Turk. J. Fish.& Aquat. Sci. 26(3), TRJFAS26139
https://doi.org/10.4194/TRJFAS26139

PROOF
RESEARCH PAPER

Turkish Journal of FISHERIES and AQUATIC SCIENCES

Development of Fish Metabolome Database: Expanding Access and Advancing Management of Metabolite Data Resources

Andre Sihombing[1], Ira Maryati[3,*], Giri Rohmad Brokah[2], Lia Sadita[3], Harry B. Santoso[4], Anne Parlina[1]

[1]Research Center for Data and Information Science, National Research and Innovation Agency, Bandung, Indonesia.
[2]Research Center for Food Technology and Processing, National Research and Innovation Agency, Serpong, Indonesia.
[3]Directorate of Scientific Collection Management, National Research and Innovation Agency, Jakarta, Indonesia.
[4]Faculty of Computer Science, Universitas Indonesia, Depok, Indonesia.

How to Cite

Sihombing, A., Maryati, I., Brokah, G.R., Sadita, L., Santoso, H.B., Parlina, A. (2026). Development of Fish Metabolome Database: Expanding Access and Advancing Management of Metabolite Data Resources. *Turkish Journal of Fisheries and Aquatic Sciences*, *26(3)*, *TRJFAS26139.* https://doi.org/10.4194/TRJFAS26139

Article History
Received 24 June 2024
Accepted 31 July 2025
First Online 28 August 2025

Corresponding Author
E-mail: iram001@brin.go.id

Keywords
Fish metabolome database
Fisheries data
Preferred reporting items for systematic reviews and meta-analyses
Data management

Abstract

This study designs the Fish Metabolome Database (FMDB) as a repository of fish metabolite data. The data is valuable for fisheries research, such as fishery product safety, water pollution, and fish interactions with the environment. The database was designed with the requirements generated from previous research related to metabolite data management and adapted to the datasets. The Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) strengthened by benchmark at the synthesis stage are used to gather the requirements. The proposed data fields and features have been implemented in a prototype database containing 91 datasets of fish metabolites. Its simple functionality makes it easy to search for databases on species keywords or metabolite names. The limited amount of data is a research limitation that will continue to be developed to improve the database. However, as there is currently no existing database in that field, FMDB becomes a foundation for managing fish metabolite data that is useful for further research and education in the field of fisheries. This research also provides insight for professionals and practitioners in database management with requirements based on literature reviews.

Introduction

Indonesia, as a maritime country, produces fishery products in large quantities. Statistical data released by the Ministry of Marine Affairs and Fisheries Republic of Indonesia shows total marine catch fishery production of 6.4 million tons in 2020 (Ministry of Marine Affairs and Fisheries Republic of Indonesia, 2023). According to data released by FAO, the top ten global fisheries marine capture producers, this number is ranked second highest after China (FAO, 2022). The large amount of production creates challenges in maintaining quality and safety. Consequently, fish metabolite data and research are closely related to this issue.

Fish metabolite data is valuable information that provides insights into physiological and biochemical processes occurring in fish. The data will be very useful in various studies, such as for biomarkers or detecting deviations. Metabolite data is used to predict the effect of environmental factors on fish quality, such as changes in water temperature (Shima et al., 2023) (Ma et al., 2023), chemical contamination (Yan et al., 2023), and so on.

Metabolomic technology is also used to assess the quality and safety of seafood products (Power et al., 2023). Chemical contamination is still one of the significant causes of disease and health problems for consumers (Joshi et al., 2015). One of the chemical contaminants that needs serious attention is formaldehyde. Formaldehyde is a genotoxic compound because it can cause DNA-protein and protein-protein cross-linking when it enters the human body (Yeh et al.,

2013). The presence of natural formaldehyde in some fishery products can make it difficult to detect the illegal addition of artificial formaldehyde (contamination) (Sanyal et al., 2017). Research on biomarkers of natural formaldehyde and artificial/synthetic formaldehyde in fishery products is crucial. Metabolic biomarkers are biological marker compounds in living things that experience abnormalities and changes due to a response to foreign compounds that enter the living body and environmental changes (Smolinska et al., 2012). Metabolic biomarkers have been widely used in fishery products to characterize and identify target metabolites, such as amino acids and nucleic acids, that occur in response to toxic compounds that enter the fish body (Shumilina et al., 2015). With different properties and biological systems, the biomarker metabolites that can be used to distinguish contaminants in fish species, such as natural formaldehyde from artificial formaldehyde in other types of fish will likely be different. In addition, the fact that fishery products contain various natural contaminants like formaldehyde is confusing when positive samples contain formaldehyde in the field. A database of biomarker metabolites that is more extensive and can represent fish product species based on their habitat characteristics needed to be used as a reference data bank in distinguishing natural contaminants, especially formaldehyde from synthetic fishery products.

Studies related to the development of species-specific metabolite databases have been carried out. The database development that is widely used as a reference is the Human Metabolome Database (HMDB) 5 (Wishart et al. 2007, 2013, 2018, 2022). Animal metabolite databases have also been developed for research (Yanshole et al., 2023). For plant metabolites, a database has been created in the three-dimensional structure of all metabolic compounds (Sakurai et al., 2023). Unfortunately, a database that manages and provides public fish metabolite data is unavailable.

A database of biomarker metabolites that is more extensive and can represent fish product species based on their habitat characteristics is needed to be used as a reference data bank in distinguishing natural contaminants, especially formaldehyde, from synthetic fishery products. Designing a fish metabolome database is crucial to cope with that need. Therefore, this research developed a fish metabolome database (FMDB) based on primary fish metabolite data.

Related Works

In previous research, metabolome databases have been developed to manage metabolite data of various species with their respective uniqueness and chemically distinct metabolites. The objectives of this study closely align with the earlier research. The aim is to leverage metabolite profiling to shed light on the functions of genes or metabolites of interest. Those efforts resulted in the development of the metabolite profiling database, which serves as a valuable resource for functional genomics. In line with the study, it seeks to address a pressing need in fish metabolic research by creating a comprehensive database that can assist researchers in characterizing and understanding the metabolites present in various fish species. This study also recognizes the importance of freely accessible data and tools in advancing their respective fields, underlining the critical role of databases in hypothesis generation and gene/metabolite annotation. Several following metabolite databases can offer valuable insight into the development of FMDB.

The database development that is widely used as a reference is the Human Metabolome Database (HMDB). This database was first released in 2007 and continues to be developed until its updated version in 2022 (Wishart et al., 2007)(Wishart et al., 2013), (Wishart et al., 2018) (Wishart et al., 2022). This database has very comprehensive coverage regarding human metabolites and their associated biological, physiological, and chemical properties. Its massive development makes this database valuable to follow both in terms of its interface and features. Related to human metabolites, other research is developing a database the BioM2MetDisease, a manually curated resource that focuses on establishing a comprehensive and experimentally supported collection of associations between metabolic diseases and various bio molecules. The database offers a user-friendly interface for easy data exploration and retrieval, and it even provides a submission page for researchers to contribute new associations between bio molecules and metabolic diseases (Xu et al., 2017).

