Available online at http://www.journalcra.com

International Journal of Current Research
Vol. 9, Issue, 02, pp.46537-46541, February, 2017

INTERNATIONAL JOURNAL OF CURRENT RESEARCH

ISSN: 0975-833X

RESEARCH ARTICLE

PDTDB – AN INTEGRATIVE STRUCTURAL DATABASE AND PREDICTION SERVER FOR PLANT METABOLITES AND THERAPEUTIC DRUG TARGETS

[1,*]Ashok Kumar T. and [2]Rajagopal B.

[1]Department of Bioinformatics, Noorul Islam College of Arts and Science, Thuckalay, Tamil Nadu, India
[2]PG Department of Zoology, Government Arts College, Dharmapuri, Tamil Nadu, India

ARTICLE INFO

Article History:
Received 16th November, 2016
Received in revised form
10th December, 2016
Accepted 28th January, 2017
Published online 28th February, 2017

Key words:

PDTDB,
Structure Activity Relationship,
Integrative Structural Database,
Structure Binding,
Target-Ligand Interaction,
Prediction Server.

ABSTRACT

Understanding the biological phenomena of sequence, structure, function, metabolism, and molecular interactions of species is crucial to identify or analyze the biological problems in the ideal way. Integration of all data facilitates unified access to the key problem. PDTDB (Phytochemical and Drug Target DataBase) is anintegrative structural database, which contains information about the plants (secondary metabolites, plant anatomy, side effects, and medicinal properties), ligands (3D molecular structure, SMILES string, side effects, and medicinal properties), therapeutic drug targets (3D molecular structure, sequence, mechanism of the drug target, disease nature, and disease symptoms), molecular dockings (target-ligand complex interaction, drug action, home remedies, and docking results), and structure activity relationship between diseases, therapeutic targets, plants, and phytochemicals. The database provides keyword or accession number search, similar phytochemical structures search using SMILES string or by sketching the structure in the chemical structure editor, phytochemical structures with similar molecular formula search, binding structure search, and browse and/or search database entries from the sortable table. It supports predicting physicochemical properties of sequence and structures, and interactive visualization of structures in various models. PDTDB is freely accessible at https://pdt.biogem.org.

Citation: Ashok Kumar and Rajagopal, 2017. "PDTDB – An integrative structural database and prediction server for plant metabolites and therapeutic drug targets", *International Journal of Current Research*, 09, (02), 46537-46541.

INTRODUCTION

Throughout the ages, natural products have formed as the major and valuable resource for health and medicine to the humans (Bagetta, 2012; Buss and Butler, 2010; Lahlou, 2007). Most of the drug discoveries of modern medicines were either extracted from natural products or inspired by them (Prasad and Tyagi, 2015; Ji *et al*., 2009; Li and Vederas, 2009). Natural product consists of compounds that are derived from natural sources such as plants, minerals, and animals (Natural Products Research, 2016; Lahlou, 2007; What Are the Pharmaceutical Sources of Drugs, 2016). Phytonutrients or phytochemicals are the organic bioactive chemical compounds present in the plants that have disease preventive properties (What Are Phytonutrients, 2016; Lahlou, 2013; Omotayo and Borokini, 2012; A-Z of Natural Foods and Nutritional Benefits, 2016). Natural products have the tendency to improve the body's vital ability to balance and to heal itself (A Close Look at Naturopathy, 2016). Many research studies show that consuming phytonutrient rich foods and beverages prevent from disease (Phytochemicals' Role in Good Health, 2016; Health Benefits of Fruits and Vegetables, 2016; A-Z of Natural Foods and Nutritional Benefits, 2016). But some of the phytochemicals loss or change its chemical structure during cooking foods, due to thermal decomposition and chemical reactions (Palermo *et al*., 2014). The success of natural products in drug discovery is, it consists of unique and more complex molecular structures while compared to synthetic molecules (Lahlou, 2013; Krishnamoorthy and Subramaniam, 2014). Due to its nature, selection of binding regions in the drug target for docking with phytochemicals is more (Lahlou, 2007; Lahlou, 2013).

