

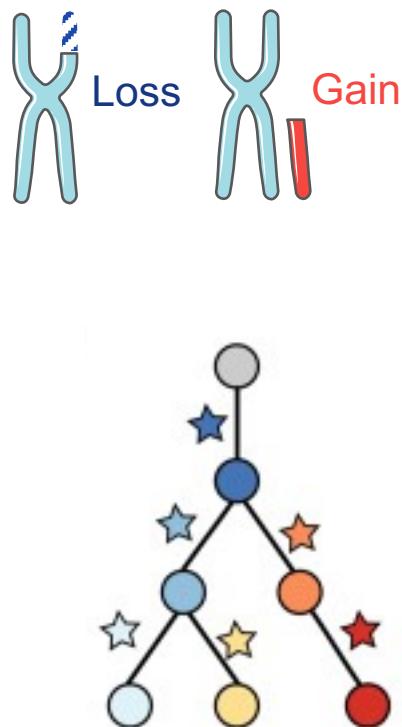
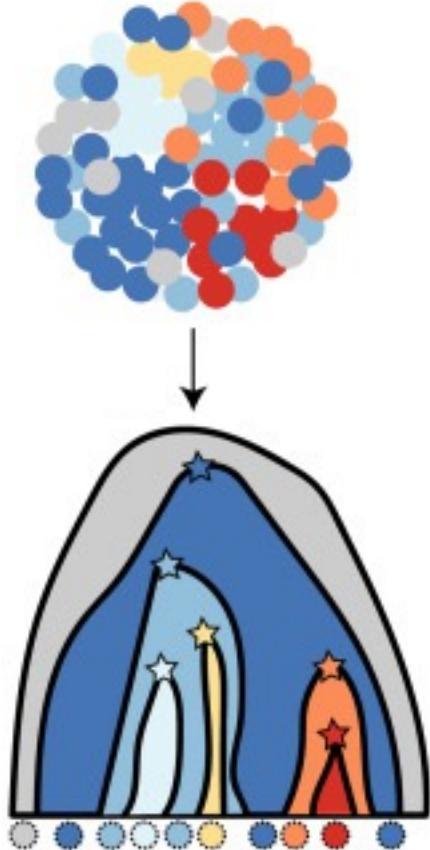
muscadet: An R Package for Detecting Somatic Copy Number Alterations from Single-Cell Multiomics

2025.11.25 · Marie Denoulet

CRCI2NA · Team 8 ICAGEN

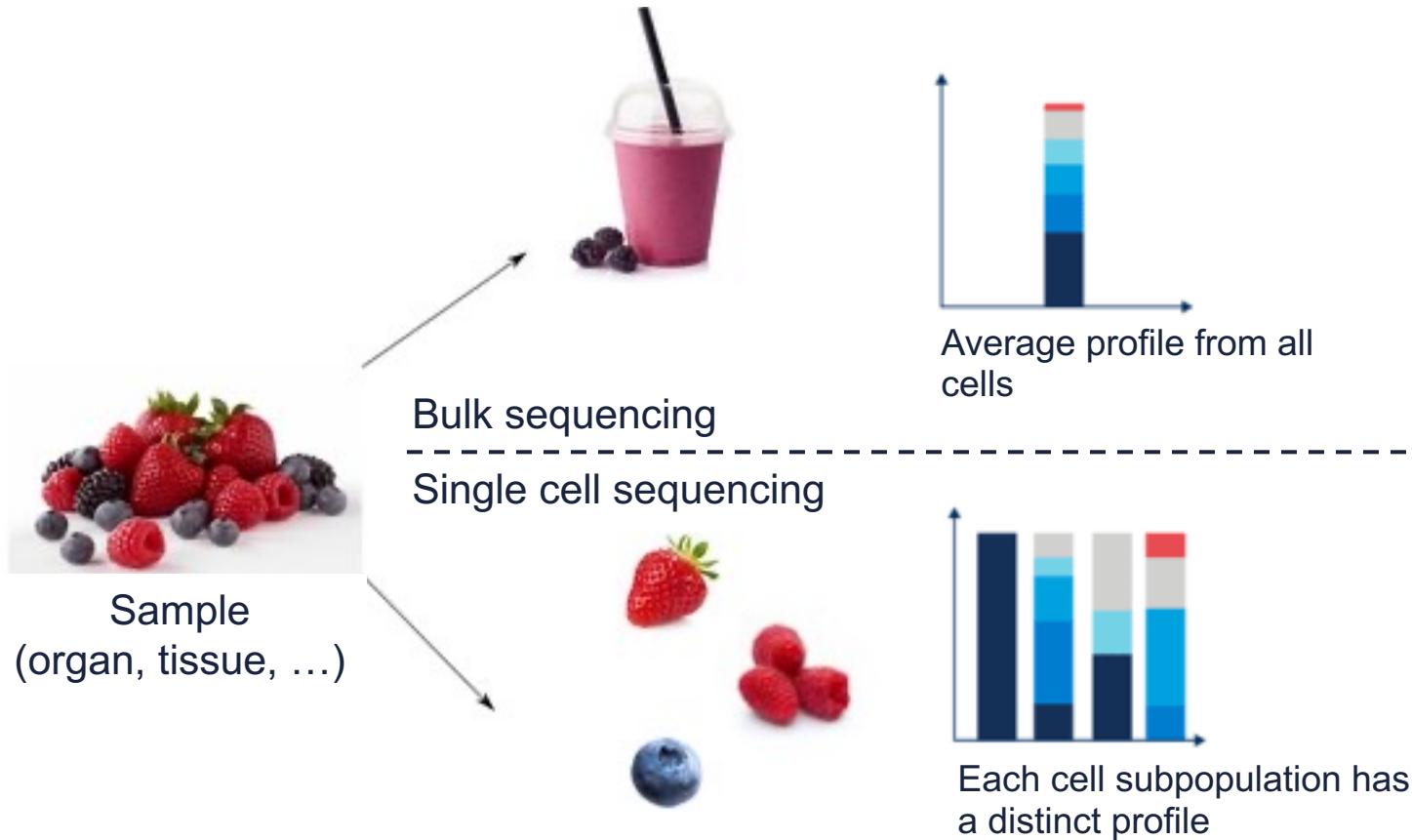


Cancer genomics: why study Copy Number Alterations?



- Cancer arises through the **accumulation of genomic alterations** that disrupt cellular regulation.
- **Copy number alterations (CNAs)** are large-scale genomic changes **gains or losses** of DNA segments from several genes to entire chromosome arms.
- CNAs can:
 - Amplify oncogenes
 - Delete tumor suppressors
 - Create dosage imbalance affecting hundreds of genes
- These events drive tumor initiation, progression, drug resistance, and metastasis.

