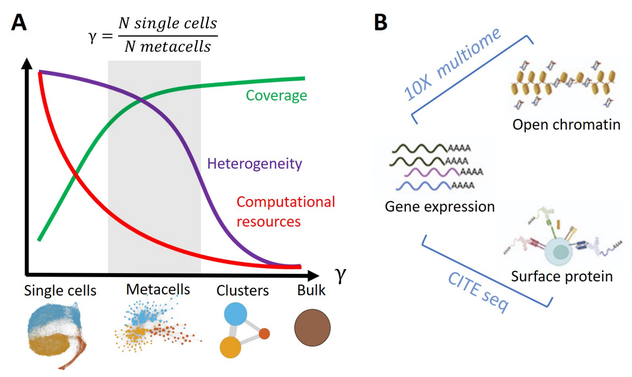
Reconciling single-cell multiomics modalities using metacells

Léonard Hérault[1,2]
leonard.herault@unil.ch
Aurélie Gabriel[1,2] **Mariia Bilous**[1,2] **David Gfeller**[1,2]
[1] *Department of Oncology, Ludwig Institute for Cancer research, University of Lausanne*
[2] *Swiss Institute of Bioinformatics*

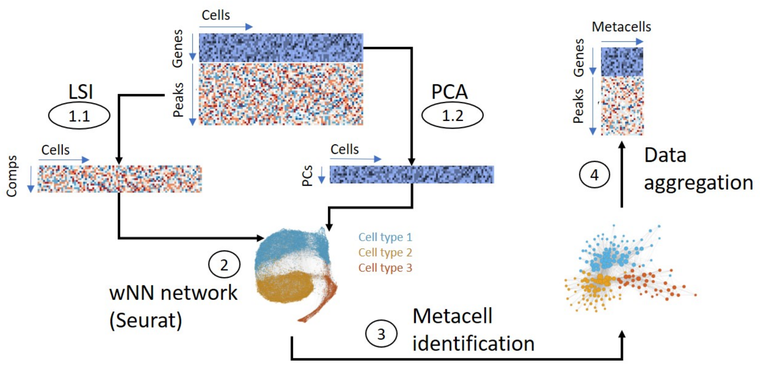
Introduction

The increased throughput of single-cell omics technologies enables researchers to study cell type-specific gene regulation at an unprecedented resolution. These promises depend on the development of computational methods to cope with both large size and the high sparsity of these data. Thus, we and others developed methods to identify metacells, disjoint and homogeneous groups of cells, in scRNA-seq data **Bilous et al. (2022)**. We are now developing a new version of our method **SuperCell** to identify metacells in single-cell multiomics (sc-multiomics) data that combine the measurement of different types of molecules (modalities) in the same single-cell.



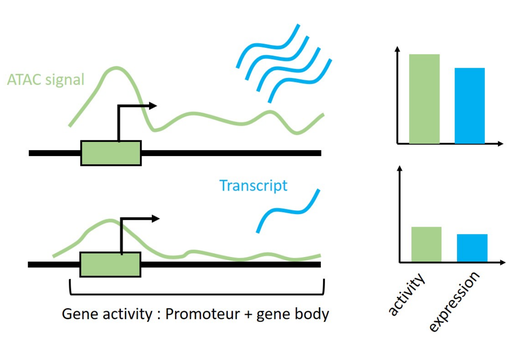
Methods

For sc-multiomics, Supercell used a weighted Nearest Neighbor (wNN) graph **(Hao et al. 2021)**), on which the metacell are identified. Single-cell raw data from each modality are aggregated per metacells that can be used for downstream analyses.



Datasets

- 10X multiome data of Peripheral Blood Mononuclear cells (PBMCs) : scRNA-seq & ATAC-seq (12'000 cells). We analysed the correlation between gene expression and gene activity.
- CITE-seq data of Bone marrow cells: scRNA-seq & 25 protein levels (30'000 cells). We analysed the correlation between gene expression and corresponding protein level.



Results

10X multiome dataset of pbmc

SuperCell identified robuste metacell in multiomic space of PBMC cells (Fig.1).

