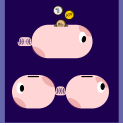
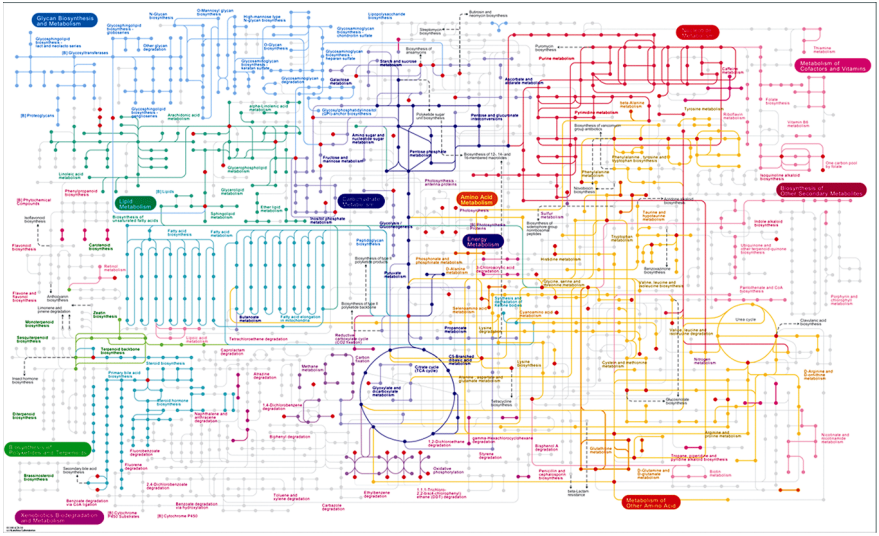


Diversity of metabolic flux distributions

Marcelo Rivas-Astroza & Roberto Mulet

cecam
Centre Européen de Calcul Atomique et Moléculaire
universität wien
Département MathNum

The Problem


Glycan Biosynthesis and Metabolism
Lipid Metabolism
Nucleotide Metabolism
Amino Acid Metabolism
Energy Metabolism
Carbohydrate Metabolism
Xenobiotics Biodegradation and Metabolism
Biosynthesis of Other Secondary Metabolites
Metabolism of Cofactors and Vitamins
Metabolism of Other Amino Acids
Glycan Biosynthesis and Metabolism
Lipid Metabolism
Nucleotide Metabolism
Amino Acid Metabolism
Energy Metabolism
Carbohydrate Metabolism
Xenobiotics Biodegradation and Metabolism
Biosynthesis of Other Secondary Metabolites
Metabolism of Cofactors and Vitamins
Metabolism of Other Amino Acids

The simplest math

$$\frac{ds_i}{dt} = \sum_j N_{ij} v_j(\vec{s})$$

- s_i concentration of metabolite $i \in [1, ..M]$
- v_j velocity of reaction $j \in [1, ..N]$
- N_{ij} Stoichiometric Matrix
- $N > M$

