

Comparison of methods and resources for cell-cell communication inference from single-cell RNA-Seq data

Daniel Dimitrov¹, Dénes Türei¹, Martin Garrido-Rodriguez ¹, Paul L. Burmedi ¹, James S. Nagai^{2,3}, Charlotte Boys¹, Ricardo O. Ramirez Flores ¹, Hyojin Kim¹, Bence Szalai⁴, Ivan G. Costa ^{2,3}, Alberto Valdeolivas ^{5,6}, Aurélien Dugourd^{1,6} & Julio Saez-Rodriguez ¹ 

The growing availability of single-cell data, especially transcriptomics, has sparked an increased interest in the inference of cell-cell communication. Many computational tools were developed for this purpose. Each of them consists of a resource of intercellular interactions prior knowledge and a method to predict potential cell-cell communication events. Yet the impact of the choice of resource and method on the resulting predictions is largely unknown. To shed light on this, we systematically compare 16 cell-cell communication inference resources and 7 methods, plus the consensus between the methods' predictions. Among the resources, we find few unique interactions, a varying degree of overlap, and an uneven coverage of specific pathways and tissue-enriched proteins. We then examine all possible combinations of methods and resources and show that both strongly influence the predicted intercellular interactions. Finally, we assess the agreement of cell-cell communication methods with spatial colocalisation, cytokine activities, and receptor protein abundance and find that predictions are generally coherent with those data modalities. To facilitate the use of the methods and resources described in this work, we provide LIANA, a Ligand-receptor ANalysis frAmework as an open-source interface to all the resources and methods.

¹Heidelberg University, Faculty of Medicine, and Heidelberg University Hospital, Institute for Computational Biomedicine, BioQuant, Heidelberg, Germany. ²Institute for Computational Genomics, Faculty of Medicine, RWTH Aachen University, Aachen 52074, Germany. ³Joint Research Center for Computational Biomedicine, RWTH Aachen University Hospital, Aachen, Germany. ⁴Faculty of Medicine, Department of Physiology, Semmelweis University, Budapest, Hungary. ⁵Roche Pharma Research and Early Development, Pharmaceutical Sciences, Roche Innovation Center Basel, Basel, Switzerland. ⁶These authors contributed equally: Alberto Valdeolivas, Aurélien Dugourd. ✉email: pub.saez@uni-heidelberg.de

Single-cell RNA sequencing (scRNA-Seq) data has become a driving force in the analysis of the cellular heterogeneity of tissues. Furthermore, Spatial Transcriptomics has recently emerged as a technology to measure gene expression while preserving the spatial distribution of cells in a sample, thus providing an unprecedented opportunity to decipher tissue architecture¹. These advancements have in turn led to an increased interest in the development of tools for cell-cell communication (CCC) inference. CCC events are essential for homeostasis, development, and disease, and their estimation is becoming a routine approach in scRNA-seq data analysis². CCC commonly 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 junctions³. For simplicity, we refer to all of these events involving protein-protein interactions as CCC.

A number of computational tools and resources have emerged that can be further classified as those that predict CCC interactions alone^{4–17}, and those that additionally estimate intracellular activities related to CCC^{18–24}. Here, we focus on the former (Table 1). These CCC tools typically use gene expression information obtained by scRNA-Seq. In general, single cells are clustered by their gene expression profile and cell type identities are assigned to the clusters based on known gene markers. Then, CCC tools can 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 a transmitter and receiver protein, accordingly expressed by the source and target cell clusters. The information about which transmitter binds to which receiver is extracted from diverse sources of prior knowledge. Roughly, CCC tools then estimate the likelihood of crosstalk based on the

expression level of the transmitter and the receiver in the source and target clusters, respectively. 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, regularisations, and scaling, to prioritise interactions according to the input datasets (Table 1). In turn, these different approaches result in diverse scoring systems that are challenging to compare and evaluate. The difficulties are further exacerbated by the lack of an appropriate gold standard to benchmark the performance of CCC methods^{2,25}. Nevertheless, different strategies have been used to indirectly evaluate the methods' performance, including a presumed correlation between CCC predictions and spatial adjacency^{14,22}, recovering the effect of receptor gene knockouts²², robustness to subsampling¹⁴, agreement with proteomics¹², simulated scRNA-Seq data⁹, and the agreement among methods^{10,12,14,22}.

The available prior knowledge resources, largely composed of ligand-receptor, extracellular matrix, and adhesion interactions, are typically distinct but often show partial overlap^{3,26}. Some of these resources also provide additional details for the interactions such as information about subcellular localisation^{3,14}, classification into signalling pathways and categories^{14,27} (Supplementary Table 1). Notably, some resources^{3,8,14,27,28} (Supplementary Table 1), and consequently their corresponding methods, focus on protein complexes as the functional units of CCC, which are crucial for the coordination of signalling as different subunit combinations may induce distinct responses⁸. Despite the fact that CCC inference is constrained by the prior knowledge used,

Table 1 Tools included in the framework.

Tool/Method	Resource	Methods' scoring systems
CellChat ^{#14}	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
CellPhoneDBv2 ^{#8}	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 ¹⁰	Ramilowski	(1) weight_norm —derived via the product (by default) of the normalised expression of transmitter and receiver genes (2) weight_scale †—derived from a function (mean, by default) of the z-scores of the transmitter and the receiver, scaled according to cell cluster specificity
Crosstalk scores	-	(1) Crosstalk score †—Cytotalk-inspired ²² crosstalk scores were derived from the expression of transmitters and receivers, weighted by the likelihood of autocrine signalling between the source and target cell types.
logFC Mean	-	(1) logFC Mean †—iTALK-inspired ⁶ logFC means, derived using the mean of the logged one-versus-all fold change of receiver and transmitter gene expression
NATMI ¹¹	ConnectomeDB	(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 ^{#12}	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.
Consensus	-	(1) Robust Rank Aggregate ⁶⁵ —preferentially highly-ranked interactions are obtained from a distribution generated from the interaction rankings of other methods

Each method considers expression at the cell cluster level, and all of the scoring systems presented here use the expression of transmitters and receiver genes in the source and target cells, respectively. In addition to the seven methods, we included their consensus.

In bold are the names of cell-cell communication inference methods and their scoring functions.

Dagger (†): Explicitly incorporates communicating cell-pair specificity in interaction predictions

Hashtag (#): CellPhoneDB, CellChat, and SingleCellSignalR provide explicit thresholds to control for false positive interaction predictions. In the case of the former two, these are permutation-based p-values, whereas SingleCellSignalR's LRscore has a suggested threshold of 0.5.

Methods that additionally infer intracellular processes, such as NicheNet⁹, Cytotalk²², and SoptSC²⁰ are not directly comparable but instead provide complementary analyses.

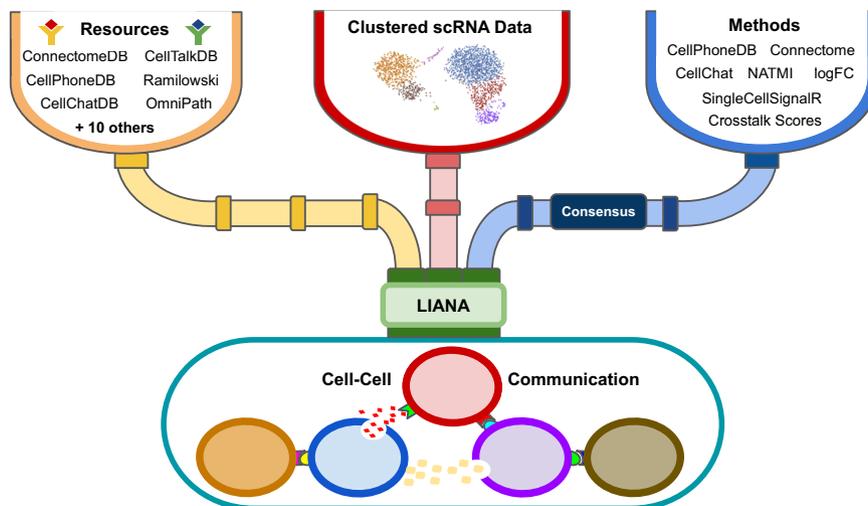


Fig. 1 LIANA—a Ligand-receptor ANalysis frAamework. LIANA takes any annotated single-cell RNA (scRNA) dataset as input and establishes a common interface to all the resources and methods in any combination. LIANA also provides a consensus ranking for the method’s predictions.

