

# Metabolic Engineering with OptFlux

**Paulo Maia**

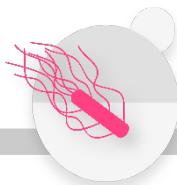
pmaia@silicolife.com

12-SEP-2017

DSM Rosalind Franklin Biotechnology Center  
Delft, The Netherlands



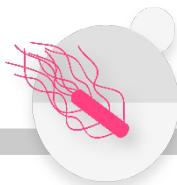
This project has received funding from the European Union's Horizon 2020 research and innovation programme  
under grant agreement No 686070



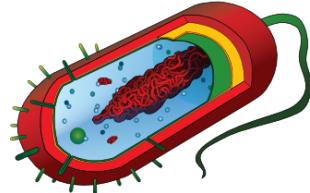
[www.optflux.org](http://www.optflux.org)

## Materials:

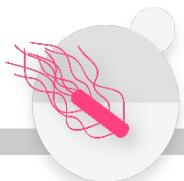
<http://darwin.di.uminho.pt/decaf/materials.zip>

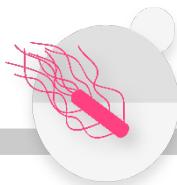


# *in silico* design of optimal cell factories



- **Founded in 2010** and privately owned by its founders
- Enabling the design of **optimized microbial strains** for the production of biofuels, chemicals or biopolymers
- Bridging computer sciences, life sciences and bioengineering
- Working with **leading chemical, materials and synthetic biology companies**





## Agenda

15:50 – 17:50

### Metabolic Engineering with OptFlux

15:50 Brief introduction to OptFlux

16:05 Model Handling:

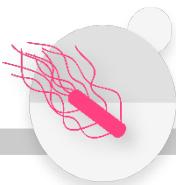
- Model Formats;
- Loading and exporting;
- Model information/representation.

16:35 Phenotype prediction/simulation:

- Wild-type / mutant prediction;
- Flux Variability Analysis (FVA) and solution robustness;
- Calculating maximum theoretical yields;
- Defining and using environmental conditions;
- Layout visualization.

17:20 Strain Optimization:

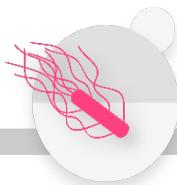
- Configuration and execution;
- Interpreting results.



# Introduction to OptFlux

***OptFlux*** software is a reference open-source software tool for **Metabolic Engineering *in silico*** applications.

It allows the use of **metabolic models** for phenotype prediction and computational strain optimization purposes.



# OptFlux – Main Features

## Open-source

- allows all users to use the tool freely
- invites the contribution of other researchers

## User-friendly

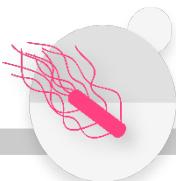
- facilitates its use by users with no/little background in modeling/informatics
- integrates a tool that allows the visualization of the pathway layouts and results

## Modular

- facilitates the addition of specific components by computer scientists / bioinformaticians
- based on the plug-in based AlBench platform

## Compatible with standards

- SBML- Systems Biology Markup Language
- Cell Designer layouts



# OptFlux – Functional Overview

## Functional Modules

### Model Handling

Model Loading  
Model Simplification  
Model Visualization  
Model Export

### Simulation

Wild -Type  
Mutant Strains: gene/  
reaction knockouts / over-  
under expression  
Env. Conditions

### Analysis

Critical reactions/genes  
Zero-valued fluxes  
Equivalent reactions Flux  
Variability Analysis

### Optimization

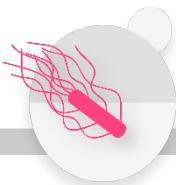
Gene / Reaction  
knockouts  
Over/Under expression  
Multiobjective

Flat Files / CSV  
SBML  
Cell Designer Layouts  
OptFlux Models Repository

FBA/pFBA  
MOMA/LMOMA  
ROOM  
MiMBI

OptGene (EA,SA)  
MOStrain (SPEA2)  
OptKnock (via COBRA)  
GDLS (via COBRA)

## Methods & Technologies



# OptFlux – Model Formats

System Biology  
Markup  
Language (SBML)

JSON (BiGG  
models)

Internal model  
repository

Flat files

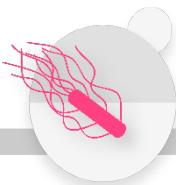
Table format  
(CSV, TSV) - Excel

Metatool

Wizard for new project creation involves specifying format, but also other options such as drain/ external **metabolite identification** and **biomass equation**

Heuristic methods are provided by default but the user can specify otherwise

Models in the internal repository can be “safely” loaded having default parameters



## Wild type simulation

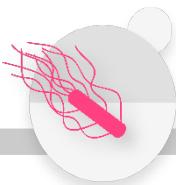
- Methods: FBA or parsimonious FBA

## Mutant simulation

- Gene or reaction changes
- Deletions or under-overexpression
- Methods: FBA, pFBA, MOMA, LMOMA, ROOM, MiMBI

## Environmental conditions

- Can be used in any simulation
- Allows defining media through available external metabolites



## Simulation based analysis

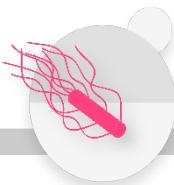
- Critical genes and reactions
- Flux variability analysis: flux variation graph, flux limits, zero value fluxes

## Structural analysis

- Zero valued fluxes
- Equivalent reactions
- Balance checker

## Model visualization

- Uses an internal library based on Prefuse
- Can show the model, model parts and **overlap** of simulation results
- Supports layouts from multiple formats (including Escher .json maps)



## Algorithms

- For single objective: Evolutionary Algorithms (OptGene) and Simulated Annealing
- For multiobjective: SPEA 2 or adaptations of EAs and SA using fitness aggregation
- OptKnock is available via CoBRA connection (Matlab)

