

Studying cell-cell communication across transcriptomic modalities



By Daniel Herrero Saboya
21.09.2023

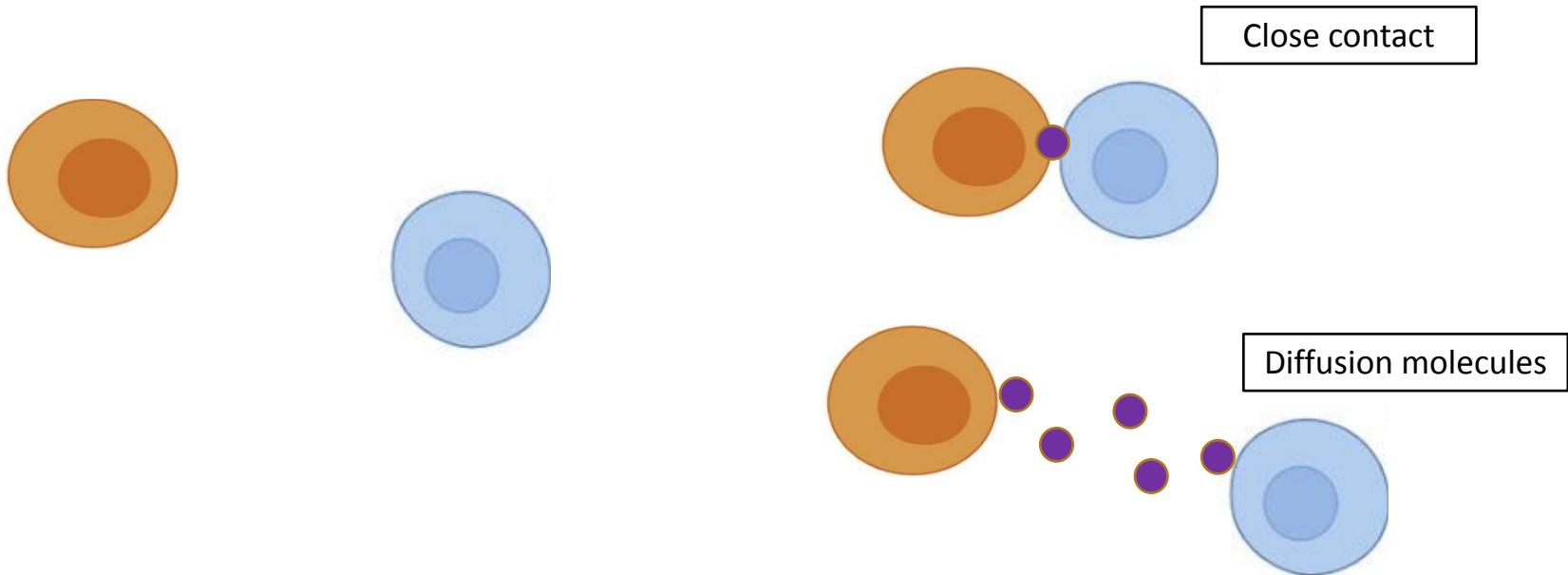
Cell-cell communication

Action in one cell that generates a change in the behaviour of another cell

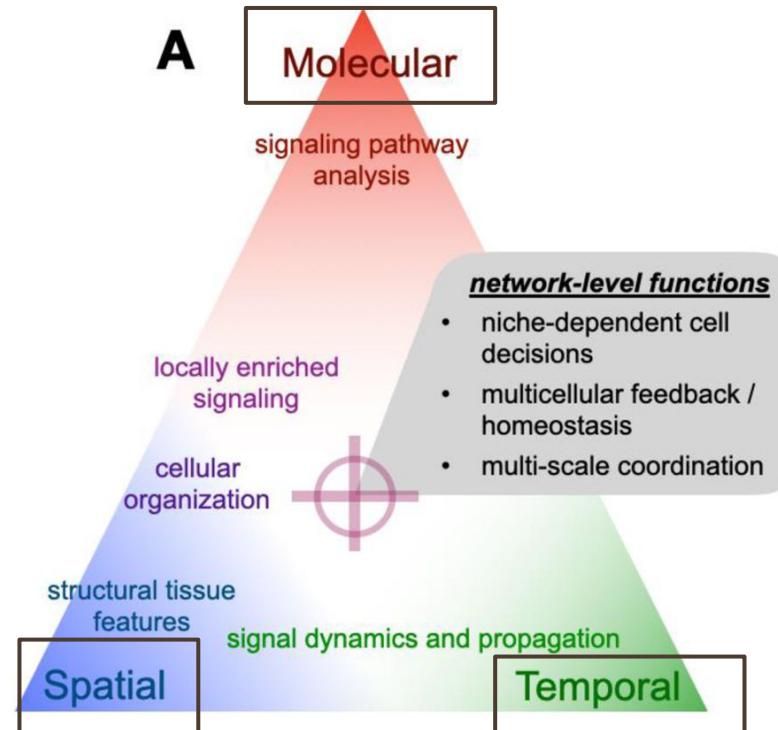


Cell-cell communication

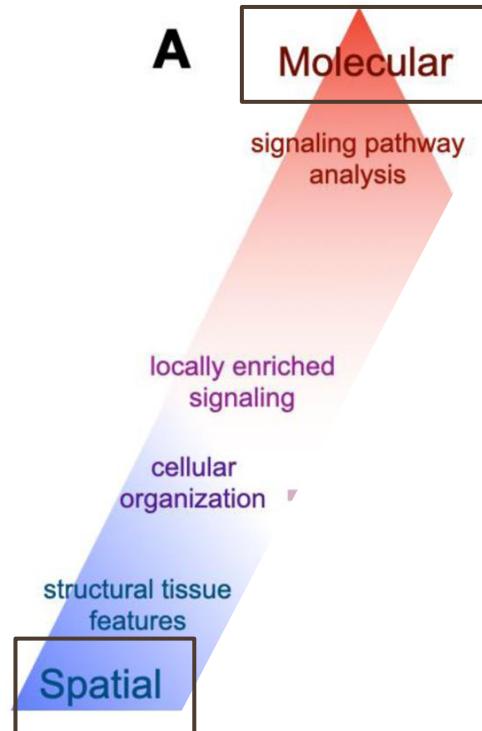
Action in one cell that generates a change in the behaviour of another cell



A holistic picture of CCC

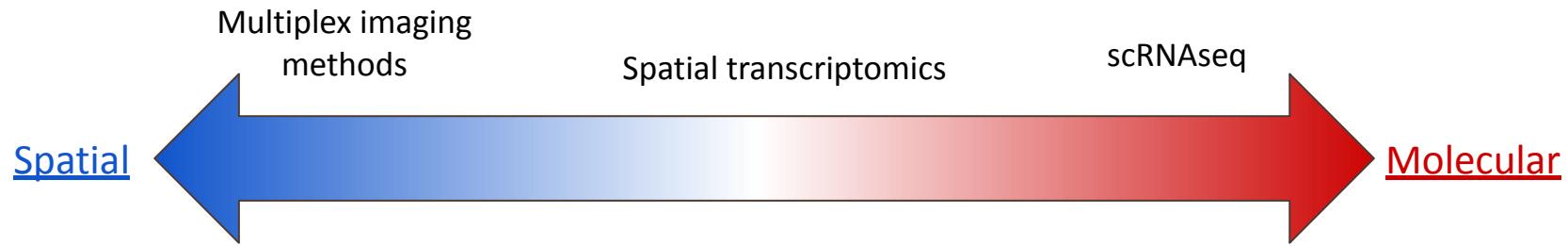


A holistic picture of CCC (in omics)



Single time-point omics omit one of the main axes of CCC:
