Med Chem Res (2012) 21:47–52
DOI 10.1007/s00044-010-9506-z

MEDICINAL CHEMISTRY RESEARCH

ORIGINAL RESEARCH

PMDB: Plant Metabolome Database—A Metabolomic Approach

M. Udayakumar · D. Prem Chandar · N. Arun · J. Mathangi · K. Hemavathi · R. Seenivasagam

Received: 29 July 2010 / Accepted: 4 November 2010 / Published online: 21 November 2010

Abstract Plant Metabolome Database (PMDB) is a database of secondary metabolites of plants in the three dimensional structures available in the biological data banks and databases. PMDB can be searched via a user-friendly interface and it integrates with JME Editor for sketching and querying metabolite structures which can be viewed in both textual and graphical format using the freely available graphics plug-in Jmol. A navigation link has been constructed which enables the users to browse the database based on metabolite name and use the help page for the detailed description about the fields in the database. This dynamically compiled database contains internal and external links to other databases (KEGG, PUBCHEM, and CAS NUMBER) to provide additional information about the metabolites. This database is intended to serve the scientific community working in the areas of molecular and structural biology and is freely available to all users, around the clock at http://www.sastra.edu/scbt/pmdb.

Keywords Secondary metabolites · JMOL · Metabolomics · Biomolecular database · Molecular structure · Plant metabolites

M. Udayakumar · D. Prem Chandar · N. Arun · J. Mathangi · K. Hemavathi
Department of Bioinformatics, School of Chemical and Biotechnology, SASTRA University, Thanjavur 613402, Tamilnadu, India
e-mail: uthay@bioinfo.sastra.edu

D. Prem Chandar
e-mail: prem@zampbioworld.org

N. Arun
e-mail: arun_aki2002@yahoo.co.in

J. Mathangi
e-mail: mathangi@bioinfo.sastra.edu

K. Hemavathi
e-mail: hema.bioinfo@gmail.com

R. Seenivasagam (✉)
Department of Pharmacoinformatics, Outsource Bioinformatics, Development Center, Grand Residency, DG 4, Kanakapura Rd, Bangalore 560 078, India
e-mail: seenivasagam.pharma@gmail.com

Introduction

Metabolism is the vital process occurring within a living cell or organism (Smith and Morowitz, 2004). Metabolites are plant biomolecules and the intermediate byproducts of metabolism. Metabolome refers to the complete set of small-molecule metabolites (such as metabolic intermediates, hormones and other signaling molecules and secondary metabolites) found within a biological sample, such as single organism (Oliver *et al.*, 1998). Plant primary metabolite is directly involved in the normal growth, development and reproduction of the plant while the secondary metabolites influence them indirectly. Small molecules which mainly include plant hormones and signaling molecules are crucial for the growth, development and differentiation of cells in plants. Databases in the ambit of plant biomolecules could be of great use to bioinformaticians and plant molecular biologists (Altman, 2004). Taking into account the absence of a comprehensive database for plant metabolites, we have excogitated the PMDB as a structurally and functionally annotated plant metabolite database. We have purported to create this library as a representation of an exchange platform for experimental researchers and bioinformaticians to develop and improve metabolomics by multidisciplinary cooperation (Kopka *et al.*, 2005).

Recent innovations in small-molecule analysis has created the ability to rapidly identify and quantify numerous compounds and these data have created new opportunities for understanding plant metabolism and its further engineering (Last *et al.*, 2007). The purpose intended to be served by the database is several fold: (a) to demystify the process of acquisition, processing and display of data by various databases pertaining to plant research, the current limitations and caveats of these resources and the future directions of these resources, (b) to bring forward the general issues of databases and data management to the larger plant research community, (c) to engage the general readership of *Plant-cell Physiology* in thinking about large datasets and their application to research problems, (d) to encourage researchers in the use and development of databases to further their research goals (Rhee and Crosby, 2005).

