OPTFLUX 3
Beginner's tutorial

Supported by
SILICOLIFE
Created by
Universidade do Minho
IBB
Institute for Biotechnology and Bioengineering
Computer Science and Technology Center

Optflux

OptFlux is an open-source software created by BiSBII (Bioinformatics and Systems Biology Interdisciplinary Initiative), a joint initiative by

IBB/CEB – Institute for Biotechnology and Bioengineering
Centre of Biological Engineering (University of Minho)
CCTC – Computer Science and Technology Center (University of Minho)

OptFlux is supported by
SilicoLife Lda.

Optflux

LICENSES

For this tutorial:

For the OptFlux software:

Created inside the SysBio Research Group (http://sysbio.di.uminho.pt)
Supported by SilicoLife Lda. (http://www.silicolife.com)

Optflux

CONTENTS

LICENSES **3**

CONTENTS **4**

FIRST THINGS FIRST! **5**

1. CREATING A NEW PROJECT **6**

1.1 FROM A SBML FILE 6

1.2 FROM OPTFLUX INTERNAL REPOSITORY 8

1.3 PROJECT CREATED IN THE CLIPBOARD 9

2. PERFORMING PHENOTYPE SIMULATIONS **12**

2.1 PERFORMING A WILD-TYPE SIMULATION 12

2.2 CREATING ENVIRONMENTAL CONDITIONS 15

2.3 PERFORMING A MUTANT SIMULATION – REACTION DELETIONS 17

2.4 PERFORMING A MUTANT SIMULATION – GENE KNOCKOUT 20

2.5 PERFORMING SIMULATIONS WITH OVER/ UNDER EXPRESSION 21

3. PERFORMING STRAIN OPTIMIZATION **24**

4. EXPLORING ANALYSIS TOOLS **27**

4.1 CRITICAL GENES/ REACTIONS **27**

4.2 FLUX VARIABILITY ANALYSIS **28**

5. VISUALIZATION **30**

5.1 SUCCINATE PRODUCTION WITH *E. COLI* 30

5.2 GLYCINE PRODUCTION WITH *E. COLI* 33

5.3 VISUALIZATION PREFERENCES 37

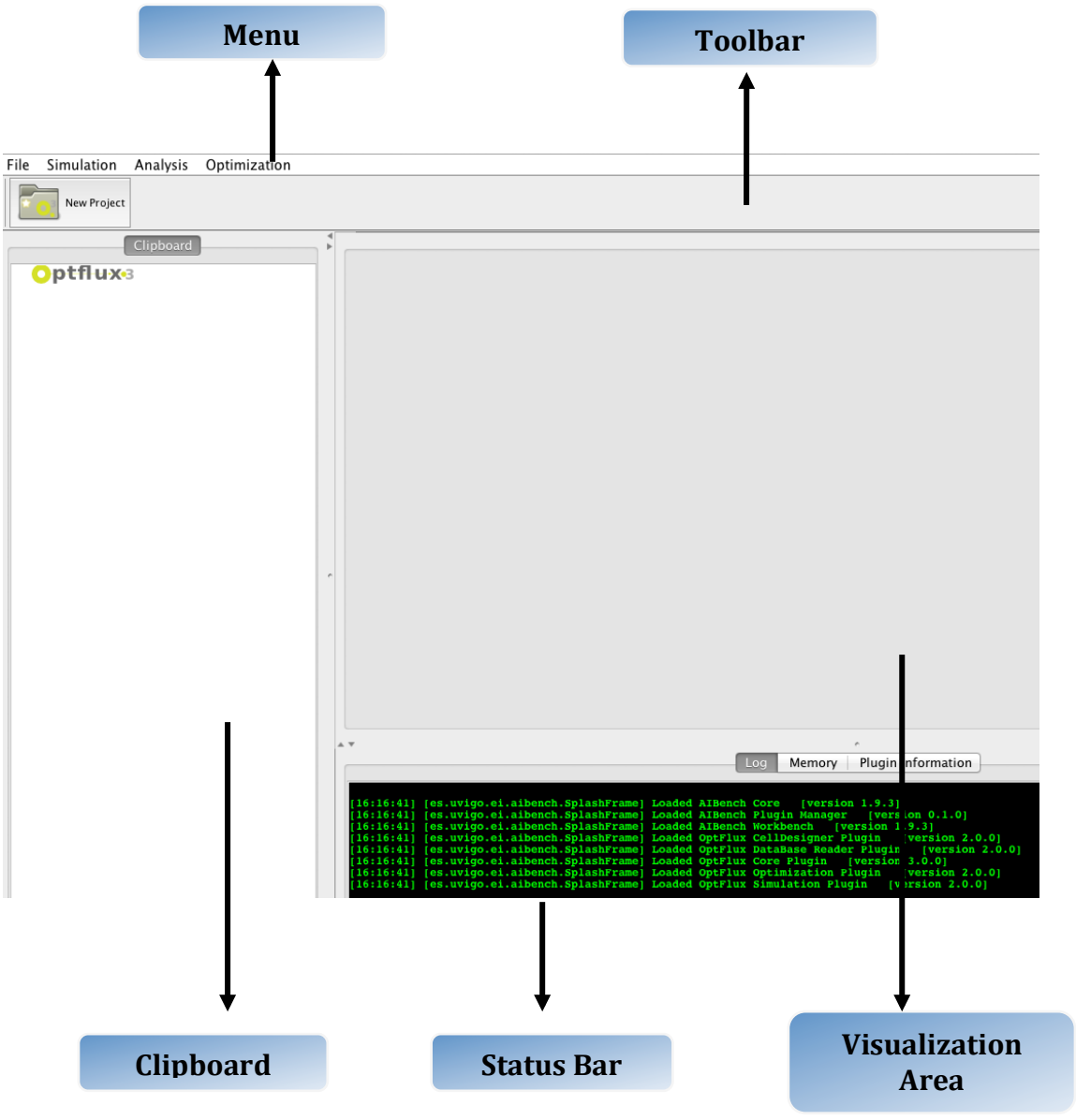
WHAT'S NEXT ? **40**

FIRST THINGS FIRST!

Hello and welcome to the OptFlux 3 beginner's tutorial. If you haven't already downloaded the software please do it here: www.optflux.org.

After launching the software you'll be presented with the layout depicted in the image below. Most of OptFlux main features and operations will be accessible to you either through the **Menu** or the **Toolbar**. You can also have access to them by right-clicking in the **Clipboard** area. Your data types i.e., the project, metabolic models, environmental conditions, simulation/optimization results, layouts for visualization, etc., will always be placed in the **Clipboard** area**.** The **Visualization Area** is the place where you can examine those data types in greater detail. When you click a data type, the different views for that object will be available in this area.

Click around to get familiar with the environment and, after that, jump to the next step. In the document, we will cover the major operations of OptFlux 3, their configuration and results. Over the next sections, phenotype simulation, strain optimization, analysis tools and visualization will be addressed.



1. CREATING A NEW PROJECT

To start a new project, you have to launch the **New Project** wizard available through the *File Menu* or the *Toolbar.*

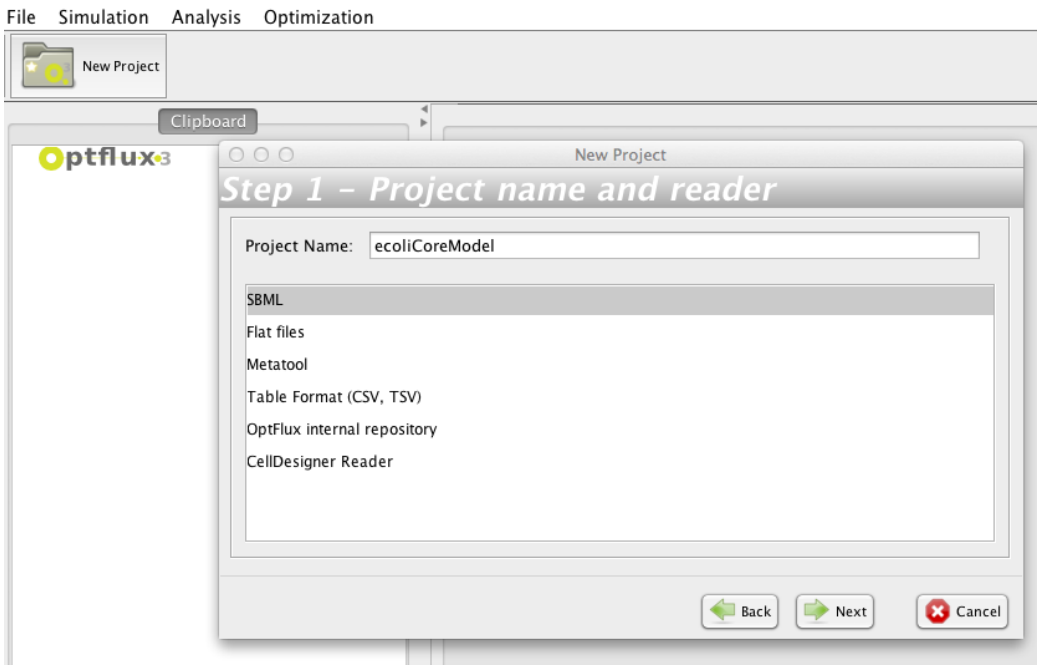
You have the option to create a new project from several different sources/ formats. In this tutorial, we will cover three of those: from a file in the Systems Biology Markup Language (SBML) format, from the OptFlux's model repository and, finally, in the last section of this document, from a Cell Designer SBML format that includes layout information.

1.1 FROM A SBML FILE

In this first example, the SBML format will be used. SBML is an XML dialect created for the representation of Systems Biology models. Check the site www.sbml.org to know more. To follow the steps in this section, you need to download the file *ecoli-core-model.xml*, available in www.optflux.org/tutorial/ecoli-core-model.xml and save it to a folder of your choice. This model is a simplified model of the metabolism of *Escherichia coli*, proposed in *EcoSal Chapter 10.2.1 - Reconstruction and Use of Microbial Metabolic Networks: the Core* Escherichia coli *Metabolic Model as an Educational Guide* by Orth, Fleming, and Palsson (2010). It is also available in the web site http://gcrg.ucsd.edu/Downloads/EcoliCore.

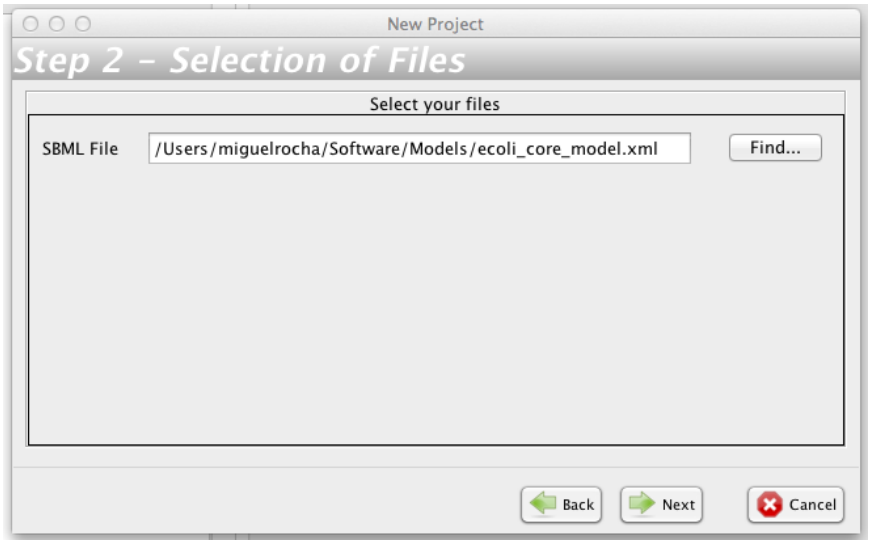
Step 1

In the first step, you must input a valid project name and select the SBML option as the model format, as shown below.



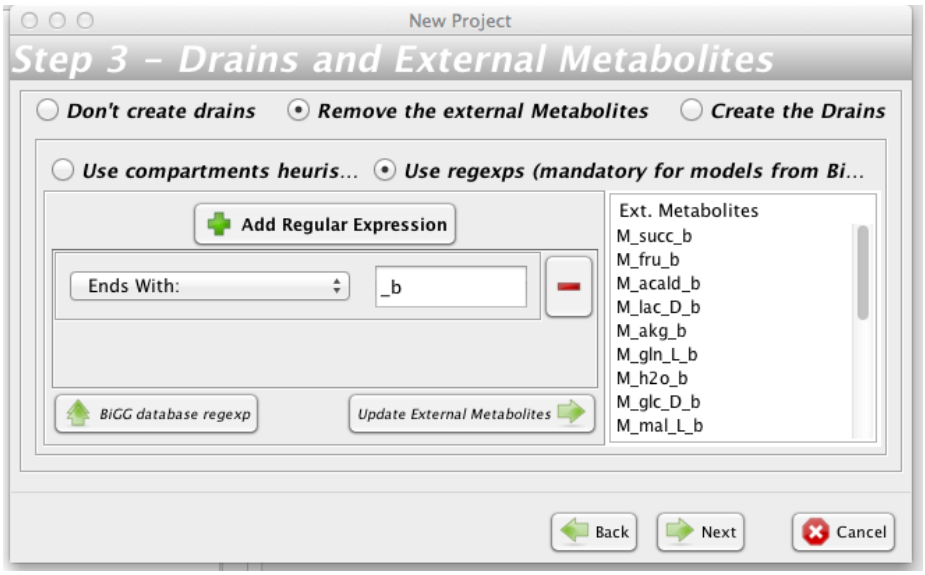
Step 2

In the second step, you must select the file to load. In this example, the file is the one provided, as stated above (*ecoli-core-model.xml*).



Step 3

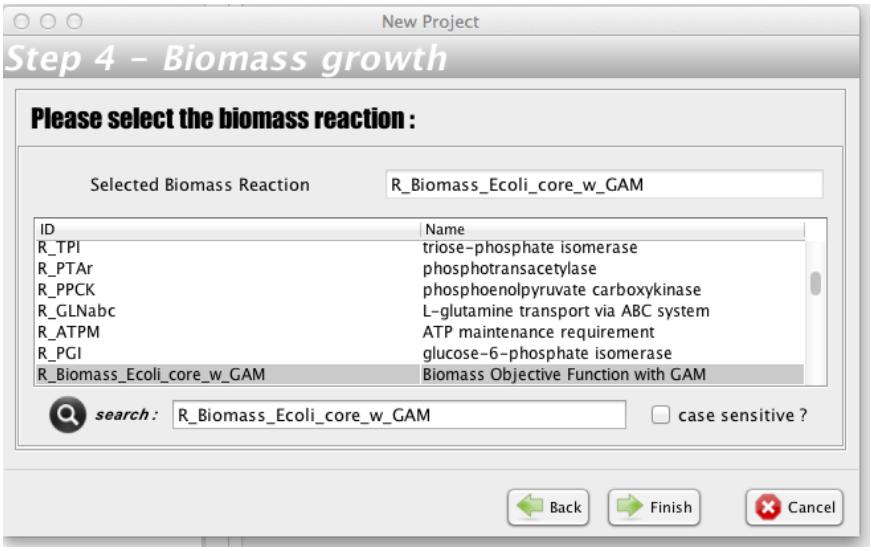
The third step is relative to the extra-cellular environment and identification of drains. This provides a set of options depending on the file format. *OptFlux* will automatically try to find the appropriate method using heuristics. In this case, it should propose to identify external metabolites using a regular expression, by considering all metabolites whose identifier ends with "_b". These will be removed from the model. In this case, this is indeed the best option, and therefore you can accept OptFlux's suggestion and proceed.



