Biochemical reconstructions, assembled with rBioNet, can be high-quality bio
toolbox (Becker relying on many QC/QA measures already present in the COBRA
genome-scale biochemical networks (Thiele and Palsson, 2010). (QC/QA), which are crucial for the construction of high-quality
implements important quality control and assurance measures
reconstruction effort more efficient and traceable.

quality control modules. A quality check-as-you-go would make the
reconstruction content, which have the disadvantage of missing
for the reconstruction process to assemble and monitor the
et al.
© The Author 2011. Published by Oxford University Press. All rights reserved. For Permissions, please email: journals.permissions@oup.com
2009
∗
list (Thiele and Palsson, 2010). The reconstruction process is well
information associated in addition to an organism-specific reaction
target organism (Palsson, 2006). A reconstruction often has detailed
gene information, literature references and reaction/metabolite
reactions, which may be expanded prior to or during any reconstruction
effort.

Motivation: Genome-scale metabolic networks are widely used in
molecular systems biology as they provide a structured format
for genomic, genetic and biochemical information available for a
target organism (Palsson, 2006). A reconstruction often has detailed
gene information, literature references and reaction/metabolite
information associated in addition to an organism-specific reaction
list (Thiele and Palsson, 2010). The reconstruction process is well
established for metabolic networks and has been applied to >45
organisms on a genome scale. While this approach is widely
used, no comprehensive, easy-to-use and free-of-charge software
package exists that ensures quality of the mostly manually curated
reconstruction content. Existing tools for metabolic reconstructions
include MetAnnoGen (Gille et al., 2007) and MEMOSys (Pabinger
et al., 2011). However, many researchers still use spreadsheets
for the reconstruction process to assemble and monitor the
reconstruction content, which have the disadvantage of missing
quality control modules. A quality check-as-you-go would make the
reconstruction effort more efficient and traceable.

Here, we present a freely available software package that
implements important quality control and assurance measures
(QC/QA), which are crucial for the construction of high-quality
genome-scale biochemical networks (Thiele and Palsson, 2010).

rBioNet is embedded in the Matlab programming environment
relying on many QC/QA measures already present in the COBRA
toolbox (Becker et al., 2007; Schellenberger et al., 2011). Biocchemical reconstructions, assembled with rBioNet, can be
readily converted into mathematical models, and they can then be interrogated using constraint-based methods (Palsson, 2006).

2 METHODS

rBioNet consists of three parts: (i) a metabolite creator with associated metabolite database (MetDB); (ii) a reaction creator with reaction database
(RxnDB); and (iii) a reconstruction creator (Fig. 1). To ensure consistency and
quality, reactions can only contain metabolites present in MetDB and they
can only be added to a reconstruction when they are present in RxnDB.

MetDB and RxnDB are stored in metab.mat and rxn.mat, respectively.
The provided files within this package contain only sample metabolites and
reactions, which may be expanded prior to or during any reconstruction
effort.

Metabolite creator: there are three ways to create metabolites: (i) manual
addition of metabolite information; (ii) loading of a tab-delimited file
containing all necessary information; or (iii) loading metabolite information
from other COBRA reconstructions (model structure). Required and optional
information is described in the manual. When saving a new metabolite,
entered information will be checked against MetDB for potential duplicate
entries. This check will be done based on metabolite abbreviation and
charged formulae. No entry can have the same metabolite abbreviation but
duplicated formulae may exist (e.g. for isoforms). Metabolites within MetDB
are organism- and compartment independent. A new metabolite can also
be created by opening, modifying and saving an existing entry as a new
metabolite (under a new abbreviation).

Reaction creator: there are three ways of populating the RxnDB: (i)
manually entering corresponding information; (ii) loading a tab-delimited
file; or (iii) loading another COBRA reconstruction (model structure). Also,
a new reaction can be added to the database by copying and modifying
an existing entry. Note that in all cases the reactions can only contain metabolic entities that exist in the MetaDB. Reactions are organism independent but compartment specific. When adding a new reaction to the MetaDB, the reaction will be checked for mass and charge balance. An unbalanced reaction will provoke a warning and a detailed report of unbalanced elements/charge. However, the user may choose to proceed as certain reactions are unbalanced even in high-quality reconstructions (e.g. biomass reaction, exchange reactions) (Thiele and Palsson, 2010). Subsequently, the new reaction is checked for uniqueness. This step may require some time depending on the size of the MetaDB. The check includes uniqueness of reaction abbreviations and of reaction formulae. No duplicated reactions with different abbreviations are permitted. However, a reaction may re-occur in a different cellular compartment. New compartments can be readily created.

