Bioinformation
www.bioinformation.net

open access
Database

InPACdb - Indian plant anticancer compounds database

Umashankar Vetrivel*, Nandhitha Subramanian, Kalabharath Pilla

School of Biosciences, Department of Bioinformatics, SRM University, Ramapuram Campus, Chennai, India; Umashankar Vetrivel – Email: vumashankar@gmail.com; *Corresponding author

Received February 19, 2009; Revised April 12, 2009; Accepted June 08, 2009; Published September 05, 2009

Abstract:
Indian Plant Anticancer Compounds Database (InPACdb) is a web-based open access database of phytochemicals. The objective of this initiative is to project the potential of anticancer phytochemicals from Indian pharmacopoeia in an integrated environment. This database is unique in providing comprehensive information covering cancer type, molecular target, 3D Stereochemical structures (tautomers, stereoisomers, conformers and resonance structures) and Chemical descriptors etc. for each entry, enabling effective cheminformatics analysis. The complete dataset of InPACdb encompasses 32 descriptive fields for each entry, and is freely available for download at http://www.inpacdb.org.

Keywords: Biological database; anticancer compounds; Indian plants.

Background:
Ancient Indian medicinal science was based on the natural products of plant origin. Recently, numerous bioactive compounds have been isolated from plant sources and several of them are currently in clinical trials. Plant-derived compounds have been a vital source of varied clinically useful anti-cancer agents: Camptothecin derivatives **[1]**, Topotecan and Irinotecan, Etoposide, derived from Epipodophyllotoxin **[2]**, and Paclitaxel (taxol) **[3]**. Furthermore, other potent molecules include Vinca alkolloids (Vinblastine, Vincristine) **[4]**, Flavopiridol, a semi-synthetic analogue of the chromone alkaloid and Rohitkine, a pyridoindole alkaloid derived from leaves of Ochrosia species **[5]**. Presently, research on anticancer drug development is largely dependent on exploring potential phytochemicals. Indian sub-continent is rich in its diversity of flora, being a tropical country with a large spread of rain forests and river basins. It is floristically rich with about 33 percent of its botanical wealth (over 15,000 species of higher plants) being endemic **[6]**. Hence, analyzing this potential, InPACdb initiative is taken to expose the potent anticancer phytochemicals of Indian Origin in international arena.

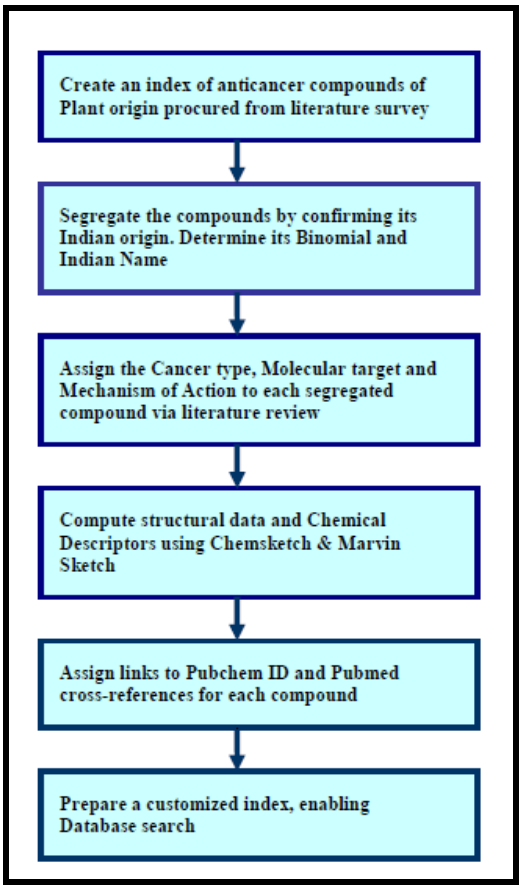

Figure 1: Flowchart showing the methodology implemented in database development

ISSN 0973-2063 (online) 0973-8894 (print)
Bioinformation 4(2): 71-74 (2009)

