Research Square

Preprints are preliminary reports that have not undergone peer review.
They should not be considered conclusive, used to inform clinical practice,
or referenced by the media as validated information.

Network preservation analysis to identify transcriptional biomarkers related to flowering in Crocus sativus

Mahsa Eshaghi
Tarbiat Modares University
Sajad Rashidi-Monfared (rashidims@modares.ac.ir)
Tarbiat Modares University

Article

Keywords: Crocus sativus, preservation module analysis, hub genes, flowering process, WGCNA, NetRep

Posted Date: September 11th, 2023

DOI: https://doi.org/10.21203/rs.3.rs-3325141/v1

License: This work is licensed under a Creative Commons Attribution 4.0 International License.
Read Full License

Abstract

Crocus sativus L. is known as an ornamental geophyte and a source of valuable spice and secondary metabolites. Network preservation module analysis is one of the best approaches to revealing special features of different conditions. It can determine patterns of divergence and conservation between transcriptome data. Herein, we explored the regulatory genes of the flowering process by RNA-Seq data containing flowering and non-flowering samples in gene expression profiles. Persevered module analysis revealed three significant non-persevered modules related to the flowering process, namely pink, green, and blue. Several hub genes associated with non-preserved modules such as PIA1, NAC90, ALY3, Sus3, MYB31, ARF5/MP, MYB31, HD-ZIP, SEP3d, OR_B, AGL6a, bZIP(TGA1) and GRAS were identified. These candidate genes can be considered key diagnostic biomarkers for the flowering process. Here, we also compare two approaches, WGCNA and NetRep for module preservation analysis. Totally, the results of these methods were consistent with non-preserved modules. NetRep was a faster (11 times) and more efficient (run more than 10000 permutations for each comparison) method than WGCNA module preservation. Differential expression genes (DEGs) screening showed that many hub genes were downregulated in non-flowering than flowering samples. Our finding revealed regulatory mechanisms of the flowering process in *C. sativus* as can be developed transcriptional biomarkers which could pave the way for promoting saffron yield via flowering induction.

Highlights

- Network perseveration analysis was performed in order to identify hub genes and transcriptional biomarkers related to the flowering process in *Crocus sativus*.
- Three special non-preserved modules related to the flowering process were identified.
- The results of NetRep and WGCNA module preservation approaches were consistent with non-preserved modules.

Introduction

Saffron is the highest-priced spice in the world. The flower is the most treasured part of saffron and among the flowering component, the stigma of saffron is widely applied as a spice or coloring and flavoring agent and medicine plant in different industries. There are six main growth stages in saffron including sprouting, cataphylls, flowers appearance, plant appearance and development, replacement corms development, plant senescence, and dormancy of corm [1].

The flowering process is associated with the expression of a group of genes that interact cooperatively within and across different biological pathways. In recent decades, the challenges of withering, rotting, and delayed flowering in saffron flowers have severely affected the quality and quantity of its stigmas and the development of the saffron industry. Therefore, research on the molecular regulatory mechanisms of saffron flowering factors is an important subject to solve these problems. Some of the important factors like sugars, hormones and temperature are involved in flowering processes in *C.*

sativus[2–5]. Transcriptomic studies have illustrated that photoperiodism and vernalization pathways in saffron affect floral induction [6,7].

Undoubtedly, the production of saffron from signaling pathways and genetic factors play important roles in the development of saffron flower as well. However, the accurate molecular and regulatory genetic mechanisms are not fully clear yet. Detecting flowering regulatory genes play an essential role in increasing and developing flowers, thereby gaining in high saffron yield.

WGCNA is a useful methodology for systems biology that investigates the co-expression of genes based on correlation. WGCNA can identify modules as a group of genes with similar expression and highly correlated, and associated hub genes in each module. [8]. NetRep is a rapid and computationally effective pattern that uses a permutation approach. NetRep was performed to generate and evaluate the practical null distributions of module preservation statistics [9]. Module preservation is the approach for differential network analysis[10]. Module preservation seeks to reveal patterns of conservation and divergence between transcriptomes [11]. Several studies have effectively applied module preservation analysis to compare and contrast the effective gene networks[11–14].

In order to recognize co-regulatory modules and related hub genes of the flowering process, weighted gene co-expression network analysis of saffron was performed at both flowering and the non-flowering groups, and then a network preservation statistics approach was employed to contrast and compare the candidate gene networks by measuring the preservation of flowering modules in the non-flowering co-expression network. This analysis facilitates the distinction of factors involved in flowering and non-flowering specific features in the flowering plants for developing related biomarkers.

Results

WGCNA analysis of transcriptome data of C. sativus flowering

Overall, 267345823 raw reads from two groups of RNA-Seq samples related to flowering and non-flowering of *C. sativus* were analyzed (Supplementary Table S1), and after trimming processes, 261998906 clean reads were obtained. Gene expression analysis and normalization were performed. 997 genes, and TFs were detected related to flowering through Blastx. The gene expression matrix of the samples in flowering and non-flowering groups was obtained. After filtering and removal of outliers, co-expression networks were constructed in order to detect the regulatory mechanism and biomarkers involved in flowering in *C. sativus*. The suitable soft threshold power beta was determined as 16 for scale-free network (Supplementary Fig. 1).

Hierarchical clustering based on the topological overlap matrix (TOM) dissimilarity analysis detected 10 modules of highly expressed genes, ranging from 33 in purple to 214 nodes in turquoise (Fig. 1). The average module size was reported about 87 genes. Using the plantTFDB database, a total of 716 TFs were recognized in the modules, in which the turquoise module with 188 TFs and blue module with 124

TFs had the most significant numbers of TFs, and the pink module with 36 TFs and magenta modules with 67 TFs had the minor numbers of TFs. 32 genes with no specific module assignment were known as the grey module.

