
IGD: a multi-omics database for *Ipomoea pes-caprae* genomic and biological research

Received: 5 June 2025

Accepted: 9 January 2026

Published online: 26 January 2026

Cite this article as: Cai J., Chen X., Li X. *et al.* IGD: a multi-omics database for *Ipomoea pes-caprae* genomic and biological research. *BMC Plant Biol* (2026). <https://doi.org/10.1186/s12870-026-08161-1>

Jiahao Cai, Xingguang Chen, Xueli Li, Mingyue Zhang, Xinqiang Lin, Yile Hu, Haoran Feng, Xinyu Li, Jinbin Hu, Shuqi Yang, Lulu Wang, Xiaoping Niu, Gang Wang, Boping Tang, Sheng Wang, Yuan Qin & Yan Cheng

We are providing an unedited version of this manuscript to give early access to its findings. Before final publication, the manuscript will undergo further editing. Please note there may be errors present which affect the content, and all legal disclaimers apply.

If this paper is publishing under a Transparent Peer Review model then Peer Review reports will publish with the final article.

ARTICLE IN PRESS

IGD: a multi-omics database for *Ipomoea pes-caprae* genomic and biological research

Jiahao Cai^{1, #}, Xingguang Chen^{1, #}, Xueli Li¹, Mingyue Zhang¹, Xinqiang Lin¹, Yile Hu¹, Haoran Feng¹, Xinyu Li¹, Jinbin Hu¹, Shuqi Yang¹, Lulu Wang¹, Xiaoping Niu¹, Gang Wang², Boping Tang², Sheng Wang³, Yuan Qin^{1, *} and Yan Cheng^{1, *}

¹State Key Laboratory of Ecological Pest Control for Fujian and Taiwan Crops, Fujian Provincial Key Laboratory of Haixia Applied Plant Systems Biology, College of Plant Protection, College of Life Sciences, School of Future Technology, Haixia Institute of Science and Technology, Fujian Agriculture and Forestry University, Fuzhou 350002, China

²Jiangsu Provincial Key Laboratory of Coastal Wetland Bioresources and Environmental Protection, Jiangsu Collaborative Innovation Center for Coastal Biology and Agriculture, School of Wetlands, Yancheng Teachers University, Yancheng City, 224002, China

³Department of Biochemistry, Microbiology and Immunology, University of Saskatchewan, Saskatoon, SK S7N 5E5, Canada

[#]These authors contributed equally to this work.

*Correspondence: Yuan Qin, yuanqin@fafu.edu.cn or Yan Cheng, chengyan1220@hotmail.com

Abstract

Ipomoea pes-caprae (IPC) is a perennial halophytic vine with remarkable salt and drought

tolerance, playing a critical ecological and medicinal role in tropical and subtropical coastal ecosystems. Despite the availability of a high-quality chromosome-level reference genome and abundant transcriptome data, the absence of an integrated data platform has hindered in-depth functional gene discovery and genomic research in IPC. To address this gap, we developed the IPC Genome Database (IGD), the first comprehensive multi-omics database dedicated to IPC. IGD provides high-quality genomic sequences, gene structure annotations, and functional annotations, along with transcriptomic expression profiles under salt stress across different tissues. It also includes time-course expression data of roots and leaves under salt stress treatment, as well as leaf expression profiles under cold and heat stress conditions. The platform also integrates a suite of user-friendly, interactive bioinformatics analysis and visualization tools to facilitate efficient data exploration and functional mining. In particular, we established a specialized subset of stress-responsive genes to support rapid candidate gene screening and evolutionary analysis related to IPC's salt and stress tolerance traits. By implementing automated data integration and quality control pipelines tailored for non-model species, IGD overcomes technical challenges such as fragmented data and delayed updates. As a centralized and scalable resource, IGD significantly lowers the barrier to functional genomics research in IPC and other halophytes, promoting their application in genetic improvement and ecological sustainability. In the future, IGD will continue to incorporate more omics data from halophytic species, further enhancing its value in salt tolerance research and molecular breeding efforts.

Key Word: *Ipomoea pes-caprae*, comparative genomics, database, halophytes

1. Introduction

Halophytes, due to their unique physiological mechanisms and molecular regulatory networks for salt and alkaline tolerance, have become important research subjects for elucidating plant stress adaptation mechanisms and evolutionary strategies [1]. These plants not only provide key gene resources for molecular breeding of stress-resistant crops but also offer theoretical foundations for the ecological management and sustainable utilization of saline-alkali lands. *Ipomoea pes-caprae* (IPC), a typical representative, is a perennial vine species of the genus *Ipomoea* in the Convolvulaceae family, widely distributed in tropical and subtropical coastal ecosystems [2]. This species possesses multiple stress tolerance traits, including salt and drought resistance, and plays an irreplaceable ecological role in coastal sand fixation, slope protection, and the restoration of degraded ecosystems. Additionally, its bioactive compounds show significant potential in medicinal applications such as anti-inflammatory and analgesic treatments.

