Malaysian Applied Biology (2022) 51(5): 37-42
https://doi.org/10.55230/mabjournal.v51i5.2346

Research Article

A Systematic Database for Microbial Data Collection Management

Norsyahima Azizi*, Suhaila Sulaiman, Farah Nini Othman and Lee Yang Ping

FGV R&D Sdn. Bhd., FGV Innovation Centre (Biotechnology), PT. 23417, Lengkuk Teknologi, 71760 Bandar Enstek,
Negeri Sembilan
*Corresponding author: norsyahima@fgvholdings.com

ABSTRACT

Microbiology research is one of the research fields that deal with a huge volume of microbial culture collection data which needs to be preserved, due to its importance as a key resource center for obtaining, identifying, producing, preserving, cataloging, and distributing culture data of various microbes. Currently, the collection and storing of microbial culture data is performed manually via various data transfers into record books, logbooks, or spreadsheets. However, the manual process seems to be impractical as it is highly prone to human error and time-consuming, aside from the struggle of managing and retrieving the data. Therefore, to address this issue, we have successfully designed and developed a customized database for microbial data collection that allows users to store, access, visualize and retrieve information in a single platform. The agile development approach has been implemented during the software development process allowing more flexibility during the development phase yet managing to fulfill the functional requirements of the end-users. The front end is created using the Bootstrap framework, an open-source and powerful platform for developing user-friendly web pages and web applications. It is connected to the back-end MySQL database as a relational database management system to organize structured data. As a result, the developed microbial culture collection has eliminated manual data entry and accelerated the data retrieval process. The platform is a one-stop platform for microbiologists when dealing with microbial culture collection data.

Key words: Microbial culture collection, software development, Bootstrap

Article History
Accepted: 28 November 2022
First version online: 26 December 2022

Cite This Article:
Azizi, N., Sulaiman, S., Othman, F.N. & Ping, L.Y. 2022. A systematic database for microbial data collection management. Malaysian Applied Biology, 51(5): 37-42. https://doi.org/10.55230/mabjournal.v51i5.2346

INTRODUCTION

Microbiology is the study of microscopic organisms, that are too small to be seen with the naked eye. Bacteria, viruses, archaea, fungi, and protozoa are among the microscopic organisms being studied which are formally referred to as microbes. The study of microbes is pivotal to the study of all living things because of their vital roles in regulating nutrient cycling, biodegradation, climate change, food spoilage, disease causation and control, and biotechnology.

In microbial research, the five basic lab procedures (5 "I's") used by microbiologists to examine and characterize microbes are 1) Inoculation (a technique for introducing a specific microorganism into culture media), 2) Incubation (a process in which inoculum-containing agar or broth is kept in an incubator.), 3) Isolation (a process by which a strain is separated from its natural, mixed environment), 4) Inspection (observation of cultures for obvious growth characteristics) and 5) Identification (determination of microbial type, typically to the level species) (Cowan, 2012). For instance, in the agricultural industry, raw data of various microbes such as code identifier, species and strain name, sampling date, sequence, and pictures are collected from field sampling and research labs.

Ideally, the data should be collected and stored systematically to avoid executing unnecessary and incorrect procedures that may produce pointless or misleading results. However, not all experimental data can be deposited in public databases as some organizations have to comply with their data security policies. In addition to that, the publicly available database is not tailored to the organization's operations. Outsourcing for system development can be costly, and because of this, manual data processing and transferring into record books, logbooks, or spreadsheets have been the standard practice for data storage and management. Spreadsheets are unquestionably useful tools for data aggregation and analysis but organizing and retrieving data can be cumbersome when the spreadsheet contains several fields and a huge volume of data. Aside from the difficulty in managing and retrieving data, the manual process is impractical as it is prone to human error and is time-consuming.

