Supplementary Text with Tables and Figures Embedded

Supplementary Information 1 - msb4100155-s3.xls

The content of *i*AF1260 in a spreadsheet format (Microsoft Excel workbook). The file contains:

- The metabolites, reactions, thermodynamic data and references for the reconstruction.
- The reconstruction content used to generate Figure 1 with summarized lists of reactions form each COG class for each of the 6 reconstructions considered analysis of the reaction to gene ratio in *i*AF1260
- The regulatory FBA analysis parameters used to determine which reactions are not available to the network during aerobic, glucose, and aerobic + glucose minimal media conditions. Also, comments and adjustments made to the rules established by Covert *et al* (Covert *et al*, 2004).
- Results from the gene essentiality analysis comparing results using *i*AF1260 and experimental screens data (Baba *et al*, 2006; Joyce *et al*, 2006).

Microsoft Excel workbook

Supplementary Information 2 - msb4100155-s4.xls

Results comparing the predicted *in silico* carbon, nitrogen, phosphorous and sulfur sources for *E. coli* using iAF1260 and *i*JR904 and those determined experimentally (http://www.biolog.com). These results are summarized in Table 3. The compounds that do not have an exchange reaction in *i*AF1260 and those currently not in the SimPheny database are sources for additional reconstruction expansion.
Microsoft Excel workbook

Supplementary Information 3 - msb4100155-s5.xls

The biomass objective function used in conjunction with the constraint-based methods utilized in the project.
Microsoft Excel workbook

Supplementary Information 4 - msb4100155-s6.zip

*i*AF1260 in SBML format (level 2, version 1). See below and http://sbml.org/documents/ for SBML format documentation.
Two SBML files are presented with the reconstruction, each containing a different flux distribution
XML files

Supplementary Information 5 - msb4100155-s7.zip

mol files (.mol) of the compounds in *i*AF1260 for which mol files could be created.
Zip File - containing .mol files

Supplementary Information 6 - msb4100155-s8.zip

Results from the comparison of the content of *i*AF1260 and the EcoCyc and MetaCyc databases.
The files contain mappings between the compounds and reactions in the reconstruction and databases.
Zip File - containing text files for the compound and reactions mappings, definition of analysis results (read_me.txt) and a Microsoft Excel workbook containing the same data.

Legend to the Supplementary Material - msb4100155-s1.pdf
Feist *et al.* - A genome-scale metabolic reconstruction for *Escherichia coli* K-12 MG1655 that accounts for 1260 ORFs and thermodynamic information

2

SBML file will contain the following reconstruction content

- reactions (format: 'R_*<reaction abbreviation >*')
  - reaction name, reversibility, reaction stoichiometry, gene-protein-reaction (GPR) association, subsystem, E.C. number
- metabolites (format: 'M_*<metabolite abbreviation>*_*<compartment abbreviation>*')
  - metabolite name, compartment, charge, formula (appended to the end of the name, *<metabolite name>*_FORMULA)
- a flux distribution associated with a steady-state modeling simulation
  - lower bound, upper bound, objective coefficient, flux value, reduced cost

SBML File Properties	
file name	Ec_iAF1260_flux1.xml
organism	E. coli K-12 MG1655
model	iAF1260
Biomass Objective Function (BOF)	Ec_biomass_iAF1260_core_59p81M (E. coli biomass objective function (iAF1260) - core - with 59.81 GAM estimate)
flux balance analysis objective	maximize BOF
Growth Associated Maintenance (GAM)	59.81 mmol ATP gDW-1
Non-Growth Associated Maintenance (NGAM)	8.39 mmol ATP gDW-1 hr-1
media conditions	computational minimal media
carbon source	8 mmol glucose gDw-1 hr-1
aerobic or anaerobic	18.5 mmol O2 gDw-1 hr-1
additional constraints	
reactions constrained to zero	CAT, SPODM, SPODMpp, FHL
flux split between reaction pairs	none

SBML File Properties	
file name	Ec_iAF1260_flux2.xml
organism	E. coli K-12 MG1655
model	iAF1260
Biomass Objective Function (BOF)	Ec_biomass_iAF1260_core_59p81M (E. coli biomass objective function (iAF1260) - core - with 59.81 GAM estimate)
flux balance analysis objective	maximize BOF
Growth Associated Maintenance (GAM)	59.81 mmol ATP gDW-1
Non-Growth Associated Maintenance (NGAM)	8.39 mmol ATP gDW-1 hr-1
media conditions	computational minimal media
carbon source	11.0 mmol glucose gDw-1 hr-1
aerobic or anaerobic	18.2 mmol O2 gDw-1 hr-1
additional constraints	
reactions constrained to zero	152 reactions identified to be unavailable to the cell under glucose aerobic conditions
flux split between reaction pairs	NDH-1:NDH-2 is 1:1 (3 pairs of reactions - different quinone usage) NADH10:NADH17pp, NADH5:NADH16pp, NADH9:NADH18pp

Legend to the Supplementary Material - msb4100155-s1.pdf   3
Feist *et al.* - A genome-scale metabolic reconstruction for *Escherichia coli* K-12 MG1655 that
accounts for 1260 ORFs and thermodynamic information

References

Baba T, Ara T, Hasegawa M, Takai Y, Okumura Y, Baba M, Datsenko KA, Tomita M, Wanner BL,
Mori H (2006) Construction of *Escherichia coli* K-12 in-frame, single-gene knockout mutants: the Keio
collection. *Mol Syst Biol* **2:** 2006.0008.

Covert MW, Knight EM, Reed JL, Herrgard MJ, Palsson BO (2004) Integrating high-throughput and
computational data elucidates bacterial networks. *Nature* **429:** 92-96.

Joyce AR, Reed JL, White A, Edwards R, Osterman A, Baba T, Mori H, Lesely SA, Palsson BO,
Agarwalla S (2006) Experimental and Computational Assessment of Conditionally Essential Genes in
Escherichia coli. J Bacteriol **188:** 8259-8271.