SCIENCE MEETS LIFE

Artificial Intelligence Flanders

Understanding cellular symphonies: from (spatial) atlases to cell-cell communication models

Yvan Saeys

yvan.saeys@ugent.be

VIB

Saeyslab
Data mining and Modelling for Biomedicine

GHENT UNIVERSITY

https://github.com/saeyslab/
@Saeyslab

Ellixir Spatial Omics Data Analysis
Lausanne 2025

Spatial (multi)-omics technologies

Important aspects for modelling:

- Single-cell resolution or not ?
- Targeted or not ?
  - Panel designed adequately ?
- **What's the biological question we want to answer ?** … and is the data suitable for that…

Untargeted spatial omics

10X Visium

10X Visium HD

BGI StOmics

Targeted single-cell spatial omics

Resolve MC

10X Xenium

Vizgen MERSCOPE

NanoString CosMx

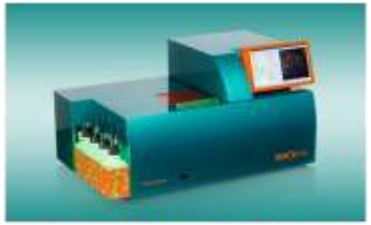
ACD RNAScope

Miltenyi MACSima

Akoya Phenocycler

IMC

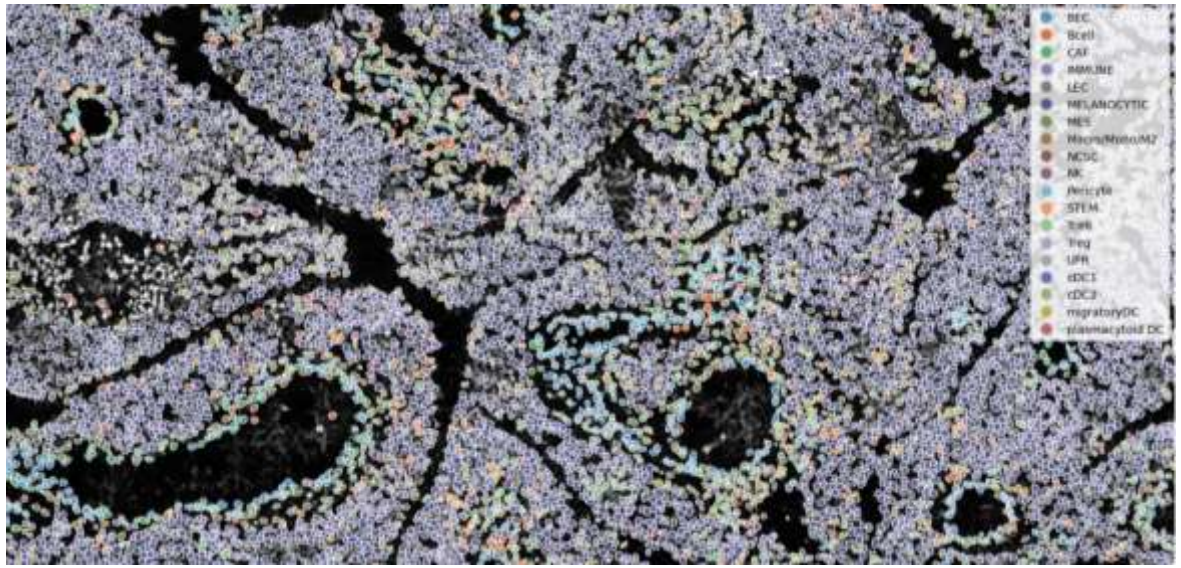
Targeted single-cell spatial multi-omics



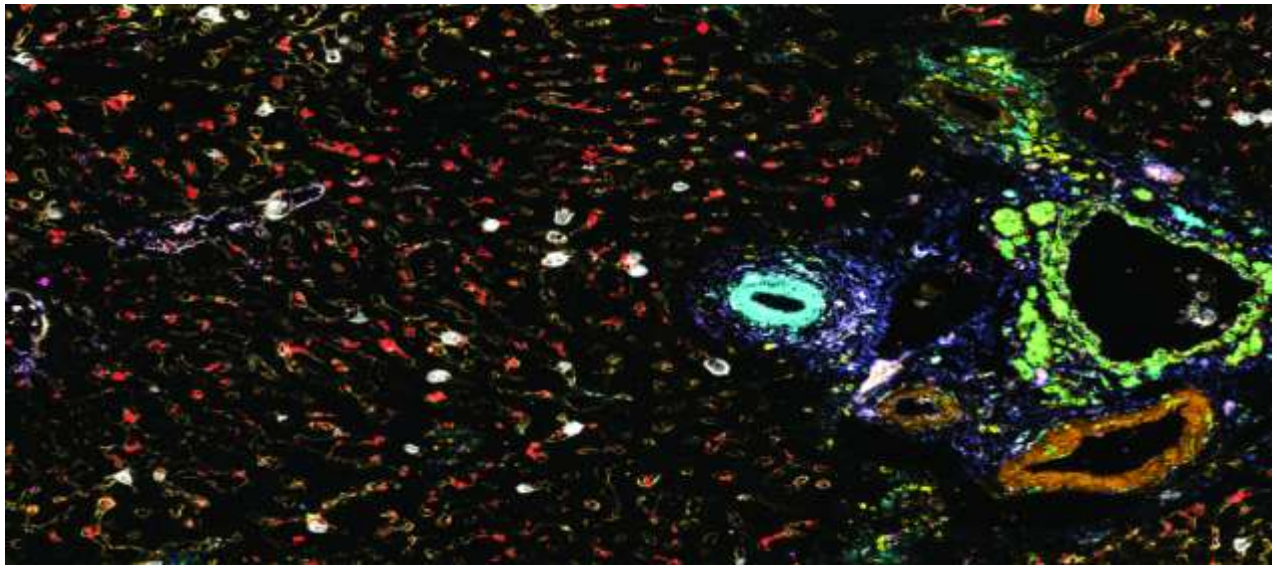
Miltenyi MACSima+RNAsky

Akoya Phenocycler-Fusion

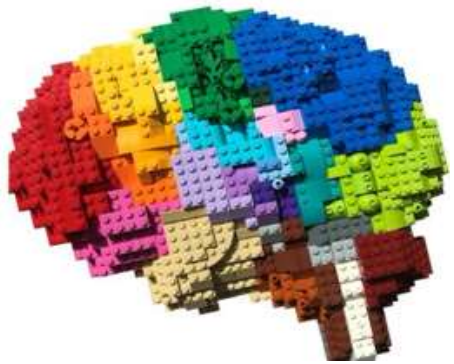
Next generation microscopes…again



Pozniak J et al. A TCF4-dependent gene regulatory network confers resistance to immunotherapy in melanoma. Cell. 2024 Jan 4;187(1):166-183.e25



Guilliams M et al. Spatial proteogenomics reveals distinct and evolutionarily conserved hepatic macrophage niches. Cell. 2022 Jan 20;185(2):379-396.e38



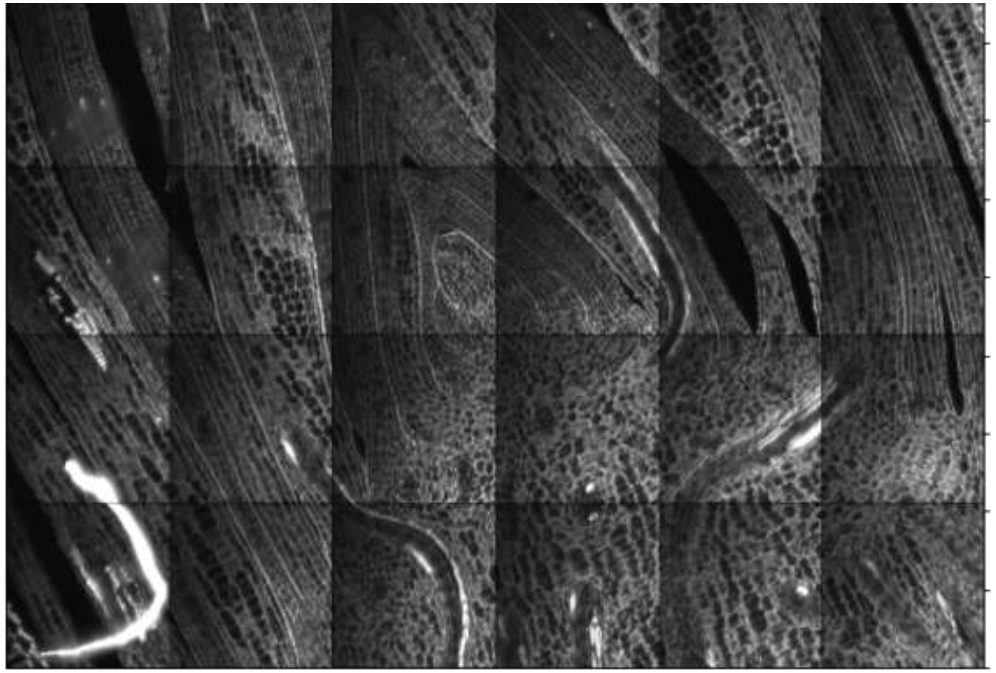
Functional spatial “omics”
multiple modalities
+ spatial context
+ AI models of
cellular interactions
gene regulation

Sometimes the sum
of the PARTS
is **GREATER** than the WHOLE

The reality of spatial omics

We get this !

Text files: up to 5 billion lines



Images: up to 150 GB per stain

We want this !

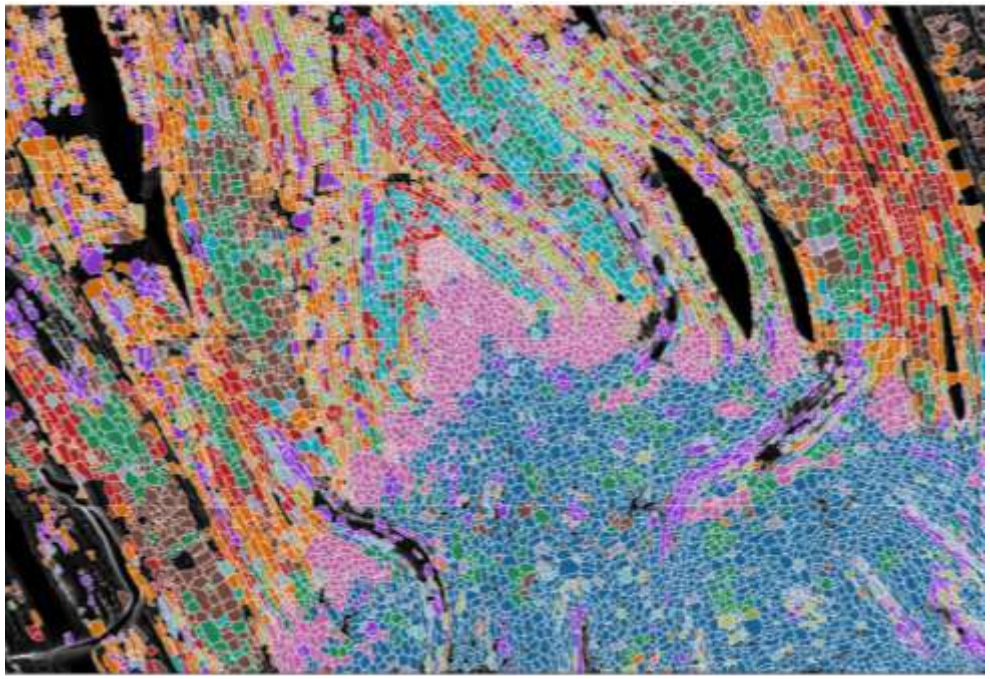


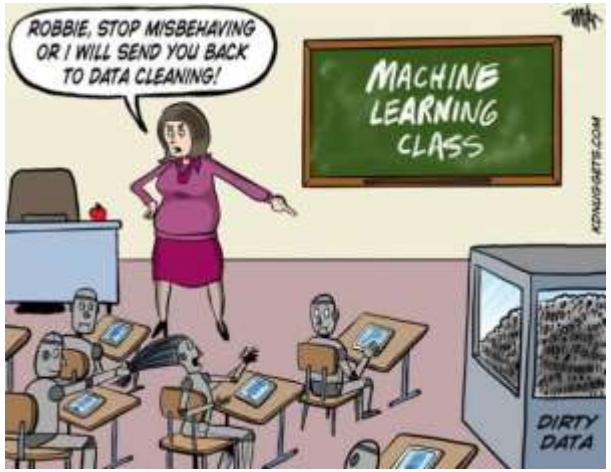
Image: Hilde Nelissen lab

Joossens J., et al. Transit amplifying cells balance growth and differentiation in above-ground meristems bioRxiv 2024.05.04.592499

As always in (biological) data science…

Quality control and preprocessing …

… and nothing else matters

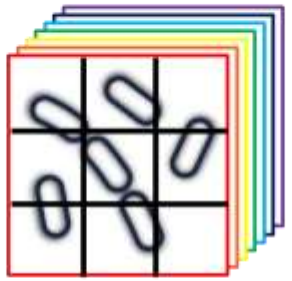


Credit: KDNuggets

Development of novel QC and preprocessing tools for spatial omics data is essential

MULTIPLEXED IMAGING

Raw image tiles



PROCESSING PIPELINE

steinbock

MCMICRO

SPArrOW

Whole-slide processed images

SEGMENTATION PIPELINE

Cellpose

DeepCell

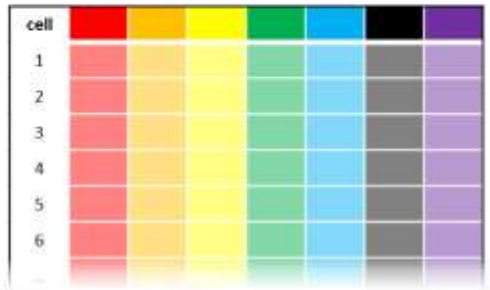
StarDist

Segmented images

SIGNAL SUMMARY

cell×feature table

features



AnnData

SpatialData

Transcriptomics
Resolve MC,
Vizgen MERSCOPE,
BGI Stereo-seq,
Nanostring CosMx

Proteomics
Akoya Phenocycler,
Miltenyi MACSima

e.g. 40 000× 40 000 pixels
~100 transcripts or
~40 protein channels

1 sample =
~25GB OME TIFF image

1 dataset =
~20 samples

Optimize models on HPC

dask

HYDRA

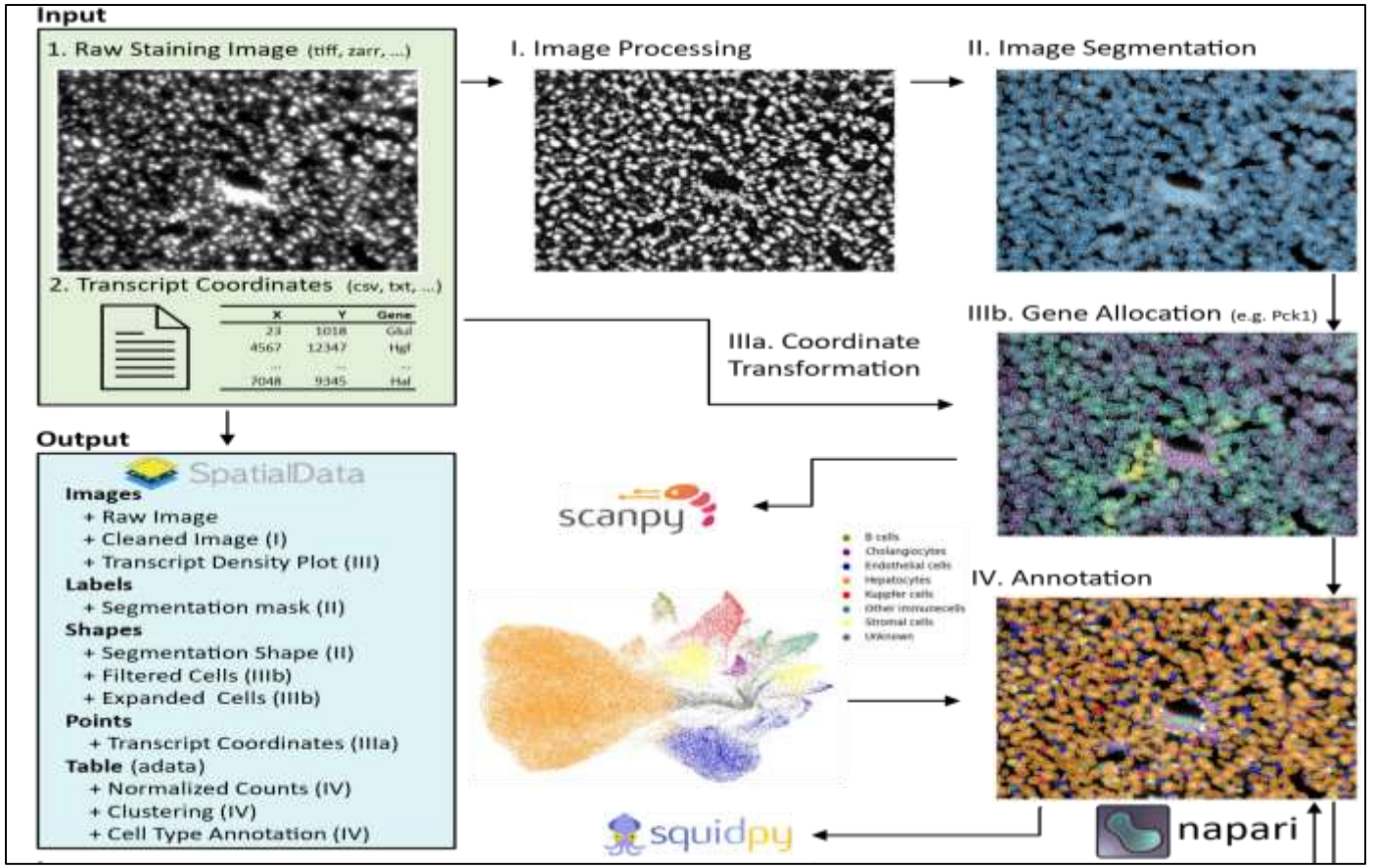
OPTUNA

VLAAMS SUPERCOMPUTER CENTRUM

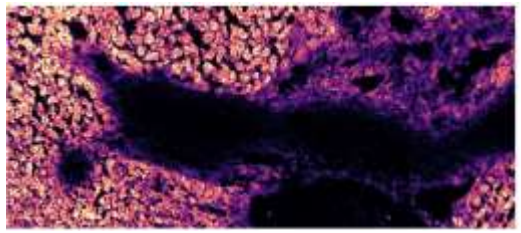
Vlaanderen is supercomputing

Sparrow: a versatile and scalable pipeline for spatial transriptomics

Lotte Pollaris

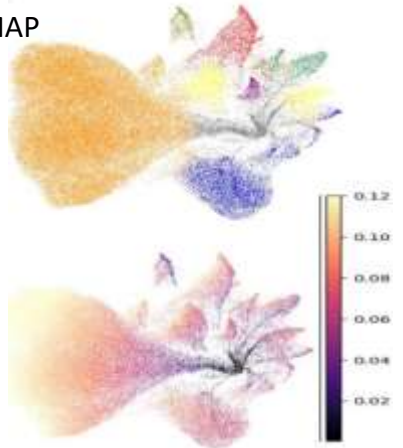


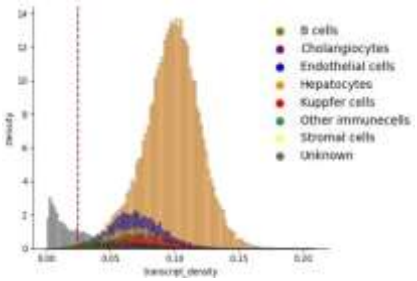
QC metrics and plots



gene	proportion_kept	raw_counts
Dpt	0.717878	110555
Nes	0.770773	22144
Alas2	0.778949	144541
Lhfp	0.783647	112201
Sfrp1	0.784296	40092
Sdc3	0.789354	253928
Cald1	0.793574	1616618
Cybb	0.801593	78828
Bpgm	0.801965	104204
Myh10	0.802167	59813

