Flux Balance Analysis (FBA)

Author(s): Vanja Vlasov, Marouen Ben Guebila, Systems Biochemistry Group, LCSB, University of Luxembourg,

Thomas Pfau, Systems Biology Group, LSRU, University of Luxembourg

Reviewer(s): Ines Thiele, Catherine Clancy, Systems Biochemistry Group, LCSB, University of Luxembourg

Thomas Pfau, Systems Biology Group, LSRU, University of Luxembourg

INTRODUCTION

Flux balance analysis (FBA) evaluates the metabolic flux distribution[1], and is one of the most used modelling approaches for metabolic systems.

The applications of FBA for molecular systems biology include prediction of the growth rates, uptake rates, knockout lethality and product secretion. In FBA, the solution space is constrained by the assumption of a steady-state, under which each internal metabolite is consumed at the same rate as it is produced.

For the quantitative estimation of the metabolic fluxes, linear programming (LP) can be used to solve the stoichiometric matrix for a given objective function under different constraints. The constraints of the problem depict the space of all eligible possibilities from which an optimal solution can be selected;

$$\begin{aligned} \min_{v} \quad & c^T v \\ \text{s.t.} \quad & Sv = b, \\ & l \le v \le u, \end{aligned}$$

Equation 1: Formula of standard FBA.

where $c \in \mathfrak{R}^n$ is a parameter vector that linearly combines one or more reaction fluxes to form what is termed the objective function, and where a $b_i < 0$, or $b_i > 0$, represents some fixed output, or input, of the ith molecular species. $S \in \mathfrak{R}^{m \times n}$ is a stoichiometric matrix for m molecular species and n reactions, and b is a vector of known metabolic exchanges. The output of FBA is a particular flux distribution, v , which maximises or minimises the objective function and stands between upper and lower bounds, u and l, respectively.

There are multiple different variants of FBA which will be discussed here:

1. **Standard FBA**
2. **Sparse FBA**
3. **Metabolite dilution FBA (mdFBA)**
4. **Geometric FBA**
5. **Parsimonious enzyme usage Flux Balance Analysis (pFBA)**
6. **Dynamic FBA**
7. **Relax FBA**

8. **Flux enrichment analysis (FEA)**

EQUIPMENT SETUP

Initialise The Cobra Toolbox and set the solver.

If necessary, initialise the cobra toolbox:

```
initCobraToolbox
```

```
      _____   _____   _____   _____     _____     |
     / ___| / _   \ |  __  \ |  __  \   / ___ \   |   COnstraint-Based Reconstruction and Analysis
    | |     | | | | | |__| | | |__| |  | |___| |  |   The COBRA Toolbox - 2017
    | |     | | | | |  __  { |  _  /   |  ___  |  |
    | |___  | |_| | | |__| | | | \ \   | |   | |  |   Documentation:
     \_____| \_____/ |_____/ |_|  \_\ |_|   |_|  |   http://opencobra.github.io/cobratoolbox
                                                  |

> Checking if git is installed ...  Done.
> Checking if the repository is tracked using git ...  Done.
> Checking if curl is installed ...  Done.
> Checking if remote can be reached ...  Done.
> Initializing and updating submodules ... Done.
> Adding all the files of The COBRA Toolbox ...  Done.
> Define CB map output... set to svg.
> Retrieving models ...   Done.
> TranslateSBML is installed and working properly.
> Configuring solver environment variables ...
  - [----] ILOG_CPLEX_PATH :  --> set this path manually after installing the solver ( see instructions
  - [----] GUROBI_PATH :  --> set this path manually after installing the solver ( see instructions )
  - [---*] TOMLAB_PATH: C:\Program Files\tomlab\
  - [----] MOSEK_PATH :  --> set this path manually after installing the solver ( see instructions )
  Done.
> Checking available solvers and solver interfaces ... Done.
> Setting default solvers ... Done.
> Saving the MATLAB path ... Done.
  - The MATLAB path was saved in the default location.

> Summary of available solvers and solver interfaces

    Support          LP   MILP     QP   MIQP    NLP
----------------------------------------------------------------------
cplex_direct  active          0      0      0      0      -
dqqMinos      active          0      -      -      -      -
glpk          active          1      1      -      -      -
gurobi        active          1      1      1      1      -
ibm_cplex     active          0      0      0      -      -
matlab        active          1      -      -      -      1
mosek         active          0      0      0      -      -
pdco          active          1      -      1      -      -
quadMinos     active          0      -      -      -      0
tomlab_cplex  active          1      1      1      1      -
qpng          passive         -      -      1      -      -
tomlab_snopt  passive         -      -      -      -      1
gurobi_mex    legacy          0      0      0      0      -
lindo_old     legacy          0      -      -      -      -
lindo_legacy  legacy          0      -      -      -      -
lp_solve      legacy          1      -      -      -      -
opti          legacy          0      0      0      0      0
----------------------------------------------------------------------
Total         -               6      3      4      2      2
```

```
+ Legend: - = not applicable, 0 = solver not compatible or not installed, 1 = solver installed.


> You can solve LP problems using: 'glpk' - 'gurobi' - 'matlab' - 'pdco' - 'tomlab_cplex' - 'lp_solve'
> You can solve MILP problems using: 'glpk' - 'gurobi' - 'tomlab_cplex'
> You can solve QP problems using: 'gurobi' - 'pdco' - 'tomlab_cplex' - 'qpng'
> You can solve MIQP problems using: 'gurobi' - 'tomlab_cplex'
> You can solve NLP problems using: 'matlab' - 'tomlab_snopt'

> Checking for available updates ...
--> You cannot update your fork using updateCobraToolbox(). [0b7c0f @ tut_FBA].
    Please use the MATLAB.devTools (https://github.com/opencobra/MATLAB.devTools).
```

For solving LP problems in a FBA analysis, certain solvers are required and can be set using the `changeCobraSolver` function:

```
% solverOK = changeCobraSolver(solverName, solverType, printLevel, unchecked)
```

The present tutorial can run with glpk package, which does not require additional installation and configuration. Although, for the analysis of large models is recommended to use the GUROBI package.