For plant metabolites, a database has been created that provides the three-dimensional structure of all metabolic compounds. This research uses Balloon software and Chem3D software to convert 2D structures into 3D. This database is very interesting because the 3D structure is more informative and downloadable (Nakamura et al., 2013). More specific plant metabolite databases have also been developed, such as for Arabidopsis (Fukushima et al., 2014) and for metabolites in wheat fusarium head blight disease (Surendra & Cuperlovic-Culf, 2017).

Metabolome databases were also developed for microbial metabolites specific to certain microbes. The Escherichia coli Metabolome Database (ECMDB) collects all E. Coli metabolite data curated from the literature (Guo et al., 2013). Another study developed a database for the chemical fingerprint of marine bacteria based on the LC-MS analytical platform (Lu et al., 2014).

Another study developed The Metabolomics Workbench, which serves as a public repository for metadata and experimental data in metabolomics, encompasses various species and experimental platforms, metabolite standards, metabolite structures, protocols, tutorials, and other educational resources. The Metabolomics Workbench provides a computational platform to integrate, analyze, track,

deposit, and disseminate large and diverse volumes of data from various metabolomics studies, including mass spectrometry (MS) and nuclear magnetic resonance spectrometry (NMR). Data spanning over 20 different species covering all major taxonomic categories, including humans and other mammals, plants, insects, invertebrates, and microorganisms (Sud et al., 2016).

These studies provide insight and several principles that can be applied in developing a fish metabolite database to enhance the management and understanding of heterogeneous fish metabolite data. A deeper study was carried out on the systematic literature review to identify features in the previous metabolomics database development as the input to requirements gathering.

Materials and Methods

The research as a whole consists of two main activities: collecting fish metabolite datasets and database development. The metabolite datasets were obtained by spectrophotometer UV-Vis (Thermo Scientific Multiscan GO 200-800 nm) and Nuclear Magnetic Resonance (NMR) (Jeol ECS 400 MHz) from several fresh fish samples in Indonesia. The project involves laboratory research activities at the Science Center Area, National Research and Innovation Agency. This data will be inputted into the database that has been developed. In the current study, the focus of discussion is on the development of the database rather than the dataset collection process itself. Thus, this study provides a comprehensive understanding of the metabolome database development process.

Requirement Elicitation

This stage combines a literature review, benchmark, and observations of available datasets to gain comprehensive insight regarding requirements and functionalities. Literature reviews have been popularly used for requirements gathering in information systems development. This method has advantages in terms of saving time and costs, providing updated requirements, and opportunities to extend requirements (Kumar & Goundar, 2023). Then, the snowball method is utilized to determine several metabolome databases used as benchmarks to enrich insight related to requirements. The final step is adjustment and comparison with the existing datasets.

The Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) method is used as a guide in conducting a literature review (Cacciamani et al., 2023), (PRISMA, 2020). Mendeley is a utilized tool for managing works of literature. A combination of several keywords was used for literature searches in several databases (ScienceDirect, SpringerLink, Emerald, ProQuest, IEEE Xplore, and open-access databases), traced through the Scopus indexer. The keyword combinations are “metabolite AND database”, “metabolome AND database”, “metabolomic AND database”, “metabolite AND repository”, “metabolome AND repository”, and “metabolomic AND repository”. We identified 109 pieces of literature related to metabolite database development. Identical screening left 106 pieces of literature included. The rigorous assessment process resulted in 29 records that were analyzed and synthesized. The criteria used for literature synthesis and analysis are data fields and functionalities. Data field analysis focuses on identifying the data elements stored in the database. As for functionalities analysis, it is related to how the database is used in the system.

From the reviewed literature, several metabolome databases emerged that were determined to be further explored and used as benchmarks in FMDB development. Benchmarking has been widely applied in requirements gathering, such as comparing tools and algorithms. According to Beyer, et al. (2019), two things that must be paid attention to in a suitable benchmarking method are resource determination and resource presentation. The resources that will be used as benchmarks in this research are several metabolome databases whose development processes have been published so that comprehensive insights can be obtained. Refinement requirements are carried out to ensure that the review results are in accordance with the objectives of database development and whether they can be implemented properly. Several aspects are used in the refinement of requirements. First, the validation aspect is carried out by ensuring that the data field is the same or in accordance with the data owned. The second aspect is technical feasibility; this aspect examines the extent to which functionalities can be implemented in accordance with the scope of research and available resources so that priority is given to the list of functionalities obtained. Refinement requirements are also carried out by looking at user feedback through system evaluation.

Documentation of requirements is carried out in detail by modeling data using an Entity Relationship Diagram (ERD), which can describe the relationship between entities and attributes in the database (Figure 1). The functionality is depicted in a use-case diagram model (Figure 2).

Design and Implementation

At this stage, the database and user interface are designed based on the requirements established in the previous stage. The actual development of the database commences. This phase includes the creation of database tables, data storage setup, and implementation of data entry and retrieval mechanisms. The design of the database table has completed the normalization process. For instance, we made the transaction table "fmdb" to ensure consistency and avoid redundancy. Primary keys are defined by using auto-increment integer types to ensure unique

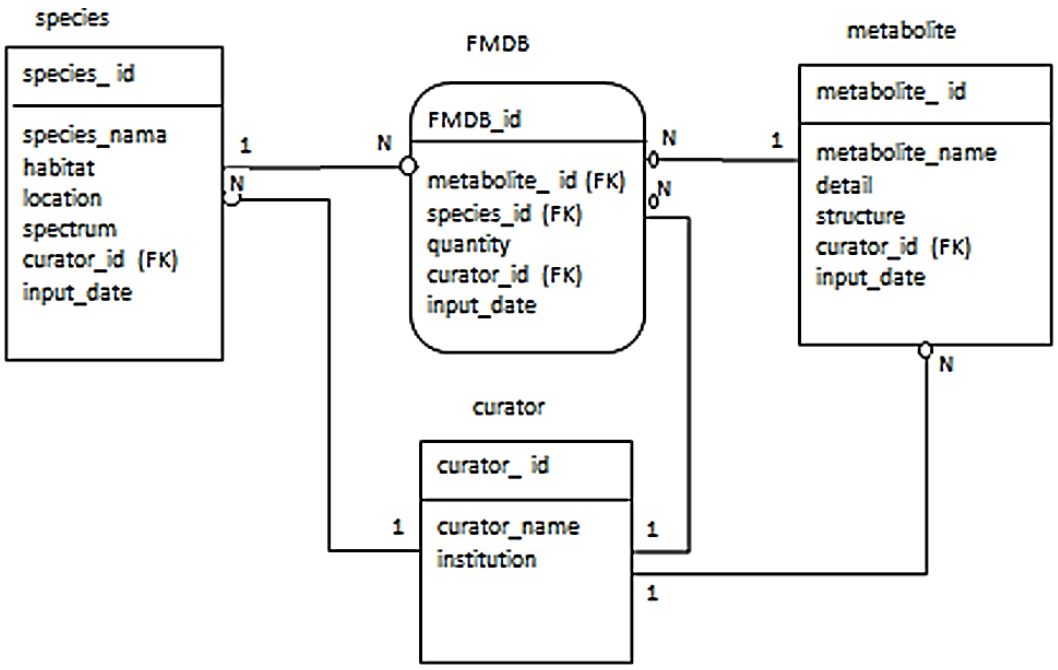

Figure 1. Entity relational diagram.

