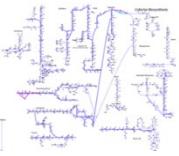


Omics

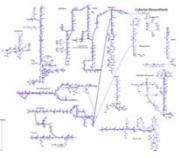
Genomics, Transcriptomics,
Proteomics, & Metabolomics



Omics

- Introduction
- Biofuel Example - Metabolomics & Proteomics
 - ✓ Creating an *in silico* COBRA model
 - ✓ Creating the plasmids for the cell factory
 - ✓ Extracting metabolomics & proteomics data
 - ✓ Creating context-sensitive *in silico* models
 - ✓ Extracting data from context-sensitive model
 - Flux balance analysis
 - Sampling analysis
- Conclusion

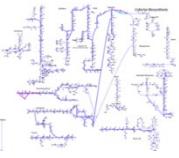
<https://en.wikipedia.org/wiki/Metabolomics>



Omics

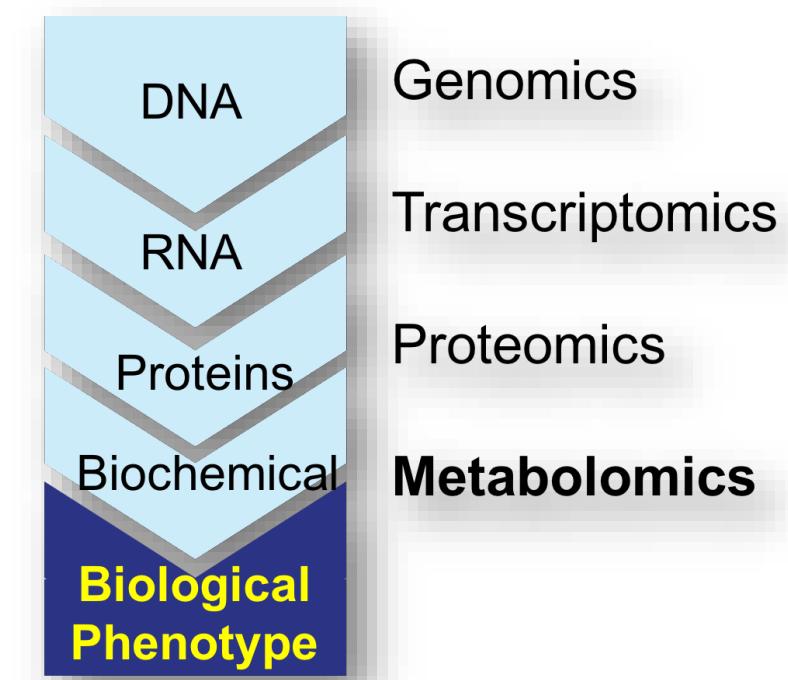
- The branches of science known informally as “omics” are various disciplines in biology whose names end in the suffix - “omics,” such as genomics, proteomics, metabolomics, and transcriptomics.
- Omics aims at the collective characterization and quantification of pools of biological molecules that translate into the structure, function, and dynamics of an organism or organisms.

<https://en.wikipedia.org/wiki/Omics>



Omics Definitions

- **Genomics** - An interdisciplinary field of biology focusing on the structure, function, evolution, mapping, and editing of genomes. A genome is an organism's complete set of DNA, including all of its genes as well as its hierarchical, three-dimensional structural configuration.
- **Transcriptomics** - the study of an organism's transcriptome, the sum of all of its RNA transcripts.
- **Proteomics** - the large-scale study of proteins.
- **Metabolomics** - the scientific study of chemical processes involving metabolites, the small molecule substrates, intermediates and products of cell metabolism.



<https://en.wikipedia.org/wiki/Metabolomics>

Wikipedia



Applications of Omics

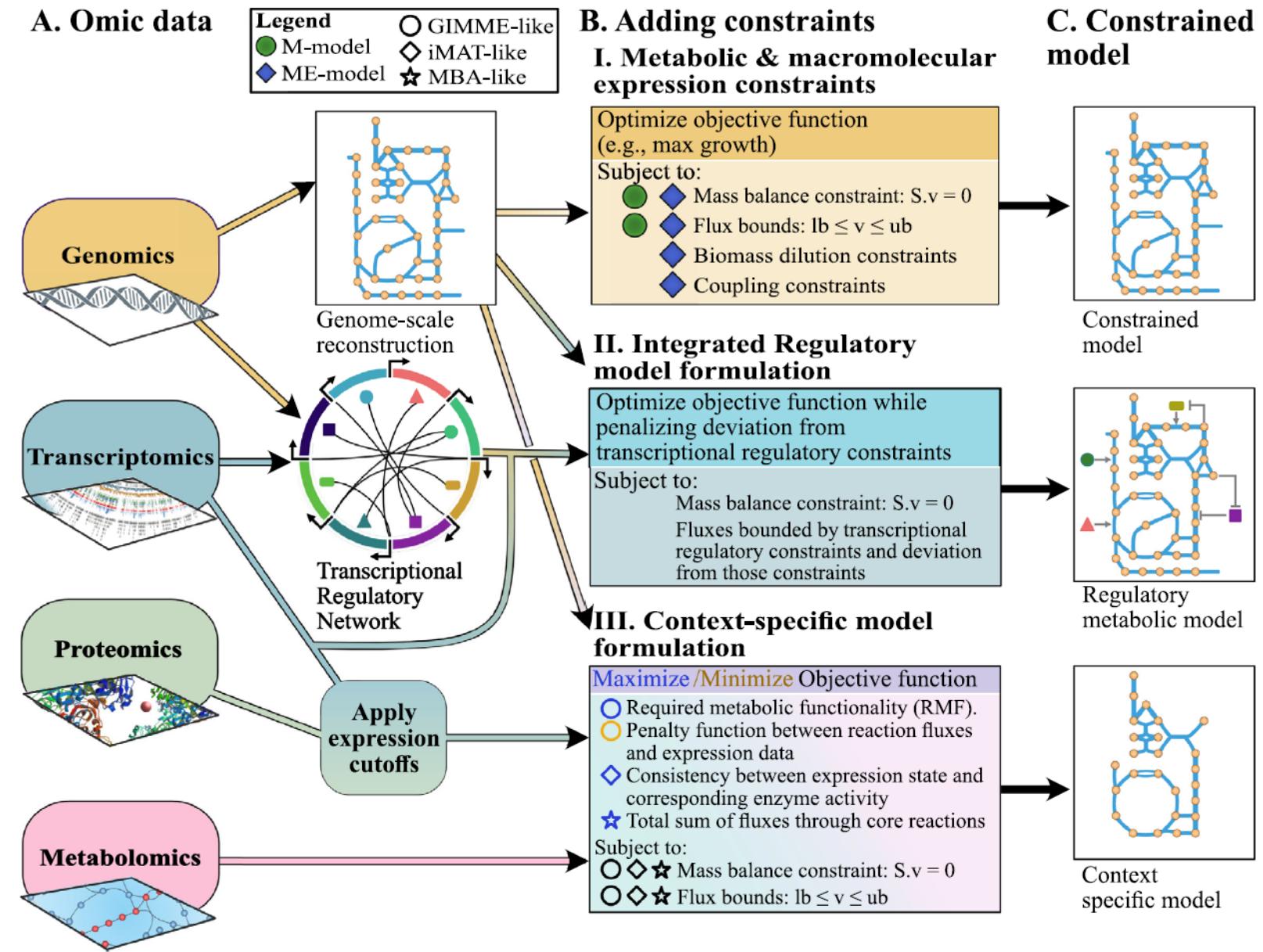
- Use annotated genome sequences to create genome-scale models of cell metabolism
- Use omic data to refine the genome-scale models
- Use transcriptomics and metabolomics to build context-specific models
- Integrate metabolomics data with genome-scale models
- Measure and predict proteome allocation
- Integrate multi-omic data with genome-scale models
- Use meta-omic data to build and refine microbial community models
- Provide a systems context for protein structures (structural genomics)

Dahal, Sanjeev, et al. "Synthesizing Systems Biology Knowledge from Omics Using Genome-Scale Models." *Proteomics* 20.17-18 (2020): 1900282.

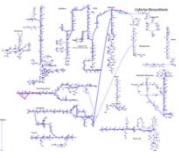


Integrating GEMs with OMIC Data

- A. Building genome-scale models (GEMs) and integrating them with various omic data types as constraints.
- B. GEMs must be constrained to obtain biologically relevant information.
- C. The resulting models can be simulated to investigate the genotype-phenotype-environment relationship in biological systems.



Dahal, Sanjeev, et al. "Synthesizing Systems Biology Knowledge from Omics Using Genome-Scale Models." *Proteomics* 20.17-18 (2020): 1900282.



Omics

- Introduction
- Biofuel Example - Metabolomics & Proteomics
 - ✓ Creating an *in silico* COBRA model
 - ✓ Creating the plasmids for the cell factory
 - ✓ Extracting metabolomics & proteomics data
 - ✓ Creating context-sensitive *in silico* models
 - ✓ Extracting data from context-sensitive model
 - Flux balance analysis
 - Sampling analysis
- Conclusion

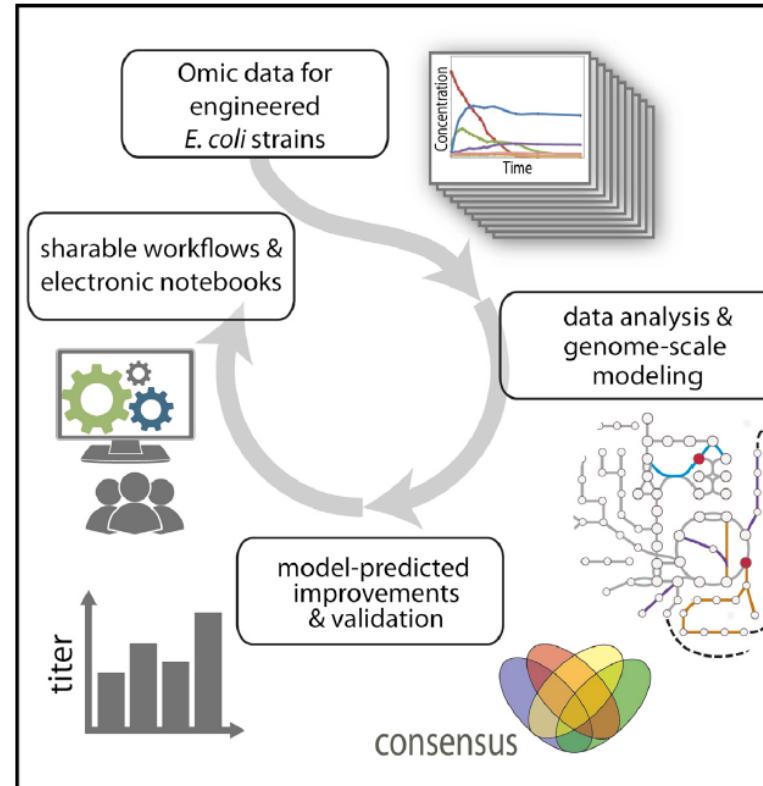


Article

Cell Systems

Characterizing Strain Variation in Engineered *E. coli* Using a Multi-Omics-Based Workflow

Graphical Abstract



Authors

Elizabeth Brunk, Kevin W. George,
Jorge Alonso-Gutierrez, ...,
Jay D. Keasling, Bernhard O. Palsson,
Taek Soon Lee

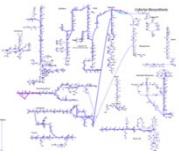
Correspondence

palsson@eng.ucsd.edu (B.O.P.),
tslee@lbl.gov (T.S.L.)

In Brief

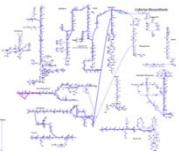
Brunk et al. develop a workflow to assess and interpret multi-omics data and use it to characterize strain variation in biofuel-producing *E. coli*.

Brunk, Elizabeth, et al. "Characterizing strain variation in engineered *E. coli* using a multi-omics-based workflow." *Cell systems* 2.5 (2016): 335-346.



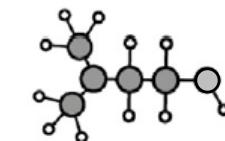
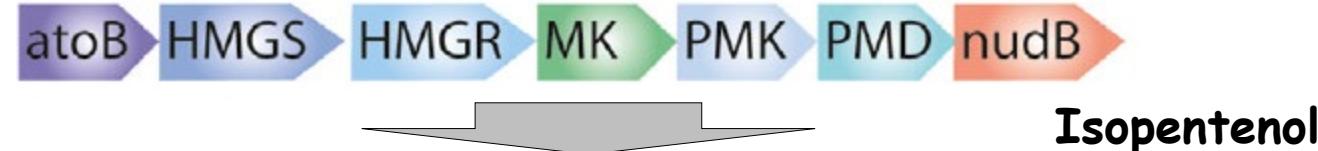
Omics

- Introduction
- Biofuel Example - Metabolomics & Proteomics
- ✓ Creating an *in silico* COBRA model
 - ✓ Creating the plasmids for the cell factory
 - ✓ Extracting metabolomics & proteomics data
 - ✓ Creating context-sensitive *in silico* models
 - ✓ Extracting data from context-sensitive model
 - Flux balance analysis
 - Sampling analysis
- Conclusion



