BIOINFORMATICS APPLICATIONS NOTE Vol. 30 no. 4 2014, pages 584–585
doi:10.1093/bioinformatics/btt723

Systems biology

Advance Access publication December 13, 2013

ORCA: a COBRA toolbox extension for model-driven discovery and analysis

Longfei Mao* and Wynand S. Verwoerd

Department of Molecular Biosciences, Centre for Advanced Computational Solutions, Lincoln University, Lincoln 7647, New Zealand

Associate Editor: Martin Bishop

ABSTRACT

Summary: Over past decades, constraint-based modelling has emerged as an important approach to obtain referential information about mechanisms behind biological phenotypes and identify physiological and perturbed metabolic states at genome-scale. However, application of this novel approach to systems biology in biotechnology is still hindered by the functionalities of the existing modelling software. To augment the usability of the constraint-based approach for various use scenarios, we present ORCA, a Matlab package, which extends the scope of established Constraint-Based Reconstruction and Analysis metabolic modelling and includes three unique functionalities: (i) a framework method integrating three analyses of multi-objective optimization, robustness analysis and fractional benefit analysis, (ii) metabolic pathways identification with futile loop elimination and (iii) a dynamic flux balance analysis framework incorporating kinetic constraints.

Availability and implementation: ORCA is freely available to academic users and is downloadable from https://sourceforge.net/projects/exorca/; a mini-tutorial is supplied in the package for training purposes as well as a software manual.

Contact: Longfei.mao@lincolnuni.ac.nz

Supplementary information: Supplementary data are available at *Bioinformatics* online.

Received on October 3, 2013; revised and accepted on December 9, 2013

1 INTRODUCTION

The recent advance in high-throughput technology has vastly increased the availability of the biological information, for reconstruction of genome-scale metabolic networks (GSMs). These can serve as a detailed representation of biological reaction networks and their functional states. For investigating GSMs in both academic and industrial settings, the Constraint-Based Reconstruction and Analysis (COBRA) software (Schellenberger *et al.*, 2011) has been established as a popular tool since 2007. Nevertheless, the systematic application of GSMs for model-driven discovery has much scope to be further developed. For example, although the original COBRA toolbox does include a variety of methods for modelling GSMs, it still lacks ready-made functions for users to model the competition between multi-objectives, commonly encountered in bioengineering. This indicates that there is an urgent need to develop novel simulation methodologies and tools, to interrogate and interpret the information in the GSMs in integrative, systemic and meaningful ways. To facilitate this emerging direction in research, we have developed a number of MATLAB functions, collectively packaged and named as ORCA (mOdel-dRiven disCovery and Analysis) toolbox, supplementary to the existing COBRA functions. These ORCA functions are intended to take advantage of existing COBRA models, for us to (i) obtain information on the capability of an interesting microorganism for desired metabolite production, (ii) understand the correlation between the perturbation of the surrounding metabolism, and the preferable metabolite synthesis, and (iii) investigate how dynamic changes of external metabolite concentrations in the environment affect the yield of desirable metabolites in a reactor. The tool has been used to elucidate the inherent potential of several promising biocatalysts for microbial fuel cell electricity production and some of the base mathematics of ORCA have been discussed in the literature (Mao and Verwoerd, 2013a, b, c).

*To whom correspondence should be addressed.

2 FEATURES

ORCA conducts flux balance analysis (FBA)-based methods with multi-objective formulation and extends the robustness analysis of COBRA into a framework context (Fig. 1). It incorporates a simple algorithm excluding the futile fluxes encountered during FBA and allows identification of the pathways pertaining to the metabolite of interest. ORCA also includes an updated version of the 'dynamicFBA (dFBA)' of COBRA. The incorporation of kinetic properties of the environment nutrient uptake was proposed in the literature (Mahadevan *et al.*, 2002). In our updated version, limitations on biomass growth resulting from, for example, nutrient depletion can be incorporated by rate-of-change constraints and substrate uptake rates modelled by Michaelis–Menten or Hill-equation kinetics. The new 'dFBA' function allows for unlimited number of substrates and products to be taken into account by automatically forming corresponding ordinary differential equations (ODEs). A listing of individual functions and a full discussion of how ORCA complements COBRA are provided in the Supplementary file.

2.1 Reconciliation functions

Because several modifications are to be accommodated in the COBRA model before the ORCA modelling, we have included a number of reconciliation functions to help set up modelling environments, such as the formulation of multi-objectives and

setting proper uptake rates of external metabolites for different cultivation conditions. In addition, ORCA includes several functions that compare models between COBRA and Optflux, another highly recognized FBA modelling software (Lakshmanan *et al.*, 2012). This helps to resolve a different tolerance in accepting various SBML formats, as we encountered in the use of Optflux and COBRA.

