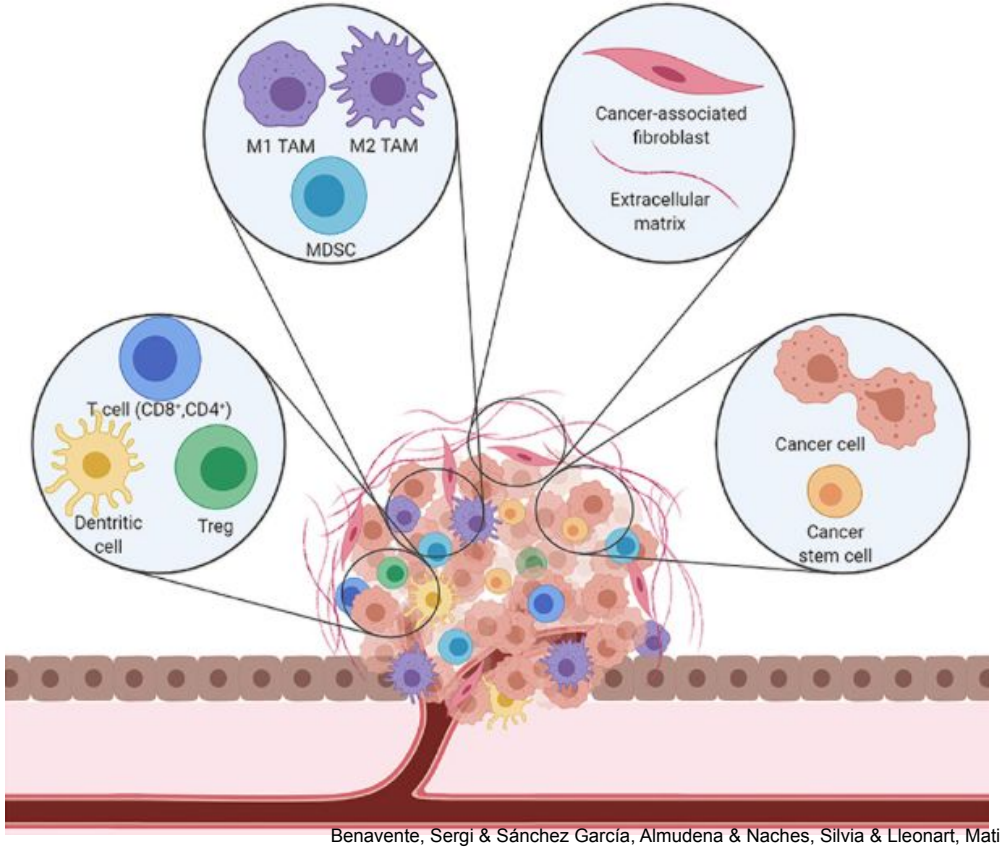
Cell-cell communication in spatial transcriptomics.

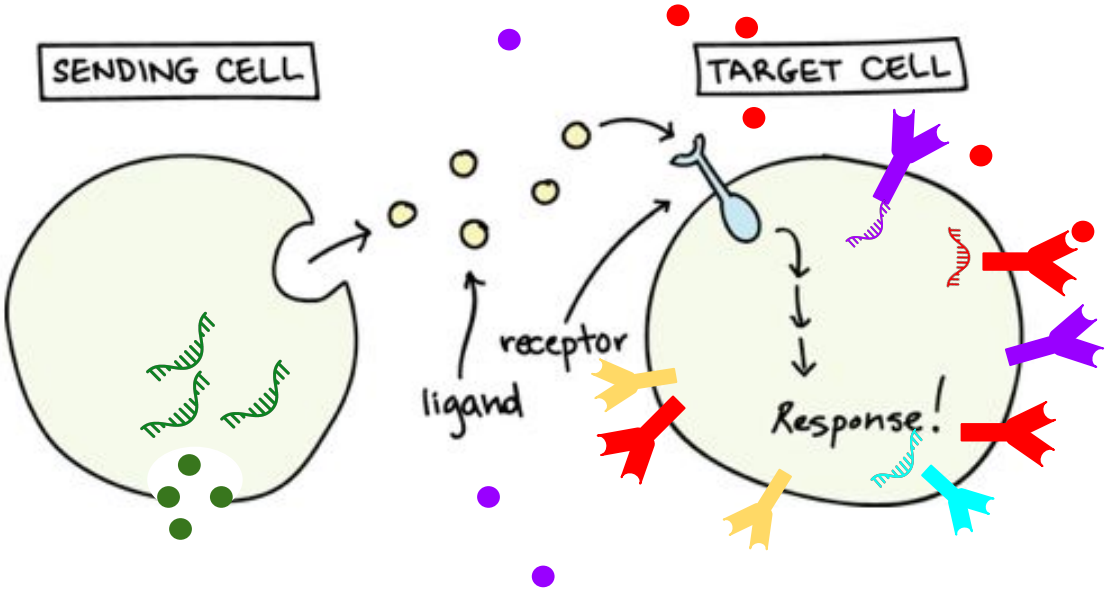
Alec Stear - 26th June 2024

Understanding cell cell communication is key for understanding tumours



Benavente, Sergi & Sánchez García, Almudena & Naches, Silvia & Lleonart, Matilde & Lorente, Juan. (2020). Therapy-Induced Modulation of the Tumor Microenvironment: New Opportunities for Cancer Therapies. 10.3389/fonc.2020.582884.

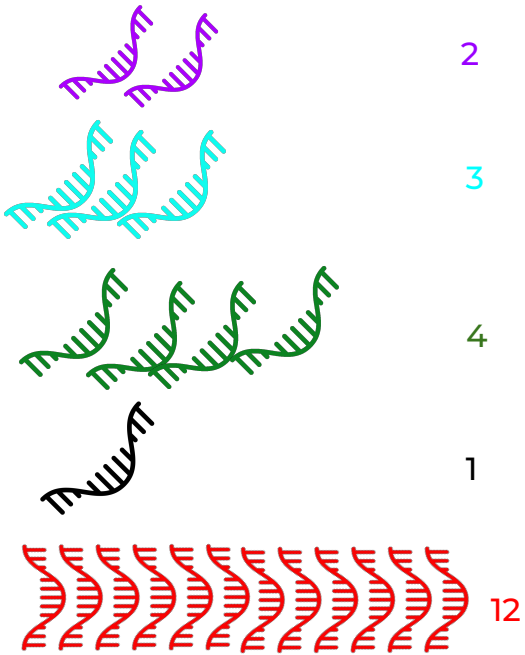
Understanding cell cell communication is key for understanding tumours



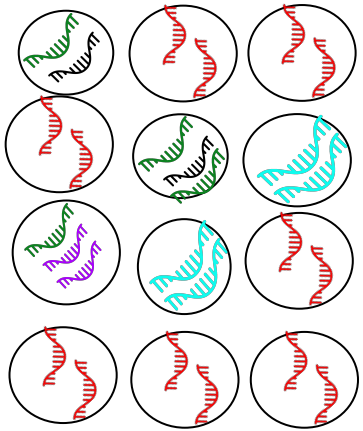
https://www.khanacademy.org/science/ap-biology/cell-communication-and-cell-cycle/cell-communication/a/introduction-to-cell-signaling

Spatial transcriptomics adds spatial localization information to infer cell-cell communication

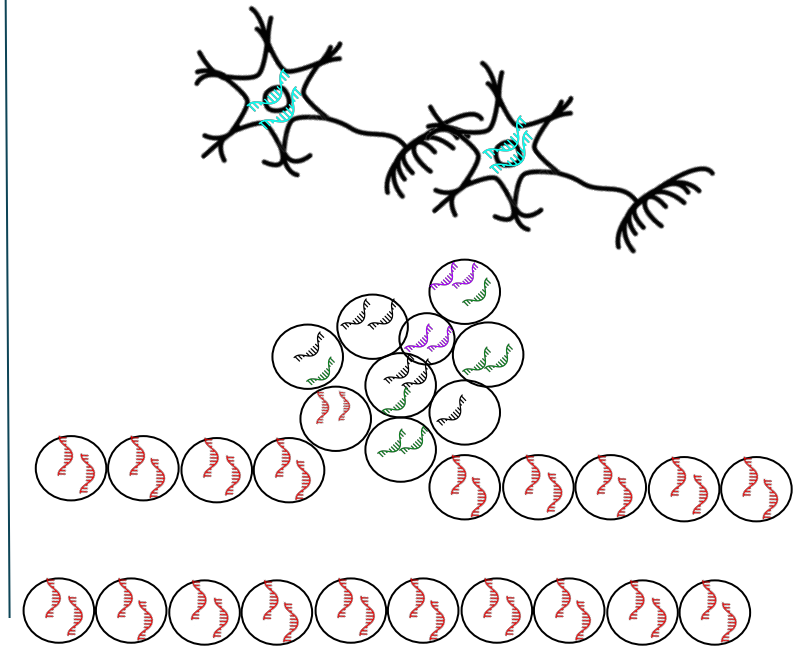
Bulk transcriptomics



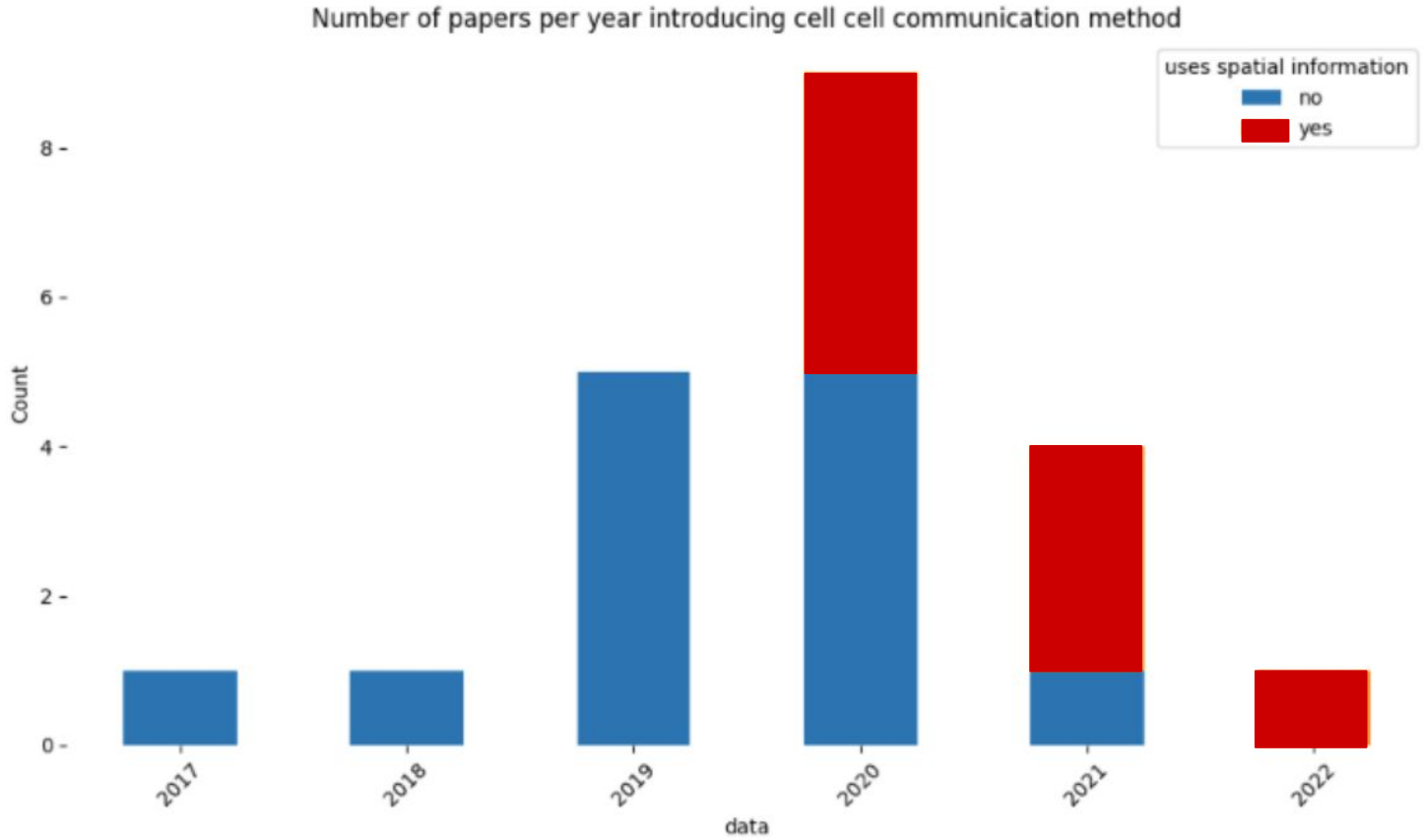
Single cell transcriptomics



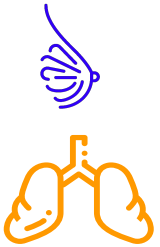
Spatial transcriptomics

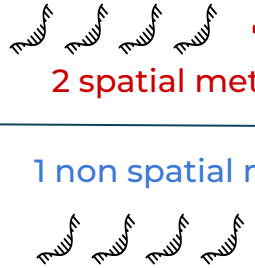


How do spatial & non spatial methods compare for inferring cell cell communication ?