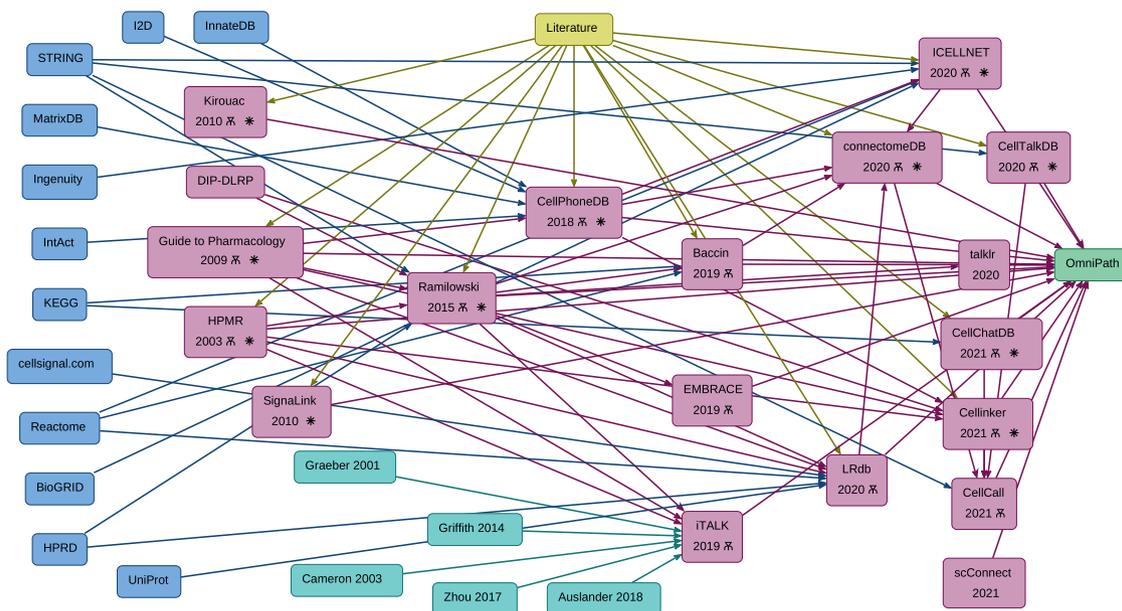


Fig. 2 Dependencies and overlap between CCC resources. The lineages of CCC interaction database knowledge. General biological knowledge databases (blue), CCC-dedicated resources (magenta), manual literature curation effort (yellow), additional resources included in iTALK (cyan), and OmniPath (green). Arrows show the data transfers between resources. The yus symbol (⌘) indicates the manual-curation of resources, defined by explicitly mentioning that these resources are ‘manually’ or ‘expert’ curated. The asterisk (*) indicates that the resource was included in the analyses presented here.

yet the impact of resource choice is largely unexplored, with the exception of a descriptive comparison of 4 resources with one method²⁶. Thus, it remains unclear how the choice of resource and method affects the results and thereby the biological interpretation of the scRNA-seq data.

In this work, we systematically compared all combinations of 16 resources and 7 CCC methods, plus their consensus (Fig. 1). First, we explored the degree of overlap among resources and whether certain resources are biased toward specific biological terms, such as pathways and tissue-enriched proteins. Then, we analysed how different combinations of resources and methods influence CCC inference by decoupling the methods from their corresponding resources and applying all method-resource combinations on six different datasets. Finally we evaluated the agreement of the different CCC methods with additional

modalities, including spatial adjacency, cytokine activities, and protein abundance. All results were generated using LIANA—a LIgand-receptor ANalysis frAamework (Fig. 1; available at <https://github.com/saezlab/liana>).

Results

Resource uniqueness and overlap. To investigate the lineages of CCC resources, we manually gathered information about the origins of every resource. Many of these resources share the same original data sources, including general biological databases such as KEGG^{29,30}, Reactome³¹, and STRING³² (Fig. 2). Moreover, interactions from the Guide to Pharmacology³³, CellPhoneDB⁸, HMPR³⁴, and in particular Ramilowski (FANTOM5)³⁵, which are manually curated, were commonly incorporated into subsequently published resources (Fig. 2; Supplementary Table 2). All

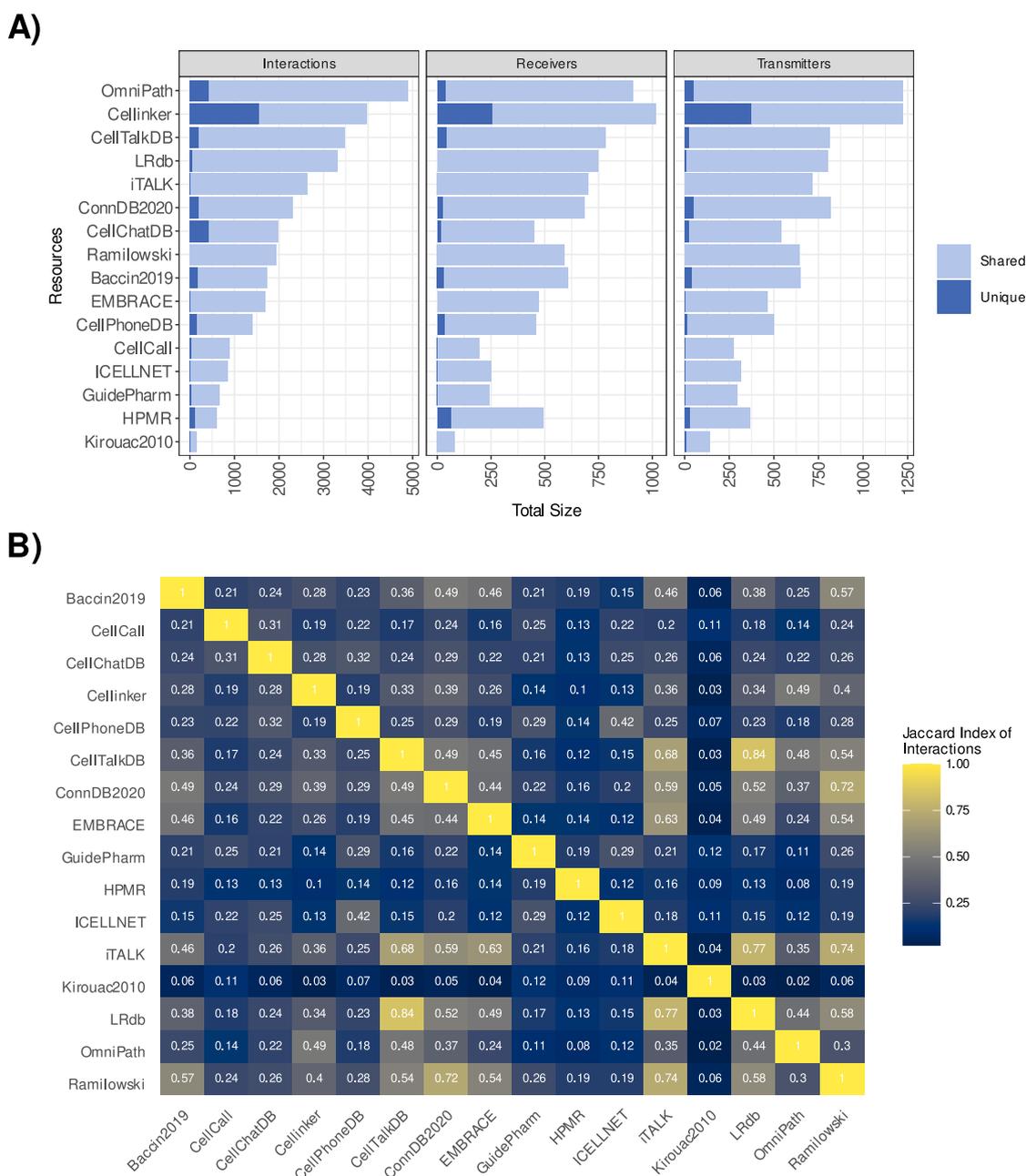


Fig. 3 Cell-cell communication resources—uniqueness and overlap. A Shared and unique Interactions, Receivers and Transmitters for each resource. **B** Similarity between the different resources based on the interactions (Jaccard Index). Source data are provided as a Source Data file.

the resources included in this analysis are integrated into OmniPath’s CCC resource³, along with additional CCC interactions from other sources (e.g. SIGNOR³⁶, Adhesive³⁷, SignalLink³⁸). A part of the OmniPath CCC resource, also referred to as ‘OmniPath’ and used in this work, was filtered by curation and protein localisation quality (“Processing of CCC resources” Methods).

As a consequence of their common origins, we noted limited uniqueness across the resources, with mean percentages of 6.4% unique receivers, 5.7% unique transmitters, and 10.4% unique interactions (Fig. 3A; Supplementary Table 1). One notable exception was Cellinker’s resource¹⁶, as 39.3% of its interactions were not present in any other resource. Despite the fact that few components were unique to any given resource, the pairwise overlap between the resources varied and was often limited (Fig. 3B; Supplementary Fig. S1). Yet, high similarity was

observed between CellTalkDB²⁶, ConnectomeDB¹¹, iTALK⁶, LRdb¹², and Ramilowski (Fig. 3B). Each of these resources, together with OmniPath and Cellinker, contained an average of at least 60% of the interactions present in other resources, largely explained by each containing a large proportion (>80%) of the interactions present in Ramilowski (Supplementary Fig. S2). Baccin²⁸, CellPhoneDB, CellChatDB, and EMBRACE showed limited similarity with other resources, as each included on average ~40–50% of the interactions present in any other resource. These latter resources, except EMBRACE, include protein complexes, which were dissociated and treated as distinct protein subunits in our resource analyses. The relatively smaller resources CellCall²³, ICELNET¹³, Guide to Pharmacology, HMPR and Kirouac2010³⁹ were the most dissimilar from the remainder. Finally, the similarity among the resources was generally higher when considering transmitters and receivers