From bulk to single-cell genomics



Bulk sequencing

- Average signal of cells
- Copy-number estimation on total population

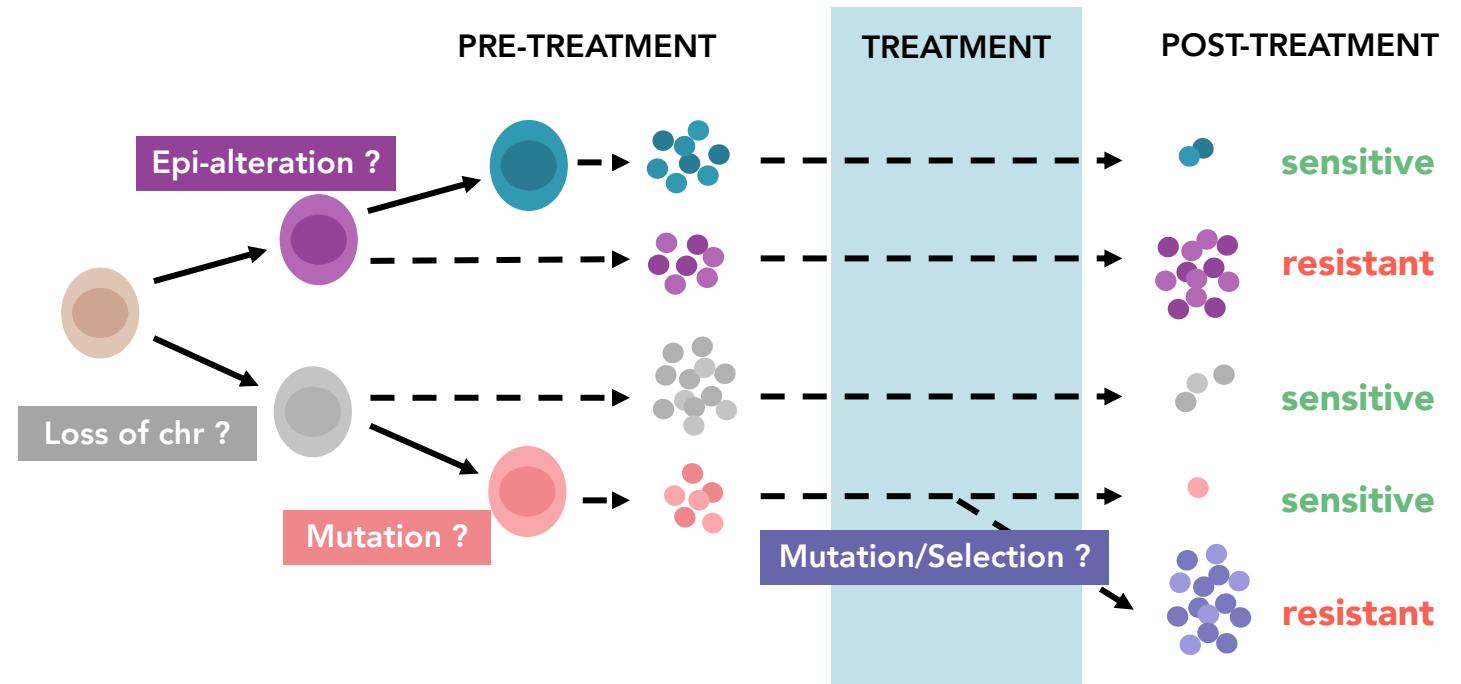
Single-cell sequencing

- Profiles individual copy-number genomes
- Heterogeneity & subclonality: reconstruction of tumor evolution, clonal architecture

Why detecting CNAs at single-cell resolution matters

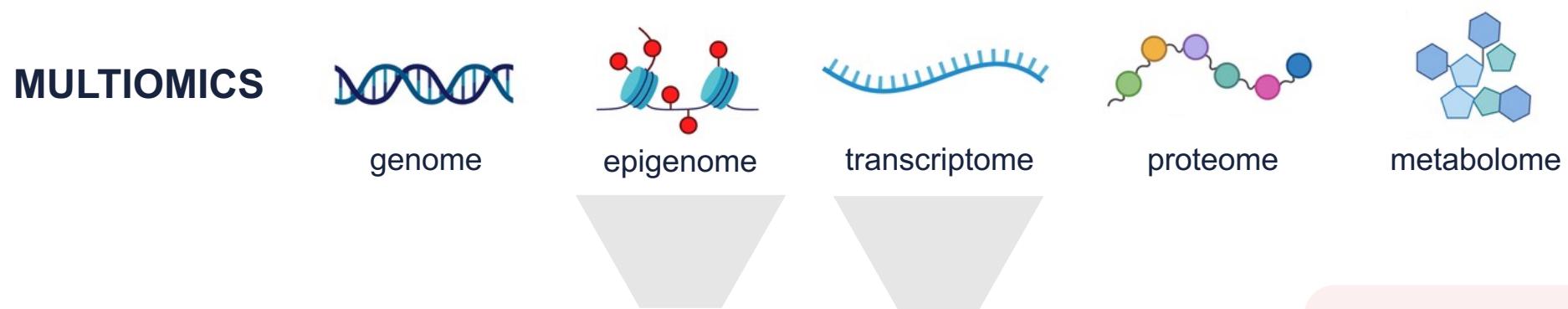
- Identify **genetic subclones** sharing CNA profiles
- Reconstruct temporal CNA acquisition and **tumor evolution**
- Relate subclones and CNAs to transcriptional/epigenetic changes and treatment responses
- Identify minor resistant clones before therapy

Integrated (epi)genomic evolution of tumor cells



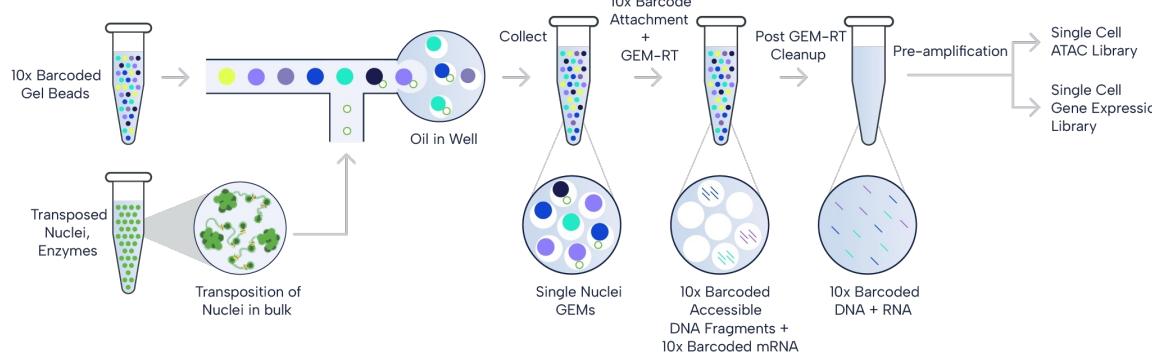
Single-cell multiomics

Different single-cell sequencing technologies capture complementary layers of cellular information



