25th of April 2025

SciLifeLab AI Seminar Series

Online

Towards an interpretable deep learning model of cancer cells

Avlant Nilsson

Assistant Professor, DDLS fellow

SciLifeLab, Department of Cell and Molecular Biology

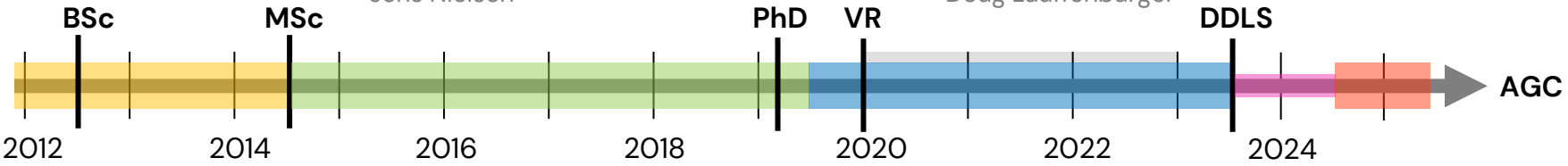
Karolinska Institutet

avlant.nilsson@ki.se

My Background

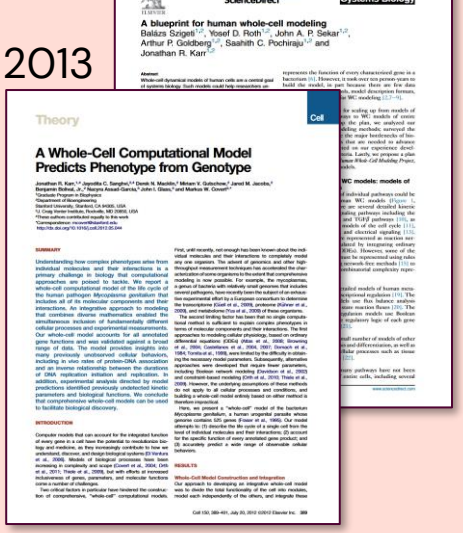
Jens Nielsen

Doug Lauffenburger



Towards AGC: an Artificial Generic Cell

2013


Theory

A Whole-Cell Computational Model Predicts Phenotype from Genotype

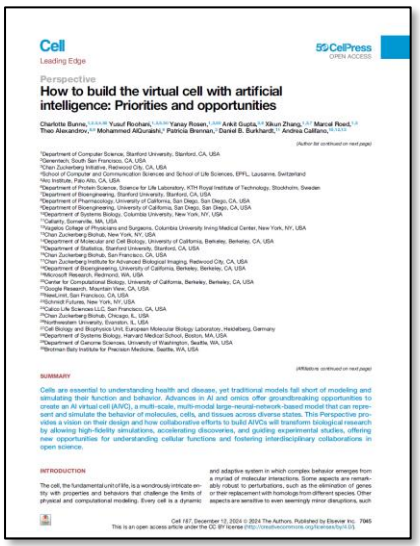
2018

ScienceDirect

Systems Biology

A blueprint for human whole-cell modeling

Dec 2024

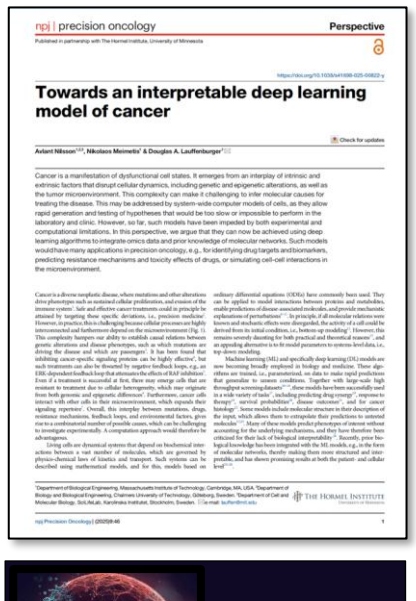

Cell

Leading Edge

Perspective

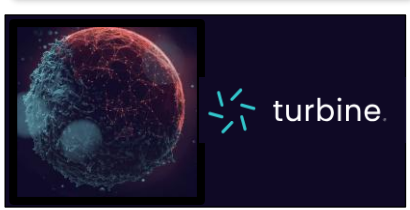
How to build the virtual cell with artificial intelligence: Priorities and opportunities

2025

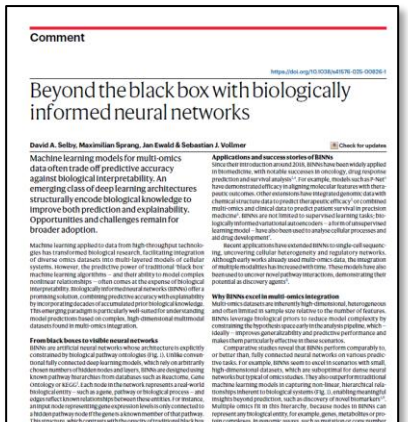

npj | precision oncology

Perspective

Towards an interpretable deep learning model of cancer



2025


Comment

Beyond the black box with biologically informed neural networks

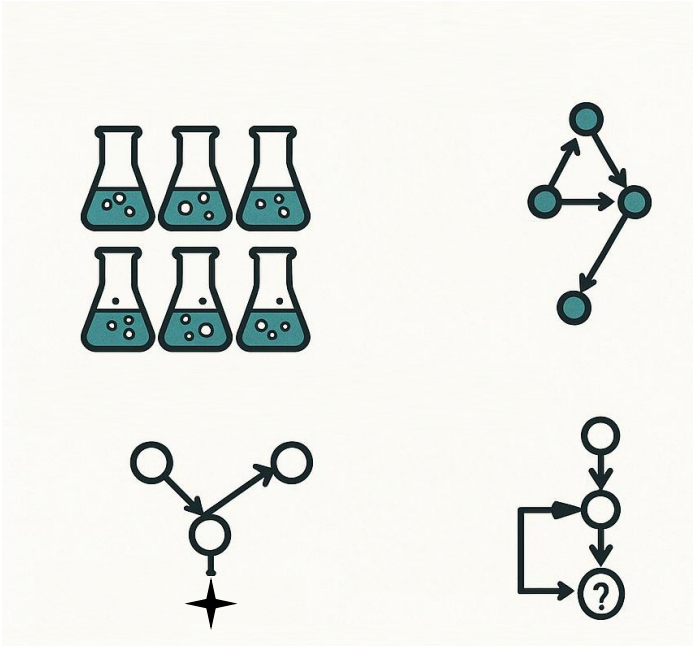
Pre-AI

AICell Lab

AlphaCell

Why AGC?

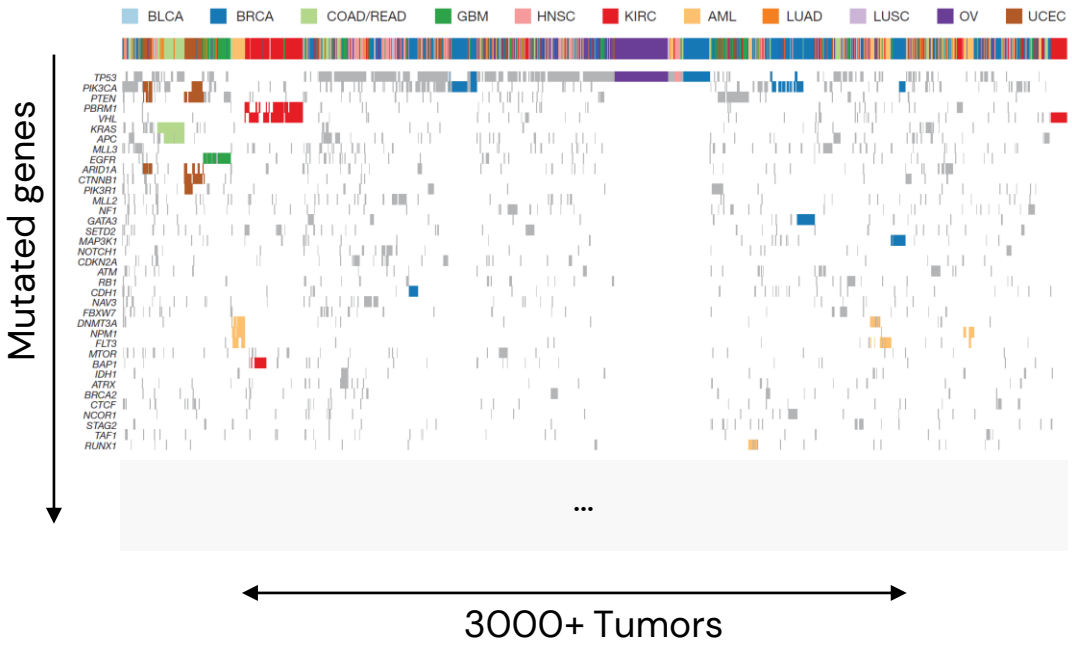
1) Fast experimentation
2) Dynamics, causal chains
3) Prediction