Temporal/causal relations become harder to study →
Leverage on previous knowledge
Hypothesis-generating methods

The spectrum of CCC analysis



What is your surrogate?



Key idea: variety of assumptions for what CCC is generates a variety of surrogates

CCC is not a specific concept for which you need to develop a specific algorithm (e.g. cats vs dogs, 1y OS)

The way you define CCC will define the way you prioritize surrogates

Methods for imaging

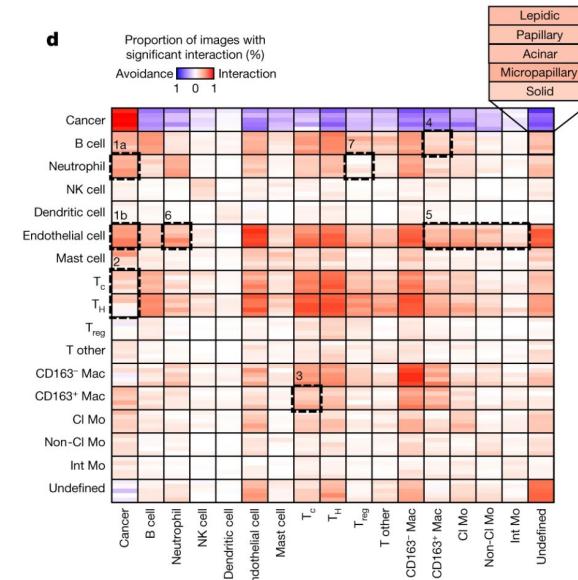
Key notion

Cells close (or far) in space are communicating (or avoiding communication)