Recently, many big pharmaceutical companies and R&D organizations have renewed their interest in discovering potential lead compounds from the natural products for life threatening diseases (Lahlou, 2007). In fact, it is estimated that about 49% of modern medicines is either natural products or directly derived (semi-synthetic) therefrom (Newman and Cragg, 2016; Newman and Cragg, 2007; Cragg and Newman, 2013). With the advances in modern drug discovery, three-dimensional molecular structure and pharmacokinetic properties of many phytochemicals were identified. Despite there are still many thousands out there which are yet to be

Corresponding author: Ashok Kumar, T.
Department of Bioinformatics, Noorul Islam College of Arts and Science, Thuckalay, Tamil Nadu, India.

discovered (Lahlou, 2007). The environmental changes over years has destroyed many valuable medicinal plants and mineral resources.

Over the past decade, several commercial and open access 3D structural databases for natural products, drug targets, and molecular interactions were released with different specifications and large entries. Even though the databases have quantitative information, there no solid database containingboth integrated data and properties prediction servers. For this purpose, we have been developing an integrativestructural database which overcomes the problem.

MATERIALS AND METHODS

Data Collection

Information and/or image of the herbal plants/natural products, phytochemical compounds, drug targets, symptoms and causes of the diseases, and natural remedies to cure the diseases were collected and manually curated from the WebMd (https://www.webmd.com), eBooks, journal articles, blogs and other web pages through Google search. All information in the database are properly cross referenced by hyperlinks to the corresponding original source web page. The 3D phytochemical structures were retrieved from the PubChem (Kim *et al*., 2016) / DrugBank (Law *et al*., 2014) / ChEBI (Degtyarenko *et al*., 2008) / ChemSpider (Pence and Williams, 2010) / HMDB (Wishart *et al*., 2013), and 2D phytochemical images were obtained from Marvin JS version 16.7.4 (https://www.chemaxon.com) tool by manually drawing the structure or importing the structure file retrieved from the database. Some of the information about rare phytochemical compounds was retrieved from the book "*Encyclopedia of Traditional Chinese Medicines*: Molecular Structures, Pharmacological Activities, Natural Sources and Applications, (Vol. 1, 2, 3, 4, 5, and 6)" (Zhou *et al*., 2011). The unavailable 3D structure of phytochemical compounds was obtained manually by drawing the phytochemical structure in the JSDraw v4.0.5 (http://www.scilligence.com) tool and then the SMILES string is translated to the 3D molecule using CACTUS (https://cactus.nci.nih.gov/translate)/Balloon (Vainio and Johnson, 2007) /Open Babel (O'Boyle *et al*., 2011) /CORINA (Sadowski *et al*., 1994) tool.

Based on several literature studies, the therapeutic targets and binding phytochemical compounds were chosen. The 3D structure of drug targets was retrieved from the Protein Data Bank (PDB) (Berman *et al*., 2000). Details of binding affinities, interactions, drug-like properties, and drug action of drug target and phytochemicals were gathered from the BindingDB (Gilson *et al*., 2016), KEGG (Kanehisa and Goto, 2000), DrugBank (Law *et al*., 2014), PharmGKB (Hewett *et al*., 2002), STITCH (Kuhn *et al*., 2014), and SuperTarget (Hecker *et al*., 2012) database. Target-ligand interaction studies are carried out using CLC Drug Discovery Workbench version 3.0.1 (https://www.qiagenbioinformatics.com) and LigPlot+ (Laskowski and Swindells, 2011) tools.

Database Construction

The PDTDB was constructed on CentOS Enterprise v5.8 x86_64 Linux server with Apache HTTP Server v2.2.31, MySQL Community Server v5.5.50, PHP v5.4.45, and phpMyAdmin v4.0.10.14. The interactive web interface was designed and implemented using jQuery, AJAX, JavaScript, Dojo, HTML5, and CSS3. A JavaScript molecule viewer tool JSmol v14.2.9 (Hanson *et al*., 2013) is embedded in web pages to interactively display the phytochemical compounds in ball-and-stick model, drug targets in ribbon model, and target-ligand complex in docking interaction model. A Java chemical structure editor JSDraw version 4.0.5 (http://www.scilligence.com) is used for searching similar phytochemical structures.