## Mutant simulation / optimization type

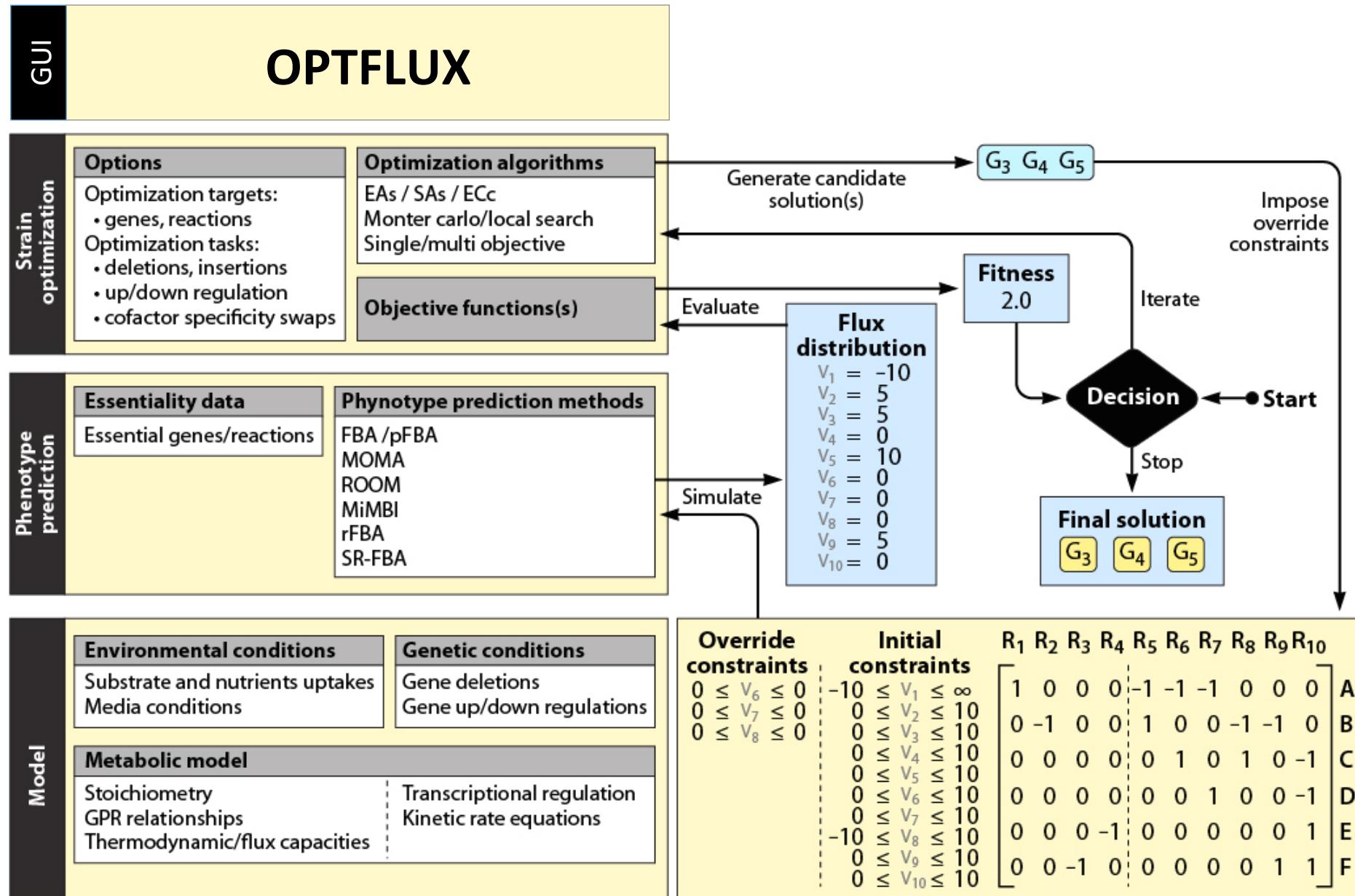
- Can use any of the options stated before (genes/ reactions; deletions/ over-under expression)
- Specific environmental conditions can be used

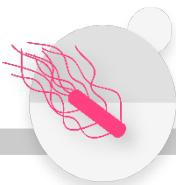
## Parameters

- Solution size can be fixed or variable; maximum size is defined
- Termination defined by the maximum number of function evaluations



# OptFlux – Strain optimization





# OptFlux – Other plug-ins

## Minimal Cut Set (MCS) enumeration (MCSEnumerator)

- Java implementation of MCSEnumerator (Klamt, 2014)

## Constraint-based Flux Analysis (CBFA)

- Phenotype prediction integrating metabolic models with constraints derived from experimental data (Carreira, 2014)

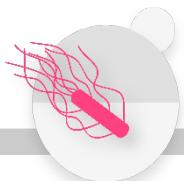
## Network Topology Analysis

- Network-centric methods for metabolic models analysis (Pinto, 2012)

## Regulatory Tool

- Integrated metabolic/regulatory phenotype prediction and strain optimization (Rocha, 2014)

...