To this end, there are a limited range of web servers and databases [MoTo DB (Grennan, 2009; Moco *et al.*, 2006), GMD@CSB.DB (Kopka *et al.*, 2005) and Phenol-Explorer (Neveu *et al.*, 2010) available for researchers to view the information on metabolite structures. Although the above-mentioned resources have quantitative information on metabolites and their functions, an interactive web query interface to search and study all the plant metabolite structures in the biological databases is not available. MoTo DB is a metabolite database dedicated to liquid chromatography-mass spectrometry (LC-MS)-based metabolomics (Grennan, 2009; Moco *et al.*, 2006). The online resource MoTo DB provides functional information only on tomato fruit (*Solanum lycopersicum*) on a few of the metabolite classes. GMD@CSB.DB is, again, a web accessible database which provides more information about plant metabolites but it is not updated by 3D structures of metabolites (Kopka *et al.*, 2005). The package Phenol-Explorer is a well-curated, online comprehensive database on polyphenol contents in foods. It contains more than 37,000 original data points collected from 638 scientific articles published in peer-reviewed journals. The quality of these data has been evaluated before they were aggregated to produce final representative mean content values for 502 polyphenols in 452 foods. This database contains no more plant metabolites details with 3D structures, because this is especially for polyphenols in foods (Neveu *et al.*, 2010). The proposed database (PMDB) provides all possible geometric information about all the plant metabolites available in 3D structures in all the biological databases. The primary goal of PMDB is to extract and consolidate information on plant metabolite by means of a user-friendly interface. The database PMDB uses a versatile PHP search engine coupled with a fast database engine (MySQL) for information retrieval and knowledge discovery.

Materials and methods

Database implementation and access

Recently, standards in data description and exchange for plant metabolomics have been proposed (Bino *et al.*, 2004) and a computer-readable data model has been developed (Jenkins *et al.*, 2004, 2005). In creating the PMDB, the metabolites were integrated into the web-based MySQL database as the backend using open source Apache Web server. An interactive web interface was created using Hypertext Preprocessor (PHP) and JavaScript (Meloni, 2000). Figure 1 gives the schematic design of the database

