BT5240: Computational Systems Biology

Assignment 3 - Report

By :
Shreya Rajagopalan
BE21B038

1 Knockout and overexpression targets to improve ethanol production in Saccharomyces pastorianus

Ethanol is a valuable byproduct with diverse applications, and our primary objective is to identify potential targets to enhance the alcoholic content of the yeast strain using the GSMM, Yeast 8.5.0 model (Download the model in '.xml' format from https://zenodo.org/records/5062615).

a. Growth characterization of S.pastorianus on other carbon substrate is presented. Analyze the conditions in the Yeast 8.5.0 model and report whether the yeast model can be used for the metabolic analysis in S.pastorianus

		Growth [h]
	Consumption rate [mmol g−1 DCW h−1]	**Experimental value**
Galactose	3.983182	0.3154
Sucrose	2.084552	0.3324
Trehalose	2.142704	0.3628
Fructose	4.406842	0.3577
Glucose	4.187956	0.3670

Figure 1: Growth rate of *S.pastorianus*

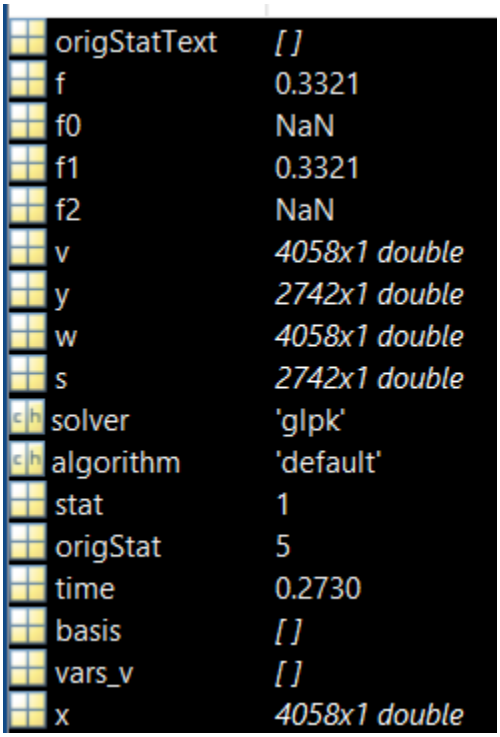
Solution:

For each carbon substrate, on performing FBA with the given consumption rates as lower bounds and setting the other 4 carbon substrates to zero, we get the following results.

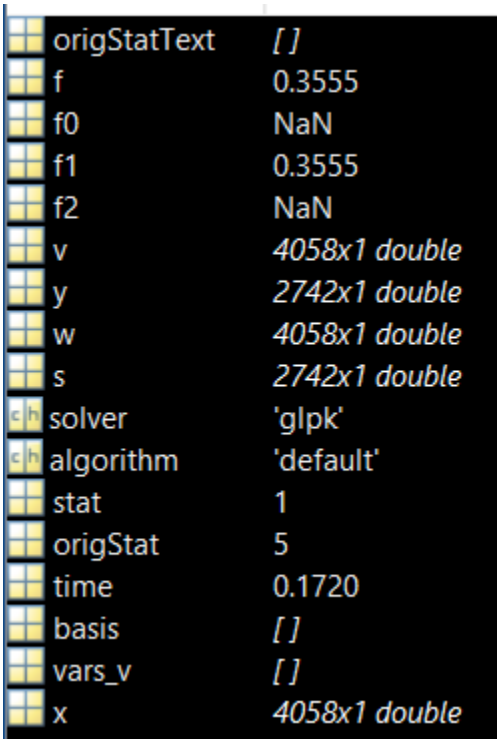
Here 'f' in the struct refers to the biomass flux indicating growth rate, which is similar to the experimental values given in all the cases.

Hence we can say that the Yeast 8.5.0 model can be used for metabolic analysis in *S. pastorianus*.

