

Inaugural dissertation
for
obtaining the doctoral degree
of the
Combined Faculty of Mathematics, Engineering and
Natural Sciences
of the
Ruprecht - Karls - University
Heidelberg

Presented by

MSc Daniel Boykov Dimitrov

born in: Varna, Bulgaria

Oral examination: 28th of June 2024

*Modelling Cell-cell Communication using
Prior Knowledge with Single-cell and
Spatially-resolved Omics Data*

Referees:

Prof. Dr. Julio Saez-Rodriguez

Prof. Dr. Dr. Peer Bork

Acknowledgements	5
Abstract	6
Zusammenfassung	6
Chapter 1: Systematic Cell-cell Communication inference enabled by high-throughput omics data	7
Transcriptomics-enabled Systems-level CCC inference.....	8
Assumptions and Limitations.....	9
Promises and Challenges of Emerging Technologies.....	9
Motivation for an All-in-one Framework.....	11
Chapter 2: Cell-cell Communication Inference from Single-cell RNA-Seq Data: a Comparison of Methods and Resources	12
1. Introduction.....	12
2. Results.....	15
2.1 Resource Uniqueness and Overlap.....	15
2.2 Resource Prior Knowledge Bias.....	17
2.2.1 Subcellular Localisation.....	18
2.2.2 Functional Term Enrichment.....	18
2.3 Agreement in CCC predictions using a Colorectal Cancer dataset.....	19
2.3.1 Interaction overlap.....	19
2.3.2 Communicating cell types.....	20
3. Discussion.....	22
4. Methods.....	23
4.1. Descriptive analysis of resources.....	23
4.2. Single-cell Transcriptomics data.....	24
4.3. Framework.....	24
4.4. Overlap Analysis.....	24
4.5. Method Specifics.....	25
4.5.1 CellChat.....	25
4.5.2 Connectome.....	25
4.5.4 SingleCellSignalR.....	26
4.5.5 NATMI.....	26
4.5.6 Squidpy.....	26
Supplementary Materials.....	27
Chapter 3: Evaluating Cell-cell Communication Inference from Single-cell Transcriptomics Data Using Alternative Modalities	43
Background.....	43
Main.....	44
Agreement with Cell-surface Protein Expression (Target-Receptor Task).....	44
Agreement with Cytokine Activities (Ligand-Target task).....	46
Agreement with Cell-type Colocalization (Source-Target task).....	46
Formalised and Open Benchmarks.....	48
Methods.....	50
Tasks Setting.....	50
Metrics.....	51
Data Availability.....	51
Code Availability.....	51
Supplementary Material.....	52
Chapter 4: LIANA+: an all-in-one cell-cell communication framework	54

1. Background.....	55
2. Main.....	58
2.1 LIANA+ enables the modelling of CCC from Spatially-Resolved Data across distinct modalities	
As a consequence of the diverse array of spatial omics technologies and the wide range of molecules and resolutions they encompass, diverse computational approaches are required to make most of the data. In this context, LIANA+ includes several strategies to flexibly analyse CCC from diverse spatially-resolved technologies.....	58
2.1.1 LIANA+ jointly models global associations across modalities.....	58
2.1.2 LIANA+ infers local interactions at the individual spot or cell locations.....	59
2.2 LIANA+ Combines Dissociated Single-cell with Spatial Data to identify deregulated Intercellular and Intracellular Signalling across Conditions.....	62
2.2.1 LIANA+ extracts Disease-specific Communication Patterns from Spatial Transcriptomics Data.....	63
2.2.2 LIANA+ extracts Condition-specific Communication Patterns from Dissociated Single-cell data.....	64
2.2.3 LIANA+ identifies deregulated intercellular and intracellular signalling events.....	64
3. Discussion.....	67
4. Methods.....	70
Bivariate Spatially-informed Metrics.....	70
Learning Spatial Relationships across Multi-views.....	72
Estimation of Spatial Connectivities.....	73
Ligand-Receptor Pathway Enrichment.....	74
Hypothesis testing for deregulated CCC across Conditions.....	75
Sign-consistent Intracellular Networks.....	75
NMF on ligand-receptor local scores.....	77
LIANA+ in multimodal single-cell & spatial data.....	77
Intercellular Communication Factorization with Tensor-cell2cell and MOFA+.....	77
LIANA+'s Prior Knowledge Component.....	78
Spot Calling using Local metrics.....	78
Spatial Colocalisation Evaluation.....	78
Sample Label Classification.....	79
Analysis of Spatially-resolved Multi-omics Data from Murine Parkinson's disease model.....	80
Analysis of Human Myocardial Infarction.....	81
Analysis of Murine Acute Kidney Injury Data.....	82
Data Availability.....	83
Code Availability.....	83
Appendix.....	84
Bibliography.....	104

Acknowledgements

Four years ago I wasn't sure what being a PhD candidate meant, nor was I particularly interested in being one. I applied to PhD positions as a compromise that would allow me to reunite with my girlfriend (now wife). Yet, as it stands, I cannot imagine doing anything else. Applying and joining the lab of Prof. Dr. Julio Saez-Rodriguez was the luckiest decision I've made.

As such, first, I wish to express my gratitude to Julio for giving me a chance to work in his group. I also want to thank him for his kindness and liberal leadership, qualities which define the "saezlab". An environment that allowed me to discover a career which I'm deeply interested in. Julio, I couldn't have wished for a better mentor, thank you!

I want to thank all saezlab members for being friendly and collaborative. Also, I am grateful to all the co-authors and collaborators that helped with the development of LIANA+.

I also thank my previous supervisors Dr. Sabine Matallana-Surget and Dr. Quan Gu without whom I wouldn't have considered an academic career. Similarly, I am grateful to Dr. Georg Zeller for hosting me as a guest in his lab and allowing me to get a broader view on science. As well as Prof. Dr. Griet Gloriux for her dedication and kindness during my stay in her lab. I also thank Prof. Dr. Peer Bork for his advice and guidance.

I express my gratitude to the European Union's Horizon 2020 research and innovation program (860329 Marie-Curie ITN "STRATEGY-CKD") that financially supported my work.

Finally, my family has shaped me into the person that I am and I thank them for their love. Especially my grandma (Parashkeva) as without her support during the most challenging times I wouldn't be here.

То Катрин: всичко е за теб и от скоро за нея. Благодаря ти всичко. И където и да сме за мен нищо няма да се промени.

Abstract

Cell-cell communication (CCC) is a dynamic process which governs and coordinates diverse biological functions. The popularity of single-cell and spatially-resolved transcriptomics has recently sparked an array of computational methods that model CCC to be developed. My thesis describes the development of LIANA+ - an all-in-one framework for the inference of CCC from single-cell and spatial (multi-) omics data. In the first chapter of my thesis, I summarise the current state of the CCC field and state my motivation for the development of LIANA. In the second chapter, I describe the initial study that led to the development of the first iteration of LIANA. I also show that the choice of CCC method and resource can impact biological insights. In the third chapter, I evaluate CCC methods using alternative data modalities, and show that most CCC methods are generally coherent with those. In the final chapter, I summarise the current challenges of the CCC field and I showcase how the modularity of LIANA+ provides a comprehensive answer to those. I further see LIANA+ as a step towards modelling host-microbiome interactions, with already some pilots in place.

Zusammenfassung

Die Zell-Zell-Kommunikation (CCC) ist ein dynamischer Prozess, der verschiedene biologische Funktionen steuert und koordiniert. Die Popularität der Single-cell und Spatially-resolved Transcriptomics hat in letzter Zeit zur Entwicklung einer Reihe von Computational Methoden geführt, die CCC modellieren. Meine Dissertation beschreibt die Entwicklung von LIANA+ - einem All-in-One Framework für die Inferenz von CCC aus Einzelzell- und räumlichen (Multi-)Omics-Daten. Im ersten Kapitel meiner Arbeit fasse ich den aktuellen Stand des CCC-Feldes zusammen und erkläre meine Motivation für die Entwicklung von LIANA. Im zweiten Kapitel beschreibe ich die erste Studie, die zur Entwicklung der ersten Iteration von LIANA führte. Ich zeige auch, dass die Wahl der CCC-Methode und -Ressource die biologischen Erkenntnisse beeinflussen kann. Im dritten Kapitel bewerte ich CCC-Methoden unter Verwendung alternativer Datenmodalitäten und zeige, dass die meisten CCC-Methoden im Allgemeinen mit diesen kohärent sind. Im letzten Kapitel fasse ich die aktuellen Herausforderungen im Bereich CCC zusammen und zeige auf, wie die Modularität von LIANA+ eine umfassende Antwort auf diese Herausforderungen bietet. Darüber hinaus sehe ich LIANA+ als einen Schritt in Richtung der Modellierung von Wirt-Mikrobiom-Interaktionen, für die es bereits einige Pilotprojekte gibt.

Chapter 1: Systematic Cell-cell Communication inference enabled by high-throughput omics data

Cell-cell communication (CCC) is a process in which cells emit and react to stimuli from their cellular niche, as well as from themselves. CCC spans diverse cell types, molecular layers, and spatial scales, encompassing both structural interactions, such as those mediated by cell adhesion molecules, as well as signalling events facilitated by various secreted molecules. The signalling molecules themselves include heteromeric protein receptors, integrins, extracellular matrix proteins, metabolites, and ions ¹. Thus, CCC represents a multifaceted and dynamic system coordinating many biological processes, such as apoptosis and cell migration, and is consequently essential in homeostasis and disease ¹. CCC is further interlinked with intracellular signalling, as external signalling molecules are commonly thought to induce a downstream response (**Figure 1**), such as canonical pathways and downstream transcription regulators in the cells receiving the signal, or target cells. In turn, this alters the internal state of the target cells, potentially further propagating this alteration to their niche or microenvironment.

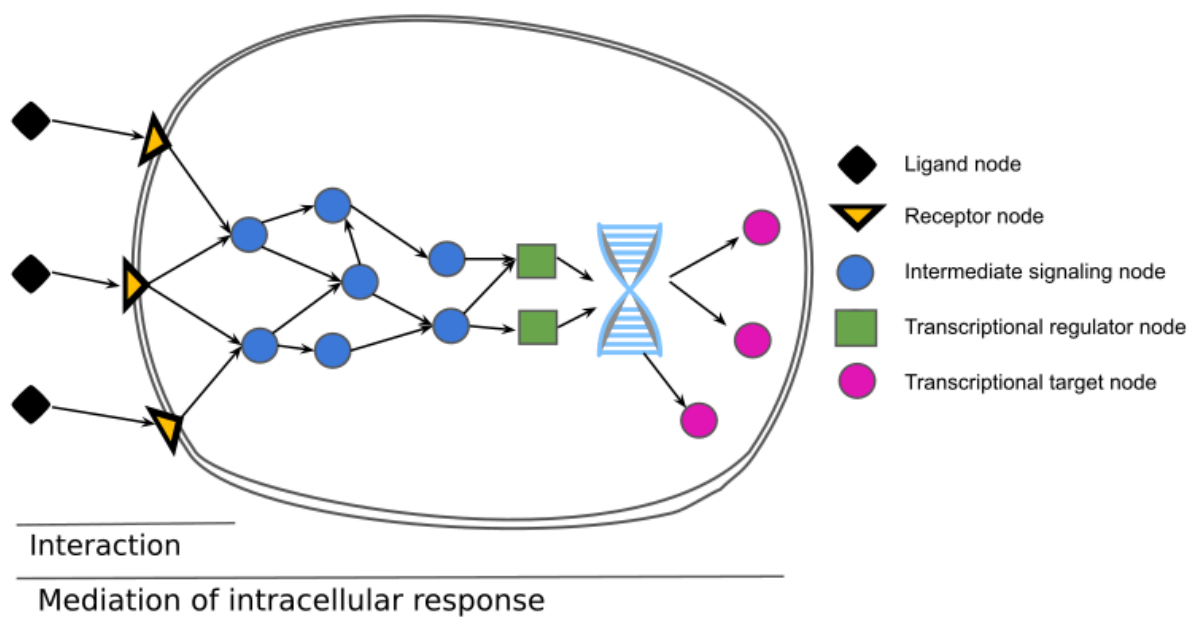












Figure 1. *Simplistic overview of CCC (intercellular) events and their propagation to intracellular response.*

Transcriptomics-enabled Systems-level CCC inference

Traditionally, the direct measurement of CCC required specialised biochemical assays, including co-immunoprecipitation, proximity labelling, and yeast two-hybrid screening ¹. Yet, these technologies are not always applicable *in vivo* and are relatively low-throughput ¹. To address these limitations, and driven by the rapid developments of transcriptomics data, specifically single-cell transcriptomics data, computational methods have emerged to infer CCC across thousands of cells and genes. Thus, enabling the study of CCC at the systems level ^{1,2}. The emergence of these methods and their popularity has in turn played a role in a currently-ongoing paradigm shift in the single-cell field. A shift away from investigating which types of cells are present, but rather focusing on the coordination and dependencies among them ^{2,3}. As a consequence, CCC inference from single-cell data is now becoming a routine approach, with an ever-increasing plethora of methods ^{1,2}.

When considering protein-mediated CCC inference from single-cell data, early methods could be classified as those that predict CCC interactions alone, referred to as ligand-receptor inference methods ⁴⁻⁶, and those that further predict CCC-mediated intracellular response ^{7,8}. Regardless of their task, most methods rely on information about the interacting proteins, obtained from prior knowledge resources ⁹. In the case of ligand-receptor methods, the interactions are commonly represented by heteromeric protein complexes, as different subunit combinations can induce distinct responses ^{4,6}. The methods which further infer intracellular signalling also require additional information such as intracellular protein-protein interaction network and/or gene regulatory interactions ^{7,8}.

Despite methodological diversity, the multifaceted and dynamic nature of CCC has often been represented in a tabular format which comprises an interaction of two proteins, the corresponding genes of which are expressed, or changed, in source and target cell types (**Figure 2**).

Ligand	Receptor	Source	Target	Score
TGF- β 1	TGF β R1			1
IL-10	IL-10R			0.99
BMP6	ALK2			0.98
IL-10	IL-10R			0.96
TGF- β 1	TGF β R1			0.91
...

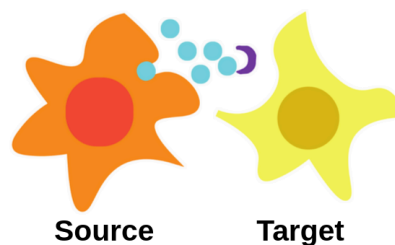


Figure 2. Typical tabular representation of CCC inference output. Ligand-receptor interactions, informed by prior knowledge, assigned a score reflecting the relative importance of the interaction between a source and target cell type.

Assumptions and Limitations

Every computational representation is an oversimplification of biological reality, hence the generation of systems-level of CCC unsurprisingly comes with certain assumptions and limitations. First, CCC methods typically work with single-cell transcriptomics data, and hence assume that co-expression of genes that encode proteins reflects intercellular interactions⁹. Such an assumption inherently implies that all events leading up to the interaction, including protein translation, processing, secretion, and diffusion, have occurred^{1,9} (**Figure 3**). Moreover, due to the almost exclusive use of transcriptomics data, other modes of CCC, such as small molecule signalling, have been typically neglected⁹. Similarly, transcriptomics data captures only a snapshot of expression at a specific sampling site; thus, long-range signalling events, such as endocrine signalling and system gradients like calcium, are not directly captured⁹. Nevertheless, there were some attempts to infer metabolite-mediated CCC events from transcriptomics data^{10–13}. Yet, such inference is limited by the challenges of estimating metabolite abundance from gene expression¹⁴. Also, recent methods have emerged to model external signals, even if they were not measured in the data¹⁵.

While clustering cells into biologically-meaningful groups is a common heuristic in the single-cell field, also used by many CCC methods, in reality CCC events occur among individual cells within a community. Therefore, any interactions inferred between cell types are influenced by the cell type annotation process¹⁶. To address this issue another set of tools predicts communication at the individual cell level, or meta-cell level^{16,17}.

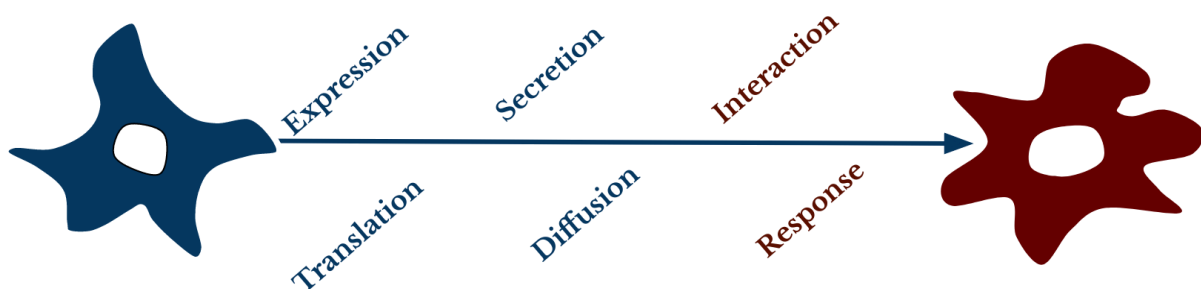


Figure 3. *Limitations of CCC inference from transcriptomics data*

Promises and Challenges of Emerging Technologies

When inferring CCC, it is noteworthy that many interactions depend on direct contact, while others are driven by secreted molecules, the effect of which diminishes over distances due to diffusion and absorption¹⁸. Yet, all the aforementioned approaches rely on transcriptome information from dissociated cells, and hence lack information about cellular neighbours. In recent years, a set of spatial omics technologies have emerged which enable high-throughput measurements while also providing physical proximities^{19,20}. As such, spatial technologies offer a promising opportunity to predict CCC, in a manner informed by the spatial proximity of cells and molecules.

Nevertheless, spatially-resolved technologies currently exhibit a trade-off between observation (spot) resolution and feature (e.g. gene) coverage¹⁹. Most single-cell resolution technologies, constrained by this trade-off, permit the measurement of only tens to hundreds

of features, providing a limited view for the systems-level inference of CCC ¹⁹. Therefore, CCC inference from such data is typically focused on the quantification of cell group co-localization, or cellular neighbourhoods ^{21,22}.

On the contrary, spatial technologies, such as the commercially-available 10x Visium platform ²³, provide transcriptome-wide coverage, yet they are constrained in a resolution ¹⁹, and each spatial location often includes transcripts from multiple cells from various groups. Despite this limitation, some spatially-informed methods have used such data to infer ligand-receptor interactions between cell types, under the assumption that each spot corresponds to a single cell type ^{24,25}, or by making use of cell type fractions per spot ^{26,27} - as inferred by cell deconvolution ²⁸. Another approach is to bypass cell-type identities altogether, and instead directly quantify the spatial relationships between proteins alone ^{29,30}. Moreover, some of the spatially-informed CCC methods identify the specific locations at which interactions occur ^{17,30-32}, while others summarise spatial dependencies across the whole slide ^{26,33} (**Figure 4**).

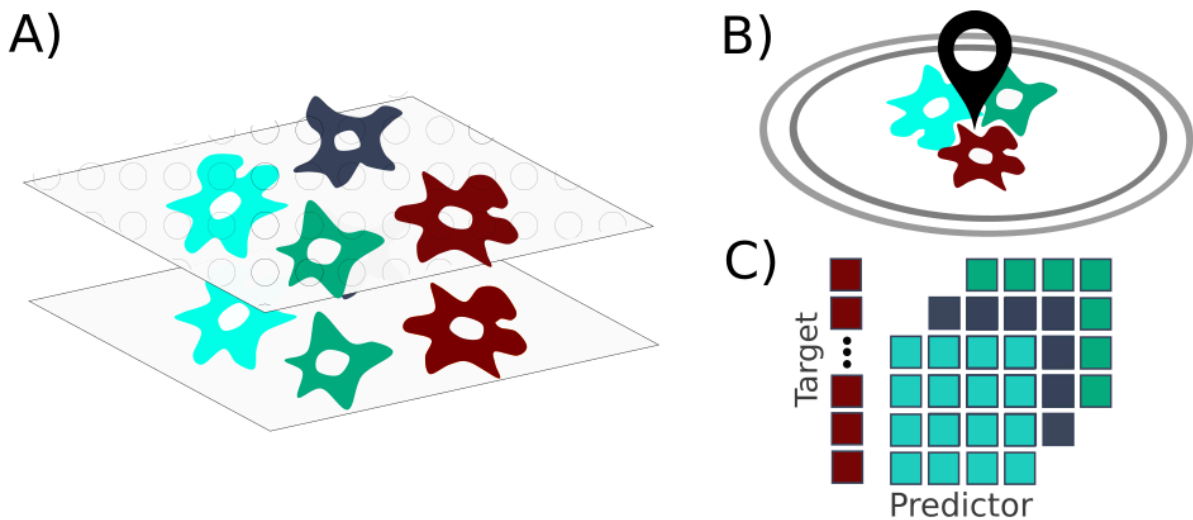


Figure 4. Schematic representation of **A)** spatially-resolved data, being modelled as a **B)** local interaction and **C)** spatial dependency across the slide.

Single-cell, and to an extent spatial, transcriptomics data is defined by its large feature space, along with a continuously growing number of cells (observations) being sampled. Yet, a third dimension of complexity is driven by a surge in sample numbers due to the generation of “cross-conditional” atlases, encompassing patient cohorts with contexts, such as diverse disease severity states, or time points ^{3,34}. This complexity has thus prompted the emergence of many approaches to obtain disease-relevant insights from single-cell data ^{3,34-36}. Similarly, there are also many diverse approaches that disentangle intercellular insights across contexts, which for the sake of simplicity one can classify as hypothesis-driven ^{7,37,38} and hypothesis-free categories ^{6,16,39}.

Another emerging axis of complexity is the expansion of feature-space by recent single-cell and spatial multi-omics technologies ²⁰. However, the vast majority of CCC methods are focused on inferring interactions from a single modality, typically transcriptomics ^{1,2,9}, with few methods considering multi-omics data ⁴⁰. Consequently, in order to transform such complex data into biologically-meaningful hypotheses, contemporary CCC methods are expected to address these emerging challenges (**Figure 5**).

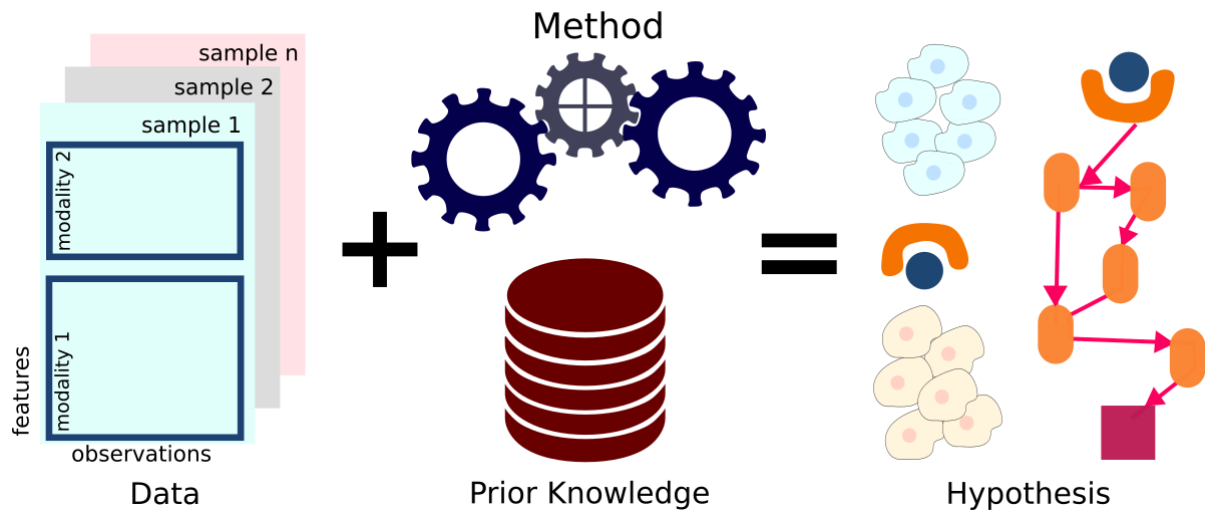


Figure 5. Modelling cell-cell communication from multi-sample single-cell omics data using prior knowledge

Motivation for an All-in-one Framework

While anticipated to provide a more informed view of CCC events, emerging technologies also prompt methodological developments - mostly done via publishing new tools. This is reflected by the ever-growing number of tools for the analysis of single-cell data which, at the time of writing, stands above 1,600 according to <https://www.scrna-tools.org/>⁴¹, of which > 100 are CCC inference tools⁴². This large number of tools, combined with sporadic development and often questionable coding practices, creates challenges for researchers when choosing the tool for their analysis. An issue that is further exacerbated by the lack of gold standard and standardised platform for benchmarking novel methods in the CCC field. In this thesis, I describe my attempt to evaluate and subsequently standardise contemporary CCC inference into a single, user-friendly framework - LIANA+, which at the time of submission of my thesis encompasses the vast majority of the aforementioned developments.

Chapter 2: Cell-cell Communication Inference from Single-cell RNA-Seq Data: a Comparison of Methods and Resources

This chapter is a preliminary and modified version of a preprint, which at the time of writing served as a comparative overview of the cell-cell communication inference field from single-cell transcriptomics data ⁴³. In this work, I describe the conceptualisation and initial usage of a framework which will later become LIANA (<https://github.com/saezlab/liana>) and subsequently LIANA+.

An edited and revised version of this work was later published at Nature Communications: “Comparison of methods and resources for cell-cell communication inference from single-cell RNA-Seq data” ⁹. Explicit permission was granted by the editorial board of Nature Communication to adapt and use any text and figures presented here. All figures in this section are shared under CC licence ^{9,43}.

1. Introduction

The growing availability of single-cell RNA sequencing (scRNA-Seq) data has enhanced our understanding of the cellular heterogeneity of tissues. Moreover, spatially-resolved transcriptomics has recently emerged as a technology to measure gene expression while preserving the spatial organisation of cells ⁴⁴. These advancements have in turn led to an increased interest in the development of tools for cell-cell communication (CCC) inference. CCC typically refers to interactions between secreted ligands and plasma membrane receptors. This picture can be broadened to include secreted enzymes, extracellular matrix proteins, transporters, and interactions that require the physical contact between cells, such as cell-cell adhesion proteins and gap junctions ⁴⁵. For simplicity, all of these events involving protein-protein interactions are referred to as CCC. CCC events are essential for homeostasis, development, and disease, and their estimation is becoming a routine approach in scRNA-seq data analysis ⁴⁶.

A number of computational tools have emerged that can be classified as those that predict CCC interactions alone ^{4,6,47–54}, and those that estimate intracellular pathway activities related to CCC ^{7,8,55–57}. This work focuses on the former (**Table 1**). Typically, using gene expression information obtained by scRNA-Seq, cells are clustered according to their gene expression profiles and cell type identities are assigned to the clusters based on known cell-type markers. Then, CCC tools are used to predict intercellular crosstalk between any pair of clusters, one cluster being the source and the other the target of a CCC event. CCC events are thus typically represented as a one-to-one interaction between ‘transmitter’ and ‘receiver’ proteins, expressed by the source and target cell clusters, respectively. The information about transmitter-receiver binding is extracted from diverse sources of prior knowledge. Every tool has two major components: a resource of prior knowledge on CCC (interactions), and a method to estimate CCC from the known interactions and the dataset at hand. Most tools have been published as the combination of one resource and one method, but in principle any resource could be combined with any method.

Despite the aforementioned common premises to explore CCC events, each tool uses a different method, such as permutation of cluster labels, regularizations, and scaling, to prioritise interactions according to the input datasets (**Table 1**). In turn, these different approaches result in diverse scoring systems that are difficult to compare and evaluate. The difficulties are further exacerbated by the lack of an appropriate gold standard to benchmark the performance of CCC methods^{2,46}. Nevertheless, different strategies have been used to indirectly evaluate the methods' performance, including a presumed correlation between CCC activity and spatial adjacency^{6,8}, recovering the effect of receptor gene knockouts⁸, robustness to subsampling⁶, agreement with proteomics⁵³, simulated scRNA-Seq data⁵⁰, and the agreement among methods^{6,8,51,53}.

The available prior knowledge resources, largely composed of ligand-receptor, extracellular matrix, and adhesion interactions, are typically distinct but often show partial overlap^{45,58}. Some of these resources also provide additional details for the interactions such as information about protein complexes^{4,6,45,59,60}, subcellular localisation^{6,45}, and classification into signalling pathways and categories^{6,59} (**Supp. Table 2**). CCC resources are often manually curated and/or built from other resources, with varying proportions of expert curation and literature support^{45,58}. Some databases gather and harmonise the information contained in the individual resources⁴⁵. Despite the fact that CCC inference is constrained by the prior knowledge used, the impact of resource choice is largely unexplored. It thus remains unclear how the choice of resource and method affects the results and thereby the biological interpretation of CCC inferred from the scRNA-seq data.

Here, I systematically describe the comparison of all combinations of 15 resources and 6 CCC methods (**Figure 1**). First, I report the degree of overlap among resources and whether certain resources are biased toward specific biological terms, such as pathways. Then, I show how different combinations of resources and methods influence CCC inference, by decoupling the methods from their corresponding resources.

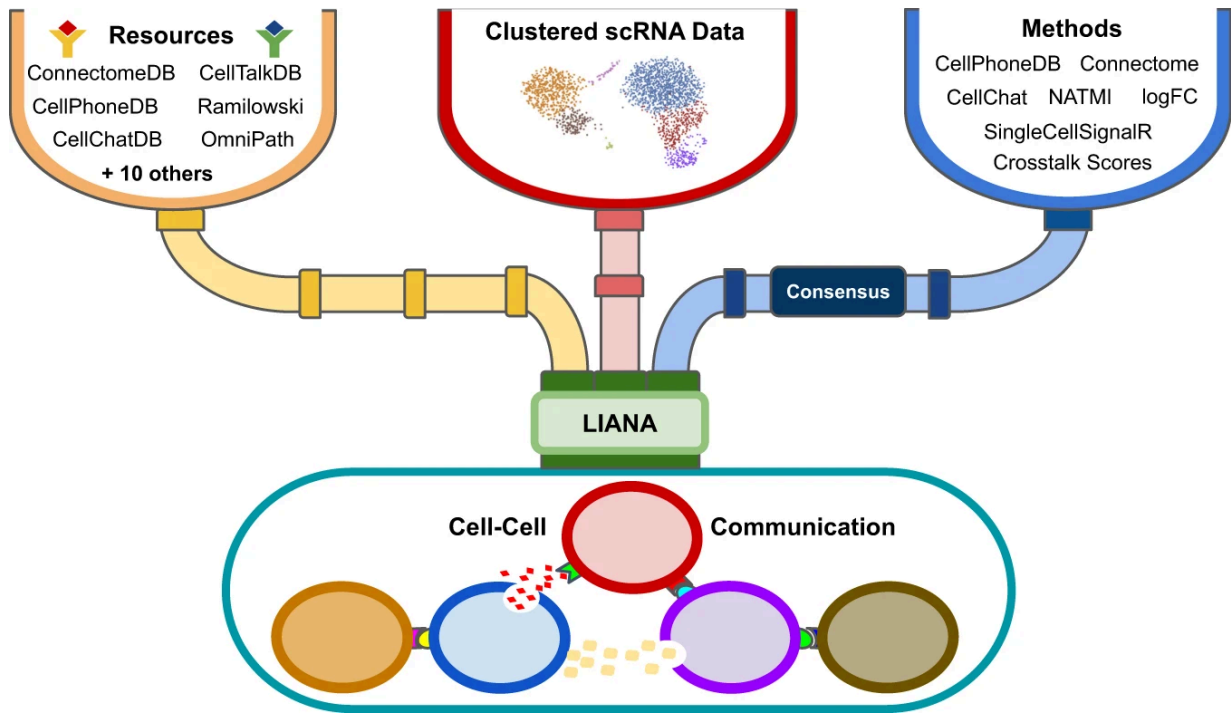


Figure 1. LIANA: a Cell-Cell Communication Framework (adapted from Dimitrov et al., 2022⁹).

Table 1. The tools included in the framework.

Each tool uses a resource and a method with specific score(s). Each method considers expression at the cell cluster level, and all of the scoring systems presented here are based on the expression of transmitters and receiver genes in the source and target cells, respectively.

Tool	Resource	Methods' scoring systems
CellChat ⁶	CellChatDB	1) Probability - based on the expression of differentially expressed transmitter and receiver genes and their mediators, calculated with the law of mass action 2) P-values† - significance identified via permutation of cell cluster labels and recalculating the probabilities for each cell pair and each transmitter-receiver interaction
Squidpy# ⁶¹ (CellPhoneDB ^{4,62,63})	OmniPath or CellPhoneDB	1) Truncated Mean - average expression of transmitter and receivers, the minimum expression (by default) of heteromeric complex of subunits 2) P-values† - significance identified via permutation of cell cluster labels to determine a null distribution of means for each receiver-transmitter interaction
Connectome ⁵¹	Ramilowski	1) weight_norm - derived via the product (by default) of the normalised expression of transmitter and receiver genes 2) weight_scale† - derived from the function (mean, by default) of the z-scores of the transmitter and the receiver, scaled according to cell cluster specificity
iTALK * ⁴⁸	iTALK	1) Expression logFC mean - required to make the scores of iTALK comparable to the other methods, and based on the differentially expressed transmitter and receiver genes between clusters. * Since the adaptation of logFC deviated from the intended usage of iTALK, iTALK was later replaced with a simple logFC mean method.
NATMI ⁵²	ConnectomeDB2020	1) Mean-expression edge weight - transmitter and receiver gene expression product 2) Specificity-based edge weight† - the mean expression of the transmitter and receiver are divided by the sum of the means of the same transmitters/receivers across all cell clusters
SingleCellSignalR ⁵³ (SCA)	LRdb	1) LRScore - a regularised score calculated using the squared expression of the transmitter and receiver (sqTRE) divided by sum of the mean of the count matrix and sqTRE
† Explicitly measures cell-cluster specific communication (referred as "cluster-specific measures") # The re-implementation of the CellPhoneDB method was run using 'Squidpy', even though Squidpy is a spatial transcriptomics framework with a much broader range of functionalities.		

2. Results

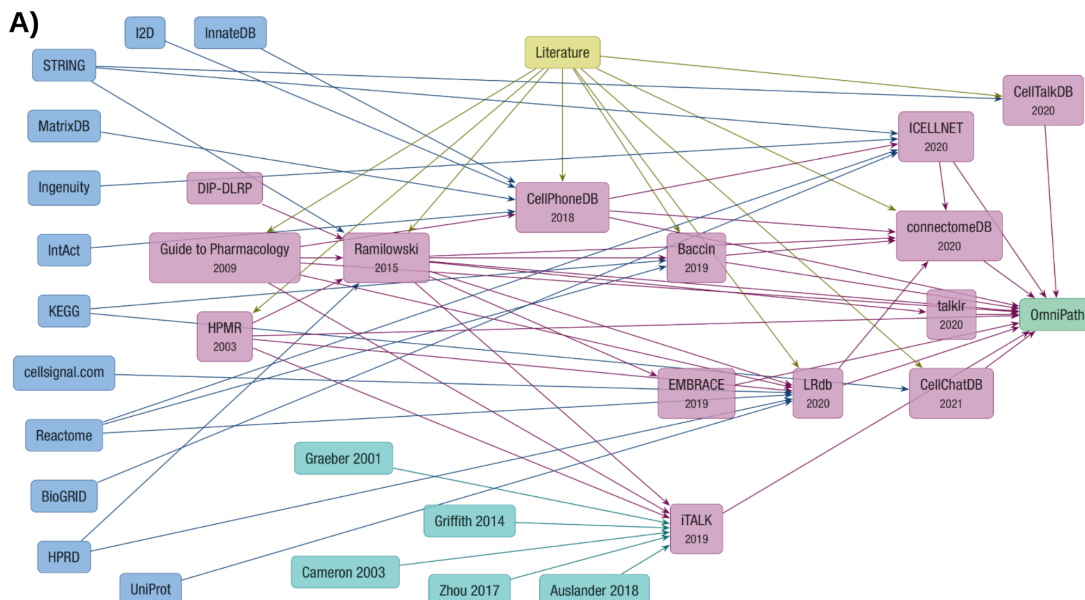
2.1 Resource Uniqueness and Overlap

To investigate the lineages of CCC resources, information about the origins of each resource were manually gathered. Many of these resources share the same original data sources, including general biological databases such as KEGG^{64,65}, Reactome⁶⁶, and STRING⁶⁷ (**Figure 2**). Moreover, interactions from Guide to Pharmacology⁶⁸, CellPhoneDB⁴, and in particular Ramilowski⁶⁹, were incorporated into subsequently published resources. All these

resources are integrated into OmniPath's CCC resource ⁴⁵, along with additional CCC interactions from other sources (e.g. SIGNOR ⁷⁰, Adhesome ⁷¹, Signalink ⁷², and others). A filtered version of OmniPath was also included ([4.1 Methods](#)).

As a consequence of their common origins, limited uniqueness across the resources was observed, with mean percentages of 4.6 unique receivers, 5.3 unique transmitters, and 16.8% unique interactions, for all resources (**Figure 2B**; **Supp. Table 2**). OmniPath and CellChatDB ⁶ had the largest degree of uniqueness, with 4, 16, and 46% for OmniPath and 17, 12, and 50% for CellChatDB in terms of receivers, transmitters, and interactions, respectively. Despite the low uniqueness among the resources, the pairwise overlap between them varied (**Figure 2C**; **Supp. Figure 1**). Particularly high similarity was observed between CellTalkDB ⁵⁸, ConnectomeDB ⁵², talklr ⁷³, iTALK* ⁴⁸, LRdb ⁵³, and Ramilowski (**Figure 2C**). The aforementioned resources, together with OmniPath, contained on average more than 65% the interactions present in the other resources (**Supp. Figure 2**), largely explained by each including a large proportion (>80%) of the interactions from Ramilowski. Heteromeric complex-containing CellChatDB, CellPhoneDB and Baccin ⁶⁰ showed limited similarity with other resources, as each included ~45% of the interactions present in any other resource, on average. The smaller resources ICELLNET ⁵⁴, Guide to Pharmacology, HMPR ⁷⁴ and Kirouac2010 ⁷⁵ were most dissimilar with the remainder of the resources and included on average only 21, 28, 17, and 7% of the interactions present in the other resources, respectively. The similarity among the resources was generally higher when considering transmitters, and receivers in particular (**Supp. Figures 1&2**).

In summary, these results indicated that due to the common origins of the resources, many of the transmitters, receivers, and interactions are not unique to any single resource. Yet, different resources include varying proportions of the collective CCC prior knowledge.



