

Barcelona Supercomputing Center Centro Nacional de Supercomputación



Personalised Medicine Centre of Excellence: how can HPC celllevel simulations help us fight against cancer and COVID

> Arnau Montagud Computational Biology group Life Science Department

7 October 2019

#### Central Dogma and the omics data



#### To infer mechanisms from omics data we need networks and models





Barcelona Supercomputing Center Centro Nacional de Supercomputación

### CORE APPLICATIONS

PerMedCoE optimises key software for cell-level simulations to the new preexascale platforms

The PerMedCoE four core applications are:

- COBRA for the simulation of cellular metabolism at genome-scale
- CellNOpt for modelling signal transduction
   networks
- MaBoSS for stochastic simulations of Boolean models
- PhysiCell an agent-based modelling framework for simulating cell-cell interactions





#### Data leads to Structure ... and Structure leads to Modelling



Le Novere, Nat Rev Genet. 2015 Mar;16(3):146-58.

# **Metabolic modelling**



Barcelona Supercomputing Center Centro Nacional de Supercomputación

#### Omics data are used to build metabolic networks and metabolic models

- ( Network reconstruction
  - From the list of genes in the genome
- ( Flux balance analysis
  - Provides information on reactions' fluxes
  - Used to study cell growth and metabolite productivity
- ( Constraint-based metabolic model



$$\frac{dX_i}{dt} = S_{ij} \cdot v_j, \qquad \forall i \in M, \forall j \in N$$



Stoichiometric matrix

n > m

 $\sum_{i=1}\sum_{j=1}S_{ij}\cdot v_j=0$ 

Time (cell growth) >> Time (metabolic reactions)

2. Impose constraints

1.

Consider steady state

Stephanopoulos et al. (1998). San Diego: Academic Press

#### Flux Balance Analysis explained graphically



COBRA

### Alternatives to modelling cellular metabolism

**Kinetic modelling** 

- **Complete description**
- Solution is a unique point •

EXCELENCIA SEVERO OCHOA



**Constraint-based modelling** 

**COBRA** 

- **Incomplete Information**
- Solution space



Slide from M. Ponce de L, BSC

#### Context-specific metabolic modelling







Slide from M. Ponce de L, BSC

### Constraint-based modelling applications in biomedicine

- ( Production of by-products
  - Antibiotics
  - Flavoured synthetic milk
- ( Data integration
- (Cell maximal growth rate
- ( Single gene knockout
- ( Double gene knockout
  - Synthetic lethality
- Essential media components
  - Amino acids
  - Vitamins
- ( Biomarkers
- ( Combinatorial therapy
  - Diet restriction + drugs that target a given signalling pathway







#### If you want more information





**cobrapy** is a python package that provides a simple interface to metabolic constraint-based reconstruction and analysis.

```
>>> import cobra
>>> model = cobra.io.read_sbml_model('Ec_core_flux1.xml')
>>> model.metabolites[:3]
[<Metabolite 13dpg_c at 0x112b2d160>,
    <Metabolite 2pg_c at 0x1024eb048>,
    <Metabolite 3pg_c at 0x112b2d748>]
```

#### https://opencobra.github.io/cobrapy/



- **(Constraint-Based Reconstruction and Analysis** 
  - Matlab, Python, Julia
- ( Practical course "Constraint-based modelling: a primer on basics and examples"
  - <u>https://github.com/ArnauMontagud/cobra\_cbm\_tut</u>
     <u>orial</u>
  - Jupyter notebook
  - Adapted from Miguel de Ponce León

# **Logical modelling**



Barcelona Supercomputing Center Centro Nacional de Supercomputación



- ( The questions are qualitative
- ( The data are discrete (mutations, copy number, etc.)
- **(C)** Expression data are not absolute values
- ( No information over time
- ( No details about the precise biochemical reactions



### Translation of an influence network into Boolean logic



A = !B & CB = AC = input

#### **Solutions**

Barcelona



Each variable can take two states: 0 or 1 

#### Boolean logic:

- Connectors: AND (&), OR (|), NOT (!), XOR (/)
- Logic depends on incoming arrows
- **((** Set of discrete variables as abstractions of activity level linked by logical rules as signed interactions
- Attractors are subgraphs of the state transition graph with no outgoing arrows: can be stable states & cyclic attractors
- Stable states = cell fates = phenotypes
- Updating dynamics can be
  - synchronous: all variables that can be updated are updated
  - asynchronous: only one variable is updated at a time

#### Two examples of models: TNF response and cell fates and Metastasis model



#### What insights can we get from the mathematical model



Letort et al., Bioinformatics, 2018, bty766



- ( What are the solutions of the model that can be interpreted biologically?
- **((** What are the **important nodes** of the network?
- ( How robust/sensitive is the model?
- (Can we predict genetic interactions (epistasis, synthetic lethality) from the model?
  - Knock outs & overexpression
- (Can we simplify/reduce the model to highlight the most important processes?