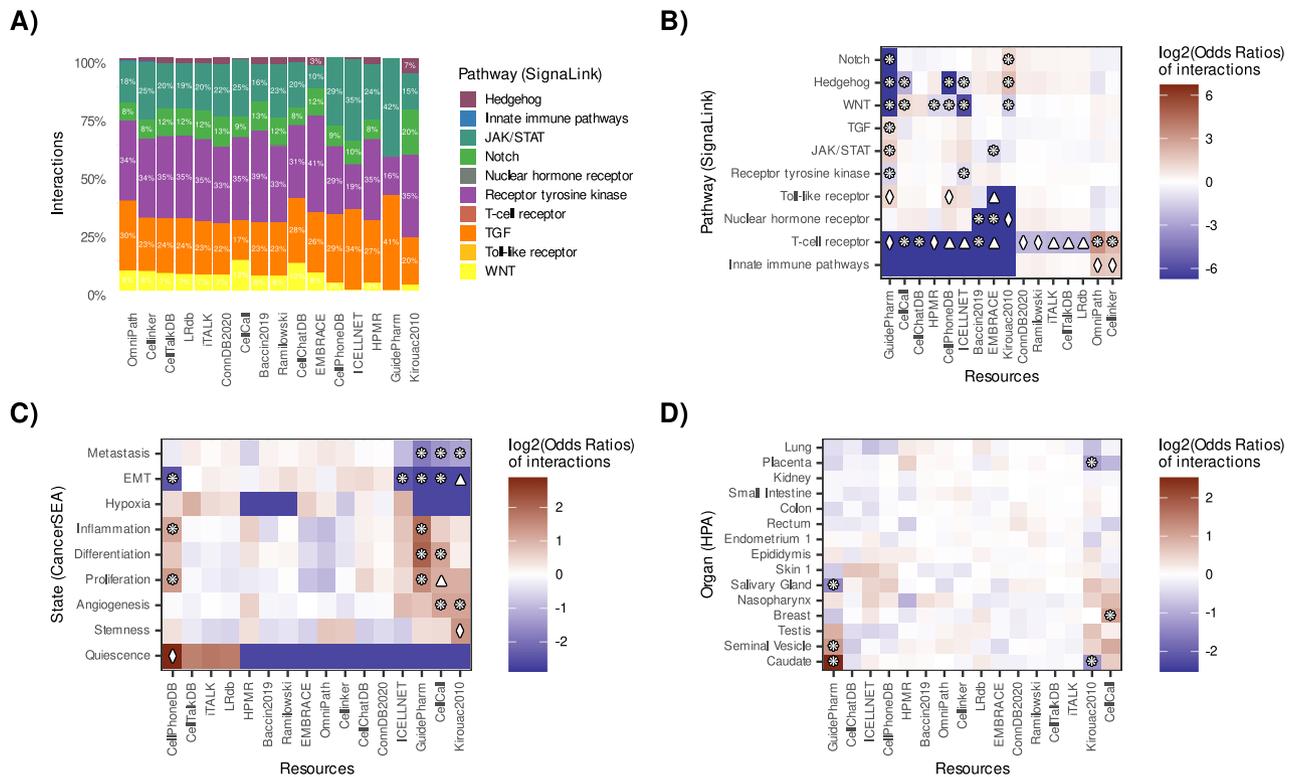


Fig. 4 Representation of functional categories in CCC resources. CCC resources distributions in terms of number of interactions (A) and relative abundance (B) matched to the SignalLink database. Relative abundance of interactions categorised by (C) CancerSEA’s cancer-related gene sets, and (D) organ-enriched proteins from the Human Protein Atlas (HPA). Fisher’s exact test was used to estimate the differentially-represented categories. Differentially represented (absolute(log₂(Odds ratio)) >1) categories were marked according to FDR-corrected *p*-values = < 0.05 (diamond, ◇), 0.01 (triangle, △), and 0.001 (8-pointed asterisk; *). Source data are provided as a Source Data file.

(Supplementary Figs. S1, 2), rather than the interaction themselves, suggesting that different resources account for different interactions between the same proteins.

Resource prior knowledge bias. Since CCC inference relies heavily on prior knowledge to estimate intercellular communication events, the choice of resource and any potential bias in it is expected to impact the results. We therefore explored whether the coverage of each CCC resource, when compared to the collection of all resources, is biased toward specific functional categories, tissue-enriched proteins, disease-associated genes, or subcellular locations.

To examine whether specific pathways and biological functions are unevenly represented in CCC resources, we matched the interactions, receivers and transmitters from each resource to well-known pathways and functional categories from SignalLink³⁸, NetPath⁴⁰, and CancerSEA⁴¹ (“Descriptive analysis of resources” Methods) and compared the resulting distributions across 16 CCC-dedicated resources (Supplementary Table 2).

The Receptor tyrosine kinase (RTK), JAK/STAT, TGF, WNT, and Notch pathways covered the largest proportions of interactions matched to SignalLink (Fig. 4A), with analogous results observed for receivers and transmitters (Supplementary Fig. S3). The interactions from Ramilowski, ConnectomeDB, CellTalkDB, LRdb, and iTALK showed a highly similar patterns, explained by the high overlap of these resources, with all of them showing significant underrepresentation of the T cell receptor pathway (Fig. 4B). A more pronounced underrepresentation of the same pathway was observed in Guide to Pharmacology, ICELLNET, CellPhoneDB, CellCall, CellChatDB, HMPR, Baccin2019, EMBRACE, and Kirouac2010. On the contrary, the T-cell

receptor pathway was significantly overrepresented in OmniPath and Cellinker. When we used NetPath instead of SignalLink to define the T-cell receptor pathway, we also observed underrepresentation in HMPR, CellCall, EMBRACE, and Kirouac2010 and overrepresentation in OmniPath (Supplementary Fig. S4A). Moreover, the Signalink WNT pathway was underrepresented in Guide to Pharmacology, ICELLNET CellPhoneDB, HMPR, and Kirouac2010, and on the contrary overrepresented in CellCall. We saw similar results when using NetPath’s WNT pathway (Supplementary Fig. S4A). We also observed uneven representations across the resources, in particular for the Hedgehog, Notch, and Innate Immune pathways (Fig. 4A; Supplementary Fig. S4A).

We then matched interactions to cancer-related gene sets from CancerSEA⁴¹, which were also unevenly represented. For example, interactions from the CellPhoneDB resource were overrepresented in gene sets associated with inflammation, proliferation, and quiescence (Fig. 4C; Supplementary Fig. S5). Gene sets associated with epithelial-mesenchymal transition were underrepresented in CellPhoneDB, Guide to Pharmacology, CellCall, ICELLNET, and Kirouac2010. This observation was further supported by the underrepresentation of direct-contact signalling in the latter two resources (See Supplementary Note 1; Supplementary Fig. S6).

We also examined the coverage of tissue-enriched proteins and disease markers from the Human Protein Atlas⁴² and DisGeNet⁴³, respectively. Organ-enriched proteins were largely uniformly distributed across the CCC resources, with some exceptions, such as organ-associated proteins from the Breast, Bone Marrow, Lymph Nodes, and the Hypothalamus (Fig. 4D; Supplementary Figs. S7, S8). Similarly, tissue-enriched proteins were generally distributed evenly across most CCC resources, with some exceptions including the underrepresentation of interactions associated with cardiomyocyte proteins in

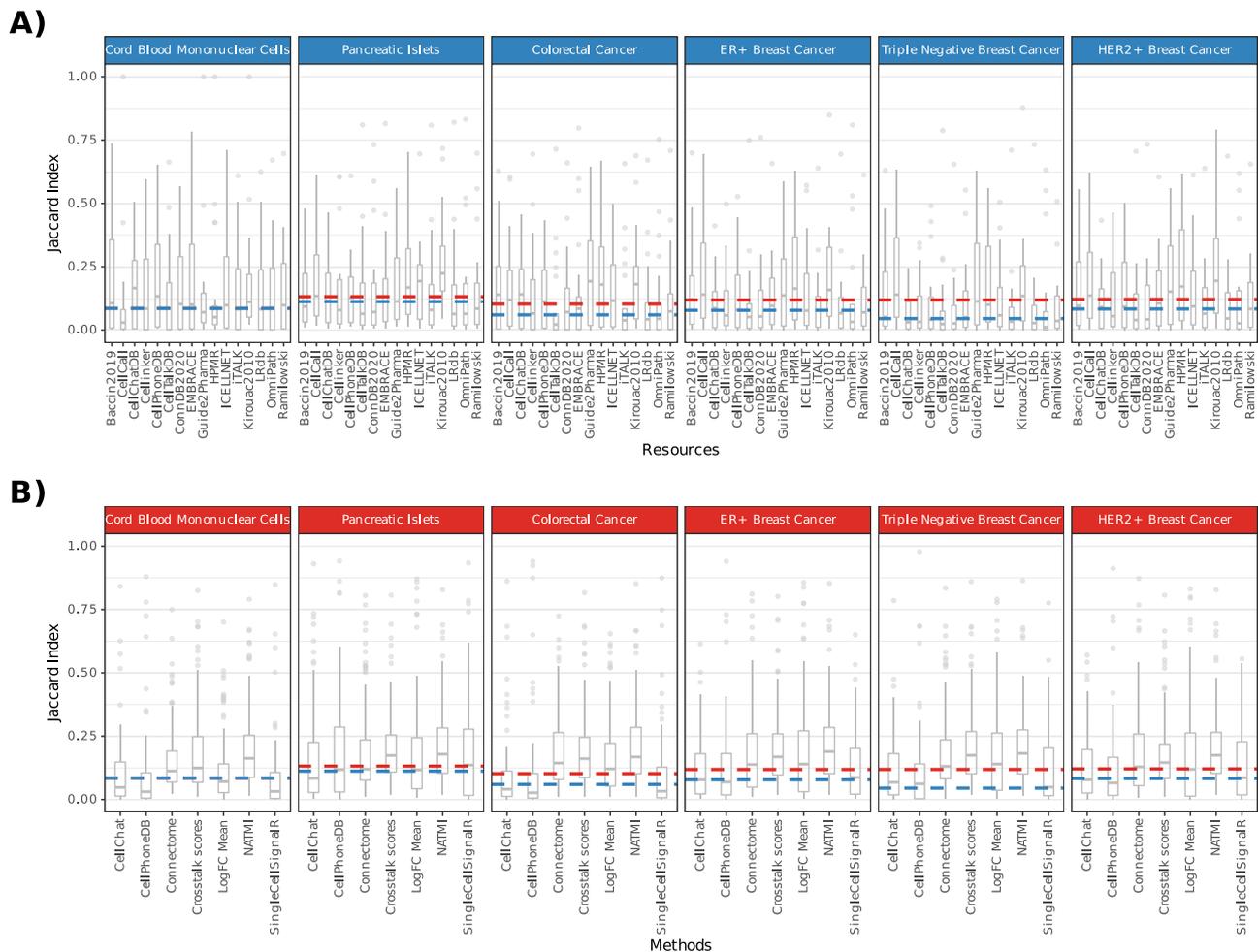


Fig. 5 Overlap of predictions using any combination of CCC methods and resources. Overlap (Jaccard index) in the 1000 highest ranked (**A**) when using the same Resource with different Methods (Blue; $n = 7$) and (**B**) when using the same Method with different Resources (Red; $n = 16$). Boxplots represent the median pairwise jaccard index with hinges showing the first and third quartiles and whiskers extending 1.5 above and below the interquartile range. The dashed lines represent the median when using different resources (red) and methods (blue); the lines overlap for the CMBCs dataset. Source data are provided as a Source Data file.