Given the significant scientific and industrial value of IPC, our research team conducted de novo transcriptome sequencing and assembly of IPC leaves under heat and cold stress conditions in 2021 [3]. Subsequently, in 2023, we completed and published a chromosome-level high-quality reference genome of IPC, along with transcriptomic data from different tissues under salt stress treatment [4]. Based on this reference genome, we re-mapped the time-course RNA-seq data under salt stress previously published by Liu et al. to the new genome and integrated the datasets to construct an updated expression matrix [5]. This integrated analysis provided deeper insights into the molecular responses of IPC to salt stress. Currently, most model plants, such as *Arabidopsis thaliana*, *Oryza sativa*, and *Zea mays* [6], have established functional genomics databases that provide standardized data platforms for related research. To address current research limitations in IPC, we developed the first integrated functional genomics database for

IPC — IGD (<https://ipcgenome.org.cn>). Through a web browser, users can conveniently access and query a variety of data types provided by the platform, including genome information, transcriptome data, and functional annotation. This database not only enables efficient management and easy access to multi-omics data, but also integrates a variety of online analysis tools, providing a one-stop data platform for functional genomics research and evolutionary analysis of IPC. During the construction of the database, we implemented automated pipelines to enable regular data updates and quality control, ensuring the accuracy and timeliness of the data. For IPC's unique traits such as salt tolerance and stress resistance, a dedicated functional gene subset database was established to support rapid candidate gene screening and evolutionary analysis. The creation of this platform not only provides a data infrastructure for basic research on IPC but also lowers the threshold for bioinformatics analysis, facilitating the transition of non-model species research from “data accumulation” to “functional interpretation”. In the future, the database will continuously incorporate more omics data from halophytes, providing core data support for elucidating adaptive evolution mechanisms in salt-tolerant plants and for breeding stress-resistant crops.

2. Methods

2.1. Data sources

The genomic data used for building the IPC genome database were obtained from our previously published genome sequencing and assembly project [4]. The raw sequencing reads (CRA013080) and the final assembled genome (GWHEQBK00000000) have been deposited in the China National Center for Bioinformation (<https://www.cncb.ac.cn>), and serve as the foundational

dataset for the current database construction. In addition, part of the RNA-seq data was derived from our previously published transcriptomic study of IPC leaves under heat and cold stress (Project ID: PRJNA656146) [3], while the other part was obtained from the publicly available time-course salt stress dataset published by Liu et al. (Project ID: PRJNA656933) [5]. We integrated these datasets and analyzed gene expression patterns across different tissues under salt stress [4]. These publicly available datasets serve as fundamental resources for the construction of this database.

2.2 Construction of GO and KEGG Annotation Database

To achieve functional annotation of transcriptome genes, this study employed the eggNOG-mapper tool [7] to construct an annotation database. Protein sequences of the IPC genome were uploaded to the eggNOG-mapper online platform (<http://eggnog-mapper.embl.de/>) for automated annotation. The resulting annotation data were then used for GO and KEGG enrichment analyses with the R package clusterProfiler [8–11].

2.3 Database System Architecture

The IGD was developed based on a decoupled front-end and back-end architecture to ensure scalability, modularity, and responsive user experience. The back-end system integrates Spring Boot for handling core biological data operations and Flask [12] for lightweight tool interfacing and data routing. The front-end framework was constructed using Vue.js and Bootstrap [13], with Element UI enhancing dynamic interface components and overall interactivity. The data layer is built upon a MySQL [14] relational database, used to store and manage genomic sequences, gene models, functional annotations, and expression profiles. To further optimize performance, Redis [15] was deployed as a caching layer, significantly accelerating data retrieval and minimizing

redundant queries. The entire platform is hosted on an Ubuntu-based Linux server environment, ensuring high stability and extensibility for concurrent user access. The architecture supports RESTful APIs for cross-module communication and enables integration with external bioinformatics tools and resources.

2.4 Visualization and Interactive Interface

To enhance data interpretation and user interaction, the IGD platform employs diverse visualization techniques and intuitive graphical interfaces. Several visualization tools were developed using the R Shiny framework [16], enabling real-time parameter adjustment and rendering of interactive charts. The expression heatmap module supports log-transformation, z-score normalization, and hierarchical clustering. It allows users to explore gene expression dynamics across multiple treatments and time points. The enrichment analysis module, powered by the clusterProfiler R package [8], produces bubble plots and bar graphs that illustrate the statistical significance and functional categories of enriched GO terms and KEGG pathways. The phylogenetic tree module integrates multiple sequence alignment and tree-building tools such as MAFFT [17] and FastTree [18]. It provides tree visualization with customizable branch labels and color-coded clades, enhancing evolutionary analysis. Chromosomal localization of genes is implemented using in-house scripts that generate dynamic SVG plots based on genome coordinates, while genome-wide navigation is supported through JBrowse2 [19], enabling track selection, zooming, and annotation browsing in a user-friendly environment. Each visual component is precomputed for large datasets and rendered dynamically based on user selection to ensure both efficiency and flexibility. Together, these features provide a powerful platform for the interactive exploration of IPC genomic resources.

