manuscripts should not be exposed to light for long duration. Some of these traditional materials act as a catalyst in the process of deterioration on palm leaf manuscripts. Thus, it can be concluded that in Indian palm leaf illustrations, mostly mineral colours have been used.


1. Kesavan, B. S., *The Book in India: A Compilation*, National Book Trust, New Delhi, 1985.
2. Dharmalingam, U. K., *Black and White in Indian Typography*, Association Typographique Internationale, Hongkong, Hongkong, 2012, pp. 1–11.
3. Naik, B. S., *Typography of Devanagari – Volume 1*, Directorate of Languages, Mumbai, 1971.
4. Freemen, R., Turning over old leaves: palm leaves used in South Asian manuscripts. *The Book and Paper Group Annual*, Amerian Institute of Conservation, Minneapolis, Minnesota, 2005, pp. 99–102.
5. Agarwal, O. P., *Conservation of Manuscripts and Paintings of South-East Asia*, Butterworth, London, UK, 1984.
6. Agus, N. A., Mydin, M. O. and Ulang, N. H., Biomimcry of palm tree leaves form and pattern on building form. *EDP Sciences*, Web of Conferences 3, 01024, 2014, 1–6.
7. Akca, D. and Gruen, A., Re-sequencing a historical palm leaf manuscript with boundary-based shape descriptors. In *International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, 19th CIPA International Symposium, Antalya, Turkey, 2003, vol. XXXIV, part 5/C15, pp. 55–60.
8. Das, J. P., *Chitra Pothi: Illustrated Palm Leaf Manuscripts from Orissa*, Niyogi Books, New Delhi, 2007.
9. Kumar, D. U., Sreekumar, G. V. and Athvankar, U. A., Traditional writing system in southern India – palm leaf manuscripts. *Design Thoughts*, 2009, pp. 2–7.
10. Patnaik, D. P., *Palm-Leaf Etchigs of Orissa*, Abhinav Publications, New Delhi, 1989.
11. Sahoo, J., A selective review of scholarly communications on palm leaf manuscripts. *Libr. Philos. Pract. (e-journal)*, 2016, 1397.
12. Dyke, Y. V., Sacred leaves: the conservation and exhibition of early Buddhist manuscripts on palm leaves. *The Book and Paper Group Annual*, American Institute of Conservation, Washington DC, USA, 2009, vol. 28, pp. 83–97.
13. Guy, J., *Palm Leaf and Paper: Illustrated Manuscripts of India and South-east Asia*, National Gallery of Victoria, Melbourne, Australia, 1982.
14. Mactaggart, P. and Mactagggart, A., *A Pigment Microscopist's Notebook*, Mactaggart, Chard, UK, 1985.
15. Spring, M. and Grout, R., The blackening of vermillion: analytical study of the process in paintings. *Natl. Gallery Tech. Bull.*, 2004, **23**, 50–61.
16. Frison, G. and Brun, G., Lapis lazuli, lazurite, ultramarine 'blue', and the colour term 'azure' up to the 13th century. *J. Int. Colour Assoc.*, 2016, **16**, 41–55.


ACKNOWLEDGEMENTS. We thank Dr B. R. Mani (National Museum Institute, New Delhi) for support. We also thank Dr Tanja Bayerova, Institute of Conservation, University of Applied Arts Vienna for constant guidance in drafting the scientific analysis for my research work during my fellowship period (Ernst-Mach Scholarship) at Vienna. We are also grateful to Prof. Rudolf Erlach, Institute of Art and Technology, Department of Archaeometry, University of Applied Arts Vienna for support while carrying out the SEM-EDX; Dr Satish Pandey and Dr Aditya Kanth (National Museum Institute, New Delhi) for invaluable inputs and support.

Received 4 December 2017; revised accepted 12 October 2019

doi: 10.18520/cs/v118/i2/285-292

# Mtb-HID: a unified database of host–pathogen interaction for various *Mycobacterium tuberculosis* strains

**Kumud Pant*, Devvret, Neema Tufchi, Ajay Pandaya and Ashish Thapliyal**

Department of Biotechnology, Graphic Era Deemed to be University, Dehradun 248 002, India