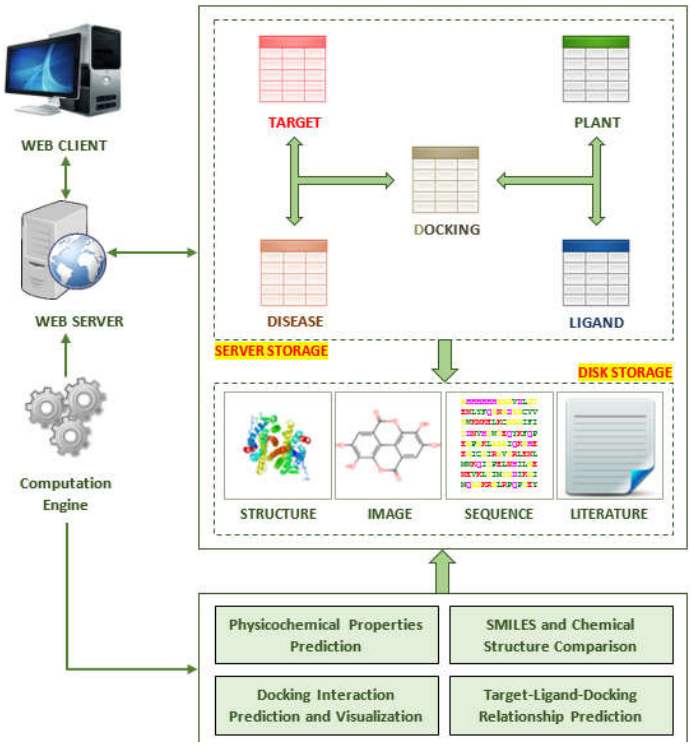


Figure 1. Workflow diagram and database schema of the PDTDB

A simple workflow diagram representing the database architecture, database schema, entity relationship model of the database, entity integration, content in the database, and computation methods are shown in the Figure 1. The architecture of PDTDB is broadly categorized into graphical user interface (front-end) and server storage interface (back-end). In a graphical user interface, it consists of user interactive web pages, which are accessible for the users (can search, browse, and download data), group members (can add data, and predict data), and administrator (can add, modify and delete data; enable or upgrade modules; control user access; and server monitoring). The site administrator has direct access to the server storage interface and several modules in the PDTDB, through administrator portal. In a server storage interface, it consists of MySQL server which is connected to the phpMyAdmin. MySQL server stores the plant, ligand, target, disease, and docking information; and links to the files stored on the local disk.

RESULTS AND DISCUSSION

PDTDB is a tri-functional and integrated database, it serves as a plant database, drug target database, and docking database. Currently the PDTDB has published 61 records of medicinal plants/natural products (including major entities such as secondary metabolites, description and health benefits and image), 237 records of phytochemical ligands (including major entities such as SMILES, description and health benefits,

image, and 3D structure), 8 records of targets (including major entities such as description, sequence, 3D structure, and image), 8 records of diseases (including major entities such as description, symptoms, and image), and 12 records of target-ligand docking (including major entities such as score, no. of interactions, herbal recipe, drug action, 3D structure, and image). Each plant image represents leaf, fruit, stem, and/or seeds. Also there are nearly 307 therapeutic drugtarget and disease data, 1,657 plant data, 49,589 ligands, and 25 docking data under review. PDTDB is still under development, as we are currently focused on data curation and construction of several automatic prediction and validation servers to increase data accuracy and provide more information.We are constantly striving to complete the process within a short span. Also, server scripting for automatic integration and data extraction from the external databases or servers is under progress.

