1 Introduction

MicrobesFlux is a web-based platform for building, modifying, and analyzing metabolic models of multiple organisms. MicrobesFlux uses LIGAND database and KGML files from Kyoto Encyclopedia of Genes and Genomes (KEGG). MicrobesFlux is designed by Dr. Xueyang Feng and You Xu at Washington University in St. Louis.

MicrobesFlux uses the Google Web Toolkit for Web interface and uses the Django web framework for server logic and database management. Metabolic models are solved in MicrobesFlux by IPOPT (Interior Point OPTimizer), a nonlinear optimization solver.

1.1 Availability, licenses and contacts

MicrobesFlux is freely accessed via [http://tanglab.engineering.wustl.edu/static/MicrobesFlux.html](http://tanglab.engineering.wustl.edu/static/MicrobesFlux.html). It supports modern Web browsers such as Google Chrome, Mozilla Firefox, or Safari. You are welcome to contact any of us for questions and comments:

Xueyang Feng, Ph.D.
Dept. of Energy, Environmental & Chemical Engineering,
Washington University in St. Louis
fengx@seas.wustl.edu
Tel: 314-348-6578

You Xu
Dept. of Computer Science & Engineering,
Washington University in St. Louis
youxu@wustl.edu
Tel: 314-935-9187
1.2 Systems architecture

MicrobesFlux is designed with three high-level components: the logic level, the application level, and the achievement level (Figure 1).
In the logic level, the basic principles for metabolic model reconstruction and constraint-based flux analysis are summarized. In additional, the KEGG LIGAND database and KGML files are used as fundamental database in MicrobesFlux.

In the application level, organism-specific metabolic networks are loaded by reading KGML files of the designated organism. Metabolic reactions in the metabolic networks are cross-referred to KEGG LIGAND database. A seed metabolic model is generated and pursued for customized reconstruction and constraint-based flux analysis.

In the achievement level, a reconstructed metabolic model is formulated into an AMPL model files as either a problem, by designating the objective function and the boundary conditions of intracellular fluxes. We can conduct both flux balance analysis (FBA) or a dynamic FBA (dFBA) from the same metabolic model.

The FBA or dFBA problem is solved by IPOPT, a nonlinear optimization solver running on a cluster of computers. The achievement level can provide the reconstructed metabolic networks recorded in the SBML format (The Systems Biology Markup Language), and visualize the metabolic networks using the Scalable Vector Graphics (SVG) format. The calculated flux distributions are sent to users through email.

1.3 Caveats

When using MicrobesFlux to generate and reconstruct a metabolic model, please be aware that the reactions extracted from KEGG LIGAND database and KGML files are not always balanced. For example, the protons (H+, KEGG ID: C00080) sometimes will be absent in certain reactions. If the reactions imbalance is detrimental to the metabolic model, users do can manually curate the corresponding reactions in MicrobesFlux. However, automatic curation of un-balanced reactions in the KEGG database is beyond the current scope of MicrobesFlux. We plan to address it in the future versions.

Since the locations of metabolic reactions are not specified in KEGG LIGAND database or KGML files, the information about metabolic compartment (e.g., mitochondrial) cannot be assigned in the generated metabolic models.

2 Basic functions of MicrobesFlux

2.1 Welcome page

The landing page of MicrobesFlux (Figure 2) has five tabs that provide users the basic information of the functionality, architecture, and frequently asked questions about MicrobesFlux. A demo video is also available as a good start for beginners to learn how to use MicrobesFlux.
2.2 Register an account and login

Users are required to login to use MicrobesFlux. New users can click the “Register” link right below the MicrobesFlux logo and follow the instructions to register an account (Figure 3).

Registered users can login MicrobesFlux by clicking the “Login” link below the MicrobesFlux logo (Figure 4).

In order to change his/her password, a user can login to MicrobesFlux and click the Account Management stack to input the new password (Figure 5).

2.3 Troubleshooting

Question: Why do I see an “you have to log in to create a new model” error when I click “New Model”?

Answer: In MicrobesFlux, a user can save his/her customized model and get optimization results sent to his/her email address. In order to do that, MicrobesFlux needs to associate models to user accounts. Thus, a user has to log in to MicrobesFlux first.
Figure 3: User registrations form in MicrobesFlux

Figure 4: User login form in MicrobesFlux
3 Model reconstruction from KEGG

3.1 Create a new model

In order to generate a seed model for the interested organism, click the New model link in the Build/Load/Save a Model stack. You can give a name to the model so that it can be saved and loaded in the future. An email address is needed for getting the various model results (e.g. SBML file of reconstructed models). In the “Input KEGG Organisms” box, type the three-letter organism code (e.g. eco for *Escherichia coli* K-12 MG1655) or the organism name. Click “Run” to start reconstructing the metabolic model (Figure 6).

