Canadian Bioinformatics Workshops

www.bioinformatics.ca

bioinformaticsdotca.github.io

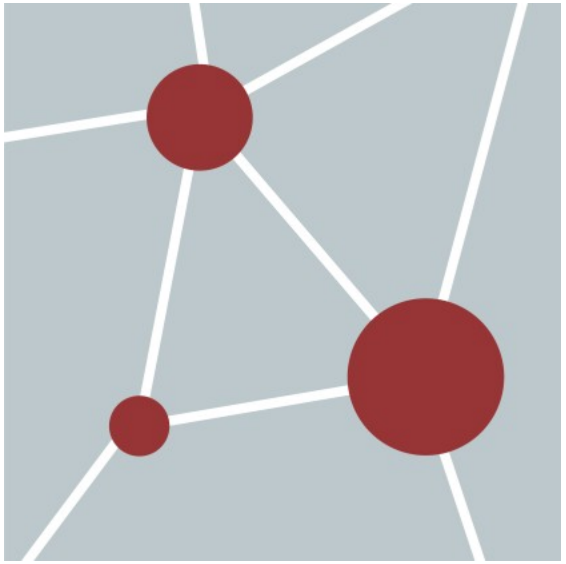
Module #6: Cell-Cell Communication

Gregory Schwartz

Pathway and Network Analysis

June 26-28, 2024

bioinformatics.ca



UHN Princess Margaret Cancer Centre

Medical Biophysics
UNIVERSITY OF TORONTO

SCHWARTZ LAB

What you will learn

The role and context of cell-cell communication

Single-cell RNA sequencing for cellular transcriptomes

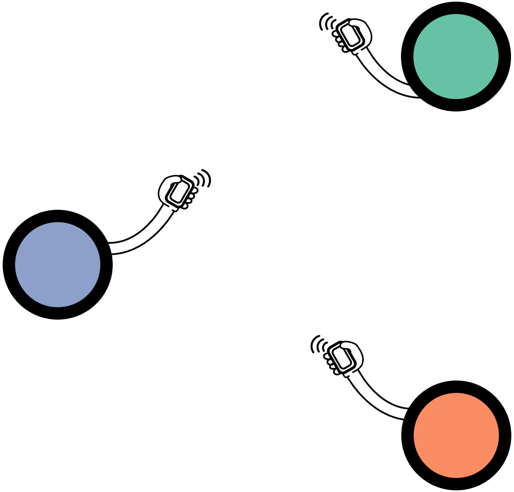
General principles of cell-cell communication detection

Inner workings of select single-cell RNA sequencing methods

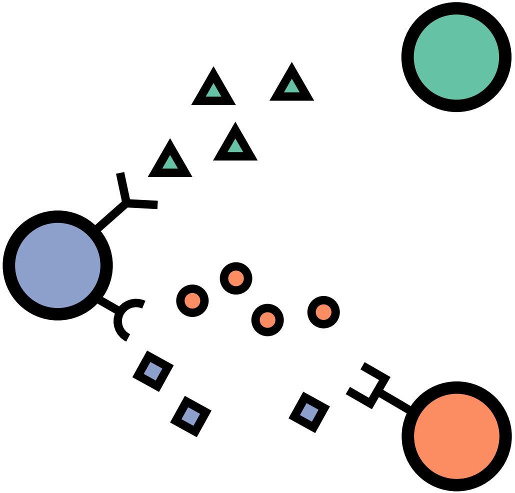
Spatial transcriptomics for communication detection

Spatially-resolved cell-cell communication method

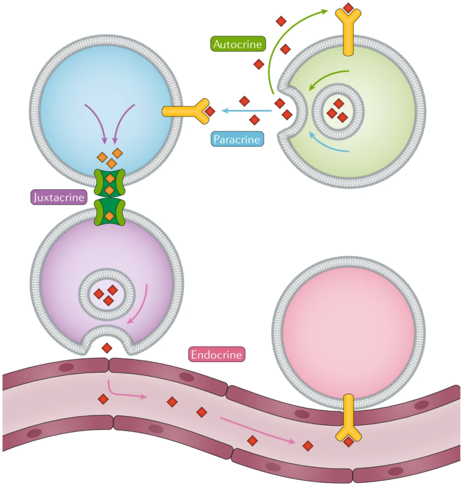
Cells communicate through intercellular signaling mechanisms



Cells communicate through intercellular signaling mechanisms

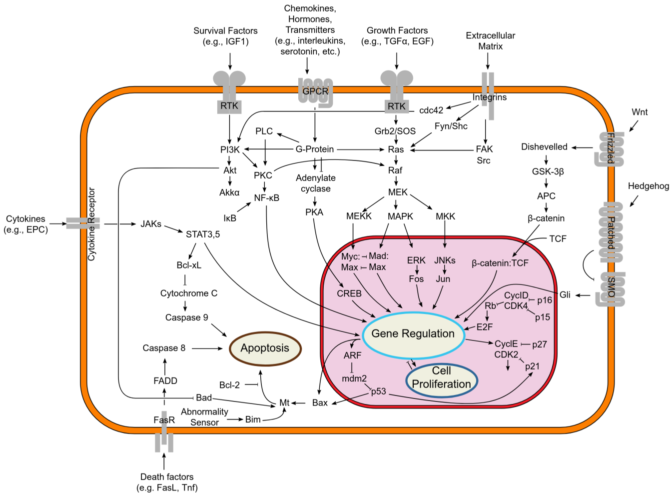


Ligands initiate different cell-cell communication mechanisms

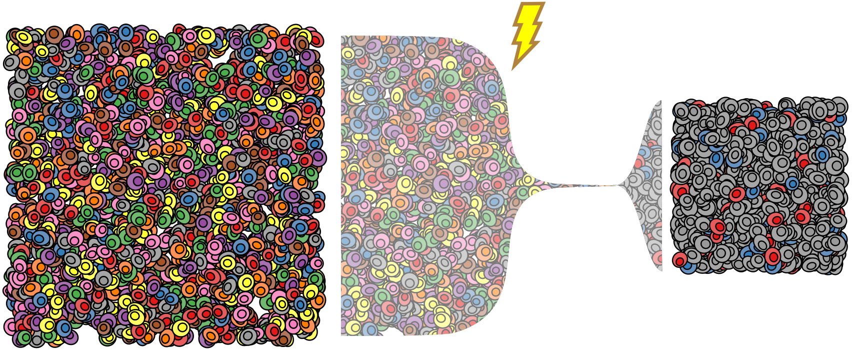


Armingol et al., *Nature Reviews Genetics*, 2021

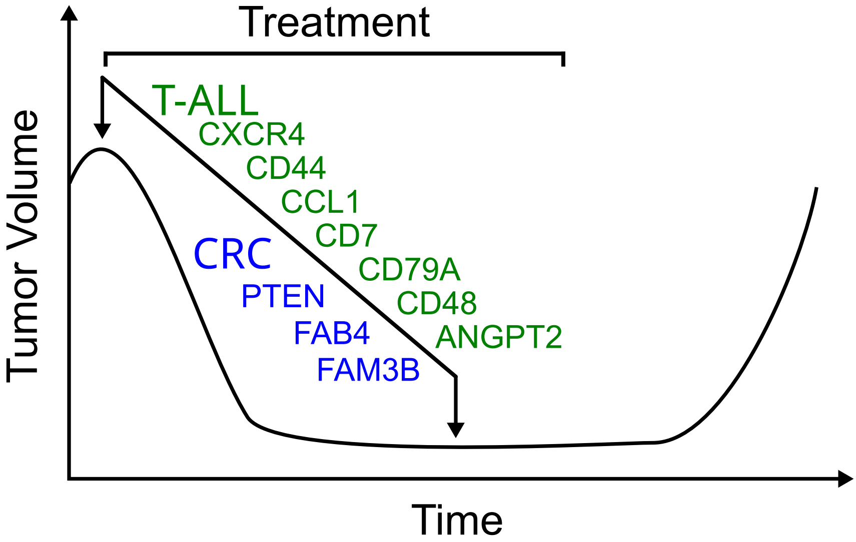
Communication triggers downstream cellular functionality

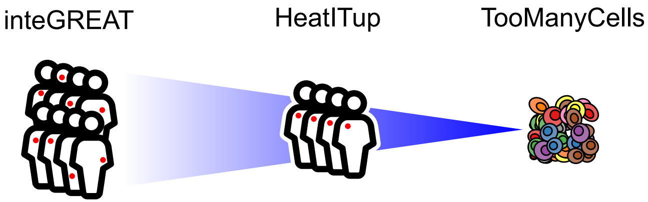


Cancer heterogeneity can reduce the efficacy of therapy

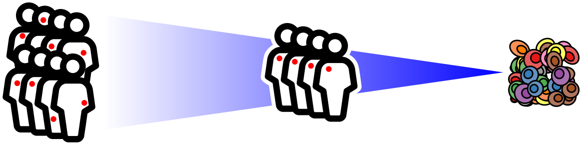


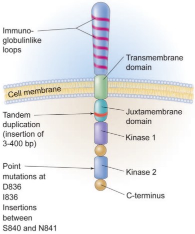
Genes involved in cell-cell communication are differentially expressed