Stationarity

$$\frac{ds_i}{dt} = \sum_j N_{ij} v_j(\vec{s}) = 0$$

Constraint modelling

$$\mathbf{N}\vec{v} = 0$$

Stationarity

$$\frac{ds_i}{dt} = \sum_j N_{ij} v_j(\vec{s}) = 0$$

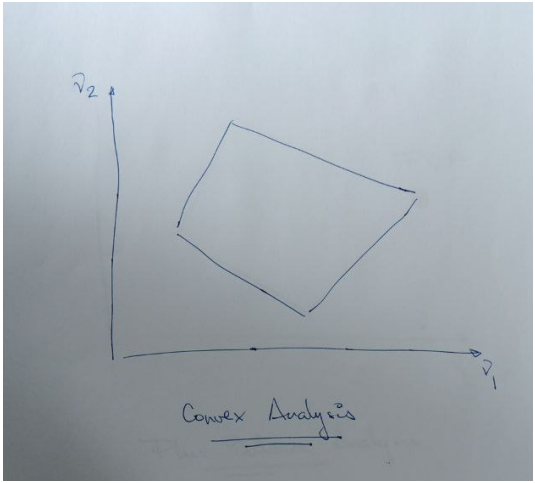
Constraint modelling

$$\mathbf{N}\vec{v} = 0$$

Constraint modelling

$$\mathbf{N}\vec{v} = \vec{b}$$

Graphical representation



Additional Assumption

- Maximize: $E = \sum_j h_j v_j$

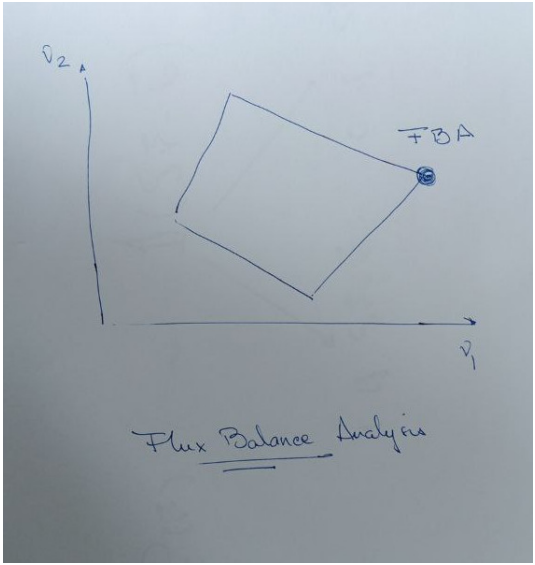
Additional Assumption

- Maximize: $E = \sum_j h_j v_j$

Flux Balance Analysis = Linear Programming

$$\mathbf{N}\vec{v} = \vec{b}$$
$$\max_{\vec{v}} E$$

Graphical representation



Experimental Support

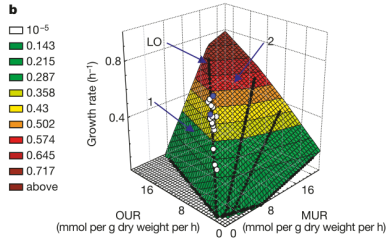


Figure 1 Growth of *E. coli* K-12 on malate. **a**, The malate–oxygen phenotype phase plane (PPP) Phase 1 is characterized by metabolic futile cycles, whereas phase 2 is characterized by acetate overflow metabolism. The line of optimality (LO, in red) separates phases 1 and 2 (ref. 21.) Data points (open circles) represent malate concentrations ranging from 0.25–3 g l^{-1}; and temperatures ranging from 29–37 °C. The two data points in blue represent the starting point (day 0) and endpoint (day 30) of adaptive evolution respectively, at a malate concentration of 2 g l^{-1} and a temperature of 37 °C. These data points represent a span of 500 generations. **b**, Three-dimensional representation of growth rates. The *x* and *y* axes represent the same variables as in **a**. The *z* axis represents the cellular growth rate (h^{-1}). OUR, oxygen uptake rate; MUR, malate uptake rate.

J.S. Edwards, R.U. Ibarra and B.O. Palsson, In silico predictions of Escherichia coli metabolic capabilities are consistent with experimental data,

Nature Biotechnology 2001, 19, 125-130

But Life is more complex than that

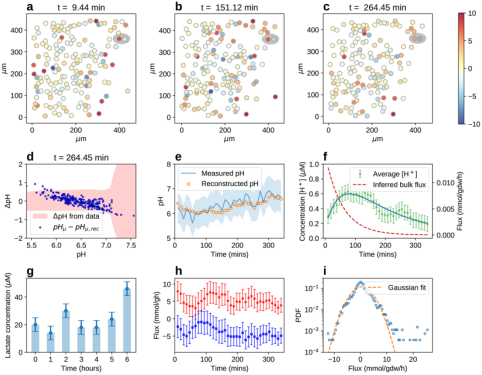
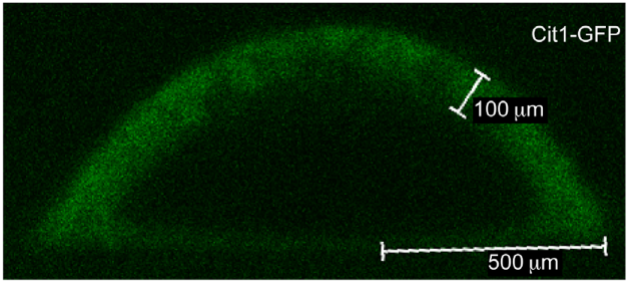


