BMC Genomics

RESEARCH

Open Access

LoniComp: a platform for gene function comparison and analysis between *Lonicera japonica* and *Lonicera macranthoides*

Jingjie Zhang[1], Bingbing Pan[1], Jiangxin Yang[1], Qi Pan[1], Panpan Zhu[1], Jiaotong Yang[1], Mian Zhang[1] and Qiaoqiao Xiao[1*]

Abstract

Lonicera japonica and *L. macranthoides* are popular medicinal plants used for treating various diseases. Recently, new chromosome level genomes of *Lonicera* have provided a huge resource for understanding gene function. Although LjaFGD was created for analyzing *L. japonica* gene functions, it is now outdated due to updated genomes and more transcriptome data. Utilizing new chromosome-level genomic and transcriptomic data, we developed co-expression networks of *L. japonica* and *L. macranthoides*. Gene annotations were performed by comparing sequences with NR, TAIR, Swissprot, and trEMBL databases. GO and KEGG annotations were predicted using InterProScan and GhostKOALA software, while gene families were identified with iTAK, HMMER, and InParanoid. To fully leverage the utilization value of public resources and data, we developed LoniComp (www.gzybioinformatics.cn/LoniComp) as a newer and information-rich alternative, a platform for gene function comparison and analysis by integrating genomic, transcriptomic data and processed functional annotations. It features tools like BLAST, Extract Sequence, Enrichment, Heatmap, DEG, and JBrowse2. We demonstrated its use with examples like *LjFT* and *LjMYB12*. It offers superior genomic data, transcriptomic resources, and analysis tools compared to LjaFGD, aiding researchers in gene function studies and comparison.

Keywords *Lonicera japonica*, *Lonicera macranthoides*, Genome, Transcriptome, Co-expression network, Gene family, Functional annotation, Functional Analysis, Platform

Introduction

Lonicera Linn. is the largest genus in the Caprifoliaceae family, with around 200 species primarily found in North Africa, North America, Europe and Asia. Many species of the genus are valued for their ornamental beauty, medicinal properties, and health benefits. In countries like Australia, the US and China, *Lonicera* is often cultivated vegetatively as a hardy ornamental [1]. Research on *Lonicera* has gained worldwide popularity [2–5]. For instance, in the United States, researchers inferred the phylogeny of *Lonicera* using restriction site-associated DNA sequencing (RADSeq) [2]. In Japan, Masaaki et al. explored the therapeutic effect of *L. japonica* flower bud extract (LJFE) on digestive tract infections induced by *Citrobacter rodentium*, a pathogen that mimics human intestinal infections, in a mouse model [3]. In addition, about 98 species of genus *Lonicera* are distributed in various provinces of China, with the largest number of species in southwest China. *L. japonica* and *L. macranthoides* are commonly used medicinal plant of genus *Lonicera* [6]. From the 1963 edition of the Chinese Pharmacopoeia, *L. japonica*'s dried buds and flowers were officially recorded as "Jinyinhua" [7]. *L. macranthoides*

*Correspondence:
Qiaoqiao Xiao
xqqiao2021@163.com
[1] Guizhou University of Traditional Chinese Medicine, Guizhou 550025, China

was identified as the primary source of 'Shanyinhua' in the 2005 Chinese Pharmacopoeia, separate from *L. japonica* [8].

L. japonica is a key herb in traditional Chinese medicine, extensively used and cultivated across China [9]. Shen Nong Ben Cao Jing states that *L. japonica* has been used since ancient times for its anti-inflammatory, antipyretic, and antibacterial effects. Recent research indicates that phenolic acids, flavonoids, iridoids and saponins are the primary active components in *L. japonica* [10], which collectively offer anti-inflammatory, bacteriostatic and antimicrobial benefits [11], antiviral properties, antioxidant potentiality, and even anti-tumor activity [12]. Consequently, *L. japonica* is mainly used in daily diets and clinical prescriptions for the prevention and treatment of inflammation and bacterial or viral infections. In response to market demands, new varieties are continually developed. The *L. japonica* (sijihua) is a high-yield, long-blooming, and stress-resistant variety created through hybridizing traditional *L. japonica* types [12]. Primarily cultivated in Pingyi County, Shandong Province, *L. japonica* (sijihua) is identified as the hardiest honeysuckle variety through regional studies [12]. *L. macranthoides*, native to southwestern China, is used in traditional Chinese medicine, with its dried flower buds treating fever, inflammation, and infections [13]. *L. macranthoides* primarily contains phenolic acids, flavonoids, and carotenoids, notably chlorogenic acid [14].

Advancements in sequencing technology have generated extensive data on *L. japonica*, *L. japonica* (sijihua), and *L. macranthoides*. In 2020, Pu et al. created a high-quality genome sequence of *L. japonica* at the chromosomal level with a genome size of 843.2 Mb and nine pseudochromosomes [15]. Using this, we developed LjaFGD, a platform for analyzing gene functions in *L. japonica* [16]. Huang et al. recently assembled a chromosome-level genome of *L. japonica* (sijihua), sized at 886.04 Mb with a scaffold N50 of 79.5 Mb, this genome is a valuable resource for exploring the genetic basis of its high stress resistance, aiding in the study of genetic diversity and the breeding improvement of *L. japonica* [12]. Yin et al. reported that the chromosomal-level genome of *L. macranthoides* consists of nine pseudochromosomes, and evolutionary analysis indicates that *L. japonica* and *L. macranthoides* diverged 1.30 to 2.27 million years ago [17]. The refinement of the three genomes of the two species of honeysuckle provides a valuable resource for studying the biosynthesis of their active components. For LjaFGD contains only one species of *L. japonica*, with the update of genome and transcriptome data, we built a new gene function comparison and analysis platform named LoniComp based on the published chromosome level genome of *L. japonica*, *L. japonica* (sijihua) and *L. macranthoides*. It provides reference for domestic and foreign users to study the comparison and analysis of gene function.