**Databases developed till now for studying human–*Mycobacterium tuberculosis* (MTB) host–pathogen interactions are scarce and pertain to specific strains of pathogen. Therefore, in the present study, a database (Mtb-HID) has been developed to serve as a unified host–pathogen interaction platform with information on interaction between various strains of MTB and humans. Mtb-HID is different from other databases since it houses information about multiple strains and is an extensive integrated repository of experimental host–pathogen protein–protein interaction (HP-PPIs) data imported from various public databases and derived from interolog-based HP-PPIs. Currently, Mtb-HID hosts records of interaction between five MTB strains and human host. It is freely accessible at http://www.pantlab.co.in/mtb-hid/.**

**Keywords:** Database, host–pathogen interaction, humans, *Mycobacterium tuberculosis*.

*MYCOBACTERIUM TUBERCULOSIS* (MTB) is the causative agent of tuberculosis (TB), one of the most devastating diseases. Around 10 million people are infected with TB worldwide, of which 2.7 million cases belong to India alone[1]. Though extensive efforts have been made by the World Health Organization (WHO), it is difficult to predict whether the disease will be eradicated by 2050 or not. The literature suggests that various strains of MTB show different patterns of infection in the host[2]. Virulence and pathogenesis of MTB lead to disease development in the host. Several mechanisms of action from the pathogen as well as host are involved in establishing the disease. Molecular interactions between host cells and specific microbial products play an important role in gaining pathogenesis. As a result, various changes occur in the host cell functioning, thus promoting the pathogen to invade the host cell and tissue[3].

MTB is a member of the Mycobacteriaceae family. It is a Gram-positive, acidfast bacterium[4]. It is considered to be acid fast because of its cell wall which is made up of lipids. It resists the Gram stain and thus acid is used for staining process[5]. The genome of MTB comprises approximately 4.4 million base pairs, containing around 4000 genes with high GC content[6]. Though the genome is highly conserved due to the co-evolutionary process,

*For correspondence. (e-mail: pant.kumud@gmail.com)

strains of MTB from different geographical regions may possess different levels of virulence, thus resulting in varied epidemiological dominance[7,8]. An understanding of the key characteristics of the pathogen like virulence, etiology and nature of the molecular interactions with its host is essential to prevent the emergence of disease.

There are various TB bacterium strains that exist globally and vary in their ecological niche. Studies suggest that the most widespread lineage of the etiological agent of human TB comprises both globally distributed (generalists) and geographically restricted (specialists) sublineages. The globally distributed sublineages infect a broader human host range, whereas the restricted ones tend to infect only specific populations[9,10]. Strains of MTB have co-evolved with the ancient hominids and are still evolving. Based on *in silico* genome-wide analysis, MTB strains are clustered into seven phylogenetic lineages that are adapted to humans and each lineage is associated with specific geographical locations[11]. Lineage 1 is Southeast Asian and Oceanian clades, lineage 2 originated in East Asia comprising 'Beijing' strains and is highly prevalent. Lineage 2 has been reported for its repeated association with drug resistance[12]. Lineage 3 represents Central Asian (CAS)/Delhi clades, distributed in some countries of East Africa and widely spread in the Indian Subcontinent. Lineage 4 covers the EuroAmerican strains and is commonly found in America, Europe, the Middle East and some parts of Africa. Lineages 2 and 4 are the most globally distributed of all human MTB lineages. Lineages 5 and 6 are localized to West Africa, and lineage 7 has recently been observed in Ethiopia or in recent Ethiopian emigrants[13].

Host–pathogen protein–protein interactions (HP-PPIs) are often involved in adhering to the target, invading the immune defences of the host, colonization, altering the molecular events of the host, replicating and persisting within the host. Studying HP-PPIs uncovers the nature of infection and helps researchers develop armamentarium against infectious diseases[14,15]. There are two experimental approaches for detecting protein interactions, i.e. (i) co-complex methods such as co-immunoprecipitation coupled with mass spectrometry, and (ii) binary approaches such as luminescence-based mammalian interactome mapping and yeast two-hybrid system[16]. These processes are expensive and time-consuming. Therefore, computational methods have been developed to improve the accuracy and efficiency in identifying interactions between host and pathogen proteins. These include phylogenetic profile, interolog method, gene neighbour and gene cluster methods[17].

Readily available data on HP-PPIs are scattered and cannot be obtained from a single source. Thus there is a need for incorporating all information regarding protein–protein interactions at a central shareable place. Realizing the need for developing a more comprehensive resource and up-to-date and organized central repository, the Mtb-HID (*Mycobacterium tuberculosis* and human protein–protein interaction database) has been developed. It is a manually curated database of experimental data as well as data obtained from HP-PPIs between human host and MTB based on interolog method.