Step 4

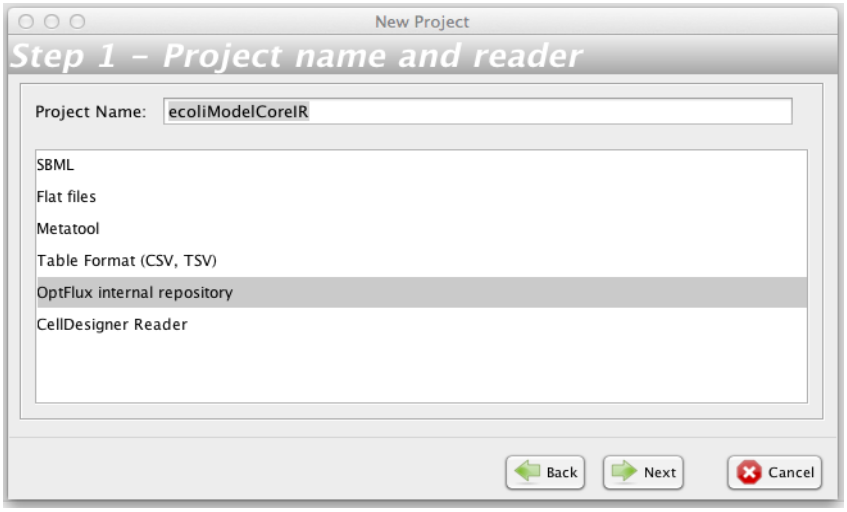
In the fourth step, OptFlux automatically tries to find the biomass growth associated

flux, since this information is essential for both simulation and optimization procedures. A heuristic method will automatically identify the appropriate reaction as it is easy to check.



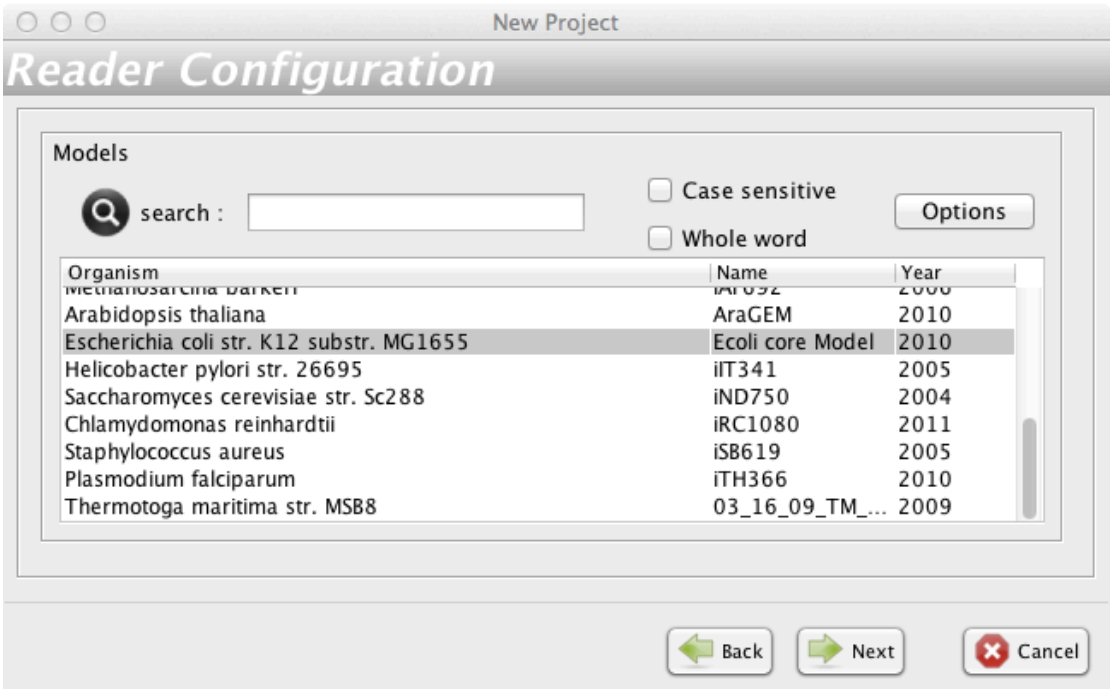
1.2 From OptFlux internal repository

One alternative option for creating a project is to import the model from the OptFlux's repository. This is an option that guarantees the importation of models collected from literature that have been validated by OptFlux's team, regarding the file format. To load a model from the repository you need to have an Internet connection working. If that is the case, choose the **New Project** option as before. In step 1, choose the option "OptFlux internal repository", as shown below.



The following step (2) is to select the model from the list, which contains the organism and the model name. Among the options, you can find the same model that

was loaded in the last section (with the name “Ecoli core Model” for *Escherichia coli*, strain K12, substrain MG1655).

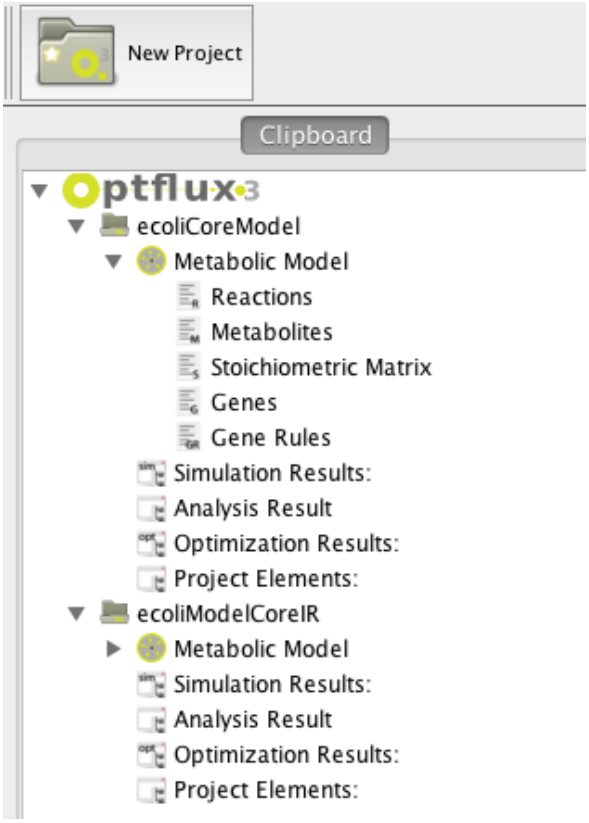


The final step corresponds to the step 4 (step 3 is not necessary in this case), regarding biomass flux selection. As before, the option proposed in the correct one.

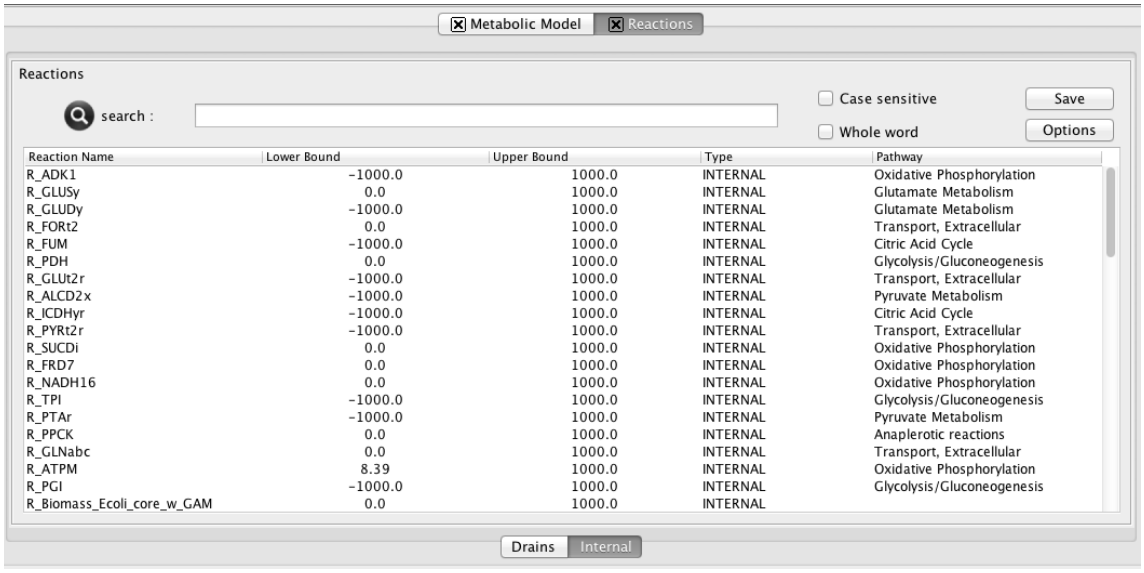
1.3 PROJECT CREATED IN THE CLIPBOARD

After following the steps in either sections 1.1 or 1.2, you will now have the model loaded, corresponding to a project in OptFlux (as shown in the following screenshot). In the image, one can see the default structure of the clipboard, and of an OptFlux project (in this case, the two projects loaded are the ones resulting from following the steps in sections 1.1 and 1.2, respectively).

The core data type is named “Metabolic Model”. Inside, one can access information about reactions (internal and drains), metabolites and also the stoichiometric coefficients in a human-readable fashion. Also, genes and gene rules (specifying gene-rule associations) are shown if the model contains such information.

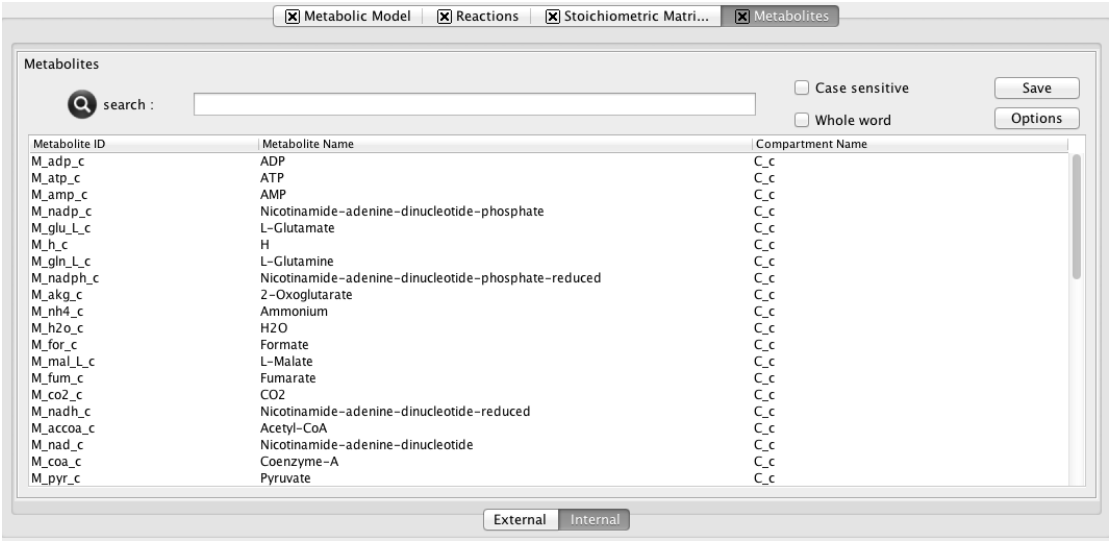


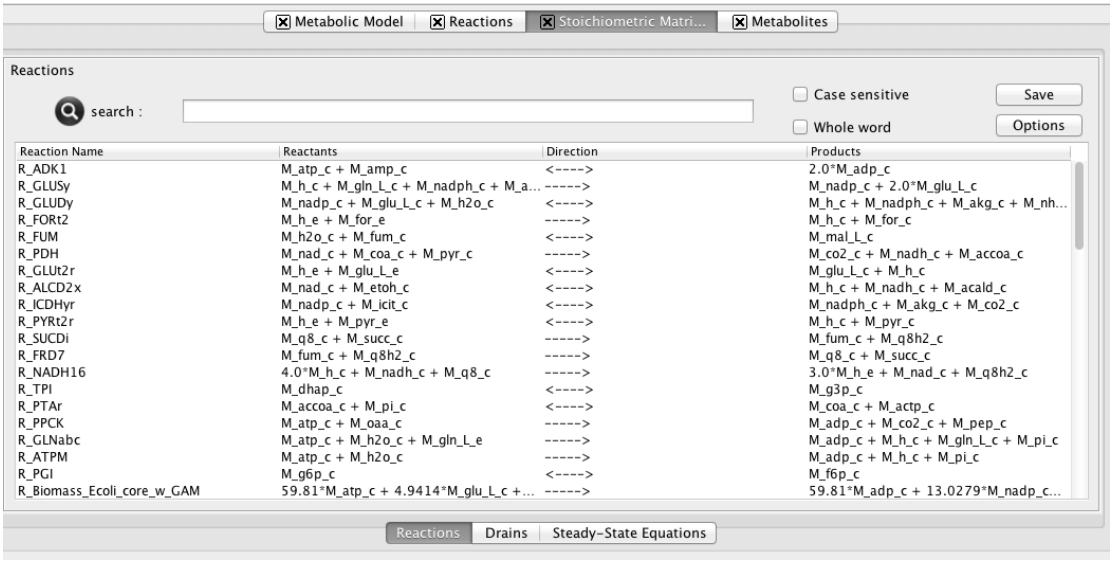
At this point, you should click around a bit to get familiar with this structure and the information therein contained. The viewers for Reactions, Metabolites, the Stoichiometric Matrix and Gene-reaction rules are depicted in the screenshots below.