Example

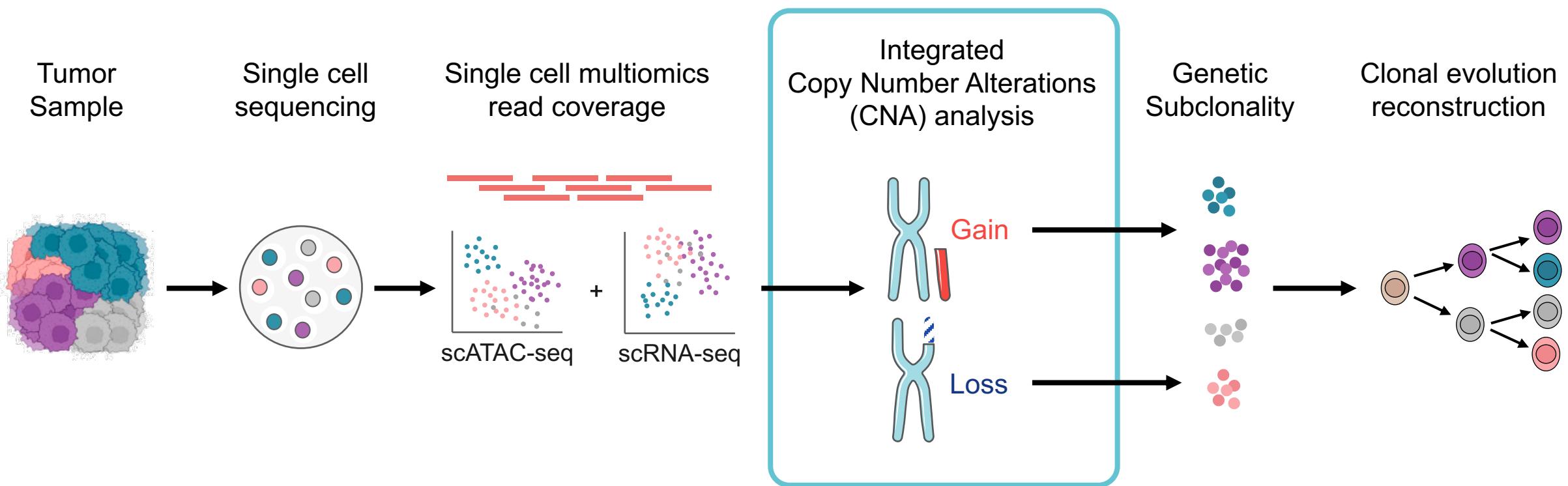
10X Multiome: scATAC-seq + scRNA-seq



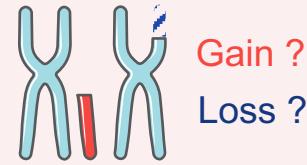
Challenges:

- Sparsity
- Noise
- Heterogeneous modalities
- Missing modalities
- Scalability

Framework to infer tumor evolution



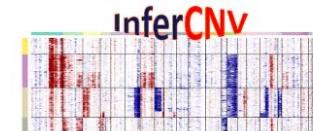
Copy Number profiling



Existing tools limited to **ONE** single-cell omic

scRNA-seq only

✓ inferCNV



✓ Numbat



scATAC-seq only

✓ Alleloscope

✓ AtaCNA

✓ epiAneufinder

Multomics

e.g. joint scRNA-seq & scATAC-seq

✗ None

► Leverage integrated information from multiomics datasets ?

Introducing muscadet: goals and scope



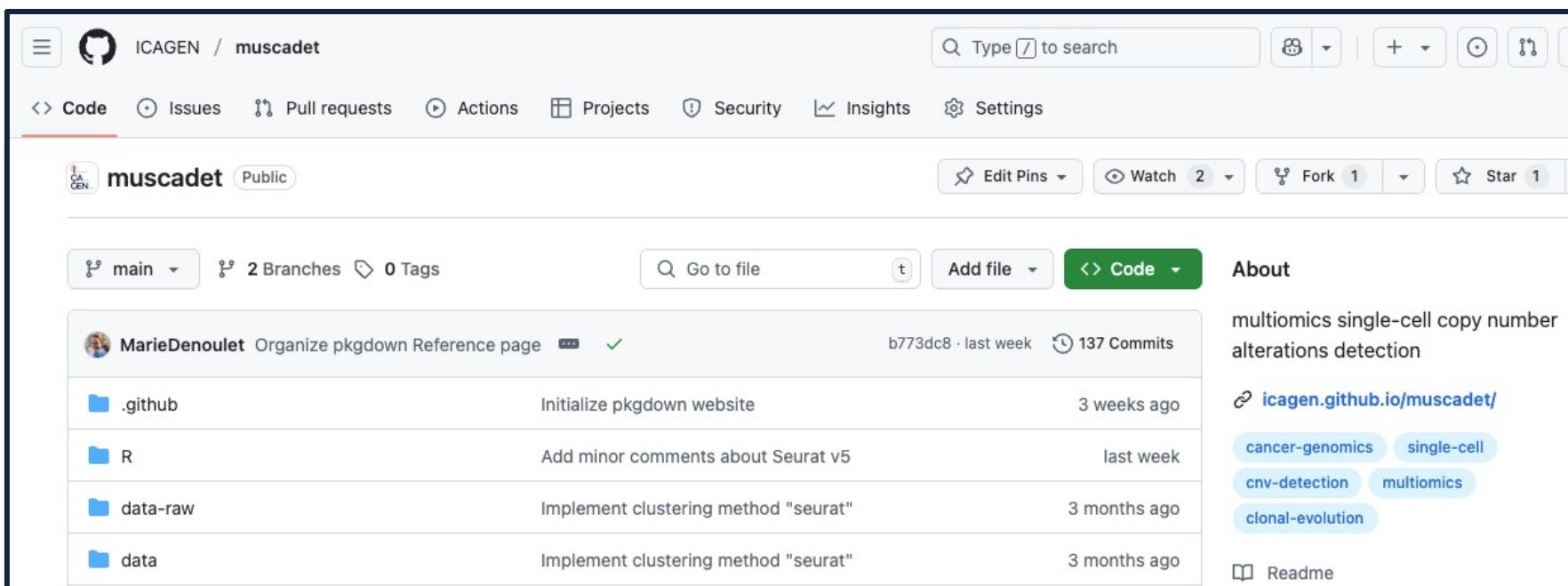
muscadet: multiomics single-cell copy-number alterations detection

1. Integrate **multiple omic** signals at single-cell level
2. Cluster cells into **subclones**
3. Enable allele-specific **CNA inference**
4. Provide a **modular R framework** for reproducible single-cell CNA analysis

Introducing muscadet: goals and scope



muscadet is functional, stable, and available on GitHub  [ICAGEN/muscadet](https://github.com/ICAGEN/muscadet)



The screenshot shows the GitHub repository page for 'muscadet'. The repository is public and has 2 branches and 0 tags. The main branch is active. The repository was last updated 3 months ago. The 'About' section describes it as a multiomics single-cell copy number alterations detection tool. It uses pkgdown for documentation and Seurat v5 for clustering. The repository is associated with the URL icagen.github.io/muscadet/ and has labels for cancer-genomics, single-cell, cnv-detection, multiomics, and clonal-evolution.