Chat-GPT generated artwork

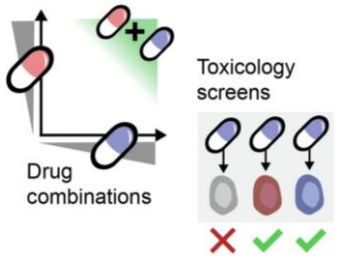
AGC = Artificial Generic Cells

Every cancer is unique



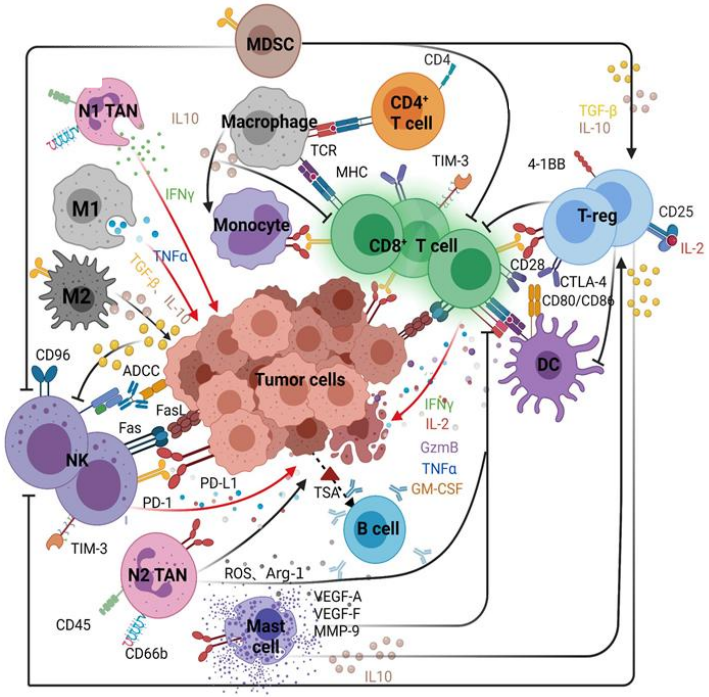
Which alterations are **driving** the cancer?

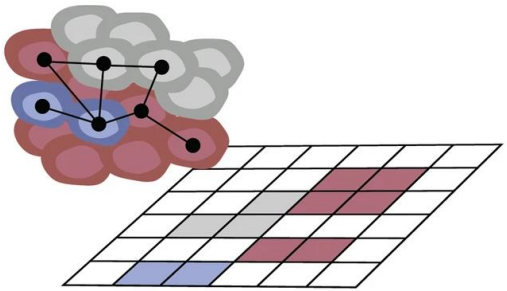
How can the cancer develop **resistance**?



Kandoth, C. et al (2013). Mutational landscape and significance across 12 major cancer types. *Nature 2013 502:7471, 502*(7471), 333–339.

Cancers interact with other cells

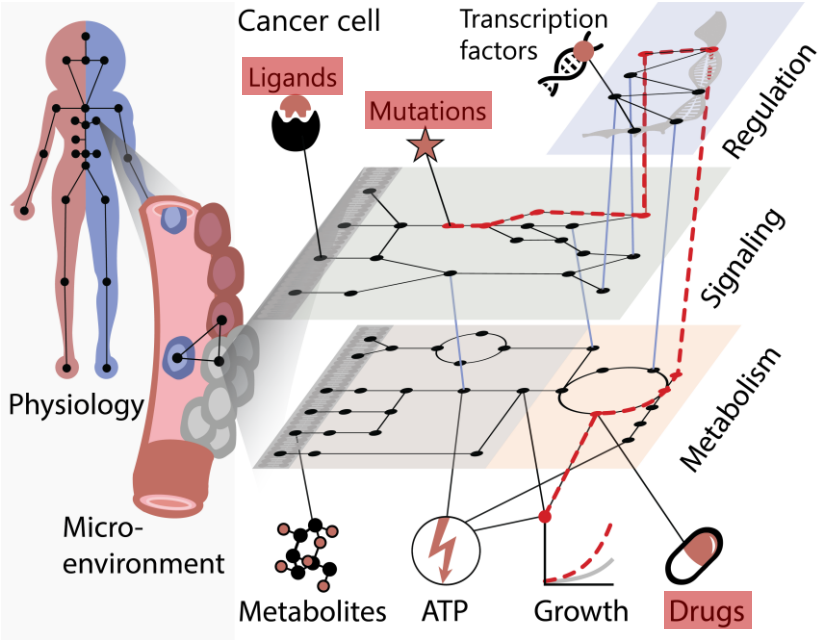




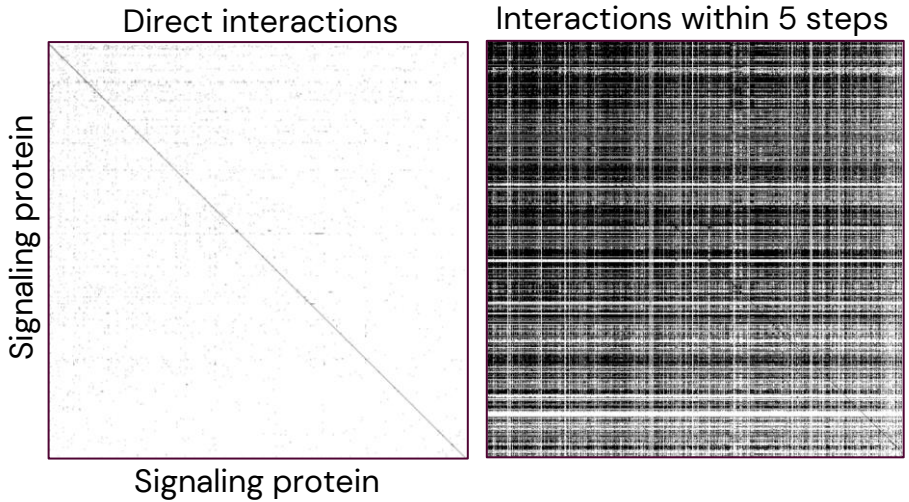
Could elucidate cell-cell interactions

Li, X. *et al.* Crosstalk Between the Tumor Microenvironment and Cancer Cells: A Promising Predictive Biomarker for Immune Checkpoint Inhibitors. *Frontiers in Cell and Developmental Biology* vol. 9 (2021).

Network-based approach



- Interpretable by construction
- Straight-forward mapping of data
- Reduced data requirements

