Metaflux User guide

Overview

Metaflux is a web-based platform that brings together genome-scale metabolic simulations and protein function prediction tools for metabolic engineering.

This page explains:

- Section 1: What Metaflux is and how it fits into your workflow
- Section 2: How to run analyses with each Metaflux tool
- Section 3: How to interpret the results

Section 1: What Metaflux is and how it fits into your workflow

Metaflux supports two major categories of analysis:

1. Network-level analysis using genome-scale metabolic models
    - FBA: Predicts growth rate and target reaction flux by simulating individual reaction knockouts.
    - FSEOF: Predicts reactions whose flux increases as flux toward product formation is progressively enforced.
    - FVSEOF: Scans how the variability of each metabolic flux changes in response to an enforced flux toward target product formation, enabling the identification of gene amplification targets.
    - iBridge: Predicts overexpression and downregulation targets by identifying bridge reactions that convert negative metabolites to positive metabolites as overexpression targets, and the opposites as downregulation targets.

2. Sequence- and reaction-level deep learning tools
    - DeepTFactor: Predicts whether a protein sequence encodes transcription factor (TF), supporting regulatory network reconstruction.
    - DeepECtransformer: Annotates enzyme commission (EC) numbers from protein sequences, allowing the functional annotation of uncharacterized genes.
    - DeepRFC: Evaluates whether a proposed reaction is feasible based on the use of reactant pairs.

How Metaflux fits into your design process

A typical workflow might proceed as follows:

1. Upload a genome-scale metabolic model, a protein sequence, or the SMILES strings of a substrate–product pair that you want to analyze.
2. Use FBA / FSEOF / FVSEOF / iBridge to explore production capabilities and identify promising engineering targets.
3. Use DeepTFactor and DeepECtransformer to functionally annotate proteins relevant to your design.
4. Use DeepRFC to evaluate whether newly proposed reactions are chemically plausible.
5. Incorporate the selected targets and validated reactions into your experimental strain design workflow.

Section 2: How to run analyses with each Metaflux tool

2.1. Running genome-scale metabolic model-based analyses (FBA, FSEOF, FVSEOF, iBridge)

Step 1. Enter your job name.

Step 2. (Optional) Enter your email address.

You may provide an email address to receive updates when your analysis:

- begins,
- finishes successfully, or
- encounters an error.

Step 3. Upload your metabolic model (xml format).

Upload a genome-scale metabolic model in XML format. After uploading, Metaflux will automatically parse the model and display:

- the biomass reaction ID (auto-filled but editable)
- all reaction IDs available for searching and selection.

If the model structure is invalid, Metaflux will return an error before proceeding.

Step 4. Search and select reaction IDs.

Use the built-in search bar to look up reaction IDs included in the model. This step is required to select:

- Target reaction
- Reaction knockouts (you may specify one or more reactions)

If an entered reaction ID does not exist in the uploaded model, Metaflux will display an error message prompting correction.

Step 5. Submit your analysis.

Once all required fields are completed, click Submit. Your analysis will be placed in the processing queue.

2.2. Running deep learning models for sequence analyses (DeepTFactor and DeepECtransformer)

Step 1. Enter your job name.

Step 2. (Optional) Enter your email address.

You may provide an email address to receive updates when your analysis:

- begins,
- finishes successfully, or
- encounters an error.

Step 3. Provide your protein sequence (FASTA format).

You may paste the sequence directly into the input box, or upload a FASTA file. A sample FASTA file is available for reference to ensure correct formatting. For DeepECtransformer, you may additionally request sequence logo generation when a single protein sequence is entered. Moreover, the amino acid length that can be analysed by DeepECtransformer is limited to 2,500 amino acids, whereas DeepTFactor supports the analysis of protein sequences up to 1,000 amino acids in length.

Step 4. Submit your analysis.

Once all required fields are completed, click Submit. Your analysis will be placed in the processing queue.

2.3. Running reaction feasibility checks with DeepRFC

Step 1. Enter your job name.

Step 2. (Optional) Enter your email address.

You may provide an email address to receive updates when your analysis:

- begins,
- finishes successfully, or
- encounters an error.

Step 3. Provide your SMILES strings of a substrate–product pair (txt format).

You may type the SMILES strings of the reactant and product directly into the input table, or upload a text file. Before uploading, please ensure the following requirements are met:

Required input format

- Each line must contain one reactant and one product, forming a 1:1 pair.
- DeepRFC does not support reactions with multiple reactants or multiple products (e.g., "A + B → C").
- Reactant and product SMILES strings must be separated by a single tab character (\t).
- Using spaces, commas, or multiple tabs will cause the system to return an error.
- Refer to the provided sample input file to ensure correct formatting.

Step 4. Submit your analysis.

Once all required fields are completed, click Submit. Your analysis will be placed in the processing queue.

Section 3: How to interpret the results

3.1. Interpreting results from FBA

Output file: fba_knockout_result.txt

This file summarizes the predicted metabolic flux behavior after systematically removing each reaction from your model. Each row corresponds to a single reaction knockout.

Column	Meaning
Reaction(s)	Reaction ID removed from the model.
Biomass flux	Predicted growth rate of the knockout strain. Values near zero indicate essential reactions whose deletion severely impairs viability.
Target reaction flux	Flux through your specified product formation reaction under the knockout scenario. Higher values suggest beneficial knockouts that redirect fluxes toward your product.
Carbon source uptake flux	Carbon source uptake constraint used during simulation.

3.2. Interpreting results from FSEOF

FSEOF identifies reactions whose fluxes increase as the flux toward the target reaction is enforced upward. Such reactions are strong candidates for gene amplification.