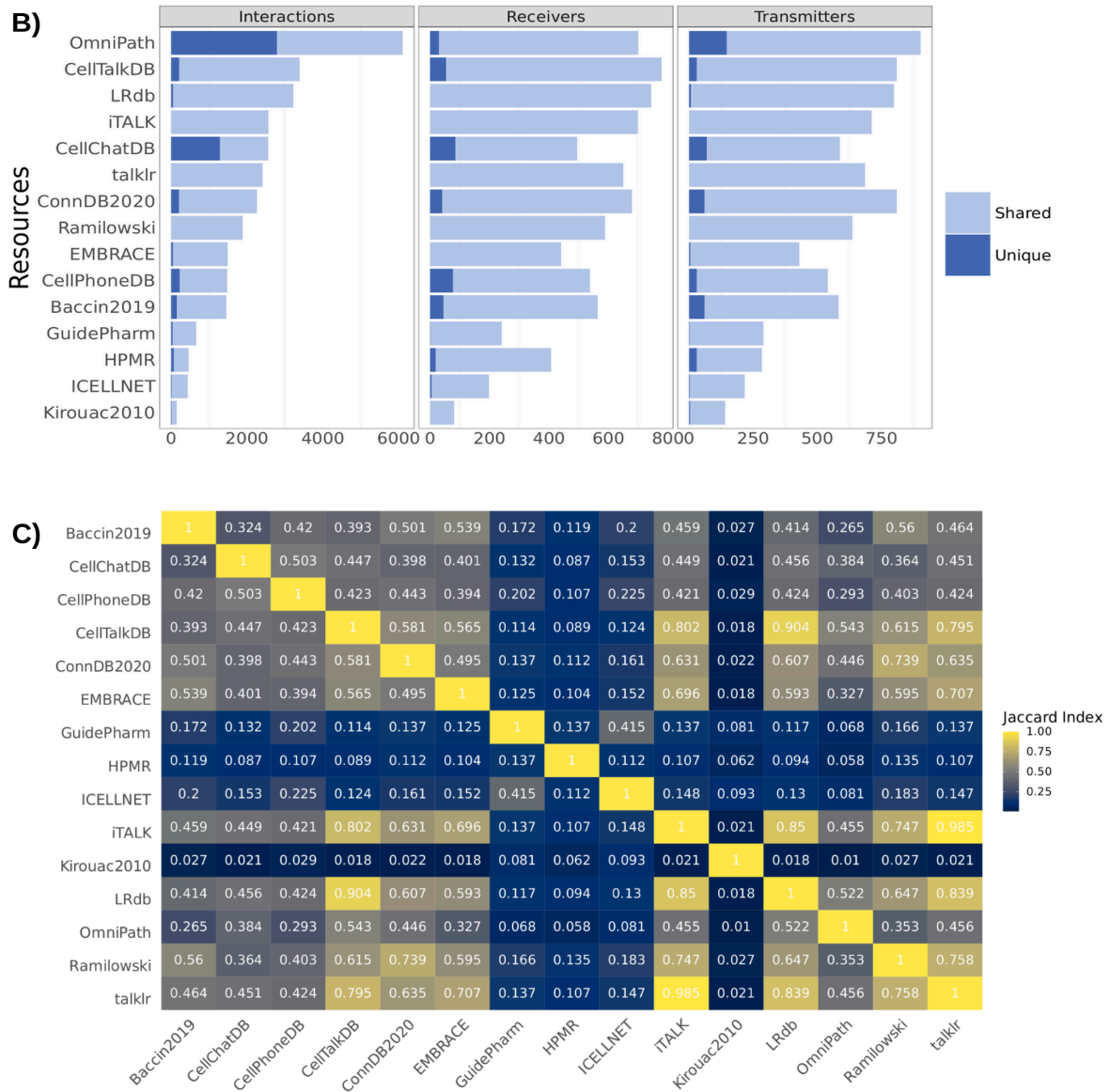


Figure 2. Dependencies and overlap between CCC resources. A) The lineages of CCC interaction database knowledge. General biological knowledge databases, CCC-dedicated resources used in this work, Literature curation, resources included in iTALK, and OmniPath are in blue, magenta, yellow, cyan, and green respectively. Arrows show the data transfers between resources. B) Shared and Unique Interactions, Receivers and Transmitters. C) Similarity between the interactions from different resources (Jaccard Index).

2.2 Resource Prior Knowledge Bias

Since CCC inference methods rely on prior knowledge to estimate intercellular communication events, the choice of resource and any potential bias in it is expected to impact the results. I therefore describe whether the resources are biased to specific subcellular locations or functional categories when compared to the collective knowledge contained in all of the resources.

2.2.1 Subcellular Localisation

Protein subcellular localisation annotations were obtained from OmniPath ⁴⁵, which combines this information from 20 resources. Then these annotations were matched to receivers and transmitters from each resource with the aim to assess the localisation profile of different resources. On average 90% of transmitters and 79% of receivers were annotated as secreted and transmembrane proteins, respectively (**Supp. Figure 3**). Localisations of transmitters and receivers were further used to categorise the interactions as secreted or direct-contact signalling, under the assumption that interactions between transmitters annotated as secreted and receivers annotated as membrane-bound represent secreted signalling events. On the contrary, an interaction between two membrane-bound proteins was assumed to be a direct-contact interaction. Building on these assumptions, I observed that all resources were predominantly (74% on average) composed of interactions associated with secreted signalling, while direct-contact signalling constituted a substantially smaller (16% on average) proportion of interactions (**Figure 3A**; **Supp. Figure 4**). Interactions categorised as neither secreted nor direct-contact were labelled as 'Other' and made up the remainder of the interactions. The proportions of secreted and direct-contact signalling varied between resources, as some of them, such as Baccin, ConnectomeDB, CellPhoneDB, HPMR, and OmniPath had an over-representation of direct-contact signalling when compared to the collective, while the opposite was noted for the case of secreted signalling (**Figure 3B**). Direct contact interactions were particularly under-represented in Guide to Pharmacology (4%), which was more focused on secreted signalling (87%). CellChatDB showed an overrepresentation of interactions matched to the category Other - a result consistent with the inclusion of interaction mediators within the database.

These results suggested that localisations of transmitters and receivers were largely uniformly distributed and that secreted signalling was predominant across all resources, with some differences between the relative abundances of secreted and direct-contact signalling.

2.2.2 Functional Term Enrichment

To examine whether specific pathways and biological functions are unevenly represented in specific resources, interactions, receivers and transmitters from each resource were matched to pathways and functional categories from Signalink ⁷² and NetPath ⁷⁶.

I saw that the Receptor tyrosine kinase (RTK), JAK/STAT, TGF, and WNT pathways covered the largest proportions of interactions matched to Signalink, with analogous results observed for receivers and transmitters (**Supp. Figure 5**). The interactions from Ramilowski, ConnectomeDB, LRdb, iTALK and talklr showed a similar pattern, which can be explained by the high overlap of these resources. On the contrary, interactions associated with innate immune pathways and T-cell receptor categories were under-represented in Guide to Pharmacology, Baccin2019, EMBRACE, Kirouac2010, ICELLNET, CellPhoneDB, and HMPR (**Figure 3C**). The innate immune pathway category was also diminished in OmniPath. In contrast, when using NetPath instead of Signalink to define the T-cell receptor pathway, the under-representation in Baccin2019 and OmniPath was not observed, and an over-representation was instead noted for ICELLNET and CellPhoneDB (**Figure 3D**; **Supp. Figure 6**). Moreover, a considerable over-representation was observed for the RTK pathway in OmniPath. The Signalink WNT pathway was under-represented in ICELLNET and Guide to Pharmacology, while for the NetPath WNT pathway this was only true for

Guide to Pharmacology. In contrast, CellChat showed a relative abundance for both the SignalLink and NetPath WNT pathways.

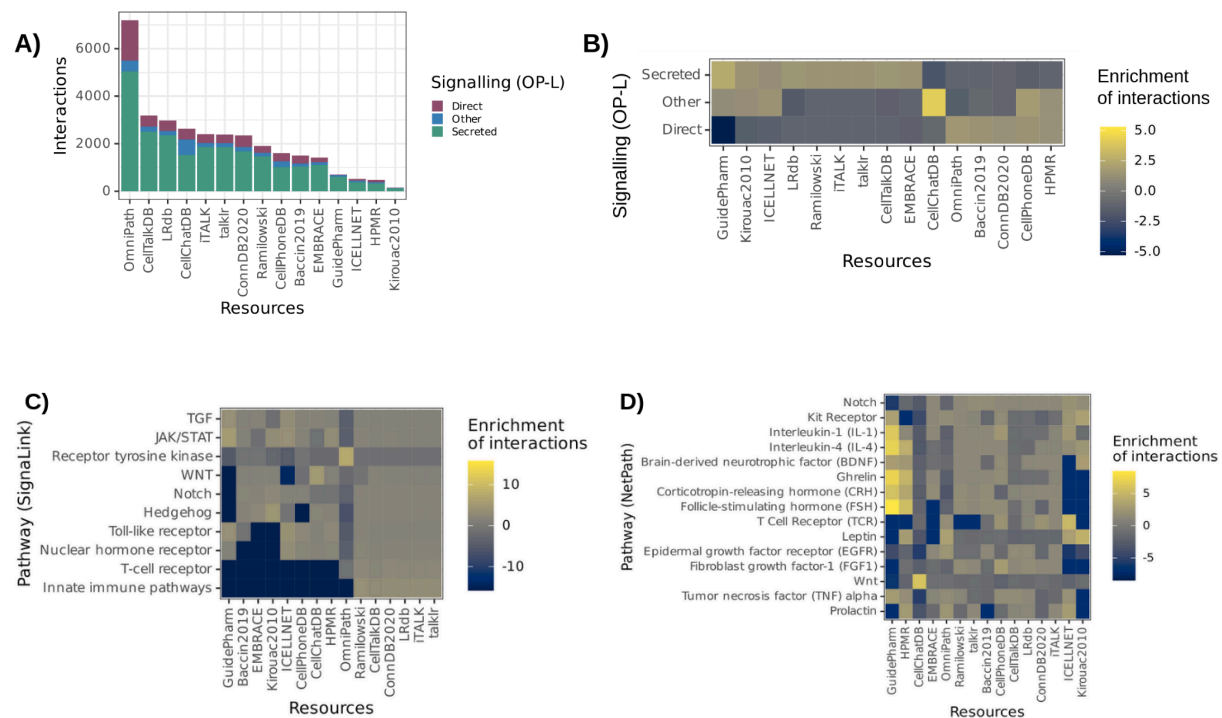


Figure 3. Localisation and Functional term abundance in CCC resources. A) Number and B) Relative abundance of signalling categories based on OmniPath-derived protein locations (OP-L) by resource. Relative abundance of C) SignalLink and D) NetPath annotations matched to interactions from each resource.

2.3 Agreement in CCC predictions using a Colorectal Cancer dataset

To examine the agreement between CCC methods and the importance of resources, I developed a framework to decouple tools from their inbuilt resources (LIANA). I used an annotated colorectal cancer (CRC) scRNA-Seq dataset ⁷⁷ with 65,362 cells from a heterogeneous cohort of 23 Korean CRC patients. I focused on the interactions between tumour cells subclassified by their resemblance of CRC consensus molecular subtypes (CMS) and immune cells from tumour samples (**Supp. Table 3**), reasoning that this subset of cell types represents a complex example where CCC events are known to have an important role. In addition to the 15 CCC resources reported in the descriptive resource analysis (**Supp. Table 1**), I also included the default or inbuilt resource for each of the tools, if available (**Table 1**), as well as a reshuffled control resource (4.3 Framework).

2.3.1 Interaction overlap

I then used each method-resource combination to infer CCC interactions, assuming that different methods should generally agree on the most relevant CCC events for the same resource and expression data. To measure the agreement between method-resource combinations, I examined the overlap between the 500 highest ranked interactions as predicted by each method. Whenever available, author recommendations were used to filter out the false-positive interactions (4.4 Method-Resource Specifics). This analysis showed

considerable differences in the interactions predicted by each of the methods regardless of the resource used, as the mean Jaccard index per resource ranged from 0.01 to 0.06 (mean = 0.024) when using different methods. These large discrepancies in the results were further supported by the pairwise comparisons between methods using the same resource, with mean Jaccard indices ranging from 0.063 (CellChat-SingleCellSignalR) to 0.110 (Connectome-NATMI). The overlap among the top predicted interactions was slightly higher when using the same method but with different resources, as Jaccard indices ranged from 0.113 to 0.203 per method (mean = 0.167) (**Supp. Figure 7**). Consequently, the highest ranked interactions for each method-resource combination largely showed stronger clustering by method than resource (**Figure 4**), with similar results observed when considering the highest ranked 100, 250, and 1000 interactions (**Supp. Figure 8**). In particular, method-resource combinations involving Squidpy, SingleCellSignalR, and Connectome clustered exclusively by method, suggesting that overlap is predominantly seen when using the same method regardless of the resource. The combinations involving NATMI also clustered by method, with the only exceptions being the Kirouac2010⁷⁵ and ICELLNET⁵⁹ resources, which were the smallest resources (**Supp. Table 2**).

These results suggested that the overlap between methods when using the same resource was low. In the revised version of this work, I showed that these results were consistent across six different datasets⁹. Hence, my results indicated that both the method and the resource had a considerable impact on the predicted interactions.

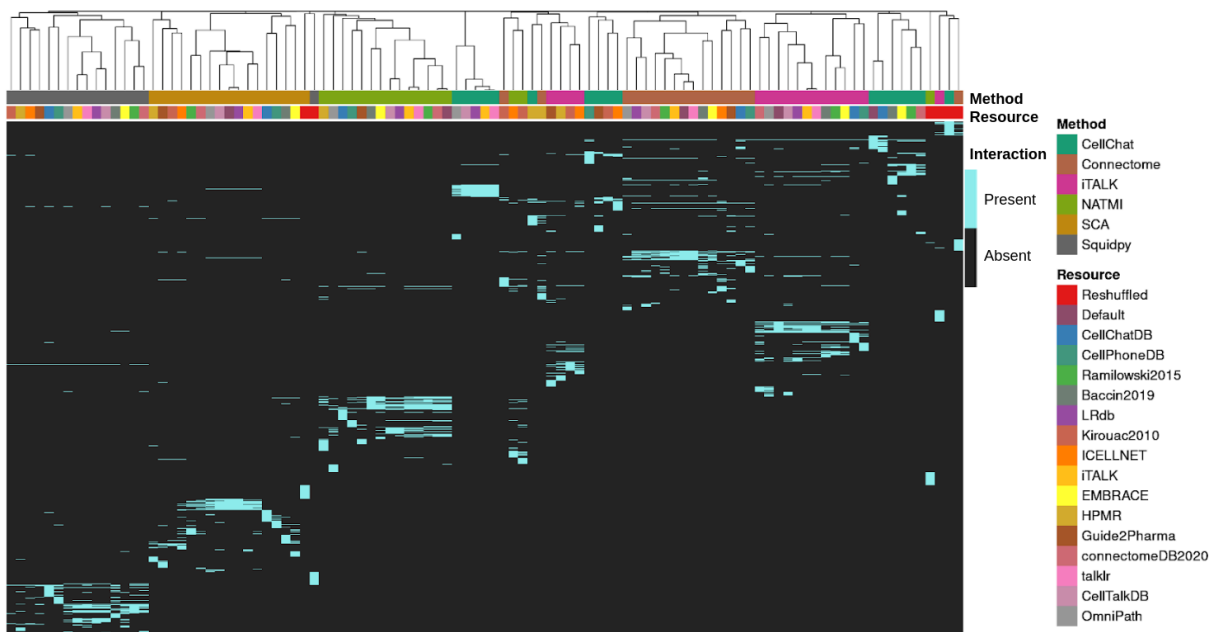


Figure 4. Overlap in the 500 highest ranked CCC interactions between different combinations of methods and resources. Method-resource combinations were clustered according to binary (Jaccard index) distances. SCA refers to the SingleCellSignalR method.

2.3.2 Communicating cell types

Next, I examined whether the discrepancies observed between the methods stem from the differences in the cell types inferred as most active in terms of CCC interactions. To this end, I used the 500 highest ranked interactions to examine the cell type activities, defined as the

proportion of interactions per cell type, separately as a source and a target of CCC events (**Figure 5**). The results largely reiterated our observations from the CCC interaction overlap analysis above, as each method largely clustered by itself, regardless of the resource used, including the reshuffled resource. These results were further supported by the average interaction ranks per communicating cell type pairs also grouping by method (**Supp. Figure 9**).

I reasoned that the observed disagreement in regards to the most actively communicating cell types was likely caused by the methods' distinct approaches to handle cell cluster specificity. I thus performed a complementary analysis using the alternative, non-cell-type-specific (magnitude) scoring systems of the methods. The higher agreement when using magnitude scores, instead of specificity, suggested that different specificity approaches are in part responsible for the observed disagreement (**Supp. Note 2; Supp. Figure 10**).

The analysis of activities per cell type largely reiterated the results from the interaction overlap analysis, particularly as each method largely clustered by itself, regardless of the resource. As a consequence, the disagreement between the methods in which cell types are the most active can have a major impact on the biological interpretation of CCC communication predictions.

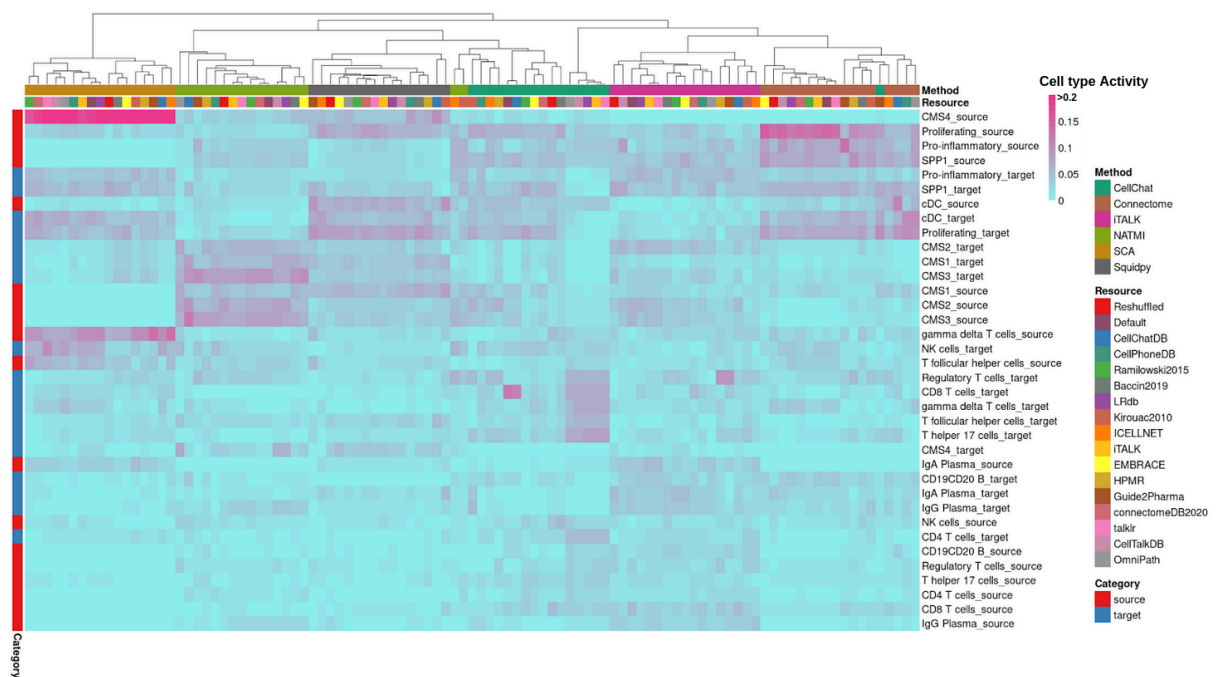


Figure 5. Activity per Cell type, inferred as the proportion of interaction edges that stem from Source Cell clusters or lead to Target Cell clusters in the highest ranked interactions.

3. Discussion

The growing interest in CCC inference has led to the recent emergence of multiple methods and prior knowledge resources to study intercellular crosstalk. To shed light on the impact of the choice of method and resource on the inference of CCC events, I built a framework to systematically combine 15 resources and 6 methods. Using this framework I described in detail the content of the different resources and to estimate cell-cell communication from scRNA-Seq in a colon cancer case study. The results from our work suggested that both the method and resource can considerably impact CCC inference.

Despite their common origins, different resources cover varying proportions of the collective prior knowledge. Particularly, a large share of the observed overlap among resources stemmed from the inclusion of Ramilowski⁶⁹ into other resources. Moreover, across the resources, the WNT, RTK, T-cell receptor and Innate immune pathways, among others, were present in varying proportions. The high abundance of interactions associated with the RTK pathway in OmniPath could be due to the ~1,600 expert curated RTK ligand-receptor interactions from SIGNOR⁷⁰ and the large size of RTK pathway in SignaLink⁷². The results presented here highlight an inherent limitation of knowledge-based inference, and hence of CCC methods, as any prior knowledge resource has its own biases and only represents a limited proportion of biological actuality. Consequently, these inherent limitations should be kept in mind for the interpretation of CCC predictions.

As a further step, I carried out a systematic analysis of the impact of resources and methods on CCC inference results using a public colorectal cancer dataset⁷⁷, and found that both resources and methods had a considerable effect on the predicted interactions, with the impact of methods outweighing that of the resource.

A potential explanation for the disagreement among the methods were the distinct approaches they use to identify the most relevant interactions (**Table 1**). A common assumption among the methods is that cluster-specific interactions are more informative than those related to multiple clusters^{4,6,51,52}. An experimental proof of this assumption and an evaluation of the distinct approaches is yet to be carried out. By focusing on the cluster-specific interactions in the dataset, these methods report the most specifically-interacting cell types⁵³, rather than the most actively communicating ones. Hence, the predicted CCC events typically do not capture processes that are common between multiple cell types. Collectively, these results suggest that the common practice to highlight the most actively communicating cell clusters based on the CCC inference^{77,78} should be considered with caution.

The disagreement between the methods possibly highlights certain limitations of the CCC inference methods. In particular, CCC events are mainly predicted based on the average gene expression at the cluster or cell type/state level. Such an assumption inherently suggests that gene expression is informative of the activity of transmitters and receivers. However, gene expression provided by scRNA-Seq is typically limited to protein-coding genes and the cells within the dataset, and hence does not capture secreted signalling events driven by non-protein molecules or long-distance endocrine signalling events. Further, CCC inference from scRNA-Seq data assumes that the product of the gene expression of a transmitter and a receiver is a good proxy for their joint activity, and thus does not consider any of the processes preceding transmitter-receiver interactions, including protein translation and processing, secretion, and diffusion.

The results presented here highlight the need to establish a benchmark to comprehensively assess the predictive power of CCC methods. However, a gold standard for benchmarking is currently not available and the biological ground truth is largely unknown^{2,46}. The field needs to identify experimental settings capable of establishing the biological ground truth. So far, intercellular interactions were mainly supported by the spatial colocalization of proteins and the functional deregulation of intracellular signalling¹, as well as the physical-interaction of cell types⁷⁹. Yet these approaches are only applicable for the post-hoc and indirect validation of CCC interactions. Thus, until an experimental gold standard becomes available, simulated datasets might be used instead. However, any in silico benchmark is by definition only a simplified approximation of reality, with its own biases⁸⁰. To our knowledge, appropriate benchmarks for resources and methods used in CCC inference are yet to be defined, although some proposals exist². In a direction similar to these proposals, I carried out some evaluations⁹, presented in **Chapter 3**.

Considerable efforts have been made to develop CCC inference, and further advancements will be key for the systems-level analysis of single-cell data. This will likely further increase by the rapidly emerging spatial transcriptomics⁴⁴ and single-cell proteomics⁸¹, and the future applications of CCC inference approaches to interspecies communication^{82,83}. Acknowledging the limitations of my work, it pointed at the interpretation inconsistencies that could arise as a consequence of the method and resource of choice.

4. Methods

4.1. Descriptive analysis of resources

The connections between resources shown in the dependency plot were manually gathered from the publications and the web pages of each CCC resource. The CCC resources used in the analyses were queried from the OmniPath database⁴⁵. The contents of the resources are identical to their original formats, apart from minor processing differences (**Supp. Table 1**). The OmniPath CCC resource is a composite resource which contains interactions from all of the CCC dedicated resources compared here, along with some additional resources⁴⁵. OmniPath's interactions were filtered according to the following criteria: i) we (co-authors and I) only retained interactions with literature references, ii) we kept interactions only where the receiver protein was plasma membrane transmembrane or peripheral according to at least 30% of the localisation annotations, and iii) we only considered interactions between single proteins (interactions between complexes are also available in OmniPath). OmniPath's intra- and intercellular components are both available via the OmnipathR package (<https://github.com/saezlab/OmnipathR>).

We defined unique and shared interactions, receivers and transmitters between the CCC resources if they could be found in only one or at least two of the resources, respectively. We used pheatmap⁸⁴ to generate the heatmaps.

To identify uneven distributions of transmitters, receivers, and interactions toward biological terms or protein localisations, we used Fisher's exact test to compare each individual resource to the collection of all the resources. We obtained protein localisations from OmniPath which collects this information from 20 databases⁴⁵. Then we kept consensus

protein localisations where at least 50% of the annotations agreed. We classified CCC interactions using the localisation combinations of proteins involved in the interactions, which included secreted, plasma membrane peripheral and transmembrane proteins. Interactions, receivers and transmitters were independently matched to the 10 pathways from SignaLink ⁷², and the 15 largest categories from and NetPath ⁷⁶. Each of the aforementioned general functional annotation databases was also obtained via OmniPath. In case of signalling pathway databases (SignaLink and NetPath), we focused on the enrichment of annotations matched to interactions. Annotation matches for transmitters and receivers were examined independently using the aforementioned functional annotation databases. Our approach allowed the same protein or interaction to be matched to multiple pathways or functional categories from the same database.

To enable a comparison of annotations across resources, we expanded protein complexes from Baccin2019 ⁶⁰, CellChatDB ⁶, CellPhoneDB ⁴, and ICELLNET ⁵².

4.2. Single-cell Transcriptomics data

The processed single cell RNA-Seq data ⁷⁷ for 23 Korean colorectal cancer patients is available at [GSE132465](https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE132465). The analysis presented here focused on the CCC interactions between colorectal cancer subtypes and immune cells, and the remainder of cell types, including unknown immune cell subtypes, were filtered out. This resulted in a subset of 18 cell types and 42,544 cells. I kept the original subtype labels, reformatted the names to work with each CCC method (**Supp. Table 3**), and sparsified the counts into a Seurat ⁸⁵ object.

The labelled scRNA-Seq data for pancreatic islet ⁸⁶ and cord blood mononuclear cells ⁸⁷ were obtained via SeuratData, normalised with Seurat ⁸⁵, and used for CCC inference without any further formatting and filtering.

4.3. Framework

For the method-resource comparison, I used Seurat ^{85,86} objects which were converted into the appropriate data format when calling each method. I used the recommended conversion method or wrapper whenever available.

The resources were obtained from OmniPath and then converted to the appropriate format for each method. A reshuffled version of ConnectomeDB2020 was generated with BiRewire ⁸⁸ and referred to as the reshuffled control resource. Each tool was run with its default or inbuilt resource, except Squidpy. The Default resource of Squidpy's **ligrec** function is OmniPath, which is already part of our benchmark set. The LIANA framework enabling the use of any resource and method combination is available at <https://github.com/saezlab/liana>.

4.4. Overlap Analysis

To compare the overlap between the interactions predicted by each method-resource combination, as a default I kept the 500 highest ranked interactions. I also considered the highest ranked 100, 250, and 1000 interactions. In the case of ties, I considered all ties. I then generated a presence-absence matrix of predicted interactions with method-resource combinations. These matrices were subsequently used to calculate Jaccard indices and to cluster the results.

Activity per cell type was calculated using the highest ranked 500 hits for each method-resource combination. Cell type activity represented the proportion of interactions (or edges) that stem from or lead to a Source or Target cell type, respectively. In other words, a Source cell with a high cell type activity, in the broadest terms, can be inferred as an active 'secretor of ligands'. I used the z-normalised average interaction rank for each possible combination of communicating cell types to estimate the cell pair ranks for each method-resource combinations. These patterns of pairwise communication activities I presented in a PCA plot. I created the heatmaps with pheatmap⁸⁴ (v1.0.12), using binary distances for the overlap heatmaps and euclidean distances for the other heatmaps.

Connectome, NATMI, and iTALK* do not provide an explicit threshold to control for false positives and the highest ranking 500 hits were kept for each without any preceding filtering. For methods where a threshold was proposed by the authors, as in the case of CellChat, Squidpy, and SingleCellSignalR, I first filtered their results accordingly and the highest ranked interactions were obtained afterwards. Further, I used cluster-specific interaction measures for each method whenever available (**Table 1**).

The same analysis was also carried out using the cluster-unspecific measures from each method. The scaling done in Connectome (weight_scale) and NATMI (Specificity-based edge weight), and in particular the cluster label permutation of CellChat (p-values) and CellPhoneDB (p-values), explicitly reflect cell-cluster specific communication, thus I used their alternative measures. SingleCellSignalR and iTALK* provide a single measure each and were hence excluded from this analysis.

Some methods, namely CellChat⁶ and the CellPhoneDB algorithm⁴, as well as resources, such as Baccin, CellChatDB, CellPhoneDB, and ICELLNET, take protein complexes into account. This largely complicates the conversion of the resources and hence the comparison with methods and resources which do not consider complexes. Furthermore, CellChat, and hence CellChatDB, goes a step further than other methods and resources, as it considers interaction mediator molecules, which are absent in the remainder of the resources⁶. Thus, even though any resource can be used with any method, some combinations put certain methods at a disadvantage.

4.5. Method Specifics

4.5.1 CellChat

CellChat was run using default settings with 1000 permutations and the gene expression diffusion-based smoothing process was omitted. CellChat returned a number of significant interactions with p-values of 0 ranging from 221 to 12,208 depending on the resource (2,988 with its inbuilt resource), these made a considerable proportion of the significant hits (p-value ≤ 0.05), as they ranged from 237 to 12,971 (3,041 with its inbuilt resource). As such, because obtaining the highest ranked interactions based on p-values was infeasible, CellChat results were filtered according to p-values (p-value ≤ 0.05) and the highest probability scores were instead used in the method-resource analysis.

4.5.2 Connectome

Connectome was run with its default settings using a Seurat object with processed gene expression counts. Results were filtered for differentially expressed genes (p-value ≤ 0.05),

as identified via a Wilcoxon test, and Connectome's scaled weights were used in the method-resource analysis.

4.5.3 iTALK

iTALK was run with its default settings using the 'DEG' option which returns corrected p-values and logFold changes for each gene. Then transmitters and receivers with q-value ≤ 0.05 were kept. A differential expression product was calculated using z-scores of transmitters and receivers and subsequently used in the method-resource analysis.

4.5.4 SingleCellSignalR

SingleCellSignalR was run with the processed gene counts, considering differentially expressed genes with a log2 fold change threshold of 1.5 or above. The highest LRscores which passed the recommended threshold of 0.5 were used in the method-resource comparison. The number of interactions predicted by SingleCellSignalR ranged between 159 to 7,240 (LRscore ≥ 0.5). The source code of SingleCellSignalR was modified to work with external resources (available at https://github.com/CostaLab/SingleCellSignalR_v1).

4.5.5 NATMI

NATMI's implementation is command-line based, thus a system command was invoked via R that calls the NATMI python module and passes the appropriate command line arguments. NATMI was run with its default settings using the processed gene expression matrix, converted from Seurat, and the specificity-based edge weights were used in the method-resource comparison. NATMI's lrc2p resource was used as the default.

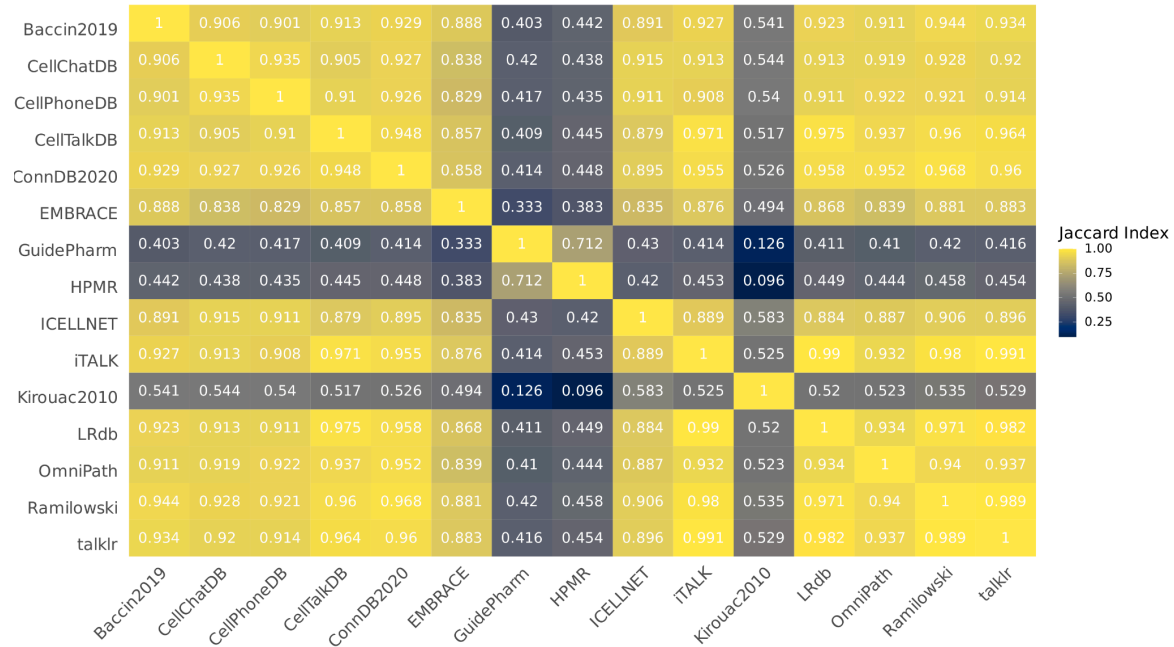
4.5.6 Squidpy

Squidpy is called via reticulate⁸⁹ (<https://rstudio.github.io/reticulate/>) and the Seurat object is converted to anndata⁹⁰ (<https://anndata.readthedocs.io/>) format in Python. The CellPhoneDB algorithm implementation was run via the Squidpy framework with 10,000 permutations, threshold of cells expressing transmitters and receivers of 0.1, and the minimum component expression was considered for complexes. For the method-resource comparison, I used the rank of p-values (p-value ≤ 0.05). Squidpy's number of significant hits ranged between 60 to 2,927 depending on the resource.

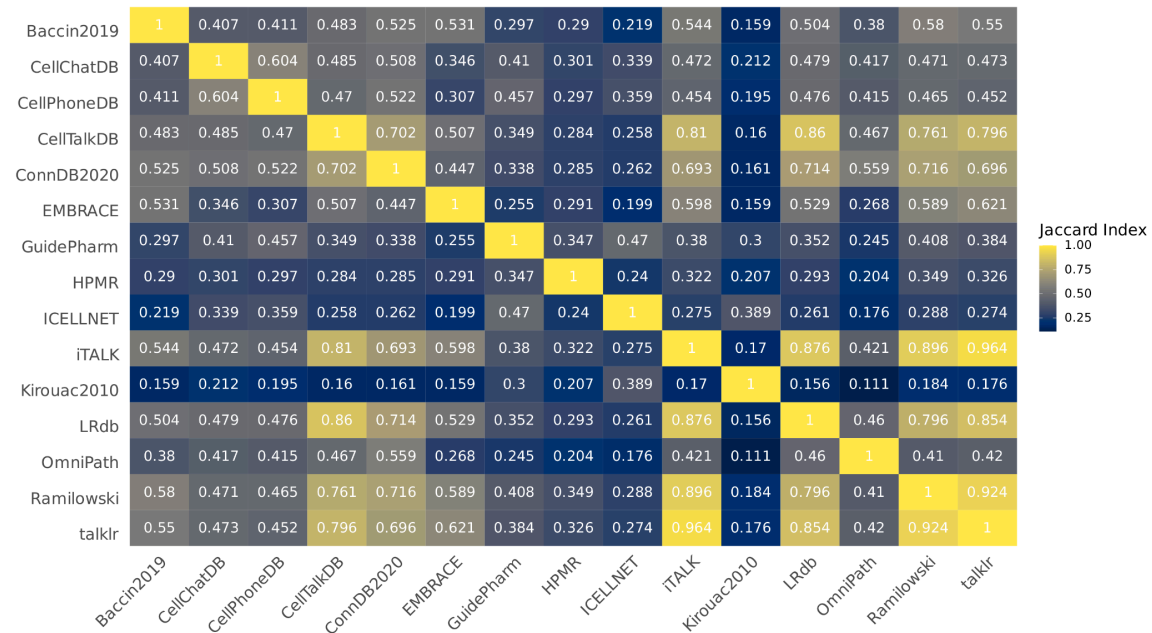
Supplementary Materials

Supplementary Figure S1. Jaccard Indices of A) Receivers and B) Transmitters from different resources.

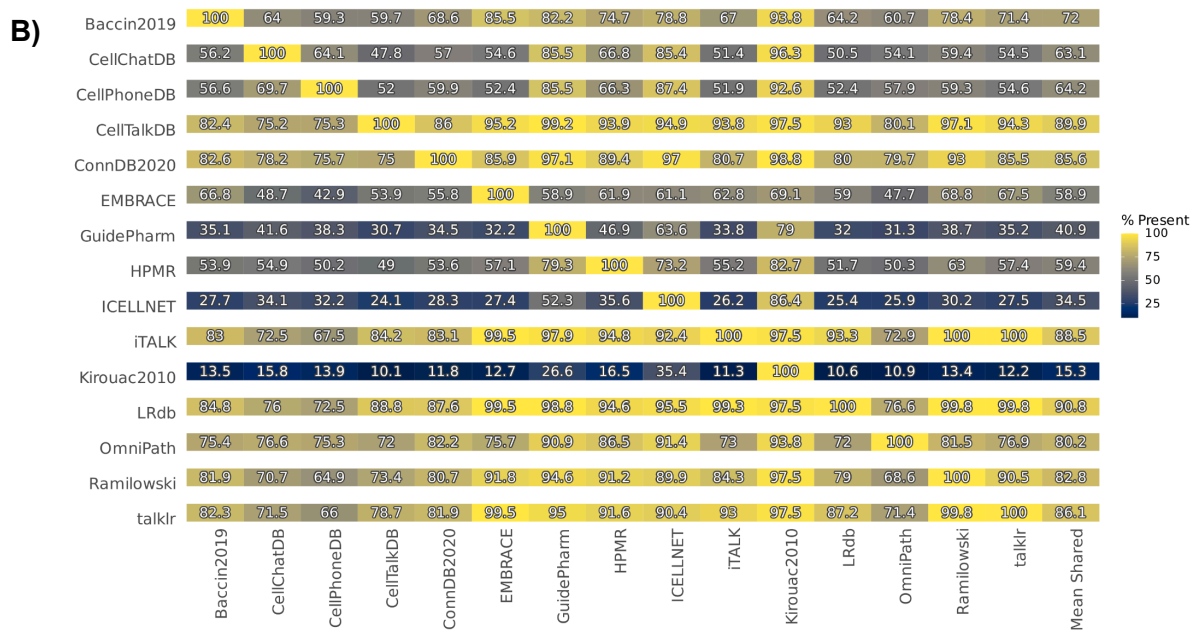
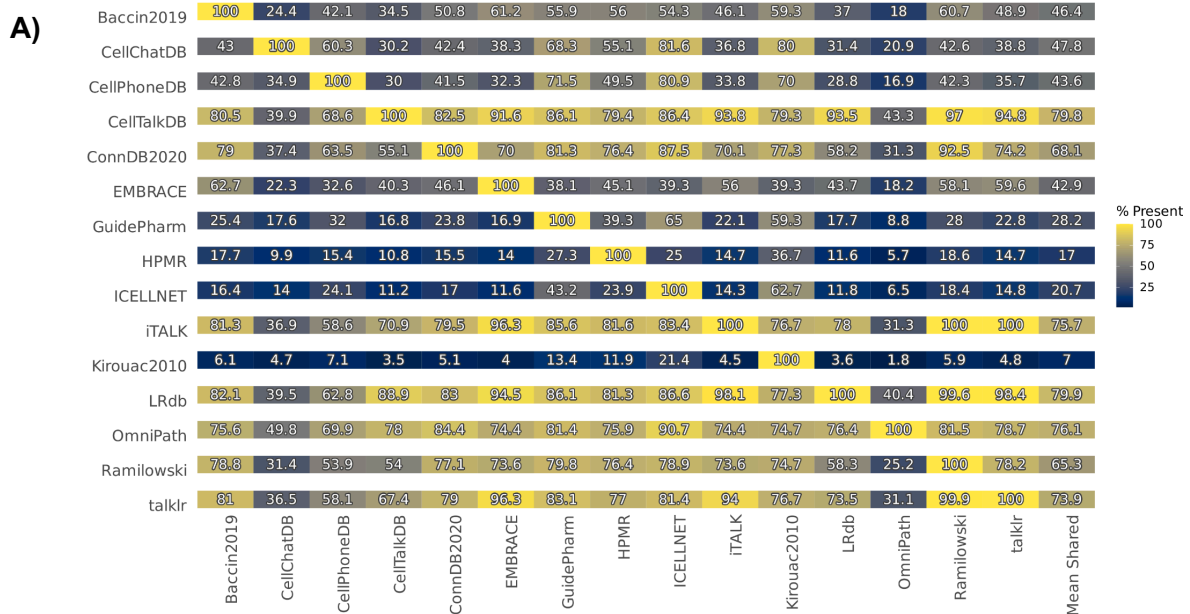
A) Receivers



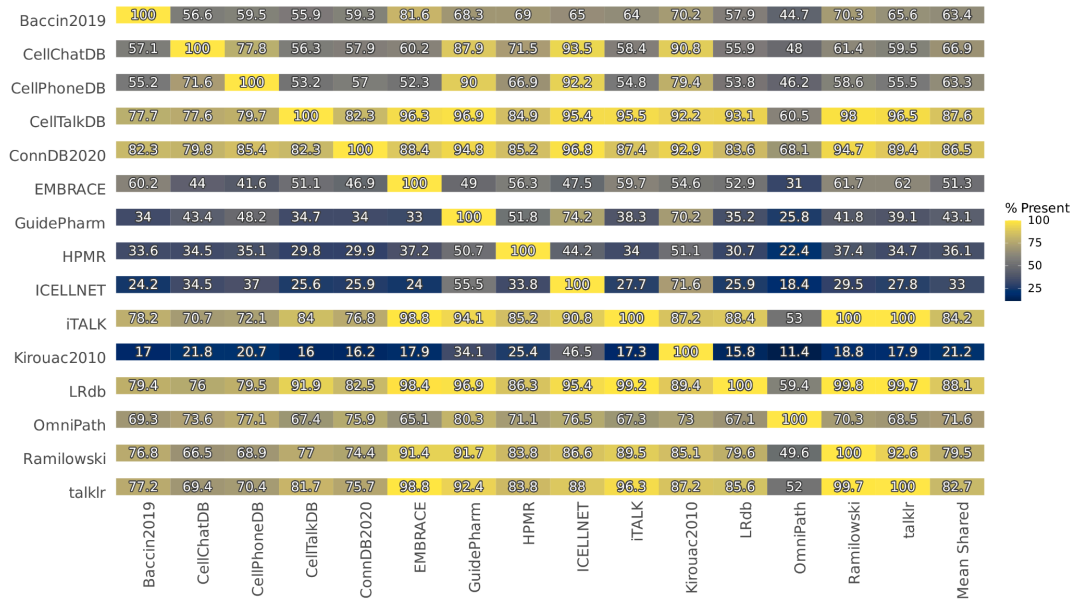
B) Transmitters



Supplementary Figure S2. A) Interactions B) Receivers and C) Transmitters present in each resource when taken from the rest of the resources. Note these plots are asymmetric and represent the % of interactions from the resources on the X axis found in each resource on the Y axis.

