Automatic identification and restoration of reaction gaps in the consensus reconstruction network for yeast metabolism

Lu, Chuan; King, Ross Donald

Publication date:
2010

Citation for published version (APA):
Lu, C., & King, R. D. (2010). *Automatic identification and restoration of reaction gaps in the consensus reconstruction network for yeast metabolism*. 1-71. Poster session presented at Systems Biology of Microorganisms, Paris, France. http://hdl.handle.net/2160/4634

Automatic identification and restoration of reaction gaps in the consensus reconstruction network for yeast metabolism

Chuan Lu, Ross D. King

Dept. of Computer Science, Aberystwyth University, Wales, UK

Introduction

Motivation

- An automated procedure of identifying and filling of reaction gaps in genome-scale metabolic networks within the framework of flux balance analysis.
- To identify the non-producible metabolites in the network
  - constraint-based optimisation techniques
  - graph traverse algorithm
- Search for reactions to add into the model to restore the reachability of the metabolites or clusters of metabolites
- This is part of an iterative process of converting a genome-scale reconstruction into an executable computational model:
  - representing the reactions in mathematical form,
  - validating and refining the mathematical model.

Consensus network reconstruction for yeast metabolism

- Yeastnet1.0: community driven, rigorously evidenced and well annotated [1] http://www.comp-sys-bio.org/yeastnet/
- Yeastnet2.0: a recent expanded network reconstruction that includes a detailed and evidence description of lipid metabolism.
- Yeastnet2.0: 1834 unique chemical reactions, 886 ORFs and 1418 metabolites located in 15 different compartments.
- **Need for automated procedure for network validation.**

Method

Background

- Flux Balance Analysis (FBA)
- Structural Gaps in metabolic networks
  Reaction gaps, orphan enzymes, …
- Mechanisms to rescue reaction gaps in the networks
  => reversibility; transportation; cell consumption
  => adding missing reactions from reference model
  => metabolite exchange (uptake or secretion)
- Focus on bridging gaps that block the cell from producing some metabolites: assuming all metabolites are all consumable, all reactions reversible.

Procedure

- Gap identification: flux range analysis
- Blocked metabolite clustering:
  - Check metabolite reachability with blockage network
  - Blockage network: formed by blocked reactions.
- Finding pseudo root blocked metabolites: graph traverse.
- Gap filling: mixed integer linear programming, principle of minimum metaboli…

$$\begin{aligned}
&\text{Minimize} \sum \mathbf{a} \\
&\text{s.t. } \mathbf{Sv} + \mathbf{Uy} = 0, \\
&v_{\min,i} \le v_i \le v_{\max,i}, \forall i \in \mathrm{R} \text{ in current model} \\
&a_j v_{\min,j} \le y_j \le a_j v_{\max,j}, \forall j \in \mathrm{R} \text{ in database} \\
&v_k \ge \varepsilon - M(1-b_k), \forall k \in \mathrm{R} \text{ with metabolite m as product} \\
&v_l \le -\varepsilon + M(1-b_l), \forall l \in \mathrm{R} \text{ with metabolite m as reactant and reversible} \\
&\sum \mathbf{b} > 0 \mathbf{a}, \mathbf{b} \in \{0,1\} \\
&\text{where } \varepsilon, M > 0,\ \varepsilon \to 0,\ \mathrm{M} \to \infty; \\
&\mathbf{S} \text{ and } \mathbf{U} \text{ are stoichiometric matrices for model and database, respectively.}
\end{aligned}$$

❖ Computational tools implemented in python, LP solver lpsolve5.

