Causal Discovery and Network Inference for Biological and Biomedical Data

Hervé Isambert

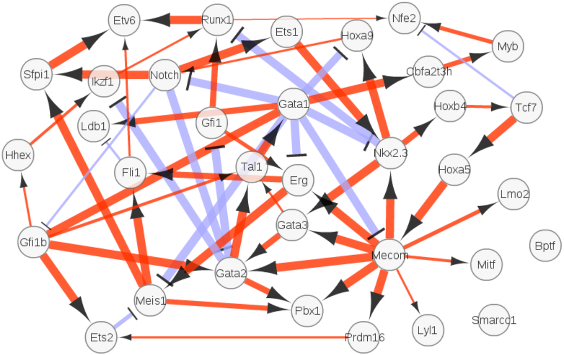
CNRS, Institut Curie & Prairie Institute, PSL University, Paris, France

Eclaire GT, Kickoff Meeting, LIP6, Sorbonne University

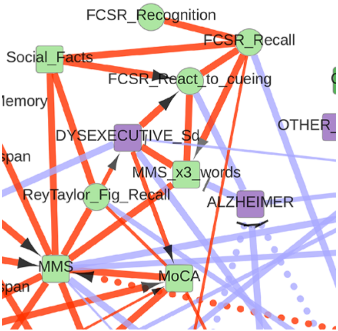
24 March 2026

MIIC: combining constraint- and score-based methods

Gene regulatory networks from single-cell expression data



Clinical networks from medical records



MIIC computes *distinct* endpoint orientation probabilities:

$p_{\text{tail}} = p'_{\text{tail}}$ $p_{\text{tail}} \Longrightarrow p'_{\text{head}}$ $p_{\text{head}} \Longleftrightarrow p'_{\text{head}}$ $(p_{\text{head}} + p_{\text{tail}} = 1)$

⇒ MIIC server https://miic.curie.fr

Verny *et al.* PloS Comput Biol 2017
Sella *et al.* Bioinformatics 2018
NeurIPS 2019, 2021 & 2024, ICML 2025
Cabeli *et al.* PloS Comput Biol 2020
Sella *et al.* npj Digit Med 2022 & 2025
Ribeiro-Dantas *et al.* iScience 2024 Dupuis *et al.* NAR 2025
Simon *et al.* eLife 2025 Fusilier *et al.* Science Immunol 2026

Correlation *versus* Information

$$\sigma_{x,y} = \frac{1}{N}\sum_{i}^{N}(x_i - \bar{x})(y_i - \bar{y})$$

X, Y need to be continuous

$$I(X;Y) = \sum_{x,y} p(x,y) \log_2 \frac{p(x,y)}{p(x)p(y)}$$

X, Y can be categorical, continuous or even high-dimensional (sequences, images,...)

Identical binary variables: $I(X;Y) = 1$ bit

Correlation *versus* Information

$$\sigma_{x,y} = \frac{1}{N}\sum_{i}^{N}(x_i - \bar{x})(y_i - \bar{y})$$

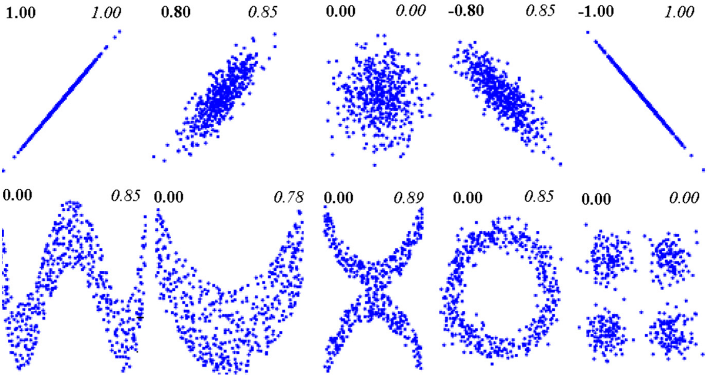
X, Y need to be continuous

$$I(X;Y) = \sum_{x,y} p(x,y)\log_2 \frac{p(x,y)}{p(x)p(y)}$$

X, Y can be categorical, continuous or even high-dimensional (sequences, images,...)

Identical binary variables: $I(X;Y) = 1$ bit

Correlation *Information*



How much is $I(\mathit{master}\,;\mathit{dog})$ in bits ?

Or... how much is $I(\textit{cell image}\,;\textit{cell omics})$ in bits ?

Or... how much is $I(\textit{cell image}\,;\textit{cell omics})$ in bits ?

but... why should we care about information ??

Likelihood, Cross-Entropy and Cross-Information

Likelihood $q(\boldsymbol{x}^{(1)}, \cdots, \boldsymbol{x}^{(N)})$ and **Cross-Entropy** $\boldsymbol{H(p, q)}$

Given some Data, N i.i.d. samples $\sim p(\boldsymbol{x})$, *and a Model distribution,* $q(\boldsymbol{x})$,
The probability that the sequence $\boldsymbol{x}^{(1)}, \boldsymbol{x}^{(2)}, \cdots, \boldsymbol{x}^{(N)}$ *was generated by* $q(\boldsymbol{x})$ *is:*

$$q(\boldsymbol{x}^{(1)}, \cdots, \boldsymbol{x}^{(N)}) \rightarrow e^{-N\, H(p,q)} = e^{N \sum_{\boldsymbol{x}} p(\boldsymbol{x}) \log q(\boldsymbol{x})} = e^{-N\left(H(p) + D_{KL}(p\|q)\right)}$$

where $D_{KL}(p||q) = \sum_{\boldsymbol{x}} p(\boldsymbol{x}) \log \frac{p(\boldsymbol{x})}{q(\boldsymbol{x})} \geqslant 0$ *is the Kullback-Leibler divergence*

Likelihood, Cross-Entropy and Cross-Information

Likelihood $q(\boldsymbol{x}^{(1)}, \cdots, \boldsymbol{x}^{(N)})$ and **Cross-Entropy** $\boldsymbol{H(p, q)}$

Given some Data, N i.i.d. samples $\sim p(\boldsymbol{x})$, and a Model distribution, $q(\boldsymbol{x})$,
The probability that the sequence $\boldsymbol{x}^{(1)}, \boldsymbol{x}^{(2)}, \cdots, \boldsymbol{x}^{(N)}$ was generated by $q(\boldsymbol{x})$ is:

$$q(\boldsymbol{x}^{(1)}, \cdots, \boldsymbol{x}^{(N)}) \rightarrow e^{-N\,H(p,q)} = e^{N \sum_x p(\boldsymbol{x}) \log q(\boldsymbol{x})} = e^{-N\left(H(p)+D_{KL}(p\|q)\right)}$$

where $D_{KL}(p||q) = \sum_x p(\boldsymbol{x}) \log \frac{p(\boldsymbol{x})}{q(\boldsymbol{x})} \geqslant 0$ *is the Kullback-Leibler divergence*

Proof: $\frac{1}{N} \log q(\boldsymbol{x}^{(1)}, \boldsymbol{x}^{(2)}, \cdots, \boldsymbol{x}^{(N)}) = \frac{1}{N} \sum_i^N \log q(\boldsymbol{x}^{(i)}) \quad \rightarrow \quad \sum_x^r p(\boldsymbol{x}) \log q(\boldsymbol{x})$ ☐

Likelihood, Cross-Entropy and Cross-Information

Likelihood $q(\boldsymbol{x}^{(1)}, \cdots, \boldsymbol{x}^{(N)})$ and Cross-Entropy $H(p, q)$

Given some Data, N i.i.d. samples $\sim p(\boldsymbol{x})$, *and a Model distribution,* $q(\boldsymbol{x})$,
The probability that the sequence $\boldsymbol{x}^{(1)}, \boldsymbol{x}^{(2)}, \cdots, \boldsymbol{x}^{(N)}$ *was generated by* $q(\boldsymbol{x})$ *is:*

$$q(\boldsymbol{x}^{(1)}, \cdots, \boldsymbol{x}^{(N)}) \rightarrow e^{-N\,H(p,q)} = e^{N \sum_x p(\boldsymbol{x}) \log q(\boldsymbol{x})} = e^{-N\big(H(p)+D_{KL}(p\|q)\big)}$$

where $D_{KL}(p||q) = \sum_x p(\boldsymbol{x}) \log \frac{p(\boldsymbol{x})}{q(\boldsymbol{x})} \geqslant 0$ *is the Kullback-Leibler divergence*

Proof: $\frac{1}{N} \log q(\boldsymbol{x}^{(1)}, \boldsymbol{x}^{(2)}, \cdots, \boldsymbol{x}^{(N)}) = \frac{1}{N} \sum_i^N \log q(\boldsymbol{x}^{(i)}) \quad \rightarrow \quad \sum_x^r p(\boldsymbol{x}) \log q(\boldsymbol{x})$ □

Multivariate (cross-)entropy $\quad H(X_1, \cdots, X_n) \equiv H(p, q) = -\sum_{\{x_i\}_n} p(\{x_i\}_n) \log q(\{x_i\}_n)$

Likelihood, Cross-Entropy and Cross-Information

Likelihood $q(\boldsymbol{x}^{(1)},\cdots,\boldsymbol{x}^{(N)})$ and Cross-Entropy $\boldsymbol{H(p,q)}$

Given some Data, N i.i.d. samples $\sim p(\boldsymbol{x})$, and a Model distribution, $q(\boldsymbol{x})$,
The probability that the sequence $\boldsymbol{x}^{(1)},\boldsymbol{x}^{(2)},\cdots,\boldsymbol{x}^{(N)}$ was generated by $q(\boldsymbol{x})$ is:

$$q(\boldsymbol{x}^{(1)},\cdots,\boldsymbol{x}^{(N)}) \;\to\; e^{-N\,H(p,q)} = e^{N\sum_x p(\boldsymbol{x})\log q(\boldsymbol{x})} = e^{-N\left(H(p)+D_{KL}(p\|q)\right)}$$

where $D_{KL}(p||q)=\sum_x p(\boldsymbol{x})\log\frac{p(x)}{q(x)}\geqslant 0$ *is the Kullback-Leibler divergence*

Proof: $\frac{1}{N}\log q(\boldsymbol{x}^{(1)},\boldsymbol{x}^{(2)},\cdots,\boldsymbol{x}^{(N)})=\frac{1}{N}\sum_i^N\log q(\boldsymbol{x}^{(i)}) \;\to\; \sum_x^r p(\boldsymbol{x})\log q(\boldsymbol{x})$ □

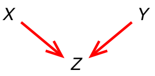
Multivariate (cross-)entropy $H(X_1,\cdots,X_n)\equiv H(p,q)=-\sum_{\{x_i\}_n}p(\{x_i\}_n)\log q(\{x_i\}_n)$

Multivariate (cross-)information $I(X_1;\cdots;X_n)$ are defined from $H(X_1,\cdots,X_n)$ and vice versa

$$\begin{aligned}
I(X;Y) &= H(X)+H(Y)-H(X,Y)\\
I(X;Y;A) &= H(X)+H(Y)+H(A)-H(X,Y)-H(X,A)-H(Y,A)+H(X,Y,A)\\
I(X_1;\cdots;X_n) &= -\sum_{\boldsymbol{S}\subseteq\boldsymbol{V}}(-1)^{|\boldsymbol{S}|}H(\boldsymbol{S})\\
H(X_1,\cdots,X_n) &= -\sum_{\boldsymbol{S'}\subseteq\boldsymbol{V}}(-1)^{|\boldsymbol{S'}|}I(\boldsymbol{S'}) \qquad [\text{Inclusion}-\text{Exclusion Principle}]
\end{aligned}$$

Link between constraint-based and score-based methods

Constraints $\equiv$ Cancelling (combinations of) multivariate cross-informations


V–structure


Latent variables L

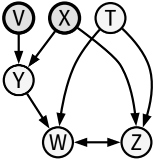
$I(X;Y)=0$
$I(X;Z)>0$
$I(Z;Y)>0$
$I(X;Z;Y)<0\,*$

$I(X;Y)=I(X;W)=I(W;Z)=0$
$I(X;Z)>0$
$I(Z;Y)>0$ $\quad I(Y;W)>0$
$I(X;Z;Y)<0\,*$ $\quad I(Z;Y;W)<0\,*$
$I(X;Z;W)=0$ $\quad I(X;Y;W)=0$
$I(X;Z;Y;W)>0$

* Signature of Causality from MI decomposition: $I(X;Y)=I(X;Z;Y)+I(X;Y|Z)$

Constraints in Ancestral Graphs

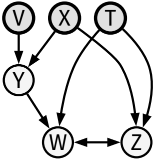
Information mutuelle nulle



$$I(V;X) \quad = \quad 0$$

Constraints in Ancestral Graphs

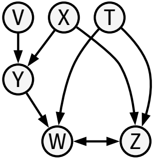
Interaction nulle



$$
\begin{aligned}
I(V;X) &= 0 \\
I(V;T) &= 0 \\
I(V;Z) &= 0 \\
I(X;T) &= 0 \\
I(Y;T) &= 0
\end{aligned}
\qquad I(V;X;T) = 0
$$

Constraints in Ancestral Graphs

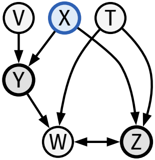
Information et interaction nulle



$$
\begin{aligned}
I(V;X) &= 0 & I(V;X;T) &= 0\\
I(V;T) &= 0 & I(V;X;Z) &= 0\\
I(V;Z) &= 0 & I(V;T;Z) &= 0\\
I(X;T) &= 0 & I(T;V;Y) &= 0\\
I(Y;T) &= 0 & I(T;X;Y) &= 0
\end{aligned}
$$

Constraints in Ancestral Graphs

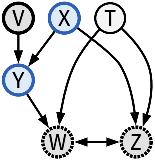
Information mutuelle conditionnelle nulle



$$I(Y;Z \mid X) \;=\; I(Y;Z) - I(Y;Z;X) \;=\; 0$$

Constraints in Ancestral Graphs

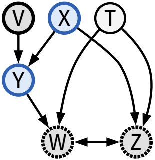
Interaction conditionnelle nulle



$$I(V;\{W,Z\} \mid X,Y) = I(V;W \mid X,Y) + I(V;Z \mid W,X,Y)$$

Constraints in Ancestral Graphs

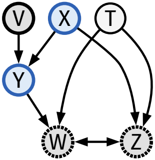
Interaction conditionnelle nulle



$$\begin{aligned} I(V;\{W,Z\}\mid X,Y) &= I(V;W\mid X,Y) + I(V;Z\mid W,X,Y) \\ &= I(V;W) - I(V;W;X) - I(V;W;Y) + I(V;W;X;Y) \\ &\quad + I(V;Z) - I(V;Z;X) - I(V;Z;Y) - I(V;Z;W) \\ &\quad + I(V;Z;X;Y) + I(V;Z;X;W) + I(V;Z;Y;W) - I(V;Z;X;Y;W) \end{aligned}$$

