



www.bioinformation.net  
Volume 22(4)



Research Article

Received April 1, 2026; Revised April 30, 2026; Accepted April 30, 2026, Published April 30, 2026

DOI: 10.6026/973206300222316

SJIF 2026 (Scientific Journal Impact Factor for 2026) = 8.478  
2022 Impact Factor (2023 Clarivate Inc. release) is 1.9

**Declaration on Publication Ethics:**

The author's state that they adhere with COPE guidelines on publishing ethics as described elsewhere at <https://publicationethics.org/>. The authors also undertake that they are not associated with any other third party (governmental or non-governmental agencies) linking with any form of unethical issues connecting to this publication. The authors also declare that they are not withholding any information that is misleading to the publisher in regard to this article.

**Declaration on official E-mail:**

The corresponding author declares that lifetime official e-mail from their institution is not available for all authors

**License statement:**

This is an Open Access article which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly credited. This is distributed under the terms of the Creative Commons Attribution License

**Comments from readers:**

Articles published in BIOINFORMATION are open for relevant post publication comments and criticisms, which will be published immediately linking to the original article without open access charges. Comments should be concise, coherent and critical in less than 1000 words.

**Disclaimer:**

Bioinformation provides a platform for scholarly communication of data and information to create knowledge in the Biological/Biomedical domain after adequate peer/editorial reviews and editing entertaining revisions where required. The views and opinions expressed are those of the author(s) and do not reflect the views or opinions of Bioinformation and (or) its publisher Biomedical Informatics. Biomedical Informatics remains neutral and allows authors to specify their address and affiliation details including territory where required.

Edited by P Kanguane

Citation: Wang *et al.* Bioinformation 22(4): 2316-2319 (2026)

# LIDB: A multi-omics database of *Lagerstroemia indica*

Lei Wang, Bai Li & Gaojun Chong\*

Department of Biotechnology, Jiaying Academy of Agricultural Sciences, Jiaying, Zhejiang, China, 314016; \*Corresponding author

**Affiliation URL:**

<https://www.jxnky.com/>

**Authors contacts:**

Lei Wang - E-mail: [leiw1215@163.com](mailto:leiw1215@163.com)

Bai Li - E-mail: [libaia@yeah.net](mailto:libaia@yeah.net)

Gaojun Chong - E-mail: [chonggaojun@126.com](mailto:chonggaojun@126.com)

**Abstract:**

The *Lagerstroemia indica* (Crepe Myrtle) is an economically important ornamental and medicinal plant, with significant potential for functional genomics and breeding. However, there is no comprehensive, centralized database integrating multi-omics data. Therefore, it is of interest to the development of the *Lagerstroemia indica* MultiOmics Database (LIDB), a web-based platform designed to integrate and analyze genomic, transcriptomic and proteomic data. The LIDB is a valuable resource for researchers, breeders and biotechnologists, facilitating data-driven discoveries in *Lagerstroemia indica* genomics.

**Keyword:** *Lagerstroemia indica*, multi-omics, database

**Background:**

*Lagerstroemia indica* is a widely cultivated ornamental plant, native to Southeast Asia and has been extensively used in landscaping, as well as in traditional medicine [1]. Various promising omics technologies have emerged in the past few decades, these omics-based approaches have proven to be valuable for exploring the genetic and molecular basis of crop development by modifying DNA, transcript levels, proteins, metabolites and mineral nutrition in the context of environmental and physiological stress responses [2]. Several omics approaches, such as genomics, transcriptomics, proteomics, metabolomics and phenomics, have revealed each corresponding molecular biological aspect integrated with plant systems [3, 4]. The advent of next-generation sequencing (NGS) technology has brought about high-throughput and rapid data generation of genomes, transcriptomes, proteomes, metabolomes and phenomes [5]. The integration of multiple omics approaches can elucidate gene functions and networks under physiological and environmental stress conditions. Comprehensive multi-omics approaches with powerful technologies have been used to identify and decipher the essential components of stress response, senescence and yield in various important economic crops [6]. These omics approaches have been implemented in several important crops, including wheat, soybean, tomato, corn, millet, cotton and other economic crops [7]. However, there are relatively few reports on omics studies in *Lagerstroemia indica*. Given the growing need for centralized resources that enable easy access to genomic and multi-omics data. Therefore, it is of interest to describe the development of the *Lagerstroemia indica* Multi-Omics Database (LIDB).