ICELNET and Kirouac2010, as well as the overrepresentation of proteins associated with Glial cells in Guide2Pharma (Supplementary Figs. S9, S10).

Finally, no differentially represented disease markers were noted in any of the CCC resources (Supplementary Fig. S11).

In summary, our results indicated biases towards certain pathways, functional categories, and tissue-enriched proteins across the different CCC resources, implying that resource choice can influence the functional interpretation of CCC predictions.

Using LIANA to systematically compare CCC predictions. To estimate the relative agreement between CCC methods and the importance of the resources, we built LIANA—a framework to decouple tools from their inbuilt resources. LIANA enabled us to combine the 16 CCC resources detailed in the descriptive resource analysis above (Supplementary Table 2), with 7 CCC methods used to prioritise ligand-receptor interactions from scRNA-Seq data (Table 1). We then predicted the interactions from all possible method-resource combinations for 6 single-cell RNA datasets from three different subtypes of breast cancer⁴⁴, cord blood mononucleated cells⁴⁵, Pancreatic Islets⁴⁶, and colorectal cancer⁴⁷ (Methods “Data availability”).

We first looked at the overlap between the 1000 highest ranked interactions predicted for every method-resource combination.

Whenever available, we used the recommended scoring functions (Supplementary Table 3), each tailored for predicting relevant interactions. We found consistently low overlap in the top predicted interactions when using either different methods or different resources (Fig. 5). The median pairwise Jaccard index when using different methods ranged from 0.045 to 0.112 across datasets (median = 0.080) (Fig. 5A). The overlap when using different resources was slightly higher, as the median pairwise Jaccard index ranged from 0.085 to 0.132 (median = 0.119) (Fig. 5B). We found similar results when considering the top 1% predicted interactions instead of the top 1000 (Supplementary Fig. S12; Supplementary Note 2). These analyses revealed substantial discrepancies in the highest-ranked predicted interactions by the different methods under study.

These discrepancies reflect the diverse nature of the scoring systems used to prioritise interactions of interest, and in particular, the different approaches used to assign communication cell cluster pair specificity to the interactions (marked with a dagger (†) in Table 1; used by all methods except SingleCellSignalR). The low overlap between the results of the different methods was also reflected by dissimilarities in the relative importances assigned to different cell types (See Supplementary Note 2).

On the other hand, the low overlap between the highest ranking interactions using different resources was largely

expected due to the limited overlap between the CCC resources as described in “Resource Uniqueness and Overlap”.

Taken together, our results suggest that both the choice of method and the resource can have a considerable impact on the predicted interactions.

Robustness to noise in resources and data. We then analysed the sensitivity of the methods to the addition of noise in the data and resource (“Robustness analyses” Methods). We found that most were fairly robust to subsampling of the total number of cells (Supplementary Fig. S13A), while erroneous annotation of cell types had a stronger effect, highlighting the importance of preprocessing and proper cluster annotation (Supplementary Fig. S13B). The methods were also adequately robust to the selective replacement of original canonical resources interactions with spurious putatively false interactions (“Robustness analyses” Methods), in which highest ranked interactions for each method were preserved (Supplementary Fig. S13C). The non-selective replacement of interactions, meant to simulate the change of resource (Supplementary Fig. S13D), had a strong effect on all methods, reflecting the low overlap when using different resources observed in the overlap analysis above.

Overall, our analysis showed that all methods, especially CellChat, CellPhoneDB, and SingleCellSignalR, were fairly robust to noise in both the data and the resource.

Association between CCC predictions and cytokine expression signatures. Next, given the lack of a ground truth, we used other data modalities to indirectly evaluate the methods using Omni-Path, the resource with the largest coverage.

First, we noted that all methods appropriately detect specifically-expressed receptor proteins across seven CITE-seq datasets (See Supplementary Note 3). Since protein levels of receptors do not necessarily imply activity, we evaluated the methods’ agreement with predicted cytokine activities using 43 cytokine expression signatures⁴⁸ on two datasets coming from two subtypes of breast cancer⁴⁴ (Methods “Agreement with cytokine signatures”). To show the association between CCC predictions and cytokine activities, we calculated the odds ratios between preferentially ranked interactions and positively enriched cytokines across a range of ranks. We found generally positive trends between cytokine activities and the most prioritised CCC interactions across all methods. The observed trends largely converged toward the random baseline as the number of considered interactions increased (Fig. 6A). Connectome, the Crosstalk scores, and NATMI showed a consistent trend across both datasets, while SingleCellSignalR, logFC Mean, CellChat, CellPhoneDB, and the consensus of the methods (Table 1) showed negative or lack of signal for the higher ranks of the HER2 + dataset (Fig. 6A; Supplementary Fig. S14). Notably, a high agreement with cytokine activities was observed for CellChat and CellPhoneDB in the HER2 + dataset, when considering all of their predictions subsequent to false-positive filtering (vertical line in Fig. 6A), highlighting the value of the false-positive control steps of these methods.

These results suggest that the interactions identified as relevant by all methods were largely concordant with cytokine activities, confirming the agreement of predicted CCC interactions with downstream signalling events.

Enrichment of predicted interactions between spatially adjacent cell types. Next, we leveraged spatial information as a way to support the methods’ predictive potential, under the assumption that, while many other factors are involved, colocalized cell populations are expected to have a higher chance to interact with

each other than other non-adjacent cell types^{14,22,49,50}. That is, the highest ranked interactions predicted between various cell populations are expected to be positively associated in interactions between pairs of adjacent cell types (Methods “Agreement with spatially adjacent cell types”).

We used the spatial mapping information from eight 10× Visium slides (see Methods), corresponding to a murine brain cortex⁵¹ and triple negative breast cancer⁴⁴ datasets, to identify the colocalized cell types in the tissues. We observed a positive trend of increased colocalisation of cell types in Visium and prioritisation of CCC interactions in the scRNA datasets (Fig. 6B). This trend was particularly consistent for the well-structured, murine brain cortex dataset, where all methods, except the Crosstalk scores, showed an association between cell type spatial adjacency and CCC predictions, with Connectome, LogFC Mean, and the consensus displaying the most positive associations. In the case of the triple negative breast cancer dataset, only the predictions by the consensus and LogFC Mean showed a consistent, positive association with spatial adjacency (Fig. 6B).

We conducted a similar analysis with seqFISH⁵² and merFISH⁵³ datasets (“Agreement with spatially adjacent cell types” Methods). In this case, we made use of the single-cell resolutions of these datasets to identify both the spatially adjacent cell types and to obtain the interaction predictions. For the seqFISH dataset, we found a clear association between the predicted CCC interactions and the spatial adjacency of their corresponding cell-types for NATMI, and moderate associations for logFC Mean and Connectome, while the other methods showed inconsistent trends or lack of signal (Supplementary Fig. S15). There was no trend in the merFISH dataset, likely due to the lower gene space of that dataset (Supplementary Fig. S15).

In summary, our results showed a positive association of interactions predicted by most methods and spatially-adjacent cell types in the well-structured brain cortex, while the associations were less consistent in the breast cancer subtypes. This positive association suggests that, despite the dissociation of single-cells and their grouping into cell types, CCC predictions partly reflect the expression patterns encoded by tissue spatial context.

Discussion

The growing interest in CCC inference has led to the recent emergence of a number of methods and prior knowledge resources. To shed light on the impact of the choice of method and resource on the inference of CCC events, we built a framework to systematically combine and compare 16 resources and 7 methods, plus their consensus. We used this framework to explore in detail the content of the different resources, to compare the predictions on six different datasets when using all combinations of methods and resources, and to assess the agreement of the methods with other data modalities. Our results suggest that both the method and resource can considerably impact CCC inference predictions, and that most methods generally capture the biological signals from other data modalities.

Resource overlap and bias. Despite their largely common origins, different resources covered varying proportions of the collective prior knowledge. A large share of the observed overlap among resources was a result of the frequent inclusion of certain resources^{8,31,33,34}, particularly Ramilowski et al.³⁵.