In this digital transformation era, a database, a computer program to store data, is commonly used as a platform for data storage due to its capability to hold a large volume of data. The spreadsheet on another hand can only hold a limited number of data. According to Microsoft Support 2021, Microsoft Excel can only hold up to 1,048,576 rows and 16,384 columns (Microsoft, 2016). Besides, a database offers a degree of flexibility in data sorting and presentation that two-dimensional spreadsheets cannot do through extensive querying functionalities such that; it can retrieve all records meeting select criteria, cross-reference records in different tables, and perform intricate aggregate computations across several tables.

Therefore, in this paper, we demonstrate a way to develop a systematic microbial culture data collection database for storing, accessing, and retrieving microbial data in a single platform.

MATERIALS AND METHODS

The overall program flow of database development involves five phases, including data collection, database operation, data processing, information, and user view (Figure 1).

Data collection and cleaning

In the data collection phase, raw data of various microbes such as code identifier, species and strain name, sampling date, sequence, and pictures are collected from field sampling and laboratory studies by researchers. These records are kept in spreadsheet documents and typically have issues such as redundant information and incorrect formatting. Thus, the raw data is cleaned and formatted before populating into the database.

Database operation

The database structure design is initiated from the clean data collection by identifying all tables (entities) in this database. The collected data is categorized into biocontrol (M1), bioaugmentation (M2), biofertilizer (M3), bioinsecticide (M4), bioherbicide (M7), *Ganoderma* (M6), and arbuscular mycorrhizal fungi (AMF)(M5). Other entities that can be derived from these categories are sequence and collector. Each entity carries a set of attributes including code identifier, species name, strain name, description, sampling date, collector name, sample type, culture date, sampling location, and GPS location. The entity-relationship diagram (ERD) is then created to illustrate how each entity in the database is related to each other (Figure 2).

The relational database management system (RDBMS) uses structured query language (SQL) queries to interact with the database. Once the tables are created, the cleaned data is populated, and data retrieval is done by using SQL queries.

Interface development

The user interface (UI) is developed using Bootstrap, a well-known open-source CSS framework that contains HTML, CSS, and JavaScript-based design templates for typography, forms, buttons, navigation, and other interface components, as reviewed by Gaikwad and Adkar (2019). The UI development is divided into six modules: user authentication, homepage, search engine, microbial passport, information download, and admin page. A mixture of Bootstraps components has been implemented to build the template, layout, content, component, and utilities such as container, container-fluid, navbar, button, drop-down item, and carousel, with some customization to suit the user's requirement.

Backend development

Back-end development is the combination of three components: a server, an application, and a database (Amza *et al*., 2003). This database is built using object-oriented programming (OOP) languages such as PHP and JavaScript, a server-side scripting language to execute and interpret user requests. Besides, this OOP is responsible for powering web application functions like user validation, data retrieval, and page navigation.

In addition, to make the web application functional and responsive, SQL is used to interact the web application with the database manager (Supaartagorn, 2011), MySQL to create, remove, rename, clone, perform processing and manipulate the databases. It also performs maintenance tasks of the databases like data backup and restoration.

RESULTS AND DISCUSSION

Database system to store and safeguard the data

Databases are becoming increasingly important for

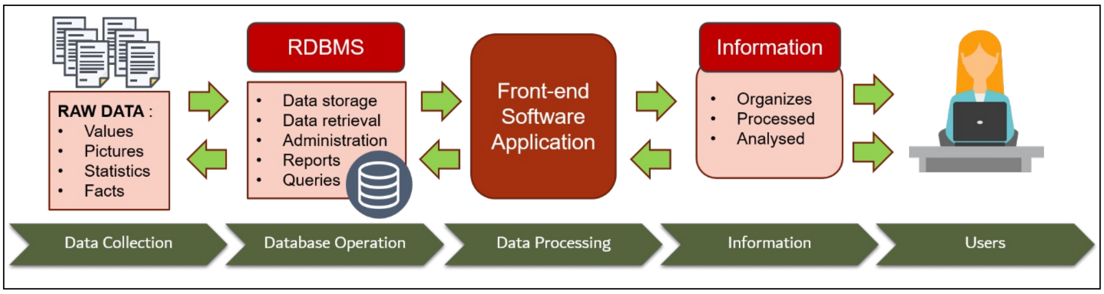


Fig. 1. The overall program flows of the database for the microbial culture collection data management system. It consists of data collection, database operation, data processing, and information and user for data retrieval.

