

FDBC: A Comprehensive Platform of the Fungi Databases Collection.

Shahid Ullah^{2*}, Wajeeha Rahman², Farhan Ullah², Riffat Jehan², Anees Ullah³, Hameed Ullah², Muhammad Ijaz², Gulzar Ahmad² and Tianshun Gao^{1*}.

¹Research Center, The Seventh Affiliated Hospital of Sun Yat-sen University, Shenzhen, Guangdong, China

²S Khan Lab Mardan, Khyber Pakhtunkhwa, Pakistan

³Kyrgyz State Medical Academy (KSMA), Kyrgyzstan

Corresponding Author*

Tianshun Gao
Research Center,
The Seventh Affiliated Hospital of Sun Yat-sen University,
Shenzhen, Guangdong, China,
E-mail: gts.hust@gmail.com
Shahid Ullah
S Khan Lab Mardan,
Khyber Pakhtunkhwa, Pakistan,
E-mail: drskbioch@gmail.com

Copyright: © 2022 Ullah S, et al. This is an open-access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Received: 17-May-2022, Manuscript No. JBTW-22-64005; **Editor assigned:** 20-May-2021, PreQC No. JBTW-22-64005(PQ); **Reviewed:** 03-Jun-2022, QC No. JBTW-22-64005; **Revised:** 10-Jun-2022, Manuscript No. JBTW-22-64005(R); **Published:** 17-Jun-2022, DOI: 10.35248/2322-3308-11.3.005

Abstract

Background: High-throughput sequencing technologies are increasingly generating data of all species including fungi. Therefore the first challenge of omics data was the design and use of electronic databases to store and manage the stormy amount of knowledge. In biological experiments, biological databases for structure and function analysis and estimation are immensely beneficial. A multitude of such databases are currently available, which help in the comprehension of complicated problems relating to fungal species.

Objectives: As a result, we have assembled almost all fungal databases from previously published works into one platform, with the aim of saving and providing an easy, coordinated platform to the entire fungi research community.

Methods: 82 fungal databases were compiled and grouped into 5 categories based on their various properties, including the Genomic database, Proteomic database, Transcriptomics database, Pathway database, and Structure database.

Results: Under numerous conditions, the available resources have led to the efficient understanding of fungal databases. Each category is given with a brief introduction; users can query the appropriate information in two ways, by clicking image expression icon, or by typing directly name in the query-bar which is on the top of the database. FDBC is developed in PHP, HTML, CSS, MySQL, and JavaScript.

Conclusion: We have provided a set of fungal databases from previously published work, and have built the database of these databases named "FDBC (Fungal Databases Collection)" which is accessible on <https://www.habdsk.org/fdbc> and will be modified in due time.

Keywords: FDBC • Fungi • Fungal database • Genomic • Fungal research community

Introduction

The fungal kingdom contains approximately 2.2-3.8 million heterotrophic eukaryotes, the majority of which are observable and substrate-dwelling [1]. Massive study has been done in various areas of fungal research, such as the use of genomic (DNA sequencing) data in barcoding and Meta barcoding attempts to examine fungal species [2-5]. Incorporating molecular phylogenetic data allowed for more precise genealogical findings, which revolutionized fungi classification [6,7], and next-generation sequencing technologies are increasingly generating data on all species, including fungi

and yeast [8], therefore the first challenge of omics data was the design and use of electronic databases to store and manage the stormy amount of knowledge.

For that manual database curation is made to make sure that biological information and knowledge are communicated in modern times [2,9]. A complete computer representation of the cell and the organism, which will allow computational prediction of higher-level complexity of cellular processes and behavior of organisms from omics data, is a major challenge in the post-genomic period [10]. Scientists have been enthralled by the definition and potential of system biology, which aims to bring together data from a variety of omics approaches (such as genomics, Transcriptomics, proteomics, and metabolomics) to provide a more integrative picture of cells, species, and populations in terms of their development, growth, evolution, and infection resistance [8,11,12].

so, a key role in bioinformatics is played by the biological databases, they allow scientists to access a large number of relevant biological data, having a broader variety of species in the genetic sequences [13], to organize and categorize these multi-omics data in the kingdom of fungi. Huge databases have been published in multiple research point of the fungi, such as "Genomic databases", OriDB [14], MycoBank [15], "Proteomic databases" (YAAM) [16], "Structure databases" PROPHECY [17], "Pathway databases", Yeast Net [18], and "Transcriptome databases" YeastSS [19], "AGD 3.0" [20]. To make it simpler, we've gathered all of these databases into one platform called FDBC, where we've thoroughly curated and integrated 82 Fungal databases based on scientific work. There were multiple ways to access the data, as well as original links and brief explanations for each database. Our archive, when combined, can be useful and modified over time for further Fungal research analysis.