**Reconstruction creator:** a user can choose to create a reconstruction from scratch or to load an existing reconstruction for expansion and/or modification. Again, only reactions that are present in the MetaDB can be added to the reconstruction. All reaction information present in the MetaDB will be associated with the model when a reaction is added to the working model. The user can edit and add further information (e.g. references, notes, subsystems) to a reconstruction reaction. Furthermore, global information, such as reconstructor, reconstructed organism, version, and source of gene index, can be added to the reconstruction.

**GPR creator:** each reconstruction can be associated with genes via gene–protein–reaction–associations (GPRs). Therefore, a gene index needs to be imported into the reconstruction creator, a tab-delimited file describing key properties of the organism’s gene annotations. A GPR is created using Boolean logic, ‘AND’, or ‘OR’.

**Reconstruction analyzer:** basic information about the reconstruction is provided and end-dead metabolites are highlighted. This analyzer provides real-time feedback on reconstruction progress and potential missing reactions. A detailed description can be found in the manual.

**Export a reconstruction:** a reconstruction can be saved in Matlab format as a model structure. All associated information will be stored. This model structure can be used for further analysis within the COBRA toolbox or exported to other formats, including spreadsheet and Systems Biology Markup Language (SBML) using the COBRA toolbox (writeSBmodel).

**Import a reconstruction:** a reconstruction can be imported into rBioNet using the reconstruction creator from Matlab format (model structure), spreadsheet or SBML (via COBRA toolbox). If a reconstruction had been constructed with BioNet, all associated information will be loaded when loading an existing reconstruction into the reconstruction creator, the user is asked to provide a file location of the gene index. This file is optional but if provided, information is accessible through the reconstruction creator.

**Reconstruction versus model:** a reconstruction provides a detailed list of biochemical functions encoded by an organism’s genome and serves as a basis for condition-specific models. rBioNet is an environment to assemble a reconstruction. The constraints on the exported model structure have to be adjusted as desired using the COBRA toolbox to obtain a condition-specific model. We do not consider reaction directionality, which are contained in the exported structure, as condition-specific constraints for this purpose.

### 3 IMPLEMENTATION

rBioNet was encoded in Matlab programming environment (MathWorks, Inc.) requiring the COBRA toolbox, version 1.3 or higher (Becker et al., 2007; Schellenberger et al., 2011), to be installed. A user-friendly interface facilitates the use of this tool by novices in programming and/or metabolic network reconstruction. A comprehensive manual is provided including installation information and frequently asked questions.

### 4 DISCUSSION

rBioNet is freely available permitting the reconstruction of high-quality, genome-scale biochemical networks consistent with established standard procedures (Thiele and Palsson, 2010). In particular, it ensures QC/QA measures required for publication-level reconstructions. rBioNet is embedded within the COBRA toolbox enabling iterative approach of reconstruction, validation and debugging. In particular, we put emphasis on an intuitive interface that allows novices to Matlab to perform high-quality metabolic reconstructions. The database allows customized inclusion of metabolites and reactions in the corresponding databases but thanks to its import function, online databases, e.g. KEGG ligand database (Kanehisa et al., 2010), may be incorporated into MetaDB and RxnDB.

A current shortcoming of rBioNet includes missing database binding for the metabolite, reaction and reconstruction creator. Nevertheless, the current setup enables simultaneous use and modification by multiple users. Future extensions may include the availability of a globally available and accessible database for metabolites and reactions, which could be directly used by any rBioNet user and thus, would permit ultimately compatibility between biochemical models. The creation of such database would require significant work and would need to be coordinated with the existing biochemical databases and efforts.

### ACKNOWLEDGEMENT

The authors would like to thank M.K. Aurch, R.M.T. Fleming, A. Heinken, J. Nogales, O. Rolfsson and S. Sahoo for testing and valuable discussion.

**Funding:** Department of Energy, Offices of Advanced Scientific Computing Research and the Biological and Environmental Research as part of the Scientific Discovery through Advanced Computing program (grant DE-SC0002009); European Research Council (grant proposal no. 232816); Rannis Nyskopunarsjodur namsonama.

**Conflict of Interest:** none declared.

### REFERENCES


2010