UMAP

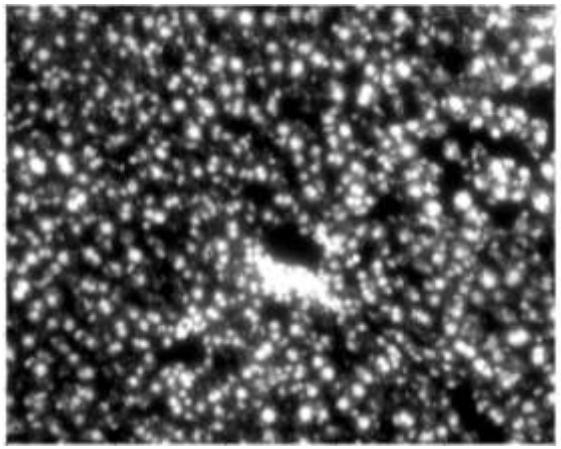


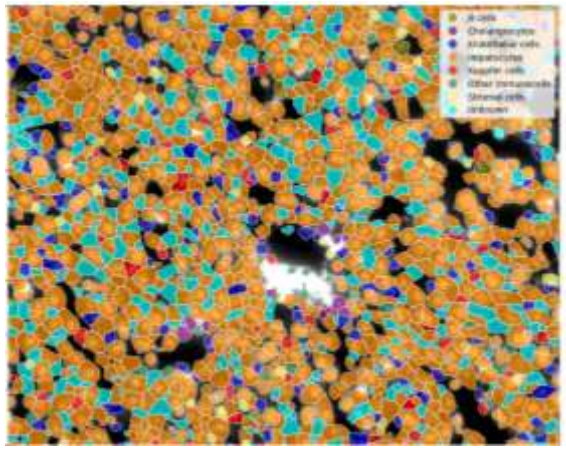


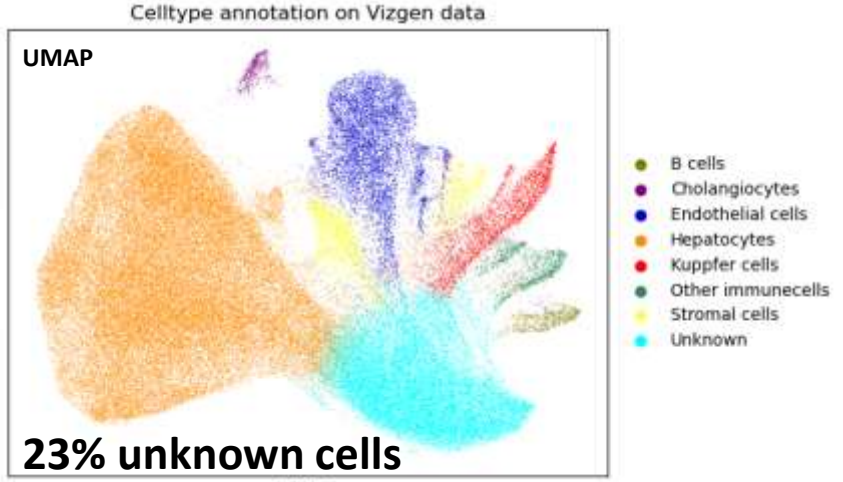
Pollaris L. et al. SPArrOW: a flexible, interactive and scalable pipeline for spatial transcriptomics analysis. bioRxiv 2024.07.04.601829

SPArrOW improves the segmentation and annotation of liver cells

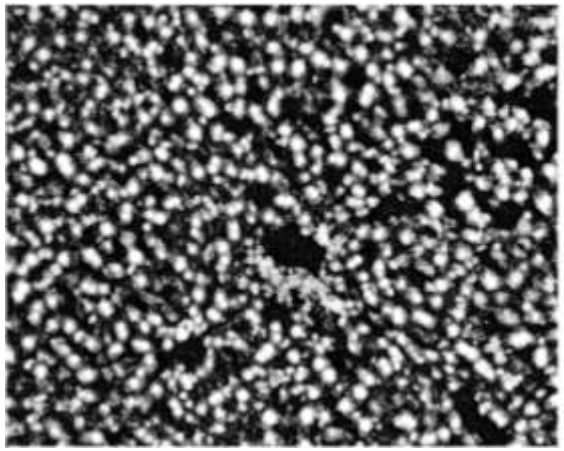
VizGen pipeline

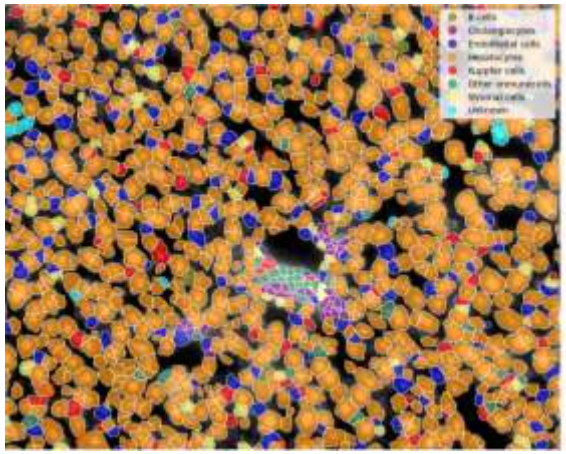


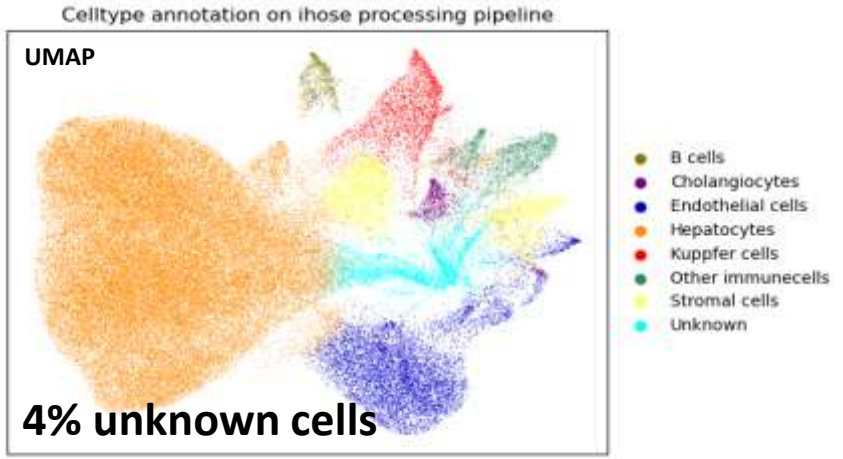




Sparrow pipeline

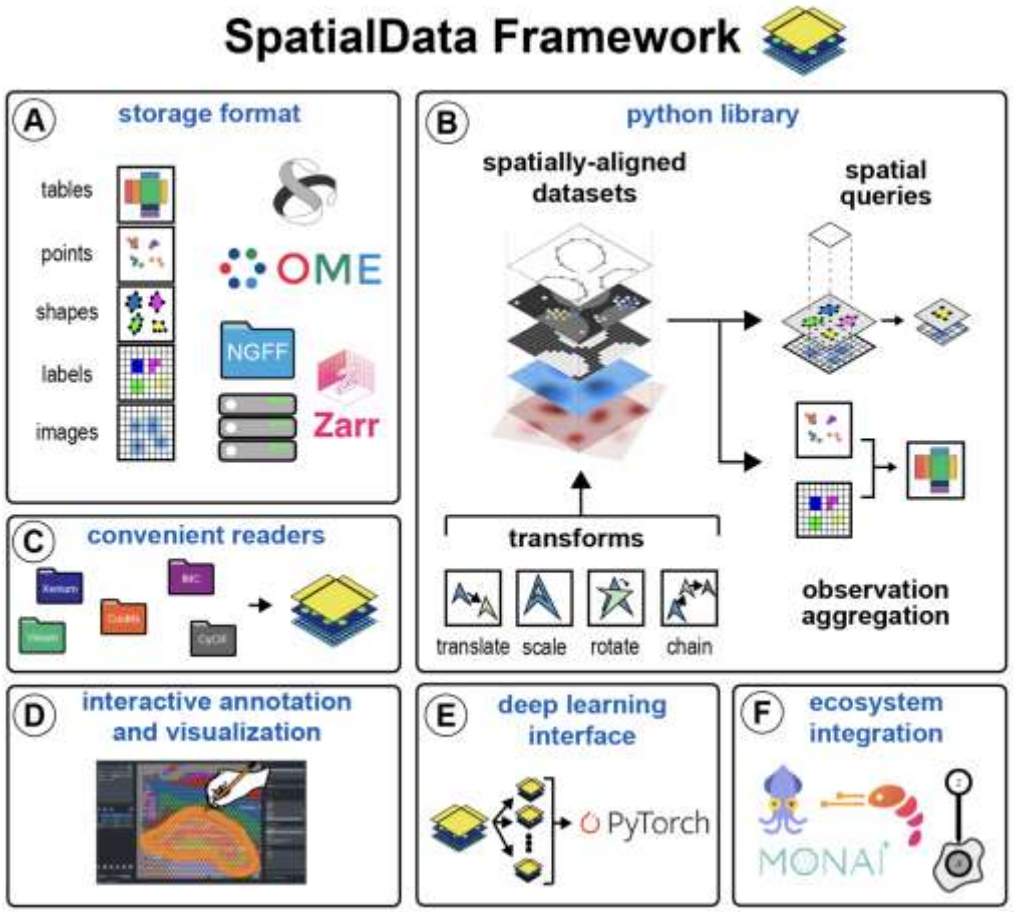


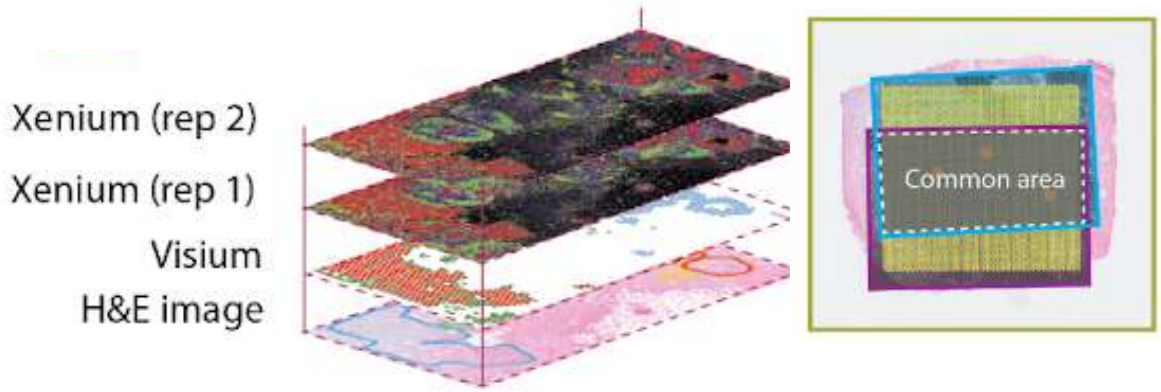




https://github.com/saeyslab/napari-sparrow

Spatial (multi-) omics infrastructure development



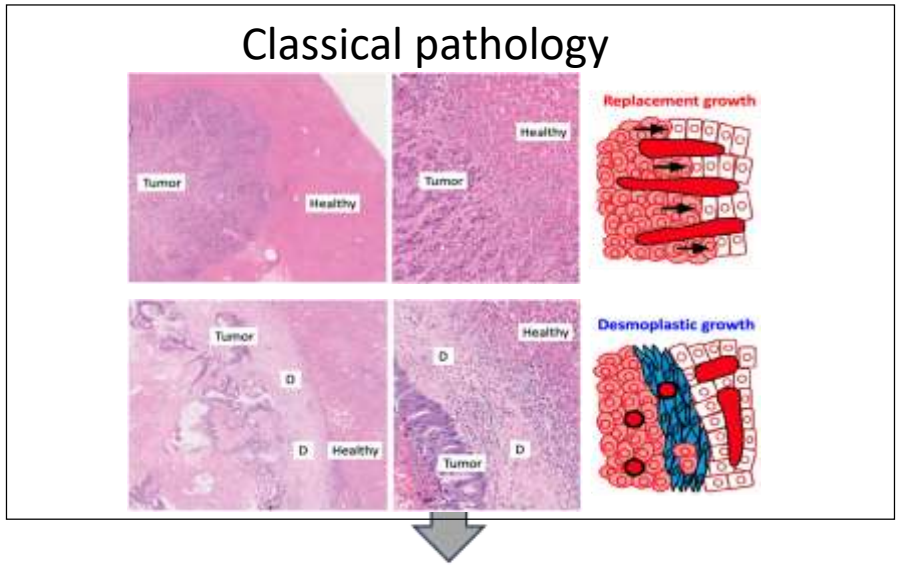


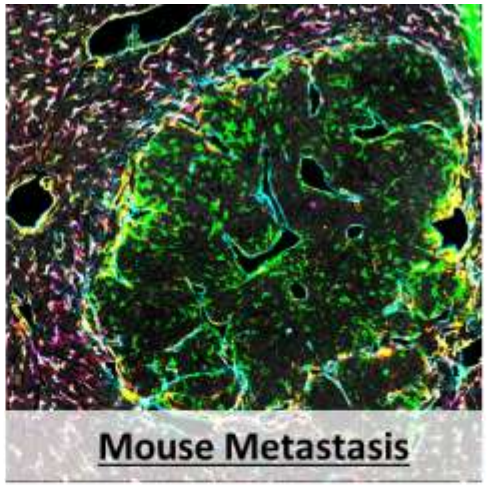
Lotte Pollaris

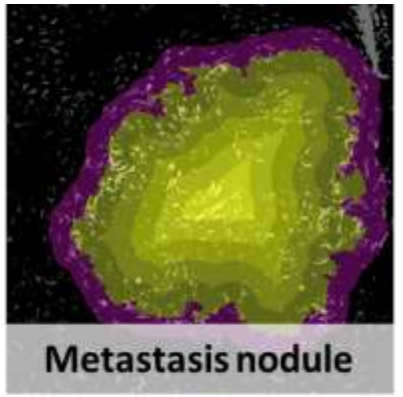
Benjamin Rombaut

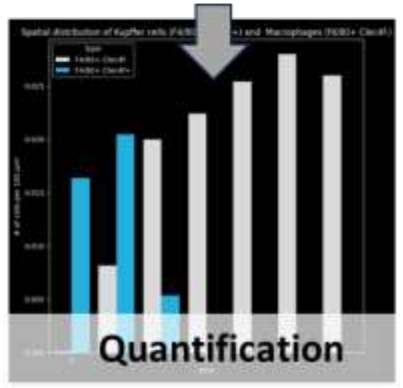
Marconato, L., Palla, G., Yamauchi, K.A. *et al.* SpatialData: an open and universal data framework for spatial omics. *Nat Methods* (2024)

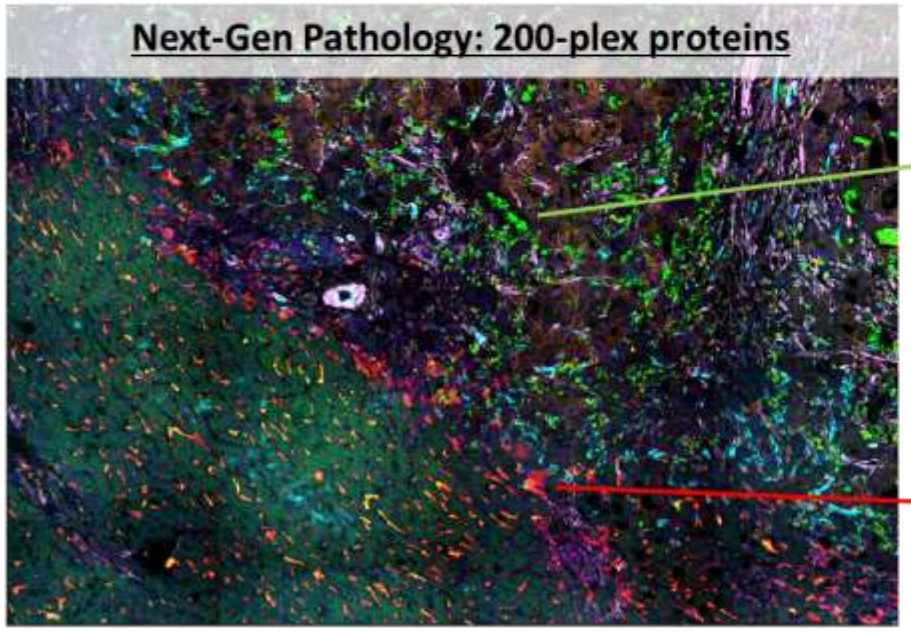
Building the foundations for next-generation pathology