Setup the appropriate solver for the machine you are using by removing the "%" (comment) sign for only the desired solver.

```
changeCobraSolver('glpk','all');
```

```
> Solver for LP problems has been set to glpk.
> Solver for MILP problems has been set to glpk.
> Solver glpk not supported for problems of type MIQP. Currently used: tomlab_cplex
> Solver glpk not supported for problems of type NLP. Currently used: matlab
> Solver glpk not supported for problems of type QP. Currently used: qpng
```

```
% changeCobraSolver('tomlab_cplex','all');
% changeCobraASolver('ibm_cplex','all');
% changeCobraSolver ('gurobi', 'all');
```

Model Setup

This tutorial will use the generic model of the human cellular metabolism[2], Recon 2.0. Other COBRA models, including Recon 3, may also be run with this tutorial. For information on metabolites structures and reactions, and to download the latest COBRA model releases, visit the Virtual Metabolic Human database (VMH, http://vmh.life).

Before proceeding with the simulations, load the model into the workspace:

```
% check if Recon3 exists:
% pathModel = '....'; enter the path to the Recon 3 model file
% filename = '2017_04_28_Recon3d.mat';
% load([pathModel, filename])
% model = modelRecon3model;
% clear modelRecon3model
% and if not select your own model, or use Recon2.0model instead filename='Recon3.0model';
global CBTDIR
load([CBTDIR filesep 'test' filesep 'models' filesep 'Recon2.0model.mat']);
model = Recon2model;
model.rxns = strrep(model.rxns, '(', '[');
model.rxns = strrep(model.rxns, ')', ']');
clear Recon2model
```

In this tutorial we assume, that the cellular objectives include energy production or optimisation of uptake rates and by-product secretion for various physiological functions of the human body.

PROCEDURE

1. Standard FBA

Standard FBA predicts an optimal solution for a cellular objective within a given set of constraints on a metabolic network (see Equation 1). Constraints on the network are set by assigning limits on the uptake, consumption or production of metabolites in reactions.

Timing:

The time to determine a FBA solution depends on the size of the genome-scale model and is commonly less than a second for a medium sized model.

Calculating maximal ATP energy production under aerobic conditions:

For each new simulation, the original model will be copied to a new variable. This preserves the constraints of the original model to perform further simulations with new constraints. Additionally, this method of renaming the model avoids confusion while performing multiple simulations at the same time.

```
modelaerobic = model;
```

The ATP demand reaction, i.e., `DM_atp_c_` within the model is a reaction that involves hydrolysis of ATP to ADP, Pi and proton in the cytosol.

```
 printRxnFormula(model, 'DM_atp_c_');
```

```
DM_atp_c_ h2o[c] + atp[c]  -> adp[c] + h[c] + pi[c]
```

We will set this reaction as our objective with the `'changeObjective'` command. Maximising the flux through the ATP demand reaction will result in the network producing a maximal amount of ATP (up to the limit of the reaction).

```
modelaerobic = changeObjective (modelaerobic, 'DM_atp_c_');
```

The glucose and oxygen, in this case, are provided in high amounts for calculating the flux through ATP demand.

The `'changeRxnBounds'` function changes the flux constraints of the lower ('`l`'), upper ('u'), or both the bounds ('b'), of the specified reaction. Here, we will change the maximal uptake of glucose to 20 μmol/min/gDW and of oxygen to 1000 μmol/min/gDW. The uptake of oxygen is effectively unconstrainted (i.e. infinity).

```
% modelaerobic = changeRxnBounds (modelaerobic, 'EX_glc_D[e]', -20, 'l'); % For Recon 3.0 mode
modelaerobic = changeRxnBounds (modelaerobic, 'EX_glc[e]', -20, 'l'); % For Recon 2.0 model
modelaerobic = changeRxnBounds (modelaerobic, 'EX_o2[e]', -1000, 'l'); % For both models, Reco
```

The function `optimizeCbModel` calculates one of the optimal solutions for a (maximum or minimum) objective reaction within the defined solution space. In the above example, the maximal flux through the `DM_atp_c_` is desired.

```
FBAaerobic = optimizeCbModel (modelaerobic, 'max')
```

```
FBAaerobic = 
         full: [7440×1 double]
          obj: 1000
        rcost: [7440×1 double]
         dual: [5063×1 double]
       solver: 'glpk'
    algorithm: 'default'
         stat: 1
     origStat: 5
         time: 0.8220
        basis: []
            x: [7440×1 double]
            f: 1000
            y: [5063×1 double]
            w: [7440×1 double]
            v: [7440×1 double]
```

- **Anticipated results**

When oxygen and all carbon sources (internal and external) are provided the flux through ATP demand reaction can reach its maximum rate of 1000 μmol/min/gDW.

- **Troubleshooting**

If there are multiple carbon sources available in the model, it may be necessary to specify more constraints in order to examine the effect of a single carbon source on ATP production.