Biofuel Pathways

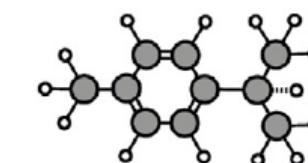
Brunk, 2016



BIGG Database



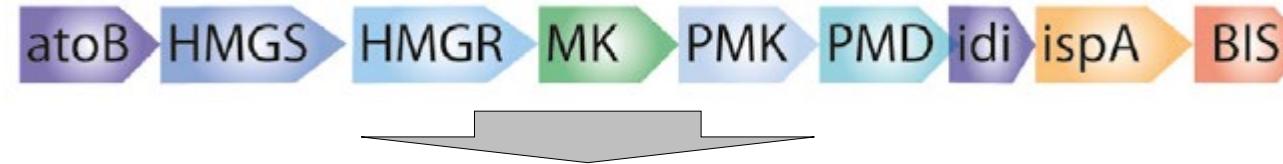
Brunk, 2016



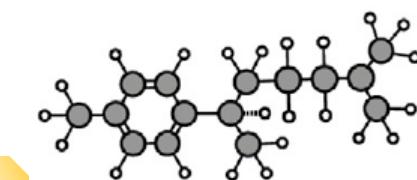
BIGG Database

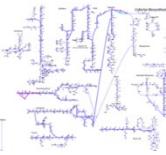


Brunk, 2016

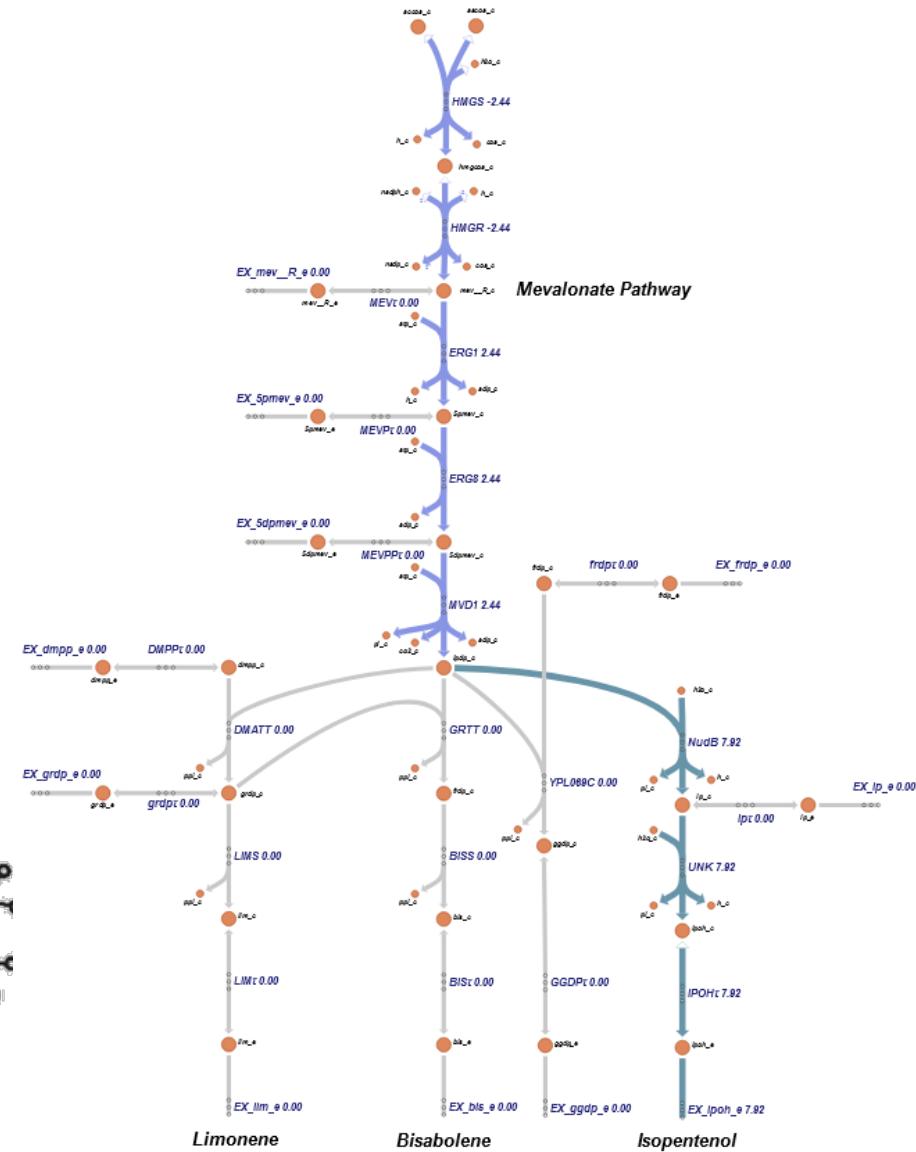
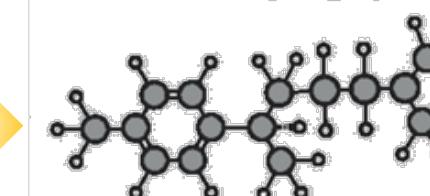
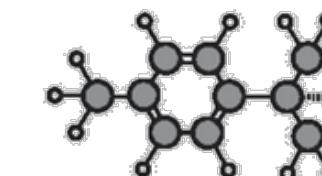
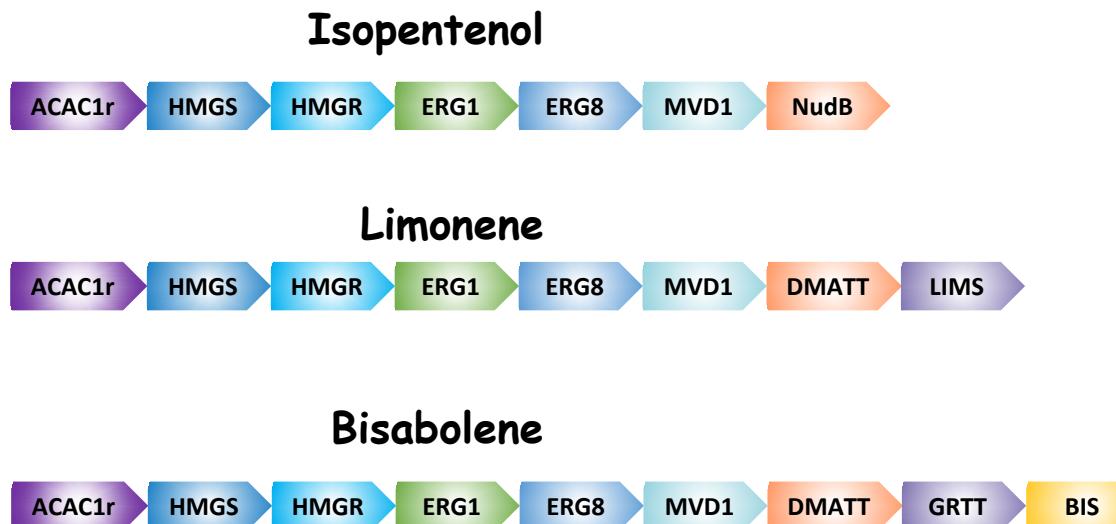


BIGG Database

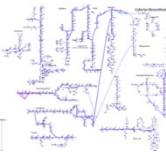




Biofuel Production in *E.coli*: Isopentenol, Limonene, and Bisabolene



Brunk, Elizabeth, et al. "Characterizing strain variation in engineered *E. coli* using a multi-omics-based workflow." *Cell systems* 2.5 (2016): 335-346.

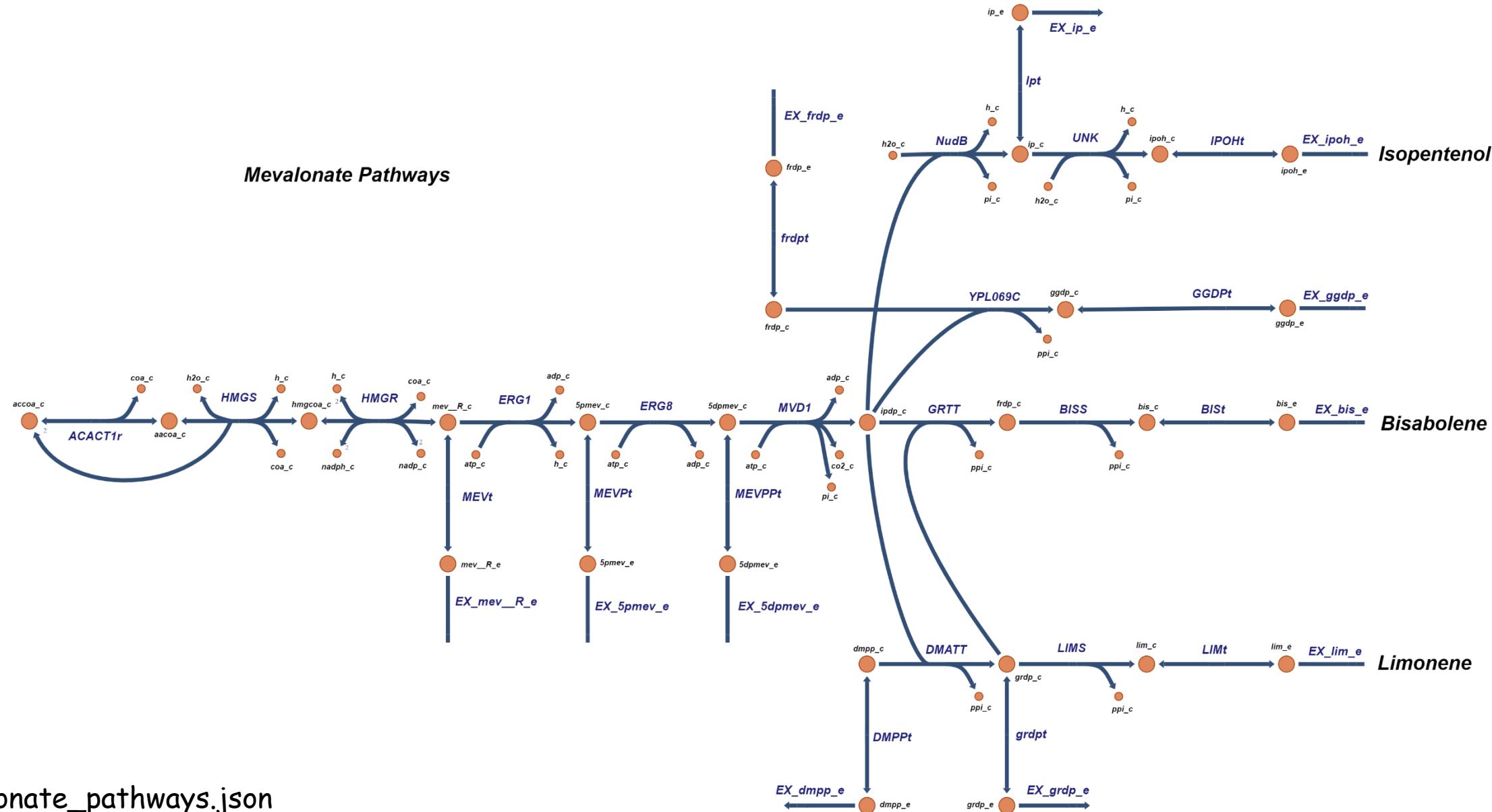


Constraint-based Metabolic Reconstructions & Analysis

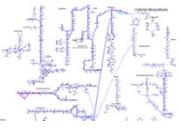
H. Scott Hinton, 2022

12

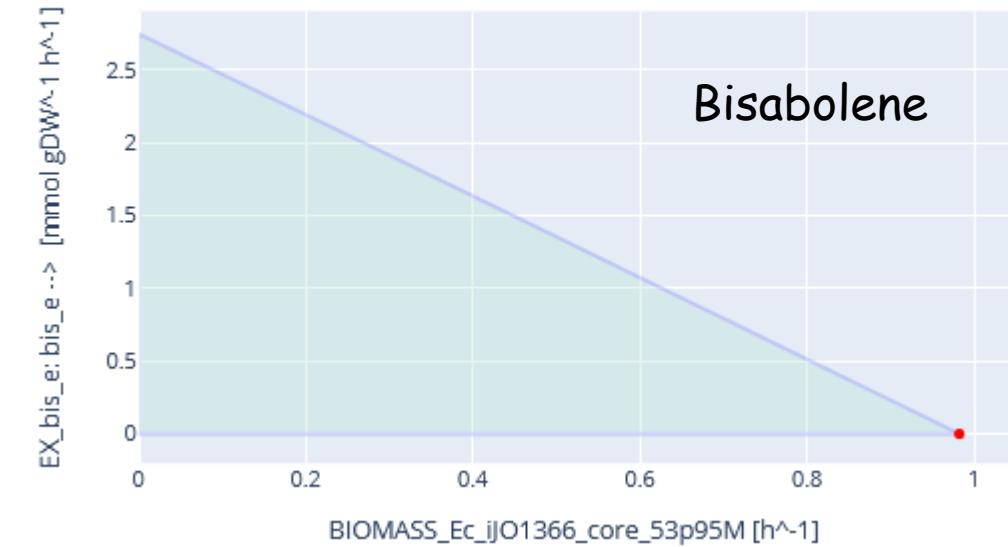
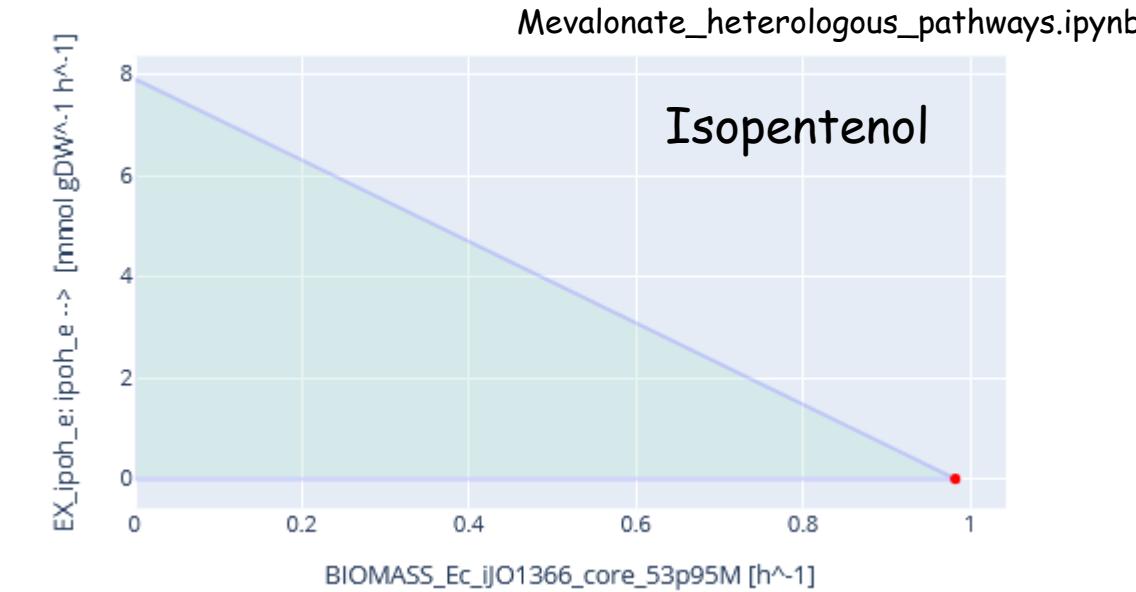
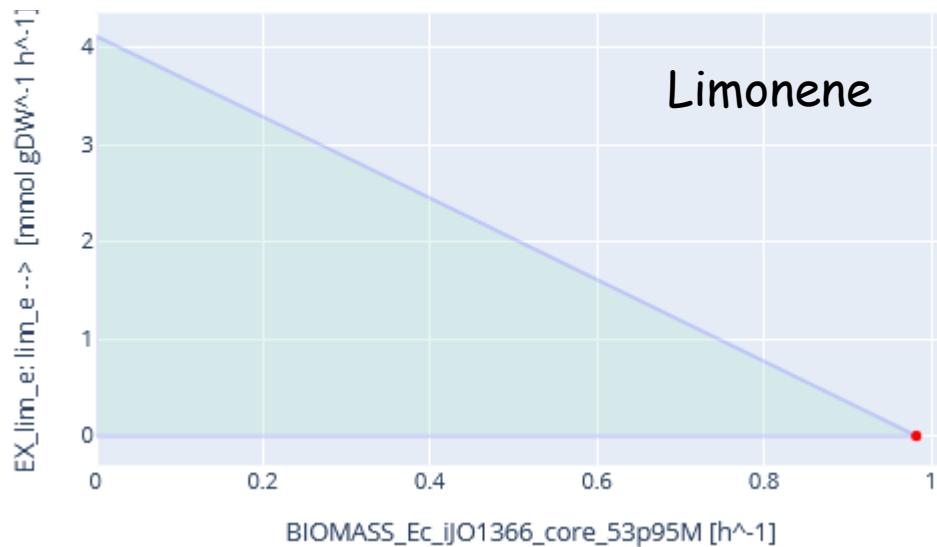
Mevalonate Pathway:



Mevalonate_pathways.json



Biofuel Production Envelopes





Heterologous Pathways in *E. coli* for Isopentenol, Limonene, and Bisabolene

Create the mevalonate-based pathways in *E.coli* (iJO1366) to create isopentenol, limonene, and bisabolene.

```
In [1]: from cameo import models
from cameo.strain_design import pathway_prediction
from cobra import Model, Reaction, Metabolite
from IPython.display import display
from cobrapy_bigg_client import client
import re
```

Load the iJO1366 *E.coli* model

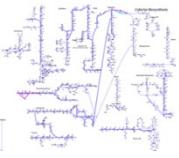
```
In [2]: model = models.bigg.iJO1366
model.solver = 'glpk'
```

Set parameter Username
Academic license - for non-commercial use only - expires 2022-10-10

Identify the reactions that can be used to create pathways in the iJO1366 model that can produce isopentenol, limonene and bisabolene [1].