Output files: fseof_result_df.csv and fseof_summary_result.csv

1) fseof_result_df.csv – raw data for whole reactions

Column	Meaning
Reaction	Reaction ID whose flux trajectory was analyzed.
Pearson correlation coefficient	Pearson correlation between reaction flux and target reaction flux across enforced flux levels. Values near 1 indicate a strong positive linear relationship between the reaction flux and the target reaction flux.
Covariance	Statistical measure of how reaction flux and target reaction flux co-vary relative to their means. Positive covariance indicates that the reaction flux and target flux increase or decrease together, whereas negative covariance indicates opposite directional changes. Values near 0 suggest minimal linear co-variation.
Status	“UP” indicates fluxes increases with enforced target reaction flux. These reactions are the primary candidates for overexpression. Reactions without “UP” are excluded from further consideration.

2) fseof_summary_result.csv – simplified list of targets

A streamlined summary that includes only UP-regulated reactions to identify key engineering targets more easily.

3) FSEOF graph visualization

Metaflux provides an interactive plot to help you visualize flux behavior intuitively. The system automatically reads fseof_summary_result.csv and displays only the UP-regulated reactions for plotting.

- x-axis: Enforced target flux values
- y-axis: Flux through each selected reaction.
- Reaction curves: Each curve represents how a reaction's flux changes as the target flux increases.

You may select one or multiple reactions for comparison, allowing you to distinguish between strong and weak flux-response behaviors and prioritize the most promising amplification targets.

3.3. Interpreting results from FVSEOF

FVSEOF scans how the variability of each metabolic flux changes in response to an enforced product formation flux, enabling the identification of gene amplification targets.

Output files: fvseof_result_df.csv, fvseof_result_summary.csv, and fvseof_up_targets.csv

1) fvseof_result_df.csv – raw variability data for each reaction

Column	Meaning
Target reaction	Target product reaction you selected
Step	Index of the enforced flux level
Target reaction flux	Value of the enforced product flux at this step
Biomass flux	Predicted biomass flux under this enforced level
Reaction	Reaction ID whose flux variability is being evaluated
Minimum reaction flux	Minimum feasible flux at this enforced level
Maximum reaction flux	Maximum feasible flux at this enforced level
Average reaction flux	Mean of minimum and maximum reaction fluxes

2) fvseof_result_summary.csv – statistical analysis

This file aggregates variability trends across all steps and highlights how strongly and consistently each reaction responds to enforced product flux.

Key columns

Column	Meaning
Average flux's slope	Average slope describing how the reaction's flux change across enforced steps. A positive slope suggests the reaction flux tends to increase as product flux increases.

Pearson correlation coefficient for the minimum reaction flux (MIN_R)	Minimum correlation coefficient observed across all enforced levels. If MIN_R > 0.9, the reaction consistently aligns with product formation.
Average flux's p-value	p-value for the linear relationship between the target reaction flux and the reaction's absolute average flux. Lower values indicate higher statistical confidence that the observed trend is meaningful rather than random.

3) fvseof_up_targets.csv – final list of up-regulated candidates

This file contains only reactions that pass Metaflux's filtering criteria, which is MIN_R > 0.9, ensuring consistent positive response across all enforced steps.

These are the reactions most suitable for gene amplification targets.

4) FVSEOF graph visualization

Metaflux provides an interactive plot to help you visualize flux behavior intuitively. The system automatically reads fvseof_result_summary.csv and fvseof_result_df.csv, and displays only UP-regulated reactions.

- x-axis: Enforced target flux values
- y-axis: Flux values for each reaction.
- Reaction curves: For each reaction, three curves (minimum, average, and maximum reaction fluxes) are plotted.

You may select one or multiple reactions for comparison, allowing you to see which one exhibits stronger or more stable flux-response trends.

3.4. Interpreting results from iBridge

Metaflux provides a list of candidate reactions, along with key statistics describing whether they positively or negatively affect flux toward the target reaction.

Output file: result_ibridge.txt

1) result_ibridge.txt

Column	Meaning
Target reaction	Candidate reaction predicted to influence the target reaction
Negative metabolite	Metabolite predicted to give negative influences on target production
Positive metabolite	Metabolite predicted to give positive influences on target production
SoC (Sum of Covariance)	Quantifies how strongly metabolite perturbations correlate with target reaction flux changes. Positive metabolite if it receives a positive SoC and negative metabolite if it receives a negative SoC.
Regulation type	Indicates whether the model predicts Up (overexpression) or Down (downregulation) for the corresponding reaction.
Normalized SoC	SoC normalized to the range 0–1, enabling easier comparison across reactions.

3.5. Interpreting results from DeepTFactor

Output file: DeepTFactor_1.csv

1) DeepTFactor_1.csv

Column	Meaning
Sequence ID	Identifier for the input protein sequence
Prediction	Predicted as Transcription Factor (TF) or non-TF
Score	Prediction score

3.6. Interpreting results from DeepECtransformer

Output file: DeepECtransformer.csv

1) DeepECtransformer.csv

Column	Meaning
Sequence ID	Identifier for the input protein
Prediction	Predicts EC number
Score	Prediction score. If DeepECtransformer cannot assign an EC number, the sequence is listed as unassigned, as no DIAMOND-based annotation is added in such cases.

3.7. Interpreting results from DeepRFC

Output file: DeepRFC_1_result.csv

1) DeepRFC_1_result.csv

Column	Meaning
Reaction	Molecular structure of the reactant and product in SMILES format
Predictive mean	Obtained from the ten predictive values of DeepRFC via Monte Carlo dropout
Standard deviation	Standard deviation of ten predictive values
Feasibility	Classification whether the reaction is feasible