Preservation Analysis

Modules preservation analysis was performed to detect the connectivity patterns between the two flowering and non-flowering samples. ModulePreservation function was used in order to calculate density and connectivity based on the preservation score. Results indicated that three modules including pink (Zsummary = 0.53), green (Zsummary = 1.50) and blue (Zsummary = 2) were highly non-preserved, while brown (Zsummary = 7.90) and turquoise (Zsummary = 6.20) modules were semi-preserved. The green and pink modules have the highest median rank statistic among all modules. No modules showed strong evidence of preservation (Zsummary > 10) in this method (Fig. 2). Network properties and connectivity patterns in the non-preserved modules were different in flowering and non-flowering samples, so they might be related to the genes involved in flowering. Also, the NetRep permutation test was performed. We considered a module to be non-preserved in the test set if all the statistics had a permutation test P-value > 0.01. The non-parametric approach has been used to produce unbiased P-values in NetRep. The permutation p values were computed based on null distributions. The pink module (P-value = 0.97), green module (P-value = 0.95) and blue module (P-value = 0.64) were the most non-preserved modules at this threshold and the red module (P-value = 0.0086) was preserved module. The cor.cor, cor.degree, and cor.contrib parameters were very small (0.00019998, 0.00459954, and 0.00629937, respectively) in red module. The permutations were 10000 and the run took place 11 times faster in NetRep than WGCNA (Fig. 3).

Hub detection

Genes with higher connectivity inside the module are known as hub genes. It is suggested that these genes are more informative [15]. The results of hub genes analysis in the most important non-preserved modules considering the worth of the hub in the network illustrated that genes including PIA1(PISTILLATA-like MADS box protein), NAC90, MYB_related (ALY3) were identified as the hub genes in the pink module. PISTILLATA/GLOBOSA-like (PI/GLO-like) MADS-box genes regulate floral organ identity in *C. staivus*. ALY3 takes part in plant growth and development [16]. MYB related, HB-PHD, Sus3 (sucrose synthase 3), MYB31, ARF5/MP hub genes were identified in the green module. Sucrose synthase can enhance primary growth and alter flower morphology. The blue module contained bZIP, 4HD-ZIP GRAS, SEP3d (SEPALLATA3-like MADS-box protein), MADS-box transcription factor AGL6a (AGL6a) and orange isoform X2(OR_B) hub genes (Fig. 4, Supplementary Table S2). The family of GRAS proteins plays essential roles in regulating plant growth, the DELLA proteins, are related to a transcription factor involved in phytochrome signaling.[17,18]. Normalized read count data and statistical analysis were used to find quantitative changes in expression levels between flowering and non-flowering groups in the EdgeR package. The expression of NAC90 was characterized by a five-fold decrease in the non-flowering groups while Sus3 and bZIP were expressed at low levels in non-flowering samples. The results of DEG show

that the most of hub genes in pink, green, and blue modules were downregulated in the non-flowering group compared with the flowering sample (Supplementary Table S3), so these genes can be suggested as suitable transcriptional diagnostic biomarkers.

Functional analysis in modules

The modules were annotated based on the functional enrichment and pathways analysis. The results illustrated that the most enriched pathways belong to metabolism, genetic information processing and signaling, and cellular processes in non-preserved modules. The green module was enriched by plant hormone signal transduction (ko04075). Metabolic pathways (ko01100) and signaling pathways (ko04350) were enrichment pathways in blue and pink modules, respectively. Go enrichment analysis showed that binding (GO:0005488), metabolic process (GO:0008152), nucleic acid binding (GO:0003676) and cellular process (GO:0009987) were the most enriched in modules and biological processes were the most enriched in non-preserved modules. The biological processes of the pink module were detected to focus on binding (GO:0005488), metabolic process (GO:0008152), and cellular process (GO:0009987). At the same time, the green module which is enriched for binding (GO:0005488), metabolic process (GO:0008152), and cellular process (GO:0009987). The biological processes of the blue module were detected to focus on binding (GO:0005488), metabolic processes (GO:0008152) and primary metabolic process (GO:0044238), cellular metabolic process (GO:0044260), and macromolecule metabolic process (GO:0043170). The list of the functional enrichment analysis for non-preserved modules is existing in Supplementary Table S4.

Discussion

C. staivus is one of the ornamental and commercial geophytes. Some of the studies of this plant related to the effects of environmental factors on physiological aspects of flowering. However, the regulatory mechanisms of flowering and cause-effect relationships between flowering and non-flowering conditions are still limited. The network analysis and system biology approach can help to improve production. In this study, we used RNA-Seq data related to the flowering and non-flowering of *C. sativus* to establish a co-expression network and preservation analysis for genes involved in the flowering process by using WGCNA and NetRep packages. The goal of this study was to characterize the main molecular signature of the flowering process and identify commonly regulated specific gene modules which could describe the correlated incidence of the two conditions. Next, we concentrated on the non-preserved modules. It is remarkable that the loss of connectivity among genes in the non-preserved modules could be ascribed to the unusual expression of some genes under non-flowering conditions, which probably are signature factors. Different approaches such as module assignment, functional analysis, and hub genes detection were performed to determine genes related to the occurrence and development of the flowering process. We compared two approaches ModulePreservation function in WGCNA and NetRep analysis. The results of these approaches were nearly consistent but the runtime comparison between these approaches showed that NetRep was faster than WGCNA and red module was found as preserved module. In this study, three important non-preserved modules associated with the flowering process, namely pink, green,

and blue were identified. Moreover, the outcomes of the functional enrichment analysis of these modules determined that some terms were related to most of the features of the developmental, metabolic process, and signaling associated with flowering. Hub gene features associated with the regulation of flower development make them noteworthy candidates to be used as diagnostic transcriptional biomarkers of the flowering factors. Remarkably, some of the genes in the non-preserved modules were reported to be related to the flowering process in the previous studies that examined the flowering regulatory genes in *C. sativus*[2,6].