2.5 Integration of Third-Party Tools

To enhance the functional breadth and analytical depth of the platform, the IGD system integrates a variety of mainstream open-source bioinformatics tools, enabling online access through a unified interface. The sequence alignment feature is powered by SequenceServer [20], allowing users to submit nucleotide or protein sequences online for BLAST analysis. The results are presented with both graphical visualization and annotated hyperlinks for further exploration. The genome browser module is built with JBrowse2, supporting dynamic multi-track loading and interactive browsing of gene structures, annotation information, and RNA-seq coverage. The primer design tool integrates PrimerServer2 [21, 22], enabling efficient primer screening based on gene IDs or custom sequences. It outputs detailed parameters including primer sequences, melting temperatures (T_m), and expected product lengths. In addition, the platform provides preprocessed expression matrices, annotation files, and sequence data for user download, facilitating further in-depth analysis with local visualization tools. All integrated tools are independently deployed on the platform server and accessed via a web-based interface to ensure fast response times and stable performance (Figure 1).

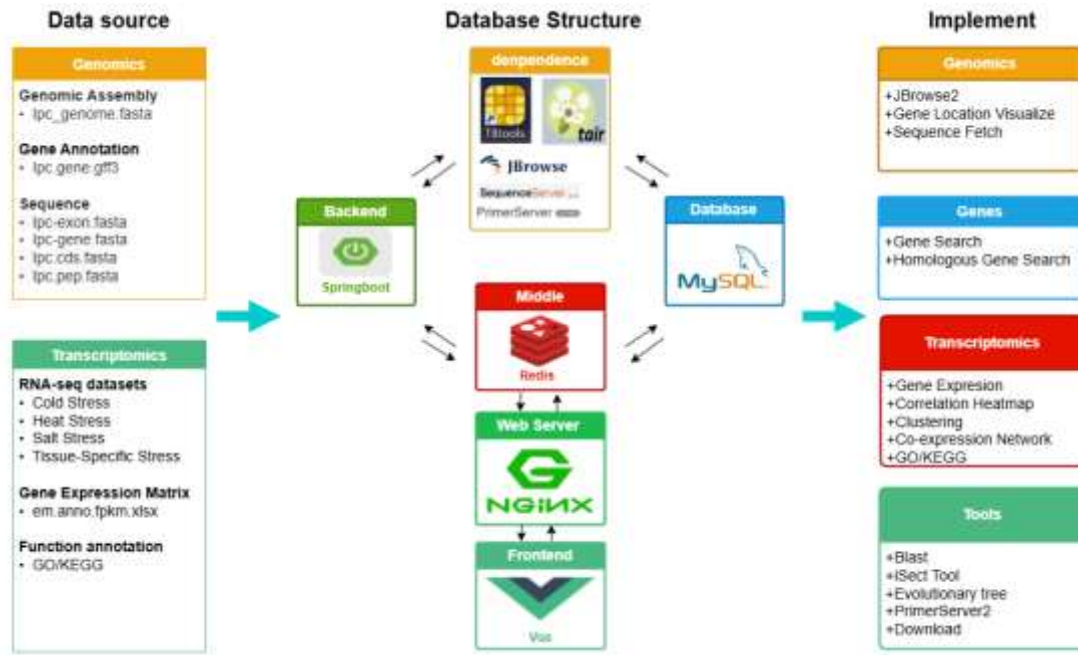


Figure 1. Database Architecture Design Diagram. Data Sources, Implementation Methods, and Results of Database Construction in IGD.

3. Results

3.1 IPC Database Content

To advance research on IPC in the fields of genomics, molecular biology, and responses to abiotic stresses, we developed the IPC Genome Database (IGD), an interactive web-based platform that integrates data retrieval, functional analysis, and visualization. This platform is designed to provide researchers with efficient and convenient access to a wide range of genomic resources and analytical tools. IGD consists of seven core modules: Home, Genomics, Genes, Transcriptomics, Tools, Downloads, and Help. Currently, the database incorporates a high-quality IPC genome assembly along with multiple transcriptome datasets derived from various tissues and treatment conditions, including cold and heat stress responses in leaves, time-series salt stress treatments in roots and stems, and tissue-specific transcriptional responses under salt stress. These resources offer valuable data for investigating the molecular mechanisms of halophytic adaptation to salt and

temperature stresses. In addition, the Help module provides comprehensive guidance on the functionalities, application scenarios, and usage instructions of each module and tool, enabling users to efficiently navigate and utilize the database platform (Figure 2).