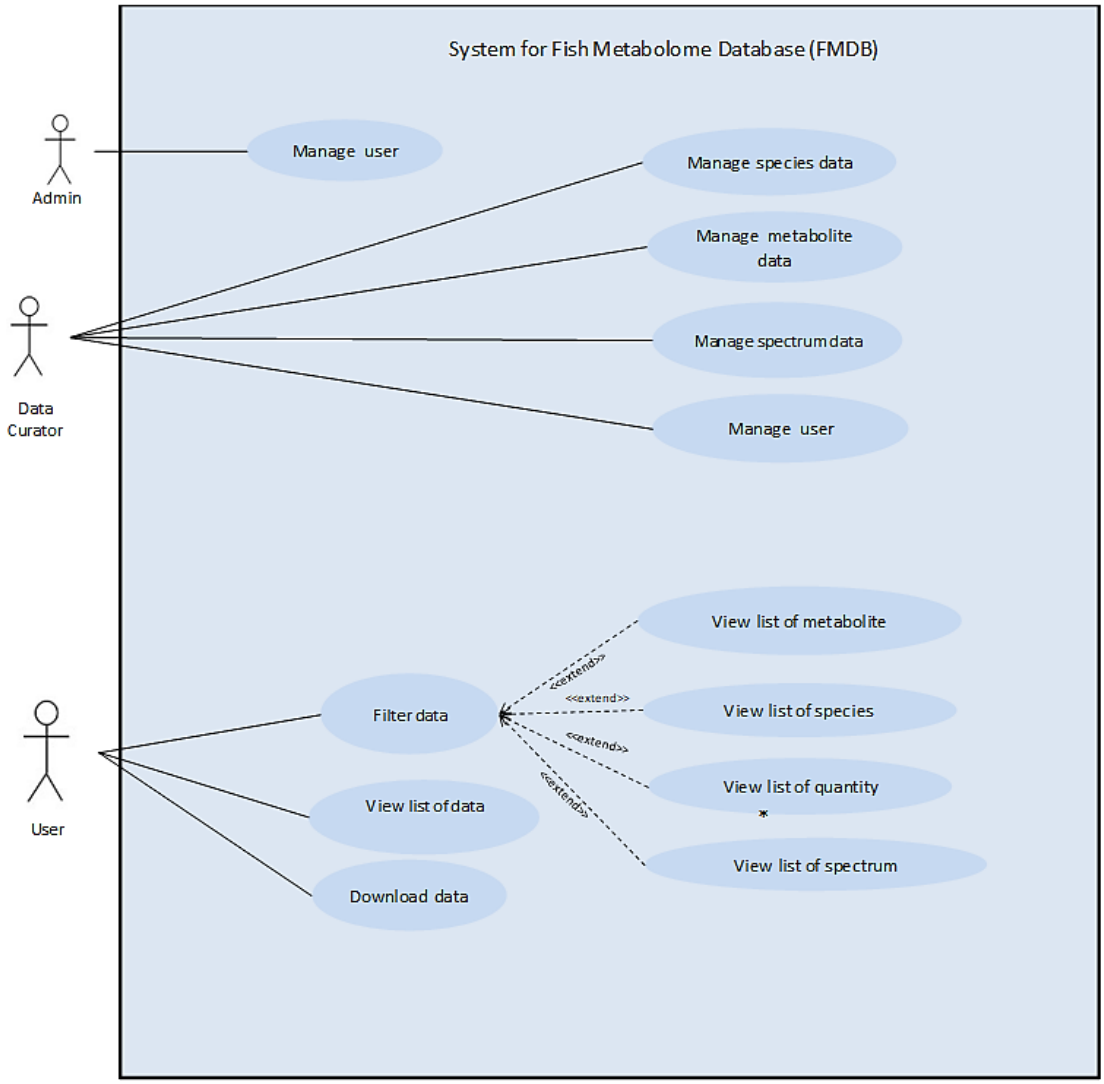

Figure 2. Use-case diagram of Fish Metabolome Database (FMDB).

identification and foreign keys are used to maintain integrity between related tables. Based on the nature of the data stored, specific fields, such as metabolite_id, species_id, timestamps, and username, are assigned appropriate data types (e.g., INT, VARCHAR, DATETIME).

To enhance query performance and ensure scalability, the primary keys and frequently queried columns implemented B-tree indexes provided by MySQL indexes columns. In addition, for search functionality, we have applied full-text indexing in text-based fields where applicable. The front end adopted the Bootstrap framework for user convenience, automatically adapting to various screen sizes using its grid system. The Input forms also incorporate real-time validation via JavaScript to provide immediate feedback and reduce user errors. PHP and MySQL are used in database development. PHP and MySQL were chosen because of their proven stability in developing web-based applications and their compatibility with the infrastructure used in this study. This study uses data with structured relationships (with foreign keys), so MySQL is more appropriate compared to other databases, such as NoSQL which are more suitable for semi-structured data. In addition, the development team in this study was more experienced in this technology, as well as the need for a more relational transaction-oriented system.

Evaluation

The evaluation consists of functionality testing and usability evaluation. The system usability scale (SUS) questionnaire which has been adapted to Indonesia is used in usability evaluation (Sharfina & Santoso, 2016). The questionnaire consists of ten questions with a point Likert scale with odd numbers for positive items and even numbers for negative (Table 1). The SUS score is calculated in the following way, for odd numbers minus 1 while for even numbers, which is 5 minus the score obtained. The sum of item score contributions multiplied by 2.5, ranging from 0 to 100 [6]. A product is considered to have good usability if the overall SUS score is equal to or above 68. The evaluation involved ten student participants from the Faculty of Fisheries. While there is no strict limit on the number of participants, (Sauro & Lewis, 2012) suggest that at least 10-15 participants can provide useful insight, which is a practical and effective minimum for the early evaluation (Sauro & Lewis, 2012).

Results

This study presented the Fish Metabolome Database (FMDB) development process. The proposed data fields and features extracted from previous research are implemented in a web-based system as an initial prototype.

System Requirements

The requirements extracted from the literature are synthesized based on two criteria, namely data fields and functionalities. A total of 29 works of literature passed the assessment stage and were determined to be reviewed. The results of literature extraction are presented in Table 2.

The literature extracted related to the data field produced seven groups, as shown in Table 2. These data fields are presented with user interface features that make retrieving the data easier. The literature extraction also generates eight groups of features. Some literature is explored more deeply for website benchmarking via the included links. This study involved sixteen databases that were explored and used as benchmarking, as shown in Table 2. Benchmarking provides an obvious representation of the criteria extracted in the literature.

The results are sufficient to produce insight regarding the data field, which is ideally provided in managing metabolite data. Although not all data fields can be implemented in the initial prototype due to research limitations, the rest can be used as input for subsequent development. The requirements of FMDB are modeled in a Use Case Diagram (Figure 2.), and the data fields are modeled in an Entity Relationship Diagram (ERD) (Figure 1.).

Design and Implementation of FMDB

Implementation of FMDB development begins with creating interface designs. The FMDB website is developed using PHP programming language as the core

Table 1. The questions in the S ystem Usability Scale (SUS) questionnaire

	Question	Strongly disagree				Strongly agree
		1	2	3	4	5
1	I thought I would use this FMDB database again.					
2	I found this FMDB database complicated to use.					
3	I thought the FMDB database was easy to use.					
4	I need help from other people or technicians to use this FMDB database.					
5	I found the features on this FMDB database work well.					
6	I thought there was too much inconsistency in this FMDB database.					
7	I felt like other people would quickly understand how to use this FMDB database.					
8	I found this FMDB database confusing.					
9	I felt very confident using the FMDB database					
10	I needed to learn a lot of things before I could get going with the FMDB database.					