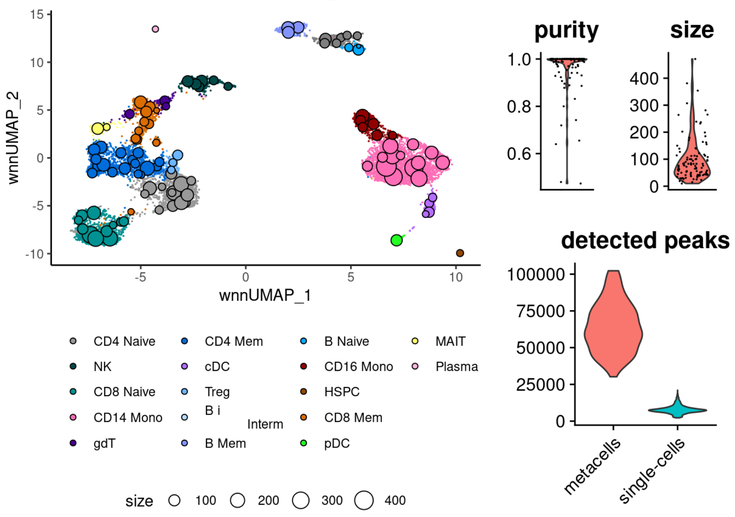


Figure 1: identified metacell in the multimomic space of PBMCs

Gene activity and expression for key TF of lymphoid cells (TCF7) and myeloid cells (SPI1) appeared more correlated at the metacell level (Fig.2).

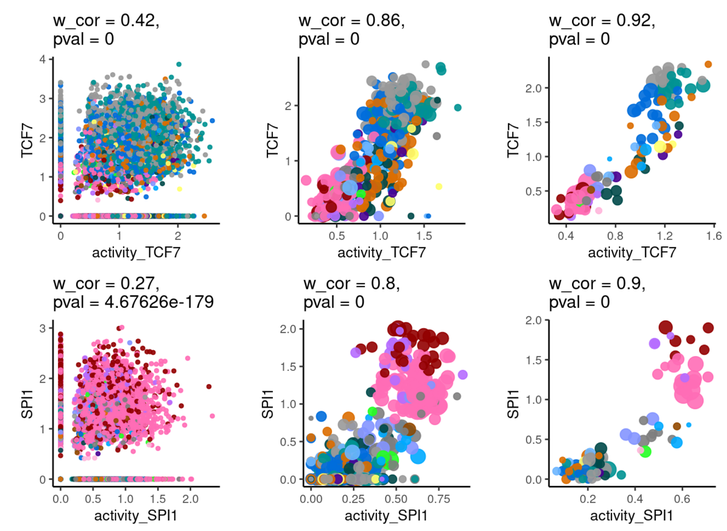


Figure 2: gene expression-activity correlation for key TF of immune cells

Analysis of this correlation for the most variable genes at the transcriptomic level (top 2000) show the same trend (Fig.3).

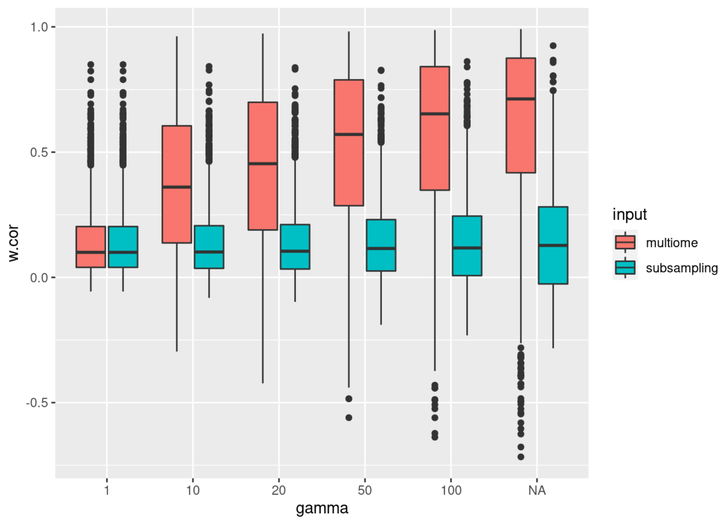


Figure 3: Global gene expression-activity correlation

SuperCell identified robuste metacell in CITE-seq space of BM cells (Fig.4).

