Analytical Biochemistry 494 (2016) 31–36

Contents lists available at ScienceDirect

Analytical Biochemistry

journal homepage: www.elsevier.com/locate/yabio

Integrated pathway-based and network-based analysis of GC-MS rice metabolomics data under diazinon stress to infer affected biological pathways

Vahideh Mahdavi [a], Faezeh Ghanati [b], Alireza Ghassempour [c, *]

[a] Pesticide Research Department, Iranian Research Institute of Plant Protection (IRIPP), P.O. Box 1475744741, Tehran, Iran
[b] Department of Phytochemistry, Medicinal Plants and Drugs Research Institute, Shahid Beheshti University, G.C. Evin, P.O. Box 19835-389, Tehran, Iran
[c] Department of Plant Biology, Faculty of Biological Science, Tarbiat Modares University, P.O. Box 14115154, Tehran, Iran

ARTICLE INFO

Article history:
Received 19 June 2015
Received in revised form
27 October 2015
Accepted 30 October 2015
Available online 12 November 2015

Keywords:
Diazinon
Rice
Metabolomics
MetaboAnalyst
MetaboNetworks
Reporter reactions

ABSTRACT

Diazinon insecticide is widely applied in rice *(Oryza sativa* L.*)* fields in Iran. However, concerns are now being raised about its potential adverse impacts on rice. In this study, a time-course metabolic change in rice plants was investigated after diazinon treatment using gas chromatography-mass spectrometry (GC-MS) and subsequently three different methods, MetaboAnalyst, MetaboNetwork, and analysis of reporter reactions, as a potential multivariate method were used to find the underlying changes in metabolism with stronger evidence in order to link differentially expressed metabolites to biological pathways. Results clearly showed the similarity of acetylcholinesterase (AChE) of rice plants to that of animals in terms of its inhibitability by diazinon and emphasized that subsequent accumulation of AChE mainly affects the metabolism of osmolites and tricarboxylic acid intermediates subsequent accumulation of ACh mainly affects the metabolism of osmolites and TCA intermediates.

Introduction

Rice *(Oryza sativa)*, a staple food crop, is often subjected to prolonged flooding events, resulting in stunted growth and death [1], but production yield of this crop is limited by several factors, mainly abiotic stresses and biological concerns such as pests. Rice is one of the most important irrigated crops in Iran and its culture involves a large consumption of pesticides. A traditional approach to management of pests is to apply an insecticide [2]. Organophosphorus pesticides (OPPs), such as diazinon, are one of the most frequently encountered groups of pesticides. In recent years, many studies have proven OPPs to be mutagenic, carcinogenic, cytotoxic, genotoxic, teratogenic, and immunotoxic [3]. In addition, OPPs have a tendency to bind to the AChE enzyme and to disrupt nerve functioning, which further results in paralysis and death. Diazinon (O,O-diethyl-O-[6-methyl-2-(1-methylethyl)-4-pyrimidinyl] phosphorothioate) is one of the free important widely used OPPs, especially in rice fields.

As rice is one of the most popular crops worldwide and based on reports of toxic properties and potential risk to human health of diazinon, it seems essential to probe the influence of diazinon on metabolite profiling of rice. Modern research in metabolomics in food science and nutrition is moving from classical methodologies to advanced analytical strategies in which MS-based techniques play a crucial role [4].

In spite of the effective results that can be obtained from conventional strategies of data processing in metabolomics studies, e.g., cluster analysis and principal components analysis, there is still a need for methods that link the processed data to the corresponding biological events in the cell. The most important challenge in inferring affected biological pathways from metabolomics data is that every metabolite participates in more than one pathway; therefore developing a method for connecting metabolite levels with the activity of biological pathways is a complex and difficult task. Existing methods do this task from different aspects. In all methods there are lists of metabolites with significant changes in concentration after a genetic or environmental perturbation, and the question is which biological pathways are affected following the perturbation. Some of the methods consider pathways from traditional biological databases such as KEGG, and try to

* Corresponding author.
E-mail address: a-ghassempour@sbu.ac.ir (A. Ghassempour).

http://dx.doi.org/10.1016/j.ab.2015.10.017
0003-2697/

find pathways with significant changes in activity or pathways that are enriched in given metabolites [5]. Others consider the metabolic network of the organism as a bipartite graph, with metabolites and reactions as nodes, and try to find subgraphs that link the desired lists of metabolites to each other [6]. Additionally, there are methods that consider the metabolic network of the organism as a constraint-based model. Fluxes of the reactions in these models are constrained by applying stoichiometric, thermodynamic, and capacity constraints. Methods in this category try to find reactions in which the metabolites produced or consumed are significantly changed [7] or try to find flux distributions that can explain the behavior of the organism under perturbation more accurately [8].

In a previous paper [9], we developed a chemometric strategy based on Random Forest (RF) applied to GC–MS metabolomics data. In this study the results of the previous study are used with different methodologies against chemometrics tools.