Commit	Author	Message	Date
b773dc8 · last week	MarieDenoulet	Organize pkgdown Reference page	137 Commits
		Initialize pkgdown website	3 weeks ago
		Add minor comments about Seurat v5	last week
		Implement clustering method "seurat"	3 months ago
		Implement clustering method "seurat"	3 months ago

muscadet 0.1.3

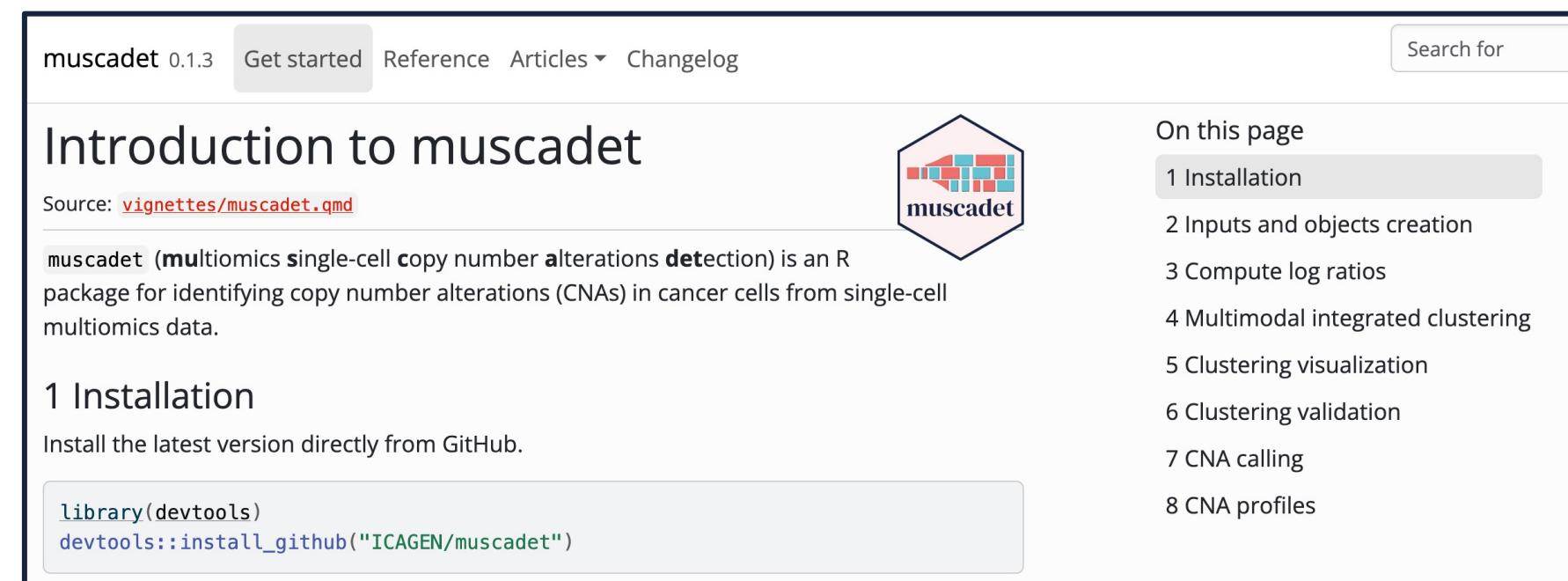
 R-CMD-check.yaml passing

Documentation website at <https://icagen.github.io/muscadet/>

Introducing muscadet: goals and scope



muscadet is functional, stable, and available on GitHub  [ICAGEN/muscadet](https://github.com/ICAGEN/muscadet)



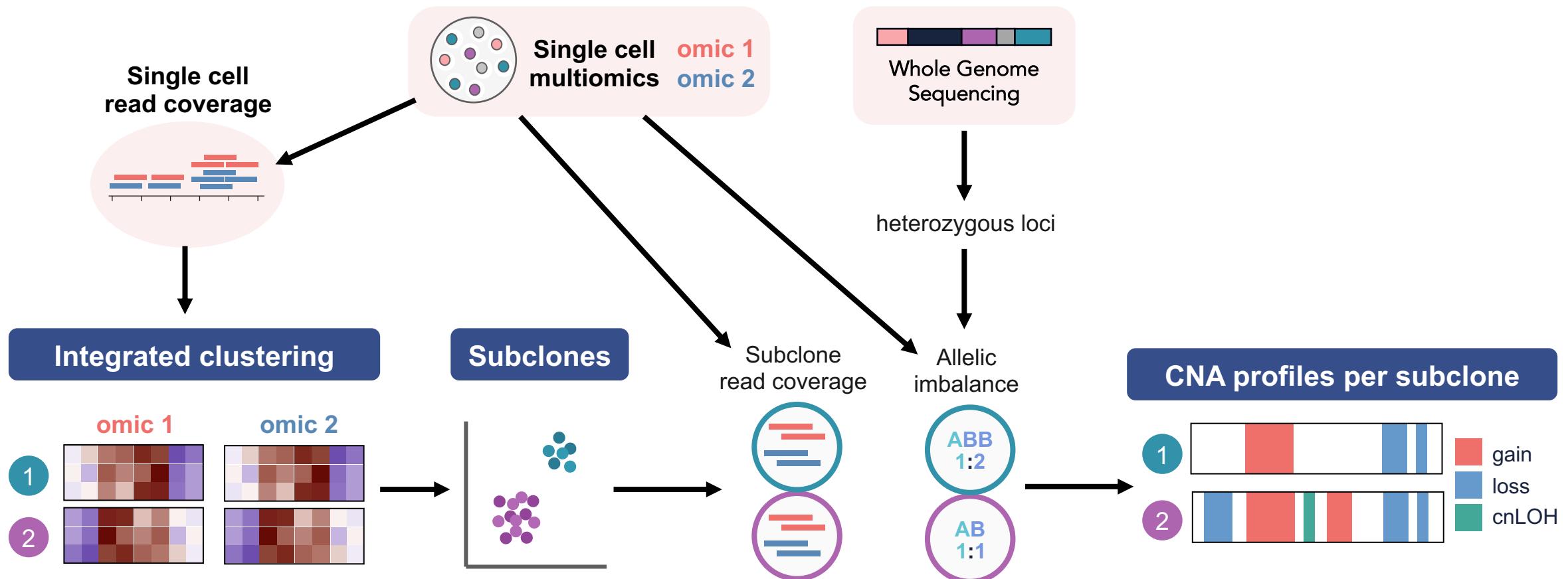
The screenshot shows the muscadet documentation website. At the top, there's a navigation bar with links for 'muscadet 0.1.3', 'Get started', 'Reference', 'Articles ▾', 'Changelog', and a search bar. The main content area features a large heading 'Introduction to muscadet' with a subtext 'Source: vignettes/muscadet.qmd'. Below this, a paragraph describes the package: 'muscadet (multiomics single-cell copy number alterations detection) is an R package for identifying copy number alterations (CNAs) in cancer cells from single-cell multiomics data.' To the right of the main content, there's a sidebar titled 'On this page' containing a numbered list of topics: 1 Installation, 2 Inputs and objects creation, 3 Compute log ratios, 4 Multimodal integrated clustering, 5 Clustering visualization, 6 Clustering validation, 7 CNA calling, and 8 CNA profiles. A small 'muscadet' logo is also present in the sidebar.