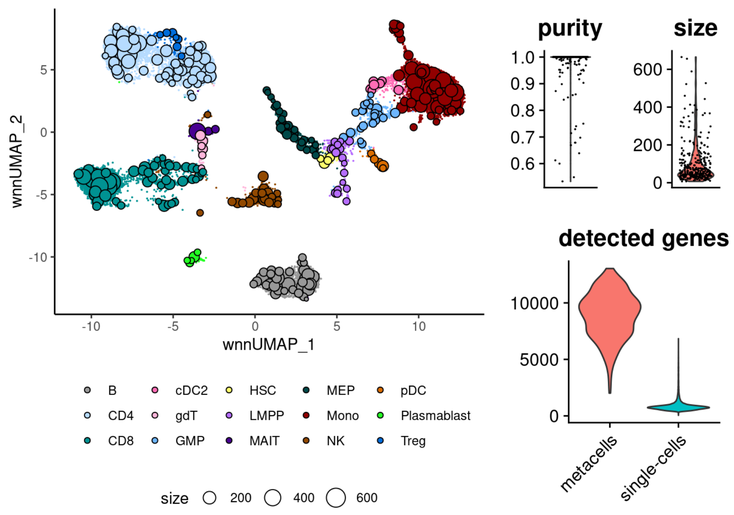


Figure 4: identyfied metacell in the multimodal single cell CITE-seq space of bone marrow cells

Metacell analysis increase the correlation between RNA and protein level of key markers of HSC (CD34) and CD8 T cells (CD8A) Fig.5.

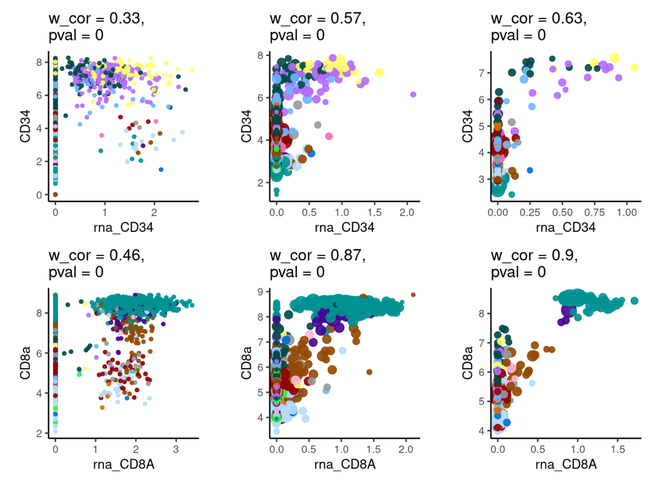


Figure 5: RNA protein correlations for key markers of BM cells

The analysis of the RNA-protein correlation for the all 25 gene-protein couples exhibits the same trend Fig.6.

Figure 6: Global RNA-Protein correlations

References

Baran, Yael, Akhiad Bercovich, Arnau Sebe-Pedros, Yaniv Lubling, Amir Giladi, Elad Chomsky, Zohar Meir, Michael Hoichman, Aviezer Lifshitz, and Amos Tanay. 2019. "MetaCell: Analysis of Single-Cell RNA-Seq Data Using K-Nn Graph Partitions." *Genome Biology* 20 (1): 206. https://doi.org/10.1186/s13059-019-1812-2.
Bilous, Mariia, Loc Tran, Chiara Cianciaruso, Aurélie Gabriel, Hugo Michel, Santiago J. Carmona, Mikael J. Pittet, and David Gfeller. 2022. "Metacells Untangle Large and Complex Single-Cell Transcriptome Networks." *BMC Bioinformatics* 23 (1): 336. https://doi.org/10.1186/s12859-022-04861-1.
Hao, Yuhan, Stephanie Hao, Erica Andersen-Nissen, William M. Mauck, Shiwei Zheng, Andrew Butler, Maddie J. Lee, et al. 2021. "Integrated Analysis of Multimodal Single-Cell Data." *Cell* 184 (13): 3573–3587.e29. https://doi.org/10.1016/j.cell.2021.04.048.

CITE-seq dataset of BM cells

posterdown

posterdown