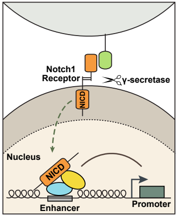

inteGREAT
HeatITup
TooManyCells

inteGREAT HeatITup TooManyCells





FLT3



Notch1

Hypothesis

Cell-cell communication contributes to progression and poor therapy response

Single-cell sequencing enables cell resolution analysis

Single-cell sequencing enables cell resolution analysis

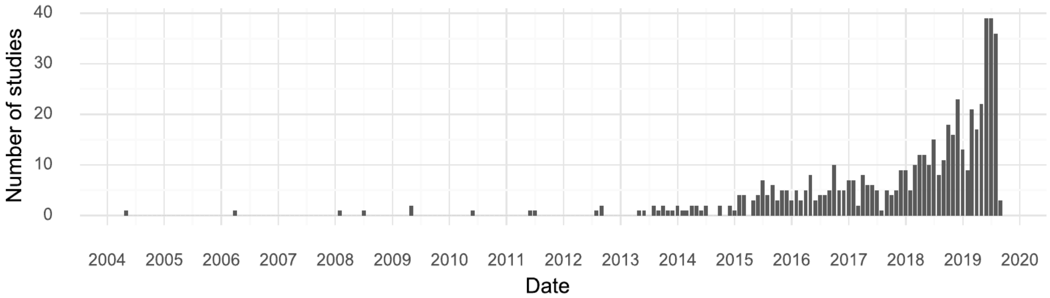
Bulk Sequencing

Single-cell sequencing enables cell resolution analysis

Bulk Sequencing

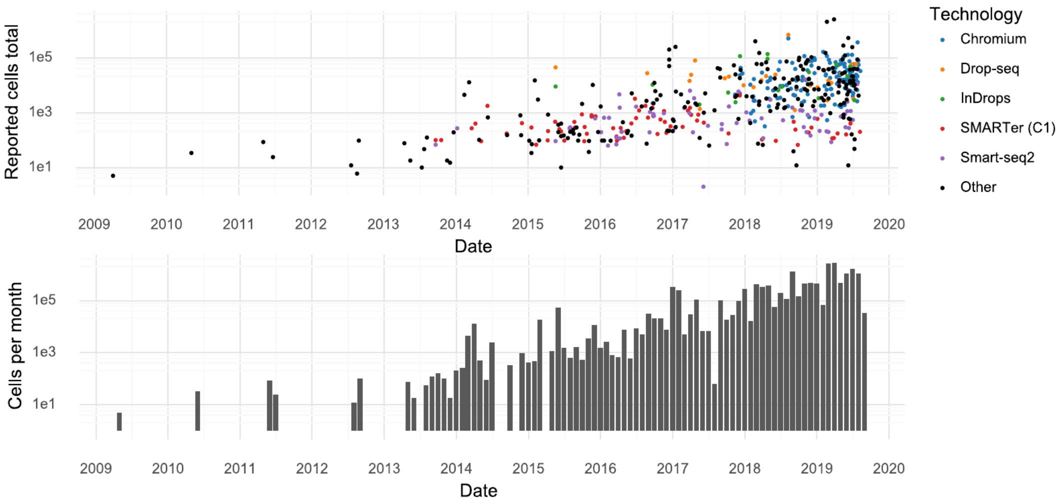
Single-cell Sequencing

Single-cell transcriptomics is exponentially increasing

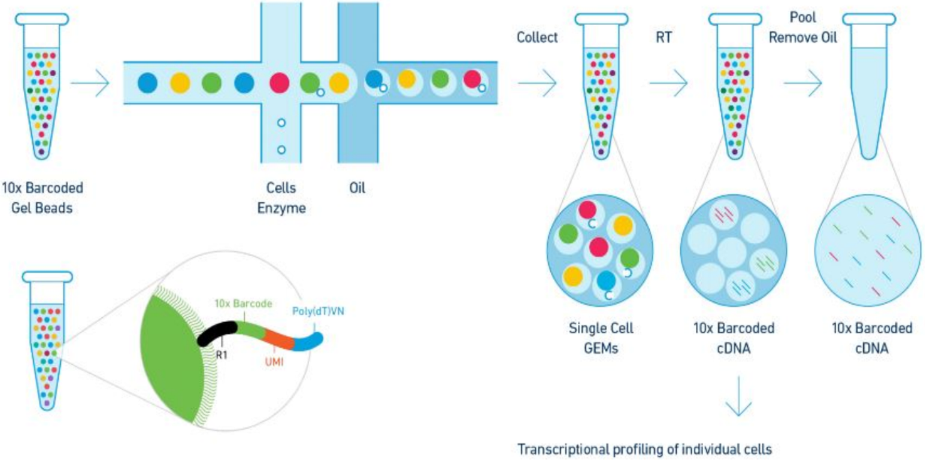


Svensson et al. *bioRxiv*, 2019.

Single-cell transcriptomics is exponentially increasing



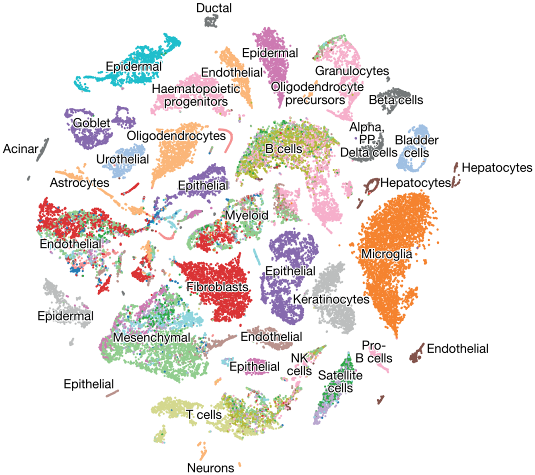
Microfluidics has high accuracy and throughput



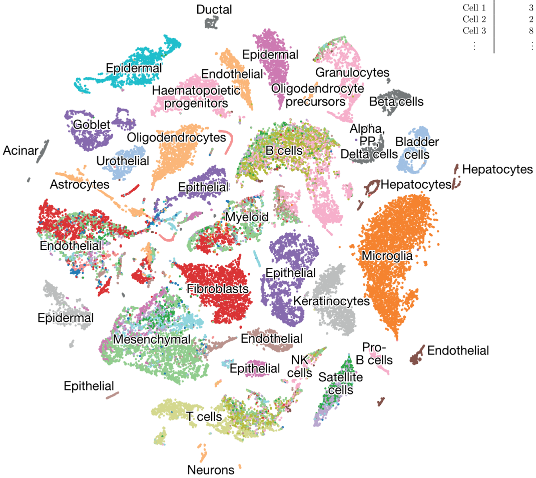
Microfluidics has high accuracy and throughput

Drop-seq (HyDrop)

Typical workflow collapses many dimensions to only two

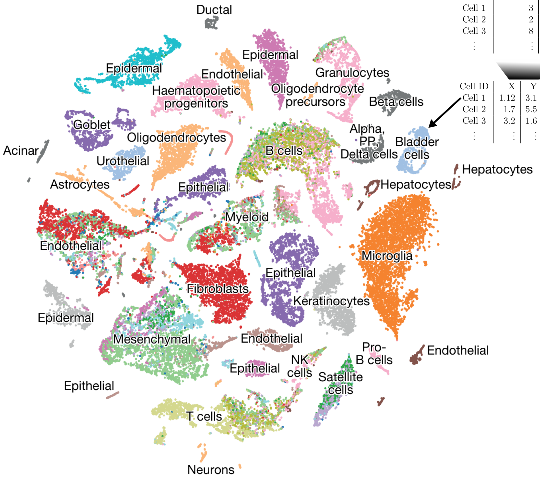


Typical workflow collapses many dimensions to only two



Cell ID	CASC11	MYC	PVT1	AC026248.1	LINC00976	...
Cell 1	3	7	5	1	8	...
Cell 2	2	1	4	1	3	...
Cell 3	8	3	5	3	2	...
⋮	⋮	⋮	⋮	⋮	⋮	...

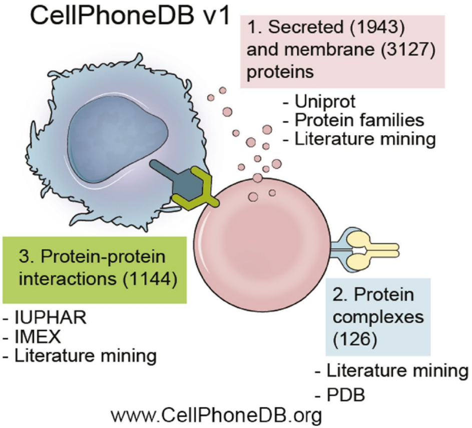
Typical workflow collapses many dimensions to only two



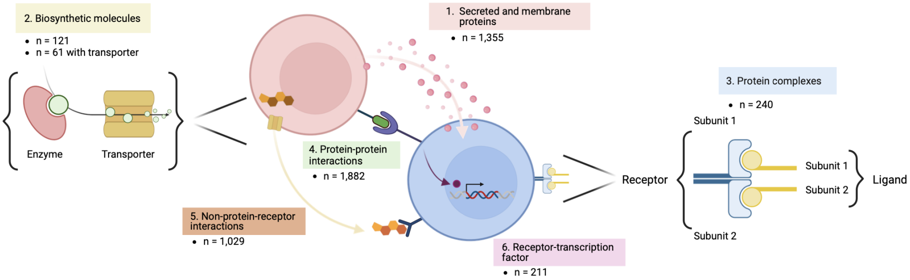
Cell ID	CASC11	MYC	PVT1	AC026248.1	LINC00976	...
Cell 1	3	7	5	1	8	...
Cell 2	2	1	4	1	3	...
Cell 3	8	3	5	3	2	...
⋮	⋮	⋮	⋮	⋮	⋮	...

