Incorporating patient-specific molecular data into a logic model of prostate cancer

Pauline Traynard[1], Jonas Béal[1], Luis Tobalina[2], Emmanuel Barillot[1], Julio Saez-Rodriguez[2,3], Laurence Calzone[1]

pauline.traynard@curie.fr

1. Institut Curie, PSL Research University, Mines Paris Tech, Inserm, U900, Paris, France
2. RWTH Aachen University, Faculty of Medicine, Joint Research Centre for Computational Biomedicine, MTI2 Wendlingweg 2 D-52074 Aachen, Germany
3. European Molecular Biology Laboratory, European Bioinformatics Institute, Wellcome Trust Genome Campus CB10 1SD Hinxton, United Kingdom

Introduction

The aim of this work is to build **Boolean models** to simulate specifically prostate cancer in individual patients or clones. That way, we seek to understand by which means the cells become cancerous and propose appropriate drug interventions to **personally treat** the patient.

This involves the following steps:

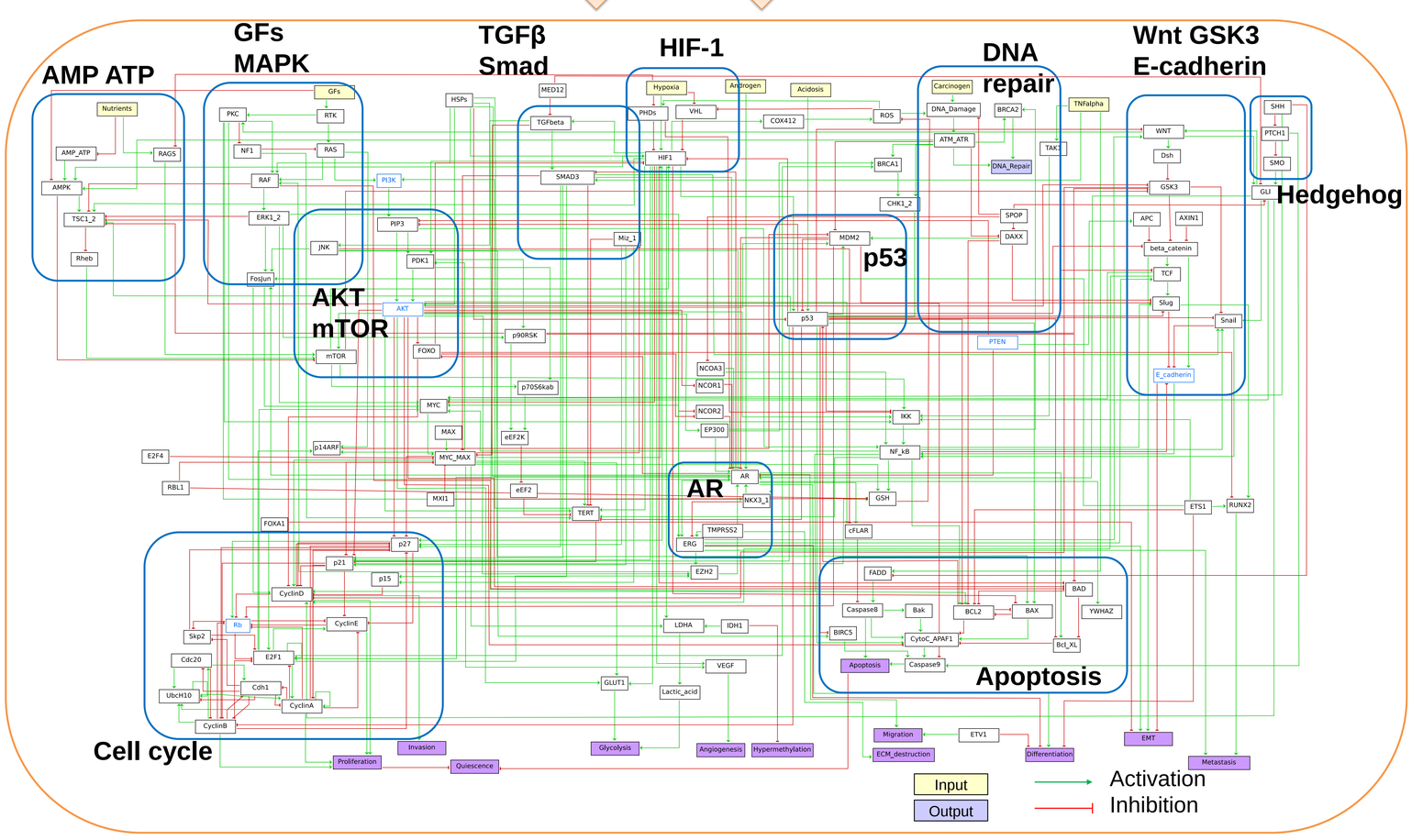
- build a **network of the signaling pathways** that include frequently altered genes in prostate cancer,
- identify **new molecular processes** to complement the network, in particular with **gene set quantification**, and **search detailed interactions** supported by literature relevant in these pathways,
- derive **a Boolean model** describing the network dynamics in specific contexts, and estimate **phenotypic probabilities** with MaBoSS,
- incorporate **multi-omics patient-specific molecular data** for personalized simulations,
- correlate the phenotypic outputs for individual patients to **clinical data**, and stratify patients.