Applications

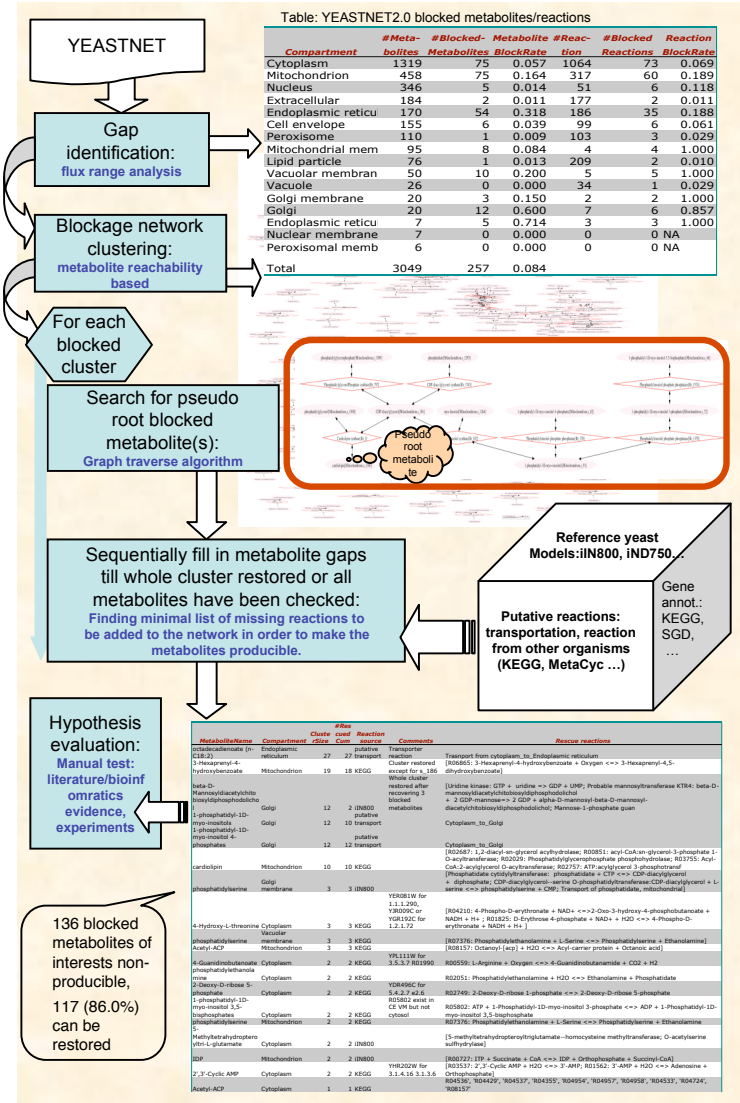

Table: YEASTNET2.0 blocked metabolites/reactions

Compartment	#Metabolites	#Blocked-Metabolites	Metabolite BlockRate	#Reaction	#Blocked Reactions	Reaction BlockRate
Cytoplasm	1319	75	0.057	1064	73	0.069
Mitochondrion	458	75	0.164	317	60	0.189
Nucleus	346	5	0.014	51	6	0.118
Extracellular	184	2	0.011	177	2	0.011
Endoplasmic reticu	170	54	0.318	186	35	0.188
Cell envelope	155	6	0.039	99	6	0.061
Peroxisome	110	1	0.009	103	3	0.029
Mitochondrial mem	95	8	0.084	4	4	1.000
Lipid particle	76	1	0.013	209	2	0.010
Vacuolar membran	50	10	0.200	5	5	1.000
Vacuole	26	0	0.000	34	1	0.029
Golgi membrane	20	3	0.150	2	2	1.000
Golgi	20	12	0.600	7	6	0.857
Endoplasmic reticu	7	5	0.714	3	3	1.000
Nuclear membrane	7	0	0.000	0	0	NA
Peroxisomal memb	6	0	0.000	0	0	NA
Total	3049	257	0.084			

