

BITS :: Call for Abstracts 2024 - Oral communication

<i>Type</i>	Oral communication
<i>Session</i>	Bioinformatics AI, Models and Tools
<i>Title</i>	From Gene to Cluster: COTAN v2 for Improved scRNA-seq Analysis and Unsupervised Clustering
<i>All Authors</i>	Silvia Giulia Galfrè(1), Marco Fantozzi(2), Irene Testa(1), Matteo Tolloso(1), Andrea Alberti(1), Alina Sirbu(1), Francesco Morandin(2), Corrado Priami(1)

Affiliation

(1) Department of Computer Science - University of Pisa, Pisa
(2) Department of Mathematics, Physics and Informatics - University of Parma

Motivation

The COTAN library [1] provides a unique approach in analyzing scRNA-seq data. Its emphasis on gene co-expression at the single-cell level, including for lowly expressed genes, promises to offer significant advantages in extracting meaningful insights from sparse datasets. However, a complete evaluation of the platform in different settings, and comparison with other tools was lacking. Furthermore, the algorithmic performance of the previous version was not fully optimized, often resulting in large computational times for the analysis. That version also lacked some user-friendly plotting tools.

Methods

COTAN (v2) was rigorously evaluated across various stages of the scRNA-seq workflow: from its core gene-gene correlation analysis to cluster marker identification. It employs a generalized contingency table approach, focusing on zero unique molecular identifier (UMI) counts to assess correlations and anti-correlations between gene pairs. The COTAN package now integrates Torch-optimized functions, significantly enhancing the speed of gene co-expression evaluations. Moreover, many different plotting functions have been added, which can help the user analyze and interpret results.

Results

1. Efficiency in Sparse Data: COTAN effectively handles the intrinsic sparsity of scRNA-seq data by bypassing the need for imputation and normalization, which can introduce biases.
2. Robustness Against False Positives: COTAN shows robustness against false positive correlations in gene expression, which is crucial for accurate data interpretation.
3. Unsupervised Clustering: the tool employs unsupervised clustering that is biologically informed, focusing on creating clusters that are uniform in their transcriptomic profiles. To define these clusters we use a specific score. We have improved the assessment of this score's sensitivity, a crucial aspect of clustering, demonstrating that even a minimal presence of different cells within a cluster can significantly influence its values.
4. Best performance on cluster marker identification: we tested COTAN against three of the most used tools over gene marker identification and COTAN outperformed the other tools in terms of ROC curves and Type I errors.

Conclusions: Our analysis positions COTAN as a versatile, innovative, and useful tool for scRNA-seq data analysis and interpretation.

[1] Galfrè, Silvia Giulia, et al. "COTAN: scRNA-seq data analysis based on gene co-expression." NAR Genomics and Bioinformatics 3.3 (2021): lqab072.

Info

Repository
<https://github.com/seriph78/COTAN>
<https://www.bioconductor.org/packages/release/bioc/html/COTAN.html>

Analysis website
https://seriph78.github.io/COTAN_Datasets_analysis/

filename -

Figure

-

Availability -

Dissemination Material

Social

[@GalfrèSilvia](https://www.linkedin.com/in/silvia-giulia-galfrè-1796556)

Summary

We're excited to share COTAN v2, the latest update in our single-cell analysis tool.

What's New?

- Quicker analyses.
- Better tools for viewing your data.
- More accurate detection of gene patterns.

- Unsupervised cell clustering.
- Top performance in finding cell markers.

Corresponding Author

<i>Name, Surname</i>	Silvia, Galfrè
<i>Email</i>	silvia.galfrè@di.unipi.it
<i>Submitted on</i>	01.05.2024

Società Italiana di Bioinformatica

C.F. / P.IVA 97319460586

E-mail bits@bioinformatics.it

Sede legale Viale G. Mazzini, 114/B - 00195 Roma

Website bioinformatics.it

message generated by sciencedev.com for <https://bioinformatics.it> 20:26:42 01.05.2024
