

A novel approach to find the missing links in genome-scale metabolic models: The BridgeIt method

Marianne Seijo*, Keng Cher Soh, Vassily Hatzimanikatis

Laboratory of Computational Systems Biotechnology (LCSB), EPFL, Lausanne 1015 Switzerland

* Presenting author: marianne.seijo@epfl.ch

Genome-scale metabolic reconstructions (GSMRs) are valuable resources in the analysis and understanding of cellular metabolism. They are based on genome sequence and annotation, and they are used to develop bottom-up mathematical models of metabolic networks. These models are used in a wide variety of studies ranging from metabolic engineering to evolutionary studies. However, there are incomplete pathways and orphan metabolites in all GSMRs, even for the most well studied organisms. These knowledge gaps are due to the lack of experimental or homologous information, as current methods rely on a database of known reactions to generate possible pathways for bridging these gaps, and they fall short when there is no sequence homology.

We present a novel computational framework called BridgeIT that is able to generate hypothetical reactions and pathways that bridge gaps in reconstructed pathways. The novel reactions generated are based on the third level of enzyme commission classification system (EC), which is consistent with known biochemical reactions, protein structures, genomic sequences, and enzyme properties that follow the EC classification. Within the BridgeIT framework, we generate all biochemically plausible reactions and pathways, which can link two or more metabolites. These pathways are then ranked according to their length, thermodynamic feasibility, and network feasibility. We next use chemical similarity metrics to link the generated hypothetical reactions with known reactions through their substrate and product similarity. The protein and gene sequences of the linked known reactions are used to identify possible sequences within the GSMR to further refine and improve the annotation of the existing GSMR. We demonstrate the ability of this method to identify gaps that can be easily filled by known reactions and also gaps that require novel reactions which existing methods fail to do so.

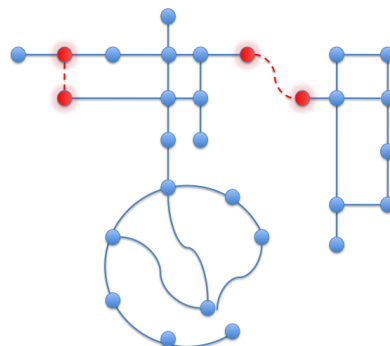


Figure 1. BridgeIT is able to generate hypothetical reactions and pathways that bridge gaps in reconstructed pathways with genome information.