When inspecting the relative compositions of the resources, we noted biases towards certain organ- or tissue-enriched proteins and functional terms. Some resources are predominantly manually-curated^{8,11,16,26,27,54}, while others^{6,12,28,55} are composed of sites which also import non-curated interactions. Thus, this

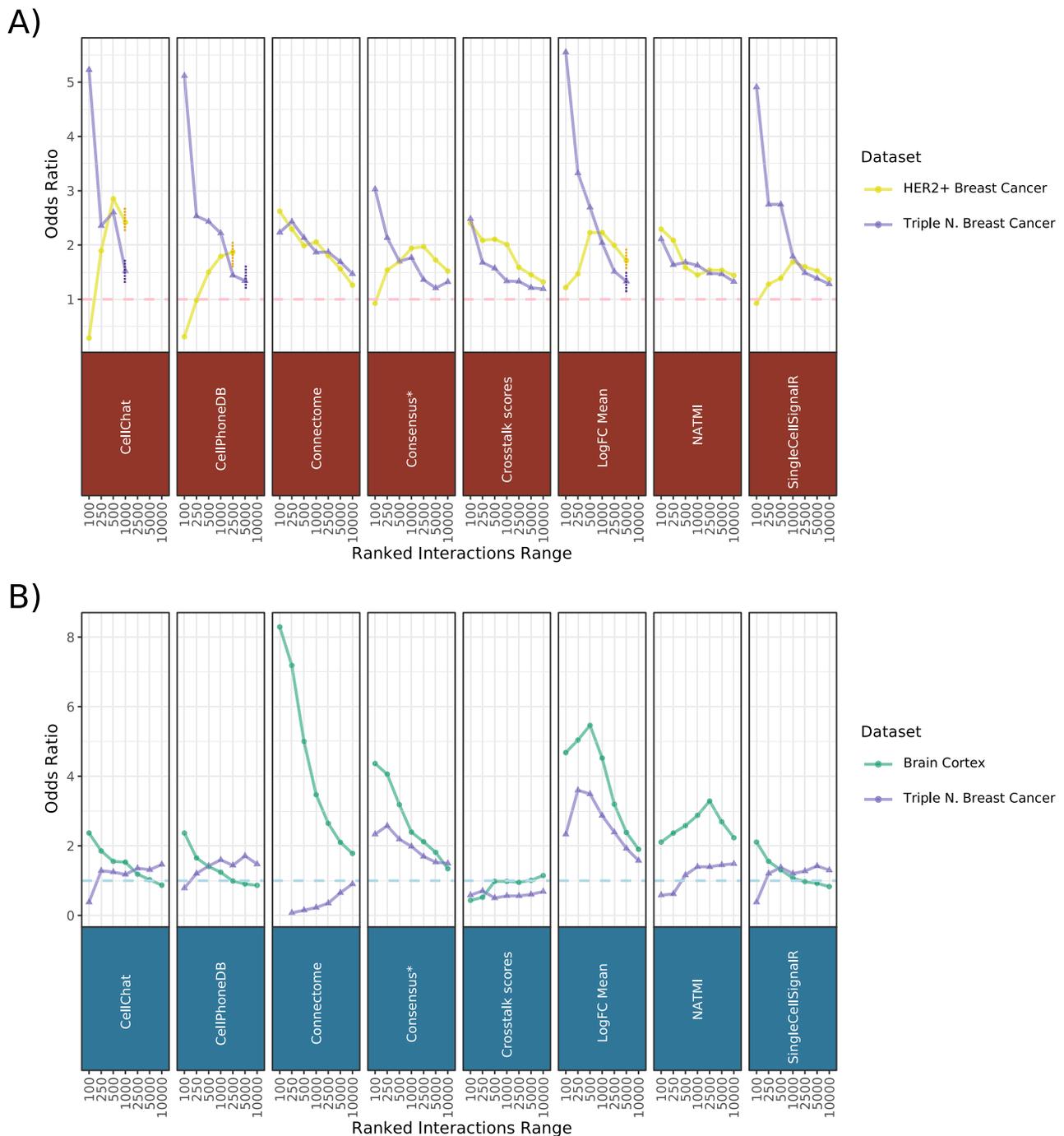


Fig. 6 Agreement of CCC predictions with other modalities. Odds ratios of **(A)** active cytokines and **(B)** colocalized cell types among the highest ranked interaction predictions, across a ranked range between 100 and 10,000. Odds ratios representing the association of preferentially ranked CCC predictions and **(A)** cytokine activities and **(B)** spatial adjacencies were calculated using Fisher's exact test. Asterisk (*): Consensus represents the aggregated ranks of all interactions predicted by all the methods. Dashed horizontal line is the baseline represented by an odds ratio of 1. The dashed vertical lines represent the truncated ranges of CellChat, CellPhoneDB, and LogFC Mean, arising from their relatively stricter preprocessing steps. Source data are provided as a Source Data file.

suggests a quality-coverage trade-off, as is commonly the case for biological prior knowledge. Of note, the literature-support reported by different authors for the same resources do not always agree^{23,26}, suggesting different interpretations of what defines a curated interaction.

These findings highlight an inherent limitation of knowledge-based inference, as any prior knowledge resource has its own biases and only represents a limited proportion of biology. Taken together, the variable overlap between the resources, their uneven

functional distributions, and the reported curation disagreements are a call for further large-scale curation efforts.

Impact of methods and resources. Our systematic analysis using different combinations of resources and methods revealed that both had a considerable effect on the predictions. In the case of the resources, the disagreements were largely expected as a consequence of their varying overlap. However, this was not necessarily the case for the methods, given their conceptually common

aim, similar assumptions, and previously reported agreement among some of them^{10–12,14}.

A major reason for the low overlap between the methods was their distinct approaches to identify the most relevant interactions. Hence, the common practice of using the number of interactions reported between two cell types as a proxy for their communication intensity is likely biased by the choice of CCC inference method. Reassuringly, our robustness analyses highlighted that the methods are fairly robust to cluster subsampling, as well as the introduction of noise to both the dataset and the resource. Collectively, these results indicate that while the methods are fairly robust to technical noise, the choice of method and resource is likely to have a major impact on the results. Therefore, downstream analyses and biological interpretation of the predicted ligand-receptor interactions should be considered with caution.

Agreement with other data modalities. Motivated by the observed discrepancies, we supported the methods' performance using complementary data modalities. We found concordance of the CCC predictions with receptor protein specificity and with cytokine activities estimated from downstream gene-expression signatures⁴⁸. Of note, the cytokine activities and receptor proteins, presented in this work as an evaluation, could also be used to improve the confidence in predictions⁵⁶. Similarly, other analyses such as pathway¹⁴ or transcription factor activities^{15,57}, as well as other types of cell-communication dedicated methods, including NicheNet¹⁹, CytoTalk²², and SoptSC²⁰, could be utilised to provide further confidence in the predicted ligand-receptor interactions.

Furthermore, similarly to previous efforts, we used spatial information to support the methods' predictions^{14,22}. We saw that most methods prioritise interactions between colocalized cell types, and this was clearer in the well-structured brain cortex than in breast cancer tissue. These results suggest that the performance of the methods depends on the type of tissue, and that, if available, spatial information should be used to inform^{58,59} or constrain⁶⁰ the predictions.

Our agreement analyses are based on assumptions that are only approximations of reality. The limitations include the restricted coverage of the cytokine activity signatures and receptor proteins, and the technical shortcomings of current spatial transcriptomics technologies. Furthermore, such benchmarks cannot distinguish simple co-expression from actual CCC events, and do not capture complex relationships between CCC events. Since a gold standard is currently not available and the biological ground truth is largely unknown^{2,25}, our analyses cannot give a definitive answer of what method is best. However, we believe that these results are useful to indirectly support the methods' predictive potential.

Overall, our results suggest that despite their relatively low agreement, the CCC methods are generally able to capture relevant biological signals, and that leveraging information from additional modalities and analyses could help to refine the predictions.

CCC inference assumptions and limitations. The shared purpose of the methods considered in this work is to predict the most relevant interactions, commonly between a secreted ligand and its receptor, each expressed by a particular cell type. All methods work under the assumption that the expression of a pair of genes at the cell type level is informative of CCC events. Some of the methods such as CellChat¹⁴, CellPhoneDB⁸, and others^{16,27,28}, go a step further by considering heteromeric complexes. Ensuring that all subunits of a protein complex are expressed to consider a cell-cell interaction valid has been shown to reduce false positive

predictions, and can thus impact significantly downstream interpretation and validation^{8,14}. CellChat additionally accounts for interaction mediator proteins¹⁴. Another common assumption among the CCC methods is that cell-type-specific interactions are more informative than those shared by multiple cell types^{8,10,11,14}. Yet by focusing on the cluster-specific interactions, the predictions may not capture biologically relevant processes that are common between multiple cell types¹².

Gene expression provided by scRNA-Seq is typically limited to the cells within the dataset, and hence does not capture long-distance endocrine signalling events. In addition, CCC inference from scRNA-Seq data assumes that gene expression of a transmitter and a receiver is a good proxy for their joint activity, without considering any of the processes preceding transmitter-receiver interactions, including protein translation and processing, secretion, and diffusion². Furthermore, gene expression is a proxy of protein levels alone, yet recent efforts attempt to capture signalling events mediated by other molecules such as neurotransmitters^{15,16}. Finally, current methods are limited to single species although some information about interspecies communication can be inferred^{61,62}.

Conclusions

Considerable efforts have been made to develop CCC inference tools and resources, and we expect that further advancements will be key for the systems-level analysis of single-cell data. The popularity of CCC inference is anticipated to increase as spatial transcriptomics¹ and single-cell proteomics⁶³ continue their rapid development. We regard the results presented here as steps towards an understanding of the strengths and weaknesses of CCC methods, and LIANA as a framework for their further analysis, benchmark, use and development.

Methods

Processing of CCC resources. The connections between resources shown in the dependency plot were manually gathered from the publications and the web pages of each CCC resource.

OmniPath is a comprehensive knowledge database with more than 100 intracellular and intercellular resources³. The OmniPath intercellular component is a composite resource which contains interactions from all of the CCC dedicated resources compared here, along with some additional resources³. All the CCC resources used in the analyses presented in this work were queried from OmniPath³, with the exception of CellCall which was processed to OmniPath format separately. The contents of the resources are identical to their original formats, apart from minor processing differences (Supplementary Table 2), such as removal of duplicates, updating to the latest gene symbols, or removal of genes lacking reviewed Uniprot IDs. All complex-containing resources were dissociated into individual subunits for the resource-focused analyses presented in this work.