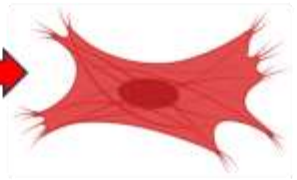
Suppressive Macrophages

Same Macrophages found in Mouse

Pro-fibrotic Macrophages

TGF-b

Activated Fibroblasts

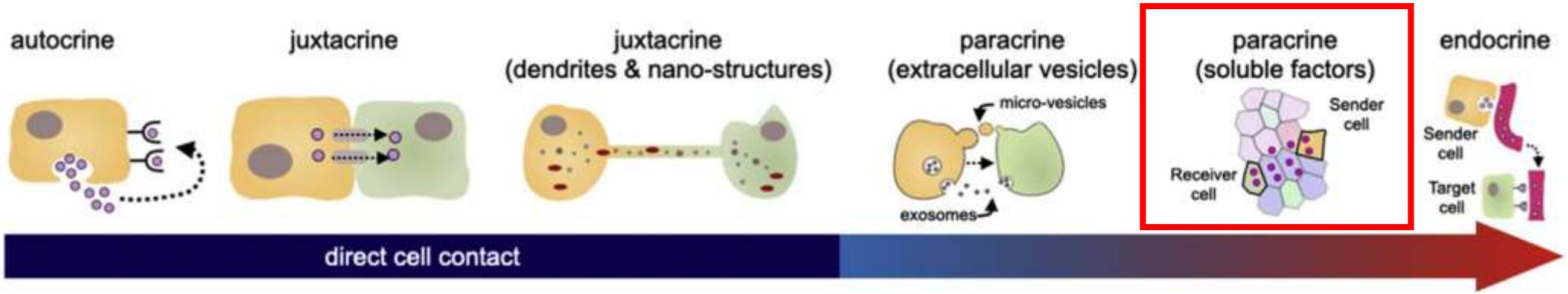


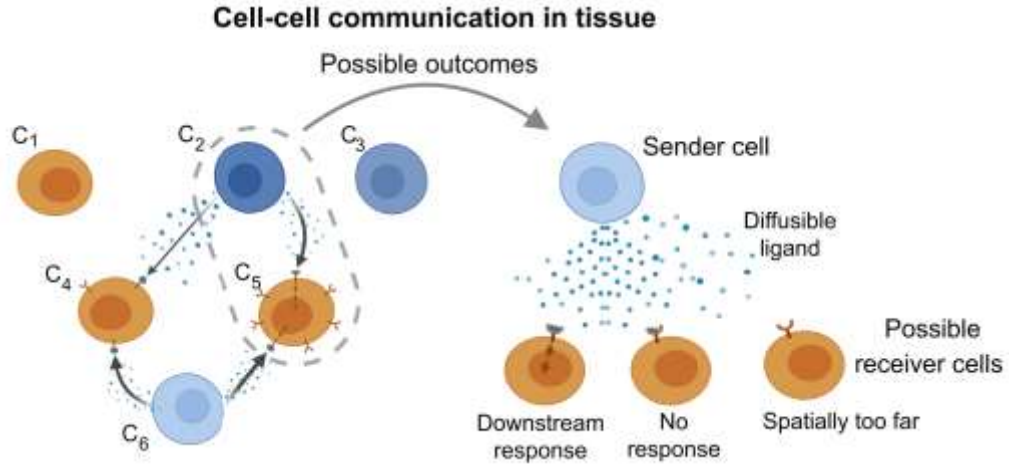
The basics of cell-cell communication (CCC) modelling

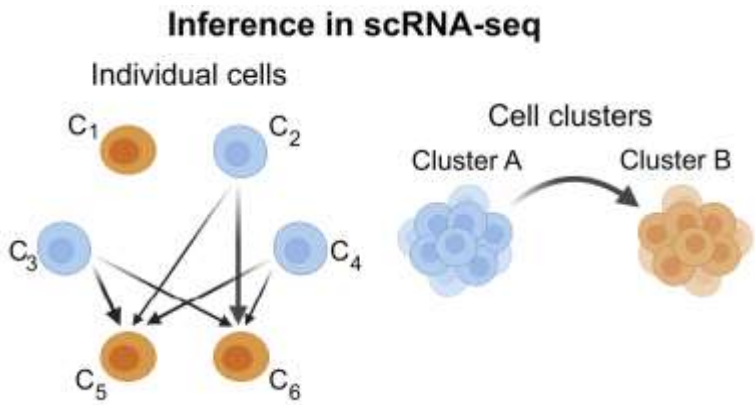
VIB

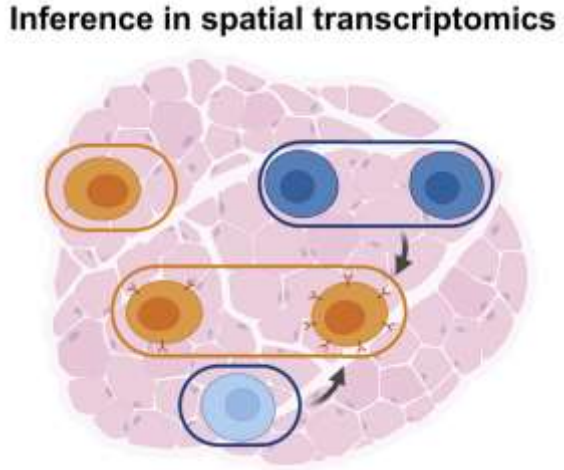
SCIENCE MEETS LIFE

Various modes of CCC exist



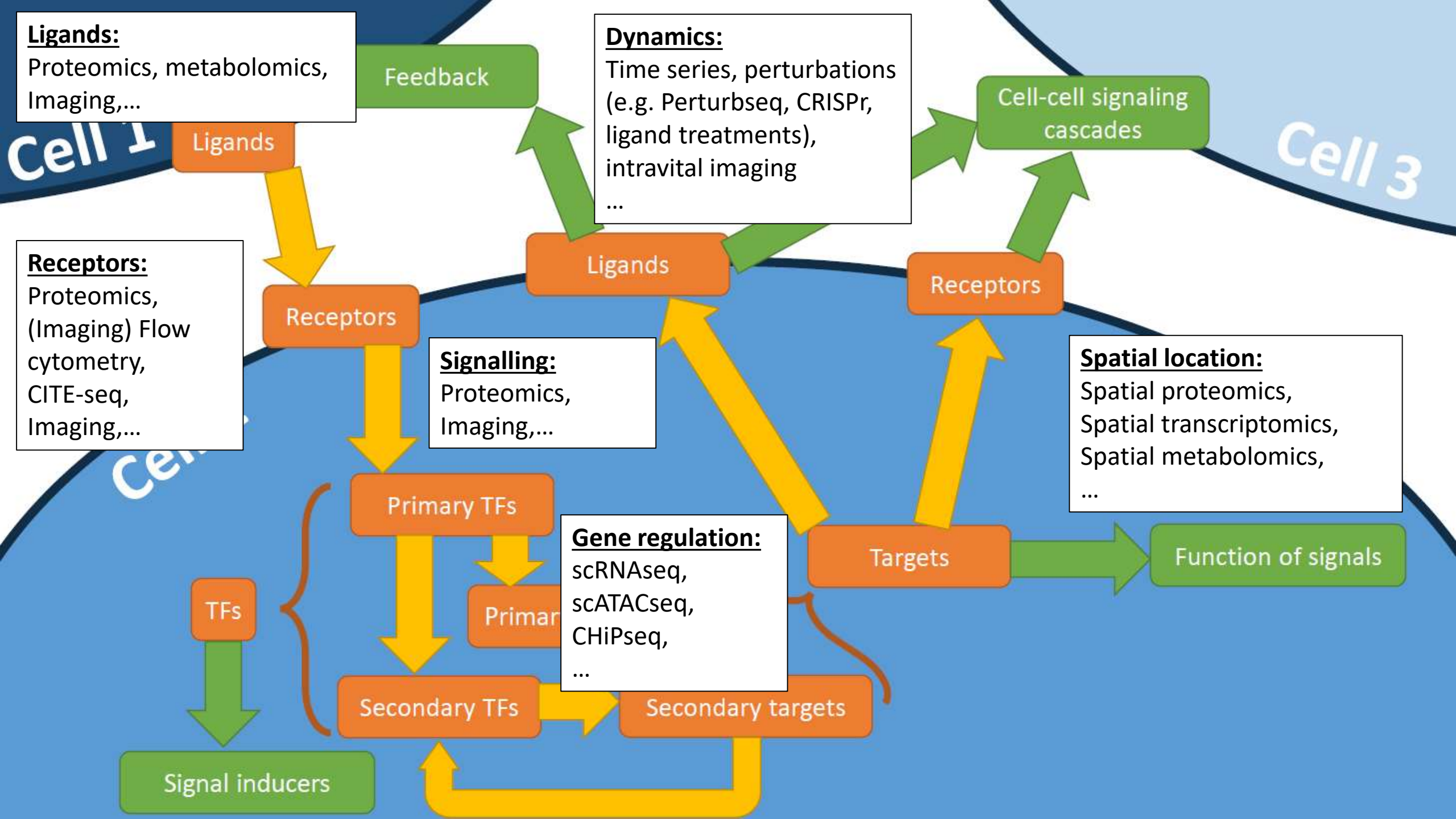






Luthria G, et al. Cell-Cell Communication Networks in Tissue: Toward Quantitatively Linking Structure with Function. Curr Opin Syst Biol. 2021 Sep;27:100341
Almet AA, Cang Z, Jin S, Nie Q. The landscape of cell-cell communication through single-cell transcriptomics. Curr Opin Syst Biol. 2021 Jun;26:12-23

What types of data do we need to study ***functional*** cell-cell communication ?


Ligands:
Proteomics, metabolomics, Imaging,…
Feedback
Cell 1
Ligands
Dynamics:
Time series, perturbations (e.g. Perturbseq, CRISPr, ligand treatments), intravital imaging
…
Cell-cell signaling cascades
Cell 3
Receptors:
Proteomics, (Imaging) Flow cytometry, CITE-seq, Imaging,…
Receptors
Ligands
Receptors
Signalling:
Proteomics, Imaging,…
Spatial location:
Spatial proteomics, Spatial transcriptomics, Spatial metabolomics, …
Primary TFs
Gene regulation:
scRNAseq, scATACseq, CHiPseq, …
Targets
Function of signals
TFs
Primar
Secondary TFs
Secondary targets
Signal inducers

“Who says what to whom, how, why, when and where ?”

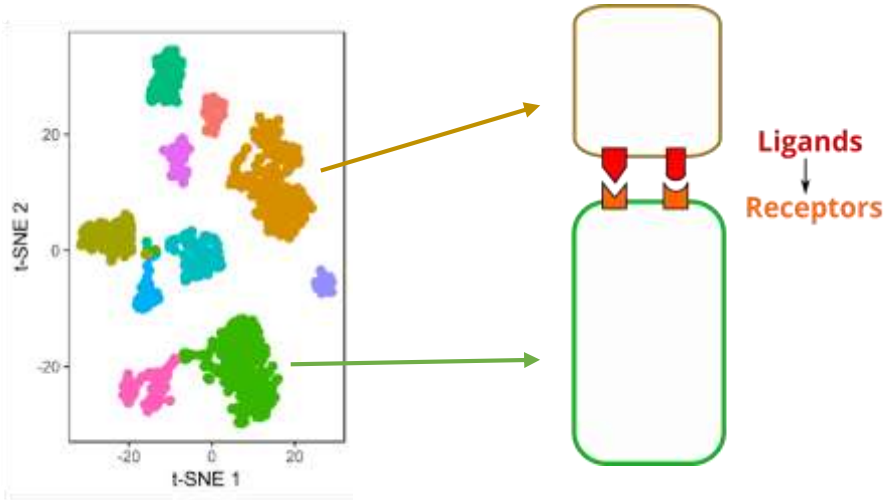
Typically studied in natural language analysis

We want to study the language of cells !!

(and if we see how NLP has evolved,
that will probably lead to Large Cellular Language models soon)

Most current methods focus on CCC inference from ***transcriptomics*** data

ligand-receptor
Inference (LRI)



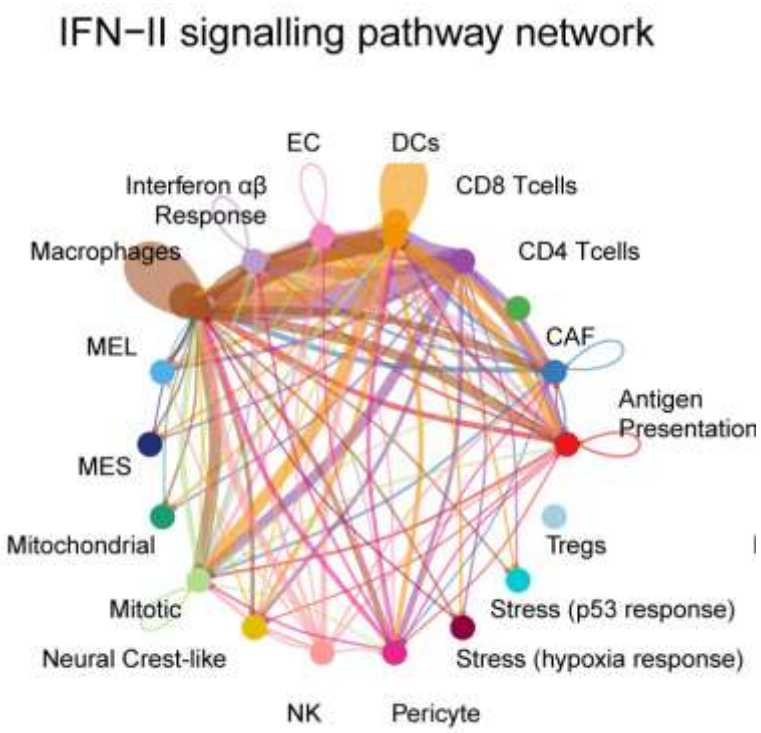
Combine expression data with
database of known ligand-receptor pairs

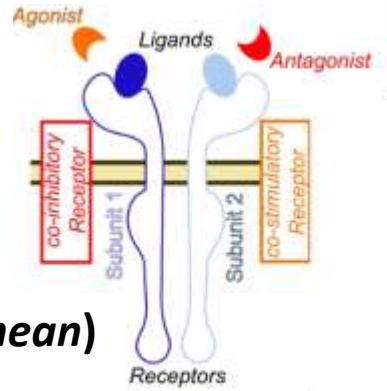
CellPhoneDB, CellChat, …, LIANA+

Key steps for LRI methods:

1. Filter gene expression matrix for ligands and receptors expression
2. Aggregate expression level of each gene across all single cells of a specific cell type
3. For each pair of cell types:
   evaluate Ligand-Receptor interaction by ligand expression in the sender and receptor expression in the receiver
4. Calculate a communication score for each LRI in each pair of cell types and rank
5. Filter using statistical analyses to identify significant interactions.

Example: CellChat to visualize the tumor microenvironment in melanoma





- Modelling of LR pairs with complex architecture (***multi-subunit***)
- Conservative modelling: multi-subunit complexes are required to show expression of all components (***geometric mean***)

$$L_i = \sqrt[m1]{L_{i,1} \cdots L_{i,m1}} \qquad R_j = \sqrt[m2]{R_{j,1} \cdots R_{j,m2}} \cdot \frac{1 + RA_j}{1 + RI_j}$$

i : sender cell type $\qquad$ j : receiver cell type

- Estimation of LR interaction activity (*agonist/antagonist and co-inhibitory/stimulatory subunits*) via Hill functions leveraging for the presence of agonists and antagonists

$$P_{i,j}^k = \frac{L_i R_j}{K_h + L_i R_j} \times \left(1 + \frac{AG_i}{K_h + AG_i}\right) \cdot \left(1 + \frac{AG_j}{K_h + AG_j}\right) \times \frac{K_h}{K_h + AN_i} \cdot \frac{K_h}{K_h + AN_j}$$

k : ligand − receptor pair $\quad AG_{i/j}$: agonist
i : sender cell type $\quad AN_{i/j}$: antagonist
j : receiver cell type

- Significance of each $P_{i,j}^k$ is evaluated by a permutation test via label switching

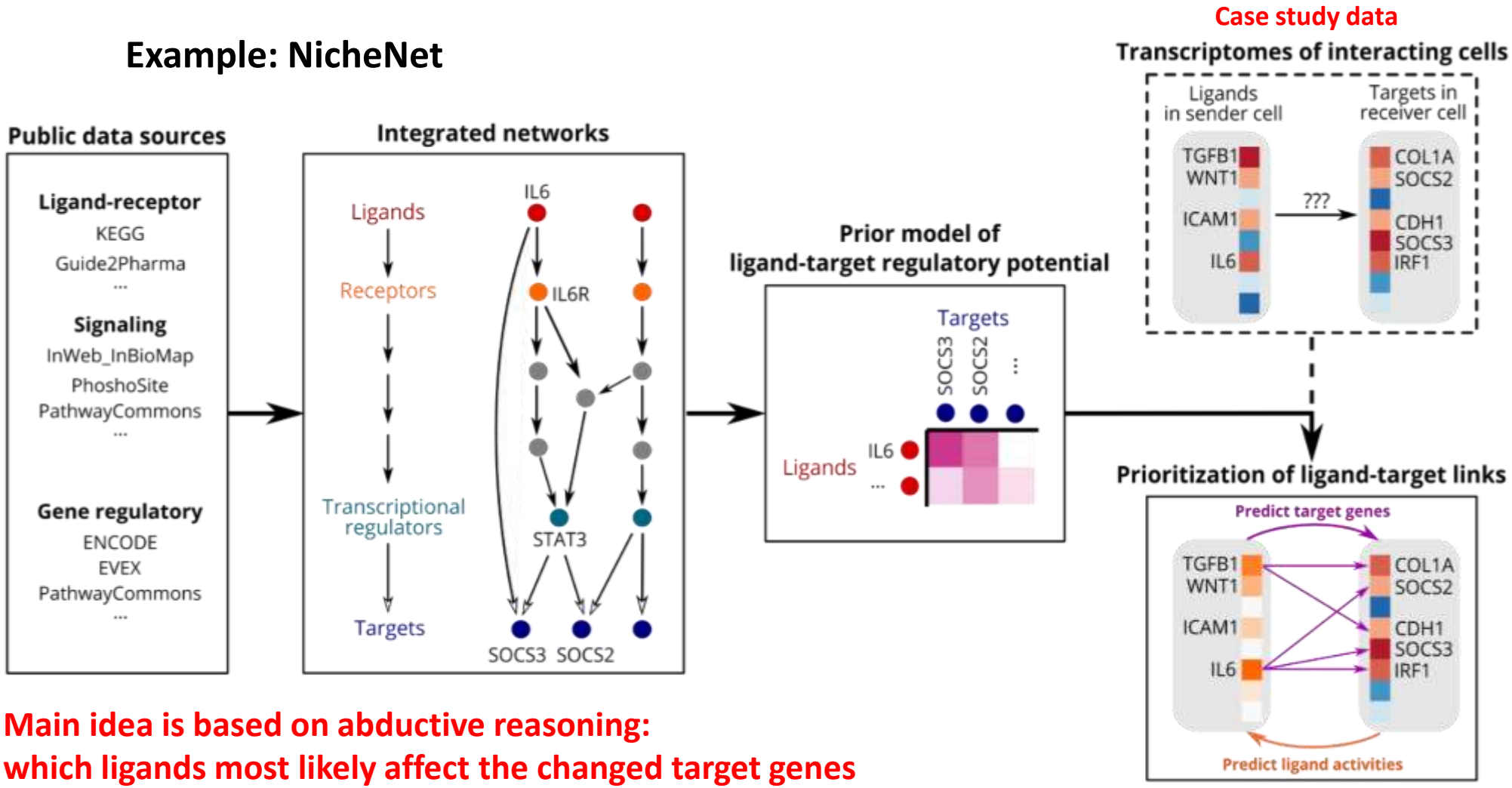
Pozniak J et al. A TCF4-dependent gene regulatory network confers resistance to immunotherapy in melanoma. Cell. 2024 Jan 4;187(1):166-183.e25
Jin, S. et al. Inference and analysis of cell–cell communication using CellChat. Nat. Commun. 12, 1088 (2021)