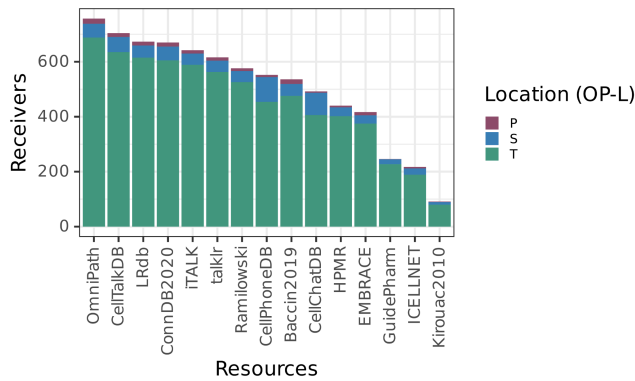


C)

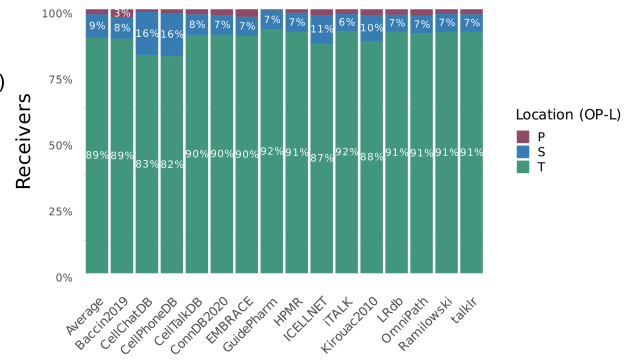


Supplementary Figure S3. Numbers and Percentages of Subcellular locations annotations of Receivers (A-B) and Transmitters (C-D) for each CCC resource. S, P and T stand for Secreted, Peripheral plasma membrane, and Transmembrane plasma membrane proteins, respectively.

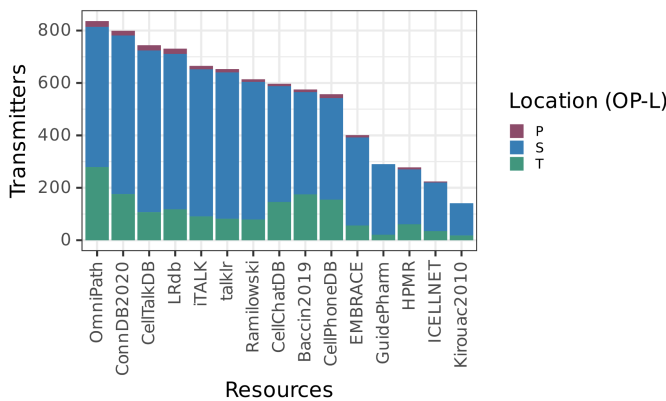
A)



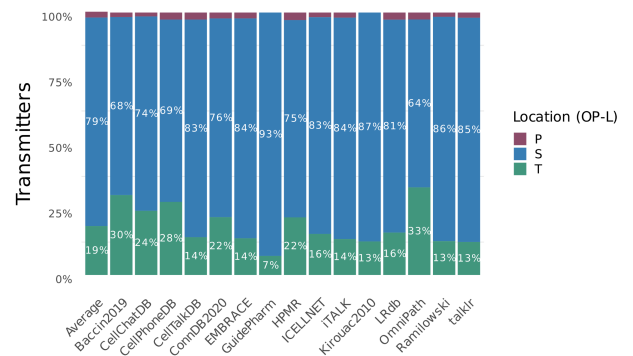
B)



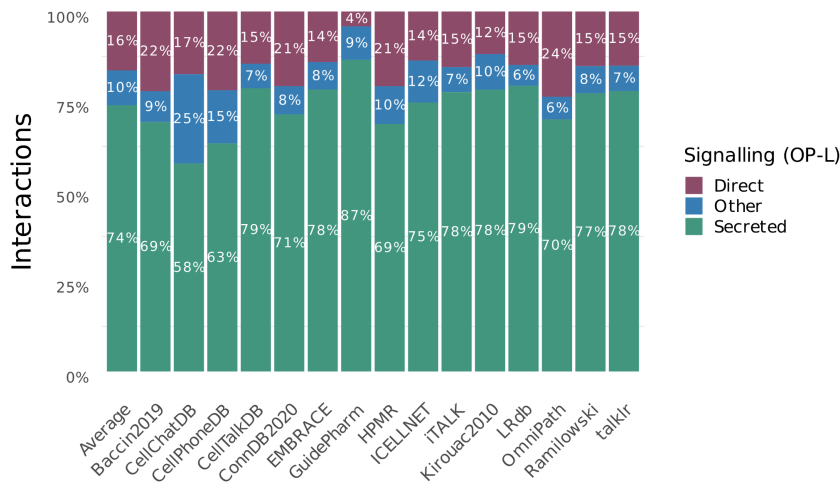
C)



D)

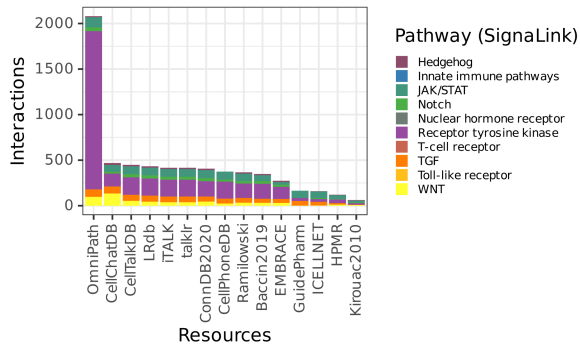


Supplementary Figure S4. Percentages per Signalling category according to OmniPath locations (OP-L) for each CCC resource.

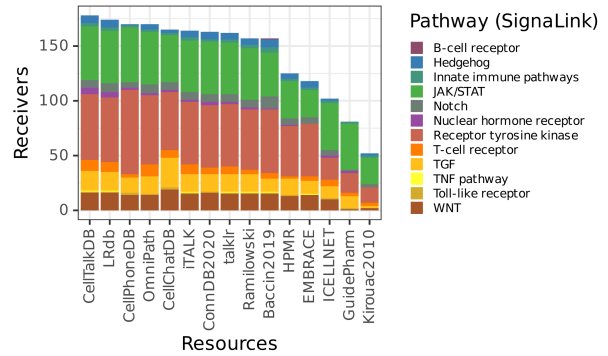


Supplementary Figure S5. Number of Interactions, Receivers and Transmitters, Enrichment Scores for their Receivers and Transmitters (D-E), and the Percentages of Interactions, Receivers and Transmitters (F-H) matched to the Signalink database per resource.

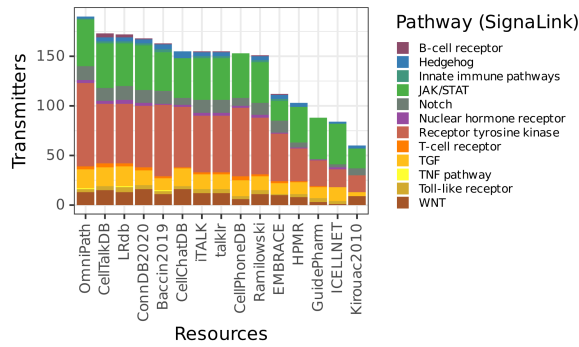
A)



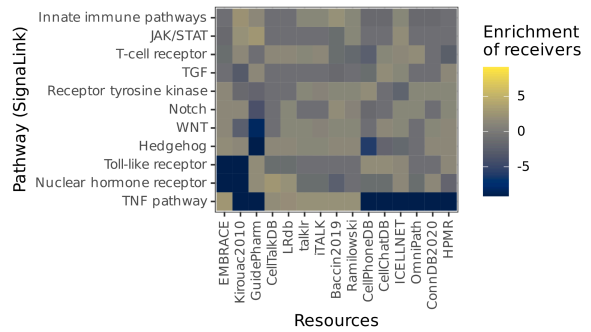
B)



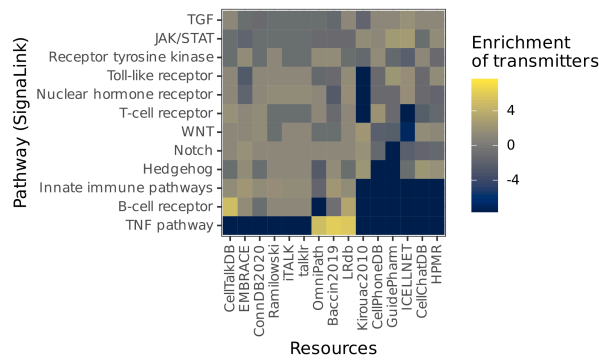
C)



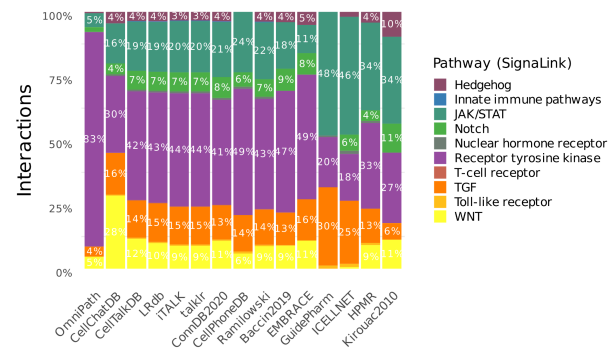
D)



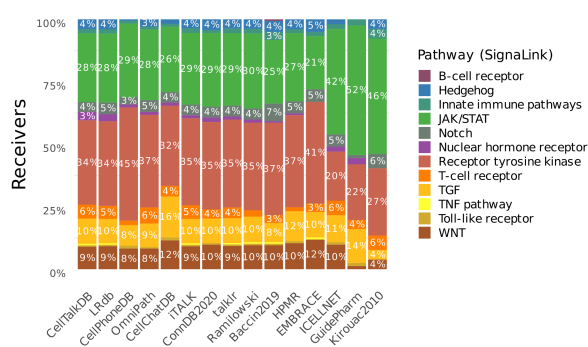
E)



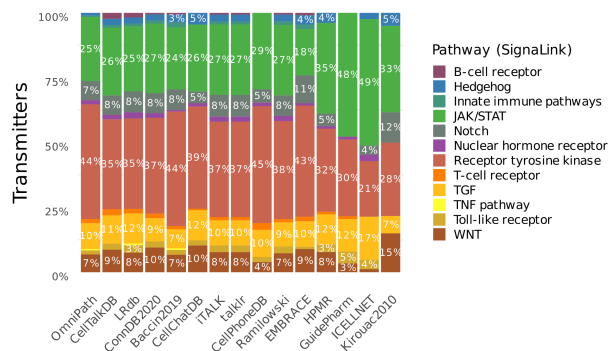
F)



G)

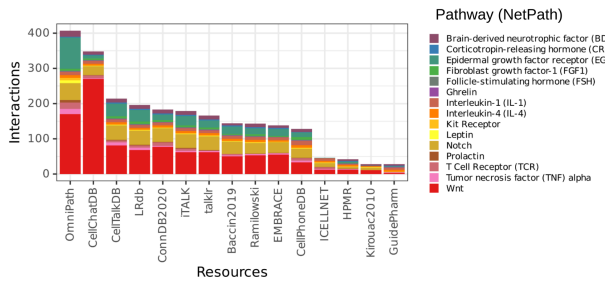


H)

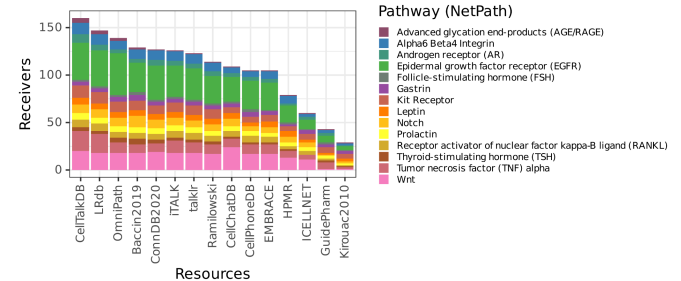


Supplementary Figure S6. Number of matches to A) Interactions, B) Receivers and C) Transmitters, Enrichment Scores for their Receivers and Transmitters (D-E), and the Percentages of Interactions, Receivers and Transmitters (F-H) matched to the NetPath database per resource.

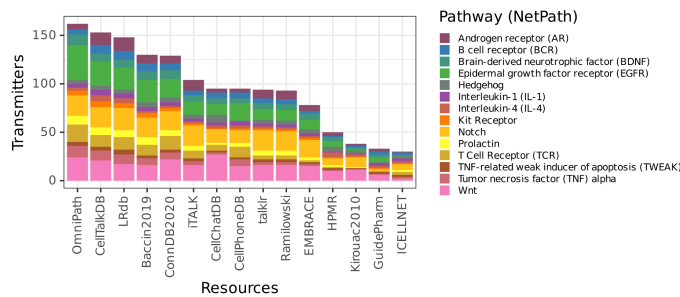
A)



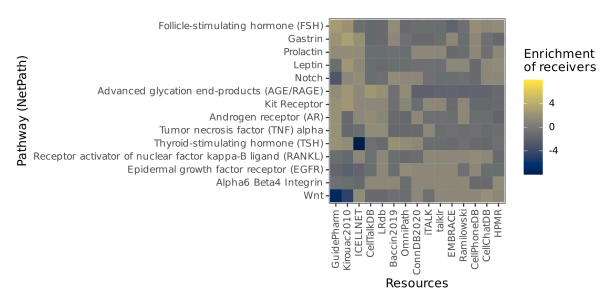
B)



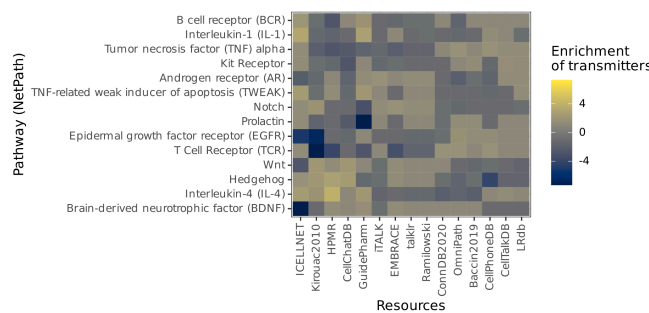
C)



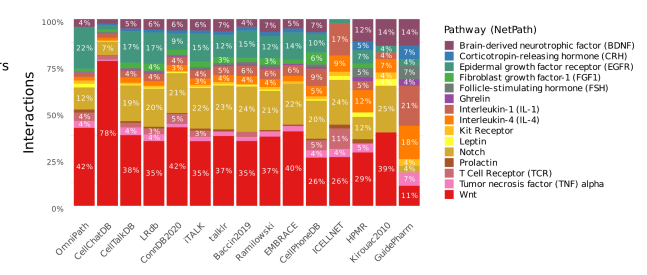
D)



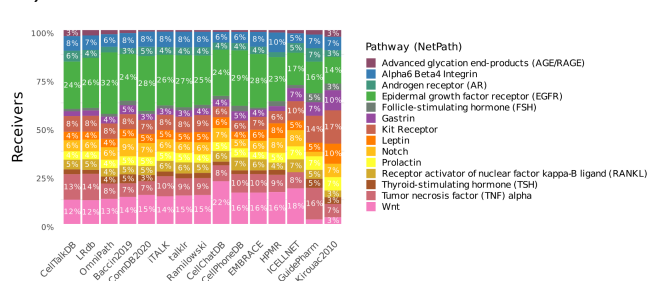
E)



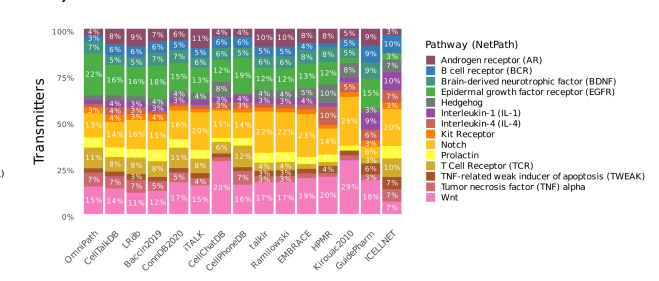
F)



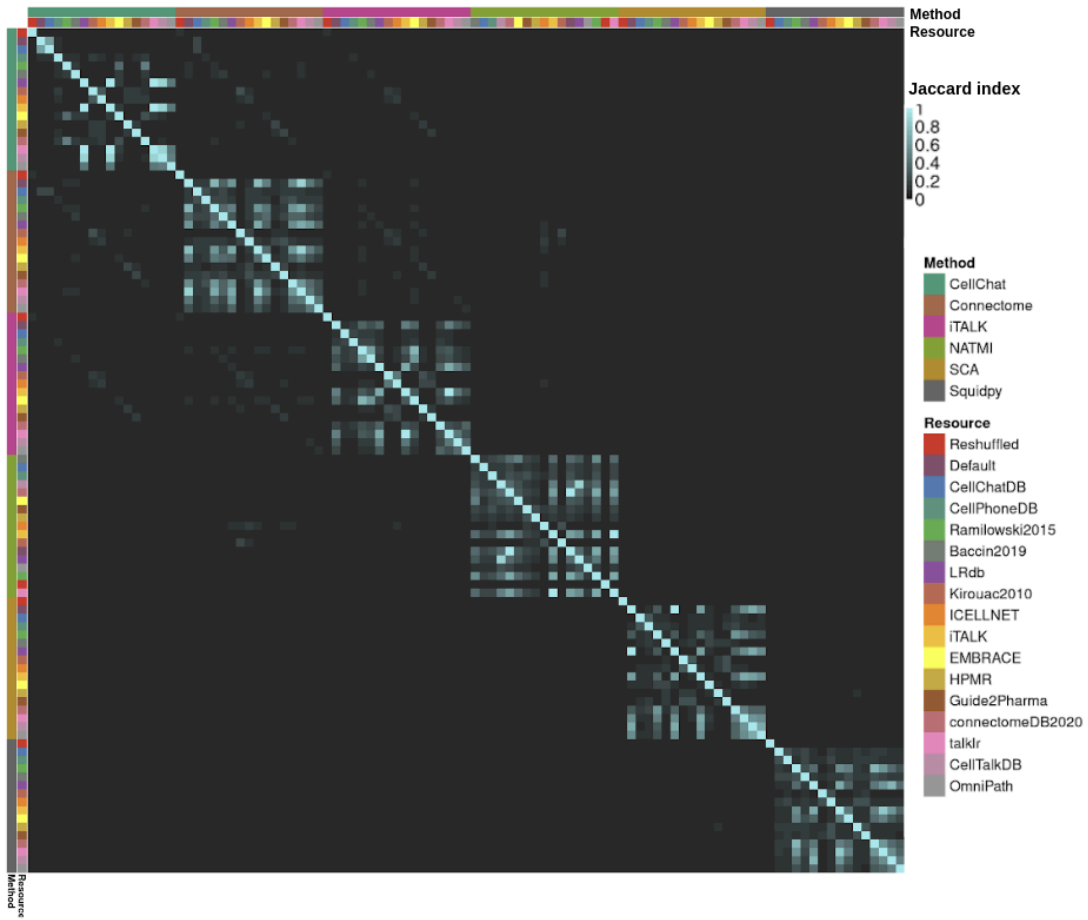
G)



H)

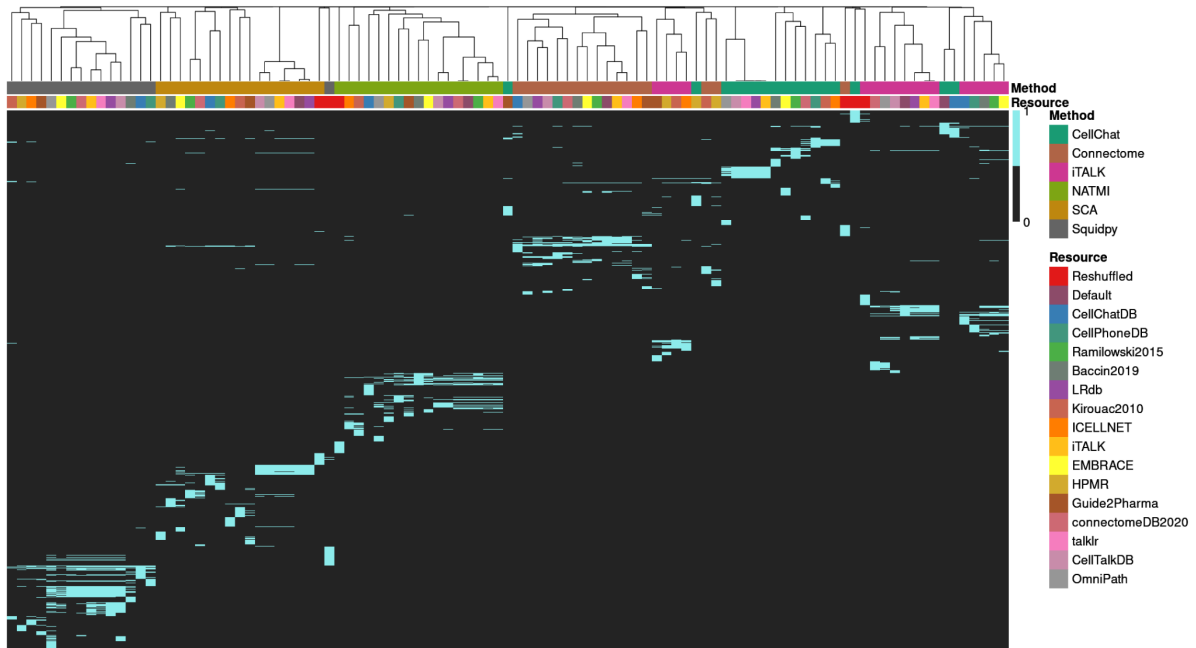


Supplementary Figure S7. Jaccard indices for the 500 highest ranked interactions obtained from each method-resource combination.

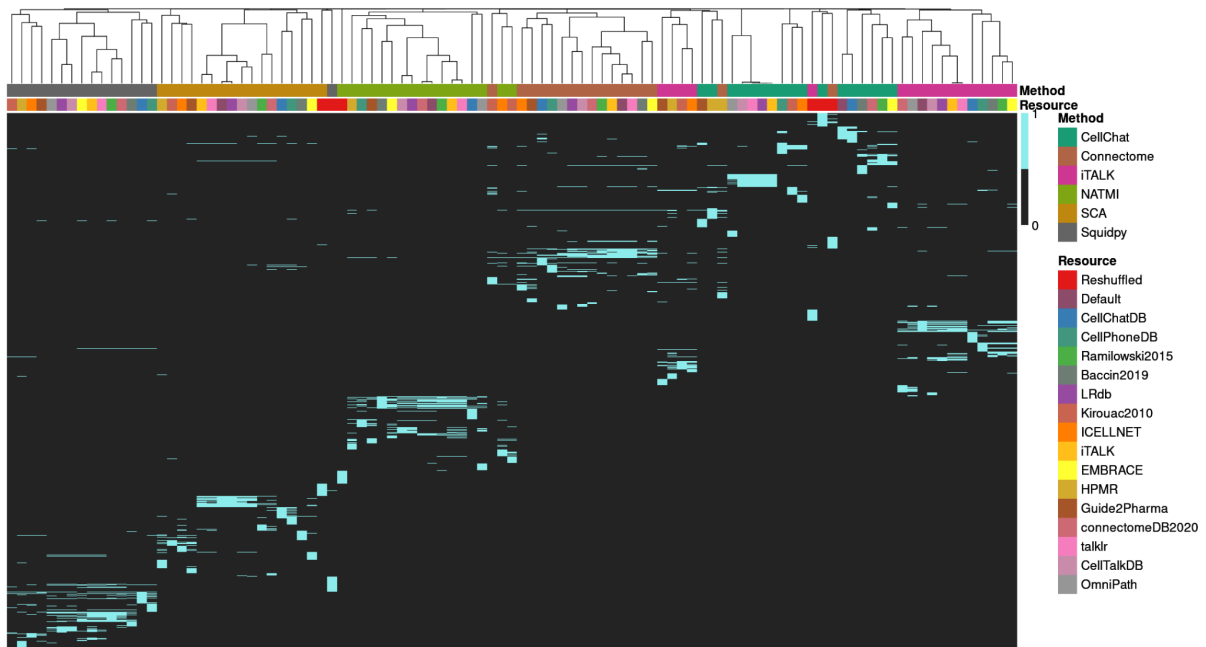


Supplementary Figure S8. *Overlap in the A) 100, B) 250, and C) 1000 highest ranked CCC interactions between different combinations of methods and resources. Method-resource combinations were clustered according to binary (Jaccard index) distances. SCA refers to the SingleCellSignalR method.*

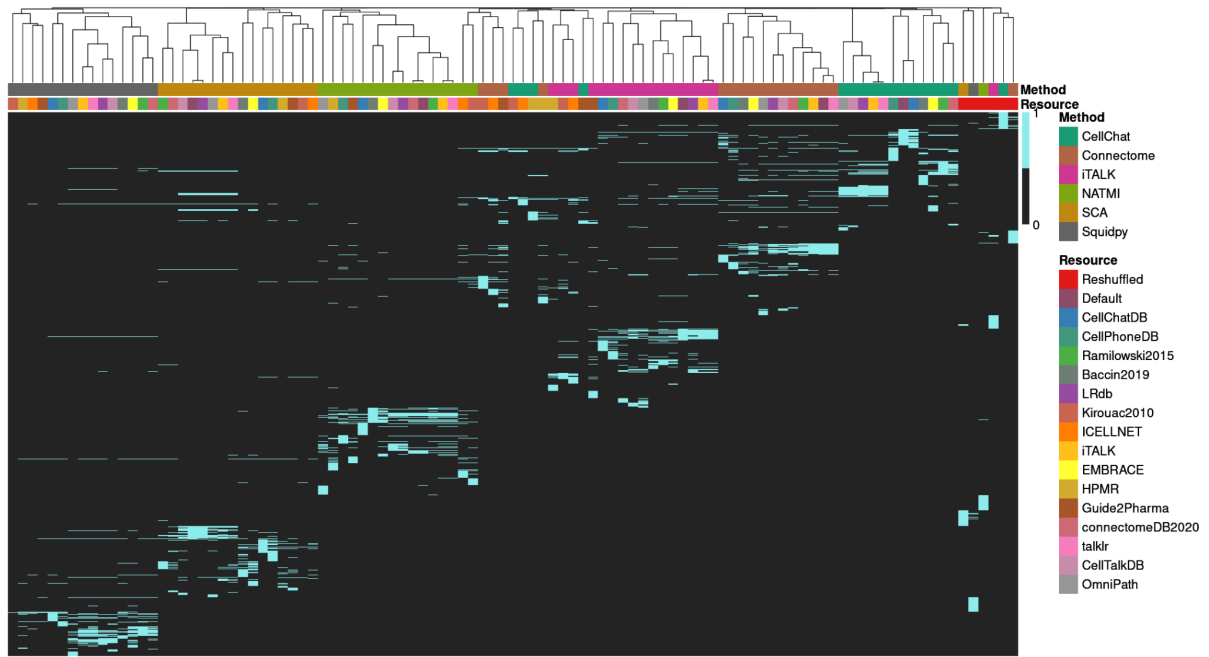
A)



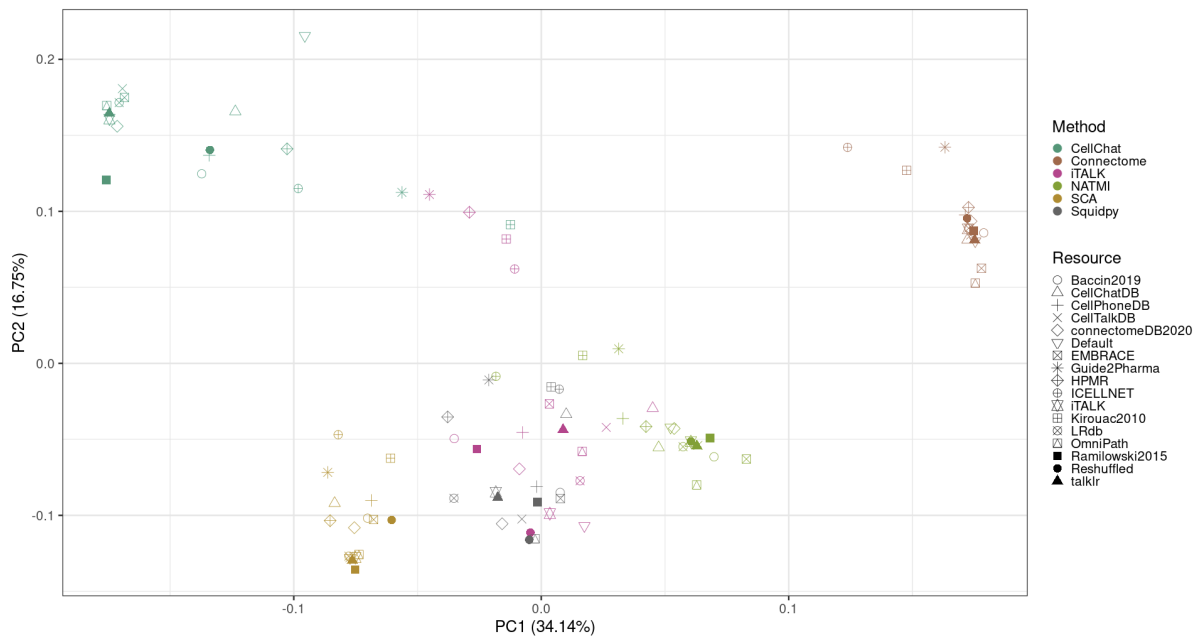
B)



C)

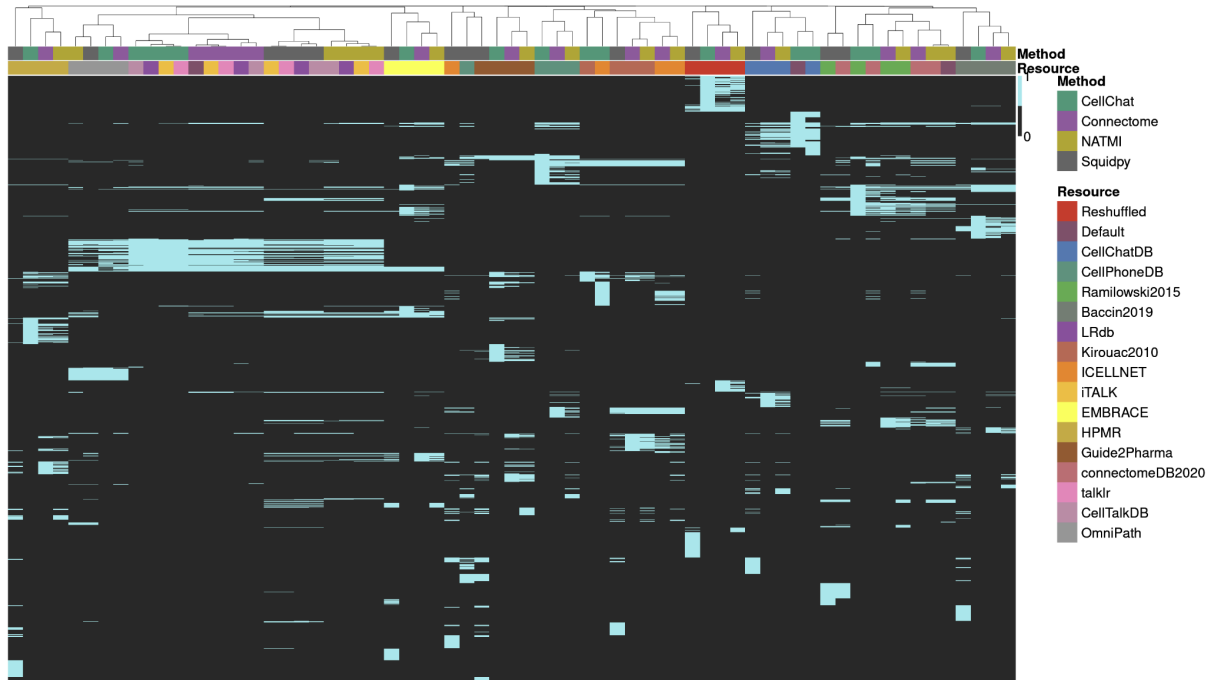


Supplementary Figure S9. PCA of normalised average interaction rank frequencies per cell pair

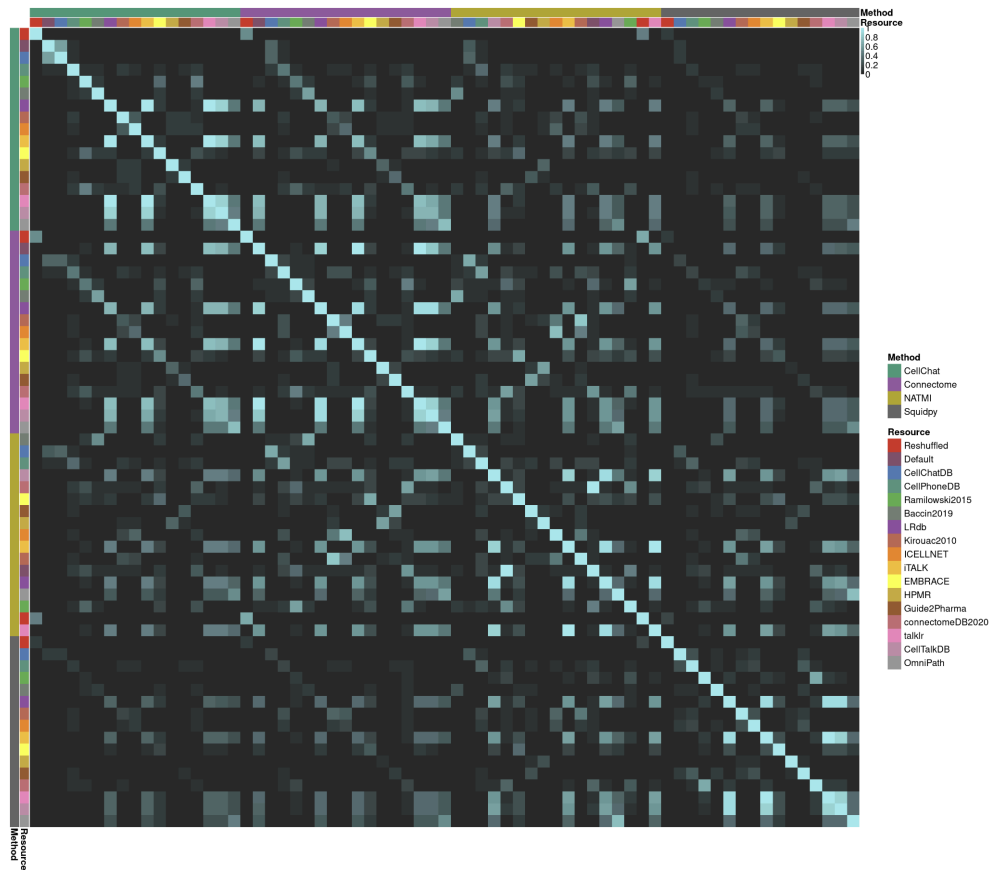


Supplementary Figure S10. Communication agreement among Magnitude-focused scores. A) *Overlap in 500 highest ranked interactions between different combinations of methods and resources. B)* *Similarity among the highest ranked interactions for each method-resource combination, as measured by Jaccard index. C)* *Activity per Cell type, inferred as the proportion of interaction edges that stem from Source Cell types or lead to Target Cell types in the highest ranked interactions.*

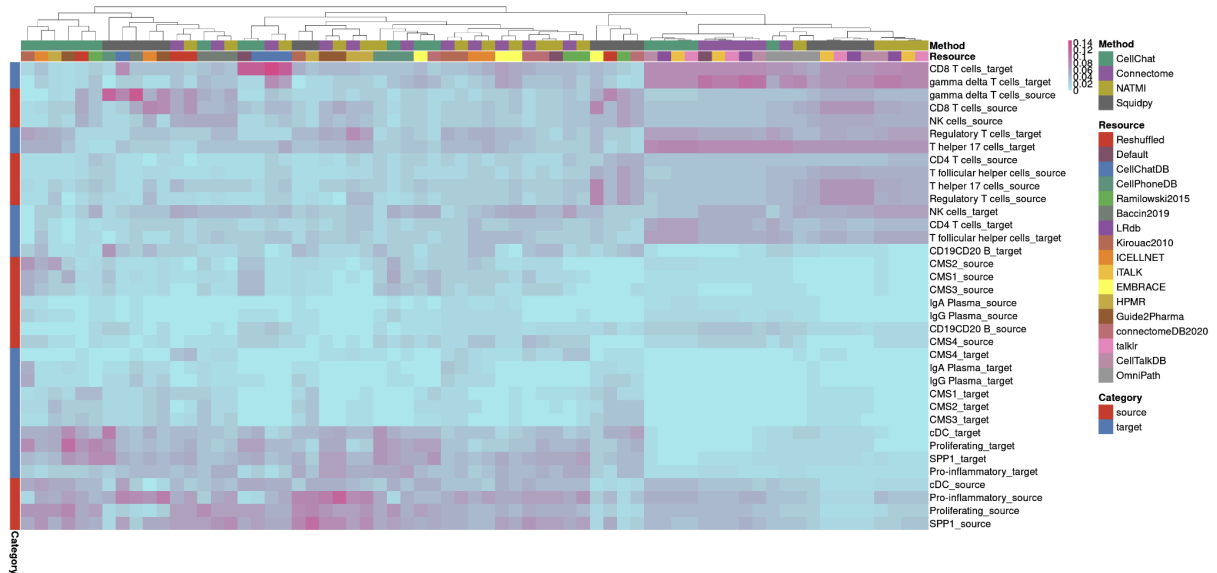
A)



B)



C)



Supplementary table 1. Description of existing resources for measuring cell–cell communication.