# OptFlux – The Environment

OptFlux File Simulation Analysis Optimization Help

OptFlux 3.3.5

MENU - OPERATIONS

Clipboard

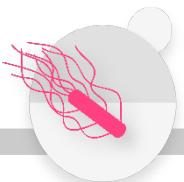
Optflux•3

CLIPBOARD - DATA

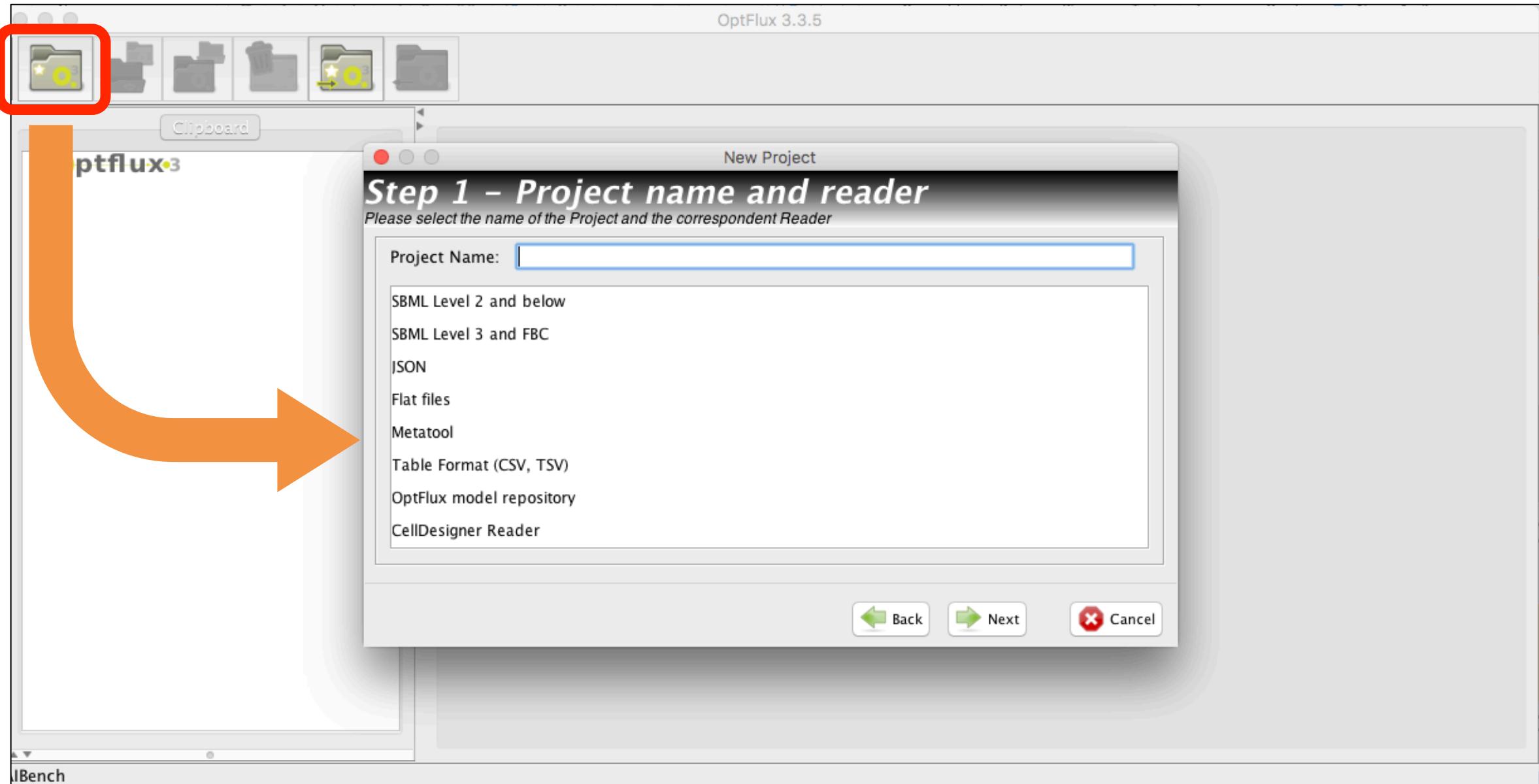
VIEWS

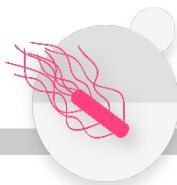
AlBench

The screenshot displays the OptFlux software environment. At the top is a menu bar with the following items: Apple (Mac OS X icon), OptFlux, File, Simulation, Analysis, Optimization, and Help. To the right of the menu bar is the text "OptFlux 3.3.5". Below the menu bar is a toolbar with five icons representing different file operations. The main workspace is titled "MENU - OPERATIONS" and contains two large sections: "CLIPBOARD - DATA" on the left and "VIEWS" on the right. The "CLIPBOARD - DATA" section shows a history of clipboard entries, with the most recent one being "Optflux•3". The "VIEWS" section is currently empty. At the bottom of the interface, there is a small tab labeled "AlBench".



# OptFlux – Loading models



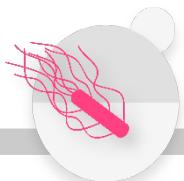


# OptFlux – Loading models

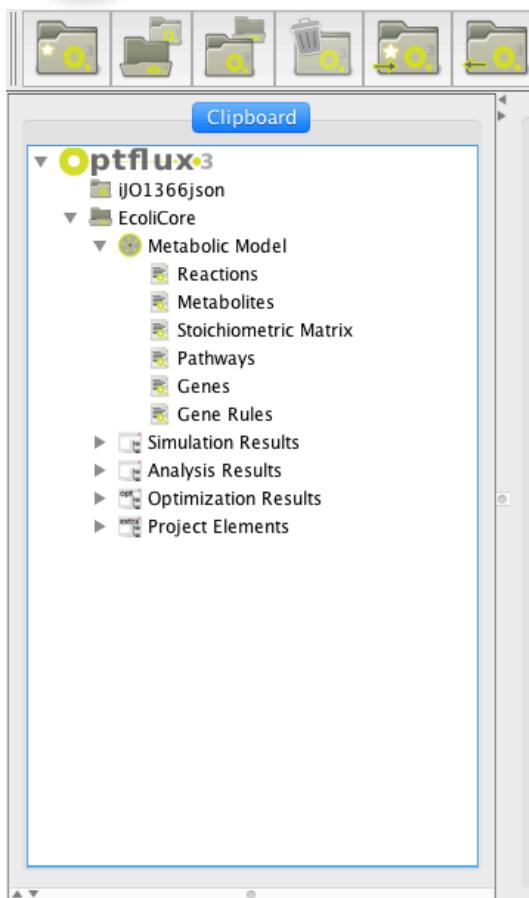
The image displays three sequential steps in the OptFlux software:

- Step 2 - Selection of Files:** A file selection dialog box titled "New Project". It shows an "SBML File" field containing the path "/Users/miguelrocha/Software/Models/ecoli\_core\_model.xml". There is also a "Find..." button.
- Step 3 - Drains and External Metabolites:** A configuration dialog box. It has radio buttons for "Don't create drains", "Remove the external Metabolites" (which is selected), and "Create the Drains". Below these are two radio buttons: "Use compartments heuris..." and "Use regexps (mandatory for models from Bi...)" (which is selected). A "Add Regular Expression" button is available. A dropdown menu "Ends With:" is set to "\_b". On the right, a list of "Ext. Metabolites" is shown: M\_succ\_b, M\_fru\_b, M\_acald\_b, M\_lac\_D\_b, and M\_akg\_b. Buttons for "BiGG database regexp" and "Update Ext..." are present.
- Step 4 - Biomass growth:** A selection dialog box. It asks "Please select the biomass reaction:". A table lists "Selected Biomass Reaction" as R\_Biomass\_Ecoli\_core\_w\_GAM. The table includes columns for "ID" and "Name". The row for R\_Biomass\_Ecoli\_core\_w\_GAM is highlighted. Other rows include R\_TPI (triose-phosphate isomerase), R\_PTAr (phosphotransacetylase), R\_PPCK (phosphoenolpyruvate carboxykinase), R\_GLNabc (L-glutamine transport via ABC system), R\_ATPM (ATP maintenance requirement), R\_PGI (glucose-6-phosphate isomerase), and R\_Biomass\_Ecoli\_core\_w\_GAM (Biomass Objective Function with GAM). A search bar at the bottom left contains "R\_Biomass\_Ecoli\_core\_w\_GAM" and a "case sensitive ?" checkbox. Navigation buttons at the bottom right include "Back", "Finish", and "Cancel".

**Model details:**  
external metabolites,  
biomass



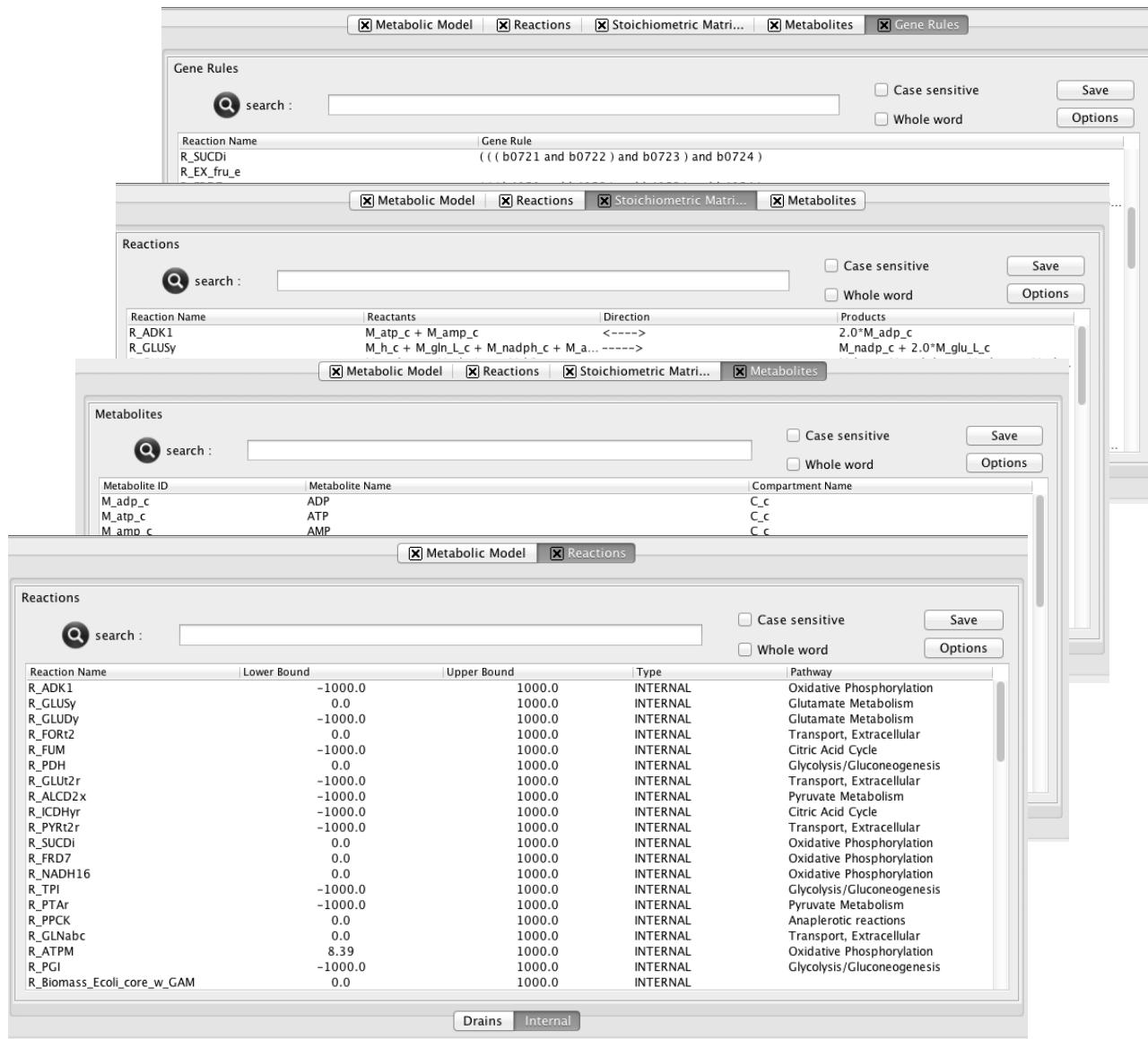
# OptFlux – Model Information



The screenshot shows the OptFlux software interface with a sidebar on the left containing a tree view of model components:

- Clipboard
- Optflux3
  - iJO1366.json
  - EcoliCore
    - Metabolic Model
      - Reactions
      - Metabolites
      - Stoichiometric Matrix
      - Pathways
      - Genes
      - Gene Rules
    - Simulation Results
    - Analysis Results
    - Optimization Results
    - Project Elements