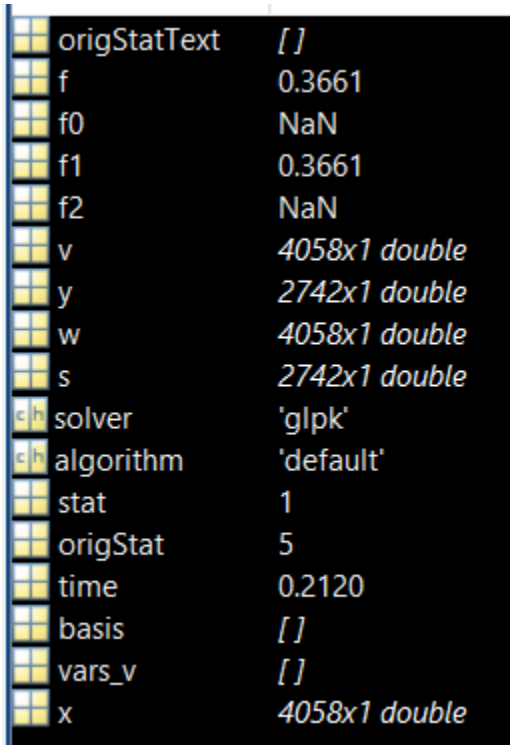
a. Galactose



b. Sucrose

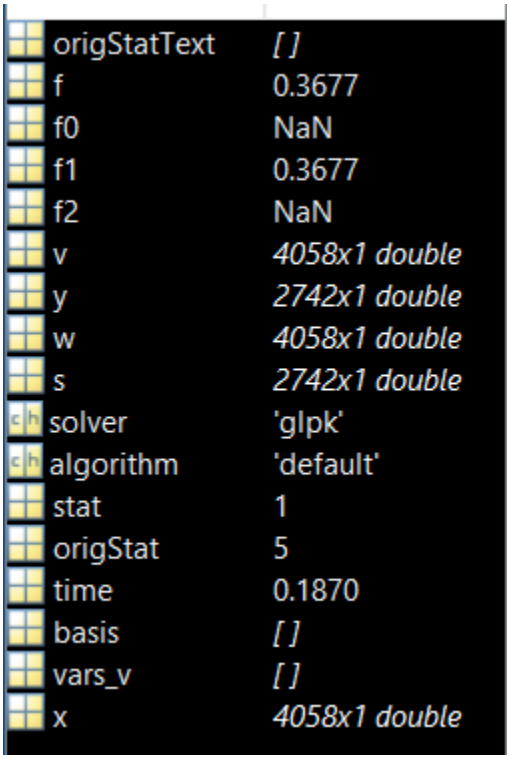


c. Trehalose



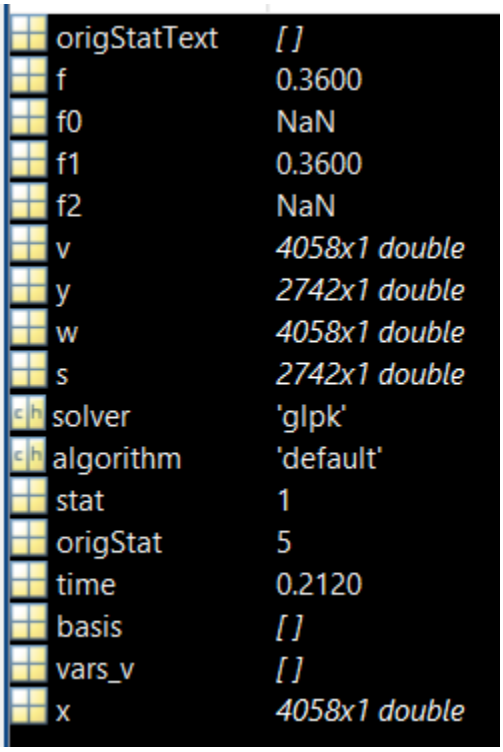
origStatText	[ ]
f	0.3661
f0	NaN
f1	0.3661
f2	NaN
v	4058x1 double
y	2742x1 double
w	4058x1 double
s	2742x1 double
solver	'glpk'
algorithm	'default'
stat	1
origStat	5
time	0.2120
basis	[ ]
vars_v	[ ]
x	4058x1 double

d. Fructose



origStatText	[ ]
f	0.3677
f0	NaN
f1	0.3677
f2	NaN
v	4058x1 double
y	2742x1 double
w	4058x1 double
s	2742x1 double
solver	'glpk'
algorithm	'default'
stat	1
origStat	5
time	0.1870
basis	[ ]
vars_v	[ ]
x	4058x1 double

e. Glucose



b. On performing the FSEOF algorithm to identify overexpression and knockout targets, the following reactions were found:

