A mathematical method and software for spatially mapping intercellular communication

Communication between cells is crucial for coordinated cellular functions in multicellular organisms. We present an optimal transport theory-based tool to infer cell–cell communication networks, spatial signaling directions and downstream targets in multicellular systems from spatial gene expression data.

This is a summary of:
Cang, Z. et al. Screening cell–cell communication in spatial transcriptomics via collective optimal transport. *Nat. Methods* https://doi.org/10.1038/s41592-022-01728-4 (2023).

Published online: 24 January 2023

The problem

Cells communicate with each other to instruct cell function and fate. A major process in cell–cell communication is through molecular interactions in which signaling molecules (ligands) interact with compatible host molecules (receptors), resulting in the activation of downstream intracellular targets[1]. Coordinated functions of cells and the mechanisms responsible for the organization of a multicellular system are directly linked to its cell–cell communication networks. Emerging spatial transcriptomics technologies[2] provide unprecedented opportunities for dissecting these complex processes. With spatial constraints on molecular interactions, competition between cells, and varying promiscuity of signaling interactions, inferring cell–cell communication from spatial data is a challenging task. Moreover, tools for evaluating and benchmarking computational methods are lacking. New methods are needed for effective representations and visualizations of spatial signaling and analysis of downstream targets. Broadly applicable methods for inferring cell–cell communication are a pressing need.

The solution

We reformulated the cell–cell communication problem as an optimal transport (OT)[3] system wherein ligands are 'transported' to receptors and the 'optimal' transport connections coupling them are calculated. Specifically, the measured ligand and receptor expression levels are considered as spatial mass distributions and the transportation cost is estimated from spatial distances. To deal with modeling challenges associated with spatial constraints and competition among cells and molecules, we developed a new OT method named collective OT, along with an efficient numerical implementation. We developed a user-friendly collective-OT-based software platform (Fig. 1a,b), communication analysis by optimal transport (COMMOT), for cell–cell communication inference using spatial gene expression data, along with intuitive tools for visualization and novel analysis of downstream targets.

We demonstrated the general applicability of COMMOT by studying spatial gene expression datasets generated using various state-of-art technologies, obtaining results consistent with prior knowledge in systems such as tumors and the brain cortex. For example, fibroblast growth factor signaling is directed toward the cerebellar cortex in the mouse brain (Fig. 1c). In a case study of human skin, COMMOT identified new signaling molecules that regulate skin development. To develop benchmarking tools, we constructed a partial differential equation model of diffusive ligands and their interactions with receptors. COMMOT accurately reconstructed cell–cell communication from the synthetic data generated by the model. Using machine learning models, COMMOT screens affected genes on the basis of the inferred signaling networks. Convenient visualizations in COMMOT, such as for spatial signaling directions, enable further exploration of the inference outputs.

Future directions

Our work provides a mathematical framework, a computational tool, and the COMMOT open-source software for inferring cell–cell communication using various sources of spatial gene expression data. COMMOT is equipped with multiple visualization options and downstream analysis utilities, has extensive documentation and tutorials, and is particularly powerful for comprehensive screening of cell–cell communication connections in spatial systems. Our study should spark interest in developing new OT methods that can handle different constraints or types of competition in biology or other applications[4].

Spatial transcriptomics data are related to, but do not directly represent, the actual abundance of proteins that mediate cell–cell communication. Multiple biophysical processes, including the translation of mRNA into protein, need to be considered for more accurate inference. A lack of such details in COMMOT might lead to false positive links for some systems. An extension and refinement of the method would be to use other spatial or non-spatial multi-omics data to fill the gap between transcriptomics and proteomics. The inclusion of prior knowledge of a specific biological system, either directly to the data or after the inference step, can further improve robustness and accuracy in discovering communication links by COMMOT.

Zixuan Cang[1] & Qing Nie[2]
[1]North Carolina State University, Raleigh, NC, USA. [2]University of California Irvine, Irvine, CA, USA.