Table 2. Literature extraction results for requirements gathering

No	Article	Species/ Sample	Data Fields							Features								Benchmark included		The Objective
			a	b	c	d	e	f	g	1	2	3	4	5	6	7	8	Yes	No	
1	(Sakurai et al., 2023)	Food, Plant, and Thing metabolite	√	√	√	√	√			√	√	√	√						√	Create a database for comparison in determining components based on their metabolite similarities.
2	(Neveu et al., 2023)	Human gut microbial	√	√	√		√	√					√	√				√		Develop a database on the gut microbial exposome that can be used to identify microbial metabolites in metabolomics datasets or to develop specific assays.
3	(Yanshole et al., 2023)	Animal	√	√	√		√	√	√		√	√	√	√				√		Provides a metabolite database of various animal species.
4	(Wishart et al., 2023)	Human microbial	√	√	√		√	√	√	√	√	√		√		√	√	√		Develop a comprehensive, multi-omic, microbiome resource that connects all microbial metabolites to human health.
5	(Cheng et al., 2022)	Gut microbes (Human & Mouse)	-	√	√	-	√	√	-	-	√	√	√	-	√	√	-	√	-	Provides a comprehensive resource of target genes of gut microbes and microbial metabolites in humans and mice
6	(Wu et al., 2022)	Lung cancer metabolome	√	-	√	-	√	√	√	-	√	√	√	√	-	-	-	√	-	Provides a database of metabolites related to lung cancer
7-9	(Wishart et al., 2022), (Wishart et al., 2013), (Wishart et al., 2018)	Human Metabolome	√	√	√	-	√	√	√	√	√	√	-	√	-	-	-	√	-	Provides comprehensive data on human metabolome
10	(Ara et al., 2021)	Tomato	-	√	-	√	-	√	-	√	√	-	-	-	-	-	-	-	√	Provides manual curation of metabolites from 25 mature tomato cultivars.
11	(Vivek-Ananth et al., 2021)	Medicinal Fungi	-	√	√	-	√	√	√	√	√	√	√	√	√	-	-	√	-	Provide secondary metabolites and therapeutic uses of medicinal fungi.
12	(Jones et al., 2021)	Cyanobacteria	-	√	√	-	√	-	-	-	-	√	-	√	-	-	-	-	√	Provide a comprehensive, publically-accessible database detailing cyanobacterial secondary metabolites
13	(Gil-De-la-fuente et al., 2021)	Aspergillus	-	-	√	-	√	√	-	√	√	√	-	√	-	-	-	-	√	Provide a free online resource to perform metabolite annotation in mass spectrometry studies devoted to the genus Aspergillus
14	(Li et al., 2020)	Mulberry	√	√	√	√	√	√	-	√	√	-	-	√	√	-	-	√	-	Provides important and comprehensive metabolome data for scientists working with mulberry
15	(Blin et al., 2019)	Microbe	-	√	√	√	√		√	√	√	√	√	√	-	-	√	√	-	Provide a database that contains annotations for full bacterial genomes and bacterial draft genomes.
16	(Karu et al., 2018)	Human fecal	√	√	√	√	-	√	√	√	√	√	-	√	-	-	-	√	-	Provides comprehensive data on the human fecal metabolome and the methods to characterize it.
17	(Huang et al., 2018)	Pseudomonas aeruginosa	-	√	√	-	-	√	-	√	√	√	√	√	-	-	-	√	-	Develop a searchable, richly annotated metabolite database specific to P. Aeruginosa.
18	(Xu et al., 2017)	Metabolite disease	√	√	√	-	√	√	√	√	√	√	-	√	-	-	√	-	√	Develop a resource of associations between metabolic diseases and biomolecules.
19	(Vadlapudi et al., 2017)	Aspergillus genus	√	√	√	-	-	√	√	√	√	√	-	√	-	-	-	-	√	Provide a curated compendium of information on Aspergillus & its secondary metabolome.
20	(Ramirez-Gaona et al., 2017)	Yeast (Saccharomyces cerevisiae)	-	√	√	-	√	√	-	√	√	√	-	√	-	√	-	√	-	Provide a database containing extensive information on the genome and metabolome of Saccharomyces cerevisiae.
21	(Blin et al., 2017)	Microorganism secondary metabolites	-	√	√	√	√	-	-	√	√	√	-	-	-	-	√	√	-	Developed the antiSMASH database to browse the bacterial genomes and perform advanced search queries combining multiple search criteria.
23	(Surendra & Cuperlovic-Culf, 2017)	Fusarium head blight	-	√	√	-	√	√	√	√	√	-	-	-	-	-	√	-	√	Provide a database to determine possible targets and roles of metabolite in resistance to FBH and aid in the development of related synthetic antifungals.
24	(Sud et al., 2016)	Heterogeneous species	√	√	√	√	√	√	-	√	√	√	√	√	√	-	-	√	-	Provide a public repository for metabolomics metadata and experimental data spanning various species.
25	(Fukushima et al., 2014)	Arabidopsis	√	√	√	-	√	√	-	√	√	√	√	-	√	-	-	-	√	Developed the Metabolite Profiling Database for Knock-OutMutants in Arabidopsis (MeKO).
26	(Deborde & Jacob, 2014)	Plant	√	√	√	-	√	√	-	√	√	√	-	-	√	-	-	√	-	Provide a platform for plant H-NMR metabolomic profiles.
27	(Lu et al., 2014)	Marine bacteria	√	√	√	-	√	-	√	√	√	-	√	-	√	-	√	-	√	Develop a chemical fingerprint database of marine bacteria based on their secondary metabolite profiles.
28	(Guo et al., 2013)	Escherichia coli	√	√	√	-	√	√	√	√	√	√	-	√	-	-	-	√	-	Develop a comprehensive annotated metabolomic database of E. Coli.
29	(Nakamura et al., 2013)	Plant	√	√	√	-	-	√	-	√	√	√	-	-	-	-	-	-	√	Develop a database containing three–dimensional (3D) structures of all of the plant metabolic compounds.

Notes: Data fields: a. Sample-specific localization (tissue specificity and taxonomic relationships), b. Visualization for the data (mass spectral, chart, and coverage of the metabolite peaks), c. Ion/metabolite information, d. Retention time, e. Treatment (analytical method, test equipment, sample treatment, etc), f. MetaboCard (Concentration/ quantity page), g. Association with disease.
Features: 1. Browse, 2. Search (more specific search), 3. Download data or figure, 4. Comparison visualization, 5. Bibliographic reference, 6. Upload data (for comparison with the existing data), 7. Auto-suggest search, 8 Clustering data (Tree browser, map, histogram).

business logic and functionality. The website consists of a front end and a back end. The back-end website consists of administrator pages and search pages that can be accessed only by registered users (data curator). The data curator's task is to perform the CRUD (Create, Read, Update, Delete) operations for the fish metabolite data. The front-end page is designed for users to find related information about the fish metabolite database. On these pages, users can use simple and advanced search.

