Commentary

(Special Topic)

KaPPA-View for integrating quantitative transcriptomic and metabolomic data on plant metabolic pathway maps

Nozomu Sakurai and Daisuke Shibata*

Kazusa DNA Research Institute, 2–6–7 Kazusa-kamatari, Kisarazu, Chiba 292–0818, Japan

With advanced technologies in DNA array, vast amounts of transcriptome data have been produced and state-of-the-art mass spectrometers have contributed to the massive production of metabolome data. Interpretation of the metabolome data in conjunction with the transcriptome data is one of the major concerns associated with research into identifying metabolism-related gene function. The web-based tool KaPPA-View was developed for representing quantitative data for individual transcripts and/or metabolites on plant metabolic pathway maps. The presentation of data in this manner facilitates a good grasp of the transcripts and metabolites, leading to hypotheses of gene function in the metabolic pathway being examined. Here, we comment on a practical use of this omics tool. © Pesticide Science Society of Japan

Keywords: metabolome, metabolomics, transcriptome, transcriptomics, omics, Arabidopsis.

Introduction

Chemistry for plant natural products has identified a number of metabolites, leading to an assumption that over 200,000 metabolites exist in the plant kingdom.[1] Recently, holistic approaches for analyzing transcripts and metabolites have been applied to plant metabolism research. Given the complete genome sequence of the model plant *Arabidopsis thaliana*, Arabidopsis DNA chips have been widely used to analyze gene expression and have consequently produced a vast amount of digitized qualitative data. The state-of-the-art technology of mass spectrometry also leads to massive production of qualitative metabolite data sets. In recent metabolomics approaches, compound separation techniques coupled with mass spectrometry (MS), such as gas chromatography (GC)-MS, liquid chromatography (LC)-MS, and capillary electrophoresis (CE)-MS metabolites[2,3] have been used. The holistic approaches are expected to facilitate findings of new metabolic rules and the development of new biotechnology for improving production of metabolites in diverse technology areas.

Facing the burst of data from such omics approaches, databases and tools have come to be indispensable for researchers who wish to identify gene function in each metabolism (see

review, Ref. 4). As an omics tool, KaPPA-View (http://kpv.kazusa.or.jp/kappa-view/) was developed for displaying quantitative data of transcripts and/or metabolites on metabolic pathway maps.[5] A good grasp of the transcripts and metabolites in this manner facilitates the generation of hypotheses of gene function in plant metabolism. Because a review article of KaPPA-View is already available,[6] here we comment on a practical use of the KaPPA-View tool.

Function of the KaPPA-View Tool and User Setup

The KaPPA-View tool is designed to display comparative quantitative data for individual transcripts and/or metabolites of two experiments on the same set of metabolic pathway maps (Fig. 1). Using quantitative values for transcripts and/or metabolites submitted by the user through the Internet, the KaPPA-View server inserts colored symbols corresponding to a defined metabolic process at that site on the maps and returns them to the user's browser. Enzymatic reactions and transcripts involved in the biosynthesis are depicted with symbols: circles for metabolites, arrows for reactions, and squares for transcripts involved in or putatively assigned to the reaction. The comparative data for individual transcripts and/or metabolites are represented as different colors on the symbols according to the color charts defined. Because in many cases several transcripts may be assigned to a single enzymatic reaction, which is a relatively common characteristic of plants, the average comparative value of the transcripts as-