Study Design





(x, y)

2 spatial methods

1 non spatial methods

Liana :	Commot :	SpatialDM :
Lig - Receptor	Lig - Receptor	Lig - Receptor
Lig - Receptor	Lig - Receptor	Lig - Receptor
Lig - Receptor	Lig - Receptor	Lig - Receptor
Lig - Receptor		Lig - Receptor
Lig - Receptor		
Lig - Receptor		

X 10

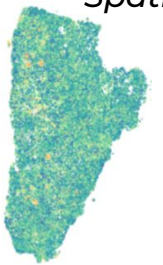
genes

Lig - Receptor : 2
Lig - Receptor : 3
Lig - Receptor : 1
Lig - Receptor : 3

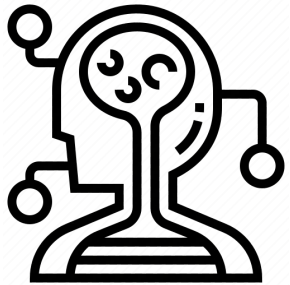
Diversity among methods ?

$$\text{Jaccard Index} = \frac{|A \cap B|}{|A \cup B|}$$

Spatial correlation ?



Biological significance ?



Non spatial method for CCC :

4 Permutation based methods

$$LRmean_{k,i,j} = LC_i + RC_j$$

k is the k-th ligand-receptor interaction
L - expression of ligand L
R - expression of receptor R
C - cell cluster
i - cell group i
j - cell group j

4 Dot product based methods

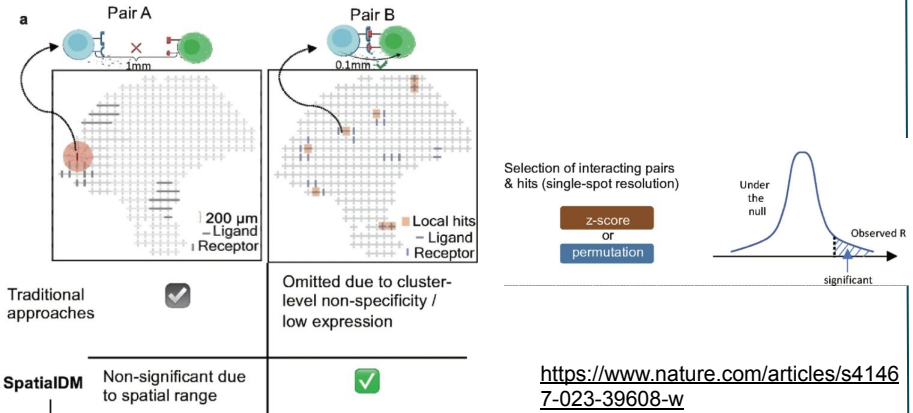
$$LRproduct_{k,i,j} = L_{C_i} R_{C_j}$$

Rank Aggregate Method

Liana :
Lig - Receptor
Lig - Receptor
Lig - Receptor
Lig - Receptor
Lig - Receptor
Lig - Receptor

Spatial methods for CCC : SpatialDM, Commot

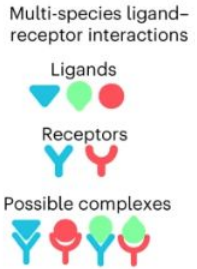
SpatialDM : spatial co-expression



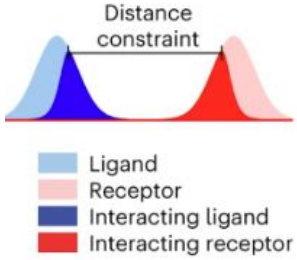
https://www.nature.com/articles/s41467-023-39608-w

Commot : Collective Optimal transport

Multiple ligand receptor interactions

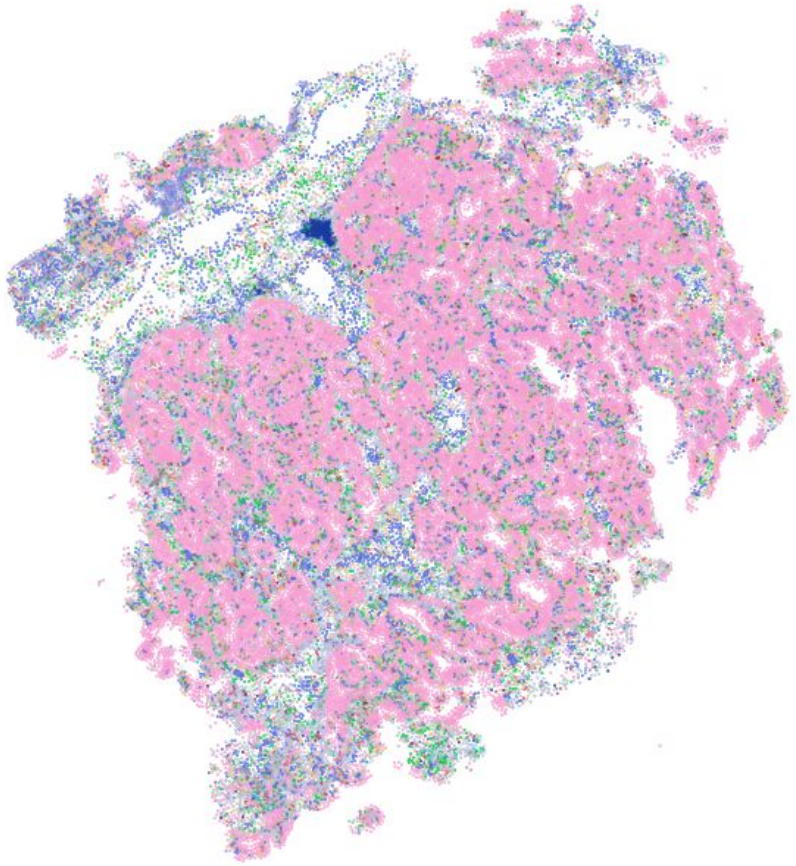


Spatial constraints



https://www.nature.com/articles/s41592-022-01728-4

What do these tissue samples look like ?



Robust decomposition of cell type mixtures in spatial transcriptomics

Dylan M. Cable, Evan Murray, Luli S. Zou, Aleksandrina Goeva, Evan Z. Macosko, Fei Chen & Rafael A. Irizarry

- B_cell
- B_plasma
- DC_1
- DC_2
- DC_activated
- DC_pc
- Endothelia_vascular
- Epi_lung
- Fibroblast
- Granulocyte
- Macrophage
- Mast_cell
- Monocyte
- Muscle_smooth
- NK
- Neutrophil
- Pericyte
- TNK_dividing
- T_CD4
- T_CD8_exhausted
- T_CTL
- T_CXCL13
- T_reg
- Tu_L1

← Spatial coordinate of each cell

Data set

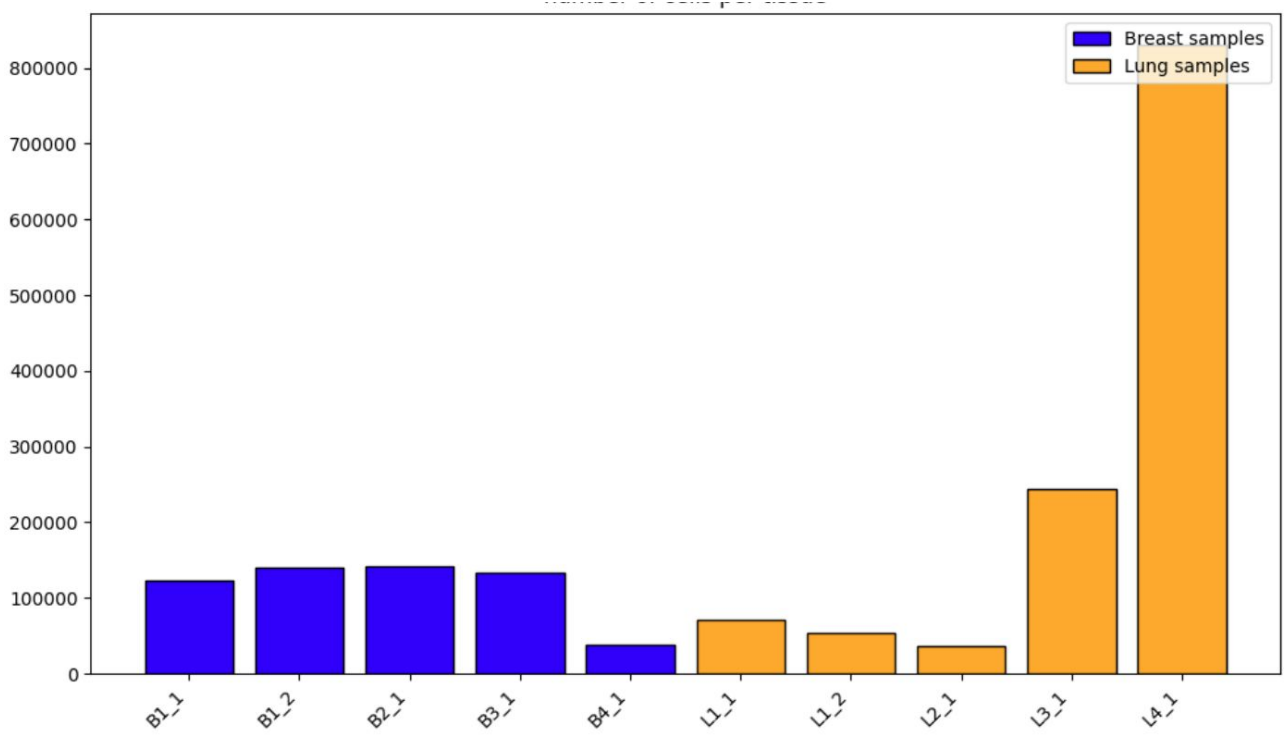
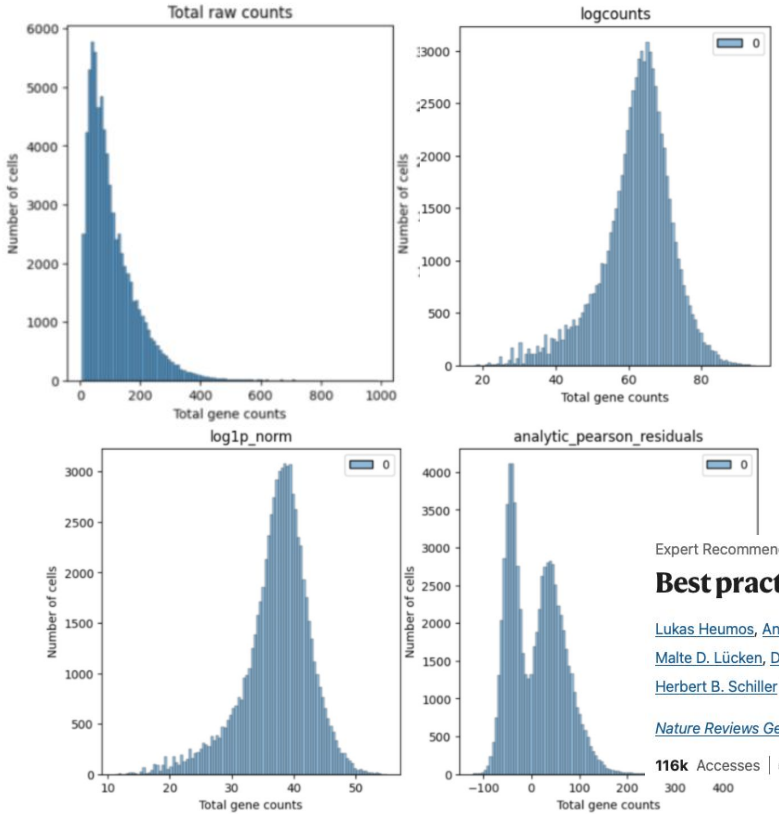