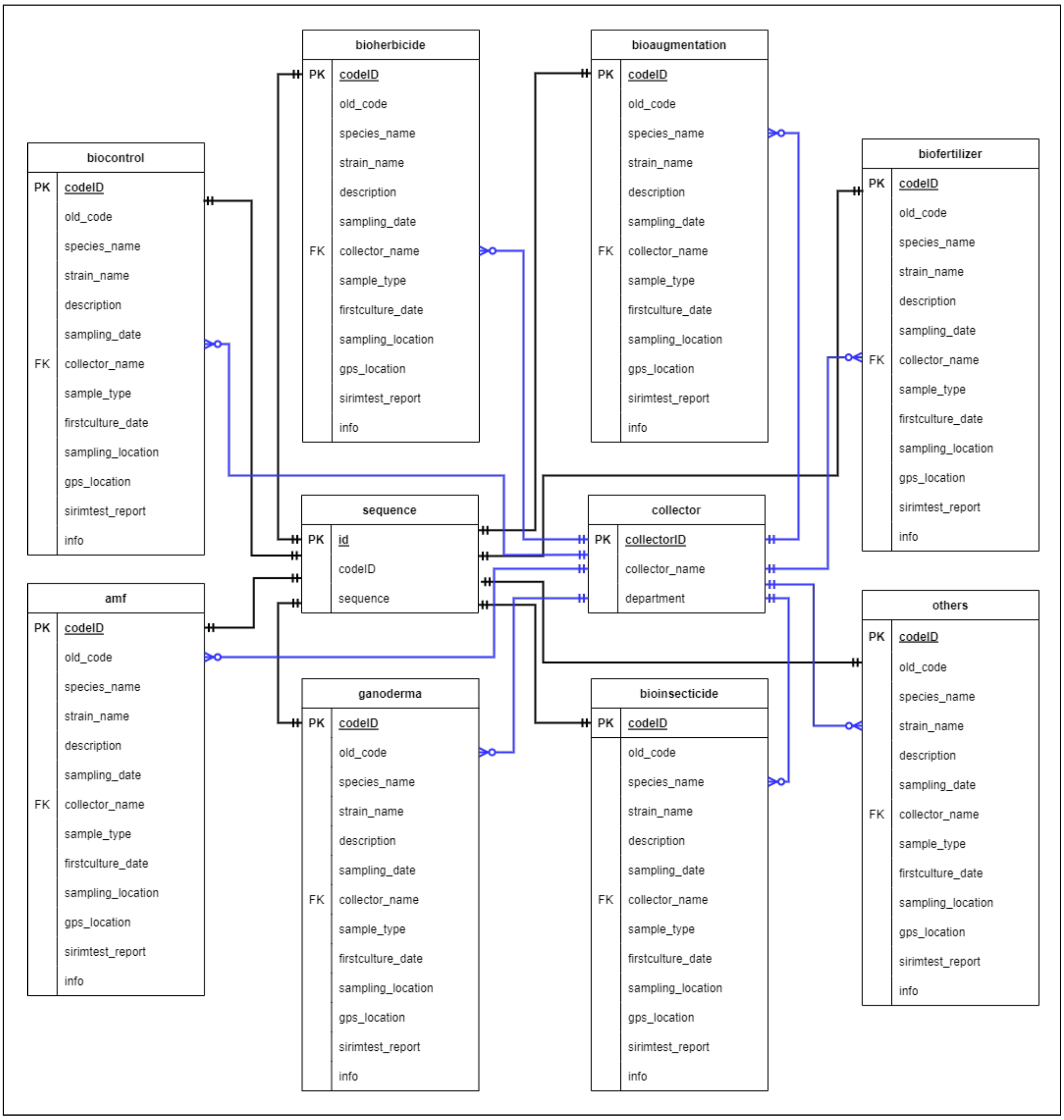


Fig. 2. Entity-relationship diagram (ERD) structure. The black line indicates the relationship between code identifier (ID) as the primary key (PK) in each table with code ID in table 'sequence' as the foreign key (FK). The blue line indicates the relationship between table 'collector' and all tables except table 'sequence' where the 'collector_name' attribute appears in each table.

microbiologists because they frequently interact, access a wide variety of biological data, and exchange information with each other. With the presence of this systematic database for microbial data collection management, it will transform the data into a platform for microbiologists to easily share information among themselves (Zhulin, 2015). Besides, the collected data will be systematically stored and preserved in one platform.

All collected data are stored in tables, columns, and rows based on the microbe's category for easy processing. As a result, the process of obtaining data will be simpler, quicker, and more visually appealing to the user. In addition, the presence of this database can reduce the chances of data loss and redundancy.

Simple and user-friendly interface

The example of the database for the microbial data collection management system comes with UI (Figure 3). The UI is important for meeting user expectations and supporting the system's effective functionality. Through contrasting visuals, clean design, and responsiveness, a well-executed user interface facilitates effective interaction between the user and the application. The UI was designed using the Bootstrap framework to make it simple, responsive, and user-friendly in helping users understand and simplify navigation and accessibility (Gaikwad & Adkar, 2019).

The homepage interface is shown in Figure 3. It consists of three parts; the main menu bar, a panel of selections based on microbial category, and a search engine. The first part comprises several menus; the '*home*' menu returns users to the landing page, the '*about*' menu displays a summary of the microbial culture collection, and the '*collection*' menu navigates the user to the list of microbial collection in the database, the '*sample deposition*' menu describes the information on how to deposit user sample in this database and '*contact us*' menu displays the location, address and other information to contact the system owner. Meanwhile, the second part is the shortcut button to navigate users to the microbial collection page and the last part is a search engine where users can find information about desired microbes.

Data retrieval in this database system will be displayed by microbe's categories shown in Figure 4. On this page, users can retrieve information about the selected microbes, search the microbe's code identifier in the search field, and able to download the retrieval data by clicking the download button (Figure 5). The result will be displayed in the subtab menu: *general*, *sequence*, *culture method*, *summary*, and *reference* under the search field. The general subtab displays general information such as species name, strain name, sampling date, collector name, etc., while the sequence subtab displays sequence information. The culture method, summary, and reference