muscadet 0.1.3

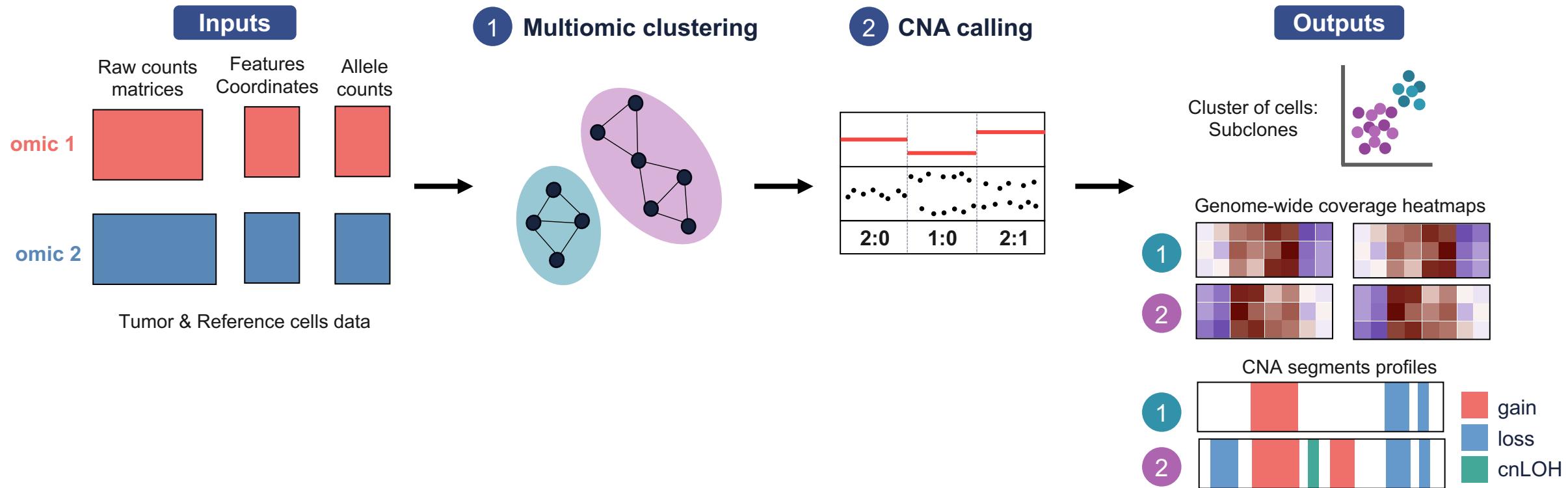


Documentation website at <https://icagen.github.io/muscadet/>

Analysis workflow



Analysis workflow

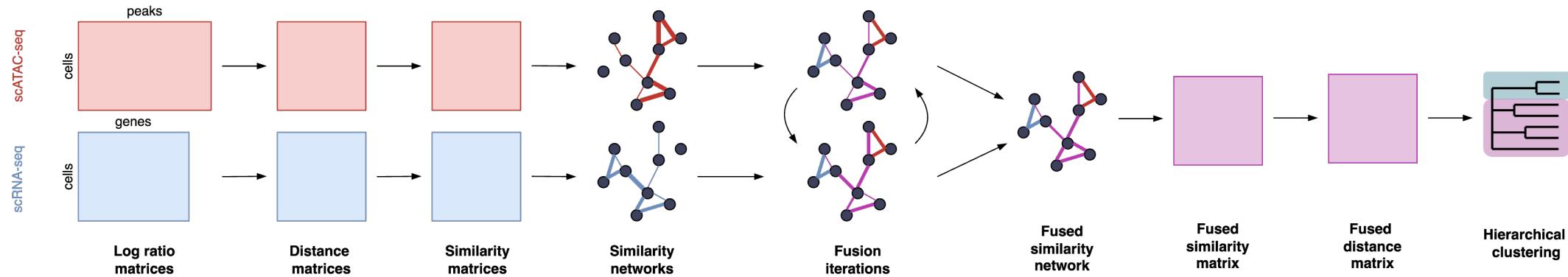


Analysis workflow

1 Multiomic clustering

Two available methods:

1. Integration by **Similarity Network Fusion (SNF)** followed by **hierarchical clustering**

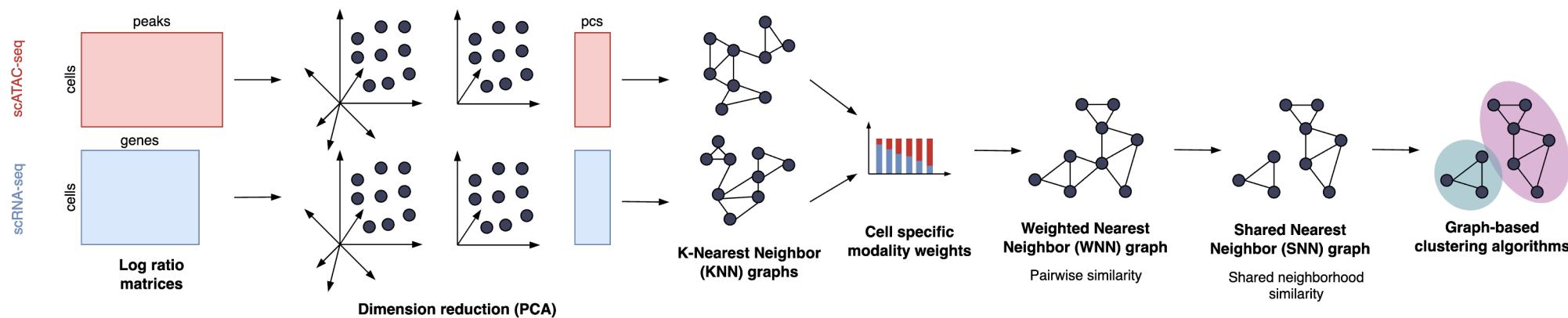


Analysis workflow

1 Multiomic clustering

Two available methods:

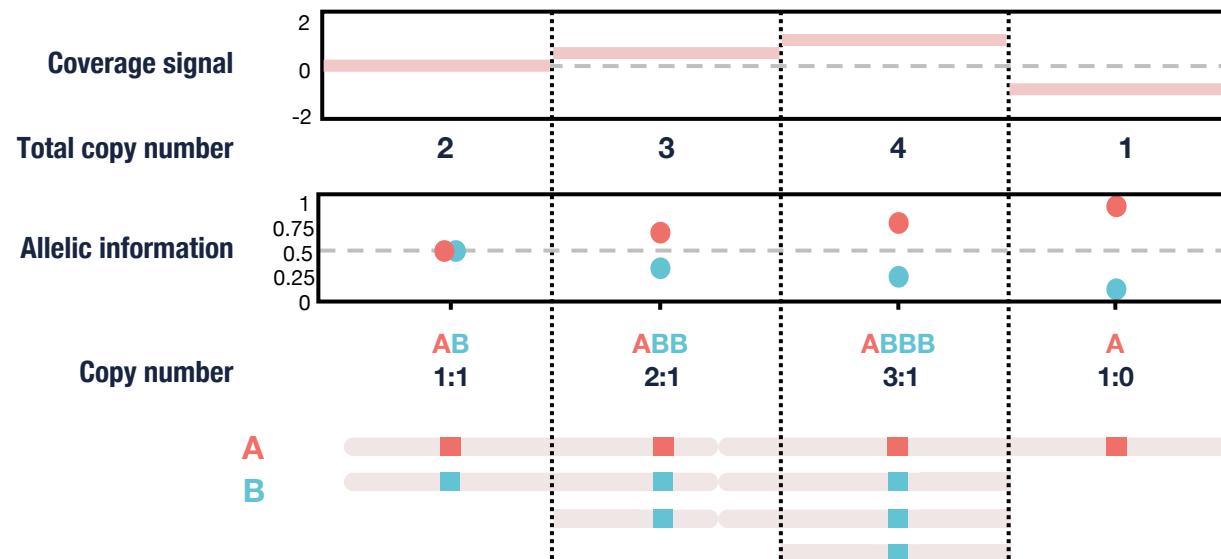
2. Seurat-based : **Weighted Nearest Neighbor (WNN) graph** construction followed by **graph-based clustering**