FIGURE 2 – Number of cells per tissue

Normalising RNA count



Expert Recommendation | Published: 31 March 2023

Best practices for single-cell analysis across modalities

Lukas Heumos, Anna C. Schaar, Christopher Lance, Anastasia Litinetskaya, Felix Drost, Luke Zappia, Malte D. Lücken, Daniel C. Strobl, Juan Henao, Fabiola Curion, Single-cell Best Practices Consortium, Herbert B. Schiller & Fabian J. Theis

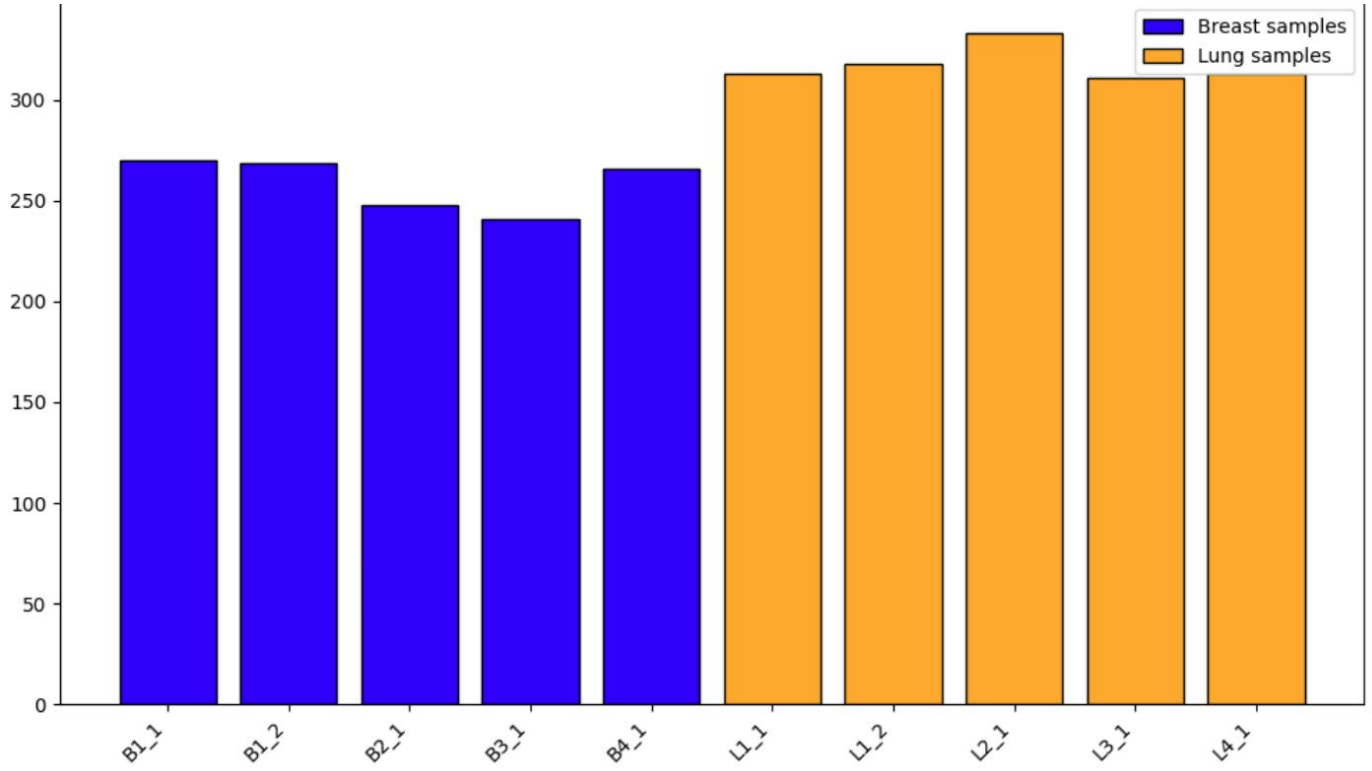
Nature Reviews Genetics **24**, 550–572 (2023) | Cite this article

116k Accesses | **56** Citations | **337** Altmetric | Metrics

FIGURE 4 – RNA counts distribution of Lung sample L1 *1* before and after different pre-processings

Data set

Genes per tissue after Quality control



RESULTS : How many Ligand receptors are significant in each tissue ?

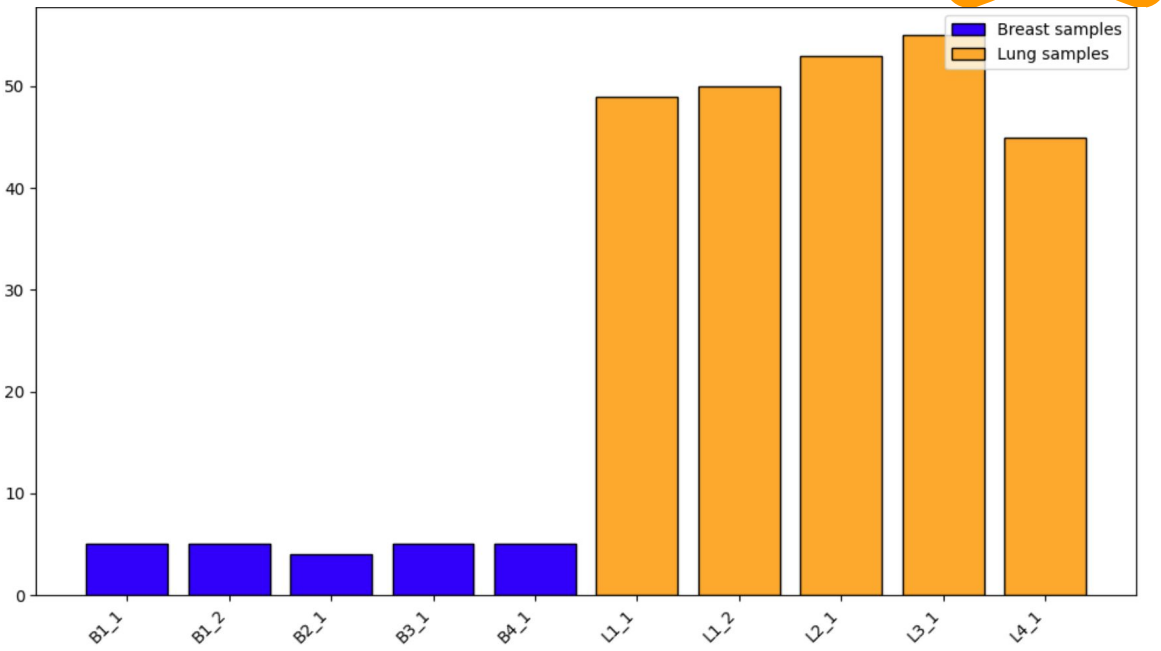
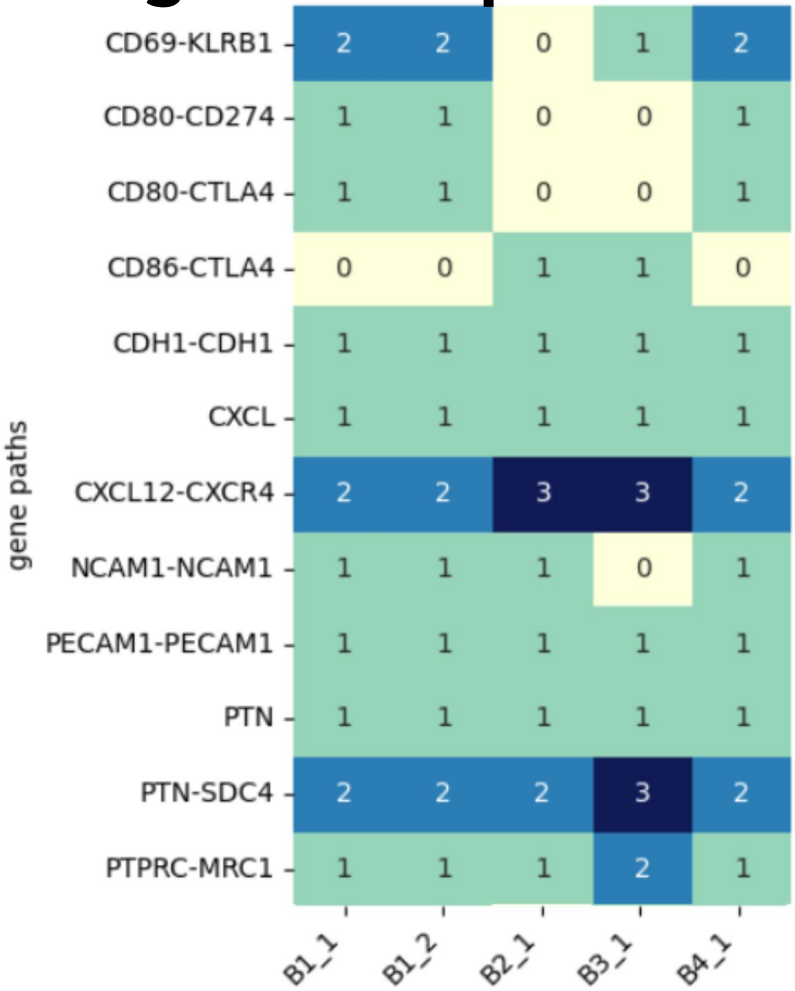
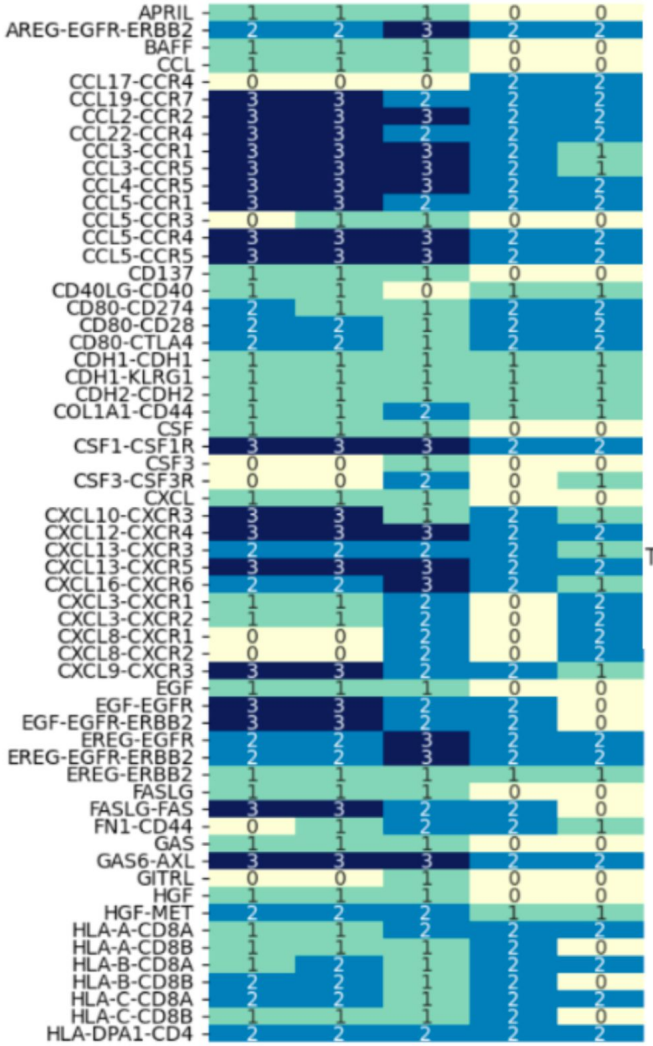
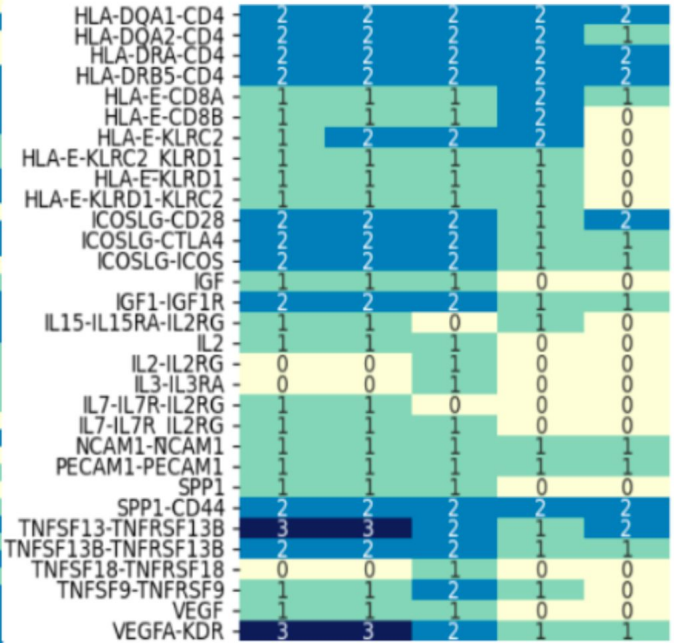