Accessing the Database

The web interface of PDTDB was designed for the benefit of both biologists and non-biologists, so that users can easily access the targeted information without much technical skill. The hyperlinks in the PDTDB web pages are annotated to provide additional information in the form of thumbnail images or tooltip text by simply moving the mouse pointer over the hyperlinked text. In the 'Help' menu of PDTDB, 'Troubleshoot', 'Tutorial', 'PDTDB Entries', and 'Resources and Tools' web pages are included to provide more information about accessing the database.The data entities in PDTDB comprises of 'Target', 'Plant', 'Ligand', and 'Docking'. Information from the PDTDB can be accessed through the 'Search' page, 'Browse' page, 'Binding' page, and 'Download' page. A brief description about the web page interfaces and accessing methods are given in below sections.

1. Search page

In 'Search' page, there are various types of search methods available to retrieve the data from PDTDB, which are listed as follows:

- *Keyword Search* method allows users to perform text-based search the 'keywords' field of the database.
- *Unique ID Search* method allows users to perform accession number search on the 'id' field of the database. The format for PDTDB accession numbers are 5 letters (database name 'pdtdb') + 1 letter (types of entries 't' – target, 'p' – plant, 'l' – ligand, and 'd' – docking) + 5 digits (serial number of the database entry in left padding with zero format). For instance, pdtdbl00282.
- Simply *Structure Search* method allows users to perform SMILES string search on the 'smiles' field of a ligand table in the database. On pressing the 'JSDraw' button, a chemical structure editor is displayed. By drawing the 2D chemical structure in the editor, users can search on the 'smiles' field of a ligand table in the database or PubChem database.
- *Chemical Formula Search* method allows users to perform molecular formula search on the 'formula' field of a ligand table in the database. Chemical formula search is included in the ligand/docking result page near the section "Molecular Formula :". By clicking the hyperlinked formula, it searches for matching records in the ligand table with similar chemical formula.

2. Browse page

The 'Browse' page contains four accordion panels with sortable tables inside. Each table in the accordion panel consists of the list of entries with the following fields: accession number, name, and alternative names/keywords. The rows in the table can be sorted in ascending/descending order according the fields, by clicking the column header in the table. Also, users can filter or shorten the list of entries in the table by typing a text in the search box.From the ligand table, to facilitate the formula search or sort, the superscript and subscript of the molecular formula were removed. For this reason, users can able to search or sort the molecular formula by simple text search. For example, $C_{29}H_{50}O_2$ can be searched as linear text C29H50O2 or C29/H50/O2.

3. Binding page

Binding search allows users to predict the structure activity relationship (SAR) of the biomolecules. Currently PDTDB supports target-wise relationship and ligand-wise relationship prediction of biomolecules. The result page of binding search is displayed as a sortable table containing target-ligand interaction, target details, ligand detail, a list of plants containing the ligand, docking score, and target-ligand interaction bonds.

