Horticulture Research

Horticulture Research, 2024, **11**: uhae238

https://doi.org/10.1093/hr/uhae238

Article

CGD: a multi-omics database for *Chrysanthemum* genomic and biological research

Jingxuan Ye[1,2], Chun Wang[1,2], Ye Liu[1,2], Shaocong Chen[1,2], Jinyu Jin[1,2], Lingling Zhang[1,2], Peixue Liu[1,2], Jing Tang[1,2], Jing Zhang[1,2], Zhenxing Wang[1,2], Jiafu Jiang[1,2], Sumei Chen[1,2], Fadi Chen[1,2] and Aiping Song[1,2,*]

[1]State Key Laboratory of Crop Genetics and Germplasm Enhancement and Utilization, Sanya Institute of Nanjing Agricultural University, Key Laboratory of Landscaping, Ministry of Agriculture and Rural Affairs, Key Laboratory of Biology of Ornamental Plants in East China, National Forestry and Grassland Administration. College of Horticulture, Nanjing Agricultural University, Nanjing 210095, China
[2]Zhongshan Biological Breeding Laboratory, No.50 Zhongling Street, Nanjing 210014, China
*Corresponding author. E-mail: aiping_song@njau.edu.cn

Abstract

Asteraceae is the largest family of dicotyledons and includes *Chrysanthemum* and *Helianthus*, two important genera of ornamental plants. The genus *Chrysanthemum* consists of more than 30 species and contains many economically important ornamental, medicinal, and industrial plants. To more effectively promote *Chrysanthemum* research, we constructed the CGD, a *Chrysanthemum* genome database containing a large amount of data and useful tools. The CGD hosts well-assembled reference genome data for six *Chrysanthemum* species. These genomic data were fully annotated by comparison with various protein and domain data. Transcriptome data for nine different tissues, five flower developmental stages, and five treatments were subsequently added to the CGD. A fully functional 'RNA data' module was designed to provide complete and visual expression profile data. In addition, the CGD also provides many of the latest bioinformatics analysis tools, such as the efficient sgRNA search tool for *Chrysanthemum*. In conclusion, the CGD provides the latest, richest, and most complete multi-omics resources and powerful tools for *Chrysanthemum*. Collectively, the CGD will become the central gateway for *Chrysanthemum* genomics and genetic breeding research and will aid in the study of polyploid evolution.

Introduction

As the largest dicotyledonous family, Asteraceae (also called Compositae) consists of more than 25 000 species in approximately 1600 genera and is widely distributed worldwide [1]. The family includes many important genera of industrial and ornamental plants, such as *Chrysanthemum* and *Helianthus*. Genus *Chrysanthemum*, as an important part of the Compositae family, contains many plants with the ornamental, medicinal and industrial value. *Chrysanthemum morifolium*, a perennial herbaceous plant that originated in China, has a history of cultivation comprising more than 3000 years [2, 3]. Since the eighth century AD, Japan and Europe have introduced cultivated chrysanthemum cultivars from China and bred many varieties [4]. At present, chrysanthemums have become one of the most economically important flowers in the world [5]. In October 2018, the genome of the first wild diploid species of the genus *Chrysanthemum*, *Chrysanthemum nankingense*, was completely sequenced and resolved [6]. In the following 6 years, the genomes of the diploid wild *Chrysanthemum* species *Chrysanthemum seticuspe* [7], *Chrysanthemum lavandulifolium* [8], *Chrysanthemum makinoi* [9], and *Chrysanthemum indicum* [10] were obtained and reported. In April 2023, the first genome of cultivated chrysanthemum (*C. morifolium* Ramat., Chinese name 'Ju Hua') was resolved and formally published [11]. These valuable data have laid a solid foundation for basic research and industrial development of genus *Chrysanthemum*.

In recent years, the development of multi-omics technologies has provided new opportunities for researchers to analyse crop growth and development from multiple dimensions [12, 13]. Based on these multi-omics data, various multi-omics analysis tools, such as genome-wide association studies (GWASs), BSA-seq and weighted correlation network analysis (WGCNA) have been developed and widely used in genetic analysis and breeding. These techniques have demonstrated the power of multi-omics in assisted breeding [14–16]. To better utilize multi-omics data, multi-omics platforms have been developed for numerous important crops, such as Brassicaceae [17], Cucurbitaceae [18], the genus *Camellia* [19], rice (*Oryza sativa* L.) [20], wheat (*Triticum aestivum* L.) [21], maize (*Zea mays* L.) [22], tomato (*Solanum lycopersicum*) [23], and rapeseed (*Brassica napus*) [24]. The lack of a multi-omics platform that integrates multi-omics data and bioinformatics tools has greatly slowed the progress of basic research and genetic breeding in chrysanthemum.