Materials and methods

Data resource

Genomic data and genome annotation of *L. japonica* was obtained from the National Genomics Data Center (NGDC, https://ngdc.cncb.ac.cn/) (Accession number: GWHAAZE00000000), *L. japonica* (sijhua) Genomic data was obtained from the China National Gene Bank database (CNGBdb, https://db.cngb.org/) (Accession number: SAMN24662184), and the genome annotation file was obtained from the Figshare platform (https://figshare.com/). Transcriptomic data for *L. japonica* and *L. japonica* (sijihua) were retrieved from the NCBI Sequence Read Archive (SRA, http://www.ncbi.nlm.nih.gov/sra), the CNGBdb, and the NGDC. Data for *L. macranthoides*, comprising reads of PacBio and Illumina RNA sequence, were sourced from the SRA database under Accession Number PRJNA800599. Annotated public protein sequences were gathered from non-redundant protein sequence database (NR, https://ftp.ncbi.nlm.nih.gov/blast/db/FASTA), the *Arabidopsis* information resource (TAIR, https://www.arabidopsis.org/) [18], and Universal Protein (UniProt, https://www.uniprot.org/). InterProScan software[19] was used to predict Gene Ontology (GO, https://www.ebi.ac.uk/QuickGO/) terms and The Protein Families Database (Pfam, http://pfam.xfam.org/) domain (Table S1).

Genome assembly, structural annotation

Using public third-generation sequencing data generated by the PacBio platform (SRA accession number: SRP357305), a genome assembly of *L. macranthoides* was developed using Canu softwore [20]. We firstly annotated the gene structure by aligning RNA-seq reads to the genome with hisat2 (v2.1.0) [21], reconstructing transcripts using stringtie (v2.1.4) [22], and predicting coding regions with TransDecoder (v5.1.0) to identify coding genes. Secondly, the protein sequences of *L. japonica* and *L. japonica* (sijihua) were mapped to the genome using the miniprot [23], and predict coding region. Following that, Augustus [24] was employed for predicting gene structure de novo. Lastly, EVidenceModeler [25] software was used to integrate the genetic annotation results predicted by various software.

Functional annotation

For gene function annotation, we utilized diamond blastp software to compare protein sequences with public databases like NR, TAIR, Swissprot, and translated EMBL nucleotide sequence data library (TrEMBL) database

(https://www.uniprot.org/downloads). InterProScan software [19] was used to predict GO terms and Pfam domain. GO annotations were obtained from the GO Consortium [26], the Kyoto Encyclopedia of Genes and Genomes (KEGG, https://www.kegg.jp/) annotation was predicted by GhostKOALA (https://www.kegg.jp/ghostkoala/) website [27] (Table S1).

Classification and identification of gene families

Initially, the hidden Markov model with ubiquitin and ubiquitin-like conjugation database (iUUCD 2.0, http://uucd.biocuckoo.org/) and its recommended threshold effectively identifies the ubiquitin family [28]. OrthoFinder [29] was employed with default settings to predict orthologs between *Arabidopsis* and *Lonicera* species. Using its *Arabidopsis* homolog, we identified the transporter family (TPs) and Carbohydrate-Active EnZymes (CAZy, http://www.cazy.org/) in three *Lonicera* species [30]. The iTAK software (http://itak.feilab.net/cgi-bin/itak/index.cgi) [31] was used to identify transcription factors/regulators (TFs/TRs) and protein kinases (PKs) in *Lonicera* species. Protein sequences were subjected to KEGG pathway annotation using GhostKOALA. Additionally, Cytochrome P450 (CYP450, http://drnelson.utmem.edu/CytochromeP450.html) genes were identified using KEGG annotations (Table S1).

The construction of co-expression network

The transcriptome data was first aligned to the reference genome with Hisat2 [21], and TPM for each sample was calculated using StringTie [22]. The Pearson Correlation Coefficient (PCC) algorithm is a method used to calculate the correlation coefficient between two genes. It determines the correlation between two continuous variables by measuring the linear relationship between them. In the construction of the co-expression network, the PCC algorithm was employed to calculate the correlation between the expressions of every pair of genes. The Mutual Rank (MR) algorithm is a method to further increase the confidence of common representation relationships. It constructs the co-expression network by calculating the geometric mean of the ranking of gene A in gene B and the ranking of gene B in gene A. The MR algorithm was used to rank the resulting gene correlations, applying the following formula:

$$PCC = \frac{\sum_{i=1}^{n}(X_i - \overline{X})(Y_i - \overline{Y})}{\sqrt{\sum_{i=1}^{n}(X_i - \overline{X})^2}\sqrt{\sum_{i=1}^{n}(Y_i - \overline{Y})^2}}$$

$$MR(AB) = \sqrt{Rank(AB) \times Rank(BA)}$$

In the provided formulas, 'n' denotes the total sample count in the RNA-seq dataset, and 'X_i' stand for the TPM value of gene X in the i sample, and 'Y_i' stand for the TPM value for gene Y in the i sample. $\overline{X}$ represents the average value of gene X in all samples, and $\overline{Y}$ represents the average value of gene Y in all samples. For the calculation of MR, we first sort the PCC values between gene B and other genes from largest to smallest. The ranking of A in this sequence is Rank (AB), and conversely, it is Rank (BA).

Protein–protein interaction (PPI) network construction

The protein–protein interaction networks for *L. japonica*, *L. japonica* (sijihua), and *L. macranthoides* were created based on their homologous relationship with *Arabidopsis* and its PPI network.

Differentially expressed genes (DEGs) in different transcriptome

For the analysis of DEGs, we used the Student's t-test (significance threshold: $P < 0.05$) combined with a $\log_2$ fold change threshold of > 1 for upregulation and < -1 for downregulation to identify significant differences in gene expression between experimental and control groups, as these thresholds align with cutoffs used in previous studies [32–35]. Genes that met both criteria were considered significantly differentially expressed.

Construction of LoniComp

The platform was developed with the LAMP stack (Linux, Apache, MySQL, PHP). A MySQL database was populated with data on gene structure, function, co-expression, PPI networks, and gene family classification. Dynamic web pages for data presentation and analysis were created using HTML, PHP, JavaScript, and CSS.

Toolkit for gene function analysis in Platform

For the visualization of genome, transcriptome data, we employed JBrowse2 software [36]. The identification of similar sequences was done using the blast tools developed by Deng et al. [37]. We used the ClusterProfiler package in R language to integrate Gene Set Enrichment Analysis [38]. The background contained four kinds of gene sets: KEGG pathway, GO, TF and PK. Among these, *L. japonica*, *L. japonica* (sijihua), and *L. macranthoides* contained 1511, 1578, and 1557 background gene sets, respectively (Table S2). Additionally, the DEG tool is employed to facilitate the comparative analysis of various samples within the same transcriptome, with the results ultimately depicted in the form of a volcano plot. Furthermore, we developed an extract sequence tool utilizing a Perl script and implemented a heatmap using the pheatmap package in the R programming language.

