Session 2
6th May
15:15 h

Revising lipid chemical structures in genome-wide metabolic models with BOIMMG

Capela, J.[1], Liu, F.[2], Dias, O.[1]

[1] Centre of Biological Engineering, University of Minho, Campus de Gualtar, 4710–057 Braga, Portugal
[2] Mathematics and Computer Science Division, Argonne National Laboratory, Argonne, USA

An important step in the reconstruction of Genome-Scale Metabolic (GSM) models is the integration of biochemical data. Such information is often incomplete or generic, lacking in completely defined chemical structures for several molecules, including lipids. The inumerous combinations of fatty acids in the side chains of lipids, hinder their storage in databases and integration into GSM models. Generic representations are commonly used to circumvent such limitation. However, lipid specificity is likely lost, and data integration problems arise, as several models contain lipids with completely defined structures and others with their generic versions. Such clash of versions is addressed by the Biochemical cOmplex data Integration in Metabolic Models at Genome-scale (BOIMMG). BOIMMG is an open-source framework that accelerates the swapping of different molecular versions (mainly lipids, structurally defined or not) in GSM models. Upon integration into a Neo4j graph database (http://neo4j.com/), lipid-specific data from LIPID MAPS Structure Database (LMSD), Swiss Lipids (SLM) and Model SEED were processed for biosynthetic contextualization within the curated pathways of MetaCyc. Several algorithms were developed to integrate this information in GSM models, afterwards.

Over 30 generic reactions were fully and 27 partially expanded, resulting in 557392 new reactions, in which 557252 were not integrated, nor listed in Model SEED. These reactions were inferred from the previously contextualized biosynthetic relationships between structurally defined compounds.

BOIMMG's information was applied to GSM models, tackling the conflict of molecules' versions. The whole glycerolipids and phospholipids' metabolic network within *E. coli* iJR904 model was expanded by our approach. The comparison between the altered model and one of its manually-expanded published iterations (iAF1260b), has shown that 53 and 38 more matching lipids and reactions, respectively, were found. Besides the new biochemical set, *BioISO*'s analysis demonstrated that biomass lipids were correctly produced, corroborating the correct expansion of the whole biosynthetic network. In conclusion, BOIMMG (available at https://boimmg.bio.di.uminho.pt/) can establish relevant relationships between complex macromolecules, within their biosynthetic context, and provide automated procedures for their integration into GSM models.