#### TNF response and cell fates



Letort et al., *Bioinformatics*, 2018, bty766

Barcelona Supercomputing Center Centro Nacional de Supercomputación Calzone et al., PLoS Comput Biol, 2010, 6(3): e1000702

### Analysing stable states probabilistically

- **(Continuous time Markov process** on the Boolean transition state space
  - Each Boolean state has an associated probability
  - Rate of change associated to each transition
    - rate up and rate down
  - Stochasticity, time, probabilities, ...
- ( Perturbations can be studied in a probabilistic manner
  - Transient effects, such as knock downs
  - Dosage experiments
- ( More information:
  - Stoll et al., BMC Syst Biol, 2012, DOI: 10.1186/1752-0509-6-116
  - Stoll et al., *Bioinformatics*, 2017, DOI: 10.1093/bioinformatics/btx123



MaBoSS

Boolean state transition graph









### TNF response and cell fates' probabilities





### Modelling patient-specific Boolean networks

Béal et al., Frontiers in Physiology, 2019

#### Data types:

RNA

Prot.

- Copy Number Alterations
- Mutations
- Expression: RNA and/or proteins

#### Modelling framework's variables:

- Node states = mutants
- Initial conditions = growth media conditions or experimental setup
- Transitions rates = Gene's ability to activate or deactivate



MaBoSS

#### Boolean state transition graph



# Phenotypes correlate with clinical data: METABRIC models vs PAM50 and survival data



MaBoSS

#### MaBoSS

### **Cell-line-specific Boolean models**

Supercomputing

tro Nacional de Supercomputación

Center



WT is non-personalised model

BPH-1 is a benign prostate

hyperplasia

#### If you want more information

#### Logical modelling pipeline

- Extensive and comprehensive studies can be done with a validated model and tools on the field
- Experimentally friendly hypotheses can come out of these studies
- ( Available at GitHub

#### Flexible pipeline of methods

- Generates data-tailored models
  - Cell lines
  - Patients
- Correlates with clinical data
  - Highly dependent on available clinical data
- ( Available at GitHub

#### https://github.com/sysbio-curie/Logical\_modelling\_pipeline



Briefings in Bioinformatics, 2017, 1–12 doi: 10.1093/bib/bbx163 Paper

#### Conceptual and computational framework for logical modelling of biological networks deregulated in diseases

Arnau Montagud, Pauline Traynard, Loredana Martignetti, Eric Bonnet, Emmanuel Barillot, Andrei Zinovyev and Laurence Calzone

#### https://github.com/sysbio-curie/PROFILE

#### Personalization of Logical Models With Multi-Omics Data Allows Clinical Stratification of Patients

Jonas Béal, Arnau Montagud, Pauline Traynard, Emmanuel Barillot\* and Laurence Calzone\*

Institut Curie, PSL Research University, Mines Paris Tech, Inserm, U900, Paris, France



### If you want more information – CellNOpt

### (CellNOpt

- ( Identification of signaling models based on perturbation data and prior knowledge
- ( <u>https://saezlab.github.io/CellNOptR/</u>
- ( <u>https://github.com/saezlab/cellnopt</u>



An illustration of how we use our logic modeling method CellNOpt to better understand deregulation of signal transduction in disease. Left: simple pathway model; right: experimental data and match between model simulations and data.



# **Agent-based as multiscale modelling**



Barcelona Supercomputing Center Centro Nacional de Supercomputación

### Agent-based modelling

- ( Agent-based models are composed of:
- 1. numerous agents;
- 2. decision-making heuristics;
- 3. an interaction topology; and
- 4. a description of the environment.

### ( Examples:

- Ecology
- Environmental Science
- Artificial Intelligence
- Tissue Biology





**PhysiCell** 

#### Ch'ng. IGI Global: Hershey, PA, 2009



BSC Barcelona Supercomputing Center Centro Nacional de Supercomputación

Hoehme et al, PNAS, 2010

#### Agent-based model for multicellular modelling

Different ABM approaches for multicellular modelling



"... An **agent-based model** is a class of computational models for simulating the actions and interactions of autonomous agents (...) It combines elements of game theory, complex systems, emergence, (...). Monte Carlo methods are used to introduce randomness..." (\*\*)



Ê (\*

(\*) Modified from Osborne, J.M. et al. (2017). PLOS Comp Bio.