These enhancements expanded the platform's capabilities and improved data visualization and analysis.

Results

Gene structure and functional annotation

L. japonica genome data from NGDC comprised 33,939 genes, 33,961 transcripts, and 33,961 proteins. *L. japonica* (sijihua) genome data, sourced from GenBank, comprises 39,320 genes, 39,320 transcripts, and 39,320 proteins. *L. macranthoides* was structurally annotated with 44,841 genes, 44,841 transcripts, and 44,841 proteins (Fig. 1A). To ensure accurate annotation, we compared these resources with renowned protein sequence databases like NR, TAIR, Swissprot, and trEMBL. The annotated gene counts for *L. japonica* were 31,784, 28,333, 24,634, and 31,815, respectively. For *L. japonica* (sijihua), annotated gene counts were 36,432, 31,222, 26,814, 36,450, respectively. For *L. macranthoides*, the counts were 40,904, 37,093, 31,923, and 40,938, respectively. InterProScan annotated 17,574, 17,399, and 19,390 genes by GO in *L. japonica*, *L. japonica* (sijihua), and *L. macranthoides*, respectively, and predicted Pfam for 25,344, 29,481, and 31,926 genes. We used GhostKOALA online tools to map KEGG annotations onto 5,975, 6,106, and 8,676 genes in *L. japonica*, *L. japonica* (sijihua), and *L. macranthoides* (Fig. 1A).

In terms of the proportion of genes annotated, among the three genomes, *L. japonica* had the highest proportion of genes annotated in trEMBL, NR, TAIR10, Swissprot and GO databases, which were 93.74%, 93.65%, 83.48%, 72.58% and 51.78%, respectively. L. *japonica* (sijihua) accounted for the largest proportion of genes annotated in the Pfam database with 74.98%. *L. macranthoides* was responsible for the largest proportion of genes annotated in the KEGG database (19.35%) (Table S3).

Gene family classification

Utilizing the iTAK software, we identified 7,313 transcription factors/regulators and 4,601 protein kinases across three *Lonicera* species. Additionally, employing the iUUCD v2.0 tool, we predicted a total of 3,861 ubiquitin protease genes within these species. Through comparative analysis with the TransPortDB and CAZy databases, we discerned 2,683 transport family genes and 2,857 CAZy family genes in the three *Lonicera* species. Furthermore, our analysis predicted 243 genes belonging to the Cytochrome P450 family, as illustrated in Fig. 1B.

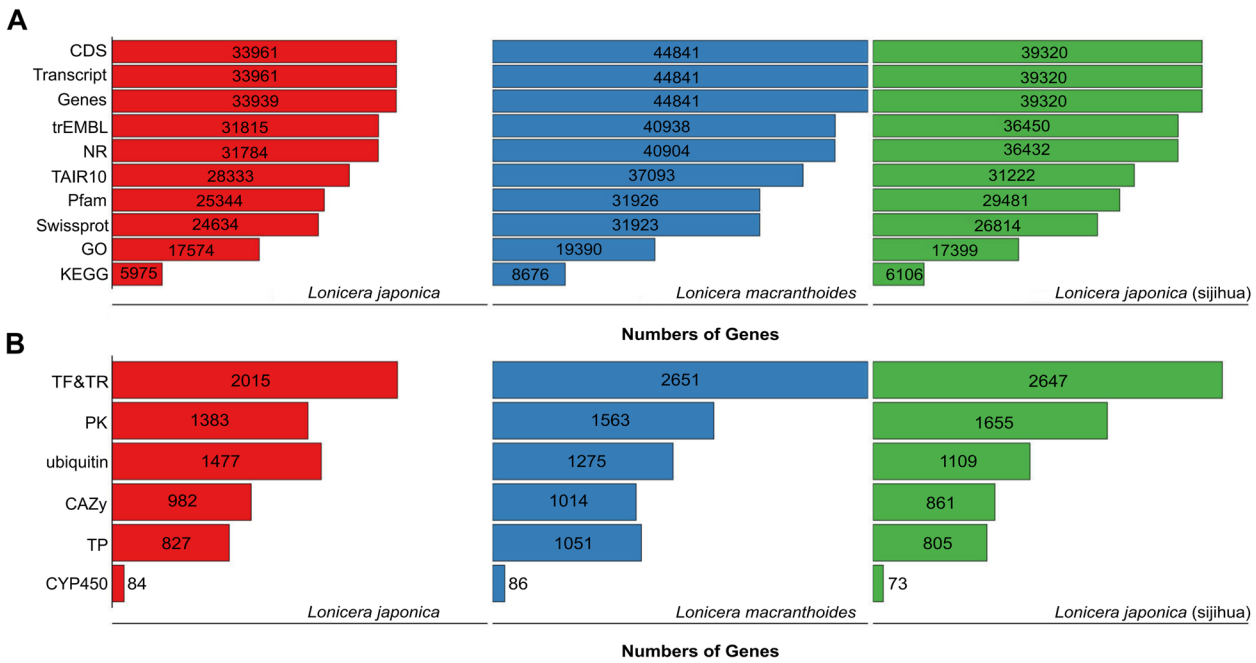

Fig. 1 Overview of the *L. japonica*, *L. japonica* (sijihua) and *L. macranthoides* functional genomics database. **A** Gene function annotation information. From top to bottom, the full names are Coding sequence (CDS), Transcript Variant (Transcript), Genes, translated EMBL nucleotide sequence data library (trEMBL), Non-Redundant Protein Sequence Database (NR), the Arabidopsis information resource (TAIR10), The Pfam protein families database (Pfam), Swiss-Prot, Gene Ontology (GO), Kyoto Encyclopedia of Genes and Genomes (KEGG). **B** Gene family classification. From top to bottom, the full names are respectively Transcription Factors/Regulators (TF&TR), Protein Kinases (PK), Ubiquitination (ubiquitin), Carbohydrate-Active EnZymes (CAZy), Transporter Family (TP), Cytochrome P450 (CYP450)

Co-expression network

RNA samples from various plant tissues, developmental stages, and stress conditions were mapped to the reference genome (Table S4). To ensure data reliability, only RNA-seq samples with over 80% mapping rate were used for TPM calculation, which then helped compute the PCC. Most gene pairs showed weak correlation in expression patterns (Fig. 2A). The MR method, using PCC ranking values, was used to identify closely linked gene pairs within networks.

