BITS :: Call for Abstracts 2023 - Oral communication

Type	Oral communication
Session	Single-cell data analysis
Title	Harmonizing the annotation of single cells in normal and aberrant hematopoiesis with the Cell Marker Accordion
All Authors	Busarello E(1), Biancon G(2), Ibnat Z(1) , VanOudenhove J(2), Lauria F(3), Aass K(4), Standal T(4), Viero G(3), Halene S(2), Tebaldi T(1,2)

Affiliation

(1) Department of Cellular, Computational and Integrative Biology (CIBIO), University of Trento, Trento, Italy
Section of Hematology, Department of Internal Medicine, Yale Comprehensive Cancer Center, Yale (2) University School of Medicine, New Haven, CT, USA
(3) Institute of Biophysics, CNR Unit at Trento, Italy
(4) Norwegian University of Science and Technology (NTNU), Trondheim, Norway

Motivation

In the past few years, single-cell and spatial technologies have significantly improved our knowledge of cellular heterogeneity and architecture in health and disease, providing unprecedented insight into hematopoiesis.
A crucial and challenging step in single-cell data analysis is the annotation of cell types and states. Marker genes provide an effective way to achieve accurate cell type identification. The recent increase in databases collecting markers potentially simplifies the automatic annotation of cells. Nevertheless, each database contains dissimilar marker sets for the same cell type, leading to inconsistent annotations depending on the source choice. Furthermore, the lack of standard classification of cell types and validation of marker sets in the databases increases the annotation discrepancies.

Methods

Here, we developed the Cell Marker Accordion, a novel R Shiny web app and R package addressing the need for robust and reproducible identification of hematopoietic cell types in single-cell datasets. As data sources, we considered multiple databases collecting human and mouse gene markers and standard collections of widely used cell sorting markers for normal and aberrant hematopoiesis. Annotations were standardized by mapping initial cell types to the Cell Ontology, and pathologies to the Disease Ontology. Next, databases were integrated obtaining a comprehensive set of 7650 markers associated with 145 cell types. After the integration, marker genes are weighted according to their specificity scores, indicating whether they are markers for different cell types, and also to their evidence consistency scores, measuring the agreement of different annotation sources.

Results

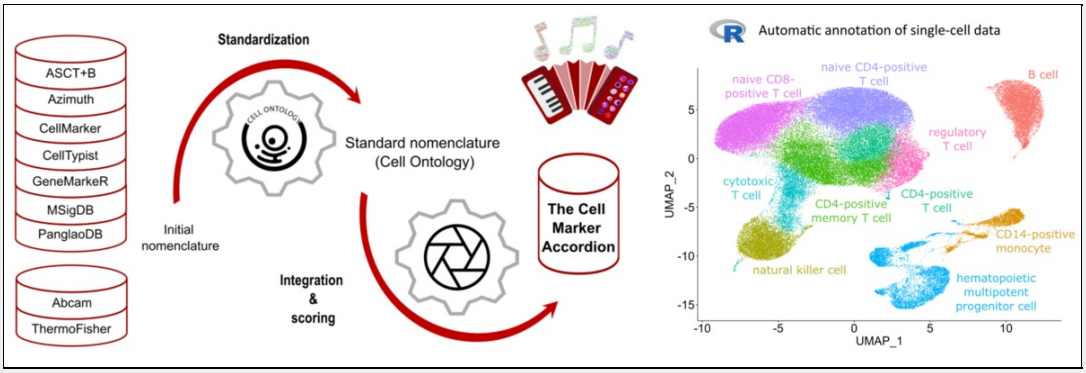
The Cell Marker Accordion web interface permits to easily retrieve lists of marker genes associated with input cell types and also viceversa, starting from a list of candidate genes to obtain the matching cell types. Hierarchies of hematopoietic cell types can be easily browsed following the Cell Ontology structure in order to obtain the desired level of resolution in the markers. The R package allows to automatically annotate single-cell data exploiting marker genes weighted according to their specificity and evidence consistency scores. The top influential markers guiding the annotation can be easily explored for each cell population.
Importantly, we validated the Cell Marker Accordion on flow cytometry sorted blood cell populations and human bone marrow samples profiled with multi-omics single-cell methods (CITE-seq and Abseq), using surface markers as the ground truth. In all cases, annotation with the Cell Marker Accordion significantly improved identification accuracy of cell types with respect to any of the single source databases.
Furthermore, we demonstrated the efficacy of the Cell Marker Accordion in identifying disease critical cells in acute myeloid leukemia and multiple myeloma patient-derived single cell datasets.
The Cell Marker Accordion is indeed a user-friendly and flexible tool that can be exploited to improve the annotation of hematopoietic populations in single-cell datasets focused on the study of human pathologies.

Info

-

filename	Accordion_workflow.png

Figure



Availability	https://github.com/TebaldiLab/TheCellMarkerAccordion

Corresponding Author

Name, Surname	Emma, Busarello
Email	emma.busarello@unitn.it
Submitted on	21.04.2023