MetaboliteName	Compartment	Cluster Size	#Blocked Cum	Reaction source	Comments	Rescue reactions
cis-decaprenyl diphosphate (n-C10:2)	Endoplasmic reticulum	27	27	putative transport	Transporter reaction	Transport from cytoplasm_to_Endoplasmic reticulum
3-Hexaprenyl-4-hydroxybenzoate	Mitochondrion	19	18	KEGG	Cluster restored except for s_186	[R05885: 3-Hexaprenyl-4-hydroxybenzoate + Oxygen <=> 3-Hexaprenyl-4,5-dihydroxybenzoate]
beta-D-mannosyldiacetylchitobiosyldiphosphodolicho	Golgi	12	2	iIN800	Whole cluster restored after recovering 3 blocked metabolites	[Uridine kinase: GTP + uridine => GDP + UMP; Probable mannosyltransferase KTR4: beta-D-mannosyldiacetylchitobiosyldiphosphodolichol + 2 GDP-mannose=> 2 GDP + alpha-D-mannosyl-beta-D-mannosyl-diacetylchitobiosyldiphosphodolichol; Mannose-1-phosphate guan
1-phosphatidyl-1D-myo-inositola	Golgi	12	10	putative transport		Cytoplasm_to_Golgi
1-phosphatidyl-1D-myo-inositol 4-phosphates	Golgi	12	12	putative transport		Cytoplasm_to_Golgi
cardiolipin	Mitochondrion	10	10	KEGG		[R02647: 1,2-diacyl-sn-glycerol acylhydrolase; R00851: acyl-CoA:sn-glycerol-3-phosphate 1-O-acyltransferase; R02029: Phosphatidylglycerophosphate phosphohydrolase; R03755: Acyl-CoA:2-acylglycerol O-acyltransferase; R02757: ATP:acylglycerol 3-phosphotransf
phosphatidylserine	Golgi membrane	3	3	iIN800		[Phosphatidate cytidylyltransferase; phosphatidate + CTP <=> CDP-diacylglycerol + diphosphate; CDP-diacylglycerol--serine O-phosphatidyltransferase:CDP-diacylglycerol + L-serine <=> phosphatidylserine + CMP; Transport of phosphatidate, mitochondrial]
4-Hydroxy-L-threonine	Cytoplasm	3	3	KEGG	YER081W for 1.1.1.290, YJR009C or YGR192C for 1.2.1.72	[R04210: 4-Phospho-D-erythronate + NAD+ <=>2-Oxo-3-hydroxy-4-phosphobutanoate + NADH + H+ ; R01825: D-Erythrose 4-phosphate + NAD+ + H2O <=> 4-Phospho-D-erythronate + NADH + H+]
phosphatidylserine	Vacuolar membrane	3	3	KEGG		[R07376: Phosphatidylethanolamine + L-Serine <=> Phosphatidylserine + Ethanolamine]
Acetyl-ACP	Mitochondrion	3	3	KEGG		[R08157: Octanoyl-[acp] + H2O <=> Acyl-carrier protein + Octanoic acid]
4-Guanidinobutanoate	Cytoplasm	2	2	KEGG	YPL111W for 3.5.3.7 R01990	R00559: L-Arginine + Oxygen <=> 4-Guanidinobutanamide + CO2 + H2
phosphatidylethanolamine	Cytoplasm	2	2	KEGG		R03051: Phosphatidylethanolamine + H2O <=> Ethanolamine + Phosphatidate
2-Deoxy-D-ribose 5-phosphate	Cytoplasm	2	2	KEGG	YDR496C for 5.4.2.7 e2.6	R02749: 2-Deoxy-D-ribose 1-phosphate <=> 2-Deoxy-D-ribose 5-phosphate
1-phosphatidyl-1D-myo-inositol 3,5-bisphosphates	Cytoplasm	2	2	KEGG	R05802 exist in CE VM but not cytosol	R05802: ATP + 1-Phosphatidyl-1D-myo-inositol 3-phosphate <=> ADP + 1-Phosphatidyl-1D-myo-inositol 3,5-bisphosphate
phosphatidylserine	Mitochondrion	2	2	KEGG		R07376: Phosphatidylethanolamine + L-Serine <=> Phosphatidylserine + Ethanolamine
5-Methyltetrahydropteroyltriglutamate	Cytoplasm	2	2	iIN800		[5-methyltetrahydropteroyltriglutamate--homocysteine methyltransferase; O-acetylserine sulfhydrylase]
IDP	Mitochondrion	2	2	iIN800	YHR202W for 3.1.4.16 3.1.3.6	[R00727: ITP + Succinate + CoA <=> IDP + Orthophosphate + Succinyl-CoA]
2',3'-Cyclic AMP	Cytoplasm	2	2	KEGG		[R03537: 2',3'-Cyclic AMP + H2O <=> 3'-AMP; R01562: 3'-AMP + H2O <=> Adenosine + Orthophosphate]
Acetyl-ACP	Cytoplasm	1	1	KEGG		R04536', 'R04429', 'R04537', 'R04355', 'R04954', 'R04957', 'R04958', 'R04533', 'R04724', 'R08157'

Conclusions

- One-step further computational effort over the initial manual curation towards a gapless network reconstruction model.
- We can systematically decrease the inconsistency of the model and potentially improve the accuracy of the model simulation.
- This approach can generate hypotheses (suggesting good candidate reactions) for manual verification or further robot experimental tests.

Acknowledgements

The YEASTNET project
Paul D. Dobson, Kieran Smallbone, Pedro Mendes *et. al.*: Manchester center for integrative Systems Biology, University of Manchester, UK
Pinar Pir, Stephen G. Oliver: Cambridge Systems Biology Center, University of Cambridge, UK
The EU FP7 project of UNICELLSYS.

Further information

Chuan Lu
Dept. of Computer Science
Aberystwyth University
Aberystwyth,
SY23 3DB, Wales, UK
cul@aber. ac.uk

References

[1] Herrgård MJ et al. A consensus yeast metabolic reconstruction obtained from a community approach to systems biology, Nature Biotechnol. 2008, 26, 1155-1160

[2] Nookaew I et al. The genome-scale metabolic model iIN800 of Saccharomyces cerevisiae and its validation: a scaffold to query lipid metabolism. BMC Syst Biol. 2008, 7;2:71.

[3] Reed JL et al. Systems approach to refining genome annotation. Proc Natl Acad Sci U S A, 2004, 103:46, 17480--17484.

[4] Burgard AP et al. Flux coupling analysis of genome-scale metabolic network reconstructions, Genome Res., 2004, 14:2, 301--312.