Strengths and limitations of LRI based CCC models

- Strengths
  - Different levels of complexity of modelling LR interaction potential based on curated databases (e.g. multi-subunit modelling)
  - Simple and easy to calculate
  - Databases can be expanded, curated,…, to improve models
  - Potential for many downstream analyses
- Weaknesses
  - Takes LR gene expression as a proxy for functional proteins
  - Depends on the curated databases of LR interactions
    - Many tools use different databases -> different results
  - Do not take into account downstream functional effects
  - Databases are general and not (yet) cell type specific

Intracellular CCC methods rank ligand-receptor interactions based on their affected target genes **(= their effect on gene expression)**

Example: NicheNet



Main idea is based on abductive reasoning:
which ligands most likely affect the changed target genes

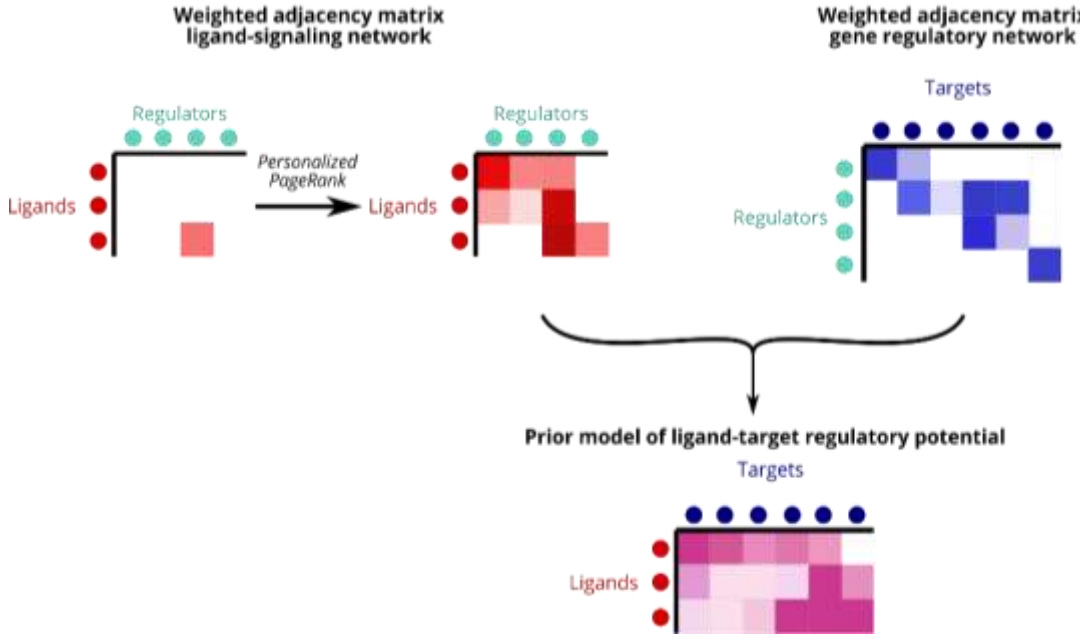
Browaeys, R., Saelens, W. & Saeys, Y. NicheNet: modeling intercellular communication by linking ligands to target genes. *Nat. Methods* **17**, 159–162 (2020)

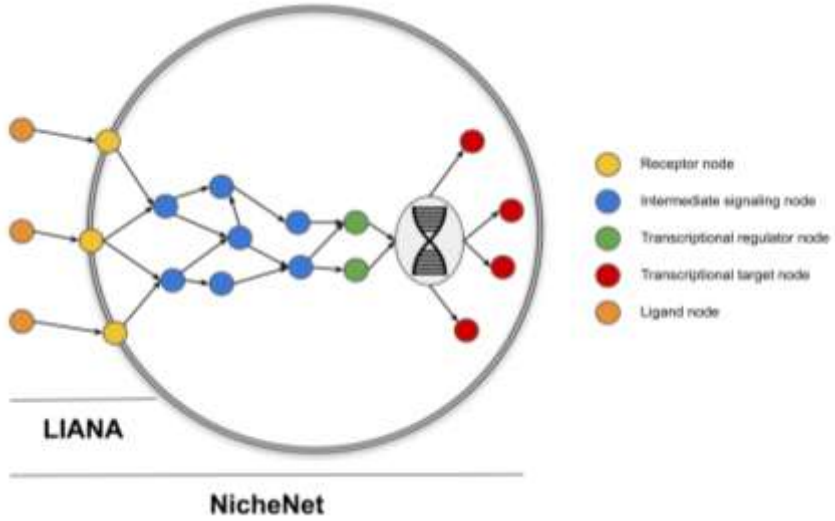
NicheNet: a flexible modelling framework

Add whole OmniPath
+
New data sources
= NicheNet 2.0

Add specific proteomics data
Add specific spatial omics data

Add specific epigenetic data
Add GRN inference results (e.g. GENIE3/SCENIC+)

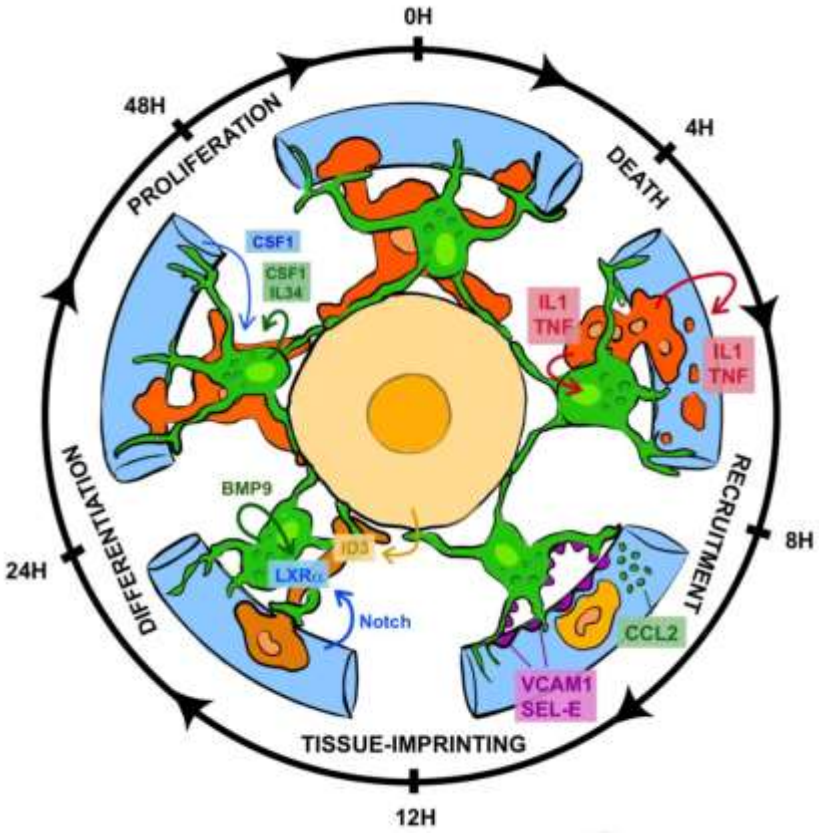


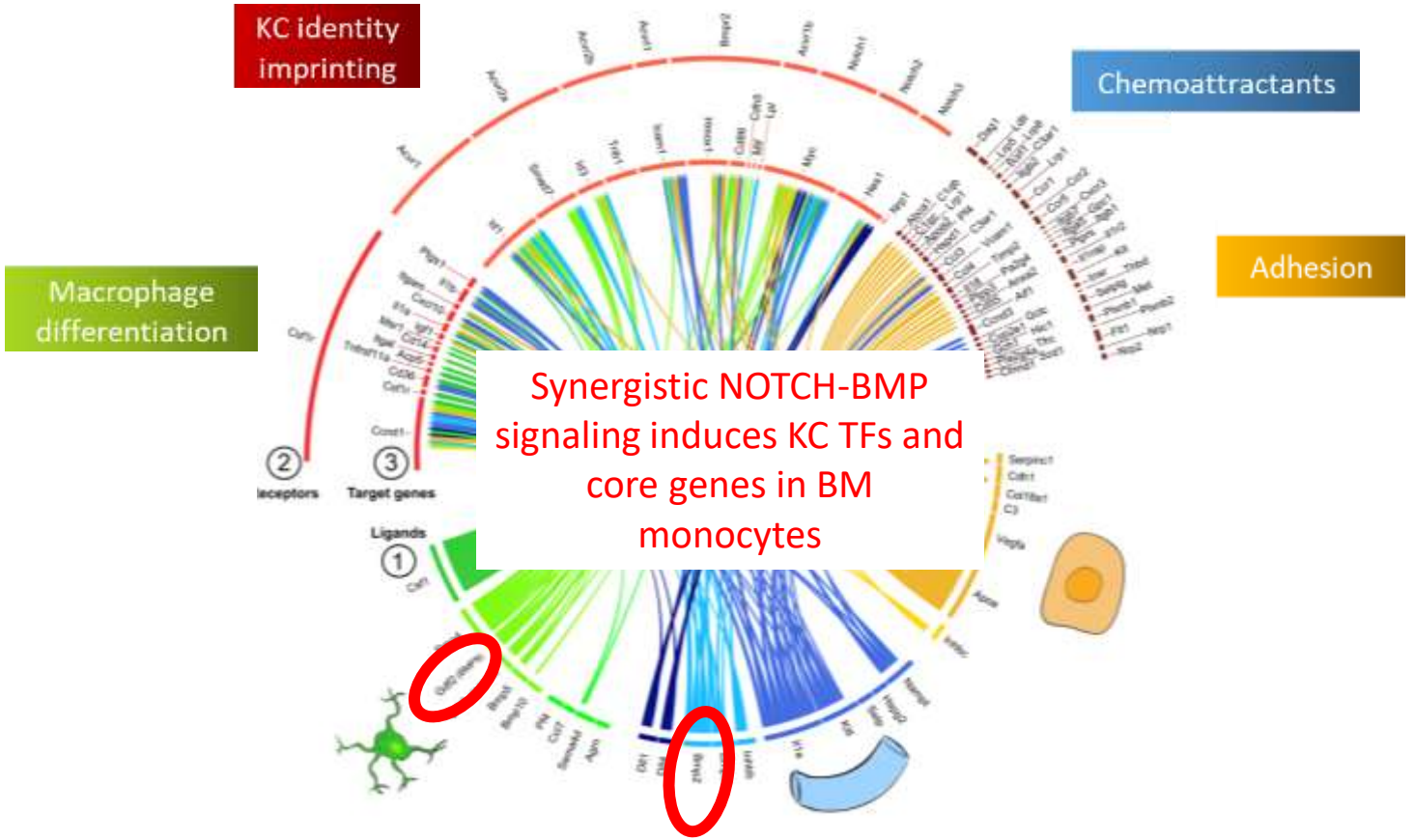


https://saezlab.github.io/liana/articles/liana_nichenet.html

NicheNet identifies the niche signals that imprint the Kupffer cell identity on engrafted monocytes

The Kupffer cell "niche" in the liver

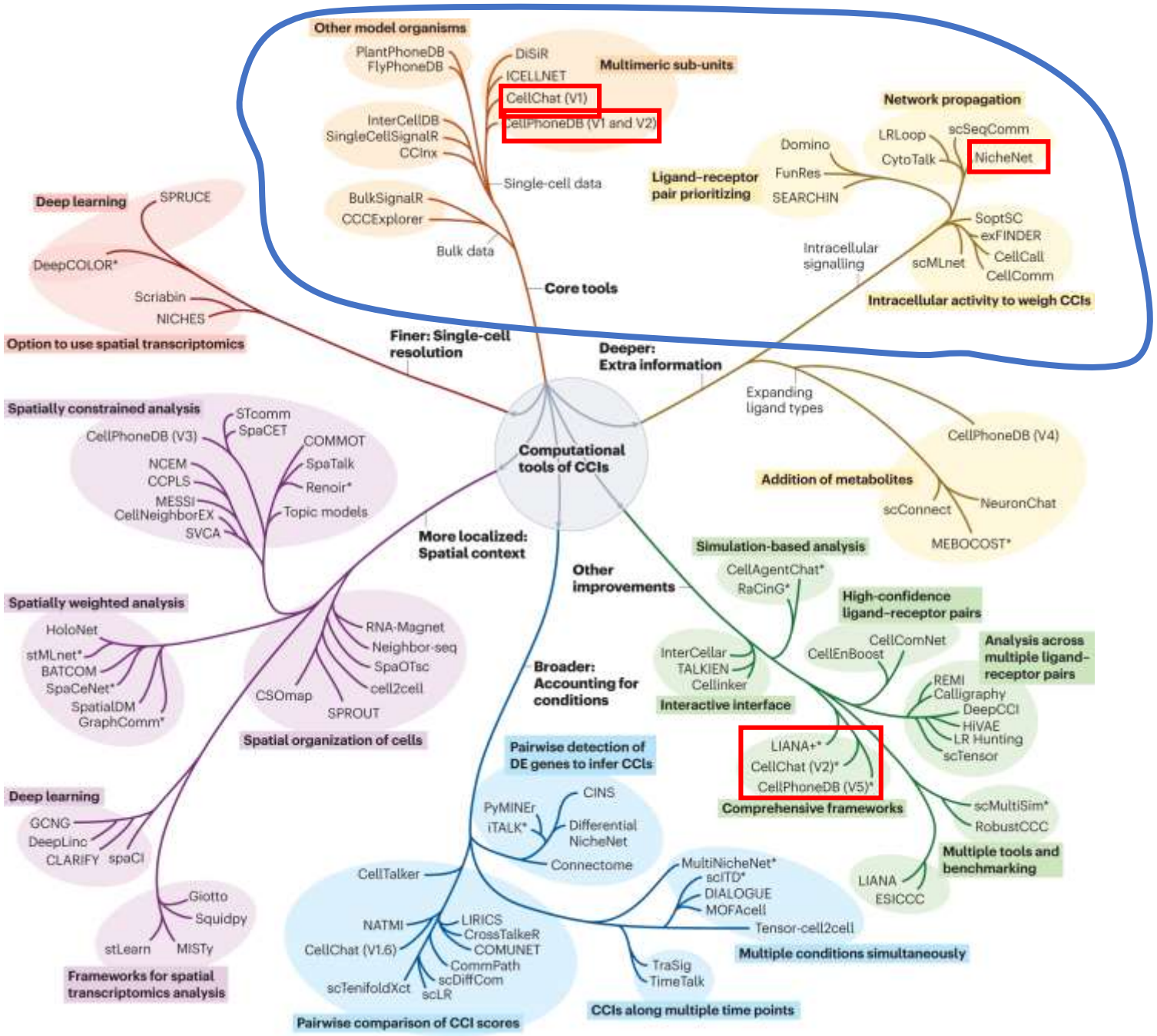




Bonnardel J, et al. Stellate Cells, Hepatocytes, and Endothelial Cells Imprint the Kupffer Cell Identity on Monocytes Colonizing the Liver Macrophage Niche. Immunity. 2019 15;51(4):638-654

Strengths and limitations of intracellular CCC models

- Strengths
  - Go beyond the mere expression of ligands and receptors and also considers downstream effects of CCC
  - Databases can be expanded, curated,…, to improve models
  - Potential for many downstream analyses
  - Interpretable by investigating the signaling pathways connecting receptors and target genes
- Weaknesses
  - No multi-meric subunit approaches
  - More complex analysis involving ligands, receptors and downstream target genes
  - Depends on the curated databases of LR interactions, signaling and GRN
  - Databases are general and not (yet) cell type specific



Core tools:

- **CellPhoneDB**
- **CellChat**
- **NicheNet**
- **Liana(+)**

Armingol, E., Baghdassarian, H. M. & Lewis, N. E. The diversification of methods for studying cell–cell interactions and communication. *Nat. Rev. Genet.* (2024)

Tool complementarity

- The many available tools can be seen as a large “toolbox” of components, and it depends on your own creativity how you combine parts of the different toolboxes
- Example:

  CellChat & NicheNet as complementary tools to study CCC events:

  - LRIs can be selected based on the orchestration of LR components (CellChat)
  - Selected ligands can be ranked for their capacity to explain the differential expression induced by the active CCC event (NicheNet)