**Materials and Methods:**

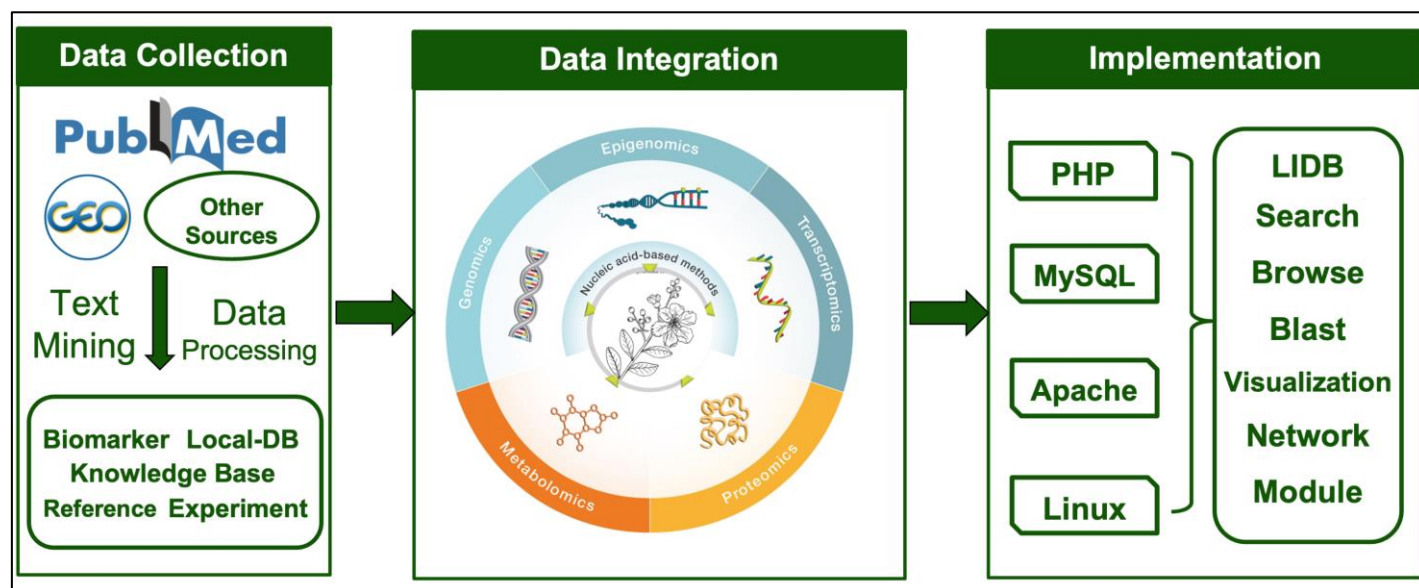
The main procedures of *Lagerstroemia indica* Multi-Omics Database (LIDB) construction are illustrated in **Figure 1**, including data collection, data integration. Regarding the database, we have extracted most of the related genome and transcriptome data from NCBI Gene Expression Omnibus. Data manually retrieved from the literature account for a minor portion of the total multi-omics data, which was obtained by querying PubMed using keywords and handpicking several essential findings from recent years. We curated and annotated the collected data with the help of publicly available knowledge. For the annotation of genes, we considered such databases as Gene [8] as well as R packages include biomaRt [9] and clusterProfiler [10]. Multi-omics integration analysis involves systematic data processing and modeling across genomic, transcriptomic, proteomic and metabolomic datasets. The

workflow begins with quality control using platform-specific tools (FastQC for sequencing data, MaxQuant for proteomics and XCMS for metabolomics) followed by batch effect correction via ComBat [11]. Data normalization employs Z-score transformation to enable cross-omics comparison. For pairwise integration (e.g., transcriptome-proteome), Spearman correlation identifies concordant molecules, while sparse canonical correlation analysis (sCCA) reveals co-varying modules between metabolomic and microbiome data [12]. Higher-order integration utilizes MOFA+ for dimensionality reduction to extract latent factors shared across omics layers [13]. Machine learning approaches like MOVICS [14] integrate NMF and consensus clustering for patient stratification [15]. Pathway mapping is performed using MetaboAnalyst with KEGG enrichment (FDR <0.05) [16].

**Database and data statistics:**

In the current version, LIDB documents 97 manual processed datasets of *Lagerstroemia indica*. Each entry is associated with the PMID and gene symbol. Meanwhile, we also collected 6 other economic crops for users to use as the reference. LIDB provides a user-friendly interface with interactive plots and searching bars in each module. In the 'Search' page. Users can retrieve data by location, gene symbol or keywords to retrieve related data. LIDB provides brief description of search results in the 'Search Result' page. To gain more detail information of a specific entry, users can click the 'Details' button. Additional information such as PubMed ID, GO annotations are displayed in the single-record page. It also integrates multi-omics data into the genome browser, JBrowse, that provides reference sequence, annotation, population statistics and genome-wide-associated analysis results for agronomic traits.