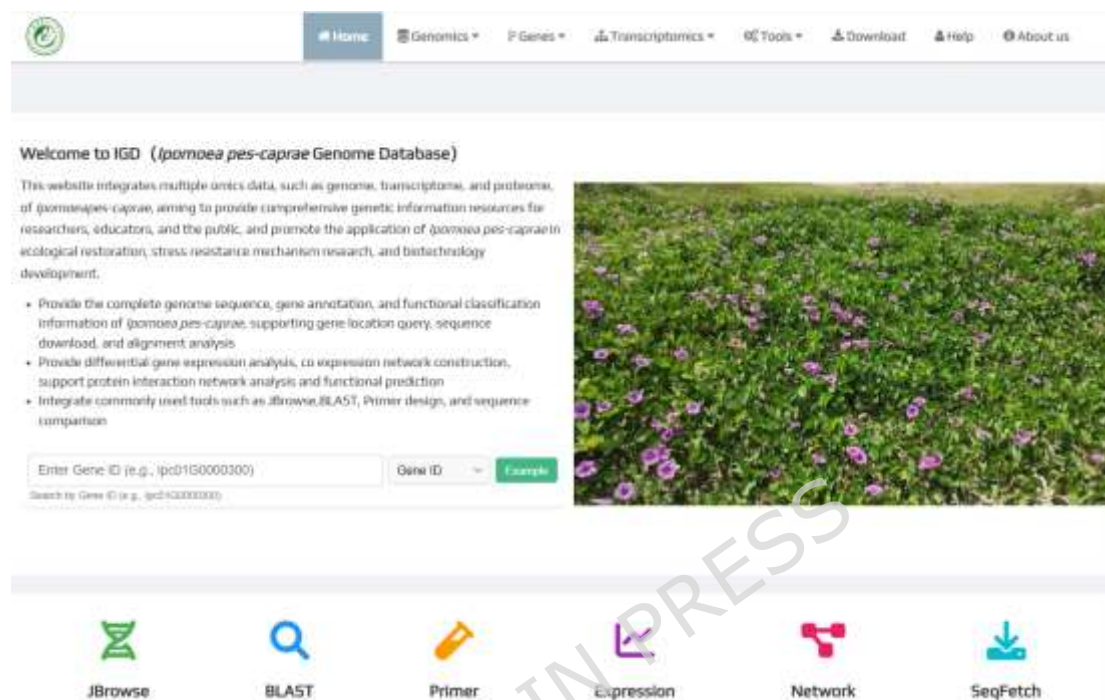


Figure 2. The homepage of the *Ipomoea pes-caprae* Genome Database.

3.2 Genomics Module

The genome module provides fundamental support for structural and functional genomic studies of IPC. This module consists of three sub-functions: JBrowse, Gene Location Visualization, and Sequence Fetch. The “JBrowse” module of the database uses JBrowse, a genome visualization tool, to display genome of IPC. Users can select a target chromosome via the search bar at the top or input specific genomic coordinates for browsing. For example, by selecting the 5046276–5214277 bp region on chromosome Chr1, users can view multiple genes within this interval, such as Ipc01G0001500. Upon clicking the gene, the system displays its detailed transcript structure, including the positions of all exons and introns, along with annotation information and expression

levels (Figure 3A). In addition, users can retrieve the DNA sequence of the selected gene. The zoom function enables multi-scale navigation, from macro-level genome structure to individual nucleotide resolution, facilitating in-depth analysis. The Gene Location Visualization function allows for precise mapping of target genes onto reference chromosomes and rapid generation of high-resolution images, which is particularly useful for gene family analysis (Figure 3B). The Sequence Fetch tool supports retrieval of user-defined genomic regions, enabling applications such as chromatin interaction analysis (e.g., Chip-seq) (Figure 3C).

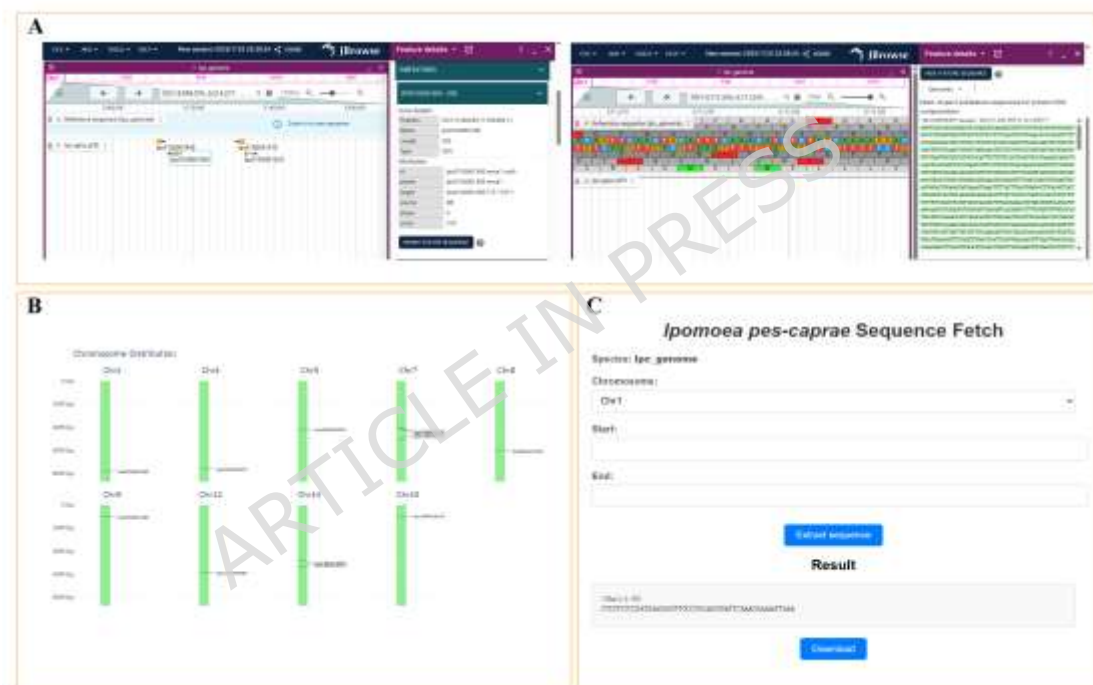


Figure 3. Genomics Module. (A) JBrowse2: A modern, interactive genome browser for visualizing and exploring genomic data. (B) Gene Chromosome Localization Visualization Tool. (C) Sequence Extraction Tool on Chromosomes.