Materials and Methods

Construction of the database

The FDBC database is constructed in a very easy way; computer platform has been used, such as Php, HTML, CSS, JavaScript, and MySQL. The data has been collected from many search engines, e.g., Google, Google Scholar and PubMed. Many keys words have been used to search the data, such as, "Fungal databases", "databases of Fungi", "fungus huge databases", and so on. For confirmation, we have manually checked every database carefully and have removed all outdated and dead links. Further our database has updated data and will be updated timely. For more and easy understanding, we have given the flow diagram of FDBC (Figure 1).

Usage of the databases

We have offered the usage of the databases very friendly and easy manner, user can search the interesting database in 3 ways, first, for a single query can direct type the name of the needed databases in the given search bar, Which is highlighted in Figure 2A, for category wise search can click on the Image

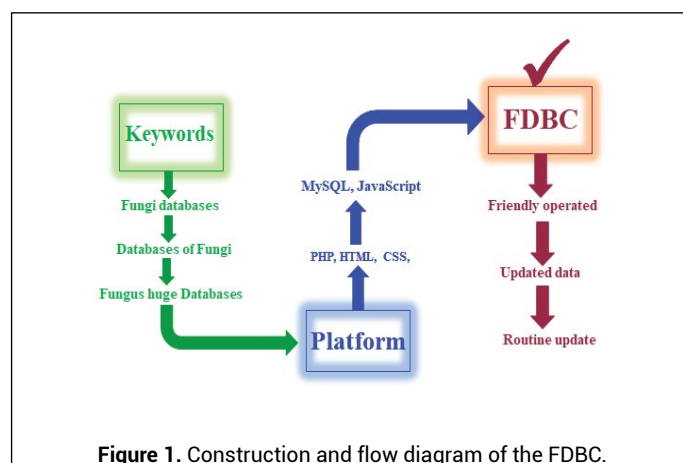


Figure 1. Construction and flow diagram of the FDBC.

expression icon or name of the category Figure 2A, that will direct leads to all the category databases list Figure 2B, for example by clicking genomic image, will give 54 genomic databases on single click, user can choose their needed database, further clicking will lead to a short introduction and official link of the needed query (Figure 2A).

Results and Discussion

Classification of the database

According to previously published work [21], many articles have been online with their classification based on species or components [22-24], e.g., human [25], plant [26]. As a result, fungal databases can be categorized based on a number of requirements, including "Genomic databases," "Proteomic databases," "Transcriptomics databases," "Pathway databases," and "Structure databases." Some of the more well-known database's examples are presented in the form of figures in each category, with information and implementations provided below.

Genomic databases

A genome is all genetic material of an organism. To ensure a uniform representation of biological information across different organisms [27], most of the genomic databases are used, to controlled vocabularies to annotate various attributes of genes and gene products [28]. Various databases have been published in fungi genomic research area, such as "The

Candida Genome Database" [29] was started in 2004 to serve as an online compendium of genomic, genetic, and molecular biology information about *Candida albicans*, shown in Figures 3A and 3B [29] is a well-known genomic database named "SGD" that allows the exploration of functional interactions between sequence and gene products in fungi and other species [30], and "Fungi DB" [31].

Proteomic databases

The main challenge in fungal biology is to understand the expression, function, and regulation of the whole set of proteins encoded by fungal genomes [32], Which is a wide-ranging numerical description of protein expression and its variation under the inspiration of biological concerns such as drug treatment, or diseases [33]. To overcome these challenges there are several fungal proteomics databases have been published, such as Figure 3C is the "Fungal Secretome Database (FSD)" which includes secretome data on 525 species with 587 datasets [34], Figure 3D "e-Fungi" provides genomic and functional data for the 34 fairly diverse fungal genomes and oomycetes [35], "FGMP" [36], which describes the study of fungal proteome and is shown in (Figure 3D).

Transcriptomics databases

Transcriptomics is the study of global change of gene expression at mRNA level [32], which regulates the major process to control location, timing, and quantity of gene expression [37]. To save these all data several transcriptomic databases have been built and giving such huge information on friendly