OmniPath's version used in this work was filtered according to the following criteria: (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 the 51st consensus percentile of the localisation annotations, and (iii) we only considered interactions between single proteins (interactions between complexes are also available in OmniPath). Tutorials on how to customise OmniPath as well as how to make use of the intracellular functional information available at OmniPath are available at <https://saezlab.github.io/liana/>. OmniPath's intra- and intercellular components were both obtained and are both available via the OmnipathR package (<https://github.com/saezlab/OmnipathR>).

Descriptive analysis of resources. 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.

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. The test *p*-values were FDR corrected. We performed the analysis using the aforementioned functional annotation databases in 3 distinct categories. For the overrepresentation of interactions, we considered annotations when both the transmitter and receiver were matched to the same category, while annotations matched to transmitters and receivers enrichments were examined independently. We allowed the same protein or interaction to be matched to multiple pathways or functional categories from the same database. Interactions, receivers, and transmitters were independently

matched to the 10 pathways from SignalLink³⁸, and the 15 largest categories from CancerSEA⁴¹, and NetPath⁴⁰. The same procedure was also applied to organ- and tissue-enriched proteins from the Human Protein Atlas⁴², accessible at <https://proteatlas.org>, and disease-associated genes from DisGeNet⁴³. Pathology-associated, uncertain, and unsupported proteins with a low/non-representative level of expression were excluded from Human Protein Atlas database, while DisGeNet gene-disease associations were filtered to include only literature-supported associations (GDA Score ≥ 0.3). Each of the aforementioned general functional annotation databases was obtained via OmniPath and their protein complexes, if present, were also dissociated.

We also obtained protein localisations from OmniPath which collects this information from 20 databases³. Then we kept consensus protein localisations above the 51st percentile. We classified CCC interactions using the localisation combinations of proteins involved in the interactions, which included secreted, plasma membrane peripheral and transmembrane proteins.

Input specifics. For the method-resource comparisons and evaluations, we used Seurat^{46,64} objects which were converted to the appropriate data format when calling each method. Whenever available, we used the recommended conversion method or wrapper for each method. Log-transformed counts were used when this was not done internally by the method.

The complex-containing interactions, if present in a given resource, were dissociated for the methods which do not take complexes into account, namely the original implementations of NATMI, SingleCellSignalR, and Connectome.

Method specifics

CellChat. CellChat was run using its default settings with 1000 permutations and the gene expression diffusion-based smoothing process was omitted.

CellPhoneDBv2. CellPhoneDB's algorithm⁸ was re-implemented in LIANA and used throughout this manuscript with 1000 permutations. Identical to the original implementation, cluster labels were reshuffled and an one-sided empirical *p*-value was calculated for the interactions with a mean expression higher than random. Only interactions whose transmitter and receiver genes were expressed in at least 10% of the cells were considered, and the subunit with the minimum expression was used for complexes.

Connectome. Connectome was run with its default settings and filtered for differentially expressed genes (*p*-value ≤ 0.05), as identified via a Wilcoxon test.

logFC Mean. The LogFC Mean score implemented in LIANA, was inspired by iTALK⁶, and it represents the average of one-versus-the-rest log₂FC expression changes for the transmitter and receiver cell types. The logFC Mean score uses LIANA's default filtering settings, namely both the transmitter and receiver genes of any interaction evaluated must be expressed in at least 10% of the cells, and it considers the subunit with the minimum expression for complex-containing interactions.

SingleCellSignalR. SingleCellSignalR was run with the processed gene counts, considering differentially expressed genes with a log₂ fold change threshold of 1.5 or above, and we filtered LRscores ≥ 0.5 for the evaluations. The "int.type" parameter was set to "autocrine". We noted that this option returned both paracrine and autocrine signalling interactions. The source code of SingleCellSignalR was modified to work with external resources (available at https://github.com/saezlab/SingleCellSignalR_v1).

NATMI. NATMI's implementation is command-line based, thus a system command is 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. The source code of NATMI was modified to be path-agnostic and to work with integers as cluster names (available at <https://github.com/saezlab/NATMI>).

Crosstalk scores. Crosstalk scores, inspired by CytoTalk²², were implemented in LIANA. CytoTalk's crosstalk scores are composed of two metrics: the preferential expression measures (PEMs) and the non-self talk scores (NSTs). The first one reflects the specific expression for quantified genes across all the cell types. The latter is defined on the basis of information theoretic measures and quantifies the mutual information (Shannon entropy) for a pair of genes (ligand and receptor) within the same cell type, and is thus designed to penalise autocrine signalling. Once NST and PEM are calculated, the crosstalk score is calculated for each ligand-receptor pair and for each cell type pair as the product of the minmax normalised PEM and NST values. To enable the comparison to the rest of the methods, and in contrast to the crosstalk scores implemented in CytoTalk, we calculated the crosstalk scores by cell type pairs and used the inverse of the non-self-talk scores for autocrine signalling interactions. Moreover, our implementation considers complexes, and interactions with transmitters or receivers with preferential expression measures of 0 are also assigned 0.

Robust-rank aggregate. A consensus rank is generated across all methods using Robust Rank Aggregation⁶⁵. These aggregated ranks can in turn be interpreted as a probabilistic distribution for interactions that are preferentially highly-ranked. The aggregate ranks are built across the universe of all interaction predictions, after independent filtering by each method. By default, missing interactions are imputed as the max ranks.

Overlap analysis. To compare the overlap between the interactions predicted by each method-resource combination, we kept the 1,000 highest ranked interactions by default, including ties. We also considered the highest ranked 1% of interactions for each method, including ties. We then generated a presence-absence matrix of predicted interactions with method-resource combinations. These matrices were subsequently used to calculate the reported Jaccard indices.

Unless explicitly mentioned, and if available, we used the scoring functions for each method recommended for single-condition interaction predictions (Supplementary Table 3).

Frequencies of interactions per cell type were calculated using the highest ranked hits for each method-resource combination. These frequencies represent the proportion of top predicted interactions (or edges) that stem from or lead to a source or target cell type, respectively. In other words, interaction frequencies represent the relative number of interactions per cell type within the highest ranked 1000 interactions.

The relative interaction strength by cell type was calculated using the regularised scores from each method, i.e. all scoring functions were scaled between 0 and 1. Then the mean regularised score per cell type, categorised as source or target, was divided by the average score of all interactions predicted.

Agreement with other modalities and robustness. All of the comparisons with other modalities were performed using the OmniPath CCC resource. For murine datasets, we converted the OmniPath to murine symbols using the biomaRt package⁶⁶.

For the binary categorisations used in the agreement with cytokine activity analysis and spatial adjacencies, we performed Fisher's exact test, sequentially in rank intervals ranging from 100 to 10,000, to obtain the Odds ratios of the positive and negative classes against a background universe. In the case of the spatial adjacency analysis, the background universe contained all predicted interactions, while for the cytokine activities, we only considered those matched to cytokines from CytoSig⁴⁸.

Agreement with cytokine signatures. CytoSig provides a collection of consensus, data-driven, cytokine-activity signatures compiled using a compendium of transcriptomic profiles⁴⁸. We used CytoSig's 43 high-quality signatures to infer which cytokines induce signalling activities in each cell type. We then used this information to assess if a cytokine-receptor interaction reported by the different CCC methods was supported by the corresponding cytokine downstream signalling activities.

We computed the cytokine activity scores for all cell types with the multivariate linear regression model ("mlm") method of decoupleR at the pseudobulk level. We chose the mlm method as an approach that models the effect of multiple cytokines and that performed best in a recent footprint-focused analysis benchmark⁶⁷.

To build the pseudobulk profiles, we log₂-transformed the summed counts within each cell type, and kept only genes which were expressed in at least 10% of the cells and with a summed raw count above 5.

In this evaluation, we used both the autocrine and paracrine CCC predictions, calculated using expression counts at the cell-type level for all cell types, from the HER2+ and triple negative breast cancer subtype datasets⁴⁴. We considered any cytokine signature with a positive score and FDR-corrected *p*-value ≤ 0.05 in the target cell types as an active cytokine. We considered all CCC predictions with a ligand corresponding to a CytoSig signature, including the same ligand to multiple receptors, matched to any of the aliases of the cytokines. Odds ratios were then calculated as the ratio between any CCC prediction with corresponding active cytokine in a given receiver cell type, and those assigned to the negative class—i.e. the remainder of the cytokine signatures.

Agreement with spatially adjacent cell types. We used the SPOTlight⁶⁸ deconvolution method with default parameters to spatially map the cell types present in our scRNAseq datasets into their corresponding 10× Visium slides. SPOTlight provides cell type proportions per spot that were subsequently used to identify colocalized cell types by computing Pearson's correlation. The Pearson coefficients were scaled to create a distribution of correlations, and only considered the most strongly correlated cell type densities (*z*-score ≥ 1.645) as colocalized, while the remainder of the cell pairs were considered as non-colocalised.

The mer- and seqFISH datasets were already annotated and provide single-cell spatial resolution, hence the same dataset was used to obtain CCC predictions and spatial information. To identify the enriched neighbouring cells for each cell type mer- and seqFISH datasets, we used Squidpy's⁶⁹ Neighbourhood Enrichment analysis with its default parameters. In accordance with the approach followed with the 10× VISIUM slides, we considered significantly colocalized cell type pairs with a normalised neighbourhood enrichment score ≥ 1.645 as spatially adjacent.