The strategy can be summarized as follows: (a) Collection of experimental HP-PPIs data. (b) Application of interolog method to determine the unknown HP-PPIs of different strains of MTB. (c) Curation of database. (d) Removal of redundancy. (e) Creating a user-friendly interface for the list of interacting proteins of various MTB strains. (f) Direct access to UniProt for further information about any protein.

Various databases on host–pathogen interactions have been developed by several research groups as summarized below. However, the Mtb-HID (http://www.pantlab.co.in/mtb-hid/) is a dedicated database that provides information about host–pathogen interaction of various strains of MTB not incorporated till now by the existing repositories.

A concise account of the available resources is presented as: (1) PHI-Base (http://www.phi-base.org/) is a host–pathogen interaction database which carries information on interactions between various infectious species and their host[18]. (2) PHIDIAS, The Pathogen–Host Interaction Data Integration and Analysis System (http://www.phidias.us/), is a server which contains information curated from the literature, gene expression data and host–pathogen interaction data for pathogens with high priority in public health and biological defence[19]. (3) PATRIC database, the Pathosystems Resource Integration Center database (https://www.patricbrc.org/), provides several analysis tools to support biomedical research on infectious diseases. Besides various infectious agents, it also provides information on host–pathogen interaction of MTB H37Rv strain (reference strain) with human host[20]. (4) PATH (obsolete) was the first specialized database for host–pathogen interaction on MTB[17]. (5) The generalized HPI databases contain only the information on the laboratory reference strain H37Rv[17,21–23].

Therefore, Mtb-HID has incorporated information on strains not covered in the currently available databases. The comparative analysis and statistics shown in Table 1 will provide a clear picture for the need of a new database.

The HP-PPIs data obtained from experimental as well as interolog method have been manually curated in Mtb-HID. The proteome data of the studied MTB strains as well as *Homo sapiens* have been collected from UniProt Proteome database and used for primary analysis (https://www.uniprot.org/proteomes/)[23]. The unknown HP-PPIs were identified through Biana Interolog Prediction Server (BIPS; http://sbi.imim.es/web/index.php/research/servers/bips)[24]. The proteins of MTB and *H. sapiens* were considered homologous if they shared at least 30% identity and 80% sequence overlap with

**Table 1.** Human–*Mycobacterium tuberculosis* (MTB) specific host–pathogen protein–protein interactions (HP-PPIs) in different databases

| Database/server | Data of strains | Number of HP-PPIs |
|---|---|---|
| PHI-Base | Information on interactions between various infectious species and their hosts.<br>For MTB, only the strain H37Rv was studied. | Sixteen interactions in case of MTB H37Rv. |
| PHIDIAS | Information on interactions of 42 pathogens with their hosts.<br>A centralized source to search, compare and analyse integrated genome sequences, conserved domains and gene expression data. | Phinet PHIDIAS does not contains PHI network information for MTB. |
| PATRIC | Information of various infectious agents.<br>For MTB H37Rv strain was studied. | Twenty-four interactions of MTB H37Rv. |
| PATH | Host–pathogen interaction-based on sequence motifs.<br>Dedicated to MTB H37Rv strain. | 118 interactions of MTB H37Rv. |
| Mtb-HID | HPI-based on experimental and interlog based method protein–protein interaction study.<br>Dedicated to strains of MTB.<br>Currently contains HP-PPIs of strains H37Rv, H37Ra, ATCC 35801 / TMC 107/Erdman, ATCC 35801 and CAS_NITR204. | Total 3198 interactions (H37Rv-698, H37Ra-698, Erdman – 698, ATCC 35801–676. CAS_NITR204–428). |