To avoid this issue, all external carbon sources need to be closed with the exception of the single carbon source of interest.

```
%Closing the uptake of all energy and oxygen sources
idx=strmatch('Exchange/demand reaction',model.subSystems);
c=0;
for i=1:length(idx)
    if model.lb(idx(i))~=0
        c=c+1;
        uptakes{c}=model.rxns{idx(i)};
    end
end

% If you use Recon3.0 model, then:
% modelalter = model;
% modelalter = changeRxnBounds(modelalter, uptakes, 0, 'b');
% modelalter = changeRxnBounds(modelalter, 'EX_HC00250[e]', -1000, 'l');

% The alternative way to do that, in case you were using another large model,
% that does not contain defined Subsystem is
% to find uptake exchange reactions with following codes:
% [selExc, selUpt] = findExcRxns(model);
% uptakes1 = model.rxns(selUpt);

% Selecting from the exchange uptake reactions those
% which contain at least 1 carbon in the metabolites included in the reaction:
 subuptakeModel = extractSubNetwork(model, uptakes);
 hiCarbonRxns = findCarbonRxns(subuptakeModel,1);
% Closing the uptake of all the carbon sources
```

```
 modelalter = model;
 modelalter = changeRxnBounds(modelalter, hiCarbonRxns, 0, 'b');
% Closing other oxygen and energy sources
 exoxygen = {'EX_adp'
    'EX_amp[e]'
    'EX_atp[e]'
    'EX_co2[e]'
    'EX_coa[e]'
    'EX_fad[e]'
    'EX_fe2[e]'
    'EX_fe3[e]'
    'EX_gdp[e]'
    'EX_gmp[e]'
    'EX_gtp[e]'
    'EX_h[e]'
    'EX_h2o[e]'
    'EX_h2o2[e]'
    'EX_nad[e]'
    'EX_nadp[e]'
    'EX_no[e]'
    'EX_no2[e]'
    'EX_o2s[e]'};
modelalter = changeRxnBounds (modelalter, exoxygen, 0, 'l');
```

Calculating maximum ATP energy production under anaerobic and glucose only conditions:

```
modelanaerobic = modelalter;
% modelanaerobic = changeRxnBounds (modelanaerobic, 'EX_glc_D[e]', -20, 'l'); % For Recon3.0 m
modelanaerobic = changeRxnBounds(modelanaerobic, 'EX_glc[e]',-20,'l');
modelanaerobic = changeRxnBounds (modelanaerobic, 'EX_o2[e]', 0, 'l');
modelanaerobic = changeObjective(modelanaerobic,'DM_atp_c_');
FBAanaerob = optimizeCbModel(modelanaerobic,'max')
```

```
FBAanaerob =
         full: [7440×1 double]
          obj: 82.6176
        rcost: [7440×1 double]
         dual: [5063×1 double]
       solver: 'glpk'
    algorithm: 'default'
         stat: 1
     origStat: 5
         time: 0.3160
        basis: []
            x: [7440×1 double]
            f: 82.6176
            y: [5063×1 double]
            w: [7440×1 double]
            v: [7440×1 double]
```

- **Anticipated results**

Compared to the aerobic condition, anaerobic condition with only glucose as an energy source has reduced flux through ATP demand (82 μmol/min/gDW), signifying the need to oxygen to run the oxidative phosphorylation. The results are dependant on the model you are using. For Recon 3.0, under anaerobic conditions with only glucose as an energy source, the flux for ATP demand is 40 μmol/min/gDW.

2. Sparse FBA

Sparse FBA calculates the optimal solution of an objective function and finds the smallest set of reactions that can carry flux to achieve the objective. Sparse FBA minimises the number of reactions by keeping same maximal objective;

$$\begin{array}{ll} \min\limits_{v} & \\ \text{s.t.} & Sv = b, \\ & l \leq v \leq u, \\ & c^T v = \rho^* \end{array}$$

Equation 2: Formula of Sparse FBA.

where the last constraint is optional and represents the requirement to satisfy an optimal objective value ρ^* derived from any solution to a FBA problem. This approach is used to check for minimal sets of reactions that either should be active or should not be active in a flux balance model that is representative of a biochemical network.

```
% [vSparse, sparseRxnBool, essentialRxnBool]  = sparseFBA(model, osenseStr,...
%  checkMinimalSet, checkEssentialSet, zeroNormApprox)
```

As an optional input, there are different appoximation types of zero-norm (only available when `minNorm = 'zero'`). Default is `cappedL1`.

```
% Other types of zero-norm:
%  * 'cappedL1' : Capped-L1 norm
%  * 'exp'      : Exponential function
%  * 'log'      : Logarithmic function
%  * 'SCAD'     : SCAD function
%  * 'lp-'      : :math:`L_p` norm with :math:`p < 0`
%  * 'lp+'      : :math:`L_p` norm with :math:`0 < p < 1`
%  * 'l1'       : L1 norm
%  * 'all'      : try all approximations and return the best result
```

Timing:

The time to determine a `sparseFBA()` solution depends on the size of the genome-scale model and is taking from < 1 second for a 1,000 reaction model, to < 2 seconds for a model with more than 10,000 reactions.

Calculating maximal ATP energy production under anaerobic and glucose only conditions:

```
modelspar = modelalter;
% For Recon3.0 model
% modelspar = changeRxnBounds (modelspar, 'EX_glc_D[e]', -20, 'l');
modelspar = changeRxnBounds(modelspar, 'EX_glc[e]',-20,'l');
modelspar = changeRxnBounds (modelspar, 'EX_o2[e]', 0, 'l');
modelspar = changeObjective(modelspar, 'DM_atp_c_');
[vSparse, sparseRxnBool, essentialRxnBool] = sparseFBA(modelspar, 'max');
```

```
---FBA---
   82.6176 FBA objective.
       878 reactions above epsilon = 1e-09
  0.343202 computation time (sec)
---Non-convex approximation---
   82.6176 = Sparse FBA objective.
         0 = ||c^T*v - f*||^2.
       878 reactions above epsilon = 1e-09
   7.76885 computation time (sec)
93 of these are heuristically minimal rxns.
93 of these are essential rxns.
```

- **Anticipated results:**

Commonly, a sparse FBA solution will have much smaller number of active reactions compared to a standard FBA on the same model with same objective function. The outputs `sparseRxnBool` and `essentialRxnBool` return vectors with 1 and 0's, with sparse and essential reactions respectively.