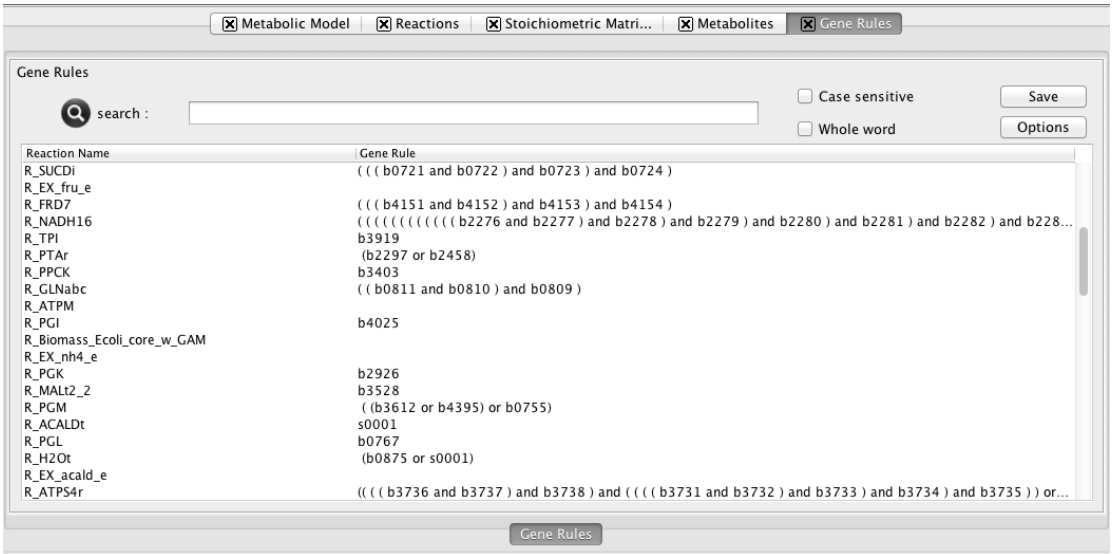


Reaction Name	Lower Bound	Upper Bound	Type	Pathway
R_ADK1	-1000.0	1000.0	INTERNAL	Oxidative Phosphorylation
R_GLUSy	0.0	1000.0	INTERNAL	Glutamate Metabolism
R_GLUDy	-1000.0	1000.0	INTERNAL	Glutamate Metabolism
R_FORt2	0.0	1000.0	INTERNAL	Transport, Extracellular
R_FUM	-1000.0	1000.0	INTERNAL	Citric Acid Cycle
R_PDH	0.0	1000.0	INTERNAL	Glycolysis/Gluconeogenesis
R_GLUt2r	-1000.0	1000.0	INTERNAL	Transport, Extracellular
R_ALCD2x	-1000.0	1000.0	INTERNAL	Pyruvate Metabolism
R_ICDHyr	-1000.0	1000.0	INTERNAL	Citric Acid Cycle
R_PYRt2r	-1000.0	1000.0	INTERNAL	Transport, Extracellular
R_SUCDi	0.0	1000.0	INTERNAL	Oxidative Phosphorylation
R_FRD7	0.0	1000.0	INTERNAL	Oxidative Phosphorylation
R_NADH16	0.0	1000.0	INTERNAL	Oxidative Phosphorylation
R_TPI	-1000.0	1000.0	INTERNAL	Glycolysis/Gluconeogenesis
R_PTAr	-1000.0	1000.0	INTERNAL	Pyruvate Metabolism
R_PPCK	0.0	1000.0	INTERNAL	Anaplerotic reactions
R_GLNabc	0.0	1000.0	INTERNAL	Transport, Extracellular
R_ATPM	8.39	1000.0	INTERNAL	Oxidative Phosphorylation
R_PGI	-1000.0	1000.0	INTERNAL	Glycolysis/Gluconeogenesis
R_Biomass_Ecoli_core_w_GAM	0.0	1000.0	INTERNAL	

Optflux


Metabolic Model
Reactions
Stoichiometric Matri...
Metabolites
Metabolites
search :
Case sensitive
Whole word
Save
Options
Metabolite ID
Metabolite Name
Compartment Name
M_adp_c ADP C_c
M_atp_c ATP C_c
M_amp_c AMP C_c
M_nadp_c Nicotinamide-adenine-dinucleotide-phosphate C_c
M_glu_L_c L-Glutamate C_c
M_h_c H C_c
M_gln_L_c L-Glutamine C_c
M_nadph_c Nicotinamide-adenine-dinucleotide-phosphate-reduced C_c
M_akg_c 2-Oxoglutarate C_c
M_nh4_c Ammonium C_c
M_h2o_c H2O C_c
M_for_c Formate C_c
M_mal_L_c L-Malate C_c
M_fum_c Fumarate C_c
M_co2_c CO2 C_c
M_nadh_c Nicotinamide-adenine-dinucleotide-reduced C_c
M_accoa_c Acetyl-CoA C_c
M_nad_c Nicotinamide-adenine-dinucleotide C_c
M_coa_c Coenzyme-A C_c
M_pyr_c Pyruvate C_c
External
Internal


Metabolic Model
Reactions
Stoichiometric Matri...
Metabolites
Reactions
search :
Case sensitive
Whole word
Save
Options
Reaction Name
Reactants
Direction
Products
R_ADK1 M_atp_c + M_amp_c <----> 2.0*M_adp_c
R_GLUSy M_h_c + M_gln_L_c + M_nadph_c + M_a... -----> M_nadp_c + 2.0*M_glu_L_c
R_GLUDy M_nadp_c + M_glu_L_c + M_h2o_c <----> M_h_c + M_nadph_c + M_akg_c + M_nh...
R_FORt2 M_h_e + M_for_e -----> M_h_c + M_for_c
R_FUM M_h2o_c + M_fum_c <----> M_mal_L_c
R_PDH M_nad_c + M_coa_c + M_pyr_c -----> M_co2_c + M_nadh_c + M_accoa_c
R_GLUt2r M_h_e + M_glu_L_e <----> M_glu_L_c + M_h_c
R_ALCD2x M_nad_c + M_etoh_c <----> M_h_c + M_nadh_c + M_acald_c
R_ICDHyr M_nadp_c + M_icit_c <----> M_nadph_c + M_akg_c + M_co2_c
R_PYRt2r M_h_e + M_pyr_e <----> M_h_c + M_pyr_c
R_SUCDi M_q8_c + M_succ_c -----> M_fum_c + M_q8h2_c
R_FRD7 M_fum_c + M_q8h2_c -----> M_q8_c + M_succ_c
R_NADH16 4.0*M_h_c + M_nadh_c + M_q8_c -----> 3.0*M_h_e + M_nad_c + M_q8h2_c
R_TPI M_dhap_c <----> M_g3p_c
R_PTAr M_accoa_c + M_pi_c <----> M_coa_c + M_actp_c
R_PPCK M_atp_c + M_oaa_c -----> M_adp_c + M_co2_c + M_pep_c
R_GLNabc M_atp_c + M_h2o_c + M_gln_L_e -----> M_adp_c + M_h_c + M_gln_L_c + M_pi_c
R_ATPM M_atp_c + M_h2o_c -----> M_adp_c + M_h_c + M_pi_c
R_PGI M_g6p_c <----> M_f6p_c
R_Biomass_Ecoli_core_w_GAM 59.81*M_atp_c + 4.9414*M_glu_L_c +... -----> 59.81*M_adp_c + 13.0279*M_nadp_c...
Reactions
Drains
Steady-State Equations


Metabolic Model
Reactions
Stoichiometric Matri...
Metabolites
Gene Rules
Gene Rules
search :
Case sensitive
Whole word
Save
Options
Reaction Name
Gene Rule
R_SUCDi (((b0721 and b0722) and b0723) and b0724)
R_EX_fru_e
R_FRD7 (((b4151 and b4152) and b4153) and b4154)
R_NADH16 (((((((((((b2276 and b2277) and b2278) and b2279) and b2280) and b2281) and b2282) and b228...
R_TPI b3919
R_PTAr (b2297 or b2458)
R_PPCK b3403
R_GLNabc ((b0811 and b0810) and b0809)
R_ATPM
R_PGI b4025
R_Biomass_Ecoli_core_w_GAM
R_EX_nh4_e
R_PGK b2926
R_MALt2_2 b3528
R_PGM ((b3612 or b4395) or b0755)
R_ACALDt s0001
R_PGL b0767
R_H2Ot (b0875 or s0001)
R_EX_acald_e
R_ATPS4r ((((b3736 and b3737) and b3738) and ((((b3731 and b3732) and b3733) and b3734) and b3735)) or...
Gene Rules

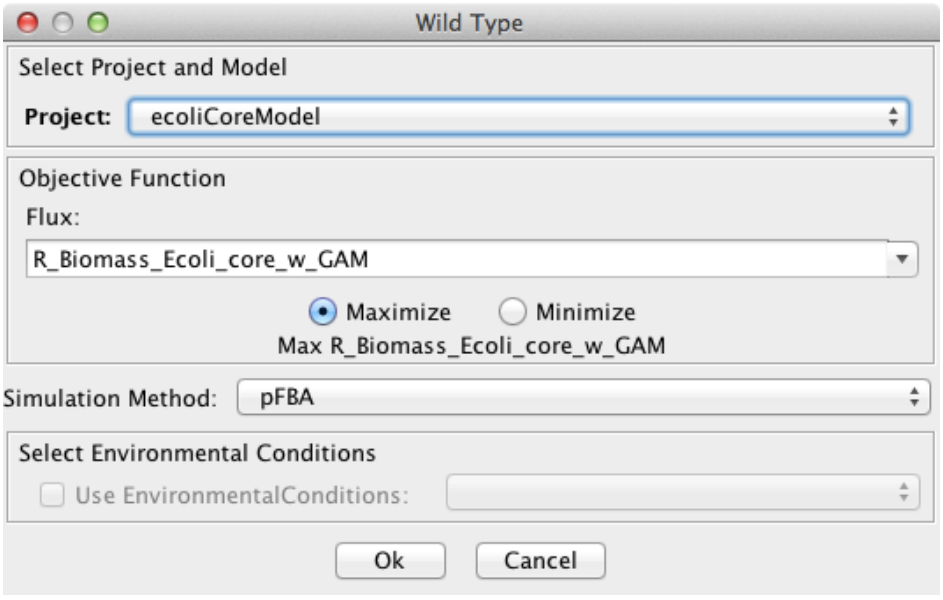
2. PERFORMING PHENOTYPE SIMULATIONS

OptFlux allows to perform phenotype simulations, i.e. determining the values of the fluxes for the reactions present in the model, given environmental conditions and genetic conditions. The former allow defining fluxes for uptake reactions and, thus, defining the medium where the cell is growing. The latter define genetic modifications over the original strains, allowing to simulate both the wild type (no genetic modifications are defined) and mutants (where different types of modifications can be imposed including gene deletions and over/underexpression).

2.1 PERFORMING A WILD-TYPE SIMULATION

OptFlux allows the user to perform a phenotype simulation of the "wild-type" strain, i.e. of the model with no genetic modifications. Access the "Wild Type" option either through the "Simulation" menu or right clicking on the model icon in the clipboard.

When the operation is launched, a panel is provided to configure its options. You must select the Project to which the simulation will refer. This step is necessary since OptFlux supports multiple-projects and each project contains its own model.



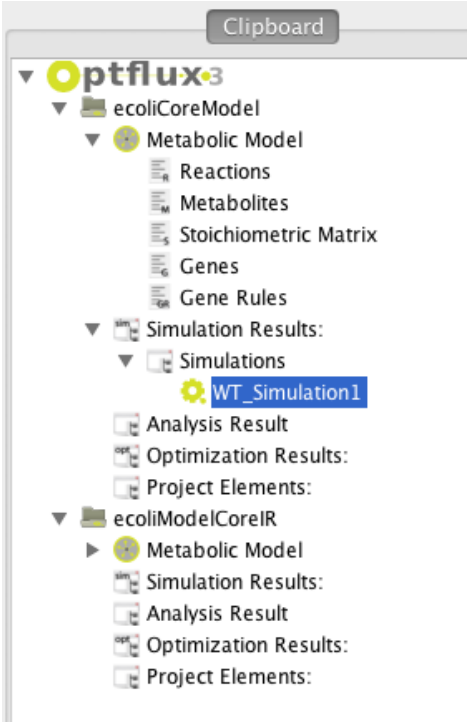
The other options in the panel allow you to define the method to run the simulation and the flux that will be maximized, which is by default filled with the model's biomass reaction, the most natural choice. Also, the sign of the optimization is defined, being by default maximization.

The method used to run the simulation, in this case, is the parsimonious Flux Balance Analysis (pFBA), described in Lewis et al, Omic data from evolved E. coli are consistent

with computed optimal growth from genome-scale models, *Molecular Systems Biology* 2010; 6:390. In the current version, the only alternative would be to use the original Flux Balance Analysis (FBA) method. The main difference between both is that the latter simply maximizes the selected flux, while the former does the same but does a second step where it minimizes the sum of fluxes constraining the target flux to its maximum value.

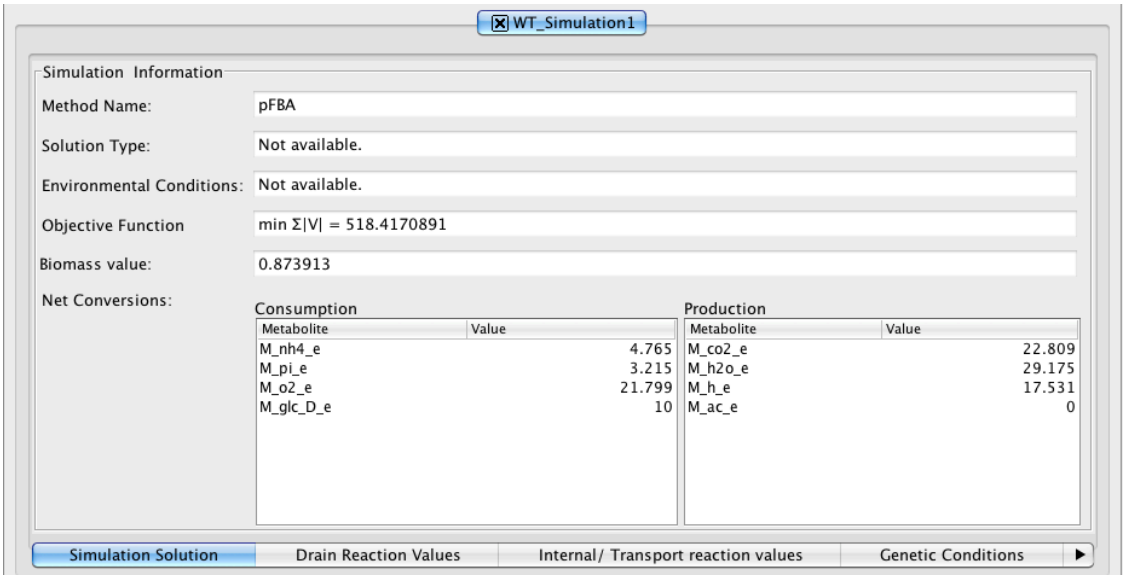
In the bottom, the panel allows for the selection of an environmental condition that can be used for instance to define media conditions different from the ones included in the model. This is described in the next section of this tutorial and, at this stage, we will leave this not selected (as default). In this case, we assume that all uptake reactions have their limits as defined in the original model.

After completing all the previous steps, a new object will be placed in the clipboard within the "Simulation Results" list, under "Simulations", with the name "WT_Simulation1" (this can be changed by right clicking the object and renaming it).

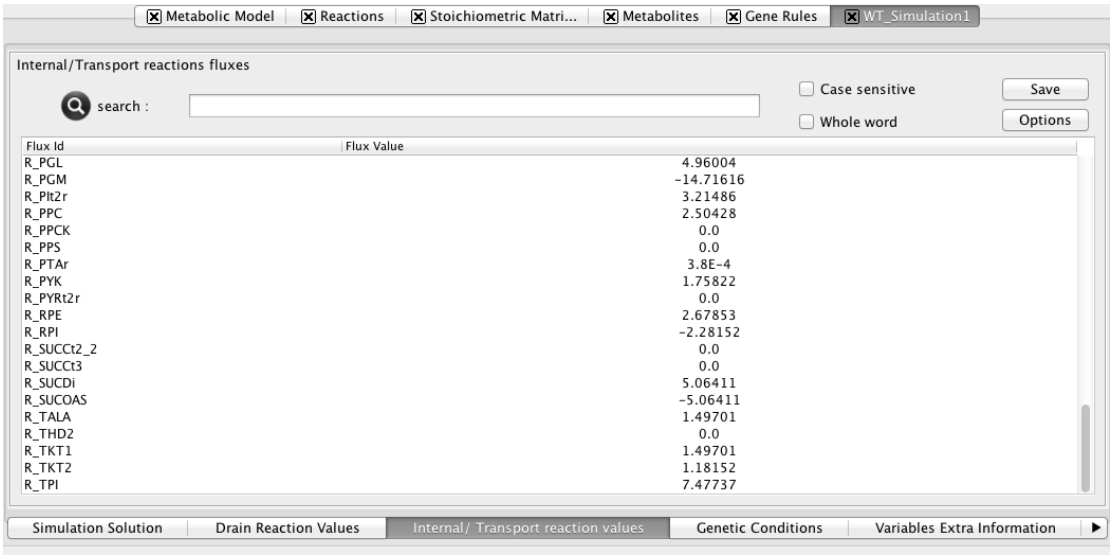


By left-clicking this object, you have access to detailed information about the performed simulation, organized in several tabs:

- **Simulation solution**: presents the simulation configuration, the value of the objective function, the net conversion, etc. (as shown below)



- **Drain reaction values and Internal/ transport reaction values**: these tabs allow to check the flux values returned by the simulation for all the fluxes, divided into two groups of reactions: internal and transport on the one hand, and drains (i.e. metabolite uptake or excretion fluxes) on the other; these values can be easily exported in a table format (e.g. CSV). The screenshot shows part of the flux values for internal reactions.

