unsteady-state Flux Balance Analysis (uFBA)

James T. Yurkovich

Department of Bioengineering and the Bioinformatics and Systems Biology Program, University of California, San Diego USA

Reviewed by Aarash Bordbar

INTRODUCTION

In this tutorial, we will use unsteady-state Flux Balance Analysis (uFBA) [1] to integrate exo- and endo-metabolomics data [2] into a constraint-based metabolic model for the human red blood cell [3]. The uFBA method allows for bypassing the steady-state assumption for intracellular metabolites that are measured.

We can model the flux through a metabolic network using a set of linear equations defined by

$$\mathbf{S} \cdot \mathbf{v} = \mathbf{b}$$

where **S** is the stoichiometric matrix, **v** is a vector of fluxes through the chemical reactions defined in **S**, and **b** represents constraints on the change of metabolite concentrations; at steady-state, **b** = 0. If the metabolomics measurements are non-linear (i.e., Fig. 1), then the first step of the uFBA workflow is to identify discrete time intervals which represent linearized metabolic states (Fig. 1). Once discrete states are identified (the raw data if linear), we proceed to estimating metabolite concentration rates of change. For each metabolic state, we can use linear regression to calculate the rate of change of each metabolite concentration. If the rate of change is significant, the model is updated by changing the steady-state constraint from 0 to

$$\mathbf{S} \cdot \mathbf{v} \geq \mathbf{b}_1$$
$$\mathbf{S} \cdot \mathbf{v} \leq \mathbf{b}_2$$

where $[\mathbf{b}_1, \mathbf{b}_2]$ represents the 95% confidence interval for each significantly changing metabolite. All unmeasured metabolites are assumed to be at steady-state (i.e., $\mathbf{b}_1 = \mathbf{b}_2 = 0$).

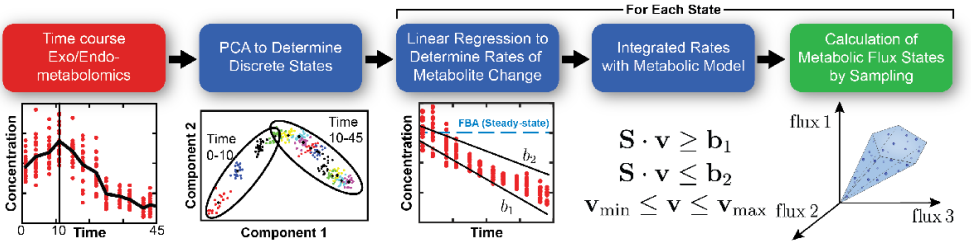

Fig. 1 | Overview of the uFBA workflow.

MATERIALS

Equipment Setup

Running uFBA requires the installation of a mixed-integer linear programming solver. We have used Gurobi 7.0.0 (http://www.gurobi.com/downloads/download-center) which is freely available for academic use (this workflow has only been tested with Gurobi solvers; use other solvers at your own risk). This tutorial uses the Statistics Toolbox to perform linear regression (if the Statistics Toolbox is not installed, compute linear regression manually; see `testUFBA.m`).

PROCEDURE

Initialize

Running uFBA requires the use of several functions from the COBRA Toolbox.

```
initCobraToolbox
```

We first load in sample data. This data is absolutely quantified and has already been volume adjusted such that intracellular and extracellular metabolite concentrations have compatible units.

```
tutorialPath = fileparts(which('tutorial_uFBA.mlx'));
load([tutorialPath filesep 'sample_data.mat']);
% We load the model by readCbModel to make sure it fits to the specifications.
model = readCbModel([tutorialPath filesep 'sample_data.mat'],'modelName','model')
```

The `sample_data.mat` file contains the following variabels:

- `met_data`: a matrix containing the exo- and endo-metabolomics data
- `met_IDs`: a cell array containing the BiGG ID for each of the metabolites in `met_data`
- `model`: a modified version [3] of the iAB-RBC-283 COBRA model structure
- `time`: a vector of the time points (in days) at which the metabolite concentrations were measured
- `uFBAvariables`: a struct containing the variables necessary for input into the uFBA algorithm

In this tutorial, the use of Gurobi is mandatory.

```
solverLPOk = changeCobraSolver('gurobi', 'LP');
solverMILPOk = changeCobraSolver('gurobi', 'MILP');
```

Estimate Metabolite Rates of Change (<1 sec.)

Next, we run linear regression to find the rate of change for each metabolite concentration.

```
changeSlopes = zeros(length(met_IDs), 1);
changeIntervals = zeros(length(met_IDs), 1);
for i = 1:length(met_IDs)
    [tmp1, tmp2] = regress(met_data(:, i), [time ones(length(time), 1)], 0.05);
    changeSlopes(i, 1) = tmp1(1);
    changeIntervals(i, 1) = abs(changeSlopes(i, 1) - tmp2(1));
end
```

The variables `changeSlopes` and `changeIntervals` contain the metabolite rates of change and 95% confidence intervals, respectively. We will create a new vector, `ignoreSlopes`, which contains a 0 for the metabolites whose slopes change significantly and a 1 otherwise:

```
tmp1 = changeSlopes - changeIntervals;
tmp2 = changeSlopes + changeIntervals;
ignoreSlopes = double(tmp1 < 0 & tmp2 > 0);
```

Integration of Metabolomics Data (<10 min.)

Finally, we need to input the data into the uFBA algorithm which is encapsulated in the function `buildUFBAmodel`. This function takes as input a COBRA model structure and a struct containing the required input variables (see Table 1).

