Uttarakhand Medicinal Plants Database (UMPDB): A Platform for Exploring Genomic, Chemical, and Traditional Knowledge

Anuj Kumar, Rohit Kumar, Mansi Sharma, Upendra Kumar, M. N. V. Prasad Gajula, and Krishna Pal Singh

Abstract: Medicinal plants are the main natural pools for the primary health care system, ethno-medicine, as well as traditional Indian system of several medicines. Uttarakhand also known as ‘Herbal State’, is a rich source of medicinal plants and traditional medicinal knowledge. A great deal of information about medicinal plants of Uttarakhand is scattered in different forms. Although many medicinal plant databases are available, currently there is no cohesive manually curated database of medicinal plants widely distributed in Uttarakhand state. A comprehensive database has been developed, known as the Uttarakhand Medicinal Plants Database (UMPDB). UMPDB provides extensive information on botanical name, common name, taxonomy, genomic taxonomy id, habit, habitat, location in Uttarakhand, part use, medicinal use, genomic information (including number of nucleotides, proteins, ESTs), chemical information, and scientific literature. Annotated medicinal plants integrated in the current version of the database were collected from the existing books, databases, and available literature. The current version of UMPDB contains the 1127 records of medicinal plants which belong to 153 plant families distributed across 13 districts of Uttarakhand. The primary goal of developing this database is to provide traditional, genomic, and chemical descriptions of the medicinal plants exclusively found in various regions of Uttarakhand. We anticipate that embedded information in the database would help users to readily obtain desired information.

Keywords: Himalaya; Uttarakhand; medicinal plants; herbal drugs; database; NCBI

1. Introduction

The Himalayas are well known for the richest hot spots of biodiversity in the world. This offers immense opportunities in various fields of advanced biological domains that are associated with sustainable life support systems [1,2]. The Indian Himalaya Region harbors approximately 8644 plant species belongs to 1748 families. These plant species are known as medicinal plants and maximum
species have been reported around up to 1800 m altitudinal range [3]. Uttarakhand, a natural environment of the Himalayas also known as ‘Dev Bhoomi’, is rich with majestic natural beauty, great wealth of medicinal plants, and traditional medicinal knowledge. Since prehistoric times, medicinal plants and their products are the primary source of medicine and highly nutritional resources across the globe [4,5]. Medicinal plants are one of the most important components of the forests of Himalaya and are well known for their efficacy in coping with various diseases [6]. Several ancient Indian scriptures—such as the Rig-Veda, Athurveda, and Charka Sanhita—revealed abundant benefits of plants of the Himalayas [7]. Even today, several plant parts—including roots, stems, leaves, etc.—are directly used as traditional medicine by a majority of communities in rural and suburban areas in India as well as rest of the world as there are no adverse side effects, unlike allopathic medicine [8].

According to the World Health Organization (WHO), 80% of world’s population still depends on traditional medicines [9,10]. Plants are a rich and natural source of alkaloids, flavonoids, phenols, chalcones, coumarines, lignans, polyketides, alkanes, alkenes, simple aromatics, peptides, terpenes, and steroids with therapeutic potential [11,12]. Due to these enormous medicinal properties, plants are exclusively used to discover drug-like molecules in the current era of drug discovery [13,14]. Phytomedicines are playing a crucial role in the health management systems in developing countries including India [15–17]. In recent decades, medicinal plants have been getting wide attention and recognition in the global biotechnology market [18–20].