Building a regulatory map of prostate cancer

A published model [1] is extended using both pathway databases and the literature.

Published networks and models

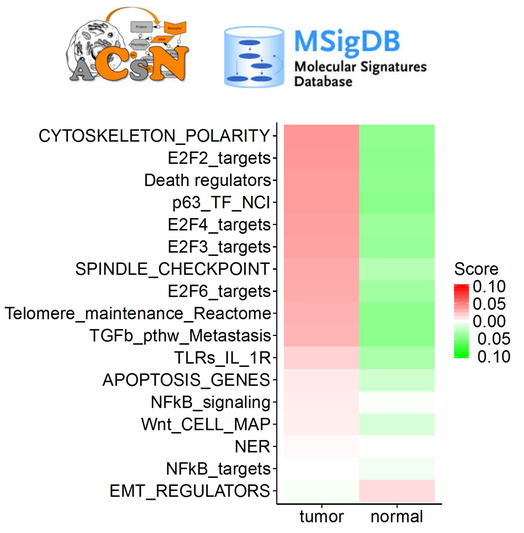
Public pathway databases, literature

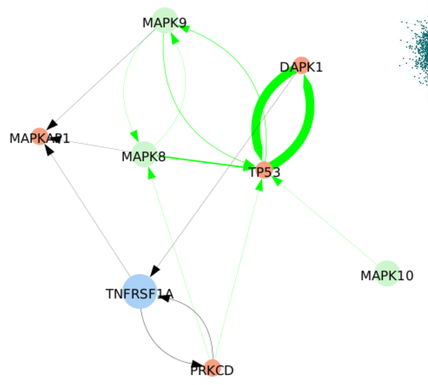


Regulatory network for prostate cancer

The model is made more specific to our cohort of patient data by applying gene set quantification methods that identify representative pathways.

ROMA [2] is a gene set quantification tool that can estimate pathway activity score and dispersion in multiple samples. Pathway activity computed with **ROMA** from expression datasets (public and private datasets) in prostate samples are used to identify significant pathways from a predefined list of gene sets and **hot spot** (top contributing) **genes.**





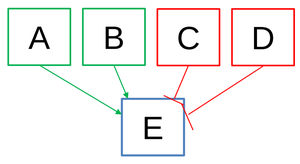
Network built with **Pypath**, representing shortest paths between the node TNFRSF1A (blue) to the nodes MAPK8, MAPK9 and MPAK10 (green) in Omnipath, with intermediary nodes colored in orange

Interactions linking novel nodes with the map are retrieved with **Pypath,** a Python module for building molecular interaction networks. It is linked to **OmniPath** [3], a comprehensive collection of literature curated human signaling pathways.

Dynamical model of prostate cancer

- Building a logic model based on the regulatory map

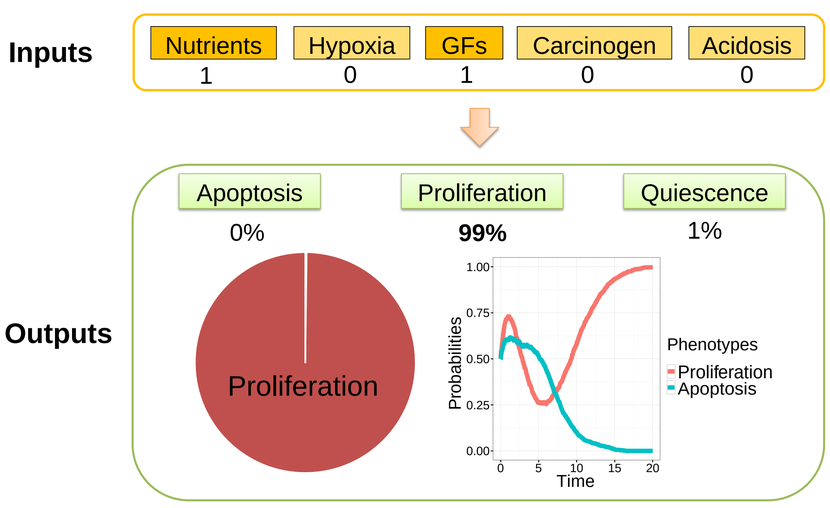
The influence network is translated into a Boolean model. Each node is a Boolean variable and is updated according to a logic rule.



Standard logic rule:
E ← (A | B) & !(C | D)

- Stochastic simulations

From a set of initial conditions, the logic model is simulated stochastically using **MaBoSS** software [4]. The solutions can be represented in different forms: a probability of network states equivalent to **the asymptotic solutions** of the stochastic simulations, or **time evolution** of the nodes of interest.



