Fcoex: an R package for detecting co-expression modules in single-cell RNA-Seq data

Tiago Lubiana, HELDER T I NAKAYA

Institute of Mathematics and Statistics, University of São Paulo

Abstract

The boom of single-cell transcriptomics was followed by a growth in methods for the analysis of single-cell data. Currently, standard pipelines (such as Seurat and Biocondutor's OSCA) do not include co-expression network building and modules detection methods. Modern systems biology uses co-expression networks both for exploratory data analysis and gene regulatory network inference. Current methods for building these networks, such as WGCNA, were developed for bulk RNA-Seq and do not perform as well in single-cell data. In the present work, we show how a feature selection algorithm, the Fast Correlation-Based Filter (FCBF), can be used to detect co-expression modules in single-cell data via an R package called fcoex. The package is awaiting reviews for the Bioconductor repository and is available at https://github.com/csbl-usp/fcoex. We applied it to single-cell data from human, mice, and zebrafish, detecting o co-expression modules with known biological partners and putative associations. The presence of anticorrelated genes in the same modules allowed the detection, in the zebrafish dataset, of a module containing both a ligand (apela) and its receptors (aplnra/aplnrb), yielding insights into the biology of vertebrate development. Also, fcoex enables module-based reclustering of the datasets for multilevel labeling of cells, uncovering new populations, and avoiding trade-offs of the current label-determination methods. In parallel, we detected new candidates for subpopulations of zebrafish embryo cells and human blood monocytes, demonstrating the usefulness of our tools for exploratory data analysis of single cells.

Funding: This work was supported by the grant 2018/10257-2, São Paulo Research Foundation (FAPESP)