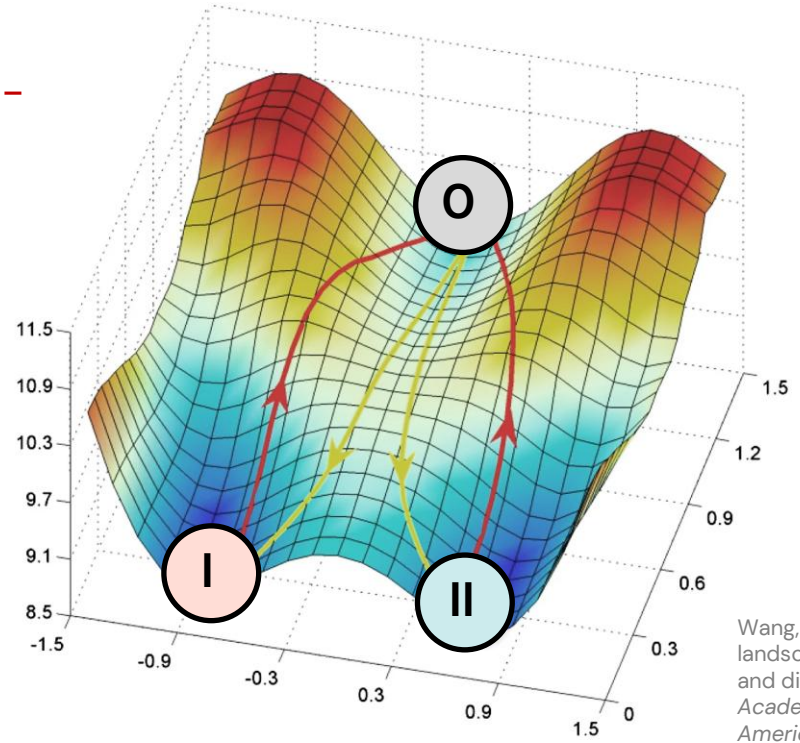


1000 protein signaling network from OmniPath

Generic model: cell types/cancers are cell states

A: each cell type/cancer – one model

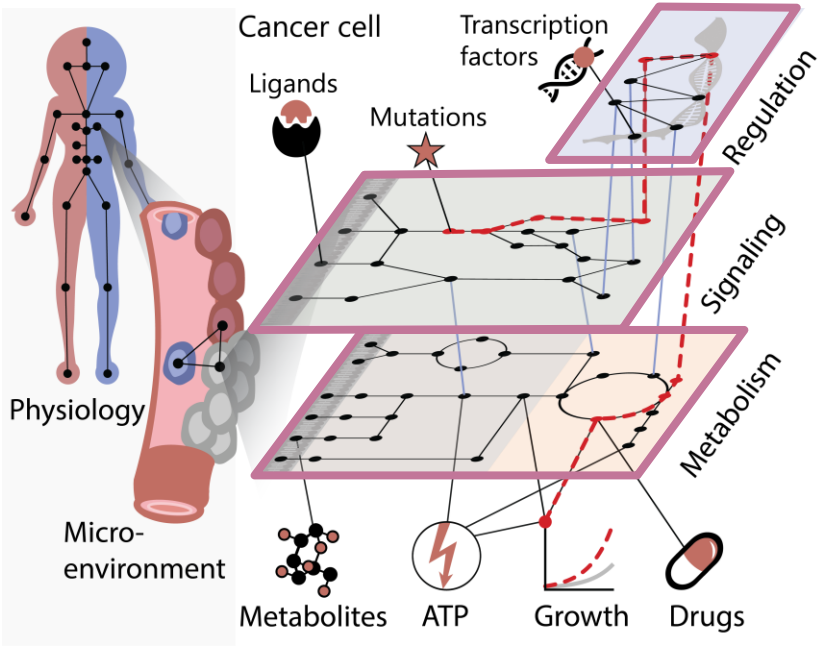
B: one model for all cells



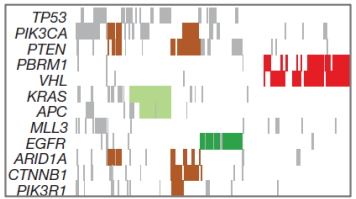
What differentiates cells is (primarily) the molecular quantitates.

Wang, J. *et al* (2011). Quantifying the Waddington landscape and biological paths for development and differentiation. *Proceedings of the National Academy of Sciences of the United States of America, 108*(20), 8257–8262.

Causes and effects transcend sub-networks



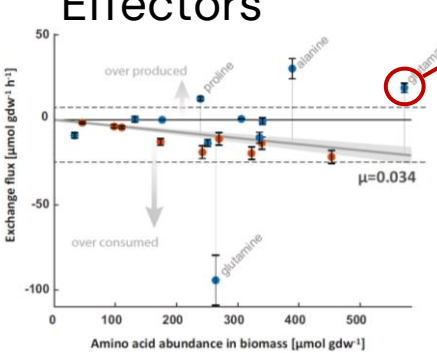
Causes

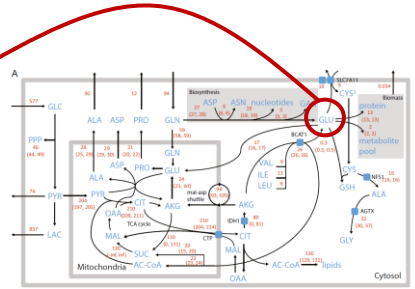


Kandoth, C. et al (2013). Mutational landscape and significance across 12 major cancer types. *Nature 2013 502:7471, 502*(7471), 333–339.

1000s of interactions

Effectors

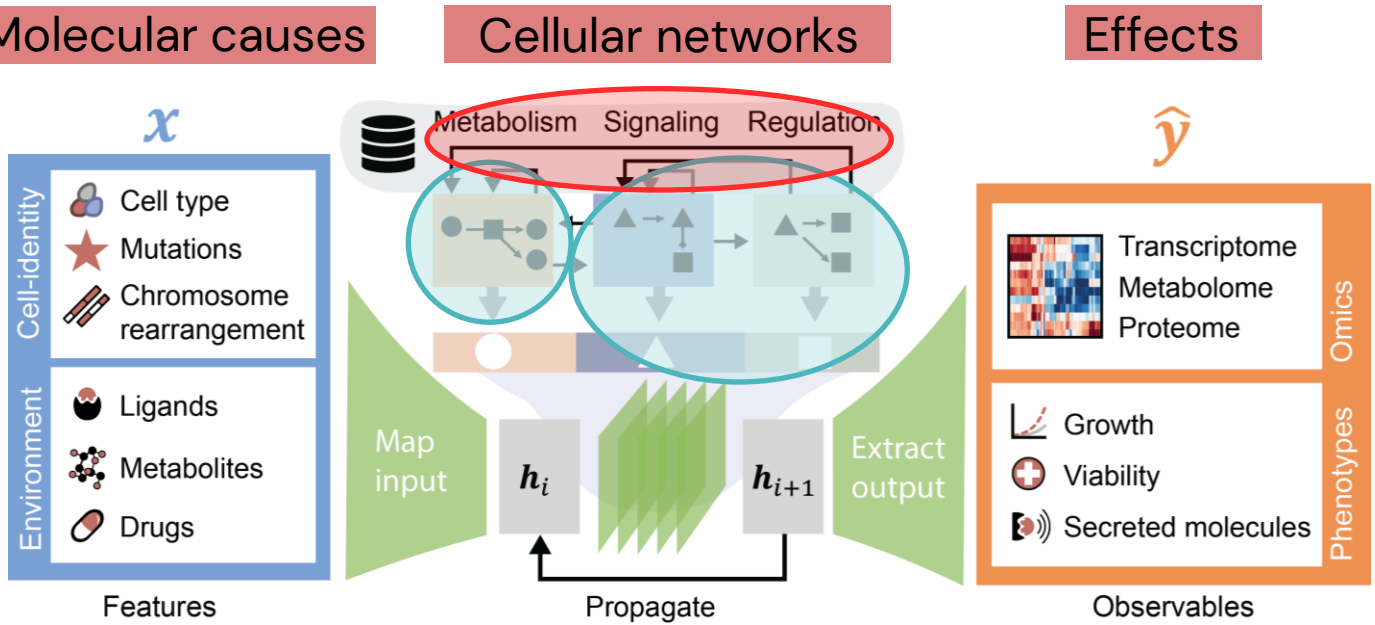




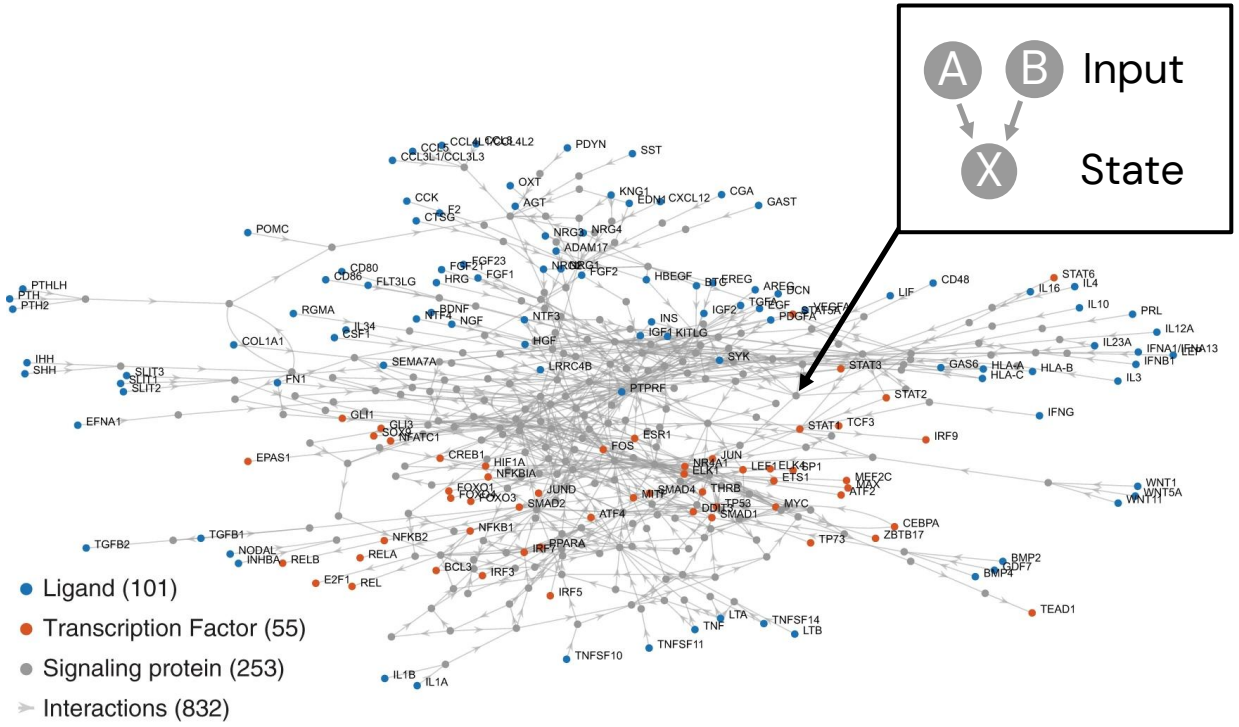
1000s of reactions

Nilsson, A. *et al.* "Quantitative Analysis of Amino Acid Metabolism in Liver Cancer Links Glutamate Excretion to Nucleotide Synthesis." *Proceedings of the National Academy of Sciences* 117, no. 19 (2020): 10294–304.

