BITS :: Call for Abstracts 2021 - Oral communication

Type	Oral communication
Session	Multiomics and Single Cell Analysis
Title	A new application for an old metric: Bray Curtis dissimilarity as tool to associate clusters among independent single cell RNAseq experiments.
All Authors	Piacenti G (1), Alessandri L (2), and Calogero RA (2)
Affiliation	

(1) Department of Life Sciences and System Biology
(2) Department of Molecular Biotechnology and Health Sciences

Motivation

Single-cell RNA sequencing (scRNAseq) has emerged as an essential tool to investigate cellular heterogeneity. Individual cells of the same phenotype are commonly viewed as identical functional units of a tissue or organ. However, single cell sequencing results suggest the presence of a complex organization of heterogeneous cell states producing together system-level functionalities. Furthermore, scRNAseq provides an extraordinary resolution into the molecular states present within a complex biological system at any given moment. However, it is a major challenge to integrate single-cell sequencing data across experiments, conditions, batches, time points and other technical considerations. New computational methods are required that can simultaneously preserve biological signals, while also integrating samples.

Methods

We evaluated the use of Bray Curtis dissimilarity score (BC), as tool to identify association between clusters derived from the analysis of different experiments. BC is an ecological population metric frequently used in metagenomics to evaluate the similarity between different samples on the basis of the bacterial represented in each sample.

Results

We have assimilated cell clusters to metagenomic populations and, after detecting cluster specific genes, we have identified the GO terms specific of each cluster. Then, we measured similarity between clusters in different experiments using BC applied both on GO terms or genes. We observed that BC works much better using cluster specific genes than GO term specific genes, but the sensitivity and specificity of BC was relatively poor. Thus, we developed a new approach in which we measure BC multiple times, gradually converting the cells of a cluster into the cells of the compared cluster, by progressive gene expression substitutions.
This approach showed a better performance that the simple BC comparison. To improve sensitivity and specificity we tested different data normalization and batch effect removal methods. The best results were obtained removing technical noise. Upon removal of technical noise BC sensitivity and specificity became optimal.
Taken together our data indicate that our BC-based algorithm can be efficiently used to integrate single-cell sequencing data across experiments.

Info

-

Figure

-

Availability	https://github.com/kendomaniac/BCsctutorial

Corresponding Author

Name, Surname	Raffaele, Calogero
Email	raffaele.calogero@unito.it

Submitted on	15.04.2021