A huge amount of traditional information—such as taxonomy, common names, location, medicinal uses, and part use—about Uttarakhand medicinal plants are scattered in the form of texts. Moreover, the genomic information about these invaluable medicinal plants is not meticulously documented and digitalized. Only recently has a huge effort been made for the development of databases and resources in this domain. Interestingly, within last four to five years, a huge number of public repositories have been established for medicinal plants and their derived phytochemicals. The list includes, but is not limited to MPD3 [21], Phytochemica [22], SerpentinaDB [23], FERN Ethnomedicinal Plant Database [24], Naturally Occurring Plant-based Anti-Cancer Compound-Activity-Target Database (NPACT) [25], SuperNatural [26], Herb Ingredients’ Targets (HIT) [27], Cancer Resource-dataset of compound-target interactions [13], Traditional Chinese medicine database (TCMD) [28], China natural products database (CNPD) [29], Chinese herb constituents database (CHCD) [30], Bioactive plant compounds database (BPCD) [30], Nuclei of Bioassays, Ecophysiology and Biosynthesis of Natural Products Database (NuBBEDB) [31,32], Northern African Natural Products Database (NANPDB) [33], Highly potent and diverse natural product library from African medicinal plants (AfroDb) [34], Molecular modeling of potential anticancer agents from African medicinal plants (AfroCancer) [35], Database of Volatile Organic Compounds in Cancer (VOCC) [36], and Nutrient Use Efficiency (NtUE) Web-Resource [37]. However, a similar comprehensive database for Uttarakhand medicinal plants is still not available despite a recent database development by the ENVIS center on Forestry with very limited scope and information.

Keeping in view of the need for medicinal values of the plants widely distributed in 13 districts of Uttarakhand, a comprehensive database called Uttarakhand Medicinal Plants Database (UMPDB) is developed. This database is based on available traditional and genomic information manually curated from existing books, databases, and literature combined with the data available in PubMed. The current version of UMPDB contains about 1127 medicinal plants species, which are widely distributed in different regions of Uttarakhand state. The taxonomic and genomic information embedded in this database is manually derived from different literature published in reputed journals and the National Center for Biotechnology Information (NCBI). Whereas, the chemical information is obtained from PubChem. The database contains the following entries for individual plant species including: botanical name, common name, taxonomy, genomic taxonomy id, habit, habitat, location in Uttarakhand, part use, medicinal use, genomic information including number of nucleotides, proteins, ESTs (available on NCBI), chemical information (available on PubChem), and scientific literature (as available on PubMed), all of these entries have been curated manually and annotated. Availability
of this comprehensive database may enable direct use of plants extracted natural compounds in in-silico drug discovery. This study may help to understand and design herbal drugs against various dangerous diseases including cancer, HIV, and diabetes. Compilation and accessibility of comprehensive information on medicinal plants and their natural chemical compounds embedded in UMPDB can be a source of considerable advantage to conservationists, resource managers, nature enthusiasts, as well as industry.

2. Results and Discussion

UMPDB is a comprehensive database. It provides traditional, genomic, and chemical information on medicinal plants widely distributed across the Uttarakhand state. In the current release, we have compiled 1127 plant species records belonging to 153 plant families. The extensive information of medicinal plants and their chemical compounds were manually curated and annotated using literature based survey and public repositories. Each entry embedded in this comprehensive database has been arranged in a plant centric manner for easy access and retrieval.

2.1. PlantCentric Portal

Each entry embedded in database provides the following extensive information on Uttarakhand medicinal plants: (i) botanical name; (ii) family name; (iii) common name; (iv) taxonomy; (v) genomic taxonomy id; (vi) habit; (vii) habitat; (viii) location in Uttarakhand; (ix) part use; (x) medicinal use; (xi) number of nucleotides; (xii) proteins and ESTs (available on NCBI); (xiii) with an active hyperlink which provides chemical information (available on PubChem); and (xiv) scientific literature (available on PubMed) as shown in Figure 1.

