**pFBA analysis and comparisons with FBA and FVA-sampling**

This folder contains the core scripts and representative outputs used for parsimonious flux balance analysis (pFBA) and comparative evaluation of differential fluxes against standard FBA and flux variability sampling (FVA-sampling).

All analyses were performed using MATLAB, and visualization or overlap analyses were completed in both MATLAB and R.

**Contents**

* MATLAB scripts: Core scripts for performing pFBA simulations, comparing with FBA and FVA-sampling, and generating annotated reaction-level results and visualizations.
* R scripts: Visualization and overlap analysis of FBA, pFBA, and FVA-sampling results.
* Example data files: Representative comparison results (e.g., FBA\_pFBA\_comparison.mat) used for validation and figure generation.
* Summary sheets: Consolidated Excel summaries and Venn-ready datasets for visualizing method overlaps and flux differences (e.g., Complete\_Analysis\_Results.xlsx, venn\_diagram\_data.xlsx).

**Notes**

Due to file size constraints, the sampling result files (HSD.mat, NSD.mat) and the full comparison dataset (fba\_pfba\_sampling\_comparison.mat) are not included.

Only representative .mat files and summary tables are provided for illustration.

For access to the full dataset or additional information, please contact: [sunjin\_moon@hms.harvard.edu](mailto:sunjin_moon@hms.harvard.edu)

**Summary of scripts**

1. **script\_01\_run\_pFBA.m**
2. Purpose:  
   Perform parsimonious flux balance analysis (pFBA) on NSD and HSD muscle-GEMs.
3. Input:
   1. model\_out\_cbra\_u.mat (NSD and HSD models)
   2. Reaction constraints and reduction factors
4. Output:
   1. pFBA\_results.mat (optimal fluxes, objectives, total flux)
   2. model\_constrained\_out.mat (constrained models)
5. **script\_02\_0\_comp\_FBA\_pFBA.m**
6. Purpose:  
   Compare standard FBA and parsimonious FBA (pFBA) results in HSD and NSD muscle-GEMs.
7. Input:
   1. pFBA\_results.mat (from Script\_01)
   2. model\_out\_cbra\_u.mat (NSD and HSD models with constraints)
8. Output:
   1. FBA\_pFBA\_comparison.mat (summary of fluxes, objectives, and reductions)
   2. Console summaries of flux/objective differences and reaction-level changes
9. **script\_02\_1\_comp\_pFBA\_NAD.m**
10. Purpose:  
    Analyze NAD/NADH/NADPH-associated fluxes in FBA vs. pFBA models to identify reduction patterns and potential NADH cycling reactions.
11. Input:
    1. FBA\_pFBA\_comparison.mat (from Script\_02)
    2. model\_out\_cbra\_u.mat (muscle-GEMs)
12. Output:
    1. NAD\_flux\_analysis.mat (detailed NAD flux results)
    2. NAD\_summary.csv (summary of total NAD flux, reduction %, and cycling reactions)
13. **script\_03\_vis.m**
14. Purpose:  
    Generate comprehensive visualizations and export tables summarizing FBA vs. pFBA and NAD flux analyses.
15. Input:
    1. model\_constrained\_out.mat
    2. FBA\_pFBA\_comparison.mat
    3. NAD\_flux\_analysis.mat
16. Output:
    1. Figures (01–11\_\*.fig/png/pdf) showing total flux, NAD flux, proportions, distributions, and reaction categories
    2. cycling\_reactions.csv and cycling\_reactions\_detailed.csv
    3. all\_nad\_reactions\_detailed.csv
    4. Complete\_Analysis\_Results.xlsx (11 sheets summarizing all results and statistics)
17. **script\_03\_vis\_nad\_cycling\_sampling.m**
18. Purpose:  
    Visualize manually curated NAD(P)-dependent cycling reactions using FBA, pFBA, and FVA-sampling flux data.
19. Input:
    1. cycling\_reactions\_detailed\_analz.csv (manually curated list)
    2. nsd.mat, hsd.mat (flux sampling results)
20. Output:
    1. Figures (01\_Cycling\_Reactions\_Comparison.fig/png/pdf) showing FBA, pFBA, and FVA-sampling fluxes for cycling reactions
    2. cycling\_reactions\_with\_sampling.xlsx (combined data and summary sheets)
21. **script\_04\_compare\_all\_export.m**
22. Purpose:  
    Compare flux distributions and correlations among FBA, pFBA, and flux sampling results for NSD and HSD muscle-GEMs.
23. Input:
    1. model\_constrained\_out.mat
    2. FBA\_pFBA\_comparison.mat
    3. nsd.mat, hsd.mat (sampling results)
24. Output:
    1. fba\_pfba\_sampling\_comparison.mat (includes flux statistics, correlations, and disease difference analyses)
    2. Console summary of flux ranges, correlation coefficients, and potential cycling reactions
25. **script\_05\_compare\_all\_export.m**
26. Purpose:  
    Generate annotated reaction-level comparisons among FBA, pFBA, and sampling results to identify reactions increased or decreased in HSD relative to NSD, and export R-ready Venn diagram data.
27. Input:
    1. fba\_pfba\_sampling\_comparison.mat
    2. model\_constrained\_out.mat
28. Output:
    1. venn\_diagram\_data.xlsx with six sheets: Increased\_Reactions, Decreased\_Reactions, Summary\_Counts, Increased\_Lists, Decreased\_Lists, R\_Binary\_Matrix (for R Venn or UpSet plots with full annotations)
29. **script\_05\_02\_compare\_all\_vis.m**
30. Purpose:  
    Visualize and summarize comparisons among FBA, pFBA, and flux sampling for NSD and HSD models.
31. Input:
    1. fba\_pfba\_sampling\_comparison.mat
    2. model\_constrained\_out.mat
32. Output:
    1. Figures showing inactive reactions, active distributions, total flux, and correlations
    2. Grouped\_ Results.xlsx (correlation matrices)
    3. Figures\_Results.xlsx (five sheets summarizing zero flux, active distribution, total flux, reduction, and manuscript summary)
33. **script\_06\_create\_venn.R**
34. Purpose:  
    Generate and analyze three-way overlaps among FBA, pFBA, and FVA-sampling results (HSD vs NSD) using Venn, Euler, and UpSet plots; perform subsystem- and gene-level enrichment analyses for commonly altered reactions.
35. Input:
    1. venn\_diagram\_data.xlsx (from MATLAB script 05)
36. Output:
    1. Folder: 06\_venn/ containing
       1. Venn & Euler plots: 1–2\_\* (increased/decreased, PNG/SVG)
       2. Overlap summaries: 3\_Venn\_Overlap\_Summary.csv, filtered versions (6–8\_\*)
       3. Subsystem analyses: 4–9\_\*\_Subsystem\_\*.png/.csv/.pdf
       4. Gene analyses: 5–10\_\*\_Gene\_\*.png/.csv
       5. UpSet plots: 10\_UpSet\_\* (increased/decreased, raw & filtered)
       6. Panel D network, table, and bar visualizations linking top subsystems and key genes