Analysis workflow

2 CNA calling

Adaptation FACETS method for CNA inference from bulk sequencing data



FACETS: allele-specific copy number and clonal heterogeneity analysis tool for high-throughput DNA sequencing

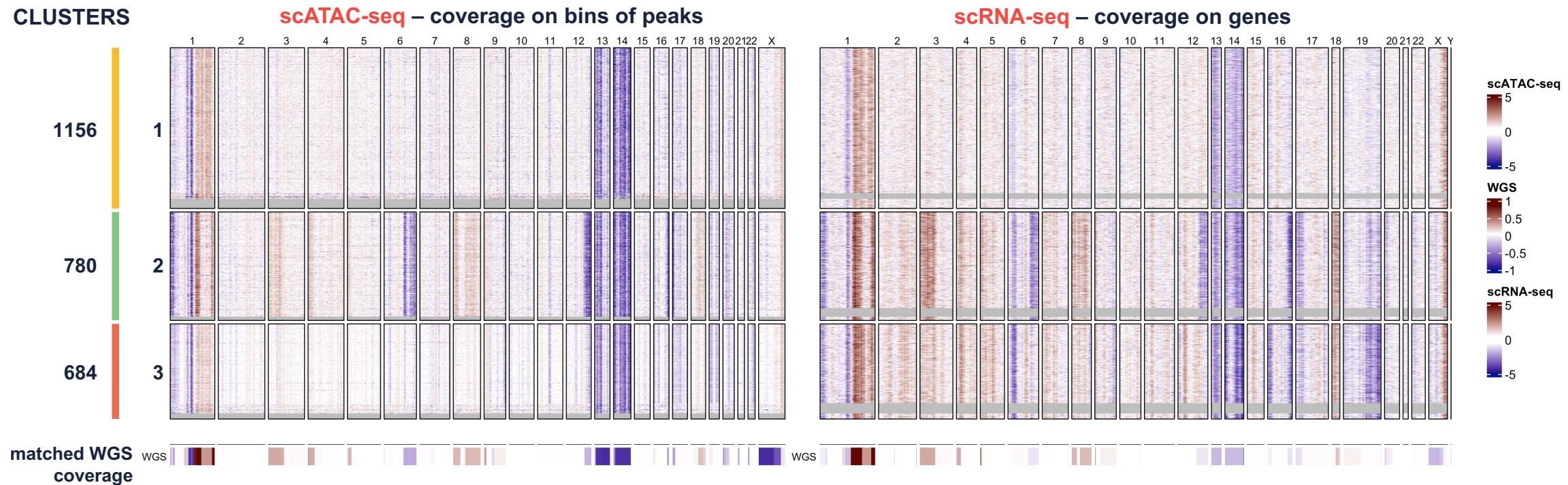
Ronglai Shen^{*} and Venkatraman E. Seshan^{*}

2016

- Segmentation of the genome based on coverage and allelic signals
- Clustering of segments with similar signals
- Identification of diploid coverage level
- Expectation-Maximization algorithm fit a mixture of genotypes to explain the observed signals
- Determination of consensus segments across clusters