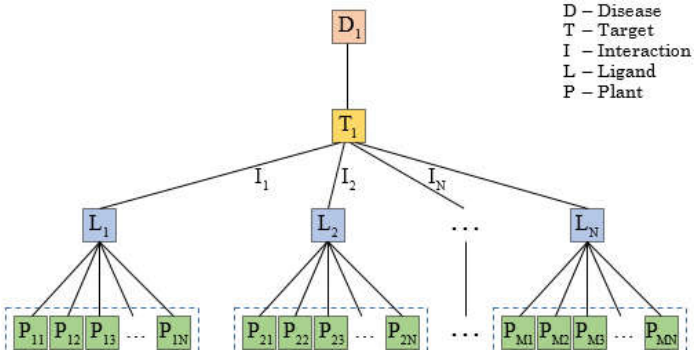


Figure 2. Network model of data integration in the PDTDB

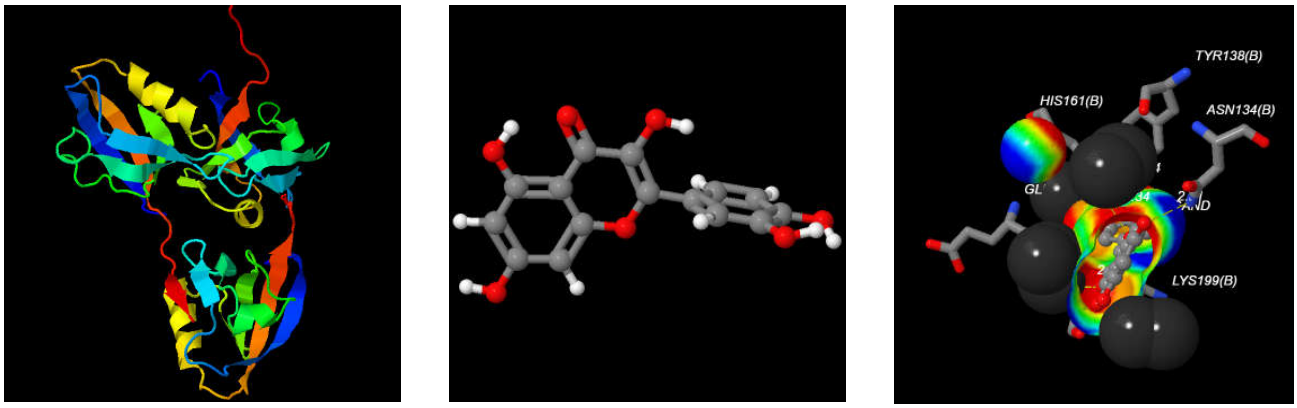
The binding networkmodel (Figure 2) represents the data integration of entities in PDTDB and structure activity relationship between diseases, therapeutic targets, plants, and phytochemicals.

4. Download page

The 'Download' page consists of four download options, namely 'Summary', 'Structures', 'Ligands', and 'Sequences'. In each section, AJAX enabled autocomplete text box is used to fetch the result instantly while typing the query. The allowed query search is Comma Separated Value (CSV) or Space Separated Value (SSV) formatted accession numbers. Users are allowed to download large numbers of sequence/structure/text files from the PDTDB as batches in compressed file format (.zip).

Features of PDTDB

Apart from serving as a database, the PDTDB allows users to predict molecular interactions of target-ligand complex, predict physicochemical properties of sequence/structure, and search for similar 3D structures in PubChem and PDB database. The physicochemical properties prediction of sequences and



(A) (B) (C)

Figure 3. Screenshot of various models of structures visualization in 3D Molecular Structure Viewer of PDTDB. (A) Ribbon Model. Target structure in cartoon model with group colours in black background, **(B) Ball-and-stick model.** Ligand structure in ball-and-stick model with CPK colours in black background, and **(C) Iso-surface model.** Docking structure in iso-surface model with ROYGB colours in black background. The dotted line represents the target-ligand interaction, the number over the dotted line represents the distance between two atoms (Å), ball-and-stick model represents the ligand molecule, the stick model represents the side chain interacting the ligand molecule, sphere model represents the atoms with weak interaction, and the surface model represents the surface area between the target and ligand

structures includes: (i) length of the protein sequence, (ii) molecular weight of protein and phytochemical compound, (iii) isoelectric point of the protein, (iv) amino acid composition, (v) total number of negatively and positively charged amino acids, (vi) atomic composition of the protein, (vii) molecular formula of the protein and phytochemical compound, (viii) total number of atoms in the protein, (ix) extinction coefficientsof the protein, (x) estimated half-life of the protein, (xi) instability index of the protein, (xii) aliphatic index of the protein, (xiii) grand average of hydropathicity (GRAVY) of the protein, and (xiv) molecular interactions of protein-ligand complex.

The build-in three-dimensional molecular visualization tool enables interactive visualization of structures in various models, based on the type of the molecules (Figure 3).

Conclusion

PDTDB is an integrated, comprehensive, and open access resource for medicinal herbs/natural products, phytochemicals, drug targets, docking results, symptoms and causes of the diseases, and natural remedies to treat the diseases. All entities in the PDTDB are integrated by the related entries to provide more information. The 3D structure of all phytochemical and interacting target proteins can be accessed for each medicinal herbs/natural products. Also, other related interactions of target and phytochemical compounds can be achieved.In future, the PDTDB will be enhanced with more automatic prediction/ validation server for physicochemical/pharmacokinetic properties, docking evaluation, visualization tool for target-ligand-plant-docking interaction network, and integrate with the external databases and prediction servers.