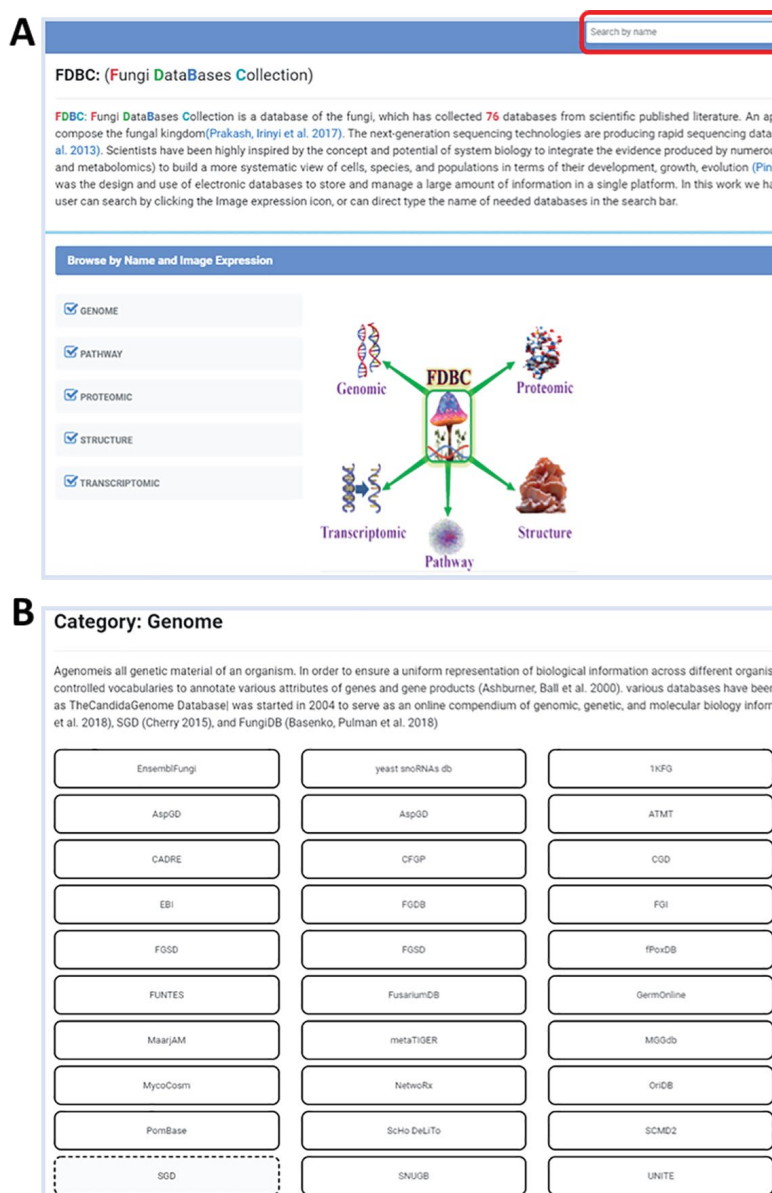


Figure 2. The usage of the FDBC, (A) Searching by name of image expression, (B) The category list of the databases.

finding ways, for example, Figure 4A, The Fungal Transcription Factor Database (FTFD) is developed to integrate the putative fungal Transcription Factors (TFs) with references, provides sequences, the taxonomic and phylogenomic background and the overall sketch of all fungal TFs [38], "YeTFaSCo" is designed for motif analysis, including sequence scanning and pre-computed gene-browser detection of motif occurrences across the entire yeast genome, shown in Figure 4B [39].

Pathway databases

The pathway is the framework for molecular biology. Fungal pathway databases attempt to demonstrate the relations between genes, proteins, and metabolites, which is crucial for understanding the various mechanisms that exist within the cell [40-42]. Fungi pathway diagram databases are useful for researchers who want to contribute, interpret, and visualize omics