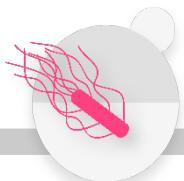
Clipboard



The image displays four windows from the OptFlux interface:

- Gene Rules**: Shows a search bar and a table with reaction names R\_SUCDI and R\_EX\_fru\_e, and a gene rule: (((b0721 and b0722) and b0723) and b0724).
- Reactions**: Shows a search bar and a table for reactions R\_ADK1 and R\_GLUSy. R\_ADK1 converts M\_atp\_c + M\_amp\_c to 2.0\*M\_adp\_c. R\_GLUSy converts M\_h\_c + M\_gln\_L\_c + M\_nadph\_c + M\_a... to M\_nadp\_c + 2.0\*M\_glu\_L\_c.
- Metabolites**: Shows a search bar and a table for metabolites M\_adp\_c, M\_atp\_c, and M\_amp\_c, all located in compartment C\_c.
- Reactions**: Shows a search bar and a detailed table of reactions. The table includes columns for Reaction Name, Lower Bound, Upper Bound, Type, and Pathway. Key entries include R\_ADK1 (Oxidative Phosphorylation), R\_GLUSy (Glutamate Metabolism), R\_FORT2 (Transport, Extracellular), R\_FUM (Citric Acid Cycle), R\_PDH (Glycolysis/Glucoseogenesis), R\_GLUT2r (Transport, Extracellular), R\_ALCD2x (Pyruvate Metabolism), R\_ICDHyd (Citric Acid Cycle), R\_PYRt2r (Transport, Extracellular), R\_SUCDI (Oxidative Phosphorylation), R\_FRD7 (Oxidative Phosphorylation), R\_NADH16 (Oxidative Phosphorylation), R\_TPI (Glycolysis/Glucoseogenesis), R\_PTAr (Pyruvate Metabolism), R\_PPCK (Anaplerotic reactions), R\_GLNabc (Transport, Extracellular), R\_ATPM (Oxidative Phosphorylation), R\_PGI (Glycolysis/Glucoseogenesis), and R\_Biomass\_Ecoli\_core\_w\_GAM (Glycolysis/Glucoseogenesis).

Model views



# OptFlux – Phenotype prediction – Wild Type

File Simulation Analysis Optimization

Wild Type... Knockout Under Over Expression Clipboard

WT\_Simulation1

Simulation Information

Method Name: pFBA

Solution Type: Not available.

Environmental Conditions: Not available.

Objective Function:  $\min \Sigma|V| = 518.4170891$

Biomass value: 0.873913

Net Conversions:

Consumption		Production	
Metabolite	Value	Metabolite	Value
M_nh4_e	4.765	M_co2_e	22.809
M_pi_e	3.215	M_h2o_e	29.175
M_o2_e	21.799	M_h_e	17.531
M_glc_D_e	10	M_ac_e	0

Simulation Solution Drain Reaction Values Internal/ Transport reaction values Genetic Conditions ▶

Wild Type

Select Project and Model

Project: ecoliCoreModel

Objective Function

Flux: R\_Biomass\_Ecoli\_core\_w\_GAM

Maximize  Minimize  
Max R\_Biomass\_Ecoli\_core\_w\_GAM

Simulation Method: pFBA

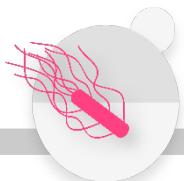
Select Environmental Conditions

Use EnvironmentalConditions: ▾

Ok Cancel



## Simulation results



# OptFlux – Phenotype prediction – Mutants

**Reaction Simulation**

Select Project and Model  
Project: ecoliCoreModel

**Reactions**

Available:

Selected: R\_ACKr, R\_ACr2r

**Select Simulation Method**  
ROOM  
Control Flux Distribution Source  
 FBA  Flux Measures

**Flux:**  
R\_Biomass\_Ecoli\_core\_w\_GAM  
 Maximize  Minimize  
Max R\_Biomass\_Ecoli\_core\_w\_GAM

**Select Environmental Conditions**  
 Use EnvironmentalConditions:

**Ok** **Cancel**

## Reaction Knockouts

**Gene Simulation**

Select Project And Model  
Project: ecoliCoreModel

**Genes Knockout Definitions**

Available: b2926, b2935, b2975, b2976, b2987, b3114, b3115, b3212, b3213, b3236, b3386, b3493, b3528, b3603, b3612, b3731  
Selected: b0114, b0115, b0116, b3403  
Inactive Reactions: R\_PPCK, R\_PDH, R\_AKGDH

**Select Simulation Method**  
MOMA  
Control Flux Distribution Source  
 FBA  Flux Measures

**Select Environmental Conditions**  
 Use EnvironmentalConditions: Env. Conditions 1

**Objective Function Definition**  
Flux:  
R\_Biomass\_Ecoli\_core\_w\_GAM  
 Maximize  Minimize  
Max R\_Biomass\_Ecoli\_core\_w\_GAM

**Ok** **Cancel**

## Gene Knockouts

**Gene Simulation**

Select Project And Model  
Project: ecoliCoreModel

**Genes UnderOver Selection**

Available: b1621, b1676, b1723, b1761, b1773, b1779, b1812, b1817, b1818, b1819, b1849, b1852, b1854, b2097, b2133, b2276  
Selected: b0451 (0.5), b0767 (2.0), b1101 (0.0), b1702 (0.25), b2029 (4.0)  
Regulated Reactions: R\_PGL 2.0, R\_GND 4.0, R\_NH4t 0.75, R\_PPS 0.25, R\_GLCpts 0.75

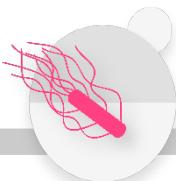
**Select Simulation Method**  
pFBA  
Control Flux Distribution Source  
 FBA  Flux Measures