To this end, we constructed the *Chrysanthemum* genome database (CGD) (CGD, http://210.22.121.250:8880/asteraceae/homePage), the first *Chrysanthemum* bioinformation center. The CGD contains all the most recently published genomic data, rich transcriptome data, and newly developed or integrated bioinformatics tools. We believe that the use of multi-omics data and tools integrated in the CGD can accelerate basic research on *Chrysanthemum* and provide effective tools for assisted genetic

Received: 18 March 2024; **Accepted:** 12 August 2024; **Published:** 21 August 2024; **Corrected and Typeset:** 1 October 2024
© The Author(s) 2024. Published by Oxford University Press on behalf of Nanjing Agricultural University. This is an Open Access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted reuse, distribution, and reproduction in any medium, provided the original work is properly cited.

Downloaded from https://academic.oup.com/hr/article/11/11/uhae238/7738309 by guest on 11 February 2026

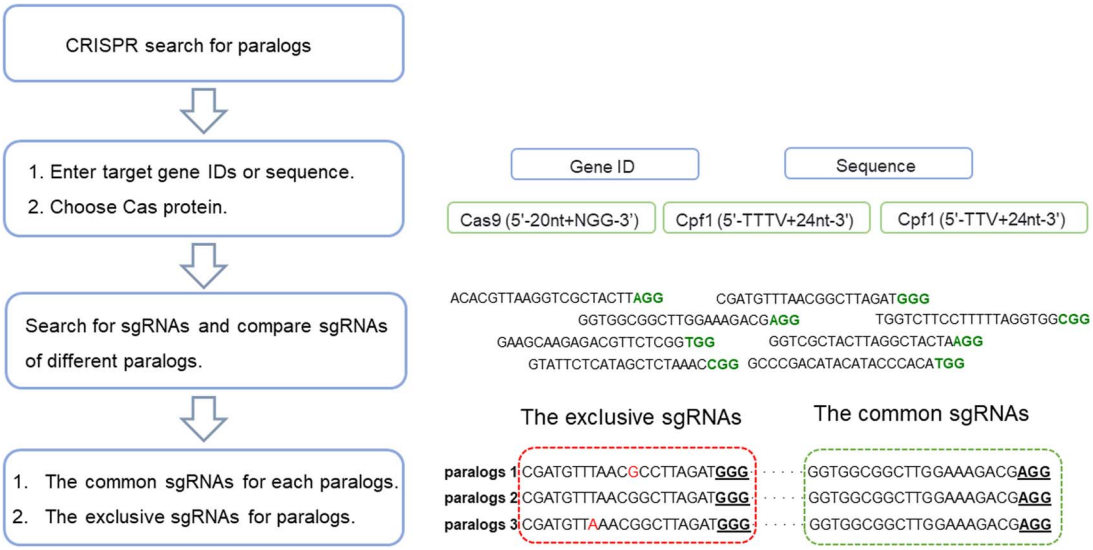


Figure 4. Workflow of the CRISPR search for paralogs.

provides an 'RNA data' module where users can obtain expression profiles by choosing either to search by gene ID (Fig. 3A) or to search by range (Fig. 3B). In addition, the expression profiles of the genes of interest in different RNA-seq items can be accessed through the 'Expression' module provided on the gene page. Furthermore, the eFP browser module was added to the CGD for visualization of all the expression profiles (Fig. 3C). Now, the expression profile data of specific genes can be accessed simply and directly.

CRISPR module

Gene editing technology plays a key role in realizing precision-directed breeding in crops. CRISPR (clustered regularly interspaced short palindromic repeats)-Cas (CRISPR associated) has evolved into one of the most advanced crop genetic engineering systems and has been widely used in ornamental crops [32, 33]. To better support the development of gene editing for *Chrysanthemum* breeding, we developed the *Chrysanthemum* CRISPR search system by integrating all the publicly released genomic data of genus *Chrysanthemum*, CRISPR-P v2.0, and CRISPR-local system [34, 35]. This function includes three Cas protein guide sequence design strategies: SpCas9, LbCas12a, and FnCas12a, which provides a variety of toolkits for *Chrysanthemum* gene editing. In 'CRISPR search' module, users can obtain the sgRNA sequence of the target gene by inputting one gene number of the target gene and selecting the sgRNA according to the score and the probability of off-target editing events given by the system. In addition, we optimized and developed a CRISPR search module for editing multiple paralogs for complex polyploid characteristics of cultivated chrysanthemums, which is called CRISPR search for paralogs (Fig. 4). Now, users can use this system to design generic targets for paralogs or specific targets for individual genes within paralogs.