To ensure network reliability, we assessed AUC values for PCCs (0.6, 0.7, 0.8, 0.9) using a binary classifier as described in supplementary information. No significant AUC differences were found among the PCC networks. To include more genes, a PCC threshold of over 0.6 was chosen (Fig. S1). The AUC values were analyzed across various MR thresholds, ensuring the PCC exceeded 0.6. This resulted in a positive co-expression network threshold of MR ≤ 30 (Fig. S2). For the negative co-expression network, thresholds were set at PCC < −0.5 and MR ≤ 30. The resulting positive co-expression network for *L. japonica*, *L. japonica* (sijihua), and *L. macranthoides* comprised, 369,804, 404,643, and 270,347 co-expression gene pairs, respectively (Fig. 2B). The resulting negative co-expression network for *L. japonica*, *L. japonica* (sijihua), and *L. macranthoides* comprised, 162,021, 163,730, and 179,021 co-expression gene pairs, respectively (Fig. 2C). The positive co-expression network comprised 31,530, 35,134, and 41,676 genes (Fig. 2B), while the negative co-expression network consisted of 16,161, 16,805, and 27,260 genes, respectively (Fig. 2C).

Protein–protein interaction (PPI) network

We predicted orthologous genes between *Arabidopsis* and *Lonicera* species, mapping *Arabidopsis*'s PPI network to *L. japonica* and *L. macranthoides*. This identified 30,416, 36,388, and 55,112 PPI gene pairs involving 6,784, 7,500, and 9,049 genes, respectively (Fig. 2D).

Network display with DEGs in LoniComp

To integrate gene co-expression and PPI networks with expression data, DEGs were identified from 21 transcriptome datasets, resulting in 524 DEG groups (Table S5). We then created a combined display contain network and DEGs, marking up-regulated DEGs in red and down-regulated DEGs in blue.

Construction of LoniComp

A platform called LoniComp has been developed for comparing and analyzing functional genomics in *L.*

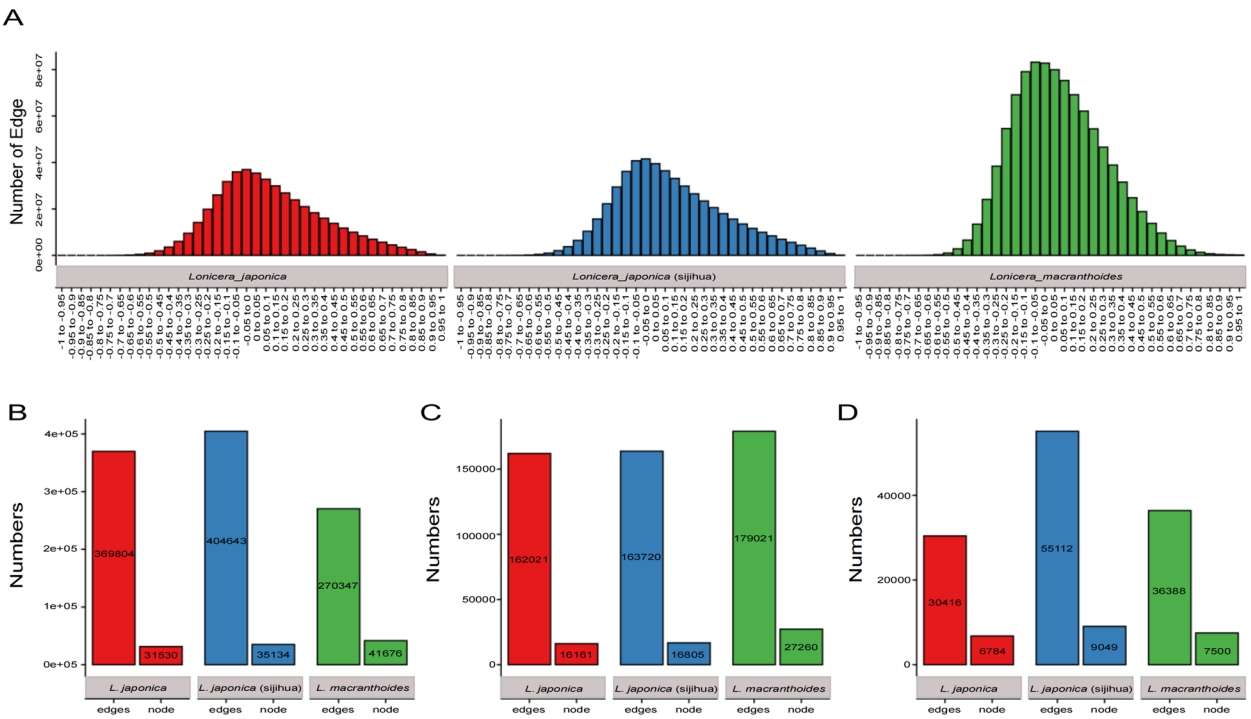

Fig. 2 Statistics of LoniComp's PCC distribution, co-expression network and PPI network data. **A** The relationship between the Pearson correlation coefficient (PCC) and the number of edges in the co-expression network. **B** Statistical analysis of nodes and edges in the positive co-expression network. **C** Statistical analysis of nodes and edges in the negative co-expression network. **D** Statistical analysis of nodes and edges in the protein–protein interaction (PPI) network

japonica, *L. japonica* (sijihua), and *L. macranthoides*, utilizing functional annotation, gene family classification, co-expression, and PPI network. There are eight sections within LoniComp, including Home, Genomes, Network, Search, Tools, Pathway, Download and Manual. The genome section consists of three genomes from *L. japonica*, *L. japonica* (sijihua), and *L. macranthoides*. Each link provides descriptions related to the species, genome, and hyperlinks to gene families. Network section includes PPI and co-expression network. To aid users in gene function search and analysis, LoniComp includes seven tools: Search, Blast, Extract Sequence, Enrichment, DEG, Heatmap, and JBrowse2. Users can search for specific genes by entering keywords or the exact accession number on the search page. The Blast tool allows for finding similar nucleic acid or protein sequences in *L. japonica*, *L. japonica* (sijihua), and *L. macranthoides*. The Extract Sequence tool extracts sequences by gene accession number and location. Enrichment performs gene set enrichment analysis, while Heatmap analysis visualizes gene expression data for the candidate genes. The DEG tool swiftly identifies differentially expressed genes in *Lonicera* species. JBrowse is integrated into LoniComp for visualizing genomic and transcriptome features. A download and manual section offers user guidance and download details (Fig. 3).