2.2. Data Access

The information for medicinal plants available in the database could be accessed through keyword search tab by assigning the botanical name (i.e., *Brachiaria ramosa* L., *Commelina benghalensis* L., *
Berberis aristata, Abutilon indicum, Acacia Arabica, etc.), common name (i.e., Jadwar, Hairy Okra Kamlya, Larkspur, Lovely Dendrobium, Dhatoora, Khaja, etc.), location in Uttarakhand (i.e., Kumaun, Bhabar Tract of Garhwal Himalaya, Kedarnath valley, Haridwar, Dehradun, etc.), and by clicking on the ‘search’ button. After submitting the keyword, a page will open presenting a list of medicinal plants with details that contain the botanical name, common name and the location in Uttarakhand. By further clicking on ‘View Details’, the web page will direct the user to complete data details of each particular medicinal plant. This scroll down based page provides extensive information that includes a summary of plant including: botanical name, common name, family, taxonomy, genomic taxonomy id, habit, habitat, location in Uttarakhand, part use, medicinal use, number of nucleotides, proteins and ESTs (available on NCBI), chemical information (available on PubChem), and scientific literature (available on PubMed). Alternatively, the search option of the UMPDB allows the users to go through the entire list of plants embedded in databases. Through full list of plants available in the database, a user can select specific plants and can search further for more detailed traditional, genomic, and chemical information. The steps involved in keyword based search with an example is shown in Figure 2.

The UMPDB database has a user-friendly entry point for each medicinal plants record embedded in database. Each plant entry of the UMPDB database has a corresponding taxonomy id, genomic information (number of nucleotides, proteins and ESTs, etc.), chemical information and up to date literature information are linked to parental databases, i.e., NCBI taxonomy portal, GenBank, PubChem, and PubMed respectively. Using external links, users can get more detailed information (if any) for genomic and chemical information, and for updated literature. In addition, high quality graphics for each plant on the UMPDB database are provided. Standardized terminologies are used throughout, which are validated and processed, making a comprehensive database for interoperability of medicinal plant data.

2.3. Comparison of UMPDB with ENVIS Medicinal Plants: Uttarakhand and Database

ENVIS Uttarakhand specific medicinal plant database listed the plants along with their traditional knowledge based details like taxa, local name, parts used, and medicinal use. This database provides the details of medicinal plants distributed only in three regions of Uttarakhand including Alaknanda valley, Chamoli, and Jaunsar-Bawar. This database does not cover many aspects of medicinal plants found in Uttarakhand. Whereas, UMPDB database is a literature curated repository for 1127 medicinal plants (Figure 3) belongs to 153 plant families (Figure 4) and widely distributed in different regions of Uttarakhand state. It is a unique repository that not only lists the medicinal plants according to their family and location, but it also provides data on various aspects like traditional, genomic, chemical, and current literature.
Figure 2. An illustration of the home page with search options and search summary. The specific example shows traditional information of plants, genomic information on NCBI, chemical information, and linked reference articles on PubMed.
Figure 3. Cloud based representation of plants species embedded in UMPDB.

Figure 4. Cloud based representation of plant families distributed across Uttarakhand. The cloud shape was generated as per geographical map of Uttarakhand.

3. Materials and Methods

In order to build a comprehensive repository for Uttarakhand medicinal plants, the data based on traditional information including botanical name, family name common name, taxonomy,
genomic taxonomy id, habit, habitat, location in Uttarakhand, part use, and medicinal use were compiled from existing literature and databases. Several research articles, five books, and two web-resources were manually curated to obtain the information on medicinal plants found in several regions of Uttarakhand along with their traditional knowledge based information. Books and web-resources utilized for manual curation of entries embedded in UMPDB are listed: ‘Medicinal Plants of Uttarakhand (Volume 1, 2, and 3)’ [3], ‘Herbal Medicine’ [38], Edible Plants of North West Himalaya (Uttarakhand)’ [39], ENVIS Medicinal Plants: Uttarakhand (http://www.frienvis.nic.in/Database/Medicinal-Plants-Uttarakhand_2150.aspx), and Indian Medicinal Plant Database (http://www.medicinalplants.in/). Genomic information—such as the number of nucleotides, proteins, and ESTs—were curated and linked to NCBI (http://www.ncbi.nlm.nih.gov). PubChem (https://pubchem.ncbi.nlm.nih.gov/) was searched for data on each plant embedded in UMPDB in order to retrieve chemical compounds information. Whereas, to address degeneracy in plant names and up to date literature, PubMed (https://www.ncbi.nlm.nih.gov/pubmed/) was searched individually for each entry of UMPDB.