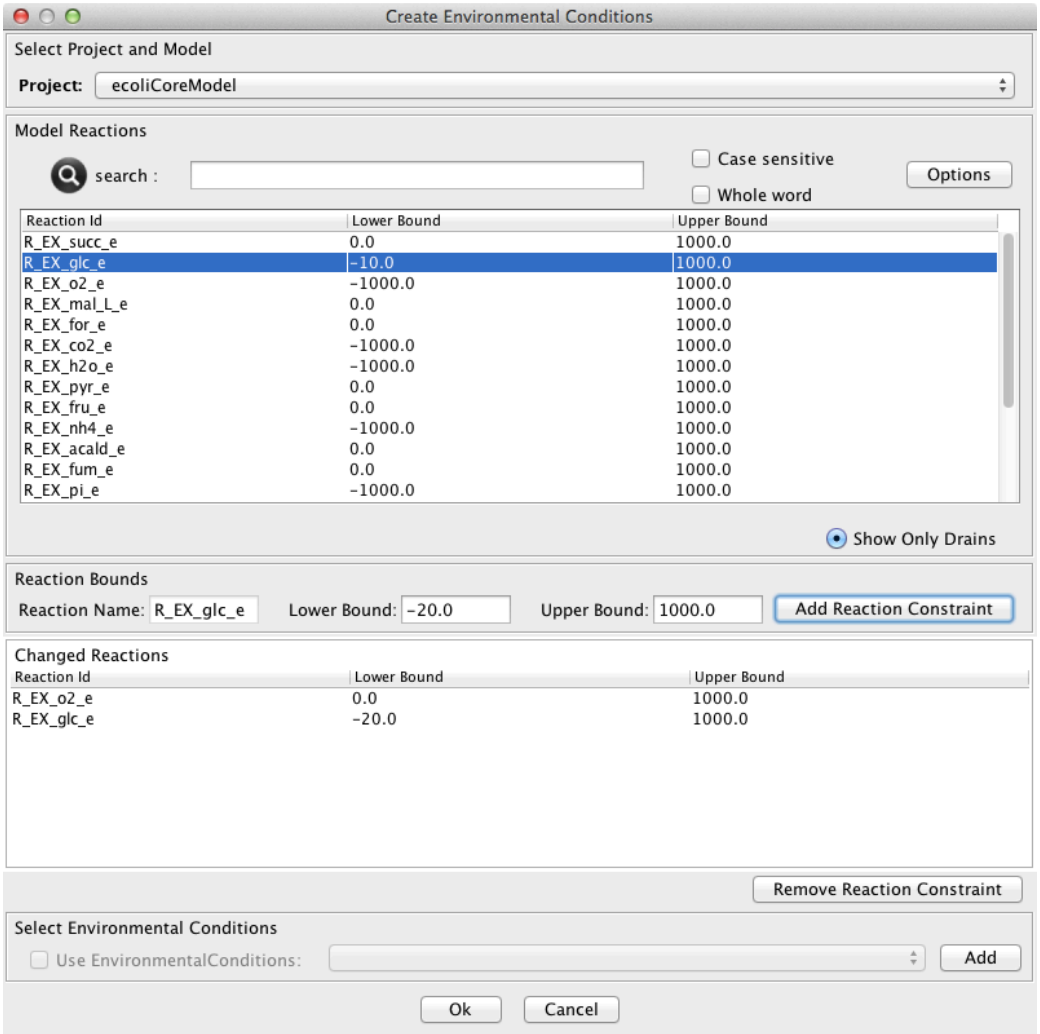


- **Genetic conditions**: allows to check the genetic conditions (including knockouts, over-under expression) applied; in this case, since it is a wild type simulation, no genetic conditions were defined.
- **Restriction and variables extra information**: allows to check complementary information (if available) returned by the simulation method (this information can change from method to method).
- **Solver output**: textual output of the solver used in the optimization.

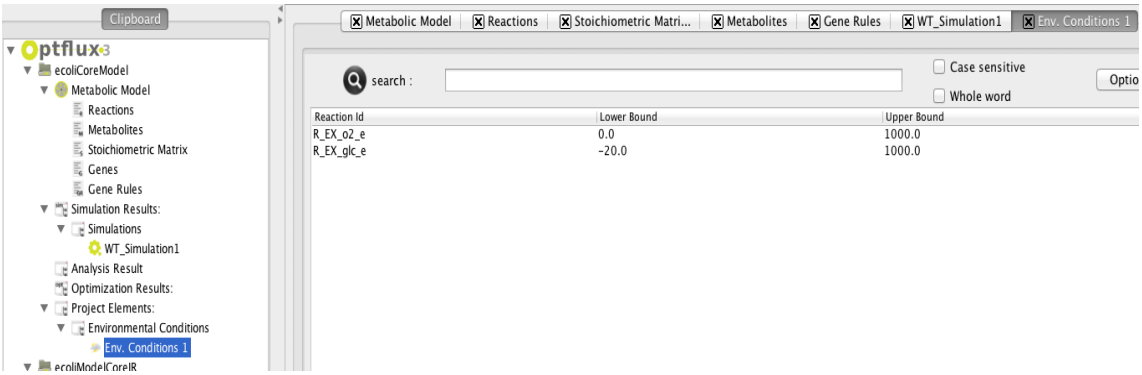
2.2 CREATING ENVIRONMENTAL CONDITIONS

Environmental conditions (ECs) can be used to define the values of uptake fluxes, i.e. the rates at which metabolites are consumed and, thus, can be used to define media for the growth of the cells. They can also be used to impose other constraints over the upper and lower limits of drains or internal/ transport reactions that arise from existing knowledge about the simulation conditions.

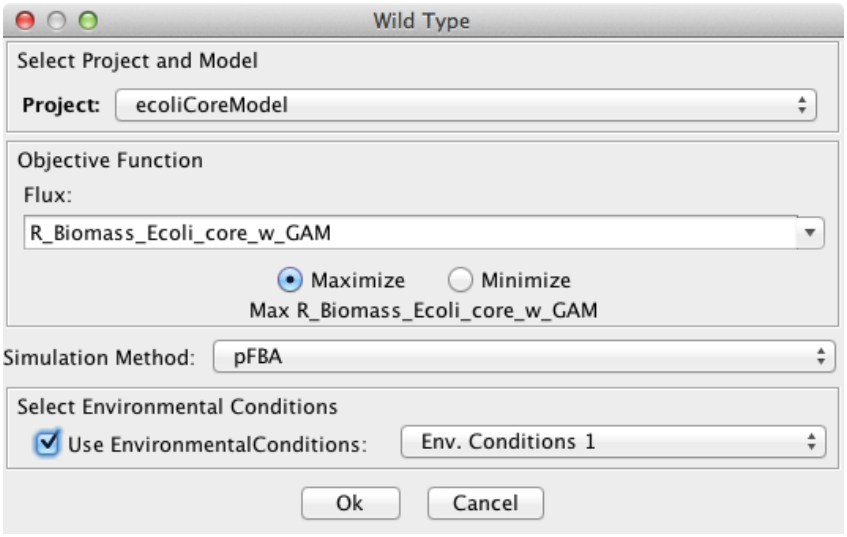
ECs are created in the option "File/ Create Environmental Conditions", as shown in the screenshot below. The interface allows the user to select drain reactions and define their limits. The upper part allows the selection of the reactions (easing its search), while the bottom part allows to add these constraints to the EC, specifying the lower and upper bounds. The list in the bottom part of the interface shows the current reactions and their limits, as they are added to the EC. In the example, the EC will define anaerobic conditions (since the lower limit for oxygen uptake is 0) and the maximum uptake for glucose is set to 20, instead of the original 10.



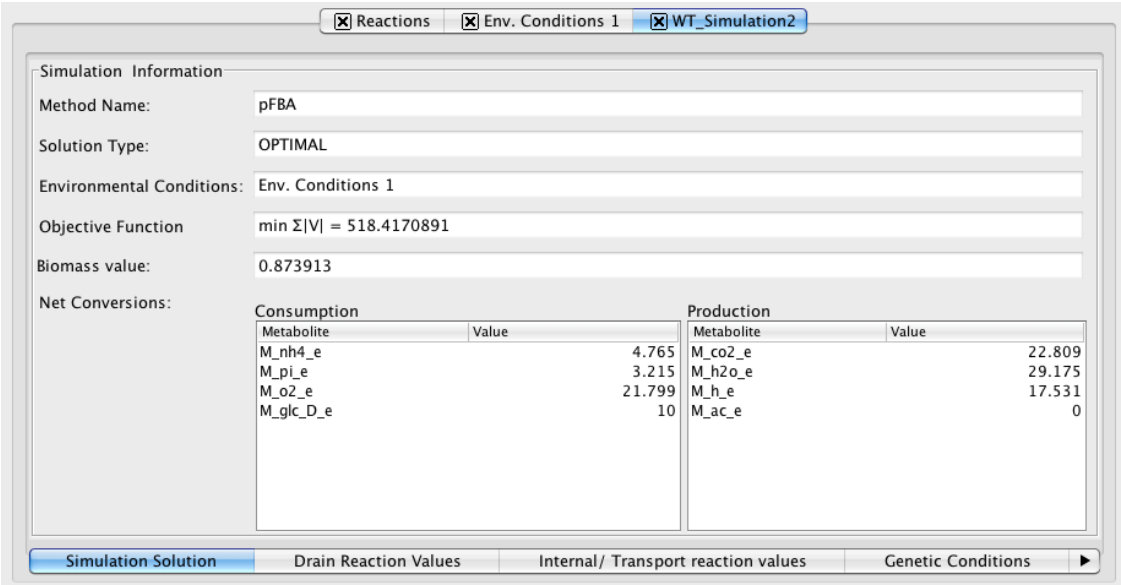
This EC is kept in the clipboard under “Project Elements/ Environmental Conditions”. The figure shows the object created in the previous operation:



ECs can be used to perform simulations. For instance, you can use the EC created above to do a different wild type simulation:



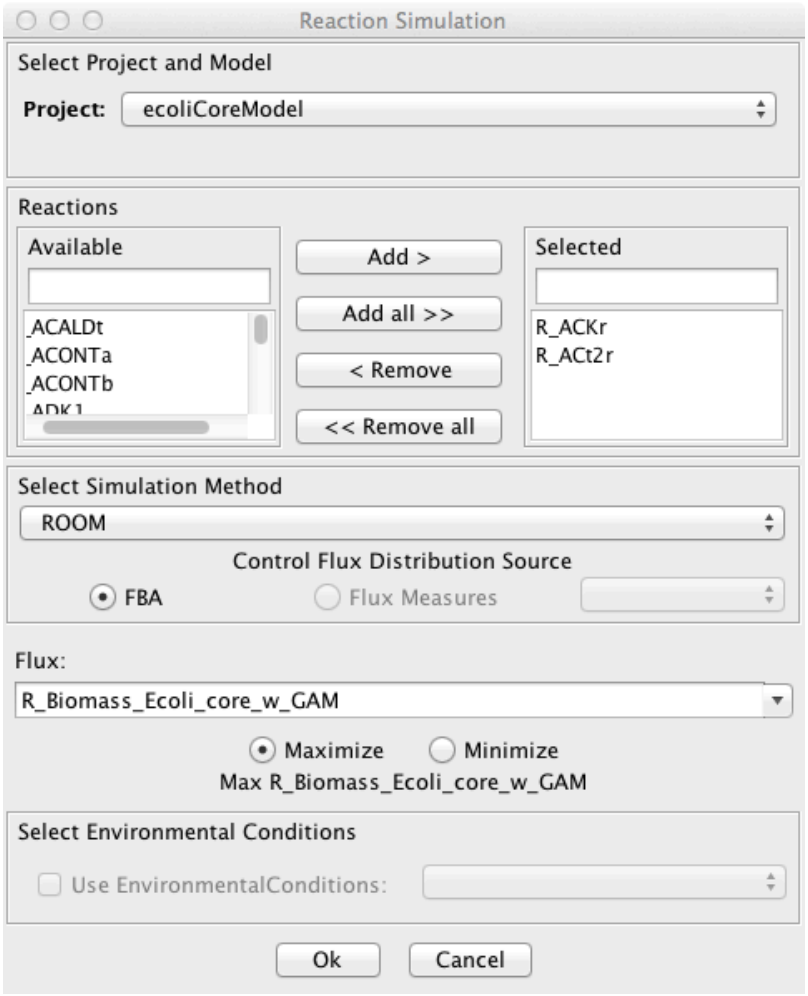
that will return the following result:



2.3 PERFORMING A MUTANT SIMULATION – REACTION DELETIONS

One important feature of OptFlux is the ability to simulate mutant strains. We will start by checking how to simulate the case where certain reactions are removed from the model, i.e. their flux is forced to be zero. You can access the "Reaction Simulation" option under the "Simulation/ Knockout" menu or right clicking on the Metabolic Model icon on the clipboard.