Constraints in Ancestral Graphs

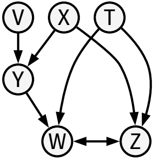
Interaction conditionnelle nulle



$$
\begin{aligned}
I(V;\{W,Z\} \mid X,Y) &= I(V;W \mid X,Y) + I(V;Z \mid W,X,Y) \\
&= I(V;W) - I(V;W;X) - I(V;W;Y) + I(V;W;X;Y) \\
&\quad + I(V;Z) - I(V;Z;X) - I(V;Z;Y) - I(V;Z;W) \\
&\quad + I(V;Z;X;Y) + I(V;Z;X;W) + I(V;Z;Y;W) - I(V;Z;X;Y;W) \\
&= 0
\end{aligned}
$$

Constraints in Ancestral Graphs

Contraintes : termes informatifs nuls



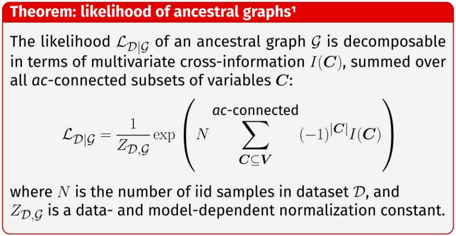
$I(V;X)=0,\ I(V;T)=0,\ I(V;Z)=0$
$I(X;T)=0,\ I(Y;T)=0$

$I(V;X;T)=0,\ I(V;X;Z)=0,\ I(V;T;Z)=0$
$I(T;V;Y)=0,\ I(T;X;Y)=0$

$I(Y;Z\mid X)=0,\ I(X;W\mid Y)=0,\ I(V;W\mid Y)=0$

$I(V;\{W,Z\}\mid X,Y)=0$

Likelihood of Ancestral Graphs (incl $\rightarrow$ and $\leftrightarrow$)



Theorem: likelihood of ancestral graphs[1]

The likelihood $\mathcal{L}_{\mathcal{D}|\mathcal{G}}$ of an ancestral graph $\mathcal{G}$ is decomposable in terms of multivariate cross-information $I(\boldsymbol{C})$, summed over all *ac*-connected subsets of variables $\boldsymbol{C}$:

$$\mathcal{L}_{\mathcal{D}|\mathcal{G}} = \frac{1}{Z_{\mathcal{D},\mathcal{G}}} \exp\left(N \sum_{\boldsymbol{C} \subseteq \boldsymbol{V}}^{ac\text{-connected}} (-1)^{|\boldsymbol{C}|} I(\boldsymbol{C}) \right)$$

where N is the number of iid samples in dataset $\mathcal{D}$, and $Z_{\mathcal{D},\mathcal{G}}$ is a data- and model-dependent normalization constant.

---

[1] Lagrange & Isambert, ICML 2025

Likelihood of Ancestral Graphs (incl → and ↔)

Theorem: likelihood of ancestral graphs[1]

The likelihood $\mathcal{L}_{\mathcal{D}|\mathcal{G}}$ of an ancestral graph $\mathcal{G}$ is decomposable in terms of multivariate cross-information $I(\boldsymbol{C})$, summed over all *ac*-connected subsets of variables $\boldsymbol{C}$:

$$\mathcal{L}_{\mathcal{D}|\mathcal{G}} = \frac{1}{Z_{\mathcal{D},\mathcal{G}}} \exp\left(N \sum_{\boldsymbol{C}\subseteq \boldsymbol{V}}^{ac\text{-connected}} (-1)^{|\boldsymbol{C}|} I(\boldsymbol{C}) \right)$$

where N is the number of iid samples in dataset $\mathcal{D}$, and $Z_{\mathcal{D},\mathcal{G}}$ is a data- and model-dependent normalization constant.

Definition: *ac*-connected subset

A subset $\boldsymbol{C}$ is said to be *ac*-connected if $\forall X, Y \in \boldsymbol{C}$, either:

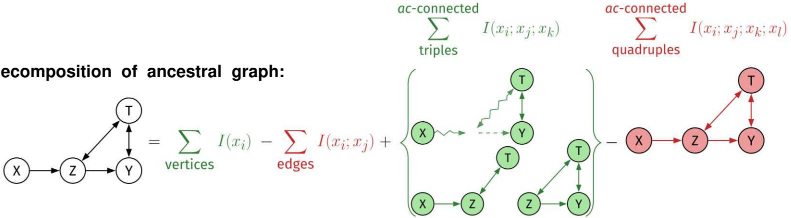
- X and Y are connected through any type of edge, or
- there exists an ***ac*-connecting path** between X and Y given $\boldsymbol{C}$.

An ***ac*-connecting path** between X and Y given $\boldsymbol{C}$ is a collider path

$$X{*}\!\rightarrow Z_1 \leftrightarrow \cdots \leftrightarrow Z_k \leftarrow\!{*}Y,$$

such that each $Z_i \in \mathbf{An}_{\mathcal{G}}(\{X,Y\} \cup \boldsymbol{C})$, that is, with $Z_i \in \boldsymbol{C}$ or through an ancestor path, *i.e.*, $Z_i \rightarrow \cdots \rightarrow T$ with $T \in \{X,Y\} \cup \boldsymbol{C}$.

Cross–entropy decomposition of ancestral graph:



---

[1] Lagrange & Isambert, ICML 2025

Likelihood of Ancestral Graphs (incl → and ↔)

Theorem: likelihood of ancestral graphs[1]

The likelihood $\mathcal{L}_{\mathcal{D}|\mathcal{G}}$ of an ancestral graph $\mathcal{G}$ is decomposable in terms of multivariate cross-information $I(\boldsymbol{C})$, summed over all *ac*-connected subsets of variables $\boldsymbol{C}$:

$$\mathcal{L}_{\mathcal{D}|\mathcal{G}} = \frac{1}{Z_{\mathcal{D},\mathcal{G}}} \exp\left(N \sum_{\boldsymbol{C}\subseteq\boldsymbol{V}}^{ac\text{-connected}} (-1)^{|\boldsymbol{C}|} I(\boldsymbol{C}) \right)$$

where N is the number of iid samples in dataset $\mathcal{D}$, and $Z_{\mathcal{D},\mathcal{G}}$ is a data- and model-dependent normalization constant.

Definition: *ac*-connected subset

A subset $\boldsymbol{C}$ is said to be *ac*-connected if $\forall X, Y \in \boldsymbol{C}$, either:

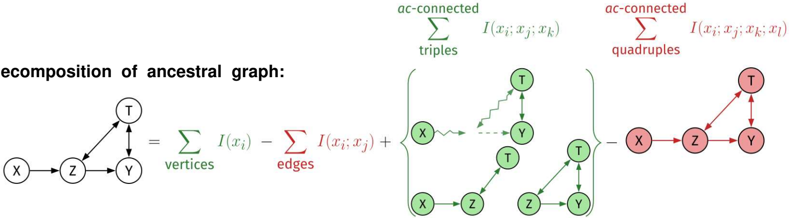
- X and Y are connected through any type of edge, or
- there exists an ***ac*-connecting path** between X and Y given $\boldsymbol{C}$.

An ***ac*-connecting path** between X and Y given $\boldsymbol{C}$ is a collider path

$$X{*}\!\rightarrow Z_1 \leftrightarrow \cdots \leftrightarrow Z_k \leftarrow\!{*}Y,$$

such that each $Z_i \in \mathbf{An}_G(\{X,Y\} \cup \boldsymbol{C})$, that is, with $Z_i \in \boldsymbol{C}$ or through an ancestor path, *i.e.*, $Z_i \rightarrow \cdots \rightarrow T$ with $T \in \{X,Y\} \cup \boldsymbol{C}$.

Cross–entropy decomposition of ancestral graph:



Proposition: estimation of cross-information for *ac*-connected subset *C*

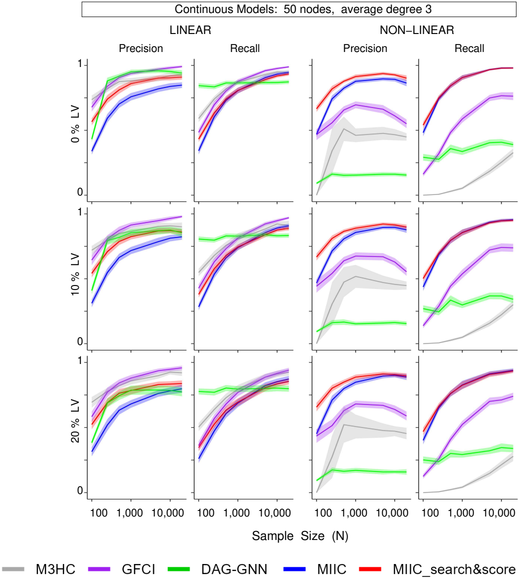
$$I(\boldsymbol{C}) = \sum p(x,y,z)\,\log\frac{q(x,y,z)\;q(x)\;q(y)\;q(z)}{q(x,y)\;q(x,z)\;q(y,z)} \equiv \sum p(x,y,z)\,\log\frac{p(x,y,z)\;p(x)\;p(y)\;p(z)}{p(x,y)\;p(x,z)\;p(y,z)}$$

---

[1] Lagrange & Isambert, ICML 2025

Benchmarks on Linear *vs* Non-linear Continuous Data

MIIC does not assume linear distribution ⇒ MIIC outperforms other methods on non-linear data

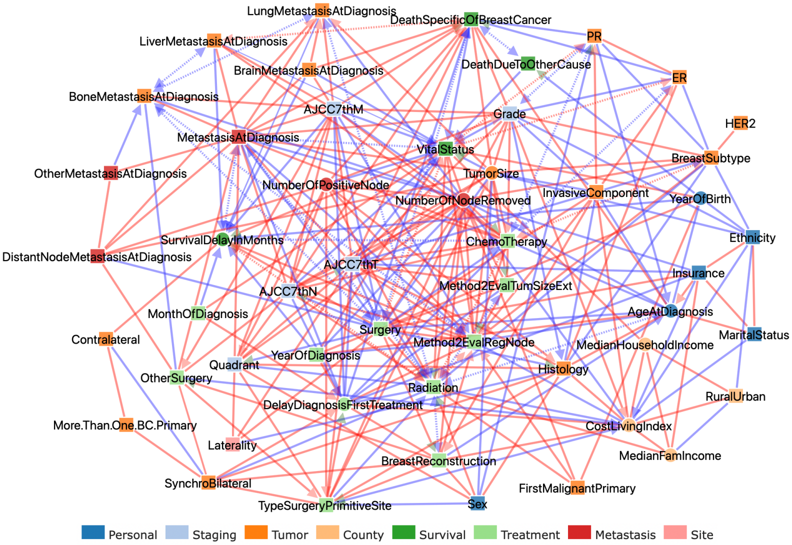


Lagrange & Isambert, ICML 2025

Learning causal networks from nationwide healthcare data

400,000 breast cancer patients from SEER database

Ribeiro-Dantas *et al.* iScience 2024



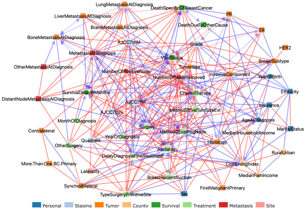
$\Longrightarrow$ Reliable causal discovery: 90% of causal edges are correct (+ 8% likely correct too)

Slide for Gregory ! ;)

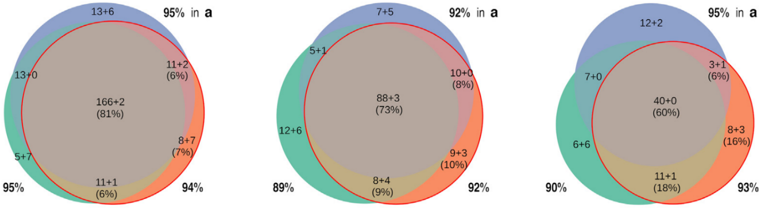
400,000 breast cancer patients from SEER database

Ribeiro-Dantas *et al.* iScience 2024

A “full graph” (400,000 patients)



B Independent subsets



100,000 patient subsets

10,000 patient subsets

1,000 patient subsets

$\Longrightarrow$ Independent resampling essentially leads to subgraphs of the “full graph”

'Putative' *versus* 'Genuine' causal edges

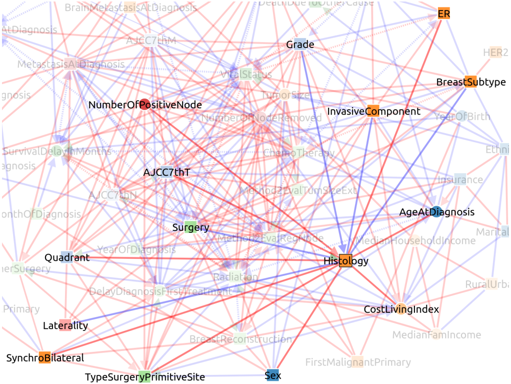
'Putative' causal edge ($p_t = p_h = 0.5 \rightarrow p_h > 0.5$) *versus* 'Genuine' causal edge ($p_t > 0.5 \rightarrow p_h > 0.5$)

ie Putative → ≡ Genuine → **or** latent common cause ↔

'Putative' *versus* 'Genuine' causal edges

'Putative' causal edge ($p_t = p_h = 0.5 \rightarrow p_h > 0.5$) *versus* 'Genuine' causal edge ($p_t > 0.5 \rightarrow p_h > 0.5$)

ie Putative → ≡ Genuine → **or** latent common cause ↔

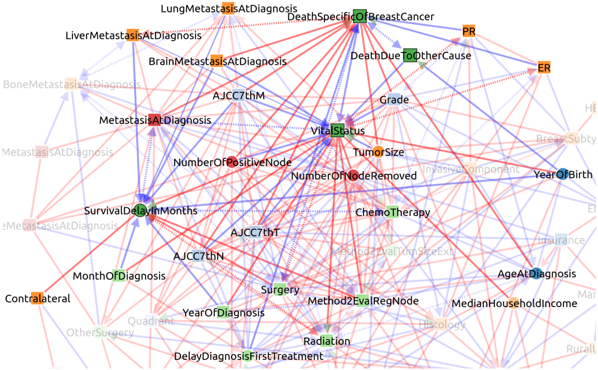


Counter-intuitive 'genuine' causal edge: Surgery → Histology

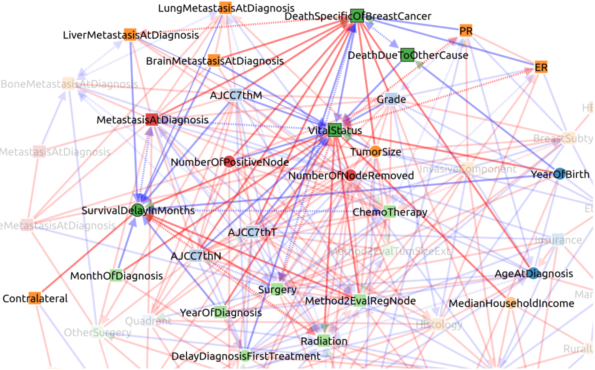
However histological types are frequently refined, *after surgery*, based on the surgical specimen:

Infiltrating duct mixed with other types of carcinoma (+77% after surgery), Infiltrating duct and lobular carcinoma (+48%), Mucinous adenocarcinoma (+19%), Infiltrating duct carcinoma, NOS (+ 7.6%), Lobular carcinoma, NOS (-11%), Carcinoma, NOS (-91%), Adenocarcinoma, NOS (-95%)

SEER Database: Survival subnetwork

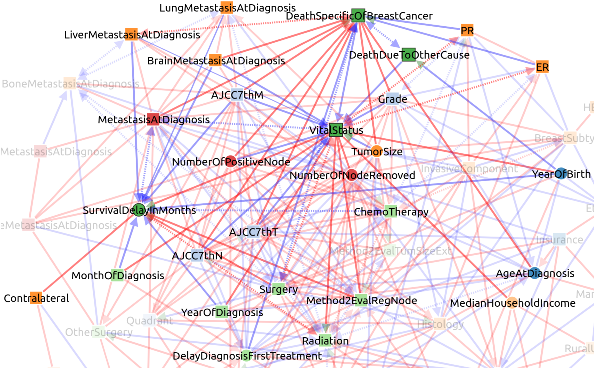


SEER Database: Survival subnetwork



Expected causal edges: Metastasis at diagnosis —▶ Death specific of breast cancer
ER —▶ Death specific of breast cancer
Year of birth —▶ Death due to other cause
Year of birth —▶ Vital status
Year of birth —▶ Survival delay in months due to 2010-2016 censoring!
Vital status —▶ Survival delay in months

SEER Database: Survival subnetwork



Expected causal edges: Metastasis at diagnosis —▶ Death specific of breast cancer

ER —▶ Death specific of breast cancer

Year of birth —▶ Death due to other cause

Year of birth —▶ Vital status

Year of birth —▶ Survival delay in months due to 2010-2016 censoring!

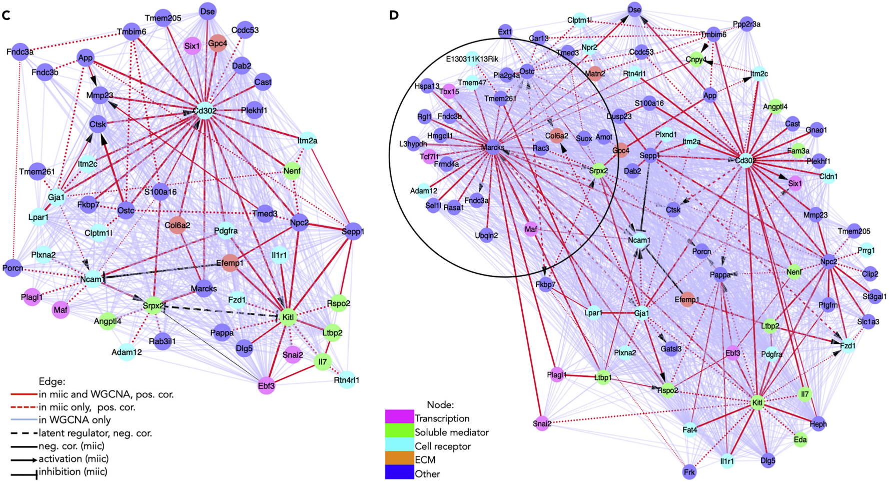
Vital status —▶ Survival delay in months

Unexpected missing edges: Insurance ···· any Survival variable! 70% explained through treatment variables

Learning Gene Regulatory Network from scRNAseq data

Gene networks in bone marrow hematopoietic stem cell-supporting stromal niche population

P Charbord, Sorbonne Université (scRNAseq data from Tikhonova *et al.* Nature 2019) Desterke *et al.* iScience 2020

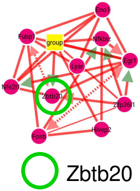


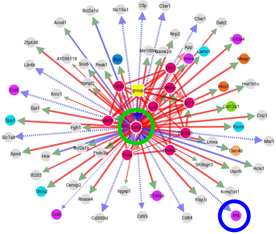
Gene regulatory networks of Leptin receptor-positive cell population at baseline (**C**, 1,712 cells) and under stress (**D**, 3,467 cells)

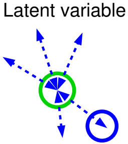
Extension to VERY LARGE Gene Regulatory Networks

scRNAseq data: Hélène Moreau, Institut Curie

10 TFs + 205 non–TFs







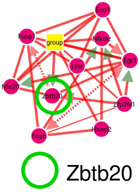
⇒ TF-MIIC can scale to hundreds TFs and thousands non-TFs

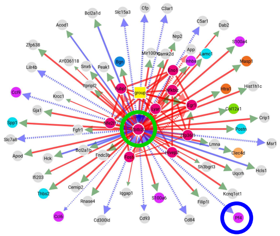
Franck Simon

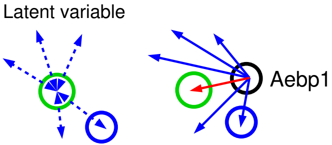
Extension to VERY LARGE Gene Regulatory Networks

scRNAseq data: Hélène Moreau, Institut Curie

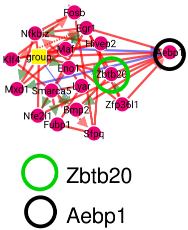
10 TFs + 205 non–TFs

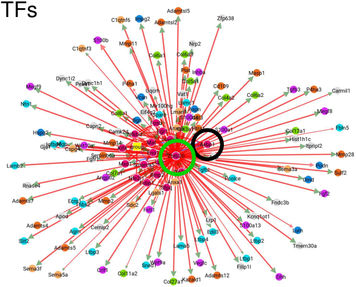


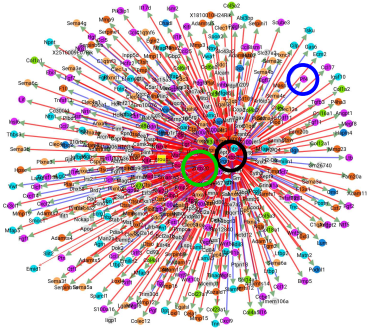




17 TFs + 703 non–TFs





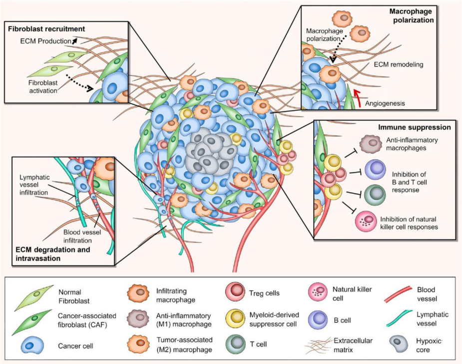


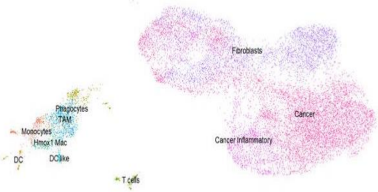
⇒ TF-MIIC can scale to hundreds TFs and thousands non-TFs

Franck Simon

Dissecting complex biological responses in scRNAseq data

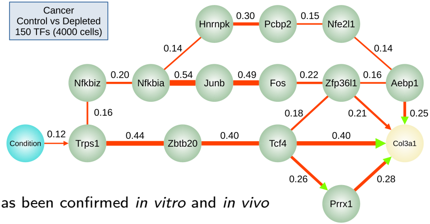
Macrophages prevent Tcell infiltration in tumors by controling ECM remodeling by cancer cells





Data from: Zoé Fusilier & Hélène Moreau (Institut Curie, U932)
Analysis: Franck Simon (Isambert's team)

Cancer
Control vs Depleted
150 TFs (4000 cells)



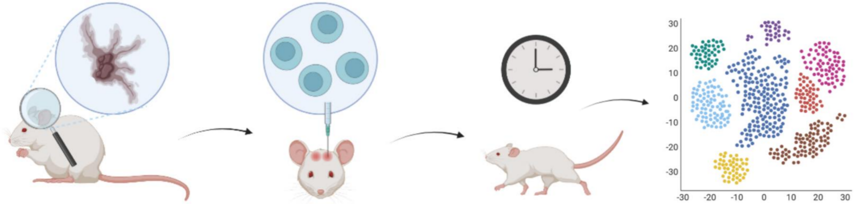
Tcf4 —▶ Col3a1 has been confirmed *in vitro* and *in vivo*

Fusilier *et al.* Science Immunology 2026

Learning Gene Network Response to Cellular Therapy

Analysis of scRNAseq data from Fillatreau Lab, Institut Necker Enfants Malades

Cellular therapy against Multiple Sclerosis
(injection of IL-10 producing Bregs):



EAE induced mice
(Multiple Sclerosis model for mice)

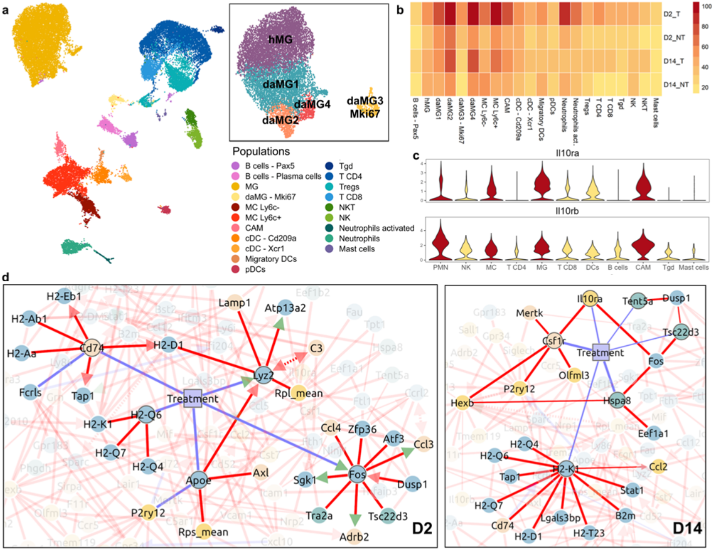
Injection of modified B cells at the peak of the disease (EAE score)

After a few days the mice is cured and behaving normally

Single-Cell analysis on Brain and Spinal Cord samples

Analysis of scRNAseq data: Louise Dupuis

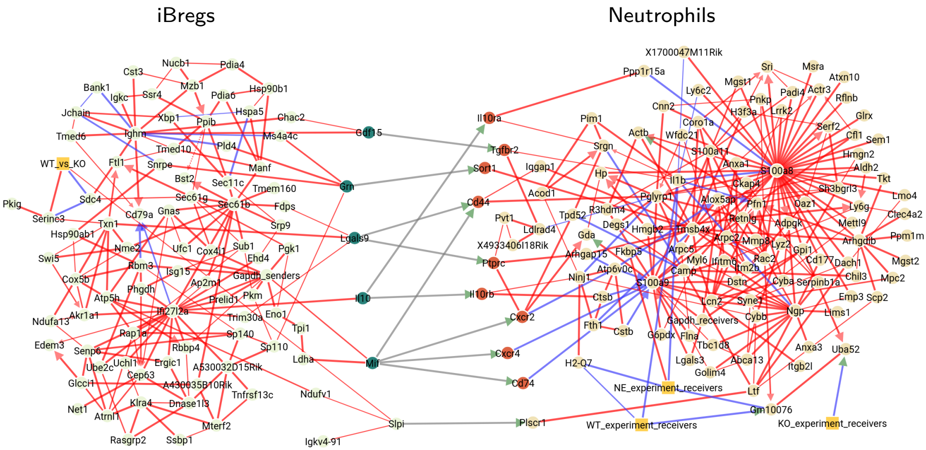
Learning Gene Network Response to Cellular Therapy



Manfroi *et al.* Induced regulatory B cells stably expressing IL-10 cure CNS autoimmunity by targeting microglia. *under revision* (2024)

Analysis of scRNAseq data
Louise Dupuis

Integrating scRNAseq and Cell-Cell Communication data



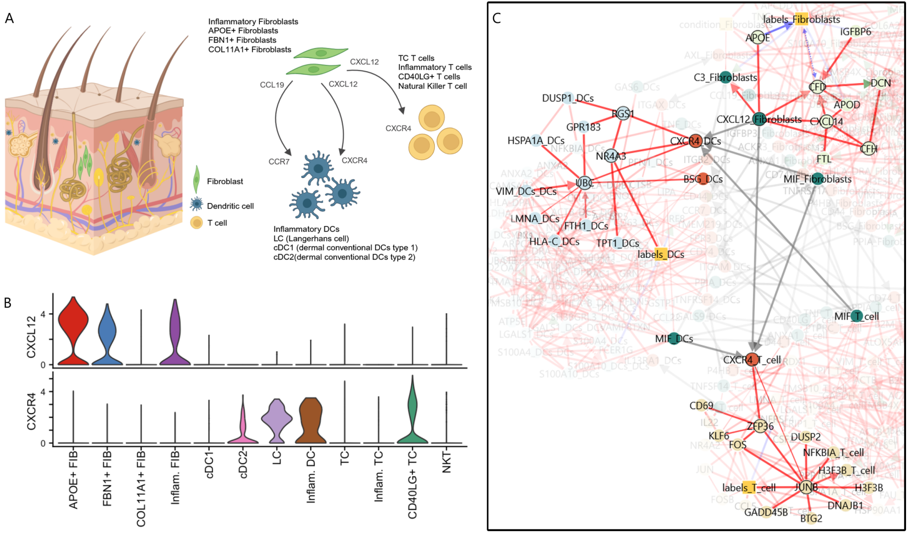
Bui-Thi *et al.* Plasmocytes inhibit CNS autoimmunity through SLPI production, *under review* (2024)

Dupuis *et al.* CausalCCC: a web server to explore intracellular causal pathways enabling cell-cell communication, NAR 2025