An interpretable deep learning model of the cell

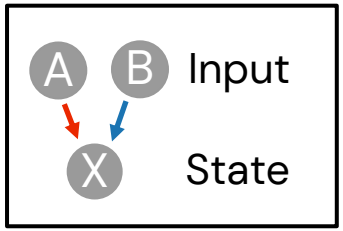


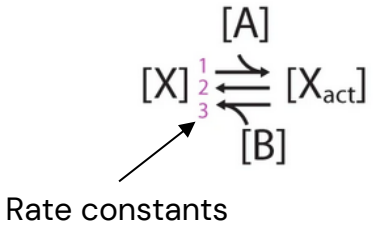
Network: effects caused by many local interactions

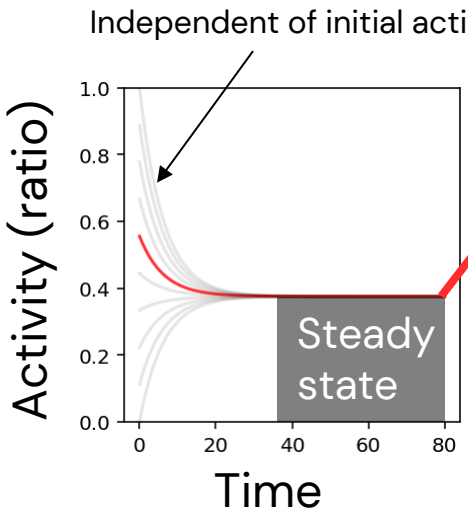


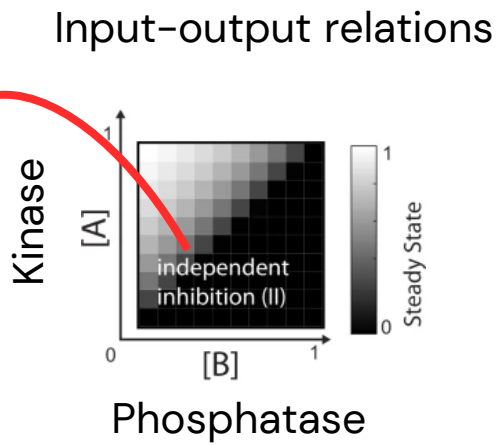
Nilsson, A, et al "Artificial Neural Networks Enable Genome-Scale Simulations of Intracellular Signaling." Nature Communications 13, no. 1 (2022): 3069.

Local steady states



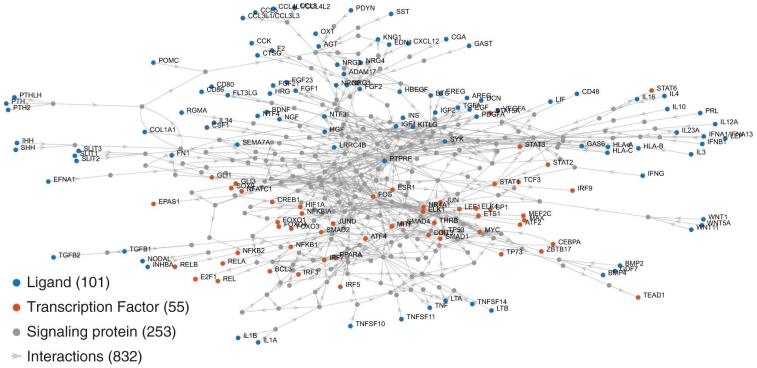




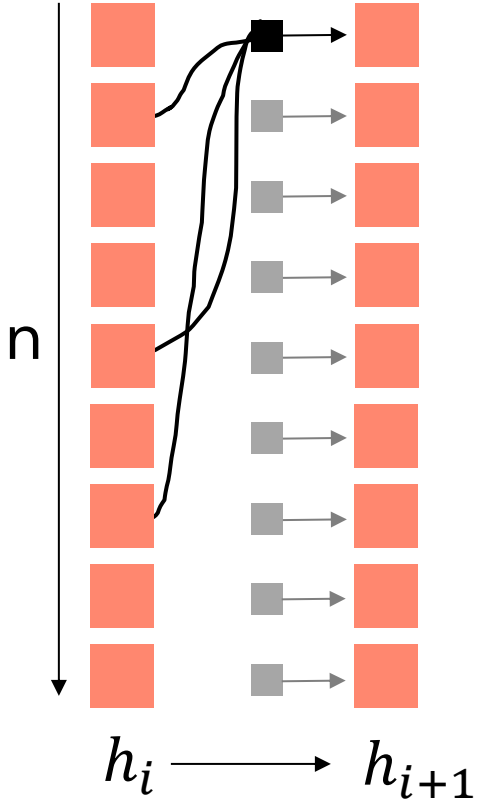


Activity of X for fixed activities of a kinase (A) and phosphatase (B)

The challenge: 1000s unique functions

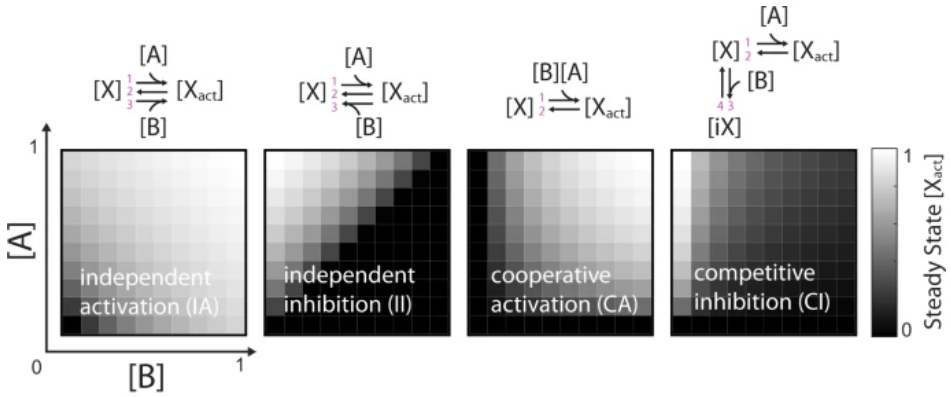


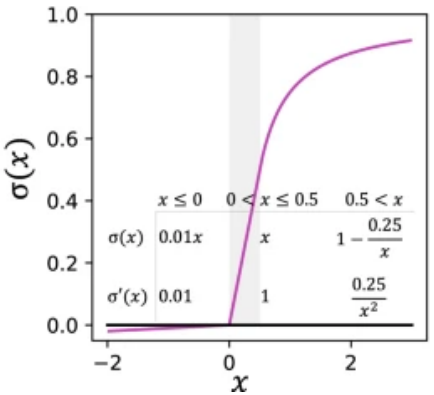
1 unique function for each signaling molecule



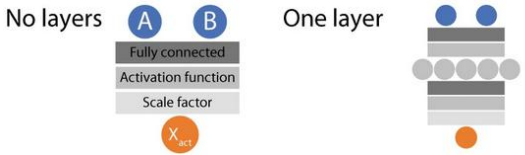
The functional form is mostly unknown

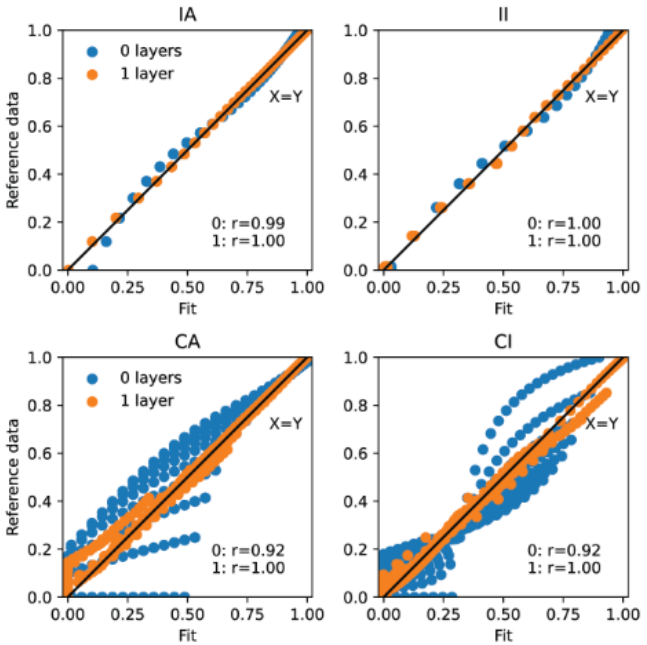
Approximation of unknown functions



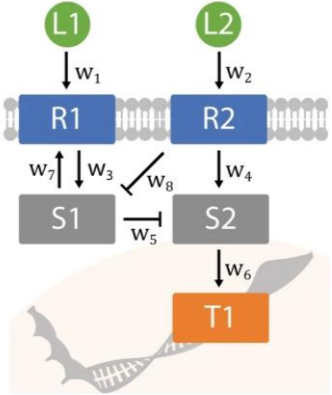


Michaelis–Menten like activation function





Biologically informed network, stable version



Network structure encoded in A, unknown learnable weights

		L1	L2	R1	R2	S1	S2	T1		
L1	L1								$bias_{L1}$	*[L1]*
L2	L2								$bias_{L2}$	*[L2]*
R1	R1	w_1				w_7			$bias_{R1}$	
R2	R2		w_2						$bias_{R2}$	
S1	S1			w_3	w_8				$bias_{S1}$	
S2	S2				w_4	w_5			$bias_{S2}$	
T1	T1						w_6		$bias_{T1}$	