Case study

Structure and functional analysis of the *LjFT* gene

The *FT* gene plays a crucial role in modulating flowering time through its interactions with various regulatory genes. It is capable of forming complexes with genes such as CONSTANS (CO) and SUPPRESSOR OF OVEREXPRESSION OF CONSTANS 1 (SOC1), thereby contributing to the initiation of the flowering process [37]. The GWHGAAZE001702 gene in *L. japonica*, identified as an *FT* member on chromosome 1 (58,252,707–58,257,150 bp), includes transcript and protein sequences. The phosphatidylethanolamine-binding protein domain is located between 60 and 159 bp in sequence PF01161 (Fig. 4A-4D).

We analyzed the positive co-expressed genes of *LjFT*. In samples SRP173429, SRP363000, and SRP417164, these genes were highly expressed in flower buds, especially in the calyx and petal, but their expression significantly decreased after flowering (Fig. 4E, Fig. 5A-5C). Furthermore, JBrowse2 read mapping revealed higher expression levels in bud samples compared to post-flowering samples (Fig. 5D-5F). Network analysis revealed that 7

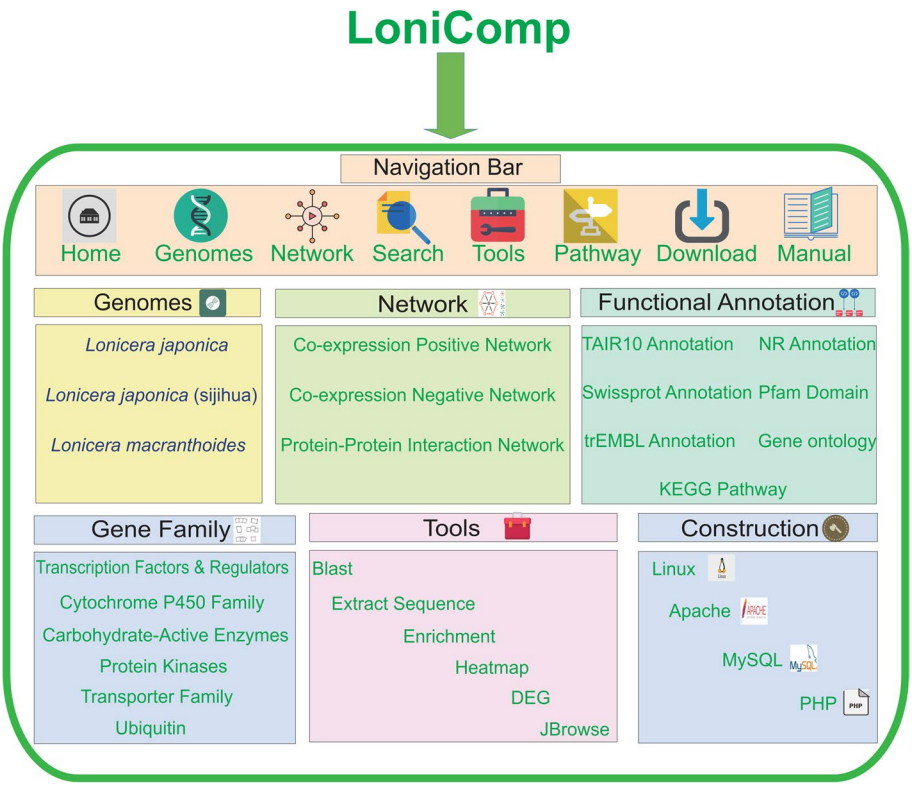

Fig. 3 Organizational chart of the LoniComp

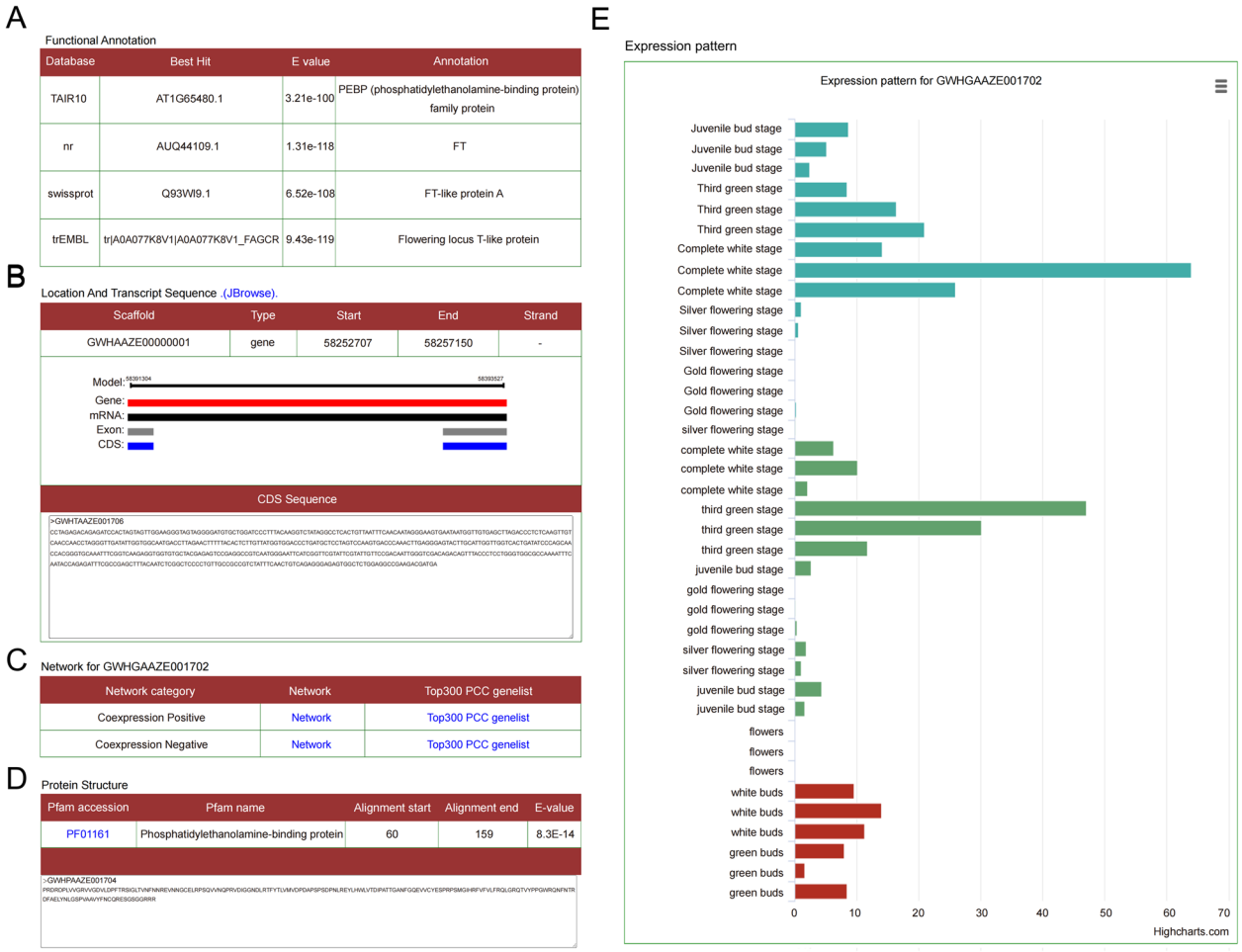

Fig. 4 Gene detail page of *LjFT* gene. **A** Gene functional annotation. **B** Location and transcript sequences. **C** Network of *LjFT*. **D** Protein structure and sequence. **E** Expression pattern in different samples

genes were positively co-expressed with *LjFT* (Fig. 6A). Among the genes that are positively co-expressed with *LjFT*, Floral-binding protein 9 (Fragment) (GWHGAAZE030453) and ABC transporter B family member 8 (GWHGAAZE025707) positively regulate flowering in many