⇒ `https://miic.curie.fr/causalCCC.php`

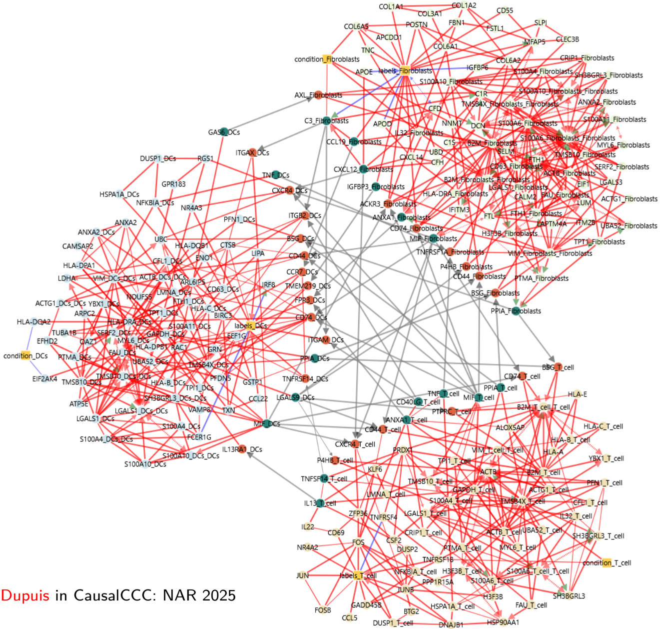
Extension to CCC Pathways across Multiple Cell Types

CausalCCC network between fibroblasts, dendritic cells, and T cells in atopic dermatitis



(**A**) Data: He *et al.* JACI 2020 (**B**) CellChat L-R pairs, Jin *et al.* Nat Commun 2021 (**C**) CausalCCC: Dupuis *et al.* NAR 2025

Extension to CCC Pathways across Multiple Cell Types



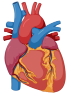
Louise Dupuis in CausalCCC: NAR 2025

CausalCCC Pathways from Spatial Transcriptomics Data

Myocardial Infarction (ST data: Kuppe *et al.* Nature 2022)

A

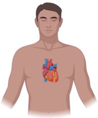
Original study: Spatial multi-omic map of human myocardial infarction (Kuppe *et al*)



31 samples from 23 patients

snRNAseq: 191,795 cells
snATACseq: 46,086 cells
26 Spatial Transcriptomics (10X)

Subset: Patient P9
10X Visium slide



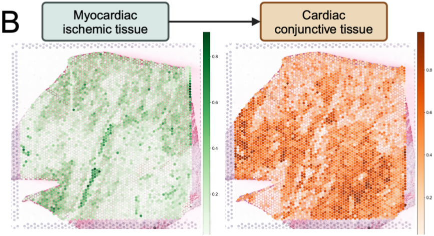
Input: n= 4,084 cells in ischemic zone

↓

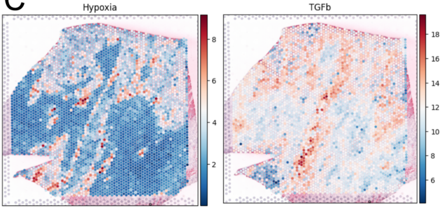
CausalCCC default pipeline

80 sender nodes, 80 receiver nodes, 17 L-R links, SPP1 as gene of interest

B



C



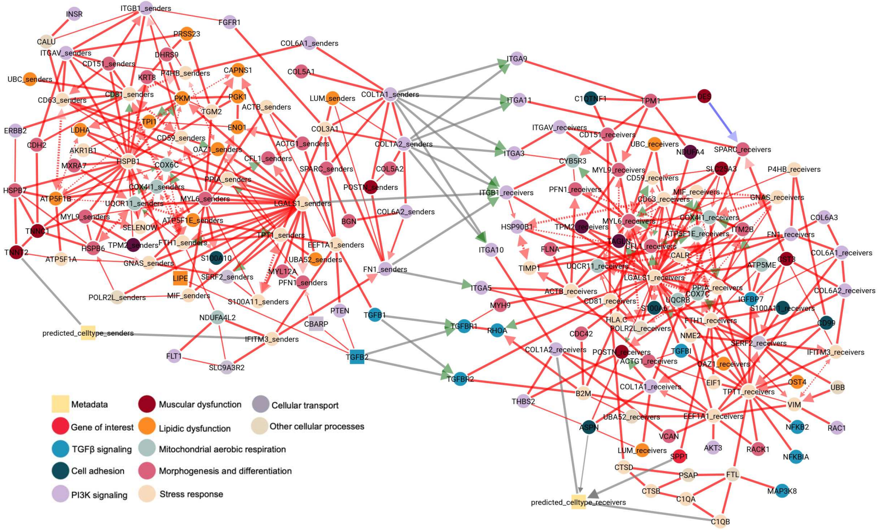
⇒ **Cell-Cell Communication** between:
ischemic myocardial cells (including cardiomyocytes, vascular smooth muscle cells, endothelial cells, pericytes) and connective tissue (including fibroblasts and myeloid cells)

Orianne Debeaupuis

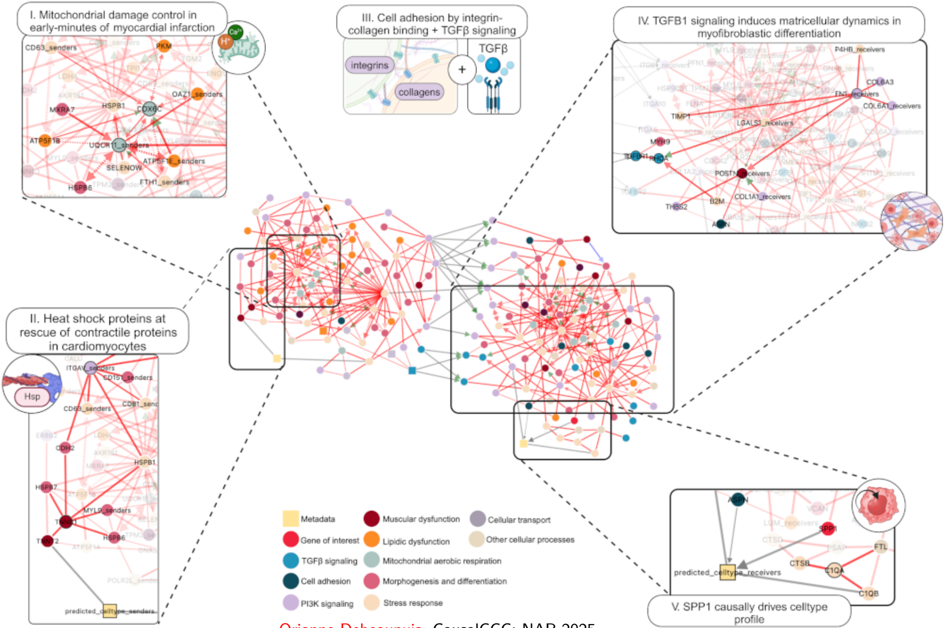
CausalCCC: NAR 2025

CausalCCC Pathways from Spatial Transcriptomics Data

Ischemic myocardial cells ⇒ cells from connective tissue



CausalCCC Pathways from Spatial Transcriptomics Data



I. Mitochondrial damage control in early-minutes of myocardial infarction

II. Heat shock proteins at rescue of contractile proteins in cardiomyocytes

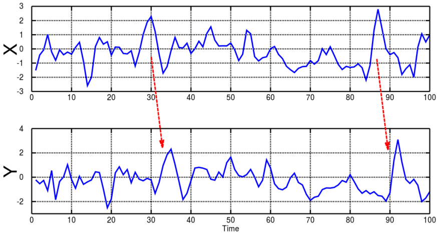
III. Cell adhesion by integrin-collagen binding + TGFβ signaling

IV. TGFB1 signaling induces matricellular dynamics in myofibroblastic differentiation

V. SPP1 causally drives celltype profile

Orianne Debeaupuis CausalCCC: NAR 2025

Extension to Time Series Data $X_{t'<t} \to Y_t$

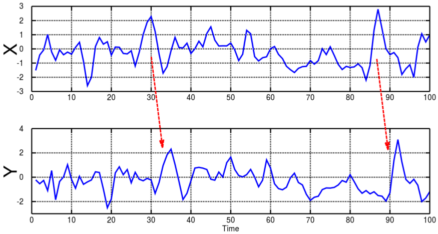


Granger Causality (Granger *Econometrica* 1969)

→ Best regression of Y_t with or without lagged values of $X_{t'<t}$?

$$Y_t = a_0 + a_1 Y_{t-1} + a_2 Y_{t-2} + \cdots + a_m Y_{t-m} + b_p X_{t-p} + \cdots + b_q X_{t-q} + \epsilon_t$$

Extension to Time Series Data $X_{t'<t} \rightarrow Y_t$



Granger Causality (Granger *Econometrica* 1969)

$\rightarrow$ Best regression of Y_t with or without lagged values of $X_{t'<t}$?

$$Y_t = a_0 + a_1 Y_{t-1} + a_2 Y_{t-2} + \cdots + a_m Y_{t-m} + b_p X_{t-p} + \cdots + b_q X_{t-q} + \epsilon_t$$

Transfer Entropy (Schreiber *PRL* 2000)

$\rightarrow$ $T_{X\rightarrow Y} = I(Y_t; \boldsymbol{X}_{t'<t} | \boldsymbol{Y}_{t'<t})$ *versus* $T_{Y\rightarrow X} = I(X_t; \boldsymbol{Y}_{t'<t} | \boldsymbol{X}_{t'<t})$

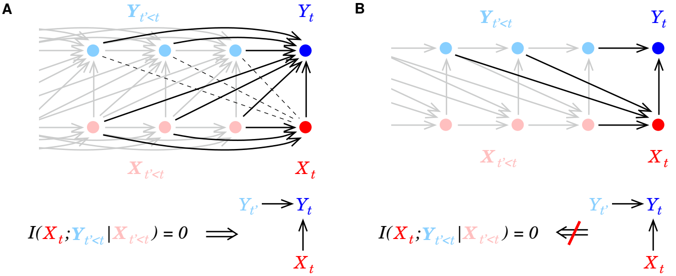
Then, if $T_{X\rightarrow Y} > 0$ & $T_{Y\rightarrow X} \simeq 0 \implies X_{t'<t} \rightarrow Y_t$

Although always the possibility of a latent variable: $X_{t'} \dashleftarrow L_{t''} \dashrightarrow Y_t$, with $t'' < t' < t$

Temporal Causality $\leftrightsquigarrow$ Graph-based Causality ??

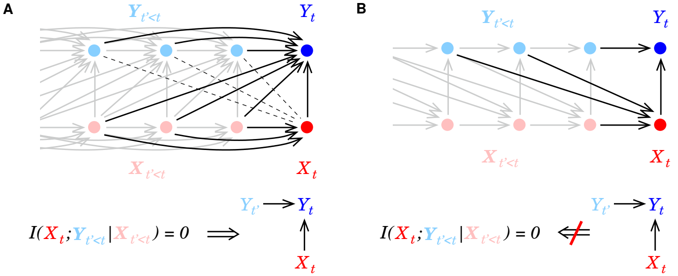
Temporal Causality $\leftrightsquigarrow$ Graph-based Causality ??

IDEA: **Information-based reasoning** on **time-unfolded graphs** $\mathcal{G}_t$



Temporal Causality ⇜⇝ Graph-based Causality ??

IDEA: **Information-based reasoning** on **time-unfolded graphs** $\mathcal{G}_t$



Theorem 1 [Temporal Causality $\implies$ Time-unfolded Graph-based Causality]

If $Y_t - X_t$ & $T_{Y\to X} = I(X_t; \boldsymbol{Y}_{t'<t} | \boldsymbol{X}_{t'<t}, \boldsymbol{V}_{t'\leqslant t}^{\setminus X,Y}) = 0$

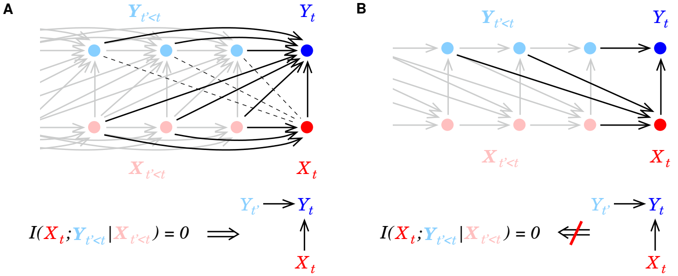
Then $\forall\, Y_{t'} - Y_t$, $Y_{t'} \to Y_t \leftarrow X_t$ [& $\forall t' < t$, $Y_{t'} \not\dashrightarrow X_t$]

But **the converse** is not true !

F Simon *et al.*, eLife 2025

Temporal Causality $\leftrightsquigarrow$ Graph-based Causality ??

IDEA: **Information-based reasoning** on **time-unfolded graphs** $\mathcal{G}_t$



Theorem 1 [Temporal Causality $\Longrightarrow$ Time-unfolded Graph-based Causality]

If $Y_t - X_t$ & $T_{Y\to X} = I(X_t; \boldsymbol{Y}_{t'<t} | \boldsymbol{X}_{t'<t}, \boldsymbol{V}^{\setminus X,Y}_{t'\leqslant t}) = 0$

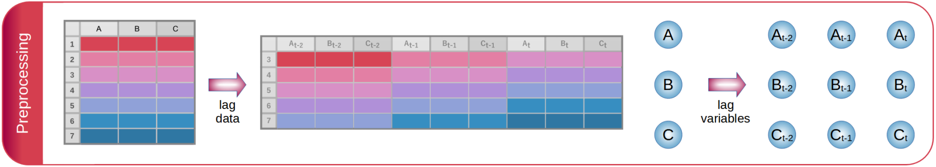
Then $\forall\ Y_{t'} - Y_t,\quad Y_{t'} \to Y_t \leftarrow X_t \quad [\ \&\ \ \forall t' < t,\ Y_{t'} \not\to X_t\ ]$

But **the converse** is not true !