Resource	Further curation	Sources	Interactions*
Baccin2019(a) ⁶⁰	Murine identifiers (only), Multimeric complexes	Ramilowski2015, KEGG reactome database, Literature	1978 (1418)
CellChatDB ⁶	Multimeric complexes, 229 signaling pathway families, agonists and antagonists, co-receptors, localisations, Murine identifiers	KEGG, Literature	2551 (2551)
CellPhoneDB ^{4,62,63}	Multimeric complexes, intercellular communication roles	Guide2Pharma, I2D, IMEx, InnateDB, IntAct, MatrixDB, MINT, UniProt, Literature	1397 (1312)
CellTalkDB ⁵⁸	Murine identifiers	STRING, Literature	3398 (3390)
ConnectomeDB2020 ⁵²	-	Ramilowski2015, CellphoneDB, Baccin2019, LRdb, ICELLNET, Literature	2293 (2264)
EMBRACE(a) ⁹¹	Murine identifiers	Ramilowski2015	1710 (1489)
Guide2Pharma(b) ⁶⁸	-	Literature	740 (662)
HPMR ⁷⁴	-	PubMed , GenBank	527 (461)
ICELNET ⁵⁹	Multimeric complexes, Signalling families, Cytokine-focus	STRING , Ingenuity , BioGRID , Reactome , CellPhoneDB	380 (371)
iTALK (c) ⁴⁸	Ligand categories	Ramilowski2015, HPMR, IUPHAR-DB , Graeber2001 , Griffith2014 , Cameron2003 , Zhou2017 , Auslander2018	2648 (2565)
Kirouac2010 ⁷⁵	-	Literature, COPE	270 (150)
LRdb ⁵³	-	cellsignal.com, Ramilowski2015, Guide2Pharma, HPMR, HPRD , Reactome, UniProt, Literature	3251 (3226)
Ramilowski2015 ⁶⁹	-	DLRP, HPMR, IUPHAR, HPRD, STRING, Literature	1894 (1888)

talklr ⁷³	-	-	2422 (2411)
OmniPath ^{#45}	Combines data from more than 100 resources and contains protein-protein and gene regulatory interactions, enzyme-PTM relationships, multimeric complexes, protein annotations, and a CCC-dedicated composite DB	Composite resource combining all of the CCC dedicated resources listed here, along with some additional interactions.	6103

All the resources above were retrieved from the OmniPath database (<https://omnipathdb.org/>). We also refer to the composite CCC resource presented here as OmniPath. The OmniPath presented in the analyses we filtered according to: **i)** we only retained interactions with literature references, **ii)** we kept interactions only where the receiver protein was plasma membrane transmembrane or peripheral according to at least 30% of the localisation annotations, and **iii)** we only considered interactions between single proteins

* The number of original interactions for each CCC dedicated resource covered in OmniPath is shown in brackets.

(a) Translated from murine identifiers to human, which accounts for the lower number of obtained interactions.

(b) Kept only the unique human-annotated interactions between transmitter and receiver proteins.

(c) Duplicates present in the original resource were excluded when imported via OmniPath

Supplementary table 2. Unique and shared Transmitters, Receivers, and interactions in each resource. Unique and shared interactions, receivers and transmitters between the CCC resources. Interactions were considered unique if they could be found in only one resource.

Resource	Transmitters	Receivers	Interactions
Baccin2019	10.29%	7.98%	10.52%
CellChatDB	11.73%	17.17%	50.29%
CellPhoneDB	5.55%	14.31%	15.27%
CellTalkDB	3.70%	6.93%	6.28%
ConnDB2020	7.41%	6.04%	9.14%
EMBRACE	1.16%	0.00%	3.42%
GuidePharm	0.69%	0.41%	5.89%
HPMR	10.56%	4.67%	15.84%
ICELLNET	1.38%	2.53%	3.64%
iTALK	0.00%	0.00%	0.08%
Kirouac2010	2.84%	0.00%	9.33%
LRdb	0.88%	0.00%	1.61%
OmniPath	16.28%	4.29%	45.70%
Ramilowski	0.00%	0.00%	0.00%
talklr	0.00%	0.00%	0.00%
Total	5.30%	4.57%	16.77%

Supplementary table 3. Formatted Korean CRC data set cell type counts and full names.

Cell type	Cell subtype	Complete Name	Cell Count
B cells	CD19CD20 B	B cells_CD19+CD20+	2,049
T cells	CD4 T cells	CD4+ T cells	3,980
T cells	CD8 T cells	CD8+ T cells	4,647
Myeloids	cDC	Conventional Dendritic Cells	353
Epithelial cells	CMS1	Consensus Molecular Subtype 1	1,201
Epithelial cells	CMS2	Consensus Molecular Subtype 2	10,771
Epithelial cells	CMS3	Consensus Molecular Subtype 3	5,486
Epithelial cells	CMS4	Consensus Molecular Subtype 4	11
T cells	gamma delta T cells	$\gamma\delta$ T cells	219
B cells	IgA Plasma	IgA+ Plasma Cells	180
B cells	IgG Plasma	IgG+ Plasma	1,661
T cells	NK cells	Natural Killer Cells	948
Myeloids	Pro-inflammatory	Pro-inflammatory Macrophages	2,325
Myeloids	Proliferating	Proliferating Macrophages	165
T cells	Regulatory T cells	Regulatory T cells	2,943
Myeloids	SPP1	SPP1+ Macrophages	3,096
T cells	T follicular helper cells	T follicular helper cells	548
T cells	T helper 17 cells	T helper 17 cells	1,961

Supplementary Note 1. Cluster Specificity and Method Dissimilarity

As a consequence of the disagreement between methods in regards to the most actively communicating cell types, I reasoned that a possible cause was the different approaches used to assign cell cluster specificity to the interactions. To this end, I conducted the same analyses presented in the main text, but instead using the measures from each method which do not explicitly reflect the cell-type specific communication (i.e. Squidpy means; unfiltered CellChat probabilities; Connectome.weight.norm; NATMI.edge.avg.expr) (**Table 1**). Since SingleCellSignalR and iTALK provide a single scoring system, they were excluded from this analysis. I observed an increase in the agreement between methods (**Supp. Figure 10A**), as the mean Jaccard index when using the same resource with different methods ranged from 0.277 to 0.618 (mean = 0.404) (**Supp. Figure 10B**). The overlap between these methods when using the same resource was hence considerably higher than that observed when using cluster-specific measures (mean = 0.0247). On the contrary, the mean Jaccard index per method when using the same resource remained relatively unchanged when compared to the scoring systems that reflect cell cluster specific communication (0.118 for non-specific measures versus 0.167 for specific). Moreover, analogously to the agreement analysis, I used the cluster-unspecific measures to estimate the active cell types. As a result, methods were observed to largely agree in terms of the most active cell types (**Supp. Figure 10C**). Thus, this analysis suggests that the distinct approaches used to assign cell cluster specificity to the interactions explain some of the disagreement between methods for our dataset. Furthermore, when using the cluster-unspecific measures, the differences in resources were the main source of dissimilarity between the results.

Chapter 3: Evaluating Cell-cell Communication Inference from Single-cell Transcriptomics Data Using Alternative Modalities

*This chapter serves as a snapshot of LIANA's development focusing on the evaluations done for its publication in Nature Communication ⁹. Here, I also present my contribution towards the formalisation of the same benchmarks into the OpenProblems live benchmarking platform (<https://github.com/openproblems-bio/openproblems>; Luecken, Gigante, Burkhardt et al., *In prep.*). The text and figures in this chapter were written and generated by me - Daniel Dimitrov.*

Explicit permission was granted by the editorial board of Nature Communication to adapt and use any text and figures presented here. All figures in this section are shared under CC licence ^{9,43}.

Background

The increased interest in modelling how cells interact from single-cell transcriptomics has drawn many computational biologists to develop tools for CCC inference ^{1,2}. Naturally, this prompts users to wonder which one generates the most reliable hypotheses; particularly, as I and others showed that the agreement between tools is limited, depending on both the resource ⁹ and method ^{9,92,93} used. Yet, the evaluation of CCC tools is hampered by the absence of a gold standard ^{1,2}, or more specifically the lack of a wide-spread, high-throughput technology capable of capturing CCC complexity as inferred from single-cell transcriptomics. Moreover, the comparison of the tools is challenging, as each comes with its own diverse scoring function(s) with distinct assumptions and preprocessing steps. To this end, I designed and developed LIANA to standardise the calling, resource, and output format of each method.

To evaluate the different methods, I used alternative data modalities, devised into three distinct tasks with the following complimentary information and assumptions (**Figure 1; Methods**):

- *Source-target task*; cell types in close proximity, as inferred from spatial transcriptomics data ²³, were anticipated to preferentially interact with each other over non-adjacent cell types
- *Ligand-Target task*; interactions which involve cytokine ligands were anticipated to reflect downstream cytokine signalling events, inferred in the target cell types ⁹⁴
- *Target-Receptor task*; receptors inferred to be involved in CCC events were anticipated to correspond to cell-surface receptor protein abundance, as measured via dual-modality CITE-Seq data ⁸⁷

For each task, I used multiple datasets, each with its own *a-priori*-defined assumed truth. Then after running LIANA, metrics such as the area under the receiver operating curve (AUROC), area under the precision recall curve (AUPRC) or standard odds ratios are calculated (**Figure 1; Methods**).

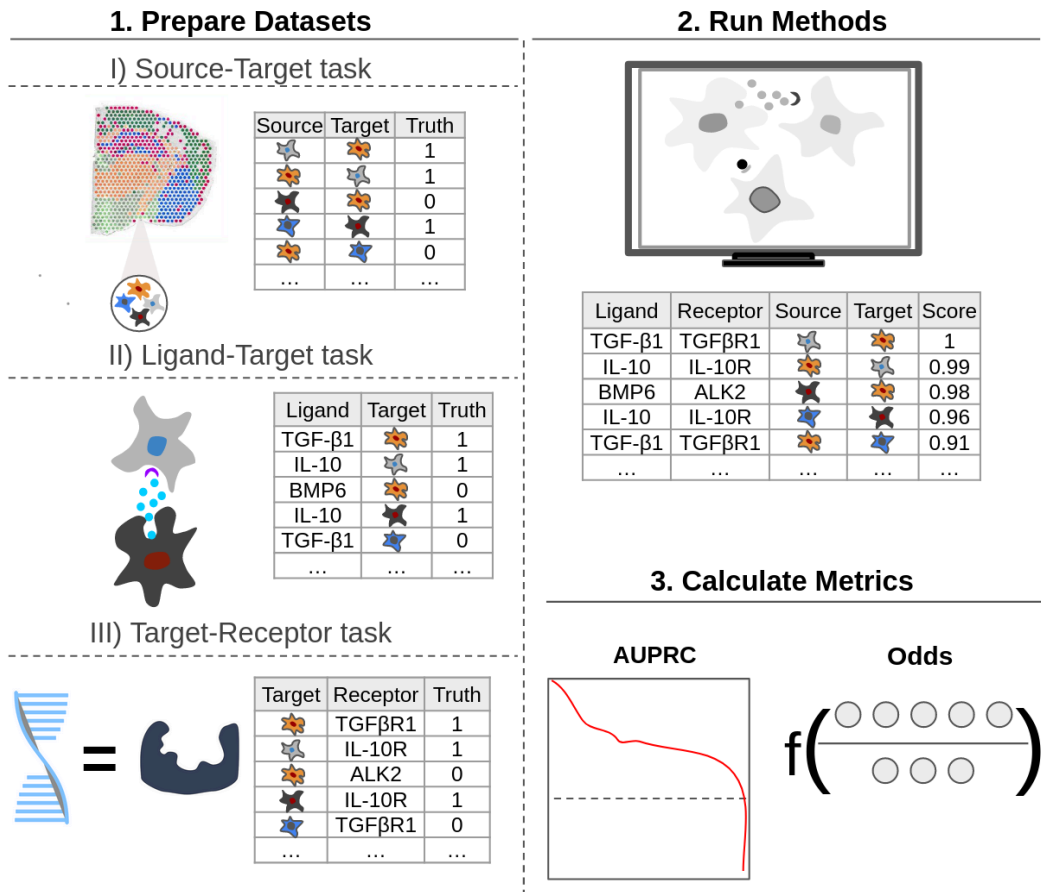


Figure 1. CCC evaluation schematic of three tasks with different types of assumed truth, pre-defined by the dataset at hand (1). For each dataset, we infer CCC interactions using LIANA (2), followed by the evaluation of each method-dataset combination by standardised metrics (3).

Main

Agreement with Cell-surface Protein Expression (Target-Receptor Task)

I used seven CITE-Seq datasets, each providing transcriptome-wide reads, along with the measurement of up to 106 antibody-tagged, surface proteins, to evaluate the agreement between CCC predictions and protein specificity (**Methods**).

I saw that all methods performed better than a random baseline when predicting CCC events associated with the most specifically-expressed proteins (**Figure 2**). NATMI⁵², Connectome⁵, and logFC performed best when considering both AUROC and AUPRC metrics - largely anticipated as the scoring functions of these methods explicitly model the specificity of ligand-receptor gene expression across cell types (**Figure 2**). CellPhoneDB⁴ and CellChat⁶ showed comparable performance with LIANA's consensus and the Crosstalk scores⁸ when using their p-values alone (**Figure 2B**), and worse performance when using their composite scores (**Figure 2A**; **Supp. table 1**) - comparable to SingleCellSignalR⁵³. SingleCellSignalR was the only method which does not explicitly model cell-type specificity, and as largely anticipated, performed worse than the other methods (in this task which reflected protein specificity) but still better than random.

Overall, this task showed that CCC predictions from single-cell transcriptomics data agree

with protein abundance, suggesting that expression is a reliable proxy for predicting CCC events at the protein level.

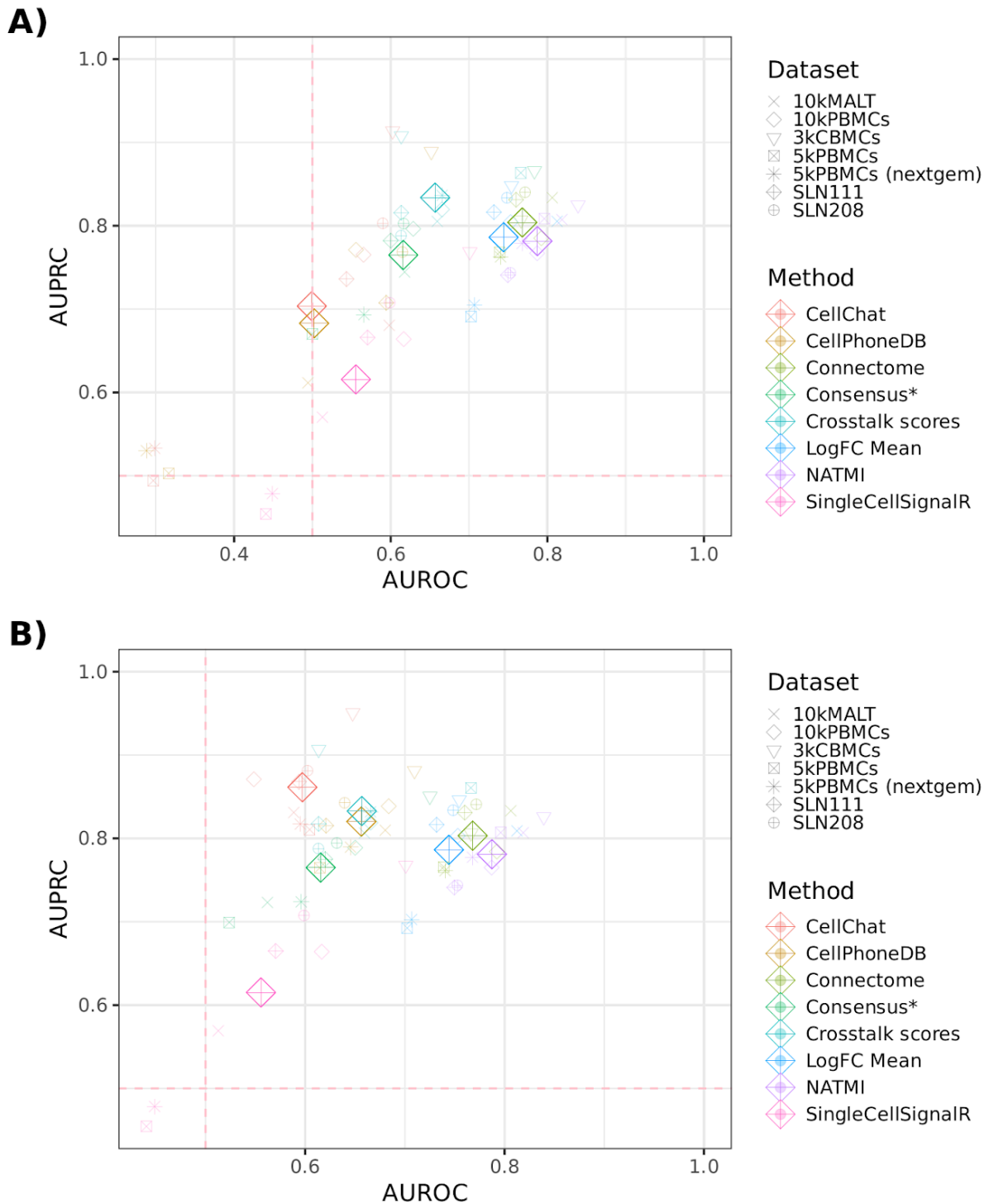


Figure 2. Agreement with cell-surface protein specificity (adapted from Dimitrov et al., 2022⁹). A) CellChat and CellPhoneDB's composite scores, and B) their p-values (specificity scores) alone; see Supp. table 1 for details.

Agreement with Cytokine Activities (Ligand-Target task)

To show the association between CCC predictions and active cytokines, I inferred cytokine activities in HER2+ and triple negative (TNBC) breast cancer subtype atlases⁹⁵ using cytokine signalling signatures⁹⁴ (**Methods**). I found a generally positive association between highly-ranked interactions and cytokine signalling (**Figure 3**). Specifically, all methods showed an enrichment of interactions with matched active cytokines when considering high-rank ranges, generally converging towards random as the number of interactions increased. In the TNBC subtype, all methods were enriched in the high-rank ranges. On the contrary, for the HER2+ subtype CellPhoneDB, CellChat, logFC, and SingleCellSignalR exhibited a negative or lack of association with cytokine activities at low ranks, but showed a positive association as the ranks considered increased. NATMI, Connectome, and the Crosstalk scores performed consistently across rank ranges across both subtypes (**Figure 3**).

These findings indicated that interactions identified as relevant by all methods generally aligned with cytokine activities, supporting the anticipated agreement between CCC and downstream intracellular signalling.

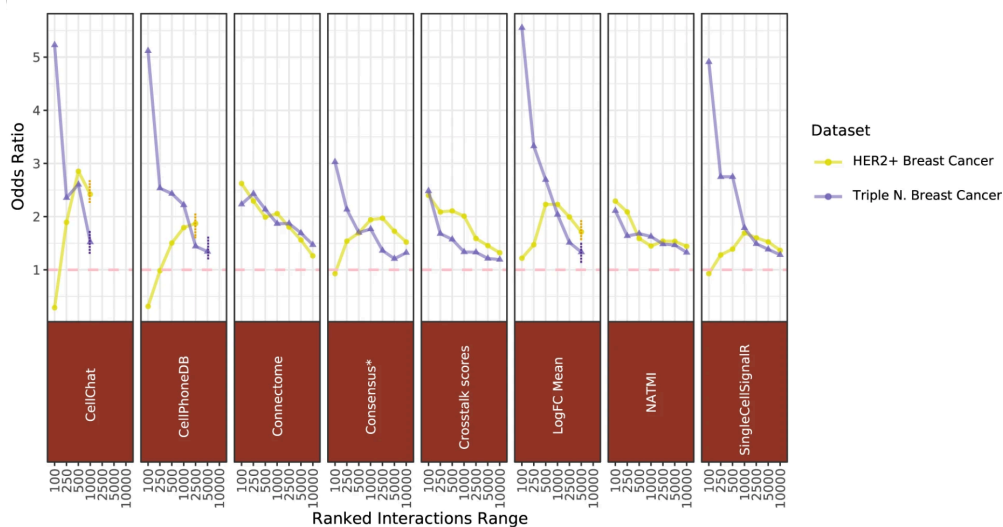


Figure 3. Odds ratios of interactions with matching active cytokines across a range between 100 and 10,000 interactions. Baseline (odds ratio of 1) represented by dashed horizontal line; vertical line represents the truncated ranges for some methods with inbuilt filtering steps - e.g. empirical p-value (adapted from Dimitrov et al., 2022⁹).

Agreement with Cell-type Colocalization (Source-Target task)

Next, I evaluated the methods' predictive performance by incorporating spatial information with the assumption that colocalized cell populations are more likely to interact. Hence, interactions prioritised by the different methods should be enriched among cell types in close proximity. The assumed truth (i.e. colocalized cell types) was generated with 10x Visium slides, matching the mouse brain cortex⁹⁶ and TNBC atlases⁹⁵, used to infer the interactions (**Methods**).

Overall, I saw a positive association between highly-ranked CCC interactions and colocalized cell types (**Figure 4**). This association was more consistent in the mouse brain

dataset, where all methods, except the Crosstalk scores, had an overrepresentation of interactions between colocalized cell types in the highly-ranked interaction ranges. In the TNBC atlas, only the consensus and LogFC methods had a consistent association with cell type adjacency (**Figure 4**).

In summary, these results supported the agreement between interactions prioritised by most methods and neighbouring cell types in the structured brain tissue. This association was also present, albeit less consistent, in the TNBC atlas. Hence, our results suggest that inferred CCC events partly reflected the enrichment of interactions between physically-adjacent cell types.

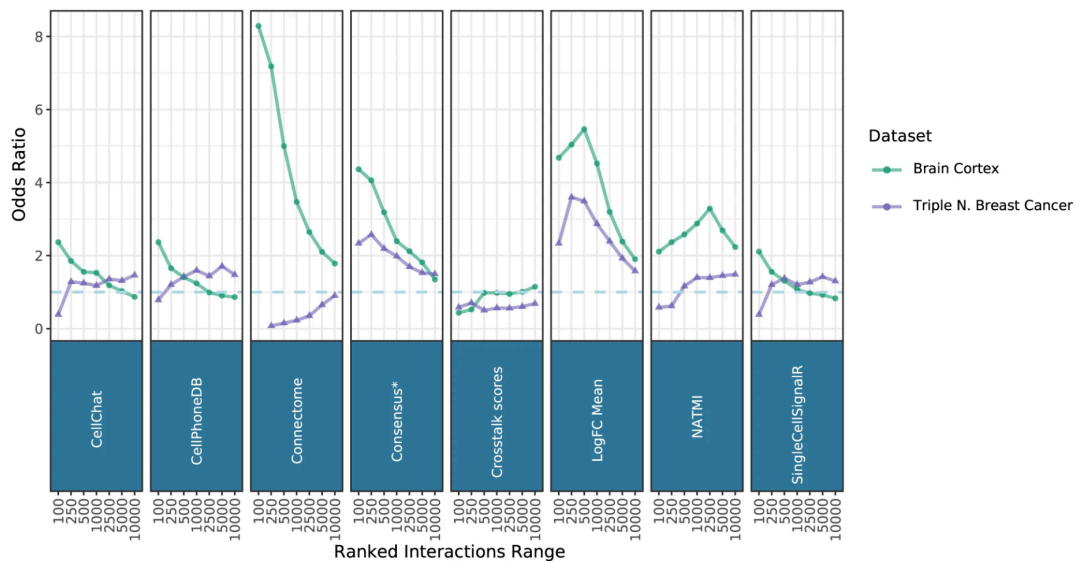


Figure 4. Odds ratios of interactions between colocalized cell types across a range between 100 and 10,000 (adapted from Dimitrov et al., 2022⁹). Baseline (odds ratio of 1) represented by dashed horizontal line.

Formalised and Open Benchmarks

Building on the colocalization (Source-Target) and cytokine activity (Ligand-Target) benchmarks shown above, we set up easily reproducible, interpretable, and open benchmark tasks in the *OpenProblems* platform. We focused on the datasets for each task, thought to better encode biologically-relevant signals. We further reformulated the tasks such that they account for the ‘partial’ assumed truth in the benchmarks. In other words, we aggregate the results from each method according to the information available in the benchmark datasets. For example, in the Ligand-Target task, we aggregated all interactions by the ‘ligand’ protein and ‘target’ cell type columns, such that there is only one prediction per assumed truth instance; using max- and sum-aggregation for each method. We additionally calculate ‘True’ and ‘Random’ baselines as references (**Methods**).

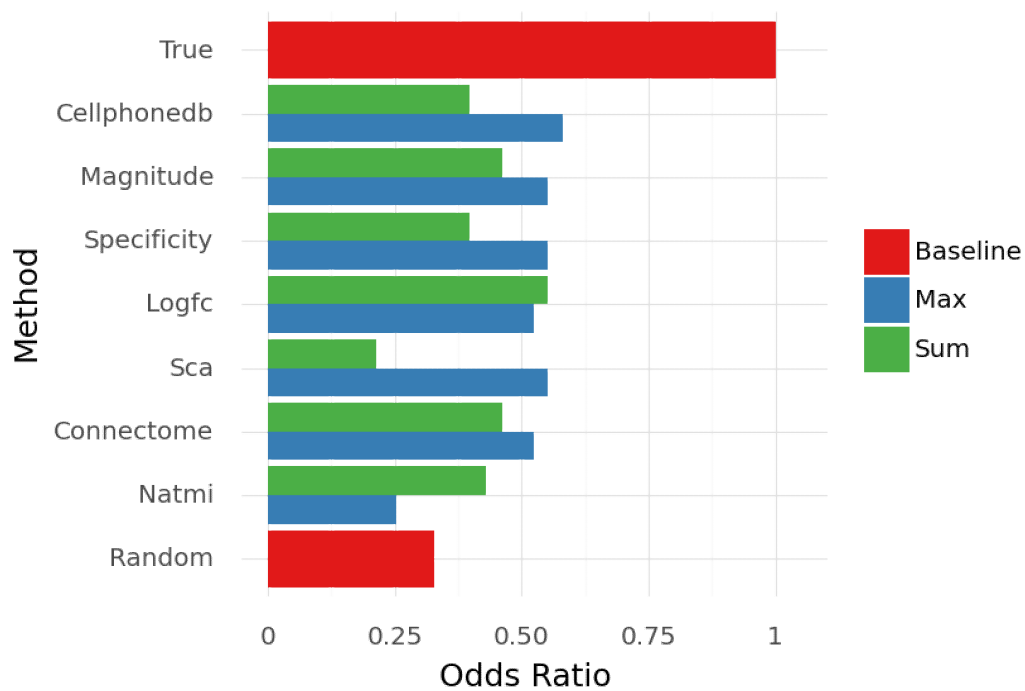
In the Ligand-Target (cytokine activity) task, methods were generally close to or worse than random when considering AUPRC (**Supp. Fig. 1**). However, all methods performed well according to the odds ratio metric (**Figure 5**), with CellPhoneDB and LIANA’s magnitude rank performing best.

In the Source-Target (colocalization) task all methods were notably better than random for both AUPRC and odds ratio metrics. CellPhoneDB and LIANA’s magnitude rank had the highest odds ratios, followed by NATMI, LIANA’s specificity rank, and log2FC (**Figure 5**). When considering AUPRC, LogFC and NATMI did better than CellPhoneDB and the magnitude rank (**Supp. Fig. 1**).

As before, we saw that methods generally performed better than a random baseline in both tasks (**Figure 5**). Yet, here we saw that magnitude-focused scoring functions generally outperformed those that focus on the specificity of the interactions across cell types. In particular, LIANA’s magnitude rank aggregate and CellPhoneDB had the best performance in both tasks (**Figure 5**).

In summary, while our results confirm that CCC methods generally capture biologically-relevant information, these evaluations represent only approximations of biological reality.

A)



B)

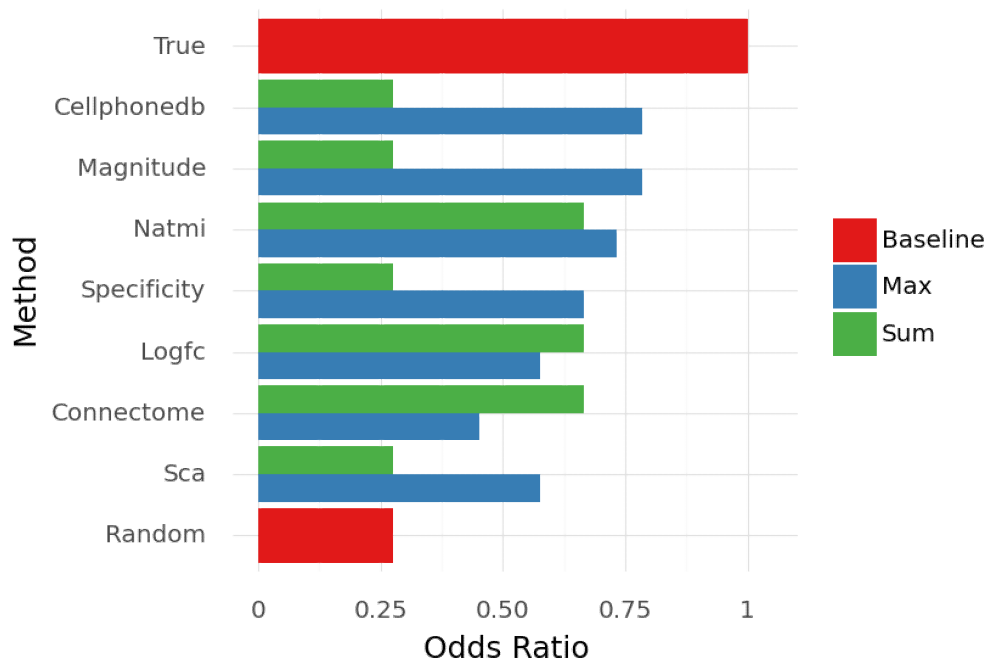


Figure 5. Min-maxed, sigmoid-transformed odds ratios for the A) Ligand-Target and B) Source-Target tasks, respectively. We consider the top 5% of interactions predicted by each method. Methods were ranked according to the aggregation function with the best performance.

Methods

Tasks Setting

In the Receptor-Target task, the abundance of receptor (cell-surface) proteins in target cell types was z-transformed across all cell types. Then receptors with abundance z-scores ≥ 1.645 are treated specifically-abundant, or the positive class. All other receptors were treated as the negative class when calculating the metrics. This task was not formalised in the *OpenProblems* benchmark, as its outcome was largely anticipated, better reflecting the nature of the scoring functions, rather than necessarily biologically-relevant performance.

In the Source-Target task, cell type deconvolution proportions per spot from 10x Visium slides were used to identify colocalized cell types. I used the SPOTlight⁹⁷ deconvolution method to spatially map the cell types present in an annotated scRNAseq atlases onto matching 10x Visium slides. Then, Pearson correlation coefficients were calculated for each cell type pair and z-scaled. Then cell type pairs with z-scaled correlations ≥ 1.645 were considered as colocalized, while the rest were thought to be non-colocalised. In other words, colocalized and non-colocalized cell types were treated as the positive and negative class, respectively. For the *OpenProblems* benchmark, we used only the murine brain dataset⁹⁶, while in Dimitrov et al., 2022⁹ I also considered the TNBC subtype atlas⁹⁵.

In the Ligand-Target task, I used downstream cytokine activities inferred for each cell type as assumed truth. Specifically, I inferred cytokine activities using CytoSig's high-quality signatures⁹⁴, together with the multivariate linear regression via decoupleR⁹⁸ for every cell type at the pseudobulk level in each single-cell atlas. I considered cytokine signatures with positive regression t-values and FDR-corrected p-value $= < 0.05$ in the target cell types as active, while the remainder were considered as inactive. In this subtask, active and inactive cytokines in the target cell types were treated as the positive and negative class, respectively. For the *OpenProblems* benchmark, I used only the TNBC subtype⁹⁵, while for Dimitrov et al., 2022⁹ I also considered the HER2+ subtype.

Five CCC methods were included in both *OpenProblems* benchmark tasks, along with specificity and expression magnitude rank aggregates. All of which were run as natively implemented in LIANA and used its consensus ligand-receptor resource. I considered only interactions for which of both the ligand and receptor, and any of their subunits, were expressed in at least 10% of the source and target cell types. For heteromeric complexes, I considered the arithmetic mean expression of the members for all methods, except CellPhoneDB, for which I used the minimum member expression. We additionally defined 'True' and 'Random' baselines, generated by copying the solutions of the data or a random selection of interactions, respectively.

In the benchmarks as done in Dimitrov et al., 2022⁹, I included seven methods, along with the consensus of their predictions (**Supp. Table 1**). The original implementations for NATMI, SingleCellSignalR, and CellChat were called via LIANA with their default settings⁹, while the remainder of the methods were re-implemented natively and used LIANA's preprocessing steps⁹.

Metrics

In the *OpenProblems* benchmark we considered 0.05 of is the number of all assumed truth interactions to calculate odds ratios which were then sigmoid-transformed. To generate the figures both AUPRC and Odds ratios were minmaxed together with the 'True' and 'Random' baselines.

In the benchmarks as done in *Dimitrov et al., 2022*, odds ratios were not sigmoid-transformed, while AUPRC/AUROC metrics were calculated with the yardstick⁹⁹ package. In this case we did not (sum- or max-) aggregate the inferred CCC predictions according to their columns with the ground truth, nor did we minmax the metrics.

Data Availability

The processed and annotated Human Breast Cancer single-cell atlas⁹⁵ is available via the GEO accession number: [GSE176078](https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE176078). The filtered breast cancer 10x Visium slides are available at <https://zenodo.org/record/4739739>.

Spatial transcriptomics datasets (10x Visium slides) on sagittal adult mouse brain anterior and posterior slices were obtained from SeuratData, available at <https://github.com/satijalab/seurat-data>, under the dataset name of `stxBrain`. The single-cell data (Allen Brain Atlas⁹⁶; GSE71585) used for the cell type deconvolution was obtained as a Seurat object, accessible at https://www.dropbox.com/s/cuowvm4vrf65pvq/allen_cortex.rds?dl=1

Processed and annotated murine spleen-lymph CITE-Seq datasets¹⁰⁰ are available at <https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE150599>.

Publicly available CITE-Seq datasets were obtained from 10X Genomics, accessible under the list of datasets at <https://tinyurl.com/10xCITEseq>. Datasets used here include:

https://support.10xgenomics.com/single-cell-gene-expression/datasets/3.1.0/5k_pbmc_protein_v3_nextgem

https://support.10xgenomics.com/single-cell-gene-expression/datasets/3.1.0/5k_pbmc_protein_v3

https://support.10xgenomics.com/single-cell-gene-expression/datasets/3.1.0/5k_pbmc_protein_v3

<https://www.10xgenomics.com/resources/datasets/10-k-cells-from-a-malt-tumor-gene-expression-and-cell-surface-protein-3-standard-3-0-0>

Code Availability

The *OpenProblems* CCC task and an in-depth task description is available at:

https://github.com/openproblems-bio/openproblems/tree/main/openproblems/tasks/_cell_cell_communication

The code for the *Dimitrov et al., 2022*⁹ benchmarks is available at https://github.com/saezlab/ligrec_decouple

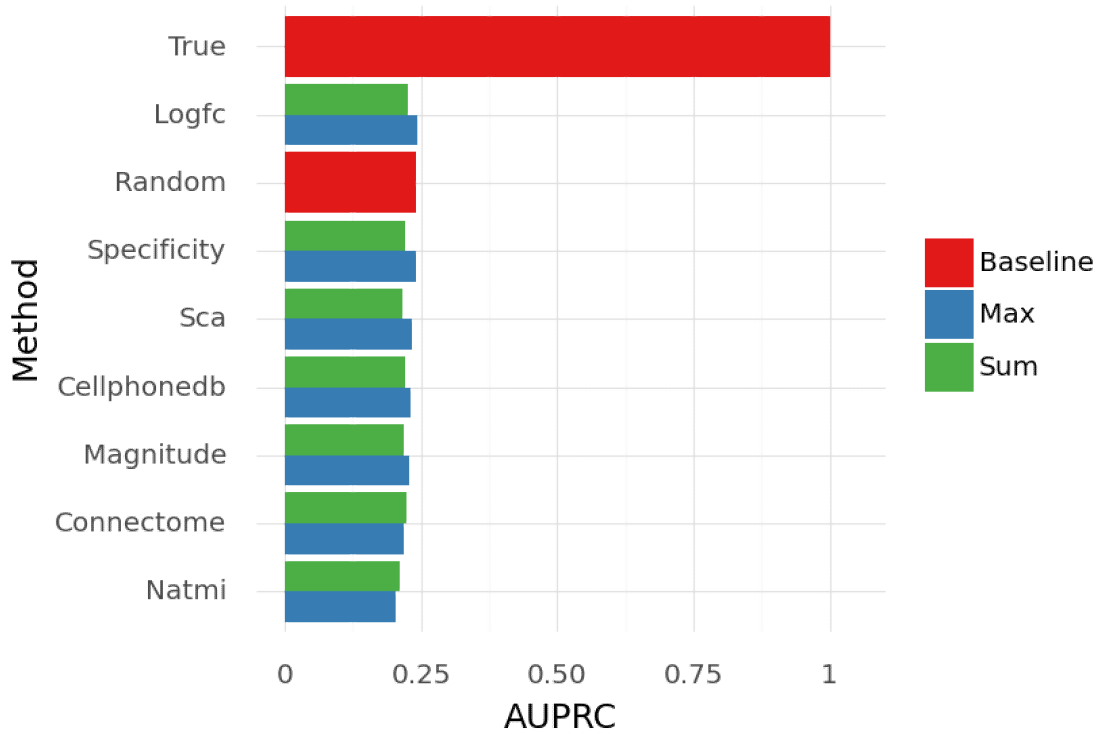
Supplementary Material

Supplementary table 1. *Scoring settings used to evaluate and compare the methods (adapted from Dimitrov et al., 2022⁹).*

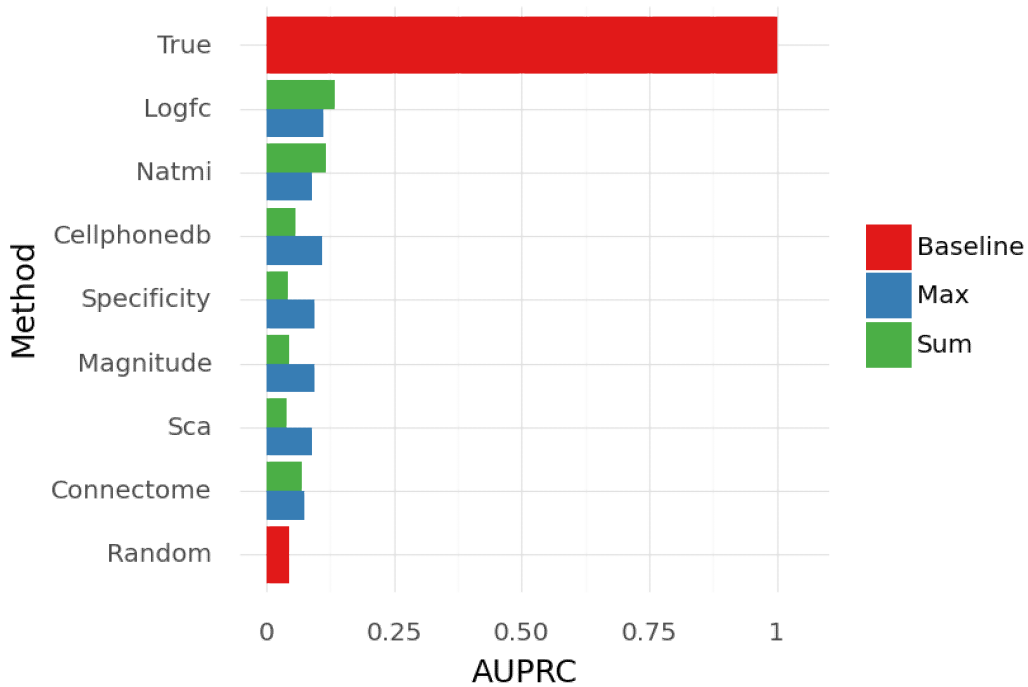
Method	Composite#		Specificity	Magnitude
CellChat	Probability	(Filtered by p-value ≤ 0.05)	<u>p-values alone</u>	Probability
CellPhoneDBv2	Truncated Mean	(Filtered by p-value ≤ 0.05)	<u>p-values alone</u>	Truncated Mean
Connectome	weight_scale		weight_scale	weight_norm
Crosstalk Scores	Crosstalk score		Crosstalk score	-
logFC Mean	logFC Mean		logFC Mean	-
NATMI	Specificity-based edge weight		Specificity-based edge weight	Mean-expression edge weight
SingleCellSignalR	LRscore		LRscore	LRscore
# Unless explicitly mentioned, the composite method settings were used.				

Supplementary Figure 1. Minmax AUPRC for the A) Ligand-Target and B) Source-Target tasks. Methods were ranked according to the aggregation function with the best performance.

A)



B)



Chapter 4: LIANA+: an all-in-one cell-cell communication framework

Since the time of writing this chapter, parts of the work presented were submitted to Biorxiv ³² (<https://www.biorxiv.org/content/10.1101/2023.08.19.553863v1.abstract>), and is currently under revision for publication. In this work, I extended LIANA to provide an all-in-one solution to contemporary cell-cell communication inference from single-cell and spatial (multi-) omics data. In the version presented here, I focus on applications that showcase the synergy between components from LIANA+. The text in this chapter was written by me. All figures in this section are shared under CC licence.