Cell ID	X	Y
Cell 1	1.12	3.1
Cell 2	1.7	5.5
Cell 3	3.2	1.6
⋮	⋮	⋮

CellPhoneDB introduced a well-curated data set of signals



CellPhoneDB introduced a well-curated data set of signals



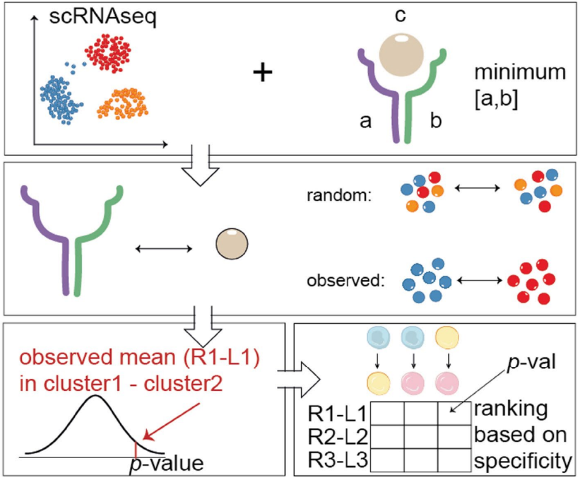
Flat table structure holds molecular interactions

id_interaction	id_cp_interaction	multidata_1_id	multidata_2_id	source	annotation_strategy	is_ppi	curator	directionality	classification
0	CPI-SC0A2DB962D	329	1507	PMID:12392763	curated	True	RVentoTormo	Adhesion-Adhesion	Adhesion by Cadherin
1	CPI-SC0B5CEA47D	716	1507	uniprot	curated	True	RVentoTormo	Adhesion-Adhesion	Adhesion by Collagen/Integrin
2	CPI-SC0C8B7BCBB	322	1507	uniprot	curated	True	RVentoTormo	Adhesion-Adhesion	Adhesion by Collagen/Integrin
3	CPI-SC0D3C12C3F	343	1507	uniprot	curated	True	RVentoTormo	Adhesion-Adhesion	Adhesion by Collagen/Integrin
4	CPI-SC0B86B7CED	930	1507	uniprot	curated	True	RVentoTormo	Adhesion-Adhesion	Adhesion by Collagen/Integrin
5	CPI-SC0FA343CEF	810	1507	uniprot	curated	True	RVentoTormo	Adhesion-Adhesion	Adhesion by Collagen/Integrin
6	CPI-SC0CCCF9A7F	720	1507	uniprot	curated	True	RVentoTormo	Adhesion-Adhesion	Adhesion by Collagen/Integrin
7	CPI-SC0E85E1FB0	1347	1507	uniprot	curated	True	RVentoTormo	Adhesion-Adhesion	Adhesion by Collagen/Integrin

Flat table structure holds molecular interactions

	id_multidata	name	receptor	receptor_desc	other	other_desc	secreted_highlight	secreted_desc	transmembrane	secreted	peripheral	integrin	is_complex
0		P03372	True		False		False		False	False	True	False	False
1		Q92753	True		False		False		False	False	False	False	False
2		O95477	False		False		False		True	False	False	False	False
3		Q13133	True		False		False		False	False	False	False	False
4		P04150	True		False		False		False	False	False	False	False
5		Q7Z5A7	False		False		True		False	True	False	False	False
6		Q9UNN8	True		False		False		False	False	False	False	False
7		Q96LR4	False		False		True		False	True	False	False	False
8		O60883	True		False		False		True	False	False	False	False
9		Q96RI1	True		False		False		False	False	False	False	False
10		P30533	False		False		False		False	False	False	False	False

CellPhoneDB compares high expression of each cell cluster



Pre-set thresholds control for noise

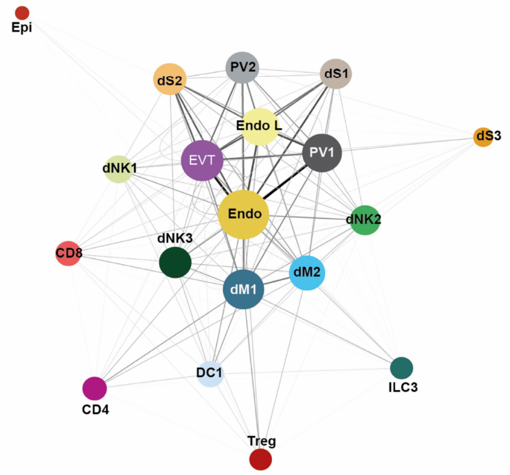
Only ligands and receptors in >10% of cells considered

Randomly permute cluster labels 1000 times

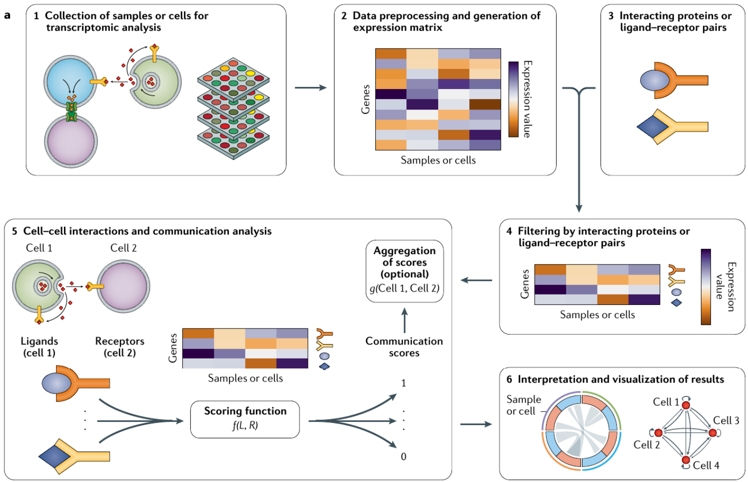
Calculate proportion of means “more extreme” than observed

The minimum of protein dimers are considered

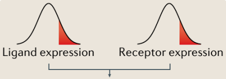
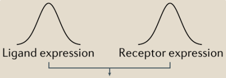
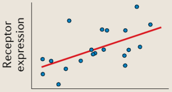
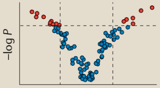
CellPhoneDB compares high expression of each cell cluster



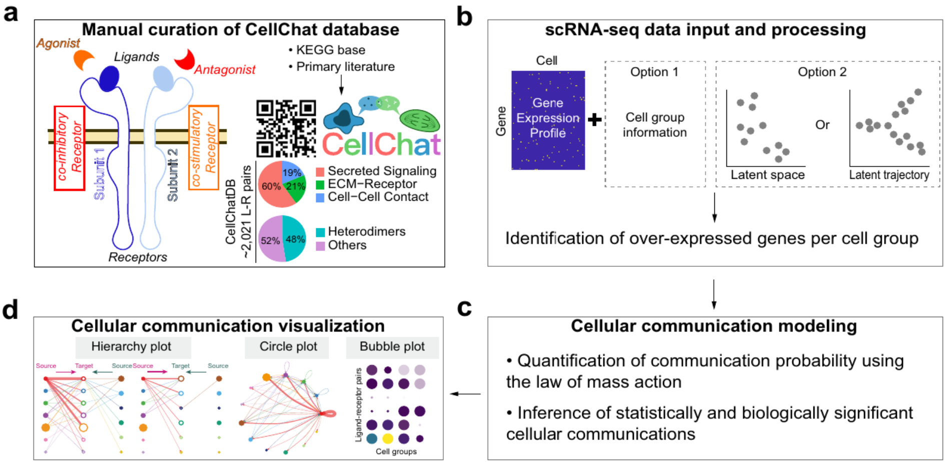
Cell-cell communication algorithms decipher ligand-receptor pairs



Cell-cell communication algorithms decipher ligand-receptor pairs

		Recommended data	Communication score
Expression thresholding	 Ligand expression; Receptor expression; **(L > x) and (R > x)**	**Bulk, single cell**	**Binary**
Expression product	 Ligand expression; Receptor expression; **$L \times R$**	**Single cell**	**Continuous**
Expression correlation	 Receptor expression; Ligand expression	**Bulk, single cell**	**Continuous**
Differential combinations	 –log P; Fold change between ligand and receptor samples	**Bulk, single cell**	**Binary**

CellChat alternatively uses law of mass action



CellChat's estimation is based on Hill equations

$$\theta = \frac{[L]^n}{K_d + [L]^n} \quad (1)$$

θ = Fraction of receptor protein bound by ligand

$[L]$ = Ligand concentration

K_d = Dissociation constant from law of mass action (half binding)

n = Hill coefficient

CellChat's estimation is based on Hill equations

$$P^k_{i,j} = \frac{L_i R_j}{K_h + L_i R_j} \times \left(1 + \frac{AG_i}{K_h + AG_i}\right) \cdot \left(1 + \frac{AG_j}{K_h + AG_j}\right) \times \frac{K_h}{K_h + AN_i} \cdot \frac{K_h}{K_h + AN_j} \times \frac{n_i n_j}{n^2},$$

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

CellChat's estimation is based on Hill equations

Ligand/
receptor law of
mass action

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

$$\times \frac{K_h}{K_h + AN_i} \cdot \frac{K_h}{K_h + AN_j} \times \frac{n_i n_j}{n^2},$$

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

CellChat's estimation is based on Hill equations

Ligand/
receptor law of
mass action

$$P^k_{i,j} = \boxed{\frac{L_iR_j}{K_h + L_iR_j}} \times \left(1 + \frac{AG_i}{K_h + AG_i}\right) \cdot \left(1 + \frac{AG_j}{K_h + AG_j}\right) \times \frac{K_h}{K_h + AN_i} \cdot \frac{K_h}{K_h + AN_j} \times \frac{n_i n_j}{n^2},$$

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

Geometric
mean of *m*1

Geometric
mean of *m*2

CellChat's estimation is based on Hill equations

Ligand/
receptor law of
mass action

$$P_{i,j}^{k} = \boxed{\frac{L_i R_j}{K_h + L_i R_j}} \times \left(1 + \frac{AG_i}{K_h + AG_i}\right) \cdot \left(1 + \frac{AG_j}{K_h + AG_j}\right)$$
$$\times \frac{K_h}{K_h + AN_i} \cdot \frac{K_h}{K_h + AN_j} \times \frac{n_i n_j}{n^2},$$