Conflicts of interest

The authors declare that they have no competing interests.

REFERENCES

A Close Look at Naturopathy. In: *Quackwatch*. http://www.quackwatch.com/01QuackeryRelatedTopics/Naturopathy/naturopathy.html. Accessed 23 Oct 2016

A-Z of Natural Foods and Nutritional Benefits. In: *Nature Cures*. http://www.naturecures.co.uk/a-zkitchencures.htm. Accessed 23 Oct 2016

Bagetta G. 2012.*Herbal Medicines*: *Development and Validation of Plant-derived Medicines for Human Health*. CRC Press, Boca Raton, FL.

Berman HM, Westbrook J, Feng Z, *et al*. 2000. The Protein Data Bank. *Nucleic Acids Res*, 28(1):235–242. doi: 10.1093/nar/28.1.235

Buss AD, Butler MS. 2010.*Natural Product Chemistry for Drug Discovery*. Royal Society of Chemistry, Cambridge. ISBN: 978-0-85404-193-0

Cragg GM, Newman DJ. 2013. Natural products: A continuing source of novel drug leads. *Biochim Biophys Acta*, 1830(6):3670–3695. doi: 10.1016/j.bbagen.2013.02.008

Degtyarenko K, de Matos P, Ennis M, *et al*. 2008. ChEBI: a database and ontology for chemical entities of biological interest. *Nucleic Acids Res*, 36(D1):D344–D350. doi: 10.1093/nar/gkm791

Gilson MK, Liu T, Baitaluk M, *et al*. 2016. BindingDB in 2015: A public database for medicinal chemistry, computational chemistry and systems pharmacology. *Nucleic Acids Res*, 44(D1):D1045–D1053. doi: 10.1093/nar/gkv1072

Hanson RM, Prilusky J, Renjian Z, *et al*. 2013. JSmol and the Next-Generation Web-Based Representation of 3D Molecular Structure as Applied to Proteopedia. *Isr J Chem*, 53(3-4):207–216. doi: 10.1002/ijch.201300024

Health Benefits of Fruits and Vegetables.In: *Reboot With Joe*. http://www.rebootwithjoe.com/eating/a-z-fruit-veggies/. Accessed 21 Aug 2016

Hecker N, Ahmed J, von Eichborn J, *et al*. 2012. SuperTarget goes quantitative: update on drug–target interactions. *Nucleic Acids Res*, 40(D1):D1113–D1117. doi: 10.1093/nar/gkr912

Hewett M, Oliver DE, Rubin DL, *et al*. 2002. PharmGKB: the Pharmacogenetics Knowledge Base. *Nucleic Acids Res*, 30(1):163–165. doi: 10.1093/nar/30.1.163

Ji H-F, Li X-J, Zhang H-Y. 2009. Natural products and drug discovery. Can thousands of years of ancient medical knowledge lead us to new and powerful drug combinations in the fight against cancer and dementia? *EMBO reports*, 10(3):194–200. doi: 10.1038/embor.2009.12

Kanehisa M, Goto S. 2000. KEGG: Kyoto Encyclopedia of Genes and Genomes. *Nucleic Acids Res*, 28(1):27–30. doi: 10.1093/nar/28.1.27

Kim S, Thiessen PA, Bolton EE, *et al*. 2016. PubChem Substance and Compound databases. *Nucleic Acids Res*, 44:D1202–D1213. doi: 10.1093/nar/gkv951

Krishnamoorthy K, Subramaniam P. 2014. Phytochemical Profiling of Leaf, Stem, and Tuber Parts of *Solena amplexicaulis* (Lam.) Gandhi Using GC-MS. *Int Sch Res Notices*, 2014:1–13. doi: 10.1155/2014/567409

Kuhn M, Szklarczyk D, Pletscher-Frankild S, *et al*. 2014. STITCH 4: integration of protein–chemical interactions with user data. *Nucleic Acids Res*, 42:D401–D407. doi: 10.1093/nar/gkt1207

Lahlou M. 2013. The Success of Natural Products in Drug Discovery. *Pharmacol Pharm*, 4(3):17–31. doi: 10.4236/pp.2013.43A003