Figure 3. (a–c) Snapshots at different time points (at $t_A = 9$ min, $t_B = 151$ min, and $t_c = 264$ min after the cell culture is settled, all frames are reported in the Supporting Information Figures S1–S5) of the same square visual field (length $L = 500\ \mu m$) during a typical experiment. Cells are represented schematically as disks of diameter $10\ \mu m$ whose color intensity scales with the flux (side bar, blue vs red for importing vs exporting flux). Probes not shown. (d,e) Quality of the reconstructed pH gradient profile. In (d), the error between the pH calculated from the inferred fluxes and the experimentally observed pH is plotted against the latter for each probe (at time $t_c = 264$ min, all frames are reported in the Supporting Information Figures S6–S10). In (e), the time trace of the pH measured by a given probe is reported alongside the reconstructed trend at that spatial point. Shaded areas represent the experimental error on the pH at the probes. (f) Time trends of the bulk $[H^+]$ concentration (experimental, dots and reconstructed, continuous line, left y scale) and inferred bulk acidic efflux (dashed line, right y scale). (g) Time trend of the experimentally measured bulk lactate concentration in a biological replicate. (h) Single-cell flux intensity (in mmol/gdw/h) as a function of time (in min, sampling every 10 min) of the cells forming the dipole motif highlighted in the upper right corner of the frames in (a–c). (i) Single-cell experimental flux distribution (in mmol/gdw/h, (dots) and its Gaussian approximation (lines) in linear-logarithmic scale. The histogram is built from all single-cell flux values (100–200 cells per frame) and time frames (36 frames resulting from a 6 h experiment sampled every 10 min) tracked in one visual field of one experiment.

V. Onesto el al, Probing Single-Cell Fermentation Fluxes and Exchange Networks via pH-Sensing Hybrid Nanofibers, ACS Nano 2023, 17, 4,

But Life is more complex than that



A. Traven et al, Transcriptional profiling of a yeast colony provides new insight into the heterogeneity of multicellular fungal communities. PLoS One. 2012;7(9):e46243.

But Life is more complex than that

Physiological heterogeneity in biofilms

Philip S. Stewart[*1] *and Michael J. Franklin*[*§]

Abstract | Biofilms contain bacterial cells that are in a wide range of physiological states. Within a biofilm population, cells with diverse genotypes and phenotypes that express distinct metabolic pathways, stress responses and other specific biological activities are juxtaposed. The mechanisms that contribute to this genetic and physiological heterogeneity include microscale chemical gradients, adaptation to local environmental conditions, stochastic gene expression and the genotypic variation that occurs through mutation and selection. Here, we discuss the processes that generate chemical gradients in

The ISME Journal (2018) 12:1199–1209
https://doi.org/10.1038/s41396-017-0036-2

ARTICLE

The emergence of metabolic heterogeneity and diverse growth responses in isogenic bacterial cells

Emrah Şimşek[1] · Minsu Kim[1,2]

Available online at www.sciencedirect.com
ScienceDirect
Current Opinion in Microbiology

Metabolic heterogeneity in clonal microbial populations

Vakil Takhaveev and Matthias Heinemann

In the past decades, numerous instances of phenotypic diversity were observed in clonal microbial populations, extreme case of subpopulations having distinctly different activities of metabolic pathways [4–6]. Furthermore,

Opening a mathematical parethesis

We must define a probability $P(v)$.

How to choose?