© 2009 Biomedical Informatics

General information	
Accession number	ACD0022
Compound Name	Curcumin
Structure	
Compound type	Curcuminoid , phenolic
Pubchem Id	SID: 49854349
IUPAC name	(1E,6E)-1,7-bis(4-hydroxy-3-methoxyphenyl)hepta-1,6-diene-3,5-dione
Origin	
Indian name	Haridrā, haldī (Turmeric)
Binomial name	Curcuma longa
Plant image	
Target	
Cancer	leukemia and lymphoma, gastrointestinal cancers, genitourinary cancers, breast cancer, ovarian cancer, head and neck squamous cell carcinoma, lung cancer, melanoma, neurological cancers and variety of cancers
Target	(cyclin D1 and cyclin E), apoptosis (activation of caspases and down-regulation of antiapoptotic gene products), proliferation (HER-2, EGFR, and AP-1), survival (PI3K/AKT pathway), invasion (MMP-9 and adhesion molecules), angiogenesis (VEGF), metastasis (CXCR-4) and inflammation (NF-kappaB, TNF, IL-6, IL-1, COX-2, and 5-LOX
Activity	Interference with the above specified target molecules
Structure information	
Tautomers	download
Resonance	download
Sterioisomers	download
Conformers	download
Mol structure	download
pdb structure	download
SMILES notation	COC1=C(C=CC(=C1)C=CC(=O)CC(=O)C=CC2=CC(=C(C=C2)O)OC)O
Properties Computed from Structure	
Molecular Weight	368.3799 [g/mol]
Molecular Formula	$C_{21}H_{20}O_6$
XLogP	3.1
H-Bond Donor	2
H-Bond Acceptor	6
Refractivity	130.81
Rotatable Bond Count	8
Tautomer Count	2
Exact Mass	368.125988
Topological Polar Surface Area	93.1
Van der Waals surface area (3D)	502.52
Complexity	507

References

Molecular Pharmacology (American Society for Pharmacology and Experimental Therapeutics) (July 2006) 70: 1664–71. Curcumin Inhibits Hypoxia-Inducible Factor-1 by Degrading Aryl Hydrocarbon Receptor Nuclear Translocator: A Mechanism of Tumor Growth Inhibition, Choi, Hyunsung; et al. doi:10.1124/mol.106.025817

Blood, 24 September 2008, Vol. 0, No. 2008, pp. 200805155. Curcumin, a cancer chemopreventive and chemotherapeutic agent, is a biologically active iron chelator. Yan Jiao, John Wilkinson; et al.

Cancer Lett. 2008 Aug 18;267(1):133-64. Epub 2008 May 6. Curcumin and cancer: an "old-age" disease with an "age-old" solution. Anand P, Sundaram C, et al.

Figure 2: Showing an entry from InPACdb

Methodology:

InPACdb's search engine is based on Google's proprietary search technology which utilizes java scripting and custom indexing **[7]**; making search comparatively faster and user friendly. The database presently comprises a dataset of 144 compounds **(Table 1 in supplementary material)** which has been manually sorted from 990 compounds; 200 from National Cancer institute drug list and 690 from Asian anticancer material databank and other 100 compounds from other miscellaneous sources. The Cancer type, Molecular target and Mechanism of action for each segregated compound were documented from elsewhere **[8].** The chemical descriptors and 3D structure for each compound were calculated using Marvinsketch **[9]** and

ISSN 0973-2063 (online) 0973-8894 (print)
Bioinformation 4(2): 71-74 (2009)

© 2009 Biomedical Informatics

Chemsketch **[10]**, respectively. Finally, each compound's entry is indexed via Google Custom Search for database information retrieval. A detailed flowchart summarizing the methodology implemented is shown in **Figure 1**.

InPACdb features:
Ayurveda is an ancient Indian herbal medicine system and is followed till date for anticancer treatment. This medicinal system is also proven to be at par with traditional Chinese Medicinal System. Currently, there are numerous herbal databases that provide information on herbal anticancer compounds: Traditional Chinese Medicine Database System **[11]**, Asian Anticancer material database **[12]** etc. These databases largely concentrate on Chinese Traditional Medicine system and are limited to data like taxonomical classification, pictures of plants and SMILES notation of bioactive compounds. InPACdb is unique in providing comprehensive information on phytochemicals of Indian origin with anticancer activity via 32 descriptive fields; encompassing the basic data like plant image, nomenclature etc. to advanced data like stereochemical properties, chemical descriptors etc. The customized search engine of this database enables wildcard querying, which shall include SMILES string, Cancer type and names of organs, Compounds and plants etc. Each entry in the database is subdivided into 6 parts: General information, origin, target, structure information, Computed Chemical Descriptors and Literature references **(Figure 2)**.

