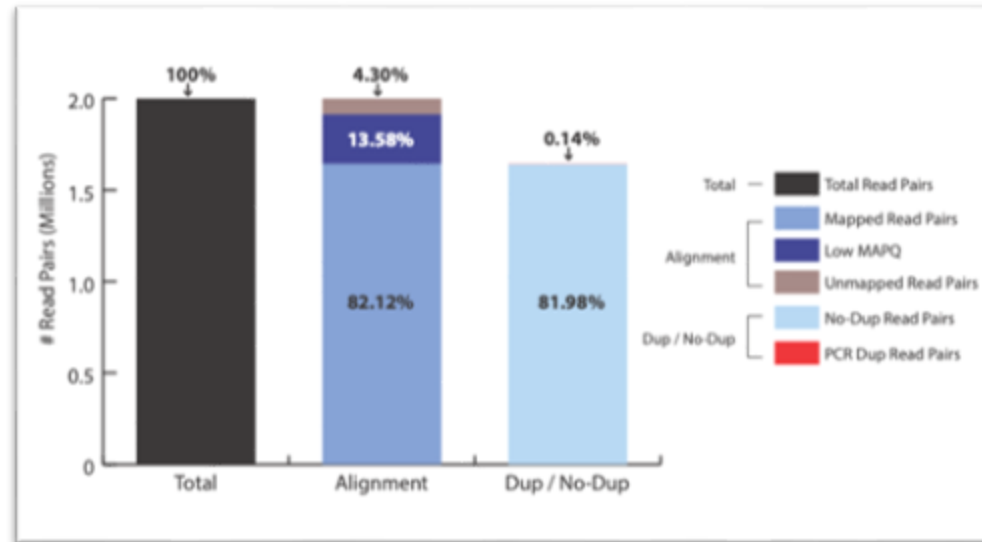
Generating the final bam file

- samtools sort

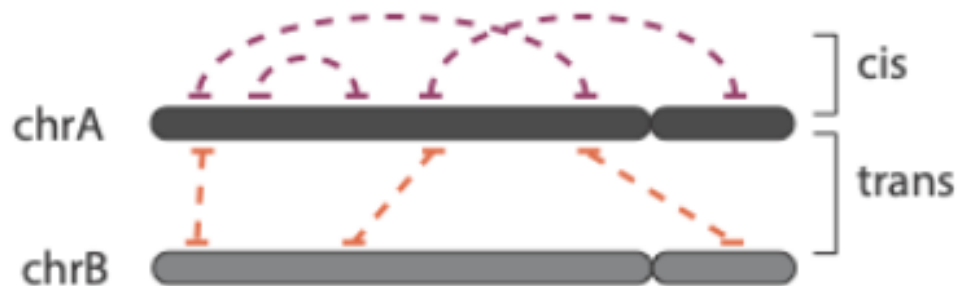
Scaffolding

Omni-C Data Analysis

Library QC



Classify cis or trans



Omni-C Data Analysis

Scaffolding

YaHS: yet another Hi-C scaffolding tool

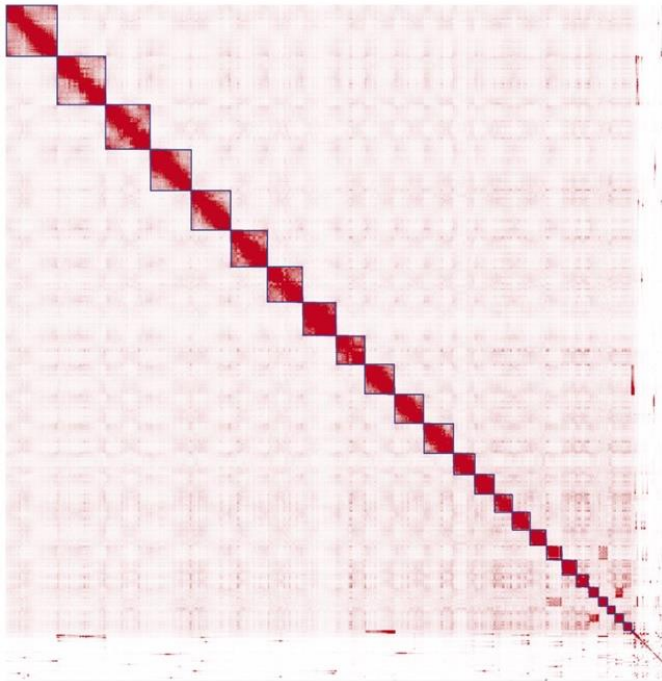
Zhou et al 2022, <https://doi.org/10.1093/bioinformatics/btac808>

YaHS scaffolding pipelines:

- map Hi-C reads to input contigs,
- break contigs where necessary to correct assembly errors,
- build a contact matrix,
- construct and prune a scaffolding graph
- output scaffolds.



Reference
Chromosome-
scale Assembly



Omni-C Data Analysis