Improving metabolite annotation through a new generation of parsing tools for genome-scale modeling

Dimitra Lappa[1], Jens B. Nielsen

Division of Synthetic Systems Biology, Facility for Computational Systems Biology, Department of Biology and Biological Engineering, Chalmers University of Technology

[1] Contact email: lappa@chalmers.se

Introduction

The Human Metabolic Reaction (HMR) [1] is a genome-scale metabolic model (GEM) [2] that stands as a scaffold for integration of omics data and flux metabolic modeling of a generic human cell. HMR contains 3765 gene associations, over 8000 reactions and more than 3000 unique metabolites. A current challenge in GEM reconstruction is metabolite annotation, particularly for metabolites in cyclic reactions of lipid metabolism (e.g. fatty acid elongation), which usually lack unique identifiers so as to catalog into a database format.

Methods

Here we describe a new strategy to correctly annotate these non-unique metabolites in the GEM modeling framework and specifically in the HMR model-derived database. We present a new generation of parsing tools in the form of software plug-ins for a series of GEM file formats. Through parsing constraint-based model files we aim to unify nomenclature and introduce annotation for non-unique metabolites involved in HMR. The parser's input is a model in a popular format (xls, sbml, COBRA-sbml, RAVEN-sbml). The parser facilitates the conversion between formats and also breaks down the input files into model components.

Results

The objective is to populate a relational database to enable the querying of the model's different parts. The database contains unique metabolites, exchange reactions, reaction directionality, subcellular compartments and gene associations [3].

The parser will cope with nomenclature inconsistencies, converting annotation if necessary. Importantly, it will detect fatty acids cyclic reactions and attempt to label them accordingly. Complementary checks will subsequently be carried to avoid structural inconsistencies and ensure the model's integrity.

Discussion

This approach is expected to help towards model standardization, while addressing the challenge of cyclic reactions in modeling the cellular metabolism and promoting the comparison and evaluation among different models.

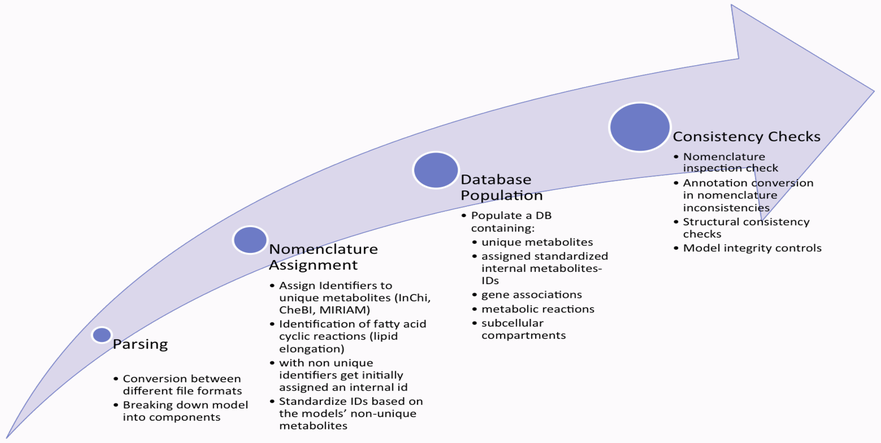
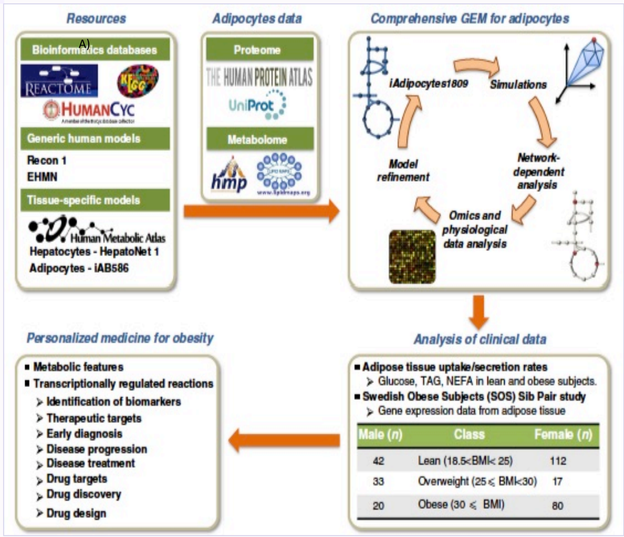


Fig1: Algorithmic overview of the parsing and nomenclature standardization process.