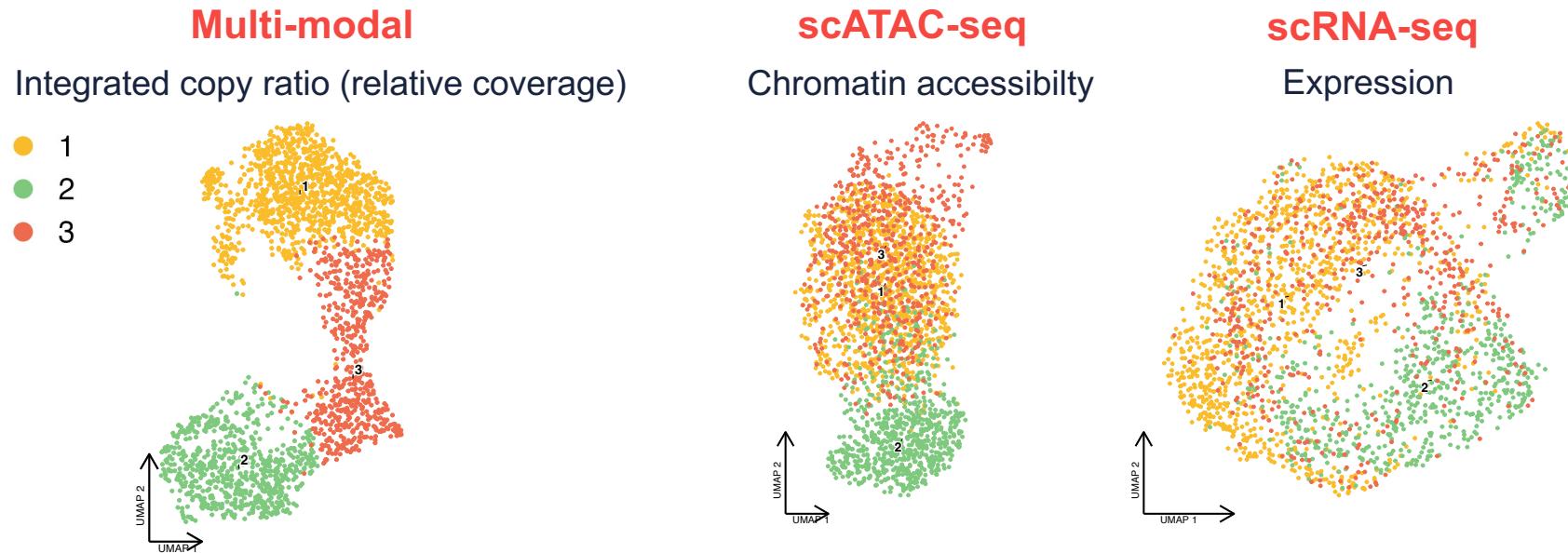
Case study on a Multiple Myeloma sample

1 Multiomic clustering



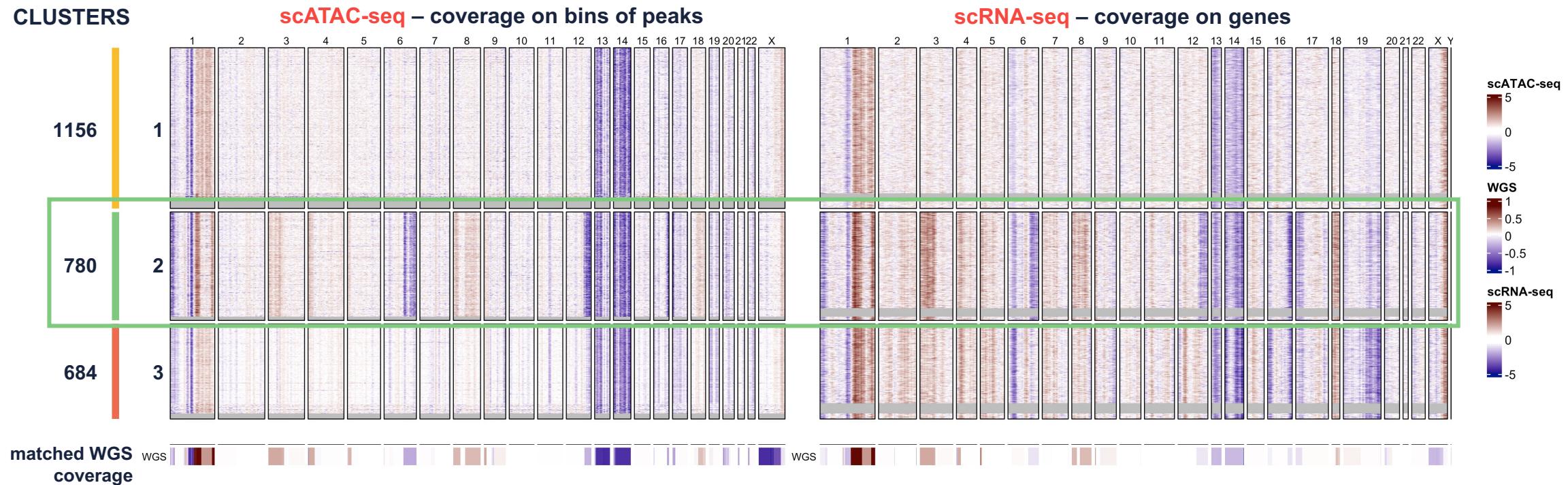
Case study on a Multiple Myeloma sample

1 Multiomic clustering



Case study on a Multiple Myeloma sample

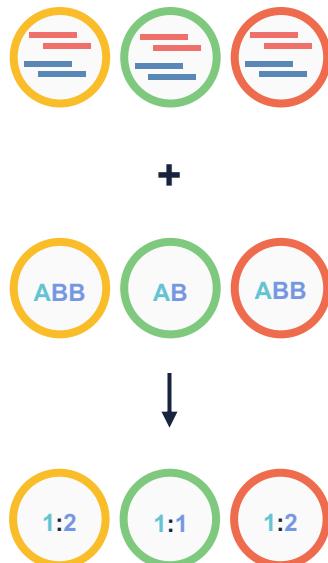
1 Multiomic clustering



Case study on a Multiple Myeloma sample

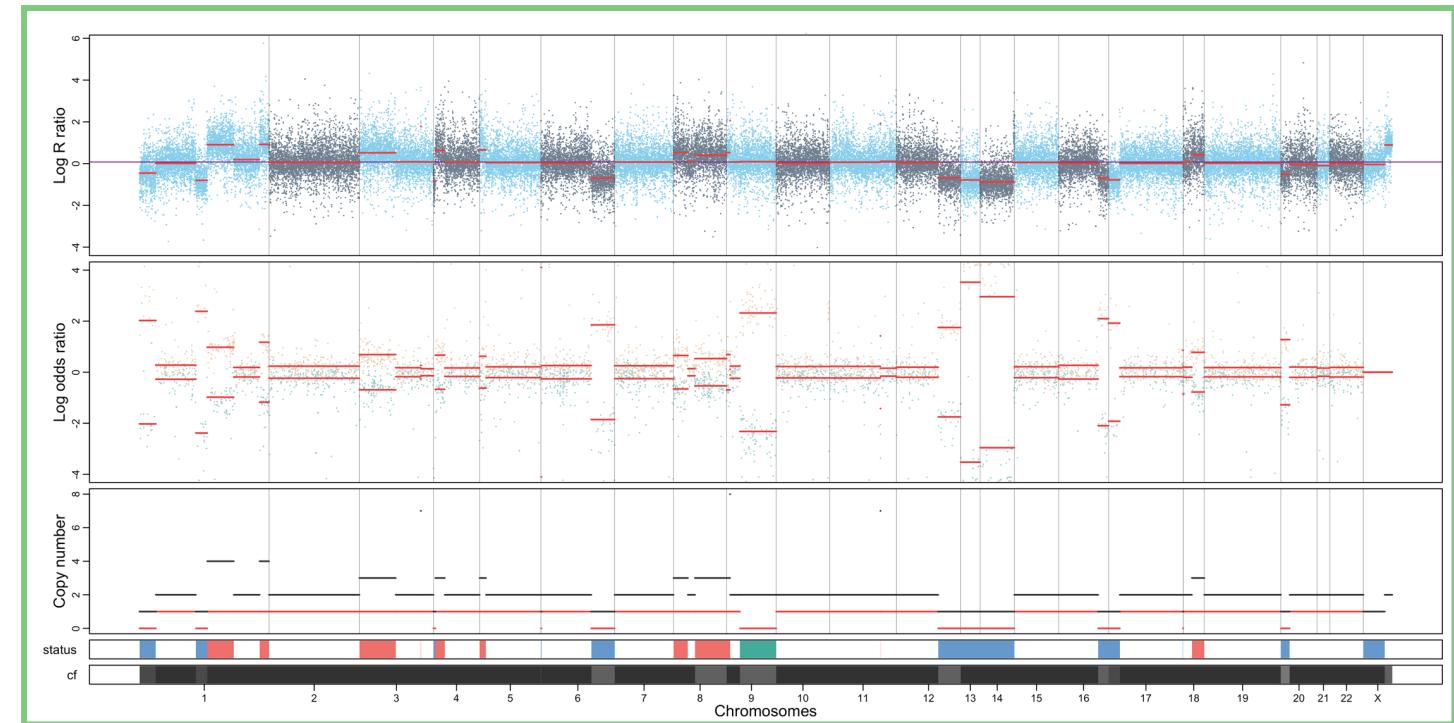
2 CNA calling

Cluster-level analysis



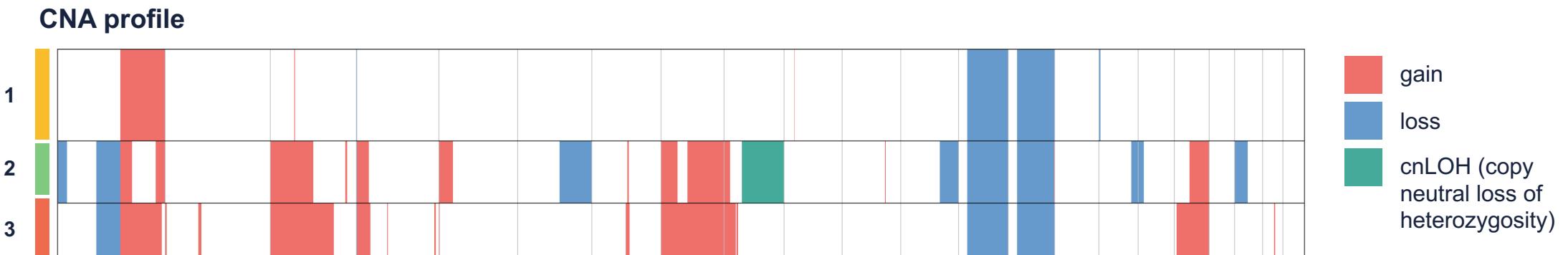
Relative read coverage
+
Allelic imbalance
↓
Copy number inference