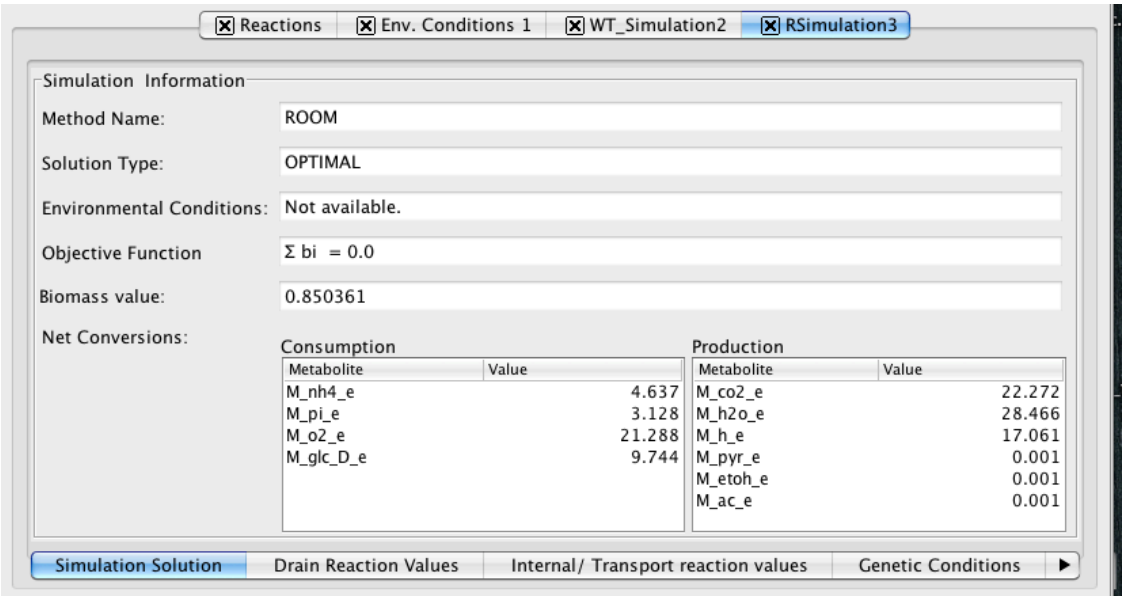
When launching the operation, the panel allows its configuration by defining the following parameters (see screenshot below):

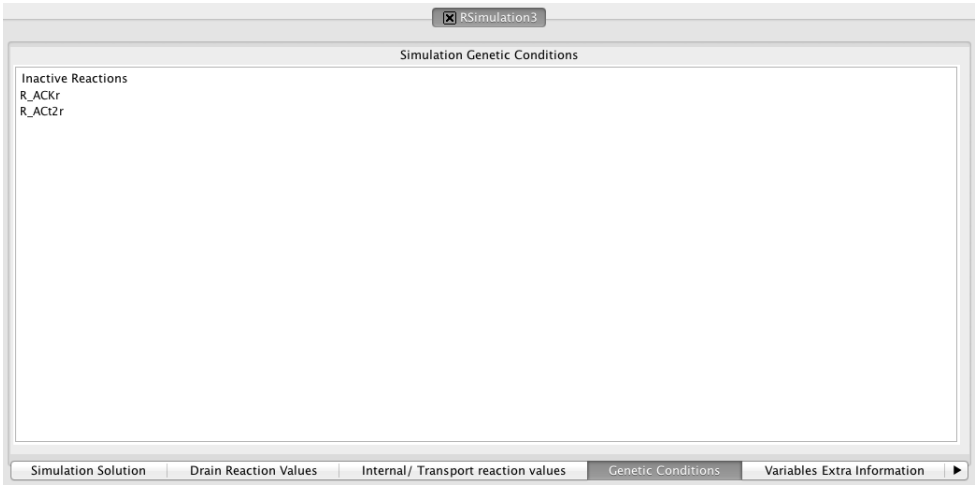


- **Project**: allows to select which project to use and, consequently, the model it contains will be used to support the simulation.
- **Reaction knockout list**: selecting in the Reaction list you can add/remove (using the arrows buttons) reactions to/from the knockout list (the list of reactions to be deleted, shown on the right box).
- **Simulation Method:** OptFlux can use several simulation methods for knockout simulations, namely: Flux-Balance Analysis (FBA), parsimonious FBA (pFBA), ROOM, MOMA, Linear MOMA. FBA and pFBA use a linear programming (LP) formulation. ROOM stands for the Regulatory On-Off Minimization Method (ROOM) and uses a Mixed Integer Linear Programming (MILP) formulation as proposed by Shlomi et at, Regulatory on/off minimization of metabolic flux changes after genetic perturbations, PNAS 2005; MOMA stands for the Minimization of Metabolic Adjustment method that uses quadratic programming, proposed by Segré et al, Analysis of optimality in natural and perturbed metabolic networks, PNAS 2002. Some of these methods require the definition of a reference (control) distribution of fluxes that can be provided by the user or calculated using one of the methods for wild type simulation.
- **Objective Function Configuration –** Here, you can select the reaction to optimize (biomass, by default), and you can also define if you will be maximizing or minimizing the flux (as in the wild type simulation).

- **Environmental Conditions (ECs) -** If you have created environmental conditions you can select one to be used in the simulation.

In this case, the method selected was ROOM, no ECs were chosen and the list of reaction deletions includes “R_ACKr” and “R_Act2r”. You can press OK and the results will be loaded into the clipboard, as was the case with the wild type simulation (i.e. under Simulation Results/ Simulations). In this case, the name is something like “RSimulation3” and the structure will be similar to the one presented above for the wild type. Indeed, this is an instance of the same data type as above. The main difference will lie on the content of the “Genetic Conditions tab”, as shown below.





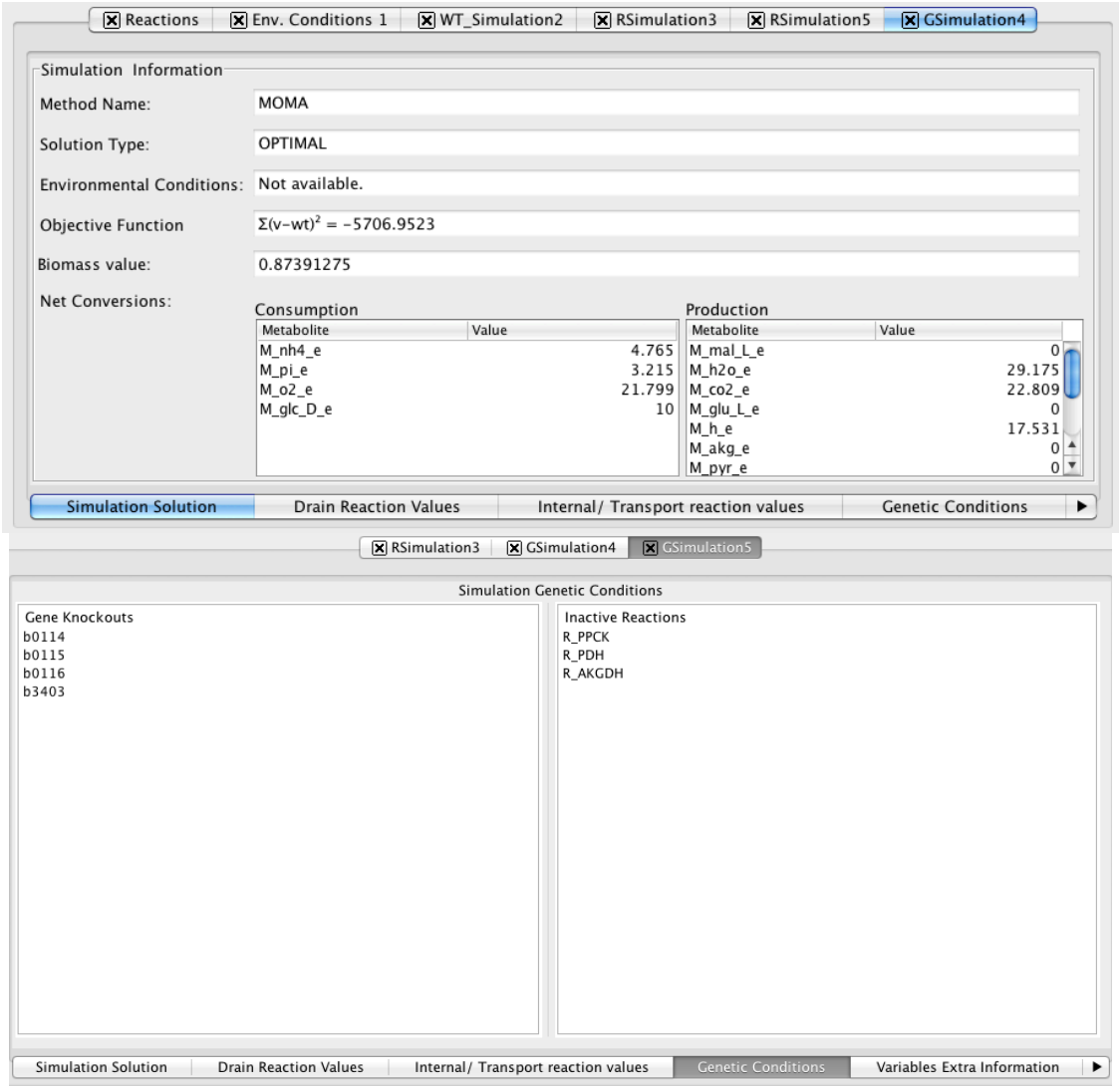
2.4 PERFORMING A MUTANT SIMULATION – GENE KNOCKOUT

If the model contains information on genes and gene-reaction associations, OptFlux allows the user to perform a simulation of mutant strains with gene knockouts. This option is available through “Simulation/ Knockout/ Gene simulation”.

The configuration of this operation is very similar to the one mentioned in the previous section, and therefore only the differences are highlighted. Indeed, in this case, you just need to select a set of genes to knock out. In the user interface, you can add/remove (using the arrows buttons) genes to/from the knockout list (the list of genes to be knocked out, in the middle box). On the right hand side, the list of reactions that will be deleted from the model as a result of the selected gene knockouts is shown.

The remaining configuration steps are equal to the ones involved for reactions, as it can be checked in the screenshot below, where the set of gene knockouts includes the genes “b0114”, “b0115”, “b0116” and “b3403”. These genes will lead to the removal of the reactions: “R_PPCK”, “R_PDH” and “R_AKGDH”.

Also, the results will be available in the same location within the clipboard, with names starting with "GSimulation" followed by a number (as before, these objects can be renamed). Again, it should be mentioned that these are instances of the same data type as in the two previous cases of phenotype simulation results. Two of the result tabs are shown below. Note that the Genetic Conditions tab includes information on the genes and reactions inactivated.



2.5 PERFORMING SIMULATIONS WITH OVER/ UNDER EXPRESSION

OptFlux allows the simulation of mutants including over/under expression of genes or reactions. The method used for the simulation is described in detail in Gonçalves et al, Optimization approaches for the in silico discovery of optimal targets for gene over/underexpression. *Journal Computational Biology* 2011. These operations can be called in the "Simulation / Under Over Expression" menu, with sub-menus for "Gene

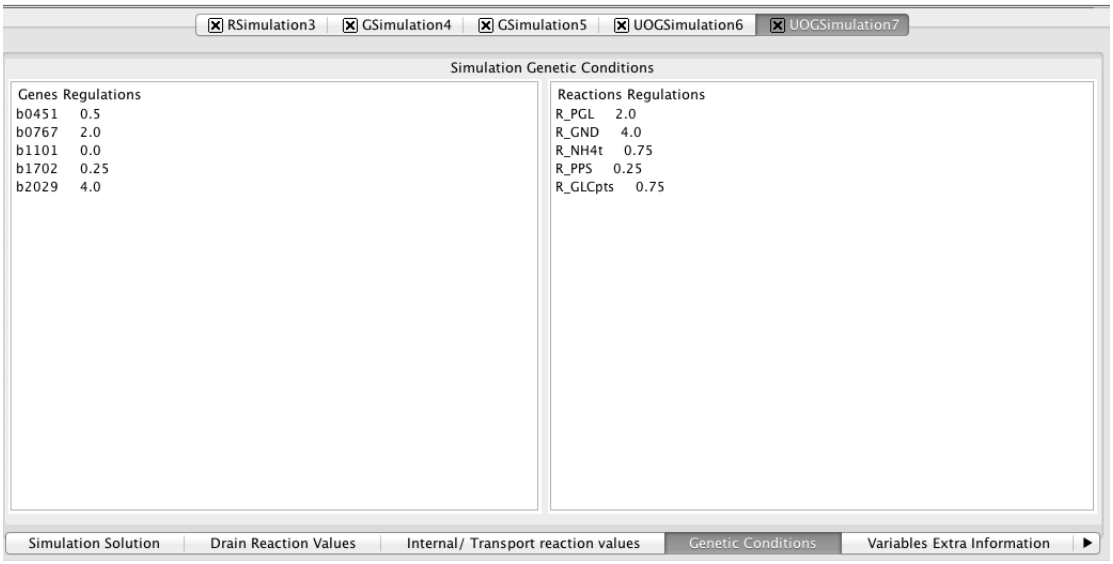
simulation" and "Reaction simulation". The user interface for these operations is similar to the ones for reaction and gene deletions, shown in the previous sections.

The main difference, in this case, lies in the need to define a value for relative expression for each gene or reaction selected. In this case, values over 1 represent overexpression, i.e. the value of the flux (in reaction simulation), or of the gene expression (in gene simulation), are constrained to values greater than the wild type by the defined factor. In the case this value is less than 1, an under-expression is defined, i.e the flux/ expression value is constrained to be smaller than the wild type. It should be mentioned that knockouts are represented by the value 0, while the value 1 represents no changes in the flux/ expression value. All genes/ reactions not included in the list are not constrained (i.e. are only constrained by the model bounds).

Let us look at an example of a gene simulation, including both overexpressed and underexpressed genes, as well as a gene deletion. As before, the table in the right shows the list of affected reactions, with the respective levels of relative fluxes. To understand these values, you need to look at the gene-reaction rules in the model and also to the method followed, where AND operators between genes are translated to the minimum function when dealing with expression levels, while OR operators are translated to the mean function.

In the example, 5 genes are affected: one knockout (“b1101”), two underexpressed genes (“b0451” and “b1702”) and two overexpressed genes (“b0767” and “b2029”). This will lead to five reactions affected, three of which will be underexpressed and the remaining two overexpressed.

The result for this operation has a structure equal to the one presented before for the knockout simulations and, as such, they are shown in the same list of Simulation Results, with a name starting by “OURSimulation” ou “OUGSimulation”, respectively. The main difference again lies in the genetic conditions tab, as shown below.



3. PERFORMING STRAIN OPTIMIZATION

OptFlux allows running operations of strain optimization using metaheuristic methods from the Evolutionary Computation field. These methods are described in detail in Rocha et al, Natural computation meta-heuristics for the in silico optimization of microbial strains. *BMC Bioinformatics* 2008.

Regarding the types of genetic conditions being optimized, OptFlux allows different variations. Indeed, you can select if genes or reactions will be optimized (assuming the model has gene-reaction associations; if not, only reactions are possible). Also, you can decide to allow over and underexpression or to simply allow gene/ reaction deletions.

OptFlux also allows both single objective and multiobjective optimization, i.e. you can choose a single objective function or combine a number of functions.

In the user interface for the operation, several parameters need to be configured, listed next:

1.Select Project
In the Project combo box select the project where you want to perform the optimization

2.Select Simulation Method
Decide which phenotype simulation method is used to simulate each solution: these include the options mentioned in the context of section 2 (FBA, pFBA, ROOM, MOMA)

***3.*Select Environmental Conditions**
If you have created environmental conditions (see 2.2) you can select one of them to be used in the simulations that are conducted to evaluate each solution of the optimization task.

4.Select Objective Functions
OptFlux can use several types of objective functions including the BPCY - Biomass-Product Coupled Yield, YIELD - Product Yield with Minimum Biomass, the maximization/ minimization of a flux value, the minimization of the sum of the flux values or the minimization of the number of knockouts. Each objective function needs its own configuration, selecting which fluxes are used in the computation of the function value. This configuration is done on the left, while in the right the set of selected functions is listed as they are defined.