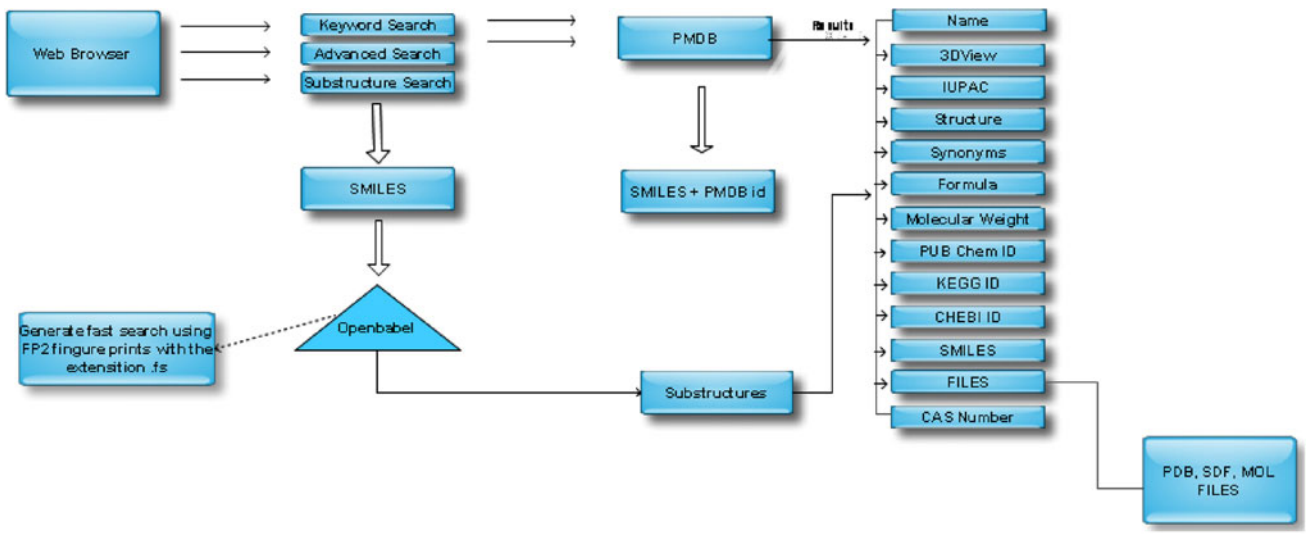

Fig. 1 Workflow diagram of PMDB database

and its working mechanism. The database primarily provides information of the metabolites on searching for them using either the keyword or structure. The graphical display is enabled by the use of Jmol to display the three-dimensional structures. The 3D atomic coordinates are updated so that users can access up-to-date information on the structures available in the Biological data banks and databases. JME Editor was integrated for sketching and querying metabolite structures. Note that Jmol works on all browsers that are enabled with Java plug-ins. The proposed database is very user friendly and has been tested extensively on various computing platforms using the Mozilla Firefox browser. The database is validated and its response time is very fast. Of course, the response time may vary depending upon the network speed and the number of users accessing the database at that particular time. This database can be freely accessed at http://www.sastra.edu/scbt/pmdb. The users of this database are requested to cite this article in their scientific reports. General comments and suggestions for additional options are welcome and should be sent to the corresponding author, Mr. R. Seenivasagam (e-mail: seenivasagam.pharma@gmail.com).

Results and discussion

Features

Plants produce a gamut of metabolites, which are of great use to mankind. Much of laboratory research has gone into studying its occurrence and properties. However, to suffice the need for a comprehensive database, we have developed the Plant Metabolome Database (PMDB). Plant Metabolome Database (PMDB) is a structurally and functionally annotated database of metabolites present in plants. The current version of the database has more than 1,000 metabolites (as of 23 Jan, 2010) three-dimensional structures; details of the numbers of metabolites are provided in Table 1. Each entry provides comprehensive details about the metabolite name, common name for the metabolite, description, canonical, as well as isomeric smiles, molecular formula, different file formats (SDF, MOL, and PDB), and 3D structure of the metabolites. The PMDB provides internal and external links to databases (KEGG, PUBCHEM, and CHEBI). The database integrates tools like JME Editor and JMOL for sketching, displaying, and manipulating metabolites structures. PMDB is a unique effort put in by us in order to integrate all the information available about small molecules under one roof so as to make it easy for researchers and scientists to access all the required information in the first sight.

PMDB provides online web tools that allow users to query, retrieve, and analyze data using: (i) keyword search which uses simple search form to allow very quick query against the database with the keywords like the common name, PMDBid and synonyms (Fig. 2a) However, on entering the starting alphabet of the metabolite, the database enlists all possible metabolites beginning with the given letter, (ii) advanced search form enables the user to search using complex queries. The search form takes Molecular formula as the input. The Molecular formula search can be done, by entering its composition, as in the following example ($C_6H_{10}O_7$). The user must fill the textboxes provided against each of the atom that could possibly be a part of the metabolite, as in, (C = 6, H = 10, O = 7). The search for options takes PMDBid, common name, synonyms; IUPAC name, Pubchem id, and CAS number as input. The database can be queried using any one of these options or both (Fig. 2f), (iii) substructure search will give

Table 1 Statistics of various plant secondary metabolites available in the PMDB database

Main class	Sub-class	No. of secondary metabolites
Alkaloids	Betalain alkaloids	28
	Indole alkaloids	129
	Peptide alkaloids	22
	Acridone alkaloids	13
	Quinazoline alkaloids	14
	Diterpenoid alkaloids	64
	Other alkaloids	176
Phenolics	Minor flavonoids	50
	Benzofurans	28
	Flavones	15
	Anthocyanins	31
	Quinones	60
	Isoflavonoids	70
	Tannins	21
Terpenoids	Diterpenoids	25
	Monoterpenoids	25
	Triterpenoid Saponins	25
	Phytosterols	27
	Steroid saponins	26
	Carotenoids	30
	Other terpenoids	17
Other metabolites	–	54
Hormones	Brassinosteroids	22
	Gibberellins	40
	Jasmonates	5
	Abscisic acid	7
	Auxins	4
	Cytokinins	2
	Ethylene	2
Total		1032