Modelling cell-cell communication (CCC) from (spatial) omics data

VIB

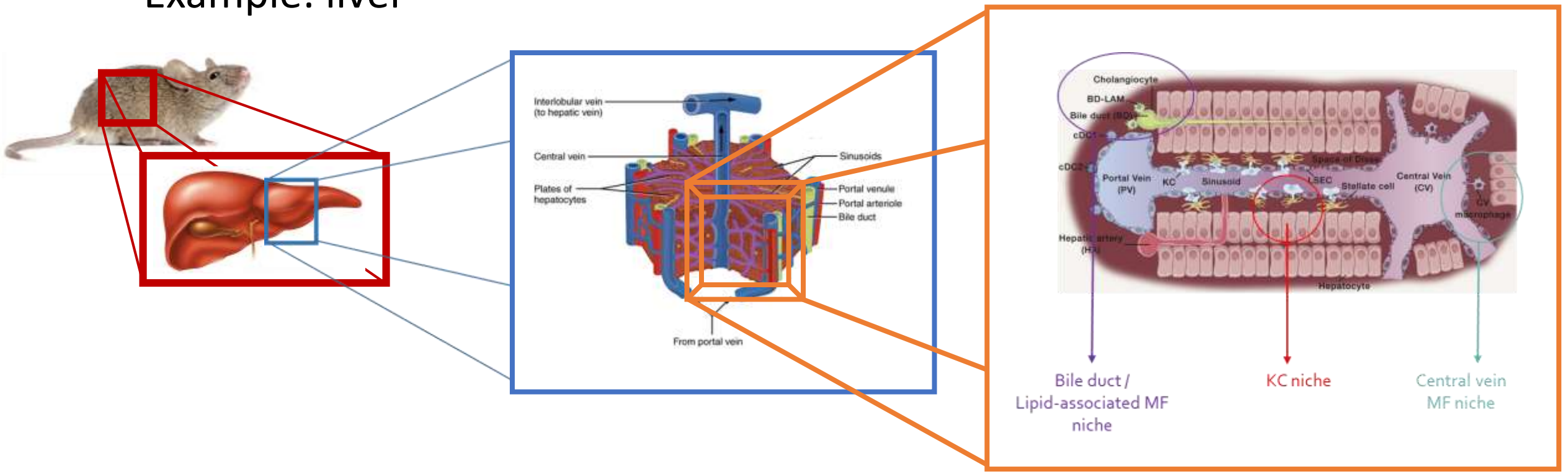
SCIENCE MEETS LIFE

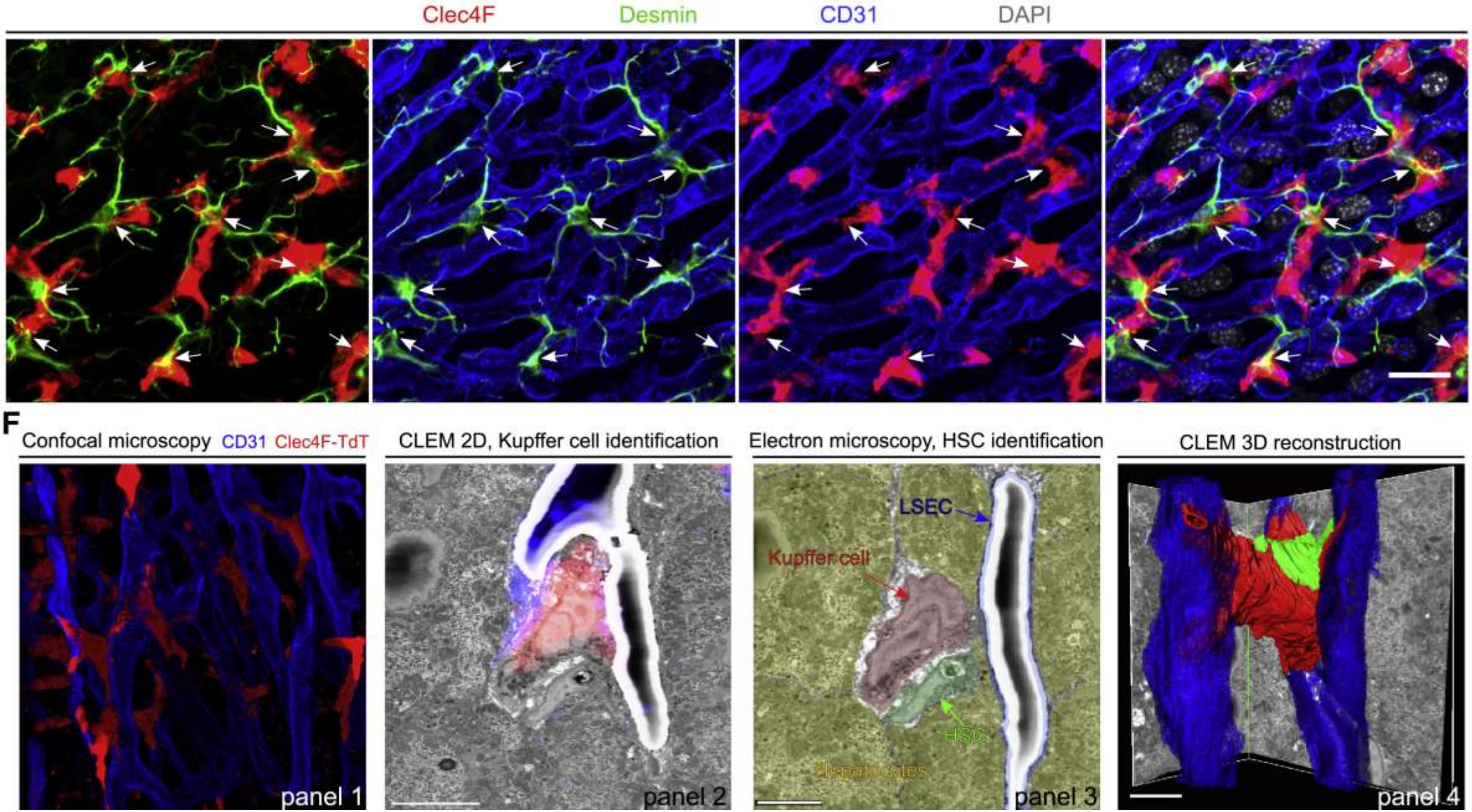
Types of spatial CCC analyses

- Definition/reconstruction of cellular neighbourhoods / niches
  - Cells within the niche likely interact
- Spatial co-localization of ligand-receptor pairs
  - Spatially enriched LRIs
  - Spatially-aware LR inference analysis
  - Spatial statistics
- Spatially informed intercellular program definition
  - LRI + intracellular signaling (+ GRN)
- Mechanistic models of communication dynamics
- Comparative analyses of all above

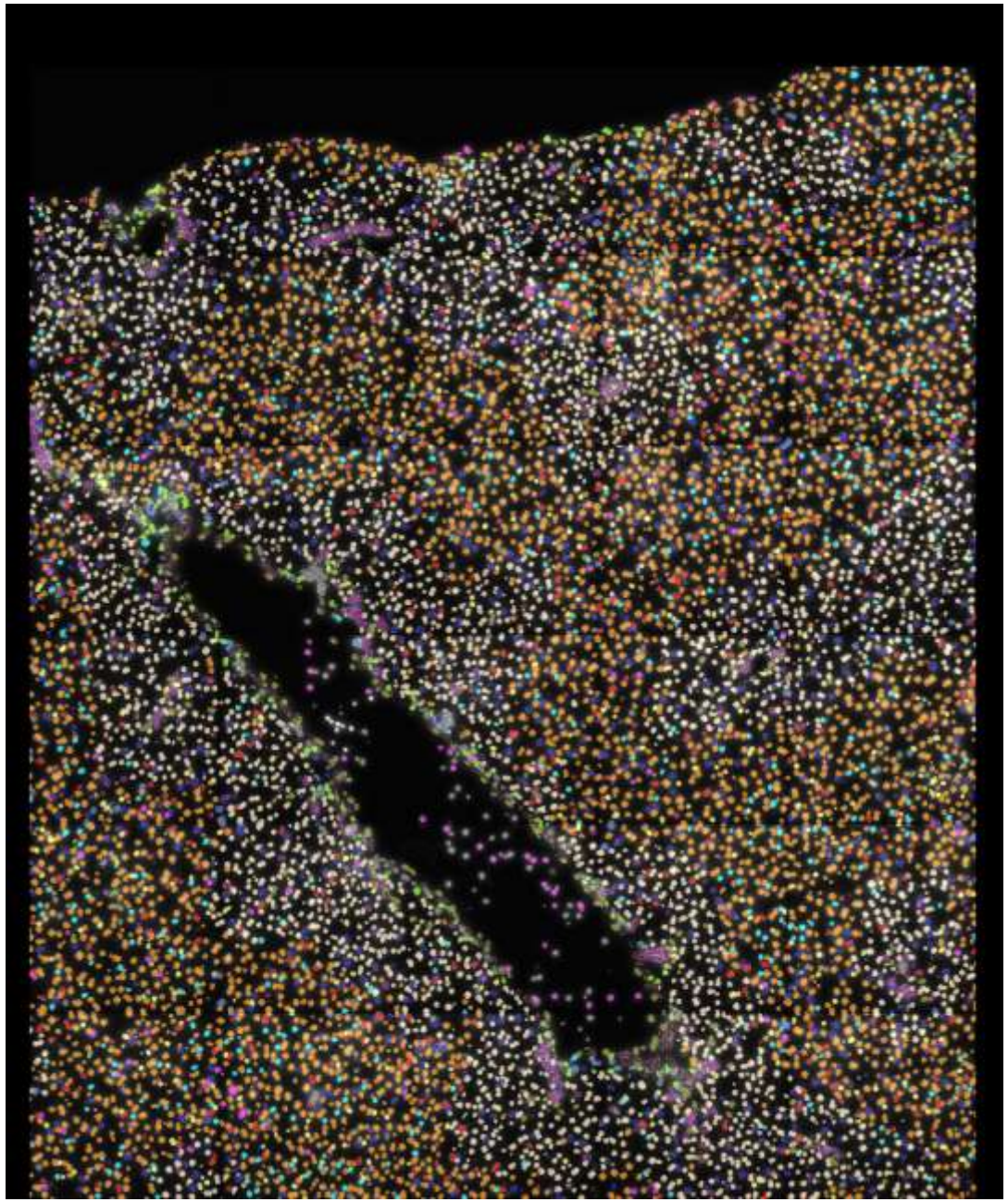
What is a niche ?

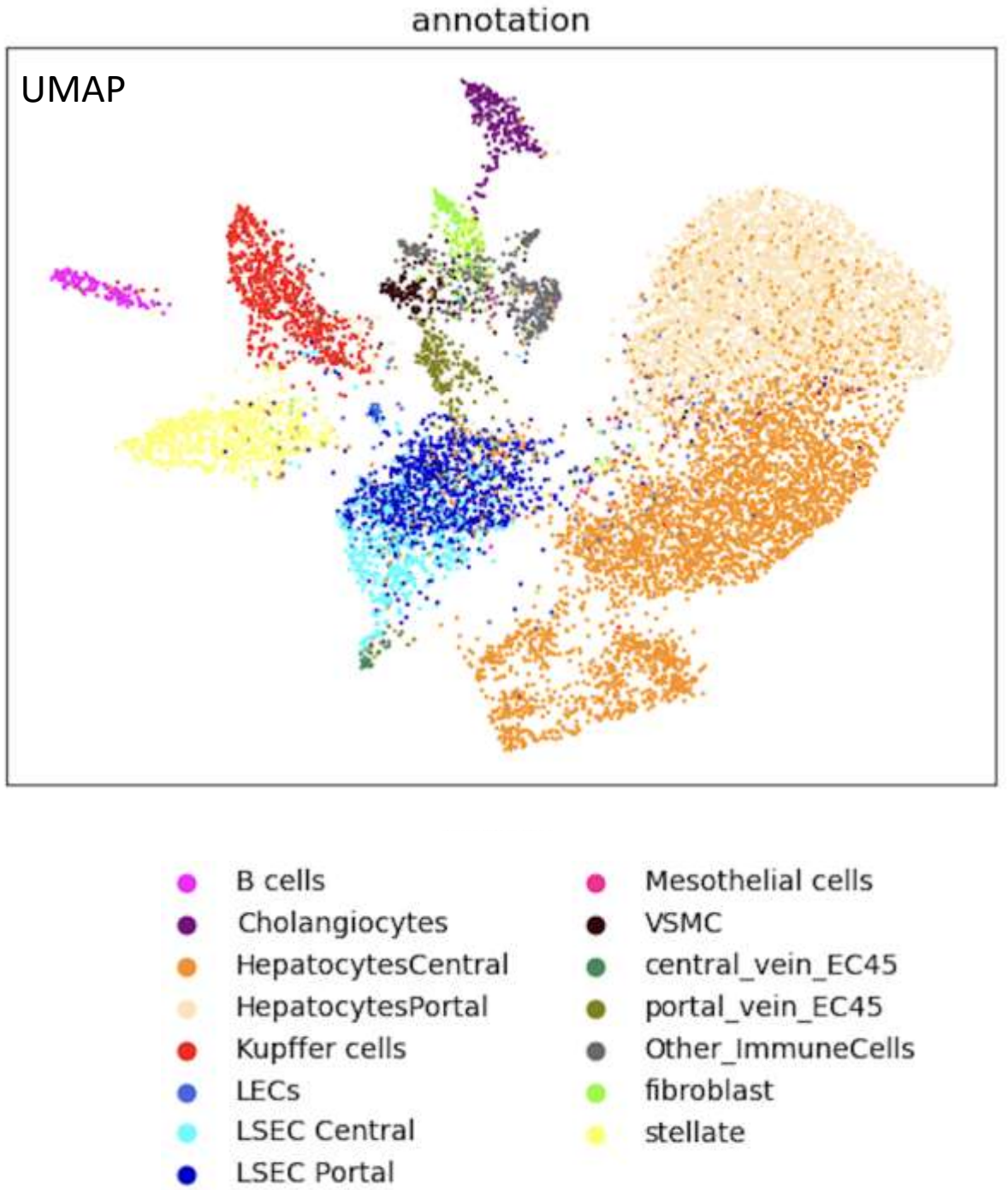
- A collection of cells that are spatially constrained to jointly implement a biological function
- Example: liver



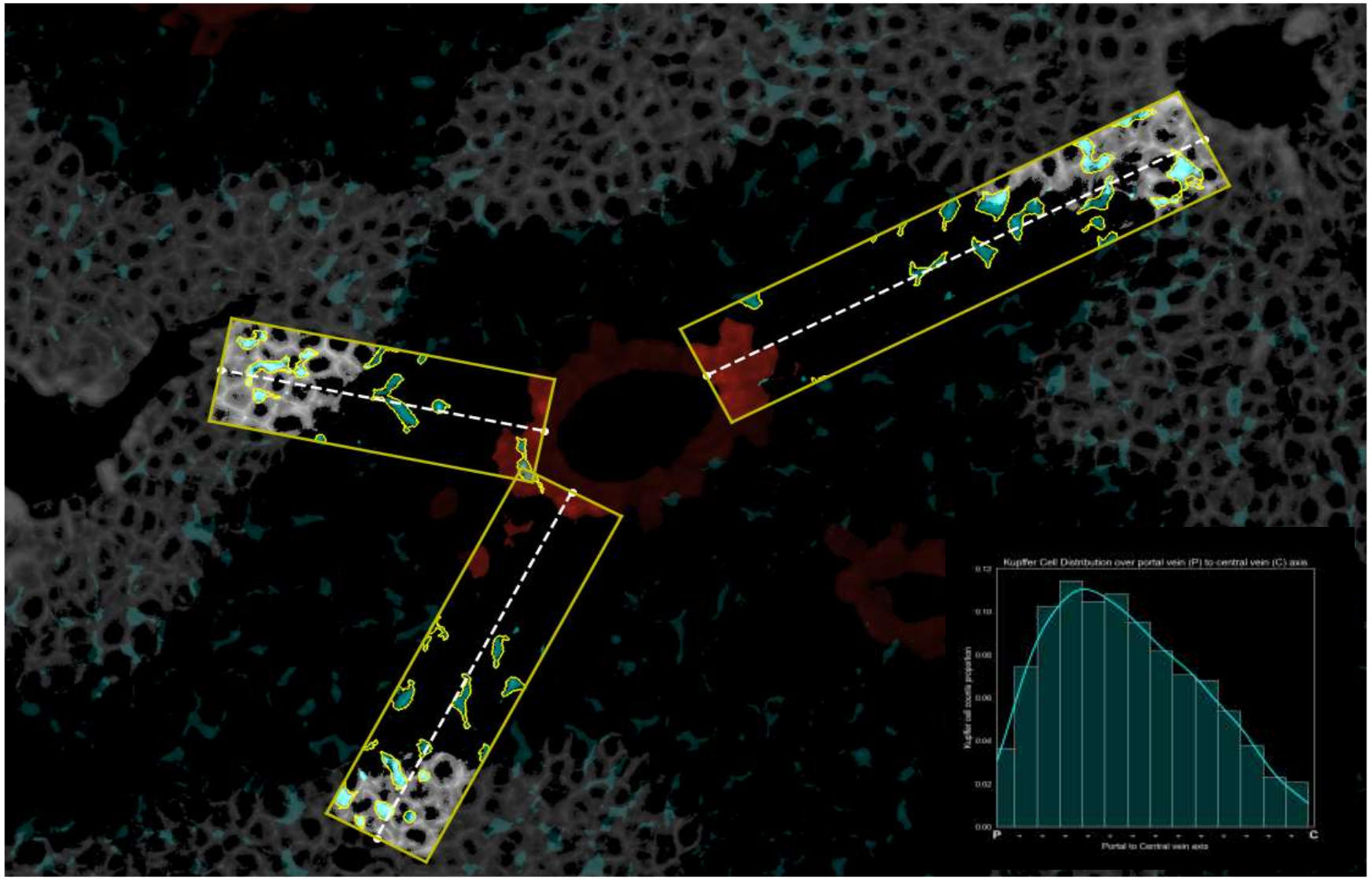


Bonnardel J, et al. Stellate Cells, Hepatocytes, and Endothelial Cells Imprint the Kupffer Cell Identity on Monocytes Colonizing the Liver Macrophage Niche. Immunity. 2019 15;51(4):638-654




annotation
UMAP
B cells
Cholangiocytes
HepatocytesCentral
HepatocytesPortal
Kupffer cells
LECs
LSEC Central
LSEC Portal
Mesothelial cells
VSMC
central_vein_EC45
portal_vein_EC45
Other_ImmuneCells
fibroblast
stellate

Ecadh (Portal Vein) Glul (Central Vein) Vsig4 (Kupffer Cells)