FIGURE 14 – Number of significant ligand - receptor pairs retrieved with CCC methods, per sample

Significant Ligand receptors across Breast tissues



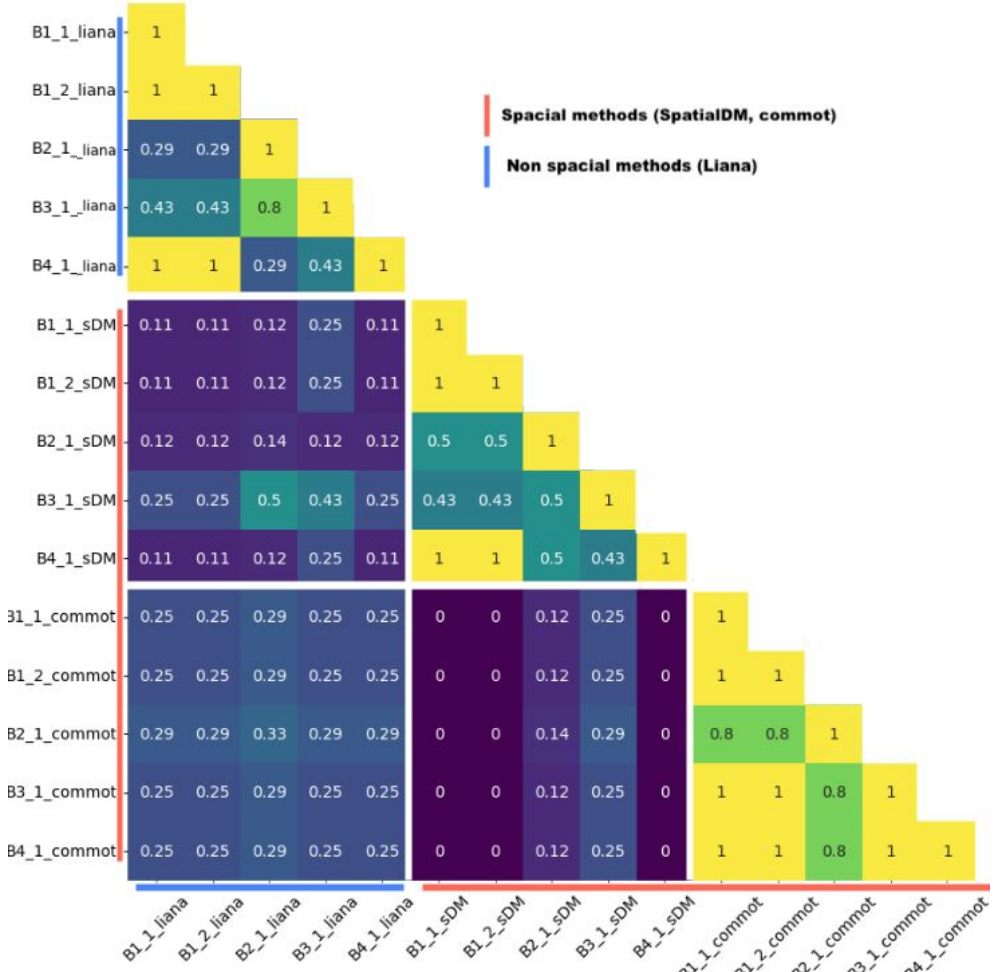


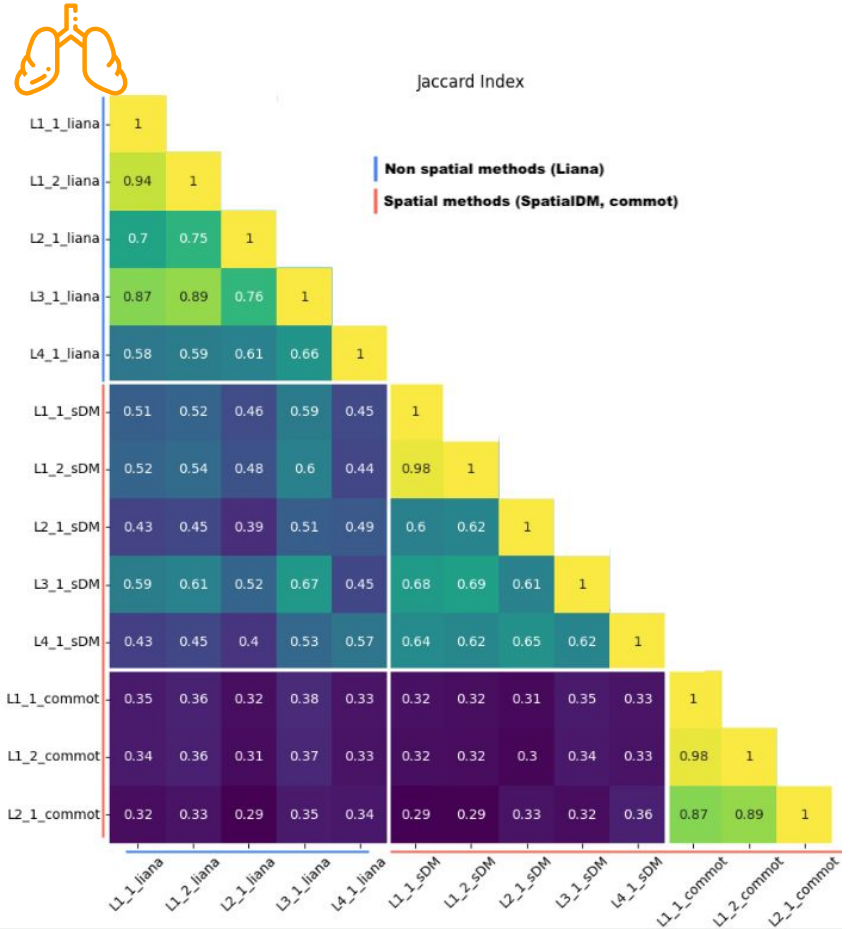


Significant Ligand receptors across Lung tissues

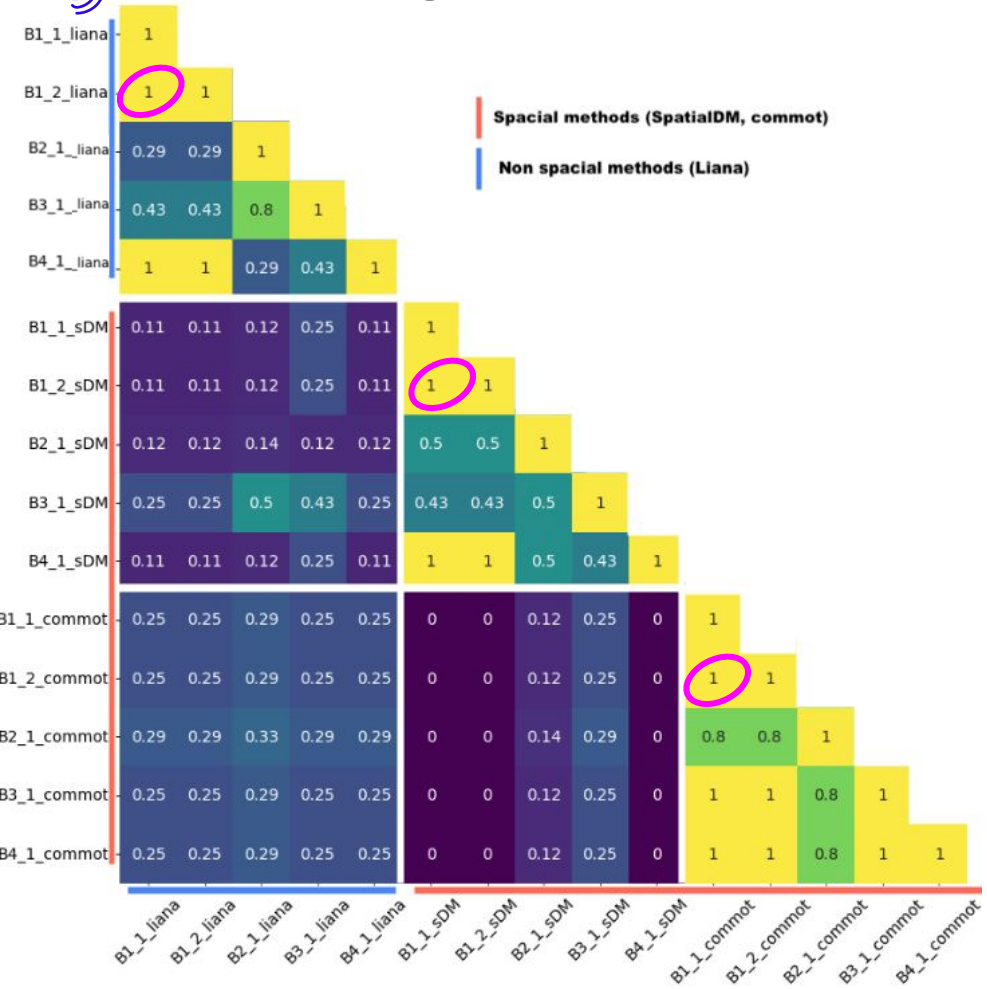