Display the sparse flux solution, but only the non-zero fluxes.

```
for i=1:length(vSparse)
    if vSparse(i)~=0
        fprintf('%10d \t %s\n', vSparse(i), modelspar.rxns{i})
    end
end
```

```
-1.158824e+01   10FTHFtm
-2.841176e+01   4ABUTtm
4.261765e+01   4MOPt2im
1.365210e-27   7DHCHSTEROLtr
-1.101885e-12   A_MANASEly
-2.213598e-12   AACOAT
-2.841176e+01   ABTArm
-7.313139e-13   ACACT1r
7.378659e-13   ACACT1x
2.527181e-28   ACGAM2E
-4.407539e-12   ACGAMK
-4.407539e-12   ACGAMtly
2.203770e-12   ACNAMlt
-4.407539e-12   ACNMLr
-6.551998e-15   ACOAD8m
4.407539e-12   ACOAHi
2.161060e-30   ACONTm
7.771523e-15   ADA
6.596416e-12   ADK1
-1.735476e-30   ADK1m
-2.085115e-30   ADNtm
-2.974168e-15   ADPRDP
-7.771523e-15   ADPT
-4.407539e-12   AGDC
2.317647e+01   AGTim
-1.101885e-12   AHEXASEly
-3.011690e-14   AKGDm
1.823522e-28   AKGtp
```

```
2.317647e+01   ALATA_L
9.838212e-13   ALDD21
-1.929902e-27   APAT2rm
-1.106799e-12   ARACHCOAtx
1.101885e-12   Asn_X_Ser_Thrtr
-2.730421e-27   ATP2ter
-2.203770e-12   ATPasel
-1.101885e-12   B_MANNASEly
-1.929902e-27   BALAtmr
-1.101885e-12   BDMT_U
1.365210e-27   C14STRr
5.533994e-13   C160CPT1
5.533994e-13   C160CPT2
6.946765e-29   C204CPT1
1.365210e-27   C3STDH1Pr
1.365210e-27   C3STKR2r
1.365210e-27   C4STMO1r
1.365210e-27   C4STMO2Pr
8.558797e-30   CLOHtex2
2.203770e-12   CMPACNAtg
-6.352941e+00   CO2t
-2.730421e-27   CO2ter
-1.158824e+01   CO2tm
-1.823522e-28   CO2tp
2.760445e-12   COAtm
-2.730421e-27   COAtr
7.378659e-13   CRNCAR3tp
2.161060e-30   CSm
-5.449493e-29   CSNAT2m
7.378659e-13   CSNAT3x
3.002187e-12   CYOR_u10m
8.820760e-12   CYTD
8.820760e-12   CYTK11
-6.616990e-12   CYTK8
6.094937e-30   DADA
-8.820760e-12   DCMPDA
-3.995737e-28   DEDOLP1_U
-3.995737e-28   DEDOLP2_U
-3.995737e-28   DEDOLR_U
6.946765e-29   DESAT20_1
6.946765e-29   DESAT22_1p
9.838212e-13   DESAT22_2p
-9.838212e-13   DESAT24_1
-8.820760e-12   DGK2m
-8.820760e-12   DGNSKm
-8.820760e-12   DGSNtm
-1.000000e+00   DHFR
-1.929902e-27   DHPM1
-5.474654e-30   DM_Asn_X_Ser_Thr_ly_
8.261765e+01   DM_atp_c_
5.203449e-43   DM_dgpi_prot_hs_r_
4.407539e-12   DM_dsT_antigen_g_
5.203449e-43   DM_gpi_sig_er_
```

```
-4.407539e-12   DM_sTn_antigen_g_
2.330847e-27   DMATT
8.835294e+01   DNDPt12m
-8.820760e-12   DNDPt18m
8.820760e-12   DNDPt42m
-1.101885e-12   DOLASNT_Uer
-1.101885e-12   DOLDPP_Uer
3.305655e-12   DOLGPP_Uer
3.995737e-28   DOLICHOL_Uter
-4.407539e-12   DOLMANP_Uter
-1.101885e-12   DOLP_Uter
-1.101885e-12   DOLPGT1_Uer
-1.101885e-12   DOLPGT2_Uer
3.995737e-28   DOLPGT3_Uer
-1.101885e-12   DOLPH_Uer
-1.101885e-12   DOLPMT_U
-1.101885e-12   DOLPMT1_Uer
-1.101885e-12   DOLPMT2_Uer
-4.407539e-12   DOLPMT3_Uer
-5.509424e-12   DOLPMT4_Uer
1.929902e-27   DURAD
8.820760e-12   DURIK1
-8.820760e-12   DURIPP
1.365210e-27   EBP1r
-1.101885e-12   ENGASEly
2.841176e+01   ENO
4.146872e-15   EX_amp[e]
6.352941e+00   EX_co2[e]
       -20   EX_glc[e]
5.047059e+01   EX_h2o[e]
-2.841176e+01   EX_nh4[e]
2.131413e-14   EX_pi[e]
        -1   EX_thf[e]
-1.101885e-12   F1PGT
-1.101885e-12   F6Tg
2.730421e-27   FACOAL160i
6.946765e-29   FACOAL203
-6.946765e-29   FAOXC2242046x
6.946765e-29   FAOXC2252053x
-6.946765e-29   FAOXC226205x
1.106799e-12   FAOXC240200x
9.838212e-13   FAOXC2452253x
-9.838212e-13   FAOXC246226x
5.533994e-13   FAS100COA
5.533994e-13   FAS120COA
5.533994e-13   FAS140COA
5.533994e-13   FAS160COA
-1.106799e-12   FAS180COA
2.000000e+01   FBA
1.823522e-28   FCOAH
-1.101885e-12   FK
4.146872e-15   FMNAT
1.823522e-28   FORMCOAtx
```

```
-1.547563e-27   FORt2m
-1.365210e-27   FORtr
-3.995737e-28   FT
1.547563e-27   FTHFLm
-1.101885e-12   FUCASEly
-1.101885e-12   FUCtly
-2.841176e+01   FUMm
-1.101885e-12   G12MT1_U
-1.101885e-12   G12MT2_U
-1.101885e-12   G13MT_U
-1.101885e-12   G14Tg
-1.101885e-12   G16MT_U
-4.407539e-12   G6PDA
-1.101885e-12   GALASE1ly
-2.203770e-12   GALtly
2.203770e-12   GALU
4.000000e+01   GAPD
-1.101885e-12   GASNASEly
2.317647e+01   GCALDD
6.551998e-15   GCC2am
6.551998e-15   GCC2bim
1.158824e+01   GCCam
1.158824e+01   GCCbim
1.158824e+01   GCCcm
-1.101885e-12   GDPFUCtg
1.101885e-12   GDPtg
-3.995737e-28   GGT_U
-1.158824e+01   GHMT2r
8.828531e-12   GK1
-1.101885e-12   GLCNACPT_U
-1.101885e-12   GLCNACT_U
5.794118e+00   GLCt1r
1.420588e+01   GLCt2_2
3.928376e-30   GLCt4
7.969962e-29   GLCter
1.385371e-14   GLUCYS
-4.000000e+01   GLUDym
5.420588e+01   GLUt2m
2.841176e+01   GLUTCOADHm
2.317647e+01   GLXtm
-2.317647e+01   GLXtp
2.317647e+01   GLYCLTtp
2.317647e+01   GLYCTO1p
-1.158824e+01   GLYtm
1.051353e-28   GMAND
-1.101885e-12   GMPtg
5.203449e-43   GPIDAer
2.330847e-27   GRTT
-8.261765e+01   GTHRDt
1.385371e-14   GTHS
-7.771523e-15   GUAD
8.828531e-12   GUAPRT
2.841176e+01   H2CO3D
```