species [39–43]. In addition, many genes in the co-expression network are significantly up-regulated in juvenile bud (Fig. 6B-6D). Therefore, our analysis shows that the *LjFT* gene plays an important role in the regulation of flowering, a conclusion that is supported by numerous relevant studies [39–43].

Characteristic and functional analysis of *LjMYB12* gene

Previous studies have shown that *LjMYB12* (GWHGAAZE009942) can promote the biosynthesis of flavonoids [44], Genetic details are in Fig. S3, with its transcriptional sequence on chromosome 2 from 120,315,361 to 120,318,481 bp (Fig. S3B). Network links were established (Fig. S3C). The MyB-like DNA binding domain, identified as PF00249, is located at 67–112 bp or 14–61 bp of the protein-coding sequence (Fig. S3D). This gene family annotation belongs to the MYB gene family (Fig. S3E). Through the analysis of expression profile, we found that the expression level of this gene was higher in flower buds than in flowers (Fig. S3F). And the results of the transcriptomic data on our website show that the transcription level of *LjMYB12* is indeed proportional to the total flavonoid content of honeysuckle during development. Furthermore, studies have shown that the expression pattern of luteoloside is higher before flowering than after flowering [45], and the expression pattern of *LjMYB12* is also consistent with that of luteoloside. In addition, KEGG enrichment analysis of *LjMYB12* and its co-expressed genes showed that these genes are associated with Naphthalene degradation, Retinol metabolism, Chloroalkane and chloroalkene degradation, AMPK signaling pathway, HIF-1 signaling pathway, Fatty acid degradation, Tyrosine metabolism, Methane