Comparing LR lists across tissues

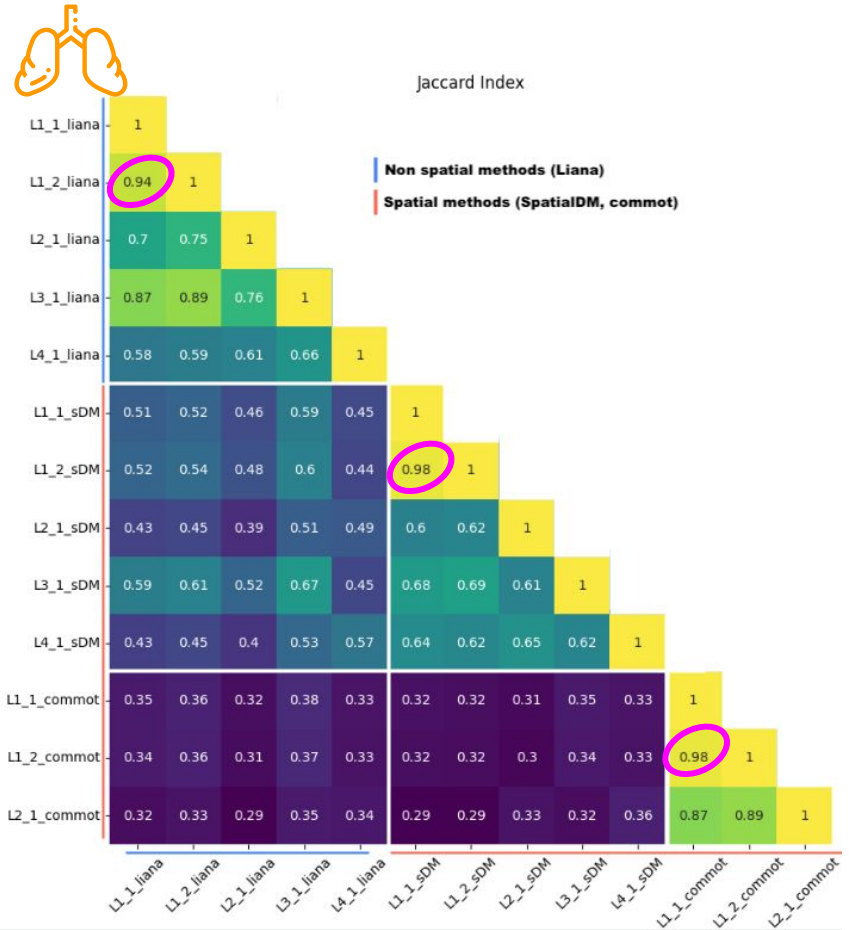
$$\text{Jaccard Index} = \frac{|A \cap B|}{|A \cup B|}$$



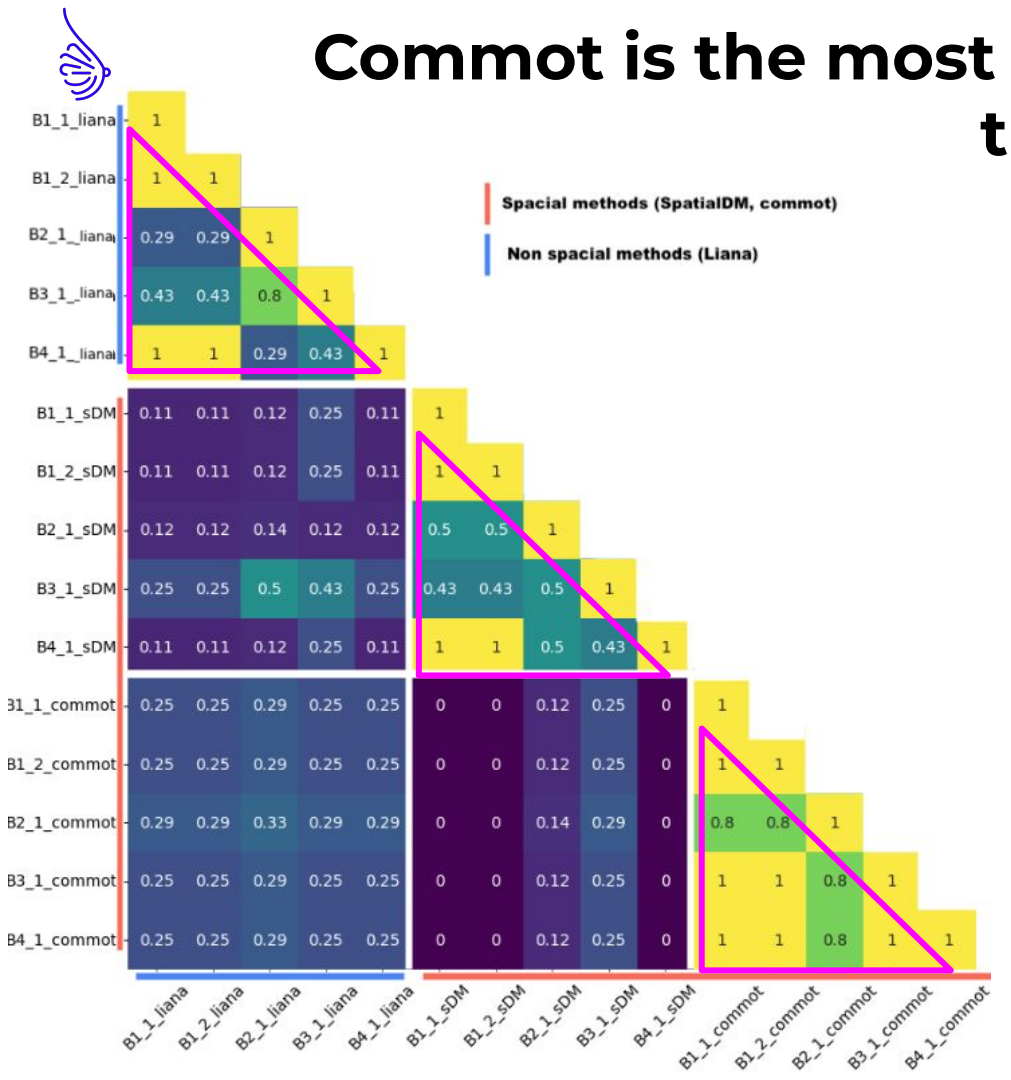


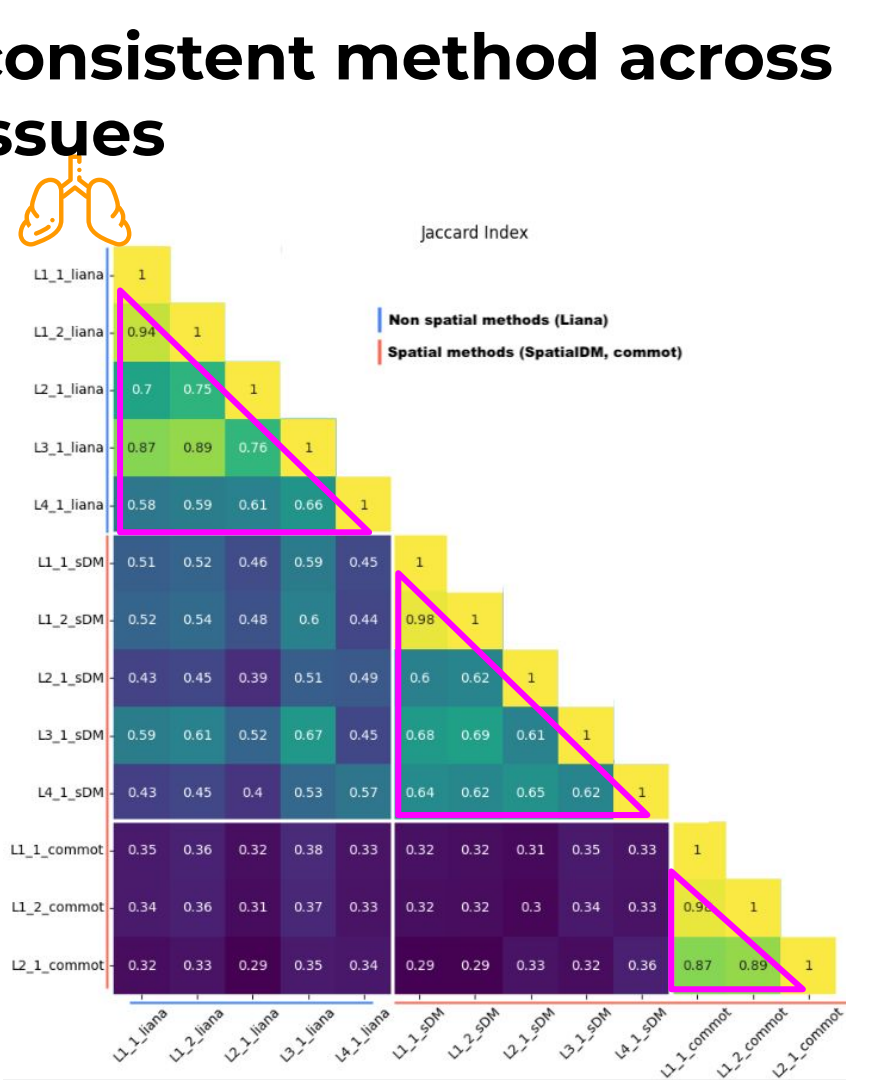
Strong robustness across replicates L1 *1* - L1 2, B1 *1* - LB1 *2*