-2.317647e+01	H2O2tp
-4.947059e+01	H2Ot
-1.101885e-12	H2Oter
-6.611309e-12	H2Otg
-1.322262e-11	H2Otly
3.117647e+00	H2Otm
-7.765036e-43	H4ETer
3.912005e-43	H7ET2er
5.203449e-43	H8TAer
2.000000e+01	HEX1
-9.916964e-12	HEX4
-2.944912e-12	HMGCOASi
-2.207046e-12	HMGCOAtm
-7.378659e-13	HMGCOAtx
-2.213598e-12	HMGLm
9.838212e-13	HPCLx
2.317647e+01	HPYRDC
1.101885e-12	Htg
6.611309e-12	Htr
2.161060e-30	ICDHxm
-4.261765e+01	LEUt5m
4.261765e+01	LEUTA
-4.261765e+01	LEUTAm
1.106799e-12	LGNCCOAtx
1.365210e-27	LNSTLSr
1.365210e-27	LSTO2r
-1.101885e-12	M1316Mg
-1.101885e-12	M13N2Tg
-1.101885e-12	M16NTg
-1.101885e-13	M4MPDOL_Uter
-1.101885e-12	M8MASNterg
5.122320e-27	MAN6PI
-6.611309e-12	MANtg
-3.305655e-12	MANtly
-6.551998e-15	MCCCrm
-2.841176e+01	MDHm
-1.101885e-12	MG1er
-1.101885e-12	MG2er
-1.101885e-12	MG3er
-6.551998e-15	MGCHrm
-1.101885e-12	MM5ag
-1.101885e-12	MM6bg
-1.101885e-12	MM7Cbg
-1.101885e-12	MM8Cg
1.220037e-28	MMEm
-1.220037e-28	MMMm
-1.929902e-27	MMSAD3m
-1.158824e+01	MTHFC
1.158824e+01	MTHFCm
-1.158824e+01	MTHFD2
1.158824e+01	MTHFD2m
4.407539e-12	N3Tg
2.155882e+01	NADH2_u10m

```
-2.459553e-13   NADHtpu
-2.974168e-15   NADN
1.365210e-26   NADPHtru
1.365210e-26   NADPtru
-2.841176e+01   NDPK1m
-8.820760e-12   NDPK6m
-2.841176e+01   NH4t3r
4.407539e-12   NS26T2g
-4.407539e-12   NS26Tg
1.762813e-11   NTD1m
-3.002187e-14   O2Stm
1.228689e-26   O2ter
1.501093e-12   O2tm
2.317647e+01   O2tp
-5.533994e-13   OCCOAtm
2.841176e+01   PCm
-5.449493e-29   PCRNtm
2.000000e+01   PFK
1.158824e+01   PGCD
2.000000e+01   PGI
-4.000000e+01   PGK
-2.841176e+01   PGM
-4.407539e-12   PGMT
4.835294e+01   PHCDm
1.823522e-28   PHYHx
-1.101885e-12   PIter
9.916964e-12   PMANM
-1.220037e-28   PPCOACm
-4.619427e-43   PPItr
8.820760e-12   PPM
9.838212e-13   PRISTANALtx
-9.838212e-13   PRISTtx
6.852941e+00   PRO1xm
8.820760e-12   PRPPS
1.158824e+01   PSERT
1.158824e+01   PSP_L
-7.771523e-15   PUNP1
-6.100933e-30   PUNP2
8.820760e-12   PUNP4
7.771523e-15   PUNP5
6.094937e-30   PUNP6
1.365210e-27   PVD3
1.764442e-11   PYK
8.820760e-12   PYNP2r
2.841176e+01   PYRt2m
4.146872e-15   RBFK
-8.735294e+01   RDH3
8.735294e+01   RDH3a
8.835294e+01   RNDR1
8.820760e-12   RNDR2
-8.820760e-12   RNDR3
-1.982779e-15   RPI
4.407539e-12   S23Tg
```

```
-1.101885e-12   S26Tg
-1.101885e-12   S2L2FN2M2MASNt
-1.101885e-12   S2L2FN2M2MASNtly
-1.101885e-12   SIAASEly
-5.203449e-43   sink_pre_prot[r]
1.715772e-29   SO4CLtex2
-3.420915e-29   SO4HCOtex
1.501093e-14   SPODM
1.365210e-27   SQLEr
1.365210e-27   SQLSr
1.823522e-28   SUCCt2m
-1.823522e-28   SUCCtp
-2.841176e+01   SUCD1m
3.011690e-14   SUCOASm
-9.913894e-16   TALA
-9.838212e-13   TETHEX3COAtx
9.838212e-13   TETPENT3COAtx
         1   THFt2
1.158824e+01   THFtm
-9.913894e-16   TKT1
-9.913894e-16   TKT2
2.000000e+01   TPI
8.835294e+01   TRDR
4.407539e-12   UAG2EMAi
4.407539e-12   UDPG4E
1.671643e-28   UDPGALt2g
2.203770e-12   UDPGALtg
-2.203770e-12   UGLCNACtg
-1.101885e-12   UMPK
-1.929902e-27   UPPN
1.365210e-27   VD3
1.906709e-30   EX_fad[e]
-1.385371e-14   EX_HC00250[e]
2.085115e-30   r0047
-2.317647e+01   r0081
2.841176e+01   r0122
2.203770e-12   r0153
2.317647e+01   r0160
2.203770e-12   r0165
-2.841176e+01   r0178
-1.385371e-14   r0193
-9.916964e-12   r0208
-2.203770e-12   r0233
-2.203770e-12   r0265
-2.203770e-12   r0268
-1.543775e-11   r0280
-1.870673e-27   r0287
5.533994e-13   r0309
-2.730421e-27   r0311
-6.551998e-15   r0386
7.771523e-15   r0394
7.345899e-13   r0400
2.644889e-11   r0413
```