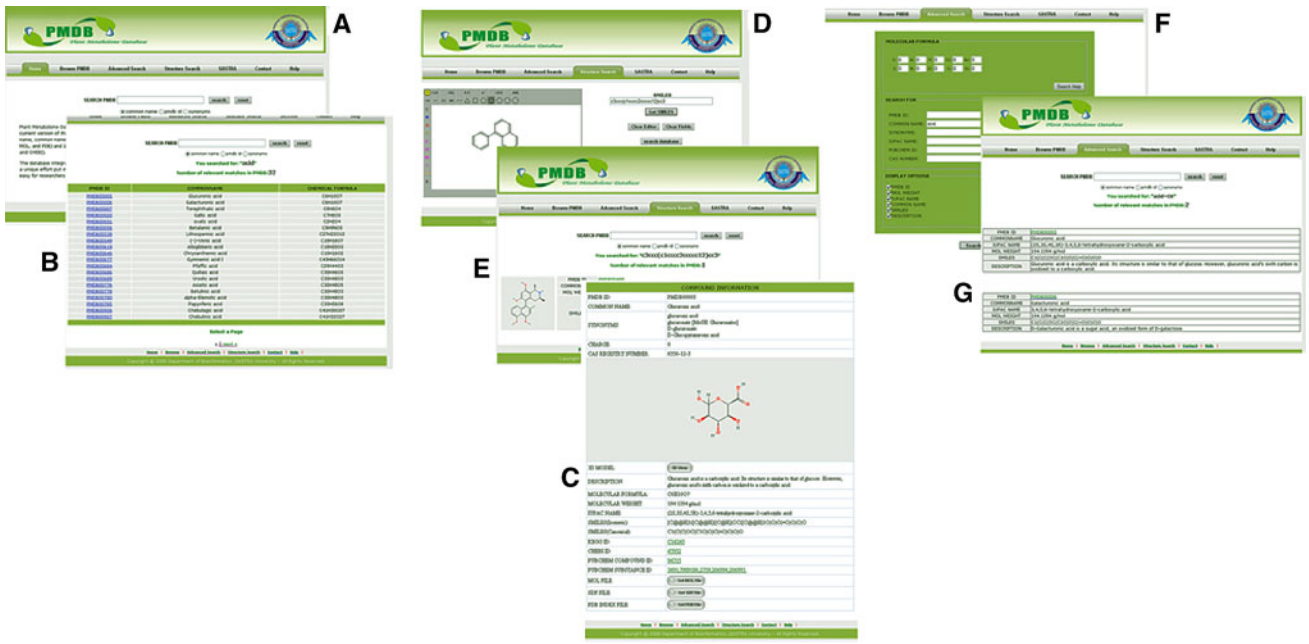

Fig. 2 This picture can help to guide the user in every stage of the search process. **a** Home Page of the PMDB database. **b** Search Page of PMDB database. The webpage is searched for the keyword "acid" in the text field. The result page builds a table of entries with short description hyperlinks displayed. **c** Detailed view of metabolite information. Clicking on the internal accession number PMDB database gives the information. **d** 3D substructure search page using integrated JME Editor. **e** List of similar structures is displayed matching the input query given in substructure search page. **f** Advanced search option enables the user to search the database using molecular formula of metabolites and list of entries matching the query are displayed in page **g**

effective database search based on the structures. User queries are captured using the JME Editor (http://www.molinspiration.com/JME) which allow user to sketch or edit molecules. The drawn structure is converted to SMILES which is given as input to OpenBabel (http://openbabel.org), a standalone program which requires prior installation. OpenBabel takes two parameters: (1) a fast index file for the entire plant metabolites database, (2) SMILES from JME. The OpenBabel generates list of substructures, and it is parsed using PHP script and results are displayed in the tabulated form (Fig. 2d). The results of the query builds a table of entries with PMDBid and a short description matching the search criteria is displayed (Fig. 2d). The detailed description about the metabolites can be retrieved by clicking the PMDBid (Fig. 2c). The information about the metabolites includes: (i) PMDBid, is a unique identifier for the metabolite, (ii) common name, (iii) synonyms, (iv) description, (v) molecular formula, (vi) canonical and isomeric smiles, (vii) file formats (SDF, MOL, PDB index file), (viii) pubchem compound id. In addition, we have generated two-dimensional structures for the metabolites (Fig. 2e).