Averages of
co-stimulatory /
co-inhibitory
receptor
modeling

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

Geometric
mean of *m*1

Geometric
mean of *m*2

CellChat's estimation is based on Hill equations

Ligand/ receptor law of mass action

Modulate with agonists

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

$$\times \frac{K_h}{K_h + AN_i} \cdot \frac{K_h}{K_h + AN_j} \times \frac{n_i n_j}{n^2},$$

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

Averages of co-stimulatory / co-inhibitory receptor modeling

Geometric mean of *m*1

Geometric mean of *m*2

CellChat's estimation is based on Hill equations

Ligand/ receptor law of mass action

Modulate with agonists

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

Modulate with antagonists

$$\times \boxed{\frac{K_h}{K_h + AN_i}} \cdot \boxed{\frac{K_h}{K_h + AN_j}} \times \frac{n_i n_j}{n^2},$$

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

Averages of co-stimulatory / co-inhibitory receptor modeling

Geometric mean of *m*1

Geometric mean of *m*2

CellChat's estimation is based on Hill equations

Ligand/
receptor law of
mass action

Modulate with
agonists

$$P^k_{i,j} = \boxed{\frac{L_i R_j}{K_h + L_i R_j}} \times \left(1 + \boxed{\frac{AG_i}{K_h + AG_i}}\right) \cdot \left(1 + \boxed{\frac{AG_j}{K_h + AG_j}}\right)$$

Modulate with
antagonists

$$\times \boxed{\frac{K_h}{K_h + AN_i}} \cdot \boxed{\frac{K_h}{K_h + AN_j}} \times \boxed{\frac{n_i n_j}{n^2}},$$

Weigh
by cell
numbers

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

Averages of
co-stimulatory /
co-inhibitory
receptor
modeling

Geometric
mean of *m*1

Geometric
mean of *m*2

Many alternative methods exist

CellPhoneDB

CellChat

Connectome

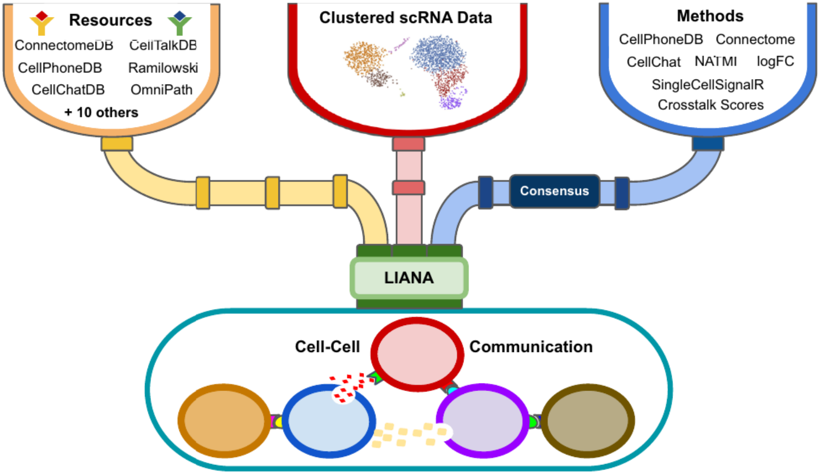
SingleCellSignalR

Scriabin

NATMI

...

LIANA aggregates many resources and methods



Daniel Dimitrov et al., *Nature Communications*, 2022.

What to do next?

Traditional differential expression analysis

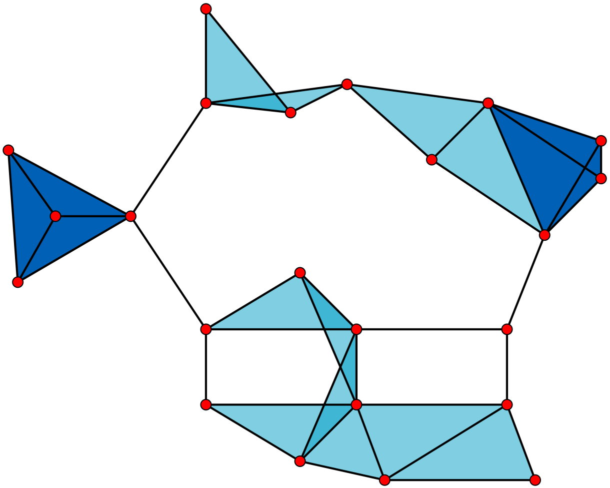
Traditional pathway analysis

Cliques

Graph centrality

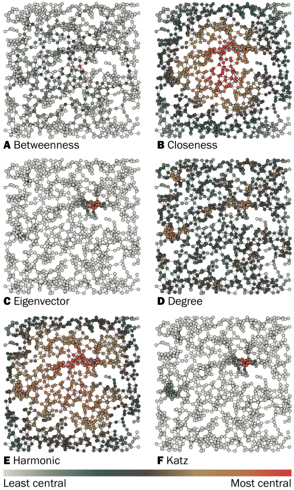
Vertex hubs

Cliques show highly-connected communities

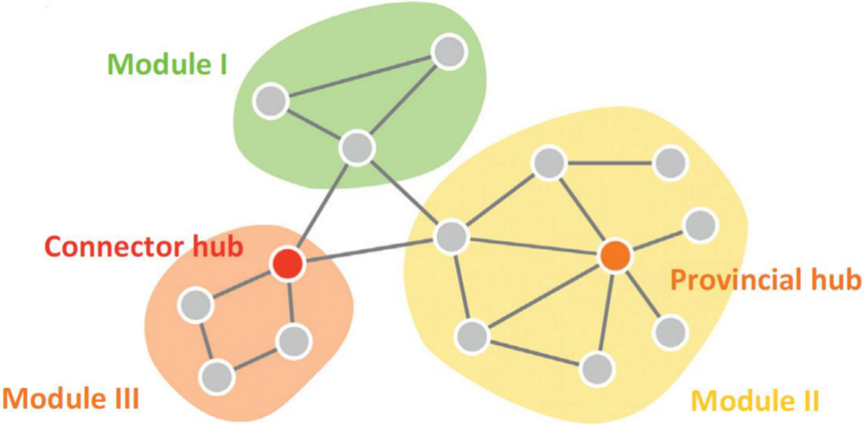


upload.wikimedia.org/wikipedia/commons/d/d0/VR_complex.svg

Graph centrality identifies vertices with easy access to most of the network



Hubs reveal highly visited vertices



Medaglia et al., *Neuroscience*, 2018

Hypothesis

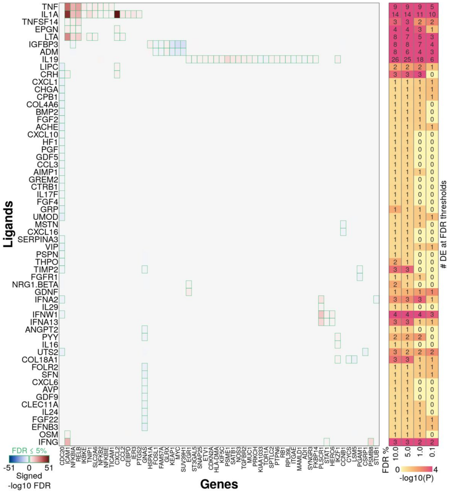
Cell-cell communication contributes to progression and poor therapy response

Hypothesis

Cell-cell communication contributes to progression and poor therapy response

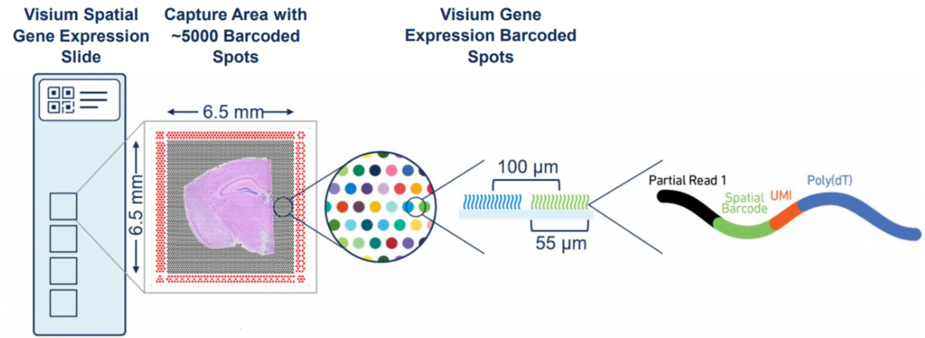
A single transcriptomic layer is insufficient to test this hypothesis

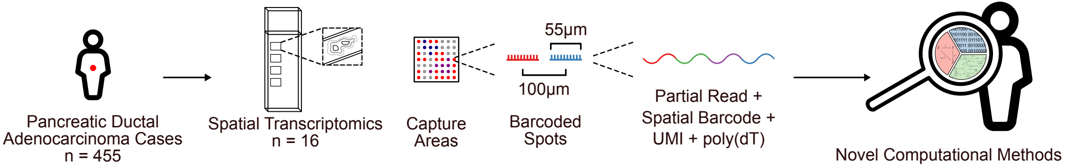
Cell lines exhibit minimal change in expression when treated with ligands



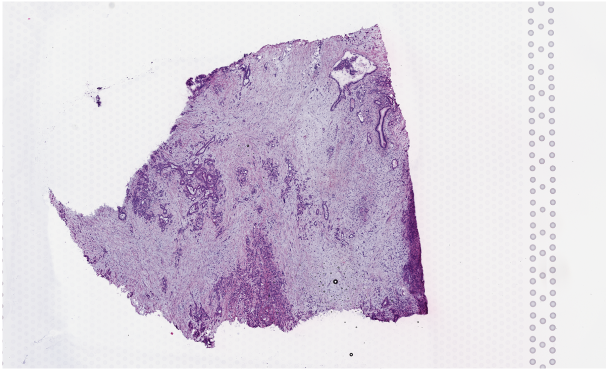
What about spatial transcriptomic data?