The UMPDB database set up is based on a three-tier architecture concept using Apache/PHP/MySQL on a linux platform [40,41]. An integrated system driven through MySQL (5.6.21), and PHP (5.6.24) was developed to handle the storage of annotated 1127 plant entries in the database. The graphical user interface (GUI) was designed to be user friendly for data query and extraction, and was also tested in all major browsers (Chrome, Firefox, Safari, and Internet Explorer) and on different OS platforms. A flowchart depicting the steps involved in preparation of UMPDB and acquisition of data is represented in Figure 5.

**Figure 5.** Flow chart depicting steps involved in preparation of UMPDB and acquisition of the stored data.
4. Conclusions

Utilizing natural resources for drug development is an important and developing area in the current era of biotechnology. Biotechnology based tools—including transcriptomics, genomics, proteomics, and metabolomics—have recently been extended to medicinal plants; however, these tools still remain underdeveloped when compared to cereal crops. While medicinal plants have been used to cope with different diseases for thousands of years, the UMPDB database can serve as a comprehensive non-redundant catalogue for studies based on medicinal plants available in the Uttarakhand Region of the Himalayas. It contains manually curated entries with an adequate level of stringency to select only those which have medicinal importance. This resource will help in providing new solutions to ethnobotanists and computational biologist/bioinformaticians who are aiming at more sustainable use of medicinal plants to discover novel targets for dangerous diseases. UMPDB database is a compendium for researchers, academicians, school students, ethnobotanists, conservationists, resource managers, nature enthusiasts to share genomic knowledge of medicinal plants.

Acknowledgments: Authors would like to give their sincere thanks to Prashanth Suravajhala, Birla Institute of Scientific Research (BISR), Jaipur for proof-reading the final version of manuscript. We are also thankful to RD Gaur, Hon. Emeritus, HNB Garhwal University, Srinagar and A.K. Mishra, Principal Scientist, IARI, New Delhi for cross-checking the taxonomic and genomic data.

Author Contributions: A.K. and K.P.S. conceived the idea. A.K. and M.S. outlined and wrote the manuscript. R.K. and M.N.V.P. developed all APIs and database search interfaces. M.S., U.K., and M.N.V.P. manually curated and annotated the database entries. All authors have reviewed and agreed upon the final version of the manuscript.

Conflicts of Interest: The authors declare no conflict of interest whatsoever.