F Simon *et al.*, eLife 2025

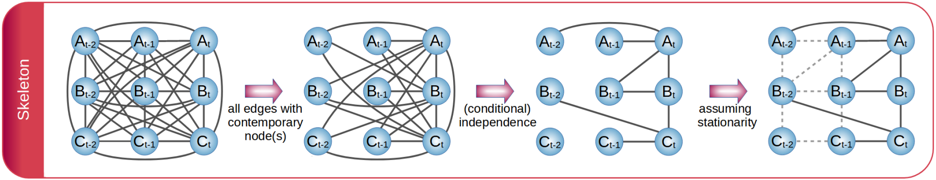
Hence, Granger-Schreiber **temporal causality** is **too restrictive** and may **miss actual causal relations** uncovered by **graph-based causal discovery** methods for time series

tMIIC: Causal discovery from time series data


Preprocessing

lag data

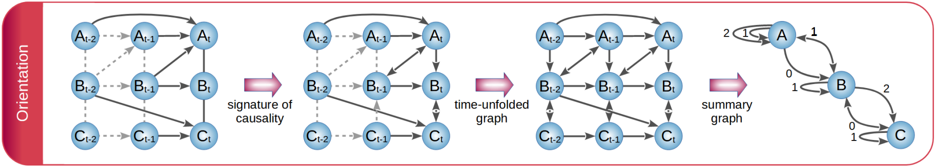
lag variables


Skeleton

all edges with contemporary node(s)

(conditional) independence

assuming stationarity


Orientation

signature of causality

time-unfolded graph

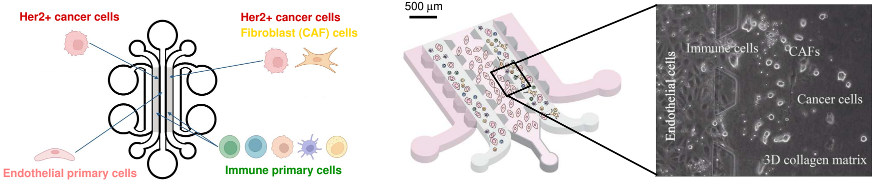
summary graph

F Simon *et al.*, eLife 2025

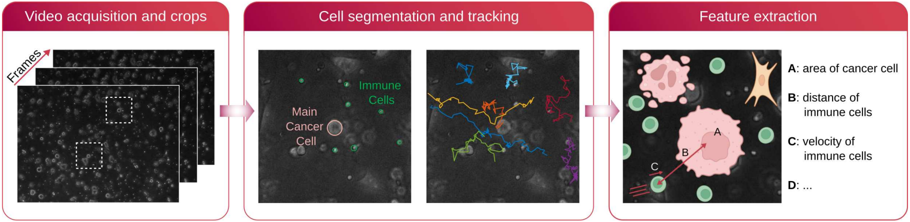
CausalXtract: Causal discovery from live cell images

Analysing ex-vivo tumor-immune response ecosystems

a Tumor–on–chip preparation



b CausalXtract’s live–cell image feature extraction module (CellHunter+)



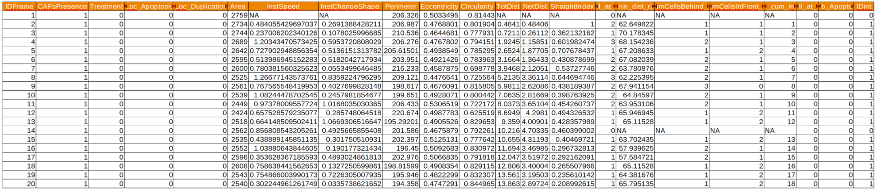
Collaboration with Maria Carla Parrini IC & Martinelli lab, Uni Rome

Simon *et al* eLife 2025

CausalXtract: Causal discovery from live cell images

From the experiments, we get a dataset with 21 variables about:

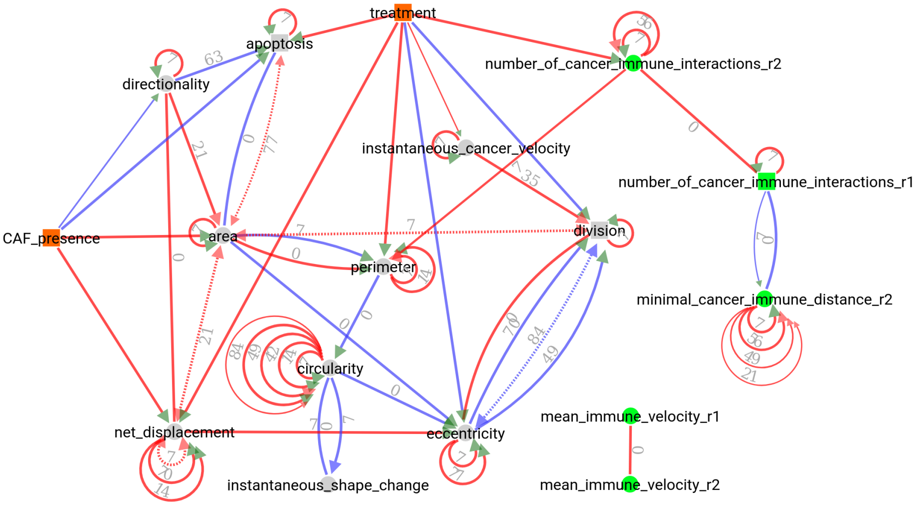
- Experimental conditions (CAF presence and Treatment)
- Cell Shape
- Cell Motility
- Cell division
- Cell apoptosis
- Cell Interactions with immune cells
- Cell Motility of immune interacting cells



IDFrame	CAFsPresence	Treatment	Loc_Apoptosis	Loc_Duplicatio	Area	InstSpeed	InstChangeShape	Perimeter	Eccentricity	Circularity	TotDist	NetDist	StraightIndex	ll_at_	min_dist_r2	umCellsBehind_	umCellsInFront_	_cum_in	ll_at_	b_Apopto	IDint
1	1	0	0	0	2759	NA	NA	206.326	0.5033495	0.81443	NA	NA	NA	0	NA	NA	NA	NA	0	0	0
2	1	0	0	0	2734	0.484055429697037	0.2691388428211	206.987	0.4768801	0.801904	0.4841	0.48406	1	2	62.649822	1	1	1	0	0	1
3	1	0	0	0	2744	0.237006202340126	0.1078025996685	210.536	0.4644681	0.777931	0.7211	0.26112	0.362132162	1	70.178345	1	1	2	0	0	1
4	1	0	0	0	2689	1.20343470573425	0.5953720808029	206.276	0.4767802	0.794151	1.9245	1.15851	0.601982474	3	68.154236	2	1	3	0	0	1
5	1	0	0	0	2642	0.727902948856354	0.5136151313782	205.61501	0.4938549	0.785295	2.6524	1.87705	0.707678437	1	67.208633	1	2	4	0	0	1
6	1	0	0	0	2595	0.513986945152283	0.5182042717934	203.951	0.4921426	0.783963	3.1664	1.36433	0.430878699	2	67.082039	2	1	5	0	0	1
7	1	0	0	0	2600	0.780381560325623	0.0553499646485	216.233	0.4587875	0.698778	3.9468	2.12051	0.53727746	2	63.780876	2	1	6	0	0	1
8	1	0	0	0	2525	1.26677143573761	0.8359224796295	209.121	0.4476641	0.725564	5.2135	3.36114	0.644694746	3	62.225395	2	1	7	0	0	1
9	1	0	0	0	2561	0.767565548419953	0.4027699828148	198.617	0.4676091	0.815805	5.9811	2.62086	0.438189387	2	67.941154	3	0	8	0	0	1
10	1	0	0	0	2539	1.08244478702545	0.2457981854677	199.651	0.4928071	0.800442	7.0635	2.81669	0.398763925	2	64.84597	2	1	9	0	0	1
11	1	0	0	0	2449	0.97378009557724	1.0168035030365	206.433	0.5306519	0.722172	8.0373	3.65104	0.454260737	2	63.953106	2	1	10	0	0	1
12	1	0	0	0	2424	0.657528579235077	0.285748064518	220.674	0.4987783	0.625519	8.6949	4.2981	0.494326532	1	65.946945	1	2	11	0	0	1
13	1	0	0	0	2518	0.664148509502411	1.0669306516647	195.29201	0.4905526	0.829653	9.359	4.00901	0.428357989	1	65.11528	1	2	12	0	0	1
14	1	0	0	0	2562	0.856808543205261	0.4925665855408	201.586	0.4675879	0.792261	10.216	4.70335	0.460399002	0	NA	NA	NA	NA	0	0	0
15	1	0	0	0	2535	0.438889145851135	0.301750510931	202.397	0.5125131	0.777642	10.655	4.31193	0.40469721	1	63.702435	1	2	13	0	0	1
16	1	0	0	0	2552	1.03880643844605	0.190177321434	196.45	0.5092683	0.830972	11.694	3.46985	0.296732813	2	57.939625	2	1	14	0	0	1
17	1	0	0	0	2596	0.353628367185593	0.4893024861813	202.976	0.5066835	0.791818	12.047	3.51972	0.292162091	1	57.584721	2	1	15	0	0	1
18	1	0	0	0	2608	0.758638441562653	0.1327250599861	198.81599	0.4908354	0.829115	12.806	3.40004	0.265507966	1	65.11528	1	2	16	0	0	1
19	1	0	0	0	2543	0.754866003990173	0.7226305007935	195.946	0.4822299	0.832307	13.561	3.19503	0.235610142	1	64.381676	1	2	17	0	0	1
20	1	0	0	0	2540	0.302244961261749	0.0335738621652	194.358	0.4747291	0.844965	13.863	2.89724	0.208992615	1	65.795135	1	2	18	0	0	1

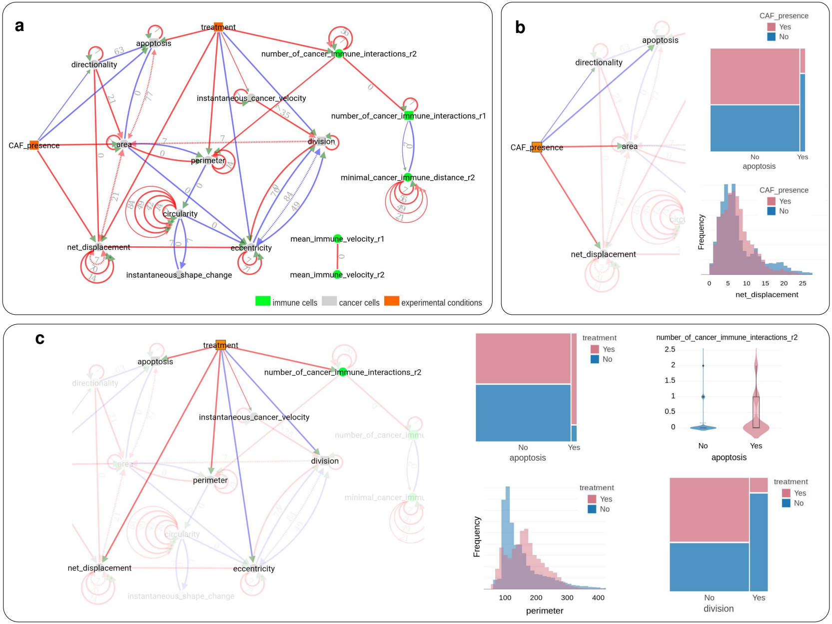
The dataset contains 36 trajectories including upto 1440 time-lapse images ($\delta\tau = 2$min), corresponding to a total of 46,935 images. $\rightarrow$ avg relaxation time $25\delta\tau$: 12 layers of $5\delta\tau$

CausalXtract: Causal discovery from live cell images

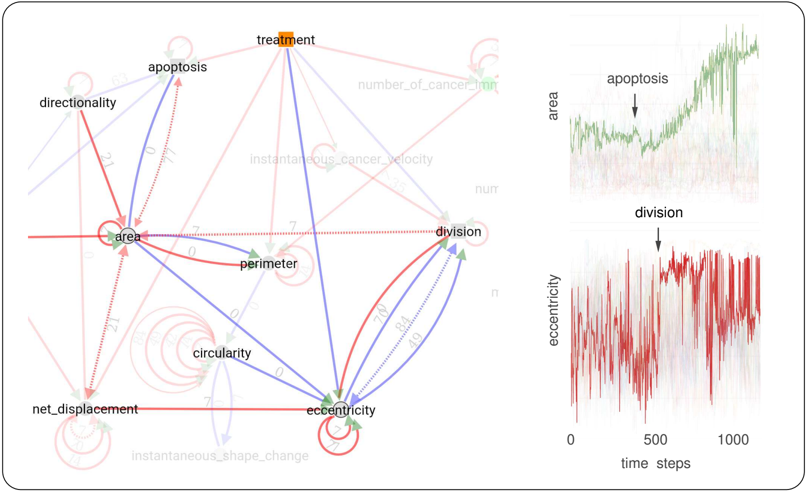


Simon *et al.* eLife 2025

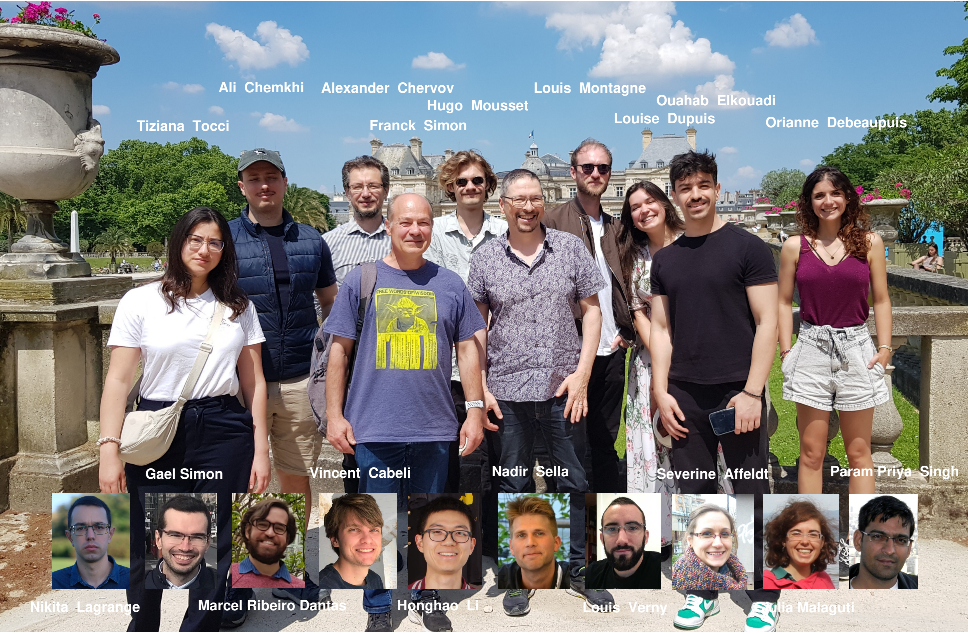
CausalXtract: Causal discovery from live cell images