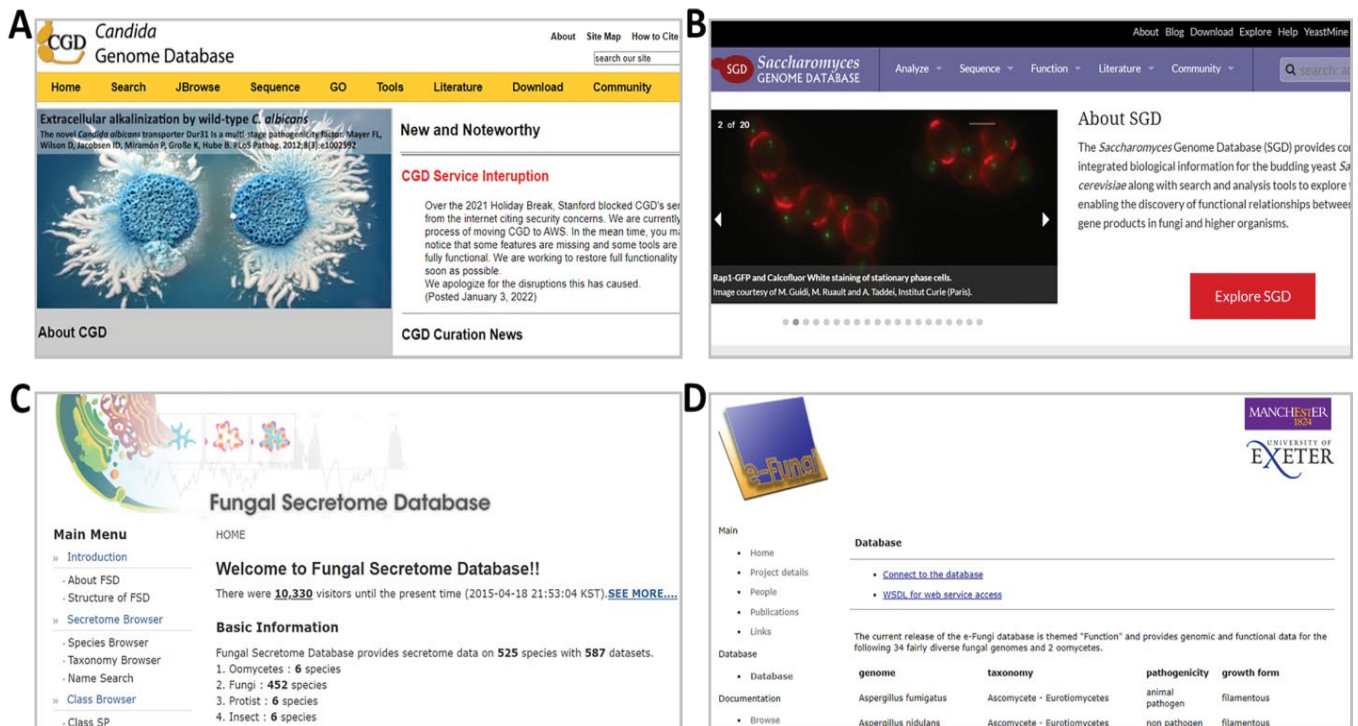


Figure 3. The containers of the genomic and proteomic databases. (A) "The Candida Genome Database", (B) The genomic database named SGD, (C) "Fungal Secretome" the proteomic database. (D) "e-Fungi" the proteomic database.

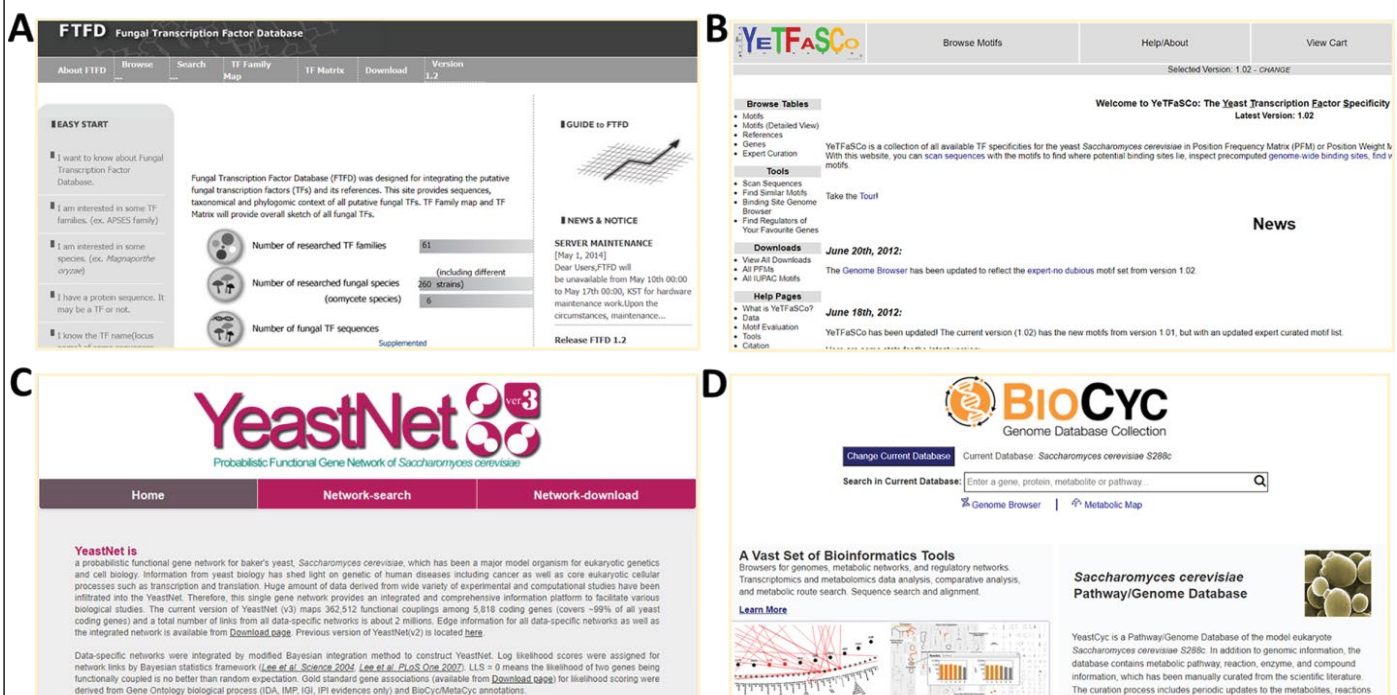


Figure 4. The main-face of the well-known Transcriptomic and pathway databases. (A) A Transcriptomic database named FTFD, (B) A Transcriptomic database named YeTFaSCo, (C) The pathway database named YeastNET, (D) The pathway database named YeastCyc.