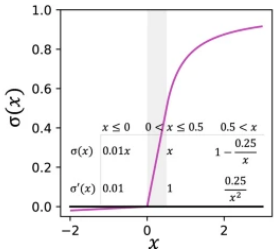
$$\mathbf{h}_n = \sigma(\mathbf{A}\mathbf{h}_{n-1} + \mathbf{b}) \qquad \mathbf{b} = \mathbf{bias} + \mathbf{x}$$

Input

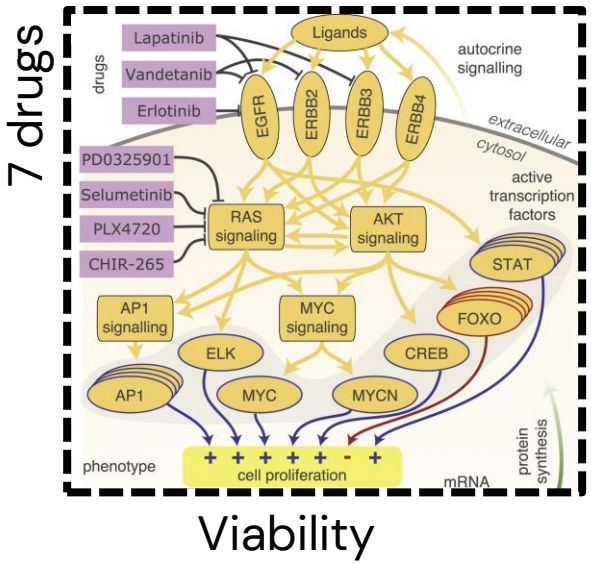
Propagate until steady state

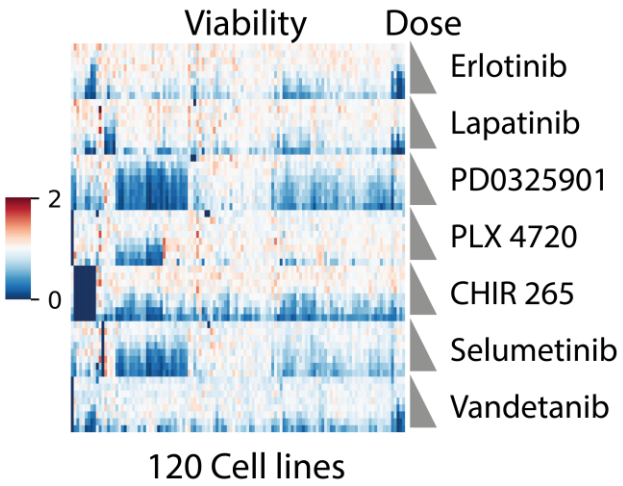
RNN
Recurrent Neural Network

The activity of each signaling node is a weighted sum of the activities of interacting nodes, which is clipped and squashed into the range 0–1

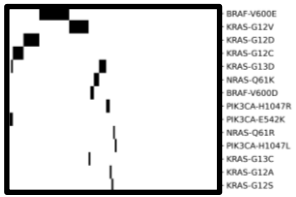


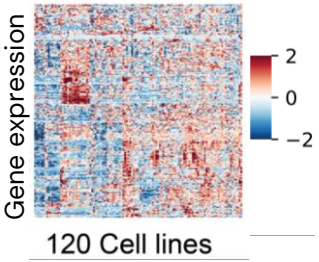
Predicting the viability effects of drugs





$X_{mutations}$

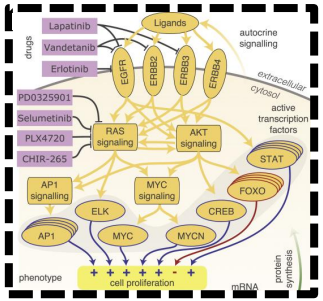




Fröhlich, F. *et al.* Efficient parameter estimation enables the prediction of drug response using a mechanistic pan-cancer pathway model. *Cell Syst.* **7**, 567-579.e6 (2018).

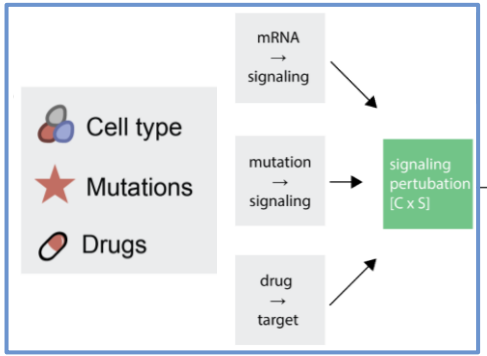
y X_{drug} $X_{expression}$

Predicting the viability effects of drugs

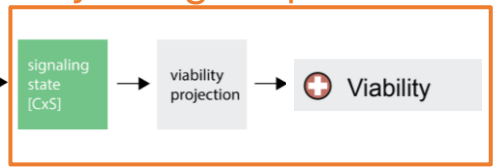


Fröhlich, F. *et al.* Efficient parameter estimation enables the prediction of drug response using a mechanistic pan-cancer pathway model. *Cell Syst.* **7**, 567-579.e6 (2018).

Mapping input



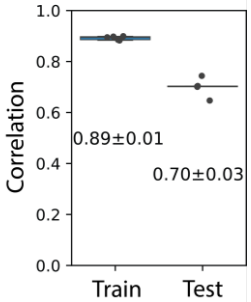
Projecting output



Nilsson, A, et al "Artificial Neural Networks Enable Genome-Scale Simulations of Intracellular Signaling." Nature Communications 13, no. 1 (2022): 3069.