OMICs Data	Reaction Name	BIGG Reaction Name	Other Names	Native Species
HMGS	hydroxymethylglutaryl-CoA synthase	HMGCOAS		Saccharomyces cerevisiae S288C (iMM904)
HMGR	hydroxymethylglutaryl-CoA reductase	HMGCOAR		Saccharomyces cerevisiae S288C (iMM904)
ERG1	mevalonate kinase	MEVK1	MK	Saccharomyces cerevisiae S288C (iMM904)
ERG8	phosphomevalonate kinase	PMEVK	PMK	Saccharomyces cerevisiae S288C (iMM904)
MVD1	diphosphomevalonate decarboxylase	DPMVD	PMD	Saccharomyces cerevisiae S288C (iMM904)
NudB	dATP pyrophosphohydrolase	NudB		Manually Created
UNK	inorganic or acidic phosphatase acting on IP	UNK		Manually Created
LIMS	limonene synthase	LIMS		Manually Created
BISS	bisabolene synthase	BISS		Manually Created
YPL069C	geranylgeranyl diphosphate synthase	GGPS_h		Phaeodactylum tricornutum CCAP 1055/1 (iLB1027_lipid)

Mevalonate_heterologous_pathways.ipynb



Omics

- Introduction
- Biofuel Example - Metabolomics & Proteomics
 - ✓ Creating an *in silico* COBRA model
 - ✓ Creating the plasmids for the cell factory
 - ✓ Extracting metabolomics & proteomics data
 - ✓ Creating context-sensitive *in silico* models
 - ✓ Extracting data from context-sensitive model
 - Flux balance analysis
 - Sampling analysis
- Conclusion



Plasmids Used in Biofuel Production

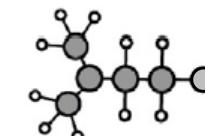
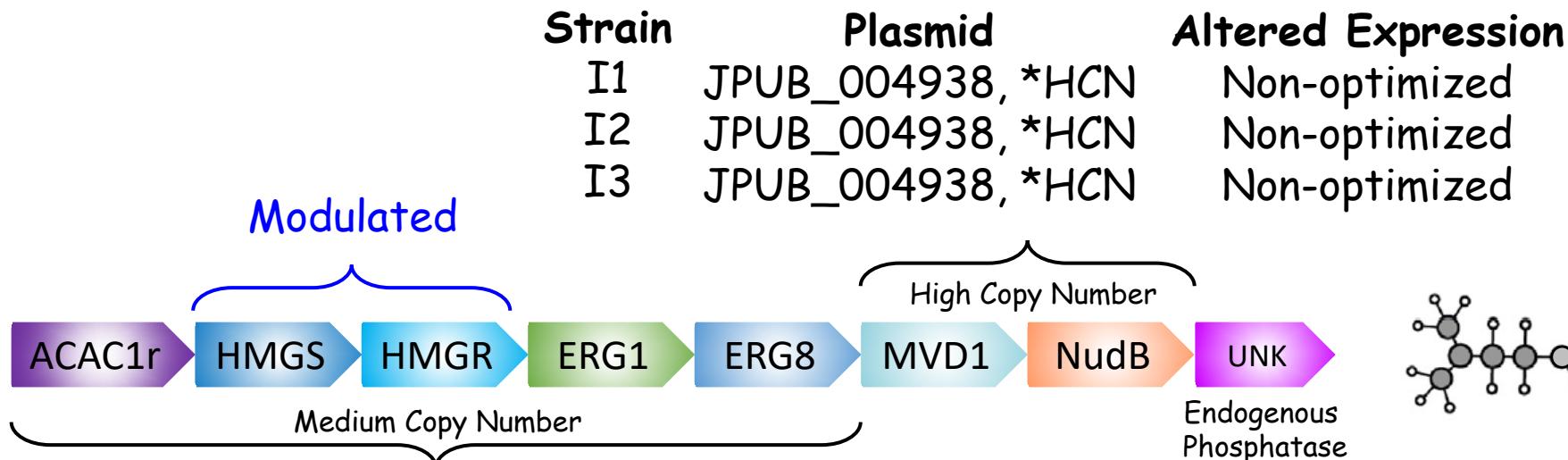
Strain	Plasmids	Description	Reference
I ₁	JPUB_006200	pBbA5c-MevTo-MK-PMK	George et al., 2014
I ₂	JPUB_006210	pBbA5c-MevTco-PMK-MK	George et al., 2014
I ₃	JPUB_004937	pBbA5c-MevTsa-MK-PMK	George et al., 2014
I ₁ , I ₂ , I ₃	JPUB_004938	pTrc99A-nudB-PMD	George et al., 2014
L ₁	JPUB_004921	pBbA5c-MevTo-MK-PMK-PMD-idi	Alonso-Gutierrez et al., 2013
L ₂	JPUB_002471	pBbE1a-GPPS-LS-atoB-HMGSsa-HMGRsa-trc-MK-PMK-PMD-idi	Alonso-Gutierrez et al., 2015
L ₃	JPUB_002470	pBbA5c-atoB-HMGSsa-HMGRsa-trc-MK-PMK-PMD-idi-trc-GPPS-LS	Alonso-Gutierrez et al., 2015
L ₁	JPUB_002464	pTrc99A-GPPS-LS	Alonso-Gutierrez et al., 2015
L ₃	JPUB_002473	pTrc99A-LS	Alonso-Gutierrez et al., 2015
B ₁	JBx_000323	pBbA5c-MevTo-MK-PMK-PMD-idi-ispA	Peralta-Yahya et al., 2011
B ₂	JPUB_002460	pBbA5c-MevTco-trc-MK-PMK-PMD-idi-ispA	Peralta-Yahya et al., 2011
B ₁ , B ₂	JPUB_002466	pTrc99A-BIS	Peralta-Yahya et al., 2011

Strains	Description	Reference
I1	JPUB_006200 + JPUB_004938	George et al., 2014
I2	JPUB_006210 + JPUB_004938	George et al., 2014
I3	JPUB_004937 + JPUB_004938	George et al., 2014
L1	JPUB_004921 + JPUB_002464	Alonso-Gutierrez et al., 2015
L2	JPUB_002471	Alonso-Gutierrez et al., 2015
L3	JPUB_002470 + JPUB_002473	Alonso-Gutierrez et al., 2015
B1	JBx_000323 + JPUB_002466	Peralta-Yahya et al., 2011
B2	JPUB_002460 + JPUB_002466	Peralta-Yahya et al., 2011
DH1		Hanahan 1983

Brunk, Elizabeth, et al. "Characterizing strain variation in engineered *E. coli* using a multi-omics-based workflow." Cell systems 2.5 (2016): 335-346.



Isopentenol OMICs Data



Strain	Plasmid	Altered Expression
I1	JPUB_006200	Non-optimized HMGS and HMGR reactions
I2	JPUB_006210	Codon-optimized HMGS and HMGR reactions
I3	JPUB_004937	HMGS and HMGR from <i>Staphylococcus aureus</i>

File Names

I1.csv, I1_prot.csv
I2.csv, I2_prot.csv
I3.csv, I3_prot.csv

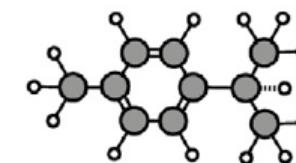
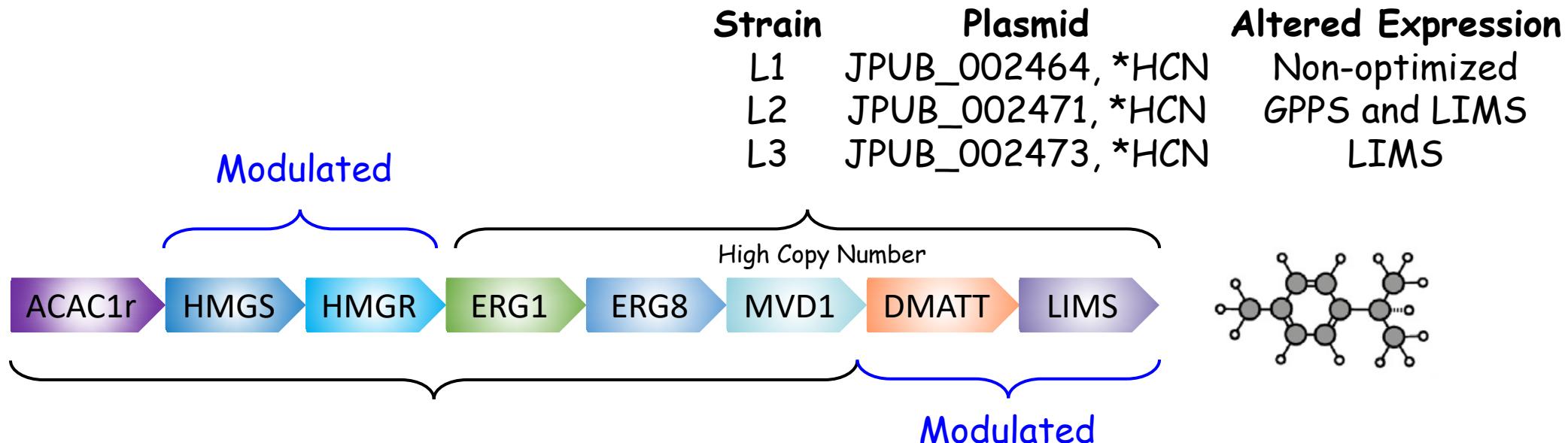
*HCN = High Copy Number

1. Brunk, Elizabeth, et al. "Characterizing strain variation in engineered *E. coli* using a multi-omics-based workflow." *Cell systems* 2.5 (2016): 335-346.

2. George, Kevin W., et al. "Correlation analysis of targeted proteins and metabolites to assess and engineer microbial isopentenol production." *Biotechnology and bioengineering* 111.8 (2014): 1648-1658.



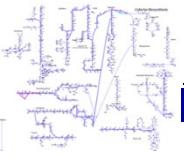
Limonene OMICs Data



Strain	Plasmid	Altered Expression	File Names
L1	JPUB_004921	Non-optimized HMGS and HMGR reactions	L1.csv, L1_prot.csv
L2	JPUB_002471	HMGS and HMGR from <i>Staphylococcus aureus</i> , *HCN	L2.csv, L2_prot.csv
L3	JPUB_002470	HMGS and HMGR from <i>Staphylococcus aureus</i> , *HCN	L3.csv, L3_prot.csv

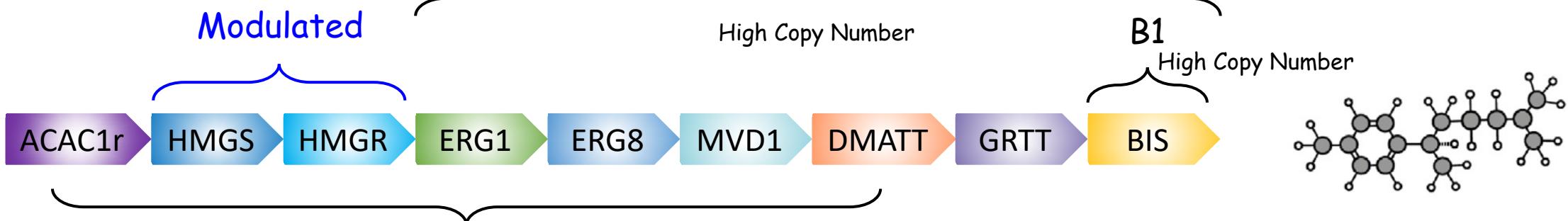
*HCN = High Copy Number

- Brunk, Elizabeth, et al. "Characterizing strain variation in engineered *E. coli* using a multi-omics-based workflow." *Cell systems* 2.5 (2016): 335-346.
- Alonso-Gutierrez, Jorge, et al. "Metabolic engineering of *Escherichia coli* for limonene and perillyl alcohol production." *Metabolic engineering* 19 (2013): 33-41.



Bisabolene OMICs Data

Strain	Plasmid	Altered Expression
B1	JPUB_002466 , *HCN	BIS
B2	JPUB_002466, *HCN	BIS
B2		



Strain	Plasmid	Altered Expression	File Names
B1	JBx_000323	Non-optimized HMGS and HMGR reactions	B1.csv, B1_prot.csv
B2	J PUB 002460	Codon-optimized HMGS and HMGR reactions, *HCN	B2.csv , B2_prot.csv

*HCN = High Copy Number

1. Brunk, Elizabeth, et al. "Characterizing strain variation in engineered *E. coli* using a multi-omics-based workflow." *Cell systems* 2.5 (2016): 335-346.
 2. Peralta-Yahya, Pamela P., et al. "Microbial engineering for the production of advanced biofuels." *Nature* 488.7411 (2012): 320-328.