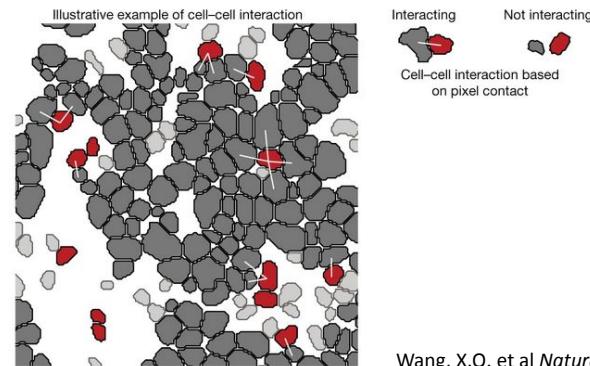
Methods for imaging

Neighborhood analysis

- How do you define neighbors/interacting cells?: graph, distance, pixels...



Sorin, M et al *Nature* 2023

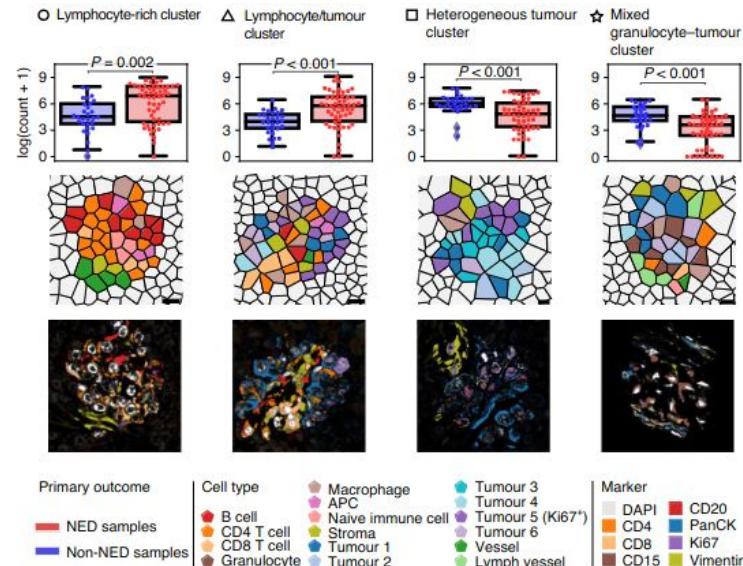


Wang, X.Q. et al *Nature* 2023

Methods for imaging

Neighborhood analysis

Niche analysis



Zhenqin Wu et al *Nat. Biomed. Eng* 2022

Methods for imaging

Neighborhood analysis

Niche analysis

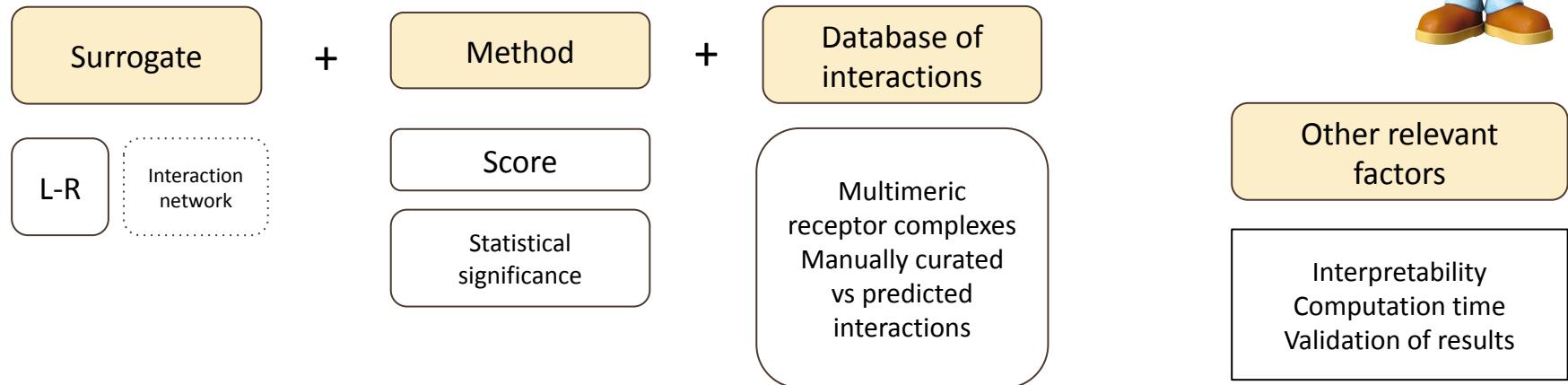
Infiltration/distance-based measures/gradients/...