CausalXtract: Opposite effects at different time lags



Simon *et al.* eLife 2025



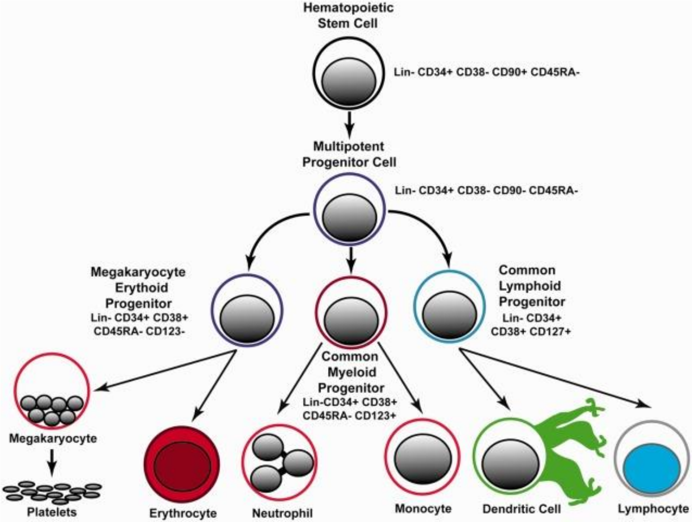
institut Curie
PSL
cnrs
PR[AI]RIE Paris School
anr
Inserm
Funded by the European Union
FRANCE 2030
Fondation pour la Recherche Médicale
HEALTH DATA HUB
Breakthrough T1D

The Team (Lab retreat, Luminy, Marseille, Sept 2023)

Franck, Gaël, Hervé, Nikita, Louis, Louise, Tiziana, Orianne, Liza

Coll: Fillatreau (INEM), Charbord (SU), Bravi (ICL), Parrini (U830), Perie, Hersen (UMR168), Hamy, Moreau, Lennon, Salmon (U932)

Application to Human Hematopoietic Stem Cells



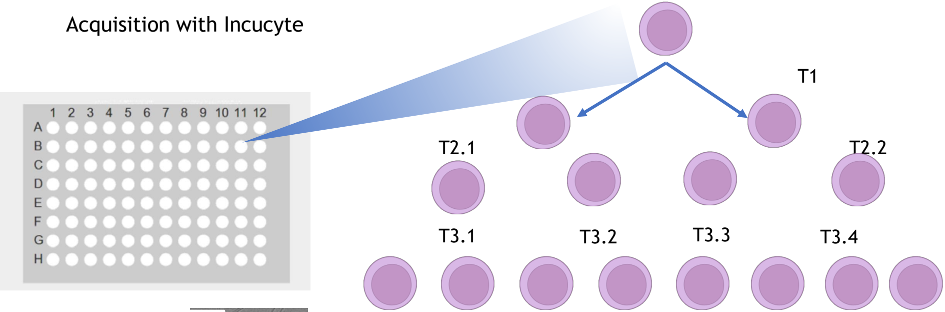
Cimato et al., 2015

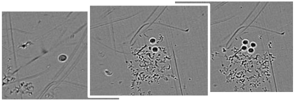
collab. with Perié lab, Institut Curie

Application to Human Hematopoietic Stem Cells

Data collection

Acquisition with Incucyte





Cell segmentation project

Frames from t0 to t=90 hours

Sequencing at 96 hours (SMART-SEQ3)

scRNASeq project

collab. with Perié lab, Institut Curie

Application to Human Hematopoietic Stem Cells

Cell counting

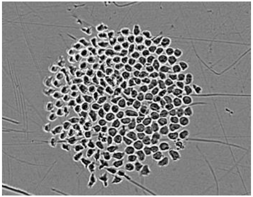
Two key properties:

- synchronicity in **division**
- similarity in fate

among progenies of individual stem and progenitor cells

Current goal:

- automatise cell counting after the first **mitosis** (= segmentation)



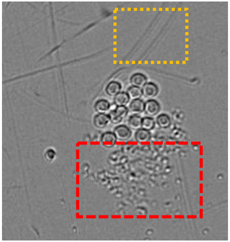
In collaboration with **Eugenio Martinelli**, **Arianna Mencattini** University of Rome Tor Vergata)

collab. with Perié lab, Institut Curie

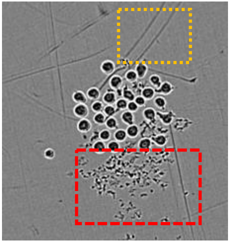
Application to Human Hematopoietic Stem Cells

Issues

- Out of focus frames
- Small cells
- Debris ⬚
- Plastic defects ⬚



Out of focus



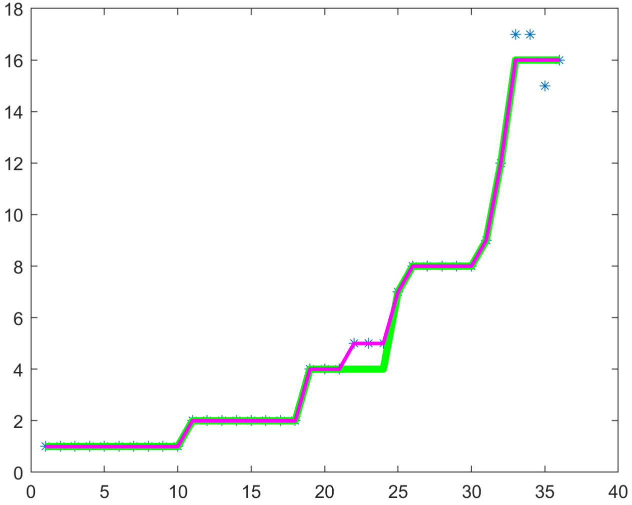
In focus

Deep Learning could handle better these issues

collab. with Perié lab, Institut Curie

Application to Human Hematopoietic Stem Cells

Stepwise correction



collab. with Perié lab, Institut Curie

Application to Human Hematopoietic Stem Cells

scRNASeq data

Raw matrixes

- Dataset1: 70 cells, 28750 genes
- Dataset2: 72 cells, 30682 genes

Metadata

- Family (or well) – 30 families
- Times of division: T1, T2.i, T3.j
- Colony size (at 90 hours)
- Fate/predicted cell type

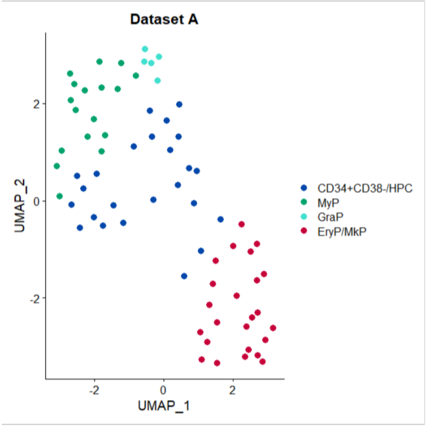
Sequencing at 96 hours

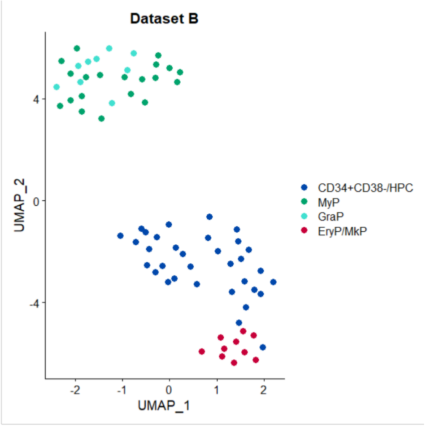
Cell type: MPP

collab. with Perié lab, Institut Curie

Application to Human Hematopoietic Stem Cells

Cell annotation

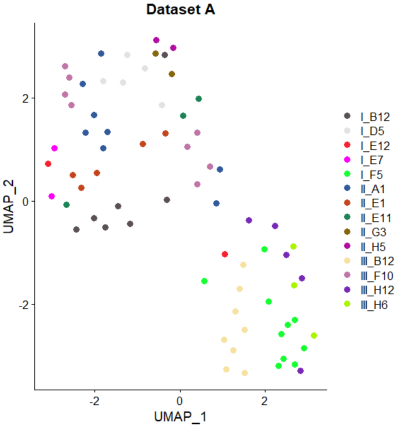


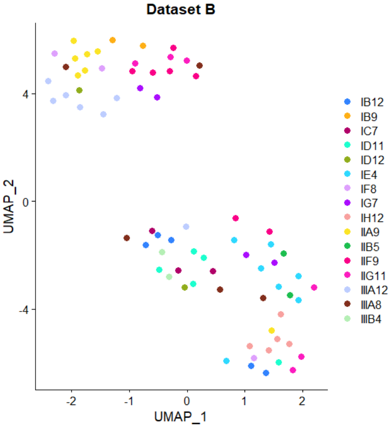


collab. with Perié lab, Institut Curie

Application to Human Hematopoietic Stem Cells

Families tend to cluster together

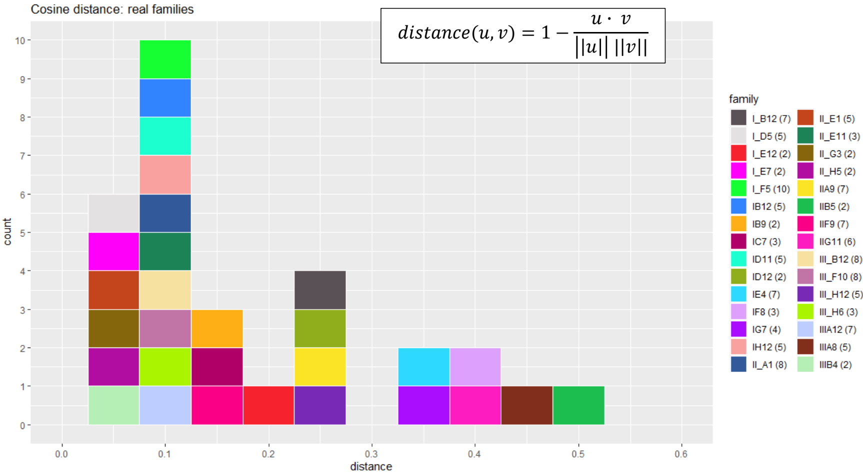




collab. with Perié lab, Institut Curie

Application to Human Hematopoietic Stem Cells

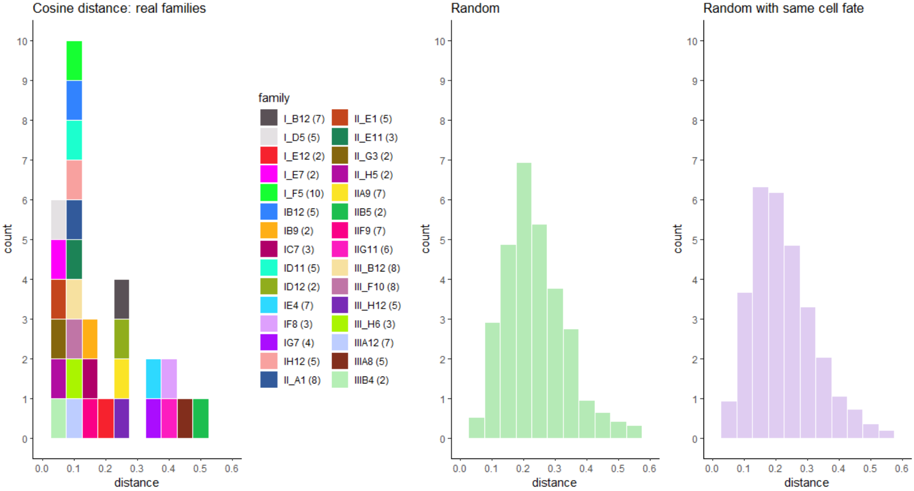
Cosine distance



$$distance(u,v) = 1 - \frac{u \cdot v}{||u||\;||v||}$$

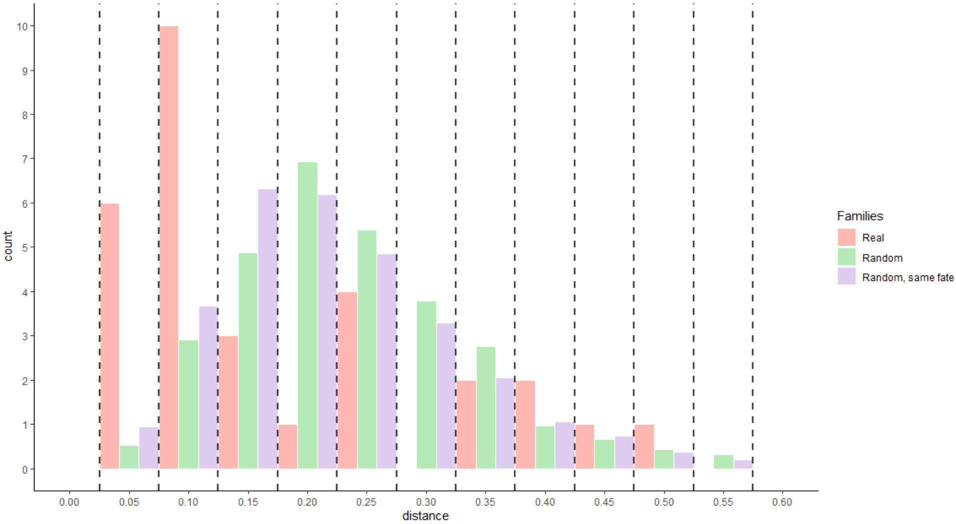
collab. with Perié lab, Institut Curie

Application to Human Hematopoietic Stem Cells



collab. with Perié lab, Institut Curie

Application to Human Hematopoietic Stem Cells



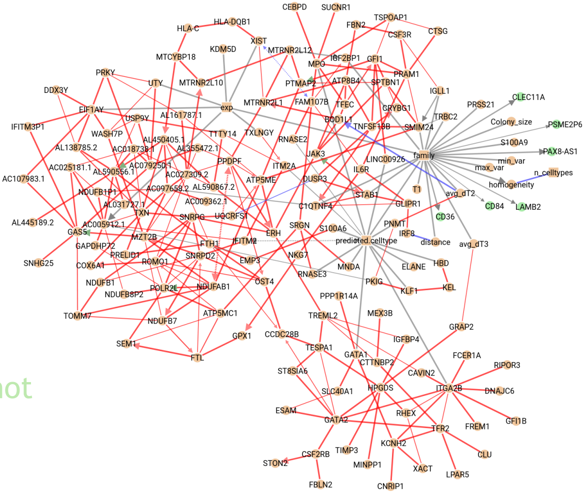
collab. with Perié lab, Institut Curie

Application to Human Hematopoietic Stem Cells

TF-MIIC

Franck Simon

- Regular MIIC: up to 500 variables
- TF (regulators)-MIIC: a few thousands of variables
- Variables: regulators or not regulators