Agreement with receptor protein abundance. To identify specifically expressed receptors across clusters, we z -transformed receptor protein abundance across cell types. Receptors with an abundance z -score ≥ 1.645 were considered specifically abundant at the protein level. These receptors were then treated as the positive class, while all others were assigned to the negative class. AUROC and AUPRC metrics were calculated using yardstick⁷⁰. For the AUPRC calculations, we downsampled the negative class 100 times to match the (lower) number of receptors assigned to the positive class. The downsampling procedure binds the expected random AUPRC to 0.5.

We allowed surface protein receptors to match multiple genes (e.g. T-cell receptors subunits), and vice versa. Gene aliases of proteins were obtained using the human and mouse gene databases from the org.Hs.eg.db⁷¹ and org.Mm.eg.db⁷² BioConductor packages. Proteins with non-standard names, or absent aliases in the aforementioned databases, were manually annotated using UniProt⁷³ as a reference.

Robustness analyses. To evaluate sensitivity of the methods to noise, we performed four distinct robustness analyses. We simulated noise in the data by subsampling the number of cells per cluster and by reshuffling the cell type labels.

Additionally, to simulate the impact of false interactions in the resource, we randomly generated interactions from the 2000 most variable genes in the dataset and randomly replaced proportions of the resource with these putative false interactions. In one scenario, we selectively replaced interactions in the resource and preserved the highest ranked interactions, while in the other scenario we non-selectively swapped any of the interactions.

All four analyses were done in an iterative manner over a range of manipulations (0–40%). We treated the highest ranked 250 interactions from the non-modified resource/data as ground truth and repeated the randomisation process 5 times.

Data processing. All 10× Genomics, including all CITE-Seq and the 3k PBMC, datasets were processed using the standard Seurat pipeline. Namely, filtered gene expression count matrices were log-normalised, and if cell type annotations were not provided, the cells were clustered, following scaling, identification of variable features, and PCA dimensionality reduction, using Seurat's⁶⁴ (v4.0.3) default settings. For 10× Genomics CITE-Seq datasets we used a clustering resolution of 0.4 and the protein abundances were centred-log-ratio transformed. In the Murine spleen-lymph CITE-Seq datasets⁷⁴, duplicated and low quality cells, as annotated by the original authors, were filtered, in agreement with the other CITE-seq datasets, gene counts were log-normalised, while protein abundances were centred-log-ratio transformed.

For the colorectal cancer dataset, we kept the original subtype labels, reformatted the names to work with each CCC method, and sparsified the counts into a Seurat⁶⁴ object. The pre-processed and labelled Pancreatic islet⁴⁶ and cord blood mononuclear cell⁴⁵ datasets were log-normalised, and subsequently used for CCC inference. In the latter dataset, any murine and doublet/multiplet cells, as annotated by the authors, were excluded.

We used ComplexHeatmap⁷⁵ to generate the heatmaps and ggplot2⁷⁶ for any of the other plots presented in this work.

Reporting summary. Further information on research design is available in the Nature Research Reporting Summary linked to this article.

Data availability

The processed and annotated Human Breast Cancer single-cell atlas⁴⁴ is available via the GEO accession number: GSE176078. The filtered breast cancer 10× Visium slides from the same publication are available at <https://zenodo.org/record/4739739>. Processed seqFISH⁷⁷ [<https://content.cruk.cam.ac.uk/jmlab/SpatialMouseAtlas2020/>] and merFISH³³ (GEO accession number: GSE113576) datasets were obtained via the spatial single-cell analysis framework—Squidpy (v1.1.0)⁶⁹ [<https://squidpy.readthedocs.io/en/latest/api.html#module-squidpy.datasets>].

Pancreatic islet⁴⁶ (GEO accession numbers: GSE84133, GSE81076, GSE85241, GSE86469; ArrayExpress: E-MTAB-5061) and cord blood mononuclear cells⁴⁵ (GEO accession number: GSE100866) scRNA-Seq datasets were obtained via SeuratData [<https://github.com/satijalab/seurat-data>].

Publicly available 5K PBMC, 5K PBMC NextGem, 10K PBCM, and 10K MALT CITE-Seq datasets were obtained from 10× Genomics (accessible under the list of datasets at <https://tinyurl.com/10xCITEseq>).

Processed and annotated murine spleen-lymph CITE-Seq datasets⁷⁴ are available via the GEO accession number: GSE150599.

The processed single cell RNA-Seq data⁴⁷ for 23 Korean colorectal cancer patients are available via the GEO accession number: GSE132465.

Spatial transcriptomics datasets (10× 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', or publicly via the 10× Genomics website under Spatial Gene Expression v1 Chemistry datasets [<https://tinyurl.com/10xVisiumDemonstration>]. The single-cell data (Allen Brain Atlas⁵¹) used for the cell type mapping (deconvolution), was obtained as a Seurat object, accessible at https://www.dropbox.com/s/cuowwm4vrf65pvq/allen_cortex.rds?dl=1, and is alternatively available via accession number: GSE71585.

The 10× Genomics' 3k PBMC dataset used in the robustness analysis is available at https://cf.10xgenomics.com/samples/cell/pbmc3k/pbmc3k_filtered_gene_bc_matrices.tar.gz. Source Data for all Supplementary Figures, along with preprocessed outputs, are available at: <https://zenodo.org/record/6531218>. Source data are provided with this paper.

Code availability

The LIANA framework is available at <https://github.com/saezlab/liana>, and the version used to generate the results presented here is available via Zenodo⁷⁸. The scripts used to generate the results presented here can be accessed at https://github.com/saezlab/ligrec_decouple.

Received: 17 June 2021; Accepted: 17 May 2022;

Published online: 09 June 2022

References

- 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).
- 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).
- Türei, D. et al. Integrated intra- and intercellular signaling knowledge for multicellular omics analysis. *Mol. Biol.* **17**, e9923 (2021).
- Kumar, M. P. et al. Analysis of Single-Cell RNA-Seq Identifies Cell-Cell Communication Associated with Tumor Characteristics. *Cell Rep.* **25**, 1458–1468.e4 (2018).
- Cillo, A. R. et al. Immune Landscape of Viral- and Carcinogen-Driven Head and Neck Cancer. *Immunity* **52**, 183–199.e9 (2020).
- Wang, Y. et al. iTALK: an R Package to Characterize and Illustrate Intercellular Communication. *BioRxiv* <https://doi.org/10.1101/507871> (2019).
- Tyler, S. R. et al. PyMINer Finds Gene and Autocrine-Paracrine Networks from Human Islet scRNA-Seq. *Cell Rep.* **26**, 1951–1964.e8 (2019).
- 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).
- Tsuyuzaki, K., Ishii, M. & Nikaido, I. Uncovering hypergraphs of cell-cell interaction from single cell RNA-sequencing data. *BioRxiv* <https://doi.org/10.1101/566182> (2019).
- Raredon, M. S. B. et al. Connectome: computation and visualization of cell-cell signaling topologies in single-cell systems data. *BioRxiv* <https://doi.org/10.1101/2021.01.21.427529> (2021).
- Hou, R., Denisenko, E., Ong, H. T., Ramiłowski, J. A. & Forrest, A. R. R. Predicting cell-to-cell communication networks using NATMI. *Nat. Commun.* **11**, 5011 (2020).
- Cabello-Aguilar, S. et al. SingleCellSignalR: inference of intercellular networks from single-cell transcriptomics. *Nucleic Acids Res.* **48**, e55 (2020).
- Noël, F. et al. ICELLNET: a transcriptome-based framework to dissect intercellular communication. *BioRxiv* <https://doi.org/10.1101/2020.03.05.976878> (2020).
- Jin, S. et al. Inference and analysis of cell-cell communication using CellChat. *Nat. Commun.* **12**, 1088 (2021).
- 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* <https://doi.org/10.1093/bioinformatics/btab245> (2021).
- Zhang, Y. et al. Cellinker: a platform of ligand-receptor interactions for intercellular communication analysis. *Bioinformatics* <https://doi.org/10.1093/bioinformatics/btab036> (2021).
- Lagger, C. et al. scAgeCom: a murine atlas of age-related changes in intercellular communication inferred with the package scDiffCom. *BioRxiv* <https://doi.org/10.1101/2021.08.13.456238> (2021).
- 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).
- Browaeys, R., Saelens, W. & Saeys, Y. NicheNet: modeling intercellular communication by linking ligands to target genes. *Nat. Methods* **17**, 159–162 (2020).
- 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).
- Cheng, J., Zhang, J., Wu, Z. & Sun, X. Inferring microenvironmental regulation of gene expression from single-cell RNA sequencing data using scMNet with an application to COVID-19. *Brief. Bioinforma.* **22**, 988–1005 (2021).