References

1. Mani, M.S. Biogeography of the Himalaya. In Ecology and Biogeography of India; Springer: The Hague, The Netherlands, 1978; pp. 664–681.
2. Samant, S.S.; Dhar, U.; Palni, L.M.S. Medicinal Plants of Himalaya, Diversity, Distribution and Potential Values; Gyanodaya Prakashan: Nainital, India, 1998.
3. Singh, K.P.; Kumar, A.; Kumar, U. Medicinal Plants of Uttarakhand; Astral International: New Delhi, India, 2017; Volume 1–3.
4. Singh, J.S.; Singh, S.P. Forest vegetation of the Himalaya. Bot. Rev. 1987, 53, 80–192. [CrossRef]
5. Gangwar, R.S.; Joshi, B.D. Some medicinal flora in the riparian zone of river Ganga at Saptrishi, Haridwar, Uttarakanchal. Himal. J. Environ. Zool. 2006, 20, 237–241.
6. Nautiyal, S.; Rajan, K.S.; Shibasaki, R. Environmental conservation vs compensation: Explorations from the Uttarakhand Himalaya. Environ. Inform. Arch. 2004, 2, 24–35.
7. Sharma, J.; Gaur, R.D.; Paliul, R.M. Conservation status and diversity of some important plant in the Shiwalik Himalaya of Uttarakhand, India. Int. J. Med. Aromat. Plants 2011, 1, 75–82.
8. Gangwar, K.K.; Deepali, G.R.; Gangwar, R.S. Ethnomedicinal plant diversity in Kmaun Himalaya of Uttarakhand, India. Nat. Sci. 2010, 8, 66–78.
9. Mukhergee, T.K. Protection of Indian traditional knowledge: In Ethno Medicinal Plants; Trivedi, P.C., Sharma, N.K., Eds.; Poiner Publishers: Jaipur, India, 2004; pp. 18–33.
10. Naini, V.; Mamidala, E. An ethnobotanical study of plants used for the treatment of diabetes in the Warangal district, Andhra Pradesh, India. Biolife 2013, 1, 24–28.
11. Panchangam, S.S.; Vahedi, M.; Megha, M.J.; Kumar, A.; Raithatha, K.; Suravajhala, P.; Reddy, P Saffron ’omics’ : The challenges of integrating omic technologies. Aevicenna J. Phytomed. 2016, 6, 604–620. [PubMed]
12. Jee, B.; Kumar, S.; Yadav, R.; Singh, Y.; Kumar, A.; Sharma, N. Ursolic acid and carvacrol may be potential inhibitors of dormancy protein small heat shock protein16.3 of Mycobacterium tuberculosis. J. Biomol. Struct. Dyn. 2016. [CrossRef] [PubMed]
13. Ahmed, J.; Meinel, T.; Dunkel, M.; Murquequito, M.S.; Adams, R.; Blasse, C.; Eckert, A.; Preissner, S.; Preissner, R. CancerResource: A comprehensive database of cancer-relevant proteins and compound interactions supported by experimental knowledge. Nucleic Acids Res. 2011, 39, D960–D967. [CrossRef] [PubMed]
14. Katiyar, C.; Gupta, A.; Kanjilal, S.; Katiyar, S. Drug discovery from plant sources: An integrated approach. *Ayu* 2012, 33, 10–19. [CrossRef] [PubMed]
15. Modak, M.; Dixit, P.; Londhe, J.; Ghaskadbi, S.; Devasagayam, T.P.A. Indian herbs and herbal drugs used for the treatment of diabetes. *J. Clin. Biochem. Nutr.* 2007, 40, 163–173. [CrossRef] [PubMed]
16. Gordon, M.; Newman, D.J. Natural products: A continuing source of novel drug leads. *Biochem. Biophys. Acta* 2013, 1830, 3670–3695.
17. Sen, S.; Chakraborty, R. Revival, modernization and integration of India traditional herbal medicine in clinical practice: Importance, challenges and future. *J. Tradit. Complement. Med.* 2017, 7, 234–244. [CrossRef] [PubMed]
18. Zhou, L.G.; Wu, J.Y. Development and application of medicinal plant tissue cultures for production of drugs and herbal medicinals in China. *Nat. Prod. Rep.* 2006, 23, 789–810. [CrossRef] [PubMed]
19. Thatoi, H.; Patra, J.K. Biotechnology and pharmacological evaluation of medicinal plants: An overview. *J. Herbs Species Med. Plants* 2011, 17, 214–218. [CrossRef]
20. Shakya, A.K. Medicinal plants: Future source of new drugs. *Int. J. Herb. Med.* 2016, 4, 59–64.
21. Mumtaz, A.; Ashfaq, U.A.; ul Qamar, M.T.; Anwar, F.; Gulzar, F.; Ali, M.A.; Saari, N.; Pervez, M.T. MPD3: A useful medicinal plants database for drug designing. *Nat. Prod. Res.* 2017, 31, 1228–1236. [CrossRef] [PubMed]
22. Pathania, S.; Ramakrishnan, S.M.; Bagler, G. Phytochemica: A platform to explore phytochemicals of medicinal plants. *Database* 2015. [CrossRef] [PubMed]
23. Pathania, S.; Ramakrishnan, S.M.; Randhawa, V.; Bagler, G. SerpentinaDB: A database of plant-derived molecules of *Rauvolfia serpentina*. *BMC Complement. Altern. Med.* 2015, 15, 262. [CrossRef] [PubMed]
24. Thakur, S.B.; Ghorpade, P.N.; Kale, M.V.; Sonawane, K.D. FERN ethnomedicinal plant database: Exploring fern ethnomedicinal plants knowledge for computational drug discovery. *Curr. Comput. Admed Drug Res.* 2015, 11, 266–271. [CrossRef]
25. Mangal, M.; Sagar, P.; Singh, H.; Raghava, G.P.; Agrawal, S.M. NPACT: Naturally occurring plant-based anti-cancer compound-activity-target database. *Nucleic Acids Res.* 2013, 41, D1124–D1129. [CrossRef] [PubMed]
26. Dunkel, M.; Fullbeck, M.; Neumann, S.; Preisser, R. SuperNatural: A searchable database of available natural compounds. *Nucleic Acids Res.* 2006, 34, D678–D683. [CrossRef] [PubMed]
27. Ye, H.; Ye, L.; Kang, H.; Zhang, D.; Tao, L.; Tang, K.; Liu, X.; Zhu, R.; Liu, Q.; Chen, Y.Z.; et al. HIT: Linking herbal active ingredients to targets. *Nucleic Acids Res.* 2011, 39, D1055–D1059. [CrossRef] [PubMed]
28. Shen, J.; Xu, X.; Cheng, F.; Liu, H.; Luo, X.; Shen, J.; Chen, K.; Zhao, W.; Shen, X.; Jiang, H. Virtual screening on natural products for discovering active compounds and target information. *Curr. Med. Chem.* 2003, 10, 2327–2342. [CrossRef] [PubMed]
29. Ehrman, T.M.; Barlow, D.J.; Hylands, P.J. Phytochemical databases of Chinese herbal constituents and bioactive plant compounds with known target specificities. *J. Chem. Inf. Model.* 2007, 47, 254–263. [CrossRef] [PubMed]
35. Ntie-Kang, F.; Nwodo, J.N.; Ibezim, M.; Simoben, C.V.; Karman, B.; Ngwa, V.F.; Sippl, W.; Adikwu, M.U.; Mbaze, L.M. Molecular modeling of potential anticancer agents from African medicinal plants. *J. Chem. Inf. Model.* **2014**, *54*, 2433–2450. [CrossRef] [PubMed]

