Bio-inspired metaheuristics for automatic synthesis of novel metabolic pathways

Matias F. Gerard, Georgina Stegmayer and Diego H. Milone

Research Institute for Signals, Systems and Computational Intelligence, sinc(i), FICH-UNL/CONICET, Argentina.

Background

Metabolic pathways synthesis is currently a necessary technique to understand and manipulate the metabolism of living beings. Different approaches, mainly based on classical search algorithms, have been proposed to find linear sequences of reactions (pathways) linking two compounds. However, such models do not take into account the availability of substrates for each reaction, frequently leading to solutions that lack biological feasibility. Moreover, current available methods have the problem of the exponential increase of search trees when the search space involves a large number of compounds and reactions. Thus, the synthesis of novel metabolic pathways is an especially challenging task due to the high number of potential solutions to explore and the feasibility restrictions that must be taken into account.

Results

We present a novel metaheuristic-based tool[1] for efficient search of metabolic pathways, able to make intelligent exploration on large spaces of solutions. This is inspired on the real ants behavior for searching food. Starting from a set of available compounds, the software ants explore the space of solutions by building only feasible metabolic pathways. On each step they select one feasible reaction and expand the set of available compounds with its products, for which new reactions can be carried out. When an ant finds a solution, it leaves pheromones over the sequence of reactions followed according to the pathway size and the proportion of the specified compounds that were linked. Thus, given a base of reactions over which to seek a solution, the initial search conditions (compounds freely available) and the set of compounds to relate, the algorithm search automatically a feasible metabolic pathway among the specified compounds. Figure 1 shows an outline of the process. As it can be seen, just basic information is provided to the algorithm, in order to perform the search and return a metabolic pathway linking the specified compounds. It has been tested on several real problems, searching for well known pathways for validation. Standard solutions were found, together with novel pathways linking the same compounds. In all cases, solutions were built only with feasible reactions. Figure 2 presents an example of a metabolic pathway linking 4 specific compounds, searched over 589 possible reactions.

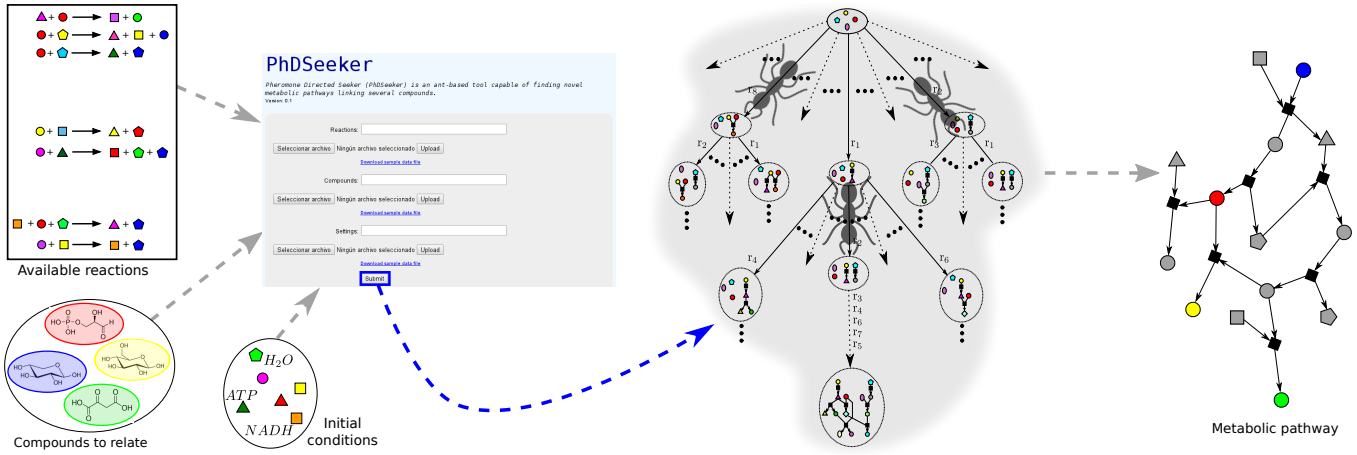

Figure 1. Outline of the steps to find a metabolic pathway linking a set of predefined compounds with PhDSeeker.