Monitored Internal Metabolites

Monitored_Metabolites_Proteins.ipynb

Cytoplasm Metabolite ID	Cytoplasm Metabolite Names	Cytoplasm Metabolite ID	Cytoplasm Metabolite Names	Cytoplasm Metabolite ID	Cytoplasm Metabolite Names
2me4p_c	2-C-methyl-D-erythritol 4-phosphate	dxy15p_c	1-deoxy-D-xylulose 5-phosphate	leu__L_c	L-Leucine
2mecdp_c	2-C-methyl-D-erythritol 2,4-cyclodiphosphate	fdp_c	D-Fructose 1,6-bisphosphate	lys__L_c	L-Lysine
3pg_c	3-Phospho-D-glycerate	frdp_c	Farnesyl diphosphate	mal__L_c	L-Malate
4c2me_c	4-(cytidine 5'-diphospho)-2-C-methyl-D-erythritol	fum_c	Fumarate	met__L_c	L-Methionine
5pmev_c	R 5 Phosphomevalonate C6H10O7P	gln__L_c	L-Glutamine	mev__R_c	R Mevalonate C6H11O4
aacoa_c	Acetoacetyl-CoA	glu__L_c	L-Glutamate	nad_c	Nicotinamide adenine dinucleotide
accoa_c	Acetyl-CoA	glx_c	Glyoxylate	nadp_c	Nicotinamide adenine dinucleotide phosphate
acon_C_c	Cis-Aconitate	gly_c	Glycine	oxalcoa_c	Oxalyl-CoA
adp_c	ADP C10H12N5O10P2	glyclt_c	Glycolate C2H3O3	pep_c	Phosphoenolpyruvate
akg_c	2-Oxoglutarate	grdp_c	Geranyl diphosphate	phe__L_c	L-Phenylalanine
ala__L_c	L-Alanine	h2mb4p_c	1-hydroxy-2-methyl-2-(E)-butenyl 4-diphosphate	pro__L_c	L-Proline
amp_c	AMP C10H12N5O7P	his__L_c	L-Histidine	pyr_c	Pyruvate
arg__L_c	L-Arginine	hmgcoa_c	Hydroxymethylglutaryl CoA C27H39N7O20P3S	ser__L_c	L-Serine
asn__L_c	L-Asparagine	icit_c	Isocitrate	succ_c	Succinate
asp__L_c	L-Aspartate	ile__L_c	L-Isoleucine	thr__L_c	L-Threonine
atp_c	ATP C10H12N5O13P3	ip_c	Isopentenol-P	trp__L_c	L-Tryptophan
cit_c	Citrate	ipdp_c	Isopentenyl diphosphate	tyr__L_c	L-Tyrosine
cys__L_c	L-Cysteine	lac__D_c	D-Lactate	val__L_c	L-Valine



Monitored Extracellular Metabolites

Extracellular Metabolite IDs	Extracellular Metabolite Names
2mecdp_e	2-C-Methyl-D-erythritol 2,4-cyclodiphosphate
4c2me_e	4-(Cytidine 5-diphospho)-2-C-methyl-D-erythritol
5pmev_e	Mevalonate-5P
ac_e	Acetate
amp_e	AMP C10H12N5O7P
bis_e	Bisabolene
dxyl5p_e	1-Deoxy-D-xylulose 5-Phosphate
for_e	Formate
frdp_e	Farnesyl-PP
ggdp_e	GeranylGeranyl-PP
glc__D_e	D-Glucose
grdp_e	GGeranyl-PP
h2mb4p_e	1-hydroxy-2-methyl-2-(E)-butenyl 4-diphosphate
ip_e	Isopentenol-P
ipoh_e	Isopentenol
lac__D_e	D-Lactate
lim_e	Limonene
mev__R_e	Mevalonate
pyr_e	Pyruvate
succ_e	Succinate

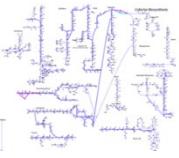
Monitored_Metabolites_Proteins.ipynb



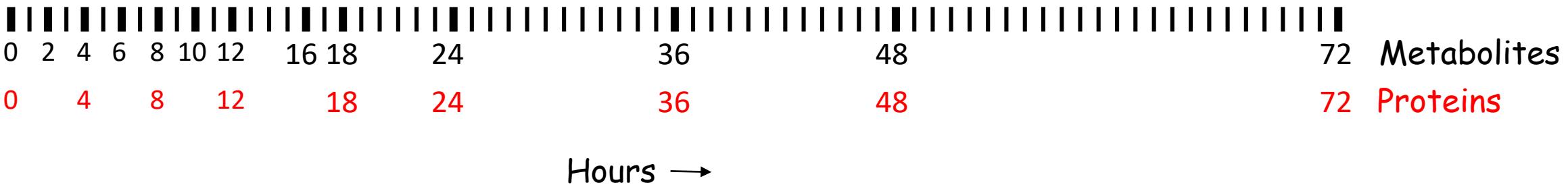
Monitored Internal Reactions

Monitored_Metabolites_Proteins.ipynb

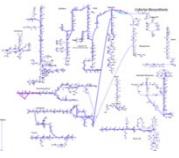
	Reaction ID	Reaction Name		Reaction ID	Reaction Name
0	ACACT1r	Acetyl-CoA C-acetyltransferase	27	IPDDI	Isopentenyl-diphosphate D-isomerase
1	ACALD	Acetaldehyde dehydrogenase (acetylating)	28	LIMS	limonene synthase
2	ACKr	Acetate kinase	29	MDH	Malate dehydrogenase
3	ACONTa	Aconitase (half-reaction A, Citrate hydro-lyase)	30	MDH2	Malate dehydrogenase (ubiquinone 8 as acceptor)
4	ACONTb	Aconitase (half-reaction B, Isocitrate hydro-lyase)	31	ME1	Malic enzyme (NAD)
5	ACS	Acetyl-CoA synthetase	32	ME2	Malic enzyme (NADP)
6	AKGDH	2-Oxoglutarate dehydrogenase	33	MVD1	Diphosphomevalonate decarboxylase
7	BISS	bisabolene synthase	34	NudB	dATP pyrophosphohydrolase
8	CS	Citrate synthase	35	PDH	Pyruvate dehydrogenase
9	ENO	Enolase	36	PDXPP	Pyridoxine 5-phosphate phosphatase
10	ERG1	Mevalonate kinase atp	37	PFK	Phosphofructokinase
11	ERG8	Phosphomevalonate kinase	38	PFL	Pyruvate formate lyase
12	FBA	Fructose-bisphosphate aldolase	39	PGI	Glucose-6-phosphate isomerase
13	FBP	Fructose-bisphosphatase	40	PGK	Phosphoglycerate kinase
14	FDH4pp	Formate dehydrogenase (quinone-8) (periplasm)	41	PGM	Phosphoglycerate mutase
15	FHL	Formate-hydrogen lyase	42	PPC	Phosphoenolpyruvate carboxylase
16	FRD2	Fumarate reductase	43	PPCK	Phosphoenolpyruvate carboxykinase
17	FRD3	Fumarate reductase	44	PPS	Phosphoenolpyruvate synthase
18	FUM	Fumarase	45	PTAr	Phosphotransacetylase
19	G6PDH2r	Glucose 6-phosphate dehydrogenase	46	PYK	Pyruvate kinase
20	GAPD	Glyceraldehyde-3-phosphate dehydrogenase	47	RPI	Ribose-5-phosphate isomerase
21	GLYCL	Glycine Cleavage System	48	SUCDi	Succinate dehydrogenase (irreversible)
22	GND	Phosphogluconate dehydrogenase	49	SUCOAS	Succinyl-CoA synthetase (ADP-forming)
23	GRTT	Geranyltranstransferase	50	TALA	Transaldolase
24	HMGR	Hydroxymethylglutaryl CoA reductase	51	TKT1	Transketolase
25	HMGS	Hydroxymethylglutaryl CoA synthase	52	TKT2	Transketolase
26	ICDHyr	Isocitrate dehydrogenase (NADP)	53	TPI	Triose-phosphate isomerase
			54	YPL069C	geranylgeranyl diphosphate synthas



Data Sample Times

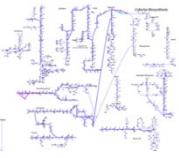


Brunk, Elizabeth, et al. "Characterizing strain variation in engineered *E. coli* using a multi-omics-based workflow." *Cell systems* 2.5 (2016): 335-346.

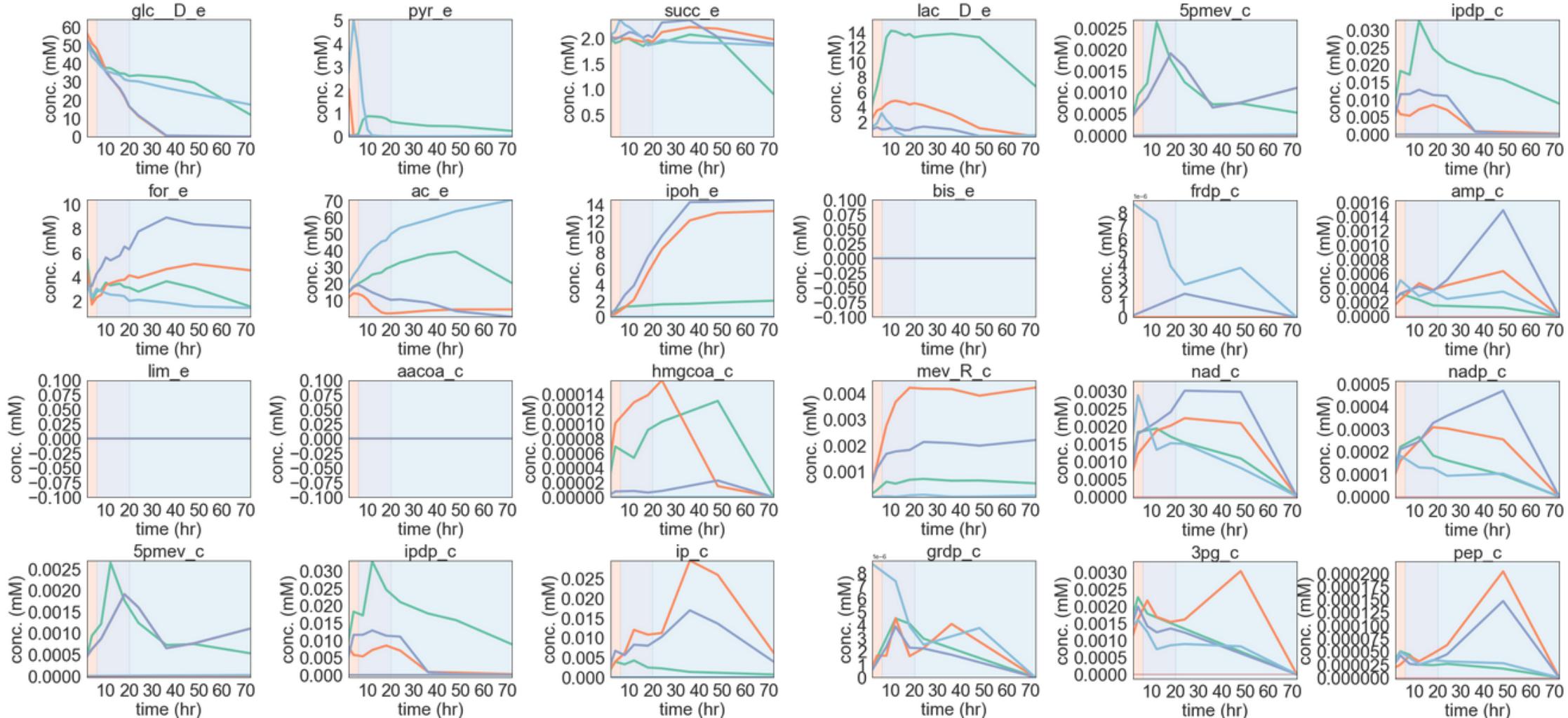


Omics

- Introduction
- Biofuel Example - Metabolomics & Proteomics
 - ✓ Creating an *in silico* COBRA model
 - ✓ Creating the plasmids for the cell factory
 - ✓ Extracting metabolomics & proteomics data
 - ✓ Creating context-sensitive *in silico* models
 - ✓ Extracting data from context-sensitive model
 - Flux balance analysis
 - Sampling analysis
- Conclusion