The PIA1 gene which regulates floral organ identity and is member of MADS-box transcription factors control plant flower development and are related to floral organ identity determination [19,20] was identified as the main regulatory hub gene of the pink module recognized in the network, which shows the highest preservation score in the flowering co-expression network. As the most connected node of the pink module, PIA1 is a fascinating regulatory hub candidate. The pink module shows genes related to the effect of the hormone factors GA2OX8 (gibberellin2-beta-dioxygenase-8-lik) and cell cycle regulation (CYCB1.2) in the flowering process. CYCB1.2 plays an important role in the cell cycle regulation pathway. CYCB1.2 is most highly expressed in flowers compared with other tissues (Day and Reddy, 1998). Sus genes are involved in carbohydrate metabolism as an important hub gene in the green non-preserved module. The alteration of carbohydrate metabolism can be impacted plant biomass production and flower development. Actually, altering the expression of sucrose metabolism-related genes such as SuSy in tobacco was reported in previous studies, so the suc3 hub gene in *C. sativus* can play essential roles in enhancing primary growth and flower morphology alteration [21]. Also, in the green module, squamosa promoter-binding-like protein 15 (SPL15) was found which regulates juvenile-to-adult growth phase transition in plants. [22] The enrichment analysis of the blue module associated with metabolic and cellular processes reveals the critical roles of genes within the module in flowering signaling pathways. Previous studies described the key role of ARF5/mp transcription factor in flower initiation and specific gene expression [23–26] so the hub gene related to this factor was important in the green module. R2R3-MYB TF participates in the flowering regulation [27]. Non-preserve modules analysis showed that MYB plays regulated roles in the flowering process in *C. sativus*. HSFs are transcription factors related to heat stress [28]. In the pink and blue non-preserved module, we found HSF TFs, which suggested that plays a regulatory role in the flowering process. The previous study demonstrated the relationship of HSF TFs in the induction of inflorescence meristem formation in *C. sativus*[6]. Totally, a group of molecular mechanisms that work together in a complex regulatory network conducts the interaction among different conditions[29]. Differential expression screening determines the levels of difference between groups or conditions. It can be able to discover candidate genes putatively involved in the metabolic processes [30–32]. Mostly, differentially expressed hub genes were downregulated in the non-flowering sample compared with the flowering group in non-preserved modules.

These findings reveal modules as well as non-preserved modules in the flowering process. We reported a list of candidate hub genes that have the potential of affecting the extensive biological network they regulate. This information can be promoted our knowledge of the regulatory mechanisms related to the

flowering process. Although our results provide greater insights into the biological process involved in flowering, further researches still require to be applied to realize the exact biological function of the proposed candidate modules and their gene members. Understanding regulatory factors in the flowering process in *C. sativus* could disclose ways to expand the harvesting period, and it would facilitate the development of industrial production.

Conclusion

C. sativus is a valuable medicinal plant and commercial source of saffron spice. The flower is a precious part of this plant. In the present study, co-expression networks related to the flowering process were constructed. Perseveration analysis was performed to determine non-preserved modules and functional enrichment analysis of these modules in *C. sativus*. We detected 10 modules associated with the flowering process and especially four important non-persevered modules such as pink, green, and blue were identified as different between flowering and non-flowering samples. The results revealed some important hub genes PIA1, NAC90, ALY3, Sus3, MYB31, ARF5/MP, MYB31, HD-ZIP, SEP3d, OR_B, AGL6a, bZIP(TGA1) and GRAS which play crucial roles in the flowering process and determined the potential of these genes regarded as biomarkers. Interestingly, the result of ModulePreservation in WGCNA analysis was nearly consistent with the NetRep results which confirmed the role of identified genes in non-preserved modules related to the flowering process in *C. sativus*. Overall, the result of the biological network analysis of *C. sativus* and understanding the regulatory mechanisms underlying the flowering process can pave the way to improve the yield of saffron.

Material and Methods

Data collection and RNA-Seq data analysis

RNA sequencing experiments of two groups including the flowering group and the non-flowering group for *C. sativus* were downloaded as fastq files from the National Center for Biotechnology Information (NCBI) Sequence Read Archive database (https://www.ncbi.nlm.nih.gov/sra) with BioProject number PRJNA524437. The accession number and details of these projects are mentioned in Supplementary TableS1.

FastQC tool (Version 0.11.9)[33] was used for quality control, and low-quality reads and adaptors were removed by fastp tool (version 0.20.1)[34]. Genes involved in the flowering process were identified through Blastx v2.6.0 + against the NR database and Arabidopsis Information Resource (TAIR) (http://www.Arabidopsis.org). The mapping of reads was applied using Bowtie 2 (version 2.4.1)[35]. The Transcripts Per Kilobase Million (TPM) values from the salmon (version 1.9.0)[36] outputs were denoted as an expression matrix. Log2 transformation was applied in order to decrease residual variability and normalization.