---

* To whom correspondence should be addressed.
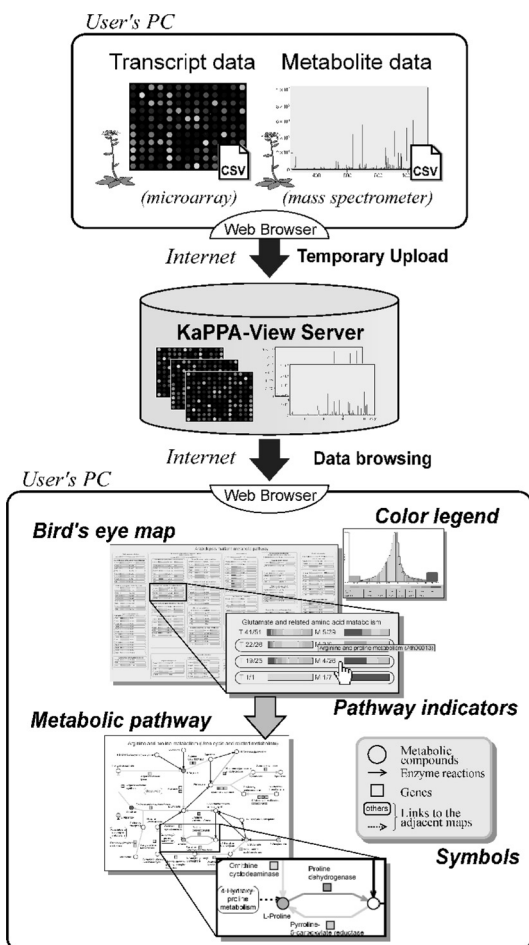E-mail: shibata@kazusa.or.jp, Fax: +81–438–52–3948

Fig. 1. Overall features of the KaPPA-View system. Through the Internet, quantitative data of transcripts (usually derived from microarray experiments) and/or metabolites are uploaded temporarily to the server of KaPPA-View for calculation. The SVG-formatted files are sent back to the user's PC browser. At the initial page, the bird's eye category map composed of category indicators is shown. The color gradation can be changed in the histogram indicator with the user's command. From the bird's eye categories, each metabolic pathway can be retrieved.

signed is represented on the arrow symbol with a defined color.

KaPPA-View uses a set of comprehensive metabolic pathway maps for Arabidopsis, which includes 1263 enzymatic reactions that are classified into seven major metabolic categories with 25 subclasses (see details in Ref. 5). Color pathway indicators for overall changes to transcripts and metabolites in the categories are also represented as category maps or the bird's eye map, allowing users to appreciate the overall picture of the quantitative changes in transcripts and metabolites. Users can access individual maps by clicking on the pathway indicators on the bird's eye map, category maps, or from the tree view (in the left frame of the browser screen) in which the maps are classified in their hierarchical order.

Users are capable of knowing the quantitative values and

other information for each transcript or metabolite listed in an element list or from popup windows that can be retrieved from each map. Using the element list, users can select the metabolite reference page, the gene reference page, or the enzyme reference page for detailed information. As the names of pathways immediately upstream or downstream are indicated on each map, related pathway maps can be displayed in popup windows upon the user's request.

Because the KaPPA-View software is a JAVA application, it can be used on popular operation systems (OS). We confirmed that at least the following computer and browser settings operated KaPPA-View properly: Windows 2000/XP and Internet Explorer 6.0 or Firefox 1.0 or higher; Mac OS X 10.3.7 and Internet Explorer 5.2, Firefox 1.0, or Netscape Navigator 7.1; and Mac OS X 9.2.2 and Netscape Navigator 7.0.2.

Before the first use, the user has to install an Adobe plug-in, Scalable Vector Graphics (SVG) Viewer 3.0 or higher, on his or her own personal computer (PC), because the KaPPA-View returns SVG-formatted files for map viewing to the user's PC. From the top page of the KaPPA-View website, users can find the site for downloading the Adobe plug-in.

Uploading User's Data for Representation

We recommend that first-time users first browse the plant metabolic pathway maps without uploading their own data sets in the "Map View" section, because it is a good exercise for knowing the KaPPA-View system.

To analyze the user's transcript and/or metabolite quantitative data, Comma Separated Value (CSV)-formatted files are required. It is practical to prepare CSV files using Microsoft Excel, as shown in the KaPPA-View manual, which can be downloadable from the KaPPA-View site. At present, the following gene identifiers are acceptable for Arabidopsis genes: Arabidopsis AGI codes (for example, At4g05160; note that only the initial letter should be capitalized) or the feature IDs of Agilent's Arabidopsis 2 or 3 oligo DNA microarray (note that available feature IDs are described in the file "feature_id.csv," which can be downloaded from the "Download" page). In a manner similar to transcript data files, CSV formatted files for metabolite data must be prepared. The KaPPA-View system recognizes metabolites as unique identifiers (IDs) that are defined in the metabolite information file "compound_info.csv," which is available from the "Download" page. The file provides information about general compound names, International Union of Pure and Applied Chemistry (IUPAC) names, Chemical Abstracts Service (CAS) IDs, and KEGG IDs. The KaPPA-View compound IDs can be searched from the "compound_info.csv" on the "Search" page.