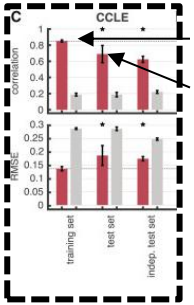
Our method

~1 hour



ODE-based method

~1 week



Train:
0.85 ± 0.01

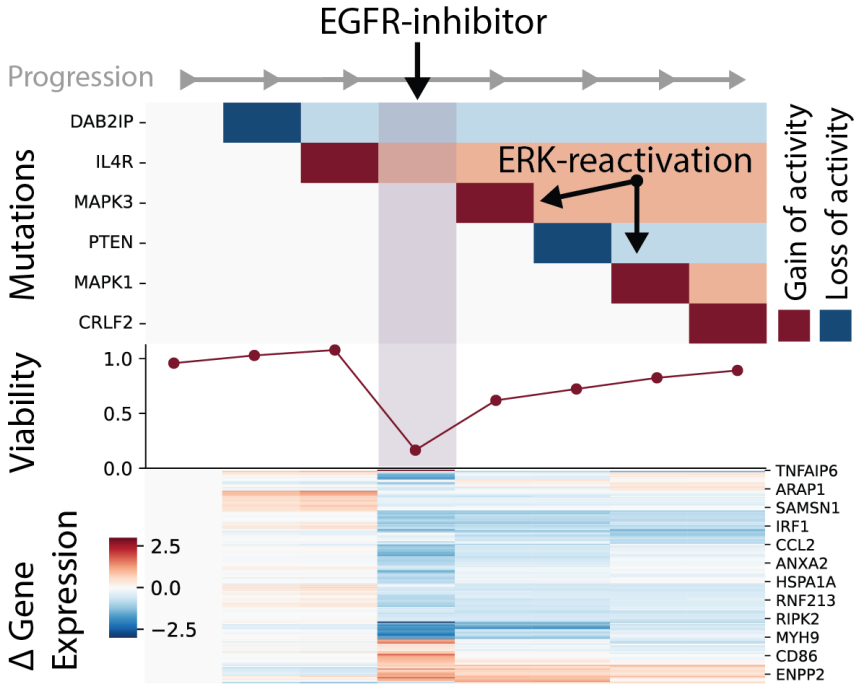
Test:
0.69 ± 0.09

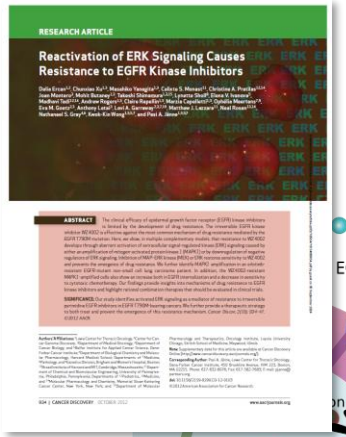
ODE, ordinary differential equation

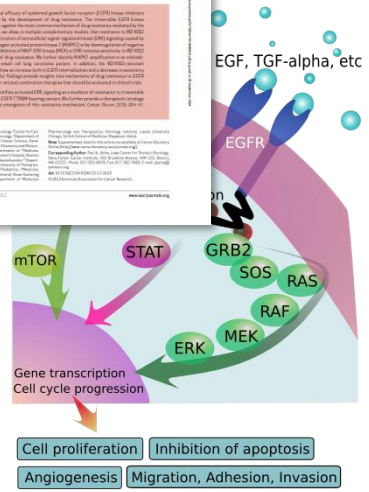
Basic simulation of cancer evolution

Unpublished results

Maximize →







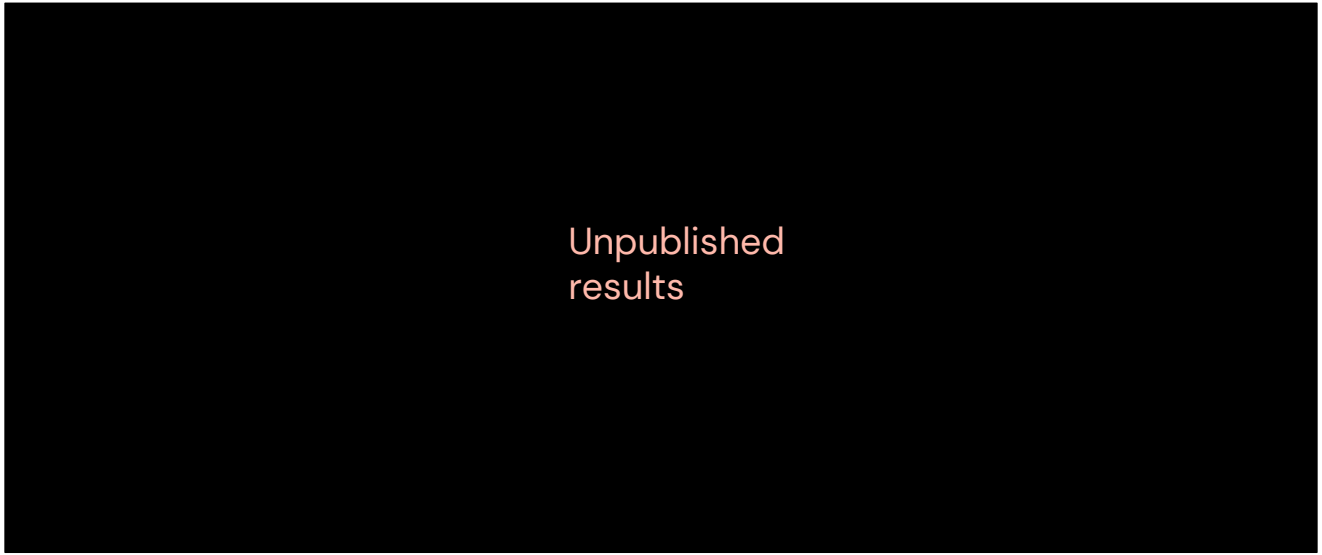
cancerindex.org/geneweb/EGFR.htm

Macrophage responses to ligand patterns

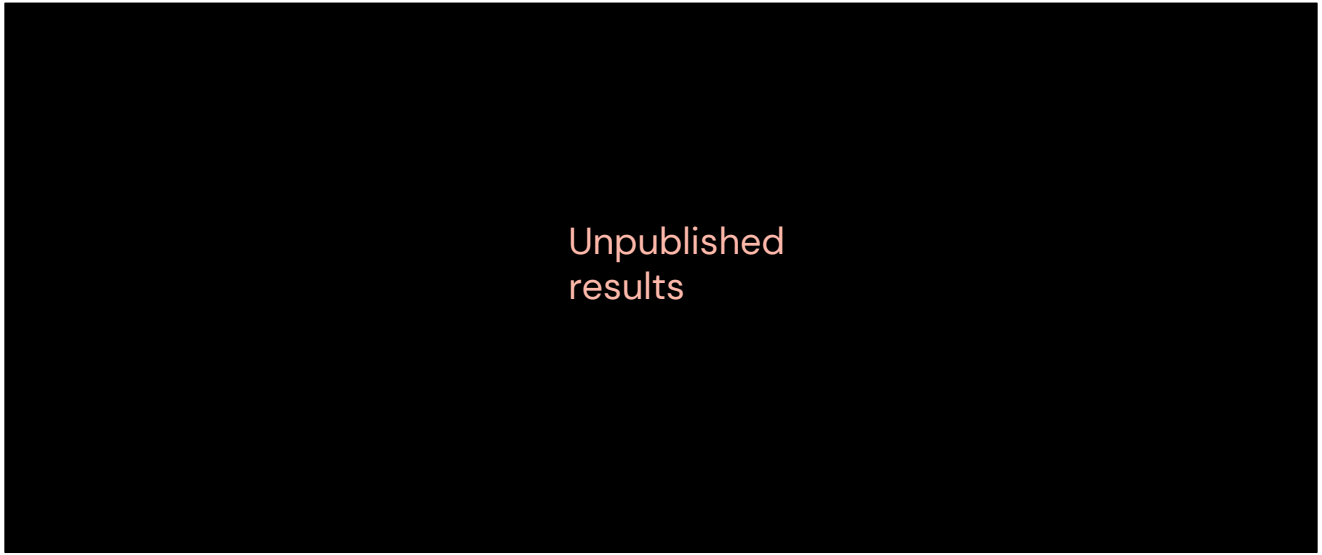
Bryan Bryson
Jacob Hochfelder
Josh Peters
Massachusetts
Institute of Technology

Unpublished
results

Sanity check: internal states

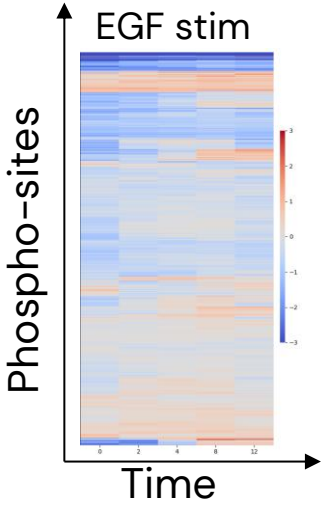


Independent validation



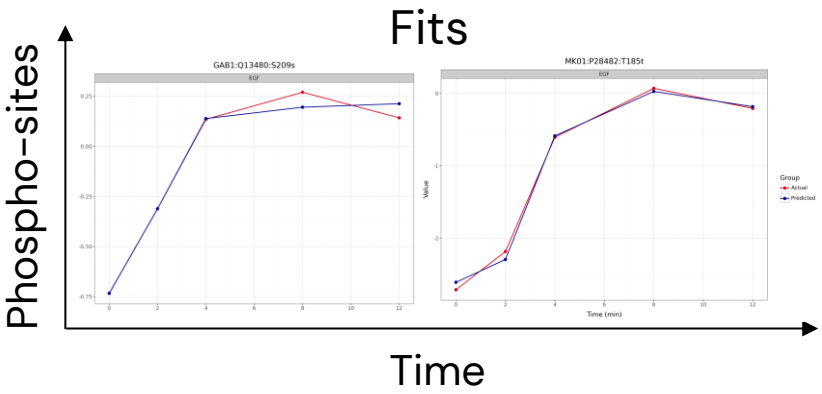
Applied to phospho-proteomics data

Unpublished results

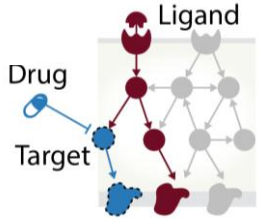


Inhibitors:
PI3K (LY294002)
SHP2 (SHP099)
RSK (SL0101)
MEK (UO126)

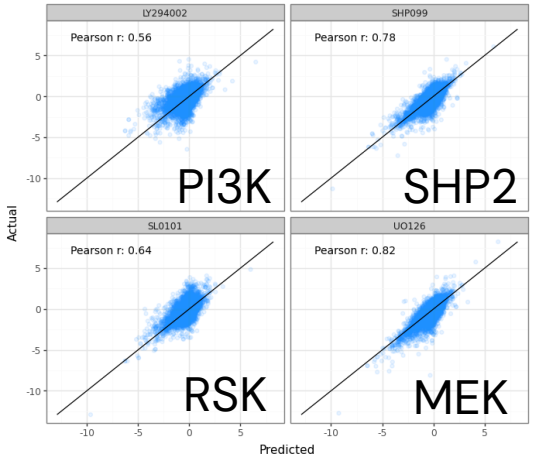
Ligand/ Drug → Signaling → Phospho-sites



0-shot learning

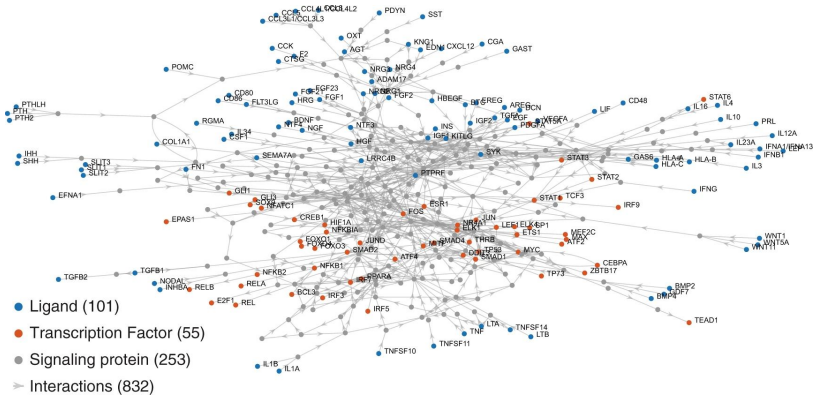


Konstantinos Antonopoulos



Feng, S., Sanford, *et al* Phosphoproteomics Data Resource for Systems-level Modeling of Kinase Signaling Networks. *bioRxiv* 2023.08.03.551714 (2023). doi:10.1101/2023.08.03.551714

How do we go deeper?