**Select Environmental Conditions**  
 Use EnvironmentalConditions: Env. Conditions 1

**Objective Function Selection**  
Flux:  
R\_Biomass\_Ecoli\_core\_w\_GAM  
 Maximize  Minimize  
Max R\_Biomass\_Ecoli\_core\_w\_GAM

**Ok** **Cancel**

## Over/under Expression



# OptFlux – Strain optimization

Project (model) →

Env. conditions →

Process setup →

Essentiality →

Algorithm

Objective function(s)

**Perform Strain Optimization**

**Select Project**  
Project: Scerevisiae\_gombert

**Select Method**  
SPEA2 Reaction Knockout

**Select Environmental Conditions**  
----[ NONE ]----

**Objective Functions Setup**  
BPCY: Biomass–Product Coupled Yield

Add ↓      ↑ Remove

**Selected Objective Functions**  
Optimization Objective Function | Simulation Method | Reference | Simulation OF

**Optimization Basic Setup**  
Maximum Number Of Solutions Evaluations: 5000  
Maximum Number Of Modifications: 6  
 Variable solution size

**Critical Reactions**  
----[ NONE ]----  
 Exclude drain reactions  
 Exclude transport reactions

Ok Cancel



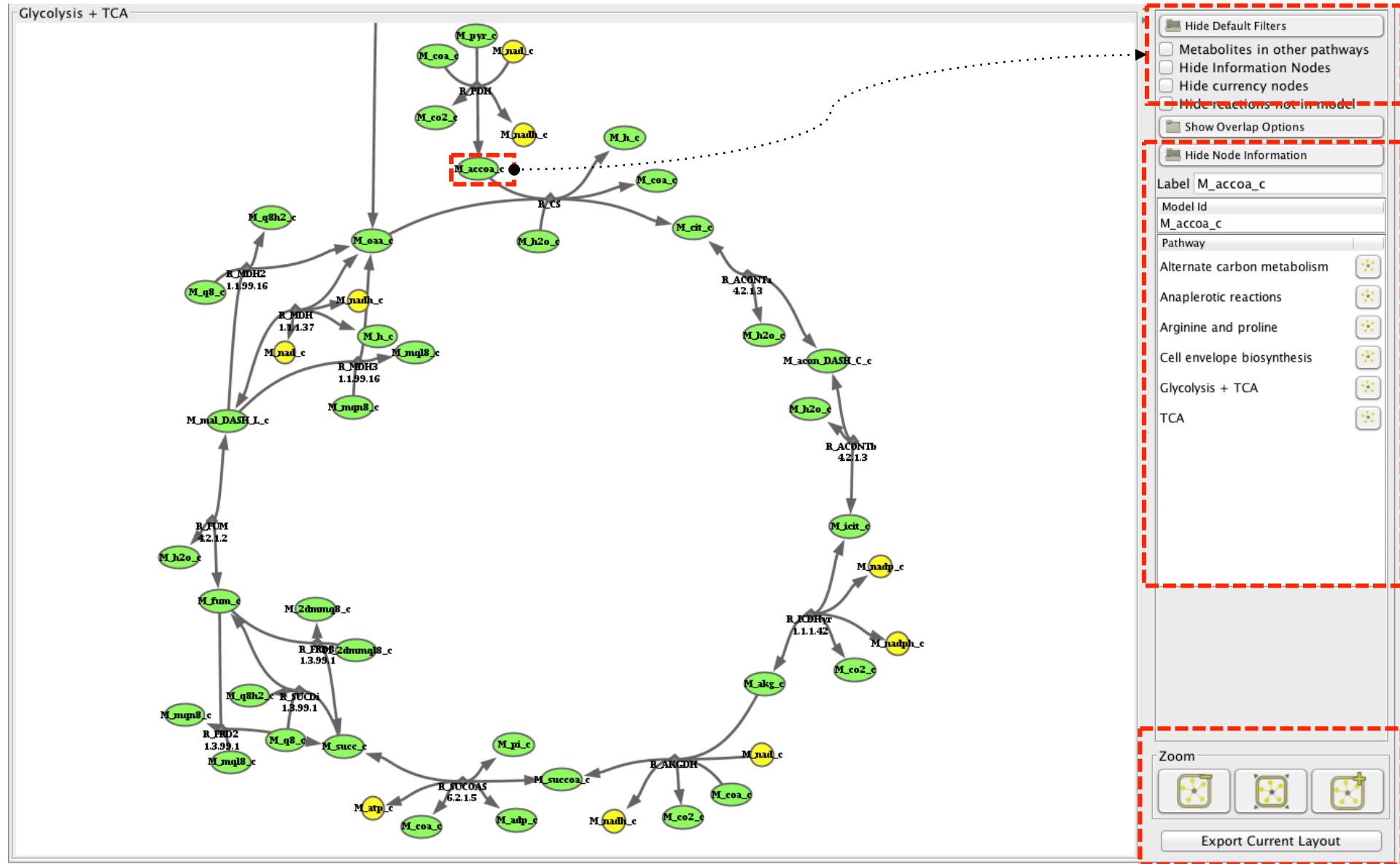
# OptFlux – Pathway visualization

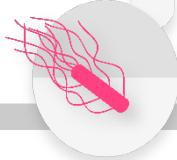
## Mouse shortcuts (on nodes):

- Left click: select node (metabolite/reaction)
- Right click: display node's available actions

## Mouse shortcuts (on layout):

- Left click and drag: area selection (multiple nodes)
- Right click and drag: pan
- Scroll up/down: zoom in/out