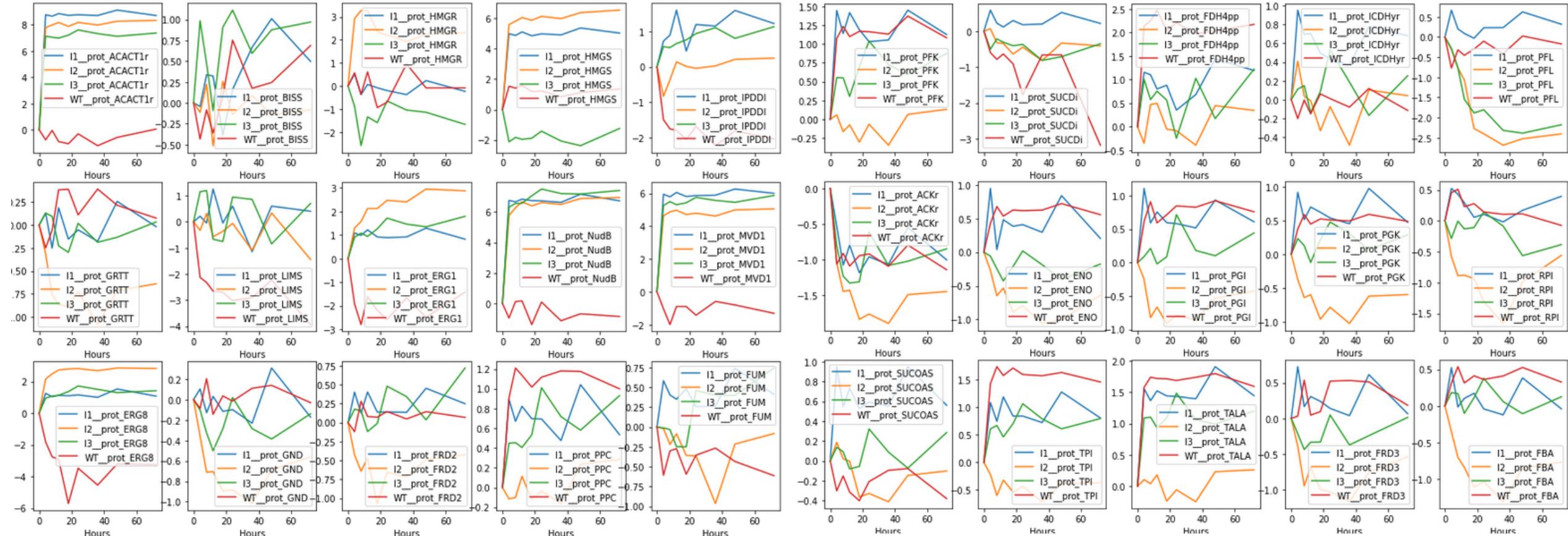
Metabolomics Plots



Stage_One_Dynamic_Differences_Revised.ipynb



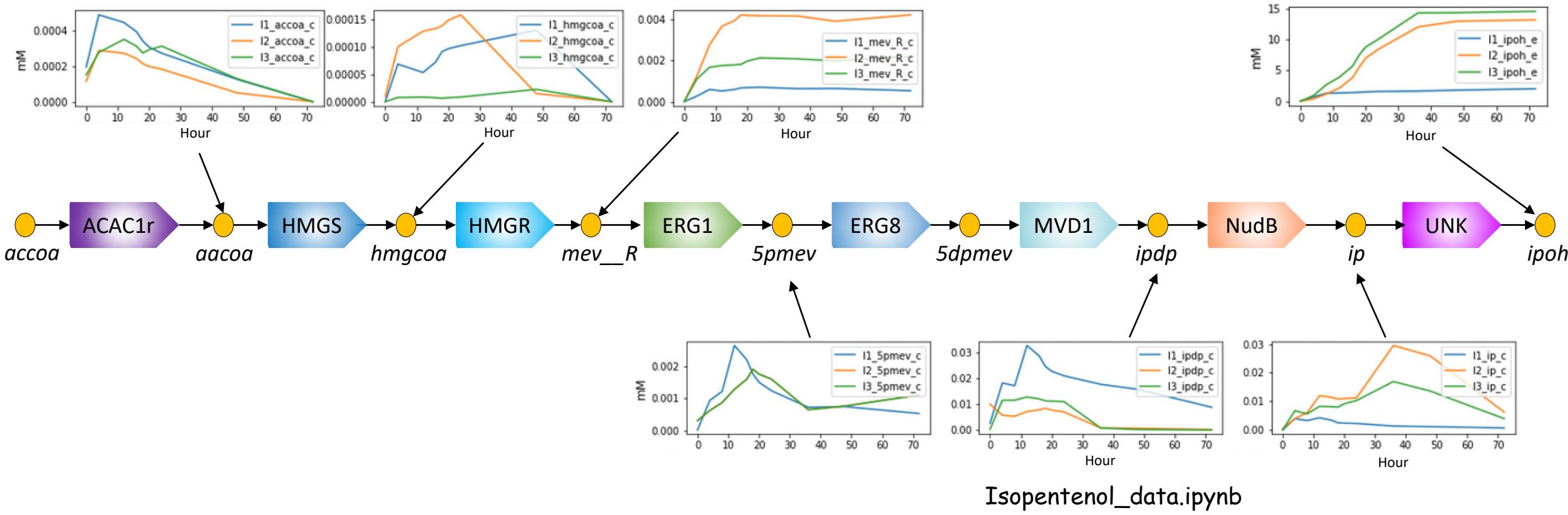
Proteomics Plots



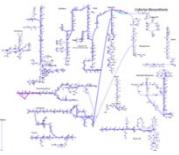
Isopentenol_OIMCS_Data.ipynb



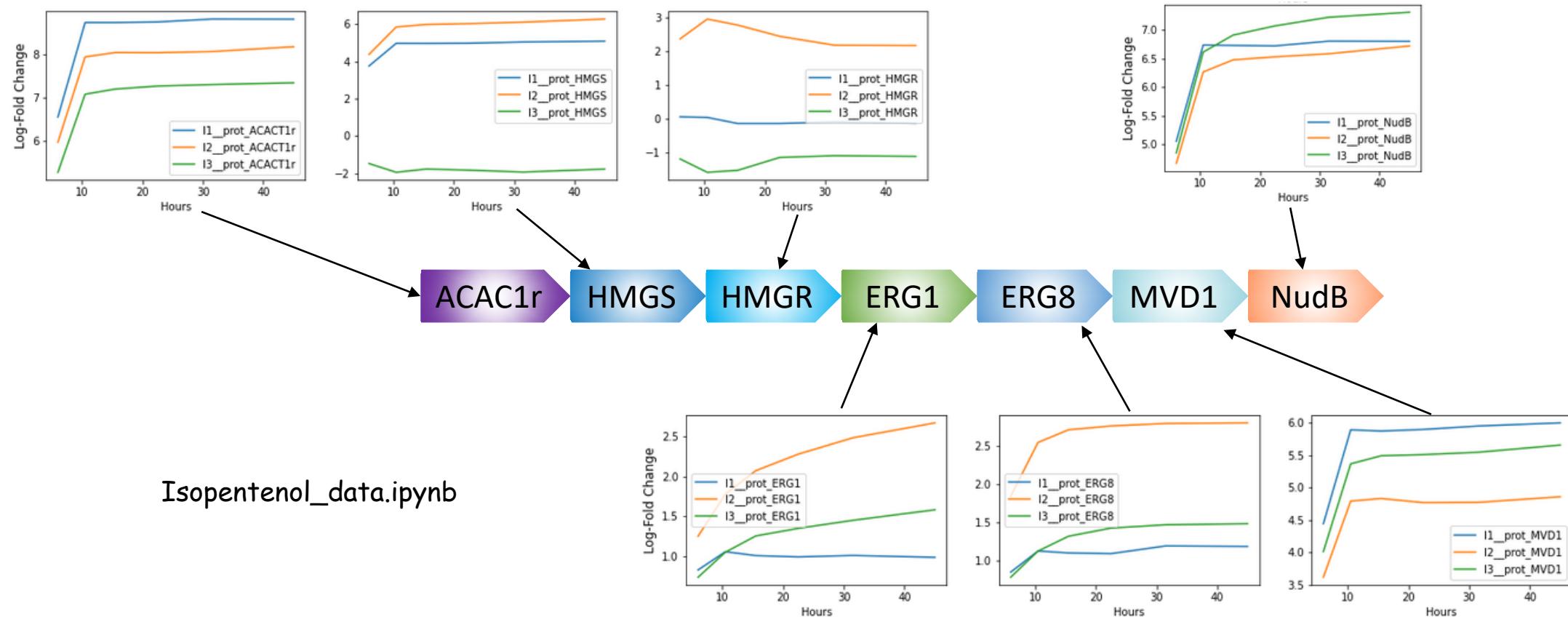
Isopentenol Metabolomics Data



Brunk, Elizabeth, et al. "Characterizing strain variation in engineered *E. coli* using a multi-omics-based workflow." *Cell systems* 2.5 (2016): 335-346.



Isopentenol Proteomics Data



Brunk, Elizabeth, et al. "Characterizing strain variation in engineered *E. coli* using a multi-omics-based workflow." *Cell systems* 2.5 (2016): 335-346.



Isopentenol OIMCS Data

Isopentenol_data.ipynb

```
In [1]: import cobra.test
from cobrapy_bigg_client import client
from cameo import phenotypic_phase_plane
from cameo.visualization.plotting.with_plotly import PlotlyPlotter
import pandas as pd
import numpy as np
import matplotlib.pyplot as mp
plotter = PlotlyPlotter()
pd.set_option('display.max_rows', 1000)
pd.set_option('display.max_columns', 100)
pd.set_option('display.width', 100)
pd.set_option('display.max_colwidth', None)
```

Load isopentenol metabolomics data

Load isopentenol data files (I1.csv, I1_prot.csv, I2.csv, I2_prot.csv, I3.csv, I3_prot.csv, DH1.csv, and DH1_prot.csv) into Panda dataframes [1]

```
In [2]: df_I1_orig = pd.read_csv('./data/I1.csv')
df_I1 = df_I1_orig.copy()
df_I1[1:13] = df_I1[1:13].replace({0:np.nan})
df_I1.round(5)
```

Out[2]:

	Unnamed: 0	Hour	Strain	Sample	OD600	Intracellular volume / sample	glc_D_e	pyr_e	succ_e	lac_D_e	for_e	ac_e	ipoh_e	bis_e	lim_e	aacoa_c	hmgc
0	0.0	0.0	I1	1.0	1.65961	0.0	58.76443	0.55966	0.09654	0.58171	1.40561	7.69692	0.00000	0.0	0.0	0.0	0.0
1	1.0	2.0	I1	10.0	NaN	NaN	52.73718	0.05809	1.98748	4.41807	5.48720	15.66065	NaN	NaN	NaN	NaN	NaN
2	2.0	4.0	I1	19.0	2.24000	0.0	48.13932	0.04464	1.90567	6.74897	2.05995	18.43014	0.69187	NaN	NaN	NaN	0.0
3	3.0	6.0	I1	28.0	NaN	NaN	44.76099	0.11082	1.93568	9.65735	2.77460	20.05595	NaN	NaN	NaN	NaN	NaN



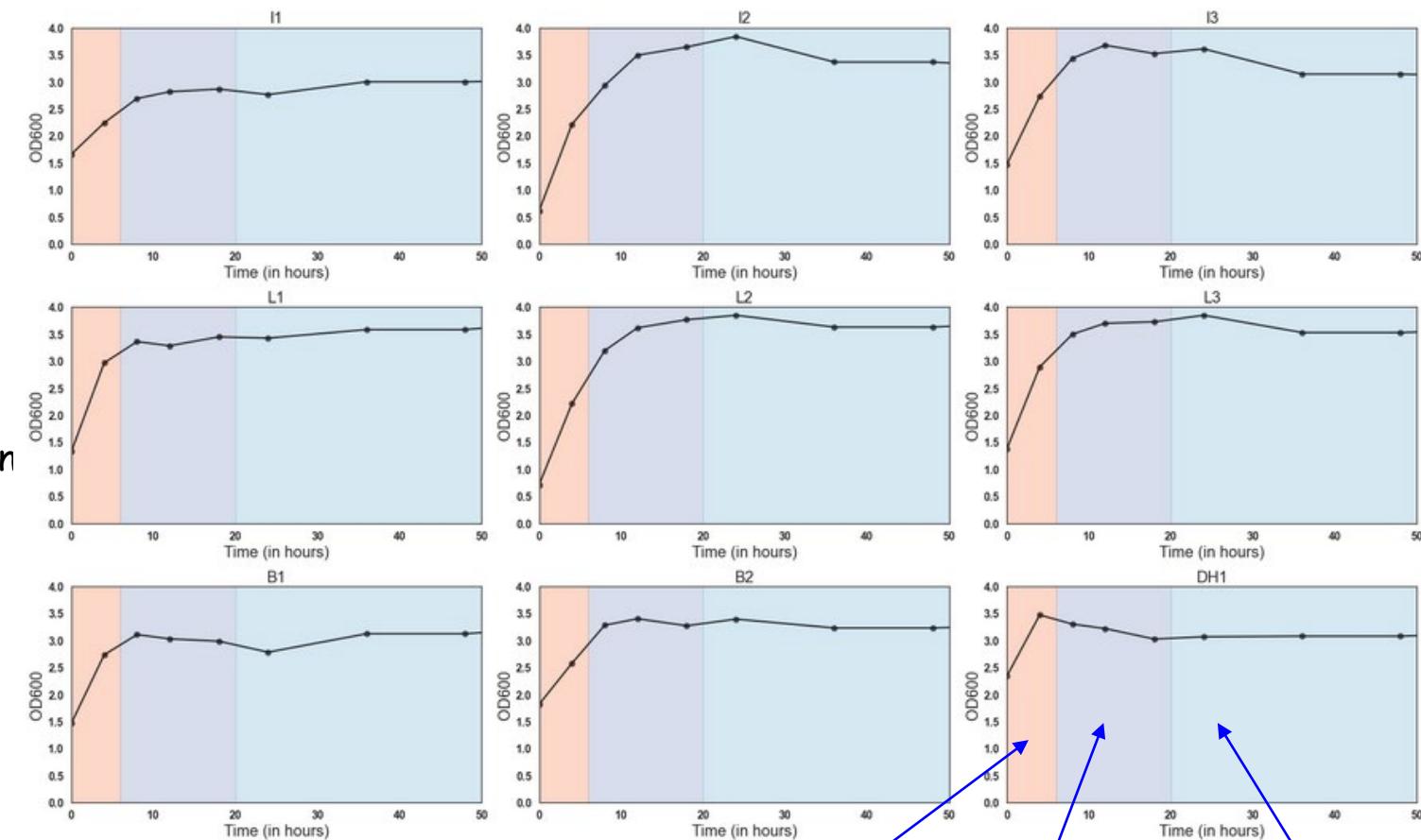
Omics

- Introduction
- Biofuel Example - Metabolomics & Proteomics
 - ✓ Creating an *in silico* COBRA model
 - ✓ Creating the plasmids for the cell factory
 - ✓ Extracting metabolomics & proteomics data
- ✓ Creating context-sensitive *in silico* models
 - ✓ Extracting data from context-sensitive model
 - Flux balance analysis
 - Sampling analysis
- Conclusion



Creation of COBRA Model Strains & Phases

- Each strain is divided into three different phases
 - Phase1 = 0,2,4,6 hours
 - Phase2 = 8,10,12,16,18,20 hours
 - Phase3 = 24,36,48 hours
- The monitored metabolomics data will be converted into a flux values (metabolite concentration averaged over each phase). The flux values will be used to create the bounds for the exchange reaction associated with the metabolite.
- A separate COBRA model will be created for the three phases of each strain (I1, I2, I3, L1, L2, L3, B1, B2, DH1)
 - Example: I1_0_6.json would be the model for phase 1 of the strain I1