Other tools

In addition to the above bioinformatics tools, many useful genome analysis tools, such as Muscle, GeneWise, GO, and KEGG enrichment, and transcriptome analysis tools, such as correlation heatmaps and co-expression networks, have been added to the CGD. To achieve a great representation of the large amount of multi-omics data, we implemented a variety of data visualization methods in addition to expression data. In the comparative genomics module, genomic syntenic blocks and syntenic gene pairs between uniform and different species can be easily obtained. These common and useful bioinformatics analysis tools embedded in databases can help users analyse *Chrysanthemum* multi-omics data. In the future, we will add more useful bioinformatics tools to the CGD and continue to optimize the suitability of existing tools in the CGD.

Discussion and perspectives

In this study, we established the CGD, the first *Chrysanthemum* database based on genomic and transcriptome data. The CGD provides the most up-to-date multi-omics dataset of *Chrysanthemum* species to date, including genomic information, gene sequences and annotations, and spatiotemporal information on gene expression. Based on these complete and rich data, we have developed and integrated numerous useful bioinformatics analysis and visualization tools. This approach will help researchers understand the evolutionary relationships of different *Chrysanthemum* species, clarify the regulatory networks of growth and development, and uncover key genes involved.

In conclusion, these data and tools fully demonstrate the powerful potential of the CGD in basic research and genetic breeding. In the future, we will strengthen our collaboration with *Chrysanthemum* researchers around the world. In addition, we will regularly collect new datasets to analyse and add to the CGD, including but not limited to genomes (for *de novo* assembly and resequencing), proteomes, transcriptomes (for single-cell transcriptomes and spatiotemporal transcriptomes), epigenomes, metabolomes, phenomes, and phenotypes. We believe that the CGD will become a central gateway for scientific research on *Chrysanthemum* species in the near future, further extending its potential and thus promoting the development of the chrysanthemum industry.

Acknowledgements

This research was funded by the Forestry Science and Technology Innovation and Promotion Project in Jiangsu Province (LYKJ [2023]22), Jiangsu Province Innovation Capacity Construction

Downloaded from https://academic.oup.com/hr/article/11/11/uhae238/7738309 by guest on 11 February 2026

Program (BM2022008-02), the National Natural Science Foundation of China (32072609, 32172609), the Achievement Transformation Fund project of Hainan Research Institute of Nanjing Agricultural University (NAUSY-CG-YB04), the 'Blue Project' of Jiangsu Higher Education Institutions and a project Funded by the Priority Academic Program Development of Jiangsu Higher Education Institutions.

Data availability

All the data hosted in the CGD are freely available at (http://210.22.121.250:8880/asteraceae/homePage).

Conflict of interest statement

The authors declare that they have no conflict of interest.

Supplementary data

Supplementary data is available at *Horticulture Research* online.