Advanced search functionalities are search features that allow users to filter, sort, or narrow search results based on two categories. FMDB advanced search provides an option to search based on specific species names or metabolite names. This aims to improve the accuracy and relevance of search results according to user needs. The advanced search interface can be seen in Figure 3.

The front-end web user interface consists of HTML, CSS, and JavaScript code. The data is managed using the database management system MySQL (https://www.mysql.com/). It is responsible for managing data storage, retrieval, and manipulation to perform CRUD operations. The front-end and the back-end interface design can be seen in Figure 4. and Figure 5.

A total of 91 fish metabolite datasets have been successfully input into the database. The unique datasets from each research sample is stored in a data field, that in some databases are analogous to MetaboCard (Deborde & Jacob, 2014). FMDB adopts this concept by creating a data field with a unique ID for each datasets. The limited number of datasets will continue to be developed by considering data curated from the literature. Several studies have demonstrated the development of a metabolite database sourced from the literature. Further research is needed to review potential data availability and appropriate data extraction techniques. This may also generate the need to add the required data fields.

Evaluation of FMDB

Functionality tests show that all assigned tasks run successfully. Tasks that were successfully carried out were creating an account, managing fish species data (add, change, and delete), managing metabolite type data (add, change, and delete), adding metabolite

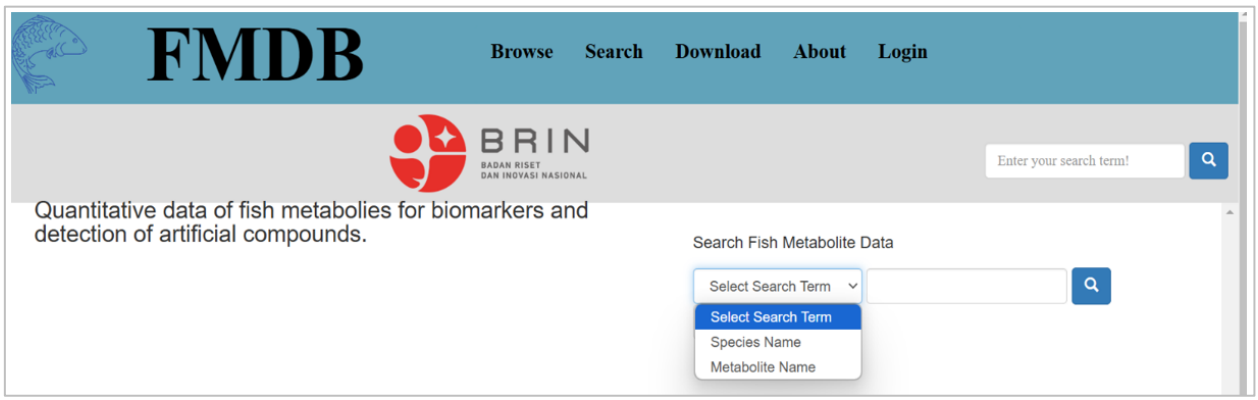

Figure 3. Advanced search of the FMDB user interface.

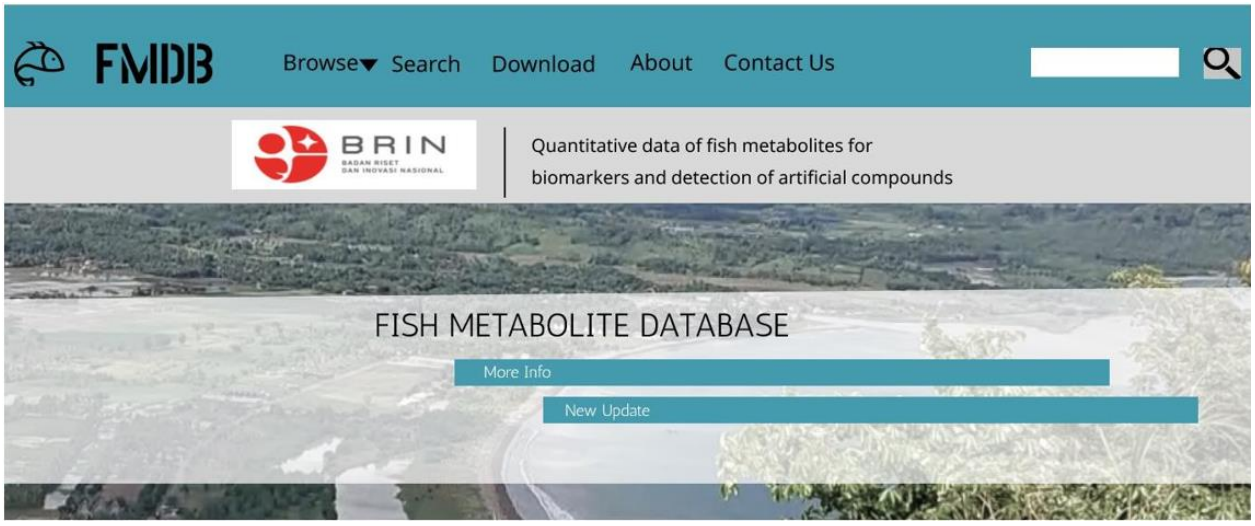

Figure 4. The front end of the FMDB user interface.

quantity data, and searching for species and metabolites on the front page and back end page. This test ensures that all the expected functions are running well and that the database can be used according to the specified requirements.

Usability testing carried out by ten students from the fisheries faculty using the SUS questionnaires gave an average score of 70.5, higher than the set standard of 68 (Sharfina & Santoso, 2016). The evaluation results indicate that FMDB can be used effectively, efficiently, and satisfactorily by users. The final part of the evaluation also provides an optional question column that accommodates the obstacles faced by users in using FMDB. The obstacle faced by users in managing metabolite data is the input of metabolite structures in the form of images. Reference metabolite data from existing databases such as PubChem and HMDB is very useful and needs to be complemented in further development.

Discussion

The adoption of data fields and features in the FMDB considering the suitability of the available fish metabolite data. The features and data fields are discussed simultaneously as both of them are interrelated. Sample-specific localization data fields are available on metabolite databases in humans, animals, and plants. Specific information regarding the location of the research sample (skin tissue, blood, etc.) will produce accurate data. In some databases, this field is also used as a specific search feature. In this study, where data is still limited to tissue samples, this field has not appeared in specific searches. However, the sample-specific localization data fields are provided (Deborde & Jacob, 2014; Wishart et al., 2022; Yanshole et al., 2023).

Some databases also present the data in visual form, image, or graph. Mass spectral is visual data that shows metabolite peaks produced by a UV-Vis spectrophotometer (Thermo Scientific Multiscan GO 200-800 nm) and Nuclear Magnetic Resonance (NMR). Data is also displayed in chart form to show the comparison of metabolite composition. Retention time data is also presented in graphical form to make it easier to compare one metabolite data with another (Sakurai et al., 2023). FMDB implements data visualization by providing mass spectral images. While the current database includes basic visualization images, future research could focus on enhancing interactive data exploration features. Implementing dynamic visualization tools like heat maps, interactive tables, and real-time filtering would allow users to analyze data more intuitively and gain deeper insights. Additionally, integrating web-based visualization frameworks could improve user engagement and facilitate pattern recognition within large datasets. These improvements would make the system more versatile and user-friendly for a broader range of applications."