3.3 Genes Module

The Genes module includes two main functions: Gene Search and Homolog Gene Search. In the

Gene Search function, users can search for genes by entering a gene ID or partial ID. The system displays detailed information about the gene, including its structure, chromosomal location, gene annotation, functional annotation, and the corresponding gene sequence, CDS sequence, and protein sequence. In addition, users can view the gene expression levels under various conditions, including expression in leaves under cold and heat treatments, expression in roots and leaves under time-course salt stress treatment, and expression across different organs under salt stress [3–5] (Figure 4A). The homologous gene search function enables rapid identification of homologous genes between IPC and *Arabidopsis thaliana*. Users can search using gene IDs, gene names, or functional descriptions from either IPC or *Arabidopsis*. For example, entering “SOS” will display all genes associated with the SOS gene family (Figure 4B), while inputting “kinase” will retrieve all genes related to kinase functions (Figure 4C). This feature offers a user-friendly interface with fast response times, facilitating efficient identification of target genes and enhancing the accuracy and efficiency of homologous gene analysis.

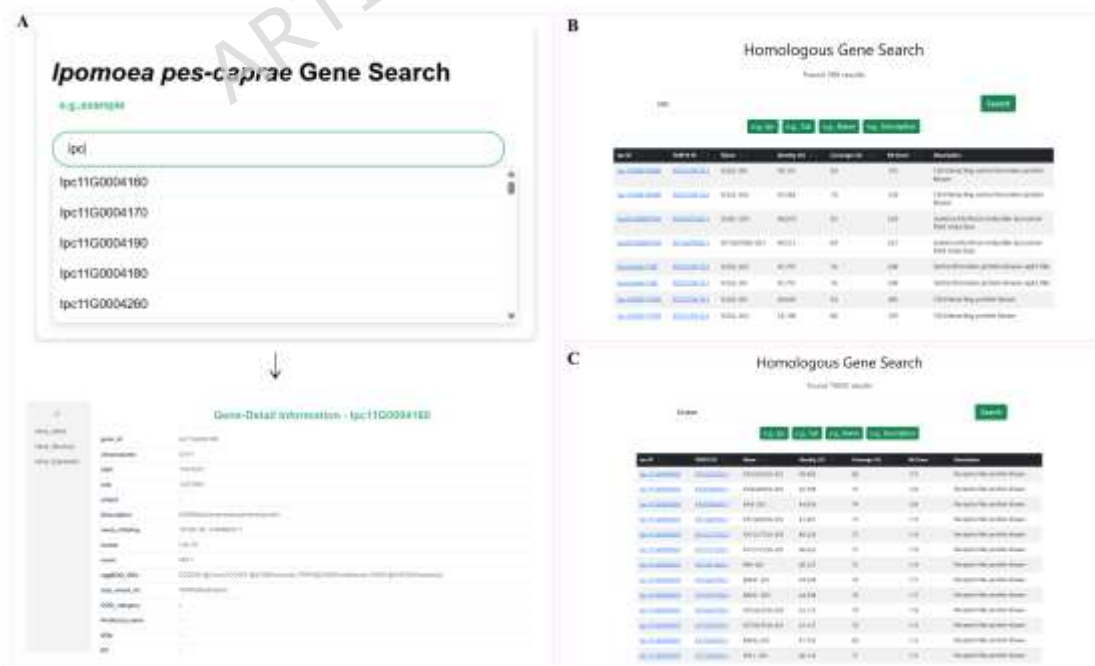


Figure 4. Genes Module. (A) Retrieve detailed gene information by querying gene IDs. (B)

Identification of homologous genes between IPC and *Arabidopsis thaliana* by gene name. (C)
Retrieval of gene sets related to specific functions, such as “kinase”, through keyword-based searches of functional descriptions.

3.4 Transcriptomics module

The transcriptomics module of the IGD database offers comprehensive and scalable solutions for transcriptome data mining, particularly in the context of stress-responsive gene regulation in IPC. It integrates both gene expression visualization and functional enrichment tools to meet diverse downstream analysis requirements. Specifically, the "Studies Detail" module provides a centralized display of transcriptomic sample information, enabling users to quickly access experimental background. The "Gene Expression Visualization" function supports the dynamic generation of heatmaps based on normalized FPKM values derived from multiple tissues and stress conditions. The system performs hierarchical clustering using Euclidean distance, allowing for bidirectional clustering of both genes and samples, thereby facilitating the identification of expression modules or co-expression clusters with potential regulatory relevance (Figure 5A). To enhance interpretability, the interactive interface allows users to customize visualization parameters such as color gradients, block shapes, and layout ratios. This functionality is particularly useful for deciphering gene expression patterns in response to abiotic stresses, such as salt and temperature. For functional annotation and enrichment analysis, the “GO and KEGG Enrichment” module (Figure 5B) integrates the latest releases of the Gene Ontology (GO) and Kyoto Encyclopedia of Genes and Genomes (KEGG) databases. Upon submission of a target gene set, the system performs statistical enrichment analysis based on the hypergeometric test and automatically generates dual-format visual outputs: bubble plots displaying the significance