Methods for single cell (& bulk) RNAseq

Key notion

Cells are more likely to be interacting if they exhibit stronger L-R signals

Build your own scRNAseq CCC model



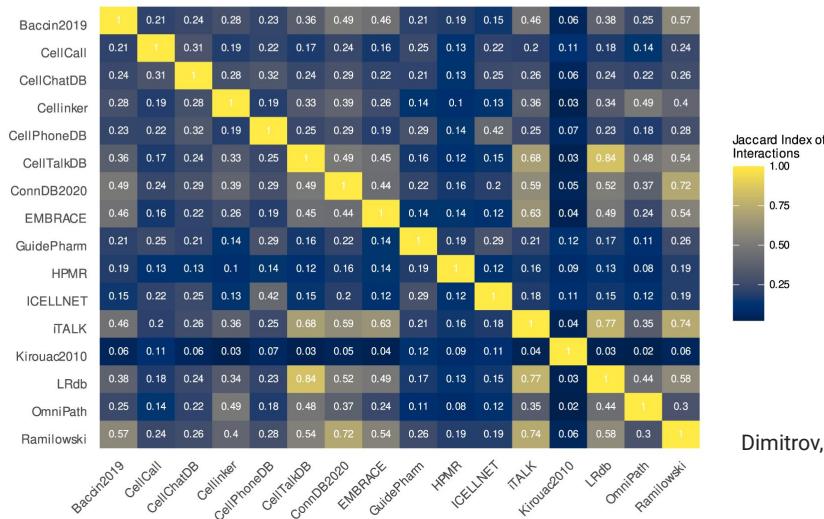
Benchmarking of methods

- 1) CCC are especially hard to validate → What do you validate?

Benchmarking of methods

- 1) CCC are especially hard to validate → What do you validate?
- 2) Results from Dimitrov et al
 - a) Practically no overlap in predictions because practically no overlap in interaction databases
 - i) *Not a straight-forward conclusion*

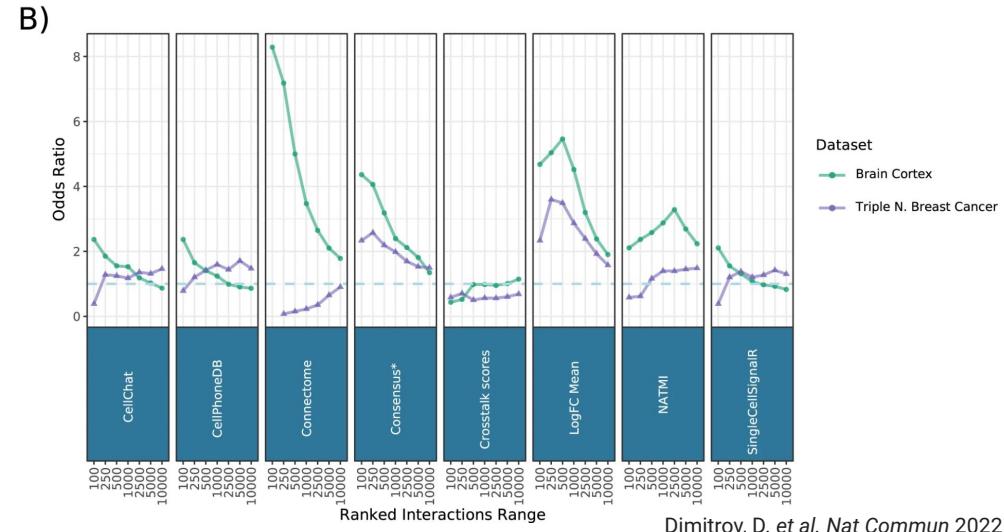
B)