Several databases provide detailed metabolite information by referring to various existing databases. Databases that are widely referenced include PubChem Compound released by the National Library of Medicine and HMDB released by the Canadian Institutes of Research. Hyperlinks are provided to refer directly to detailed information on metabolite compounds (Wishart et al., 2023).

Metabolite data in humans and animals is closely related to disease. Previous database developments have provided specific data fields related to disease relationships (Vivek-Ananth et al., 2021; Wishart et al., 2023; Wu et al., 2022; Yanshole et al., 2023). In FMDB, this data field can be used not only concerning fish disease but also for deviations caused by chemical contamination.

These data fields are presented with features on the user interface that make it easier to retrieve the

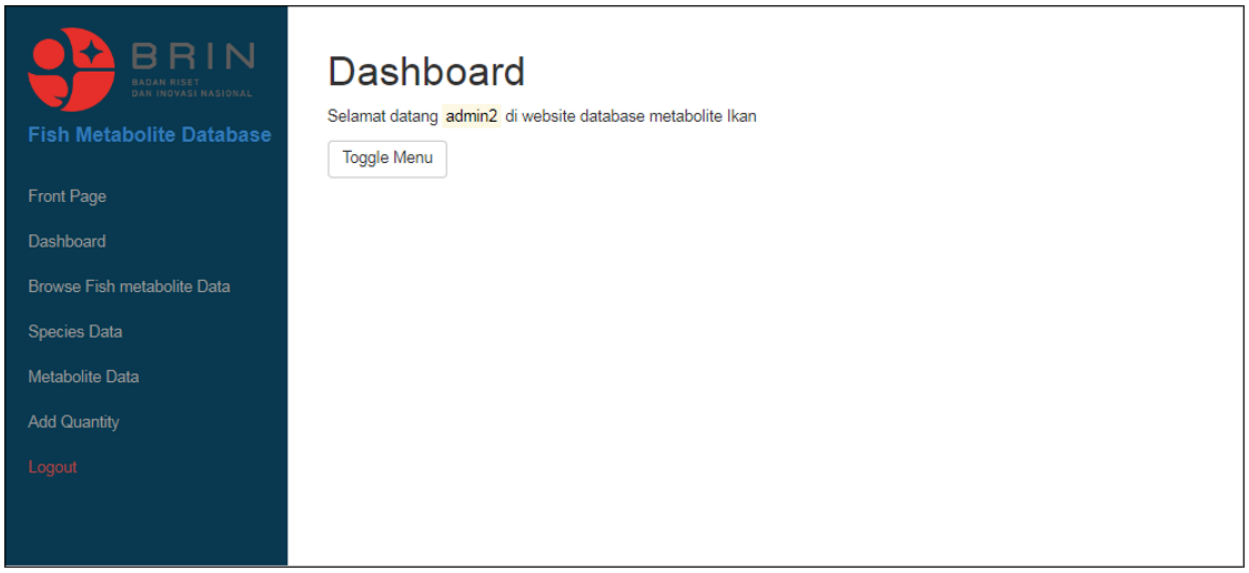

Figure 5. The back end of the FMDB interface.

data. The advanced features support the approximate' text queries utilizing tolerance and re-interpretation the typographical errors and mis-spellings. In some databases, browsing is also facilitated by data clustering (Blin et al., 2017; Wishart et al., 2022; Xu et al., 2017).

Some databases also provide a feature for downloading raw data, either in images or in other formats such as .csv and .xlsx (Cheng et al., 2022; Wishart et al., 2022; Wu et al., 2022; Yanshole et al., 2023). FMDB provides mass spectral images that can be downloaded by users. Comparing user data with available data is also a feature that has been developed previously. This is useful for various purposes, such as species identification and biomarker research. For this purpose, user data upload and data comparison visualization features are generally supported (Cheng et al., 2022; Huang et al., 2018; Neveu et al., 2023; Sakurai et al., 2023; Yanshole et al., 2023). This feature also requires quite large data involving a machine learning process.

The curation of data from the literature supports the collection of required data. The resulting data is linked to each bibliographic reference source. This provides significant input in the development of the FMDB. The limited data on fish metabolites in this study need to be enriched with data from the literature. Further research related to literature-based data development needs to be carried out.

Conclusions

FMDB, a database for managing fish metabolite data, has been developed with requirements obtained from a literature review. The extraction of literature presents seven data field groups and eight features. A deeper exploration of benchmarking and taking into account limited research data, this study selects which requirements are implemented in the initial prototypes which cannot yet be adopted are proposed for further development.

The FMDB implementation is in the form of a web-based application that consists of two interfaces, namely front-end for general users and back-end for data curator users. The features provided are adequate for managing and retrieving fish metabolite data. The evaluation has also shown good results in terms of functionality and usability. The functionality test shows that all requirements have been running well, as well as the usability test which gives a score of 70.5 above the standard SUS score of 68. Thus, it is hoped that this database will be useful for various research in the fisheries field.

FMDB still has opportunities for further development. The development of a database that combines the results of data extraction from related literature needs to be considered as subsequent development. This can expand data from research that has been conducted previously. More extensive data mining and the use of natural language processing (NLP) of text references need to be implemented for this purpose. Expanding the dataset can support advanced analyses that involve developing analytical capabilities.

Ethical Statement

The publication is approved by all authors and tacitly or explicitly by the responsible authorities where the work was carried out. And authors' institutions have no conflicts of interest.

Funding Information

The Research was funded by the Organization of Electronics and Informatics, National Research and Innovation Agency. Funding Program for Artificial Intelligence, Big Data, and Computing Technology for Biodiversity and Satellite Images (Grant No.: 1/III.6/HK/2023.

Author Contribution

Ira Maryati: research concept, methodology, and writing-original draft preparation; Giri Rohmad Barokah: research concept, data analysis, and reviewing; Andre Sihombing: software development, writing, and reviewing; Lia Sadita: writing and reviewing; Harry B. Santoso: writing and reviewing; Anne Parlina: writing and reviewing.

Conflict of Interest

The authors declare that they have no known competing financial or non-financial, professional, or personal conflicts that could have appeared to influence the work reported in this paper.

Acknowledgements

The authors would like to acknowledge the Research Organization of Electronics and Informatics, National Research and Innovation Agency. Funding Program for Artificial Intelligence, Big Data, and Computing Technology for Biodiversity and Satellite Images (Grant No.: 1/III.6/HK/2023) to provide funds for research experiments.