LIANA+: Visual abstract

1. Background

Cell-cell communication (CCC) inference has recently emerged as a major component of the analysis of single-cell and spatially-resolved transcriptomics data ^{1,42}. Many computational tools have been developed for this purpose, each contributing valuable ideas and developments.

The simplest class of CCC inference methods are those that infer protein-protein interactions from dissociated single-cell transcriptomics data, commonly referred to as ligand-receptor interaction inference methods ^{4-6,59}. Moreover, there are CCC tools that combine intercellular communication with intracellular signalling ^{7,8,16,101,102}. All methods are based on multiple assumptions, including the assumption that gene co-expression between dissociated cells, or groups of cells, reflects CCC within tissues ⁹.

In contrast to dissociated single-cell data, spatially-resolved omics technologies preserve tissue architecture and are thus thought to better reflect the intercellular events that occur within tissues. As a consequence, multiple methods that utilise spatial information have been developed to study CCC ^{2,42}. Typically, these methods infer relationships between proteins ^{26,27} or cell types (cellular neighbourhoods) ^{21,22}. Spatially-informed methods differ in the scale at which interactions are inferred, as some infer relationships globally, summarising them across slides as a whole ^{26,27,33}, while others do so locally at the individual cell or spot locations within a slide ^{29,30,103,104}.

The majority of both single-cell and spatially-informed CCC methods have focused on protein-mediated interactions, predominantly using transcriptomics data ^{1,9}, and only few methods infer CCC from multi-omics data ⁴⁰. As a consequence of the almost exclusive use of transcriptomics data, other modes of intercellular signalling, such as small molecule signalling, have been typically ignored ⁹. Recent methods have attempted to infer metabolite-mediated CCC events, again from transcriptomics data ¹⁰⁻¹³, but such inference remains largely limited by the challenges of predicting metabolite abundance from gene expression ¹⁴. Emerging multi-omics technologies ¹⁰⁵ are anticipated to provide a more informed view of CCC events, thus in turn prompting the development of new tools.

While early methods analysed CCC in individual samples or single-condition atlases, increasing sample numbers and experimental design complexity have prompted various strategies to extract differential CCC insights. These strategies include methods that (1) consider each interaction independently ^{29,37,38,106}, (2) make use of dimensionality reduction to perform pairwise comparisons between conditions ^{6,107}, or (2) model all variables, samples and cell types simultaneously ³⁹. Approaches (2) and (3) can be thought of as modelling coordinated CCC events, and I refer to such orchestrated interactions as 'intercellular programmes' from here on.

Regardless of the task at hand, CCC methods typically rely on pre-existing knowledge^{1,9}. Thus, extensive effort has been put into curating and extending prior knowledge, with a focus on gathering protein-^{4,6,45} and, to a lesser extent, metabolite-mediated CCC^{10,12,13,108}. In addition, in some resources, the interactions are further associated with pathways⁶ or transcriptional regulators^{7,109}, leading to multiple discordant databases and potential inconsistencies caused solely by the choice of resource⁹.

Finally, all these developments have been led by different groups, using various syntaxes, and most CCC tools are typically designed for a specific purpose.

Here, we (co-authors and I) introduce LIANA+ as a comprehensive framework that expands intercellular inference to cross-conditional and multi-omics data (**Fig. 1**; **Supp. Table 1**). LIANA+ unifies contemporary methods and prior knowledge, providing nine methods for the inference of CCC from single-cell transcriptomics (**Fig. 1A**) and eight methods for the inference of global and local relationships in spatially-resolved data (**Fig. 1B**), all of which are applicable to multi-omics data. We further supplement the single-cell and spatial methods in LIANA+ with different strategies to extract deregulated CCC events in both hypothesis-free and hypothesis-driven manner (**Fig. 1C**). Moreover, CCC events can be connected to intracellular signalling events via the use of (**Fig. 1D**) a causal subnetwork search, which also makes use of a rich prior knowledge base (**Fig. 1E**). All components in LIANA+ use standardised input and output formats, utilising the scverse ecosystem infrastructure¹¹⁰ (**Supp. Fig. S1**). This enables interoperability with external packages and facilitates the straightforward extensions of contemporary CCC approaches.

To highlight the flexibility of LIANA+, I applied it to a recent spatially-resolved metabolome-transcriptome dataset, where I modelled global dopamine-mediated CCC events and identified their corresponding subregions of interaction in murine Parkinson's disease model data. I also showcase the scope of our framework by jointly analysing dissociated single-nucleus and spatially-resolved human heart data with a complex cross-conditional experimental design. In this analysis, I identify intercellular programmes and corresponding intracellular signalling driving fibrosis in ischemic heart regions and cell types.

LIANA+ is an extendable, scverse-compliant¹¹⁰, open-source framework built of synergistic modules, available at <https://github.com/saezlab/liana-py>, readily applicable to a wide range of single-cell, spatial, multi-omics datasets, with any experimental design.

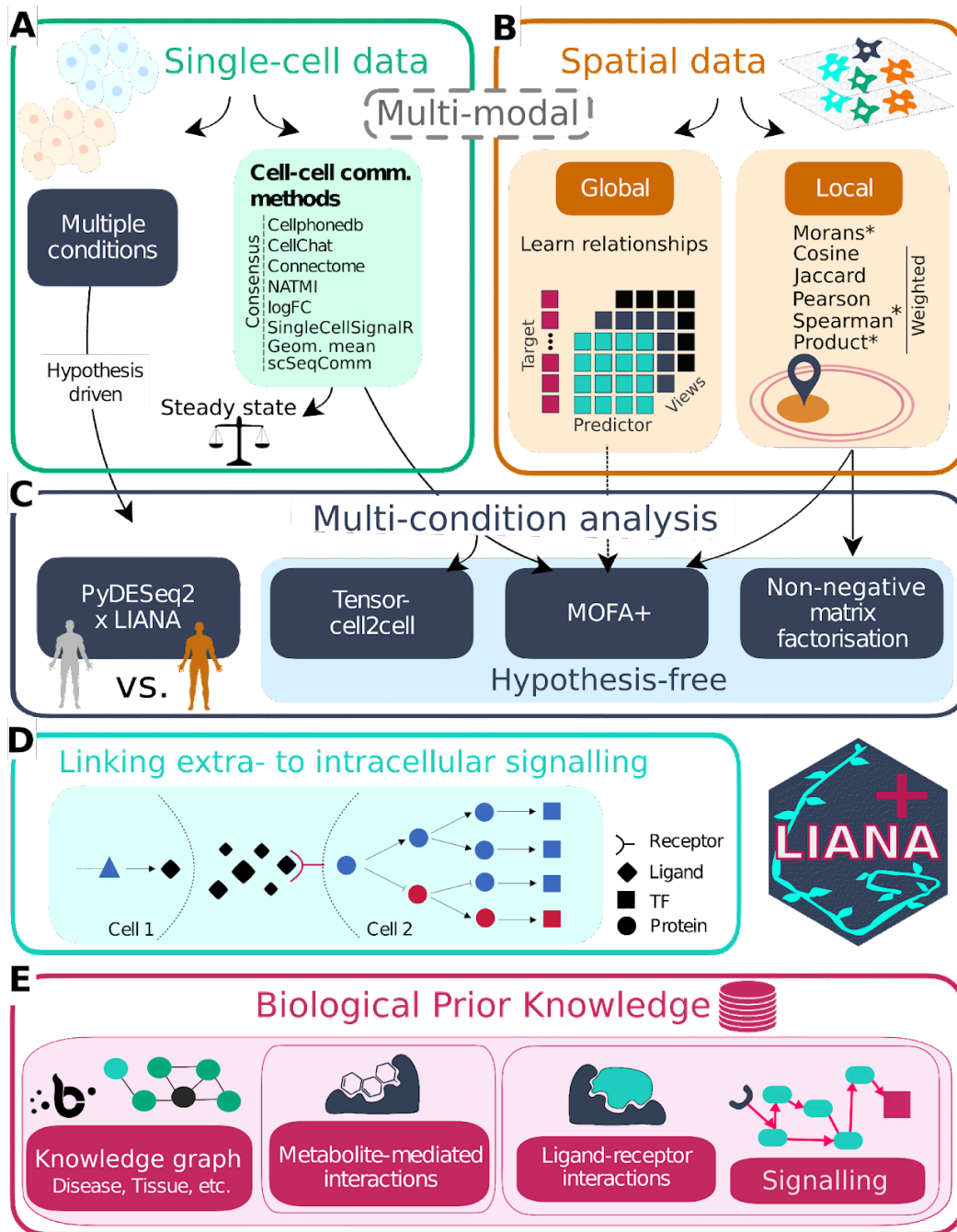


Figure 1. LIANA+ Framework Overview. A) LIANA+ natively re-implements eight ligand-receptor methods to infer interactions from single-cell transcriptomics data, along with a flexible consensus. B) LIANA+ enables multi-view learning as well as eight local metrics to respectively capture global and local interactions from spatially-resolved omics data. (*) For Moran's R I adapted both the global and local versions as in SpatialDM²⁹; for spatially-weighted Spearman correlation, I re-implemented also a masked version as in schOT¹¹¹; and for the spatially-weighted product, I additionally provide a max-normalised version. C) LIANA+ includes diverse strategies to identify deregulated CCC events across conditions, ranging from differential CCC analysis with PyDESeq2^{112,113} for hypothesis-driven exploration to more extensive, unsupervised approaches like Tensor-cell2cell³⁹, Non-negative matrix factorization, and MOFA+¹¹⁴ for hypothesis-free exploration. D) LIANA+ connects intercellular interactions to intracellular signalling pathways, providing a holistic view of signalling. E) LIANA+ is built on a rich knowledge base - OmniPath⁴⁵ and BioCypher¹¹⁵ - which comprise different ligand-receptor, metabolite-mediated, and annotation databases; alongside intracellular knowledge such as signalling pathways and transcription factors. Finally, all components (A-E) of LIANA+ are applicable to both dissociated single-cell and spatially-resolved multi-omics data.

2. Main

2.1 LIANA+ enables the modelling of CCC from Spatially-Resolved Data across distinct modalities

As a consequence of the diverse array of spatial omics technologies and the wide range of molecules and resolutions they encompass, diverse computational approaches are required to make most of the data. In this context, LIANA+ includes several strategies to flexibly analyse CCC from diverse spatially-resolved technologies.

2.1.1 LIANA+ jointly models global associations across modalities

LIANA+ adapts our recent multi-view modelling approach to learn spatial relationships across distinct types of features or spatial contexts (represented as views)³³. This enables the inference of complex relationships, such as those between ligand expressions and pathways³³ or cell types and pathways^{116,117}, jointly modelling any combination of views. Using such joint modelling, here, I learn relationships from combinations of modalities and technologies.

To showcase the ability of LIANA+ to model interactions from multi-omics data, I applied it to a recent murine Parkinson's disease model dataset that captures spatially-resolved metabolome and transcriptome measurements¹¹⁸ (**Fig. 2A**). Briefly, three mice were subjected to unilateral 6-hydroxydopamine-induced lesions in one hemisphere while the other remained intact¹¹⁸ (**Fig. 2B**).

A particular challenge of this dataset is that, while transcriptome and metabolome measurements were quantified on the same slides, the observations derived from each technology corresponded to distinct spatial locations (**Fig. 2A**). To this end, I extended our multi-view modelling approach to handle multi-omics datasets with discordant observations (Methods).

I inferred the metabolites' spatial relationships with brain-specific metabolite receptors and cell types (**Fig. 2C**; Methods). Specifically, I jointly modelled metabolite peak intensities using cell type proportions and receptors as predictor views, with the predictions from each view being combined post-hoc; allowing us to estimate joint performance of the predictor views, along with the individual contribution of each view (Methods). I carried out this process while also masking according to the lesioned and intact hemispheres on each slide (**Fig. 2B&C**; Methods). I saw that several metabolite peaks were explained relatively well by the joint model ($R^2 > 0.5$), including Dopamine and its derivative 3-Methoxytyramine, among other potentially deregulated, but unannotated, metabolite peaks (**Supp. Fig. S3A**). Focusing on Dopamine, I saw a large difference in explained variance between the intact ($R^2 = 0.535$) and lesioned hemispheres ($R^2 \approx 0$; **Fig. 2D**; **Supp. Fig. S4A**), which is expected due

to the absence of dopamine in the striatum of the lesioned hemisphere ¹¹⁸ (**Supp. Fig. S4A-C**). Looking further into the intact hemisphere, I saw that the cell type view explained a higher proportion of variance than the receptor view (**Fig. 2E**; Methods), implying that Dopamine signalling in this region was more closely associated with the signatures of specific cell types than the expression of brain-specific metabolite receptors. I noted similar contributions for per view the remainder of the well-explained metabolite peaks (**Supp. Fig. S3B**).

Specifically, in the intact striatum, I found that the three best predictors of Dopamine (Median t-value ≥ 3) were dorsal medium spiny neurons (MSN) 1 and 2 ¹¹⁹, and the Drd2 dopamine receptor (**Fig. 2F**). Notably, Dopamine's relationship with its predictors differed notably between intact and lesioned hemispheres: MSN1 (P-value=0.024), MSN2 (P-value=0.004), and Drd2 (P-value=0.08) (**Supp. Fig. S3C-E**). The association of Dopamine with the MSN1 cell state corroborated the findings of the original publication ¹¹⁸, while LIANA+ highlighted the interactions of Dopamine with MSN2 and its canonical receptor, which was not previously reported. Thus, our findings captured a broader picture encompassing both D1- and D2-type MSNs ¹¹⁹.

2.1.2 LIANA+ infers local interactions at the individual spot or cell locations

The approach described in section 2.1.1 models global spatial relationships - i.e. it considers all spots to infer a single value per interaction across the slide. As such, it provides summary statistics for each interaction in a slide, but does not provide information about the region or locations where the interactions occur. To complement identified global relationships in LIANA+, I implemented eight metrics to identify local interactions at the individual spot or cell location. Such bivariate metrics have previously been used to identify spatial co-expression patterns between genes across spatial ^{25,111} and pseudotime ¹¹¹ contexts, and recently in the context of ligand-receptor interactions ^{29,103,104}. In brief, LIANA+ includes (i) four spatially-weighted variants of commonly used similarity metrics (Cosine similarity, Pearson and Spearman correlation, and Jaccard index); (ii) a masked version of Spearman correlation ¹¹¹; (iii) simple spatially-weighted products; and (iv) a local extension of univariate spatial clustering measures such as Moran's I ^{31,120} (Methods). I evaluated the performance of these metrics in two separate tasks, and decided to use Spatially-weighted Cosine as the default local metric in LIANA+, due to its simplicity, interpretability, and consistent performance (**Supp. Fig. S2**; **Supp. Note 1**).

Using spatially-weighted Cosine similarity, I focused on identifying the specific locations at which putative interactions with Dopamine occurred. I saw that, as anticipated, the interactions of Dopamine and Drd2, highlighted above by the global multi-view learning approach, occurred only within the intact striatum regions (**Fig. 2G**), which was further

corroborated by low permutation P-values (**Fig. 2H**) and a positive association (category) between the two variables (**Fig. 2I**) in the intact hemisphere. With similar results for MSN1 and MSN2 cell states (**Supp. Fig. S4E&F**). Interestingly, our analysis revealed an asymmetry in signalling components, as while the interaction between Drd2 and Dopamine was present in the intact hemisphere, Drd2 was expressed in the striatum of both intact and lesioned hemispheres (**Supp. Fig. S4A-D**).

In conclusion, LIANA+ demonstrates robust applicability in identifying spatially-informed interactions, mediated by a variety of agents, and pinpoints the specific regions where these interactions take place. Moreover, our work illustrates LIANA+'s flexibility in innovatively integrating and extending existing elements into novel applications, as demonstrated by modelling metabolite-mediated CCC across modalities.

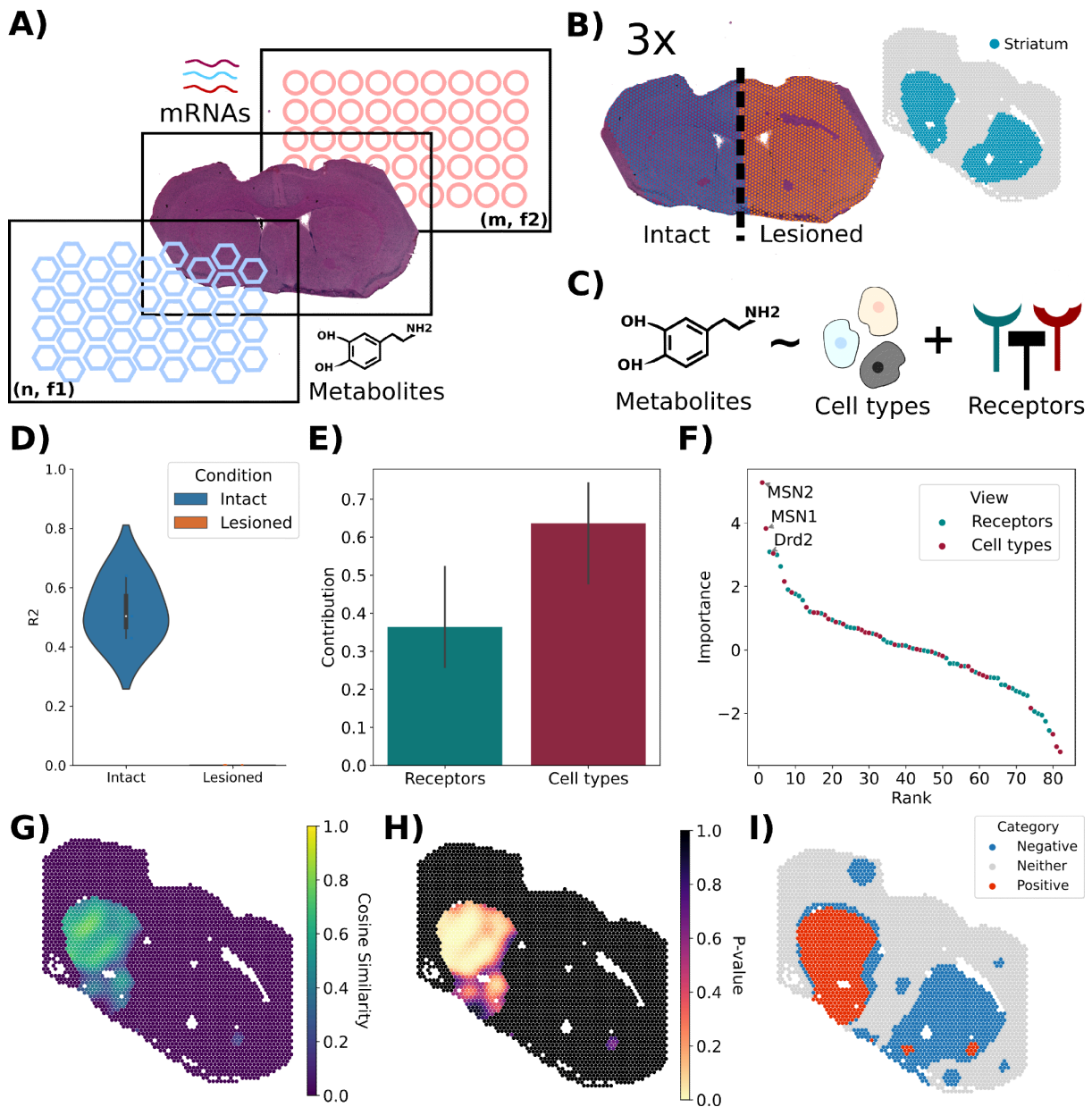


Figure 2. A) Spatially-resolved transcriptomics and metabolomics captured on the same slide via joint 10X Visium and MALDI-MSI yield distinct matrices with different observation locations. B) Parkinson's model mice annotated for striatum in intact and lesioned hemispheres. C) Multi-view modelling of metabolite peaks with brain-specific receptors and inferred cell type proportions. D) Variance explained (R^2) for Dopamine, jointly modelled by receptor and cell type views in the intact and lesioned hemispheres. The central line within each box marks the median, with the box hinges representing the first and third quartiles. The whiskers extend up to 1.5 times the interquartile range above and below the box hinges. E) Contributions of each view to the modelling of dopamine with 95% confidence intervals. F) Dopamine predictors ranked according to median importance (y-axis; ordinary least squares t-values), with names shown for the top 3 predictors: Drd2, and Medium Spiny Neurons (MSN) 1/2. Local interactions between Dopamine and its canonical Drd2 receptor as measured by spatially-weighted Cosine similarity (G), its corresponding permutation p-values (H), and categories (I). Figure panels showcase slide B1 from experiment V11L12-109 ¹¹⁹.

2.2 LIANA+ Combines Dissociated Single-cell with Spatial Data to identify deregulated Intercellular and Intracellular Signalling across Conditions

As the number of samples and experimental design complexity in single-cell and spatial datasets continue to increase, generalizable methods are required to analyse CCC across conditions. To address this, LIANA+ natively re-implements nine ligand-receptor methods for the scalable inference of CCC from dissociated single-cell data (**Fig. 1A**; **Supp. Table 2**), coupling them with factorization approaches to analyse CCC across conditions (**Fig. 1C**). In addition enabling applications with Tensor-cell2cell, as previously demonstrated^{39,121}, I propose an alternative unsupervised approach that leverages the flexibility of the MOFA+ framework¹¹⁴ (Methods; **Supp. Note 2**). With a distinguishing feature of adapting MOFA+ being that I obtain interaction importances per cell type pair (view) rather than their global importance across all cell types^{39,121}. Using five public cross-conditional atlases (**Supp. Table 3**), I show that, regardless of the ligand-receptor method, both Tensor-cell2cell and MOFA+ capture intercellular programmes that separate samples according to different conditions (**Supp. Fig. S5**; **Supp. Note 2**).

Moreover, in line with current best practices^{122,123} LIANA+ utilises differential expression analysis^{112,113} to enable the targeted (hypothesis-driven) exploration of CCC across conditions (Methods). Briefly, LIANA+ combines prior knowledge with differential statistics, calculated at the pseudobulk^{122,123} level, to identify interactions, the ligands and/or receptors of which, are deregulated across cell type pairs³⁷ (Methods).

Finally, CCC events commonly initiate or emanate from intracellular processes. Consequently, leveraging a comprehensive knowledge base LIANA+ provides various strategies to investigate them (**Fig. 1D&E**). Our framework enables the annotation of ligand-receptor interactions to relevant pathways. Moreover, it provides a causal subnetwork search to link CCC events of interest to downstream intracellular signalling (**Supp. Fig. S6**; **Supp. Note 3**).

To demonstrate the synergistic nature of LIANA+'s components, I analysed a cross-conditional dataset integrating single-nucleus and spatial transcriptomics data. This dataset comprised 29 single-nucleus and 28 10X Visium spatial samples from myogenic, fibrotic, and ischemic heart regions following myocardial infarction¹¹⁷ (**Fig. 3A**). Using this data, the role of myofibroblasts and macrophages was previously highlighted in the fibrotic process, characterised by the synthesis of extracellular matrix proteins for scar tissue formation¹¹⁷. Here, I further elucidate the intercellular and corresponding intracellular signalling mechanisms facilitating cardiac tissue repair and remodelling.

2.2.1 LIANA+ extracts Disease-specific Communication Patterns from Spatial Transcriptomics Data

To identify shared intercellular programmes across the different conditions, I first inferred local ligand-receptor interactions using spatially-weighted Cosine similarity for each of the 28 spatial transcriptomics slides. Then I concatenated the resulting interactions across all samples and applied non-negative Matrix Factorization (NMF) (Methods). As suggested by an automatic elbow selection procedure, NMF revealed five disease-relevant intercellular programmes (factors), with each resulting in factor scores per observation (spot), along with associated interaction loadings (**Fig. 3B**). I saw that Factors 1, 2, 5 were significantly enriched in ischemic samples, Factor 4 was enriched in fibrotic and downregulated in (healthy) myogenic samples, while Factor 3 was positively associated with the myogenic condition (**Fig. 3C**).

To elucidate the biological processes associated with the identified factors, I assessed whether ligand-receptor interactions were enriched in particular pathways ¹²⁴ (**Fig. 3D**; Methods). I noted that interactions in both the ischemia-associated Factor 4 and the fibrosis-associated Factor 5 were enriched in TGF β signalling interactions - a well-known driver of fibrosis ¹²⁵. On the other hand, both Factors 4 and 5, as well as myogenic-associated Factor 3, were negatively associated with pro-inflammatory pathways such as TNF α and NFK β , and the PI3K pathway. Moreover, I saw that interactions from ischemia-associated factors 1, 2 were enriched in Hypoxia, EGFR, MAPK, and JAK-STAT pathways, reflecting anticipated inflammatory patterns in ischemic regions ¹²⁶ (**Fig. 3D**).

Consistent with expectations, ischemia-associated Factor 1 localised to ischemic regions (**Fig. 3E**). Notably, among the top 30 interactions, were several integrins (*ITGB1*, *ITGAV*, *ITGA5*, and *ITGA7*) interacting with profibrotic (*FN1*, *SPP1*) genes ^{117,127} and matrix glycoproteins (*TNC*, *THBS1*) (**Supp. Fig. S7A**). Interactions with TNC aligned well with its reported RNA and protein levels in early myocardial infarction stages ¹²⁸ - concurrent with the largely early-onset of the samples in our data ¹¹⁷. *FN1* was reported as a marker of myofibroblasts ¹¹⁷, while *SPP1* ¹¹⁷ *FN1* ¹²⁹, or both ¹²⁷ as markers of pro-fibrotic macrophages. Moreover, *THBS1* and *ITGB1* were recently implicated in a self-amplifying, immune-cell recruitment loop, including *FN1*+ *THBS1*-expressing macrophages ¹²⁹.

Overall the identified intercellular programmes aligned well with literature. Thus, I hypothesise that the interactions identified within ischemic regions, specifically those between *FN1* and *SPP1* with *ITGB1*-containing complexes (**Fig. 3F&G**), are potential drivers of pro-fibrotic response.

2.2.2 LIANA+ extracts Condition-specific Communication Patterns from Dissociated Single-cell data

Identifying co-localised genes from spatially-resolved data can help us pinpoint relevant interactions driving disease; yet it remains limited by the feature-resolution trade-off in most contemporary spatial technologies^{19,20}. Current spatial technologies either capture multiple cells within spots, relying on deconvolution to identify cell type frequencies within them, or they capture a limited fraction of molecules^{19,20}. On the contrary, CCC events inferred from dissociated data show weak correspondence with the spatial colocalization⁹, pointing to an anticipated high false positive rate.

Motivated by the recent emergence of a technology capable of measuring full transcriptome single-nuc data, while also preserving spatial information¹³⁰, I updated our previous evaluation that used cell type colocalization as assumed truth⁹ (**Supp. Note 4**; Methods). Similarly, I saw weak association of ligand-receptor interactions predicted in a spatially-uninformed manner and the colocalization of their ligands, receptors, and cell types. Moreover, regardless of the method, performance was only marginally better than random (**Supp. Fig. S8**; **Supp. Note 4**).

To enhance the accuracy of our analysis and reduce false positives, I incorporated spatial information, derived from the spatial transcriptomics data, into ligand-receptor predictions from the single-nucleus data. Specifically, I considered interactions only if they were identified as condition-relevant by the NMF analysis in section 2.2.1, and if they occurred between cell type pairs observed to co-localise (**Supp. Fig. S7B**; Methods). Then, I inferred ligand-receptor interactions with LIANA+ for each of the 29 single-nucleus samples, followed by factorisation with MOFA+ on the obtained ligand-receptor interactions across all cell type pairs and all samples (**Fig. 3H**; Methods). This unsupervised analysis identified a factor (Factor 1), the sample loadings of which were significantly different across the conditions (P-value < 0.0001; **Fig. 3I**). Within the same factor, interactions with fibroblasts were prominent sources of CCC, including the fibroblast-to-myeloid axis ($R^2 = 25.3\%$) (**Fig. 3J**), also concurrent with spatial association between the two cell types (**Supp. Fig. S7B**). These findings underscore fibroblasts and myeloid cells as key actors in ischemic tissue states, highlighting their role both as drivers and targets in the fibrotic process.

In a similar application, I also used MOFA+ to identify intercellular programmes driving response to acute kidney injury in single-cell dissociated data, and further supported in spatial transcriptomics (**Supp. Fig. S9**; **Supp. Note 5**).

2.2.3 LIANA+ identifies deregulated intercellular and intracellular signalling events

To elucidate specific ligand-receptor interactions deregulated in ischemia, I performed differential expression analysis, using the myogenic and fibrotic states as baseline

references. (**Fig. 3K**; Methods). Focusing on fibroblast-to-myeloid signalling and the associated interactions identified by the unsupervised analyses in sections 2.2.1 and 2.2.2, I observed a high degree of concordance with deregulated interactions (**Fig. 3L**). Notably, *THBS1* and *TNC* were significantly deregulated (FDR < 0.05) in fibroblasts and myeloid cells, respectively, while *SPP1*, *THBS1*, and *FN1* were deregulated in both (**Fig. 3L**). Additionally, two ITGB1-containing integrin complexes interacting with FN1 and SPP1 were upregulated in myeloid cells (**Fig. 3L**).

Building on these findings, I delved into the intracellular signalling within myeloid cells triggered by the FN1/SPP1 and ITGA5&ITGB1 integrin complex interactions (**Fig. 3K**). This focus was motivated by the reported role of these genes in fibrosis^{117,127,129}; their co-deregulation in myeloid and fibrotic cell types (**Fig. 3L**); as well as the fibroblast-to-myeloid signalling axis highlighted by both the MOFA+ latent space (**Fig. 3I**) and spatial analysis (**Supp. Fig. S7B**). Using a signed and directed prior knowledge network, I computed a family of sign-consistent subnetworks linking the upregulation of FN1/SPP1 and ITGA5&ITGB1 with deregulated downstream TFs (**Fig. 3K**; Methods). This revealed a putative signalling network involving kinases MAPK1 and MAPK14, and TF co-regulators ATM, EP300, and YAP (**Fig. 3M**). Specifically, these regulatory proteins were predicted to up-regulate SMAD1/3 TFs, key members of canonical TGF β superfamily signalling¹³¹; as well as the down-regulated FOXO3 - recently implicated in TGF- β regulated myofibroblast differentiation via the inhibition of SMAD3¹³².

In summary, our analysis revealed insights of intercellular and intracellular events linked to the establishment of the myofibroblast phenotype and recruitment of pro-fibrotic SPP1+ macrophages in myocardial infarction¹¹⁷. As such, I demonstrate that LIANA+ offers a complete suite to identify novel disease-related communication patterns, along with diverse strategies to interpret the underlying biological processes.

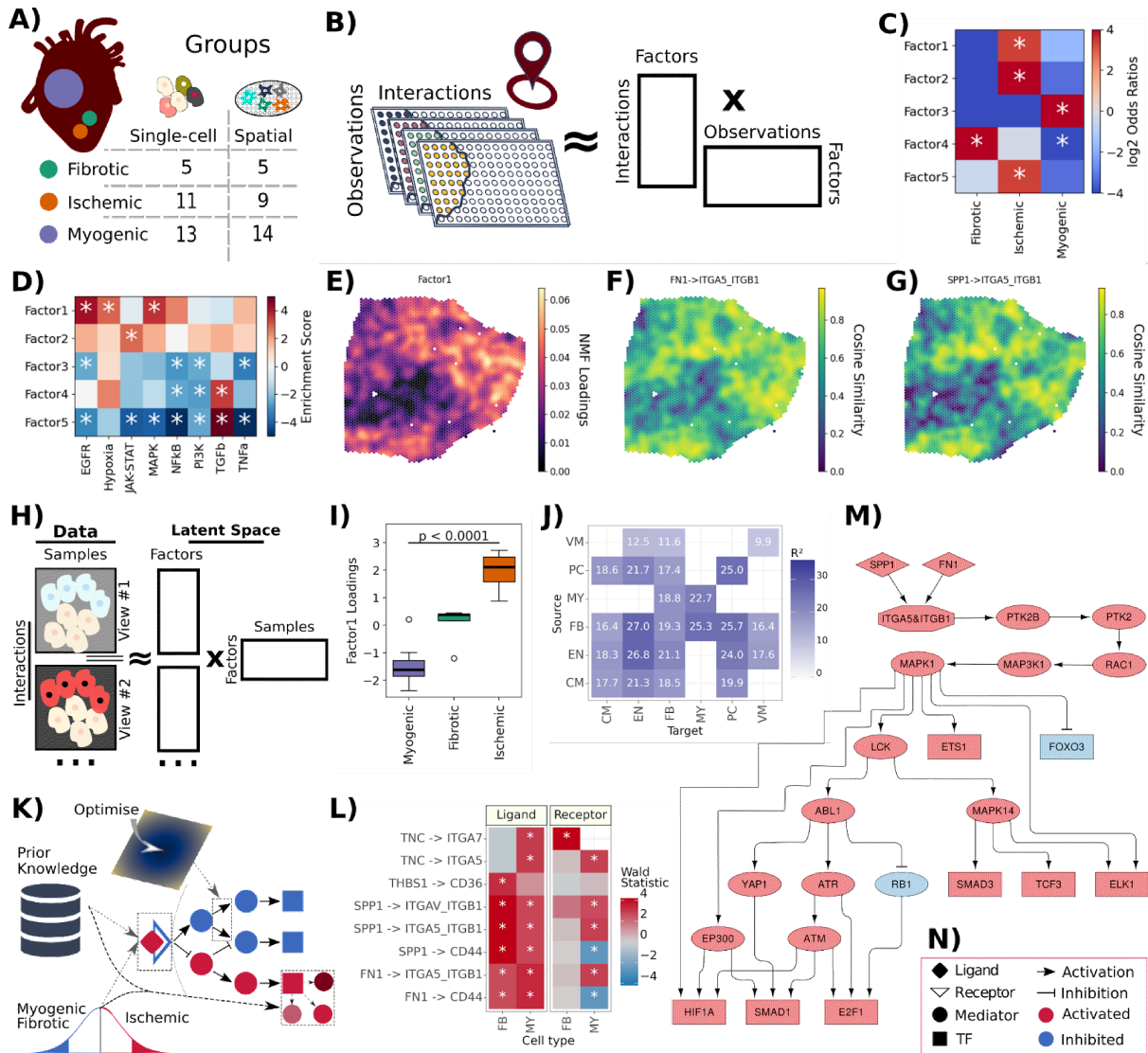


Figure 3. A) Sampling sites for human myocardial infarction single-nucleus and spatial transcriptomics data. B) Overview of NMF (Non-negative matrix factorization) applied to local interactions inferred for each location/observation across multiple slides. C) log₂-transformed Odds ratios representing the enrichment or depletion of each of the factors inferred by NMF in the fibrotic, ischemic, and myogenic labels. D) Pathway enrichment of NMF ligand-receptor loadings using pathway annotations from the PROGENy¹²⁴ resource. Asterisks indicate FDR-corrected p-values <0.05. E) NMF Factor 1 scores per observation. Spatially-weighted Cosine similarity of F) FN1 and G) SPP1 with the ITGA5_ITGB1 complex, respectively; in a selected (ACH0014) sample. H) Procedure to decompose ligand-receptor interactions inferred at the cell type level from dissociated single-cell data samples, using standard ligand-receptor methods (such as those in Figure 1C), into factors and corresponding feature sets. I) MOFA+ Factor scores following ligand-receptor score decomposition with one-way ANOVA P-value across ischemic, fibrotic, and myogenic samples. The central line within each box marks the median, with the box hinges representing the first and third quartiles. The whiskers extend up to 1.5 times the interquartile range above and below the box hinges. Outliers are depicted as individual hollow points beyond the whiskers. J) MOFA+ Factor 1 variance explained across cell type pairs (views). Abbreviations used include AD for Adipocytes, CM for Cardiomyocytes, EN for Endothelial cells, FB for Fibroblasts, PC for Pericytes, PR for Proliferating cells, VM for Vascular smooth muscle cells, NE for Neuronal cells, MY for Myeloid cells, MA for Mast cells, and LY for Lymphoid cells. K) The procedure used to identify deregulated intercellular and intracellular signalling events, includes the combination of differential statistics (in this case from a contrast between samples from ischemic versus the rest) with prior knowledge of ligand-receptors and TF regulons to identify deregulated intercellular interactions and TFs across the conditions. Followed by a network optimisation approach to identify sign-coherent subnetworks linking the deregulated ligand-receptor interactions and TFs. L) Interactions of interest deregulated in Fibroblast and Myeloid cell types with FDR <0.05 marked with an asterisk. M) Sign-coherent signalling network emanating from FN1 and SPP1 via integrin complex ITGA5_ITGB1 and propagating down to transcription factors deregulated in Myeloid cells in ischemia. N) Corresponding legend for the network figure shown in M); TF is an abbreviation of Transcription Factor.

3. Discussion

As our ability to quantify molecules at scale increases, so does the demand for comprehensive methods to generate biological insights. Building on flexible and efficient infrastructure ¹¹⁰ and a rich biological knowledge base ^{45,115}, LIANA+ expands previous methodological developments, redefining them into synergistic components, to enable flexible and novel CCC analyses from single-cell and spatial (multi-)omics data.

Single-cell technologies capture cellular heterogeneity at an unprecedented scale, yet during the dissociation process information about tissue architecture is lost. As a consequence, and as highlighted by our work, spatially-agnostic CCC inference is susceptible to high rates of false positives. Conversely, spatial technologies preserve tissue context, yet they either provide limited resolution since each spot captures multiple cells, or measure a relatively low number of genes ^{19,20}. Therefore, if spatially resolved-data is available, its integration with dissociated data is recommended to minimise false positives and offset the current limitations of both technologies. Here, I demonstrate the synergistic nature of LIANA+ by adopting such a strategy to myocardial infarction data. Specifically, I incorporate spatial information of both ligand-receptor and cell type pair interactions to identify novel ischemia-related intra- and intercellular signalling events, coherent across both single-nucleus and spatially-resolved transcriptomics data.

Moreover, the distinct components of our framework are also complementary when applied to a single technology. For example, I use spatially-informed multi-view modelling ³³ to summarise spatial relationships across the whole slide, and combine it with local spatial metrics to pinpoint interactions at the individual spot/cell level. As showcased by an application using metabolome-transcriptome ¹¹⁸ multi-omics data, LIANA+ summarises CCC interactions relevant for the whole slide, and also identifies the specific subregions within which interactions occur.

Thus, our findings demonstrate that the flexibility and synergy between components of our framework enable unique analyses for the identification of CCC events across a wide range of technologies.

CCC is a multicellular process involving coordinated CCC events across multiple cell types (intercellular programmes). In line with methods that use factorisation approaches to identify gene expression patterns coordinated across cell types ^{3,34,35}, similar approaches were used to identify intercellular programmes across samples ³⁹. While each of the ligand-receptor methods in LIANA+ is typically applied to a single slide or sample, I combine them with factorisations to enable the unsupervised identification of intercellular programmes in multi-sample and multi-condition datasets ^{39,114,121}. Here, I combined NMF with

spatially-informed local interaction metrics to identify cross-conditional intercellular programmes that arise following myocardial infarction ¹¹⁷. Similarly, by repurposing MOFA+ ¹¹⁴, I identified intercellular programmes that highlight the fibroblast-myeloid cell signalling axis as major drivers of fibrosis. Thus, in addition to identifying condition-relevant intercellular patterns, such approaches can also serve as hypothesis-generation tools in CCC. While here each factorization technique is presented with a specific application, they are interchangeable and could be replaced by alternative approaches ^{133–135}.

Moreover, LIANA+ enables deregulated CCC to be identified using simple differential expression analysis. While such an approach requires a specific hypothesis *a priori* and focuses on a single interaction at a time, it is a straightforward and interpretable alternative that can also complement the more complex unsupervised approaches. While I combine unsupervised factorisation with targeted differential analysis to uncover interactions potentially deregulated in ischemia, each approach can also be used independently, depending on the specific objectives and the nature of the dataset at hand.

Intercellular signalling events frequently initiate or are initiated by intracellular signal transduction pathways, thus functioning as intertwined networks of concerted signalling between and within cells. LIANA+ connects intercellular signalling to downstream events by combining analysis enrichment ⁹⁸ and network inference ¹³⁶ methods with existing knowledge resources ⁴⁵. First, it enables the annotation of ligand-receptor interactions to specific contexts, such as pathways, and the subsequent enrichment testing of such annotations. Second, our framework enables the inference of causal networks, which I use here to link CCC and transcription factors deregulated in ischemia. To our knowledge, this is the only approach that considers the sign of signalling and deregulation of inter- and intracellular signalling, thus providing networks better aligned with existing prior knowledge. Though in addition to our such sign-consistent network search, the flexibility of LIANA+ also enables other network approaches to be incorporated ^{7,8,16,101,102}.