Spatial transcriptomics spatially resolves gene expression

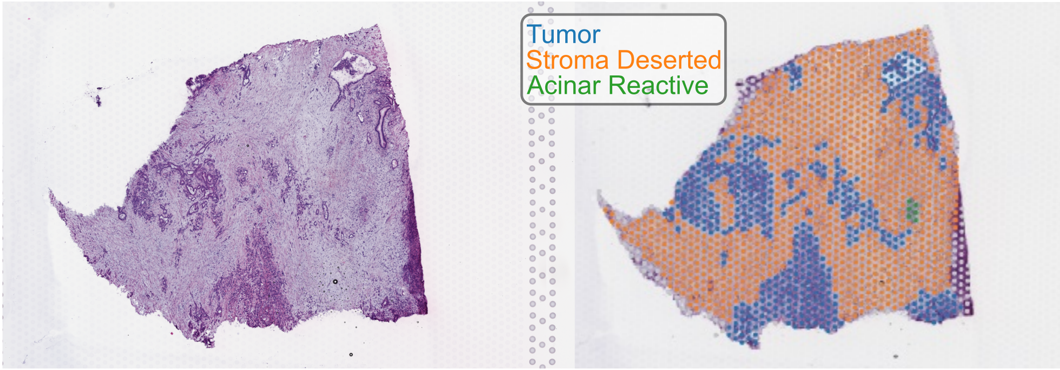



Pancreatic Ductal Adenocarcinoma Cases
n = 455
Spatial Transcriptomics
n = 16
Capture Areas
55µm
100µm
Barcoded Spots
Partial Read +
Spatial Barcode +
UMI + poly(dT)
Novel Computational Methods

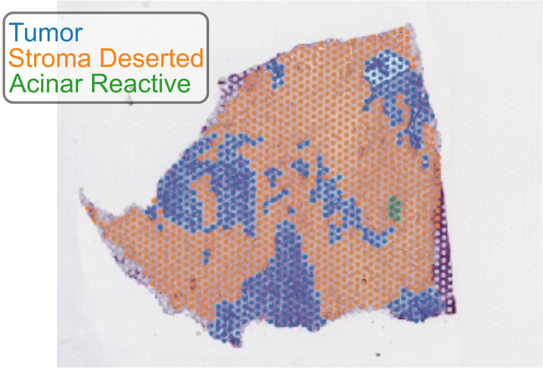
Spatial information paired with molecular profiles is a powerful tool

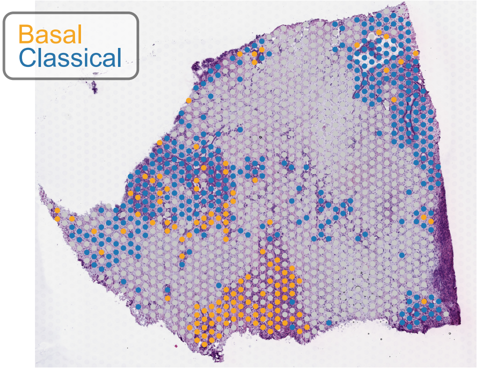


Spatial information paired with molecular profiles is a powerful tool

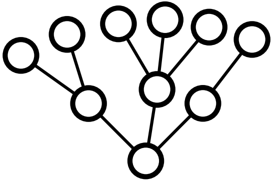


Spatial transcriptomics reveals pancreatic ductal adenocarcinoma signatures

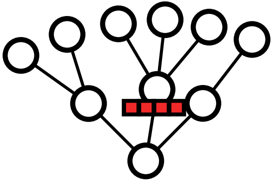




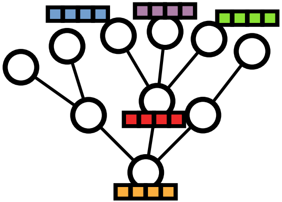
Graph convolutional networks learn node embeddings for downstream analysis



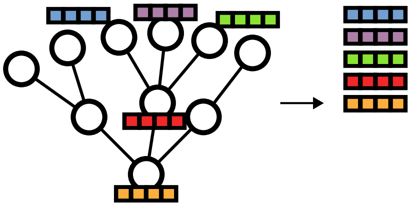
Graph convolutional networks learn node embeddings for downstream analysis



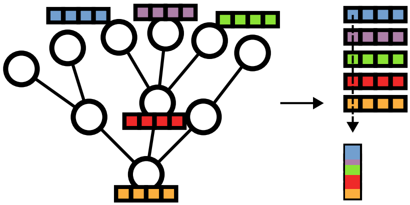
Graph convolutional networks learn node embeddings for downstream analysis



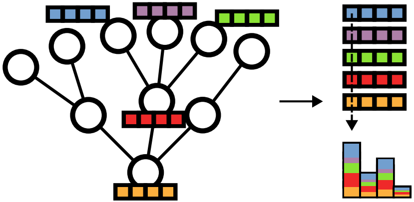
Graph convolutional networks learn node embeddings for downstream analysis



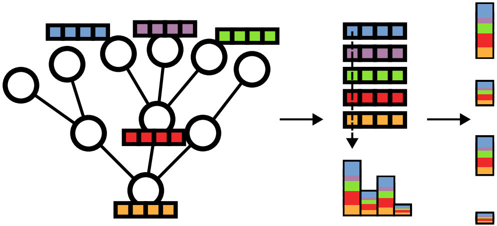
Graph convolutional networks learn node embeddings for downstream analysis



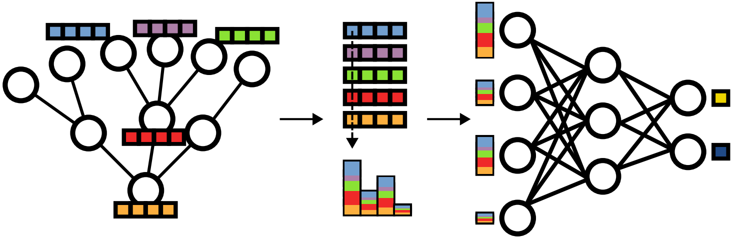
Graph convolutional networks learn node embeddings for downstream analysis



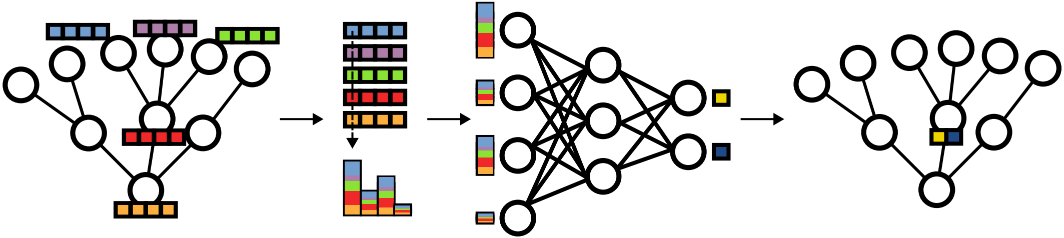
Graph convolutional networks learn node embeddings for downstream analysis



Graph convolutional networks learn node embeddings for downstream analysis

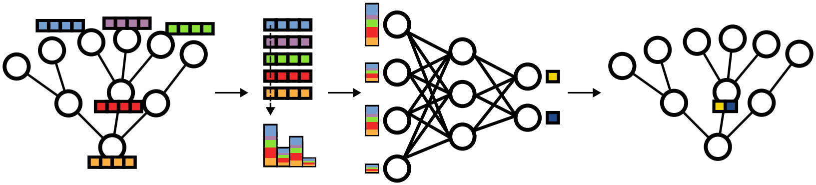


Graph convolutional networks learn node embeddings for downstream analysis

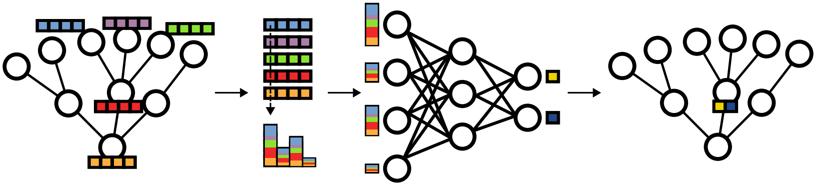


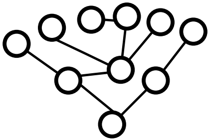
How do we learn without a training set?

Graph convolutional networks learn node embeddings for downstream analysis

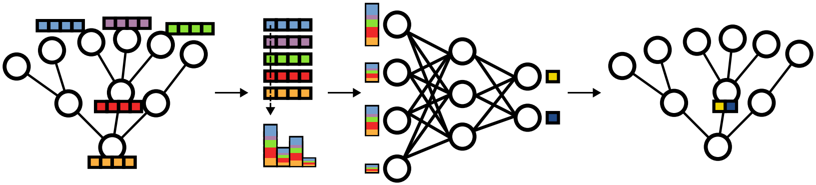


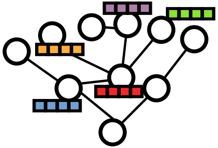
Graph convolutional networks learn node embeddings for downstream analysis