(\*\*) Source: https://en.wikipedia.org/wiki/Agent-based\_model

Slide from M. Ponce de L, BSC

### Multi-scale modeling framework: PhysiCell

An open source physics-based cell simulator for 3-D multicellular systems





#### The basic cell agent has properties

Cell Cycle Phase -Premitotic -Postmitotic -Kió7 negative -Apoptotic -Necrotic -Necrotic (swelling) -Necrotic (lysis)



# Cell agent properties

- Cell Volume
  - nucleus
  - o cytoplasm
- **Position** (x, y, z)
  - Neighborhood
  - Environment
- Cell internal state
  - Cell cycle phase (*G*<sub>0</sub>, *M*, *etc*)
  - Growth rate
  - Custom phenotype



#### The simulation domain (grid/lattice)



#### Cells can have mechanics and entities can diffuse

80

60

40

20

Υ

-20

-40

-60

-80

-100

EXCELENCIA SEVERO

Cell Cycle Phase

-Postmitotic

-Apoptotic

-Necrotic

-Ki67 negative

-Necrotic (swelling)

Necrotic (lysis)



100

100

Slide from M. Ponce de L, BSC

**PhysiCell** 

Physicell

#### Environment can be dynamic and reactive









#### Surrounding physical environment

Surrogate for extra-celular matrix

- Field with densitities that can be produced & consumed
- Inert agents that can be moved



**PhysiCell** 

Pnysicell

#### Simulation workflow



#### Simulation's main loop

while t\_current < tend update\_difussion() if  $\Delta t \% \Delta tmech == 0$  update\_cell\_mechanics() if  $\Delta t \% \Delta tcell == 0$  update\_cell\_processes()  $\Delta t = 0$   $\Delta t += t_step$ t\_current += t\_step

#### Time scales

- $\Delta t_{\rm diff}$ : (diffusion/transport): 0.01 min
- $\Delta t_{\mathrm{mech}}$ : (cell movement): 0.1 min
- $\Delta t_{cell}$ : (cell processes): 6 min



Slide from M. Ponce de L, BSC

### Modifying agent-based to have genotype-to-phenotype modelling



### Merging MaBoSS with PhysiCell: PhysiBoSS

Multi-scale ABM framework integrating physical dimension and cell signalling

#### MaBoSS (Barillot team at Institut Curie)

• Boolean model stochastic simulator for cell signalling

#### PhysiCell (Macklin team at Indiana University)

• Agent-based framework for simulating multicellular systems

#### PhysiBoSS (Barillot team at Institut Curie)

- A PhysiCell extension to include cellular signalling
- This allows to perform combined studies of:
  - Environmental perturbation
  - Genetic perturbation





### Cells have different phenotypes depending on their genes' activation

PhysiCell

FASLG

FADD



-Necrotic

-Necrotic (swelling) -Necrotic (lysis)

**Time scales** 

- 
$$\Delta t_{\rm diff} / \Delta t_{\rm mech} / \Delta t_{\rm cell}$$







### The framework allows for genetic and environmental heterogeneities

- **((** Agents can represent different cell strains
  - With different biology
    - WT and mutants
    - Patient-specific networks
  - With different physical properties
    - Cell-cell adhesion
    - Cell-matrix adhesion



#### ( Different tumour architectures





One-cell-thick monolayer

#### The framework allows for finding optimal drug regimes



Letort et al., Bioinformatics, 2018, bty766

#### PhysiBoSS experiment: TNF pulse studies





~48 h simulation time, 30 min wall time ~2500 cells

TNF pulses every 150 min



Slide from M. Ponce de L, BSC

# Examples of use of PhysiBoSS













### Ongoing work: Migrating PhysiBoSS to use MPI

### (Currently:

- Intra-node shared memory
- OpenMP is possible: 1 node
- max 48 cores in MN4
- 1B cells in several simulations



### ( Ideal:

- Inter-node shared memory
- Combine OpenMP + MPI
- 1B cells in one single simulation
- Lead by Gaurav Saxena, BSC
- Goals:
  - Expand the scope of the simulations by several orders of magnitude
  - Enable the simulation of complex behaviours

### Domain decomposition: 2D example

Full Domain (
$$\frac{\partial \rho}{\partial t} = D\nabla^2 \rho - \lambda \rho \dots$$
)  
unit of the second second

**Individual Voxel:** stores the values of each molecule concentration. Connected to other voxels through Moore neighborhood (PDE solver) **Ghost (Halo) Cells:** needed to update boundary voxels in a transparent way. Needed to exchange information between neighbour voxels

#### Domain decomposition: 2D example



#### Ongoing work: Migrating PhysiBoSS to use MPI



tro Nacional de Supercomputaciór



**PhysiCell** 

- Collaboration with Gaurav Saxena & David Vicente, HLST, BSC
  - Refactoring parts of the code to implement MPI
  - Environmental data has to be shared among different nodes