The methods implemented in LIANA+ have a number of limitations. First, they use prior knowledge, which is limited, often exhibiting biases and a trade-off between coverage and quality ^{7,9,137}. Most curation efforts have been focused on annotating ligand-receptor interactions ^{4,6}, and additional prior knowledge efforts are needed in particular for the inference of CCC beyond protein-mediated events. Moreover, contextualising prior knowledge to specific cell types, tissues, or disease can help to reduce erroneous predictions. As an example, we developed and used MetalinksDB ¹³⁸, a comprehensive resource for the inference of metabolite-mediated CCC, here customised to brain-specific metabolites. Second, CCC from dissociated single-cell data remains limited to the

co-expression of communication partners, and this co-expression at the transcript level may not translate to the protein level, let alone imply a functional interaction⁹. Likewise, while spatially-resolved data is a step further from its dissociated counterparts, it is limited to the co-localization of molecules. Finally, while I showed the ability of LIANA+ to generate CCC insights across a range of technologies, along with some preliminary evaluations, systematic benchmarks of CCC methods are still pending. Similarly to the evaluations presented in our work, others also exist but they too remain limited to the use of orthogonal modalities such as spatial data^{9,93} or downstream signalling⁹. As emerging technologies¹³⁹⁻¹⁴¹ which provide *bona fide* CCC events, become measurable at scale and widely available, LIANA+ will serve as a facilitator for such benchmarks and comparisons. Therefore, like all individual CCC methods, including those implemented in our framework, LIANA+ remains a tool for hypothesis generation, requiring validation experiments.

Overall, LIANA+ generalises the multifaceted aspects of cell-cell communication inference into synergistic components. These components can be combined in various ways, and their configurations can be tailored to address diverse and emerging questions and datasets. Given the modularity of LIANA+, new methods can be integrated into the framework and immediately tap into the established ecosystem of methods and resources, benefiting from enhanced compatibility and interoperability. I further hope that LIANA+ will be used as a versatile tool for the study of CCC driven by diverse mediators, beyond protein-mediated and metabolite-mediated interactions, expanding the range of CCC events that could be studied, such as host-microbiome interactions¹⁴²⁻¹⁴⁴. Thus, LIANA+ not only stands as a comprehensive and scalable tool for studying communication events but also serves as a catalyst for future developments in the field.

4. Methods

Bivariate Spatially-informed Metrics

Common notations:

$x \in \mathbb{R}^n$ and $y \in \mathbb{R}^n$ are vectors of two variables for spots (or cells) n ,

\bar{x} and \bar{y} is the mean of the variable values,

$w \in \mathbb{R}^{n \times n}$ is a spatial proximity weight matrix indicating the degree of spatial association between spots i and j .

Inspired by schOT¹¹¹, I implemented local weighted variants of common similarity metrics, such as Pearson and Spearman correlation:

$$\text{wPearson}\rho_i = \frac{\sum w_{ij} \sum (w_{ij} x_j y_j) - \sum (w_{ij} x_j) \sum (w_{ij} y_j)}{\sqrt{(\sum w_{ij} \sum (w_{ij} x_j^2) - \sum (w_{ij} x_j)^2)(\sum w_{ij} \sum (w_{ij} y_j^2) - \sum (w_{ij} y_j)^2)}}$$

$$\text{wSpearman}\rho_i = \frac{\sum w_{ij} \sum (w_{ij} r_{x_j} r_{y_j}) - \sum (w_{ij} r_{x_j}) \sum (w_{ij} r_{y_j})}{\sqrt{(\sum w_{ij} \sum (w_{ij} r_{x_j}^2) - \sum (w_{ij} r_{x_j})^2)(\sum w_{ij} \sum (w_{ij} r_{y_j}^2) - \sum (w_{ij} r_{y_j})^2)}}$$

where summation is performed over j and r_{x_j} , r_{y_j} are ranks of x and y for spot j .

A second masked version of Spearman correlation, as proposed and default approach in schOT, was also implemented; where I consider r_{x_j} , r_{y_j} only for spots with non-zero w .

Moreover, I provide weighted Jaccard and Cosine similarity metrics:

$$\text{wJaccard}_i = \frac{\sum_{j=1}^n \min(x_j, y_j) w_{ij}}{\sum_{j=1}^n \max(x_j, y_j) w_{ij}}$$

where:

x and y are vectors of the same length, binarized by setting values > 0 to 1, to signify presence or absence of a read out.

$$\text{wCosine}_i = \frac{\sum_{j=1}^n w_{ij} x_j y_j}{\sqrt{\sum_{j=1}^n w_{ij} x_j^2} \sqrt{\sum_{j=1}^n w_{ij} y_j^2}}$$

As well as a simple spatially-weighted product (as in NICHES¹⁷) and a max-normalised product as a scale invariant metric:

$$\text{wProduct}_i = \left(\sum_{j=1}^n w_{ij} x_j \right) \left(\sum_{j=1}^n w_{ij} y_j \right)$$

$$\text{wNormProduct}_i = \left(\frac{\sum_{j=1}^n w_{ij} x_j}{\max(|x'|)} \right) \left(\frac{\sum_{j=1}^n w_{ij} y_j}{\max(|y'|)} \right)$$

where $x' \in \mathbb{R}^n$ and $y' \in \mathbb{R}^n$ are the spatially-weighted vectors of x and y respectively

I adapted bivariate Global and local Moran's R, extensions of Moran's I¹²⁰, from SpatialDM³¹; both of which are measures of spatial co-occurrence.

Local Moran's R is defined as:

$$\text{Local}R_i = (x_i - \bar{x}) \sum_{j=1}^n w_{ij} (y_j - \bar{y}) + (y_i - \bar{y}) \sum_{j=1}^n w_{ij} (x_j - \bar{x})$$

In contrast to Local Moran's R, Global Moran's R infers the co-clustering of two variables globally, and is defined as:

$$\text{Global}R = \frac{\sum_i^n \sum_j^n w_{ij} (x_i - \bar{x})(y_j - \bar{y})}{\sqrt{\sum_i^n (x_i - \bar{x})^2} \sqrt{\sum_i^n (y_i - \bar{y})^2}}$$

When working interactions, the members of which contain heteromeric complexes, I consider the minimum expression of complex members per spot. Any interactions, the members of which are not expressed in at least 10% (by default) of the spots are excluded.

Local Score Categories

Inspired by GeoDa¹⁴⁵, I categorise local bivariate associations according to the magnitude and sign of the two variables. If spatially-weighted variables are non-negative (e.g. gene expression) then they are z-transformed. Then for each spot i , I categorise interactions according to the sign of the spatially-weighted variables (v) involved in the interaction - i.e. as positive, negative or neither:

$$\text{sign}v_i = \text{sign}\left(\sum_{j=1}^n (v_j w_{ij})\right)$$

Then to obtain a category for the interaction, I combine the sign of the two variables (x and y). If both variables are positive (high-high), then the interaction is positive (1); if one variable is positive (high) and the other negative (low) then the interaction is negative (-1); if both variables are negative (low-low), or either variable is neither (e.g. equals to 0), then the

interaction is labelled as “neither” (0). The latter enables us to distinguish relationships where both variables are highly-abundant (high-high) from those where both values are lowly-abundant (low-low).

For statistical testing of local metrics, I use spot label permutations to generate a Null distribution against which empirical local p-values are computed.

I provide a detailed tutorial on the bivariate metrics at:

<https://liana-py.readthedocs.io/en/latest/notebooks/bivariate.html>

Learning Spatial Relationships across Multi-views

To learn multivariate interactions in space that go beyond bivariate metrics, we re-implemented MISTy’s multi-view learning approach ³³. Our multi-view approach jointly models different spatial and functional aspects of the data, such that it can fit any number of views, and each view can contain any number of variables. As shown in this work, one can use it to jointly model different combinations of RNA expression, cell type proportions, or metabolite peak intensities.

In LIANA+, multi-view objects are represented as subclasses of MuData ¹⁴⁶, modified to ensure the correct format of the views and corresponding spatial connectivities. Each multi-view structure has an intrinsic view (intraview) that contains the target variables of interest for each spatial location. The other views can be considered as “extra” views, composed solely of predictor variables. Predictor variables can also represent a transformation of the variables within the intraview taking into account a specific spatial context, as well as other categories of variables.

We additionally enable the use of different single-view modelling approaches. By default, the models are based on random forests and can capture complex non-linear relationships. Here, we also implemented linear models. For both random forests and linear models, we use the implementations available from scikit-learn ¹⁴⁷.

Once the multi-view structure is defined, each target is modelled by predictors from each view independently. As such, for each target we obtain (1) relationship statistics (‘importances’) for each of the predictors from the distinct views; (2) the relative ‘contribution’ of each view to the joint prediction of each target (3); as well as the goodness of fit (e.g. R^2) of the model.

(1) The statistics for each predictor, signifying its importance in the prediction of a given target variable, are calculated depending on the modelling approach. For random forests, we use the reduction of variance explained that can be attributed to each predictor across all regression trees. For linear models, which were used throughout this manuscript, we use the ordinary least squares t-statistics of the estimated parameters under the zero value null hypothesis. The independent view-specific predictions are combined by a cross-validated

linear meta-model ³³ to obtain the contributions of view-specific models (2), along with the goodness of fit of the overall model, for each target variable (3). In particular, we can discern between the contribution of the intraview, modelled as the intrinsic variability among target variables within the same cells/spots, from the joint predictive contribution of “extra” views, which encode spatial information.

We further implemented a masking procedure according to observation/spot labels (e.g. regions or conditions) within the same slide; essentially masking each view according to labels assigned to observations in the intra view. Prior to masking, we spatially-weight each of the extra views.

To facilitate the use of our multi-view learning approach, I provide in depth tutorials on how to generate and model custom and predefined multi-view structures: <https://liana-py.readthedocs.io/en/latest/notebooks/misty.html>.

Estimation of Spatial Connectivities

As in MISTy ³³, spatial connectivity weights are calculated using families of radial basis:

$w_{ij} = e^{-\frac{d_{ij}^2}{l^2}}$, Gaussian $w_{ij} = e^{-\frac{d_{ij}^2}{2l^2}}$, linear $w_{ij} = 1 - \frac{d_{ij}}{l}$, and exponential kernels

$w_{ij} = e^{-\frac{d_{ij}}{l}}$; where w is a weight matrix $w_{ij} \in [0, 1]$ of shape $n \times n$, d_{ij} is the Euclidean distance between cells or spots i and j , l is a parameter controlling the shape, or bandwidth. We additionally provide a cutoff parameter below which spatial connectivities are set to zero.

When working with multi-modal spatial technologies, the different modalities of which have observations with distinct locations, spatial connectivity weights can be estimated according to a reference coordinate system. Essentially, this enables the application of our MISTy across any number of views, regardless of their dimensions.

In this context, let $X \in \mathbb{R}^{n \times d1}$ and $Y \in \mathbb{R}^{m \times d2}$ represent the data matrices from two distinct modalities. These matrices have the shapes $(n, d1)$ and $(m, d2)$ respectively. Here, n and m represent the number of observations in each modality with $i \neq j$, while $d1$ and $d2$ denote the number of features or dimensions specific to each modality. The spatial connectivity weights are captured in a matrix $W \in \mathbb{R}^{n \times m}$, which quantifies the spatial proximity between each observation in Y (j) to each observation in X (i).

The alignment of these modalities is achieved through the following transformation:

$$Y' = YW$$

In this equation, Y' represents the transformed version of Y , with a new shape of

$Y \in \mathbb{R}^{n \times d^2}$, aligning Y with the observation space of X while maintaining the unique feature dimensions of the modality.

Similarly, to enable the calculation of local metrics that necessitate a calculation of a covariance between features from the modalities prior to spatial weighting or transformation, I used linear grid interpolation as implemented in `scipy`¹⁴⁸. This method interpolates one matrix to the reference coordinate space - typically the modality with a lower number of observations.

Throughout the manuscript, unless otherwise specified, I used Gaussian weights with a bandwidth of 150, and a cutoff of 0.1, and the diagonal (self to self) was set to 0 for MISTy and 1 for the local scores. For both product and normalised product, I additionally apply L1-normalisation to the weights. This adjustment accounts for the variability in the number of neighbours across different spots, implicitly accounted for by the remainder of the local metrics.

Programmatically the calculation of spatial connectivities mirrors Squidpy's `spatial_neighbors` function, and thus spatial connectivities can be easily replaced with Squidpy's neighbourhood graphs¹⁴⁹.

Ligand-Receptor Pathway Enrichment

To perform ligand-receptor pathway enrichment, I first convert gene set (pathway) resources, represented as weighted bipartite graphs where each gene belongs to a gene set, into ligand-receptor sets. Specifically, I assign a weight to each ligand-receptor interaction, based on the mean weight of the ligands and receptors involved in the interaction, also taking into account the presence of heteromeric subunits. Moreover, I assign a given ligand-receptor interaction to a specific gene set (or pathway), only if all members of the interaction are part of the gene set, and in the case of weighted resources are additionally sign-consistent. Finally, once a ligand-receptor resource is generated, I use `decoupler-py` to perform enrichment with univariate linear regression⁹⁸.

In this manuscript, I used the PROGENy resource¹²⁴ to assign pathway annotations to ligand-receptor interactions. In contrast to classic pathway gene sets, PROGENy contains consensually-regulated targets of pathway perturbations¹²⁴, not genes thought to be members of the pathways. However, this resource-conversion procedure is applicable to any resource, including undirected resources, such as GO terms for which all members of a gene set will be assigned a weight of 1.

Hypothesis testing for deregulated CCC across Conditions

To enable hypothesis testing for CCC, similarly to the strategy implemented in MultiNicheNet³⁷, I first generate pseudobulk profiles by summing raw expression counts for each sample and cell type with the decoupler-py package⁹⁸. After filtering low quality genes in (e.g. considering minimum expression in terms of total counts and samples in which the gene is expressed), I perform differential analysis for each cell type independently with DESeq2¹¹³, as implemented in PyDESeq2¹¹².

Once feature statistics per cell type are generated, I transform those into a dataframe of interaction statistics by joining them to a selected ligand-receptor resource, while additionally calculating average feature expression and expression proportions per cell type, based on a user-provided AnnData object¹⁵⁰. Similarly to any other method in LIANA+, interactions expressed in less than 10% (by default) of the cells per cell type are filtered, considering the individual members of heteromeric complexes.

A detailed tutorial is available at: <https://liana-py.readthedocs.io/en/latest/notebooks/targeted.html>

Sign-consistent Intracellular Networks

By combining CCC predictions with prior knowledge networks of intracellular signalling, it is possible to recover putative causal networks linking CCC events to transcription factors. To accomplish this, we used CORNETO¹⁵¹ - a Python package that unifies network inference problems from prior knowledge - to implement a modified version of the integer linear programming (ILP) formulation implemented in CARNIVAL¹⁵².

This modified version of CARNIVAL¹⁵² takes four distinct inputs: (1) a prior knowledge graph (PKN) of signed protein-protein interactions, where nodes are proteins and edges are activating or inhibitory interactions; continuous and signed (2) starting (input) nodes and (3) end (output) node values, with negative values indicating downregulation and positive values indicating upregulation. In addition, we take (4) values for the rest of the nodes in the graph [0, 1] (e.g. gene expression proportions), with higher values incurring less penalty than genes with lower values when the gene is included in the inferred network. Then, a subnetwork, optimised for sparsity, is extracted from the PKN which connects the input (starting) nodes to the output (end) nodes, taking into account both the directionality and sign of interaction.

The resulting inferred network is a directed acyclic graph that connects the (2) input nodes to the (3) end nodes (e.g. receptor to transcription factors), including the values for each edge and node of the graph indicating if the node is upregulated (+1), or downregulated (-1). A node n_c in the graph can be upregulated only if there is at least one selected parent node n_p such that n_p is upregulated and there is an activating edge between n_p and n_c , or n_p is

inhibited and there is an inhibitory edge between n_p and n_c . Similarly, a node n_c can be downregulated if there is a parent node n_p downregulated with an activating edge between n_p and n_c , or if a parent node n_p is upregulated and there is an inhibitory edge between n_p and n_c .

These rules are encoded using linear constraints and continuous/binary variables to define a Mixed ILP problem, which is a particular type of a combinatorial problem with linear constraints. The optimization problem is defined as:

$$\arg \min_{\mathbf{x}^+, \mathbf{x}^-} L(\mathbf{x}^+, \mathbf{x}^-) = \sum_{i: v_i > 0} (1 - x_i^+) v_i + \sum_{i: v_i < 0} (1 - x_i^-) |v_i| + \sum_i \lambda_i (x_i^+ + x_i^-)$$

where x^+ is a vector of binary variables for each node in the PKN indicating whether the node i is upregulated ($x_i^+ = 1$) or not; x^- is a vector of binary variables for each node in the PKN indicating whether node i in the PKN is downregulated ($x_i^- = 1$) or not; v is a vector of values for measured nodes (input nodes and output nodes), where positive values are upregulated species and negative values are downregulated species. For example, v can be estimated as fold change, t-statistic or any other score indicating difference in activity in a protein in the PKN between two conditions.;

In this modified version of CARNIVAL, we additionally introduce λ - a vector of penalties to penalise the inclusion of protein nodes in the resulting inferred network, according to (4) node weights w $[0, 1]$ in the (1) PKN. We introduce λ specifically to discretise highly-expressed from lowly-expressed genes, thus accounting for the differences in drop-out rates across cell types in single-cell data.

We set λ to penalty_{\max} (1 as default) and penalty_{\min} (0.01 as default) according to a threshold (0.1) by default:

$$\lambda_i = \begin{cases} \text{penalty}_{\max} & \text{if } w_i \leq \text{threshold} \\ \text{penalty}_{\min} & \text{if } w_i > \text{threshold} \end{cases}$$

Linear constraints impose conditions on the variables of the ILP problem. For example, a node cannot be upregulated and downregulated at the same time ($x^+ + x^- \leq 1$). The problem includes other variables and linear constraints to guarantee that the final networks valid solutions are acyclic networks and that the rules explained before are respected. Additional information about the formulation can be found in Liu et al. (2019)¹³⁶.

I show the inference of sign-consistent networks downstream of deregulated CCC events, identified using differential expression analysis with PyDESeq2 ¹¹² here: <https://liana-py.readthedocs.io/en/latest/notebooks/targeted.html>

NMF on ligand-receptor local scores

A utility function was implemented that takes an AnnData object ¹⁵⁰ as input and uses Scikit-learn's NMF implementation to factorise the input matrix into two matrices of dimensions k , n and k , d ; where d is the number of features, n is the number of observations (cells); and k is the number of components (factors). To estimate k , I additionally provide an heuristic elbow selection procedure, in which the optimal component number (k) is chosen from a sequential range of components using elbow selection as implemented in the kneedle package ¹⁵³. Selection of optimal k is based on the mean absolute reconstruction error.

LIANA+ in multimodal single-cell & spatial data

To enable the inference of CCC across modalities, the methods implemented in LIANA+ accept MuData objects ¹⁴⁶ as input. These provide essential functionalities to load and store multimodal data ¹⁴⁶, and can be thought of as an extension of AnnData objects ¹⁵⁰, which are the default input of LIANA+ when working with unimodal single-cell or spatial data.

Feature level transformations such as z-scoring or min-max scaling are used to transform features within omics and across omics to a comparable scale to facilitate integration.

Intercellular Communication Factorization with Tensor-cell2cell and MOFA+

Inspired by the CCC factorization approach proposed in Tensor-cell2cell ³⁹ and building on our recent application of MOFA+ to dissociated cross-condition atlases ³, I use the ligand-receptor inference methods from LIANA+ across each sample independently, and then transform this into a multi-view structure of cell type pairs (views), each represented by samples and the ligand-receptor interaction scores in each. To build the multi-view structure, I use the MuData format ¹⁴⁶, and only views with at least 20 (by default) interactions in at least 3 (by default) samples are kept. Moreover, I exclude samples if they have less than 10 interactions (by default) and interactions are considered only if they are present in at least 50% of the samples (by default). Then I use the MOFA+ statistical framework ¹¹⁴ to decompose the variance of the ligand-receptor scores across samples into intercellular communication programmes.

Tutorials on extracting intercellular programmes from single-cell dissociated data with MOFA+ and Tensor-cell2cell are available at:

<https://liana-py.readthedocs.io/en/latest/notebooks/mofatalk.html>

https://liana-py.readthedocs.io/en/latest/notebooks/liana_c2c.html

LIANA+'s Prior Knowledge Component

All components of LIANA+ rely on existing biological knowledge. As such, LIANA+ draws from OmniPath's rich database of ligand-receptor resources ⁴⁵, providing access to 15 different resources, along with a consensus resource. To increase the flexibility of our CCC workflows, the knowledge in LIANA+ can be further expanded by leveraging BioCypher, which provides utilities for the modular and reproducible representation of knowledge ¹¹⁵. For instance, I used Metalinks ^{138,154} - a comprehensive and customisable resource of metabolite-protein interactions, additionally incorporating annotations such as tissues, pathways and diseases.

Spot Calling using Local metrics

To benchmark how well each local score in LIANA+ preserves biological information, I devised spot classification and regression tasks. In the spot classification task, I used four public breast cancer 10X Visium slides ⁹⁵, with annotations labelled as malignant (containing "cancer" in their annotation) or non-malignant spots (any other spot). For each slide, I calculate local ligand-receptor scores using the local metrics in LIANA+. Then for each local metric, I trained and evaluated Random Forest Classifiers, with 100 estimators, using a Stratified K-Fold cross-validation strategy (k=10). AUROC and weighted F1 were calculated on the test sets, and their average across the folds was used in visualisations.

In the regression task, I used a public dataset with 28 10x Visium slides from left-ventricle heart tissues to compare how well different local metrics capture cell-type specific ligand-receptor events. In particular, I checked how well do the local scores LIANA+ predict cell type proportions per spot, inferred using cell2location ²⁸ as done in Kuppe et al., 2022 ¹¹⁷. I used a Random Forest Regressor, with 100 estimators, utilising a K-fold cross-validation training strategy (k=5), and calculated the variance explained (R^2) and root mean squared error for each score. All classification and regression tasks were performed through Scikit-learn (v.1.3.2).

For the inference of ligand-receptor interactions throughout this work, I used LIANA's consensus resource - a resource combining the curated ligand-receptor resources in OmniPath ⁴⁵.

Spatial Colocalisation Evaluation

To benchmark how well each ligand-receptor method in LIANA+ captures potential interactions, I used five processed and recently published spatially-informed single-nuc RNA-seq datasets ¹³⁰.

Making use of spatial information for each dataset, I estimated Global Moran's R using LIANA+ to generate the assumed ground truth - i.e. ligand-receptors and cell types seen to

be colocalized better than random (Moran's $R > 0$ and $FDR < 0.05$). Then I ran all ligand-receptor methods in LIANA+ without taking spatial information into account.

AUROC was calculated for the whole distribution of each ligand-receptor method's scoring functions. To calculate Balanced Accuracy and Normalised F1 (see below), I used false positive filtering thresholds as suggested by each of the methods' authors (if available). For CellPhoneDB, CellChat, and Geometric Mean, interactions with p-values below 0.05 were filtered. For CellChat, I only kept interactions for which both ligand and receptor p-values were also under 0.05. Similarly, for Connectome and log2FC, interactions were kept only if both ligand and receptor p-values were under 0.05 and had a positive scaled weight or log-fold change, respectively. For SingleCellSignalR, I considered ligand-receptor interactions with LRscores above 0.6. For LIANA's Rank Aggregate I kept only those with a magnitude rank < 0.05 . Since NATMI and scSeqComm don't necessarily suggest a threshold, were deemed kept if they were within the top 5% of the specificity weight and inter score distributions, respectively. Similarly, when considering the individual scoring functions for each method, I considered only the top 5% for each as positive predictions across all methods.

Normalized F1 = $\frac{1}{2} \left(\frac{F1_{\text{observed}}}{F1_{\text{permuted}}} \right)$ where observed is the F1 score for the actual ligand-receptor predictions, while permuted was generated by permuting the predictions 100 times.

Balanced Accuracy = $\frac{1}{2} \left(\frac{TP}{TP + FN} + \frac{TN}{TN + FP} \right)$ where TP stands for True Positives, TN is for True Negatives, FP for False Positives, and FN for False Negatives.

Sample Label Classification

For the condition classification task, building on a similar approach ³⁹, I used public, pre-processed, cross-conditional atlases (**Supp. Table 3**), each selected such that they include more than five samples per condition following preprocessing. To ensure that only high-quality samples were used in each of the atlases, I removed any samples with less than 1,000 cells or z-transformed total counts above or below a z-score of 3 and -2, respectively. In the Carraro et al. dataset ¹⁵⁵, I kept samples with more than 700 cells. Moreover, only cell types found in at least 5 samples and with at least 20 cells in each individual sample were considered. To ensure that the samples were balanced between the conditions, if either condition had a sample ratio higher than 1.5 x the number of samples in the other condition, then the overrepresented condition was subsampled to the number of samples in the

underrepresented one. Each dataset was normalised to 10,000 total counts per cell and log_{1p}-transformed.

Subsequent to preprocessing, I inferred ligand-receptor interactions at the cell-type level using the homogenised methods in LIANA+, independently for each sample. Any interactions not expressed in at least 10% of the cells in both source and receiver cell types were filtered.

Then the output from LIANA+ was converted to the structures used by the factorization approaches employed by MOFA+ and Tensor-cell2cell - a multi-view and a 4D tensor, respectively. For running both factorisation approaches, I consider interactions only if they were present in 33% of the samples, and any interactions missing in a sample were assumed to be biologically-meaningful and assigned as zero. For all datasets, I decomposed the CCC events into 10 factors, except Reichart et al ¹⁵⁶, which was decomposed into 20 factors due to its larger sample size.

Using the factor scores for each method-factorization approach combination I then performed a classification task, similar to the one from Armingol et al. ³⁹. Specifically, a Random Forest Classifier, with 100 estimators, was trained and evaluated on the sample factor scores computed for each score-factorization combination, utilising a Stratified K-Folds cross-validation strategy (k=3), performed over 5 seeds. Then the mean Area Under the Receiver-Operator Curve (AUROC) and weighted F1 scores were then calculated on testing set's the probabilities and label predictions, respectively.

Analysis of Spatially-resolved Multi-omics Data from Murine Parkinson's disease model

I obtained a pre-processed data set of three murine brain sections following 6-hydroxydopamine perturbation in one hemisphere ¹¹⁸; with joint metabolite and transcriptome measurements generated with MALDI-MSI and 10X Visium technologies, respectively. I processed the metabolite and count matrices for each slide separately, applying standard log_{1p} normalisation and standard quality control measures to the gene expression data. For the metabolite intensities, I used total count normalisation followed by z-transformation. The observations of two modalities were manually aligned following the identification of tissue-containing observations for the metabolome data, using a procedure similar to the original publication ¹¹⁸. Cell type proportions were inferred using Tangram's cell cluster level approach ¹⁵⁷ with default parameters and an annotated single-cell data as reference ¹¹⁹.

Following the alignment of the two modalities, I modelled metabolite peaks (intraview) with cell type proportions and brain-specific receptors as predictors (extra views). I bypassed modelling the intraview - i.e. I did not model each metabolite peak by the remainder of the

peaks (as done by default), since I was interested solely in the predictive performance of the extra views (receptors and cell types). I focused on the intersection of the top 250 highly-variable metabolites (targets) across the three slides, and excluded any predictors with little-to-no variation - i.e. genes not within the top 12500 highly-variable genes; and cell types with a coefficient of variation below the 20th percentile. Brain-specific receptors were obtained from MetalinksDB ¹⁵⁸, customised to include only metabolites found in the brain or cerebrospinal fluid. Finally, I used spatially-weighted Cosine similarity on the z-transformed matrices of each modality to estimate the local scores and their corresponding P-values and categories. For all analyses I used a bandwidth of 1,000 with a cutoff of 0.1 to calculate the spatial connectivities.

Analysis of Human Myocardial Infarction

Following basic filtering and standard log₁₀p-normalisation, I estimated ligand-receptor local scores using spatially-weighted Cosine similarity on each of the 28 processed 10X Visium transcriptomics slides ¹¹⁷. I considered interactions whose members were expressed in at least 10% of the spots. Then I concatenated the resulting ligand-receptor AnnData objects (slides), and kept only those interactions present in at least 10 of the slides. Subsequently, I decomposed the concatenated object with NMF. I used Fisher's exact test to examine whether specific condition labels were enriched when considering samples with average factor scores above the 75th quantile.

Pathway activities of ligand-receptor interaction loadings were calculated using linear regression ⁹⁸ and sets of ligand-receptor pathways, annotated using the PROGENy pathway resource ¹²⁴ with all genes.

I used a pre-processed dataset of 29 single-nuc samples from the same publication ¹⁵⁹. Raw gene counts were normalised to 10,000 total counts per cell and log₁₀p-transformed. I inferred ligand-receptor interactions per sample using LIANA's magnitude rank aggregate - a consensus of multiple magnitude-focused scores (**Supp. Table 2**), considering only interactions with all members expressed in at least 10% of the cells per cell type. I further inferred ligand-interactions only if they were deemed as condition-relevant in the spatial analysis - i.e. those with at least one standard deviation above the mean per NMF factor. Moreover, I considered cell type pairs to interact, only if they had a strong spatial relationship across all 10X Visium slides (target $R^2 > 0.05$; median t-value > 1.645); as per a misty run predicting each cell type (inferred with cell2location ²⁸) by the remainder. Then I decomposed the obtained ligand-receptor interactions from the dissociated data with MOFA+, considering interactions with at least 15 interactions in 30% of the samples, and views with at least 10 samples. Any missing interaction values were filled with zeroes.

For hypothesis-testing I generated a pseudobulk profiles for cell type using decoupler-py, considering only genes with at least 10 counts across each of the samples or at least 20 counts in total, a large_n of 5, and expressed in at least 10% of the samples ⁹⁸. Then within each profile I performed differential expression analysis with PyDeSeq2^{112,113}, contrasting ischemic samples versus the rest - i.e. I treated fibrotic and myogenic samples as baseline references. The output statistics were then converted into a single dataframe of ligand-receptor differential statistics using LIANA+, keeping only interactions all members of which (including complex subunits) were expressed in at least 10% of the cells in source and target cell types.

Then I estimated TF activities using the Wald statistics from PyDeSeq2 with univariate linear regression ⁹⁸ and CollecTRI ¹⁶⁰. For the inference of the downstream signalling events, I obtained OmniPath's protein-protein interaction network, considering interactions with consensus direction and a curation effort ≥ 3 . Then using CORNETO (v0.9.1-alpha.5), I inferred the plausible causal networks propagating via the interaction between FN1/SPP1 and ITGA5&ITGB1, down to all TFs identified as significantly deregulated (FDR < 0.05) in myeloid cells. I used gene proportions with a cutoff of 0.1, such that nodes above the cutoff were assigned a penalty of 1, and those below a penalty of 0.01. An edge penalty of 0.02 was also used, and to ensure ILP space was thoroughly explored the problem was solved 100 times, each time introducing small amounts of uniform noise. I used the Gurobi ¹⁶¹ solver under an academic licence. Then the acyclic subnetworks obtained by each iteration were concatenated, such that edges from any solution were kept, and the network union was visualised with CytoScape ¹⁶².

Analysis of Murine Acute Kidney Injury Data

I first filtered the preprocessed single-cell dataset, with pre-annotated cell types, to only those cell types with at least 15 cells in at least 10 samples; additionally excluding urothelial cells as they are not expected to communicate with most of the other cell types in the kidney. Following total count and log_{1p} normalisation, I inferred ligand-receptor interactions using LIANA's consensus method (**Supp. Table 2**), excluding any interactions not expressed in at least 10% of the cells in both the source and target cell types. Then I transformed the resulting ligand-receptor interactions into views representing cell type pairs, keeping only those interactions present in at least 25% of the samples, views with at least 15 interactions and at least 5 views. Finally, I decomposed those views into 5 factors using MOFA+. Kruskal Wallis test was performed on the sample loadings for Factor 1.

For the spatial data, I filtered the preprocessed 10x slides, such that only spots with at least 400 genes expressed and genes expressed in at least 5 spots were kept. I additionally excluded any spot outliers according to mitochondrial, ribosomal and total count content,

using comparable but slide-specific thresholds. Then for the interactions of interest identified in the dissociated datasets, I calculated local cosine similarity and global Moran's R.

Data Availability

Processed myocardial infarction single-nucleus and 10X Visium data was downloaded from the Human Cell Atlas (<https://data.humancellatlas.org/explore/projects/e9f36305-d857-44a3-93f0-df4e6007dc97>), also available via Zenodo at <https://zenodo.org/records/6578047>.

Processed breast cancer 10× Visium slides⁹⁵ (GSE176078; <https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE176078>) were obtained via <https://zenodo.org/record/4739739>.

Spatially-resolved metabolome-transcriptome data¹¹⁸ was obtained from <https://data.mendeley.com/datasets/w7nw4km7xd/1>, also available under GEO repository accession number [GSE232910](https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE232910).

Annotated single-cell mouse brain data¹¹⁹, used for reference for deconvolution, was obtained from <http://mousebrain.org/adolescent/>, with GEO accession number [GSE178265](https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE178265).

Single-nuc¹⁶³ and spatially-resolved¹⁶⁴ mouse kidney injury datasets were obtained via GEO accession numbers: GSE139107 (<https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE139107>) and GSE182939 (<https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE182939>), respectively.

Code Availability

LIANA+ is available at <https://github.com/saezlab/liana-py>, along with detailed tutorials describing the distinct components presented here (<https://liana-py.readthedocs.io>). LIANA+ is regularly released on Github and stable versions are released on PyPI (<https://pypi.org/project/liana/>). The code for the analyses presented in this manuscript is available at https://github.com/saezlab/lianaplus_manuscript.

Appendix

Supplementary table 1. Feature comparison of selected CCC tools.

	LIANA+	LIANA	CellPhoneDB	CellChat	Multi-NicheNet	Tensor-cell2cell	Scriabin	SpatialDM	COMOT	NCEM
Single-cell (Dissociated) Inference										
Group-based LR	9 methods + Consensus	6 methods + Consensus	✓	✓	✓	✓	x	x	x	x
Group-free LR	x	x	x	x	x	x	✓	x	x	x
Spatial Inference										
Global										
Bivariate	✓	x	x	x	x	x	x	✓	x	✓
Multi-view	✓	x	x	x	x	x	x	x	x	✓
Local										
Bivariate	✓	x	x	x	x	x	x	✓	x	x
Multi-variate	x	x	x	x	x	x	x	x	✓	x
Multi-condition										
DEA-Based	✓	x	✓	x	✓	x	x	x	x	x
Hypothesis-free	✓	x	x	✓	x	✓	✓	*1	x	x
Multimodal										
Handles multimodal data	✓	x	x	x	x	x	x	x	x	x
Knowledge										
Protein-mediated LR	15 Resources + Consensus	15 Resources + Consensus	✓	✓	✓	✓	LIANA's Resources	✓	✓	✓
Metabolite-mediated LR	✓	x	✓	x	x	x	x	x	x	x
LR Pathway Annotations	*2	x	x	✓	x	✓	x	✓	x	x
Downstream Signalling	✓	x	x	x	✓ (extensive)	x	✓ (via NicheNet)	x	x	x
Misc										
Language	Python	R	Python	R	R	Python	R	Python	Python	Python
*1 SpatialDM uses z-scores across samples to find differentially deregulated LRs, other tools utilise factorization approaches for hypothesis-free multi-condition analysis										
*2 LIANA provides a flexible function to annotate interactions according to any pathway gene set										

Supplementary table 2. Single-cell ligand-receptor methods implemented in LIANA+

Method*	Magnitude	Specificity
CellPhoneDBv2 ⁴	$LRmean_{k,ij} = \frac{L_{C_i} + R_{C_j}}{2}$	CellPhoneDBv2's cell label permutation approach; See ρ
Geometric Mean	$LRgeometric.mean_{k,ij} = \sqrt{L_{C_i} \cdot R_{C_j}}$	See ρ
CellChat† ⁶	$LRprob_{k,ij} = \frac{L_{C_i}^* \cdot R_{C_j}^*}{Kh + L_{C_i}^* \cdot R_{C_j}^*}$ where $Kh = 0.5$ by default and L^* & R^* are aggregated using Tuckey's Trimean function.	See ρ
SingleCellSignalR ⁵³	$LRscore_{k,ij} = \frac{\sqrt{L_{C_i} R_{C_j}}}{\sqrt{L_{C_i} R_{C_j}} + \mu}$ where μ is the mean of the expression matrix M	-
NATMI ⁵²	$LRproduct_{k,ij} = L_{C_i} R_{C_j}$	$SpecificityWeight_{k,ij} = \frac{L_{C_i}}{\sum^n L_{C_i}} \cdot \frac{R_{C_j}}{\sum^n R_{C_j}}$
Connectome ⁵	$LRproduct_{k,ij} = L_{C_i} R_{C_j}$	$LRz.mean_{k,ij} = \frac{z_{L_{C_i}} + z_{R_{C_j}}}{2}$ where z is the z-score of the expression matrix M
LogFC‡	-	$LRlog2FC_{k,ij} = \frac{Log2FC_{C_i,L} + Log2FC_{C_j,R}}{2}$
ScSeqComm (intercellular scores only) ¹⁰²	$P(X) = \Phi\left(\frac{X - \mu}{\sigma/\sqrt{n}}\right)$ $LRinterscore_{k,ij} = \min(P(L_{C_i}), P(R_{C_i}))$ Where Φ is the CDF of a normal distribution, μ is the mean, σ is the standard deviation, and n is the number of observations	-
LIANA's Consensus#	Uses all of the above, except Geometric mean, <u>independently</u> for both magnitude and specificity	
Shared Notation: k is the k-th ligand-receptor interaction		

L - expression of ligand L ; R - expression of receptor R ; See \bar{X}
 C - cell cluster
 i - cell group i
 j - cell group j
 M - a library-size normalised and log1p-transformed gene expression matrix
 X - normalised gene expression vector

Permutations to calculate specificity:

$$p\text{-value}_{k,ij} = \frac{1}{P} \sum_{p=1}^P [fun_{permuted}(L_{C_i}^*, R_{C_j}^*) \geq fun_{observed}(L_{C_i}^*, R_{C_j}^*)]$$
 ; where P is the number of permutations, and L^* and R^* are ligand and receptor expressions aggregated by group (cluster) using fun ; arithmetic mean for CellPhoneDB and Geometric Mean, and Tuckey's TriMean for

CellChat:
$$TriMean(X) = \frac{Q_{0.25}(X) + 2 \cdot Q_{0.5}(X) + Q_{0.75}(X)}{4}$$

Consensus

First, a normalised rank matrix[0,1] is generated separately for magnitude and specificity as:

$$r_{ij} = \frac{rank_{ij}}{\max(rank_i)} \quad (1 \leq i \leq m, 1 \leq j \leq n)$$

where m is the number of ranked score vectors, n is the length of each score vector (number of interactions), $rank_{ij}$ is the rank of the j -th element (interaction) in the i -th score rank vector, and $\max(rank_i)$ is the maximum rank in the i -th rank vector.

For each normalised rank vector r , we then ask how probable it is to obtain $r_{(k)}^{null} \leq r_{(k)}$, where $r_{(k)}^{null}$ is a rank vector generated under the null hypothesis. The RobustRankAggregate method¹⁶⁵ expresses the probability $r_{(k)}^{null} \leq r_{(k)}$ as $\beta_{k,n}(r)$ through a beta distribution. This entails that we obtain probabilities for each score vector r as:

$$p(r) = \min_{1, \dots, n} \beta_{k,n}(r) * n$$

where we take the minimum probability p for each interaction across the score vectors, and we apply a Bonferroni multi-testing correction to the P-values by multiplying them by n .

Notes:

Δ Some differences are expected with the original implementations largely due to different preprocessing steps which LIANA+ harmonised across the different methods. Specifically, LIANA+ considers the minimum score (e.g. average expression) for complex subunits, while some methods consider the mean, geometric mean, or simply do not account for complexes at all.

† The original CellChat implementation also uses information of mediator proteins and pathways, which are specific to the CellChat resource. To enable CellChat's scores to be resource-agnostic we do not utilise this information

⌘ While we refer to the genes as ligands and receptors for simplicity, these can represent the gene expression also of membrane-bound or extracellular-matrix proteins, as well as heteromeric complexes for which the minimum expression across subunits is used.