References

Ara, T., Sakurai, N., Takahashi, S., Waki, N., Suganuma, H., Aizawa, K., Matsumura, Y., Kawada, T., & Shibata, D. (2021). TOMATOMET: A metabolome database consists of 7118 accurate mass values detected in mature fruits of 25 tomato cultivars. *Plant Direct*, *5*(4). https://doi.org/10.1002/pld3.318

Beyer, D., Löwe, S., & Wendler, P. (2019). Reliable benchmarking: requirements and solutions. *International Journal on Software Tools for Technology Transfer*, *21*(1), 1–29. https://doi.org/10.1007/s10009-017-0469-y

Blin, K., Medema, M. H., Kottmann, R., Lee, S. Y., & Weber, T. (2017). The antiSMASH database, a comprehensive database of microbial secondary metabolite biosynthetic gene clusters. *NucleicAcidsResearch*,*45*(D1), D555 D559. https://doi.org/10.1093/nar/gkw960

Blin, K., Pascal Andreu, V., De Los Santos, E. L. C., Del Carratore, F., Lee, S. Y., Medema, M. H., & Weber, T. (2019). The antiSMASH database version 2: A comprehensive resource on secondary metabolite biosynthetic gene clusters. *Nucleic Acids Research*, *47*(D1), D625–D630. https://doi.org/10.1093/nar/gky1060

Cacciamani, G. E., Chu, T. N., Sanford, D. I., Abreu, A., Duddalwar, V., Oberai, A., Kuo, C. C. J., Liu, X., Denniston, A. K., Vasey, B., McCulloch, P., Wolff, R. F., Mallett, S., Mongan, J., Kahn, C. E., Sounderajah, V., Darzi, A., Dahm, P., Moons, K. G. M., Hung, A. J. (2023). PRISMA AI reporting guidelines for systematic reviews and meta-analyses on AI in healthcare. *Nature Medicine*, *29*(1), 14–15.https://doi.org/10.1038/s41591-022-02139-w

Cheng, L., Qi, C., Yang, H., Lu, M., Cai, Y., Fu, T., Ren, J., Jin, Q., & Zhang, X. (2022). GutMGene: A comprehensive database for target genes of gut microbes and microbial metabolites. *Nucleic Acids Research*, *50*(D1), D795–D800. https://doi.org/10.1093/nar/gkab786

Deborde, C., & Jacob, D. (2014). MeRy-B, a Metabolomic Database and Knowledge Base for Exploring Plant Primary Metabolism. In Ganesh Sriram (Ed.), *Plant Metabolism: Methods and Protocols, Methods in Molecular Biology* (Vol. 1083, pp. 3–16). Humana Press Inc. https://doi.org/10.1007/978-1-62703-661-0_1

FAO. (2022). *The State of World Fisheries and Aquaculture 20 22*.https://doi.org/10.5860/choice.50-5350

Fukushima, A., Kusano, M., Mejia, R. F., Iwasa, M., Kobayashi, M., Hayashi, N., Watanabe-Takahashi, A., Narisawa, T., Tohge, T., Hur, M., Wurtele, E. S., Nikolau, B. J., & Saito, K. (2014). Metabolomic characterization of knockout mutants in arabidopsis: Development of a metabolite profiling database for knockout mutants in arabidopsis. *Plant Physiology*, *165*(3), 948–961. https://doi.org/10.1104/pp.114.240986

Gil-De-la-fuente, A., Mamani-Huanca, M., Stroe, M. C., Saugar, S., Garcia-Alvarez, A., Brakhage, A. A., Barbas, C., & Otero, A. (2021). Aspergillus metabolome database for mass spectrometry metabolomics. *Journal of Fungi*, *7*(5) https://doi.org/10.3390/jof7050387

Guo, A. C., Jewison, T., Wilson, M., Liu, Y., Knox, C., Djoumbou, Y., Lo, P., Mandal, R., Krishnamurthy, R., & Wishart, D. S. (2013). ECMDB: The E. coli Metabolome Database. *Nucleic Acids Research*, *41*(D1), D625–D630. https://doi.org/10.1093/nar/gks992

Huang, W., Brewer, L. K., Jones, J. W., Nguyen, A. T., Marcu, A., Wishart, D. S., Oglesby-Sherrouse, A. G., Kane, M. A., & Wilks, A. (2018). PAMDB: A comprehensive Pseudomonas aeruginosa metabolome database. *Nucleic Acids Research*, *46*(D1), D575–D580. https://doi.org/10.1093/nar/gkx1061

Jones, M. R., Pinto, E., Torres, M. A., Dörr, F., Mazur-Marzec, H., Szubert, K., Tartaglione, L., Dell'Aversano, C., Miles, C. O., Beach, D. G., McCarron, P., Sivonen, K., Fewer, D. P., Jokela, J., & Janssen, E. M.-L. (2021). CyanoMetDB, a comprehensive public database of secondary metabolites from cyanobacteria. *Water Research*, *196*. https://doi.org/10.1016/j.watres.2021.117017

Karu, N., Deng, L., Slae, M., Guo, A. C., Sajed, T., Huynh, H., Wine, E., & Wishart, D. S. (2018). A review on human fecal metabolomics: Methods, applications and the human fecal metabolome database. *Analytica Chimica Acta*, *1030*, 1–24. https://doi.org/10.1016/j.aca.2018.05.031

Kumar, B. A., & Goundar, M. S. (2023). Developing mobile language learning applications: a systematic literature review. *Education and Information Technologies*, *28*(5), 5651–5671. https://doi.org/10.1007/s10639-022-11377-x

Li, D., Ma, B., Xu, X., Chen, G., Li, T., & He, N. (2020). MMHub, a database for the mulberry metabolome. *Database*, *2020*. https://doi.org/10.1093/database/baaa011

Lu, L., Wang, J., Xu, Y., Wang, K., Hu, Y., Tian, R., Yang, B., Lai, Q., Li, Y., Zhang, W., Shao, Z., Lam, H., & Qian, P.-Y. (2014). A high-resolution LC-MS-based secondary metabolite fingerprint database of marine bacteria. *Scientific Reports*, *4*(6537). https://doi.org/10.1038/srep06537

Ma, S., Lv, Y., Hou, L., Jia, Z., Lin, S., Wang, S., He, X., & Hou, J. (2023). Effect of acute temperature stress on energy metabolism, immune performance and gut microbiome of largemouth bass (Micropterus salmoides). *Aquaculture and Fisheries*, *October*. https://doi.org/10.1016/j.aaf.2023.10.001

Ministry of Marine Affairs and Fisheries Republic of Indonesia. (2023). *Capture Fisheries Production Data*. https://statistik.kkp.go.id/home.php?m=prod_ikan_prov&i=2#panel-footer-kpda

Nakamura, K., Shimura, N., Otabe, Y., Hirai-Morita, A., Nakamura, Y., Ono, N., Ul-Amin, M. A., & Kanaya, S. (2013). KNApSAcK-3D: A three-dimensional structure database of plant metabolites. *Plant and Cell Physiology*, *54*(2). https://doi.org/10.1093/pcp/pcs186

Neveu, V., Nicolas, G., Amara, A., Salek, R. M., & Scalbert, A. (2023). The human microbial exposome: expanding the Exposome-Explorer database with gut microbial metabolites. *Scientific Reports*, *13*(1). https://doi.org/10.1038/s41598-022-26366-w

Power, D. M., Taoukis, P., Houhoula, D., Tsironi, T., & Flemetakis, E. (2023). Integrating omics technologies for improved quality and safety of seafood products. *Aquaculture and Fisheries*, *8*(4), 457–462. https://doi.org/10.1016/j.aaf.2022.11.005

PRISMA. (2020). *PRISMA Transparent Reporting of Systematic Reviews and Meta-Analyses*. Prisma. http://www.prisma-statement.org/

Ramirez-Gaona, M., Marcu, A., Pon, A., Guo, A. C., Sajed, T., Wishart, N. A., Karu, N., Feunang, Y. D., Arndt, D., & Wishart, D. S. (2017). YMDB 2.0: A significantly expanded version of the yeast metabolome database. *Nucleic Acids Research*, *45*(D1), D440–D445. https://doi.org/10.1093/nar/gkw1058

Sakurai, N., Yamazaki, S., Suda, K., Hosoki, A., Akimoto, N., Takahashi, H., Shibata, D., & Aoki, Y. (2023). The Thing Metabolome Repository family (XMRs): comparable untargeted metabolome databases for analyzing sample-specific unknown metabolites. *Nucleic Acids Research*, *51*(D1), D660–D677. https://doi.org/10.1093/nar/gkac1058

Sauro, J., & Lewis, J. R. (2012). *Quantifying the User Experience*. Elsevier Inc.; Scopus. https://doi.org/10.1016/C2010-0-65192-3

Sharfina, Z., & Santoso, H. B. (2016). An Indonesian Adaptation of the System Usability Scale (SUS). *2016 International Conference on Advanced Computer Science and*

Information Systems (ICACSIS), 145–148.