If you are not sure about the name of your organisms, click the link of “Input KEGG Organisms”. The link will lead you to the list of organisms in the KEGG database (Figure 7).
This is an image of a page from a document, showing a table titled "KEGG Organisms: Complete Genomes." The table lists various organisms categorized under different groups such as Mammals, Vertebrates, Eukaryotes, Birds, Reptiles, Amphibians, and Fishes. Each entry includes the abbreviation, species name, and source (RefSeq, GenBank, Ensembl). The image is labeled as Figure 7: Organisms list in KEGG database.
3.2 Save & load a model

In order to save a model, simply click the “Save model” link whenever you feel like. If you wish to save the model with another name, click “Save model As” (Figure 8).

![Image of MicrobesFlux interface]

Figure 8: Save models using a new name

In order to load a model that is previously saved, click the “Load model” link. A list of saved models can be found in the drop down list (Figure 9). Choose the one that you want to load and click the “Load” button.

3.3 Troubleshooting

**Question 1:** Why do I see a page showing “Finding records that match your criteria”?

**Answer:** It happens when there is a communication problem between the server and the client. Users can simply refresh the webpage and re-login to load the same model and continue working on it without losing any data.

**Question 2:** Why do I get the window alert (Figure 10) about saving the current model before loading another one?

**Answer:** The window alert is a simple reminder for you to save the current model. If you have already saved the model or want to discard changes to the current model without saving it, the current model, click “OK” to proceed. Otherwise, click the “Cancel” button and MicrobesFlux will keep your current model intact.
Figure 9: Load a model

Figure 10: Window alert during model loading
4 Metabolic models drafting

4.1 Genome information

After generating a seed model or load a previous model loaded in, the genome information of the model can be viewed by clicking “Genome Information” link in the “Pathway & FBA” stack (Figure 11). The number of genes and annotated genes, as well as the number of pathways and active pathways are calculated based on the genome sequencing stored in KGML files of the designated organisms.

4.2 Metabolic pathways manipulations

The metabolic pathways of the designated metabolic model can be viewed and manipulated by clicking “Metabolic Pathways” in the “Pathways & FBA” stack (Figure 12). Users can choose to regroup metabolic reactions based on the metabolic reactions can be grouped by pathways (Figure 13a), reaction directions (i.e., arrows, Figure 13b), or other features such as pathway knock-out, reactants, and products.

To reconstruct a metabolic model, diverse tools are provided in MicrobesFlux, including 1) adding inflow/outflow reactions; 2) implementing biomass production reaction; 3) pathway knock-out; and 4) introducing heterologous pathways.
Figure 12: Pathways of a metabolic models

Figure 13: Grouping metabolic reactions
4.3 Adding inflow/outflow reactions

The inflow/outflow reactions are the transport reactions between two metabolic compartments, cytosol and extracellular medium. These inflow/outflow reactions are determinant for building metabolic models. However, they are not originally included in the seed metabolic model generated from KGML files and KEGG LIGAND database.

To add inflow/outflow reactions in MicrobesFlux, in the Add reactions module, choose the reaction type as inflow (outflow), then put the metabolites that are exchanged between cytosol and the extracellular medium (Figure 4.4). The reactant will have the “.ext” suffix for inflow reactions that transport metabolite from extracellular medium to cytosol. The products will have the suffix “.ext” for outflow reactions that transport metabolite from cytosol to extracellular medium.

To upload the added inflow/outflow reactions into the metabolic model, click Validate Pathway button first to make sure that the names of metabolites are consistent with metabolites names in the MicrobesFlux database. A complete list of compound names that are recognizable can be found in http://tanglab.engineering.wustl.edu/media/valid_compounds.html.

Once the validation is done and the metabolites names are valid, click the “Add” button will add the inflow/outflow reactions to the metabolic model (Figure 14).

4.4 Implement biomass reactions

The biomass production reaction can be manually added in a similar fashion as shown above. In general, choose Biomass in Reaction Type and put the biomass compositions in the reactants column (Figure 4.4). Then validate pathway and add it into the reconstructed model. Note: The biomass production reaction is necessary for FBA studies in “Optimization” module.
4.5 Knocking-out and changing reactions from metabolic models

To manipulate a pathway, choose the designated reaction by clicking the reaction row in the column (Figure 15). A separate table will be present, which contains “Reactants”, “Products”, “Arrows”, “Pathway”, and a “KO” box. Click the “KO” box to knock out a pathway, or change the reactants and products of the reaction as your will. The arrows of the reaction can also be switched between “<==” and “==>”. 