EXPERT OPINION

"With COMMOT, Cang et al. present an elegant mathematical solution to the problem of inferring cell–cell communication from spatial transcriptomics data based on a variant of optimal transport. The method is applied to spatial datasets of different sizes and technologies, and robustness of results is shown. Further, the authors show how their method can be used in different biological contexts, including human breast cancer and mouse brain samples." **Fabian Theis and Marius Lange, Helmholtz Munich, Germany**

FIGURE

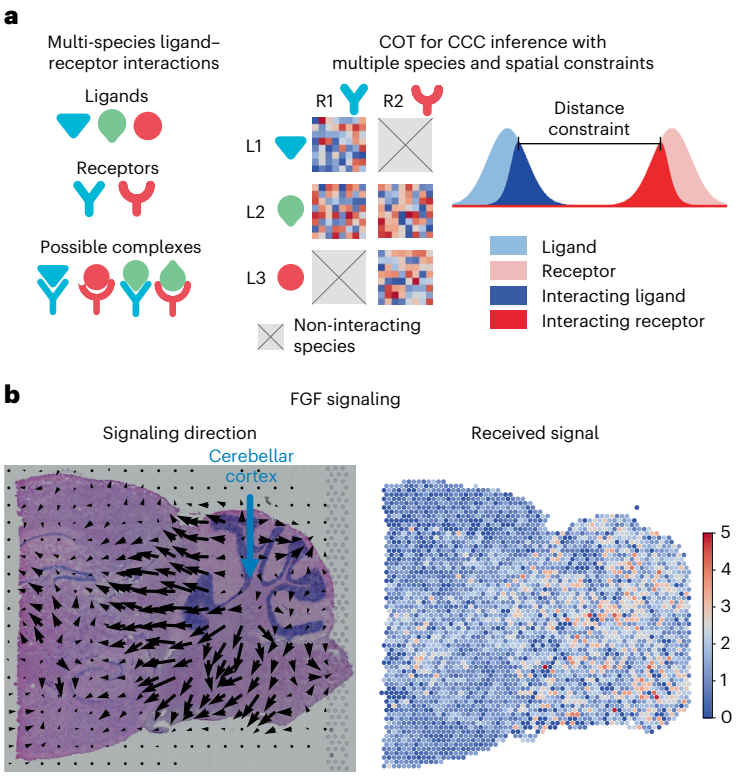

Fig. 1 | Design and illustration of COMMOT. a, Left, inferring cell–cell communication (CCC) in space by considering competition between different ligand and receptor species. Right, collective optimal transport (COT) can include multi-species distributions (such as the ligands L1–L3 and receptors R1 and R2) and enforce constraints on spatial ranges. **b**, Inferred fibroblast growth factor (FGF) signaling direction (arrows) in a sagittal section of the posterior mouse brain in Visium spatial transcriptomics data.

BEHIND THE PAPER

In our earlier work, we equipped single-cell gene expression data with spatial information by integrating it with imaging data using the optimal transport method. Now, with spatially annotated single-cell data or spatial transcriptomics data, we can better study cell–cell communication. We have a long-standing interest in intercellular communication and have developed widely used tools[5] for inferring intercellular communication from non-spatial data. With more spatial data becoming available, we thought the optimal transport method would be a great way to establish the connection between ligands and receptors represented as mass distributions to model cell–cell communication. The need to incorporate various biological constraints motivated us to develop extensions of optimal transport theory. Our study led to both a practically useful piece of software for signaling inference from spatial data and the mathematical development of a generally applicable optimal transport method. **Z. C. & Q. N.**

REFERENCES

1. Armingol, E., Officer, A., Harismendy, O. & Lewis, N. E. Deciphering cell–cell interactions and communication from gene expression. *Nat. Rev. Genet.* **22**, 71–88 (2021).
   A review article on cell–cell communication analysis from gene expression data.
2. Moses, L. & Pachter, L. Museum of spatial transcriptomics. *Nat. Methods* **19**, 534–546 (2022).
   A review article on spatial transcriptomics technologies and analysis methods.
3. Peyré, G. & Cuturi, M. Computational optimal transport: with applications to data science. *In Foundations and Trends in Machine Learning* **11**, 355–607 (2019).
   A book on optimal transport and its applications.
4. Cang, Z., Nie, Q. & Zhao, Y. Supervised optimal transport. *SIAM J. Appl. Math.* **82**, 1851–1877 (2022).
   A study of the mathematical details of the optimal transport method motivated by our study.
5. Jin, S. et al. Inference and analysis of cell–cell communication using CellChat. *Nat. Commun.* **12**, 1088 (2021).
   This paper presents a widely used tool to infer cell–cell communication from non-spatial single-cell transcriptomics data.

FROM THE EDITOR

"Studying cell communications in the context of where cells are located is of interest. This work presents a promising computational tool, COMMOT, that employs collective optimal transport to study cell–cell communication from spatially resolved transcriptomics data." **Lei Tang, Senior Editor, *Nature Methods*.**