2.2 Core function 1: optimizeM

This function examines the production potential of a GSM for a preferable metabolite. The user can specify any desirable intracellular metabolites included in the GSM. This function automates implementation of three analyses: (i) pareto optimality analysis, (ii) bivariate robustness analysis and (iii) fraction benefit analysis (Supplementary file for details of these analyses).

2.3 Core function 2: fatmin

This function iteratively implements our previously published algorithm Flux variability analysis with target flux minimization (FATMIN) (Mao and Verwoerd, 2013c). FATMIN is devised to eliminate futile loops in a metabolic network and characterize all the alternate optimal solutions or equivalent phenotypic states in a metabolism. In particular, FATMIN results characterize the network mechanisms underlying the robustness due to alternative metabolic pathways. If a series of flux levels for synthesis of desirable molecules spanning from the lowest (control state) to highest are used in running FATMIN, the function can scan for potential metabolic switches in the metabolism. These metabolic switches represent the transitions between phenotypic behaviours and can provide indirect information about a system's tolerance to different levels of perturbation. The distinct phenotypic behaviours also indicates the suitability of the microorganism for desirable chemical production. Furthermore, the resultant lists of reactions from FATMIN can be visualized in a context of pathways and subnetworks by calling a recently published COBRA extension Paint4Net (Kostromins and Stalidzans, 2012).

2.4 Core function 3: dFBA

Different from the steady state assumption that numerous properties are unchanging in time, dynamic FBA can examine time-dependent properties of a model by coupling a model to the bioreactor process in order to reveal the dynamic interactions between the cell and the environment (Mahadevan *et al.*, 2002). In ORCA, dFBA consists of a suite of functions that provide a basic dynamic flux balance modelling framework, combining a steady-state intracellular model of the metabolism with a list of selected component-linked ODEs and Michaelis–Menten or Hill function uptake kinetics. In practice, different simulation conditions demand change or addition of dynamic equations to reflect different biomass and external metabolites in the function script, as detailed in the Supplementary file.

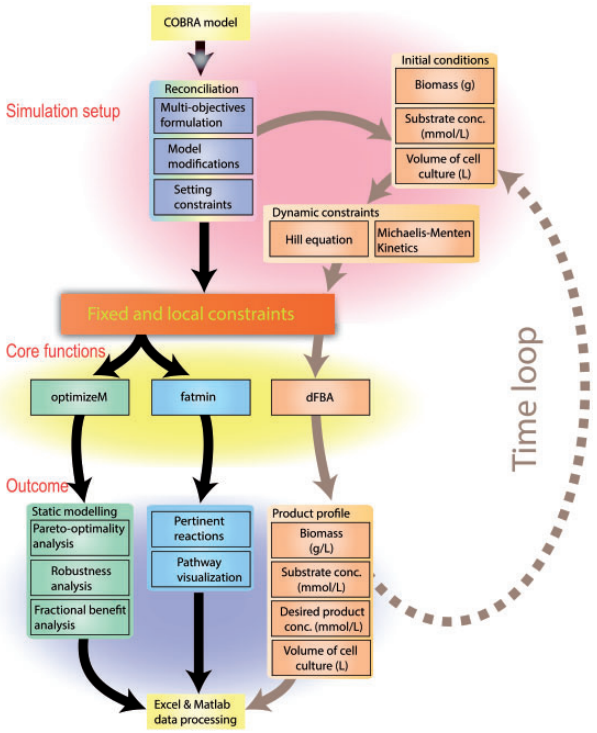

Fig. 1. The application structure and core functions of ORCA

Conflict of Interest: none declared.

REFERENCES

Kostromins,A. and Stalidzans,E. (2012) Paint4Net: COBRA toolbox extension for visualization of stoichiometric models of metabolism. *Biosystems*, **109**, 233–239.

Lakshmanan,M. *et al.* (2012) Software applications for flux balance analysis. *Brief. Bioinform.*, [Epub ahead of print.].

Mahadevan,R. *et al.* (2002) Dynamic flux balance analysis of diauxic growth in *Escherichia coli*. *Biophys. J.*, **83**, 1331–1340.

Mao,L. and Verwoerd,W.S. (2013a) Exploration and comparison of inborn microbial capacity of aerobic and anaerobic metabolisms of *Saccharomyces cerevisiae* for current production. *Bioengineered*, **4**, 420–430.

Mao,L. and Verwoerd,W.S. (2013b) Genome-scale stoichiometry analysis to elucidate the innate capability of the cyanobacterium *Synechocystis* for electricity generation. *J. Ind. Microbiol. Biotechnol.*, **40**, 1161–1180.

Mao,L. and Verwoerd,W.S. (2013c) Model-driven elucidation of the inherent capacity of *Geobacter sulfurreducens* for electricity generation. *J. Biol. Eng.*, **7**, 14.

Schellenberger,J. *et al.* (2011) Quantitative prediction of cellular metabolism with constraint-based models: the COBRA toolbox v2.0. *Nat. Protoc.*, **6**, 1290–1307.