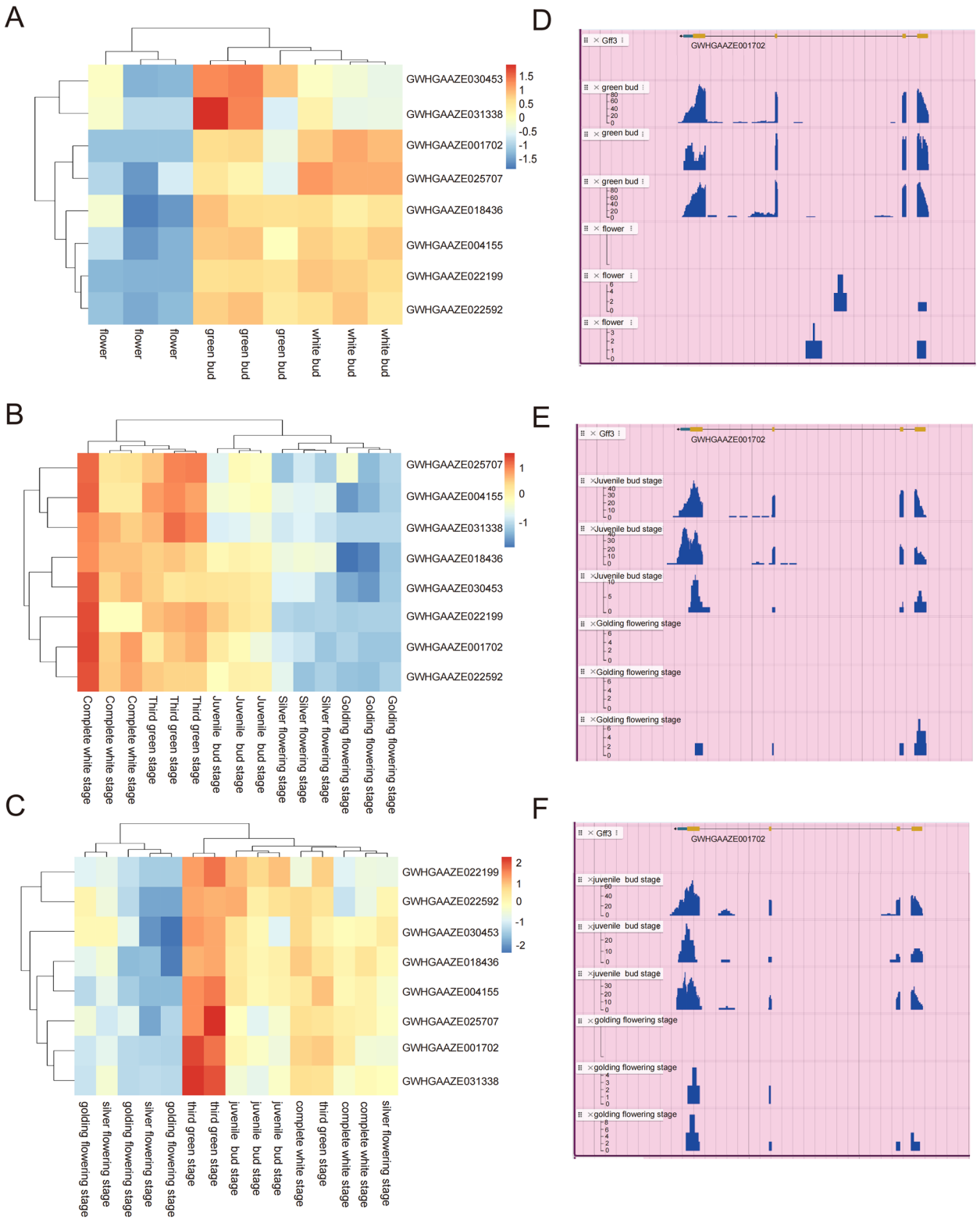

Fig. 5 Heatmaps of the *LjFT* gene and its positive co-expression genes in three sets of samples and visualization of the *LjFT* gene using JBrower2. **A** Heatmap of *LjFT* and its positive co-expressed gene in SRP417164. **B** Heatmap of *LjFT* and its positive co-expressed gene in SRP173429. **C** Heatmap of *LjFT* and its positive co-expressed gene in SRP363000. **D** The expression of *LjFT* gene in SRP417164 was visualized by JBrowse2. **E** The expression of *LjFT* gene in SRP173429 was visualized by JBrowse2. **F** The expression of *LjFT* gene in SRP363000 was visualized by JBrowse2

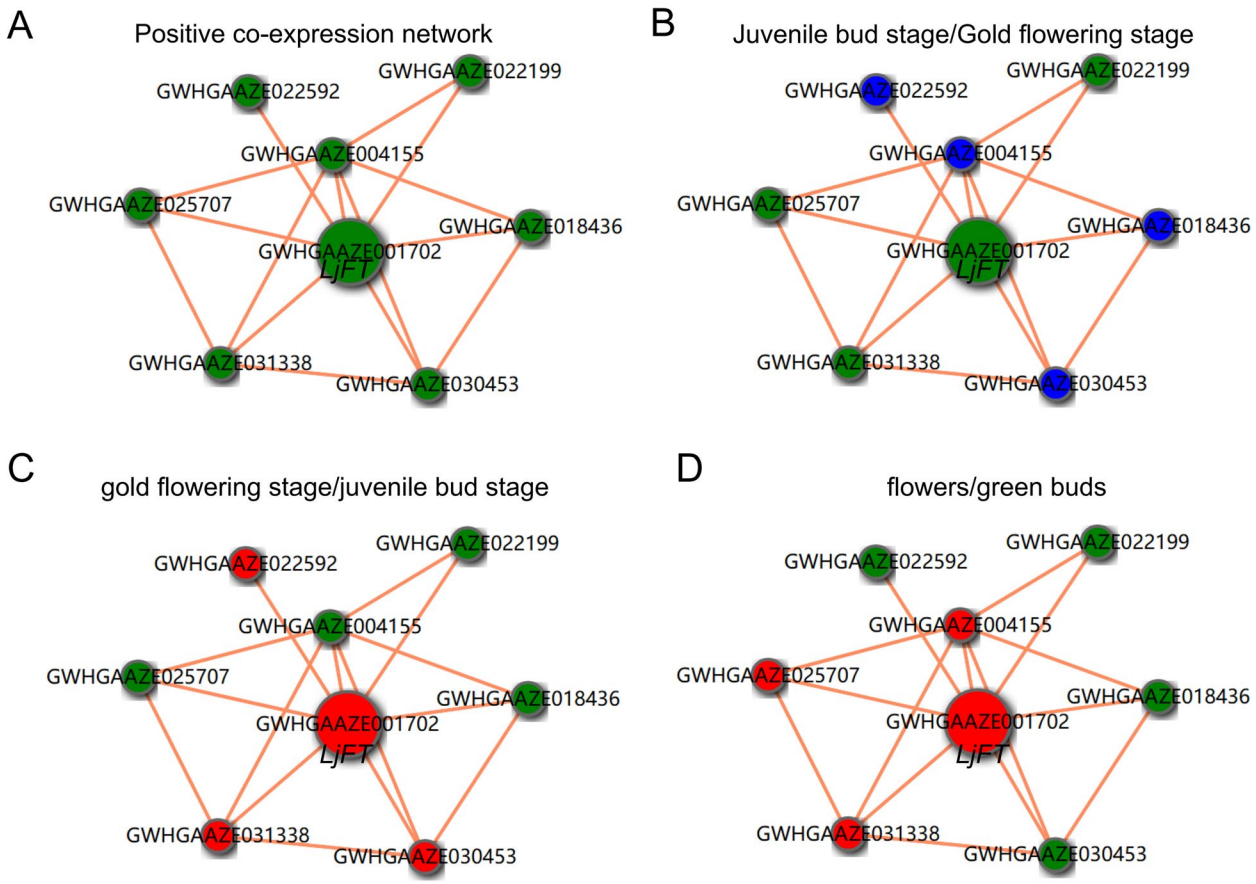

Fig. 6 Co-expression network of the *LjFT* gene. Red nodes represent up-regulated genes, blue nodes represent down-regulated genes, orange links represent positive co-expression relationships and green represent no significant change genes. **A** Positive co-expression network of the *LjFT* gene. **B** Comparison of co-expression networks of *LjFT* genes between Juvenile bud stage and Gold flowering stage in SRP173429. **C** Comparison of co-expression networks of *LjFT* genes between gold flowering stage and juvenile bud stage in SRP363000. **D** Comparison of co-expression networks of *LjFT* genes between flowers and green buds and juvenile bud stage in SRP417164

metabolism, Fructose and mannose metabolism, Drug metabolism-cytochrome P450, Metabolism of xenobiotics by cytochrome is associated with the P450 pathway (Fig. 7, Fig. S4).