Co-expressed gene network analysis

The Weighted Gene Co-expression Network Analysis (WGCNA) approach for *C. sativus* transcriptome data was used to build a co-expression network in the R language to reveal flowering-associated modules [37]. First, the unqualified genes were removed by using the goodSamplesGenes function in the WGCNA package. The outliers were filtered out, then cluster groups of highly co-expressed genes were extracted from the expression matrix. Step by step function in the WGCNA package was applied to construct modules. The adjacency matrix was transformed into a weighted correlation matrix with considered soft thresholding of the power as 16 to compute a topological overlap matrix (TOM) and make a hierarchical cluster tree. Lastly, the module detection was performed with the main parameters a minimum module size of 30, the power of 16, corType = “bicor”, networkType = “Signed Hybrid”, the “TOMType” of signed and “mergeCutHeight = 0.2”. Average linkage hierarchical clustering analysis based on a dynamic hybrid tree-cutting algorithm was used to cluster TOM-based adjacency matrices for identifying the modules. Then the modules with high similarity were merged according to the module eigengenes function.

Preservation Analysis

Two approaches were used to figure out the extent of the module preservation between the flowering group and non-flowering group samples. The first approach was the modulePreservation function in the WGCNA and the second, NetRep software, which is available on the R package and based on the permutation approach and p-value significant level to assign module preservation. The module preservation function in the WGCNA package was used in order to investigate preserved modules based on Zsummary and median rank criteria. This method is a complex approach that employs network separability, density, and preservation statistics based on connectivity [38] and needs adjacency matrices of both reference (flowering) and test networks (nonflowering) as input, however, module detection is only essential for the reference flowering network (Langfelder et al., 2011). The permutation test (number of permutations = 1000) which randomly permutes the module detection in the test nonflowering networks was applied. Permutation test p-value was assigned to understand the significant level of the observed value of the preservation module. Zsummary depends on the module size and needs modules with different sizes for comparison, however, median rank is less sensitive with a module size. Modules with the higher Zsummary and the lower median rank indicate strong preservation statistics. We considered modules with Zsummary < 10 or median rank > 8 as nonpreserved modules, The modules with the critical value of 5 < Zsummary < 10 and MedianRank > 8, and Zsummary > 10 and MedianRank < 8 were considered as semi-preserved and module preservation, respectively.

Also, we performed NetRep [9] calculate module preservation statistics. Null distributions of the module preservation statistics were examined on gene expression data for flowering and non-flowering samples of *C. sativus*. We considered flowering data as the discovery set and non-flowering data as a test set and estimated a p-value significant level to evaluate preservation for each module.

Detection of hub-hub genes modules

The correlation between the expression profile of a gene and the ME of a module was assigned by kME or module membership. The moduleEigengenes function and KME factor in WGCNA were applied to detect

the hub genes in related modules.12 topological analysis methods are suggested in CytoHubba to display the highly connected genes. All the methods were considered and compared in each module. Then to increase the accuracy of the prediction, maximum click centrality (MCC) and degree method in the cytoHubba plugin of Cytoscape software (version 3.9.1)[39] were applied to identify the highly connected genes as hub genes.

Differentially expressed genes assessment and GO and KEGG enrichment screening

EdgeR (version 3.42.4) [40] was used to screen differentially expressed genes with parameter cutoff p-value < 0.05 and log fold change ratio > 2. AgriGo web tool[41] was performed to GO enrichment analyses and the significantly enriched pathways were identified with the KEGG pathway database.

Declarations

Acknowledgements

We are grateful to Tarbiat Modares University for the support of this work.

Funding

This work was also supported by Tarbiat Modares University (Grand number: 89852).

Author information

Authors and Affiliations

Department of Plant Biotechnology, Faculty of Agriculture, Tarbiat Modares University, Tehran, Iran.

Mahsa Eshaghi

Department of Plant Biotechnology, Faculty of Agriculture, Tarbiat Modares University, Tehran, Iran.

Sajad Rashidi-Monfared

Author contributions

M.E: Investigation, Formal analysis, Software, Visualization, Writing – original draft. S.R.M: Conceptualization, Methodology, Data curation, Writing - review & editing, Supervision, Visualization, Project administration. All authors have read and approved the final manuscript.

Corresponding authors

Sajad Rashidi-Monfared

Consent for publication

Not applicable.

Competing interests

The authors declare no competing interests.

Data availability

The datasets analyzed during the current study are available in the https://www.ncbi.nlm.nih.gov/sra. More details can be found in the supplementary material.

Supplementary Information

Supplementary material related to this article can be found, in the online version

References

1. Córcoles, H. L., Ramos, A. B., Garcia, F. M., Valverde, M. R. & Riquelme, F. J. M. Phenological growth stages of saffron plant (*Crocus sativus* L.) according to the BBCH Scale. *Spanish journal of agricultural research* **13**, 18 (2015).
2. Renau-Morata, B. *et al.* Flower induction and development in saffron: Timing and hormone signalling pathways. *Industrial Crops and Products* **164**, 113370 (2021).
3. Jirage, D. B., Ravishankar, G., Suvarnalatha, G. & Venkataraman, L. Profile of polyamines during sprouting and growth of saffron (Crocus sativus L.) corms. *Journal of plant growth regulation* **13**, 69-72 (1994).
4. Farooq, S. & Koul, K. Changes in gibberellin-like activity in corms of saffron plant (Crocus sativus L.) during dormancy and sprouting. *Biochemie und Physiologie der Pflanzen* **178**, 685-689 (1983).
5. Bagri, J. *et al.* Metabolic shift in sugars and amino acids regulates sprouting in Saffron corm. *Scientific reports* **7**, 11904 (2017).
6. Qian, X. *et al.* Single-molecule real-time transcript sequencing identified flowering regulatory genes in Crocus sativus. *BMC genomics* **20**, 857, doi:doi.org/10.1186/s12864-019-6200-5 (2019).
7. Hu, J. *et al.* Transcriptome profiling of the flowering transition in saffron (*Crocus sativus* L.). *Scientific reports* **10**, 1-14 (2020).
8. Zhang, B. & Horvath, S. A general framework for weighted gene co-expression network analysis. *Statistical applications in genetics and molecular biology* **4** (2005).
9. Ritchie, S. C. *et al.* A scalable permutation approach reveals replication and preservation patterns of network modules in large datasets. *Cell systems* **3**, 71-82 (2016).
10. Langfelder, P., Luo, R., Oldham, M. C. & Horvath, S. Is my network module preserved and reproducible? *PLoS computational biology* **7**, e1001057 (2011).

