

Contribution to SBGN contest: best SBGN map - breadth, accuracy, aesthetics

SBGN-PD Map of Central Metabolism in Humans Based on the KEGG Pathways

Falk Schreiber^{1,2}, Tobias Czauderna¹, Michael Wybrow³, Kim Marriott³

¹Leibniz Institute of Plant Genetics and Crop Plant Research (IPK) Gatersleben, Germany

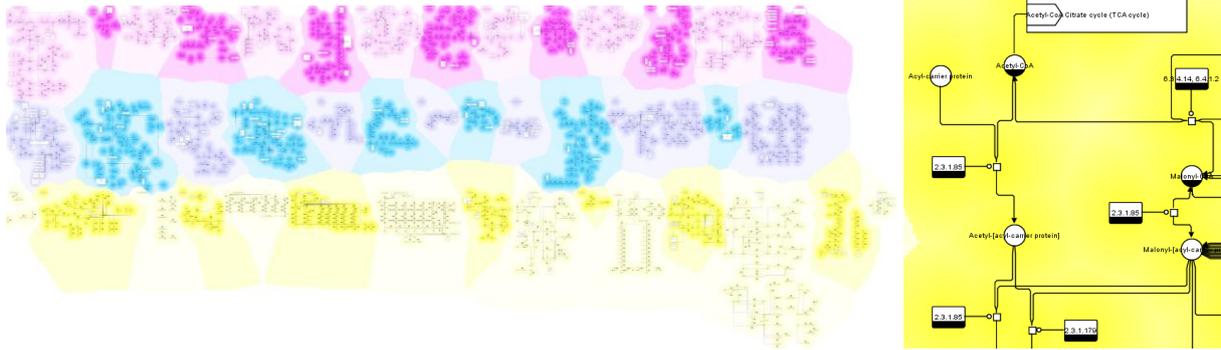
²Institute of Computer Sciences, Martin Luther University of Halle-Wittenberg, Germany

³School of Information Technologies, Monash University Melbourne, Australia

This is the graphical representation of major parts of the central metabolism in humans as a SBGN-PD map. It comprises 42 pathways grouped in three categories. The information has been derived from the KEGG database [1], a repository of metabolic pathways in many organisms. Here we used information restricted to humans. The colors represent different metabolic pathways in three categories: carbohydrate (magenta), amino acid (blue), and lipid (yellow) metabolism. The representation is correct SBGN-PD apart from the annotation nodes representing the name of each pathway.

The map was created in the following way: Maps have been downloaded from KEGG and translated into SBGN-PD notation using SBGN-ED [2]. During the translation a novel constraint layout algorithm (based on the work in [3]) has been employed to obtain images which fulfill the SBGN-PD layout specification and are similar to the KEGG images (Note that the SBGN maps contain a higher number of nodes and edges than the KEGG images). The pathways have been combined and the clusters colored using SBGN-ED / Vanted [4] functionality (SBGN-ED is an add-on of Vanted). Due to some problems in translating the KEGG maps (e.g., errors in the KGML file) some manual corrections of the SBGN map were necessary during this process such as the correction of some edge types.

A high resolution image (png) and a pdf file are attached. As the map has a huge size (~26.000 * 10.000 pixel) a small png image is also attached.



The map and an enlargement of a small part of the map

[1] M. Kanehisa, S. Goto, S. Kawashima and A. Nakaya: The KEGG databases at GenomeNet, *Nucleic Acids Research*, 30 (1): 42-46, 2002.

[2] T. Czauderna, C. Klukas and F. Schreiber: Editing, validating, and translating of SBGN maps. *Bioinformatics*, 26 (18): 2340-2341, 2010.

[3] F. Schreiber, T. Dwyer, K. Marriott and M. Wybrow. A generic algorithm for layout of biological networks. *BMC Bioinformatics* 10: 375, 2009.

[4] B. H. Junker, C. Klukas and F. Schreiber: VANTED: A system for advanced data analysis and visualization in the context of biological networks. *BMC Bioinformatics* 7: 109, 2006.