data, which is used to simplify more complex analytical activities like cellular modeling [43]. Numerous fungi pathway databases have been usable in scientific research, some of them are "KEEG Yeast" [44], "Yeast Net" Figure 4C is the probabilistic adaptive gene network for baker's yeast, *Saccharomyces cerevisiae*, was a leading model organism for eukaryotic genetics and cell biology. Massive volumes of data from a wide range of laboratory and analytical experiments have been penetrated the Yeast Net. This single gene network thus offers an interconnected and detailed knowledge portal to promote numerous biological studies [45], Figure 4D is a "Yeast Cyc" which is a Pathway Database of the eukaryote model *Saccharomyces cerevisiae* which is part of the broader Bio Cyc set of thousands of Pathway Databases for sequenced genomes comprising metabolic pathways, reactions, enzyme and compound details (Figure 4D) [46].

Structure databases

Fungal structural databases are the experimental determination of fungi structure, which aimed to organized and annotate useful information of the fungi [47]. It contains information that has been gathered in the context of

structural databases due to the distinct morphological stages and different structures. Some well-known and relevant databases that provide structural data for fungi are listed, such as in Figure 5A "PROPHECY" database tests the phenotype of the deletion strain based on growth activity during micro-cultivation. it quantifies growth aberrations by calculating growth rate, growth efficiency, and adaptation time [17]. Figure 5B is a database named "Bio Models" a resource that enables the storage and sharing of published, peer-reviewed quantitative, interactive models of biological processes. Model elements and actions are carefully tested to fit the initial publication and manually curated to ensure reliability (Figure 5B) [48].

Statistics of the database

In this work, we provided access to all fungal databases, and have shown the category-wise, Chronological order, and percentile development of the fungi databases. Figure 6A shown the categories wise growth of the fungal databases, which is the enormous growth and accomplishment for the mycological community, among these categories the genomic data is strengthened well over time. Figure 6B reflects the data chronology, which shows decreased after 2015 due to many broken and outdated links, that

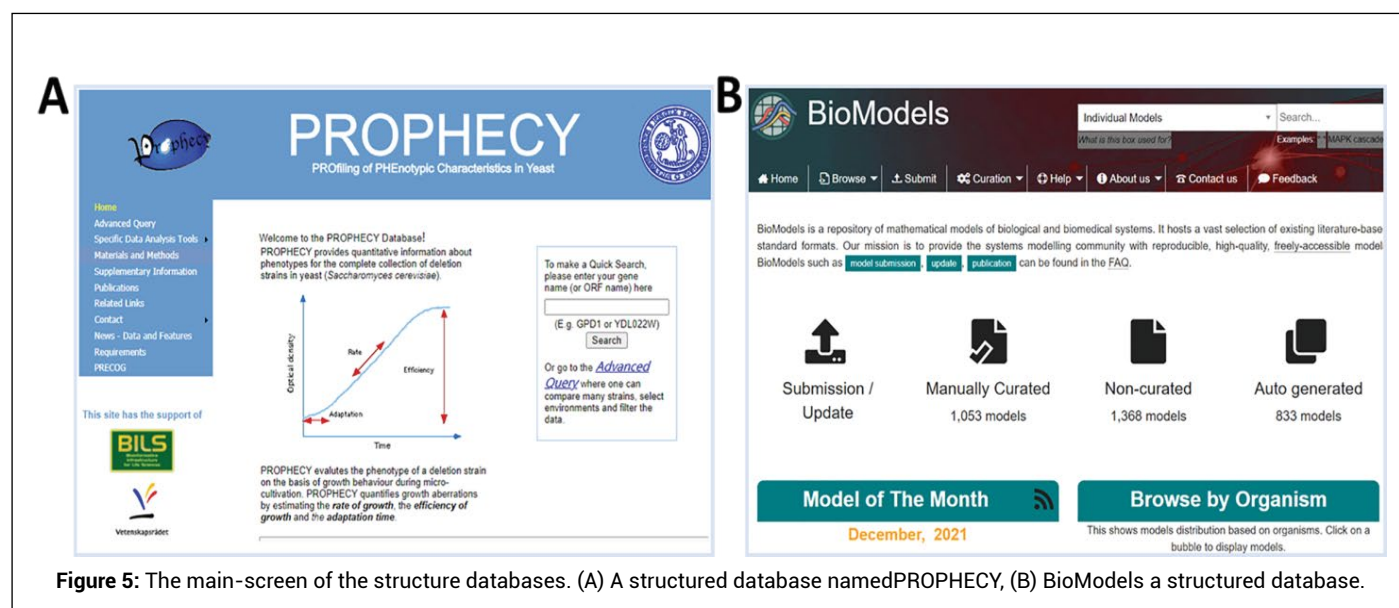


Figure 5: The main-screen of the structure databases. (A) A structured database namedPROPHECY, (B) BioModels a structured database.