‡ 1-vs-rest cell group log2FC for each gene is calculated as

$$\log_2 FC = \log_2(\text{mean}(X_i)) - \log_2(\text{mean}(X_{\text{not}_i}))$$

* LIANA considers interactions as occurring only if both the ligand and receptor, as well as all of their subunits, are expressed above a certain proportion of cells in both clusters involved in the interaction (0.1 by default). This can be formulated as an indicator function as follows:

$$I \left\{ L_{C_j}^{\text{expr.prop}} \geq 0.1 \text{ and } R_{C_j}^{\text{expr.prop}} \geq 0.1 \right\}$$

Supplementary table 3. Cross-conditional atlases used in the sample classification task

Dataset	Organ	Samples‡	Condition	Reference	Data URL
Kuppe	heart	23	Cardiac Infarction	117	https://cellxgene.cziscience.com/collections/8191c283-0816-424b-9b61-c3e1d6258a77
Reichart	heart	126	Cardiomyopathies	156	https://cellxgene.cziscience.com/collections/e75342a8-0f3b-4ec5-8ee1-245a23e0f7cb
Carraro	lung	16	Cystic fibrosis	155	https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE150674
Haberman	lung	18	Pulmonary Fibrosis	166	https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE135893
Velmeshev	brain	38	Autism Spectrum Disorder	167	https://codeocean.com/capsule/9737314/tree/v2; https://www.ncbi.nlm.nih.gov/bioproject/PRJNA434002/
‡ Following quality control and included in the classification					

Supplementary Note 1. Spot calling Evaluation

First, I evaluated the ability of the spatially-weighted local metrics to classify malignant and non-malignant spots in four breast cancer slides⁹⁵ (Methods). All scoring functions did well at classifying malignancy (AUROC > 0.9; weighted F1 > 0.85; **Supp. Fig. S2A-B**). Spatially-informed Jaccard, Cosine, and the products had slightly higher AUROCs (average ≈ 0.95) and F1 scores (average > 0.88) than other methods across the slides, generally followed by weighted Pearson and Spearman correlations, masked Spearman correlation, and finally bivariate Moran's R (**Supp. Fig. S2A-B**).

Second, using 28 spatial transcriptomics slides from myogenic, ischemic, and fibrotic heart tissue upon myocardial infarction¹¹⁷, I evaluated the ability of local ligand-receptor scores to recover cell type proportions (Methods). I noted that the spatially-weighted products, along with Cosine and Jaccard similarities, had slightly higher predictive performance in ischemic ($R^2 > 0.32$) and fibrotic tissues ($R^2 > 0.27$) than the rest of the metrics, while Moran's R did best in myogenic slide ($R^2 \approx 0.13$) (**Supp. Fig. S2C**); with similar results also observed in terms of Root Mean Squared Error.

In summary, all spatially-informed local scores in LIANA+ preserved the biological signal of gene expression. Yet, our results suggested that spatially-weighted products, Jaccard, and Cosine similarities, which performed best in both the regression and classification tasks (albeit marginally better than other methods), might all be most suitable as a default local scoring function. Since Cosine Similarity is bound between -1 and +1, and it does not require the data to be binarized, I chose it as the default metric - used throughout the manuscript.

Nevertheless, the other scoring metrics are likely suited better for other tasks. For example, spatially-weighted Jaccard index should be well suited for categorical or binary data. Spearman correlation should be more relevant when inferring relationships between ordinal, ranked, or non-linear variables. Similarly, metrics such as Moran's R, or the simple and normalised products, are useful in cases in which there is no overlap between the two variables (e.g. one-hot encodings), as they weigh the spatially- variables prior to examining any covariance between them. Thus, the choice of metric should take in consideration the data and task at hand.

Supplementary Note 2. *Combinations of ligand-receptors with MOFA+ and Tensor-cell2cell evaluated using Sample Label Classification.*

In contrast to our hypothesis-testing approach, and those proposed by other methods^{29,37,106}, unsupervised, or higher-order dimensionality reduction approaches, can simultaneously model CCC events across samples and cell types³⁹. As a consequence, rather than considering each interaction on its own, these approaches capture and summarise coordinated interactions into intercellular programmes. Moreover, these approaches also inform us about the most relevant cell types or interactions that separate the samples according to some condition, along with summaries of model performance, such as the amount of variance explained or reconstruction error.

Recently, I aligned Tensor-cell2cell and LIANA to enable the user-friendly inference of context-specific CCC patterns with any method or resource from the latter¹²¹. Here, I showcase the use of an alternative unsupervised approach leveraging the MOFA+ framework¹¹⁴.

This approach inherits the efficiency and flexibility of MOFA+ to enable factor analysis of CCC interactions by modelling pairs of cell groups as views. As such, it enables feature flexibility across the cell type pairs, and provides ligand-receptor importances per cell type pair and factor. Moreover, it allows the simultaneous decomposition of the input data and alignment with additional covariates¹³³, as also proposed by a recent CCC decomposition approach¹⁶⁸.

To evaluate the ability of both Tensor-cell2cell and MOFA+ in identifying intercellular programmes that distinguish samples from different conditions, I set up a classification task (Methods). I used five public cross-conditional atlases from the human heart, lung, and brain (**Supp. Table 3**), combining each of the ligand-receptor methods in LIANA+ with both dimensionality reductions (**Supp. Fig. S5A**).

In each dataset, I inferred interactions independently for each sample using the scoring functions from each ligand-receptor method in LIANA+, utilising magnitude-based scores if provided (**Supp. Table 2**). Then, I used MOFA+ and Tensor-cell2cell to decompose the ligand-receptor output, obtained per sample (**Supp. Fig. S5A**). Using a binary classification setup, I calculated AUROC and F1 for each method-factorisation combination to examine well each classified condition (**Supp. Fig. S5A**; Methods).

I saw that all combinations performed better than random in most datasets (**Supp. Fig. S5B**), with log2FC (AUROC=0.89) with the expression average of ligand-receptor means (CellPhoneDB; AUROC=0.87) performing best when combined with MOFA+; while the rest of the methods had an AUROC \approx 0.8, except CellChat (AUROC=0.62); though this does not necessarily reflect the performance of the CellChat. When combined with Tensor-cell2cell, methods showed slightly higher variance; with the product (AUROC=0.84) and Rank

Aggregate (LIANA's consensus; AUROC=0.82) having the highest AUROCs, while the rest of the methods' AUROCs ranged between 0.69 and 0.76, except SingleCellSignalR (AUROC=0.6). I also saw similar results when using weighted F1 (**Supp. Fig. S5C**). Moreover, methods showed on average higher AUROCs when combined with MOFA+ than Tensor-cell2cell. With the exceptions of CellChat which performed better when combined with Tensor-cell2cell, and the Carraro dataset on which methods combined with Tensor-cell2cell showed better performance. This potentially reflects an intrinsic difference of the regularizations used between the two approaches. Specifically as MOFA+ attempts to enforce orthogonality ¹¹⁴, while the non-negative tensor component analysis (PARAFAC) used by Tensor-cell2cell, which can be thought of as a higher-order extension of NMF ³⁹, does not.

Overall, our results show that both Tensor-cell2cell and MOFA+ consistently capture condition-relevant intercellular patterns, regardless of the ligand-receptor method at hand.

Supplementary Note 3. *LIANA+ leverages a rich knowledge base to link CCC events to intracellular signalling.*

Leveraging knowledge from OmniPath⁴⁵ and BioCypher¹¹⁵ (**Supp. Fig. S6A**), LIANA+ provides access to both protein-mediated and metabolite-mediated¹⁵⁸ ligand-receptor interactions, which can be further annotated and contextualised to pathways, disease, or locations of interest (**Supp. Fig. S6B**). Such annotations can also be used to perform downstream enrichment analysis on the output of any implemented CCC methods. For example, in section 2.2 I used enrichment analysis⁹⁸ to identify active pathways¹²⁴ associated with ligand-receptor loadings from ischemic and fibrotic heart regions¹¹⁷.

In addition, LIANA+ can infer signalling networks to link identified CCC events to downstream intracellular signalling pathways and transcription factors. In contrast to existing network methods used in the CCC field^{7,8}, our approach considers the direction of deregulation for nodes of interest. Specifically, it evaluates the sign of deregulation (activation or inhibition) in ligands, receptors, and transcription factors, alongside the signalling pathways linking them. To do so, I incorporate knowledge of protein-protein interactions (activating or inhibiting), ligand-receptor interactions, and transcription factors with their targets (**Supp. Fig. S6A**; Methods). Then a network-optimisation approach^{136,151} (**Supp. Fig. S6C**) is used to identify causal paths that connect deregulated CCC events (input nodes) with active transcription factors (output nodes) (**Supp. Fig. S6D**; Methods). It is further noteworthy that the approach implemented here is not bound specifically to linking protein-mediated CCC interactions to TFs, but can be flexibly applied to find links between any set of nodes, including such between metabolites and downstream signalling¹⁵², or vice versa.

As such, LIANA+ provides a comprehensive suite to decipher and interpret intracellular signalling related to CCC.

Supplementary Note 4. Evaluation of ligand-receptor methods using spatially-resolved, transcriptome-wide single-cell RNA-Seq data

I used the recent slide-tags technology¹³⁰ to compare ligand-receptor methods' performance⁹. In contrast to our previous work⁹, and that of others^{93,169}, which relied on the integration of matched spatial and single-cell RNA-Seq data, this technology enables the inference of ligand-receptor interaction and generation of assumed truth using the same data. Moreover, rather than using solely the colocalization of cell types alone⁹, here I additionally take into account the colocalization of ligand-receptor interactions (Methods).

Using this setting I compared CellPhoneDB⁴, CellChat⁶, Connectome⁵, NATMI⁵², SingleCellSignalR⁵³, scSeqComm¹⁰², log2FC, Geometric Mean, and LIANA's consensus (Rank Aggregate) across five slide-tag datasets¹³⁰. First, I generated AUROC curves, and saw that the individual scoring functions of most methods performed marginally better than random (AUROC > 0.5), with inter-dataset variance being too high to suggest which method works best (**Supp. Fig. S8A**).

Next, I evaluated the effectiveness of the false positive filtering thresholds suggested by each method (if available; Methods). To account for the unequal class distributions, I used Balanced Accuracy to compare the methods. I again saw that the methods' performances were closely aligned with random (**Supp. Fig. S8B**). Finally, as done by a previous benchmark⁹³, I computed F1 scores, but saw a notable correlation with the number of interactions retained by each method (Spearman's rho = 0.765). Moreover, the varied filtering processes across methods result in each predicting a distinct set of interactions, thereby further limiting the effectiveness of this metric. As such, I opted to use a normalised F1 score which considered the random outcome anticipated for each method (Methods). Using these normalised F1 scores, I again saw that the methods and their individual scores (when capped at the 95th quantile) performed only slightly better than randomly (**Supp. Fig. S8C&D**).

Taken together, I saw that methods generally exceeded random performance, albeit marginally. Yet, the large variance across datasets and methods limited our ability to suggest the method that works best. Moreover, such a setting can only suggest which interactions are unlikely to occur, but remains limited as expression co-localisation does not necessitate interaction. Instead, our evaluation suggests that regardless of method, ligand-receptor interaction inference, uninformed of spatial location, is anticipated to result in high false positive rates. Thus, our results highlight the need for supporting information⁹.

Supplementary Note 5. LIANA+ enables joined CCC single-cell and spatial analysis

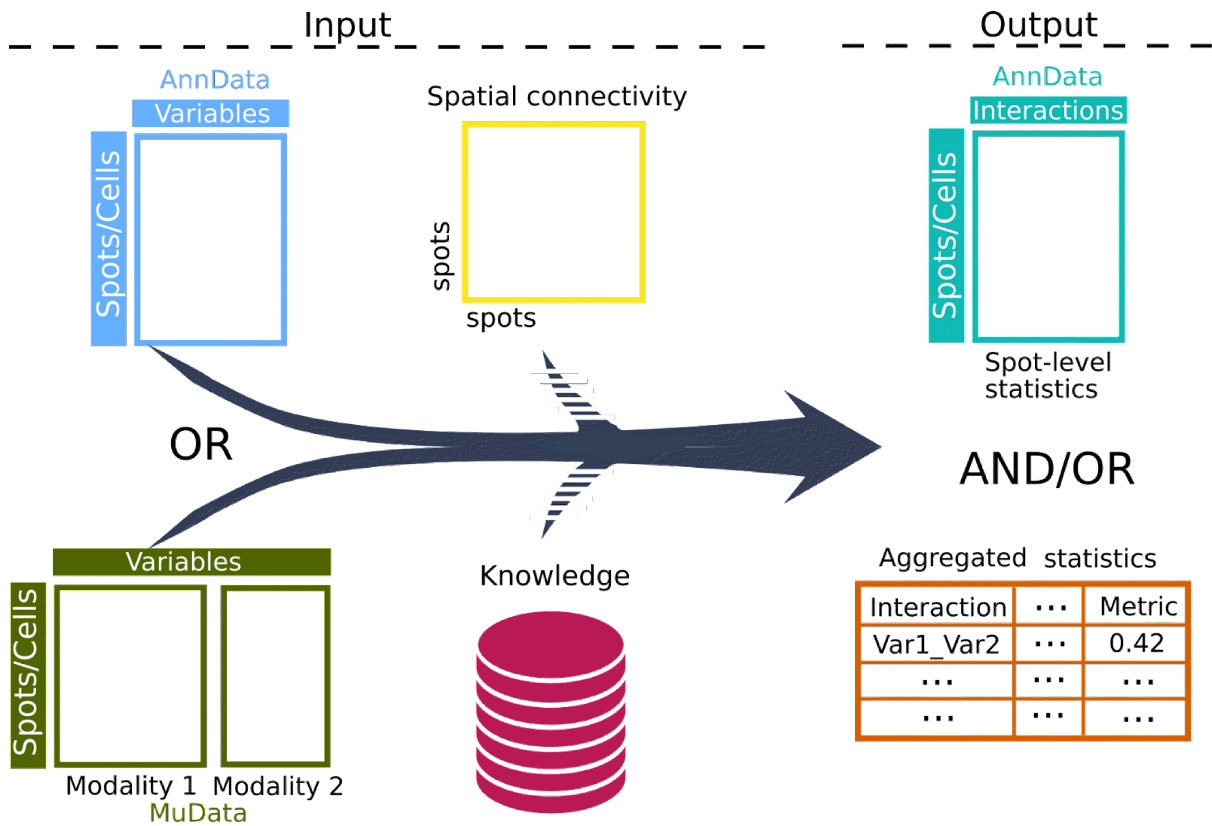
To jointly study CCC events in dissociated and spatially-resolved data, I used LIANA+ with two public datasets of mouse acute kidney injury (AKI) model ^{163,164}. Both datasets employed a time course experimental design, in which murine kidneys were sequenced following bilateral ischemia-reperfusion injury (**Supp. Fig. S9A**).

First, using a single-nucleus AKI atlas (n=24) ¹⁶³, I inferred potential ligand-receptor interactions between groups of cells at the sample level and decomposed the interactions with MOFA+ into a set of factors (Methods), with the aim to identify deregulated intercellular programmes associated with kidney injury. After quality control (Methods), I analysed CCC interactions across 88 cell type pairs and saw that Factor 1 separated early acute time points from the rest (Kruskal-Wallis P-value = 0.0069; **Supp. Fig. S9B**). Sample factor scores associated in Factor 1 were highest at the 12 hour time point, earlier than previously CCC results in a similar mouse model by Li et al., 2022 ¹⁷⁰. Factor 1 explained on average 13.5% of the variability (R^2) of ligand-receptor interactions across cell-type pairs, with Fibroblasts being involved in the several cell type pairs with variances explained > 30%, including their interaction as potential sources of communication with Proximal tubule epithelia ($R^2 = 52.7\%$; **Supp. Fig. S9C**). Fibroblasts were also the recipient cell type with the highest mean variance explained (average $R^2 = 23.5\%$; **Supp. Fig. S9C**), likely associated with their potential role in mediating the repair process following kidney injury ¹⁷¹.

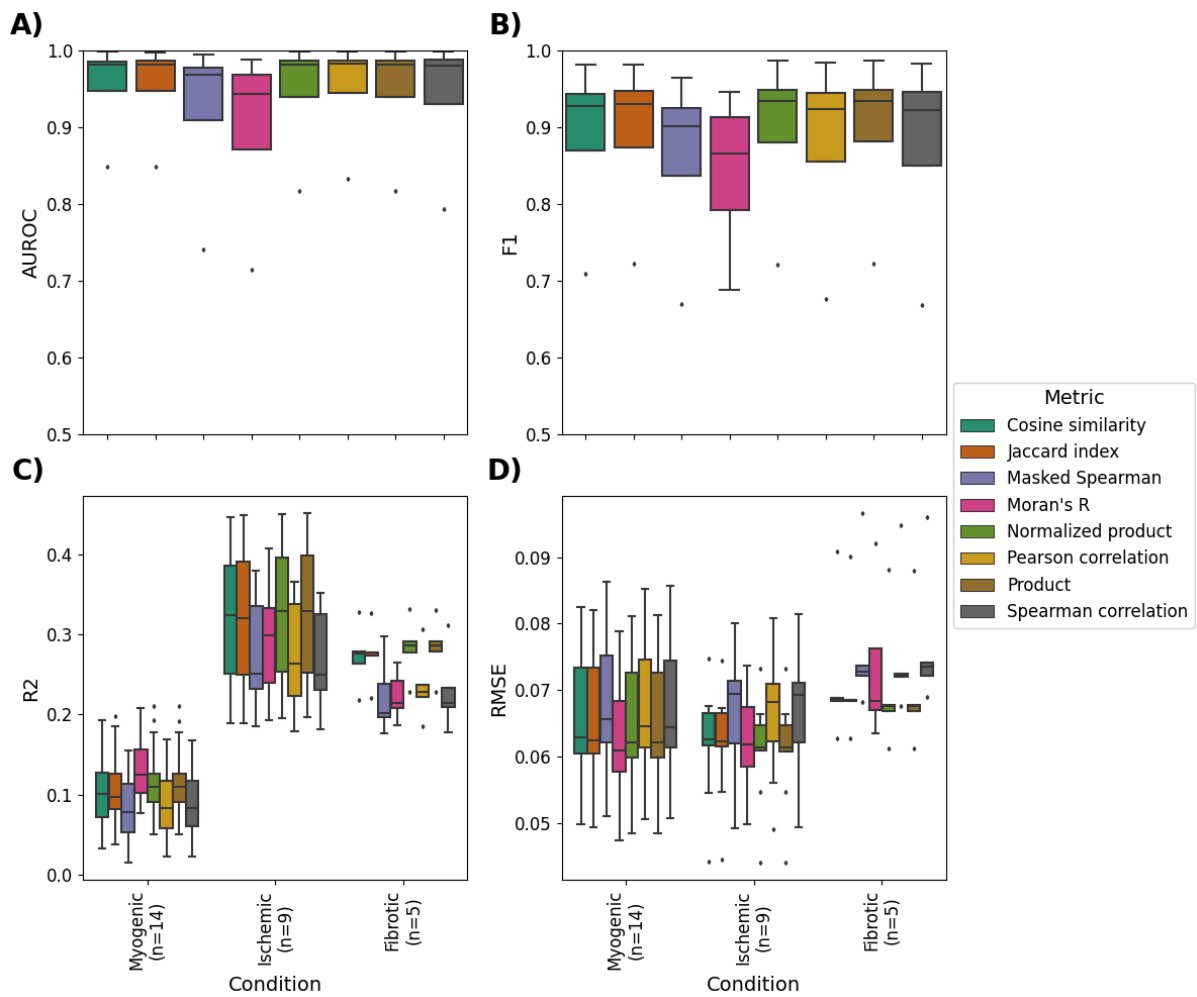
Within the top 15 interactions associated with Factor 1 (**Supp. Fig. S9D**), I noted several potential interactions that involved *Spp1* and *Tnc*, known to contribute to extracellular-matrix remodelling and tissue repair ^{172,173}. Other extracellular-matrix interactions, such as *Slit2* and *Robo1/2*, as well as *Lama2* & *Dag1*, were negatively associated with Factor 1 (**Supp. Fig. S9D**). Thus, Factor 1 potentially represents an intercellular response related to the disruption of the extracellular matrix and its remodelling.

To see if the interactions found between groups of dissociated cells are also captured in spatially-resolved data, I inferred potential interactions using LIANA+'s spatial component in five 10x Visium slides from the same AKI model ¹⁶⁴. I saw that the interactions between *Spp1* and the *Itgav/Itgb1* integrin complex increased both in spatial coverage, as well as co-clustering (Moran's R) albeit low, subsequent AKI (**Supp. Fig. S9E**); in line with findings from Li et al., 2022 in dissociated single-cell data ¹⁷⁰. Specifically, I saw that in the control slide, the interaction was localised mostly in a specific part of the kidney, the medulla, while subsequent to kidney injury the interaction was ubiquitous across the whole kidney, coherent with its ubiquitous Factor 1 loadings (**Supp. Fig. S9D**).

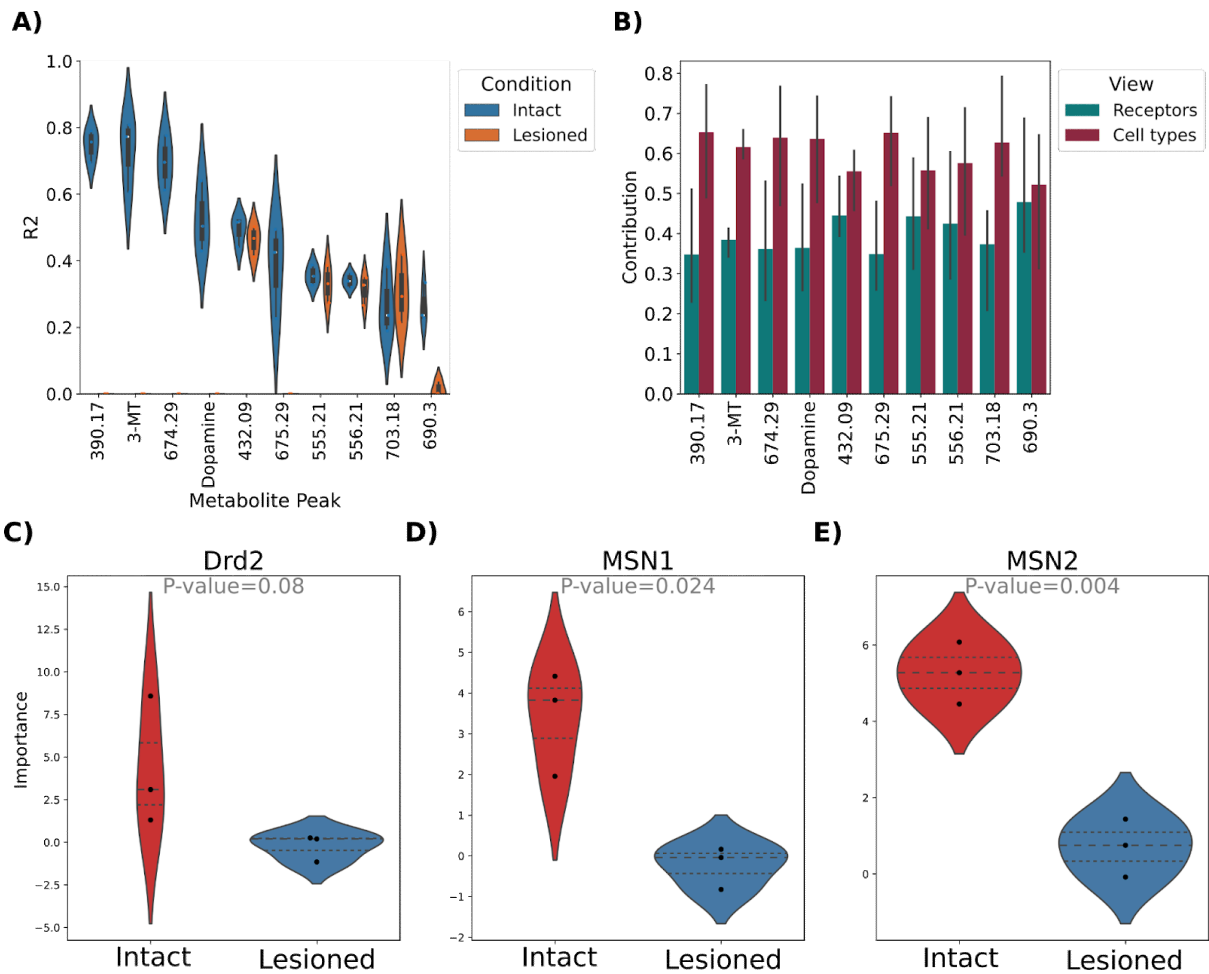
In summary, using LIANA+ I identified in a hypothesis-free manner intercellular programmes potentially involved in early response to AKI in dissociated single-cell data, and supported those using independent spatial transcriptomics samples.



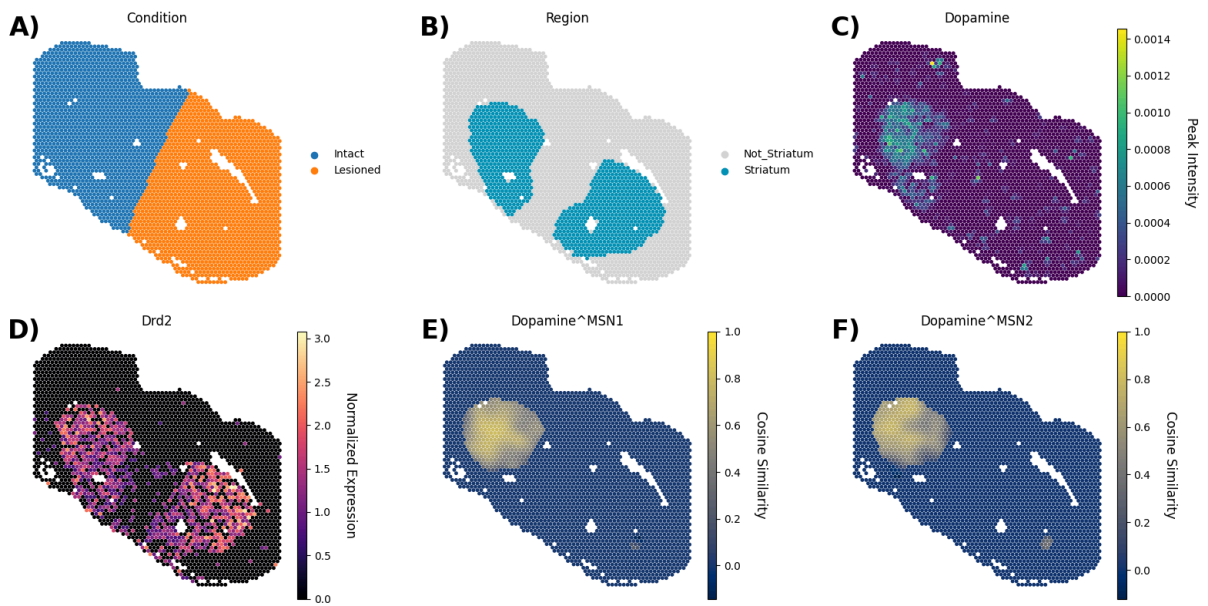
Supplementary Figure S1. LIANA+ accepts inputs as unimodal (AnnData) or multimodal (MuData) data objects together with optional prior knowledge resource and/or spatial information. These are then transformed into dataframes of aggregated interaction results or statistics at the individual spot- or cell-level.



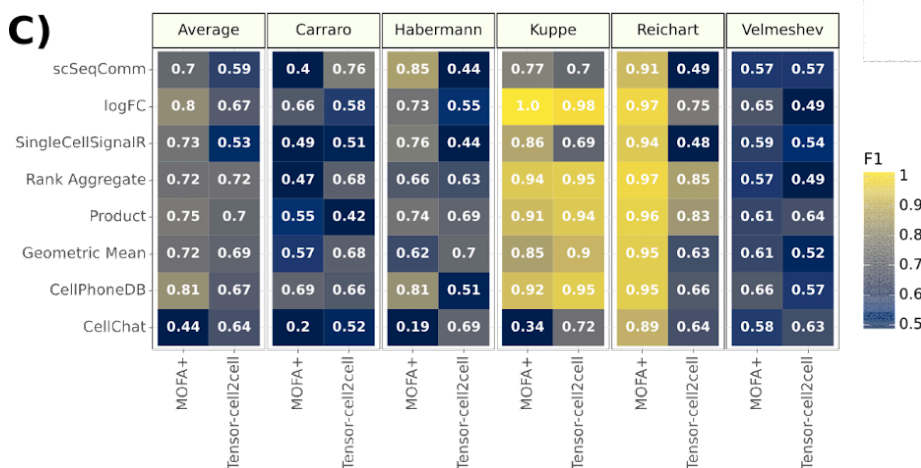
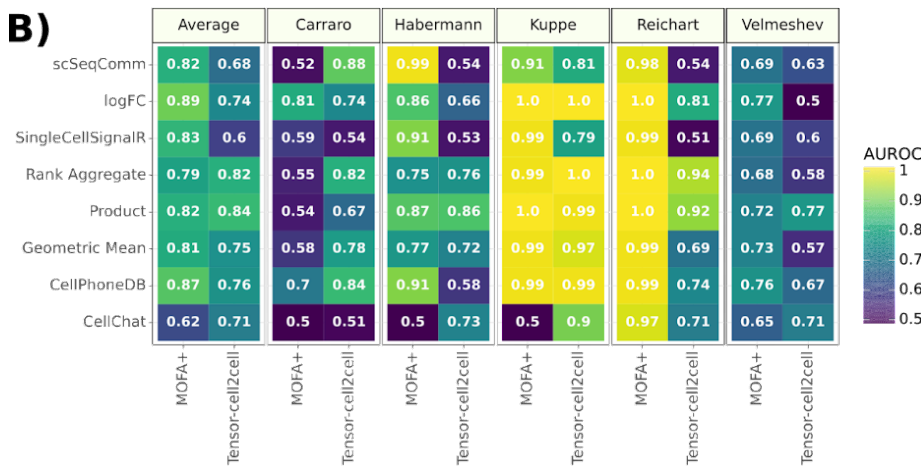
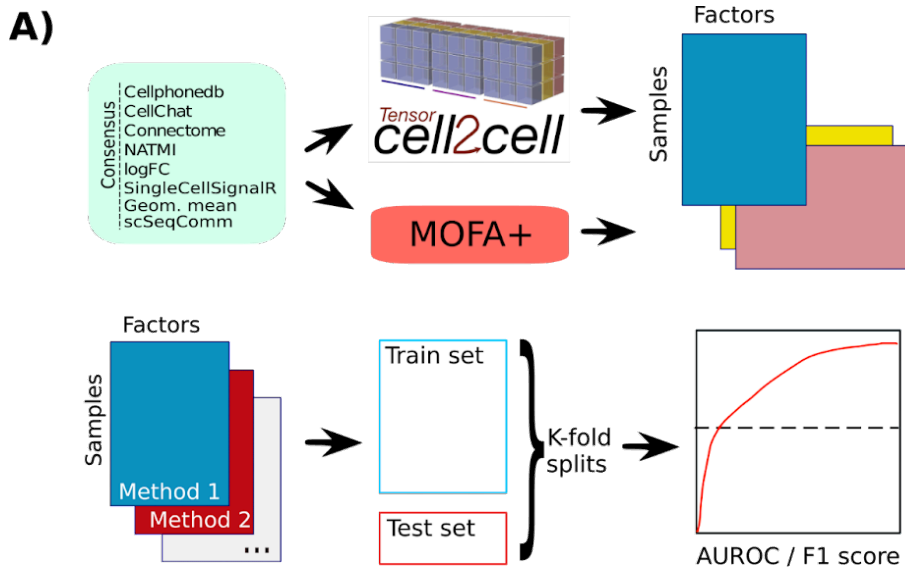
Supplementary Figure S2. **A)** AUROC and **B)** weighted F1 when using local metrics to classify malignant spots in breast cancer data; **C)** R-squared and **D)** RMSE when using local metrics to predict cell type proportions in heart data. The line in the boxplots represents the median with hinges showing the first and third quartiles and the whiskers extend up to 1.5 times the interquartile range above and below the box hinges.



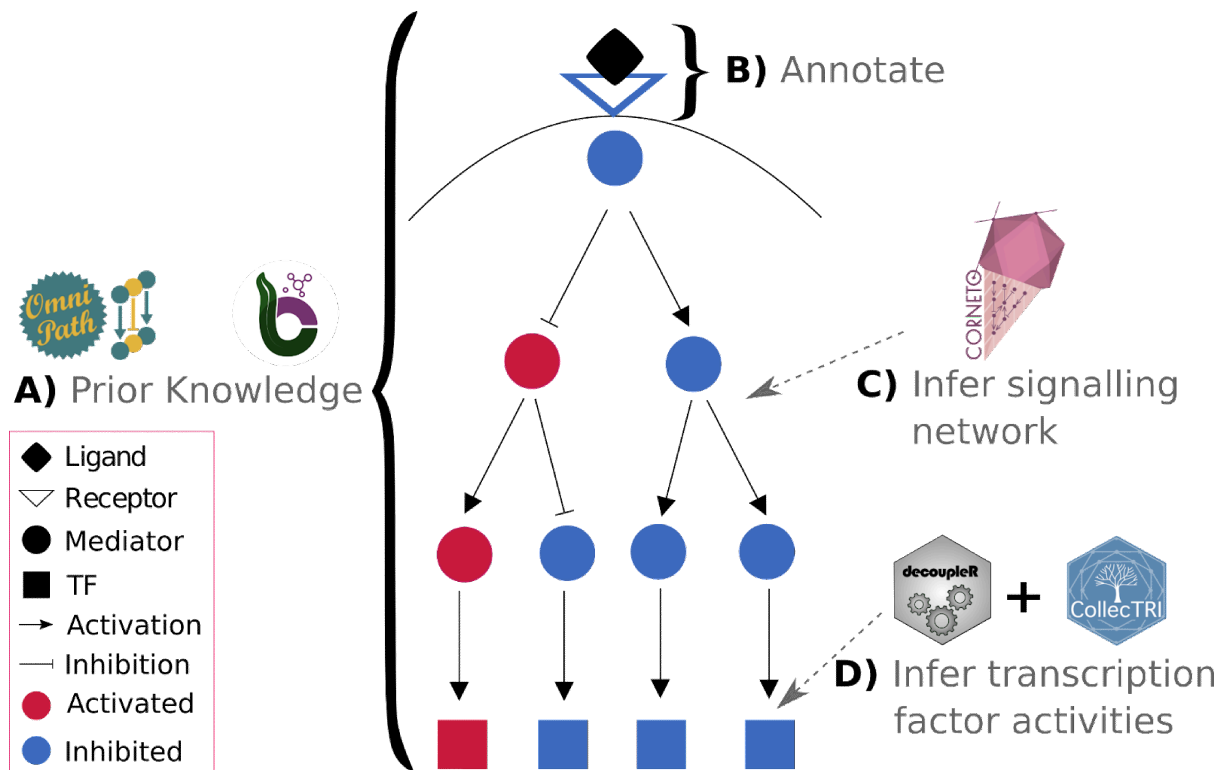
Supplementary Figure S3. A) Top 10 metabolite peaks with highest variance explained (R^2). **B)** Relative contributions of views (receptors and cell types) when jointly predicting metabolite peak intensities. **C-E)** Differences between lesioned and intact hemispheres in Dopamine's canonical Drd2 receptor, and medium spiny neuron (MSN) cell states 1 and 2. P-values were calculated with one-sided paired t-tests.



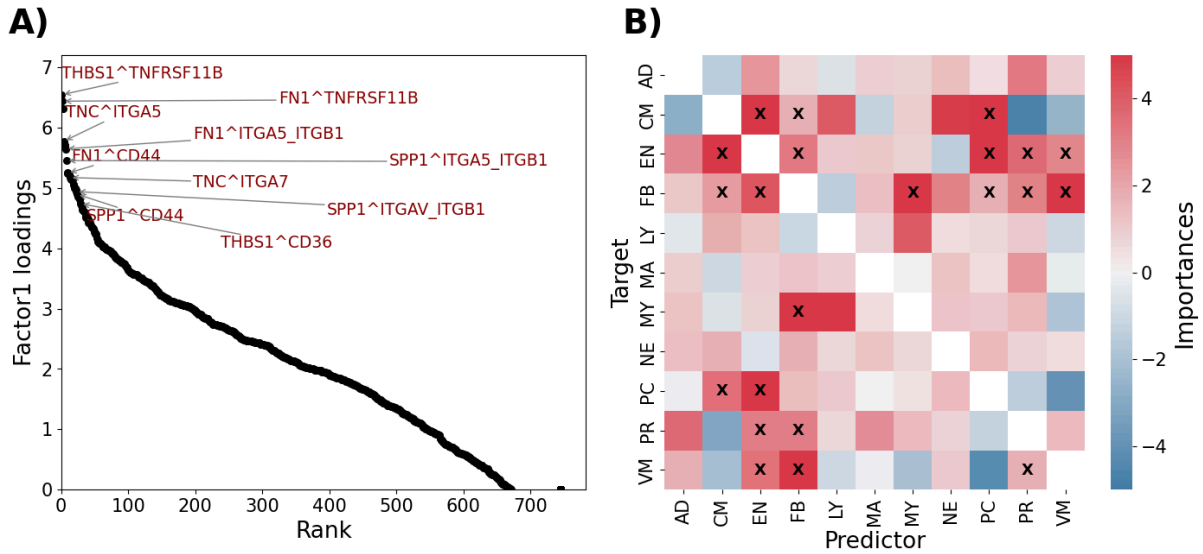
Supplementary Figure S4. **A)** Lesioned and Intact hemispheres; **B)** Striatum annotations; **C)** Dopamine Peak intensities; **D)** Drd2 receptor normalised expression levels. Interactions between Dopamine and Medium Spiny Neuron cell states **E)** 1 and **F)** 2. All panels show slide B1 from experiment V11L12-109.



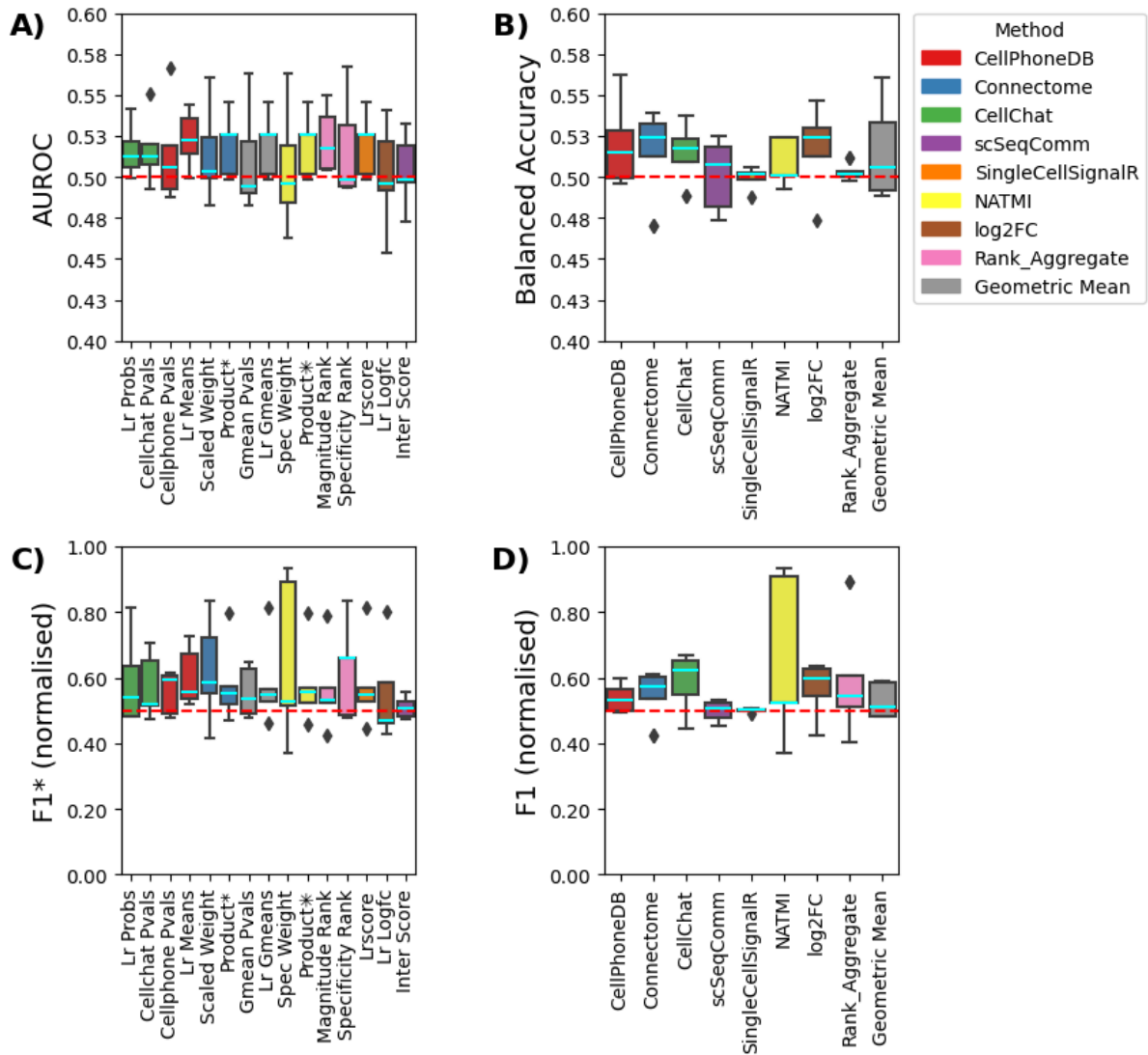
Supplementary Figure S5. A) Classification setup to evaluate the ability of ligand-receptor methods, combined with Tensor-cell2cell and MOFA+ to separate conditions from multi-condition atlases in an unsupervised manner. **B)** Average area under the receiver-operator curve (AUROC) and **C)** weighted F1 score.