Overexpression targets:

1. 'glucose-6-phosphate isomerase'
2. 'glyceraldehyde-3-phosphate dehydrogenase'
3. 'phosphoglycerate kinase'
4. 'pyruvate decarboxylase'
5. 'triose-phosphate isomerase'
6. 'water diffusion'
7. 'ethanol exchange'
8. 'ethanol transport'
9. 'water diffusion'

Knockout targets:

1. 'ferrocytochrome-c:oxygen oxidoreductase'
2. 'ubiquinol:ferricytochrome c reductase'
3. 'ribose-5-phosphate isomerase'
4. 'O2 transport'
5. 'water exchange'

2 Drug targets for Mycobacterium tuberculosis

M.tuberculosis is a unique pathogen that displays an array of complex lipids on its surface that play an important role in its pathogenesis. The genes responsible for the synthesis of these lipids are interesting drug targets that are routinely explored. The following is a list of genes that play an important role in virulence of Tuberculosis. (Download the model from http://bigg.ucsd.edu/models/iEK1008)

Gene name	Gene ID
fbpA	Rv3804c
modA	Rv1857
kasB	Rv2246
pks7	Rv1662
icl1	Rv0467
mce4	Rv3496c
mbtB	Rv2383c

Explore the following questions:

a. Are these genes essential for Tuberculosis?

b. What are the metabolites and reactions associated with these genes and what role are they playing?

c. Identify if any of these genes above form a synthetic lethal pair.

Solution:

a. On performing single gene deletion of the given genes using the function *singleGeneDeletion(),* the computed growth rate ratio between deletion strain and wild type (grRatio) is 0 for the third gene *kasB* and 1 for others. This shows that the cell stops growing when gene *kasB* is deleted but the cell survives for the other 6 single gene deletions. Thus *kasB* is an essential gene for Tuberculosis among the given genes.

b. There are a total of 19 reactions and 54 metabolites that are associated with the given genes

Reaction ID	Reaction	Reaction Name	Associated genes	Associated gene from the given list	Associated metabolites
FBPA'	h2o[c] + 2 tre6mm[c] -> 2 h[c] + tdm3[c] + tre[c] '	MNXR90624'	(Rv3804c and Rv1886c and Rv0129c and Rv3803c)'	gene_Rv3804c'	'h2o[c]' 'h[c]' 'tdm3[c]' 'tre[c]' 'tre6mm[c]'
MOBDabc'	atp[c] + h2o[c] + mobd[e] -> pi[c] + h[c] + adp[c] + mobd[c] '	Molybdate transport via ABC system'	(Rv1857 and Rv1858 and Rv1859)'	gene_Rv1857'	'atp[c]' 'h2o[c]' 'pi[c]' 'h[c]' 'adp[c]' 'mobd[c]' 'mobd[e]'
MOBDabc_rev'	atp[c] + h2o[c] + mobd[c] -> pi[c] + h[c] + adp[c] + mobd[e] '	T06232b'	(Rv1857 and Rv1858 and Rv1859)'	gene_Rv1857'	'atp[c]' 'h2o[c]' 'pi[c]' 'h[c]' 'adp[c]' 'mobd[c]' 'mobd[e]'
MYCSacp50'	33 h[c] + 24 nadph[c] + hexc[c] + 12 malACP[c] -> 13 h2o[c] + 24 nadp[c] + 11 ACP[c] + 12 co2[c] + meroacidACP[c] '	MYCSacp50'	((Rv1483 and Rv2245 and Rv2246) or (Rv0098 and Rv2245 and Rv2246))'	gene_Rv2246'	'h2o[c]' 'h[c]' 'nadph[c]' 'nadp[c]' 'ACP[c]' 'co2[c]' 'hexc[c]' 'malACP[c]' 'meroacidACP[c]'
MYCSacp56'	42 h[c] + 30 nadph[c] + hexc[c] + 15 malACP[c] -> 16 h2o[c] + 30 nadp[c] + 14 ACP[c] + 15 co2[c] + mmeroacidACP[c] '	MYCSacp56'	((Rv1483 and Rv2245 and Rv2246) or (Rv0098 and Rv2245 and Rv2246))'	gene_Rv2246'	'h2o[c]' 'h[c]' 'nadph[c]' 'nadp[c]' 'ACP[c]' 'co2[c]' 'hexc[c]' 'malACP[c]' 'mmeroacidACP[c]'
MYCSacp58'	47 h[c] + 30 nadph[c] + hexc[c] + 16 malACP[c] -> 17 h2o[c] + 30 nadp[c] + 15 ACP[c] + 16 co2[c] + kmeroacidACP[c] '	MYCSacp58'	((Rv1483 and Rv2245 and Rv2246) or (Rv0098 and	gene_Rv2246'	'h2o[c]' 'h[c]' 'nadph[c]' 'nadp[c]'