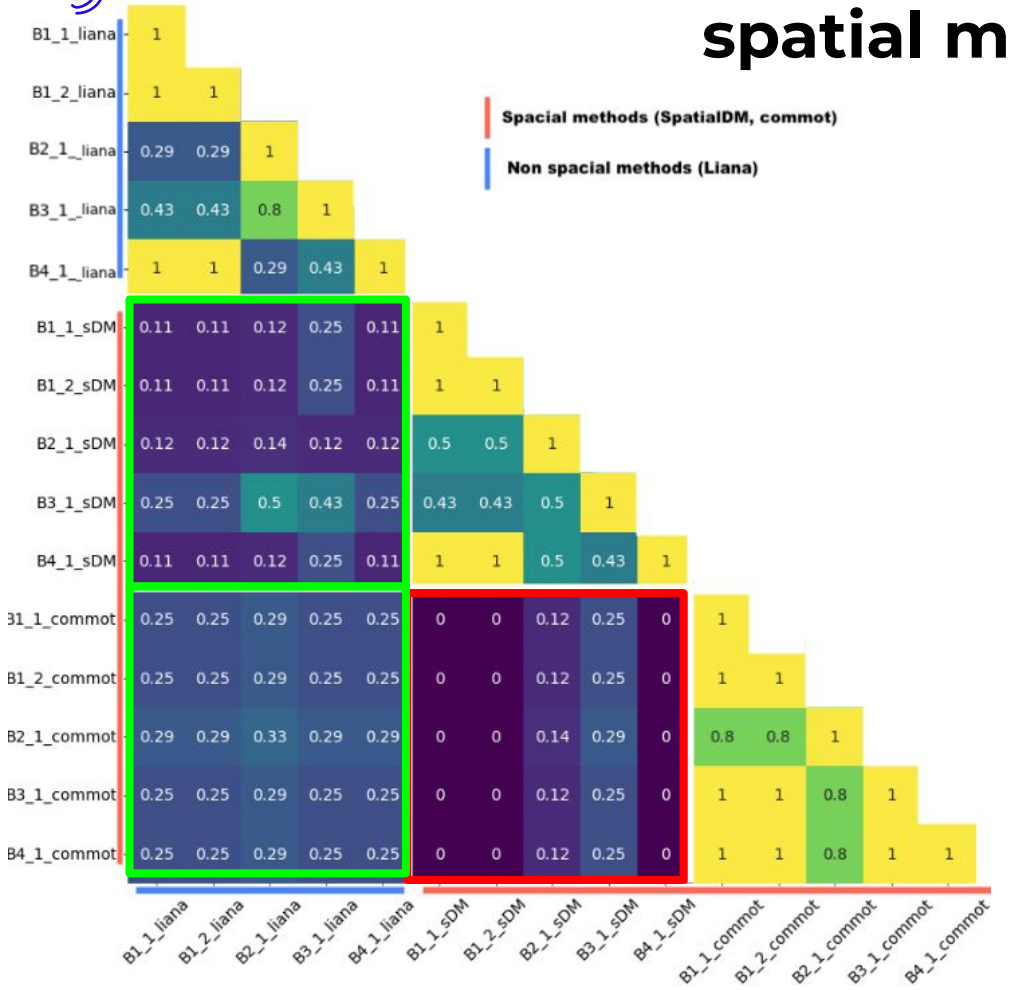


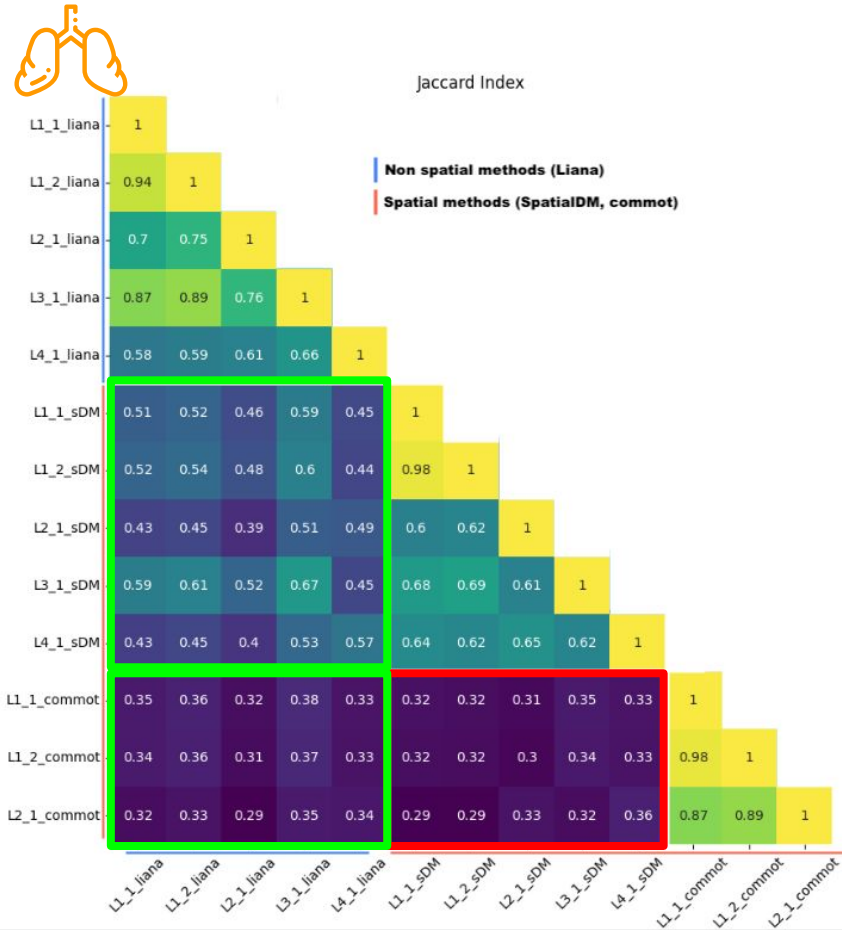
Commot is the most consistent method across tissues



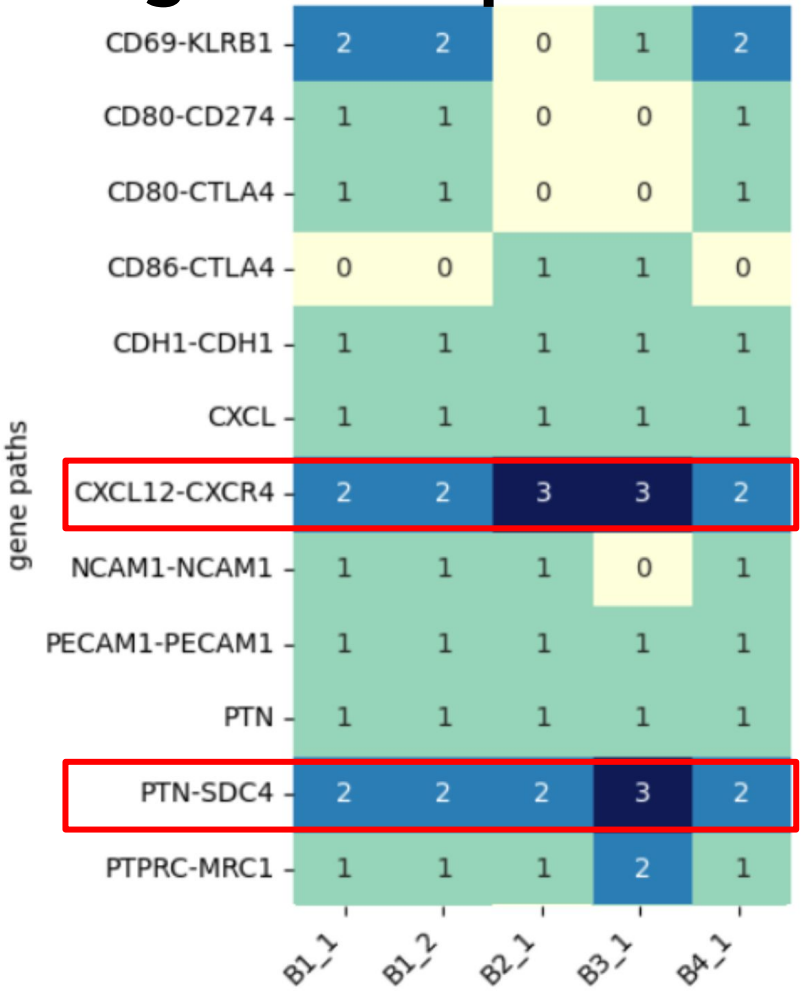


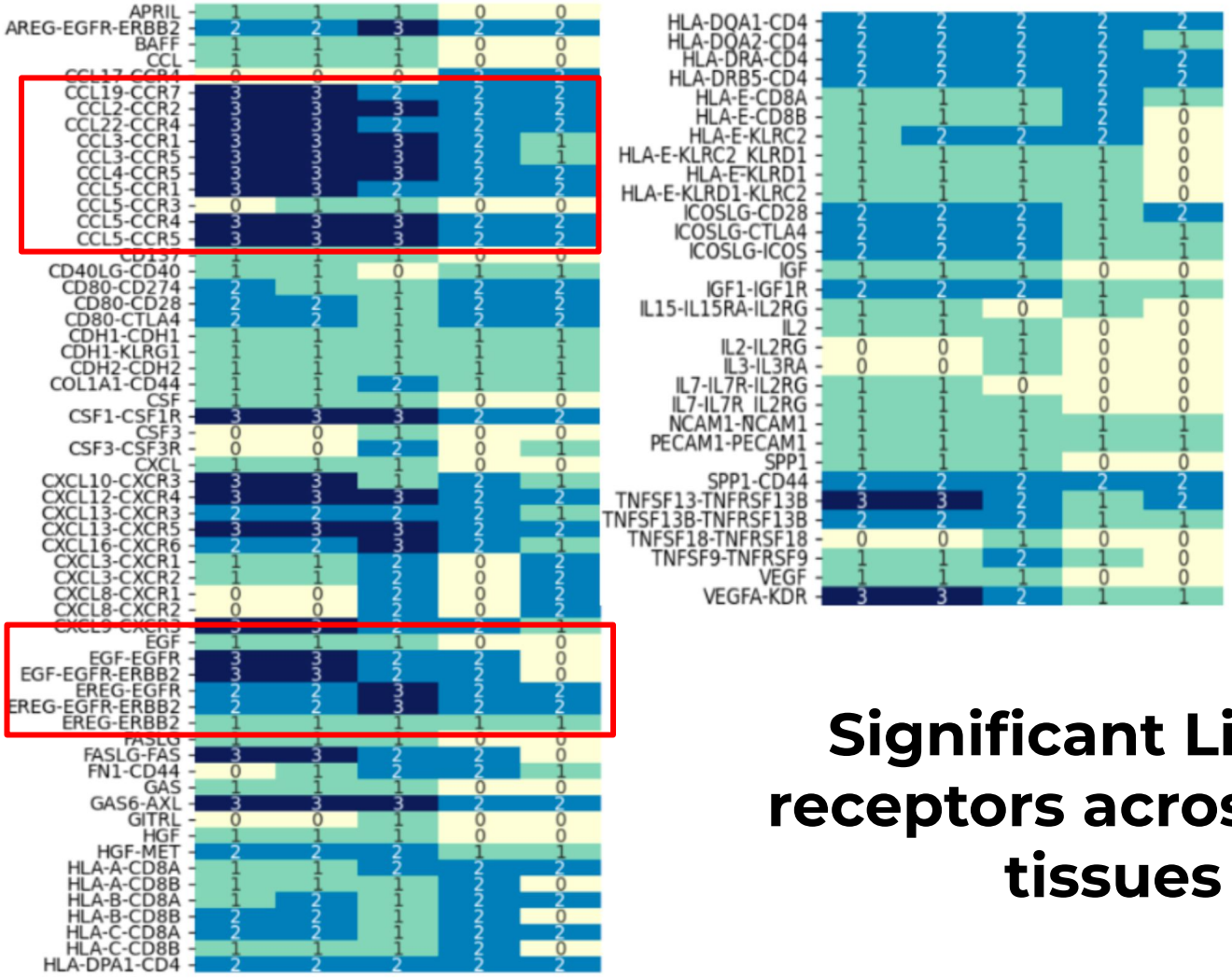
Little similarity across methods, even less across spatial methods.





Significant Ligand receptors across Breast tissues





Significant Ligand receptors across Lung tissues

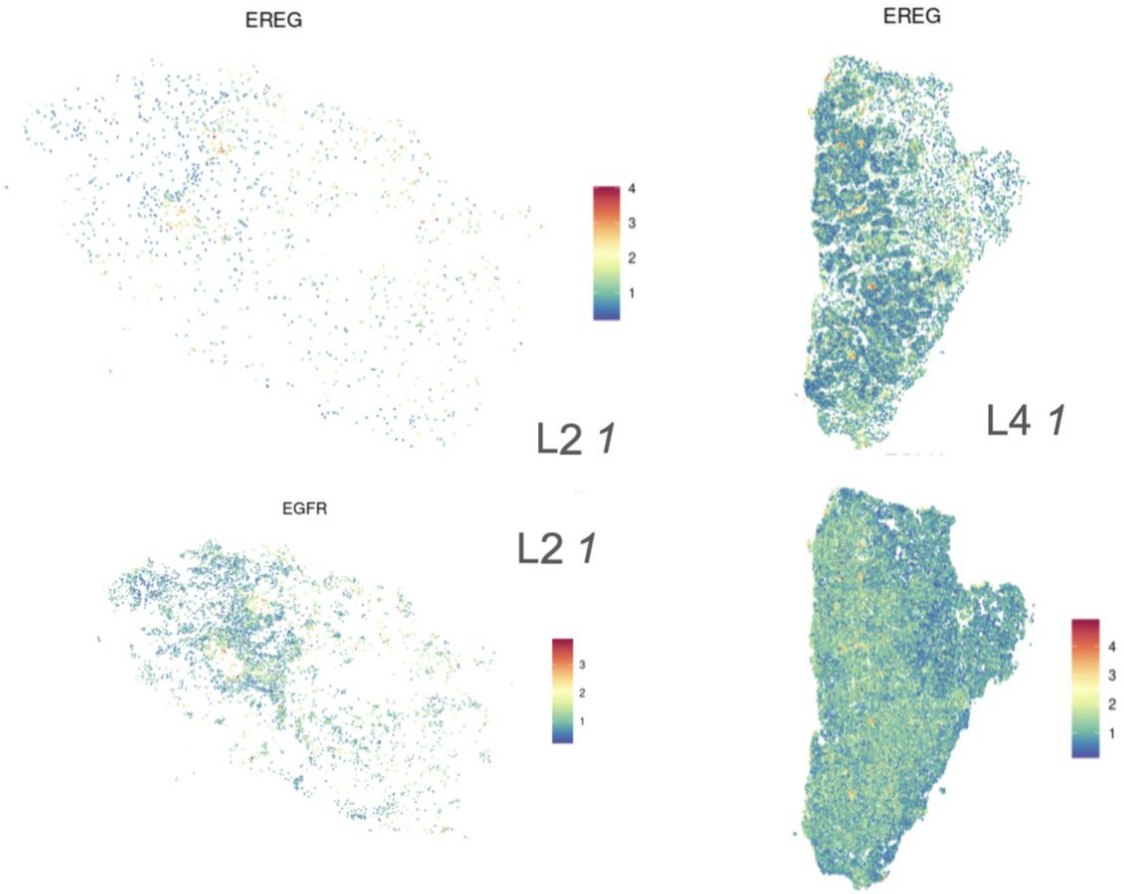


FIGURE 19 – Spatial location of RNA for EREG - EGFR ligand-receptors in samples L2, L4