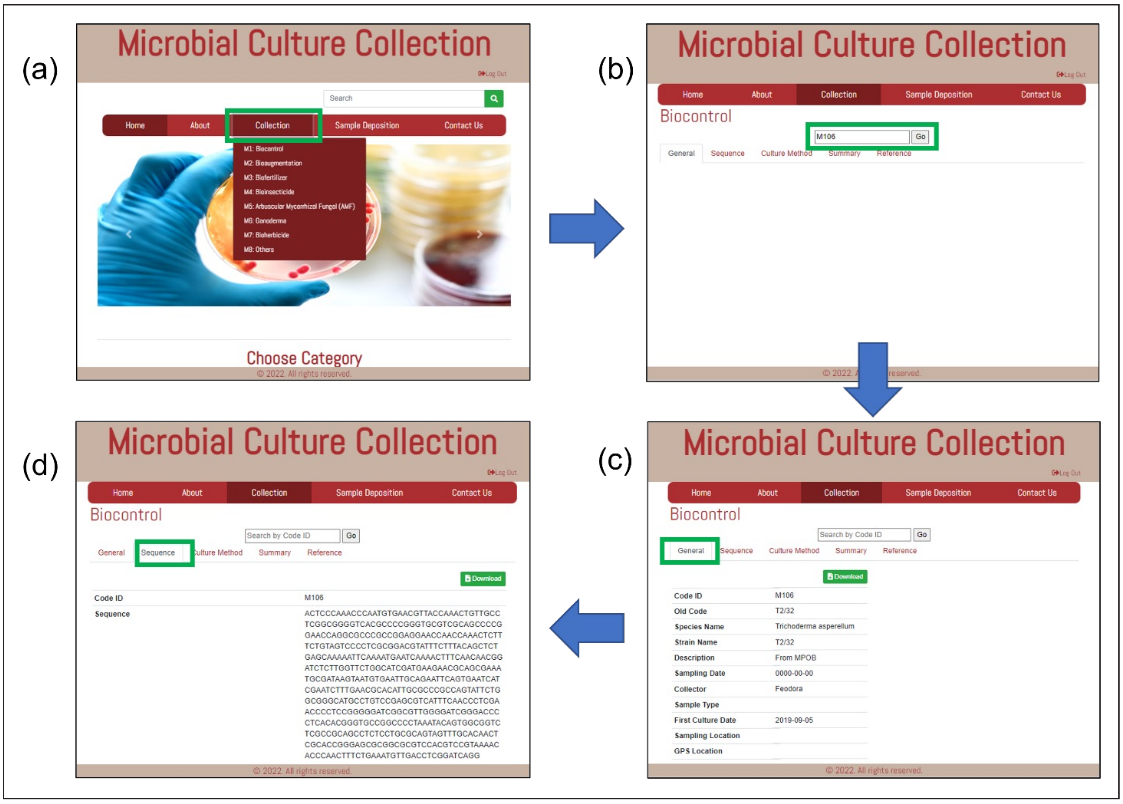
Fig. 3. The homepage of the microbial culture collection database.

subtabs show the method used to culture the microbe, a summary of all information, and references to relevant journals, scientific papers, or books.

Future development

As a dynamic database, the content of this database is kept updated and enriched from time to time, depending on the collected data. The addition of large amounts of data would lead microbiologists to more opportunities to conduct research in agriculture, medicine, and environments where the uniqueness of the microorganisms would lead to enhanced production of yield, immunization of host system, prevention and treatment of infectious disease, and reducing the infection in our environment, especially in a hospital. (Anand *et al*., 2022; Caselli, 2017). As a result, this can reinforce the importance of databases in managing large amounts of data by minimizing data redundancy to avoid storage, energy, and cost waste. It also facilitates data sharing and data accessibility among researchers and may prevent data leakage because the database is only accessible to certain groups.

Even though the built system is capable of managing large amounts of microbiological data, several improvements have been identified to improve the user. These include automated data formatting of submitted raw data directly from the system without manual editing. Aside from that, functional improvement through the integration of bioinformatics tools such as the Basic Local Alignment Tool (BLAST) (Altschul *et al*., 1990) would improve user experience by allowing the user to perform similarity searches on their sequences against biological sequence databases. In addition to that, inventory management functionality would also be introduced into the system to assist researchers in tracking and retrieving curated data.



(c)

Code ID	M106
Old Code	T2/32
Species Name	Trichoderma asperellum
Strain Name	T2/32
Description	From MPOB
Sampling Date	0000-00-00
Collector	Feodora
Sample Type	
First Culture Date	2019-09-05
Sampling Location	
GPS Location	