Conclusions

In this work we presented a novel metaheuristic-based tool for the automatic search of metabolic pathways. The model allows to find, not only linear feasible metabolic pathways, but also branched ones. Tests performed using real cases for validation showed that the algorithm can recall the standard metabolic pathways, together with other alternatives of feasible solutions. The web interface provided for this tool can facilitate searching for metabolic pathways under a wide range of initial conditions, and can be an interesting option for the study and design of novel pathways.

Supported by: CONICET, MinCyT, UNL.

[1] `http://fich.unl.edu.ar/sinc/web-demo/phdseeker/`.

sinc(i) Research Center for Signals, Systems and Computational Intelligence (fich.unl.edu.ar/sinc)
M. Gerard, G. Stegmayer & D. H. Milone; "Bio-inspired metaheuristics for automatic synthesis of novel metabolic pathways"
4th ISCB-LA Bioinformatics Conference, pp. 1-2, 2016.

sinc(i) Research Center for Signals, Systems and Computational Intelligence (fich.unl.edu.ar/sinc)
M. Gerard, G. Stegmayer & D. H. Milone; "Bio-inspired metaheuristics for automatic synthesis of novel metabolic pathways"
4th ISCB-LA Bioinformatics Conference, pp. 1-2, 2016.

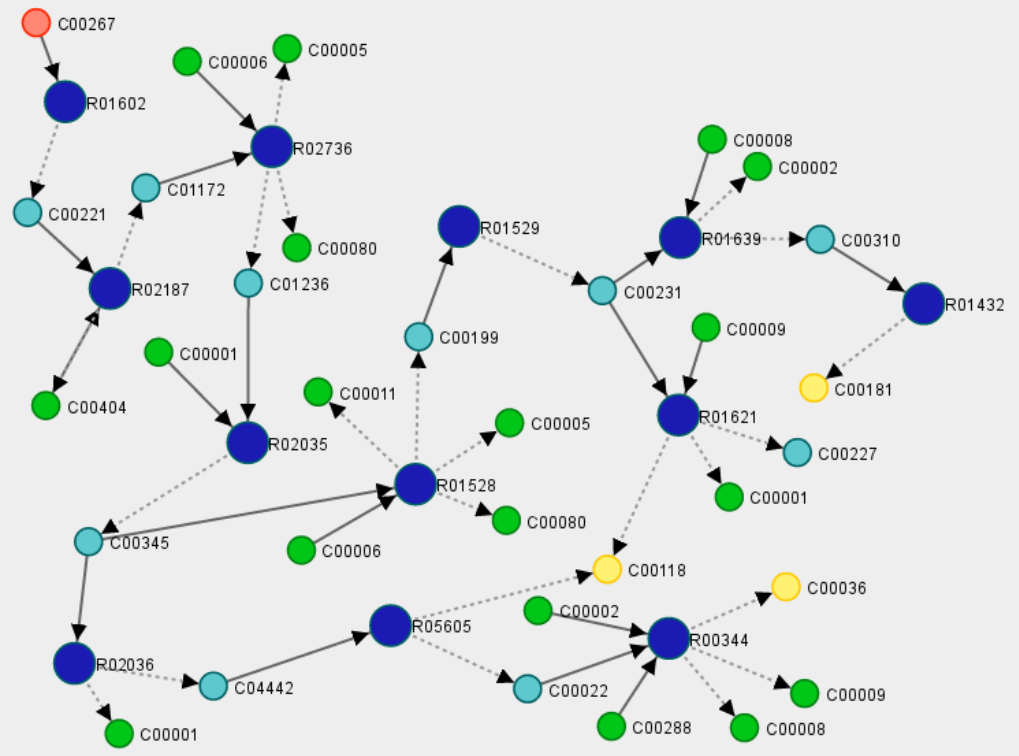

Figure 2. Example of a metabolic pathway found on a search space of 589 reactions linking compounds oxaloacetate (C00036), glyceraldehyde-3P (C00118), D-Xylose (C00181) and α-D-glucose (C00267), standard KEGG notation. Initial substrate in red. Final products in yellow. Freely available compounds in green and intermediate products in light blue. Blue circles indicate reactions.