Cluster 2 Copy number profile

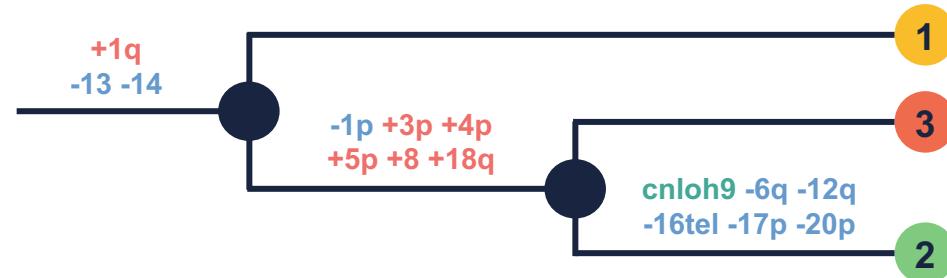


Case study on a Multiple Myeloma sample

2 CNA calling



Inferred clonal evolution



Performance and limitations



Strengths

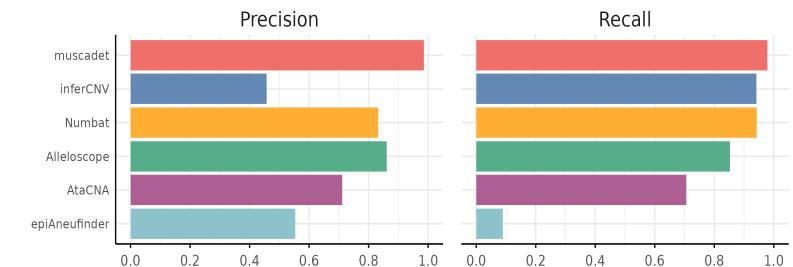
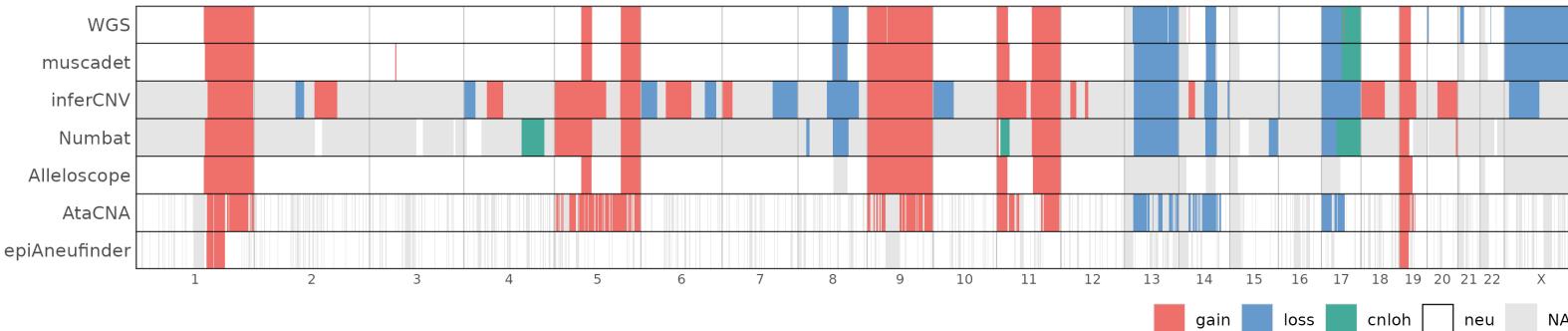
- **Multiomic integration:** Integrates multiple modalities to amplify signal and compensate for sparsity
- **Modular:** adapted for both mono-omic and multi-omic datasets
- **Flexible:** user-driven cluster definition with clustering partition exploration and validation
- **Scalable to additional omics:** designed to support other omics types and more than two modalities

Limitations

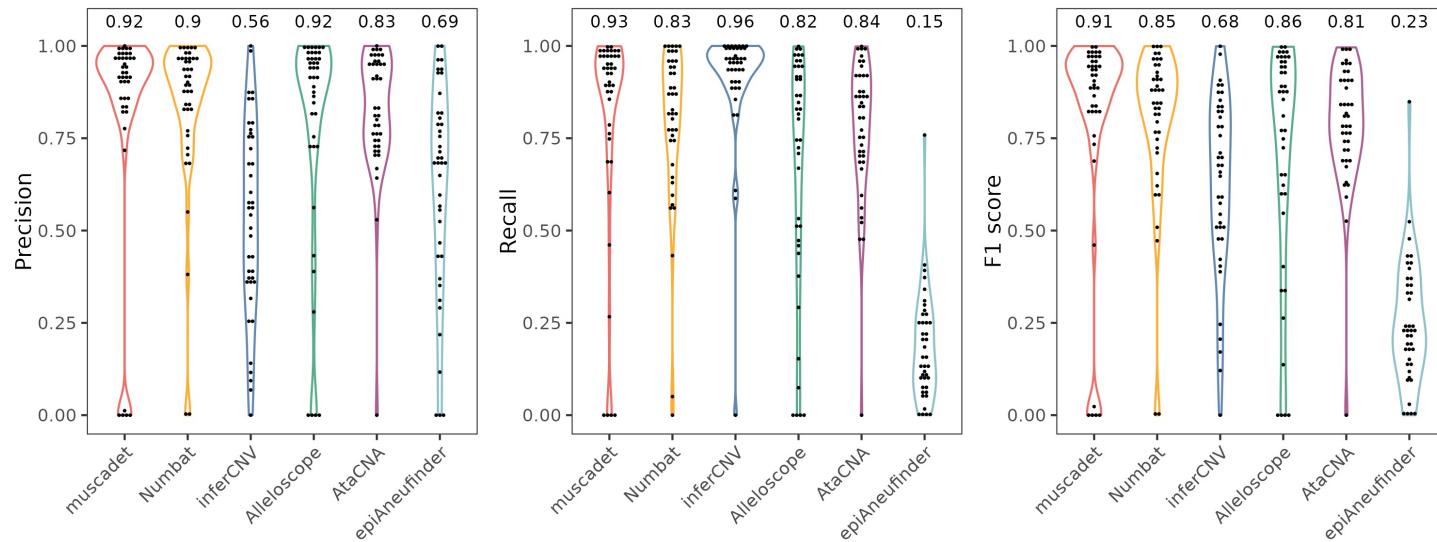
- **Low-cell-count clusters:** Small clusters with sparse data produce unstable or noisy CNA calling
- **Reference requirement:** Requires reference cells data for a diploid base level
- **Preparation of inputs:** Allele-specific CNA inference requires allele counts preprocessing

Benchmark against other CNA callers

Per-sample comparisons of CNA segments to the WGS copy number estimation



Benchmark metrics across samples compared to WGS



* preliminary results

Acknowledgments

CRCI²NA /CAGEN Team



CRCI²NA
CANCER & IMMUNOLOGY

Stéphane Minvielle
Florence Magrangeas
Philippe Moreau
Loïc Campion
Victor Bessonneau-Gaborit
Catherine Guérin-Charbonnel
Céline Chevalier
Jennifer Derrien
Louis Sanglier

Eric Letouzé
Nils Giordano
Mia Cherkaoui
Magali Devic
Elise Douillard



Curie Institute DEpiC Team



Céline Vallot
Mathias Schwartz
Marthe Laisne
Justine Marsolier
Melissa Saichi
Grégoire Jouault
Adeline Durand
Yuna Landais
Adeline Lablanche
Delia Dupre