($-\log_{10}$ P-value) and gene set coverage of enriched terms, and bar charts depicting the gene density within key metabolic pathways or molecular functions. All results are exportable in SVG/PDF formats and accompanied by comprehensive statistical tables, including adjusted P-values, enrichment factors, and other relevant metrics, providing multidimensional evidence for gene function and regulatory network interpretation. Through algorithmic optimization, this module enables efficient processing of large-scale transcriptomic data. Its interactive visualization design significantly enhances the readability of complex datasets and facilitates the translational application of transcriptomic research.

ARTICLE IN PRESS

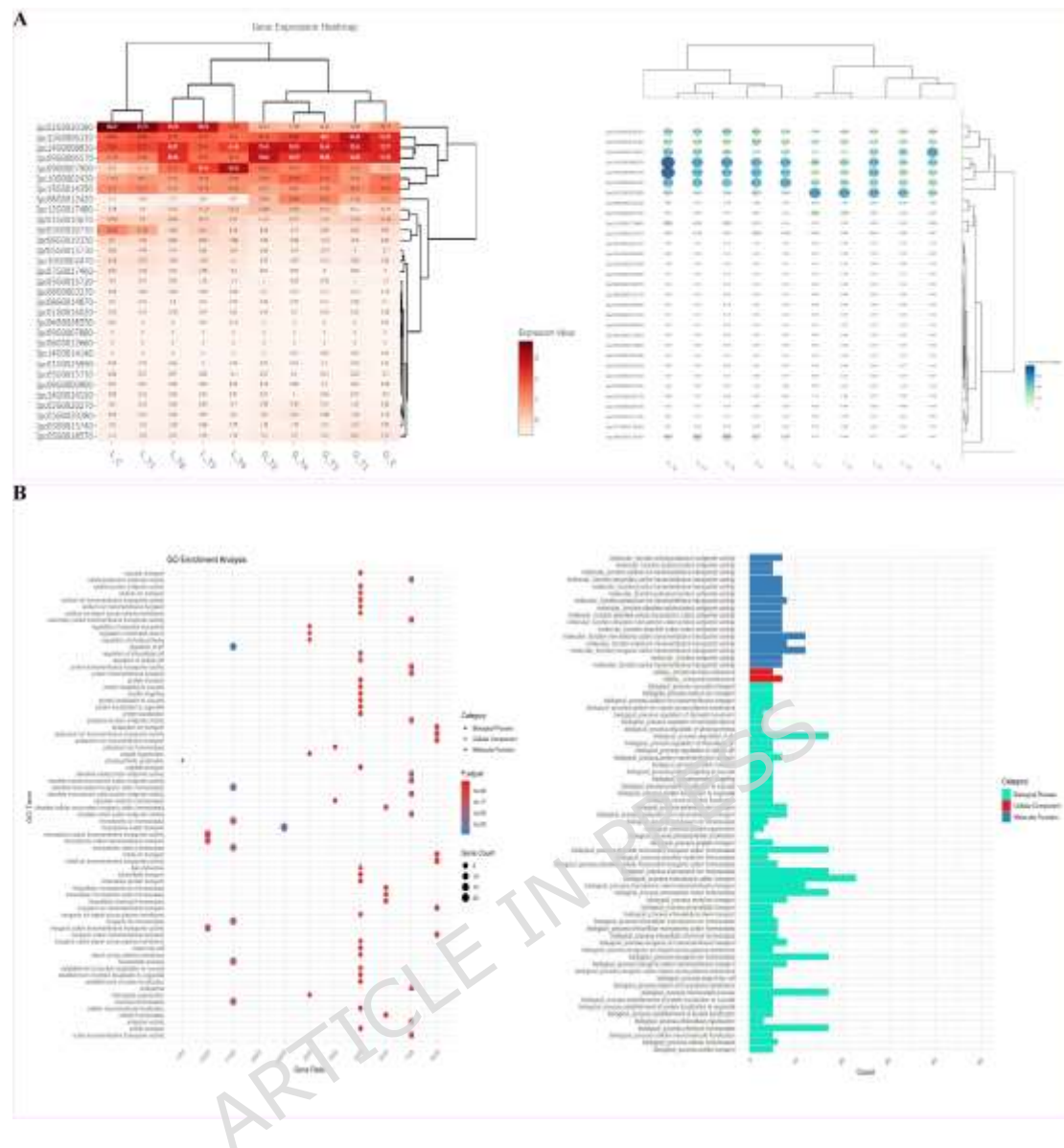


Figure 5. Transcriptomics Module. (A) Expression Heatmap Visualization. (B) GO Enrichment Analysis Bubble Plot and Bar Chart.

