BITS :: Call for Abstracts 2021 - Oral communication

Type	Oral communication
Session	Multiomics and Single Cell Analysis
Title	scSeqComm: a bioinformatic tool to identify, quantify and characterize cellular communication from single cell RNA sequencing data
All Authors	Giacomo Baruzzo (1), Giulia Cesaro (1), Barbara Di Camillo (1,2,3)

Affiliation

(1) Department of Information Engineering, University of Padova, Padova, Italy
(2) Department of Comparative Biomedicine and Food Science, University of Padova, Padova, Italy
(3) CRIBI Innovative Biotechnology Center, University of Padova, Padova, Italy

Motivation

Despite the different ways in which cells can communicate, a considerable amount of cellular crosstalk is performed through ligand-receptor interactions. As a response to the intercellular signaling through the ligand-receptor binding, the receptor activates an intracellular signaling cascade inside the receiving cell that triggers the final cell response.
Recent studies have demonstrated that scRNA-seq data can be used to infer intercellular communication through the analysis of ligands and receptors gene expression data [1–3]. Most of the methods infer ongoing intercellular communication between different cell clusters combining ligand and receptor expression data after the application of a user specified expression threshold, a step that has a critical role on results accuracy. Moreover, current methods largely ignore the intracellular signal that is activated downstream the cell and the consequent cell response.
In this work, we propose one of the first attempt to infer, quantify and characterize both the intercellular and the related intracellular signaling from scRNA-seq data. The main contributions of this work are i) a computational procedure to quantify and characterize the effect of an ongoing intracellular signaling in the receiving cells, and ii) a novel way to identify and quantify intercellular signaling avoiding user specified expression threshold.

Methods

To identify and quantify ongoing intercellular signaling, we score ligands and receptors in terms of their level of expression in comparison with a random gene distribution in the same cell cluster, accounting for both cluster variance and numerosity of cells. Using the information from ligand-receptor pairs databases, we combined ligand and receptor scores using the minimum function (“fuzzy logical AND” operator), modeling the fact that cellular communication is ongoing when both the ligand and the receptor are active.
To infer intracellular signaling we combine data from pathway (KEGG [4]) and transcriptional regulatory network (HTRIdb [5], TRRUST [6] and RegNetwork [7]) databases to build graphs of interacting genes. For each graph containing an annotated receptor, the evidence of an ongoing intracellular signaling is computed as a function of 1) the topology of the intracellular signaling graph that link the receptor to the downstream transcription factors (TF), measured as the Personalized PageRank score [8] between receptor and TF nodes; 2) the scRNA-seq gene expression data of genes regulated by TFs, using a Fisher test on TF regulated genes that are differentially expressed (DE) in the given cell cluster. The same DE genes are then used to functionally characterize the inferred intracellular signaling through a Gene Ontology enrichment analysis.
The proposed method is implemented in the R package scSeqComm available at https://gitlab.com/sysbiobig/scseqcomm.

Results

In terms of intercellular signaling, our scoring scheme: 1) does not require setting an arbitrary expression level threshold; 2) makes no assumption on the scale (e.g. log scale) and meaning (e.g. CPM, FPKM) of the data; 3) considers the characteristics and reliability of the entire dataset under analysis, since it depends not only on the ligand/receptor expression level, but also on the expression level of other genes, on the number of cells in the cluster and on the variability of the data; 4) has a continuous and finite range that facilitates the ranking of the results and their interpretability and portability across datasets/clusters. To the best of our knowledge, none of the intercellular scores proposed in literature has all the characteristics listed above.
In terms of intracellular signaling, our attempts to infer, quantify and functionally characterize the downstream effects of an ongoing intercellular communication have three main goals. First, we aimed at providing a new level of cellular communication analysis that is relevant but largely unaddressed by current methods. Second, the detection of an ongoing intracellular signaling may be used to strengthen the

evidence of an inferred intercellular signaling and vice versa, increasing the robustness of results. Last, the functional characterization of cellular communication effects further assists the interpretation of the results and the understating of ongoing biological processes.
Considering the noise typical of scRNA-seq data and the fact that RNA is only a proxy of protein activity, we think that the combined evidence of ongoing intercellular and intracellular signaling represents the most reliable way to study cellular communications from scRNA-seq data solely, enabling the prioritization and identification of the most strong and reliable signals.
To better appreciate the various facets of the method, we applied it on a publicly available scRNA-seq dataset of tumor microenvironment [9] achieving good overlap with independent bioinformatics results and in accordance with current biological knowledge.

Info

References

- [1] Kumar MP, Du J, Lagoudas G, Jiao Y, Sawyer A, Drummond DC, et al. Analysis of Single-Cell RNA-Seq Identifies Cell-Cell Communication Associated with Tumor Characteristics. Cell Rep 2018;25:1458-1468.e4. https://doi.org/10.1016/j.celrep.2018.10.047.
- [2] Puram S V., Tirosh I, Parikh AS, Patel AP, Yizhak K, Gillespie S, et al. Single-Cell Transcriptomic Analysis of Primary and Metastatic Tumor Ecosystems in Head and Neck Cancer. Cell 2017;171:1611-1624.e24. https://doi.org/10.1016/j.cell.2017.10.044.
- [3] Zhang Q, He Y, Luo N, Patel SJ, Han Y, Gao R, et al. Landscape and Dynamics of Single Immune Cells in Hepatocellular Carcinoma. Cell 2019;179:829-845.e20. https://doi.org/10.1016/j.cell.2019.10.003.
- [4] Kanehisa M, Goto S. KEGG: Kyoto Encyclopedia of Genes and Genomes. Nucleic Acids Res 2000;28:27–30. https://doi.org/10.1093/nar/28.1.27.
- [5] Bovolenta LA, Acencio ML, Lemke N. HTRIdb: an open-access database for experimentally verified human transcriptional regulation interactions. BMC Genomics 2012;13. https://doi.org/10.1186/1471-2164-13-405.
- [6] Han H, Cho JW, Lee S, Yun A, Kim H, Bae D, et al. TRRUST v2: An expanded reference database of human and mouse transcriptional regulatory interactions. Nucleic Acids Res 2018;46:D380–6. https://doi.org/10.1093/nar/gkx1013.
- [7] Liu ZP, Wu C, Miao H, Wu H. RegNetwork: An integrated database of transcriptional and post-transcriptional regulatory networks in human and mouse. Database 2015;2015:1–12. https://doi.org/10.1093/database/bav095.
- [8] Page L, Brin S, Motwani R, Winograd T. The PageRank Citation Ranking: Bringing Order to the Web. Stanford InfoLab; 1999.
- [9] Tirosh I, Izar B, Prakadan SM, Wadsworth MH, Treacy D, Trombetta JJ, et al. Dissecting the multicellular ecosystem of metastatic melanoma by single-cell RNA-seq. Science (80-) 2016;352:189–96. https://doi.org/10.1126/science.aad0501.

Source of funding

- This research was funded mainly by PROACTIVE 2017 “From Single-Cell to Multi-Cells Information Systems Analysis” (C92F17003530005, Department of Information Engineering, University of Padova).
- Part of this work was supported by MIUR (Italian Minister for Education) under the initiative "Departments of Excellence" (Law 232/2016).
- Giacomo Baruzzo was founded by Department of Information Engineering, University of Padova under the initiative “Research Grant (type B) – B senior initiative”.

Figure

-

Availability	https://gitlab.com/sysbiobig/scseqcomm

Corresponding Author

Name, Surname	Barbara, Di Camillo
Email	barbara.dicamillo@unipd.it

Submitted on	27.04.2021