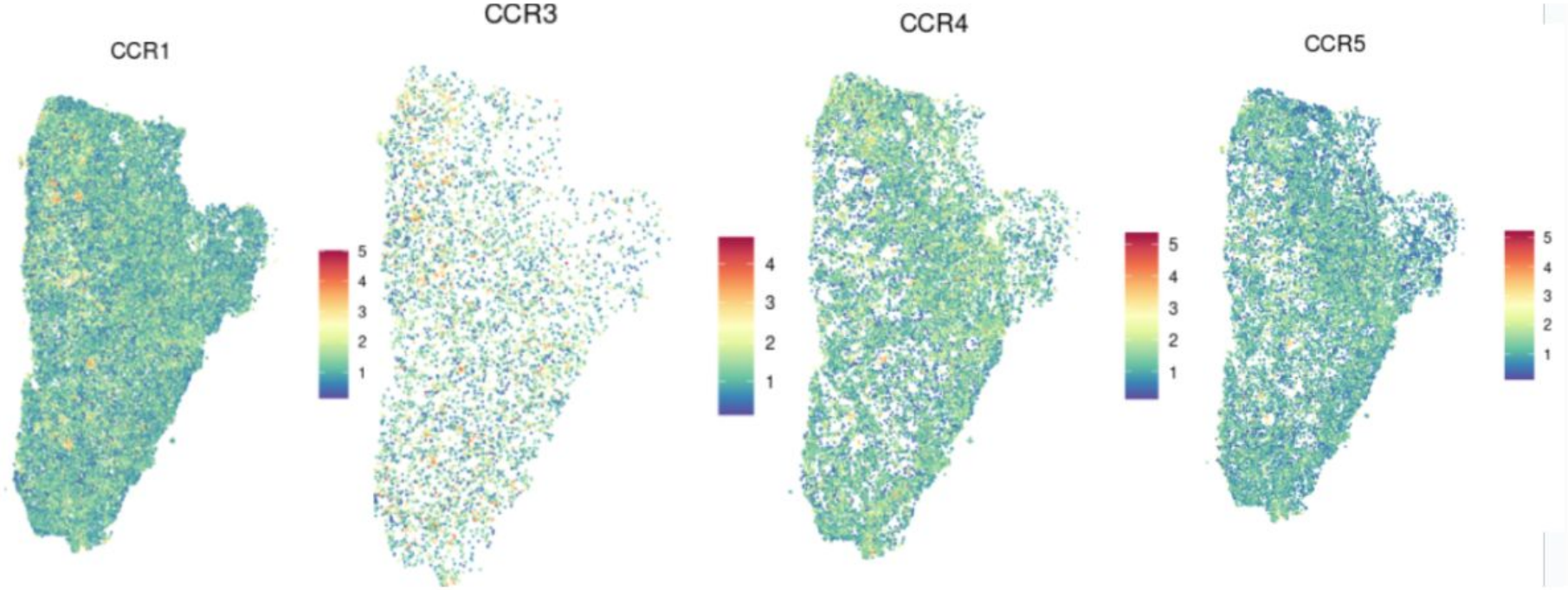


FIGURE 21 – Spatial location of RNA for CCR receptors in sample L4 *1*

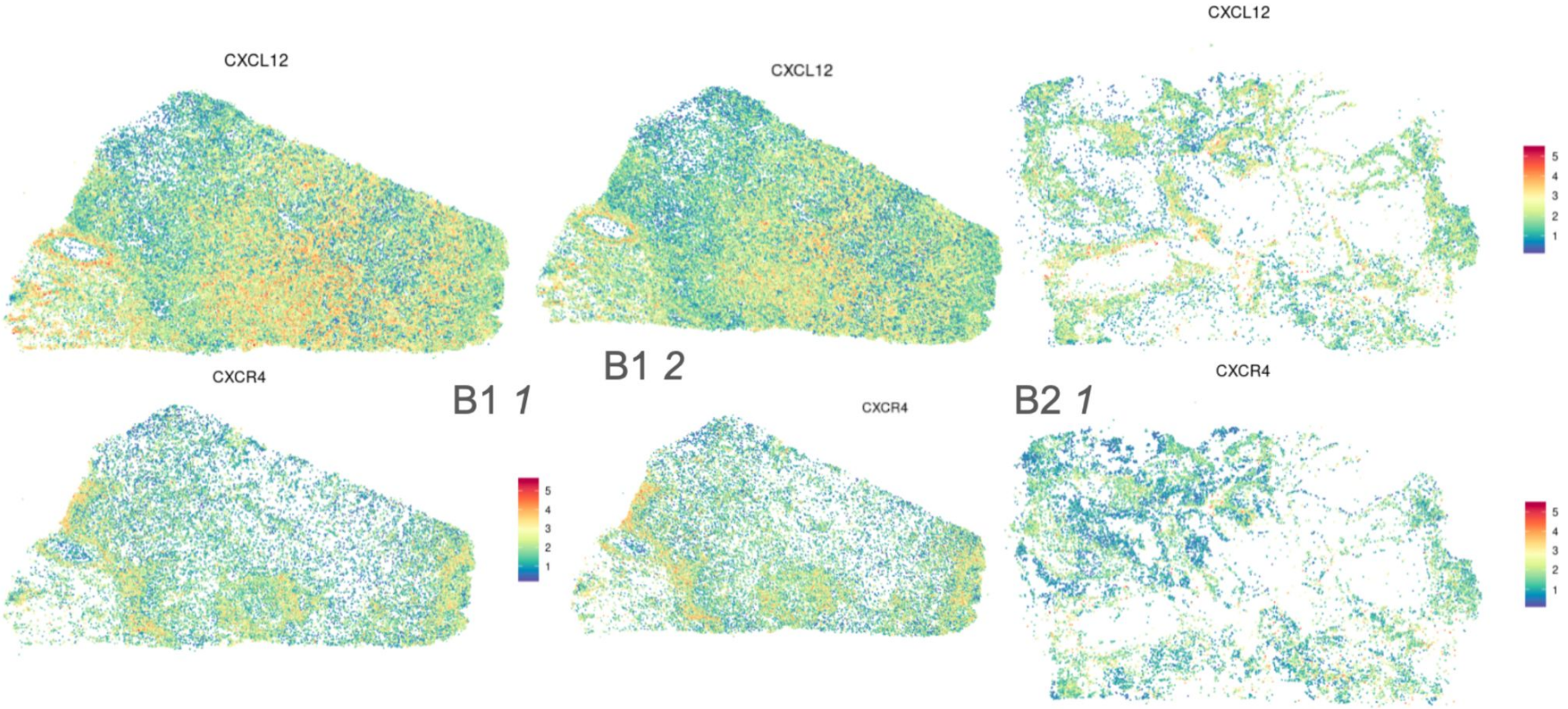
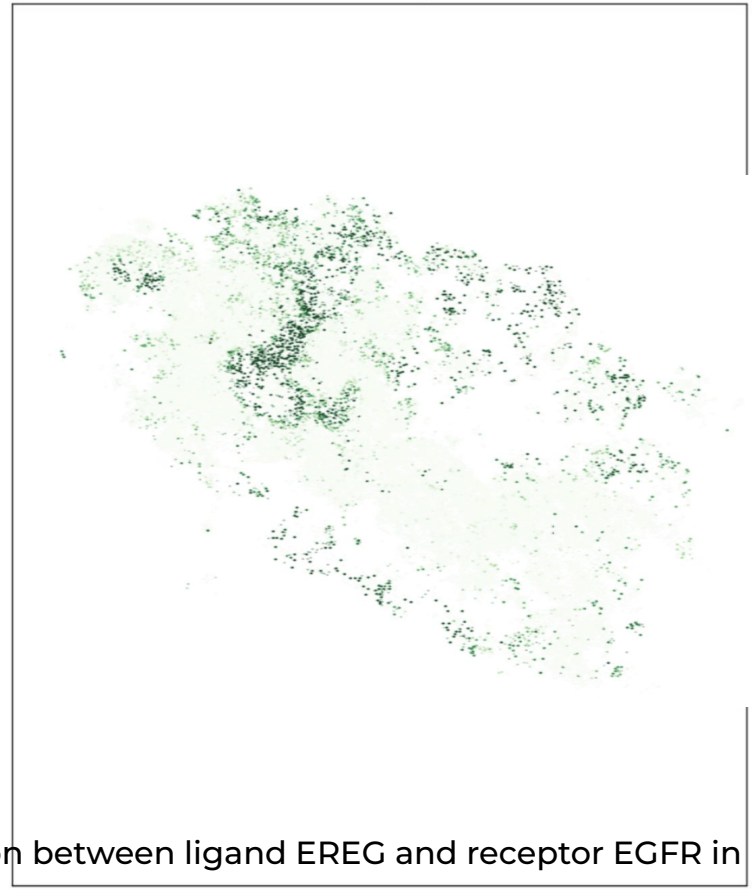


FIGURE 22 – Spatial location of RNA for CXCL12 - CXCR4 in samples B1 *1*, B1 *2*, B2

Measuring spatial correlation of Ligand - receptors



Spatial correlation between ligand EREG and receptor EGFR in Lung sample L2

Global Moran's I :

H0 : no spatial correlation between Ligand - receptor
H1 : spatial correlation

ligand receptor	tissue	Z p-value	Z
EREG_EGFR	L2	<0.001	62.03
	L4	<0.001	1053.37
CXCL12_CXCR4	B1 1	>0.05	
	B1 2	>0.05	
	B2	<0.001	390.96

Z > 0 : positive correlation between ligand receptor
Z < 0 : negative correlation between ligand receptor

$$\text{Global Moran's R} = \frac{N}{W}\sum_i \sum_j w_{ij}\,\tilde{x}_i\,\tilde{y}_j$$

$$\text{Local Moran's R} = \sum_j w_{ij}\,\tilde{x}_i\,\tilde{y}_j + \sum_j w_{ij}\,\tilde{y}_i\,\tilde{x}_j$$

What are the roles of these Ligand , receptors ?