3.5 Tools Integration

Beyond its core analytical modules, the database integrates a comprehensive suite of interoperable tools to support diverse genomic research demands. The BLAST engine supports cross-species homolog identification, allowing users to input protein or nucleotide sequences from any species for homology searches against a dedicated, locally indexed IPC database (E-value threshold

$\leq 1e-5$), thereby rapidly identifying candidate orthologs with potential functional relevance (Figure 6A). The iSect tool supports coordinate-based queries for precise extraction of gene sequences (e.g., 2,000 bp upstream of the 5' end of a gene) (Figures 6B), outputting sequences in FASTA format for downstream applications such as promoter region analysis. The Phylogenetic Tree Builder, based on the Neighbor-Joining algorithm, allows rapid construction of phylogenetic trees for user-defined sets of IPC genes. The resulting trees are visualized in rectangular or circular layouts (Figure 6C) and exported in Newick format for subsequent editing and optimization. The Primer Design Tool, powered by a dynamic programming algorithm (Primer3 core), enables customizable primer design by allowing specification of product length and annealing temperature. It performs in situ specificity validation against the IPC reference genome, automatically excluding primers that span exon-exon junctions or bind to non-target regions, thereby maximizing experimental success rates (Figure 6D).



Figure 6. Tools Integration. (A) SequenceServer: A web-based BLAST search tool for local genomic sequence analysis. (B) Gene Sequence Extraction Tool. (C) Phylogenetic Tree Construction. (D) PrimerServer: An automated web-based tool for designing PCR primers with customizable constraints.

4. Discussion and perspectives

In this study, we constructed IGD, the first database of IPC based on genomic and transcriptomic data. IGD provides a comprehensive and high-quality multi-omics dataset for IPC, including genomic information, gene sequences with functional annotations, and spatiotemporal gene

expression dynamics. Based on these rich data resources, we have developed and integrated multiple powerful and interactive bioinformatics analysis and visualization tools. This platform will facilitate researchers in deeply elucidating the regulatory networks underlying the growth, development, and stress resistance of IPC, and in identifying key functional genes, thereby advancing systematic understanding and applied research on salt tolerance mechanisms in halophytes.

In summary, the data and tools provided by IGD fully demonstrate its considerable potential in fundamental scientific research and genetic improvement breeding. In the future, we plan to regularly update and integrate more types of omics data for IPC, continuously enriching and enhancing the resources and functionalities of IGD. We will also continuously update the data resources of IPC and other halophytes, such as *Suaeda glauca* [23]. This will further expand the scientific value and service capacity of the database, supporting saline-alkali land remediation and bioresource utilization. Moreover, we will conduct in-depth investigations into gene family characteristics and transcriptomic regulatory mechanisms of IPC, promoting the discovery and validation of related functional genes, thereby providing a solid foundation for genetic improvement and functional gene research in halophytes. We firmly believe that IGD will become a critical data platform in the field of halophyte research, driving sustained progress and innovative breakthroughs in both scientific research and industrial applications.

Availability of data and materials

The database and web interface can be accessed at <https://ipcgenome.org.cn/>.

Funding

This work was supported by Natural Science Foundation of Fujian Province (2023J01442),

National Natural Science Foundation of China (32170380).

Acknowledgments

We thank the editors and reviewers for revising this manuscript.

Author Contributions

JC and XC contributed to data curation, formal analysis, software, visualization, and wrote the original draft. XL contributed to formal analysis, methodology, resources, and original draft writing. HF, YH, and MZ were responsible for software development, data visualization, and drafting the manuscript. XL and JH performed data curation, validation, and contributed to the original draft. XL contributed to validation and writing the original draft. SY assisted with visualization and writing. LW, XN, GW, BT, and SW reviewed and edited the manuscript. YQ and YC conceived and supervised the project, acquired funding, administered the project, and contributed to writing and reviewing the manuscript. All authors reviewed and approved the final version of the manuscript.

Ethics approval and consent to participate

This paper was completed within the laws of the People's Republic of China. No specific permits were required for our field research. The study species is not included in the "List of Protected Plants in China".

Consent for publication

Not applicable.

Competing interests

The authors declare no competing interests.

References

1. J B, A R-M, S S. ROS homeostasis in halophytes in the context of salinity stress tolerance. *Journal of experimental botany*. 2014;65. <https://doi.org/10.1093/jxb/ert430>.
2. Miryeganeh M, Takayama K, Tateishi Y, Kajita T. Long-distance dispersal by sea-drifted seeds has maintained the global distribution of *Ipomoea pes-caprae* subsp. *brasiliensis* (Convolvulaceae). *PLoS One*. 2014;9:e91836. <https://doi.org/10.1371/journal.pone.0091836>.
3. Cheng Y, Zhou Q, Li W, Cheng H, Mohammadi MA, Liu Y, et al. De novo transcriptome assembly and gene expression profiling of *Ipomoea pes-caprae* L. under heat and cold stresses. *Scientia Horticulturae*. 2021;289:110379. <https://doi.org/10.1016/j.scienta.2021.110379>.
4. Cheng Y, Wang Y, Sun J, Liao Z, Ye K, Hu B, et al. Unveiling the genomic blueprint of salt stress: insights from *Ipomoea pes-caprae* L. *Seed Biology*. 2023;2. <https://doi.org/10.48130/SeedBio-2023-0021>.
5. Liu Y, Dai X, Zhao L, Huo K, Jin P, Zhao D, et al. RNA-seq reveals the salt tolerance of *Ipomoea pes-caprae*, a wild relative of sweet potato. *Journal of Plant Physiology*. 2020;255:153276. <https://doi.org/10.1016/j.jplph.2020.153276>.
6. Woodhouse MR, Cannon EK, Portwood JL, Harper LC, Gardiner JM, Schaeffer ML, et al. A pan-genomic approach to genome databases using maize as a model system. *BMC plant biology*. 2021;21:385. <https://doi.org/10.1186/s12870-021-03173-5>.
7. Huerta-Cepas J, Szklarczyk D, Heller D, Hernández-Plaza A, Forslund SK, Cook H, et al.