11. Almeida-Silva, F. & Venancio, T. M. Network comparison: consensus modules and module preservation. *dim* **10802**, 28.
12. Heidari, M., Pakdel, A., Bakhtiarizadeh, M. R. & Dehghanian, F. Integrated analysis of lncRNAs, mRNAs, and TFs to identify regulatory networks underlying MAP infection in cattle. *Frontiers in Genetics* **12**, 668448 (2021).
13. Sheybani, N., Bakhtiarizadeh, M. R. & Salehi, A. An integrated analysis of mRNAs, lncRNAs, and miRNAs based on weighted gene co-expression network analysis involved in bovine endometritis. *Scientific reports* **11**, 18050 (2021).
14. Harutyunyan, A., Jones, N. C., Kwan, P. & Anderson, A. Network preservation analysis reveals dysregulated synaptic modules and regulatory hubs shared between alzheimer's disease and temporal lobe epilepsy. *Frontiers in Genetics* **13**, 821343 (2022).
15. Filteau, M., Pavey, S. A., St-Cyr, J. & Bernatchez, L. Gene coexpression networks reveal key drivers of phenotypic divergence in lake whitefish. *Molecular biology and evolution* **30**, 1384-1396 (2013).
16. Pfaff, C. *et al.* ALY RNA-binding proteins are required for nucleocytosolic mRNA transport and modulate plant growth and development. *Plant physiology* **177**, 226-240 (2018).
17. De Lucas, M. *et al.* A molecular framework for light and gibberellin control of cell elongation. *Nature* **451**, 480-484 (2008).
18. Feng, S. *et al.* Coordinated regulation of Arabidopsis thaliana development by light and gibberellins. *Nature* **451**, 475-479 (2008).
19. Parenicova, L. *et al.* Molecular and phylogenetic analyses of the complete MADS-box transcription factor family in Arabidopsis: new openings to the MADS world. *The Plant Cell* **15**, 1538-1551 (2003).
20. Nam, J., DePamphilis, C. W., Ma, H. & Nei, M. Antiquity and evolution of the MADS-box gene family controlling flower development in plants. *Molecular biology and evolution* **20**, 1435-1447 (2003).
21. Coleman, H. D., Beamish, L., Reid, A., Park, J.-Y. & Mansfield, S. D. Altered sucrose metabolism impacts plant biomass production and flower development. *Transgenic research* **19**, 269-283 (2010).
22. Schwarz, S., Grande, A. V., Bujdoso, N., Saedler, H. & Huijser, P. The microRNA regulated SBP-box genes SPL9 and SPL15 control shoot maturation in Arabidopsis. *Plant molecular biology* **67**, 183-195 (2008).
23. Przemeck, G. K., Mattsson, J., Hardtke, C. S., Sung, Z. R. & Berleth, T. Studies on the role of the Arabidopsis gene MONOPTEROS in vascular development and plant cell axialization. *Planta* **200**, 229-237 (1996).
24. Hardtke, C. S. & Berleth, T. The Arabidopsis gene MONOPTEROS encodes a transcription factor mediating embryo axis formation and vascular development. *The EMBO journal* **17**, 1405-1411 (1998).
25. Zhao, Z. *et al.* Hormonal control of the shoot stem-cell niche. *Nature* **465**, 1089-1092 (2010).
26. Yamaguchi, N. *et al.* A molecular framework for auxin-mediated initiation of flower primordia. *Developmental cell* **24**, 271-282 (2013).