References

1. The Angiosperm Phylogeny Group, Chase MW, Christenhusz MJM. *et al.* An update of the angiosperm phylogeny group classification for the orders and families of flowering plants: APG IV. *Bot J Linn Soc*. 2016;**181**:1–20
2. Chen J, Wang C, Zhao H. *et al. The Origin of Garden Chrysanthemum*. Hefei: Anhui Science & Technology Publishing House; 2012:
3. Anderson NO. *Flower Breeding and Genetics: Issues, Challenges and Opportunities for the 21st Century*. Dordrecht: Springer, 2006
4. Ohmiya A. Molecular mechanisms underlying the diverse array of petal colors in chrysanthemum flowers. *Breed Sci*. 2018;**68**:119–27
5. Boase MR, Miller R, Deroles SC. Chrysanthemum systematics, genetics, and breeding. *Plant Breeding Rev*. 2010;**14**:321–61
6. Song C, Liu Y, Song A. *et al.* The *Chrysanthemum nankingense* genome provides insights into the evolution and diversification of chrysanthemum flowers and medicinal traits. *Mol Plant*. 2018;**11**:1482–91
7. Nakano M, Hirakawa H, Fukai E. *et al.* A chromosome-level genome sequence of *Chrysanthemum seticuspe*, a model species for hexaploid cultivated chrysanthemum. *Commun Biol*. 2021;**4**:1167
8. Wen X, Li J, Wang L. *et al.* The *Chrysanthemum lavandulifolium* genome and the molecular mechanism underlying diverse capitulum types. *Hortic Res*. 2022;**9**:uhab022
9. van Lieshout N, van Kaauwen M, Kodde L. *et al.* De novo whole-genome assembly of *Chrysanthemum makinoi*, a key wild chrysanthemum. *G3 (Bethesda)*. 2022;**12**:jkab358
10. Deng Y, Yang P, Zhang Q. *et al.* Genomic insights into the evolution of flavonoid biosynthesis and O-methyltransferase and glucosyltransferase in *Chrysanthemum indicum*. *Cell Rep*. 2024;**43**:113725
11. Song A, Su J, Wang H. *et al.* Analyses of a chromosome-scale genome assembly reveal the origin and evolution of cultivated chrysanthemum. *Nat Commun*. 2023;**14**:2021
12. Lihavainen J, Šimura J, Bag P. *et al.* Salicylic acid metabolism and signalling coordinate senescence initiation in aspen in nature. *Nat Commun*. 2023;**14**:4288
13. Uhrig RG, Schläpfer P, Roschitzki B. *et al.* Diurnal changes in concerted plant protein phosphorylation and acetylation in *Arabidopsis* organs and seedlings. *Plant J*. 2019;**99**:176–94
14. Zhang X, Su J, Jia F. *et al.* Genetic architecture and genomic prediction of plant height-related traits in chrysanthemum. *Hortic Res*. 2023;**11**:uhad236
15. Wang S, Wang K, Li Z. *et al.* Architecture design of cucurbit crops for enhanced productivity by a natural allele. *Nat Plants*. 2022;**8**:1394–407
16. Ding L, Song A, Zhang X. *et al.* The core regulatory networks and hub genes regulating flower development in *Chrysanthemum morifolium*. *Plant Mol Biol*. 2020;**103**:669–88
17. Chen H, Wang T, He X. *et al.* BRAD V3. 0: an upgraded Brassicaceae database. *Nucleic Acids Res*. 2022;**50**:D1432–41
18. Yu J, Wu S, Sun H. *et al.* CuGenDBv2: an updated database for cucurbit genomics. *Nucleic Acids Res*. 2023;**51**:D1457–64
19. Gao Q, Tong W, Li F. *et al.* TPIA2: an updated tea plant information archive for *camellia* genomics. *Nucleic Acids Res*. 2023;**52**:D1661–7
20. Sakai H, Lee SS, Tanaka T. *et al.* Rice annotation project database (RAP-DB): an integrative and interactive database for rice genomics. *Plant Cell Physiol*. 2013;**54**:e6–6
21. Ma S, Wang M, Wu J. *et al.* WheatOmics: a platform combining multiple omics data to accelerate functional genomics studies in wheat. *Mol Plant*. 2021;**14**:1965–8
22. Gui S, Yang L, Li J. *et al.* ZEAMAP, a comprehensive database adapted to the maize multi-omics era. *iScience*. 2020;**23**:101241
23. Kudo T, Kobayashi M, Terashima S. *et al.* TOMATOMICS: a web database for integrated omics information in tomato. *Plant Cell Physiol*. 2017;**58**:e8–8
24. Yang Z, Wang S, Wei L. *et al.* BnIR: a multi-omics database with various tools for *Brassica napus* research and breeding. *Mol Plant*. 2023;**16**:775–89
25. Altschul SF, Gish W, Miller W. *et al.* Basic local alignment search tool. *J Mol Biol*. 1990;**215**:403–10
26. Wang Y, Tang H, DeBarry JD. *et al.* MCScanX: a toolkit for detection and evolutionary analysis of gene synteny and collinearity. *Nucleic Acids Res*. 2012;**40**:e49–9
27. Pruitt KD, Tatusova T, Maglott DR. NCBI reference sequence (RefSeq): a curated non-redundant sequence database of genomes, transcripts and proteins. *Nucleic Acids Res*. 2005;**33**:D501–4
28. UniProt Consortium. UniProt: the universal protein knowledgebase in 2021. *Nucleic Acids Res*. 2021;**49**:D480–9
29. Gene Ontology Consortium, Aleksander SA, Balhoff J. *et al.* The gene ontology knowledgebase in 2023. *Genetics*. 2023;**224**:iyad031
30. Kanehisa M, Furumichi M, Sato Y. *et al.* KEGG for taxonomy-based analysis of pathways and genomes. *Nucleic Acids Res*. 2023;**51**:D587–92
31. Blum M, Chang HY, Chuguransky S. *et al.* The InterPro protein families and domains database: 20 years on. *Nucleic Acids Res*. 2021;**49**:D344–54
32. Shan Q, Wang Y, Li J. *et al.* Targeted genome modification of crop plants using a CRISPR-Cas system. *Nat Biotechnol*. 2013;**31**:686–8
33. Tang J, Ye J, Liu P. *et al.* Ornamental plant gene editing: past, present and future. *Ornam Plant Res*. 2023;**3**:1–6
34. Liu H, Ding Y, Zhou Y. *et al.* CRISPR-P 2.0: an improved CRISPR-Cas9 tool for genome editing in plants. *Mol Plant*. 2017;**10**:530–2
35. Sun J, Liu H, Liu J. *et al.* CRISPR-local: a local single-guide RNA (sgRNA) design tool for non-reference plant genomes. *Bioinformatics*. 2019;**35**:2501–3

Downloaded from https://academic.oup.com/hr/article/11/11/uhae238/7738309 by guest on 11 February 2026