4.6 Export metabolic models

The metabolic model can be extracted in SBML format by clicking the “Get SBML” button (Figure ??). Also, the metabolic model can be visualized in SVG format by clicking the “Get Pathway Map” button. The SBML file (Figure 17) and SVG (Figure 18) file will be sent to users email.

4.7 Troubleshooting

Question 1: Why the Add button cannot be activated?
Figure 17: The SBML file of a metabolic model

Figure 18: Metabolic network visualization (SVG)
Answer: Some of the metabolites in KEGG LIGAND database cannot be recognized by MicrobesFlux. A complete list of compound names that are recognizable can be found in [http://tanglab.engineering.wustl.edu/media/valid_compounds.html](http://tanglab.engineering.wustl.edu/media/valid_compounds.html). In such case, we suggest users to use the KEGG Compound ID (e.g. C00001 for $H_2O$) instead.

**Question 2: Do I have to input biomass production reaction manually?**

Answer: Yes, the biomass production reaction has to be manually input in MicrobesFlux, based on two considerations.

First, the biomass compositions in different organisms are not the same. Therefore, a customized biomass production reaction entry is necessary.

Second, the completeness of biomass production reaction depends on the character of metabolic models. For a large-scale metabolic model, such as genome-scale metabolic model, the complete biomass composition including compositions of amino acids, lipids, nucleic acids, etc, is required. However, for a simplified metabolic model, it is not necessary to input all the biomass compositions into consideration. To cope with the diverse demands of metabolic models, we let users decide what should be input as biomass production reaction.

**Question 3: How can I delete the inflow/outflow reactions?**

Answer: Click the “KO” button before the inflow/outflow reactions to remove these reactions from the metabolic model.

### 5 Flux balance analysis (FBA) of metabolic models

The mathematical representation of FBA is

\[
\min \sum_i c_i v_i \\
\text{s.t.} \quad S v = 0 \\
\quad \quad \quad \quad lb \leq v \leq vb
\]

This linear optimization can be solved by optimization solvers such as CPLEX and IPOPT, with the set-up of three modules: objective function, flux balance equations, and boundary conditions for each metabolic reaction. The flux balance equations will be automatically generated by MicrobesFlux, while the objective function and boundary conditions will need to be identified by users.

#### 5.1 Designate the objective function

Two types of objective functions can be designated for flux balance analysis in MicrobesFlux: maximizing biomass, or maximizing a customer-defined objective function (Figure 19).
Figure 19: Setting up the objective function

If maximizing biomass is chosen as the objective function, please choose “Biomass” in the objective function column and make sure that the biomass production reaction has been included in the metabolic model (Figure 5.1). If maximizing a customer-defined objective function (i.e., $\max \sum c_i v_i$) is chosen as the objective function, the weighting factor of each metabolic reaction can be manually tuned in the “Optimization” stack (Figure 20).

Figure 20: User-defined objective function

5.2 Generate flux balance equations

The flux balance equations (i.e., $Sv = 0$) are automatically generated by MicrobesFlux, based on the metabolic model reconstructed by users (Figure 21). To change the flux balance equations, users have to make new modifications in “Metabolic Pathways” module.
5.3 Set the boundary conditions

The boundary conditions of each metabolic flux can be set (shown in Figure 22). \( l_b \) indicates the lower bound of the chosen reaction rate while \( u_b \) indicates the upper bound of the chosen reaction rate. The \( l_b \) cannot exceed \( u_b \). When \( l_b = o_b \), the metabolic flux is effectively fixed to \( l_b \) (or \( u_b \)).

5.4 Submit an FBA job

Once the objective function is designated and the boundary conditions are set, click the “Submit FBA Job” to submit the request to MicrobesFlux for solving the FBA problem. It will normally take 15 minutes to get the email from MicrobesFlux, with the result file attached.

5.5 Dynamic flux balance analysis

MicrobesFlux also provide dynamic flux balance analysis (dFBA) for the same metabolic model reconstructed by users. Click the “Submit for Dynamic FBA” button to start the setting for dFBA in MicrobesFlux (Figure 23).

The dFBA in MicrobesFlux decomposes the dynamic microbial metabolism into numerous steady state metabolisms with various inflow/outflow fluxes (Figure 23). Therefore, an additional data file (in txt format) is needed to identify
the inflow/outflow fluxes at each of the steady-state metabolism (Figure refdfba-data).

Figure 24: A sample dFBA data used in MicrobesFlux

After uploading the data file to MicrobesFlux, click “Submit DFBA Job” to send the request for solving the dFBA problem. It will normally take one to five minutes for users to get the email from MicrobesFlux, with the result file attached.