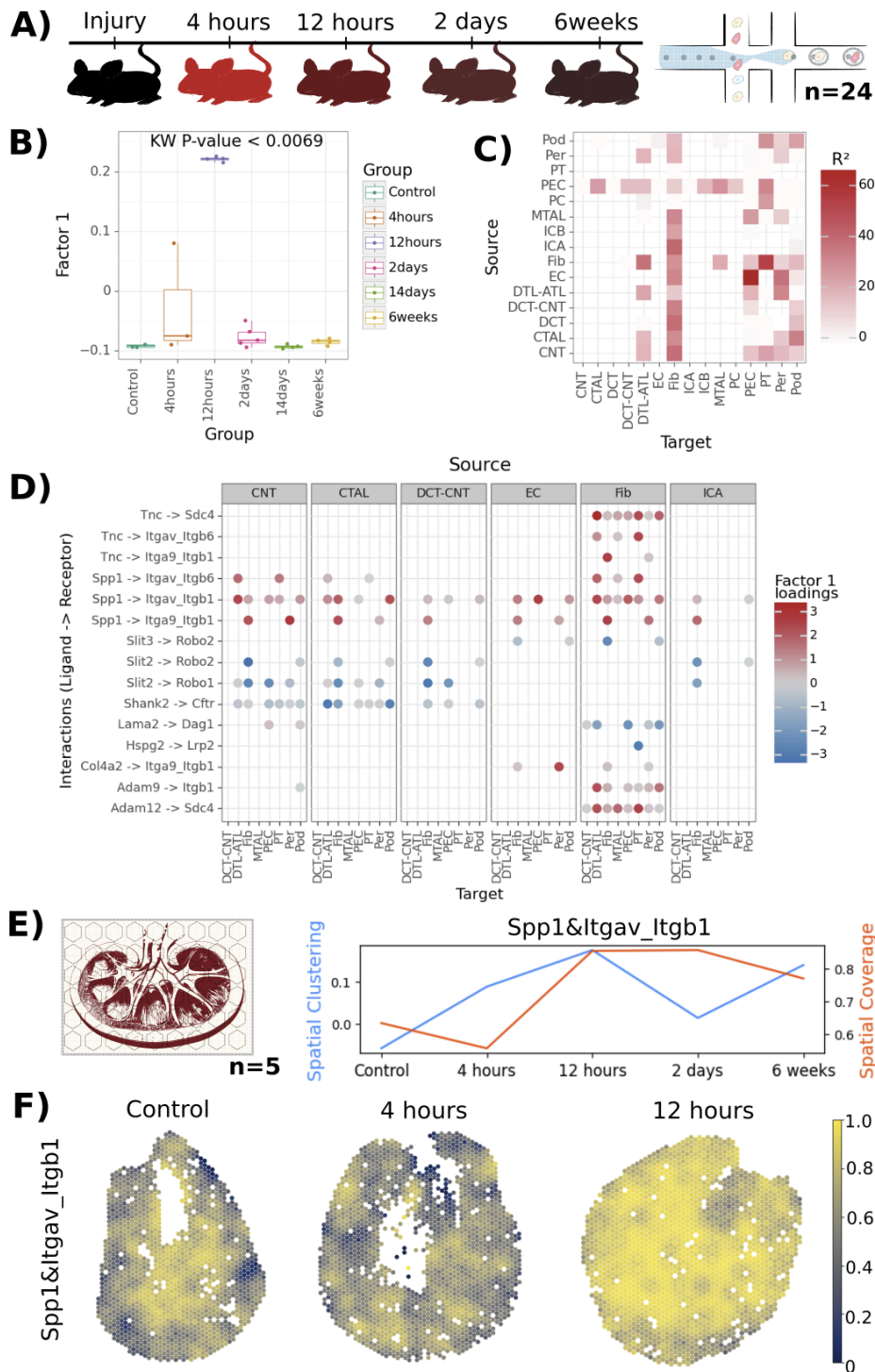
Supplementary Figure S6. *LIANA+* makes use of A) prior knowledge^{45,115} to B) annotate interactions according to e.g. pathways, disease, or location. Similarly, it uses this prior knowledge to infer putative causal (sign-coherent) signalling networks^{136,151}, emanating from ligand-receptor interactions down to active transcription factors. The transcription factor activities are themselves estimated making use of generalistic regulon prior knowledge¹⁶⁰ and enrichment analysis⁹⁸.



Supplementary Figure S7. A) Top 30 interactions (with "FN1", "TNC", "THBS1", and "SPP1" as ligands) in Factor 1 identified using NMF on local ligand-receptor metrics in spatially-resolved 10X visium heart samples. B) MISTy importances (Median t-value) modelling cell type interactions from 10X visium slides. Cell type interactions with Median t-value > 1.645 and R² > 5% are marked with X. Abbreviations used include AD for Adipocytes, CM for Cardiomyocytes, EN for Endothelial cells, FB for Fibroblasts, PC for Pericytes, PR for Proliferating cells, VM for Vascular smooth muscle cells, NE for Neuronal cells, MY for Myeloid cells, MA for Mast cells, and LY for Lymphoid cells.



Supplementary Figure S8. Comparison of ligand-receptor inference methods using the spatial colocalisation of cell types and ligand-receptors as assumed truth. Each method and its individual scoring functions are represented by a different colour. **A)** Quantifies the performance of each methods' individual scoring functions using the Area Under the Receiver Operating Characteristic Curve (AUROC). **B)** Balanced Accuracy for each method filtered according to its suggested false positive filtering thresholds. **C)** and **D)** measure normalised F1 score following filtering for each score and method, respectively. For each metric, a score of 0.5, denoted by the dashed red line, indicates random performance. Both NATMI and Connectome use expression products (Product*) as a measure of magnitude strength. The central line in cyan within each box marks the median, with the box hinges representing the first and third quartiles. The whiskers extend up to 1.5 times the interquartile range above and below the box hinges. Outliers are depicted as individual diamond-shaped points beyond the whiskers.



Supplementary Figure S9. A) Experimental design of a murine AKI model ^{163,164}. B) Distribution of Factor 1 sample scores at different time points following AKI. The line in the boxplots represents the median with hinges showing the first and third quartiles and whiskers extending 1.5 above and below the interquartile range. C) Variance explained by pairs of cell groups (views). D) Interaction loadings associated with Factor 1. E) Spatial Clustering (Global Moran's R) and Coverage (mean Cosine similarity) of *Spp1&Itgav_Itgb1* across conditions. F) Spatially-weighted Cosine similarity of *Spp1* and the *Itgav/Itgb1* complex in Control, 4 and 12 hours after injury.

Bibliography

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).
2. Almet, A. A., Cang, Z., Jin, S. & Nie, Q. The landscape of cell-cell communication through single-cell transcriptomics. *Current Opinion in Systems Biology* **26**, 12–23 (2021).
3. Ramirez Flores, R. O., Lanzer, J. D., Dimitrov, D., Velten, B. & Saez-Rodriguez, J. Multicellular factor analysis of single-cell data for a tissue-centric understanding of disease. *BioRxiv* (2023) doi:10.1101/2023.02.23.529642.
4. Efremova, M., Vento-Tormo, M., Teichmann, S. A. & Vento-Tormo, R. CellPhoneDB: inferring cell-cell communication from combined expression of multi-subunit ligand-receptor complexes. *Nat. Protoc.* **15**, 1484–1506 (2020).
5. Raredon, M. S. B. *et al.* Computation and visualization of cell-cell signaling topologies in single-cell systems data using Connectome. *Sci. Rep.* **12**, 4187 (2022).
6. Jin, S. *et al.* Inference and analysis of cell-cell communication using CellChat. *Nat. Commun.* **12**, 1088 (2021).
7. Browaeys, R., Saelens, W. & Saeys, Y. NicheNet: modeling intercellular communication by linking ligands to target genes. *Nat. Methods* **17**, 159–162 (2020).
8. Hu, Y., Peng, T., Gao, L. & Tan, K. CytoTalk: De novo construction of signal transduction networks using single-cell transcriptomic data. *Sci. Adv.* **7**, (2021).
9. Dimitrov, D. *et al.* Comparison of methods and resources for cell-cell communication inference from single-cell RNA-Seq data. *Nat. Commun.* **13**, 3224 (2022).
10. Zheng, R. *et al.* MEBOCOST: Metabolic Cell-Cell Communication Modeling by Single Cell Transcriptome. *BioRxiv* (2022) doi:10.1101/2022.05.30.494067.
11. Armingol, E., Larsen, R. O., Cequeira, M., Baghdassarian, H. & Lewis, N. E. Unraveling the coordinated dynamics of protein- and metabolite-mediated cell-cell

- communication. *BioRxiv* (2022) doi:10.1101/2022.11.02.514917.
12. Zhao, W., Johnston, K. G., Ren, H., Xu, X. & Nie, Q. Inferring neuron-neuron communications from single-cell transcriptomics through NeuronChat. *Nat. Commun.* **14**, 1128 (2023).
 13. Garcia-Alonso, L. *et al.* Single-cell roadmap of human gonadal development. *Nature* **607**, 540–547 (2022).
 14. Hrovatin, K., Fischer, D. S. & Theis, F. J. Toward modeling metabolic state from single-cell transcriptomics. *Mol. Metab.* **57**, 101396 (2022).
 15. He, C., Zhou, P. & Nie, Q. exFINDER: identify external communication signals using single-cell transcriptomics data. *Nucleic Acids Res.* **51**, e58 (2023).
 16. Wilk, A. J., Shalek, A. K., Holmes, S. & Blish, C. A. Comparative analysis of cell–cell communication at single-cell resolution. *Nat. Biotechnol.* (2023) doi:10.1038/s41587-023-01782-z.
 17. Raredon, M. S. B. *et al.* Comprehensive visualization of cell-cell interactions in single-cell and spatial transcriptomics with NICHES. *Bioinformatics* **39**, (2023).
 18. Oyler-Yaniv, A. *et al.* A Tunable Diffusion-Consumption Mechanism of Cytokine Propagation Enables Plasticity in Cell-to-Cell Communication in the Immune System. *Immunity* **46**, 609–620 (2017).
 19. Moffitt, J. R., Lundberg, E. & Heyn, H. The emerging landscape of spatial profiling technologies. *Nat. Rev. Genet.* **23**, 741–759 (2022).
 20. Vandereyken, K., Sifrim, A., Thienpont, B. & Voet, T. Methods and applications for single-cell and spatial multi-omics. *Nat. Rev. Genet.* **24**, 494–515 (2023).
 21. Chen, Z., Soifer, I., Hilton, H., Keren, L. & Jovic, V. Modeling Multiplexed Images with Spatial-LDA Reveals Novel Tissue Microenvironments. *J. Comput. Biol.* **27**, 1204–1218 (2020).
 22. Schapiro, D. *et al.* histoCAT: analysis of cell phenotypes and interactions in multiplex image cytometry data. *Nat. Methods* **14**, 873–876 (2017).
 23. Vickovic, S. *et al.* High-definition spatial transcriptomics for in situ tissue profiling. *Nat.*

- Methods* **16**, 987–990 (2019).
24. Pham, D. T. *et al.* stLearn: integrating spatial location, tissue morphology and gene expression to find cell types, cell-cell interactions and spatial trajectories within undissociated tissues. *BioRxiv* (2020) doi:10.1101/2020.05.31.125658.
 25. Dries, R. *et al.* Giotto: a toolbox for integrative analysis and visualization of spatial expression data. *Genome Biol.* **22**, 78 (2021).
 26. Shao, X. *et al.* Knowledge-graph-based cell-cell communication inference for spatially resolved transcriptomic data with SpaTalk. *Nat. Commun.* **13**, 4429 (2022).
 27. Fischer, D. S., Schaar, A. C. & Theis, F. J. Modeling intercellular communication in tissues using spatial graphs of cells. *Nat. Biotechnol.* **41**, 332–336 (2023).
 28. Kleshchevnikov, V. *et al.* Cell2location maps fine-grained cell types in spatial transcriptomics. *Nat. Biotechnol.* **40**, 661–671 (2022).
 29. Li, Z., Wang, T., Liu, P. & Huang, Y. SpatialDM for rapid identification of spatially co-expressed ligand-receptor and revealing cell-cell communication patterns. *Nat. Commun.* **14**, 3995 (2023).
 30. Cang, Z. *et al.* Screening cell-cell communication in spatial transcriptomics via collective optimal transport. *Nat. Methods* **20**, 218–228 (2023).
 31. Unable to find information for 13527485.
 32. Dimitrov, D. *et al.* LIANA+: an all-in-one cell-cell communication framework. *BioRxiv* (2023) doi:10.1101/2023.08.19.553863.
 33. Tanevski, J., Flores, R. O. R., Gabor, A., Schapiro, D. & Saez-Rodriguez, J. Explainable multiview framework for dissecting spatial relationships from highly multiplexed data. *Genome Biol.* **23**, 97 (2022).
 34. Mitchel, J. *et al.* Tensor decomposition reveals coordinated multicellular patterns of transcriptional variation that distinguish and stratify disease individuals. *BioRxiv* (2022) doi:10.1101/2022.02.16.480703.
 35. Jerby-Aron, L. & Regev, A. DIALOGUE maps multicellular programs in tissue from single-cell or spatial transcriptomics data. *Nat. Biotechnol.* **40**, 1467–1477 (2022).

36. Crowell, H. L. *et al.* muscat detects subpopulation-specific state transitions from multi-sample multi-condition single-cell transcriptomics data. *Nat. Commun.* **11**, 6077 (2020).
37. Browaeys, R. *et al.* MultiNicheNet: a flexible framework for differential cell-cell communication analysis from multi-sample multi-condition single-cell transcriptomics data. *BioRxiv* (2023) doi:10.1101/2023.06.13.544751.
38. Garcia-Alonso, L. *et al.* Mapping the temporal and spatial dynamics of the human endometrium in vivo and in vitro. *Nat. Genet.* **53**, 1698–1711 (2021).
39. Armingol, E. *et al.* Context-aware deconvolution of cell-cell communication with Tensor-cell2cell. *Nat. Commun.* **13**, 3665 (2022).
40. Kim, H. J., Lin, Y., Geddes, T. A., Yang, J. Y. H. & Yang, P. CiteFuse enables multi-modal analysis of CITE-seq data. *Bioinformatics* **36**, 4137–4143 (2020).
41. Zappia, L. & Theis, F. J. Over 1000 tools reveal trends in the single-cell RNA-seq analysis landscape. *Genome Biol.* **22**, 301 (2021).
42. Armingol, E., Baghdassarian, H. M. & Lewis, N. E. The diversification of methods for studying cell-cell interactions and communication. *Nat. Rev. Genet.* (2024) doi:10.1038/s41576-023-00685-8.
43. Dimitrov, D. *et al.* Comparison of Resources and Methods to infer Cell-Cell Communication from Single-cell RNA Data. *BioRxiv* (2021) doi:10.1101/2021.05.21.445160.
44. Chen, X., Teichmann, S. A. & Meyer, K. B. From Tissues to Cell Types and Back: Single-Cell Gene Expression Analysis of Tissue Architecture. *Annu. Rev. Biomed. Data Sci.* **1**, 29–51 (2018).
45. Türei, D. *et al.* Integrated intra- and intercellular signaling knowledge for multicellular omics analysis. *Mol. Syst. Biol.* **17**, (2021).
46. 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).

47. Cillo, A. R. *et al.* Immune Landscape of Viral- and Carcinogen-Driven Head and Neck Cancer. *Immunity* **52**, 183-199.e9 (2020).
48. Wang, Y. *et al.* iTALK: an R Package to Characterize and Illustrate Intercellular Communication. *BioRxiv* (2019) doi:10.1101/507871.
49. Tyler, S. R. *et al.* PyMINEr Finds Gene and Autocrine-Paracrine Networks from Human Islet scRNA-Seq. *Cell Rep.* **26**, 1951-1964.e8 (2019).
50. Tsuyuzaki, K., Ishii, M. & Nikaido, I. Uncovering hypergraphs of cell-cell interaction from single cell RNA-sequencing data. *BioRxiv* (2019) doi:10.1101/566182.
51. Unable to find information for 10345576.
52. Hou, R., Denisenko, E., Ong, H. T., Ramilowski, J. A. & Forrest, A. R. R. Predicting cell-to-cell communication networks using NATMI. *Nat. Commun.* **11**, 5011 (2020).
53. Cabello-Aguilar, S. *et al.* SingleCellSignalR: inference of intercellular networks from single-cell transcriptomics. *Nucleic Acids Res.* **48**, e55 (2020).
54. Noël, F. *et al.* ICELLNET: a transcriptome-based framework to dissect intercellular communication. *BioRxiv* (2020) doi:10.1101/2020.03.05.976878.
55. Choi, H. *et al.* Transcriptome analysis of individual stromal cell populations identifies stroma-tumor crosstalk in mouse lung cancer model. *Cell Rep.* **10**, 1187–1201 (2015).
56. Wang, S., Karikomi, M., MacLean, A. L. & Nie, Q. Cell lineage and communication network inference via optimization for single-cell transcriptomics. *Nucleic Acids Res.* **47**, e66 (2019).
57. Cheng, J., Zhang, J., Wu, Z. & Sun, X. Inferring microenvironmental regulation of gene expression from single-cell RNA sequencing data using scMLnet with an application to COVID-19. *Brief. Bioinformatics* **22**, 988–1005 (2021).
58. Shao, X. *et al.* CellTalkDB: a manually curated database of ligand-receptor interactions in humans and mice. *Brief. Bioinformatics* **22**, (2021).
59. Noël, F. *et al.* Dissection of intercellular communication using the transcriptome-based framework ICELLNET. *Nat. Commun.* **12**, 1089 (2021).
60. Baccin, C. *et al.* Combined single-cell and spatial transcriptomics reveal the molecular,

- cellular and spatial bone marrow niche organization. *Nat. Cell Biol.* **22**, 38–48 (2020).
61. Palla, G. *et al.* Squidpy: a scalable framework for spatial single cell analysis. *BioRxiv* (2021) doi:10.1101/2021.02.19.431994.
 62. Garcia-Alonso, L. *et al.* Mapping the temporal and spatial dynamics of the human endometrium in vivo and in vitro. *BioRxiv* (2021) doi:10.1101/2021.01.02.425073.
 63. Vento-Tormo, R. *et al.* Single-cell reconstruction of the early maternal-fetal interface in humans. *Nature* **563**, 347–353 (2018).
 64. Kanehisa, M. & Goto, S. KEGG: Kyoto encyclopedia of genes and genomes. *Nucleic Acids Res.* **28**, 27–30 (2000).
 65. Kanehisa, M., Furumichi, M., Sato, Y., Ishiguro-Watanabe, M. & Tanabe, M. KEGG: integrating viruses and cellular organisms. *Nucleic Acids Res.* **49**, D545–D551 (2021).
 66. Fabregat, A. *et al.* The Reactome Pathway Knowledgebase. *Nucleic Acids Res.* **46**, D649–D655 (2018).
 67. Szklarczyk, D. *et al.* The STRING database in 2017: quality-controlled protein-protein association networks, made broadly accessible. *Nucleic Acids Res.* **45**, D362–D368 (2017).
 68. Harding, S. D. *et al.* The IUPHAR/BPS Guide to PHARMACOLOGY in 2018: updates and expansion to encompass the new guide to IMMUNOPHARMACOLOGY. *Nucleic Acids Res.* **46**, D1091–D1106 (2018).
 69. Ramilowski, J. A. *et al.* A draft network of ligand-receptor-mediated multicellular signalling in human. *Nat. Commun.* **6**, 7866 (2015).
 70. Licata, L. *et al.* SIGNOR 2.0, the SIGNaling Network Open Resource 2.0: 2019 update. *Nucleic Acids Res.* **48**, D504–D510 (2020).
 71. Winograd-Katz, S. E., Fässler, R., Geiger, B. & Legate, K. R. The integrin adhesome: from genes and proteins to human disease. *Nat. Rev. Mol. Cell Biol.* **15**, 273–288 (2014).
 72. Fazekas, D. *et al.* SignalLink 2 - a signaling pathway resource with multi-layered regulatory networks. *BMC Syst. Biol.* **7**, 7 (2013).

73. Wang, Y. talkr uncovers ligand-receptor mediated intercellular crosstalk. *BioRxiv* (2020) doi:10.1101/2020.02.01.930602.
74. Ben-Shlomo, I., Yu Hsu, S., Rauch, R., Kowalski, H. W. & Hsueh, A. J. W. Signaling receptome: a genomic and evolutionary perspective of plasma membrane receptors involved in signal transduction. *Sci. STKE* **2003**, RE9 (2003).
75. Kirouac, D. C. *et al.* Dynamic interaction networks in a hierarchically organized tissue. *Mol. Syst. Biol.* **6**, 417 (2010).
76. Kandasamy, K. *et al.* NetPath: a public resource of curated signal transduction pathways. *Genome Biol.* **11**, R3 (2010).
77. Lee, H.-O. *et al.* Lineage-dependent gene expression programs influence the immune landscape of colorectal cancer. *Nat. Genet.* **52**, 594–603 (2020).
78. Zheng, Y. *et al.* Single-Cell Transcriptomics Reveal Immune Mechanisms of the Onset and Progression of IgA Nephropathy. *Cell Rep.* **33**, 108525 (2020).
79. Deczkowska, A. *et al.* XCR1+ type 1 conventional dendritic cells drive liver pathology in non-alcoholic steatohepatitis. *Nat. Med.* **27**, 1043–1054 (2021).
80. Weber, L. M. *et al.* Essential guidelines for computational method benchmarking. *Genome Biol.* **20**, 125 (2019).
81. Mahdessian, D. *et al.* Spatiotemporal dissection of the cell cycle with single-cell proteogenomics. *Nature* **590**, 649–654 (2021).
82. Gul, L. *et al.* Extracellular vesicles produced by the human commensal gut bacterium *Bacteroides thetaiotaomicron* affect host immune pathways in a cell-type specific manner that are altered in inflammatory bowel disease. *BioRxiv* (2021) doi:10.1101/2021.03.20.436262.
83. Westermann, A. J. & Vogel, J. Cross-species RNA-seq for deciphering host-microbe interactions. *Nat. Rev. Genet.* **22**, 361–378 (2021).
84. Kolde, R. pheatmap: Pretty Heatmaps. <https://CRAN.R-project.org/package=pheatmap> (2019).
85. Satija, R., Farrell, J. A., Gennert, D., Schier, A. F. & Regev, A. Spatial reconstruction of

- single-cell gene expression data. *Nat. Biotechnol.* **33**, 495–502 (2015).
86. Butler, A., Hoffman, P., Smibert, P., Papalexi, E. & Satija, R. Integrating single-cell transcriptomic data across different conditions, technologies, and species. *Nat. Biotechnol.* **36**, 411–420 (2018).
 87. Stoeckius, M. *et al.* Simultaneous epitope and transcriptome measurement in single cells. *Nat. Methods* **14**, 865–868 (2017).
 88. Iorio, F. *et al.* Efficient randomization of biological networks while preserving functional characterization of individual nodes. *BMC Bioinformatics* **17**, 542 (2016).
 89. Ushey, K., Allaire, J. J. & Tang, Y. Interface to “Python” [R package reticulate version 1.20]. <https://cran.r-project.org/web/packages/reticulate/index.html> (2020).
 90. Wolf, F. A., Angerer, P. & Theis, F. J. SCANPY: large-scale single-cell gene expression data analysis. *Genome Biol.* **19**, 15 (2018).
 91. Sheikh, B. N. *et al.* Systematic Identification of Cell-Cell Communication Networks in the Developing Brain. *iScience* **21**, 273–287 (2019).
 92. Wang, S. *et al.* A systematic evaluation of the computational tools for ligand-receptor-based cell-cell interaction inference. *Brief. Funct. Genomics* **21**, 339–356 (2022).
 93. Liu, Z., Sun, D. & Wang, C. Evaluation of cell-cell interaction methods by integrating single-cell RNA sequencing data with spatial information. *Genome Biol.* **23**, 218 (2022).
 94. Jiang, P. *et al.* Systematic investigation of cytokine signaling activity at the tissue and single-cell levels. *Nat. Methods* **18**, 1181–1191 (2021).
 95. Wu, S. Z. *et al.* A single-cell and spatially resolved atlas of human breast cancers. *Nature Genetics* (2021).
 96. Tasic, B. *et al.* Adult mouse cortical cell taxonomy revealed by single cell transcriptomics. *Nat. Neurosci.* **19**, 335–346 (2016).
 97. Elosua-Bayes, M., Nieto, P., Mereu, E., Gut, I. & Heyn, H. SPOTlight: seeded NMF regression to deconvolute spatial transcriptomics spots with single-cell transcriptomes.

- Nucleic Acids Res.* **49**, e50 (2021).
98. Badia-I-Mompel, P. *et al.* decoupleR: ensemble of computational methods to infer biological activities from omics data. *Bioinformatics Advances* **2**, vbac016 (2022).
 99. Kuhn, M. & Vaughan, D. *yardstick: Tidy Characterizations of Model Performance*. (CRAN, 2021).
 100. Gayoso, A. *et al.* Joint probabilistic modeling of single-cell multi-omic data with totalVI. *Nat. Methods* **18**, 272–282 (2021).
 101. Jakobsson, J. E. T., Spjuth, O. & Lagerström, M. C. scConnect: a method for exploratory analysis of cell-cell communication based on single-cell RNA-sequencing data. *Bioinformatics* **37**, 3501–3508 (2021).
 102. Baruzzo, G., Cesaro, G. & Di Camillo, B. Identify, quantify and characterize cellular communication from single-cell RNA sequencing data with scSeqComm. *Bioinformatics* **38**, 1920–1929 (2022).
 103. Rao, N. *et al.* Charting spatial ligand-target activity using Renoir. *BioRxiv* (2023) doi:10.1101/2023.04.14.536833.
 104. Pham, D. *et al.* Robust mapping of spatiotemporal trajectories and cell-cell interactions in healthy and diseased tissues. *Nat. Commun.* **14**, 7739 (2023).
 105. Baysoy, A., Bai, Z., Satija, R. & Fan, R. The technological landscape and applications of single-cell multi-omics. *Nat. Rev. Mol. Cell Biol.* **24**, 695–713 (2023).
 106. Lagger, C. *et al.* scAgeCom: a murine atlas of age-related changes in intercellular communication inferred with the package scDiffCom. *BioRxiv* (2021) doi:10.1101/2021.08.13.456238.
 107. Nagai, J. S., Leimkühler, N. B., Schaub, M. T., Schneider, R. K. & Costa, I. G. CrossTalker: analysis and visualization of ligand-receptor networks. *Bioinformatics* **37**, 4263–4265 (2021).
 108. Zhang, Y. *et al.* Cellinker: a platform of ligand-receptor interactions for intercellular communication analysis. *Bioinformatics* (2021) doi:10.1093/bioinformatics/btab036.
 109. Zhang, Y. *et al.* CellCall: integrating paired ligand-receptor and transcription factor

- activities for cell-cell communication. *Nucleic Acids Res.* **49**, 8520–8534 (2021).
110. Virshup, I. *et al.* The scverse project provides a computational ecosystem for single-cell omics data analysis. *Nat. Biotechnol.* **41**, 604–606 (2023).
 111. Ghazanfar, S. *et al.* Investigating higher-order interactions in single-cell data with scHOT. *Nat. Methods* **17**, 799–806 (2020).
 112. Muzellec, B., Telenczuk, M., Cabeli, V. & Andreux, M. PyDESeq2: a python package for bulk RNA-seq differential expression analysis. *BioRxiv* (2022) doi:10.1101/2022.12.14.520412.
 113. Love, M. I., Huber, W. & Anders, S. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol.* **15**, 550 (2014).
 114. Argelaguet, R. *et al.* MOFA+: a statistical framework for comprehensive integration of multi-modal single-cell data. *Genome Biol.* **21**, 111 (2020).
 115. Lobentanzer, S. *et al.* Democratizing knowledge representation with BioCypher. *Nat. Biotechnol.* **41**, 1056–1059 (2023).
 116. Masarapu, Y. *et al.* Spatially resolved multiomics on the neuronal effects induced by spaceflight. *Res. Sq.* (2023) doi:10.21203/rs.3.rs-2865086/v1.
 117. Kuppe, C. *et al.* Spatial multi-omic map of human myocardial infarction. *Nature* **608**, 766–777 (2022).
 118. Vicari, M. *et al.* Spatial multimodal analysis of transcriptomes and metabolomes in tissues. *Nat. Biotechnol.* (2023) doi:10.1038/s41587-023-01937-y.
 119. Zeisel, A. *et al.* Molecular architecture of the mouse nervous system. *Cell* **174**, 999-1014.e22 (2018).
 120. Anselin, L. A local indicator of multivariate spatial association: extending geary's c. *Geogr. Anal.* **51**, 133–150 (2019).
 121. Baghdassarian, H., Dimitrov, D., Armingol, E., Saez-Rodriguez, J. & Lewis, N. E. Combining LIANA and Tensor-cell2cell to decipher cell-cell communication across multiple samples. *BioRxiv* (2023) doi:10.1101/2023.04.28.538731.
 122. Heumos, L. *et al.* Best practices for single-cell analysis across modalities. *Nat. Rev.*

- Genet.* **24**, 550–572 (2023).
123. Zimmerman, K. D., Espeland, M. A. & Langefeld, C. D. A practical solution to pseudoreplication bias in single-cell studies. *Nat. Commun.* **12**, 738 (2021).
 124. Schubert, M. *et al.* Perturbation-response genes reveal signaling footprints in cancer gene expression. *Nat. Commun.* **9**, 20 (2018).
 125. Wynn, T. A. Cellular and molecular mechanisms of fibrosis. *J. Pathol.* **214**, 199–210 (2008).
 126. Ismahil, M. A. *et al.* Remodeling of the mononuclear phagocyte network underlies chronic inflammation and disease progression in heart failure: critical importance of the cardiosplenic axis. *Circ. Res.* **114**, 266–282 (2014).
 127. Hoefft, K. *et al.* Platelet-instructed SPP1⁺ macrophages drive myofibroblast activation in fibrosis in a CXCL4-dependent manner. *Cell Rep.* **42**, 112131 (2023).
 128. Imanaka-Yoshida, K., Tawara, I. & Yoshida, T. Tenascin-C in cardiac disease: a sophisticated controller of inflammation, repair, and fibrosis. *Am J Physiol, Cell Physiol* **319**, C781–C796 (2020).
 129. Wang, W., Ren, X., Chen, X., Hong, Q. & Cai, G. Integrin β 1-rich extracellular vesicles of kidney recruit Fn1⁺ macrophages to aggravate ischemia-reperfusion-induced inflammation. *JCI Insight* (2024).
 130. Russell, A. J. C. *et al.* Slide-tags enables single-nucleus barcoding for multimodal spatial genomics. *Nature* **625**, 101–109 (2024).
 131. Schmierer, B. & Hill, C. S. TGFbeta-SMAD signal transduction: molecular specificity and functional flexibility. *Nat. Rev. Mol. Cell Biol.* **8**, 970–982 (2007).
 132. Karadeniz, Z. *et al.* Novel effects of Foxo3a on immune system and cardiac remodeling after myocardial infarction. *Eur. Heart J.* **41**, (2020).
 133. Velten, B. *et al.* Identifying temporal and spatial patterns of variation from multimodal data using MEFISTO. *Nat. Methods* **19**, 179–186 (2022).
 134. Townes, F. W. & Engelhardt, B. E. Nonnegative spatial factorization applied to spatial genomics. *Nat. Methods* **20**, 229–238 (2023).

135. Turos, D., Vasiljevic, J., Hahn, K., Rottenberg, S. & Valdeolivas, A. Chrysalis: decoding tissue compartments in spatial transcriptomics with archetypal analysis. *BioRxiv* (2023) doi:10.1101/2023.08.17.553606.
136. Liu, A. *et al.* From expression footprints to causal pathways: contextualizing large signaling networks with CARNIVAL. *NPJ Syst. Biol. Appl.* **5**, 40 (2019).
137. Garrido-Rodriguez, M., Zirngibl, K., Ivanova, O., Lobentanzer, S. & Saez-Rodriguez, J. Integrating knowledge and omics to decipher mechanisms via large-scale models of signaling networks. *Mol. Syst. Biol.* **18**, e11036 (2022).
138. Farr, E. *et al.* MetalinksDB: a resource for metabolite-mediated cell-cell communication.
139. Udani, S. *et al.* Secretion encoded single-cell sequencing (SEC-seq) uncovers gene expression signatures associated with high VEGF-A secretion in mesenchymal stromal cells. *BioRxiv* (2023) doi:10.1101/2023.01.07.523110.
140. Wheeler, M. A. *et al.* Droplet-based forward genetic screening of astrocyte-microglia cross-talk. *Science* **379**, 1023–1030 (2023).
141. Zhang, S. *et al.* Monitoring of cell-cell communication and contact history in mammals. *Science* **378**, eabo5503 (2022).
142. Galeano Niño, J. L. *et al.* Effect of the intratumoral microbiota on spatial and cellular heterogeneity in cancer. *Nature* **611**, 810–817 (2022).
143. Saarenpää, S. *et al.* Spatially resolved host-bacteria-fungi interactomes via spatial metatranscriptomics. *BioRxiv* (2022) doi:10.1101/2022.07.18.496977.
144. Lötstedt, B., Stražar, M., Xavier, R. J., Regev, A. & Vickovic, S. Spatial host-microbiome sequencing. *BioRxiv* (2022) doi:10.1101/2022.07.18.500470.
145. Anselin, L., Syabri, I. & Kho, Y. Geoda: an introduction to spatial data analysis. in *Handbook of applied spatial analysis* (eds. Fischer, M. M. & Getis, A.) 73–89 (Springer Berlin Heidelberg, 2010). doi:10.1007/978-3-642-03647-7_5.
146. Bredikhin, D., Kats, I. & Stegle, O. MUON: multimodal omics analysis framework. *Genome Biol.* **23**, 42 (2022).

147. Pedregosa, F. *et al.* Scikit-learn: Machine Learning in Python. *Journal of Machine Learning Research* (2011).
148. Virtanen, P. *et al.* SciPy 1.0: fundamental algorithms for scientific computing in Python. *Nat. Methods* **17**, 261–272 (2020).
149. Palla, G. *et al.* Squidpy: a scalable framework for spatial omics analysis. *Nat. Methods* **19**, 171–178 (2022).
150. Virshup, I., Rybakov, S., Theis, F. J., Angerer, P. & Wolf, F. A. anndata: Annotated data. *BioRxiv* (2021) doi:10.1101/2021.12.16.473007.
151. saezlab/corneto: CORNETO: an optimization library for modeling biological network inference problems. <https://github.com/saezlab/corneto>.
152. Dugourd, A. *et al.* Causal integration of multi-omics data with prior knowledge to generate mechanistic hypotheses. *Mol. Syst. Biol.* **17**, e9730 (2021).
153. Satopaa, V., Albrecht, J., Irwin, D. & Raghavan, B. Finding a “kneedle” in a haystack: detecting knee points in system behavior. in *2011 31st International Conference on Distributed Computing Systems Workshops* 166–171 (IEEE, 2011). doi:10.1109/ICDCSW.2011.20.
154. biocypher/liana: Knowledge graph for cell-cell interactions using BioCypher. <https://github.com/biocypher/liana>.
155. Carraro, G. *et al.* Transcriptional analysis of cystic fibrosis airways at single-cell resolution reveals altered epithelial cell states and composition. *Nat. Med.* **27**, 806–814 (2021).
156. Reichart, D. *et al.* Pathogenic variants damage cell composition and single cell transcription in cardiomyopathies. *Science* **377**, eabo1984 (2022).
157. Biancalani, T. *et al.* Deep learning and alignment of spatially resolved single-cell transcriptomes with Tangram. *Nat. Methods* **18**, 1352–1362 (2021).
158. Farr, E. B. *et al.* MetalinksDB: a flexible and contextualizable resource of metabolite-protein interactions. *BioRxiv* (2023) doi:10.1101/2023.12.30.573715.
159. Kang, H. M. *et al.* Multiplexed droplet single-cell RNA-sequencing using natural

- genetic variation. *Nat. Biotechnol.* **36**, 89–94 (2018).
160. Müller-Dott, S. *et al.* Expanding the coverage of regulons from high-confidence prior knowledge for accurate estimation of transcription factor activities. *Nucleic Acids Res.* **51**, 10934–10949 (2023).
 161. Gurobi Optimization, LLC. Gurobi Optimizer Reference Manual. (2023).
 162. Shannon, P. *et al.* Cytoscape: a software environment for integrated models of biomolecular interaction networks. *Genome Res.* **13**, 2498–2504 (2003).
 163. Kirita, Y., Wu, H., Uchimura, K., Wilson, P. C. & Humphreys, B. D. Cell profiling of mouse acute kidney injury reveals conserved cellular responses to injury. *Proc Natl Acad Sci USA* **117**, 15874–15883 (2020).
 164. Dixon, E. E., Wu, H., Muto, Y., Wilson, P. C. & Humphreys, B. D. Spatially resolved transcriptomic analysis of acute kidney injury in a female murine model. *J. Am. Soc. Nephrol.* **33**, 279–289 (2022).
 165. Kolde, R., Laur, S., Adler, P. & Vilo, J. Robust rank aggregation for gene list integration and meta-analysis. *Bioinformatics* **28**, 573–580 (2012).
 166. Habermann, A. C. *et al.* Single-cell RNA sequencing reveals profibrotic roles of distinct epithelial and mesenchymal lineages in pulmonary fibrosis. *Sci. Adv.* **6**, eaba1972 (2020).
 167. Velmeshev, D. *et al.* Single-cell genomics identifies cell type-specific molecular changes in autism. *Science* **364**, 685–689 (2019).
 168. Dai, Q., Epstein, M. P. & Yang, J. STACCato: Supervised Tensor Analysis tool for studying Cell-cell Communication using scRNA-seq data across multiple samples and conditions. *BioRxiv* (2023) doi:10.1101/2023.12.15.571918.
 169. Luo, J., Deng, M., Zhang, X. & Sun, X. ESICCC as a systematic computational framework for evaluation, selection, and integration of cell-cell communication inference methods. *Genome Res.* (2023) doi:10.1101/gr.278001.123.
 170. Li, H., Dixon, E. E., Wu, H. & Humphreys, B. D. Comprehensive single-cell transcriptional profiling defines shared and unique epithelial injury responses during

- kidney fibrosis. *Cell Metab.* **34**, 1977-1998.e9 (2022).
171. Deng, L.-C., Alinejad, T., Bellusci, S. & Zhang, J.-S. Fibroblast Growth Factors in the Management of Acute Kidney Injury Following Ischemia-Reperfusion. *Front. Pharmacol.* **11**, 426 (2020).
172. Liaw, L. *et al.* Altered wound healing in mice lacking a functional osteopontin gene (spp1). *J. Clin. Invest.* **101**, 1468–1478 (1998).
173. Midwood, K. S., Hussenet, T., Langlois, B. & Orend, G. Advances in tenascin-C biology. *Cell. Mol. Life Sci.* **68**, 3175–3199 (2011).