Maximum Entropy Principle

$S = -\max_{P(v)} \int P(v) \log P(v)$

Among all the probability densities compatible with the data (or knowledge), the one having the largest value of S is the one that best represents our knowledge of the system

Derivation

$$\max_{P_n} -\max_{P_n} \sum_n P_n \log P_n$$

Derivation

$$\max_{P_n} -\max_{P_n} \sum_n P_n \log P_n$$

subject to: $\sum_n P_n = 1$

Derivation

$$\mathcal{L} = -\sum_n P_n \log P_n - \alpha(\sum_n P(n) - 1)$$

Derivation

$$\mathcal{L} = -\sum_n P_n \log P_n - \alpha(\sum_n P(n) - 1)$$

$$\frac{d\mathcal{L}}{dP_m} = -\sum_n \log P_n \delta_{n,m} - \sum_n P_n \frac{1}{P_m} \delta_{n,m} - \alpha \sum_n \delta_{n,m} = 0$$

Derivation

$$\mathcal{L} = -\sum_n P_n \log P_n - \alpha(\sum_n P(n) - 1)$$

$$\frac{d\mathcal{L}}{dP_m} = -\sum_n \log P_n \delta_{n,m} - \sum_n P_n \frac{1}{P_m} \delta_{n,m} - \alpha \sum_n \delta_{n,m} = 0$$

$$\frac{d\mathcal{L}}{dP_m} = -\log P_m - 1 - \alpha = 0$$

Derivation

$$\mathcal{L} = -\sum_n P_n \log P_n - \alpha(\sum_n P(n) - 1)$$

$$\frac{d\mathcal{L}}{dP_m} = -\sum_n \log P_n \delta_{n,m} - \sum_n P_n \frac{1}{P_m} \delta_{n,m} - \alpha \sum_n \delta_{n,m} = 0$$

$$\frac{d\mathcal{L}}{dP_m} = -\log P_m - 1 - \alpha = 0$$

$$P_n \sim e^{-(1+\alpha)}$$

A more general and interesting case

$$\mathcal{L} = -\sum_n P_n \log P_n$$

A more general and interesting case

$$\mathcal{L} = -\sum_n P_n \log P_n$$

subject to: $\sum_n P_n = 1$ and $\sum_n f_n P_n =< f >$

Derivation

$$\mathcal{L} = -\sum_n P_n \log P_n - \alpha(\sum_n P_n - 1) - \beta(\sum_n f_n P_n - < f >)$$

Derivation

$$\mathcal{L} = -\sum_n P_n \log P_n - \alpha(\sum_n P_n - 1) - \beta(\sum_n f_n P_n - < f >)$$

$$P_n \sim e^{\beta f_n}$$

Relationship between fitness and heterogeneity in exponentially growing microbial populations

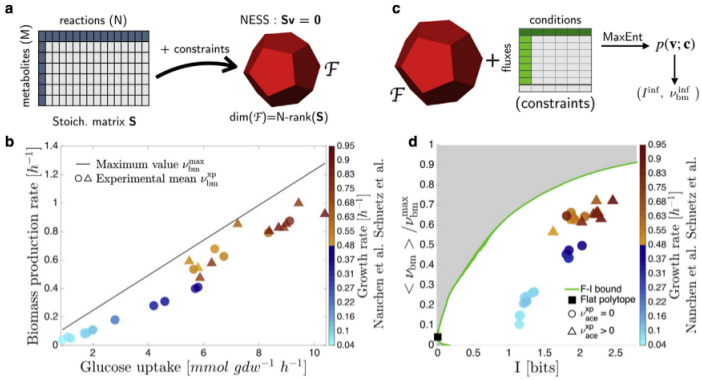
Biophysical Journal, **121**, 1919-1930 (2022)

Datasets:

- *Nanchen, A., A. Schicker, and U. Sauer. 2006. Nonlinear dependency of intracellular fluxes on growth rate in miniaturized continuous cultures of Escherichia coli. Appl. Environ. Microbiol. 72:1164–1172.*
- *Schuetz, R., N. Zamboni, ., U. Sauer. 2012. Multidimensional otimality of microbial metabolism. Science. 336:601–604*
- 33 experiments, growth rate, glucose uptake, more than 20 values of fluxes