Ideally, all metabolites in the model would be measured, resulting in a feasible model. However, experimental limitations limit the number of metabolites that can measured. Thus, when the metabolite constraints are added, the model will most likely not simulate. The uFBA algorithm reconciles the

measured metabolomics data and the network structure by parsimoniously allowing unmeasured metabolites concentrations to deviate from steady-state (i.e., $\mathbf{S} \cdot \mathbf{v} = \mathbf{b}$) in order to build a computable model. We refer to the method for deviating unmeasured metabolites from steady-state as "metabolite node relaxation." As part of this procedure, free exchange of extracellular metabolites out of the system is only allowed if (1) the metabolite concentration is measured to be increasing or (2) if the relaxation of a particular extracellular metabolite is required for model feasibility.

There are five different techniques built into the uFBA method to perform the node relaxation. The technique used in this tutorial is an MILP optimization that minimizes the number of unmeasured metabolites relaxaed from steady-state; this choice effectively minimizes the changes made to the model in order to achieve feasibilitiy. Full details for this and all other node relaxation techniques can be found in [1]. Sinks are added for each of the relaxed metabolite nodes, and the flux through each of these sinks is minimized while still allowing the model to simulate. The minimimum value is then multiplied by a relaxation factor `lambda` (Table 1) and used as the bound for the sink reaction.

Full details for the algorithm are provided in the original publication [1].

Required Inputs	**Description**
model	A COBRA model structure containing (at minimum) the following fields: S, b, lb, ub, mets, rxns
metNames	A cell array containing the model IDs of the measured metabolites that will have bounds set by the algorithm. These metabolites should correspond to model.mets. Note: measured metabolites that were not significantly changed over time should also be included.
changeSlopes	A vector (length(metNames) x 1) that contains the mean rate of change (the slope from linear regression) for each metabolite in metNames.
changeIntervals	A vector (length(metNames) x 1) that contains the difference between the mean slope of change and the upper bound of the 95% confidence interval for each slope in changeSlopes.
ignoreSlopes	A binary vector (length(metNames) x 1) that instructs which changeSlopes to be ignored (ignore if 1). Metabolites were ignored if the values of the slopes were not significant based on linear regression (i.e., if slope value +/- the interval crossed zero).
Optional Inputs	**Description**
objRxn	The objective reaction (corresponding to model.rxns) for the new uFBA model. Default is the objective reaction from the original model.
metNoSink	A cell array of metabolites (corresponding to model.mets) that should not have a sink added, typically for metabolites where the concentration is known to be zero. Default is an empty cell array.
metNoSinkUp	A cell array of metabolites (corresponding to model.mets) that should not have a sink added in the up direction (which would allow metabolite accumulation). Default is an empty cell array.
metNoSinkDown	A cell array of metabolites (corresponding to model.mets) that should not have a sink added in the down direction (which would allow metabolite depletion). Default is an empty cell array.
conflictingMets	A cell array of intracellular metabolites (corresponding to model.mets) where the intracellular rates conflict with extracellular rates, and the model cannot compensate through biosynthesis of the metabolite or use of the flux in other pathways. Typically only necessary for very simple cell types (e.g., RBCs). The intracellular rate is adjusted to the extracellular to allow the model to simulate. Default is an empty cell array.
solvingStrategy	One of {'case1','case2','case3','case4','case5'} which correspond to the 5 node relaxation techniques discussed in the methods section of [1]. Default value is the first LP technique, 'case2'.
lambda	A multiplicative relaxation away from the minimum allowed deviation from the steady-state model. Default value is 1.5.
numIterations	The number of iterations for the integer cut optimization method. Default value is 100.
timeLimit	The time limit for the solver during the numIterations optimization loop. Default value is 30 seconds.
eWeight	A weighting factor for preferential selection of extracellular sinks over intracellular during node relaxation. Default value is 1e4. If no weigthing is preferred, eWeight should be set to a value of 1.
Outputs	**Description**
model	The final uFBA model.
metsToUse	Metabolites for which metabolomics data was integrated.
relaxedNodes	A cell array which contains three columns: (1) which metabolites were relaxed from steady-state; (2) the direction of the relaxation (accumulation/depletion); and (3) the upper bound of the added sink.

Table 1 | Inputs and outputs of the `buildUFBAmodel` function.

```
uFBAvariables.metNames = met_IDs;
uFBAvariables.changeSlopes = changeSlopes;
uFBAvariables.changeIntervals = changeIntervals;
uFBAvariables.ignoreSlopes = ignoreSlopes;

uFBAoutput = buildUFBAmodel(model, uFBAvariables);
```

The output contains the resulting model (`uFBAoutput.model`):

```
model_ufba = optimizeCbModel(uFBAoutput.model)
```

References

[1] A Bordbar*, JT Yurkovich*, G Paglia, O Rolfsson, O Sigurjonsson, and BO Palsson. "Elucidating dynamic metabolic physiology through network integration of quantitative time-course metabolomics." *Sci. Rep.* (2017). doi:10.1038/srep46249. (* denotes equal contribution)

[2] A Bordbar, PI Johansson, G Paglia, SJ Harrison, K Wichuk, M Magnusdottir, S Valgeirsdottir, M Gybel-Brask, SR Ostrowski, S Palsson, O Rolfsson, OE Sigurjonsson, MB Hansen, S Gudmundsson, and BO Palsson. "Identified metabolic signature for assessing red blood cell unit quality is associated with endothelial damage markers and clinical outcomes." *Transfusion* (2016). doi:10.1111/trf.13460.

[3] A Bordbar, D McCloskey, DC Zielinski, N Sonnenschein, N Jamshidi, and BO Palsson. "Personalized Whole-Cell Kinetic Models of Metabolism for Discovery in Genomics and Pharmacodynamics." *Cell Systems* (2015). doi:10.1016/j.cels.2015.10.003.