In this paper, we have explored for the first time three methods, MetaboAnalyst, MetaboNetworks, and reporter reactions analysis, as potential multivariate methods for GC-MS analysis of metabolomics data of rice *(O. sativa* L.) plants under diazinon stress. We used three methods to find the underlying changes in metabolism with stronger evidence. Additionally, the selected methods have different points of view and their results can be complementary to each other.

Materials and methods

Seeds of Shiroodi variety rice *(O. sativa* spp. *indica)* were germinated and grown as described in a previous study. Diazinon treatment was started by adding 10% diazinon granular formulation during heading and flowering. A set of control experiments were performed. Rice leaves were taken from control and treated plants. In this study, the results of the GC-MS metabolomics data from previous study [9] were used.

Pathway analysis using MetaboAnalyst

MetaboAnalyst 2.0 is a Web-based tool for data processing, statistical analysis, and pathway analysis of metabolomics data [5]. The input is a table of the normalized peak area of the detected metabolites under different conditions—here concentrations of metabolites in control and treated plants. The data of each day were uploaded separately for the analysis. To reduce systematic bias in the data, normalization was performed with a pooled sample from normal concentrations. Then log transformation was performed to make all concentrations more comparable and normally distributed. After that, for pathway analysis, *O. sativa* was selected as the source organism. A global test was selected for enrichment analysis, once time in combination with a betweenness centrality test and one time in combination with a degree centrality test, which are tests of topological analysis. The result is a list of pathways in the KEGG database with corresponding *p*-values from enrichment analysis and a measure of the impact of each pathway based on betweenness centrality or degree centrality analysis [5].

Analysis of reporter reactions

As suggested by Cakir et al. [10], reporter reactions are the reactions in a metabolic network model whose adjacent metabolites have a significant change in concentration following a perturbation. This is an attempt to infer differential reaction significance based on metabolite measurements, and hence provides a basis for understanding the underlying cellular processes responding to the perturbations. To find reporter reactions in this study, we followed the protocol suggested by Cakir et al. [10] step by step. Briefly, a metabolic network model of rice was used [11], with 1735 reactions and 1484 metabolites. Of the 31 detected metabolites, 17 were included in the model. To enrich the model in detecting metabolites, reduction of the model was performed in four steps: removing compartmentation, merging fully coupled reactions, removing reactions with zero flux, and removing reactions that have no measured metabolites. After the reduction step, a Mann–Whitney test was performed for significance of change in the concentration of metabolites. In this study the concentrations of about 130 metabolites were measured, but only 31 of them could be identified. The *p*-values of unidentified metabolites were assigned randomly to metabolites in the model that do not have measured concentrations. This random assignment was repeated 1000 times. *p*-values were converted to *z*-scores with the inverse normal cumulative distribution function. Then a score for each reaction was calculated based on the *z*-scores of its adjacent metabolites. The mean of the scores in 1000 repetitions was considered as the final score for each reaction. Reactions with a score of 1.64 or higher were considered as reporter reactions. For each reaction, the *z*-scores in each repetition were averaged to get a final *z*-score. Those reactions with the highest *z*-scores ($z > 1.64$, corresponding to $p < 0.10$) can be defined as reporter reactions for metabolome data [10].

Shortest path analysis using MetaboNetworks

MetaboNetworks is a MATLAB-based toolbox that can be used to determine the shortest paths between a set of metabolites. Briefly, it creates a binary matrix with rows and columns of metabolites. If two metabolites can be converted to each other according to information in the KEGG database the corresponding element of the matrix is one, otherwise zero. Then the shortest path between any two metabolites of interest is calculated using the bfs algorithm. We used this toolbox to calculate the shortest paths between any pairs of metabolites with significant change in each day (determined by the Mann–Whitney test). Then we assumed a cutoff of three for the length of the shortest path in order to find only biologically meaningful paths between metabolites. In other words, we considered only the paths between those metabolites that were near enough to each other in a pathway to imply possible biological patterns [1,12].

Integration of the three mentioned methods to find the most probable biological processes affected after treatment with diazinon

After the analysis with MetaboAnalyst, pathways with significant *p*-values from enrichment analysis or high impact from topology analysis were considered as pathways whose activity is changed after treatment. Then, we searched for reporter reactions and metabolites with a distance of three or less to determine, in particular, which part of each pathway had undergone a change in activity. In other words, we considered only those pathways from MetaboAnalyst for which a relatively large section is also detected by at least one of the two other methods.

Results and discussion

According to the analysis of metabolomics data with the MetaboAnalyst Web site, we have found that activities of different biological pathways were changed each day after treatment. These suggested pathways are predefined pathways stored in the KEGG database. Usually, these pathways are too general and include a large number of enzymes, reactions, and metabolites. To study the effect of these pathways in more detail, we integrated the results of the MetaboAnalyst analysis with the results gained from two