

Centre for
Heart Lung Innovation
UBC and St. Paul's Hospital



Department of
Anesthesiology, Pharmacology
& Therapeutics
Faculty of Medicine

MESSI: Multiomics Experiments with SyStematic Interrogation

A pipeline for benchmarking multi-omics integration methods on classification problems using Nextflow

Chunqing Tony Liang

Msc student in Bioinformatics

Supervisor: Dr. Amrit Singh

March 7, 2024

Multiomics data and analysis

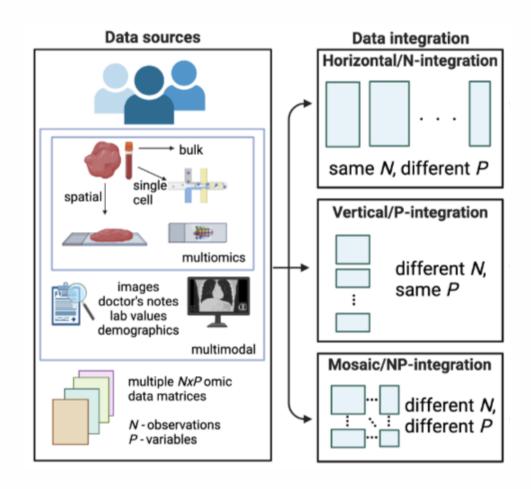


Figure 1: Overview of multiomics data and integration ¹

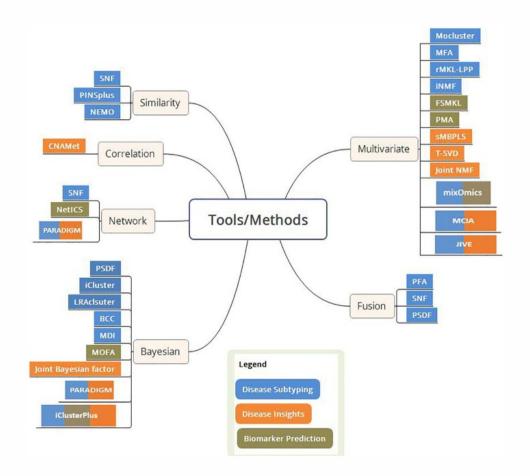
1. Shannon et al., *Manuscript Under Revision*, 2024.

\bigcirc I

Motivation

- Technological advancement and reduced costs -> studies with multiomics data
- Multiomics data -> bulk/single-cell/spatial data
- More generalized -> "multimodal data", other non-omics data
- Types of integrations: by sample (N), by omics (P), or both

Multiomics data integration methods



Question arises

- Many integration methods
 - -> Which to use, how to choose them?
- Reproducibility crisis
 - —> How to reproduce method and get reliable results?
- Existing benchmark studies are not 100% complete or allencompassing
 - —> Technical difficulty in implementation?

Figure 2: Overview of multi-omics data integration tools ¹

1. Indhupriya Subramanian et al., "Multi-Omics Data Integration, Interpretation, and Its Application," *Bioinformatics and Biology Insights* 14 (2020): 1177932219899051.

Overview of MESSI

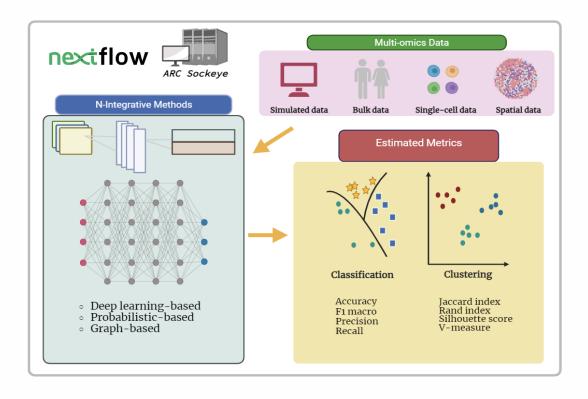


Figure 3: Overview of MESSI workflow

- Generalizable for more integration methods
- Data transformable for Python or R based methods
- Reproducible workflow, results/findings

- Multiple data types (bulk, sc, spatial)
- Help to develop new integration method
- Compare large set of independent data at once

Acknowledgements

- Dr. Amrit Singh
- Dr. Young Woong Kim
- Dr. Maryam Ahmadzadeh
- Rishika Daswani
- Roy He
- Michael Yoon
- Jeffrey Tang
- Akshdeep Sandhu
- Yovindu Don
- Raam Sivakumar
- Prabhleen Sandhu
- Mingming Zhang
- Samuel Leung