Lahlou M. 2007. Screening of Natural Products for Drug Discovery. *Expert Opin Drug Discov*, 2(5):697–705. doi: 10.1517/17460441.2.5.697

Laskowski RA, Swindells MB. 2011. LigPlot+: Multiple Ligand–Protein Interaction Diagrams for Drug Discovery. *J Chem Inf Model*, 51(10):2778–2786. doi: 10.1021/ ci200227u

Law V, Knox C, Djoumbou Y, *et al*. 2014. DrugBank 4.0: shedding new light on drug metabolism. *Nucleic Acids Res*, 42:D1091-1097. doi: 10.1093/nar/gkt1068

Li JW-H, Vederas JC. 2009. Drug Discovery and Natural Products: End of an Era or an Endless Frontier? *Science*, 325(5937):161–165. doi: 10.1126/science.1168243

Natural Products Research—Information for Researchers. In: *NCCIH*. https://nccih.nih.gov/grants/naturalproducts. Accessed 12 Aug 2016

Newman DJ, Cragg GM. 2007. Natural Products as Sources of New Drugs over the Last 25 Years. *J Nat Prod*, 70(3):461–477. doi: 10.1021/np068054v

Newman DJ, Cragg GM. 2016. Natural Products as Sources of New Drugs from 1981 to 2014. *J Nat Prod*, 79(3):629–661. doi: 10.1021/acs.jnatprod.5b01055

O'Boyle NM, Banck M, James CA, *et al*. 2011. Open Babel: An open chemical toolbox. *J Cheminform*, 3(1):1–14. doi: 10.1186/1758-2946-3-33.

Omotayo FO, Borokini TI. 2012. Comparative phytochemical and ethnomedicinal survey of selected medicinal plants in Nigeria. *Sci Res Essays*, 7(9):989–999. doi: 10.5897/ SRE09.525

Palermo M, Pellegrini N, Fogliano V. 2014. The effect of cooking on the phytochemical content of vegetables: Effect of cooking on vegetable phytochemicals. *J Sci Food Agric*, 94(6):1057–1070. doi: 10.1002/jsfa.6478

Pence HE, Williams A. 2010. ChemSpider: An Online Chemical Information Resource. *J Chem Educ*, 87(11):1123–1124. doi: 10.1021/ed100697w

Phytochemicals' Role in Good Health. In: *Today's Dietitian*. http://www.todaysdietitian.com/newarchives/090313p70.shtml. Accessed 21 Aug 2016

Prasad S, Tyagi AK. 2015. Traditional Medicine: The Goldmine for Modern Drugs. *Adv Tech Biol Med*,3:e108. doi: 10.4172/2379-1764.1000e108

Sadowski J, Gasteiger J, Klebe G. 1994. Comparison of Automatic Three-Dimensional Model Builders Using 639 X-ray Structures. *J Chem Inf Comput Sci*, 34(4):1000–1008. doi: 10.1021/ci00020a039

Vainio MJ, Johnson MS. 2007. Generating Conformer Ensembles Using a Multiobjective Genetic Algorithm. *J Chem Inf Model*, 47(6):2462–2474. doi: 10.1021/ ci6005646

What Are Phytonutrients? Types and Food Sources. In: *WebMD*. http://www.webmd.com/diet/guide/phytonutrients-faq. Accessed 13 Aug 2016

What Are the Pharmaceutical Sources of Drugs? In: *Healdove*. https://healdove.com/health-care-industry/Where-do-drugs-come-from-Sources-of-Drugs. Accessed 22 Oct 2016

Wishart DS, Jewison T, Guo AC, *et al*. 2013. HMDB 3.0—The Human Metabolome Database in 2013. *Nucleic Acids Res*, 41(D1):D801–D807. doi: 10.1093/nar/gks1065

Zhou J, Xie G, Yan X. 2011.*Encyclopedia of Traditional Chinese Medicines -Molecular Structures*,*Pharmacological Activities*, *Natural Sources and Applications*. Springer, Heidelberg. ISBN: 978-3-642-17733-0

****