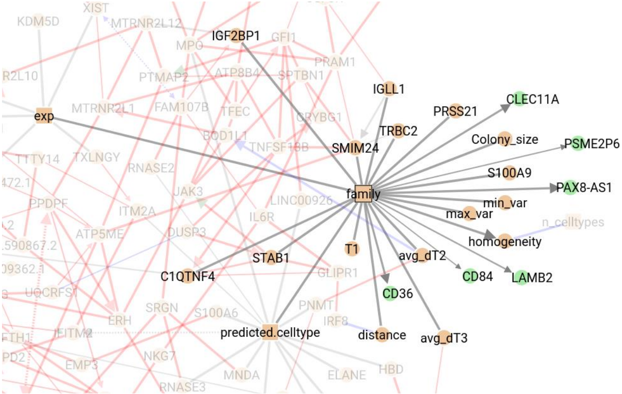


collab. with Perié lab, Institut Curie

Application to Human Hematopoietic Stem Cells

MIIC networks

- Feature selection: based on **Mutual Information**

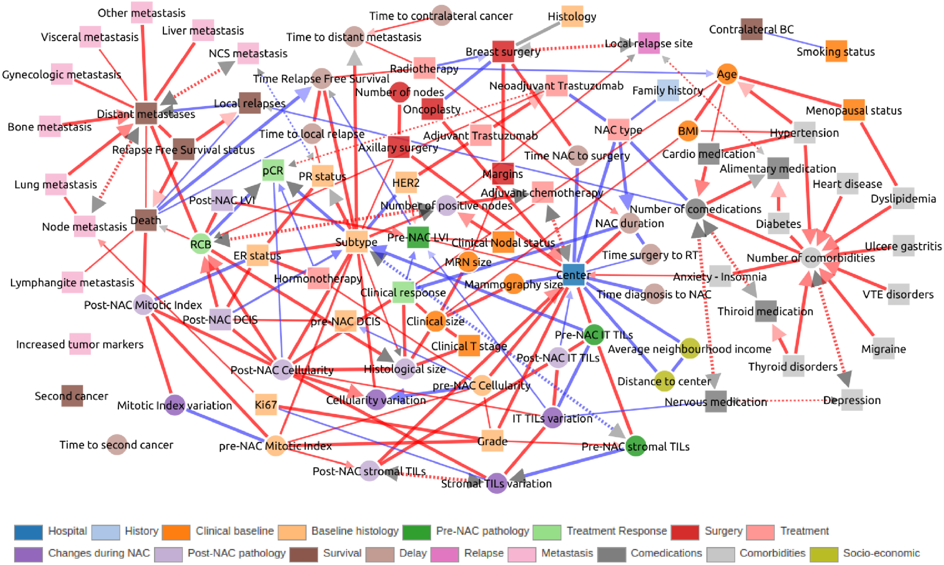


collab. with Perié lab, Institut Curie

Clinical network from a large Breast Cancer cohort

1,200 patients, NEOREP cohort, Curie Hospitals.

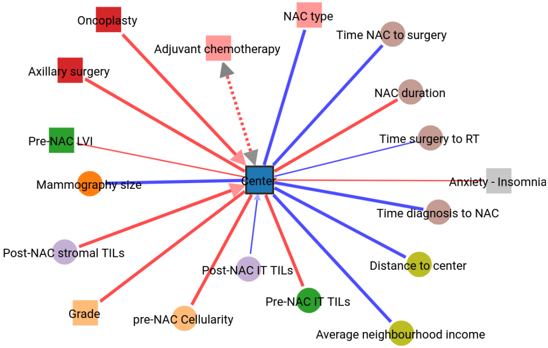
Sella *et al.* npj Digital Medicine, 2022

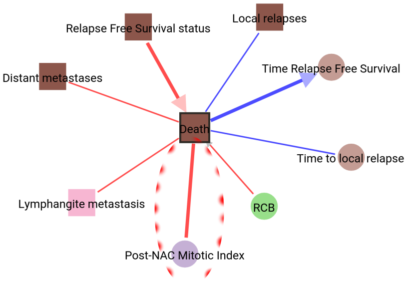


MIIC finds unexpected biases and new prognotic markers

NEOREP cohort: 1,200 patients, Reyal lab, Curie Hospitals.

Sella *et al.* npj Digital Medicine 2022



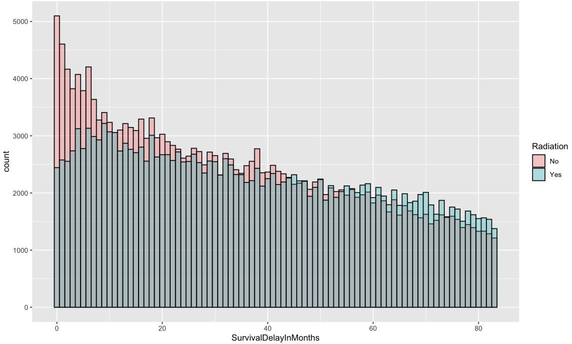


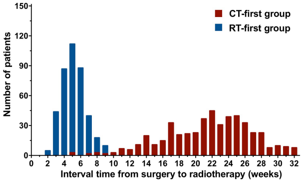
Center bias (Paris / St Cloud hospitals):
→ Different clinical profiles
Different clinical practices

Post-NAC Mitotic Index:
→ New prognotic markers
in addition to RCB

SEER Database: Survival subnetwork

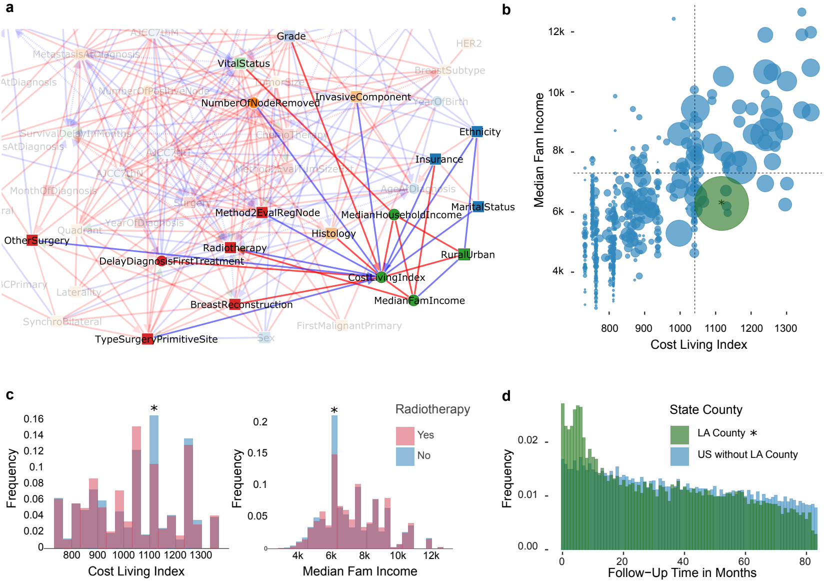
Interpretation of causal edge: Vital Status —▶ Radiotherapy in terms of ≈ 6-7 month delays



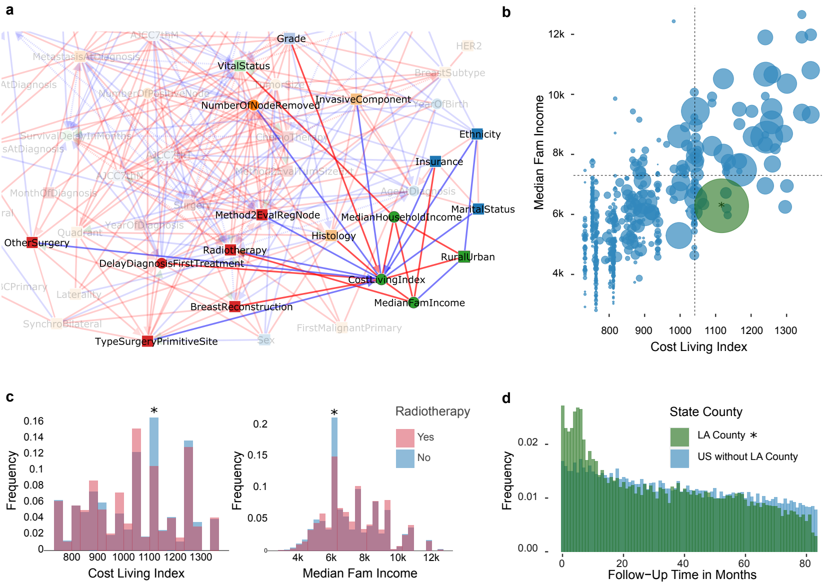


Chen et al. Front Onco 2020

SEER Database: Socio-economic subnetwork



SEER Database: Socio-economic subnetwork



∗ L.A. county (10% of dataset): 32.6% Radiotherapy in L.A. *vs* 50% rest of US; 7% of L.A. patients drop out *vs* 1.5% rest of US

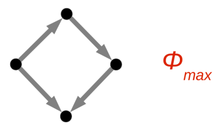
How to Learn Graphical Models from Observational Data

a one-slide primer

1. Search & Score ('Bayesian', scoring function ϕ)

Find $\mathcal{G}$ that **maximizes** Likelihood $\phi_{\mathcal{G}} = \mathcal{L}_{\mathcal{D}|\mathcal{G}}$

Super-exponential space of networks, only $\longrightarrow$



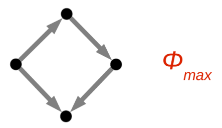
How to Learn Graphical Models from Observational Data

a one-slide primer

1. Search & Score ('Bayesian', scoring function ϕ)

Find $\mathcal{G}$ that **maximizes** Likelihood $\phi_{\mathcal{G}} = \mathcal{L}_{\mathcal{D}|\mathcal{G}}$

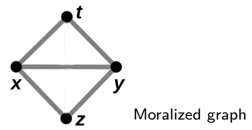
Super-exponential space of networks, only $\longrightarrow$



2. Inversion of variance-covariance matrix Σ

Compute $L = D - A \sim \Sigma^{-1}$ (assuming linearity)

Invertibility/Stability issue and only — edges



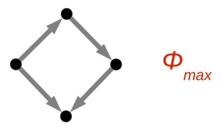
How to Learn Graphical Models from Observational Data

a one-slide primer

1. **Search & Score** ('Bayesian', scoring function ϕ)

Find $\mathcal{G}$ that **maximizes** Likelihood $\phi_{\mathcal{G}} = \mathcal{L}_{\mathcal{D}|\mathcal{G}}$

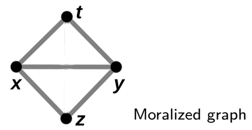
Super-exponential space of networks, only →



2. **Inversion of variance-covariance matrix Σ**

Compute $L = D - A \sim \Sigma^{-1}$ (assuming linearity)

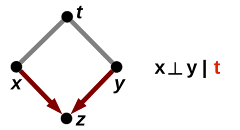
Invertibility/Stability issue and only — edges



3. **Constraint-based** (Conditional independences)

Broader network class including — → ↔

Slow and **Not robust to sampling noise**



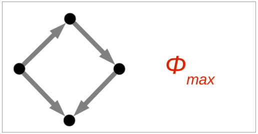
How to Learn Graphical Models from Observational Data

a one-slide primer

1. Search & Score ('Bayesian', scoring function ϕ)

Find $\mathcal{G}$ that **maximizes** Likelihood $\phi_{\mathcal{G}} = \mathcal{L}_{\mathcal{D}|\mathcal{G}}$

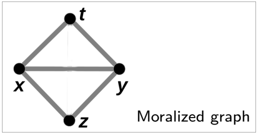
Super-exponential space of networks, only →



2. Inversion of variance-covariance matrix Σ

Compute $L = D - A \sim \Sigma^{-1}$ (assuming linearity)

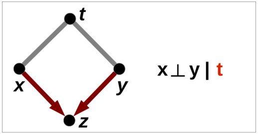
Invertibility/Stability issue and only — edges



3. Constraint-based (Conditional independences)

Broader network class including — → ↔

Slow and **Not robust to sampling noise**



MIIC: a **robust Network Inference** method combining **contraints** and **scores**

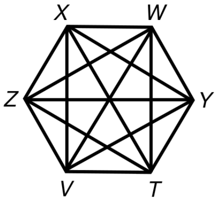
Affeldt *et al.* 2015; Verny *et al.* 2017; Cabeli *et al.* 2020; Ribeiro-Dantas *et al.* 2024; Simon *et al.* 2025; Lagrange *et al.* 2025

MIIC combines Information theory with Causal discovery

MIIC first **substracts** iteratively the **most significant contributions** to each $I(X,Y)$

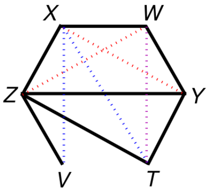
$$I(X;Y) - I(X;Y;A_1) - I(X;Y;A_2|A_1) - \cdots - I(X;Y;A_n|\{A_i\}_{n-1}) \;=\; I(X;Y|\{A_i\})$$

(0) Complete graph



(1) Remove edges

$I(X;Y|\{A_i\}) \simeq 0$



$I(X,T|Y,Z) \simeq 0$

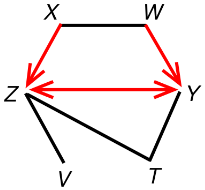
$I(X;Y|W) \simeq 0$

$I(X;V|Z) \simeq 0$

$I(W;Z|X) \simeq 0$

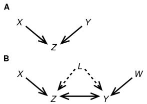
$I(W;T|Y,Z) \simeq 0$

(2) Orient V-structures (p_{head})

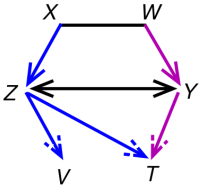


$I(X;Y;Z|W) < 0$

$I(W;Z;Y|X) < 0$



(3) Propagate orientations (p_{tail})