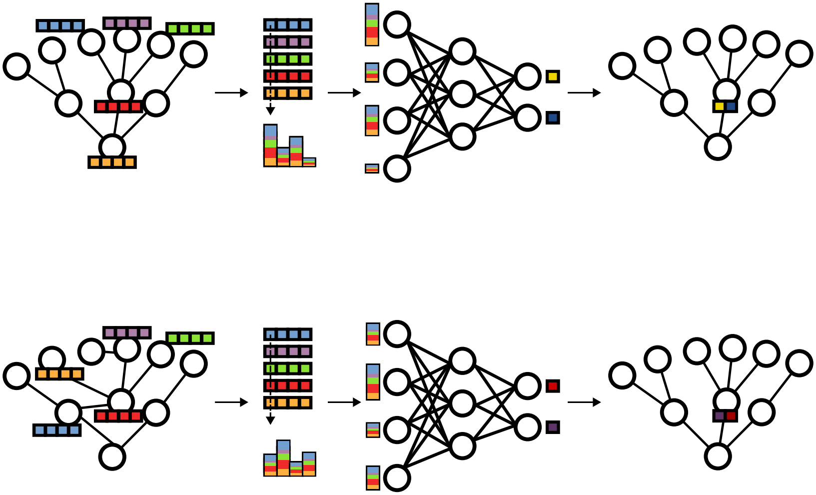


Graph convolutional networks learn node embeddings for downstream analysis

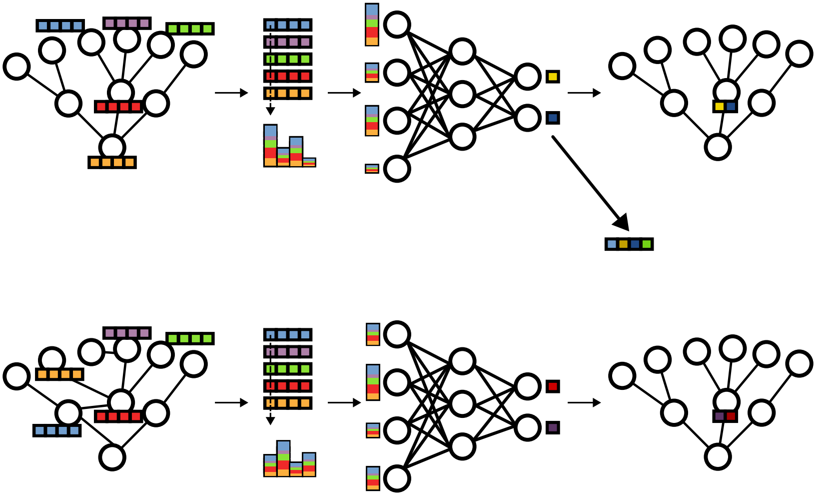




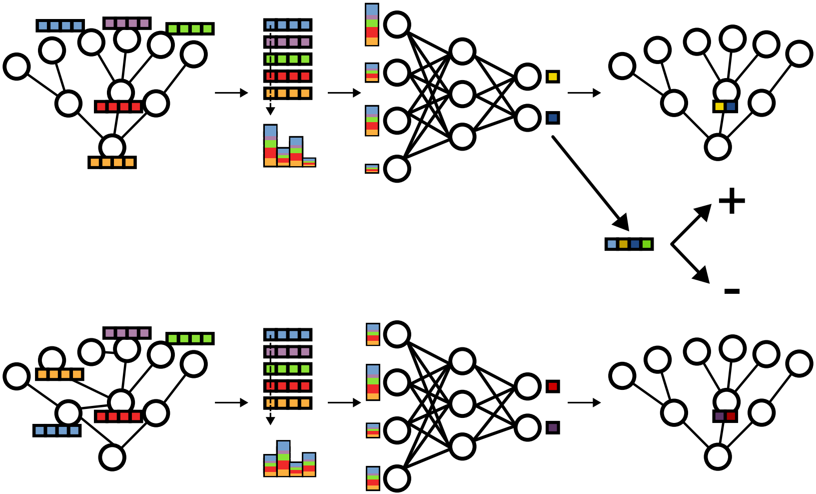
Graph convolutional networks learn node embeddings for downstream analysis



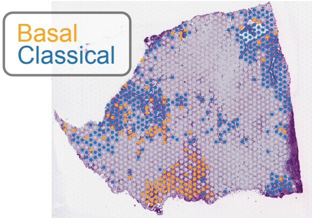
Graph convolutional networks learn node embeddings for downstream analysis



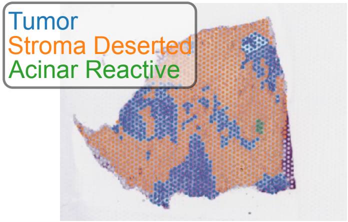
Graph convolutional networks learn node embeddings for downstream analysis

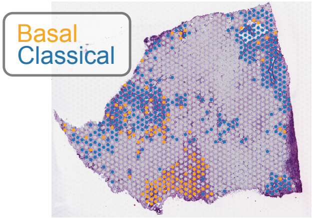


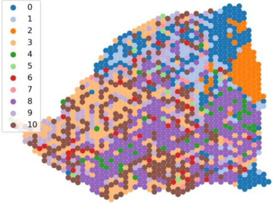
Our model better incorporates spatial cell organization



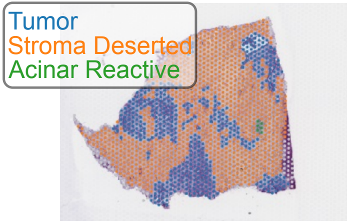
Our model better incorporates spatial cell organization

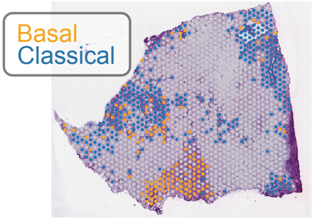


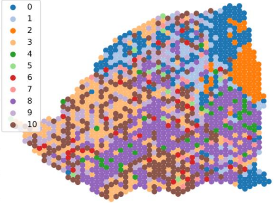


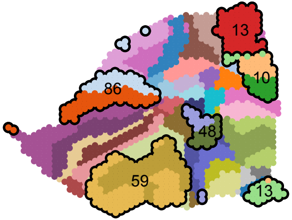


Our model better incorporates spatial cell organization

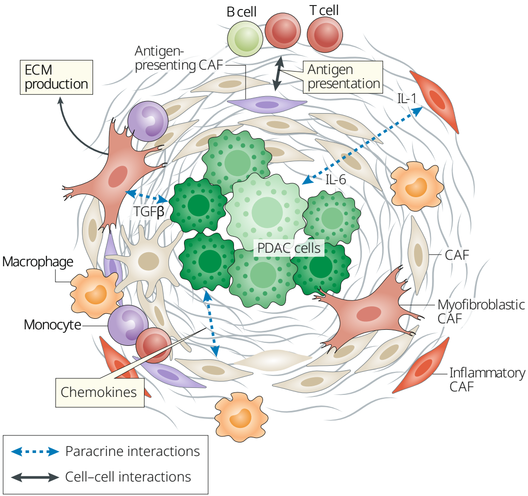




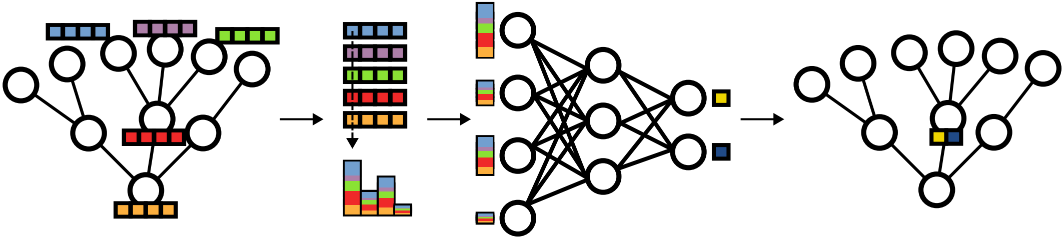




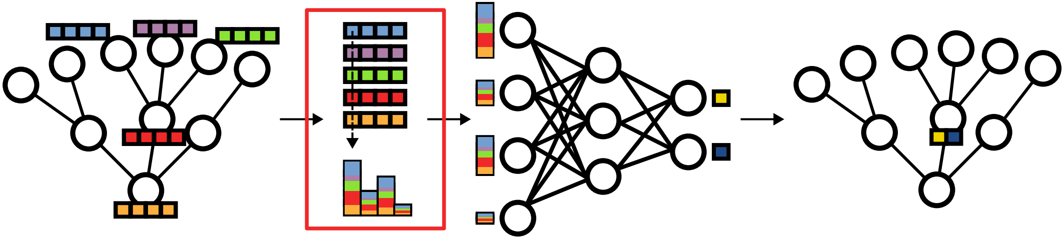
The microenvironment may influence subtype



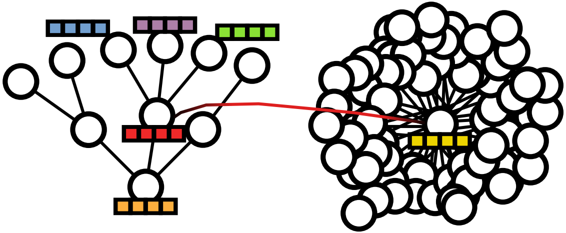
Graph attention networks learn communication



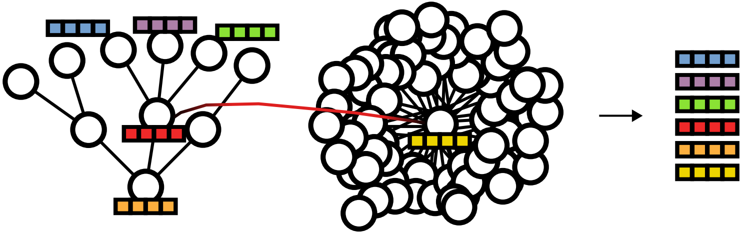
Graph attention networks learn communication



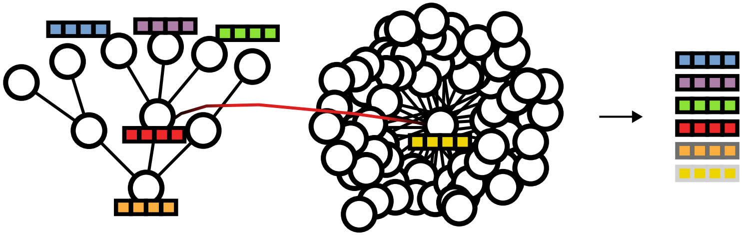
Graph attention networks learn communication



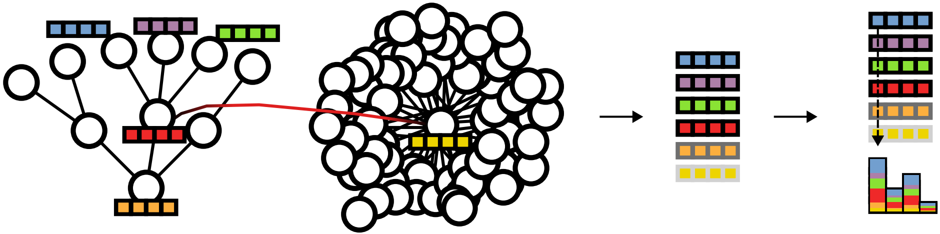
Graph attention networks learn communication



Graph attention networks learn communication



Graph attention networks learn communication

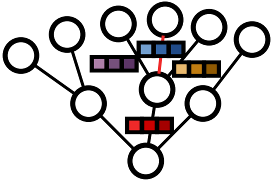


G

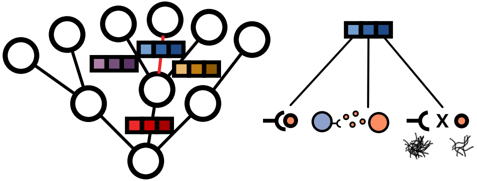
Do you use an attention model?