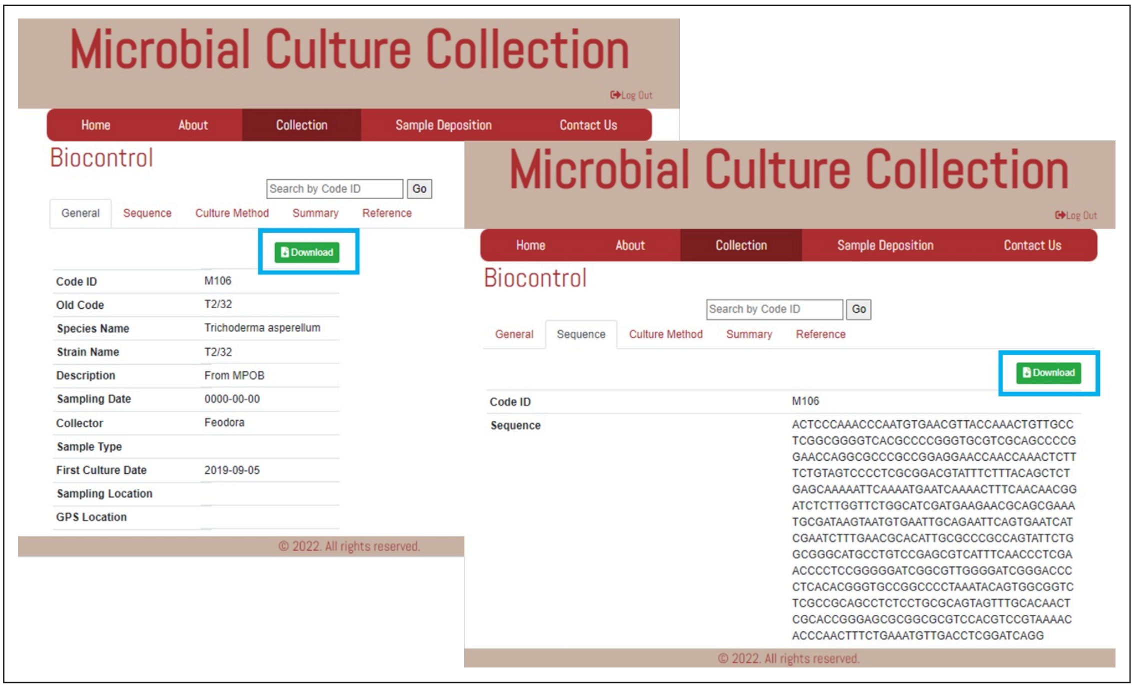
(d)

Code ID M106

Sequence

```
ACTCCCAAACCCAATGTGAACGTTACCAAACTGTTGCC
TCGGCGGGGTCACGCCCCGGGTGCGTCGCAGCCCCG
GAACCAGGCGCCCGCCGGAGGAACCAACCAAACTCTT
TCTGTAGTCCCCTCGCGGACGTATTTCTTTACAGCTCT
GAGCAAAAATTCAAAATGAATCAAAACTTTCAACAACGG
ATCTCTTGGTTCTGGCATCGATGAAGAACGCAGCGAAA
TGCGATAAGTAATGTGAATTGCAGAATTCAGTGAATCAT
CGAATCTTTGAACGCACATTGCGCCCGCCAGTATTCTG
GCGGGCATGCCTGTCCGAGCGTCATTTCAACCCTCGA
ACCCCTCCGGGGGATCGGCGTTGGGGATCGGGACCC
CTCACACGGGTGCCGGCCCCTAAATACAGTGGCGGTC
TCGCCGCAGCCTCTCCTGCGCAGTAGTTTGCACAACT
CGCACCGGGAGCGCGGCGCGTCCACGTCCGTAAAAC
ACCCAACTTTCTGAAATGTTGACCTCGGATCAGG
```

Fig. 4. Data retrieval process (a) Interface to select the microbial collection category (b) Search engine in the green box is to search the data by code identification (codeID) (c) Data display in the *general* sub-tab menu (d) The sequence of the codeID will be displayed in *sequence* sub-tab menu.



