Aberystwyth University

Constraint-based optimisation tools for semi-automated refinement of genome-scale yeast metabolic models

Pir, Pinar; Dobson, Paul D.; Smallbone, Kieran; Mendes, Pedro; King, Ross Donald; Lu, Chuan; Oliver, Stephen G.; Clare, Amanda

Publication date: 2010

Citation for published version (APA):
Constraint-based Optimisation Tools for Semi-automated Refinement of Genome-scale Yeast Metabolic Models

Chuan Lu1, Pinar Pir1, Kieran Smallbone1, Paul D. Dobson4, Amanda Clare1, Pedro Mendes3, Stephen G. Oliver1, and Ross D. King1

1 Department of Computer Science, Aberystwyth University, UK, 2Cambridge Systems Biology Centre, Department of Biochemistry, University of Cambridge, UK, 3Manchester Centre for Integrative Systems Biology, University of Manchester, UK, 4School of Chemistry, University of Manchester, UK

Introduction

Motivation
- Genome-scale metabolic network models are useful for analysing the cellular behaviour of organisms
- Semi-automated procedure for model validation and refinement are important for quality assurance in such models
- Computational tools for iterative model validation and optimisation are necessary to assist hypothesis generation and evaluation

Genome scale metabolic models for S. cerevisiae
- A consensus reconstruction: Yeast1, community driven, rigorously evidenced, well annotated [1]
- Further development: Yeast4, expanded from Yeast1, with improved representation of metabolic transport, lipid metabolism, etc. [2]
- Further development: Yeast4, expanded from Yeast1, with improved evidence, well annotated [1]
- A consensus reconstruction: Yeast1, community driven, rigorously evidenced, well annotated [1]
- Yeast4: 1102 unique metabolite reactions, and 924 metabolites located in 15 cellular compartments

Methods

Framework of Flux Balance Analysis (FBA)
- Identification of flux distribution using stoichiometry model, assuming steady states, with constraints on mass balance and thermodynamics to maximise/minimise an objective function (e.g. to max growth rate)
- Utilisation of constraint-based optimisation, linear/nonlinear programming (LP/NLP), mixed integer linear programming (MILP)

Gap filling
- Structural Gaps in metabolic networks
- Reactions gaps, missing gene-protein-reaction associations, etc.
- Mechanisms to rescue reaction gaps
- Reversibility, transport, biomasses: formation, metabolism exchange
- Addition of missing reactions from reference model
- Identification of minimal set of reactions to add on, in order to restore biomass formation or blocked reactions [4]

Optimal Metabolic Network Identification (OMNI)
- Models under-constrained:
  - Reactions absent in yeast, irreversible or unfavorable under certain conditions, suppressed due to regulatory, etc.
  - Bi-level constrained optimisation
- Minimisation of discrepancies between observations and predictions while maximising the growth rate [5]
- Converting to MILP by exploiting duality for LP

Applications

Computational tool implementation
- Implemented in Python, using CPLEX, glpk, IpSolve as LP/MILP solvers
- Read/write models in SBML format
- Model stored in bipartite graph and/or stoichiometry matrices
- Suitable for both FBA and logical model simulations
- Converting model network (bi-level) optimisation problem to constraint-based optimisation problem: LP, MILP.
- Algorithms for gap filling, OMNI
- Search algorithms for graph traverse, and identification of minimal models

Model validation with experimental data
- Single deletion data under minimal medium
- Wildtype growing under different conditions: RobotScientist’s [3] automated titration experiments on yeast utilising amino acids as sole C/N source
- Awareness of data quality issue

Gap filling procedure
- Constraint-based optimisation and literature searching
- Solutions for 14 of 16 false inviable single deletions under minimal medium
- Further curation needed to fill in the missing reaction esp. for alternative pathways of ergosterol biosynthesis

OMNI procedure
- More than 1 solution for 12 out of 48 false viable cases subject to OMNI
- Solution evaluation: in-silico simulation using phenotype data in SGD
- Application of a minimal set of revisions, resulting in:
  - True inviables increased by 12, at the cost of 1 extra false inviable
- Suggested revisions:
  - Constraining the reaction directionality
  - Removing reactions:
    - e.g. alternative pathway for quinoline synthesis absent in yeast
  - Adding regulator rules to control reaction activation
    - e.g. GALT and GALE activated only after sensing glutalate
  - Testing in vivo by robot: auxotrophy experiments

Conclusions

- Proposed computational tools can effectively search for (multiple) revision suggestions for yeast metabolic models
- Semi-automated model refinement, supported with literature search and robot scientist experiments, helps to improve the model in phenotype perception
- Future work
  - Use of logic programming to integrate models with evidence from experimental data and constraint-based analysis
  - Learning GPR associations and regulatory rules and automated suggestion of experiments, either in-silico or in-vivo

Acknowledgement
The EU FP7 project of UNICELLSYS

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

References