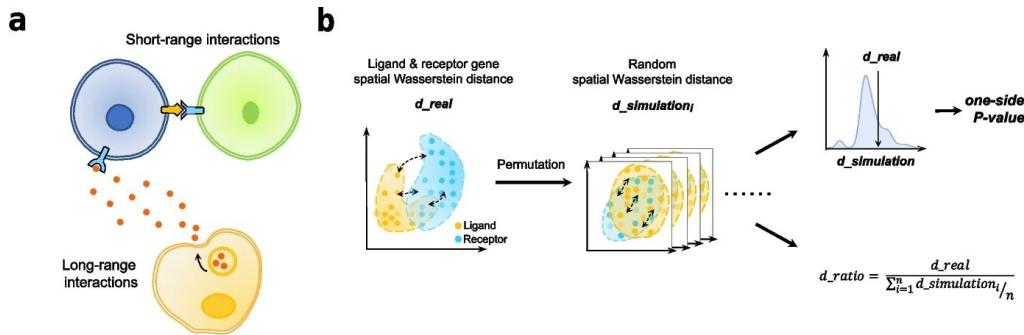
Dimitrov, D. et al. Nat Commun 2022

Benchmarking of methods

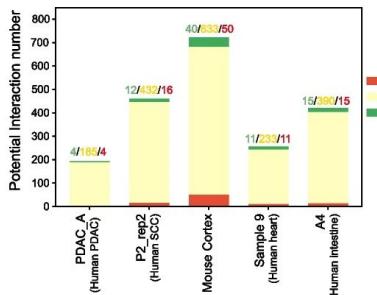
- 1) CCC are especially hard to validate → What do you validate?
- 2) Results from Dimitrov et al
 - a) Practically no overlap in predictions because practically no overlap in interaction databases
 - b) Generally, predicted CCC are neighbors in space



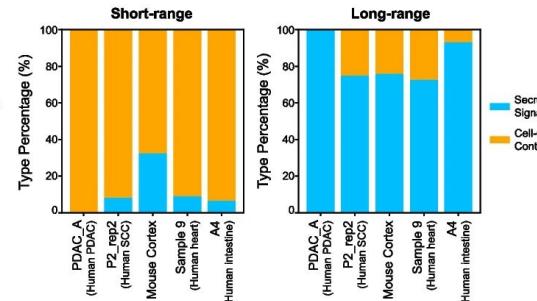
Common ground for validation



C Interaction Types & Number in each Sample



d Interaction Types Percentage



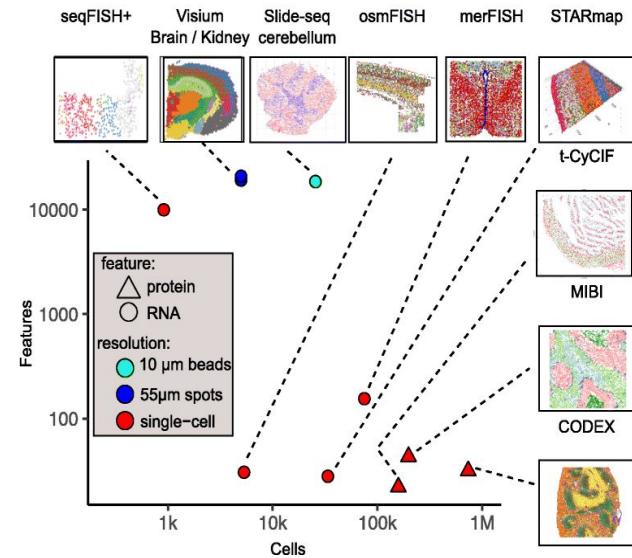
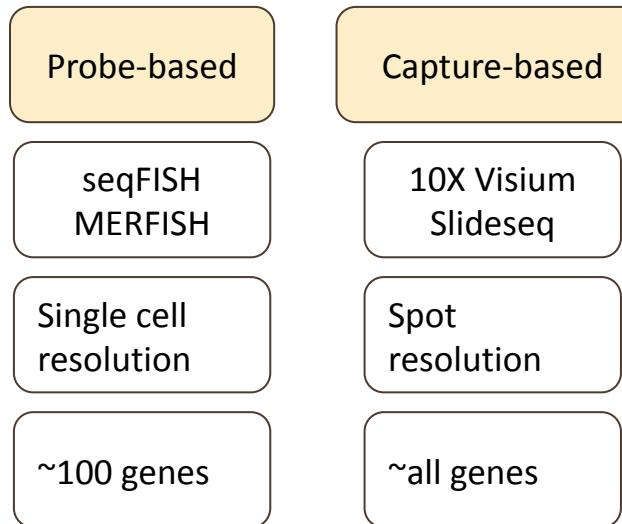
Methods for spatial transcriptomics