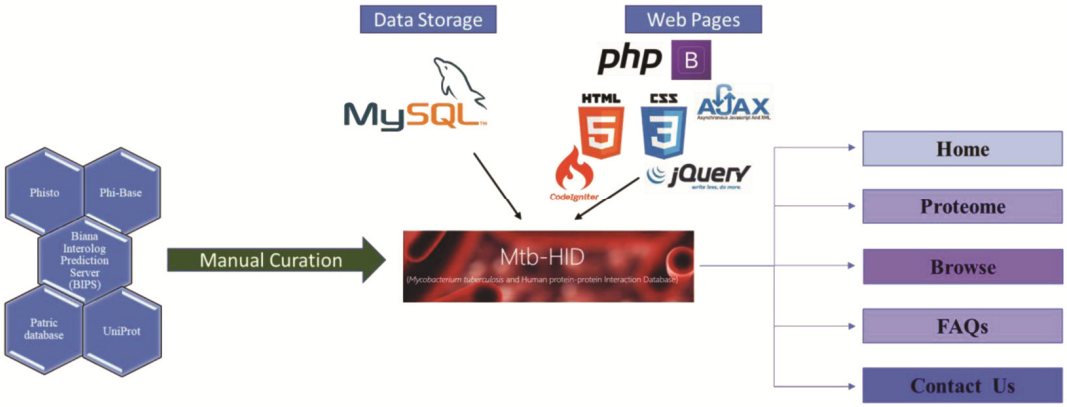


**Figure 1.** Architecture of *Mycobacterium tuberculosis* host–pathogen interaction database (Mtb-HID).

E-value less than $1 \times e^{-10}$. In this study we have set homology condition using filters, i.e. Blast and Joint *E*-value $\leq 5e^{-10}$ % and joint-identities ≥80, query sequence coverage ≥80 and template sequence coverage ≥90. The protein sequences were also filtered by predictions based on cluster of orthologous genes defined by eggNOG[25]. These protein–protein interactions were validated through the String database version 10.5 (https://string-db.org). This database shows a wide range of information on protein–protein interactions retrieved through text mining, homology prediction and other techniques. Interolog-based mapping approach was used to generate the complete interactome of all the studied strains of MTB with reference to humans[26].

Mtb-HID was developed as a relational database using MySQL. User interface of Mtb-HID has been developed on the Bootstrap framework that makes it the mobile-first website. The server scripting and back end were developed in CodeIgniter framework MVC architecture. Mtb-HID runs on a Linux system using an Apache 2 web server. The functionality and internal structure of Mtb-HID were tested using white-box and black-box respectively. After successful testing, the system has been hosted on http://www.pantlab.co.in/mtb-hid/. Ajax, jQuery JSON tools and technique were used to make the search and advanced search operations more user-friendly. The publicly accessible web server has three-tier architecture (Figure 1). Figure 2 shows the web interface of Mtb-HID. Users can retrieve HP-PPIs by searching for proteins in UniProt query format (e.g. P9WPU7) using bootstrap form (front end) of the database. Query is processed through PHP (middle tier) to extract data from

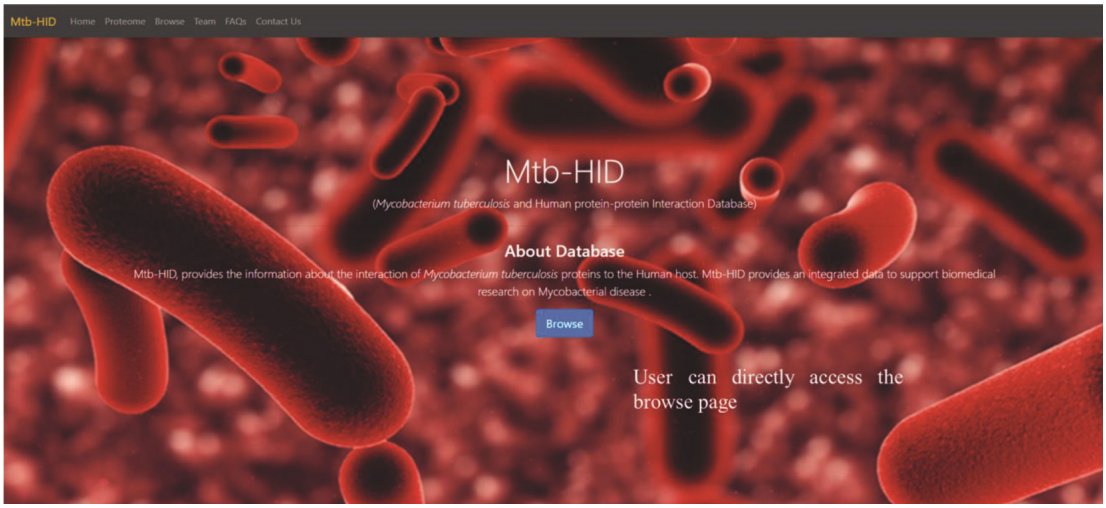


**Figure 2.** Screenshot of the web interface of Mtb-HID.


**Figure 3.** Screenshot of the browse page of Mtb-HID that allows users to access the host–pathogen interaction (HPI) between different *Mycobacterium tuberculosis* strains and human host.

the relational database. The result related to a query is then presented to the users.

