

## Metabotools tutorial I

Authors: Malke K. Aurich, Sylvain Arreckx, Systems Biochemistry Group, LCSB, University of Luxembourg.

Reviewer(s): Anne Richelle, Lewis Lab at University of California, San Diego.

## INTRODUCTION

In this tutorial, we generate contextualized models of two lymphoblastic leukemia cell lines, CCRF-CBM and Mol-4 cells. They will be generated by integrating semi-quantitative metabolomic data, transcriptomic data, and growth rates. We will afterwards analyze the solution space of these models by using a sampling analysis.

Before running a section in the tutorial, read the corresponding sections in the MiraboTools protocol and supplemental tutorial (Data sheet 2).

<http://www.tandfonline.com/doi/full/10.1080/10717825.2013.770201>

**PROCEEDINGS**

Clear workspace and initialize the COBRA Toolbox

clear  
is in combination with

**Step 6 - Define the output location and set the LP solver**

Define the output path and set the solver for LP problem

```
global CDTDIR % set path to ccdirectories (path to CDTDIR)
outputPath = pwd; outputPath = 'ADD YOUR PATH TO YOUR OUTPUT FOLDER'
solver = 'glpk'; % solver = 'ADD YOUR SOLVER'; % e.g., 'cplex_direct' for CPLEX
solverDir = char(ccdirectories)/solver; % " " " " " "
```

Check the solver setup

```
if solverDB == 1
    print(f"Solver %s is set.\n", solver)
else
    error(f"Solver %s could not be used. Check if %s is in the PATH (set PATH) or check for Typo", solver, solver)
end
```

Load and check that the input model is correctly loaded

```

TutorialPath = fileparts(which("Tutorial_metadata.txt"))
if (length(result[TutorialPath][filesep, "starting_model.txt"], 'file') == 2)
  starting_model = read.csv(result[TutorialPath][filesep, "starting_model.txt"])
  tryCatch("The model is loaded.")
else
  error("The model "starting_model" could not be loaded.")
end

```

Check output path and writing permission

```
% Test if outputPath, dir() == T
assert('output directory in "outputPath" does not exist. Verify that you type it correctly or create the directory.');
```

```
end

% Make and clear a dummy file to test the writing to output directory
A = rand(1);
try
    save([outputPath filesep 'A']);
catch ME
    assert('Files cannot be saved to the provided location: %s/obtain rights to write into %s directory or set "outputPath" to a different location.');
```

#### Step 1: Shaping the model's environment using setMediumConstraints

Constrain the model using the data related to EPTM medium composition. To this end, define the set of exchange reactions for which exometabolic data are available.

[illegible][illegible]

Define constraints on basic medium components, i.e., metabolites that are unique from the medium but not captured by the measured data

```
medianCompound = {'XX_csd(e)': 'XX_b(e)'; 'XX_b2(e)': 'XX_bcd(e)'; 'XX_bcd(e)': 'XX_bcd(e)'; 'XX_csd(e)': 'XX_csd(e)'; 'XX_csd(e)': 'XX_csd(e)'; 'XX_csd(e)': 'XX_csd(e)'};
medianCompound = 10 - 100;
```

[illegible]



```

2837812.667 2988029.667 2833212.667 1783123.667
2995908.333 3818106.333 3820408.333 2266832.333
68877.18333 67815.12 89686.49 95625.28
3888 3888 82405.9667 2281286.867
45386.86667 32977.77333 36366.26667 68759.23
181385.1 1238708.1 3818812.867 2978828.3
258828182.3 221218423 221818883 258883897.3
612568.8333 622562.3 588881.7333 888785.6
818888.8333 788812.5667 688813.8 812885.9
88888.78667 88751.96 88717.18667 68882.88333
328737.333 3588189.333 3583788.333 328886.6333
95829.78667 183886.7867 97586.78667 182878.69

```

```
};
```

Define the data associated with CCRF-CBM cell culture

```

input_R = [
% control 2 TP 1 control 3 TP 2 Cond 2 TP 1 Cond 2 TP 2
65251.86667 68888.93 23888.77 88889.88
3888 38878.786 3888 96181.77333
2838866.633 1817842.967 5232377.933 138888859.9
183882.9667 188882.82 293883.7 688876.5267
673539.8667 453297.8667 637788.5667 638887.2667
8882.527333 8788.7883 8817.588 88886.77933
28188.15 21888.78 62188.67333 1812812.38
3888 3888 9818.982 128833.8973
3888 3888 7232.238333 143587.7387
3888 3888 3888 17642.53667
612582 688887.667 682328.633 3888981.333
253882 2532723.667 2888877 1378888
68882.2667 627512.8333 388889.3 388888.2
88878.6 387589.1333 178779.1 28888.3333
888795 838512.133 823778.667 658881.667
18818.8 898875.8 188887.5 58881.6667
28823778.33 28881288.67 21128835.67 18888765.67
2123838.67 21232778.33 28788535.33 17238888
7835388.67 7525888.33 68888817.67 21238888.33
878888.8333 98323.3 888122.5667 238275.9333
18924.86 17888.2387 188275.26 68881.18333
2837812.667 2988029.667 2833212.667 1783123.667
2995908.333 3818106.333 3820408.333 2266832.333
68877.18333 67815.12 89686.49 95625.28
3888 3888 82405.9667 2281286.867
45386.86667 32977.77333 36366.26667 68759.23
181385.1 1238708.1 3818812.867 2978828.3
258828182.3 221218423 221818883 258883897.3
612568.8333 622562.3 588881.7333 888785.6
818888.8333 788812.5667 688813.8 812885.9
88888.78667 88751.96 88717.18667 68882.88333
328737.333 3588189.333 3583788.333 328886.6333
95829.78667 183886.7867 97586.78667 182878.69