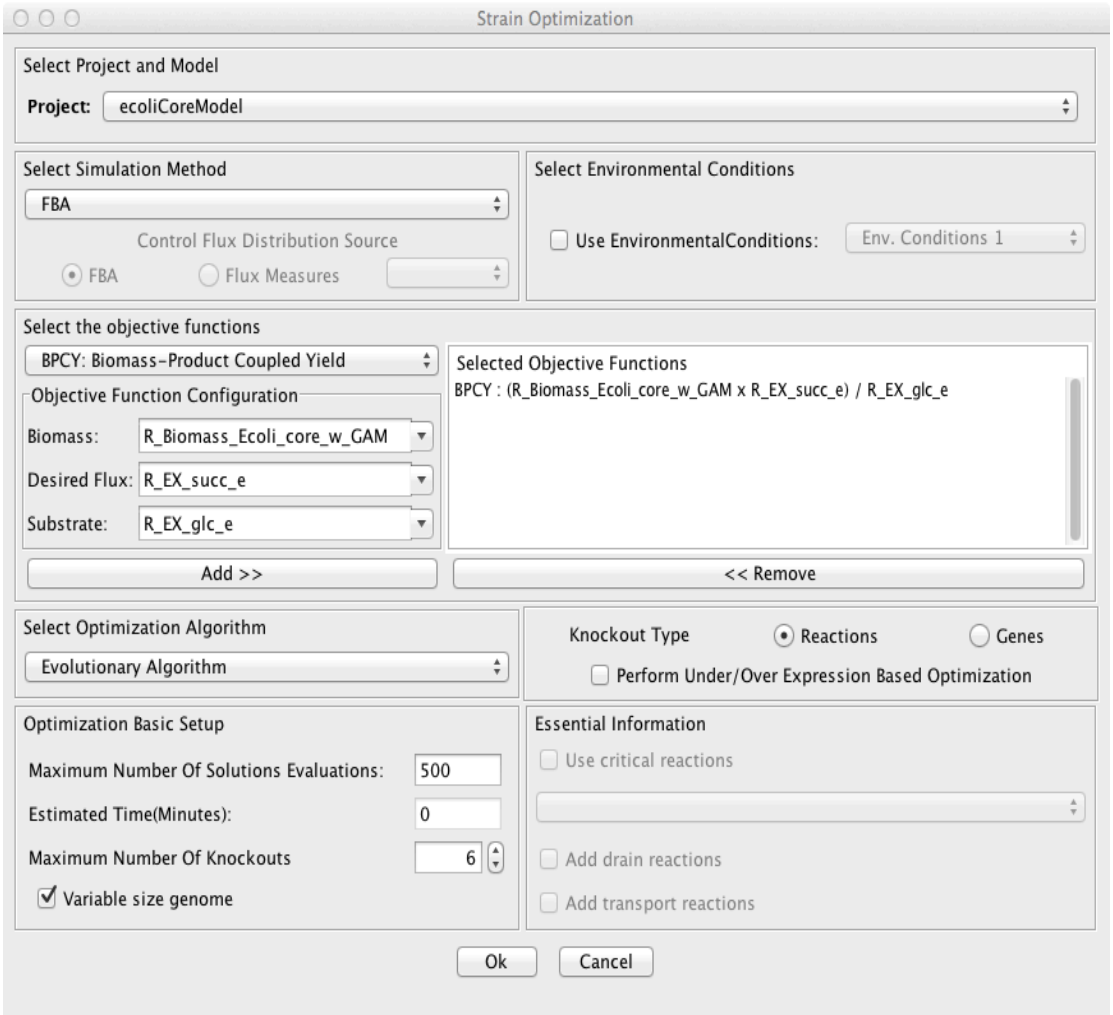
5. Algorithm selection and its configuration
In the left bottom corner, you can select one of the available algorithms: SPEA (Strengh Pareto Evolutionary Algorithm) – a multiobjective approach, and

Evolutionary Algorithms (EA) or Simulated Annealing (SA) – single objective methods. Also, you can select the number of solutions to be evaluated, the maximum number of affected genes or reactions allowed in the solutions, and decide if variable size solutions are allowed.

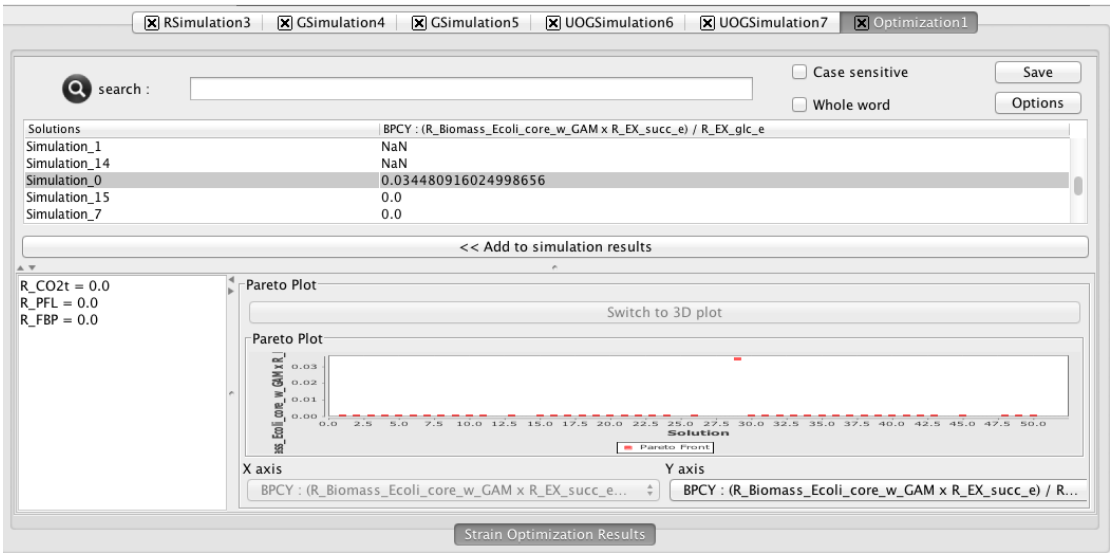
6. Type of optimization

In the right bottom corner, the type of optimization is selected (genes vs reactions, knockouts vs over/under expression). Also, here, the use of essential reactions or genes is defined (see section 4.1). In the case these are defined, the essential genes or reactions will not be used in the solutions for the strain optimization.

In the screenshot below, an example is shown, where an optimization process using EAs is configured, using the BPCY as a single objective function (used to maximize the production of succinate, while keeping a high value of the biomass flux and considering glucose as the substrate). The simulation of the solutions uses FBA. The process will look for sets of reaction deletions with a maximum of 6 reactions (but that can vary in size). The algorithm stops after 500 simulations.



The results of the optimization process are shown in the clipboard under "Optimization results". The results for this example are shown below.



The best solution found in this case is highlighted, having a BPCY value of 0.034.

4. EXPLORING ANALYSIS TOOLS

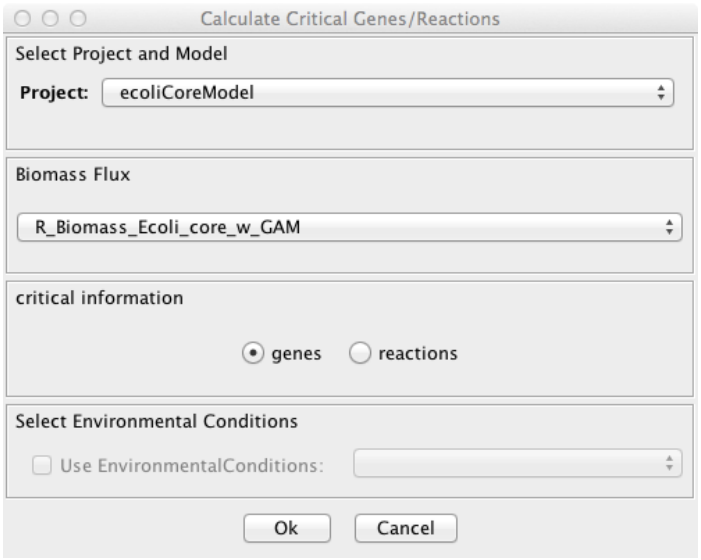
Currently, the core distribution of OptFlux 3 includes two types of analysis: determining critical (or essential) genes/ reactions and flux variability analysis. These will be covered in the next sections.

4.1 CRITICAL GENES/ REACTIONS

The determination of critical (or essential) genes (or reactions) is performed with the following rationale: a gene (reaction) is considered critical if when it is removed from the model, this leads to a value of biomass flux of less than 5% of the reference value calculated for the wild type strain (i.e. simulating the model with the full set of genes). The method used for all simulations in this case is FBA.

This operation is available in the "Analysis" menu, "Calculate Critical Genes/ Reactions" submenu. You need to choose the project to apply the operation, the biomass flux to be used (by default the biomass flux kept in the model in selected), if critical genes or reactions will be calculated and if a given environmental condition is used.

An example is provided in the following screenshot, where the model from the previous sections is used and genes are selected, using the default biomass flux.



The result is kept in the clipboard under "Analysis results/ Critical genes" as shown in the next figure.

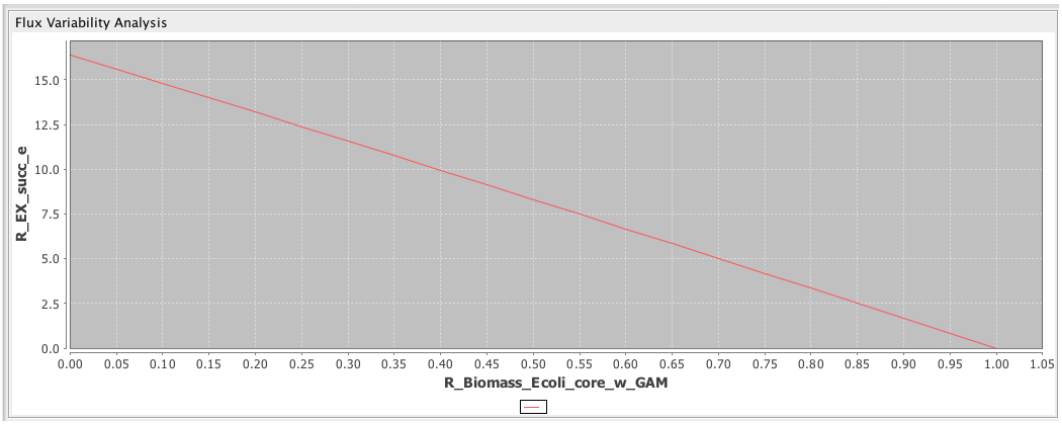
These results include the set of calculated critical genes. This interface also allows the manual addition/ removal of genes from this list, through the add / remove buttons, allowing the user to introduce genes that are known to be essential, although this was not predicted by the model, or to remove genes known not to be essential. Also, the full set of critical genes or reactions can be imported from a file in the "File/ Import" menu.

The sets of genes/ reactions created by this tool can be used to reduce the search spaces of the strain optimization methods presented in the previous section.

4.2 FLUX VARIABILITY ANALYSIS

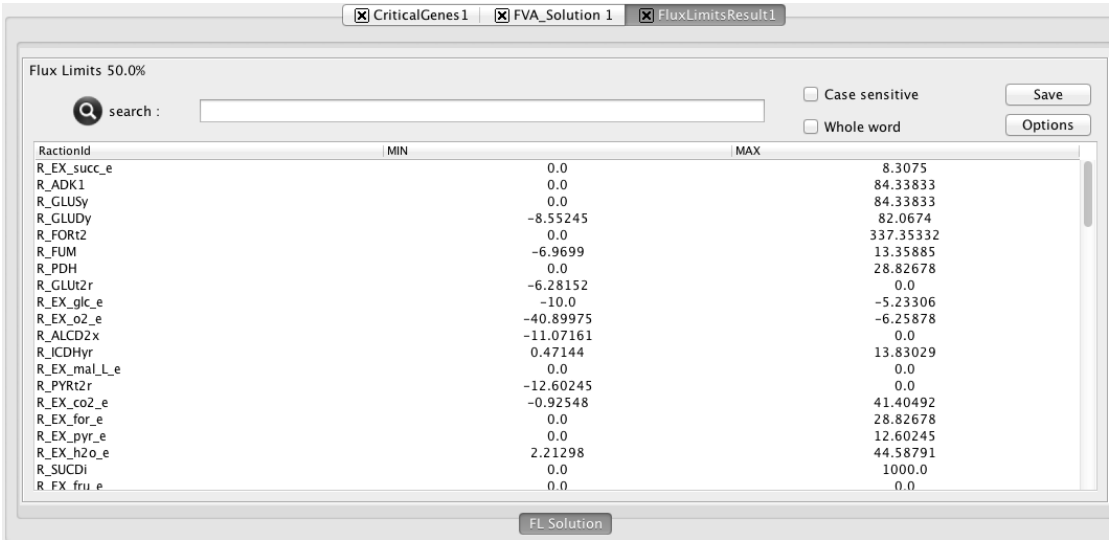
The operations of the Flux Variability Analysis menu allow to perform some studies of the way fluxes can change under some conditions. Two distinct analyses can be addressed: flux variation graphs and calculation of flux limits.

The first allows to study the maximum allowed value for a given flux, imposing constraints defining different levels of minimum biomass. An example is shown in the following screenshot, representing the results of this operation when applied to the flux of secretion of succinate (*R_EX_succ_e*), defining different levels of biomass taking as reference the wild type simulation (the values are percentages of wild type biomass value, from 0% to 100% with a step of 5%). The result shows in the y-axis the maximum possible value of the succinate flux, given the imposed level of the biomass flux (x-axis).



On the other hand, the latter allows to calculate the maximum and minimum possible values of the whole set of reactions for the model, given a defined (fixed) level of the biomass flux (relative to the wild type). In the configuration of the operation, you need only to define the project, the minimum biomass level (in percentage of the wild type value) and if an environmental condition is used. If a value of zero is selected for the biomass, this can be used to define limits for the variation of the fluxes for all reactions (also known as tight bounds).

The result of the previous method, for the same model used above, and using a biomass level of 50% is shown below.



RactionId	MIN	MAX
R_EX_succ_e	0.0	8.3075
R_ADK1	0.0	84.33833
R_GLUSy	0.0	84.33833
R_GLUDy	-8.55245	82.0674
R_FORt2	0.0	337.35332
R_FUM	-6.9699	13.35885
R_PDH	0.0	28.82678
R_GLUt2r	-6.28152	0.0
R_EX_glc_e	-10.0	-5.23306
R_EX_o2_e	-40.89975	-6.25878
R_ALCD2x	-11.07161	0.0
R_ICDHyr	0.47144	13.83029
R_EX_mal_L_e	0.0	0.0
R_PYRt2r	-12.60245	0.0
R_EX_co2_e	-0.92548	41.40492
R_EX_for_e	0.0	28.82678
R_EX_pyr_e	0.0	12.60245
R_EX_h2o_e	2.21298	44.58791
R_SUCDi	0.0	1000.0
R_EX_fru_e	0.0	0.0

5. VISUALIZATION

OptFlux allows the visualization of models (or parts of models) using an internal tool. Currently, this tool is already pre-installed and supports layouts from different formats. Basic information about the functionalities of the plugin can be found in the Optflux's wiki page for the visualization plugin.

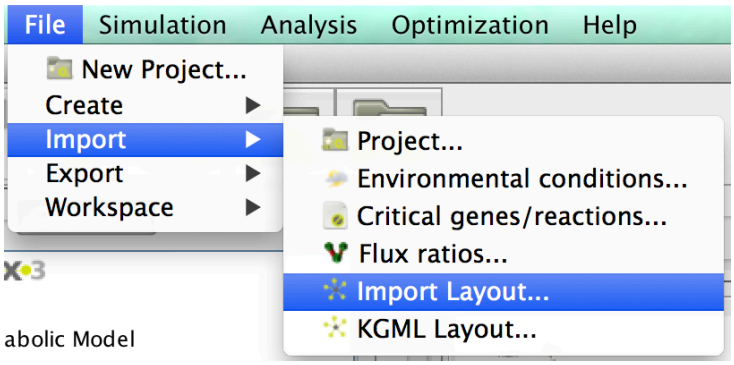
To show these features, two different use cases of the visualization plugin will be displayed. All the necessary materials for these use cases are available at:
http://www.optflux.org/suppmaterial/visualization/materials.zip

5.1 SUCCINATE PRODUCTION WITH *E. COLI*

In the first example, we address the visualization of genetic modifications to *E. coli* that improve the production of succinate.