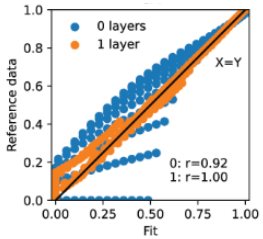


Goal:
Express 1000s of functions, with a single network

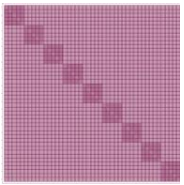

Olof Nordenstorm

Not feasible:
Loop over 1000s of unique networks

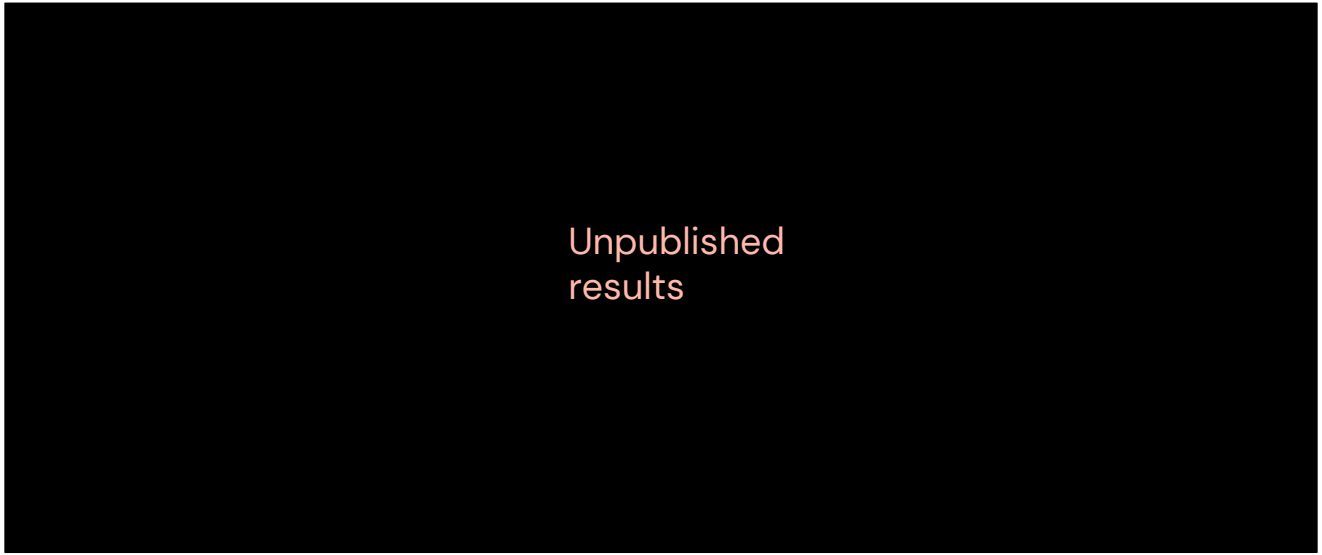
All of this was run with the weighted sum and non-linear transformation