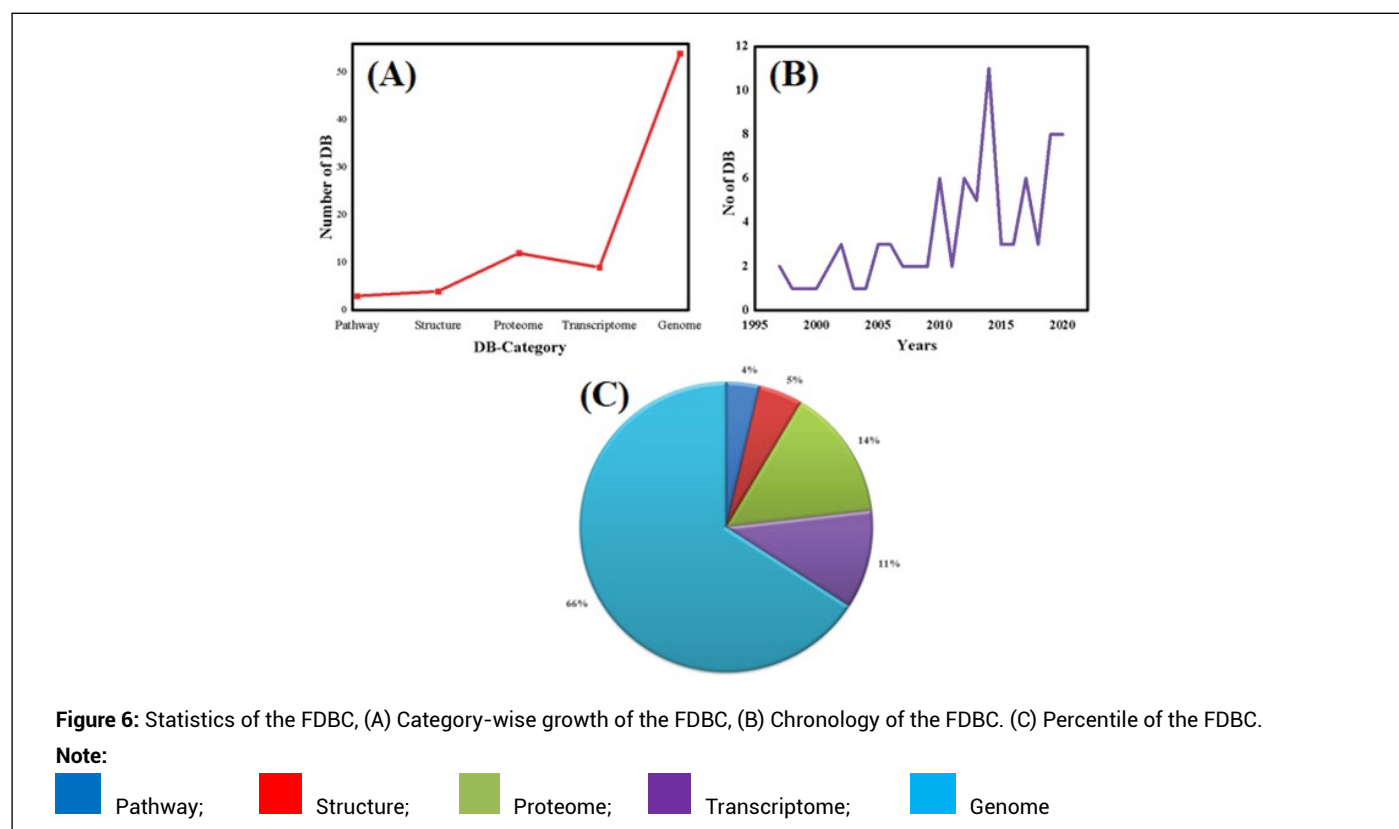


Figure 6: Statistics of the FDBC, (A) Category-wise growth of the FDBC, (B) Chronology of the FDBC. (C) Percentile of the FDBC.

Note:

Pathway; Structure; Proteome; Transcriptome; Genome

Table 1. Comparison table of the FDBC with other published work.

PMID	Species or component	Type of the data	No of DB	Journal name
FDBC	Fungi	Table+DB	82	
28179406	Fungi	Table	24	Journal of Clinical Microbiology
25712261	Human	Table	74	Genomic, Proteomic & Bioinformatics (GPB)
31906604	Nucleic acid	Table	70	Nucleic Acids Research (NAR)
7764641	DNA+Protein	Table	50	Current Opinion in Biotechnology
18265344	Protein	Table	121	Current Protocols in Molecular Biology

we have removed all the dead links and given a fresh and revised for the mycological community, while Figure 6C is a simple percentile of the FDBC data (Figure 6C).