			Rv2245 and Rv2246))'		'ACP[c]' 'co2[c]' 'hexc[c]' 'malACP[c]' 'kmeroacidACP[c]'
FASm1601'	hdca[c] + 3 h[c] + 2 nadph[c] + mmcoa__S[c] -> h2o[c] + 2 nadp[c] + coa[c] + co2[c] + m1ocdca[c] '	FASm1601'	(Rv1663 or Rv1662)'	gene_Rv1662'	'hdca[c]' 'h2o[c]' 'h[c]' 'nadph[c]' 'nadp[c]' 'coa[c]' 'co2[c]' 'mmcoa__S[c]' 'm1ocdca[c]'
FASm180'	15 h[c] + 10 nadph[c] + 4 malcoa[c] + mmcoa__S[c] + octa[c] -> 5 h2o[c] + 10 nadp[c] + mocdca[c] + 5 coa[c] + 5 co2[c] '	FASm180'	(Rv1663 or Rv1662)'	gene_Rv1662'	'h2o[c]' 'h[c]' 'nadph[c]' 'nadp[c]' 'mocdca[c]' 'coa[c]' 'co2[c]' 'malcoa[c]' 'mmcoa__S[c]' 'octa[c]'
FASm1801'	3 h[c] + 2 nadph[c] + mmcoa__S[c] + m1ocdca[c] -> h2o[c] + 2 nadp[c] + coa[c] + co2[c] + dmarach[c] '	FASm1801'	(Rv1663 or Rv1662)'	gene_Rv1662'	'h2o[c]' 'h[c]' 'nadph[c]' 'nadp[c]' 'coa[c]' 'co2[c]' 'mmcoa__S[c]' 'm1ocdca[c]' 'dmarach[c]'
FASm2001'	3 h[c] + 2 nadph[c] + mmcoa__S[c] + dmarach[c] -> h2o[c] + 2 nadp[c] + coa[c] + co2[c] + tmbhn[c] '	FASm2001'	(Rv1663 or Rv1662)'	gene_Rv1662'	'h2o[c]' 'h[c]' 'nadph[c]' 'nadp[c]' 'coa[c]' 'co2[c]' 'mmcoa__S[c]' 'dmarach[c]' 'tmbhn[c]'
FASm2002'	ocdca[c] + 3 h[c] + 2 nadph[c] + mmcoa__S[c] -> h2o[c] + 2 nadp[c] + coa[c] + co2[c] + marach[c] '	FASm2002'	(Rv1663 or Rv1662)'	gene_Rv1662'	'ocdca[c]' 'h2o[c]' 'h[c]' 'nadph[c]' 'nadp[c]' 'coa[c]'