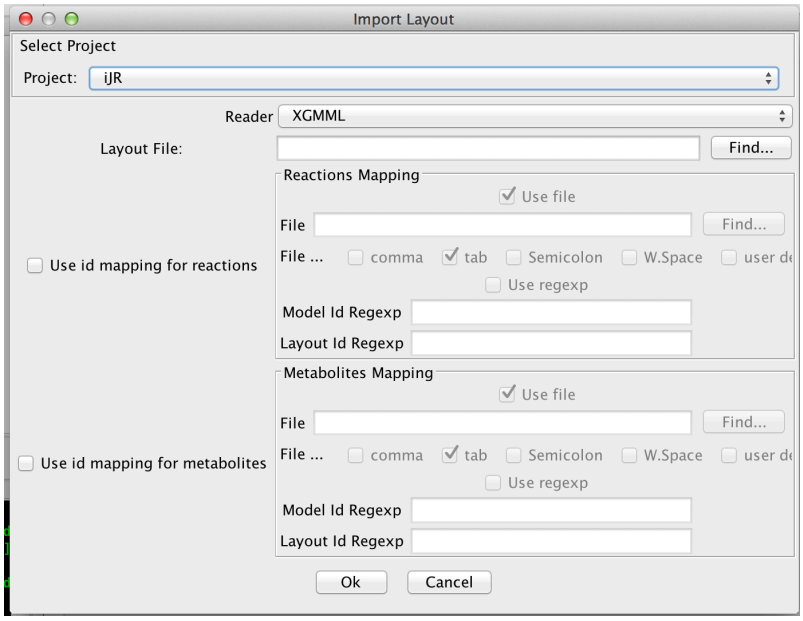
Create a new project in Optflux using the *E. coli* model from the SBML file *"iJR904.xml"* available in the "*materials/models*" folder (follow what was explained in section 1.1 above). More information on how to create a project and load a model is available at the new project wiki page.

To import the layout, select the option *"File -> Import -> Import Layout"*

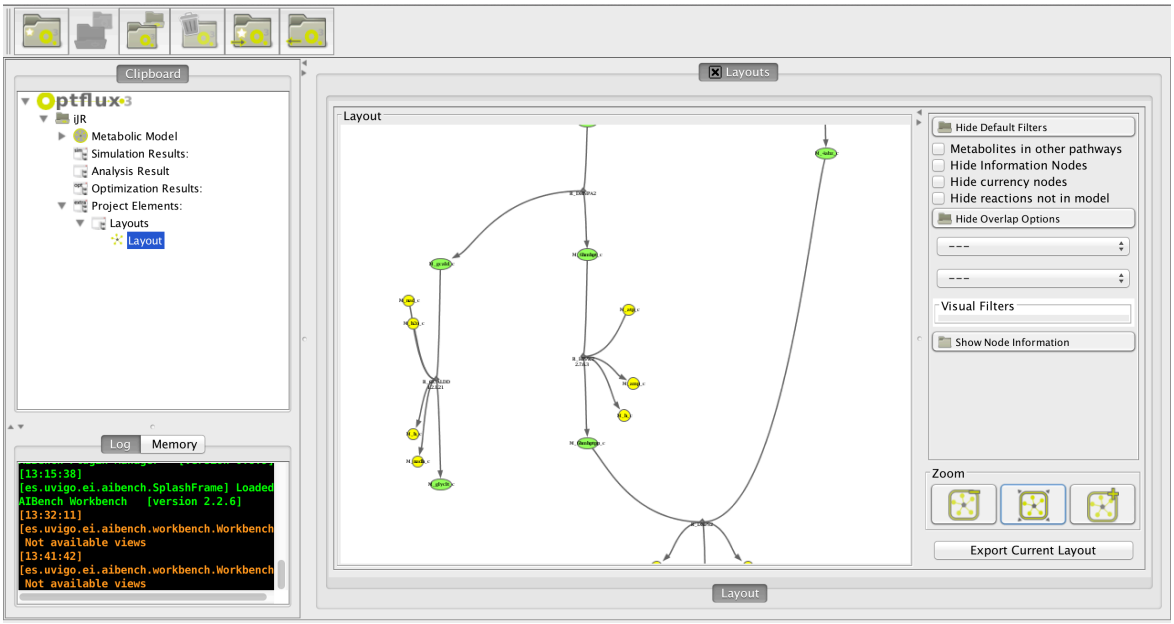


The interface depicted below will appear, where it is possible to select the layout and mapping files for its entities. To learn more about this operation, more information is available at the import layout section of the wiki.

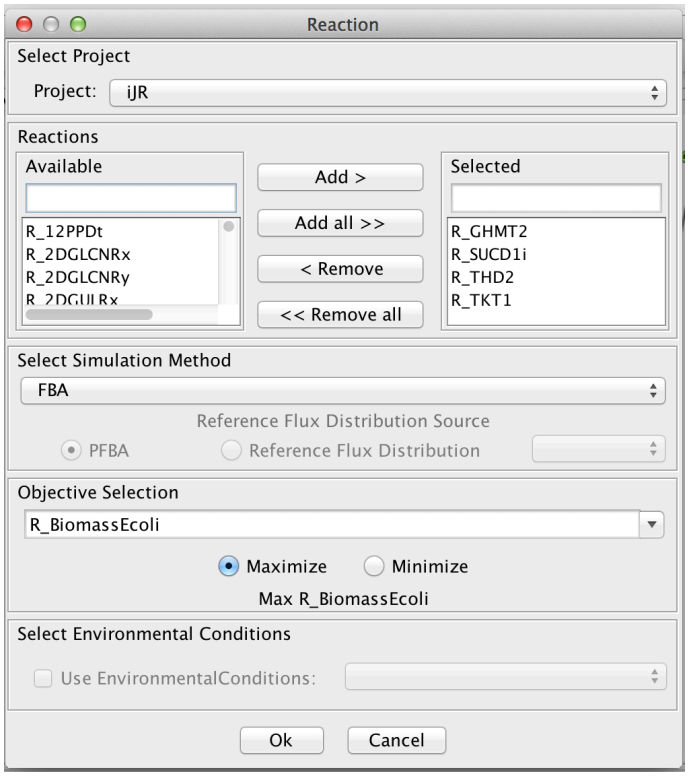
In this case, only the XGMML layout file will be necessary, being available in the folder *"materials/layouts/ijr/ijr.xgmml"*.



After loading the layout, it can be accessed through the clipboard, by clicking the items under the category “Layout”.

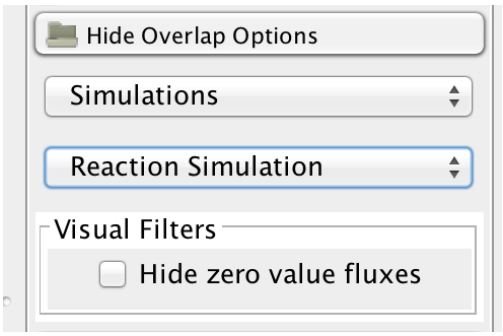


The next step requires performing a reaction knock-out simulation in Optflux as explained above (section 2.3). For this specific simulation select the following knockouts: *R_GHMT2*, *R_SUCD1i*, *R_THD2* and *R_TKT1* as shown in the picture below:



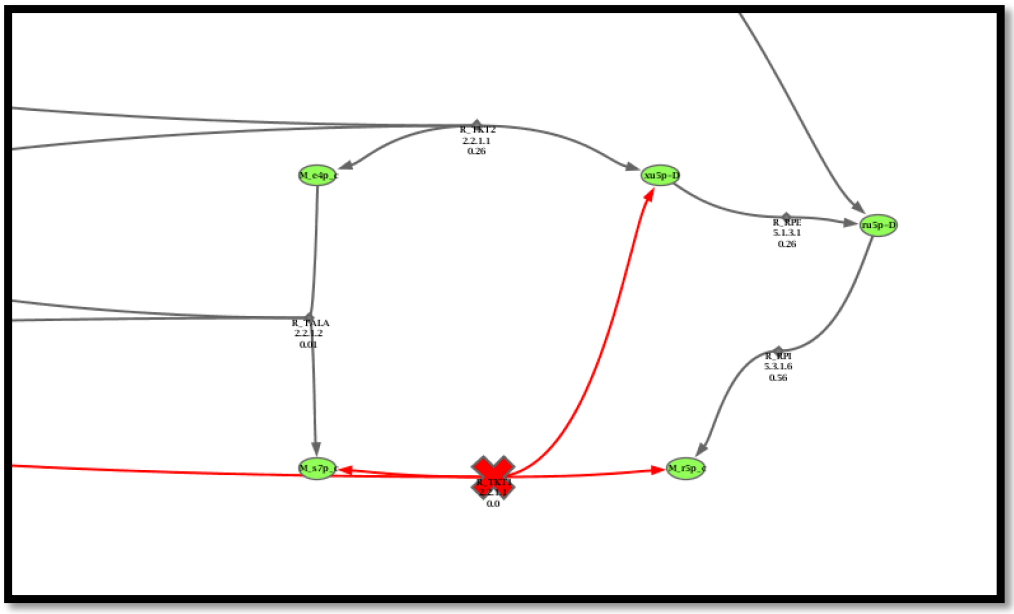
Press OK and the simulation result will be displayed in the clipboard.

In the right panel of the visualization plugin interface, there is a section for the overlaps. As soon as any simulation is performed, an item is added to the overlap panel. From there, you can select which simulation will be overlapped over the layout.



Some overlaps also have visual filters associated. If you tick the “Hide zero value fluxes” filter, all the fluxes from this specific simulation that do not carry any flux will be invisible.

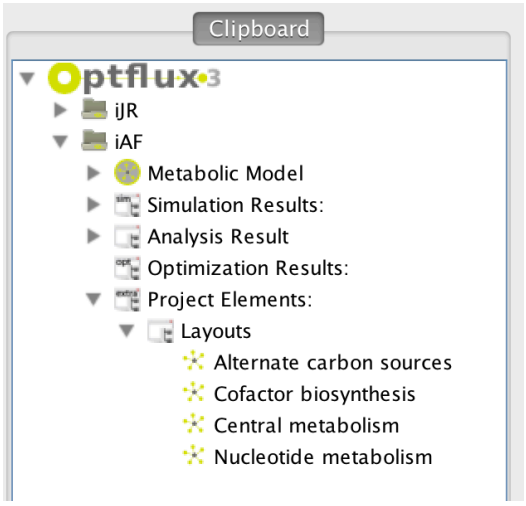
From this point, and overlapping the previously performed simulation, it is possible to visualize the different fluxes and genetic modifications in the network.



5.2 GLYCINE PRODUCTION WITH *E. COLI*

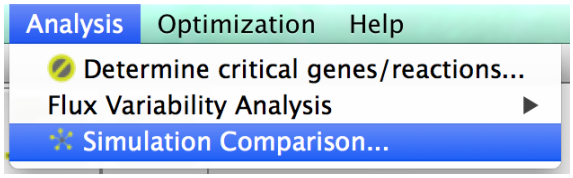
The first step is similar to the previous case. Create a new project in Optflux, using the iAD1260 model, whose SBML file is located at "*materials/models/iAF1260.xml*".

Due to the model size, it is preferable if the layouts represent parts of the model, or pathways. For this example, four layouts will be used. The different pathways are available in the folder *"materials/layouts/pathways/".* Similar to the previous case, load the *XGMML* pathway layouts (central metabolism, alternate carbon sources, co-factor biosynthesis and nucleotide metabolism). It is good practice to change the name of the loaded layouts in the clipboard (Right click -> rename element). By the end of these operations, the clipboard should look like the figure below.

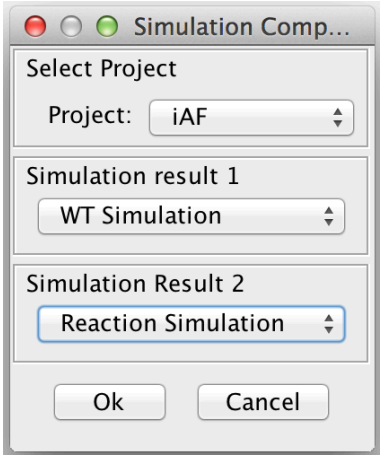


The next step is performing the simulations. First, perform a wild type simulation (see section 2.1 above). After this, and similarly to the previous case, perform a knockout simulation (see section 2.3), using the following reaction knockouts: *R_ICL, R_GLYCL, R_PPC and R_GART*.

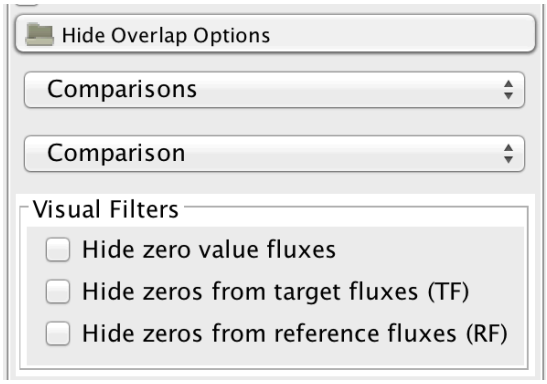
Finally, and to demonstrate the analytical capabilities of the visualization plugin, create a comparison between the wild type simulation and the knockout simulation, by accessing the menu "*Analysis -> Simulation Comparison*".



Select the simulations you want to compare (in this case only the 2 simulations performed are available).



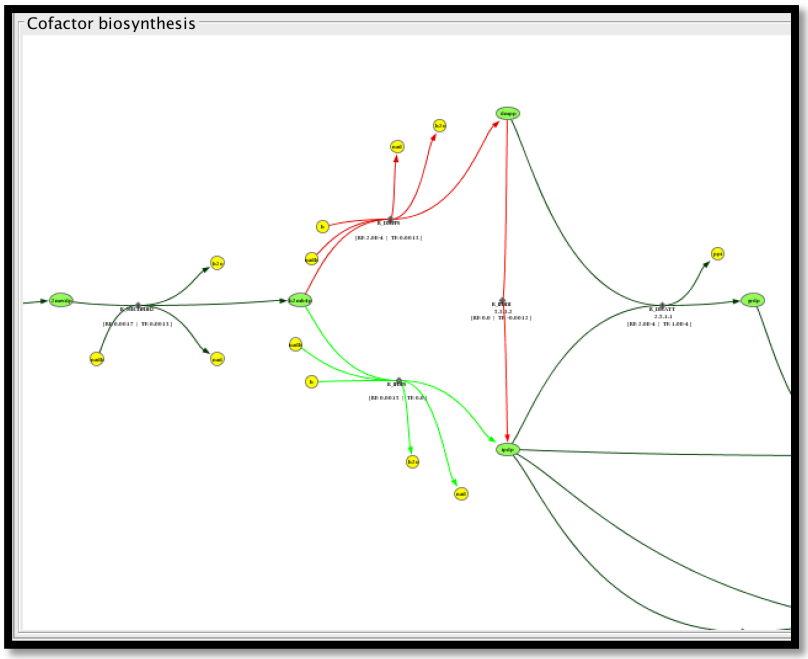
When the comparison is performed, a new overlap will be added to the visualization interface, accessible through the overlap options. This object will be under the category "Comparisons".