Cells Of Interest Segmentation

Select Gradient Area

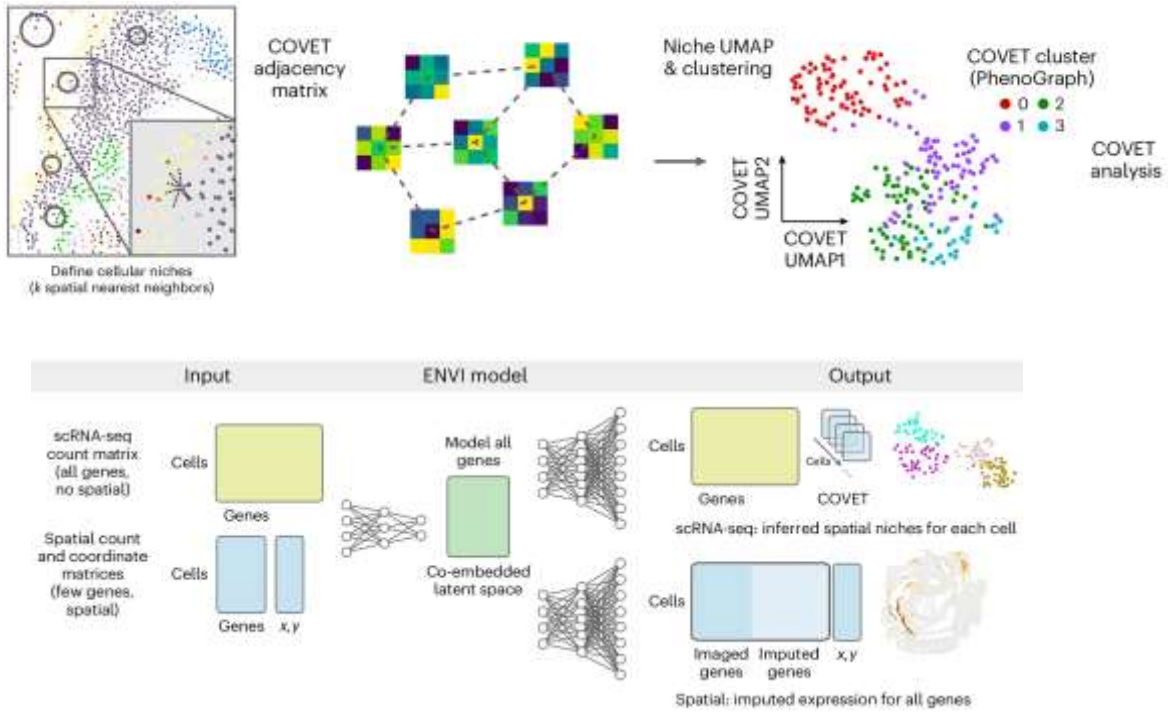
Segmentation Within Selected Areas

Orthogonal Projection Onto Resp. Axis

Spatial Analysis

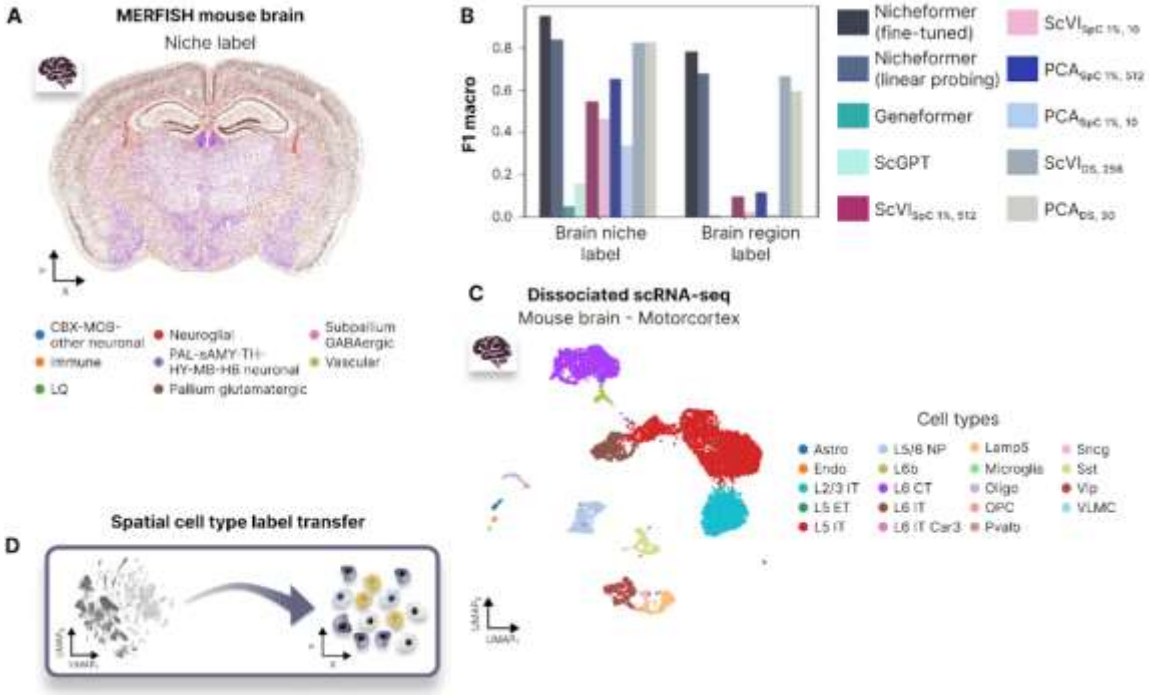
Niche reconstruction algorithms

Environmental variational inference (ENVI) using the covariance environment (COVET)



Haviv D, et al.The covariance environment defines cellular niches for spatial inference. Nat Biotechnol. 2024 Apr 2:10.1038

Transformer model for spatial label prediction (NicheFormer)



Schaar A.C. et al. Nicheformer: a foundation model for single-cell and spatial omics. Biorxiv 2024

How to turn any CCC tool into a spatial one

Approach 1: pre-filtering

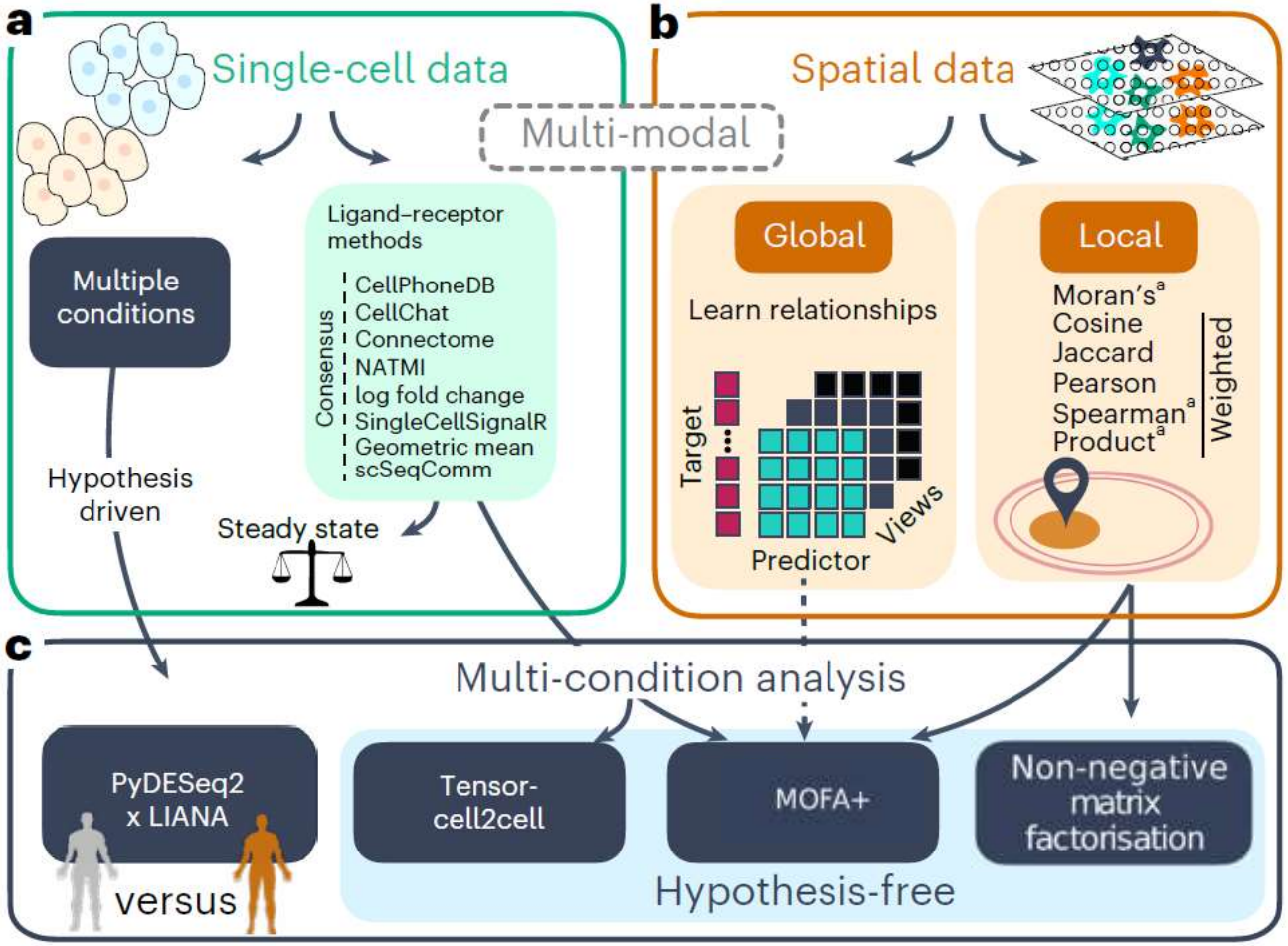
- Select a niche/neighbourhood/ domain/group of cells either manually or computationally
- Apply CCC tool of interest to the selected cells / spots

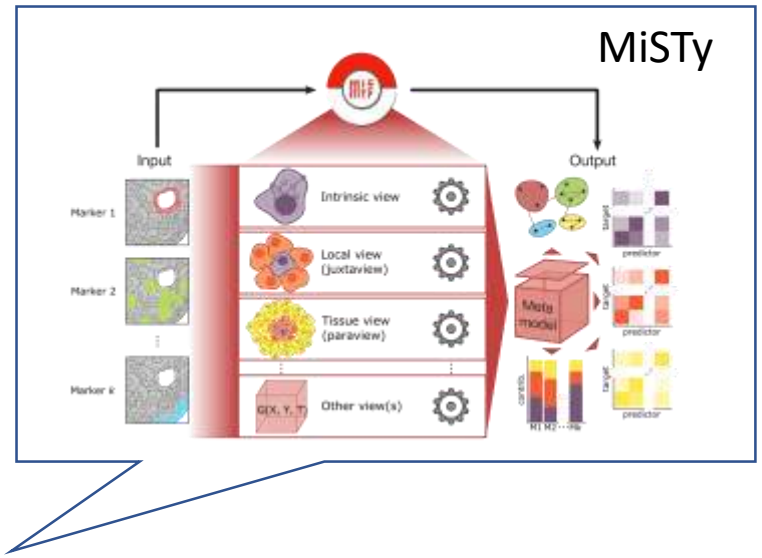
Approach 2: post-filtering

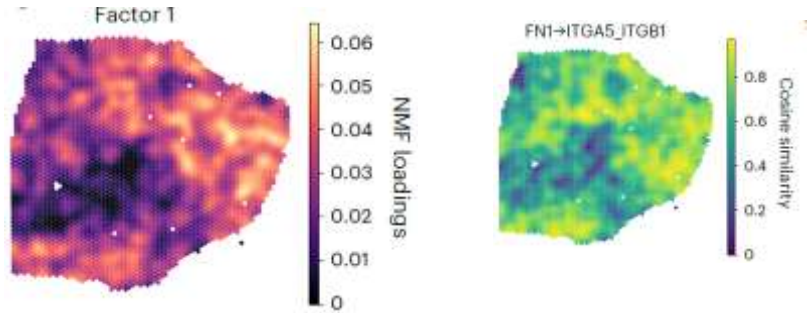
- Select cells / spots of interest
- Apply CCC tool of interest to the selected cells / spots
- Filter the ranked LRIs based on co-localization

Simple, but offers a lot of control and is easy to run and interpret

Liana+: a flexible CCC modelling framework

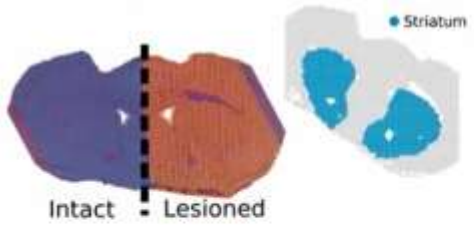


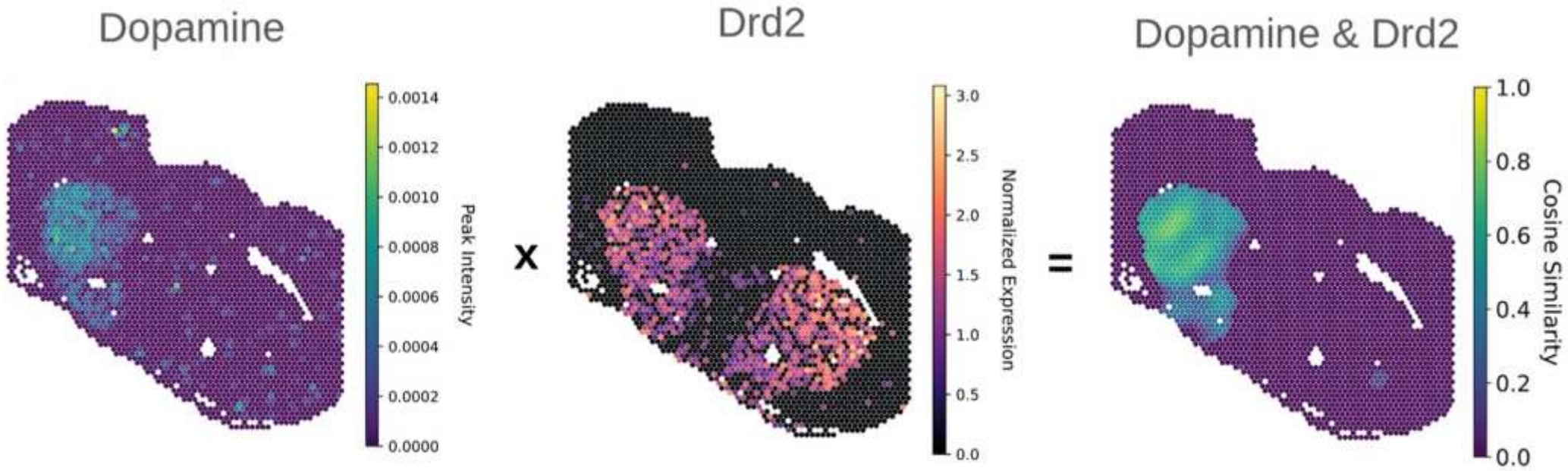




Dimitrov, D., Schäfer, P.S.L., Farr, E. *et al.* LIANA+ provides an all-in-one framework for cell–cell communication inference. *Nat Cell Biol* **26**, 1613–1622 (2024)

LIANA+ Example: spatial transcriptomics and metabolomics in brain lesions





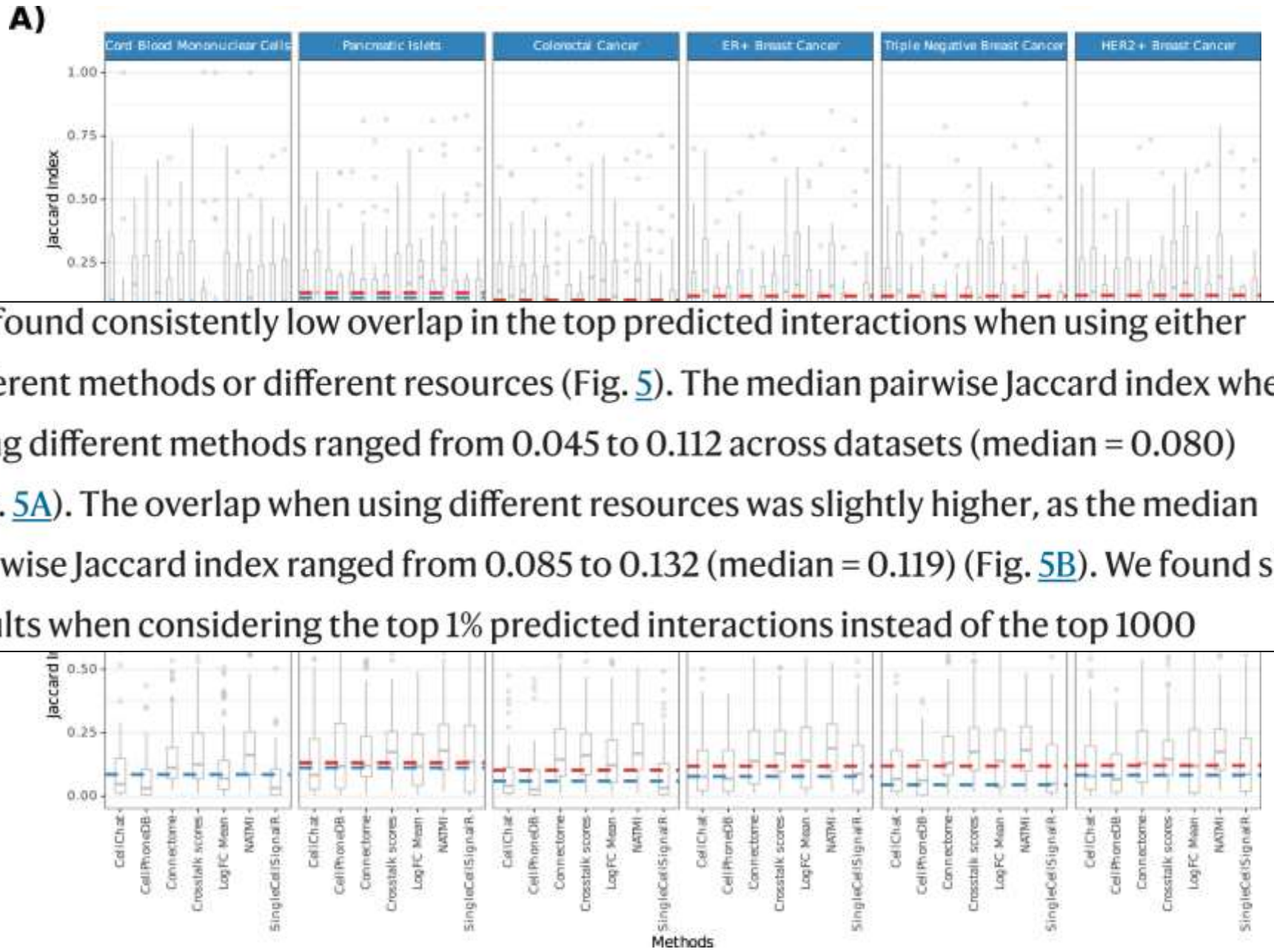
Validation and benchmarking of CCC tools and their results

VIB

SCIENCE MEETS LIFE

Same Resource with different Methods

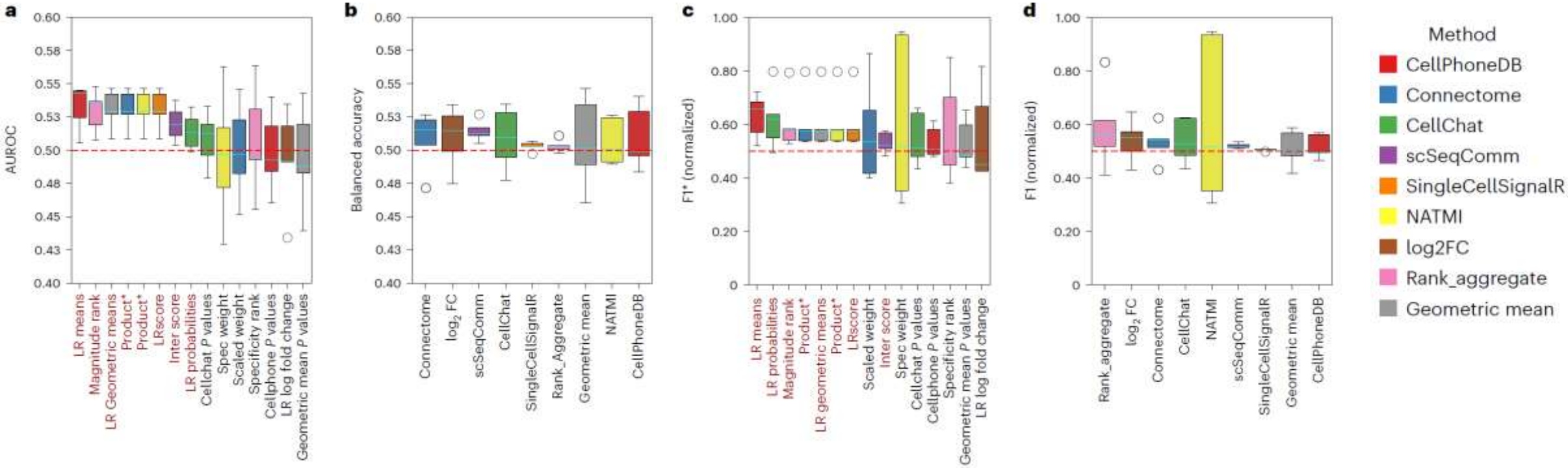
Same Method with different Resources



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

Dimitrov, D.,. *et al.* Comparison of methods and resources for cell-cell communication inference from single-cell RNA-Seq data. *Nat Commun* **13**, 3224 (2022)

Validation using spatial co-localization of cell types and ligand–receptors as assumed truth.