```

```
};
```

Use the function `defineDataSecretionProfile` to calculate the uptake and secretion rate over the time of the culture for both condition (e.g. CCRF-CBM and Mst-4 cells)

```

Tal = 8.88;
[cond2_uptake, cond2_secretion, cond2_secretion, slope_ratio] = defineDataSecretionProfile(
[input_R, input_R, data_R888, Tal, exch_excl, exclude_upt, exclude_secv, add_secv, add_upt]);

```

Step 4: Calculate the difference between the uptake and secretion profiles from the two conditions

Use `calculateQuantitativeDiffs` to calculate the sets of exchange reactions with higher uptake and secretion in condition 1 than in condition 2.

Also adapt the condition uptake and secretion for the second condition, this is sometimes necessary to allow the model to achieve a feasible flux.

```

cond2_secretion = [cond2_secretion; 'EX_asyto(s)'; 'EX_30asp'; 'EX_wlto(s)'; 'EX_asyto(s)'; 'EX_asyto(s)'; 'EX_asyto(s)'; 'EX_asyto(s)'];
cond2_secretion{member(cond2_secretion, {'EX_asyto(s)'; 'EX_30asp'; 'EX_wlto(s)'; 'EX_asyto(s)'; 'EX_asyto(s)'; 'EX_asyto(s)'; 'EX_asyto(s)'; 'EX_asyto(s)'}))} = [];
cond2_uptake = [cond2_uptake; 'EX_wlto(s)'];
cond2_uptake{member(cond2_uptake, {'EX_wlto(s)'}))} = [];

[cond2_upt_higher, cond2_upt_higher, cond2_sec_higher, cond1_sec_higher, cond2_uptake_size,...
cond2_uptake_id0, cond2_secretion_id0, cond2_secretion_id0] = calculateQuantitativeDiffs(data_R888,...
slope_ratio, ex_R888, lad_08, cond1_uptake, cond2_uptake, cond1_secretion, cond2_secretion);

```

NOTE: Sometimes, you will need to remove some metabolites from the uptake and secretion profiles, e.g. those for which you assume a different directionality as in the data or if the metabolites are not detected at a specific sampling time. Indeed, the inclusion of these extreme point could distort the results. Example of consumption slope ratio associated to `EX_asyto(s)` is 7875% higher in Mst-4 compared to CCRF-CBM cells. Therefore, these metabolites need to be removed from the input for semi-quantitative adjustment unless such large differences are justified and make sense biologically.

```

remove = {'EX_asyto(s)'; 'EX_30asp(s)'};
A = [];
for i = 1:length(cond2_upt_higher)
    if find(member(remove, cond2_upt_higher(i, 1))) > 0
        A = [A; i];
    end
end
cond2_upt_higher(A, :) = [];

```

Step 5: Enforce uptake and secretion rate using qualitative constraints



```

Tolerance = 0;
maxTime = 1000000;

Tolerance = 'modelA'% MULTIS condition specific model
performSampling(model_Mult, warmup, Tolerance, nFiles, pointPerFile, stepsPerPoint, Tolerance, maxTime, outputPath);
Tolerance = 'modelB'% CDM-CDM condition specific model
performSampling(model_CDM, warmup, Tolerance, nFiles, pointPerFile, stepsPerPoint, Tolerance, maxTime, outputPath);

```

Use the function `summarizedSamplingResults` to return the median of the flux values from the two sampled models. The analysis can be limited to a specific set of reaction defined in `show_react`. Moreover, reactions associated with genes of special interest (e.g. differentially expressed genes) can be defined in `dataGenes` to facilitate the analysis

```

Tol = 0;
nFiles = 50;
pointPerFile = 1000;
starting_Model = modelMedium;
nOut_per_pige = 0;
Size = 10;
modelA = model_Mult;
modelB = model_CDM;
dataGenes = [12;201;451;452;1537;5680;1632;5681;1737;1757;2588;2186;2218;2596];
show_react = ("PDC")("3DC218")("ATP18a")("ATP");
[STATS, STATOR] = summarizedSamplingResults(modelA, modelB, outputPath, nFiles, pointPerFile, starting_Model, dataGenes, show_react, Tol, 1

```