2.A)



2.C)

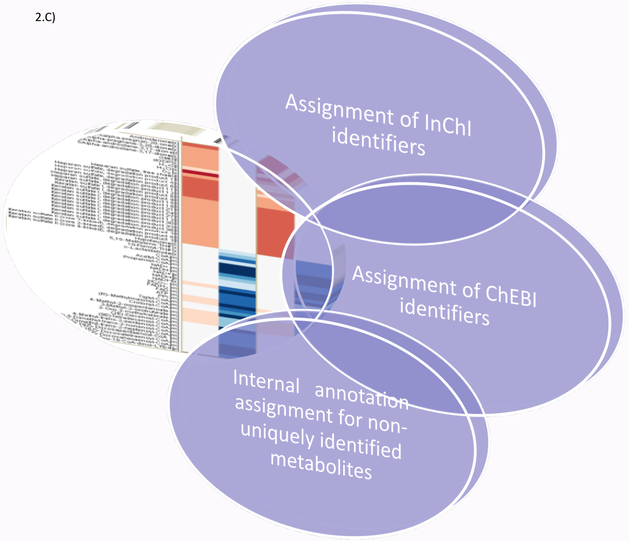


Fig3: : Algorithmic overview of the model's sanity check for structural and annotation inconsistencies. A series of inspections is performed, with respect to the sbml format and the xml output containers, the stoichiometry of the model, the gene references and association mapping, as well as the removal of non essential elements such as unused metabolites and duplicate reactions.

2.B)

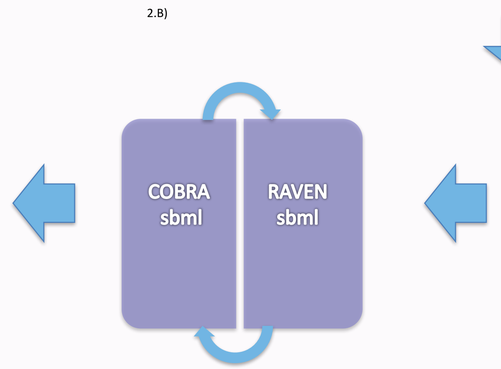
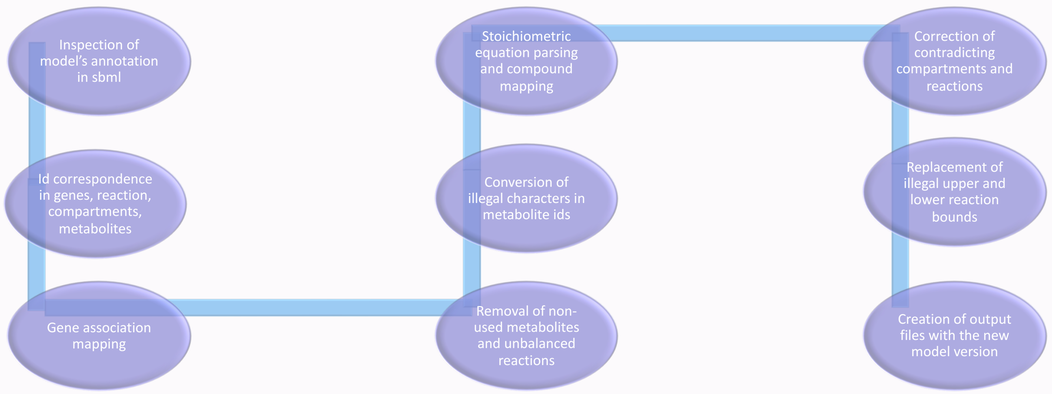


Fig2: A) Main components and resources used for drafting the prototype Genome-Scale Metabolic model of Human Metabolic Reaction, from tissue specific models and integrative multi-omics analysis to personalized medicine for metabolic conditions. [1]

B) The parser performs conversion among popular Genome-Scale Metabolic model formats, especially among different sbml Levels and releases. There are notable differences with respect to xml annotation containers between COBRA and RAVEN software for Genome-Scale modeling draft and reconstruction, that are addressed in this step of the process

C)The parser's complementary software creates a pool of metabolites and classifies them into unique or pseudo-metabolites (those participating cyclic reactions of lipid metabolism). External databse dentifiers are assigned accordingly.



References:

[1]Mardinoglu A., Agren R., Asplund A., Uhlen M., Nielsen J.B.," Genome-scale metabolic modeling of hepatocytes reveals serine deficiency in patients with non-alcoholic fatty liver disease", Nature Communications 5, 3083, 2014

[2] Lewis, N.E. et al., "Constraining the metabolic genotype–phenotype relationship using a phylogeny of in silico methods", Nature Reviews 10, 291-305, 2012

[3] Pornputtapong N., Nookaew I., Nielsen J.B., "Human Metabolic Atlas: an online resource for human metabolism", Database: The Journal of Biological Databases and Curation, In Press