27. Chen, Q. *et al.* Genome-wide identification and expression analysis of the R2R3-MYB transcription factor family revealed their potential roles in the flowering process in longan (Dimocarpus longan). *Frontiers in Plant Science* **13**, 820439 (2022).
28. Nover, L. *et al.* Arabidopsis and the heat stress transcription factor world: how many heat stress transcription factors do we need? *Cell stress & chaperones* **6**, 177 (2001).
29. Abedini, D. & Rashidi Monfared, S. Co-regulation analysis of co-expressed modules under cold and pathogen stress conditions in tomato. *Molecular biology reports* **45**, 335-345 (2018).
30. Ebrahimi, A. *et al.* Furanocoumarins from Heracleum persicum L.: Unveiling their biosynthesis and gene expression. *Industrial Crops and Products* **203**, 117160 (2023).
31. Kim, K.-H., Song, K., Park, J.-M., Kim, J.-Y. & Lee, B.-M. RNA-Seq analysis of gene expression changes related to delay of flowering time under drought stress in tropical maize. *Applied Sciences* **11**, 4273 (2021).
32. Garcia, R. S., Coronejo, S., Concepcion, J. & Subudhi, P. K. Whole-genome sequencing and RNA-seq reveal differences in genetic mechanism for flowering response between weedy rice and cultivated rice. *International journal of molecular sciences* **23**, 1608 (2022).
33. Andrews, S. (2010).
34. Chen, S., Zhou, Y., Chen, Y. & Gu, J. fastp: an ultra-fast all-in-one FASTQ preprocessor. *Bioinformatics* **34**, i884-i890, doi:doi.org/10.1093/bioinformatics/bty560 (2018).
35. Langmead, B. & Salzberg, S. L. Fast gapped-read alignment with Bowtie 2. *Nature methods* **9**, 357 (2012).
36. Patro, R., Duggal, G., Love, M. I., Irizarry, R. A. & Kingsford, C. Salmon provides fast and bias-aware quantification of transcript expression. *Nature methods* **14**, 417, doi:doi.org/10.1038/nmeth.4197 (2017).
37. Langfelder, P. & Horvath, S. WGCNA: an R package for weighted correlation network analysis. *BMC bioinformatics* **9**, 559, doi:doi.org/10.1186/1471-2105-9-559 (2008).
38. Oldham, M. C., Horvath, S. & Geschwind, D. H. Conservation and evolution of gene coexpression networks in human and chimpanzee brains. *Proceedings of the National Academy of Sciences* **103**, 17973-17978 (2006).
39. Chin, C.-H. *et al.* cytoHubba: identifying hub objects and sub-networks from complex interactome. *BMC systems biology* **8**, S11, doi:doi.org/10.1186/1752-0509-8-S4-S11 (2014).
40. Chen, Y., Lun, A. T. & Smyth, G. K. From reads to genes to pathways: differential expression analysis of RNA-Seq experiments using Rsubread and the edgeR quasi-likelihood pipeline. *F1000Research* **5** (2016).
41. Tian, T. *et al.* agriGO v2. 0: a GO analysis toolkit for the agricultural community, 2017 update. *Nucleic acids research* **45**, W122-W129 (2017).

Figures

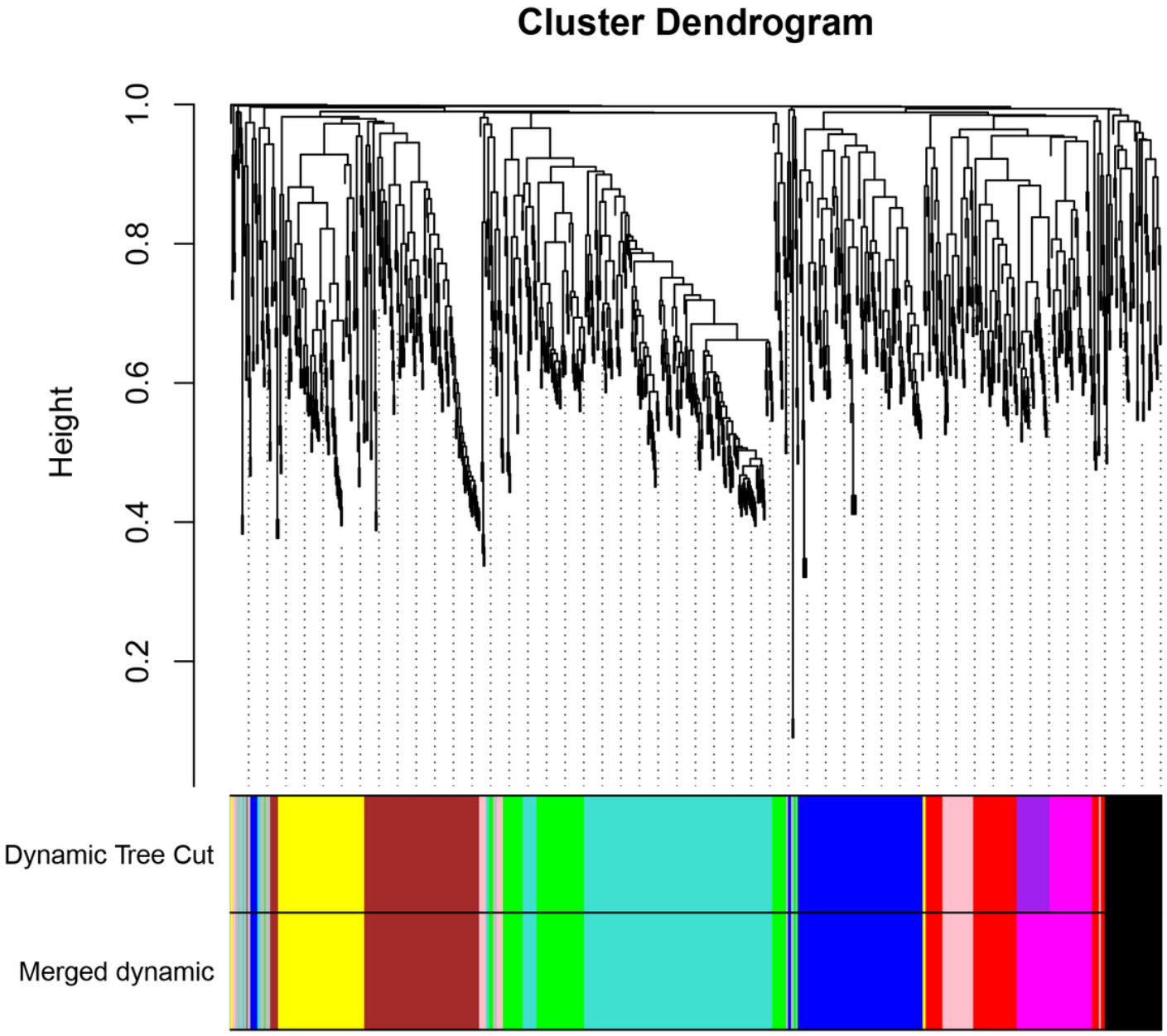

Figure 1

Average linkage hierarchical clustering by the Topological Overlap Matrix and adjacency-based dissimilarity used to understand modules in *C.sativus*. the original modules along with the as- marked merged modules are proved with various colors in the below the dendrogram.