The general information part of the database entry displays the compound's unique InPACdb accession number, followed by its nomenclature, 2D-structure image, compound type (Alkaloid, Flavonoid, Phenolic etc), IUPAC name and an external link to the compound's PUBCHEM ID (if available). Origin section provides detailed information about Indian plant origin from which the compound is isolated. The Indian nomenclature of the plant is given either in Sanskrit (if available) or general name in other Indian languages, subsequently followed by its Binomial name and respective plant's image. Target section gives information on the probable types of cancers on which the compound can act upon, which were curated based on peer-reviewed literature sources. Furthermore, the molecular targets of the respective compounds are also provided along with its bioactivity. Atomic coordinates of each compound entry is provided in the form of PDB and MOL format in structure information section. Moreover, additional information of the compounds like isomers (Sterioisomers, Tautomers, Resonance) and conformers are also given in SDF format, in addition to SMILES notation.

Properties Computed from the structure: this section of the database provides various computed descriptors like Molecular Weight, Molecular Formula, XLogP, H-Bond Donors, H-Bond Acceptors, Refractivity, Rotatable Bond Count, Tautomer Count, atomic Mass, Topological Polar Surface Area, Vander Waals surface area (3D) and Complexity. Finally, reference section lists the citations relevant to the anticancer activity of the respective compounds in the database. Moreover, the multiple links to the citations are also provided (if available).

Utility:
The data presented in InPACdb can be effectively used for Cheminformatics studies like QSAR analysis, pharmacophore search, molecular docking etc. pertaining to anticancer drug discovery. InPACdb also clearly portrays the potential role of google custom search as a Biological information retrieval system.

Caveat and Future Developments:
The database will be extended with more data on molecular interactions, embedded interactive visualization tools and additional chemical descriptors. The users are also welcome to contribute relevant data to the database via email to authors. The dataset and web interface InPACdb shall be upgraded periodically.

References

[1] JA Kepler *et al., J. Org. Chem.* **34**:3853 (1969) [PMID: 5357525]
[2] Utsugi *et al., Cancer Research.* **56**:2809 (1996) [PMID: 8665518]
[3] GM Cragg, DJ Newman, *J. Ethnopharmacol*. **100**:72 (2005) [PMID: 16009521]
[4] IS Johnson *et al., Cancer Res.* **23**:1390 (2001) [PMID: 14070392]
[5] F Arguello *et al., Blood.* **91**:2482 (1998) [PMID: 9516149]
[6] http://www.bioversityinternational.org/publications/Web_version/174/ch06.htm
[7] http://www.google.com/coop/cse
[8] http://pubchem.ncbi.nlm.nih.gov
[9] http://www.chemaxon.com
[10] http://www.acdlabs.com
[11] http://www.cintcm.ac.cn
[12] http://asiancancerherb.info

Edited by P. Kangueane
Citation: Vetrivel *et al.*, Bioinformation 4(2): 71-74 (2009)
License statement: This is an open-access article, which permits unrestricted use, distribution, and reproduction in any medium, for non-commercial purposes, provided the original author and source are credited.

ISSN 0973-2063 (online) 0973-8894 (print)
Bioinformation 4(2): 71-74 (2009)
© 2009 Biomedical Informatics

Supplementary material:

Table 1: List of database categories with number of entries under each category

S. No	Database Category	No. of entities
1	Total number of indian plant species	105
2	Total number of compounds	144
3	Total number of computed structures (PDB, MOL, tautomers, resonance, conformers, sterio-isomers)	800
4	Total number of PDB and Mol structures	288

ISSN 0973-2063 (online) 0973-8894 (print)
Bioinformation 4(2): 71-74 (2009)

© 2009 Biomedical Informatics