PMDB integrates a tool JMOL (http://jmol.sourceforge.net/) display and manipulate 3D structures of the metabolites. The browse option adds a feature to the database that will display all the plant metabolites based on the categories of secondary metabolites, hormones, and signaling molecules. By clicking the category, the option is expanded to view the metabolite information. The PMDB provides a "help page" which contains the documentation about the fields used in the database and configuration of the JMOL.

Case study

Artemisia judaica is a traditional Egyptian medicinal plant. De novo shoot organogenesis was induced by culturing etiolated hypocotyls and intact seedlings on medium supplemented with thidiazuron [*N*-phenyl-*N*′-(1,2,3-thidiazol-yl) urea] via callusing at the cotyledonary notch region (Liu *et al.*, 2003). Up to 16 shoots formed per seedling cultured on a medium containing 1 µmol, l (-1) thidiazuron for an optimal duration of exposure for 20 days. Regenerated shoots formed roots when sub-cultured onto a medium containing 1 µmol, l (-1) indole-3-butyric acid. Here, regeneration protocol developed in this study provides a basis for germplasm conservation and for further structural investigation of medicinally active biomolecule of indole-3-butyric acid (Fig. 3) of *A. judaica*. In the investigation of the anticonvulsant activity of aqueous root extract of *Securidaca longepedunculata*, strychnine is a primary biomolecule and medicinal metabolite (Fig. 4). The

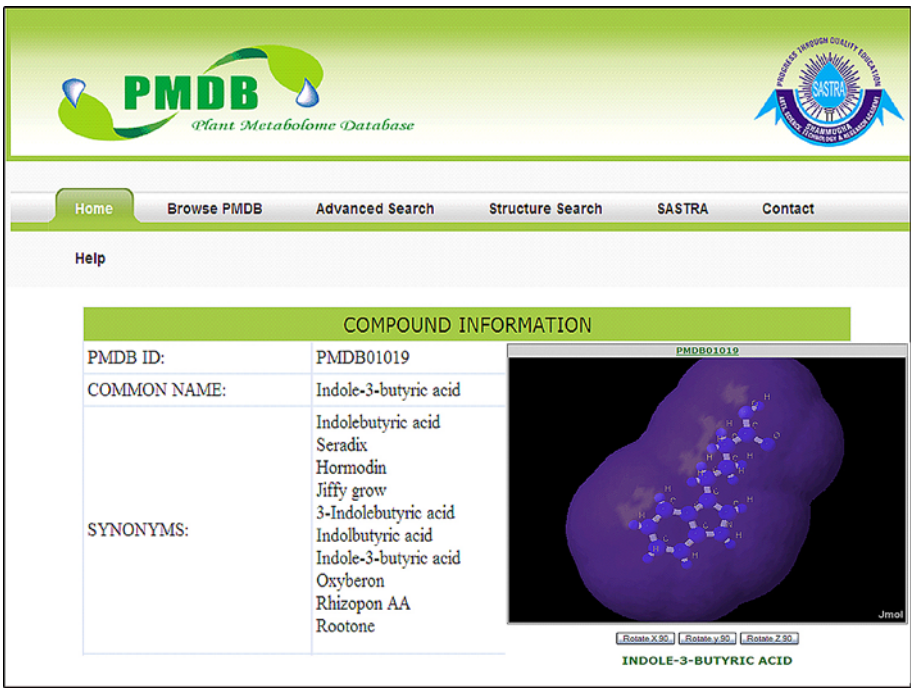

Fig. 3 A screenshot showing the structure of a indole 3-butyric acid (PMDB ID: 01019) with solvent accessible surface (VDW + 1.4 Å) in *Artemisia judaica*. The graphics panel Jmol (*bottom right*) displays the structure