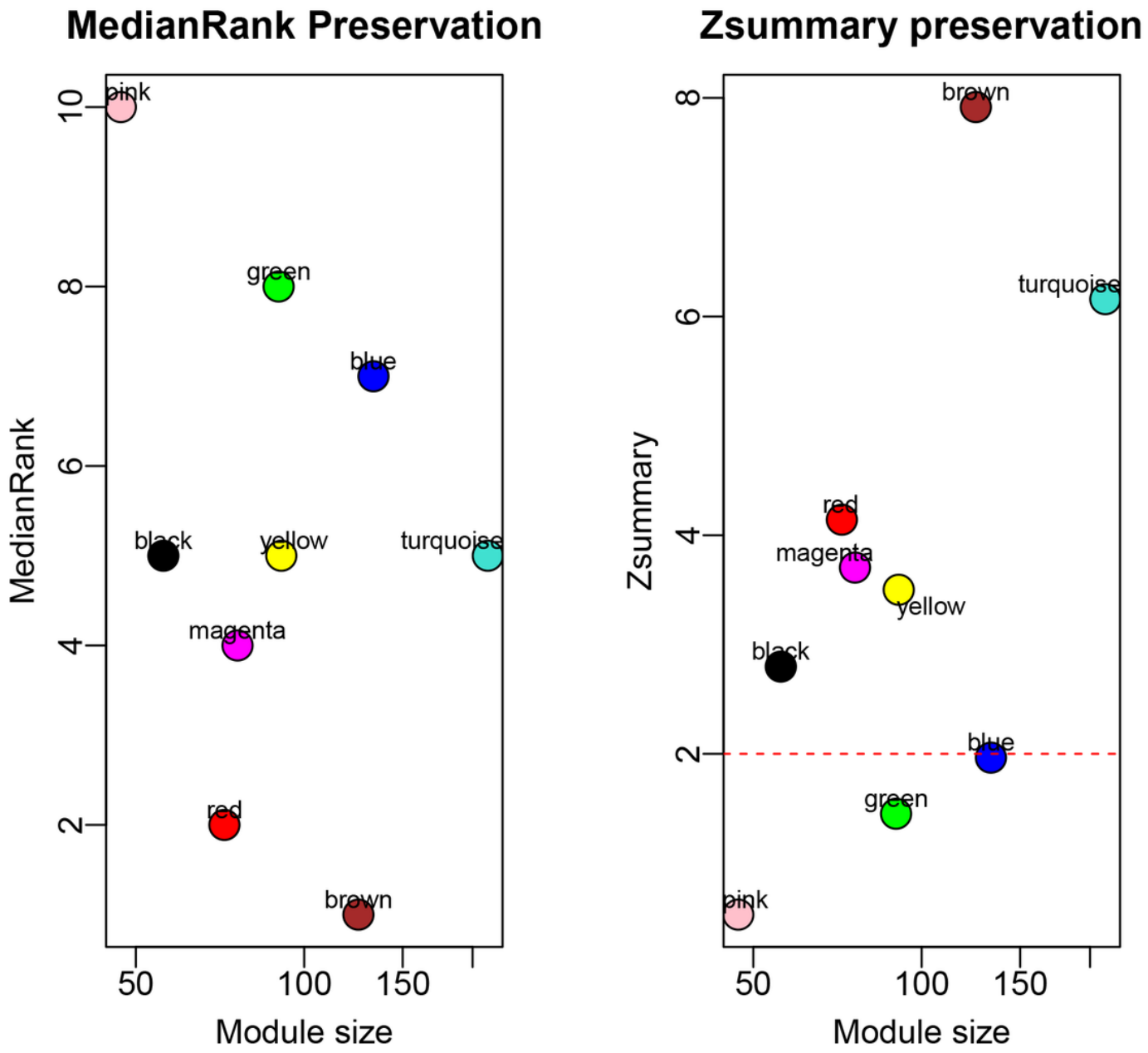

Figure 2

The medianRank preservation statistics (y-axis) of the modules(A). Each bubble shows a module, labeled by names and colors. High preservation indicates by the minimum number of the y-axis (B) the Zsummary preservation measurements (y-axis) of the modules. Each module is labeled different color in panel. The red line indicates the threshold. A Zsummary value below 2 represents a non-preserved module, and a value over 5 provides moderate evidence of module preservation.