-6.946765e-29	r0443
-7.378659e-13	r0463
1.000000e+00	r0512
-8.820760e-12	r0531
-2.841176e+01	r0541
2.317647e+01	r0558
2.317647e+01	r0559
-1.870673e-27	r0633
5.533994e-13	r0638
5.533994e-13	r0639
5.533994e-13	r0652
5.533994e-13	r0653
-6.551998e-15	r0656
-5.533994e-13	r0660
2.203770e-12	r0668
-4.835294e+01	r0686
5.533994e-13	r0714
5.533994e-13	r0716
5.533994e-13	r0718
5.533994e-13	r0720
5.533994e-13	r0722
-5.533994e-13	r0724
-1.870673e-27	r0730
-1.870673e-27	r0731
-1.870673e-27	r0732
-1.870673e-27	r0733
-1.870673e-27	r0734
5.533994e-13	r0735
1.365210e-27	r0781
1.051353e-28	r0782
-1.365210e-27	r0783
1.870673e-27	r0791
2.841176e+01	r0801
2.841176e+01	r0838
-4.146872e-15	r0870
2.841176e+01	r0885
-2.730421e-27	r0936
2.730421e-27	r0937
-1.385371e-14	r0940
2.841176e+01	r0941
-6.946765e-29	r1008
4.146872e-15	r1106
5.567988e-28	r1156
7.378659e-13	r1292
-6.750878e-29	r1400
-5.716815e-29	r1418
1.716725e-14	r1423
8.835294e+01	r1431
8.835294e+01	r1433
6.274180e-15	r1441
-6.852941e+00	r1453
2.213598e-12	r1464
-8.807375e-12	r2425

```
5.533994e-13   r2435
5.420588e+01   r2520
-1.106799e-12   RE0565C
-1.106799e-12   RE0566C
-1.106799e-12   RE0567C
-1.106799e-12   RE0568C
1.365210e-27   RE1303C
1.762813e-11   RE1530M
1.501093e-12   CYOOm3
-7.378659e-13   C30CPT1
5.533994e-13   FAOXC101C8m
-9.838212e-13   FAOXC15NADx
-1.870673e-27   FAOXC4C2m
5.533994e-13   PMTCOAFABP1tc
1.365210e-27   1a25DHVITD3TRn
1.365210e-27   25HVITD3c
-3.938571e-30   3HCO3_NAt
2.841176e+01   4ABUTtcn
-8.820760e-12   DATPtm
1.762813e-11   DGTPtm
1.365210e-27   DM_1a25dhvitd3[n]
2.841176e+01   DM_4abut[n]
1.106799e-12   DM_pmtcoa[r]
-9.833286e-30   EX_q10h2[e]
4.146872e-15   FADDPle
4.146872e-15   FMNALKPle
2.730421e-27   FRDPtcr
2.330847e-27   IPDDI
8.406606e-18   NADtm
2.459553e-13   NADtx
-9.833286e-30   Q10H2e
-9.913894e-16   RPEc
-9.833286e-30   q10h2tc
1.000000e+00   DM_fol
-2.974168e-15   DM_ncam
```

3. Metabolite dilution flux balance analysis (mdFBA)

This is a variant of FBA for predicting metabolic flux distributions by accounting for growth-associated dilution of all metabolites in a context-dependent manner[3].

A solution from the function `mdFBA` supports that all metabolites used in any reaction of the solution can either be produced by the network or taken up from the surrounding medium.

Timing:

Since this is a MIXED Integer Problem it can take a long time to solve.