The simple browse page allows users to select the MTB strains. As the strains get clicked, the detailed table on host–pathogen interacting proteins is displayed. The strain-specific interactions can be studied through the browse page (Figure 3). Currently, the web server hosts information on five strains, i.e. H37Rv, H37Ra, ATCC 35801/TMC 107/Erdman, ATCC 35801 and CAS_NITR204.

The interaction of any MTB protein to the human host protein can be searched through the advanced search page of the website by entering any MTB protein name or its UniProt ID. A list of interacting proteins matching to the query is shown. The query will be searched in the dataset of all strains. For example, if one searches for

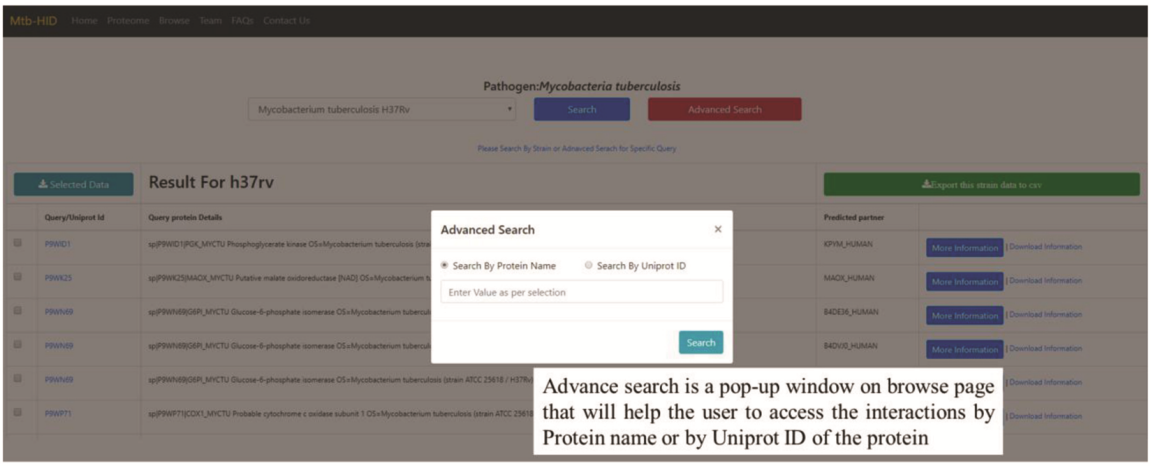


**Figure 4.** Screenshot of advanced search of Mtb-HID that allows users to access HPI (*M. tuberculosis* strains with human host) of the desired protein using protein name or its UniProt ID on a single page.

cytochrome *c* oxidase subunit 1 using UniProt ID P9WP71 (COX1_MYCTU), totally 325 results are retrieved, including Q0Z7F1_Human (cytochrome *c* oxidase subunit 2) and Q2HJS3_HUMAN (cytochrome *c* oxidase subunit 3). Details of each ID can be obtained by clicking on the result; the hyperlinks will navigate the page to UniProt Knowledgebase (Figure 4)[27].

In addition, an FAQ section has been incorporated in the web server which will help users better understand the host–pathogen interaction analysis. Users can directly contact the system administration for any query through the 'Contact Us' page.

TB is preventable and curable, but the rise of drug resistance, i.e. MDR (multi-drug resistance) and XDR (extensive drug resistance) is a major concern in the treatment of the disease. Studies suggest that an individual can harbour multiple strains of MTB[28]. It has also been reported that different strains show various patterns of infection in the host[2,12]. The initial research on drug discovery for TB was done on MTB laboratory strains H37Rv or at H37Ra. Detailed analysis of other strains is also required and needs exploration in the field of host–pathogen interactions. The study of diversified patterns of infection can be obtained through HP-PPIs. This is particularly important for drug discovery to identify potential drug targets present in all the strains of MTB, so that one drug can act on all the strains efficiently. The existing host–pathogen interaction servers house data mostly of H37Rv strain of MTB which is a laboratory strain. Mtb-HID is different since comprehensive data on the HPPPIs of five strains have been made available. We plan to extend the database by incorporating all interactions from various other MTB strains also in the future.