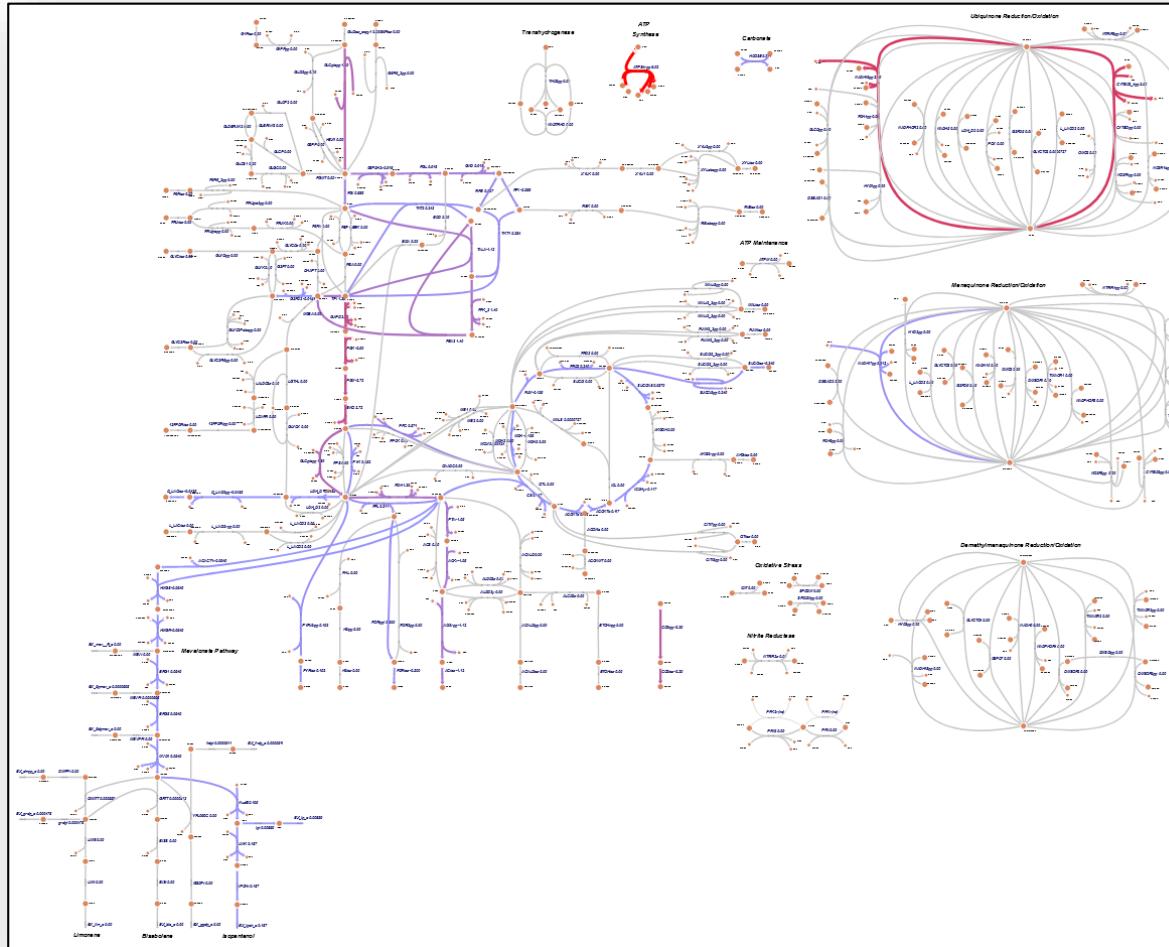


Stage_three_GEM_revised_part_1.ipynb

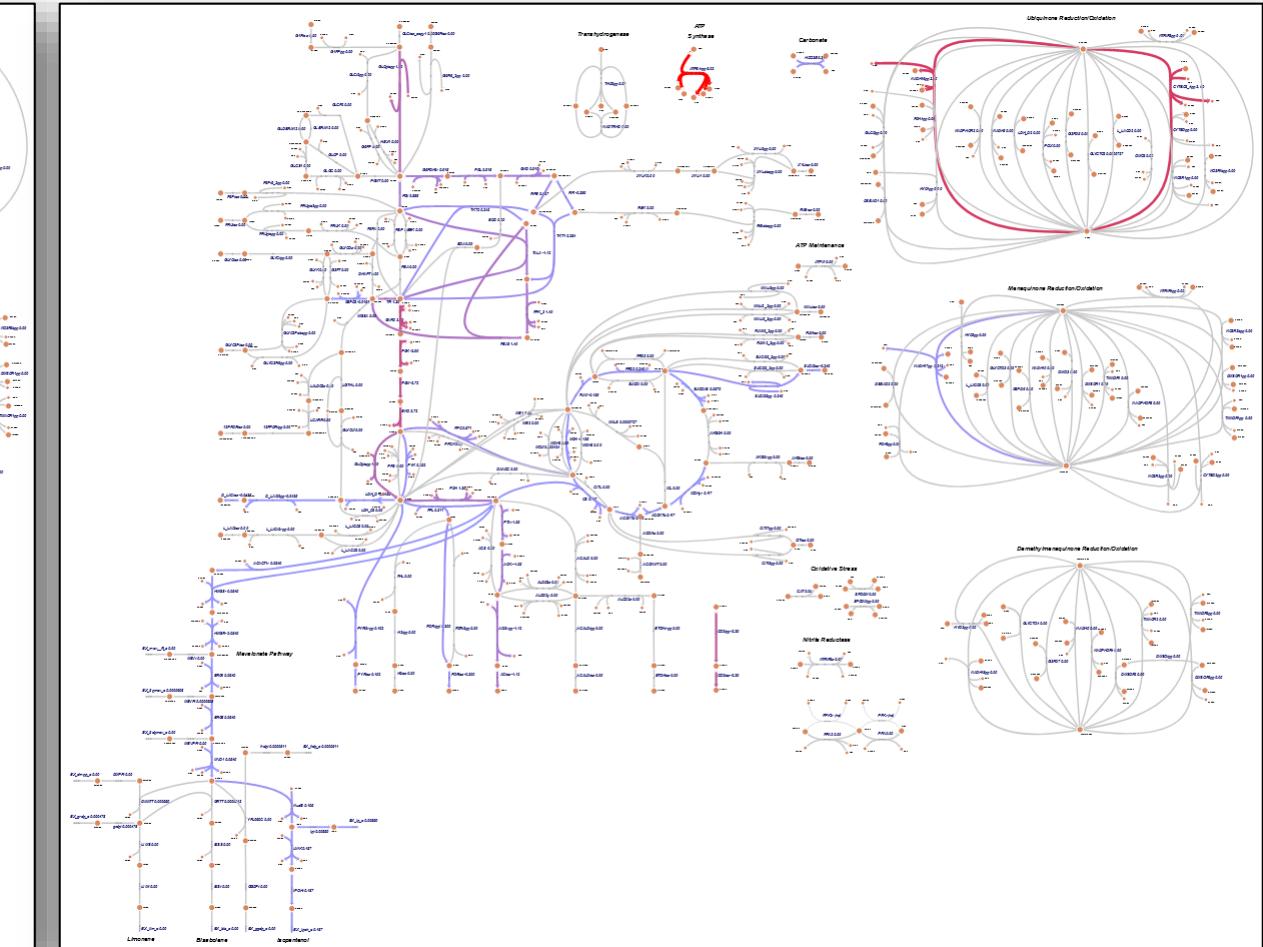
Phase I Phase II Phase III



Creating COBRA Models of Strains & Phases



Model of Strain I3 during Phase 1



Model of Strain I3 during Phase 3



Creating OMICS-based Mevalonate Biofuel COBRA Models

Adapted from "Stage_three_GEM notebook" by Brunck [1]

Load the appropriate Python packages

```
In [1]: import cobra.test
from cobrapy_bigg_client import client

# Panda python module for dataframe and data storage/manipulation
import pandas as pd
pd.set_option('mode.use_inf_as_na',True)
pd.set_option('display.max_rows', 500)
pd.set_option('display.max_columns', 999)
pd.set_option('precision', 3)

from contextlib import contextmanager
import sys, os

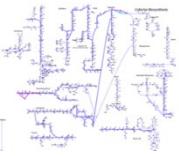
@contextmanager
def suppress_stdout():
    with open(os.devnull, "w") as devnull:
        old_stdout = sys.stdout
        sys.stdout = devnull
        try:
            yield
        finally:
            sys.stdout = old_stdout
```

Load the "iJO1366 Mevalonate Pathways" model

```
In [2]: model_orig = cobra.io.load_json_model("iJO1366_mevalonate_pathways.json");
model_orig.solver = 'glpk'
model = model_orig.copy();

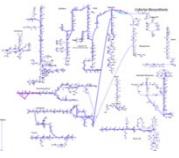
Set parameter Username
Academic license - for non-commercial use only - expires 2022-10-10
```

Creating_OMICS-based_Mevalonate_Biofuel_COBRA_Models.ipynb

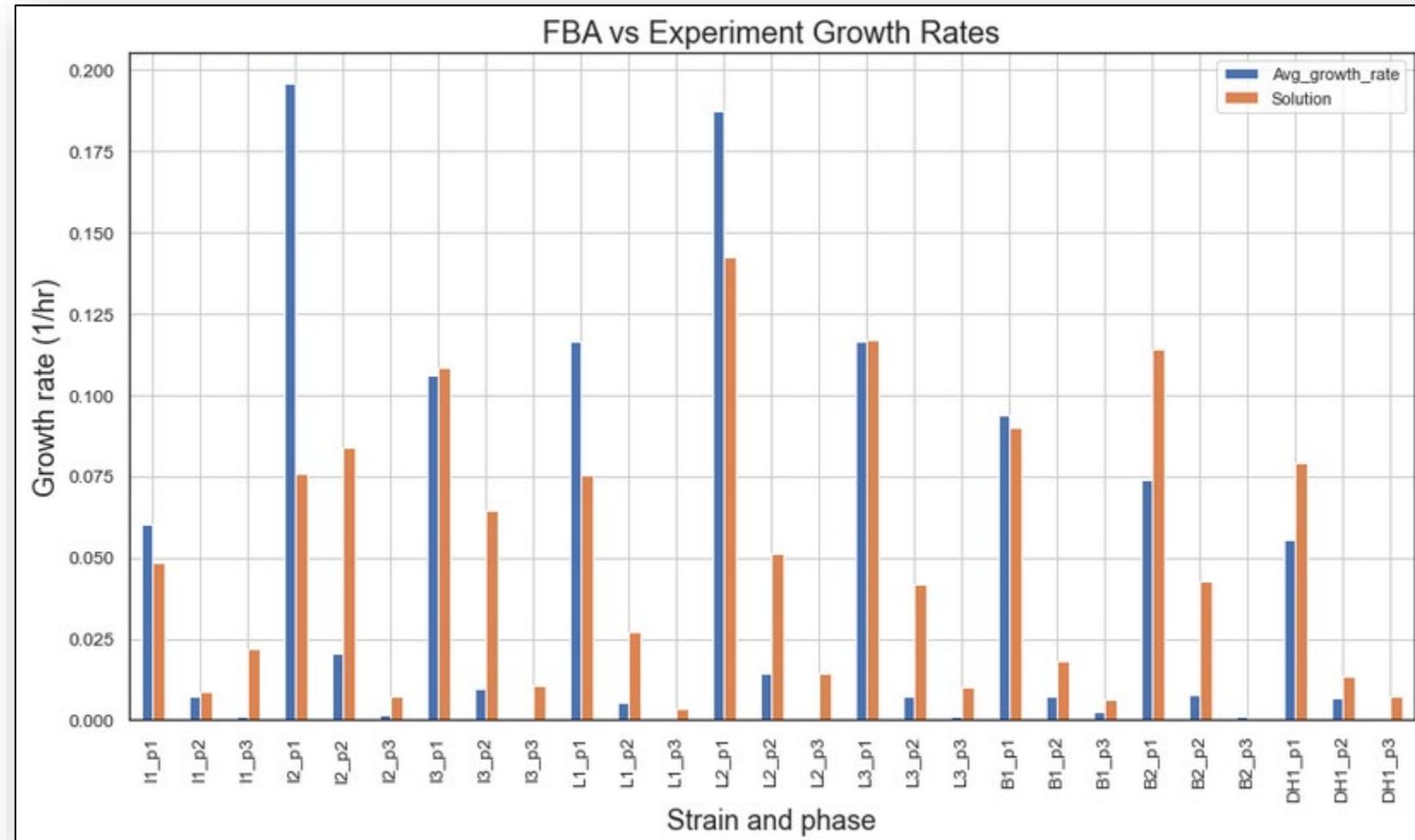


Omics

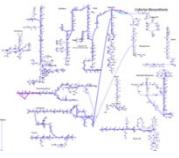
- Introduction
- Biofuel Example - Metabolomics & Proteomics
 - ✓ Creating an *in silico* COBRA model
 - ✓ Creating the plasmids for the cell factory
 - ✓ Extracting metabolomics & proteomics data
 - ✓ Creating context-sensitive *in silico* models
 - ✓ Extracting data from context-sensitive model
 - ➡ ■ Flux balance analysis
 - Sampling analysis
- Conclusion



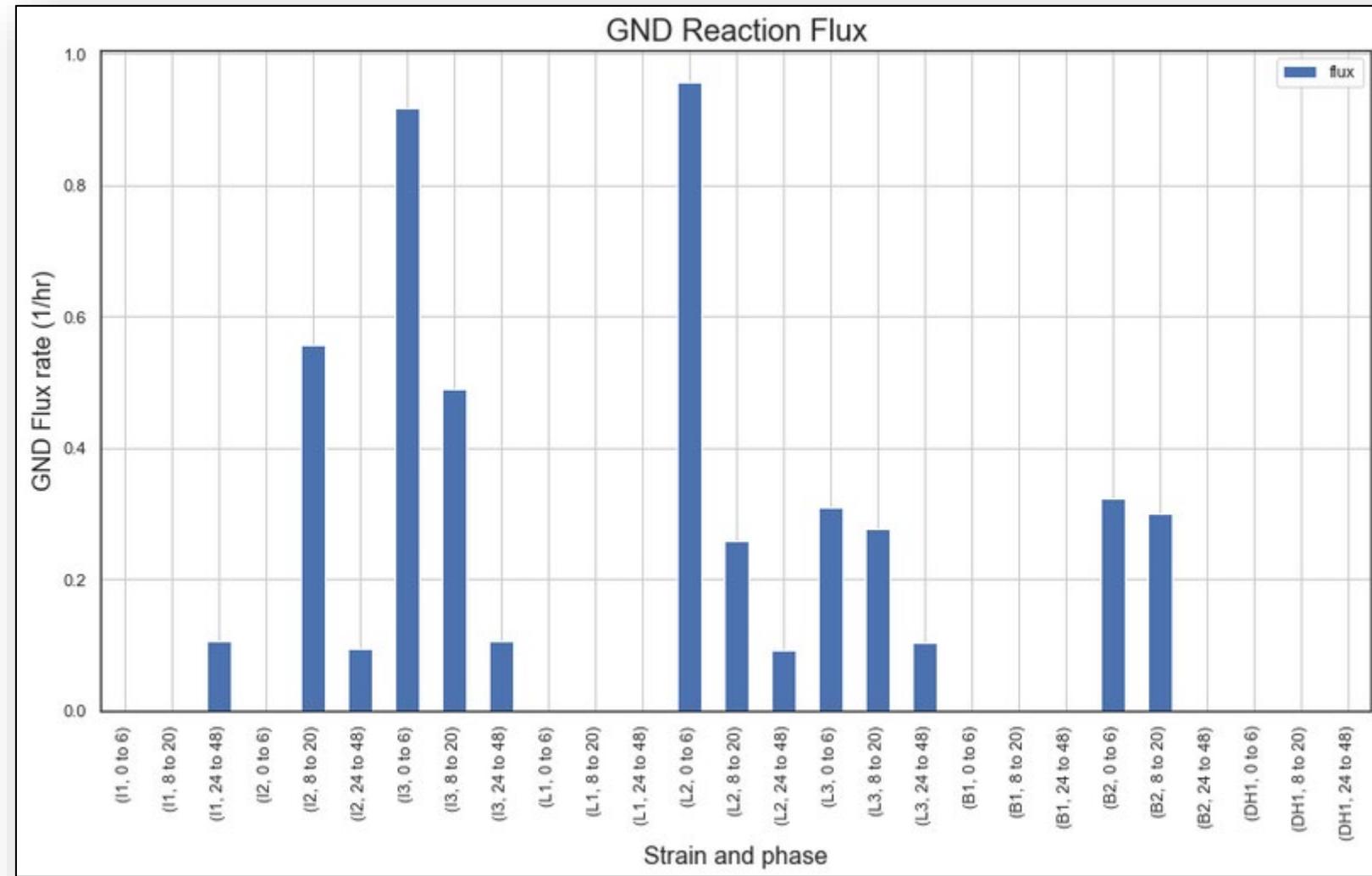
Comparing FBA Solutions to Experimental Growth Rates