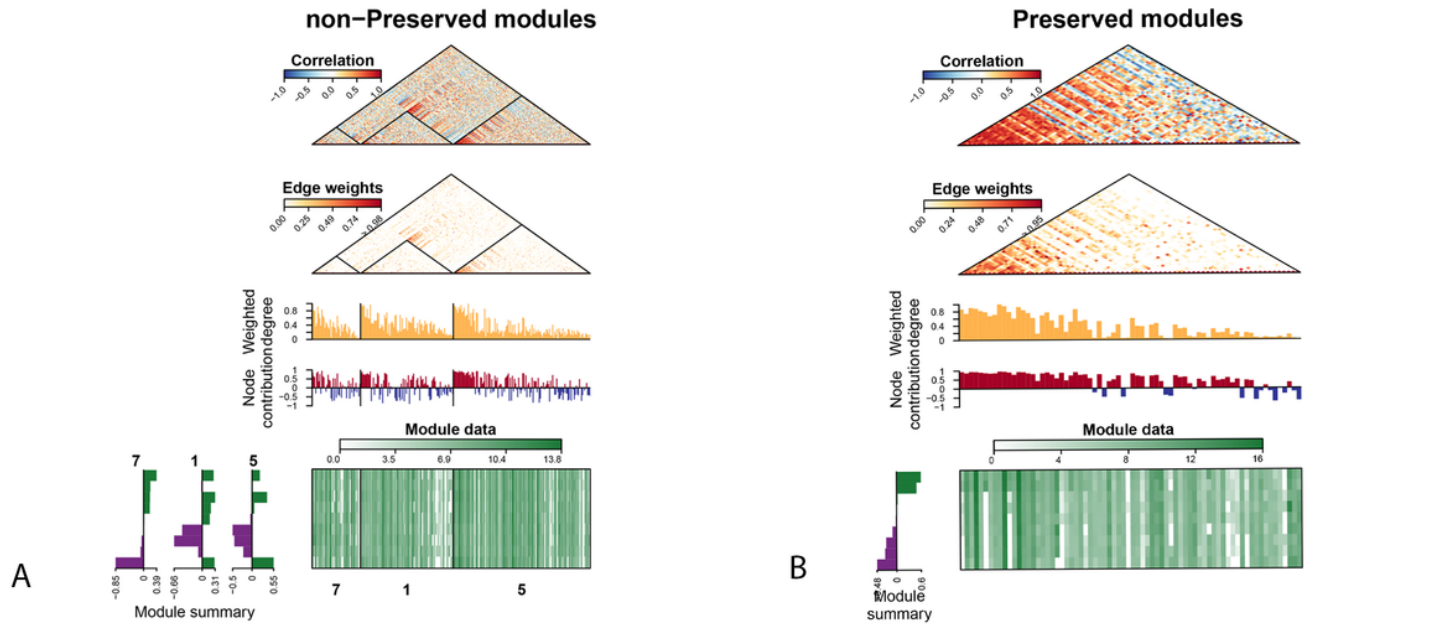

Figure 3

(A) Topological network in the flowering group (discovery). From top to bottom: the correlation structure heatmap, the interaction network edge weights heatmap, normalized weighted degree (computed within the module and normalized by the maximum value), and node contribution (Pearson correlation among gene expression profile). Genes are ordered by dropping the order of weighted degrees.

(B) Topological network in flowering and non-flowering samples. Genes are ordered as in the flowering samples.

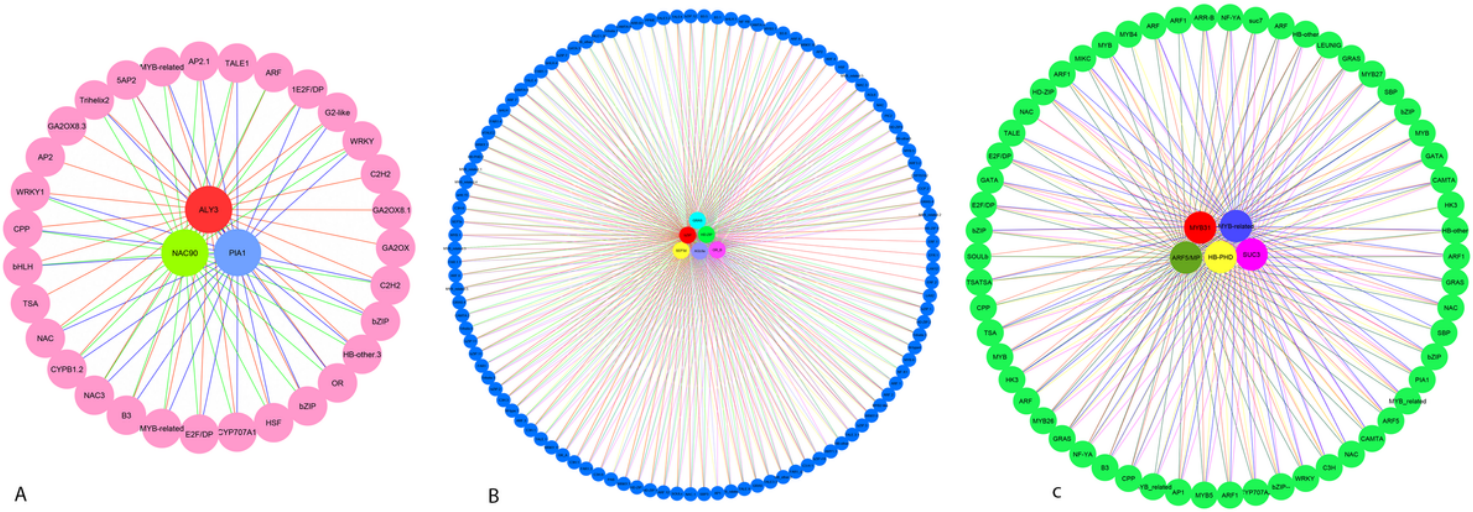

Figure 4

The non-preserved modules and hub genes related to the flowering process in *C.sativus*. Each node shows a gene, and each edge is the co-expression relationship with a specific color. The pink module(A). The blue module(B). The green module(C).

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- supplementaryinformation.docx
- TableS3.xlsx
- TableS4.xlsx