Possible:
Sparse block matrix, but not GPU friendly

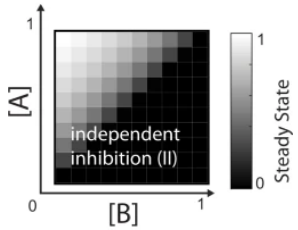


Solution: the propagator



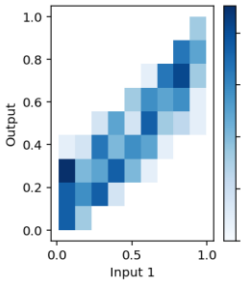
Synthetic dataset

Unpublished results

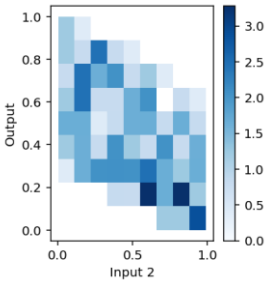


First order + several interaction terms

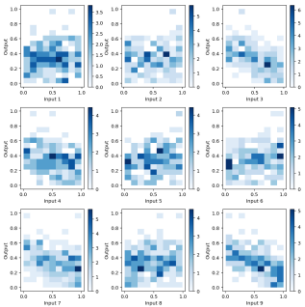
Activator

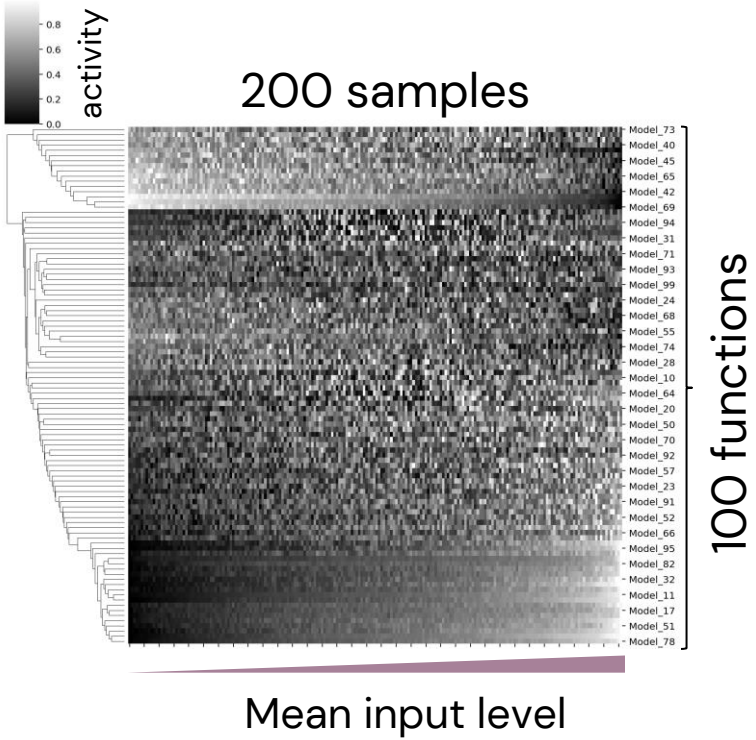


Inhibitor

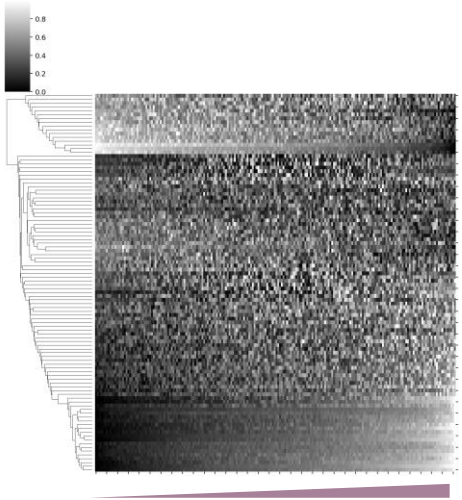


Up to 10 simultaneous inputs

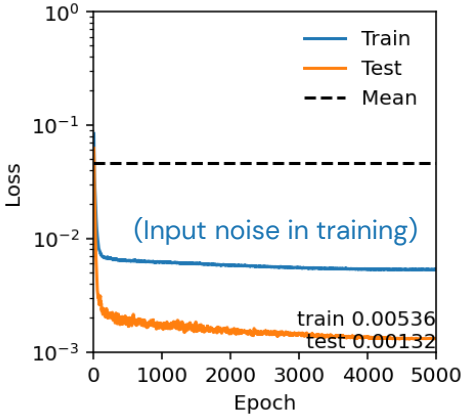


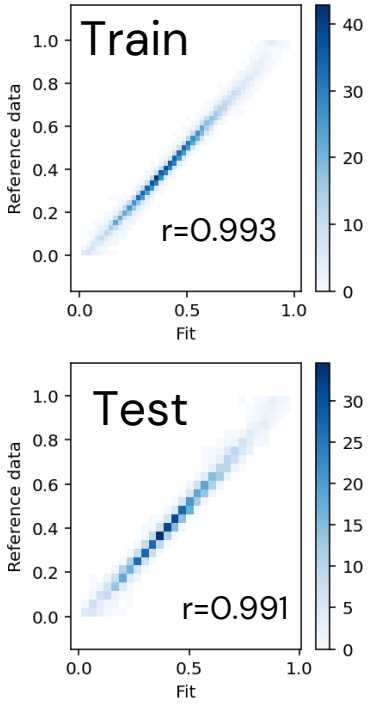


Results



Mean input level





Unpublished results

Applied to metabolism

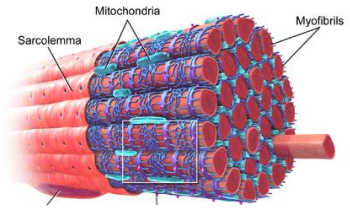
Unpublished results

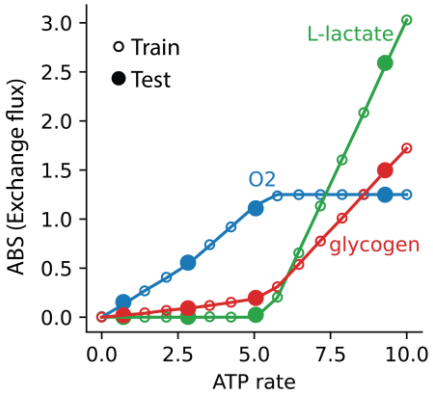
Xuechun Xu

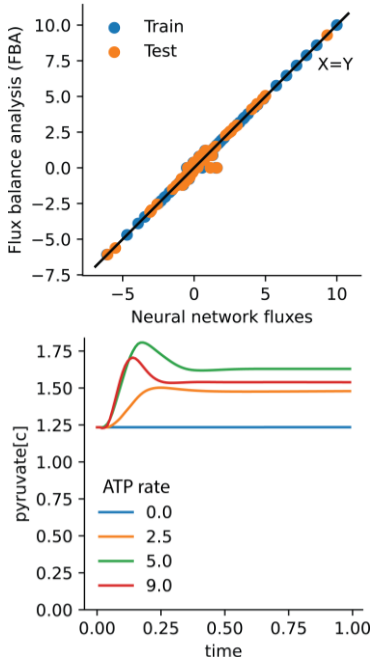
Unpublished results

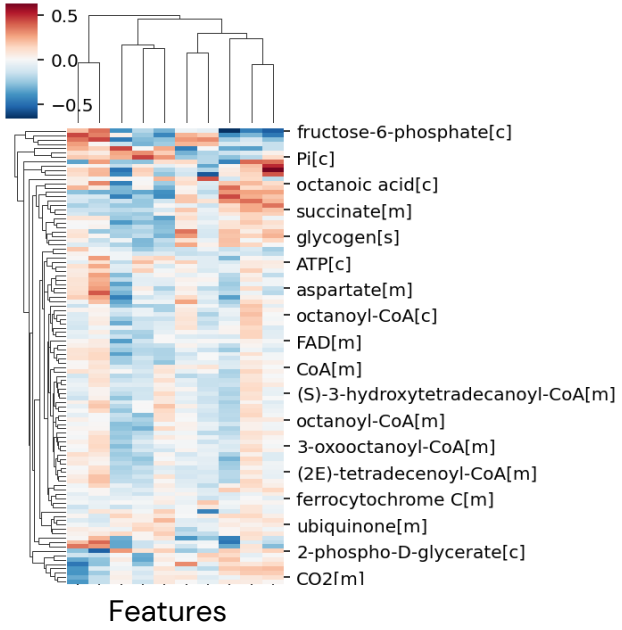
Muscle simulations

Unpublished results



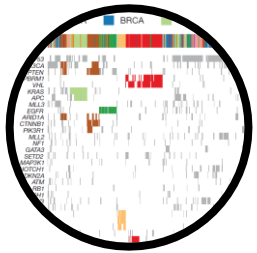






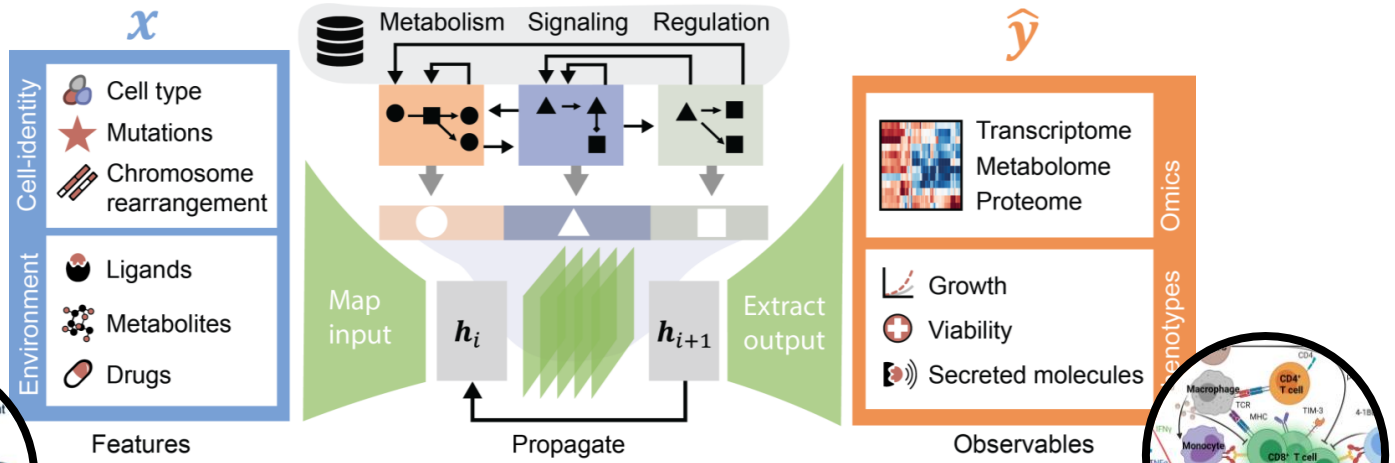
Nilsson, A., Björnson, E., Flockhart, M., Larsen, F. J. & Nielsen, J. Complex I is bypassed during high intensity exercise. *Nature Communications* **10,** 5072 (2019).

An interpretable deep learning model of the cell



1) **Predict effects of mutations**

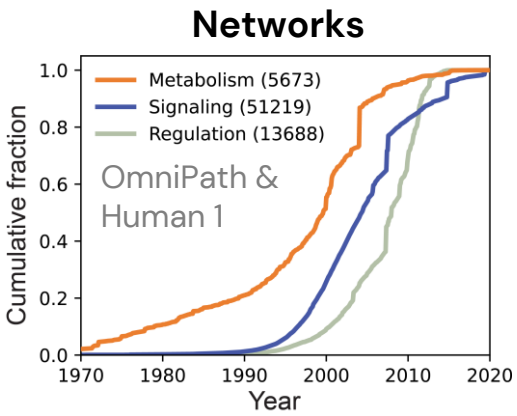
2) **Identify paths to drug resistance**

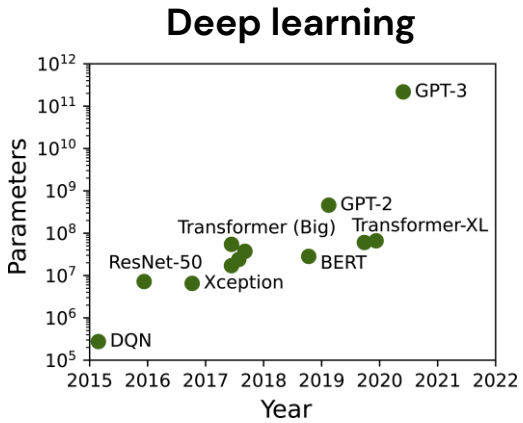


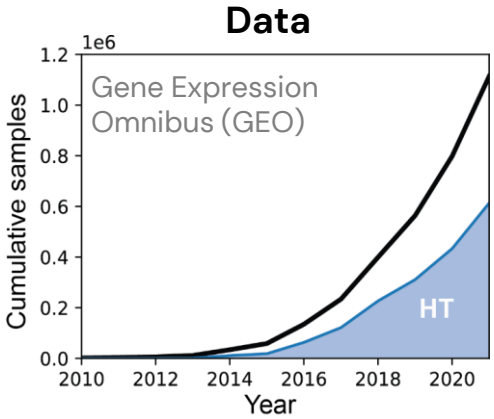
- Submodule integration
- Gene regulatory timescales
- Missing interactions (lack of knowledge)

3) **Simulate cell-cell interactions**

Now is the time for an AI model of the cell







Acknowledgments

Avlant Nilsson


Olof Nordenstorm


Xuechun Xu


Konstantinos Antonopoulos

SciLifeLab

Vetenskapsrådet
Swedish Research Council

Knut and Alice Wallenberg Foundation

CANCERFONDEN

This work was supported by the SciLifeLab & Wallenberg Data Driven Life Science Program (grant: KAW 2020.0239)

Rickard Sandberg

Kasper Karlsson

Janne Lehtiö

Thomas Helleday

Päivi Östling

Arne Lindqvist

Bryan Bryson
Josh Peters
Jacob Hochfelder

Michael Birnbaum
Caleb Perez

Claus Jorgensen

Julio Saez-Rodriguez
Attila Gabor
Pablo Rodríguez Mier
Sebastian Lobentanzer

Bas Teusink
Jurgen Haanstra

The lab is growing, talk with us about future projects!
avlant.nilsson@ki.se


KAROLINSKA INSTITUTET
ANNO 1810
Karolinska
Institutet