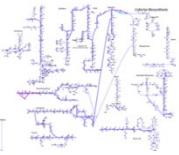
Stage_three_GEM_revised_part_1.ipynb



Comparing Reaction Flux Across Strains



Stage_three_GEM_revised_part_1.ipynb



Creating OMICS-based Mevalonate Biofuel COBRA Models

Adapted from "Stage_three_GEM notebook" by Brunck [1]

Load the appropriate Python packages

```
In [1]: import cobra.test
from cobrapy_bigg_client import client

# Panda python module for dataframe and data storage/manipulation
import pandas as pd
pd.set_option('mode.use_inf_as_na',True)
pd.set_option('display.max_rows', 500)
pd.set_option('display.max_columns', 999)
pd.set_option('precision', 3)

from contextlib import contextmanager
import sys, os

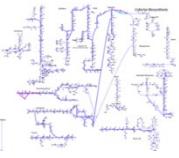
@contextmanager
def suppress_stdout():
    with open(os.devnull, "w") as devnull:
        old_stdout = sys.stdout
        sys.stdout = devnull
        try:
            yield
        finally:
            sys.stdout = old_stdout
```

Load the "iJO1366 Mevalonate Pathways" model

```
In [2]: model_orig = cobra.io.load_json_model("iJO1366_mevalonate_pathways.json");
model_orig.solver = 'glpk'
model = model_orig.copy();

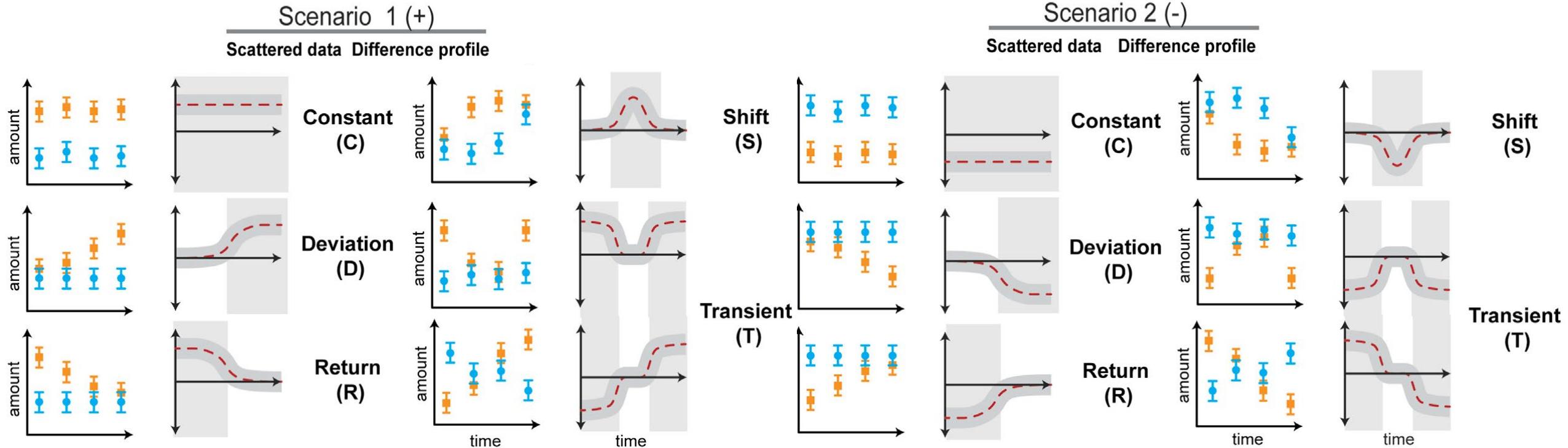
Set parameter Username
Academic license - for non-commercial use only - expires 2022-10-10
```

Creating_OMICS-based_Mevalonate_Biofuel_COBRA_Models.ipynb

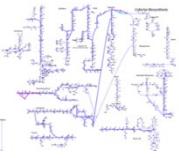


Difference Profiles

test (e.g. strain X) control (e.g. wild-type) difference (test - control)

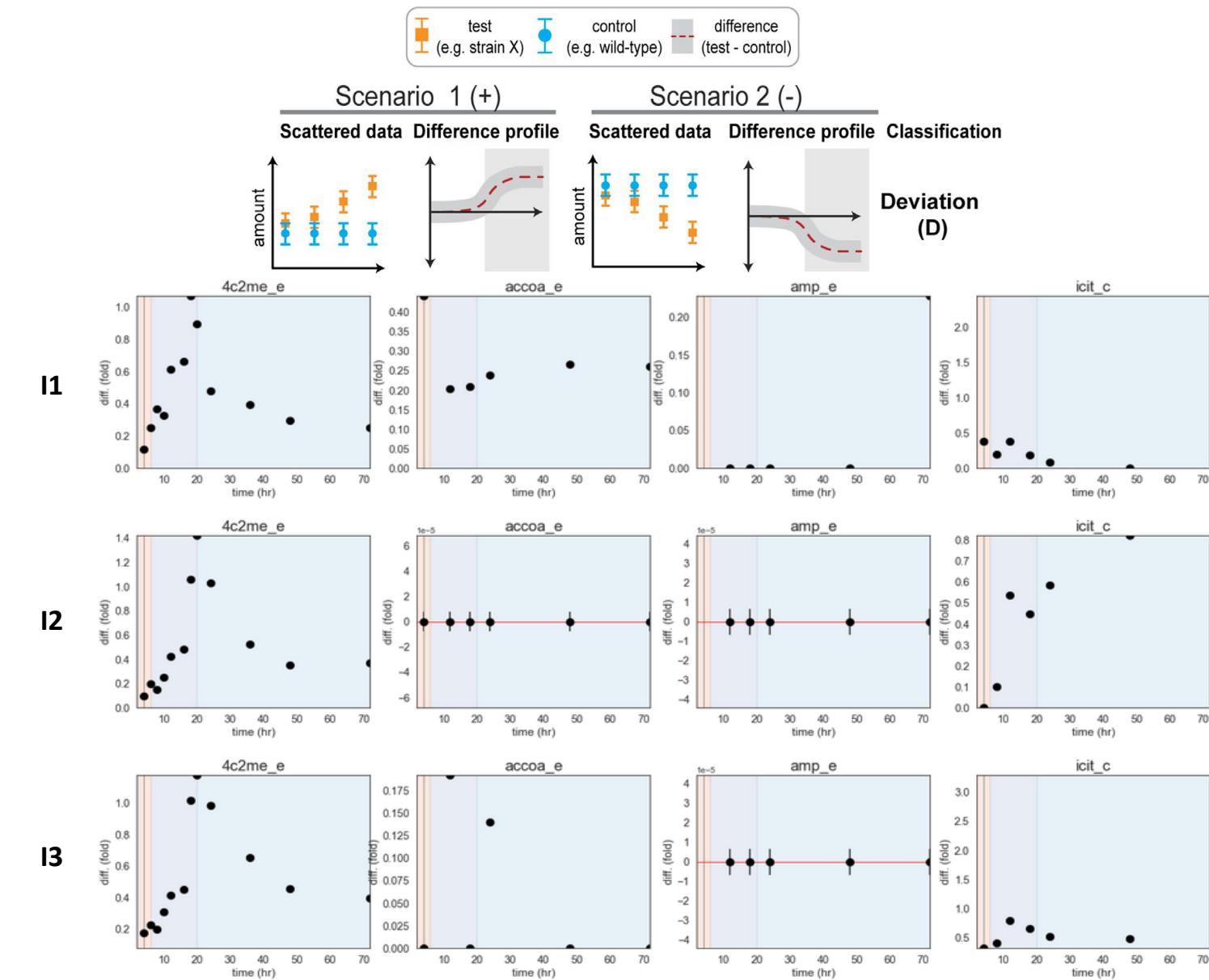


Brunk, Elizabeth, et al. "Characterizing strain variation in engineered *E. coli* using a multi-omics-based workflow." *Cell systems* 2.5 (2016): 335-346



Measured Difference Profiles

Stage_one_Dynamic_Differences_Revised.ipynb





Metabolite Z-Score Heat Maps

- Compare the z-score of a metabolite from a new strain to its WT value.
- What does a z-score mean:
 - ✓ Negative Z score indicates depletion:
 $WT \text{ flux} > \text{strain flux}$ ($\text{strain} - \text{WT} < 0$)
 - ✓ Positive Z score indicates enrichment:
 $WT \text{ flux} < \text{strain flux}$ ($\text{strain} - \text{WT} > 0$)

	Change in Metabolite Z-scores with respect to Wild Data								
	I1	I2	I3	L1	L2	L3	B1	B2	
accoa_c	0.72	0.42	0.41	1.19	0.5	0.52	0.81	0.33	
akg_c	0.89	0.92	1.0	2.29	0.29	0.28	1.61	0.47	
acon_C_c	0.42	0.3	0.56	0.69	0.12	0.4	0.66	0.37	
cit_c	0.57	0.51	0.7	0.96	0.17	0.39	0.66	1.02	
icit_c	0.39	0.0	0.32	0.38	0.0	0.13	0.29	0.19	
glx_c	1.74	1.68	1.34	0.74	0.93	0.7	0.84	0.94	
glyclt_c	0.45	0.37	0.33	0.31	0.24	1.37	0.37	0.29	
fdp_c	2.08	1.96	1.75	0.86	1.16	1.28	1.18	0.81	
3pg_c	1.41	1.01	1.24	1.03	0.73	1.01	1.26	0.85	
pep_c	1.02	0.51	0.86	0.92	0.28	0.65	0.91	0.71	
nadp_c	1.22	0.88	1.16	1.04	2.73	0.93	1.76	1.09	
nad_c	0.64	0.42	0.62	0.86	0.55	0.82	0.95	0.63	
amp_c	0.62	0.46	0.63	0.72	0.32	0.56	0.69	0.44	

Hour = 4

Stage_one_Dynamic_Differences_Revised.ipynb



Stage One: Dynamic Differences (Revised Version)

```
In [1]: import os
%matplotlib inline

# COBRA (genome-scale modeling) module and tools
import cobra
import cobra.io
from cobra import Model, Reaction, Metabolite
from cobra.io.mat import load_matlab_model
from cobra.io.mat import save_matlab_model
from scipy.io import loadmat, savemat
from cobra.flux_analysis import variability
import theseus # 

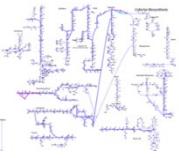
# Panda python module for dataframe and data storage/manipulation
import pandas as pd
pd.set_option('mode.use_inf_as_na',True)
pd.set_option('display.max_rows', 100)
pd.set_option('display.max_columns', 999)
pd.set_option('precision', 3)

# seaborn visualization tools
import seaborn as sns
sns.set(style="white")
c1, c2, c3, c4, c5, c6, c7, c8, c9, c10 = sns.color_palette("Set2", 10)
b1, b2, b3, b4, b5, b6 = sns.color_palette("Blues")

# other
import warnings
warnings.simplefilter('ignore', DeprecationWarning)

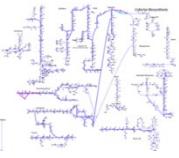
# statistical and math toolkits
import numpy as np
import scipy.io
print("I'm using scipy version: ")
scipy.__version__
```

Stage_one_Dynamic_Differences_Revised.ipynb



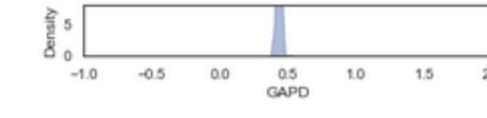
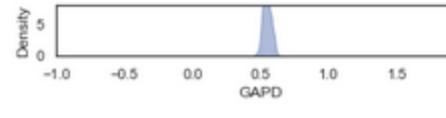
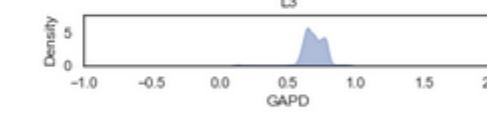
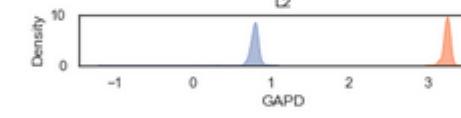
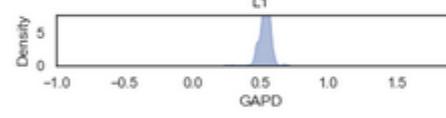
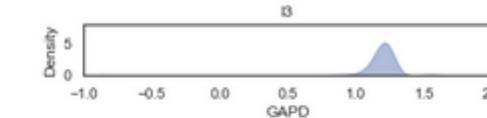
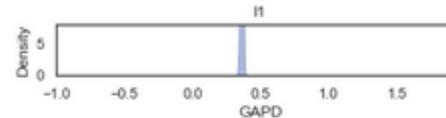
Omics

- Introduction
- Biofuel Example - Metabolomics & Proteomics
 - ✓ Creating an *in silico* COBRA model
 - ✓ Creating the plasmids for the cell factory
 - ✓ Extracting metabolomics & proteomics data
 - ✓ Creating context-sensitive *in silico* models
 - ✓ Extracting data from context-sensitive model
 - Flux balance analysis
 - ▪ Sampling analysis
- Conclusion

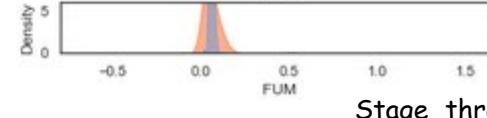
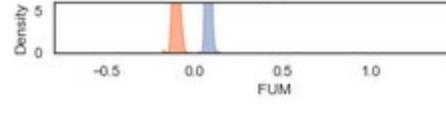
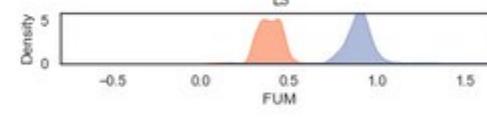
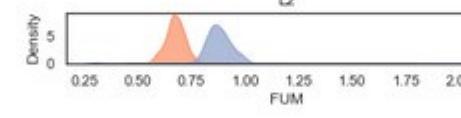
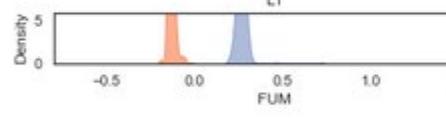
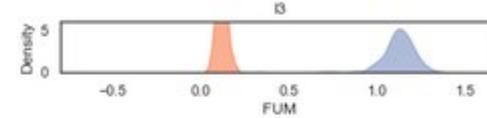
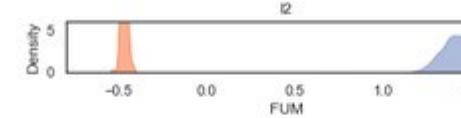
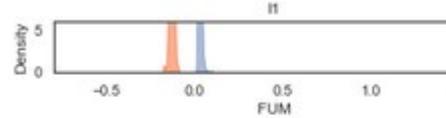