Particularly interesting because of 2 major innovations:

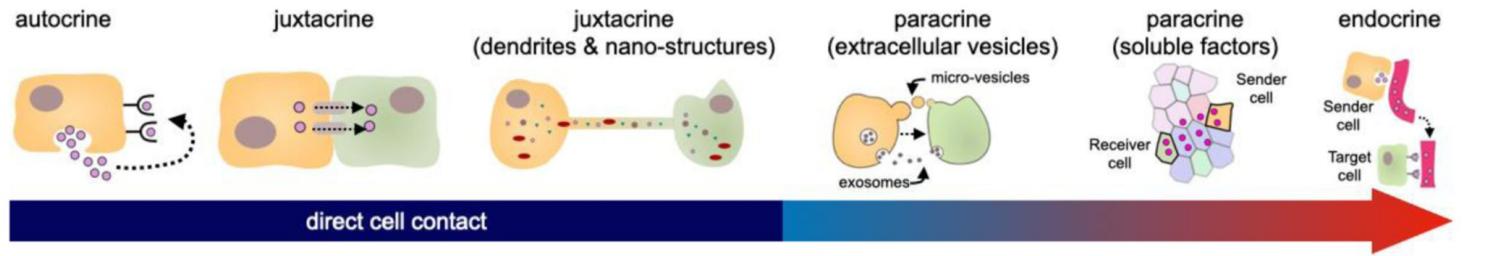
- 1) Combination of spatial and molecular components
- 2) Unsupervised/discovery (more in the molecular side)

Spatial transcriptomics technologies

Quantify mRNA expression of large numbers of genes within the spatial context of tissues and cells