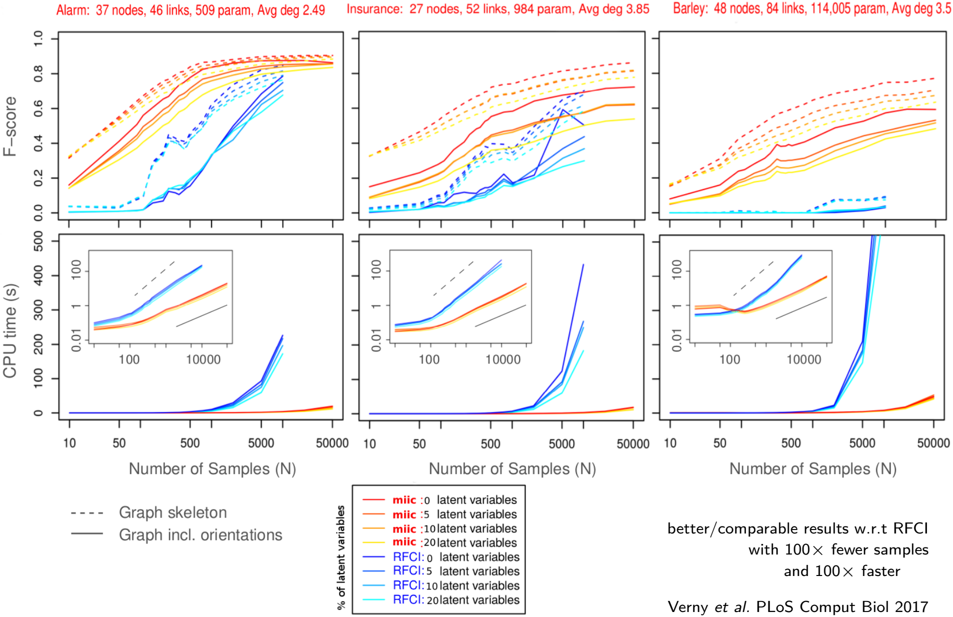


$I(X;V;Z) > 0$

$I(X;T;Z|Y) > 0$

$I(W;T;Y|Z) > 0$

Benchmark on Ancestral Graphs (including — → ↔ edges)

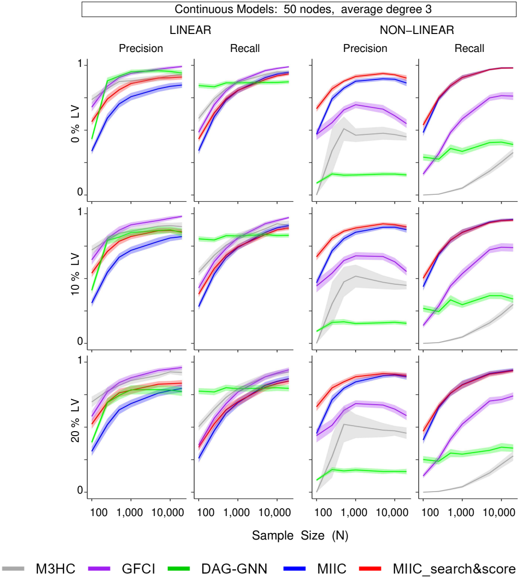


better/comparable results w.r.t RFCI
with 100× fewer samples
and 100× faster

Verny *et al.* PLoS Comput Biol 2017

Benchmarks on Linear *vs* Non-linear Continuous Data

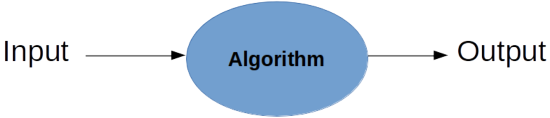
MIIC does not assume linear distribution ⇒ MIIC outperforms other methods on non-linear data



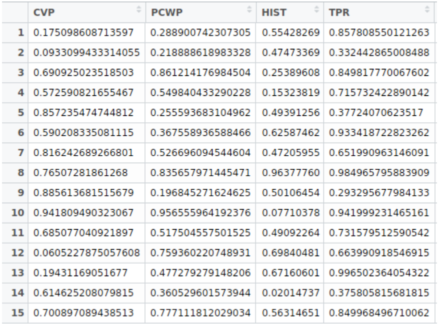
Lagrange & Isambert, ICML 2025

What is network reconstruction ?

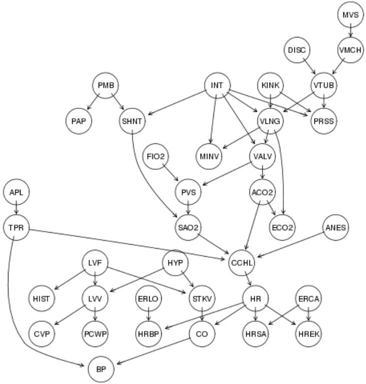
Disentangling direct from indirect relations between variables, including *or* excluding cause-effect relationships.



Input → **Algorithm** → Output

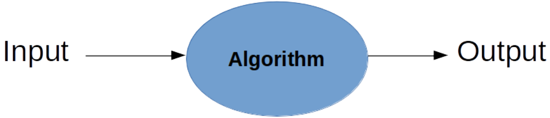


	CVP	PCWP	HIST	TPR
1	0.175098608713597	0.288900742307305	0.55428269	0.857808550121263
2	0.0933099433314055	0.218888618983328	0.47473369	0.332442865008488
3	0.690925023518503	0.861214176984504	0.25389608	0.849817770067602
4	0.572590821655467	0.549840433290228	0.15323819	0.715732422890142
5	0.857235474744812	0.255593683104962	0.49391256	0.37724070623517
6	0.590208335081115	0.367558936588466	0.62587462	0.933418722823262
7	0.816242689266801	0.526696094544604	0.47205955	0.651990963146091
8	0.76507281861268	0.835657971445471	0.96377760	0.984965795883909
9	0.885613681515679	0.196845271624625	0.50106454	0.293295677984133
10	0.941809490323067	0.956555964192376	0.07710378	0.941999231465161
11	0.685077040921897	0.517504557501525	0.49092264	0.731579512590542
12	0.0605227875057608	0.759360220748931	0.69840481	0.663990918546915
13	0.19431169051677	0.477279279148206	0.67160601	0.996502364054322
14	0.614625208079815	0.360529601573944	0.02014737	0.375805815681815
15	0.700897089438513	0.777111812029034	0.56314651	0.849968496710062

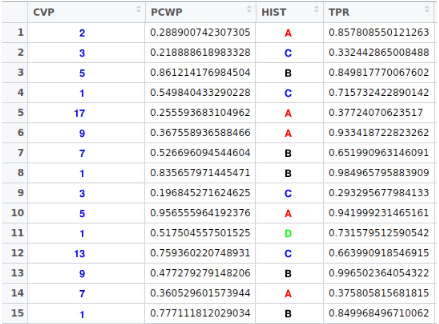


What is network reconstruction ?

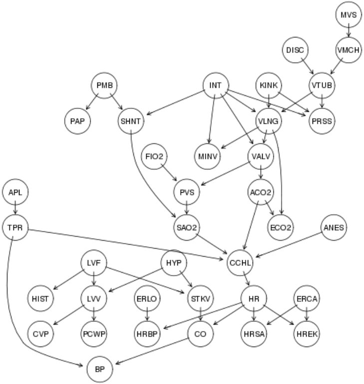
Disentangling direct from indirect relations between variables, including *or* excluding cause-effect relationships.



Input → **Algorithm** → Output

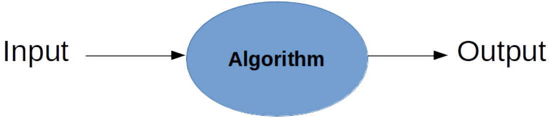


	CVP	PCWP	HIST	TPR
1	2	0.288900742307305	A	0.857808550121263
2	3	0.218888618983328	C	0.332442865008488
3	5	0.861214176984504	B	0.849817770067602
4	1	0.549840433290228	C	0.715732422890142
5	17	0.255593683104962	A	0.37724070623517
6	9	0.367558936588466	A	0.933418722823262
7	7	0.526696094544604	B	0.651990963146091
8	1	0.835657971445471	B	0.984965795883909
9	3	0.196845271624625	C	0.293295677984133
10	5	0.956555964192376	A	0.941999231465161
11	1	0.517504557501525	D	0.731579512590542
12	13	0.759360220748931	C	0.663990918546915
13	9	0.477279279148206	B	0.996502364054322
14	7	0.360529601573944	A	0.375805815681815
15	1	0.777111812029034	B	0.849968496710062



What is network reconstruction ?

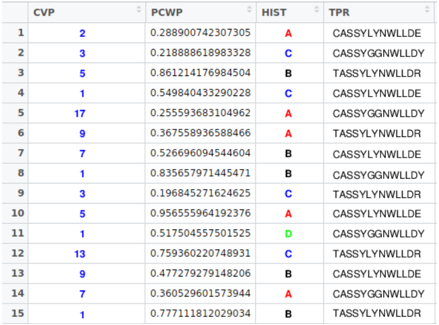
Disentangling direct from indirect relations between variables, including *or* excluding cause-effect relationships.



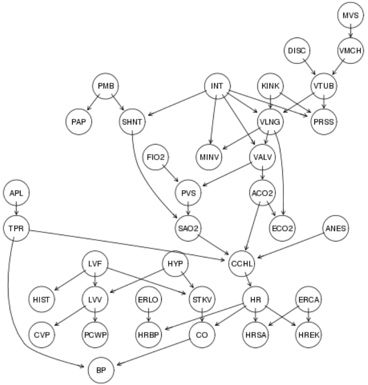
Input

Algorithm

Output

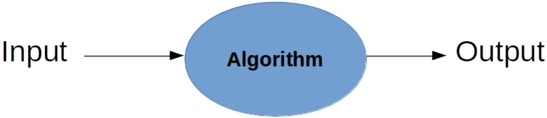


	CVP	PCWP	HIST	TPR
1	2	0.288900742307305	A	CASSYLYNWLLDE
2	3	0.218888618983328	C	CASSYGGNWLLDY
3	5	0.861214176984504	B	TASSYLYNWLLDR
4	1	0.549840433290228	C	CASSYLYNWLLDE
5	17	0.255593683104962	A	CASSYGGNWLLDY
6	9	0.367558936588466	A	TASSYLYNWLLDR
7	7	0.526696094544604	B	CASSYLYNWLLDE
8	1	0.835657971445471	B	CASSYGGNWLLDY
9	3	0.196845271624625	C	TASSYLYNWLLDR
10	5	0.956555964192376	A	CASSYLYNWLLDE
11	1	0.517504557501525	D	CASSYGGNWLLDY
12	13	0.759360220748931	C	TASSYLYNWLLDR
13	9	0.477279279148206	B	CASSYLYNWLLDE
14	7	0.360529601573944	A	CASSYGGNWLLDY
15	1	0.777111812029034	B	TASSYLYNWLLDR



What is network reconstruction ?

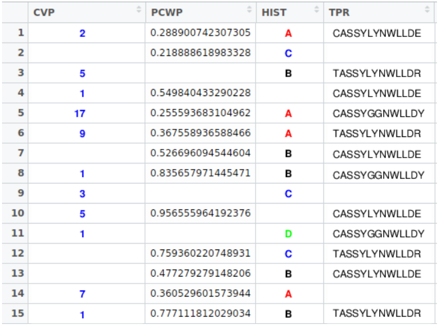
Disentangling direct from indirect relations between variables, including *or* excluding cause-effect relationships.



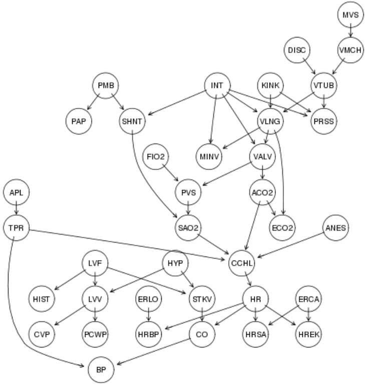
Input

Algorithm

Output



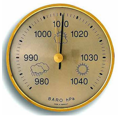
	CVP	PCWP	HIST	TPR
1	2	0.288900742307305	A	CASSYLYNWLLDE
2		0.218888618983328	C	
3	5		B	TASSYLYNWLLDR
4	1	0.549840433290228		CASSYLYNWLLDE
5	17	0.255593683104962	A	CASSYGGNWLLDY
6	9	0.367558936588466	A	TASSYLYNWLLDR
7		0.526696094544604	B	CASSYLYNWLLDE
8	1	0.835657971445471	B	CASSYGGNWLLDY
9	3		C	
10	5	0.956555964192376		CASSYLYNWLLDE
11	1		D	CASSYGGNWLLDY
12		0.759360220748931	C	TASSYLYNWLLDR
13		0.477279279148206	B	CASSYLYNWLLDE
14	7	0.360529601573944	A	
15	1	0.777111812029034	B	TASSYLYNWLLDR



Can one Infer Causation from mere Correlation ?

Can one Infer Causation from mere Correlation ?

Two-point Correlation does not imply causation



$$\underset{?}{\overset{\longrightarrow}{\longleftarrow}}$$

Can one Infer Causation from mere Correlation ?

Two-point Correlation does not imply causation

←

→

Can one Infer Causation from mere Correlation ?

Two-point Correlation does not imply causation

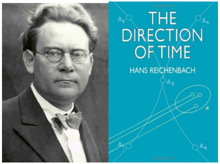
Three-or-more-point Correlation *might* imply causation

Can one Infer Causation from mere Correlation ?

Two-point Correlation does not imply causation

Three-or-more-point Correlation *might* imply causation

Common Cause Principle: $\mathbf{E_1} \leftarrow \mathbf{C} \rightarrow \mathbf{E_2}$



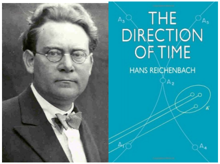
Hans Reichenbach 1891-1953

Can one Infer Causation from mere Correlation ?

Two-point Correlation does not imply causation

Three-or-more-point Correlation *might* imply causation

Common Cause Principle: $\mathbf{E_1} \rightarrow \mathbf{M} \rightarrow \mathbf{E_2}$?



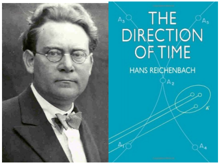
Hans Reichenbach 1891-1953

Can one Infer Causation from mere Correlation ?

Two-point Correlation does not imply causation

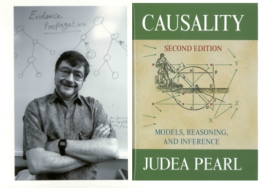
Three-or-more-point Correlation *might* imply causation

Common Cause Principle: $\mathbf{E_1} \rightarrow \mathbf{M} \rightarrow \mathbf{E_2}$?



Hans Reichenbach 1891-1953

Independent 'Cause' Principle: $\mathbf{C_1} \rightarrow \mathbf{E} \leftarrow \mathbf{C_2}$

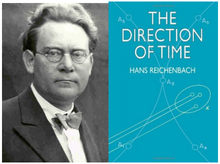


Can one Infer Causation from mere Correlation ?

Two-point Correlation does not imply causation

Three-or-more-point Correlation *might* imply causation

Common Cause Principle: $\mathbf{E_1} \to \mathbf{M} \to \mathbf{E_2}$?



Hans Reichenbach 1891-1953

But *latent variable* ?.... $\mathbf{C_1} \leftarrow L \rightarrow \mathbf{E} \leftarrow \mathbf{C_2}$