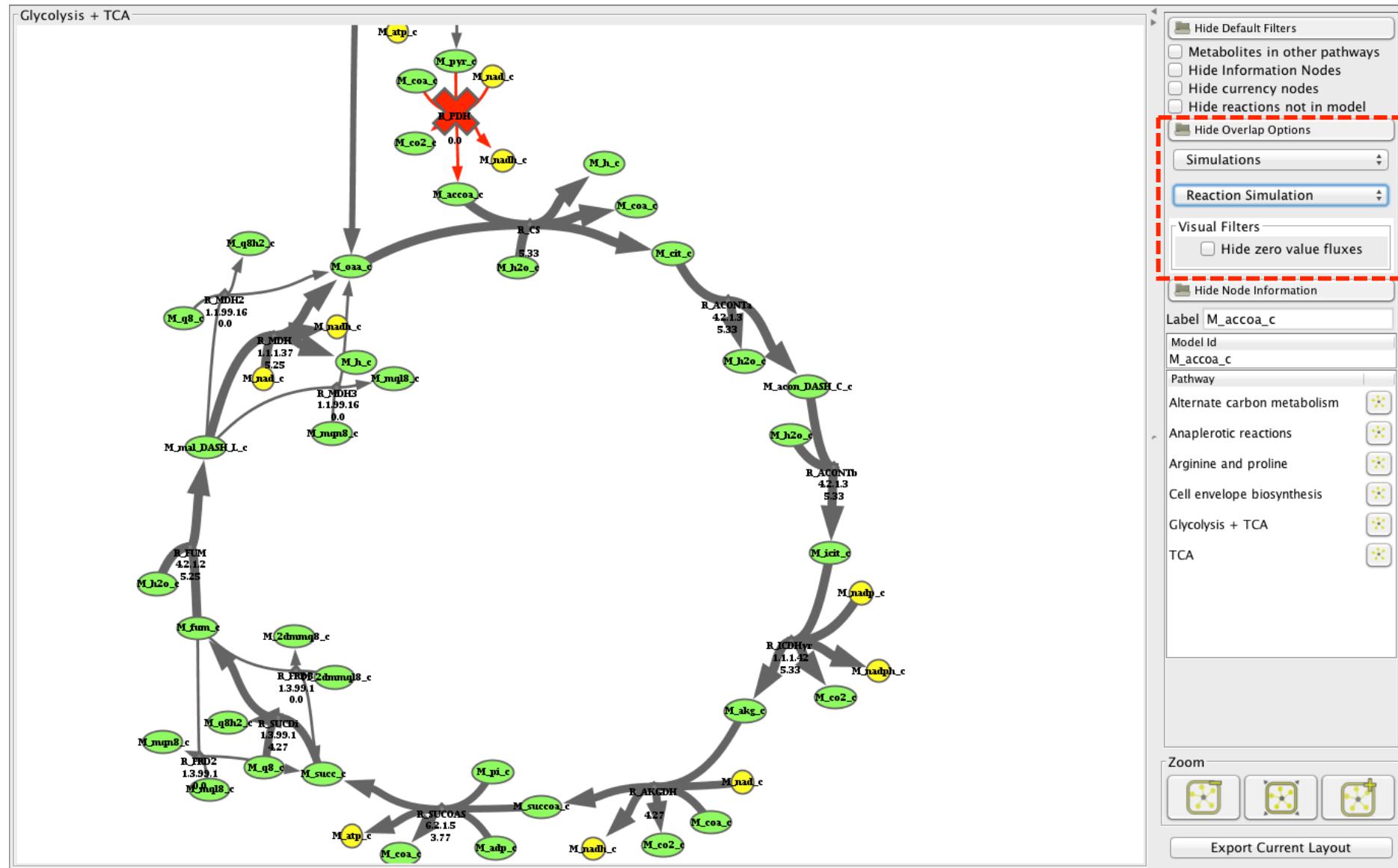




# OptFlux – Pathway visualization: overlaps

## Notes:

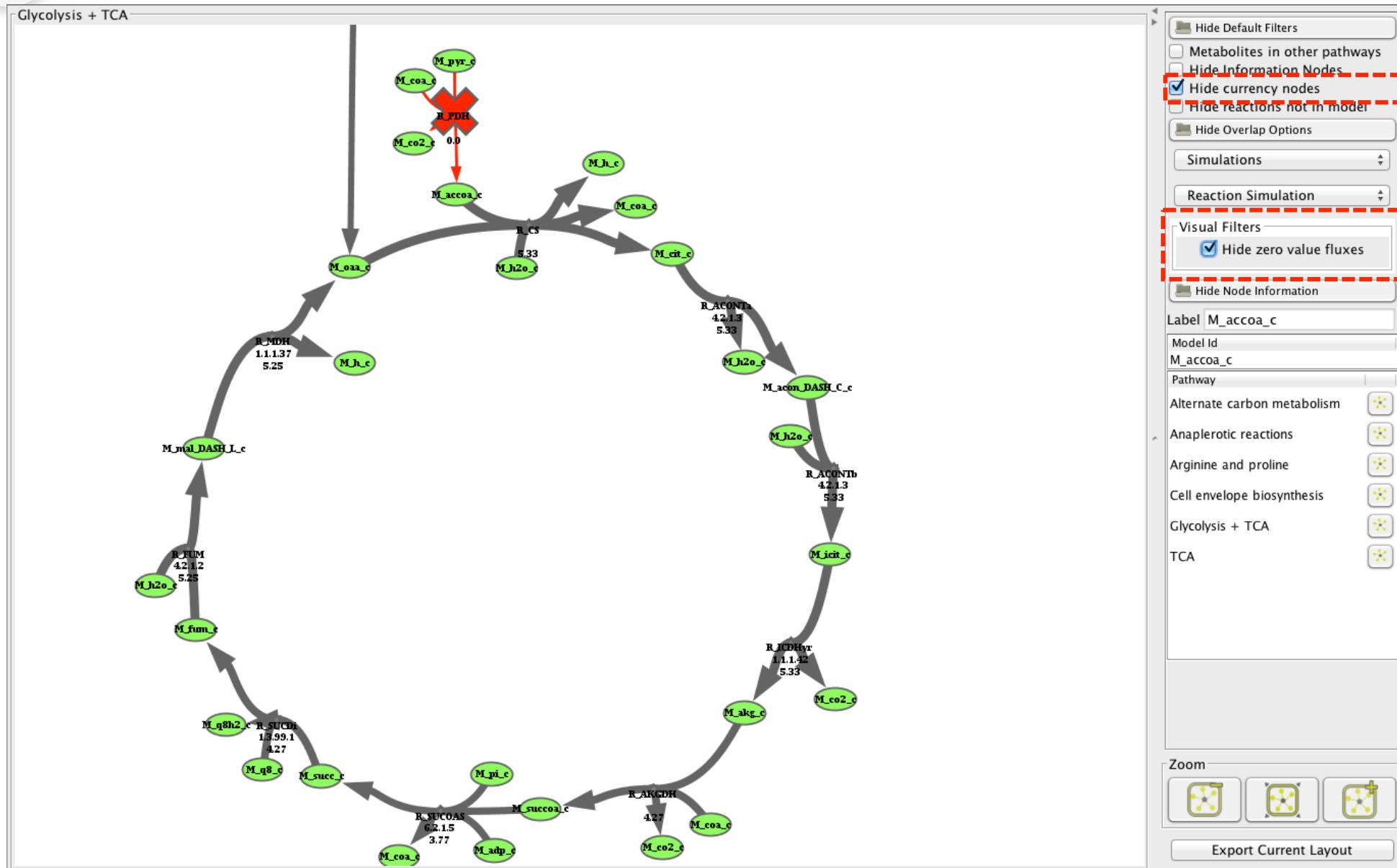
- Red crosses represent a suppression (via knockout) of the respective reaction
- Upward green arrows represent an overexpression relative to the WT of the corresponding flux
- Downward red arrows represent an underexpression relative to the WT of the corresponding flux