Relationship between fitness and heterogeneity in exponentially growing microbial populations

Biophysical Journal, **121**, 1919-1930 (2022)

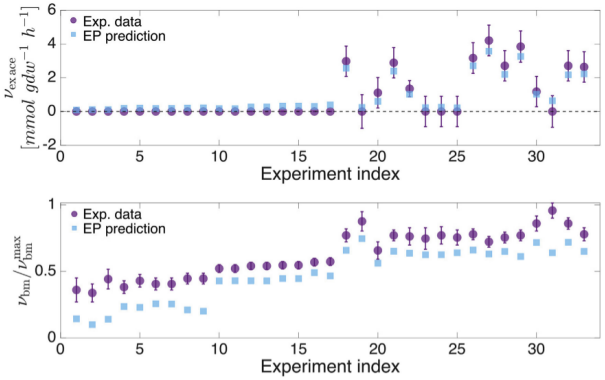


$$P(v) \sim e^{\sum_i \beta_i v_i}$$

$$\langle v_i \rangle_{exp} = \int dv v_i e^{\sum_i \beta_i v_i}$$

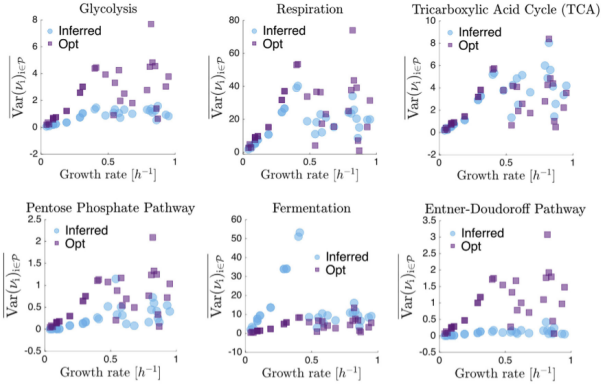
Relationship between fitness and heterogeneity in exponentially growing microbial populations

Biophysical Journal, **121**, 1919-1930 (2022)

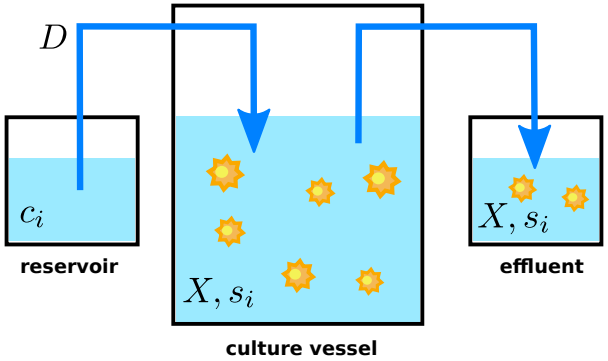


Relationship between fitness and heterogeneity in exponentially growing microbial populations

Biophysical Journal, **121**, 1919-1930 (2022)



Inference of metabolic fluxes in nutrient-limited continuous cultures: A Maximum Entropy approach with the minimum information, iScience 25, 105450 (2022)



Inference of metabolic fluxes in nutrient-limited continuous cultures: A Maximum Entropy approach with the minimum information, iScience 25, 105450 (2022)

$$\frac{dX}{dt} = (\mu - D)X$$

$$\mu = \mu(u, r) \qquad \sigma = \sigma(s)$$

$$\frac{ds_i}{dt} = -u_i X + (c_i - s_i)D$$

Inference of metabolic fluxes in nutrient-limited continuous cultures: A Maximum Entropy approach with the minimum information, iScience 25, 105450 (2022)

$$\frac{dX}{dt} = (\mu - D)X = 0$$

$$\mu(u, r) = D$$

$$u_i < \frac{c_i D}{X}$$

Inference of metabolic fluxes in nutrient-limited continuous cultures: A Maximum Entropy approach with the minimum information, iScience 25, 105450 (2022)

Datasets:

- *Kayser, A., Weber, J., Hecht, V., and Rinas, U. (2005). Metabolic flux analysis of Escherichia coli in glucose-limited continuous culture. I. Growth-rate dependent metabolic efficiency at steady state. Microbiology 151, 693–706.*
- *Nanchen, A., A. Schicker, and U. Sauer. 2006. Nonlinear dependency of intracellular fluxes on growth rate in miniaturized continuous cultures of Escherichia coli. Appl. Environ. Microbiol. 72:1164–1172.*
- *Folsom, J.P., Parker, A.E., and Carlson, R.P. (2014). Physiological and proteomic analysis of Escherichia-coli iron-limited chemostat growth. J. Bacteriol. 196, 2748–2761.*

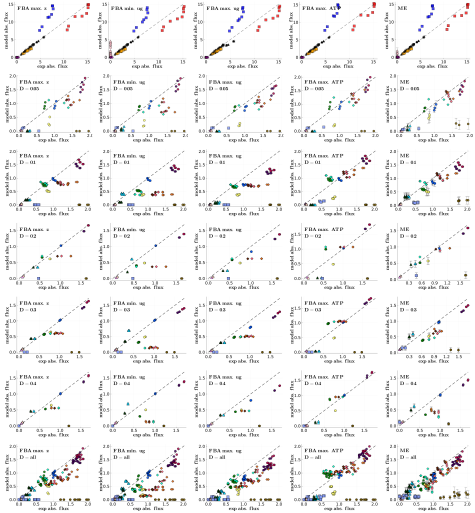
Inference of metabolic fluxes in nutrient-limited continuous cultures: A Maximum Entropy approach with the minimum information, iScience 25, 105450 (2022)

$$P(v) \sim e^{\beta_1 \mu + \beta_g u_g}$$

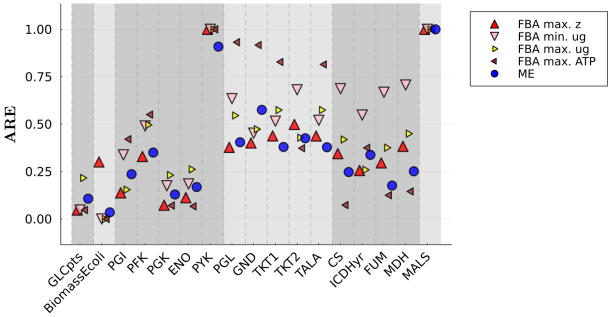
$$\langle \mu \rangle_{exp} = \int dv \mu e^{\beta_1 \mu + \beta_g u_g}$$

$$\langle u_g \rangle_{exp} < \frac{c_g}{DX} \int dv u_g e^{\beta_1 \mu + \beta_g u_g}$$

Inference of metabolic fluxes in nutrient-limited continuous cultures: A Maximum Entropy approach with the minimum information, iScience 25, 105450 (2022)



Inference of metabolic fluxes in nutrient-limited continuous cultures: A Maximum Entropy approach with the minimum information, iScience 25, 105450 (2022)



Phenotype-specific estimation of metabolic fluxes using gene expression data, iScience **26**, 106201 (2023)

Given a transcriptome, how unobserved mechanisms of reaction kinetics should be systematically accounted for when inferring the fluxome?

We have the probability distribution of the transcritome $P(g)$

Phenotype-specific estimation of metabolic fluxes using gene expression data, iScience **26**, 106201 (2023)

$$S = -\sum_{v} P(v) \log P(v)$$

Phenotype-specific estimation of metabolic fluxes using gene expression data, iScience **26**, 106201 (2023)