- *CXCL12 chemokine and it's receptor*
- *PTN - SDC4*

- *EGF - EGFR*
- *CCR1, CCR3, CCR4, CCR5*

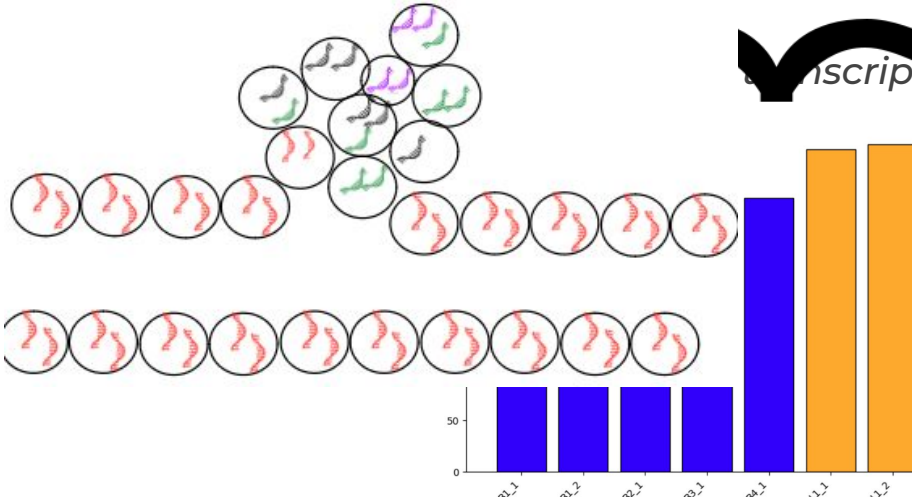
Uncontrolled cell growth, inflammation, angiogenesis, cell adhesion

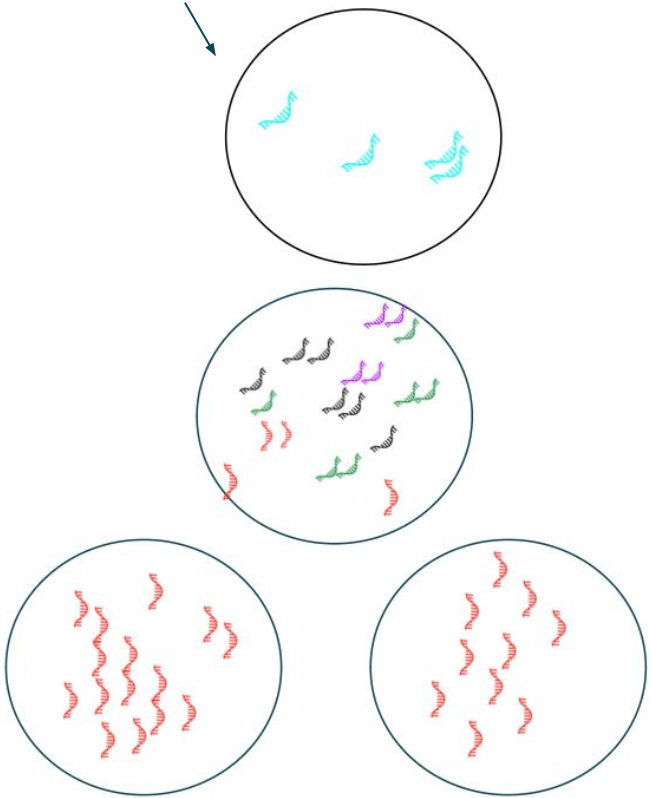
Limits in comparing cell-cell communication methods

→ *Bias of applying methods made for a lower resolution (Visium) to a higher resolution (Xenium)*

ls for map

nscrip





Conclusion

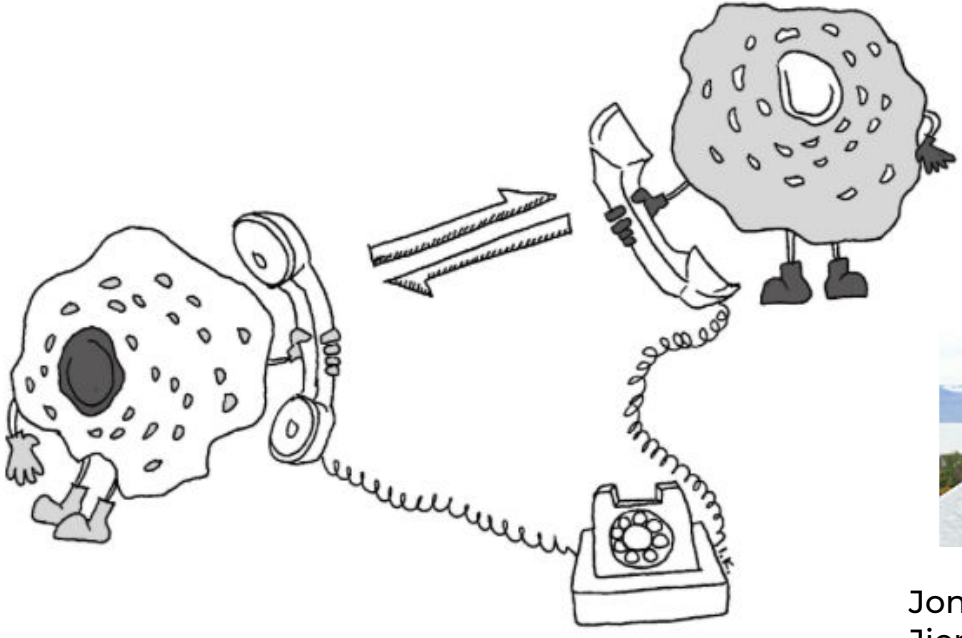
Spatial transcriptomics → deeper understanding of cell-cell communication in the tumour microenvironment :

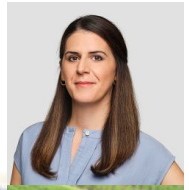
- *identify significantly expressed ligand receptor pairs, related to cancer in breast and lung tissues*
- *verify their proximal location in the sample.*

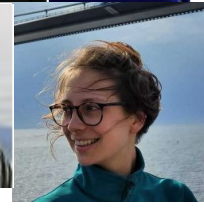
Absence of ground truth.

Possibility of developing a synthetic dataset of cells, along with their ligand-receptor interactions and benchmark methods against this dataset .

Thank you







Jonathan,
Jieran

*Irene Kim; Staff Photographer **Yale News***
irene.kim@yale.edu

Non spatial method for CCC :

Liana - ReadTheDocs
https://ccc-protocols.readthedocs.io/en/latest/notebooks/ccc_python/02-Infer-Communication-Scores.html

Permutation between L - R

CellPhoneDBv2

Magnitude:

$$LRmean_{k,i,j} = \frac{L_{C_i} + R_{C_j}}{2}$$

Specificity: CellPhoneDBv2 introduced a **permutation approach** also adapted by other methods, see permutation formulation below.

Geometric Mean

Magnitude:

$$LRgeometric.mean_{k,i,j} = \sqrt{L_{C_i} \cdot R_{C_j}}$$

Specificity: An adaptation of CellPhoneDBv2's permutation approach.

Dot product

Connectome

Magnitude:

$$LRproduct_{k,i,j} = L_{C_i} R_{C_j}$$

Specificity:

$$LRz.mean_{k,i,j} = \frac{z_{L_{C_i}} + z_{R_{C_j}}}{2}$$

where `z` is the z-score of the expression matrix `M`:

$$X_z = (X - mean(X))/std(X)$$

log2FC

Specificity:

$$LRlog2FC_{k,i,j} = \frac{\text{Log2FC}_{C_i,L} + \text{Log2FC}_{C_j,R}}{2}$$

SingleCellSignalR

Magnitude:

$$LRscore_{k,i,j} = \frac{\sqrt{L_{C_i} R_{C_j}}}{\sqrt{L_{C_i} R_{C_j}} + \mu}$$

where `mu` is the mean of the expression matrix `M`

NATMI

Magnitude:

$$LRproduct_{k,i,j} = L_{C_i} R_{C_j}$$

Specificity:

$$SpecificityWeight_{k,i,j} = \frac{L_{C_i}}{\sum^n L_{C_i}} \cdot \frac{R_{C_j}}{\sum^n R_{C_j}}$$

Spatial method for CCC : SpatialDM

SpatialDM for rapid identification of s expressed ligand–receptor and reveal communication patterns

Zhuoxuan Li, Tianjie Wang, Pentao Liu & Yuanhua Huang

Nature Communications **14**, Article number: 3995 (2023) | Cite this

Spatial method for CCC : Commot

Given a spatial transcriptomics dataset of n_s cells or spots and n_l ligand species and n_r receptor species, the collective optimal transport determines an optimal multi-species coupling $\boldsymbol{P}^* \in \mathbb{R}_+^{n_l \times n_r \times n_s \times n_s}$ where $\boldsymbol{P}^*_{i,j,k,l}$ scores the signaling strength from sender cell k to receiver cell l through ligand i and receptor j. This is achieved by solving a minimization problem, $\min_{\boldsymbol{P} \in \Gamma} \sum_{(i,j) \in I} \alpha_{(i,j)} \left\langle \boldsymbol{P}_{i,j,\cdot,\cdot}, \boldsymbol{C}_{(i,j)} \right\rangle_F$ where

$$\Gamma = \{\boldsymbol{P} \in \mathbb{R}_+^{n_l \times n_r \times n_s \times n_s} : \boldsymbol{P}_{i,j,\cdot,\cdot} = 0 \text{ for } (i,j) \notin I, \\ \sum_{j,l} \boldsymbol{P}_{i,j,k,l} \leq \boldsymbol{X}_{i,k}, \sum_{i,k} \boldsymbol{P}_{i,j,k,l} \leq \boldsymbol{X}_{j,l}\},$$

I is the index set for ligand and receptor species that can bind together, and $\boldsymbol{X}_{i,k}$ is the expression level of gene i on spot k. The species-specific cost matrix $\boldsymbol{C}_{(i,j)}$ is a modified distance matrix for between-spot distance that replaces distances exceeding the spatial range of ligand i by infinity. The competitions between molecule species and cells are considered by assuming that a given receptor species or cell has limited capacity for interactions, such that a stronger inferred interaction with one ligand species or cell reduces the potential of interaction with other ligand species or cells (see the Methods and Supplementary Note for detailed formulations and algorithm derivations).

Screening cell–cell communication in spatial transcriptomics via collective optimal transport

Zixuan Cang, Yanxiang Zhao, Axel A. Almet, Adam Stabell, Raul Ramos, Maksim V. Plikus, Sc
Atwood & Qing Nie ✉

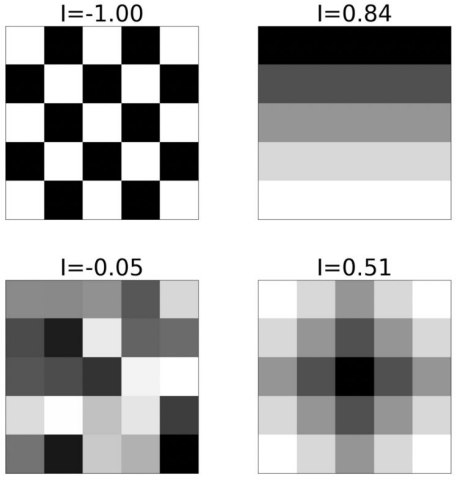
Nature Methods **20**, 218–228 (2023) | Cite this article

Global Moran's *I* is a measure of the overall clustering of the spatial data. It is defined as

$$I = \frac{N}{W} \frac{\sum_{i=1}^{N} \sum_{j=1}^{N} w_{ij}(x_i - \bar{x})(x_j - \bar{x})}{\sum_{i=1}^{N} (x_i - \bar{x})^2}$$

where

- N is the number of spatial units indexed by i and j;
- x is the variable of interest;
- $\bar{x}$ is the mean of x;
- w_{ij} are the elements of a matrix of spatial weights with zeroes on the diagonal (i.e., $w_{ii} = 0$);
- and W is the sum of all w_{ij} (i.e. $W = \sum_{i=1}^{N} \sum_{j=1}^{N} w_{ij}$).



https://en.wikipedia.org/wiki/Moran%27s_I#/media/File:Moran's_I_example.png