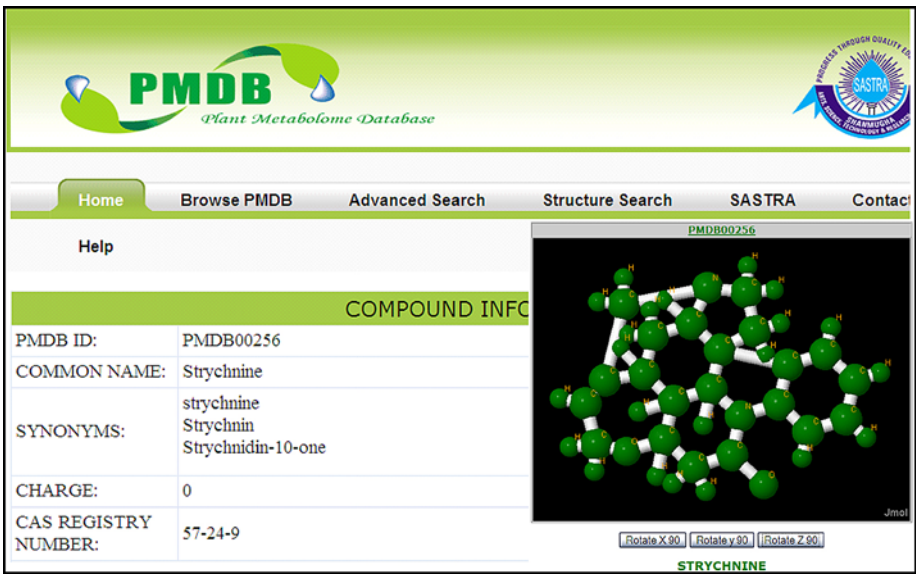

Fig. 4 The output panel depicts the coordination of C, H, N, and O atoms in the crystal structure of strychnine (PMDB ID: 00256) of aqueous root extract of *Securidaca longepedunculata*

anticonvulsant effect of the aqueous root extract (100, 200, and 400 mg/kg) was evaluated in mice using the strychnine- and picrotoxin-induced seizure models. The extract (100–400 mg/kg) produced a significant ($P < 0.01$) dose-dependent increase in onset of convulsion compared to the control for strychnine- and picrotoxin-induced seizures (Adeyemi *et al.*, 2010). It also produced a significant ($P < 0.01$) dose-dependent prolongation of the cumulative time spent in the open arms of the elevated plus maze and Y maze compared with the control. These findings justify the use of *Securidaca longepedunculata* in traditional medicine for the management of convulsion.

Perspectives of computational metabolomics

The study of a collection of metabolites as a whole (metabolome), as opposed to isolated small molecules, is a fast-growing to take us one step further toward understanding cell biology and relating the genetic capabilities of an organism to field promising its observed phenotype. The new sciences of metabolomics and metabonomics can exploit a variety of existing experimental and computational methods, but they also require new technology that can deal with both the amount and the diversity of the data relating to the rich world of metabolites (Bhalla *et al.*,

2005; Fiehn, 2001; Fiehn *et al.*, 2001). More specifically, the collaboration between bioinformaticians and chemoinformaticians promises to advance our view of cognate molecules, by shedding light on their atomic structure and properties. Modeling of the interactions of metabolites (Allwood *et al.*, 2010), with other entities in the cell and eventually complete modeling of reaction pathways will be essential for analysis of the experimental data and prediction of an organism's response to environmental challenges (Nobeli and Thornton, 2006). Identification and quantification of metabolites occurring within specific cell types or single cells of plants and other organisms is of particular interest for natural product chemistry, chemical ecology, and biochemistry in general (Moco *et al.*, 2009). The integration of studies at the gene, transcript, protein, and metabolite levels in localized regions will provide useful information for the understanding of biology as a system (micrometabolomics). However, the diverse potential of metabolomics in many fields, including cell engineering has made it a universal tool for industrial, medical, and research purposes. It is also a vital component of a 'systems biology' approach, as it is believed to be a good reflection of the phenotype of any cell or tissue. At the heart of metabolomics' growth is the issue of method development, including sample preparation, instrument analysis, data processing, and bioinformatics (Khoo and Al-Rubeai, 2007).