36. Agarwal, S.M.; Sharma, M.; Fatima, S. VOCC: A database of volatile organic compounds in cancer. *RSC Adv.* **2016**, *6*, 114783–114789. [CrossRef]

37. Kumar, A.; Pandeya, A.; Malik, G.; Kumari, P.H.; Kumar, S.A.; Sharma, M.; Gahlaut, V.; Gajula, M.N.V.P.; Suravajhala, P.; Singh, K.P.; et al. A web-resource for nutrient use efficiency related genes, QTLSs, and microRNA in important cereals and model plants. *bioRxiv* 2017. [CrossRef]

38. Sood, S.K.; Kumari, P.; Thakur, R.; Bassi, S.K.; Thakur, A. *Herbal Medicine*; Pointer Publishers: Jaipur, India, 2015.

39. Shah, R. *Edible Plants of North West Himalaya (Uttarakhand)*; Bishen Singh Mahinder Pal Singh: Dehradun, India, 2015.

40. Chaduvula, P.K.; Bonthala, V.S.; Manjusha, V.; Siddiq, E.A.; Polumetla, A.K.; Prasad, G.M.N.V. CmMDb: A versatile database for Cucumis melo microsatellite markers and other horticulture crop research. *PLoS ONE* **2015**, *10*, e0118630.

41. Bhawna, V.S.; Gajula, M.N.V. *PvTFDB*: A *Phaseolus vulgaris* transcription factors database for expediting functional genomics in legumes. *Database* **2016**. [CrossRef] [PubMed]