Development of new features

Several articles have been published in well-reputed journals such as "Path guide: A pathway resource list" published in Nucleic Acids Research, "Biological Databases for Human Research" published in genomic proteomic and bioinformatics Journal, which have collected updated databases in the form of tables, for more understanding we have compared our work with previously published research in the Table 1. In the current study, we have provided a table as well as developed a database by a novel way of representation, which hasn't been built and published before in such an easy and friendly searching way. Shortly, we have the planned future development of new features in the article, such as Figure 5A simply gives the category information by clicking image expression.

Conclusion

In modern scientific work, databases are the first phase in almost all of the biological research fields, which assists the researchers in all aspects, many useful databases have been published and going to be published for easy searching, to avoid displacement or disorders of such data, several articles have been published which compiled databases of different species or components, such as "Human", "Plant", "Proteins". Therefore, a decent collection of fungal databases is also needed to save, manage, and organized the fungal databases for future and easy searching, so that we have provided a set of fungal databases from previously published work, and have built the database of these databases named "FDBC (Fungal Databases Collection)" that have divided into five categories with a short description, and gives easy access to all the data. FDBC enables users to search a database by typing the name of the database in the search bar, or by clicking on an image category expression or a name of the category. Besides, the data in our database has been updated and will be updated over time.

Declarations

Ethics approval and consent to participate

Not applicable.

Consent for publication

Not applicable.

Availability of data and materials

These data will be available under the journal rule and regulation.

Competing interests

The authors don't have any compete of interests.

Funding

This project is supported by National Natural Science Foundation of China [32100434] and Shenzhen's introduction of talents and research start-up [392020].

Author Contributions

Dr. Shahid Ullah and Dr. Tianshn Gao designed and supervised the project, performed data analysis. Farhan Ullah, Wajeeha Rahman, Anees Ullah, Riffat Jehan, Hameed Ullah, Muhammad Ijaz, and Gulzar Ahmad, contributed to data analysis. Dr. Shahid Ullah wrote the manuscript. All authors reviewed the manuscript.

Acknowledgement

To avoid future conflict, we would like to say that, our database is uploaded on S Khan Lab website <https://www.habds.org/fdbc>, so that we have provided some content in this article.

References

- Prakash, P.Y., et al. "Online databases for taxonomy and identification of pathogenic fungi and proposal for a cloud-based dynamic data network platform." *J Clin Microbiol* 55(2017):1011-1024.
- Percudani, R., et al. "Ureidoglycolate hydrolase, amidohydrolase, lyase: how errors in biological databases are incorporated in scientific papers and vice versa." *J Biological Database* 2013.
- Snelders, N.C., et al. "Microbiome manipulation by a soil-borne fungal plant pathogen using effector proteins." *Nature Plants* 6(2020):1365-1374.
- Banuett, F. "Signalling in the yeasts: an informational cascade with links to the filamentous fungi." *Microbiol Mol Biol Rev* 62.2(1998):249-274.
- Treseder, K.K. & Lennon, J.T. "Fungal traits that drive ecosystem dynamics on land." *Microbiol Mol Biol Rev* 79(2015):243-262.
- Yarza, P., et al. "A phylogenetic framework for the kingdom Fungi based on 18 S rRNA gene sequences." *Mar Genomics* 36(2017):33-39.
- Bianchi, F., et al. "Regulation of amino acid transport in *Saccharomyces cerevisiae*." *Microbiol Mol Biol Rev* 83(2019).
- Walters, A.D., et al. "Nuclear envelope expansion in budding yeast is independent of cell growth and does not determine nuclear volume." *Mol Biol Cell* 30(2019):131-145.
- Ullah, S., et al. "EDBCO-19: Emergency Data Base of COVID-19." *J Clin Med Res* 2(2020):1-4.
- Kanehisa, M., et al. "The KEGG resource for deciphering the genome." *Nucleic Acids Res* 32(2004):D277-D280.
- Pinu, F.R., et al. "Systems biology and multi-omics integration: Viewpoints from the metabolomics research community." *Metabolites* 9(2019):76.
- Chan, K., et al. "Genomic comparison of *Salmonella enterica* serovars and *Salmonella bongori* by use of an *S. enterica* serovar Typhimurium DNA microarray." *J Bacteriol* 185(2003):553-563.
- Baxevanis, A.D. & Bateman, A. "The importance of biological databases in biological discovery." *Curr Protoc Bioinform* 50(2015):111-118.
- Siow, C.C., et al. "OriDB, the DNA replication origin database updated and extended." *Nucleic Acids Res* 40(2012):D682-D686.
- Robert, V., et al. "Mycobank gearing up for new horizons." *IMA fungus* 4(2013):371-379.
- Ledesma, L., et al. "YAAM: Yeast Amino Acid Modifications database." *Database* (2018).
- Fernandez-Ricaud, L., et al. "PROPHECY-a yeast phenome database, update 2006." *Nucleic Acids Res* 35(2007):D463-D467.
- Kim, H., et al. "YeastNet v3: A public database of data-specific and integrated functional gene networks for *Saccharomyces cerevisiae*." *Nucleic Acids Res* 42(2014):D731-D736.
- McMillan, J., et al. "YeastTSS: an integrative web database of yeast