### 5.6 Troubleshooting

**Question 1: Why cannot I receive any email from MicrobesFlux?**

**Answer:** Please check the spam box to see if there is any email. For dFBA job, please make sure that the data file uploaded follows the format as that shown in [http://tanglab.engineering.wustl.edu/media/sample.txt](http://tanglab.engineering.wustl.edu/media/sample.txt). If there is still no email from MicrobesFlux, please contact us and report it as an issue using [http://code.google.com/p/kegg-dfba/](http://code.google.com/p/kegg-dfba/).

**Question 2: Why the results of FBA/dFBA do not make any sense?**

**Answer:** Please double-check the metabolic model reconstructed and makes sure that the gaps (such as the dead-end metabolic reaction, co-factor imbalance, etc) are filled in the reconstructed metabolic model. The protocol for generating a high-quality genome-scale metabolic reconstruction can be found in “Thiele I,
6 A case study in MicrobesFlux

Here we present a case study by using MicrobesFlux to reconstruct a medium-scale metabolic model for *Thermoanaerobacter* sp. strain X514. The drafted model is consisted of 196 metabolites and 229 reactions (162 intracellular reactions, 19 inflow/outflow reactions, 39 gap-filling reactions, and 9 biomass-producing reactions). The major steps in metabolic model reconstruction are listed as below:

**Step 1**: login to MicrobesFlux and build a seed model. Choosing “tex *Thermoanaerobacter*_X514” as the organism (Figure 25).

![MicrobesFlux](image)

Figure 25: Build a seed model for *Thermoanaerobacter* sp. strain X514

**Step 2**: Input inflow/outflow fluxes in the reconstructed model, based on the literature survey of identified transporters in *Thermoanaerobacter* sp. strain X514 and gap-fill requirements of the model. To delete inflow/outflow fluxes, choose the reactions and mark as “KO” (Figure refinflow-outflow).

**Step 3**: Input biomass production reaction and heterologous pathways in the reconstructed model. Here, we use the following reaction for biomass production:

\[
0.011 \text{NAD} + 0.009 \text{NADP} + 0.01 \text{CoA} + 0.016 \text{Tetrahydrofolate} \]
\[
+ 0.528 \text{C00017} + 0.066 \text{C00046} + 0.026 \text{C00039} + 0.076 \text{C01356} \\
+ 0.101 \text{C00889} + 0.08 \text{C06707} + 0.043 \text{C00182} + 220.0 \text{ATP} \rightarrow \text{Biomass}
\]

The gaps in metabolic models can come from multiple sources. First, depends on the size and purpose of the metabolic model, some metabolic reactions need to be changed (e.g. Knock-out), which will be addressed in **Step 4**. Second, some metabolic reactions may be mis-annotated in the seed model or introduced as a foreign reaction from other biological systems. These pathways will be input in the reconstruction step as “Heterologous Pathways”.

Palsson BØ. A protocol for generating a high-quality genome-scale metabolic reconstruction. Nat Protoc. 2010 Jan;5(1):93-121
The KEGG Compound IDs are always suggested to be used due to its comparability with MicrobesFlux. Some compound, such as NADP can also be recognized in MicrobesFlux. A complete list of compound names that are recognizable can be found in [http://tanglab.engineering.wustl.edu/media/valid_compounds.html](http://tanglab.engineering.wustl.edu/media/valid_compounds.html).

Step 4: Manipulate pathways in the metabolic model. Since the draft model is a simplified model focusing on carbohydrate metabolism and amino acids biosynthesis, pathways that are not of interest in the reconstructed model can be deleted by marking the pathway as “KO”. The cofactors involved in the reactions and the pathway directions can also be manipulated (Figure 28).

Step 5: Output the reconstructed model as SBML. Once the reconstruction of metabolic model is finished, the SBML file can be generated. The SBML files can then be read and used by other fluxomics software packages (Figure 29).

Step 6: Set the objective functions and boundary conditions for FBA. Here, we choose “maximize biomass” as the objective function, and set the boundary conditions for each metabolic flux. Certain inflow/outflow fluxes are needed to be assigned with fixed values.

Step 7: Get the result file of the reconstructed model. The result file will include three parts: 1) Genome Info; 2) Pathways and 3) AMPL file of the reconstructed model. In the AMPL file, the objective function, flux balance equations and the boundary conditions are included. The simulation results can be found in the “Results” section of AMPL file.
Figure 27: Input biomass production reaction in the reconstructed model of *Thermoanaerobacter* sp. strain X514

Figure 28: Manipulate pathways in the reconstructed model of *Thermoanaerobacter* sp. strain X514
Figure 29: SBML files for the reconstructed model of *Thermoanaerobacter* sp. strain X514

Figure 30: FBA setting for the reconstructed model of *Thermoanaerobacter* sp. strain X514
Figure 31: Result file of FBA studies for reconstructed model of *Thermoanaerobacter* sp. strain X514