Yes, as a language model, I use an attention mechanism to help me understand the relationships between different words and phrases in a given text sequence. The attention mechanism allows me to assign varying levels of importance to different parts of the input text, which helps me generate more accurate and meaningful responses. This attention mechanism is a key component of many advanced language models, including the GPT (Generative Pre-trained Transformer) family of models, which I am a part of.

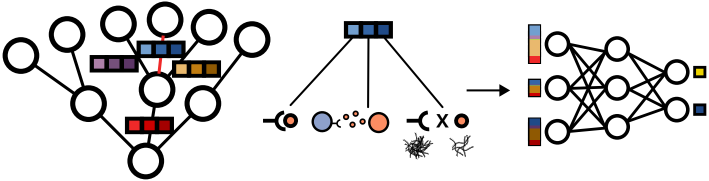
Graph attention networks learn communication



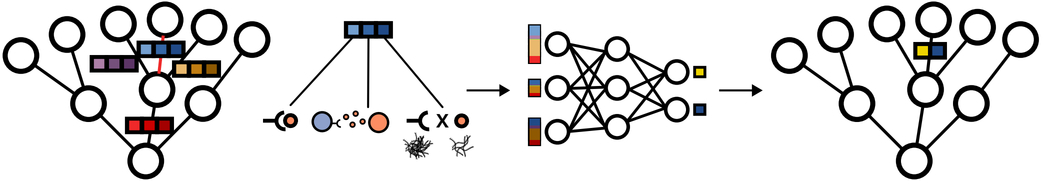
Graph attention networks learn communication



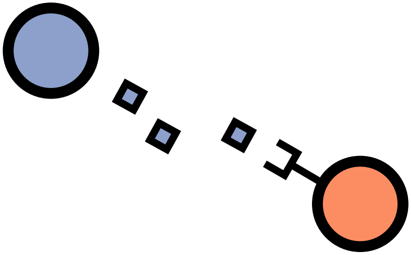
Graph attention networks learn communication



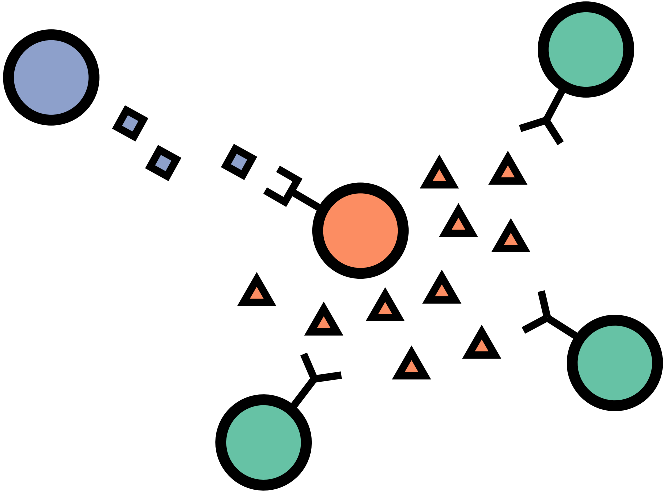
Graph attention networks learn communication



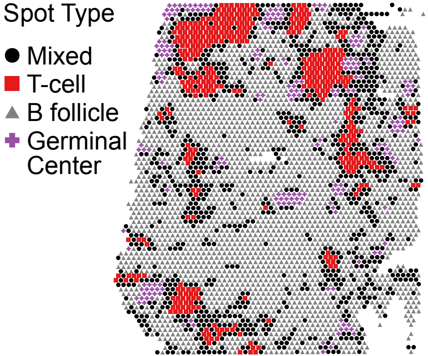
Our model detects cell-cell communication patterns



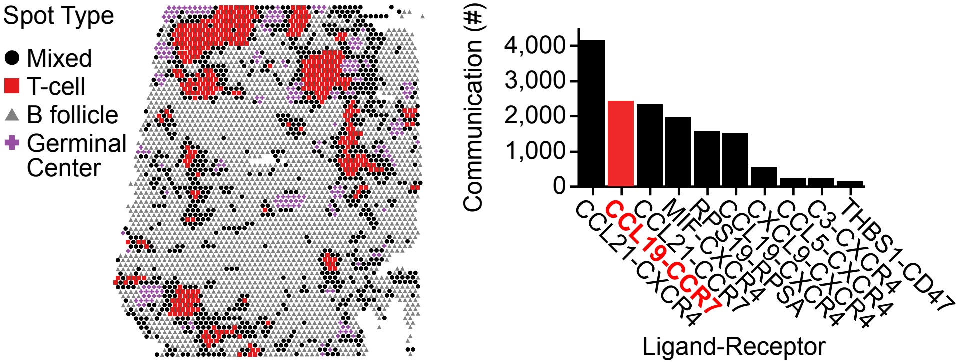
Our model detects cell-cell communication patterns



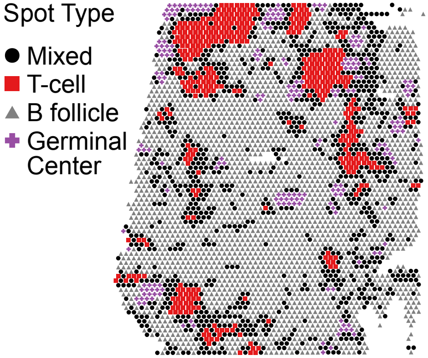
NEST identifies localized *CCL19-CCR7* in lymph node