					'co2[c]' 'mmcoa__S[c]' 'marach[c]'
ICL'	icit[c] -> succ[c] + glx[c] '	Isocitrate lyase'	(Rv0467 or (Rv1915 and Rv1916))'	gene_Rv0467'	'icit[c]' 'succ[c]' 'glx[c]'
MCITL2'	micit[c] -> pyr[c] + succ[c] '	Methylisocitrate lyase'	Rv0467'	gene_Rv0467'	'pyr[c]' 'succ[c]' 'micit[c]'
CHSTEROLt_1'	atp[c] + h2o[c] + chsterol[e] -> pi[c] + h[c] + adp[c] + chsterol[c] '	MNXR1163'	(Rv3499c and Rv3498c and Rv3497c and Rv3496c and Rv3495c and Rv3494c and Rv3500c and Rv3501c)'	gene_Rv3496c'	'atp[c]' 'h2o[c]' 'pi[c]' 'h[c]' 'adp[c]' 'chsterol[e]' 'chsterol[c]'
MCBTS'	atp[c] + h[c] + 2 o2[c] + bhb[c] + 2 lys__L[c] + odecoa[c] + salc[c] + thr__L[c] -> 5 h2o[c] + pi[c] + coa[c] + co2[c] + adp[c] + mcbts[c] '	MCBTS'	(Rv2378c and Rv2379c and Rv2380c and Rv2381c and Rv2382c and Rv2383c and Rv2384)'	gene_Rv2383c'	'atp[c]' 'h2o[c]' 'pi[c]' 'h[c]' 'coa[c]' 'co2[c]' 'o2[c]' 'adp[c]' 'bhb[c]' 'lys__L[c]' 'odecoa[c]' 'salc[c]' 'thr__L[c]' 'mcbts[c]'
MCBTS2'	atp[c] + h[c] + 2 o2[c] + bhb[c] + ser__L[c] + 2 lys__L[c] + odecoa[c] + salc[c] -> 5 h2o[c] + pi[c] + coa[c] + co2[c] + adp[c] + mcbtt[c] '	MCBTS2'	(Rv2378c and Rv2379c and Rv2380c and Rv2381c and Rv2382c and Rv2383c and Rv2384)'	gene_Rv2383c'	'atp[c]' 'h2o[c]' 'pi[c]' 'h[c]' 'coa[c]' 'co2[c]' 'o2[c]' 'adp[c]' 'bhb[c]' 'ser__L[c]' 'lys__L[c]' 'odecoa[c]' 'mcbtt[c]' 'salc[c]'
MCBTS3'	atp[c] + 2 nadp[c] + 3.5 o2[c] + bhb[c] + ser__L[c] + 2 lys__L[c] + occoa[c] + salc[c] -> 5 h2o[c] + pi[c] + 2 h[c] + 2 nadph[c] +	MCBTS3'	(Rv2378c and Rv2379c and Rv2380c and	gene_Rv2383c'	'atp[c]' 'h2o[c]' 'pi[c]'

	coa[c] + co2[c] + adp[c] + cmcbtt[c] '		Rv2381c and Rv2382c and Rv2383c and Rv2384)'		'h[c]' 'nadph[c]' 'nadp[c]' 'coa[c]' 'co2[c]' 'o2[c]' 'adp[c]' 'bhb[c]' 'ser__L[c]' 'cmcbtt[c]' 'lys__L[c]' 'occoa[c]' 'salc[c]'
MBTA3'	3 h[c] + acac[c] + nadh[c] + salc[c] + thr__L[c] + n6hlysmal[c] + n6hlys[c] -> 6 h2o[c] + nad[c] + mcbtm[c] '	MBTA3'	(Rv2384 and Rv2383c and Rv2382c and Rv2381c and Rv2380c and Rv2379c and Rv2377c)'	gene_Rv2383c'	'h2o[c]' 'h[c]' 'acac[c]' 'nad[c]' 'nadh[c]' 'salc[c]' 'thr__L[c]' 'n6hlysmal[c]' 'n6hlys[c]' 'mcbtm[c]'
MBTA1'	h[c] + acac[c] + nadh[c] + salc[c] + thr__L[c] + n6hlysmal[c] + n6hlys[c] -> 6 h2o[c] + nad[c] + mcbtt[c] '	MBTA1'	(Rv2384 and Rv2383c and Rv2382c and Rv2381c and Rv2380c and Rv2379c and Rv2377c)'	gene_Rv2383c'	'h2o[c]' 'h[c]' 'acac[c]' 'nad[c]' 'nadh[c]' 'mcbtt[c]' 'salc[c]' 'thr__L[c]' 'n6hlysmal[c]' 'n6hlys[c]'

c. On performing double gene deletions using the function *doubleGeneDeletions()* we see that the grRatio is 0 only for gene pair(*kasB, kasB)*. However since *kasB* is a single lethal, it cannot be a synthetic lethal, since for a synthetic lethal pair deletion of either of the genes should not kill the cell. Therefore, among the given genes, there are no synthetic lethal pairs.