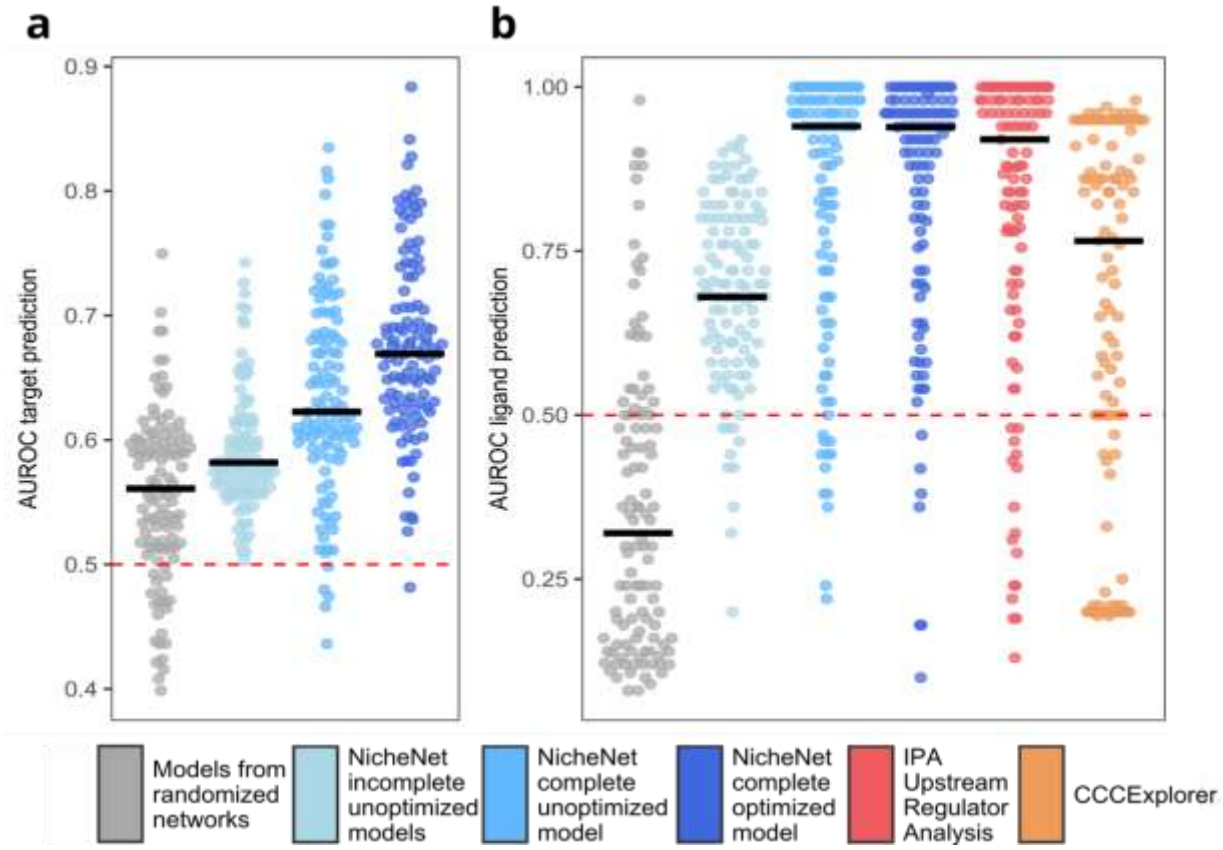


Dimitrov, D., Schäfer, P.S.L., Farr, E. *et al.* LIANA+ provides an all-in-one framework for cell–cell communication inference. *Nat Cell Biol* **26**, 1613–1622 (2024)

CCC benchmarkings are crucially lacking

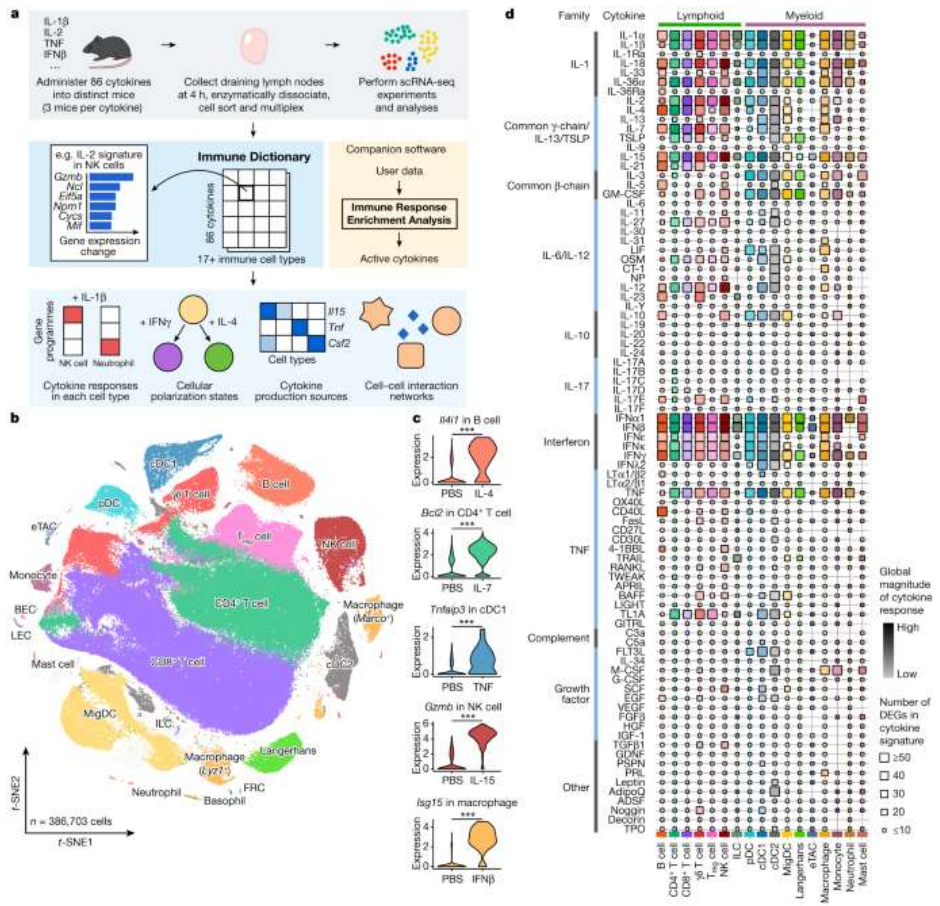
- We collected 111 transcriptome datasets of ligand-stimulated cells.
- How well does NicheNet predict all DE genes after ligand stimulation?
- How well does NicheNet predict the active ligand, given the set of DE genes?

Evaluating NicheNet's prior model



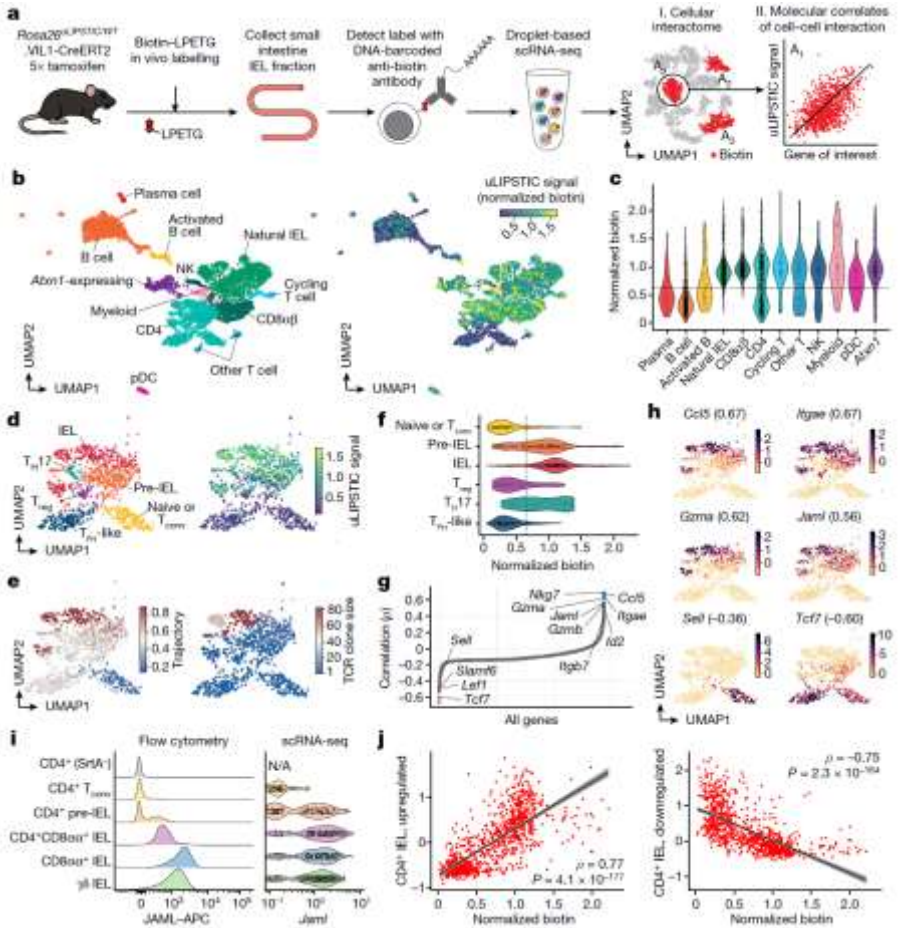
Browaeys, R., Saelens, W. & Saeys, Y. NicheNet: modeling intercellular communication by linking ligands to target genes. *Nat. Methods* **17**, 159–162 (2020)

Towards ground truth datasets to benchmark CCC models



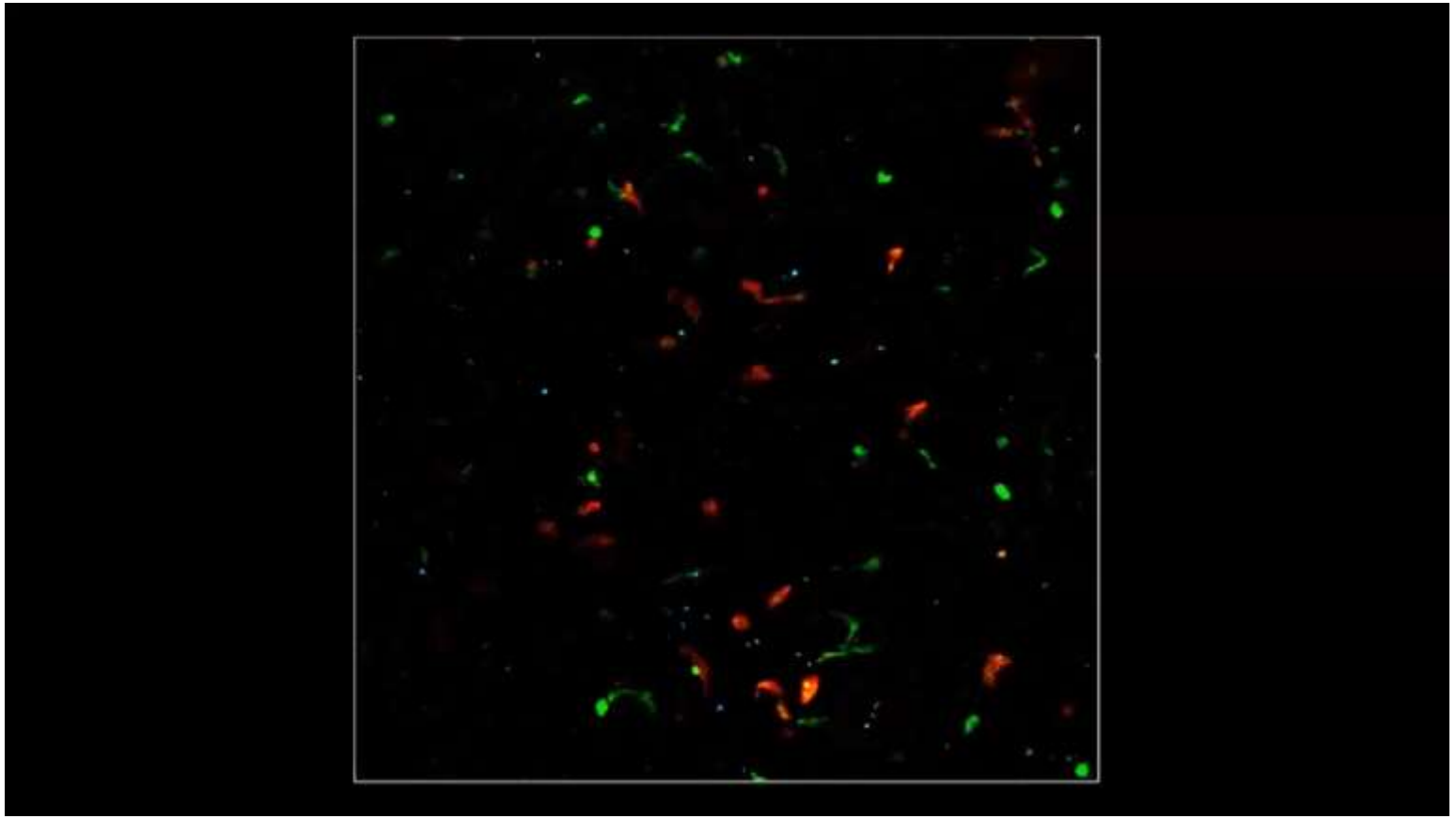
Cui, A., Huang, T., Li, S. *et al.* Dictionary of immune responses to cytokines at single-cell resolution. *Nature* **625**, 377–384 (2024)

uLIPSTIC



Nakandakari-Higa, S., Walker, S., Canesso, M.C.C. *et al.* Universal recording of immune cell interactions in vivo. *Nature* (2024)

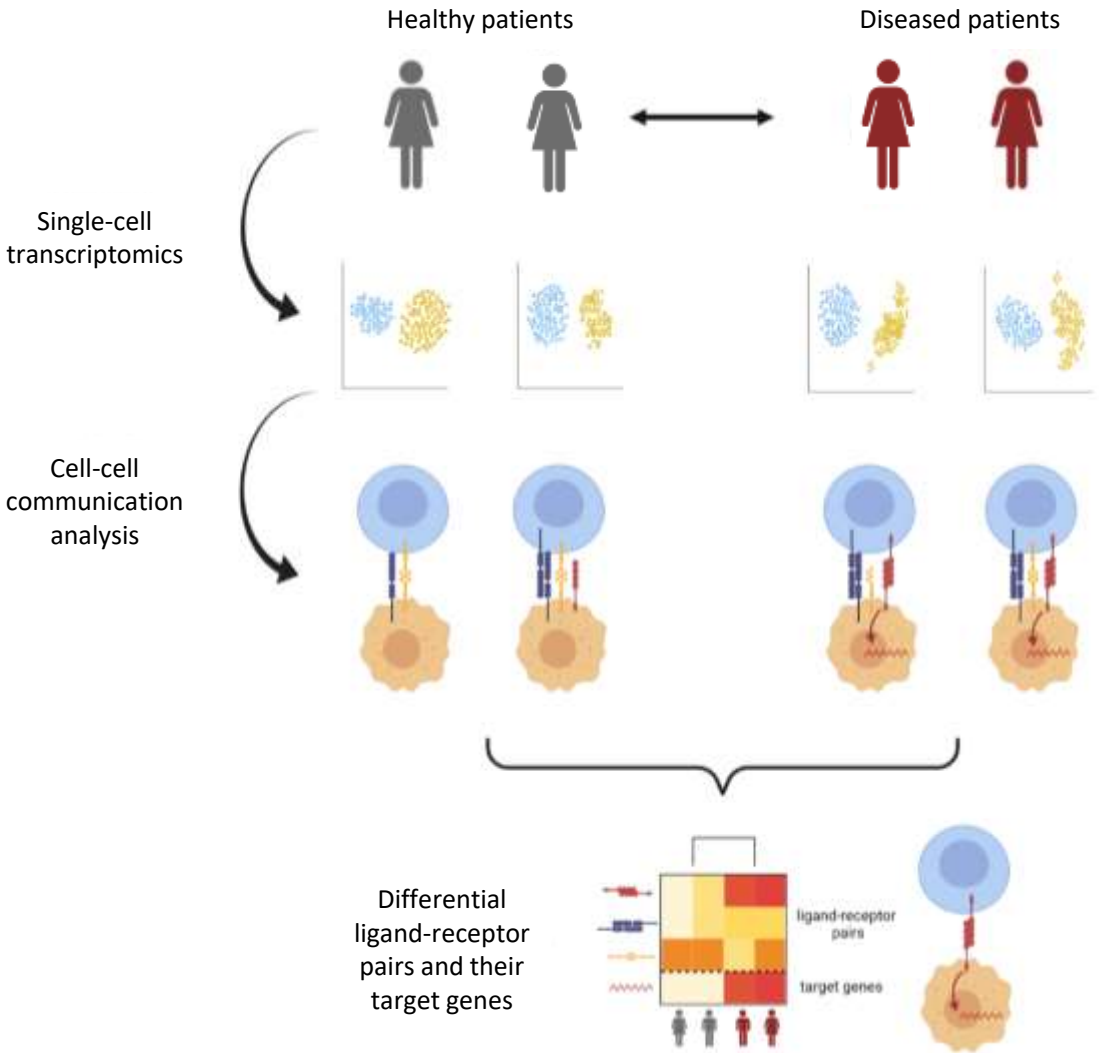
In vivo interaction dynamics using multi-photon imaging



Bonnardel J, et al. Stellate Cells, Hepatocytes, and Endothelial Cells Imprint the Kupffer Cell Identity on Monocytes Colonizing the Liver Macrophage Niche. Immunity. 2019 15;51(4):638-654

New avenues for CCC modelling

MultiNicheNet prioritizes differentially expressed and active ligand-receptor pairs between different conditions from multi-sample (spatial) transcriptomics data