### Ongoing: to have different kind of cells in the microenvironment

#### (Cells could be:

- Stromal cells
- Immune cells
- Cancer-associated cells
- We want to model cancerimmune system interaction

EXCELENC SEVERO



Barcelona

Center

Supercomputing

ro Nacional de Supercomputaciór



PhysiCell

Arnau Montagud, Computational Biology group

#### PhysiCell

#### Ongoing: Connecting metabolic models to PhysiCell



- ( Open source code for multi-scale modelling
  - Available at GitHub
- Uses agent-based modelling for physical phenomena
- ( PhysiBoSS uses Boolean modelling for biological phenomena
  - Adapted from MaBoSS

#### https://github.com/MathCancer/PhysiCell

http://physicell.org/tutorials/

RESEARCH ARTICLE

# PhysiCell: An open source physics-based cell simulator for 3-D multicellular systems

Ahmadreza Ghaffarizadeh<sup>1</sup>, Randy Heiland<sup>2</sup>, Samuel H. Friedman<sup>1,3</sup>, Shannon M. Mumenthaler<sup>1</sup>, Paul Macklin<sup>1,2</sup>\*

1 Lawrence J. Ellison Institute for Transformative Medicine, University of Southern California, Los Angeles, California, United States of America, 2 Intelligent Systems Engineering, Indiana University, Bloomington, Indiana, United States of America, 3 Opto-Knowledge Systems, Inc., Torrance, California, United States of America

\* macklinp@iu.edu

#### https://github.com/bsc-life/PhysiBoSSv2

Systems biology

PhysiBoSS: a multi-scale agent based modelling framework integrating physical dimension and cell signalling

Gaelle Letort <sup>1,2,3</sup>\*, Arnau Montagud <sup>1,2,3</sup>, Gautier Stoll <sup>4</sup>, Randy Heiland <sup>5</sup>, Emmanuel Barillot <sup>1,2,3</sup>, Paul Macklin <sup>5</sup>, Andrei Zinovyev <sup>1,2,3</sup> and Laurence Calzone <sup>1,2,3,\*</sup>



#### **USE CASES**

PerMedCoE works on five use cases to drive the development of cell-level simulations

Cancer Diagnosis Based on Omics Information



Drug Synergies for Cancer Treatment



Tumour Evolution Based on Single-Cell Omids and Imaging

Personalised Modelling of Groups of Rare-Disease Related Patients COVID-19 Multiscale Modelling of the Virus and Patients' Tissue



### Multiscale modelling for COVID-19 infection

- ( Our work builds upon the PhysiCell for COVID initiative, a multiscale model of SARS-CoV-2 dynamics in lung tissue that can integrate
  - virus infection
  - epithelial host cell demise
  - different immune cells' response
- We incorporate PhysiBoSS so we can integrate
   Boolean models
  - offers mechanistic insights of SARS-CoV-2 infection and dissemination among human host cells



#### MOI: multiplicity of infection

#### MOI = 0.10, no immune



#### MOI = 0.10, default immune



### PhysiBoSS-COVID can integrate cell-specific Boolean models

- We incorporate cell- and pathway-specific Boolean models to detail the interactions of virus and human cells.
- ( These models come from the COVID-19 Disease Map initiative
  - Integrates and formalises mechanistic knowledge of COVID-19 infection
  - Map to model converted by CaSQ
- ( These Boolean models can then be simulated using MaBoSS allowing for the study of the cells' mechanisms and their perturbations





### PhysiBoSS allows for the mechanistic exploration of COVID-19 infection

- With PhysiBoSS we can now inspect
  - mutants that affect epithelial cells' apoptosis

<nil>

Time (hours)

Time (hours)

Virus inside

Virus inside -- Apoptosis type

Virus inside -- Apoptosis type

WT

MKO,

50%

heterogeneous cell ٠ populations

1500

1500



t=0



Time (hours)



#### PARTNERS







MAX DELBRÜCK CENTER FOR MOLECULAR MEDICINE IN THE HELMHOLTZ ASSOCIATION







cnag

centre nacional d'anàlisi genòmica

centro nacional de análisis genómico



University of Ljubljana



Atos





**HEIDELBERG UNIVERSITY** HOSPITAL

54

www.permedcoe.eu

#### Contact arnau.montagud@bsc.es @ArnauMontagud

# Per Me Co

HPC/Exascale Centre of Excellence in Personalised Medicine

#### Follow us in social media:



www.linkedin.com/company/permedcoe @permedcoe



The PerMedCoE project has received funding from the European Union's Horizon 2020 research and innovation programme under the grant agreement Nº951773 www.permedcoe.eu