

BITS :: Call for Abstracts 2023 - Oral communication

<i>Type</i>	Oral communication
<i>Session</i>	Single-cell data analysis
<i>Title</i>	COTAN v2: a Comprehensive and Versatile Framework for Single-Cell Gene Co-Expression Studies and Cell Type Identification
<i>All Authors</i>	Silvia Giulia Galfrè ⁽¹⁾ , Marco Fantozzi ⁽²⁾ , Daniel Puttini ⁽³⁾ , Corrado Priami ⁽¹⁾ , Francesco Morandin ⁽³⁾
<i>Affiliation</i>	(1) Department of Computer Science - University of Pisa (2) Free-lance (3) Department of Mathematics, Physics and Informatics - University of Parma
<i>Motivation</i>	COTAN is a statistical and computational method that analyzes the co-expression of gene pairs at the single-cell level to estimate gene co-expression in single-cell RNA sequencing (scRNA-seq) data. The low efficiency of scRNA-seq methodologies makes sensitive computational approaches crucial to accurately infer transcription profiles in a cell population.
<i>Methods</i>	COTAN employs an innovative mathematical model that leads to a generalized contingency table framework. It relies on the zero unique molecular identifier (UMI) counts distribution instead of focusing on positive counts to evaluate or extract different scores and information for gene correlation studies and gene or cell clustering. COTAN assesses whether gene pairs are correlated or anti-correlated, provides a new correlation index with an approximate p -value for the associated test of independence, and checks whether single genes are differentially expressed, scoring them with a newly defined global differentiation index (GDI). It also plots and clusters genes according to their co-expression pattern with other genes to study gene interactions and identify cell-identity markers. Through the GDI, COTAN assesses whether a cell cluster is homogeneous or not, making it a valuable tool for cell clustering and assignment. The latest version of COTAN includes new functions and plots to check and clean the dataset and several visualization tools to help users explore and interpret their data.
<i>Results</i>	COTAN shows high sensitivity in extracting information from small clusters and lowly expressed genes. It leverages its contingency table framework to directly identify genes that are over-represented or under-represented in the cluster with respect to the rest of the dataset. COTAN computes an enrichment score for a given list of marker genes, which can be used to identify and merge small uniform clusters and to check a final cluster identification. The strength of COTAN is its ability to help researchers better understand scRNA-seq data, identify gene modules, cell types, and new marker genes, and gain insights into the underlying biology of their samples.
<i>Info</i>	http://www.bioconductor.org/packages/release/bioc/html/COTAN.html
<i>filename</i>	-
<i>Figure</i>	-
<i>Availability</i>	-
Corresponding Author	
<i>Name, Surname</i>	Silvia Giulia, Galfrè
<i>Email</i>	silvia.galfrè@di.unipi.it
<i>Submitted on</i>	28.04.2023

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