Hypothesis

$$P(v) = \prod_i P_(v_i)$$

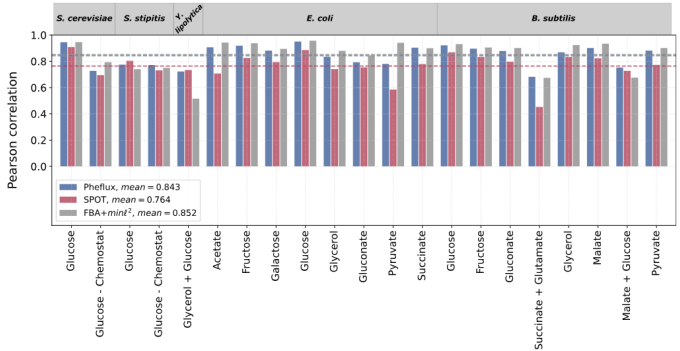
$$P(v_i) = \frac{v_i/g_i}{V}$$

where $V = \sum_i^N \sum_j^{g_i} v_i/g_i$

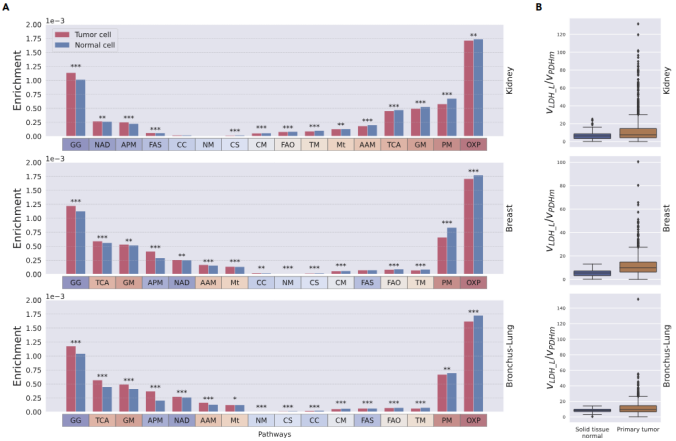
Phenotype-specific estimation of metabolic fluxes using gene expression data, iScience **26**, 106201 (2023)

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Deposited data		
S. cerevisiae GEM	Mo et al.[72]	Mo et al.,72 iMM904
S. cerevisiae RNA-seq transcriptomics	Nookaew et al.[99]	Chemostat and batch, using glucose as carbon source.
S. cerevisiae 13C fluxomics	Papini et al.[73]	Chemostat and batch, using glucose as carbon source
S. stipitis GEM	Liu et al.[105]	iTL885
S. stipitis RNA-seq transcriptomics	Papini et al.[73]	Chemostat and batch, using glucose as carbon source
S. stipitis 13C fluxomics	Papini et al.[73]	Chemostat and batch, using glucose as carbon source
Y. lipolytica GEM	Kerkhoven et al.[74]	iYali
Y. lipolytica RNA-seq tran-scriptomics	Sabra et al.[100]	Glycerol and glucose as carbon source
Y. lipolytica 13C fluxomics	Sabra et al.[100]	Glycerol and glucose as carbon source
E. coli GEM	Orth et al.[71]	iJO1366
E. coli microarray transcrip-tomic	Gerosa et al.[101]	Eight different carbon sources.
E. coli 13C fluxomics	Gerosa et al.[101]	Eight different carbon sources
B. subtilis GEM	Oh et al.[75]	iYO844
B. subtilis microarray tran-scriptomics	Nicolas et al.[103]	Eight different carbon sources.
B. subtilis 13C fluxomics	Chubukov et al.[102]	Eight different carbon sources.
H. sapiens GEM	Brunk et al.[86]	Recon3D
Kidney primary tumor and solid tissue normal FPKMs	https://portal.gdc.cancer.gov/	GDC API fields: cases.primary_site: kidney, files.analysis.workflow_type: HTSeq - FPKM
Breast primary tumor and solid tissue normal FPKMs	https://portal.gdc.cancer.gov/	GDC API fields: cases.primary_site: breast, files.analysis.workflow_type: HTSeq - FPKM
Bronchus-Lung primary tu-mor and solid tissue normal FPKMs	https://portal.gdc.cancer.gov/	GDC API fields: cases.primary_site: bronchus and lung, files.analysis.workflow_type: HTSeq - FPKM

Phenotype-specific estimation of metabolic fluxes using gene expression data, iScience **26**, 106201 (2023)



Phenotype-specific estimation of metabolic fluxes using gene expression data, iScience **26**, 106201 (2023)



Conclusions