Code ID	M106
Old Code	T2/32
Species Name	Trichoderma asperellum
Strain Name	T2/32
Description	From MPOB
Sampling Date	0000-00-00
Collector	Feodora
Sample Type	
First Culture Date	2019-09-05
Sampling Location	
GPS Location	

Code ID	M106
Sequence	ACTCCCAAACCCAATGTGAACGTTACCAAACTGTTGCC TCGGCGGGGTCACGCCCCGGGTGCGTCGCAGCCCCG GAACCAGGCGCCCGCCGGAGGAACCAACCAAACTCTT TCTGTAGTCCCCTCGCGGACGTATTTCTTTACAGCTCT GAGCAAAAATTCAAAATGAATCAAAACTTTCAACAACGG ATCTCTTGGTTCTGGCATCGATGAAGAACGCAGCGAAA TGCGATAAGTAATGTGAATTGCAGAATTCAGTGAATCAT CGAATCTTTGAACGCACATTGCGCCCGCCAGTATTCTG GCGGGCATGCCTGTCCGAGCGTCATTTCAACCCTCGA ACCCCTCCGGGGGATCGGCGTTGGGGATCGGGACCC CTCACACGGGTGCCGGCCCCTAAATACAGTGGCGGTC TCGCCGCAGCCTCTCCTGCGCAGTAGTTTGCACAACT CGCACCGGGAGCGCGGCGCGTCCACGTCCGTAAAAC ACCCAACTTTCTGAAATGTTGACCTCGGATCAGG

Fig. 5. The retrieved data can be downloaded using the download button in the blue box.

CONCLUSION

In conclusion, generating a huge amount of data while keeping it internally with proper data management system is possible in an organization, agency, or even a research group. Here, we have shared our experience in the development of a systematic database for microbial data collection management as a key resource center for preserving microbial culture. With this platform, all microbial culture data will be stored systematically, making it easy for researchers to access and retrieve data.

ACKNOWLEDGEMENTS

The work was funded by FGV Agri Services Sdn. Bhd. The authors wish to express our gratitude to the CEO and researchers of FGV R&D Sdn. Bhd. for the given support.

CONFLICT OF INTEREST

The authors declare no conflict of interest.

REFERENCES

Amza, C., Cox, A.L. & Zwaenepoel, W. 2003. Distributed versioning: Consistent replication for scaling back-end databases of dynamic content web sites. Lecture Notes in Computer Science, 2672: 282–304.

Altschul, S.F., Gish, W., Miller, W., Myers, E.W. & Lipman, D.J. 1990. Basic local alignment search tool. Journal of Molecular Biology, 215(3): 403–410. https://doi.org/10.1016/S0022-2836(05)80360-2

Anand, U., Vaishnav, A., Sharma, S.K., Sahu, J., Ahmad, S., Sunita, K., Suresh, S., Dey, A., Bontempi, E., Singh, A.K., Proćków, J. & Shukla, A.K. 2022. Current advances and research prospects for agricultural and industrial uses of microbial strains available in world collections. Science of The Total Environment, 842: 156641.

Caselli, E. 2017. Hygiene: microbial strategies to reduce pathogens and drug resistance in clinical settings. Microbial Biotechnology, 10(5): 1079–1083.

Cowan, M. K. 2012. Microbiology: A Systems Approach. McGraw-Hill.

Gaikwad, S.S. & Adkar, P. 2019. A Review Paper on Bootstrap Framework. IRE Journals, 2(10): 349–351.

Microsoft. 2016. Excel specifications and limits [WWW Document]. Microsoft Support. URL https://support.microsoft.com/en-us/office/excel-specifications-and-limits-1672b34d-7043-467e-8e27-269d656771c3 (accessed 10.9.22).

Supaartagorn, C. 2011. PHP framework for database management based on MVC pattern. International Journal of Computer Science and Information Technology, 3(2): 251–258. https://doi.org/10.5121/ijcsit.2011.3219

Zhulin, I. B. 2015. Databases for Microbiologists. Journal of Bacteriology, 197(15): 2458–2467.