Patient molecular data integration

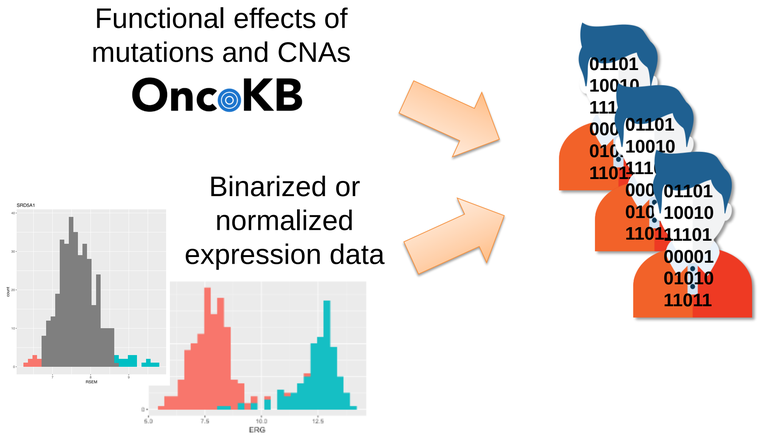
- Patient-specific model instantiation

The model provides a support to incorporate multi-omics patient-specific molecular data, encoded as **mutations** or **initial states**. This will allow to suggest optimized and **personalized strategies** for therapeutic interventions.

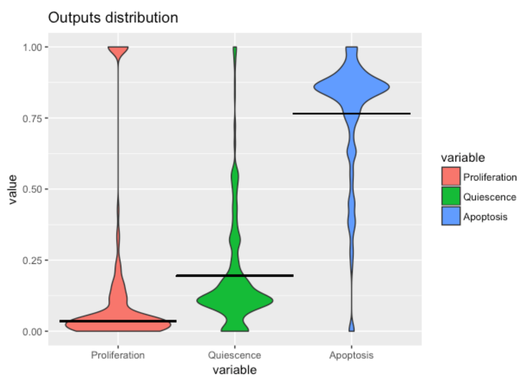
Functional effects of mutations and CNAs

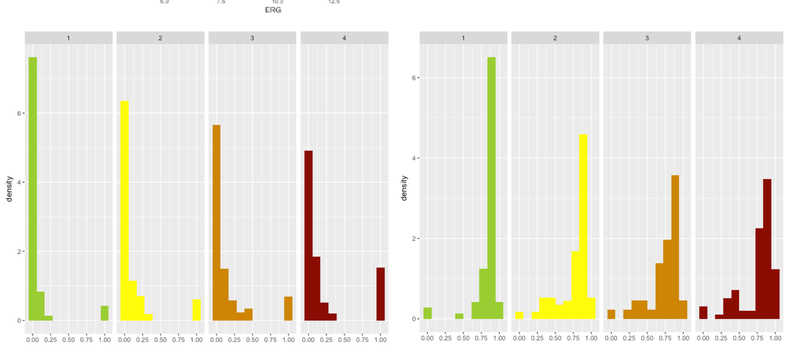
OncoKB

Binarized or normalized expression data



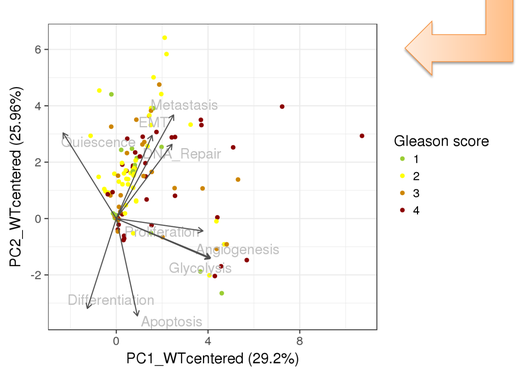
Patient-specific simulations





Proliferation & Gleason groups

Apoptosis & Gleason groups



Comparison with clinical data and patient stratification

References

[1] Fumia, H. F., & Martins, M. L. (2013). Boolean Network Model for Cancer Pathways: Predicting Carcinogenesis and Targeted Therapy Outcomes. PLoS ONE, 8(7)

[2] Martignetti, L., Calzone, L., Bonnet, E., Barillot, E., & Zinovyev, A. (2016). ROMA: Representation and Quantification of Module Activity from Target Expression Data. Frontiers in Genetics, 7, 18.

[3] Turei, D., Korcsmaros, T., Saez-Rodriguez, J.: Guidelines for literature curated signaling pathway resources *(under revision)*

[4] Stoll, G., Viara, E, Barillot, E., Calzone, L. (2012). Continuous time Boolean modeling for biological signaling: application of Gillespie algorithm. *BMC Syst Biol*, 6:116.

Acknowledgements

The project PrECISE has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 668858.
PrECISE is supported (in part) by the Swiss State Secretariat for Education, Research and Innovation (SERI) under contract number 15.0324-2. The opinions expressed and arguments employed therein do not necessarily reflect the official views of the Swiss Government.

Tools

www.pypath.omnipathdb.org
https://maboss.curie.fr
https://github.com/sysbio-curie/Roma