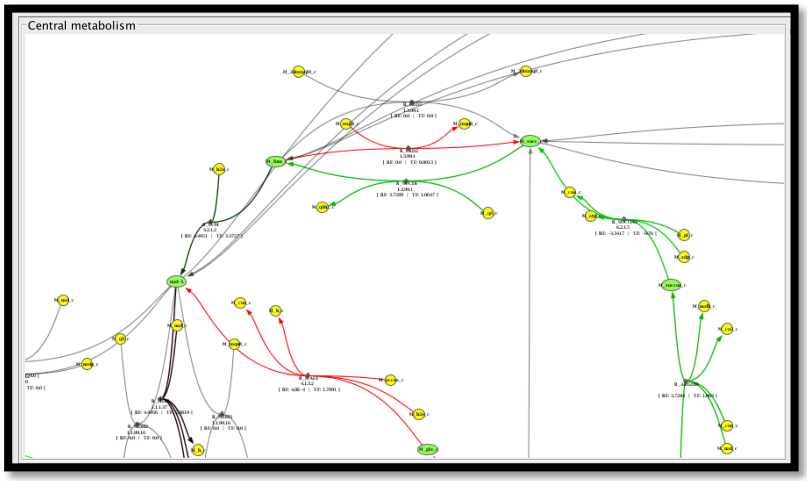


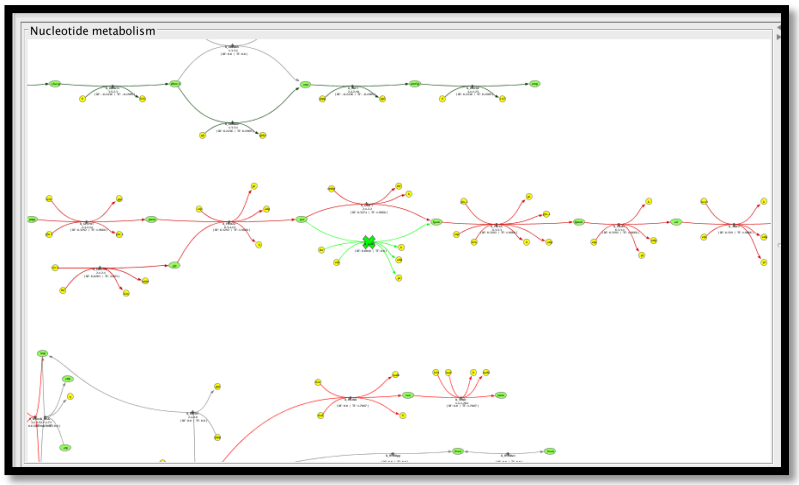
By choosing this overlap, some new filters also become available. In this case, the “Hide zero value fluxes” filter hides all reactions that have a zero flux in both simulations. The other two filters hide zero values specific to each of the simulations, allowing the user to visualize different scenarios (for instance, if the 3 filters are selected, only reactions with fluxes on both simulations will be displayed).

Below are some screenshots of the simulation comparison. More information can be obtained in the wiki page for the overlaps.

Optflux

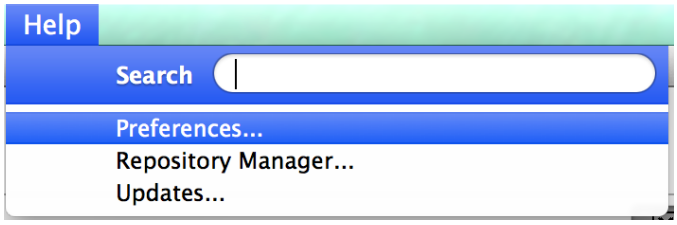

Cofactor biosynthesis


Central metabolism

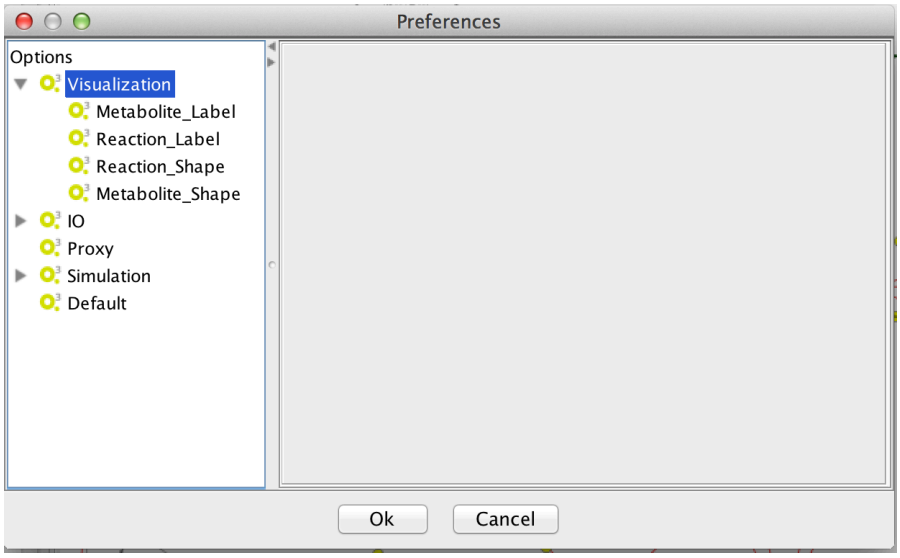

Nucleotide metabolism

5.3 VISUALIZATION PREFERENCES

Another interesting component of the visualization plugin are the *"Visualization preferences"* that can be accessed through the menu *"Help -> Preferences"*.

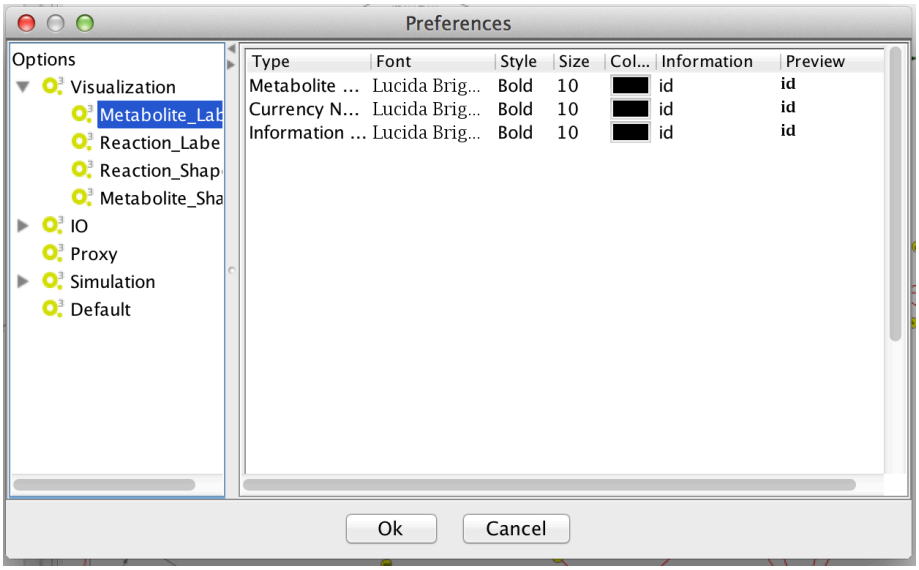


Clicking in the preferences will show a tree, where it is possible to change some of Optflux's default definitions for several components. The visualization preferences are under the branch "Visualization".

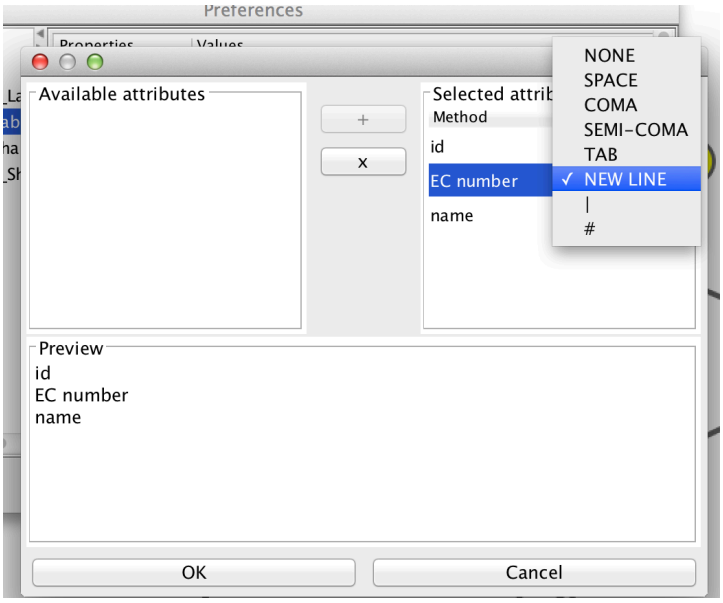


From here, it is possible to edit two main components of the visualization. One of them is the configuration of the labels. It is possible to change the font and color of the labels, both for the metabolite and reaction nodes. Also, and maybe more interestingly, it is also possible to edit the information displayed in these labels.

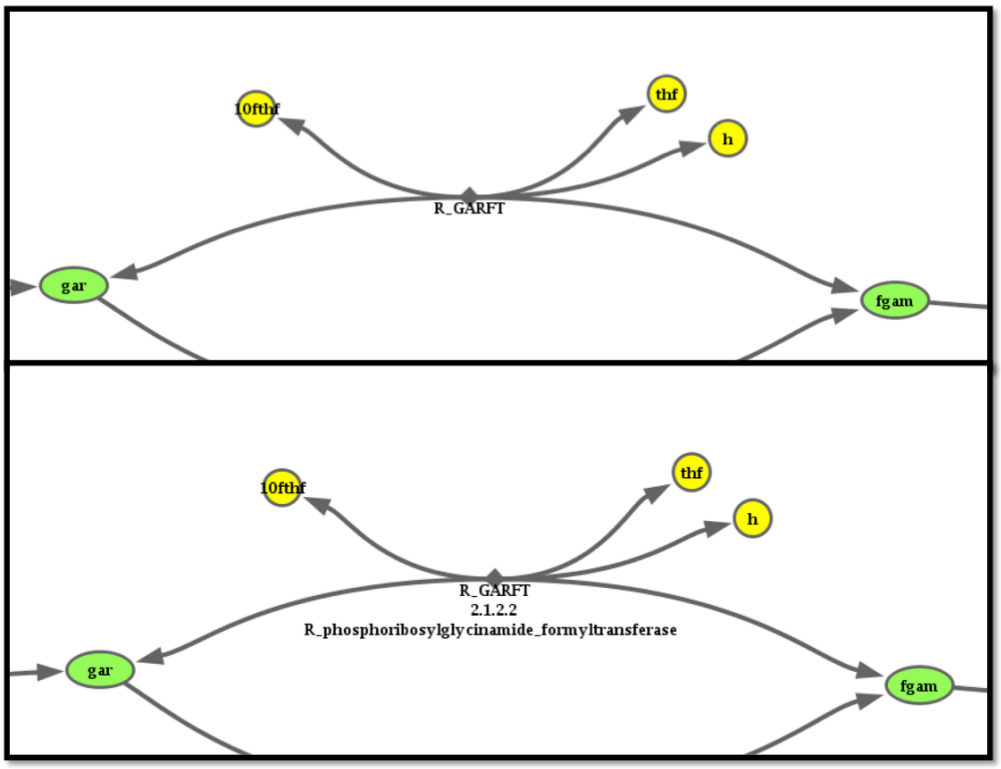
In the respective interface below, it is possible to select which of the metabolite nodes' labels you want to edit. It is possible to change the font, style, size and color of these labels, as well as the content by clicking in the field "Information". These fields are also available for the reactions' labels. Indeed, in the example below the reactions' labels content was modified.



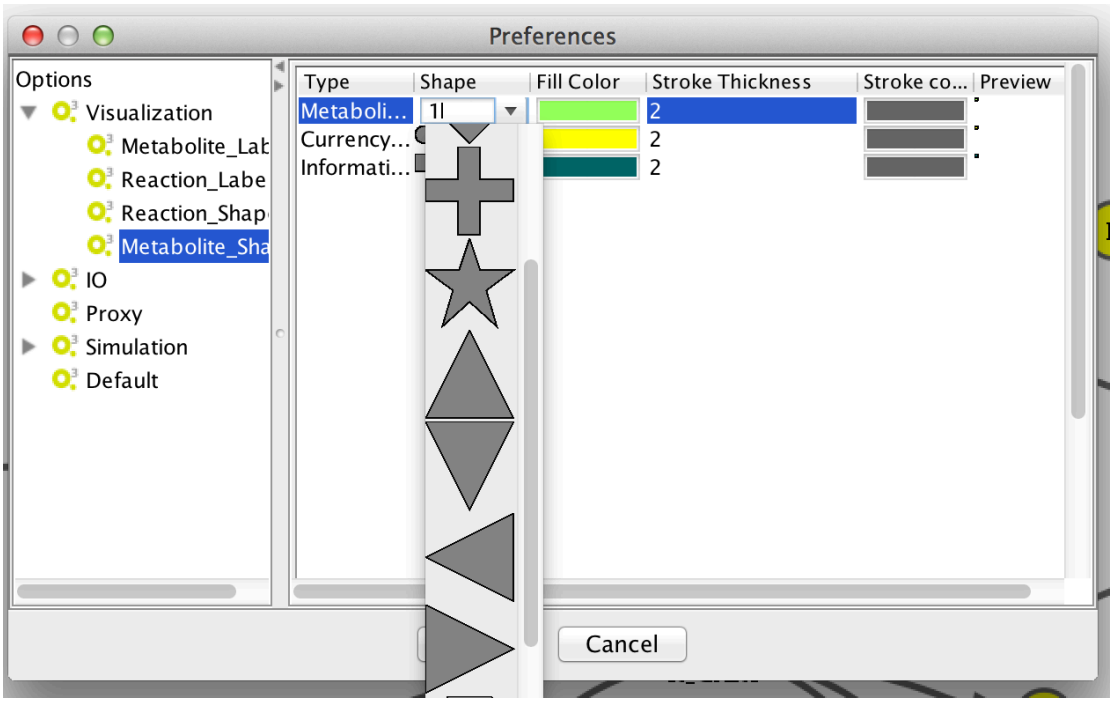
For instance, for the reaction nodes, a list of available attributes is displayed in the left box of the menu displayed in the image below. In the right panel are the selected attributes to display in the node. It is also possible to define the separators for these attributes to be displayed, and have a preview of the result in the box below (In this case, all the available attributes for this model were selected, and are all separated by NEW LINE).



The final result is displayed below.



The other set of options configurable are the shapes and colors of the nodes. For both reaction and metabolite nodes, it is possible to change the shape of the node, the fill color and the stroke color and thickness.



WHAT'S NEXT ?

Now that we have covered the most basic operations of OptFlux, it is time to mention that in the web site www.optflux.org, you will find a number of plug-ins that can be installed to enlarge OptFlux's functionalities and that are constantly been added and updated. Please visit our web site and find the options currently available.

Also, a more detailed explanation of the way OptFlux's functionalities work is available as a set of How To's in our web site. These illustrate step by step the main operations of OptFlux.

Have fun !!

The OptFlux developers