Sampling Analysis

Glyceraldehyde-3-phosphate dehydrogenase



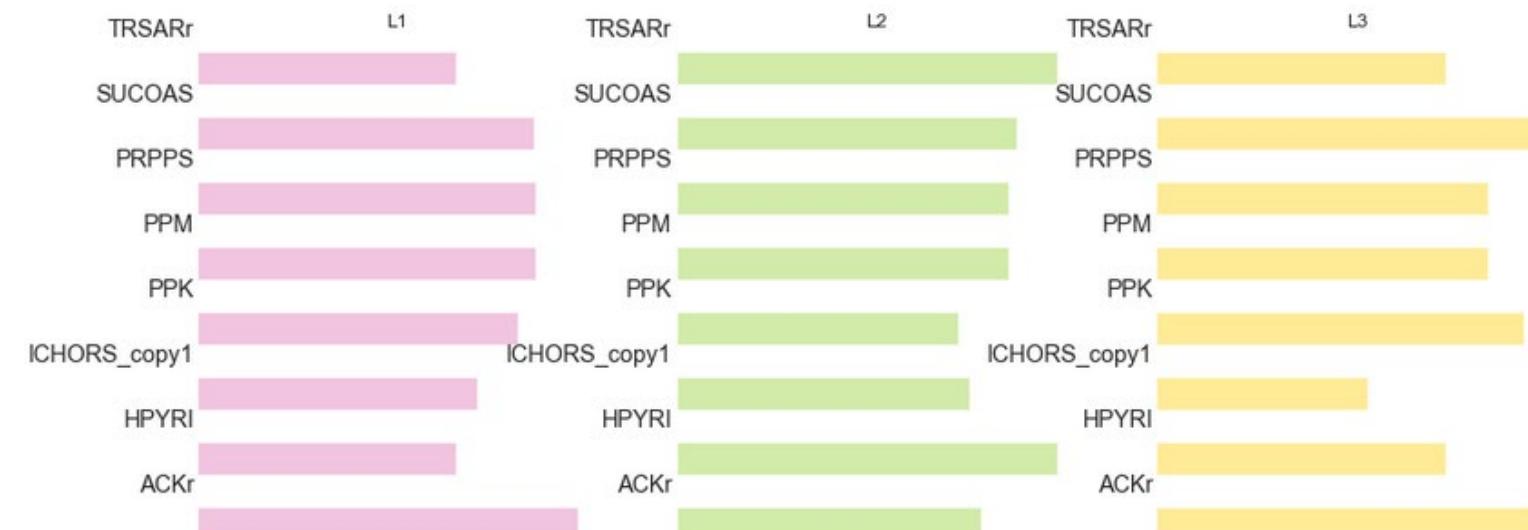
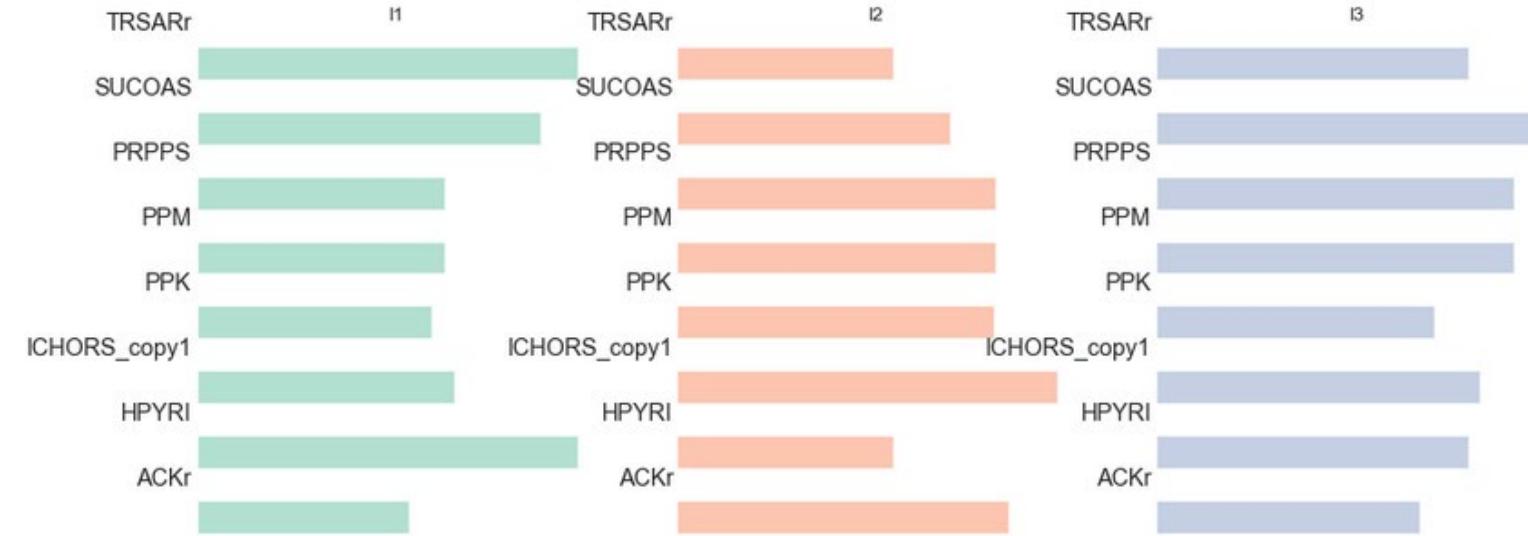
Fumarase



Stage_three_GEM_revised_part_2.ipynb



Z-scores of Reactions by Strain





Stage Three: Sampling the context-sensitive models

Modified version of supplementary material from "Brunk, Elizabeth, et al. "Characterizing strain variation in engineered *E. coli* using a multi-omics-based workflow." Cell Systems 2.5 (2016): 335-346.

```
In [1]: import cobra.test
from cobra.test import create_test_model
import numpy as np
import pandas as pd
from cobra.sampling import sample
import matplotlib.pyplot as plt
pd.set_option('display.max_rows', 500)

#model = create_test_model("ecoli")
#model.solver = 'glpk'
```

Create lists of models to be sampled

```
In [2]: model_files = ['I1_0_6.json', 'I1_8_20.json', 'I1_24_48.json',
'I2_0_6.json', 'I2_8_20.json', 'I2_24_48.json',
'I3_0_6.json', 'I3_8_20.json', 'I3_24_48.json',
'L1_0_6.json', 'L1_8_20.json', 'L1_24_48.json',
'L2_0_6.json', 'L2_8_20.json', 'L2_24_48.json',
'L3_0_6.json', 'L3_8_20.json', 'L3_24_48.json',
'B1_0_6.json', 'B1_8_20.json', 'B1_24_48.json',
'B2_0_6.json', 'B2_8_20.json', 'B2_24_48.json',
'DH1_0_6.json', 'DH1_8_20.json', 'DH1_24_48.json']

model_files_p1 = ['I1_0_6.json', 'I2_0_6.json', 'I3_0_6.json',
'L1_0_6.json', 'L2_0_6.json', 'L3_0_6.json',
'B1_0_6.json', 'B2_0_6.json', 'DH1_0_6.json']
model_files_p2 = ['I1_8_20.json', 'I2_8_20.json', 'I3_8_20.json',
'L1_8_20.json', 'L2_8_20.json', 'L3_8_20.json',
'B1_8_20.json', 'B2_8_20.json', 'DH1_8_20.json']
model_files_p3 = ['I1_24_48.json', 'I2_24_48.json', 'I3_24_48.json',
'L1_24_48.json', 'L2_24_48.json', 'L3_24_48.json',
'B1_24_48.json', 'B2_24_48.json', 'DH1_24_48.json']
```

Sampling_data.ipynb



Stage Three: Part 2 - Sampling Analysis

Adapted from "Stage_three_GEM notebook" by Brunck [1]

Background

In this notebook we provide an example of taking multi-omics data (metabolomics and proteomics) and explore 9 different strains of *E. coli* (8 engineered strains and 1 wild-type, DH1, strain) using sampling analysis.

Load the standard COBRApy and python packages

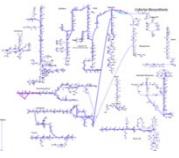
```
In [1]: import os
%matplotlib inline

import cobra
import cobra.io
from cobra import Model, Reaction, Metabolite
#from cobra.io.sbml import create_cobra_model_from_sbml_file
#from cobra.io.sbml import write_cobra_model_to_sbml_file
from cobra.io.mat import load_matlab_model
from cobra.io.mat import save_matlab_model
from scipy.io import loadmat, savemat
import pandas as pd
import matplotlib.pyplot as plt

from cobra.flux_analysis import variability

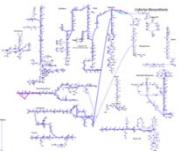
# Panda python module for dataframe and data storage/manipulation
import pandas as pd
pd.set_option('mode.use_inf_as_na',True)
pd.set_option('display.max_rows', 100)
pd.set_option('display.max_columns', 999)
pd.set_option('precision', 3)
```

Stage_three_GEM_revised_part_2.ipynb



Omics

- Introduction
- Biofuel Example - Metabolomics & Proteomics
 - ✓ Creating an *in silico* COBRA model
 - ✓ Creating the plasmids for the cell factory
 - ✓ Extracting metabolomics & proteomics data
 - ✓ Creating context-sensitive *in silico* models
 - ✓ Extracting data from context-sensitive model
 - Flux balance analysis
 - Sampling analysis
- Conclusion

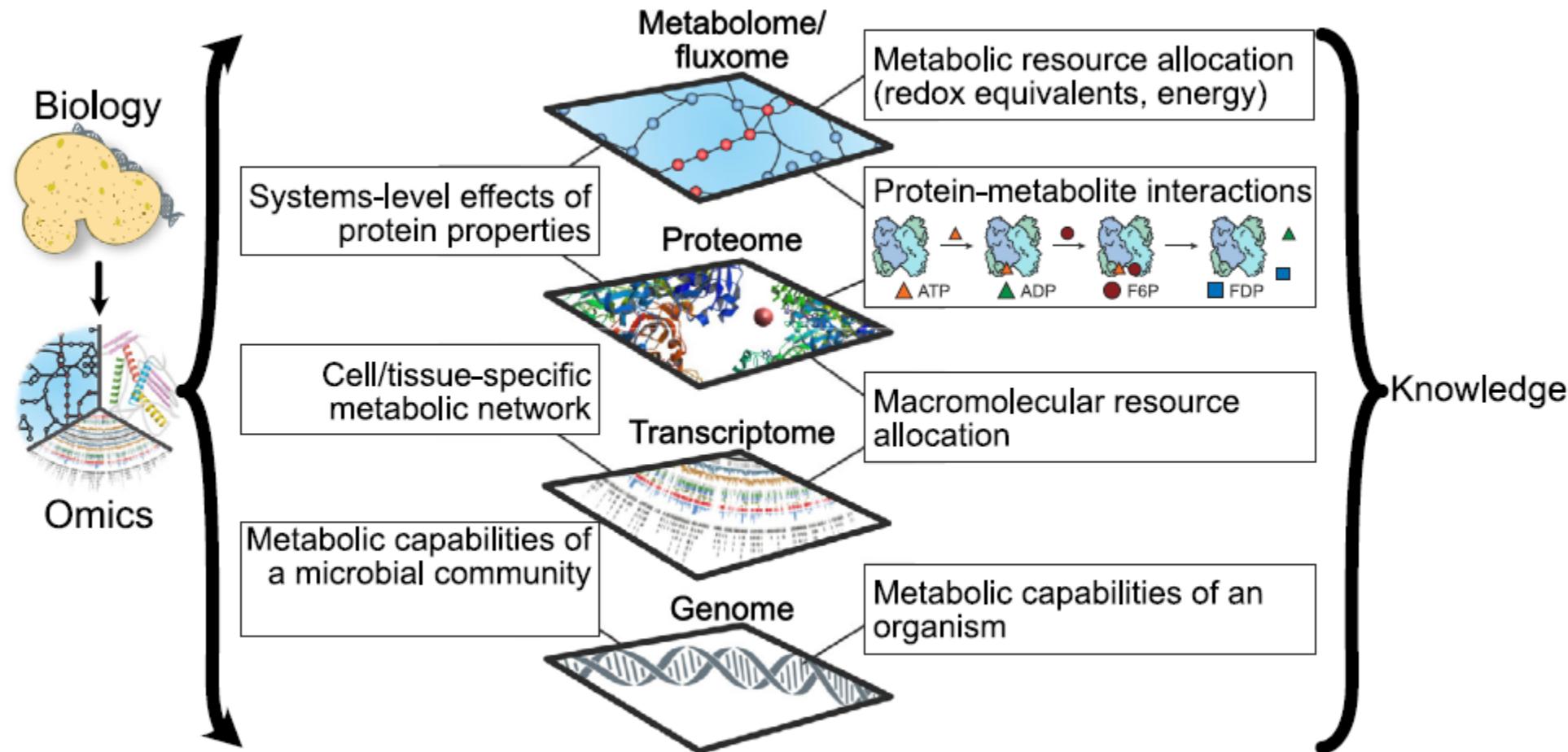


References

1. Dahal, Sanjeev, et al. "Synthesizing Systems Biology Knowledge from Omics Using Genome-Scale Models." *Proteomics* 20.17-18 (2020): 1900282.
2. Brunk, Elizabeth, et al. "Characterizing strain variation in engineered *E. coli* using a multi-omics-based workflow." *Cell systems* 2.5 (2016): 335-346.
3. Martin, Vincent JJ, et al. "Engineering a mevalonate pathway in *Escherichia coli* for production of terpenoids." *Nature biotechnology* 21.7 (2003): 796-802.
4. George, Kevin W., et al. "Correlation analysis of targeted proteins and metabolites to assess and engineer microbial isopentenol production." *Biotechnology and bioengineering* 111.8 (2014): 1648-1658.
5. Alonso-Gutierrez, Jorge, et al. "Metabolic engineering of *Escherichia coli* for limonene and perillyl alcohol production." *Metabolic engineering* 19 (2013): 33-41.
6. Peralta-Yahya, Pamela P., et al. "Identification and microbial production of a terpene-based advanced biofuel." *Nature communications* 2.1 (2011): 1-8.
7. Peralta-Yahya, Pamela P., et al. "Microbial engineering for the production of advanced biofuels." *Nature* 488.7411 (2012): 320-328.
8. Hanahan, Douglas. "Studies on transformation of *Escherichia coli* with plasmids." *Journal of molecular biology* 166.4 (1983): 557-580.



Graphical Overview of Synthesizing Knowledge Using OMIC Data and Genome-scale Models.



Dahal, Sanjeev, et al. "Synthesizing Systems Biology Knowledge from Omics Using Genome-Scale Models." *Proteomics* 20.17-18 (2020): 1900282.