Calculating ATP energy production under aerobic condition using mdFBA:

In this function, there is an optional output `newActives`, that represent reactions that are only active in this analysis.

```
% The valid solution can be produced with the Recon 3.0 model
```

```
% modelmd = model;
% modelmd = changeRxnBounds(modelmd, 'EX_glc_D[e]',-20,'l');
% modelmd = changeRxnBounds (modelmd, 'EX_o2[e]', -1000, 'l');
% modelmd = changeObjective(modelmd, 'DM_atp_c_');

% [sol, newActives] = mdFBA(modelmd)
```

- **Troubleshooting:**

When a model does not have a feasible solution, add the input: `'getInvalidSolution', true`.

```
% clear modelmd
modelnosol = modelalter;
modelnosol = changeObjective(modelnosol, 'DM_atp_c_');
[sol, newActives] = mdFBA(modelnosol,  'getInvalidSolution', true)
```

```
sol =

     []
newActives =

  0×0 empty cell array
```

Sometimes when an FBA analysis of a model with the same objective function and constraints is run many times, or using different LP logarithm, we may get different set of solutions for individual reactions. In other words, there are different sets of `'FBAsolution.x'` values (fluxes of the reactions) and still get the same objective function value `'f'`. Therefore, the opitmal solution is not unique. This can create difficulty when investigating the changes to fluxes between two different conditions. In this case a unique solution is required to compare the changes to fluxes.

This issue can be solved with wither of the following the methods

- `geometricFBA`, which provides a standard, central and reproducible solution, or
- `pFBA`, which provides a solution based on the minimal fluxes through the model, and classify each gene according to how it contributes to the optimal solution.

4. Geometric FBA

The geometric FBA solves the smallest frame that contains all sets of optimal FBA solutions and posts a set of multiple linear programming problems[4].

This FBA analysis applies iterations, where by each iteration reduces the permissible solution space. After a finite number of iterations, it resolves one single solution of the flux distribution.

```
% USAGE:
% flux = geometricFBA(model, varargin)
```

Timing:

The time to determine a geometric FBA solution depends on the size of the genome-scale model and the number of iterations. For a model with more than 10,000 reactions and several iterations takes ≥ 30 minutes.

Calculating ATP energy production under anaerobic conditions using geometric FBA:

```
modelgeo = modelalter;
% For Recon3.0 model
% modelgeo = changeRxnBounds (modelgeo, 'EX_glc_D[e]', -20, 'l');
modelgeo = changeRxnBounds(modelgeo, 'EX_glc[e]',-20,'l');
modelgeo = changeRxnBounds (modelgeo, 'EX_o2[e]', 0, 'l');
modelgeo = changeObjective(modelgeo, 'DM_atp_c_');
% WARNING: Depending on the size of the model running this function might take very long;
% FBAgeo = geometricFBA (modelgeo, 'flexRel', 1e-3);
```

Display the unique fluxes from reactions, that are non-zero in the geometric FBA solution.

```
% for i=1:length(FBAgeo)
%     if FBAgeo(i)~=0
%         fprintf('%10d \t %s\n', FBAgeo(i), modelgeo.rxns{i})
%     end
% end
```

Troubleshooting:

When the algorithm has convergence problems, change one of the optional inputs, `flexRel`, into e.g. `1e-3`. The default is 0 when there is flexibility to flux bounds.

Enter the optional parameters as parameter name followed by parameter value, for example:

`flux = geometricFBA(model, 'epsilon', 1e-9)`

5. Parsimonious enzyme usage Flux Balance Analysis (pFBA)

The pFBA method was developed to achieve higher flux levels when more enzymes are required[5].

After performing the FBA to find the optimal value for the objective function, pFBA gets the answer of an another linear program to determine the flux distribution that minimises the total flux through all metabolic reactions in the model.

Timing:

The time to determine a pFBA solution depends on the size of the genome-scale model and is taking from < 1 minute for a 1,000 reaction model, to 5 minutes for a model with more than 10,000 reactions.

The function is:

```
% [GeneClasses RxnClasses modelIrrevFM] = pFBA(model, varargin)
```

Where 'varagin' includes required inputs:

```
% * 'geneoption' - 0 = minimize the sum of all fluxes in the network,
%                   1 = only minimize the sum of the flux through
%                   gene-associated fluxes (default),
%                   2 = only minimize the sum of the flux through
%                   non-gene-associated fluxes
%
% * 'map' - map structure from readCbMap.m (no map written if empty)
%
% * 'mapoutname' - File Name for map
%
% * 'skipclass' - 0 = classify genes and reactions (default).
%                 1 = Don't classify genes and reactions. Only return
```

```
%                        model with the minimium flux set as upper bound.
```

Given outputs in this function are:

```
% OUTPUTS:
% GeneClasses:  Structure with fields for each gene class
% RxnsClasses:  Structure with fields for each reaction class
% modelIrrevFM: Irreversible model used for minimizing flux with
%               the minimum flux set as a flux upper bound
```