- eggNOG 5.0: a hierarchical, functionally and phylogenetically annotated orthology resource based on 5090 organisms and 2502 viruses. *Nucleic Acids Research*. 2019;47:D309–14. <https://doi.org/10.1093/nar/gky1085>.
8. Yu G, Wang L-G, Han Y, He Q-Y. clusterProfiler: an R package for comparing biological themes among gene clusters. *OMICS*. 2012;16:284–7. <https://doi.org/10.1089/omi.2011.0118>.
 9. Wu T, Hu E, Xu S, Chen M, Guo P, Dai Z, et al. clusterProfiler 4.0: A universal enrichment tool for interpreting omics data. *Innovation*. 2021;2. <https://doi.org/10.1016/j.xinn.2021.100141>.
 10. Yu G. Thirteen years of clusterProfiler. *Innovation*. 2024;5. <https://doi.org/10.1016/j.xinn.2024.100722>.
 11. Xu S, Hu E, Cai Y, Xie Z, Luo X, Zhan L, et al. Using clusterProfiler to characterize multiomics data. *Nature Protocols*. 2024;19:3292–320. <https://doi.org/10.1038/s41596-024-01020-z>.
 12. Copperwaite M, Leifer C. *Learning Flask Framework*. Packt Publishing Ltd; 2015.
 13. EL-Sagheer RM, Eliwa MS, Alqahtani KM, EL-Morshedy M. Asymmetric randomly censored mortality distribution: bayesian framework and parametric bootstrap with application to COVID-19 data. *Journal of Mathematics*. 2022;2022:8300753. <https://doi.org/10.1155/2022/8300753>.
 14. Salunke SV, Ouda A. A performance benchmark for the PostgreSQL and MySQL databases. *Future Internet*. 2024;16:382. <https://doi.org/10.3390/fi16100382>.
 15. Liao X, Peng L, Yang T, Li T, Zhu Z. Redis-based full-text search extensions for relational databases. *International Journal of Machine Learning and Cybernetics*. 2024;15:4475–91. <https://doi.org/10.1007/s13042-024-02160-0>.
 16. Xiao Z, Lam H-M. ShinySyn: a Shiny/R application for the interactive visualization and integration of macro- and micro-synteny data. *Bioinformatics*. 2022;38:4406–8. <https://doi.org/10.1093/bioinformatics/btac503>.
 17. Katoh K, Rozewicki J, Yamada KD. MAFFT online service: multiple sequence alignment, interactive sequence choice and visualization. *Briefings in Bioinformatics*. 2019;20:1160–6. <https://doi.org/10.1093/bib/bbx108>.
 18. Price MN, Dehal PS, Arkin AP. FastTree: computing large minimum evolution trees with profiles instead of a distance matrix. *Molecular Biology and Evolution*. 2009;26:1641. <https://doi.org/10.1093/molbev/msp077>.
 19. Diesh C, Stevens GJ, Xie P, De Jesus Martinez T, Hershberg EA, Leung A, et al. JBrowse 2: a modular genome browser with views of synteny and structural variation. *Genome Biology*. 2023;24:74. <https://doi.org/10.1186/s13059-023-02914-z>.
 20. Priyam A, Woodcroft BJ, Rai V, Moghul I, Munagala A, Ter F, et al. Sequenceserver: a modern graphical user Interface for custom BLAST databases. *Molecular Biology and Evolution*. 2019;36:2922–4. <https://doi.org/10.1093/molbev/msz185>.
 21. He W, Zhuo Y, Wang C, Huang Y, Zang X, Yang C, et al. Genome-wide primer scan (GPS): a python package for a flexible, reliable and large-scale primer design toolkit. *Frontiers of Computer Science*. 2025;19:192906. <https://doi.org/10.1007/s11704-024-40392-z>.
 22. Zhu T, Liang C, Meng Z, Li Y, Wu Y, Guo S, et al. PrimerServer: a high-throughput primer design and specificity-checking platform. *bioRxiv*. 2017;:181941. <https://doi.org/10.1101/181941>.
 23. Cheng Y, Sun J, Jiang M, Luo Z, Wang Y, Liu Y, et al. Chromosome-scale genome sequence of *Suaeda glauca* sheds light on salt stress tolerance in halophytes. *Horticulture Research*. 2023;10:uhad161. <https://doi.org/10.1093/hr/uhad161>.

ARTICLE IN PRESS