**Overlaps:**  
Allows selecting a flux distribution from any simulation and overlapping the fluxes on top of the currently loaded layout. Arrow thicknesses reflect the relative flux going through the reactions

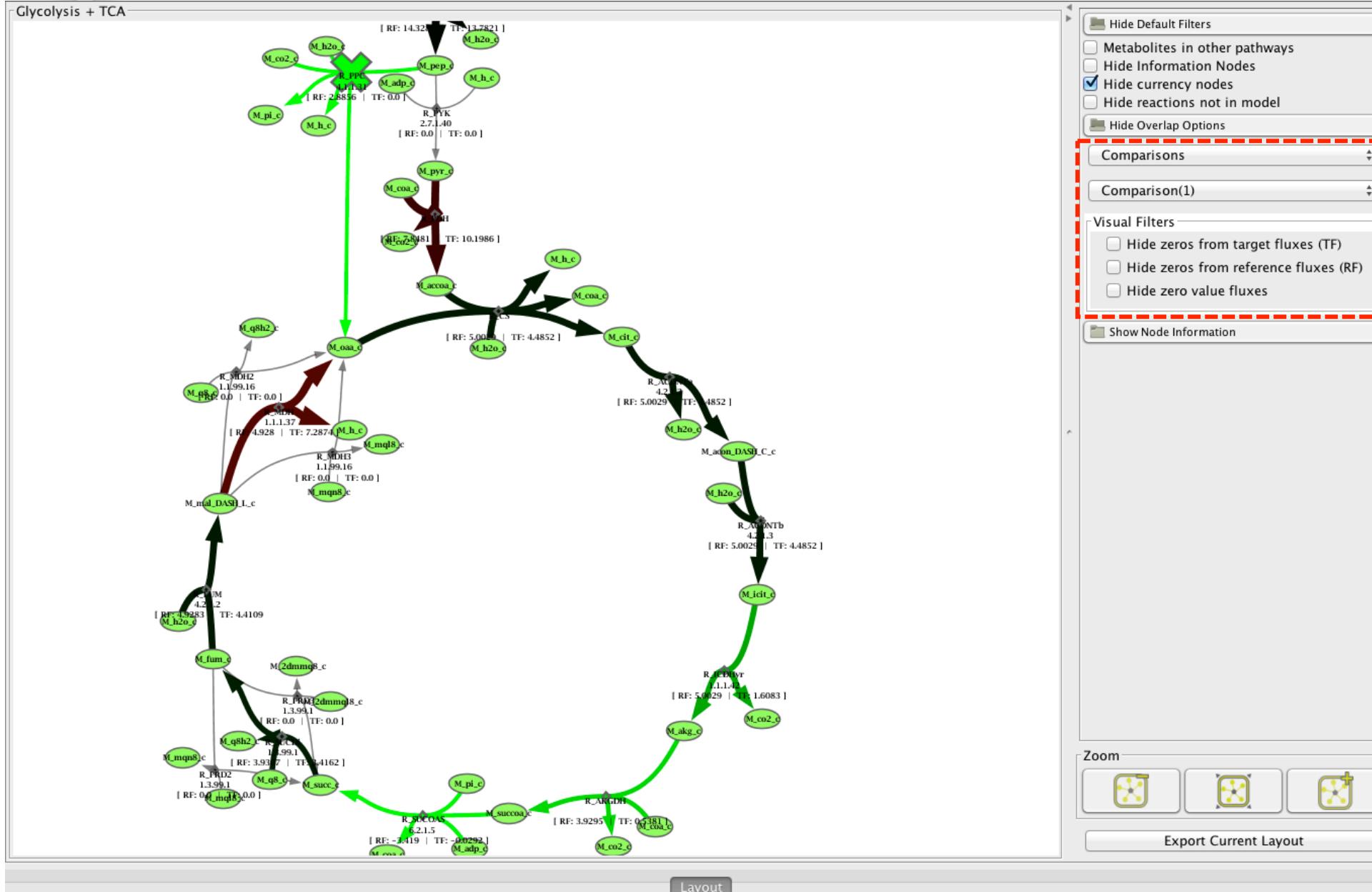


# OptFlux – Pathway visualization: overlaps





# OptFlux – Pathway visualization: comparison



## Notes:

- Green colored arrows represent reactions where the flux is higher in the reference than in the target flux distribution
- Red colored arrows represent reactions where the flux is higher in the the target reference than in the flux distribution
- The darker the colors, the closer the two reaction's fluxes are to each other
- Completely black reactions represent no change between the two reaction's fluxes.