*Conflicts of interest:* The authors declare no conflict of interest.


1. WHO, Global tuberculosis report 2018, World Health Organization, Geneva, 2018.
2. Manabe, Y. C. *et al.*, Different strains of *Mycobacterium tuberculosis* cause various spectrums of disease in the rabbit model of tuberculosis. *Infect. Immunol.*, 2003, **71**(10), 6004–6011.
3. Toossi, Z., The inflammatory response in *Mycobacterium tuberculosis* infection. In *Inflammation* (ed. Cronstein, B. N.), Springer, The Netherlands, 2001, pp. 139–151.
4. Koch, A., and Mizrahi, V., *Mycobacterium tuberculosis*. *Trends Microbiol.*, 2018, **26**(6), 555–556.
5. Farver, C. F. and Jagirdar, J., Mycobacterial diseases. In *Pulmonary Pathology* (ed. Tomashefski, J. F.), 2018, 2nd edn, pp. 201–216.
6. Cole, S. *et al.*, Deciphering the biology of *Mycobacterium tuberculosis* from the complete genome sequence. *Nature*, 1998, **393**(6685), 537.
7. Brown, T., Nikolayevskyy, V., Velji, P. and Drobniewski, F., Associations between *Mycobacterium tuberculosis* strains and phenotypes. *Emerg. Infect. Dis.*, 2010, **16**(2), 272.
8. Bouwman, A. S. *et al.*, Genotype of a historic strain of *Mycobacterium tuberculosis*. *Proc. Natl. Acad. Sci. USA*, 2012, **109**(45), 18511–18516.
9. Stucki, D. *et al.*, *Mycobacterium tuberculosis* lineage 4 comprises globally distributed and geographically restricted sublineages. *Nature Genet.*, 2016, **48**(12), 1535.
10. Hershberg, R., Human host range of *Mycobacterium tuberculosis*. *Nature Genet.*, 2016, **48**(12), 1453.
11. Yimer, S. A. *et al.*, *Mycobacterium tuberculosis* lineage 7 strains are associated with prolonged patient delay in pulmonary tuberculosis patients in Amhara region, Ethiopia. *J. Clin. Microbiol.*, 2015, **53**(4), 1301–1309.
12. Borrell, S. and Gagneux, S., Infectiousness, reproductive fitness and evolution of drugresistant *Mycobacterium tuberculosis* (state of the art). *Int. J. Tuberc. Lung Dis.*, 2009, **13**(12), 1456–1466.
13. Comas, I. *et al.*, Out-of-Africa migration and Neolithic coexpansion of *Mycobacterium tuberculosis* with modern humans. *Nature Genet.*, 2013, **45**(10), 1176–1182.
14. Sen, R., Nayak, L. and De, R. K., A review on host–pathogen interactions: classification and prediction. *Eur. J. Clin. Microbiol. Infect. Dis.*, 2016, **35**(10), 1581–1599.
15. Rana, A., Ahmed, M., Rub, A. and Akhter, Y., A tug-of-war between the host and the pathogen generates strategic hotspots for the development of novel therapeutic interventions against infectious diseases. *Virulence*, 2015, **6**(6), 566–580.