22. Hu, Y., Peng, T., Gao, L. & Tan, K. CytoTalk: De novo construction of signal transduction networks using single-cell transcriptomic data. *Sci. Adv.* **7**, eabf1356 (2021).
23. Zhang, Y. et al. CellCall: integrating paired ligand-receptor and transcription factor activities for cell-cell communication. *Nucleic Acids Res.* **49**, 8520–8534 (2021).
24. Mishra, V. et al. Systematic elucidation of neuron-astrocyte interaction in models of amyotrophic lateral sclerosis using multi-modal integrated bioinformatics workflow. *Nat. Commun.* **11**, 5579 (2020).
25. Almet, A. A., Cang, Z., Jin, S. & Nie, Q. The landscape of cell-cell communication through single-cell transcriptomics. *Curr. Opin. Syst. Biol.* **26**, 12–23 (2021).
26. Shao, X. et al. CellTalkDB: a manually curated database of ligand-receptor interactions in humans and mice. *Brief. Bioinformatics* **22**, bbaa269 (2021).
27. Noël, F. et al. Dissection of intercellular communication using the transcriptome-based framework ICELLNET. *Nat. Commun.* **12**, 1089 (2021).
28. 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).
29. Kanehisa, M. & Goto, S. KEGG: Kyoto encyclopedia of genes and genomes. *Nucleic Acids Res.* **28**, 27–30 (2000).
30. Kanehisa, M., Furumichi, M., Sato, Y., Ishiguro-Watanabe, M. & Tanabe, M. KEGG: integrating viruses and cellular organisms. *Nucleic Acids Res.* **49**, D545–D551 (2021).
31. Fabregat, A. et al. The Reactome Pathway Knowledgebase. *Nucleic Acids Res.* **46**, D649–D655 (2018).
32. 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).
33. 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).
34. 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).
35. Ramilowski, J. A. et al. A draft network of ligand-receptor-mediated multicellular signalling in human. *Nat. Commun.* **6**, 7866 (2015).
36. Licata, L. et al. SIGNOR 2.0, the SIGNaling Network Open Resource 2.0: 2019 update. *Nucleic Acids Res.* **48**, D504–D510 (2020).
37. 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).
38. Fazekas, D. et al. SignaLink 2—a signaling pathway resource with multi-layered regulatory networks. *BMC Syst. Biol.* **7**, 7 (2013).
39. Kirouac, D. C. et al. Dynamic interaction networks in a hierarchically organized tissue. *Mol. Syst. Biol.* **6**, 417 (2010).
40. Kandasamy, K. et al. NetPath: a public resource of curated signal transduction pathways. *Genome Biol.* **11**, R3 (2010).
41. Yuan, H. et al. CancerSEA: a cancer single-cell state atlas. *Nucleic Acids Res.* **47**, D900–D908 (2018).
42. Uhlén, M. et al. Proteomics. Tissue-based map of the human proteome. *Science* **347**, 1260419 (2015).
43. Piñero, J. et al. DisGeNET: a comprehensive platform integrating information on human disease-associated genes and variants. *Nucleic Acids Res.* **45**, D833–D839 (2017).
44. Wu, S. Z. et al. A single-cell and spatially resolved atlas of human breast cancers. *Nature Genetics* **53**, 1334–1347 (2021).
45. Stoeckius, M. et al. Simultaneous epitope and transcriptome measurement in single cells. *Nat. Methods* **14**, 865–868 (2017).
46. 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).
47. Lee, H.-O. et al. Lineage-dependent gene expression programs influence the immune landscape of colorectal cancer. *Nat. Genet.* **52**, 594–603 (2020).
48. Jiang, P. et al. Systematic investigation of cytokine signaling activity at the tissue and single-cell levels. *Nat. Methods* **18**, 1181–1191 (2021).
49. Armingol, E. et al. Inferring the spatial code of cell-cell interactions and communication across a whole animal body. *BioRxiv* <https://doi.org/10.1101/2020.11.22.392217> (2020).
50. Palla, G., Fischer, D. S., Regev, A. & Theis, F. J. Spatial components of molecular tissue biology. *Nat. Biotechnol.* **40**, 308–318 (2022).
51. Tasic, B. et al. Adult mouse cortical cell taxonomy revealed by single cell transcriptomics. *Nat. Neurosci.* **19**, 335–346 (2016).
52. Lohoff, T. et al. Highly multiplexed spatially resolved gene expression profiling of mouse organogenesis. *BioRxiv* <https://doi.org/10.1101/2020.11.20.391896> (2020).
53. Moffitt, J. R. et al. Molecular, spatial, and functional single-cell profiling of the hypothalamic preoptic region. *Science* **362**, eaau5324 (2018).
54. Jin, S. et al. Inference and analysis of cell-cell communication using CellChat. *BioRxiv* <https://doi.org/10.1101/2020.07.21.214387> (2020).
55. Sheikh, B. N. et al. Systematic Identification of Cell-Cell Communication Networks in the Developing Brain. *iScience* **21**, 273–287 (2019).
56. 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).
57. Jung, S., Singh, K. & del Sol, A. FunRes: resolving tissue-specific functional cell states based on a cell-cell communication network model. *Brief. Bioinformatics* **22**, bbaa283 (2021).
58. Fischer, D. S., Schaar, A. C. & Theis, F. J. Learning cell communication from spatial graphs of cells. *BioRxiv* <https://doi.org/10.1101/2021.07.11.451750> (2021).
59. Tanevski, J., Ramirez Flores, R. O., Gabor, A., Schapiro, D. & Saez-Rodriguez, J. Explainable multi-view framework for dissecting inter-cellular signaling from highly multiplexed spatial data. *BioRxiv* <https://doi.org/10.1101/2020.05.08.084145> (2020).
60. 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).
61. 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* <https://doi.org/10.1101/2021.03.20.436262> (2021).
62. Westermann, A. J. & Vogel, J. Cross-species RNA-seq for deciphering host-microbe interactions. *Nat. Rev. Genet.* **22**, 361–378 (2021).
63. Mahdessian, D. et al. Spatiotemporal dissection of the cell cycle with single-cell proteogenomics. *Nature* **590**, 649–654 (2021).
64. 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).
65. Kolde, R., Laur, S., Adler, P. & Vilo, J. Robust rank aggregation for gene list integration and meta-analysis. *Bioinformatics* **28**, 573–580 (2012).
66. Durinck, S., Spellman, P. T., Birney, E. & Huber, W. Mapping identifiers for the integration of genomic datasets with the R/Bioconductor package biomaRt. *Nat. Protoc.* **4**, 1184–1191 (2009).
67. Badia-i-Mompel, P. et al. decoupleR: Ensemble of computational methods to infer biological activities from omics data. *BioRxiv* <https://doi.org/10.1101/2021.11.04.467271> (2021).
68. 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).
69. Palla, G. et al. Squidpy: a scalable framework for spatial single cell analysis. *BioRxiv* <https://doi.org/10.1101/2021.02.19.431994> (2021).
70. Kuhn, M. & Vaughan, D. *yardstick: Tidy Characterizations of Model Performance*. (CRAN, 2021).
71. Carlson, M. org.Hs.eg.db: Genome wide annotation for Human. R package version 3.8.2. *Bioconductor* <https://doi.org/10.18129/b9.bioc.org.hs.eg.db> (2017).
72. Carlson, M. org.Mm.eg.db. *Bioconductor* <https://doi.org/10.18129/b9.bioc.org.mm.eg.db> (2017).
73. UniProt Consortium. UniProt: the universal protein knowledgebase in 2021. *Nucleic Acids Res.* **49**, D480–D489 (2021).
74. Gayoso, A. et al. Joint probabilistic modeling of single-cell multi-omic data with totalVI. *Nat. Methods* **18**, 272–282 (2021).
75. Gu, Z., Eils, R. & Schlesner, M. Complex heatmaps reveal patterns and correlations in multidimensional genomic data. *Bioinformatics* **32**, 2847–2849 (2016).
76. Wickham, H. ggplot2. *WIREs Comp. Stat.* **3**, 180–185 (2011).
77. Lohoff, T. et al. Integration of spatial and single-cell transcriptomic data elucidates mouse organogenesis. *Nat. Biotechnol.* **40**, 74–85 (2022).
78. Dimitrov, D. saezlab/liana version 0.05 (Devel). *Zenodo* <https://doi.org/10.5281/zenodo.6475164> (2022).

Acknowledgements

This work was supported in part by the European Union's Horizon 2020 research and innovation program (860329 Marie-Curie ITN "STRATEGY-CKD") and the German Federal Ministry of Education and Research (Bundesministerium für Bildung und Forschung BMBF) Computational Life Sciences LaMarck grant no. 031L0181B), awarded to J.S.R. This work was in part funded by the clinical research unit CRU344 supported by the German Research Foundation (DFG) and the E:MED Consortia Fibromap funded by the German Ministry of Education and Science (BMBF) awarded to I.C. We express our gratitude to Erick Armingol, Pau Badia i Mompel, Hratch Baghdassarian, Luz Garcia-Alonso and Suoqin Jin for their helpful feedback and discussions and to Ece Kartal for the design of LIANA's outline graphic. For the publication fee we acknowledge financial support by Deutsche Forschungsgemeinschaft within the funding programme "Open Access Publikationskosten" as well as by Heidelberg University.

Author contributions

J.S.R. conceived the project. D.D. set up the framework used in this manuscript, with the help of D.T., M.G.R., and J.S.N. D.D. performed the comparisons and evaluations

presented in this work with the support of A.D., A.V., R.O.R.F., and J.S.R. D.T. set up the resource formatting infrastructure with the help of D.D. D.T., D.D., and C.B. created the resource analysis pipeline. P.L.B. performed the robustness analysis under the guidance of D.D. J.S.R. supervised the project with the help of A.V. and A.D. H.K., R.O.R.F., and B.S. performed preliminary and supplementary analyses that helped shape the work presented here. I.C. supervised J.S.N. A.D. and A.V. contributed equally to the manuscript. All authors contributed and revised the final version of the manuscript.

Funding

Open Access funding enabled and organized by Projekt DEAL.

Competing interests

J.S.R. has received funding from GSK and Sanofi and fees from Travers Therapeutics. A.V. is currently employed by F. Hoffmann-La Roche Ltd. The remaining authors declare no competing interests.

Additional information

Supplementary information The online version contains supplementary material available at <https://doi.org/10.1038/s41467-022-30755-0>.

Correspondence and requests for materials should be addressed to Julio Saez-Rodriguez.

Peer review information *Nature Communications* thanks Qing Nie, Xiaohui Fan and the other anonymous reviewer(s) for their contribution to the peer review of this work. Peer review reports are available.

Reprints and permission information is available at <http://www.nature.com/reprints>

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2022