By clicking the "Upload" on the Main Menu, the user can upload his or her own data sets of either transcripts or metabolites or both. The data uploaded will be erased from the server memory when users return to the top page, shut

down the browser, or if no action is done for 30 min. Other users are never able to access any of another person's data.

To view the data uploaded on metabolic maps, the user has to return to the main menu and click "Analysis." Five steps for data selection are required in the "Analysis" section.

Application to Other Plant Species

Although the present version of KaPPA-View is currently only for Arabidopsis, users may apply the transcript data of other plant species to the omics tool by adapting the user's gene list to that of the AGI numbers for the Arabidopsis genes with sequence homology. However, because the number of genes assigned to a reaction in Arabidopsis differs in most cases from that of other species, the user must take the difference into account for the adjustment of gene numbers to each reaction. Another limitation of applying this tool to other plant species is that species-specific metabolic pathways, such as those of isoflavonoids and alkaloids that are not found in Arabidopsis, are not applicable for analysis with KaPPA-View. Currently, plant metabolic maps for other plant species in our project are under preparation.

Data Interpretation

The presentation of data in a manner in which quantitative values of individual transcripts and metabolites are shown on plant metabolic pathway maps, facilitate a good grasp of the transcripts and metabolites, and consequently helps to generate hypotheses of gene function in the metabolic pathway being examined. Because several homologous gene products are often assigned to a single enzymatic reaction, such data presentation on the maps helps to find which genes among homologous genes might be involved in the phenomena examined with the comparative experiments.

Users should note that the genes assigned onto the metabolic reactions are not necessarily guaranteed with experimental evidence, but rather many of them were selected with a computational homology search to a limited number of known genes assigned to each reaction, which originate not only from Arabidopsis, but also from other organisms including animals and bacteria. It is an inherent nature of databases that some faults exist in the information. Therefore, interpretation of the results from the tool should be done with caution. Users are encouraged to send email messages to kappa-view@kazusa.or.jp if they find any errors or would like to comment on the information in the tool.

Currently, it has been shown that analyzing metabolic pathways with networks of gene co-expression is powerful for the interpretation of cell wall gene expression.[7] In the near future, a new version of KaPPA-View will include an option for generating images of networks of gene co-expression, based on Arabidopsis transcriptomic data, on metabolic pathway maps.

References

1) R. A. Dixon and D. Strack: *Phytochemistry* **62**, 815–816 (2003).
2) L. W. Sumner, P. Mendes and R. A. Dixon: *Phytochemistry* **62**, 817–836 (2003).
3) S. Sato, T. Soga, T. Nishioka and M. Tomita: *Plant J.* **40**, 151–163 (2004).
4) S. Y. Rhee, P. Zhang, H. Foerster and C. Tissier: "Biotechnology in Agriculture and Forestry," Vol. 57, ed. by K. Saito, R. A. Dixon and L. Willmitzer, Springer-Verlag, Berlin, Heidelberg, Section II.4, pp. 141–154, 2006.
5) T. Tokimatsu, N. Sakurai, H. Suzuki, H. Ohta, K. Nishitani, T. Koyama, T. Umezawa, N. Misawa, K. Saito and D. Shibata: *Plant Physiol.* **138**, 1289–1300 (2005).
6) T. Tokimatsu, N. Sakurai, H. Suzuki and D. Shibata: "Biotechnology in Agriculture and Forestry," Vol. 57, ed. by K. Saito, R. A. Dixon and L. Willmitzer, Springer-Verlag, Berlin, Heidelberg, Section II.5, pp. 155–163, 2006.
7) S. Persson, H. Wei, J. Milne, G. P. Page and C. R. Somerville: *Proc. Natl. Acad. Sci. USA* **102**, 8633–8638 (2005).