16. Bonetta, L., Protein–protein interactions: interactome under construction. *Nature*, 2010, **468**(7325), 851.
17. Huo, T., Liu, W., Guo, Y., Yang, C., Lin, J. and Rao, Z., Prediction of host–pathogen protein interactions between *Mycobacterium tuberculosis* and *Homo sapiens* using sequence motifs. *BMC Bioinformat.*, 2015, **16**(1), 100.
18. Winnenburg, R., Baldwin, T. K., Urban, M., Rawlings, C., Köhler, J. and Hammond-Kosack, K. E., PHI-base: a new database for pathogen host–interactions. *Nucleic Acids Res*. (*suppl. 1*), 2006, **34**, D459–D464.
19. Xiang, Z., Tian, Y. and He, Y., PHIDIAS: a pathogen–host interaction data integration and analysis system. *Genome. Biol.*, 2007, **8**(7), R150.1–R150.15.
20. Wattam, A. R. *et al.*, PATRIC, the bacterial bioinformatics database and analysis resource. *Nucleic Acids. Res*., 2013, **42**(D1), D581–D591.
21. Rapanoel, H. A., Mazandu, G. K. and Mulder, N. J., Predicting and analyzing interactions between *Mycobacterium tuberculosis* and its human host. *PLoS ONE*, 2013, **8**(7), e67472, 1–4.
22. Zhou, H., Gao, S., Nguyen, N. N., Fan, M., Jin, J., Liu, B. and Wong, L., Stringent homology-based prediction of *H. sapiens–M. tuberculosis* H37Rv protein–protein interactions. *Biol. Direct*, 2014, **9**(1), 5, 1–30.
23. Chen, C., Huang, H. and Wu, C. H., Protein bioinformatics databases and resources. In *Protein Bioinformatics*, Humana Press, New York, USA, 2017, pp. 3–39.
24. Garcia-Garcia, J., Schleker, S., Klein-Seetharaman, J. and Oliva, B., BIPS: BIANA Interolog Prediction Server. A tool for protein–protein interaction inference. *Nucleic Acids Res.*, 2012, **40**(W1), W147–W151.
25. Powell, S. *et al.*, eggNOG v3. 0: orthologous groups covering 1133 organisms at 41 different taxonomic ranges. *Nucleic Acids Res.*, 2011, **40**(D1), D284–D289.
26. Walhout, A. J. *et al.*, Protein interaction mapping in *C. elegans* using proteins involved in vulval development. *Science*, 2000, **287**(5450), 116–122.
27. UniProt Consortium, UniProt: the universal protein knowledgebase. *Nucleic Acids Res.*, 2017, **45**(D1), D158–D169.
28. Cohen, T., van Helden, P. D., Wilson, D., Colijn, C., McLaughlin, M. M., Abubakar, I. and Warren, R. M., Mixed-strain *Mycobacterium tuberculosis* infections and the implications for tuberculosis treatment and control. *Clin. Microbiol. Rev.*, 2012, **25**(4), 708–719.

Received 5 March 2019; revised accepted 9 September 2019

doi: 10.18520/cs/v118/i2/292-297

# Effect of storage conditions on vermicompost quality

**Pankaj Kumar[1,2,*], Pawan Sharma[3], Tilak Ranjan Kumar[1], H. C. Sharma[3] and R. C. Dubey[4]**

[1]Department of Microbiology, Dolphin (PG) College of Science and Agriculture Chunnikalan, Fatehgarh Sahib 140 307, India
[2]Department of Microbiology, Dolphin (PG) Institute of Biomedical and Natural Sciences, Manduwala, Dehradun 248 007, India
[3]ICAR-Indian Institute of Soil and Water Conservation, Chandigarh 160 019, India
[4]Department of Botany and Microbiology, Gurukul Kangri Vishwavidyalaya, Haridwar 249 404, India

**To study the effects of storage conditions on vermicompost, an experiment was conducted with freshly prepared, un-dried fresh vermicompost (60% moisture), and pre-air dried vermicompost (30% moisture) stored in polythene bags for a period of four weeks under different moisture and aeration conditions. Three different storage conditions were tested by placing fresh and pre-dried vermicompost in: (1) open bags, (2) holed bags and (3) sealed bags. Vermicompost properties were analysed weekly for four weeks after storage. The moisture content declined in both fresh and pre-dried vermicompost, with a maximum decline under open bag condition, followed by holed and sealed bags. In the sealed airtight bags with fresh vermicompost, a rapid decline in total organic carbon, nitrogen and electrical conductivity was observed during the first and second week of storage, possibly due to microbially-triggered volatilization losses. However, such decline was lacking in pre-dried vermicompost. In open and holed bags, the carbon and nitrogen were retained and rather increased during storage, possibly due to ongoing aerobic decomposition and no volatilization losses. The highest nutrient quality was observed under predried holed bag conditions, possibly due to optimal microbial activity releasing nutrients, combined with no volatilization losses. It was concluded that fresh vermicompost must be air-dried before its storage in bags. Storage of air-dried vermicompost under aerobic conditions using open/ holed bags appears to be the best option for retaining nutrients and quality of vermicompost.**

**Keywords:** Carbon, electrical conductivity, nitrogen, quality, storage conditions, vermicompost.

VERMICOMPOST is widely used as an organic source of nutrients and carbon due to its high availability of nutrients and also for improving soil aeration, water-holding capacity, buffer capacity, and cation exchange capacity of soils[1,2]. Application of agrochemicals for

*For correspondence. (e-mail: guptapankaj23@gmail.com)