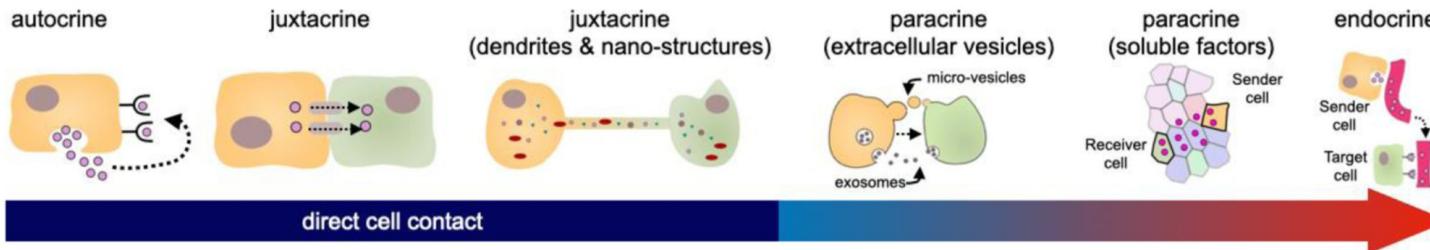


The spatial constraint

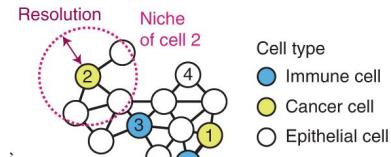


Gaurav Luthria et al *Curr Opin Syst Biol* 2021

The spatial constraint



Gaurav Luthria et al *Curr Opin Syst Biol* 2021

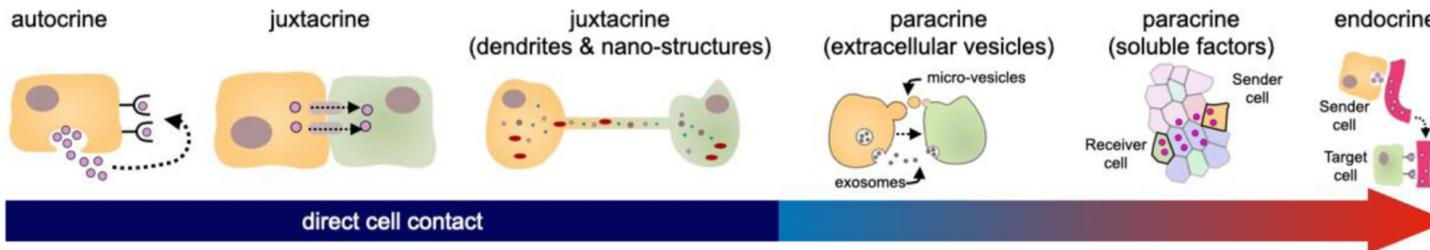


Fischer, D.S. et al. *Nat Biotechnol* 2022

$$K_{env} = \sigma_E^2 \exp\left(-d_{i,j}^2/2l^2\right)$$

Arnol, D. et al. *Cell Rep.* 2019

The spatial constraint



Gaurav Luthria et al *Curr Opin Syst Biol* 2021

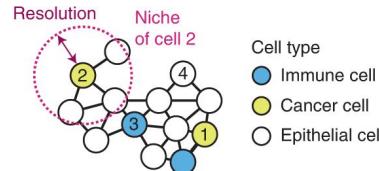
Hard



Soft



Thresholds are usually adopted in an arbitrary manner

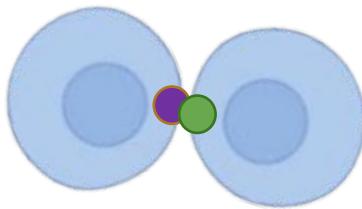


Fischer, D.S. et al. *Nat Biotechnol* 2022

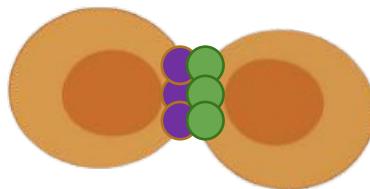
$$K_{env} = \sigma_E^2 \exp\left(-d_{i,j}^2/2l^2\right)$$

Arnol, D. et al. *Cell Rep.* 2019

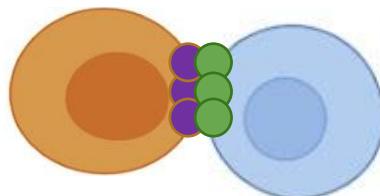
Supervised vs unsupervised



Supervised: genes X & Y are known → Evaluate
E.g. Score = $(X + Y)_{\text{close_cells}}$



Supervised vs unsupervised



Unsupervised: how is gene X affected by the environment → Model/Predict

E.g. 1) $X \sim Y_{\text{close_cells}}$

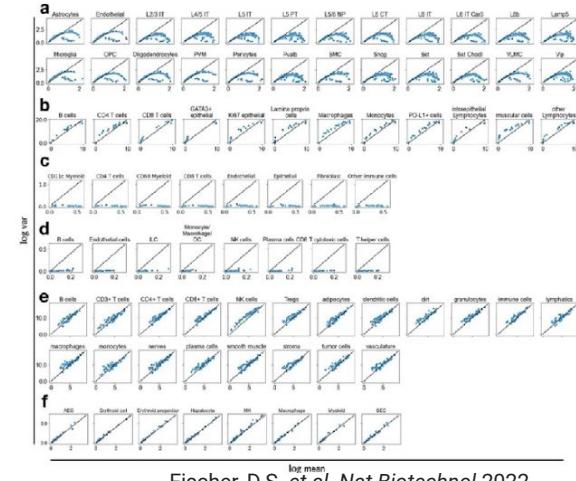
2) Look at model (e.g. accuracy in prediction)

Other considerations

Data type

seqFISH/MERFISH >>> 10X Visium > slideSeq/STARmap

However, each data type has different characteristics



Other considerations

Data type

Output

- Binary prediction

- Communication score

- Model to predict gene expression

- Statistical support?

Other considerations

Data type

Output

Multiple image integration

Barely addressed as image integration is a very recent problem

Other considerations

Data type

Output

Multiple image integration

Many other nice things

- Cell type-specific model

- Directionality

- Shape/flow/gradient

- Optimal communication distance

- Competition between ligands/receptors

Conclusions

CCC is:

- 1) Hard to study → Reductionist approaches rely heavily on assumptions
- 2) Hard to validate → Going back to the holistic phenomenon is challenging

Spatial and molecular offer complementary views, but:

Are they overlapping?

Or rather unrelated?

Can we extrapolate the conclusions from one to the other?

} My PhD project