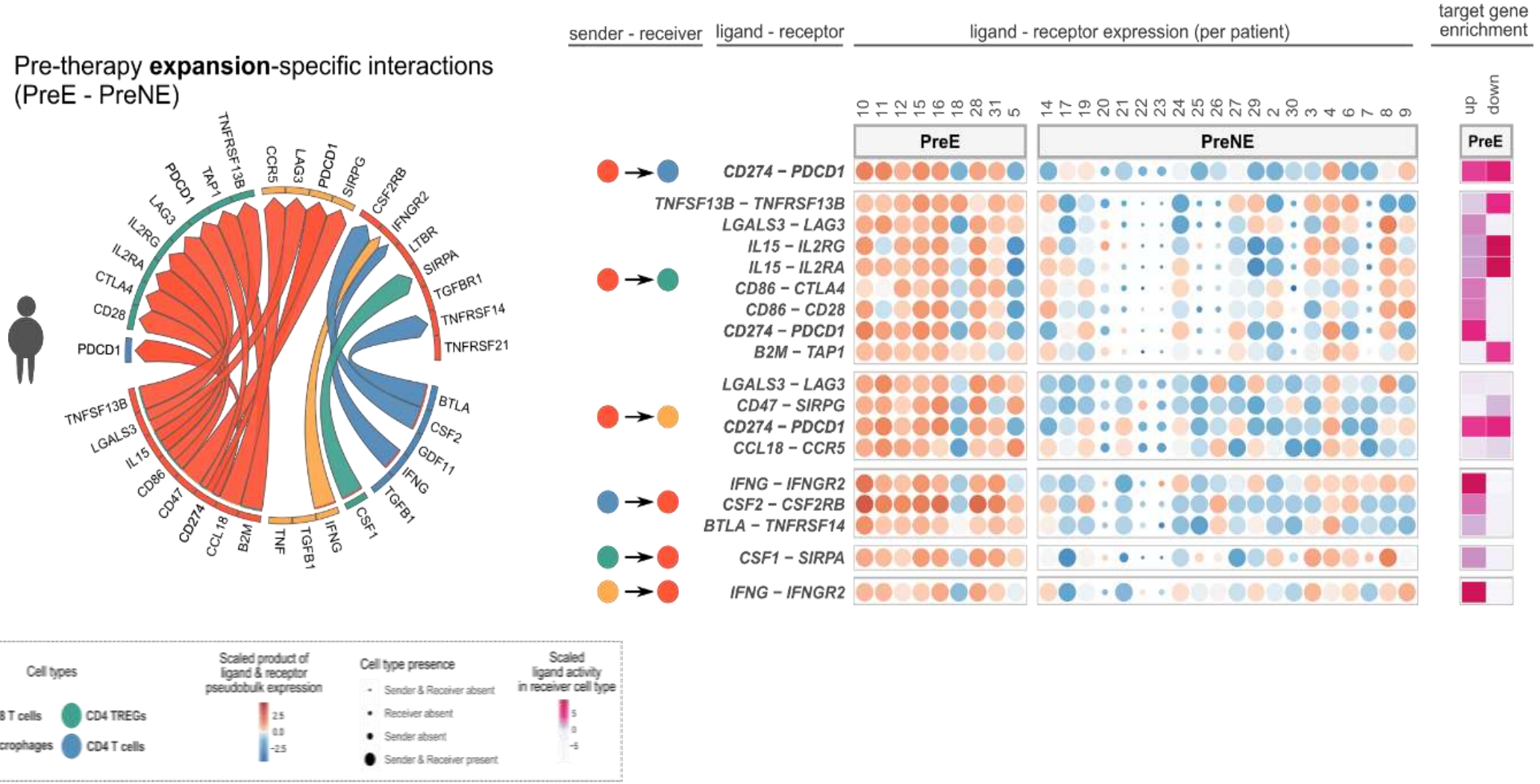


In contrast to other CCC tools, MultiNicheNet:

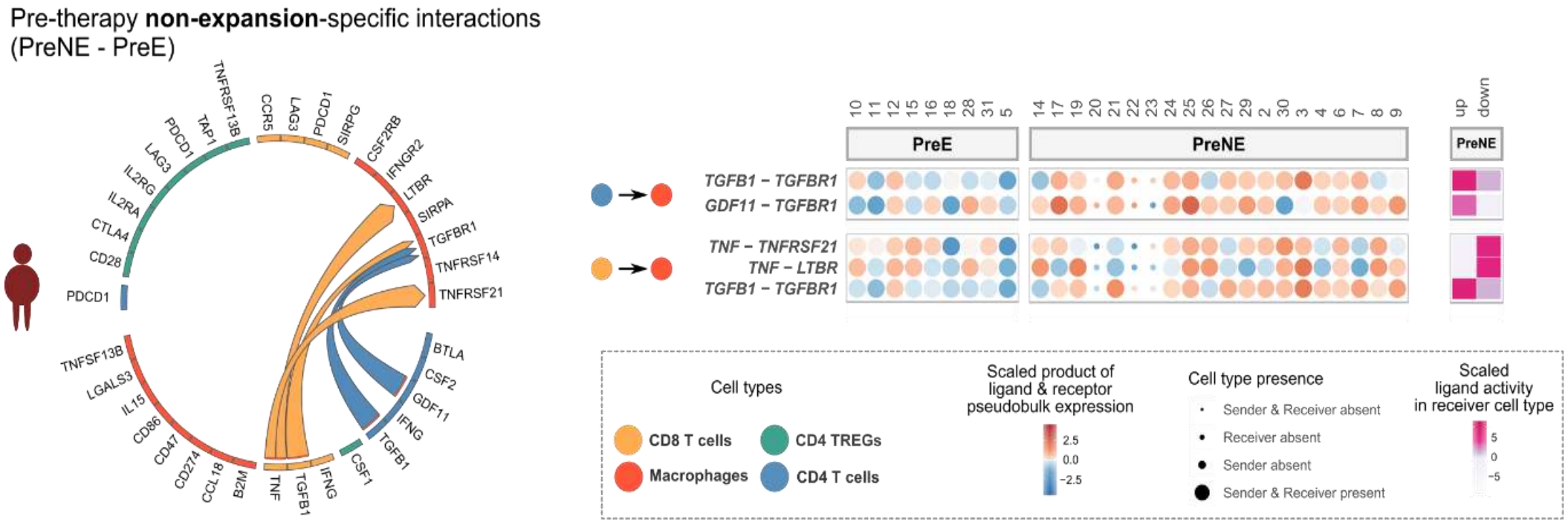
- does not pool cells across samples
- is based on a sound statistical framework
- takes into account **inter-patient heterogeneity**
- addresses **complex questions**
- corrects for **batch effects** or other covariates
- offers **flexibility in the prioritization** schemes
- provides intuitive and insightful **visualizations** to explore the top predictions

Browaeys, R. et al. MultiNicheNet: a flexible framework for differential cell-cell communication analysis from multi-sample multi-condition single-cell transcriptomics data. bioRxiv 2024

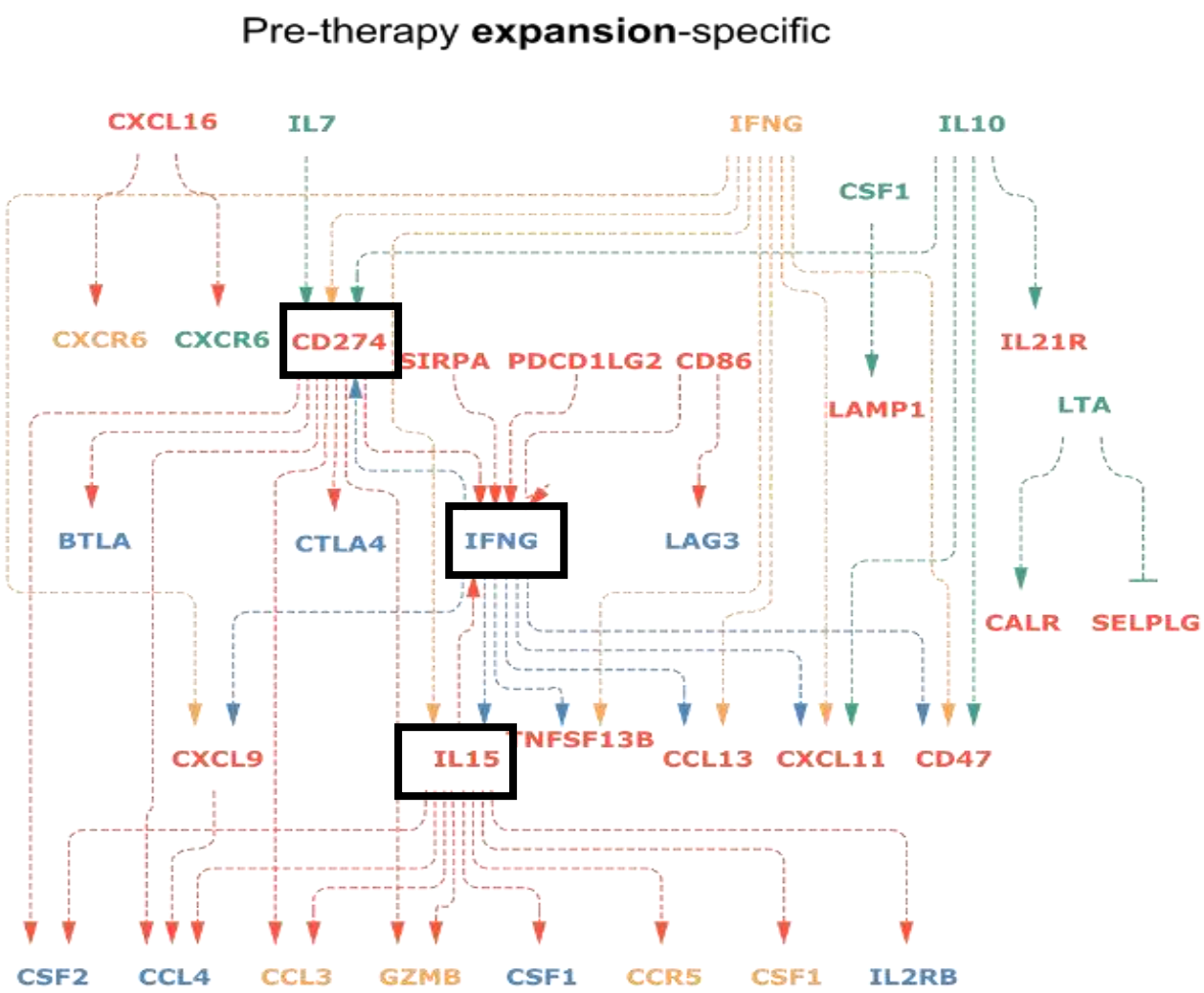
Differentially expressed and active ligand-receptor pairs between macrophages and T-cells

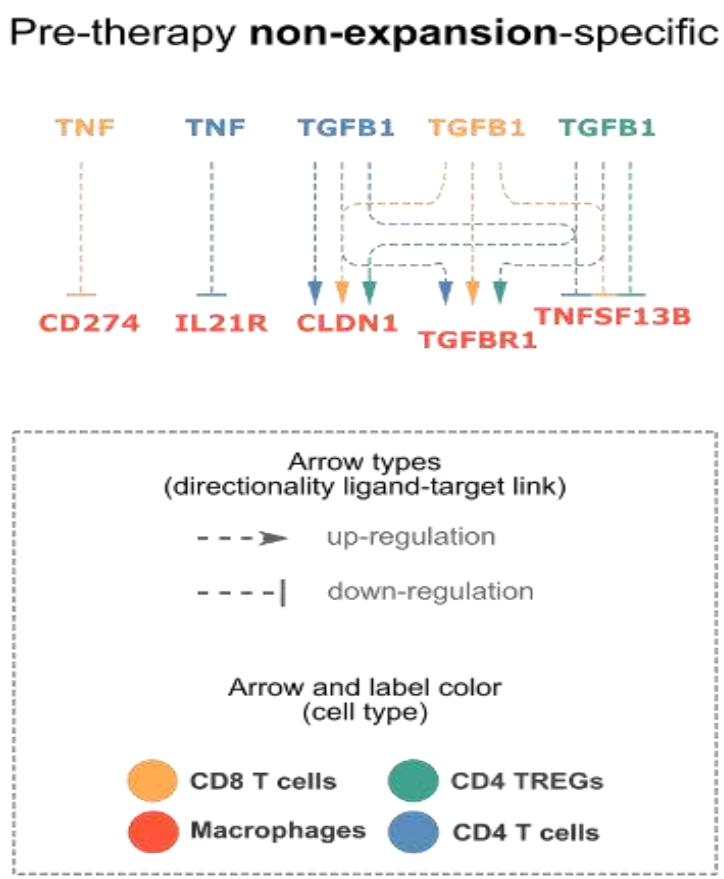


Differentially expressed and active ligand-receptor pairs between macrophages and T-cells

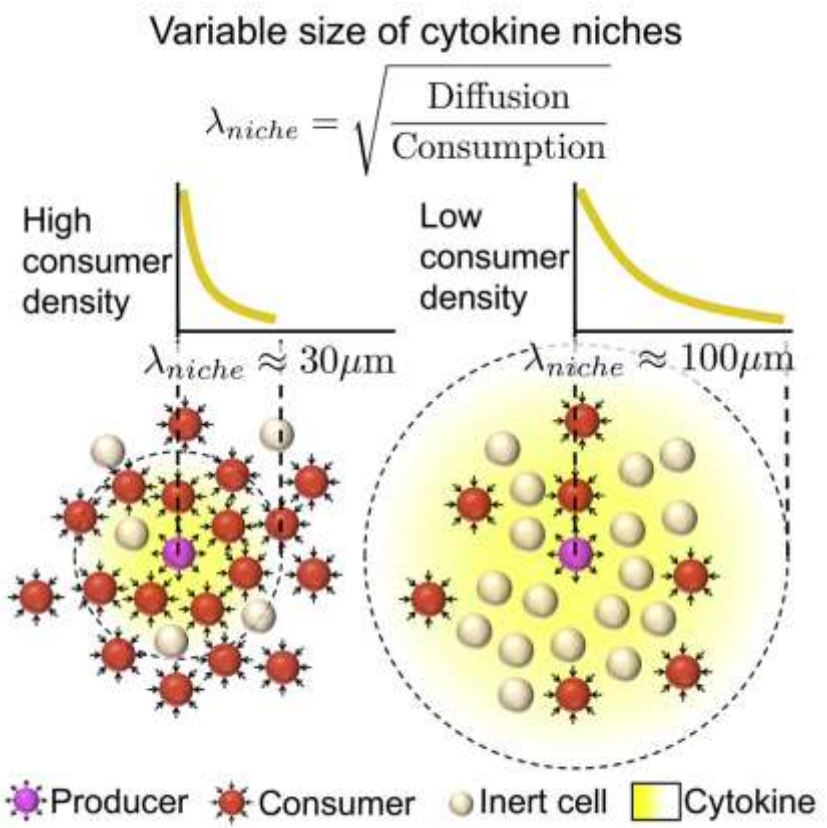


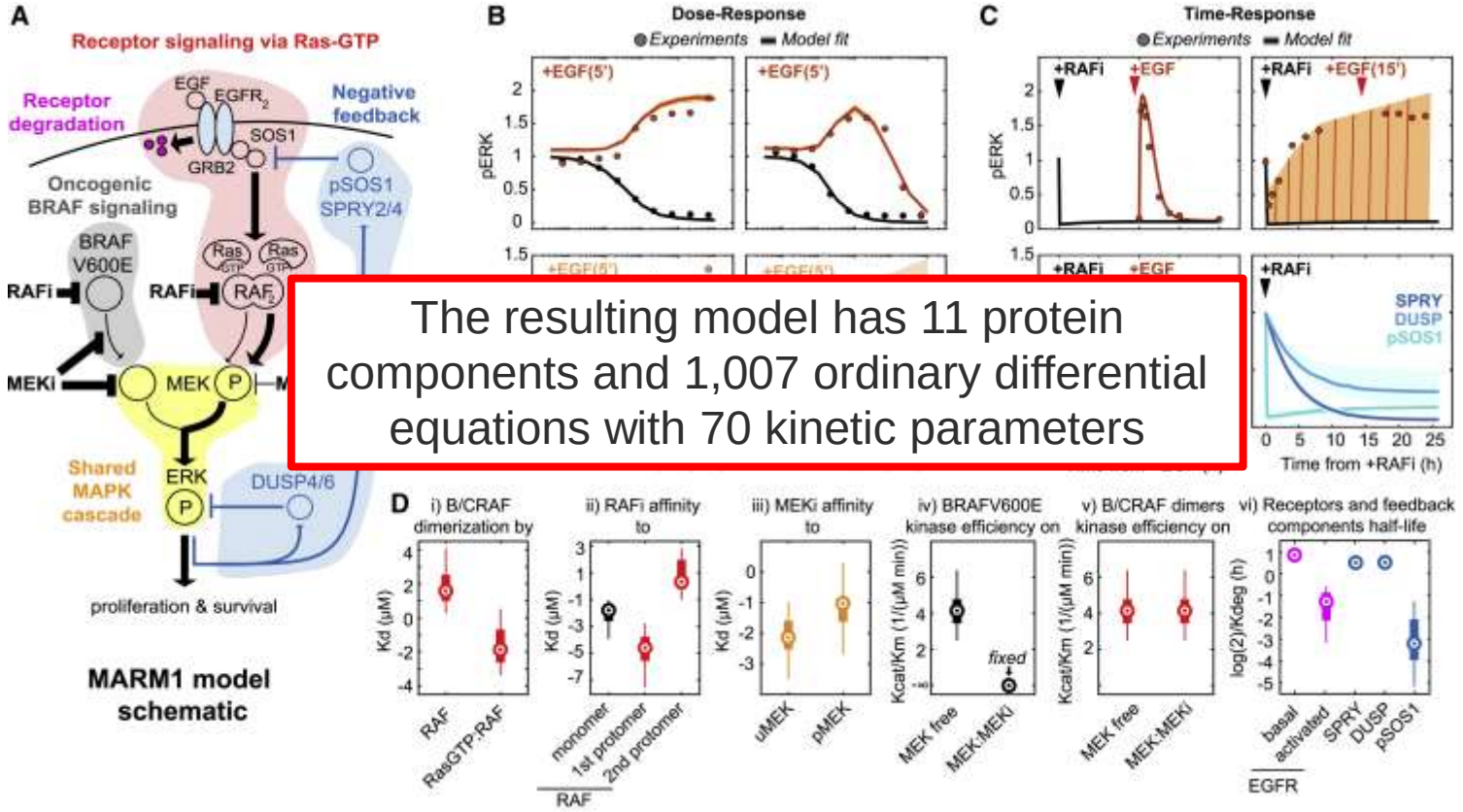
Differential intercellular signalling cascades





Dynamic models of cell-cell interactions





The resulting model has 11 protein components and 1,007 ordinary differential equations with 70 kinetic parameters

Oyler-Yaniv A, et al.. A Tunable Diffusion-Consumption Mechanism of Cytokine Propagation Enables Plasticity in Cell-to-Cell Communication in the Immune System. Immunity. 2017 Apr 18;46(4):609-620

Gerosa L, et al. Receptor-Driven ERK Pulses Reconfigure MAPK Signaling and Enable Persistence of Drug-Adapted BRAF-Mutant Melanoma Cells. Cell Syst. 2020 Nov 18;11(5):478-494

Acknowledgements

Yvan Saeys lab

Collaborators:

Martin Guilliams lab
Charlotte Scott lab
Bart Lambrecht lab
Chris Marine lab
Julio Saez-Rodriguez lab

Robin Browaeys


VIB
science meets life