439247 Assessing the Reliability of Genome-Scale Metabolic Models in the Prediction of Gene Knockout Phenotypes

Friday, September 18, 2015
Crowne Plaza Heidelberg City Centre

Helder Lopes, Centre of Biological Engineering, Deparment of Biological Engineering, University of Minho, Braga, Portugal, Paulo Maia, Silicolife Lda, Braga, Portugal, Kiran Patil, Structural and Computational Biology Unit, EMBL, Heidelberg, Germany and Isabel Rocha, Centre of Biological Engineering, Department of Biological Engineering, University of Minho, Braga, Portugal

Whole-genome sequencing paved the way to the reconstruction of genome-scale metabolic models (GSMMs), by increasing the knowledge of the genetic processes that underlie cellular functions in a target organism. These metabolic networks are built from all known metabolic genes and metabolic reactions and try to correlate the genome with molecular physiology, based on genome annotation, biochemical characterization and available literature on the organism of interest. *Saccharomyces cerevisiae* is one of the most studied microorganisms for which there is a lot of fundamental knowledge and it is one of the organisms for which there is most progression in terms of systems biology, positioning itself as a very important model for eukaryotic cells. The *in silico* representation of the interaction network between all the system components is carried out to predict a wide range of cellular metabolic functions at the genome scale, from the growth capability in different environmental conditions to the effect of gene knockouts, at steady-state and using different simulation methods. GSMMs can be hence used in a wide range of biotechnological applications, from metabolic engineering strategies to drug targeting.

In this study, we intended to assess the accuracy level of these GSSMs when predicting phenotypes based on gene deletions performed *in vivo*, by comparing experimental data retrieved from the Saccharomyces Genome Database (SGD) [1] with *in silico* simulations. The reliability of the reconstructed models is often assessed using only gene essentiality data. In our study, we used the Yeastmine tool to collect information from the SGD, a powerful web-accessible resource that contains functional structured information based on experimental evidence about budding yeast genes, namely phenotype information from single gene knockouts. SGD contains information about over 180 different

types of phenotypes from which nearly 10% of can be predicted using GSMMs, from auxotrophy to changes in compound excretion after a certain genetic perturbation.

We observed that 34% of the total number of genes associated to a certain phenotype on SGD using the iMM904 model [2] and 32% using the Yeast6 model [3], do not inactivate any reaction when the gene is deleted, meaning that no changes will be observed compared to the WT flux distribution if we delete *in silico* that particular gene. Moreover, we have simulated 220 cases of auxotroph strains upon single gene deletions reported in SGD, excluding the cases mentioned before and using 2 different GSMMs (iMM904 and Yeast6) and 2 different simulations methods (pFBA and LMoMA) in 3 different media conditions. The best results do not exceed 30% of accuracy in predicting the *in vivo* phenotype. An identical scenario was observed when assessing other types of phenotypes. Some of the phenotype information present in SGD can be integrated with the existing metabolic models in order to increase their accuracy prediction level, which might include, among other strategies, systematic changes in GPR rules and reversibility of reactions.

[1] Cherry JM, Hong EL, Amundsen C, *et al.* Saccharomyces Genome Database: the genomics resource of budding yeast. Nucleic Acids Res. 2012; 40(Database issue):D700-5.

[2] Mo ML, Palsson BØ, Herrgård MJ. Connecting extracellular metabolomic measurements to intracellular flux states in yeast. BMC Syst Biol 2009;3:37.

[3] Heavner BD, Smallbone K, Price ND, Walker LP. Version 6 of the consensus yeast metabolic network refines biochemical coverage and improves model performance. Database (Oxford). 2013; 2013: bat059.

Extended Abstract: File Not Uploaded

See more of this Session: Poster Session II
See more of this Group/Topical: Poster Submissions