Shima, H., Sakata, K., & Kikuchi, J. (2023). Prediction of Influence Transmission by Water Temperature of Fish Intramuscular Metabolites and Intestinal Microbiota Factor Cascade Using Bayesian Networks. *Applied Sciences*, *13*(5), 3198. https://doi.org/10.3390/app13053198

Shumilina, E., Ciampa, A., Capozzi, F., Rustad, T., & Dikiy, A. (2015). NMR approach for monitoring postmortem changes in Atlantic salmon fillets stored at 0 and 4 °C. *Food Chem.*, *184*, 12–22. https://doi.org/doi:10.1016/j.foodchem.2015.03.037

Sud, M., Fahy, E., Cotter, D., Azam, K., Vadivelu, I., Burant, C., Edison, A., Fiehn, O., Higashi, R., Nair, K. S., Sumner, S., & Subramaniam, S. (2016). Metabolomics Workbench: An international repository for metabolomics data and metadata, metabolite standards, protocols, tutorials and training, and analysis tools. *Nucleic Acids Research*, *44*(D1), D463–D470. https://doi.org/10.1093/nar/gkv1042

Surendra, A., & Cuperlovic-Culf, M. (2017). Database of resistance related metabolites in Wheat Fusarium head blight Disease (MWFD). *Database: The Journal of Biological atabases and Curation*, *2017*. https://doi.org/10.1093/database/bax076

Vadlapudi, V., Borah, N., Yellusani, K. R., Gade, S., Reddy, P., Rajamanikyam, M., Vempati, L. N. S., Gubbala, S. P., Chopra, P., Upadhyayula, S. M., & Amanchy, R. (2017). Aspergillus Secondary Metabolite Database, a resource to understand the Secondary metabolome of Aspergillus genus. *Scientific Reports*, *7*(1). https://doi.org/10.1038/s41598-017-07436-w

Vivek-Ananth, R. P., Sahoo, A. K., Kumaravel, K., Mohanraj, K., & Samal, A. (2021). MeFSAT: a curated natural product database specific to secondary metabolites of medicinal fungi. *RSC Advances*, *11*(5), 2596–2607. https://doi.org/10.1039/d0ra10322e

Wishart, D. S., Feunang, Y. D., Marcu, A., Guo, A. C., Liang, K., Vázquez-Fresno, R., Sajed, T., Johnson, D., Li, C., Karu, N., Sayeeda, Z., Lo, E., Assempour, N., Berjanskii, M., Singhal, S., Arndt, D., Liang, Y., Badran, H., Grant, J., Scalbert, A. (2018). HMDB 4.0: The human metabolome database for 2018. *Nucleic Acids Research*, *46*(D1), D608–D617. https://doi.org/10.1093/nar/gkx1089

Wishart, D. S., Guo, A., Oler, E., Wang, F., Anjum, A., Peters, H., Dizon, R., Sayeeda, Z., Tian, S., Lee, B. L., Berjanskii, M., Mah, R., Yamamoto, M., Jovel, J., Torres-Calzada, C., Hiebert-Giesbrecht, M., Lui, V. W., Varshavi, D., Varshavi, D., Gautam, V. (2022). HMDB 5.0: The Human Metabolome Database for 2022. *Nucleic Acids Research*, *50*(D1), D622–D631. https://doi.org/10.1093/nar/gkab1062

Wishart, D. S., Jewison, T., Guo, A. C., Wilson, M., Knox, C., Liu, Y., Djoumbou, Y., Mandal, R., Aziat, F., Dong, E., Bouatra, S., Sinelnikov, I., Arndt, D., Xia, J., Liu, P., Yallou, F., Bjorndahl, T., Perez-Pineiro, R., Eisner, R., Scalbert, A. (2013). HMDB 3.0-The Human Metabolome Database in 2013. *Nucleic Acids Research*, *41*(D1), D801–D807. https://doi.org/10.1093/nar/gks1065

Wishart, D. S., Oler, E., Peters, H., Guo, A., Girod, S., Han, S., Saha, S., Lui, V. W., LeVatte, M., Gautam, V., Kaddurah-Daouk, R., & Karu, N. (2023). MiMeDB: the Human Microbial Metabolome Database. *Nucleic Acids Research*, *51*(D1), D611–D620. https://doi.org/10.1093/nar/gkac868

Wishart, D. S., Tzur, D., Knox, C., Eisner, R., Guo, A. C., Young, N., Cheng, D., Jewell, K., Arndt, D., Sawhney, S., Fung, C., Nikolai, L., Lewis, M., Coutouly, M.-A., Forsythe, I., Tang, P., Shrivastava, S., Jeroncic, K., Stothard, P., Querengesser, L. (2007). HMDB: The human metabolome database. *Nucleic Acids Research*, *35* (SUPPL.1), D521–D526. https://doi.org/10.1093/nar/gkl923

Wu, W.-S., Wu, H.-Y., Wang, P.-H., Chen, T.-Y., Chen, K.-R., Chang, C.-W., Lee, D.-E., Lin, B.-H., Chang, W. C.-W., & Liao, P.-C. (2022). LCMD: Lung Cancer Metabolome Database. *Computational and Structural Biotechnology Journal*, *20*, 65–78. https://doi.org/10.1016/j.csbj.2021.12.002

Xu, Y., Yang, H., Wu, T., Dong, Q., Sun, Z., Shang, D., Li, F., Xu, Y., Su, F., Liu, S., Zhang, Y., & Li, X. (2017). BioM2MetDisease: A manually curated database for associations between microRNAs, metabolites, small molecules and metabolic diseases. *Database*, *2017*. https://doi.org/10.1093/database/bax037

Yan, Z., Zhou, Y., Zhang, Y., & Zhang, X. (2023). Distribution, Bioaccumulation, and Risks of Pharmaceutical Metabolites and Their Parents: A Case Study in an Yunliang River, Nanjing City. *International Journal of Environmental Research and Public Health*, *20*(4). https://doi.org/10.3390/ijerph20042967

Yanshole, V. V, Melnikov, A. D., Yanshole, L. V, Zelentsova, E. A., Snytnikova, O. A., Osik, N. A., Fomenko, M. V, Savina, E. D., Kalinina, A. V, Sharshov, K. A., Dubovitskiy, N. A., Zaikovskii, A. A., Kobtsev, M. S., Mariasina, S. S., & Tsentalovich, Y. P. (2023). Animal Metabolite Database: Metabolite Concentrations in Animal Tissues and Convenient Comparison of Quantitative Metabolomic Data. *Metabolites*, *13*.