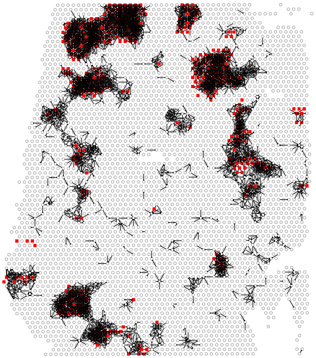


NEST identifies localized *CCL19-CCR7* in lymph node

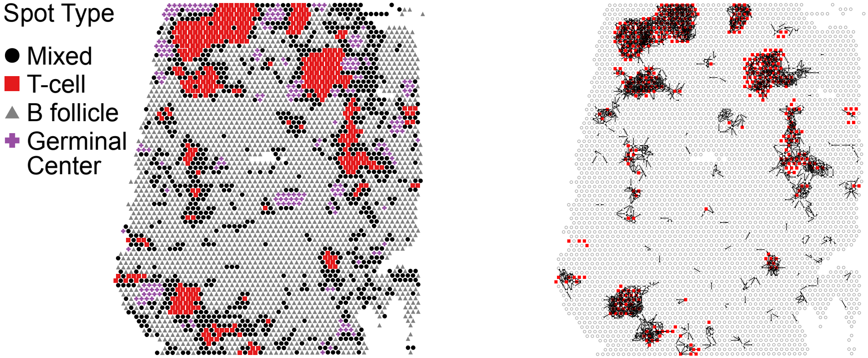


NEST identifies localized *CCL19-CCR7* in lymph node

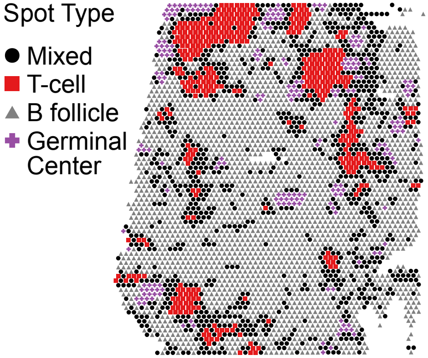


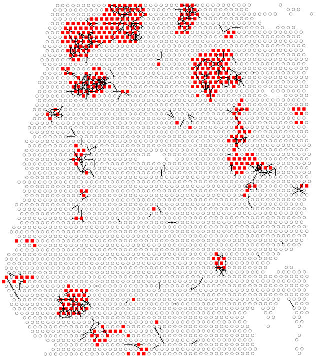


NEST identifies localized *CCL19-CCR7* in lymph node

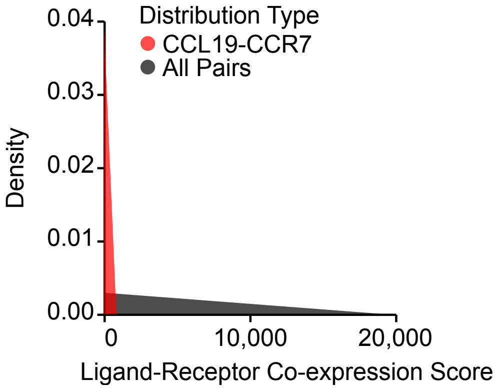


NEST identifies localized *CCL19-CCR7* in lymph node

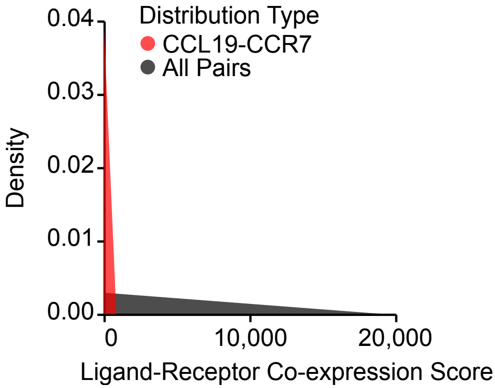


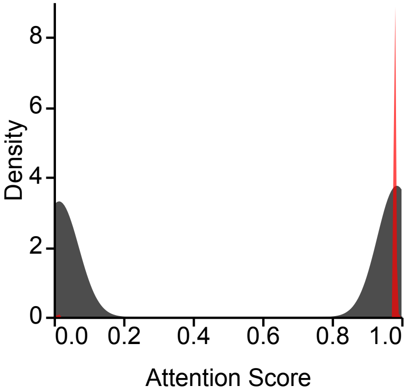


NEST identifies localized *CCL19-CCR7* in lymph node

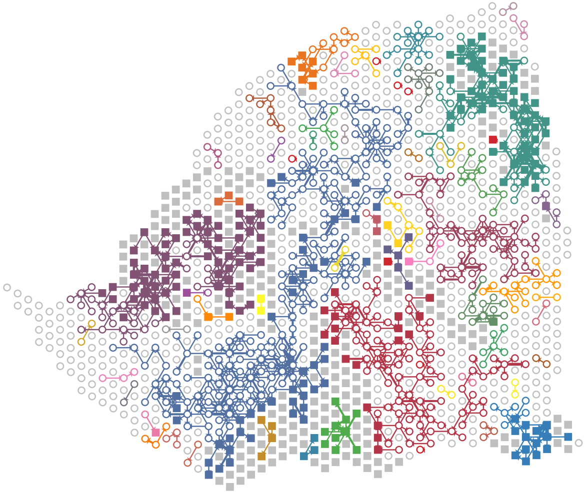


NEST identifies localized *CCL19-CCR7* in lymph node

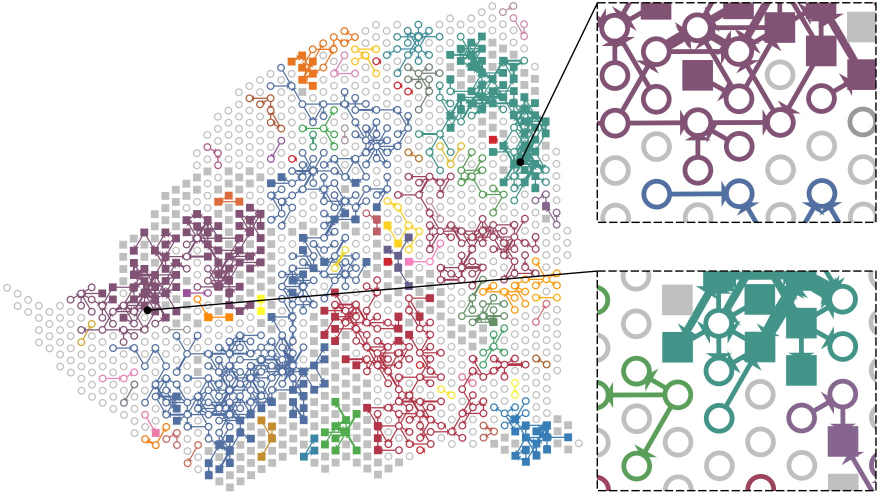




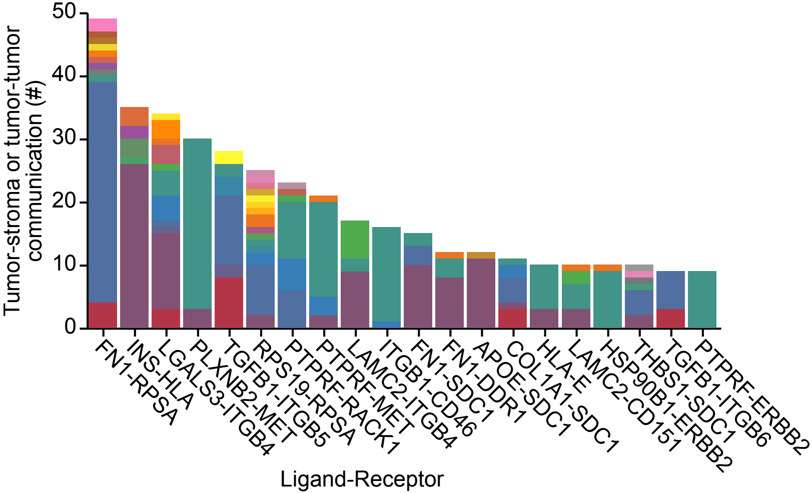
NEST reveals spatially dependent communication *in situ*



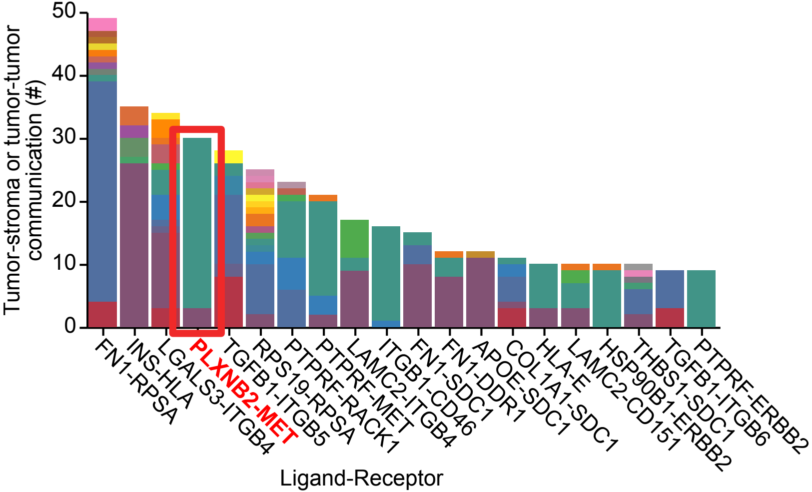
NEST reveals spatially dependent communication *in situ*



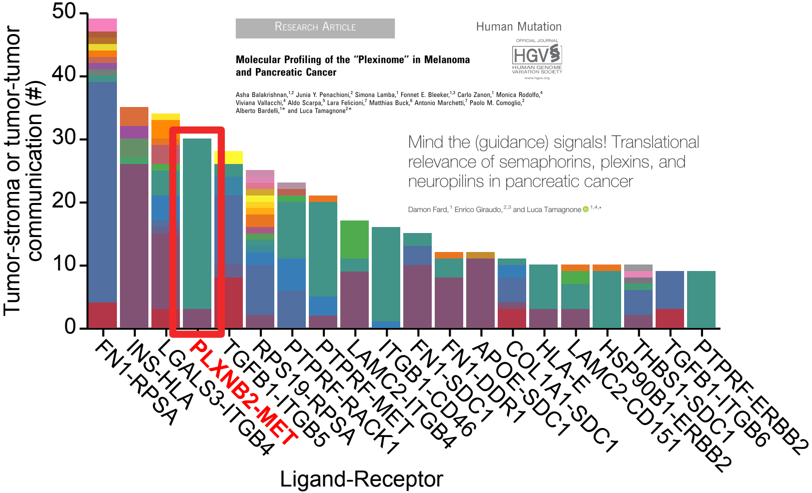
NEST reveals spatially dependent communication *in situ*



NEST reveals spatially dependent communication *in situ*

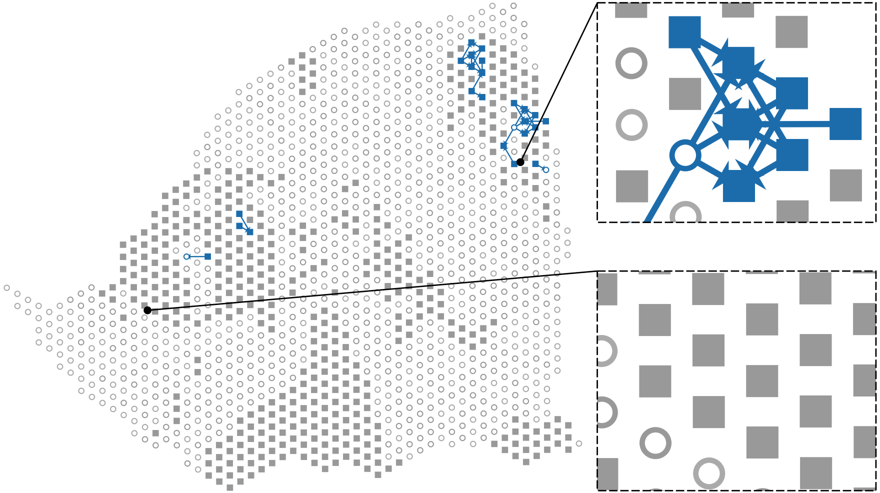


NEST reveals spatially dependent communication *in situ*



NEST reveals spatially dependent communication *in situ*

NEST reveals spatially dependent communication *in situ*



What was covered within this module

The role and context of cell-cell communication

Single-cell RNA sequencing for cellular transcriptomes

General principles of cell-cell communication detection

Inner workings of select single-cell RNA sequencing methods

Spatial transcriptomics for communication detection

Spatially-resolved cell-cell communication method

We are on a Coffee Break & Networking Session

Workshop Sponsors:

Canadian Centre for Computational Genomics

HPC4Health

OICR Ontario Institute for Cancer Research

Ontario Genomics

GenomeCanada