Conclusion

PMDB is a structurally and functionally annotated database for metabolites in plants. While plants form the major life forms, this database exhaustively gives information regarding metabolomes and small molecules in plants. This database is intended to bridge the gap between the results of various wet lab works and bioinformatics through a one-stop gateway. The database is designed to provide an online access with efficient search tools and well designed web interface. This comprehensive knowledge resource will be of great use to researchers working in the area of biochemistry, computational chemistry, and bioinformatics.

Acknowledgments The authors acknowledge Peter Ertl (Novartis et al) for the JME. The work was carried out at the Department of Bioinformatics developed with the financial support of the SASTRA University. The authors sincerely acknowledge Prof R. Sethuraman, Vice-Chancellor, SASTRA University, for providing scope to work and encouragement toward research. We thank Dr. M. Vijayalakshmi for her valuable comments and suggestions rendered to carry out the work.

References

Adeyemi OO, Akindele AJ, Yemitan OK, Aigbe FR, Fagbo FI (2010) Anticonvulsant, anxiolytic and sedative activities of the aqueous root extract of *Securidaca longepedunculata* Fresen. J Ethnopharmacol 130(2):191–195

Allwood JW, Clarke A, Goodacre R, Mur LA (2010) Dual metabolomics: a novel approach to understanding plant-pathogen interactions. Phytochemistry 71:590–597

Altman RB (2004) Building successful biological databases. Brief Bioinform 5:4–5

Bhalla R, Narasimhan K, Swarup S (2005) Metabolomics and its role in understanding cellular responses in plants. Plant Cell Rep 24:562–571

Bino RJ, Hall RD, Fiehn O et al (2004) Potential of metabolomics as a functional genomics tool. Trends Plant Sci 9:418–425

Fiehn O (2001) Combining genomics, metabolome analysis, and biochemical modelling to understand metabolic networks. Comp Funct Genomics 2:155–168

Fiehn O, Kloska S, Altmann T (2001) Integrated studies on plant biology using multiparallel techniques. Curr Opin Biotechnol 12:82–86

Grennan AK (2009) MoTo DB: a metabolic database for tomato. Plant Physiol 151:1701–1702

Jenkins H, Hardy N, Beckmann M et al (2004) A proposed framework for the description of plant metabolomics experiments and their results. Nat Biotechnol 22:1601–1606

Jenkins H, Johnson H, Kular B, Wang TL, Hardy N (2005) Toward supportive data collection tools for plant metabolomics. Plant Physiol 138:67–77

Meloni JC (2000) PHP essentials, vol 4. BPB Publications, New Delhi, pp 108–132

Khoo SH, Al-Rubeai M (2007) Metabolomics as a complementary tool in cell culture. Biotechnol Appl Biochem 47:71–84

Kopka J, Schauer N, Krueger S et al (2005) GMD@CSB.DB: the Golm Metabolome Database. Bioinformatics 21:1635–1638

Last RL, Jones AD, Shachar-Hill Y (2007) Towards the plant metabolome and beyond. Nat Rev Mol Cell Biol 8:167–174

Liu CZ, Murch SJ, EL-Demerdash M, Saxena PK (2003) Regeneration of the Egyptian medicinal plant *Artemisia judaica* L. Plant Cell Rep 21:525–530

Moco S, Bino RJ, Vorst O et al (2006) A liquid chromatography-mass spectrometry-based metabolome database for tomato. Plant Physiol 141:1205–1218

Moco S, Schneider B, Vervoort J (2009) Plant micrometabolomics: the analysis of endogenous metabolites present in a plant cell or tissue. J Proteome Res 8:1694–1703

Neveu V, Perez-Jimenez J, Vos F et al (2010) Phenol-Explorer: an online comprehensive database on polyphenol contents in foods. Database 2010:bap024

Nobeli I, Thornton JM (2006) A bioinformatician's view of the metabolome. Bioessays 28:534–545

Oliver SG, Winson MK, Kell DB, Baganz F (1998) First use of the term metabolome in the literature—systematic functional analysis of the yeast genome. Trends Biotechnol 16:373–378

Rhee SY, Crosby B (2005) Biological databases for plant research. Plant Physiol 138:1–3

Smith E, Morowitz H (2004) Universality in intermediary metabolism. Proc Natl Acad Sci USA 101:13168–13173