- transcription start sites. " *Database* (2019).
20. Gattiker, A., et al. "Ashbya Genome Database 3.0: A cross-species genome and transcriptome browser for yeast biologists. " *Bmc Genomics* 8(2007):9.
 21. KhanSA, JM., et al. "Tremendous contribution of Dr. Shahid Ullah to scientific community during COVID-19 pandemic in the form of scientific research. " *J Clin Med Res* 2(2020):1-7.
 22. Xu, D. "Protein databases on the internet. " *Curr Protoc Protein Sci* 70(2012):2611-2617.
 23. Celis, JE., et al. "Human and mouse proteomic databases: novel resources in the protein universe. " *FEBS letters* 430(1998):64-72.
 24. Harper, R. "Access to DNA and protein databases on the Internet. " *Curr Opin Biotechnol* 5(1994):4-18.
 25. Zou, D., et al. "Biological databases for human research. " *Genom Proteom Bioinform* 13(2015):55-63.
 26. Ullah, S., et al. "DBPR: Data Base of Plant Research. " (2021).
 27. Pincus, SE., et al. *Molecular biology of RNA*. (1987):175-210.
 28. Ashburner, M., et al. "Gene ontology: tool for the unification of biology. " *Nature gene* 25(2000):25-29.
 29. Skrzypek, MS., et al. "Eukaryotic Genomic Databases. " *Springer* (2018):31-47.
 30. Cherry, JM. "The saccharomyces genome database: Exploring genome features and their annotations. " *Cold Spring Harb Protoc* (2015):088922.
 31. Basenko, EY., et al. "FungiDB: an integrated bioinformatic resource for fungi and oomycetes". *J Fungi* 4(2018):39.
 32. Bhadauria, V., et al. "Advances in fungal proteomics. " *Microbiol Res* 162(2007):193-200.
 33. Anderson, NL & Anderson, NG. "Proteome and proteomics: new technologies, new concepts, and new words. " *Electrophoresis* 19(1998):1853-1861.
 34. Choi, J., et al. "Fungal secretome database: integrated platform for annotation of fungal secretomes. " *BMC Genomic* 11(2010):1-15.
 35. Hedeler, C., et al. "E-Fungi: a data resource for comparative analysis of fungal genomes. " *BMC Genomic* 8(2007):1-15.
 36. Cissé, OH & Stajich, JE. "FGMP: assessing fungal genome completeness. " *BMC Bioinformatic* 20(2019):1-9.
 37. Wu, WS., et al. "YCRD: yeast combinatorial regulation database. " *PloS one* 11(2016):e0159213.
 38. Park, J., et al. "FTFD: an informatics pipeline supporting phylogenomic analysis of fungal transcription factors. " *Bioinformatic* 24(2008):1024-1025.
 39. De Boer, CG & Hughes, TR. "YeTFaSCo: A database of evaluated yeast transcription factor sequence specificities. " *Nucleic acids res* 40(2012):D169-D179.
 40. Villaveces, JM., et al. "Tools for visualization and analysis of molecular networks, pathways, and-omics data. " *Adv Appl Bioinform Chem* 8(2015):11.
 41. Rowan-Nash, AD., et al. "Cross-domain and viral interactions in the microbiome. " *Microbiol Mol Biol Rev* 83(2019).
 42. Stulberg, ER., et al. "Genomic and secondary metabolite analyses of *Streptomyces* sp. 2 AW provide insight into the evolution of the cycloheximide pathway. " *Front Microbiol* 7(2016):573.
 43. Kuan, CS., et al. "DemaDb: an integrated dematiaceous fungal genomes database. " *Database* (2016).
 44. Kanehisa, M., et al. "The KEGG databases at Genome Net. " *Nucleic acid res* 30(2002):42-46.
 45. Herrgård, MJ., et al. "A consensus yeast metabolic network reconstruction obtained from a community approach to systems biology. " *Nat Biotechnol* 26(2008):1155-1160.
 46. Gutiérrez, C., et al. "Wine yeasts identification by MALDI-TOF MS: Optimization of the preanalytical steps and development of an extensible open-source platform for processing and analysis of an in-house MS database. " *Int J Food Microbiol* 254(2017):1-10.
 47. Dirks, AC & Jackson, RD. "Community structure of arbuscular mycorrhizal fungi in soils of switchgrass harvested for bioenergy. " *Appl Environ Microbiol* 86(2020).
 48. Chelliah, V., et al. "In Silico Systems Biology. " *Springer* (2013):189-199.

Cite this article: Shahid Ullah, Wajeeha Rahman, Farhan Ullah, Gulzar Ahmad, Muhmmad ijaz, Tianshun Gao. FDBC: A Comprehensive Platform of the Fungi Databases Collection. *J Biol Today's World*, 2022, 11(3), 001-009