Calculating ATP energy production under anaerobic conditions using pFBA:

```
modelp = modelalter;
% For Recon3.0 model
% modelp = changeRxnBounds (modelp, 'EX_glc_D[e]', -20, 'l');
modelp = changeRxnBounds(modelp, 'EX_glc[e]',-20,'l');
modelp = changeRxnBounds (modelp, 'EX_o2[e]', 0, 'l');
modelp = changeObjective(modelp, 'DM_atp_c_');
[GeneClasses RxnClasses modelIrrevFM] = pFBA(modelp,...
    'geneoption', 0, 'skipclass', 1)
```

```
test
netFlux fluxMeasure  ->
GeneClasses =

     []
RxnClasses =

     []
modelIrrevFM =
                      S: [5064×10168 double]
                   rxns: {10168×1 cell}
                     lb: [10168×1 double]
                     ub: [10168×1 double]
                    rev: [10167×1 double]
                      c: [10168×1 double]
             rxnGeneMat: [10168×2194 double]
                  rules: {10168×1 cell}
                  genes: {2194×1 cell}
                grRules: {10168×1 cell}
             subSystems: {10168×1 cell}
               rxnNames: {10168×1 cell}
              rxnKeggID: {10167×1 cell}
    rxnConfidenceEcoIDA: {10167×1 cell}
    rxnConfidenceScores: {10167×1 cell}
             rxnsboTerm: {10167×1 cell}
          rxnReferences: {10167×1 cell}
           rxnECNumbers: {10167×1 cell}
               rxnNotes: {10167×1 cell}
                   mets: {5064×1 cell}
                      b: [5064×1 double]
               metNames: {5063×1 cell}
            metFormulas: {5063×1 cell}
              metCharge: [5063×1 double]
             metCHEBIID: {5063×1 cell}
              metKeggID: {5063×1 cell}
           metPubChemID: {5063×1 cell}
         metInchiString: {5063×1 cell}
        metHepatoNetID: {5063×1 cell}
              metEHMNID: {5063×1 cell}
            ExchRxnBool: [10167×1 logical]
              EXRxnBool: [10167×1 logical]
```

```
            DMRxnBool: [10167×1 logical]
          SinkRxnBool: [10167×1 logical]
          SIntRxnBool: [10167×1 logical]
              metHMDB: {5063×1 cell}
              modelID: 'Recon2.0model'
                match: [10167×1 double]
      reversibleModel: 0
```

Display minimal fluxes of the reactions that are required for producing energy only from only glucose media.

```
for i=1:length(modelIrrevFM.lb)
    if modelIrrevFM.lb(i)~=0
        fprintf('%10d \t %s\n', modelIrrevFM.lb(i), modelIrrevFM.rxns{i})
    end
end
```

```
8.261765e+01    DM_atp_c_
2.725088e+03    netFlux
```

6. Dynamic FBA

The dynamic FBA is an extension of standard FBA that accounts for cell culture dynamics, implementing both dynamic (nonlinear programming) and static (LP) optimisation of an objective function and applying constraints to the rates of change of flux in addition to the standard FBA constraints[6].

The dynamic FBA method implemented in this function is essentially the same as the method described by Varma A. and B. O. Palsson[7].

```
modeldinamic = model;
% For Recon3.0 model
% modeldinamic = changeRxnBounds (modeldinamic, 'EX_glc_D[e]', -20, 'l');
modeldinamic = changeRxnBounds (modeldinamic, 'EX_glc[e]', -20, 'b');
modeldinamic = changeRxnBounds (modeldinamic, 'EX_o2[e]', -1000, 'l');
modeldinamic = changeRxnBounds (modeldinamic, 'EX_ac[e]', -1000, 'l');
% For Recon3.0 model
% smi = {'EX_glc_D[e]' 'EX_ac[e]'};
smi = {'EX_glc[e]' 'EX_ac[e]'};
% exchange reaction for substrate in environment

smc = [10.8]; % Glucose, Acetate concentration (all in mM)

Xec = 0.001; % initial biomass
dt = 1.0/1000.0; % time steps
time = 1.0/dt; % simulation time

[concentrationMatrix, excRxnNames, timeVec,...
    biomassVec] = dynamicFBA(modeldinamic, smi, smc, Xec, dt, time, smi );
```

7. Relax FBA

Find the minimal set of relaxations on bounds and steady-state constraint to make the FBA problem feasible.

```
modelrelax = modelalter;
FBArel = relaxFBA(modelrelax)
```

The output `FBArel` contains solution fields, where

`FBArel.v` is the reaction rate;

`FBArel.r` is set of reactions that need relaxation on steady state constraints `S*v = b`;

`FBArel.p` is relaxation on lower bound of reactions;

`FBArel.r` is relaxation on upper bound of reactions;

8. Flux enrichment analysis (FEA)

The flux enrichment analysis calculates the likelihood that a set of fluxes would belong to a subsystem or pathway.

Timing:

The time to calculate the FEA is < 1 second for any size of a model.

```
modelfea = model;
res = optimizeCbModel(modelfea,'max');
% say you are interested in enriching the active reactions
activeReactions = find(res.x)
% You can also look for e.g. positive/negative/zeros flux reactions,
% that depends pretty much on the question.
% Now you look for the enrichement of reactions per subsystems
resultCell = FEA(modelfea, activeReactions, 'subSystems')
```

REFERENCES

[1] Orth, J. D., Thiele I., and Palsson, B. Ø. What is flux balance analysis? *Nat. Biotechnol.,* 28(3), 245–248 (2010).

[2] Thiele, I., et al. A community-driven global reconstruction of human metabolism. *Nat. Biotechnol.,* 31(5), 419–425 (2013).

[3] Benyamini, T, Folger, O., Ruppin, E., Schlomi, T. Flux balance analysis accounting for metabolite dilution. *Genome Biology*., 11(4):R43 (2010).

[4] Smallbone, K., and Simeonidis, E. Flux balance analysis: A geometric perspective. *J Theor Biol*., 258: 311-315 (2009).

[5] Lewis, N.E., et al. Omic data from evolved E. coli are consistent with computed optimal growth from genome-scale models. *Mol Syst Biol*., 6:390 (2010).

[6] Mahadevan, R., Edwards, J.S., Doyle, F.J. Dynamic Flux Balance Analysis of Diauxic Growth in Escherichia coli. *Biophys J.,* 83(3):1331-1340 (2002).

[7] Varma A. and Palsson, B. Ø. Stoichiometric flux balance models quantitatively predict growth and metabolic by-product secretion in wild-type Escherichia coli W3110. *App Environ Microbiol*., 60(10):3724-3731 (1994).