The tool page offers a plethora of bioinformatics tools, including Local Alignment Search Tool (BLAST), GWAS Single-Trait for laboratory research. The database was constructed based on LAMP (Linux, Apache, MySQL and PHP) mode. MySQL was selected to store biomarker information. The front end of the webpage was written in HTML, CSS and JavaScript, while PHP was used to realize the interaction of the webpage. The database was running under the environment of Linux and Apache. The website is freely available to all without the need for registration or login. Designed for cross-platform compatibility, it ensures a user-friendly interface across computers and tablets for broad accessibility.



**Figure 1:** Workflow of LIDB

#### Utility:

*Lagerstroemia indica* is a representative species of the genus *Lagerstroemia*. The extensive application of multi-omics technology in *Lagerstroemia* research has enabled people to better understand many unique genetic bases and origin histories of *Lagerstroemia*, providing important information for a more in-depth, comprehensive and systematic understanding of *Lagerstroemia* and providing a broader space for the innovation and quality improvement of germplasm resources in the future, but there are still some limitations in the research process. First, although genomic studies of *Lagerstroemia* have taken a leading role in multi-omics research, the available chromosome-level genomes are still relatively limited in both size and diversity of studied accessions, especially when compared with well-established crop systems such as rice or maize. Moreover, several advanced high-throughput sequencing technologies—such as spatial transcriptomics and single-cell omics—have not yet been applied to *Lagerstroemia* and the use of transcriptomic, metabolomic, proteomic and integrated multi-omics analyses remains relatively scarce. Nevertheless, with the continuous accumulation of data and the emergence of more powerful integration methodologies, we anticipate that resources such as the *Lagerstroemia indica* Database (LIDB) will become increasingly valuable tools for researchers in this field.

#### Future work:

As the database continues to evolve, further integration with additional omics layers, such as epigenomics and sc-RNA-seq, will expand its capability to provide a more holistic view of plant biology. The expansion of data through collaborative efforts and the incorporation of community-driven contributions will ensure that LIDB remains at the forefront of *Lagerstroemia indica* research, helping to foster innovation in both fundamental biology and practical applications.

#### Funding:

This research was funded by the Research on Radiation Breeding and Tissue Culture Technology of *Lagerstroemia indica*, grant number 2021C02071-4-5.

**Availability:** LIDB can be visited freely via <http://lidb.org.cn>.

#### References:

- [1] Yu C *et al.* *BMC Plant Biol.* 2024 **24**:171. [PMID:38443839]
- [2] Raza A *et al.* *Trends Plant Sci.* 2025 **30**:10. [PMID: 40148151]
- [3] Huang L *et al.* *BMC Genomics.* 2014 **15**:1. [PMID: 25428615]
- [4] Neveu P *et al.* *New Phytologist.* 2019 **221**:588. [PMID: 30152011]
- [5] Großkinsky D.K *et al.* *J Exp Bot.* 2018 **4**:825. [PMID: 29444308]
- [6] Varadharajan V *et al.* *Plants (Basel).* 2025 **14**:865. [PMID: 40265800]
- [7] Yang F *et al.* *Nat Genet.* 2025 **57**:2106. [PMID:40670880]
- [8] Maglott D *et al.* *Nucleic Acids Res.* 2011 **39**:52. [PMID: 21115458]
- [9] Durinck S *et al.* *Nat Protoc.* 2009 **4**:1184. [PMID: 19617889]
- [10] Wu T *et al.* *Innovation (Camb).* 2021 **2**:100141. [PMID: 34557778]
- [11] <https://pubmed.ncbi.nlm.nih.gov/17572666/>
- [12] Argelaguet R *et al.* *Genome Biol.* 2022 **21**:111. [PMID: 32393329]
- [13] Wu G *et al.* *Genome Med.* 2020 **12**:73. [PMID: 32825850]
- [14] Lu X *et al.* *Bioinformatics.* 2021 **36**:5539. [PMID: 33315104]
- [15] Pang Z *et al.* *Nucleic Acids Res.* 2021 **49**:388. [PMID: 34019663]
- [16] Wu Q *et al.* *Signal Transduct Target Ther.* 2022 **7**:270. [PMID: 35989380]

*Caveat Emptor is applicable among the literate community where required and possible. The publisher, its journal, editors and the internal/external reviewers take adequate steps to check, evaluate, correct, edit, revise and improve content where possible and required.*