Discussion

L. japonica, *L. japonica* (sijihua), and *L. macranthoides* are widely used in traditional Chinese medicine, requiring substantial quantities. The genomes of *L. japonica*, *L. japonica* (sijihua) and *L. macranthoides* have been sequenced [12, 15, 17], which provides available resources for the study of biochemistry, genetics, molecular biology, and molecular evolution. Integrating their omics data is crucial for advancing scientific research. We developed LoniComp, a platform that integrates genomes, transcriptome data, annotations, and analytical tools for functional genomics comparison and analysis between *L. japonica* and *L. macranthoides*.

Numerous platforms have been developed to collect and analyze gene function information for various plant species, primarily focusing on crops, fruits, and vegetables. However, our platform is about medicinal plants *L. japonica*, *L. japonica* (sijihua) and *L. macranthoides*, which can provide reference for the subsequent construction of gene function platforms in other medicinal plants. Currently, many gene function platforms are outdated, and some websites are unusable. Since completing LjaFGD in 2021, we've focused on *Lonicera* research, collecting the latest genomic and transcriptome data. As a newer and more informative alternative to LjaFGD, LoniComp represents a significant advancement over its predecessor, addressing previous limitations through the integration of more comprehensive genomic datasets and broader transcriptome resources. By prioritizing access to up-to-date data and advanced functionality, LoniComp serves as a critical resource for driving innovation in genomic and transcriptomic research.

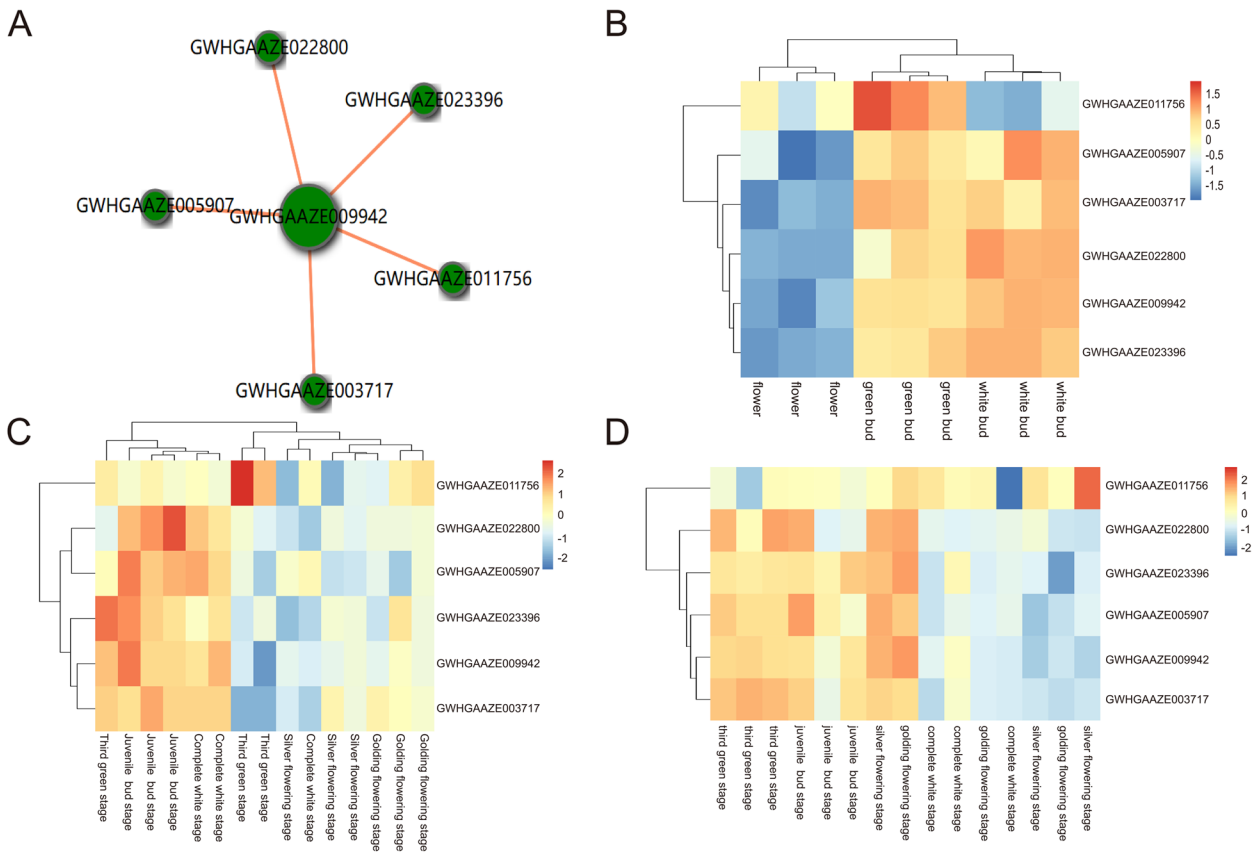

Fig. 7 Co-expression network of LjMYB12 and its associated heatmaps. **A** Positive co-expression network of the LjMYB12 gene. **B** Heatmap of LjMYB12 and its co-expressed gene in SRP417164 (RNA-Seq of L. japonica flowers at different developmental stages). **C** Heatmap of LjMYB12 and its co-expressed gene in SRP173429 (RNA-seq data of L. japonica flowers at different developmental stages). **D** Heatmap of LjMYB12 and its co-expressed gene in SRP363000 (RNA-seq was performed on the different developmental stages of L. japonica flowers)

To ensure the platform's availability, we analyzed two typical cases. As a key flowering integration gene, *FT* gene plays an important role in the process of flowering transition [46]. In *A. thaliana*, *AtFT* gene promotes the transition of reproduction and flowering [47]. The *VcFT* gene in blueberries induces early and continuous flowering by counteracting photoperiod and low-temperature stress [48]. We used the information and tools provided by LoniComp to conduct correlation analysis of *LjFT* gene function and regulation. Several groups of pre- and post-flowering transcriptome heat maps showed that the *LjFT* gene decreased rapidly after flowering (Fig. 5), so the *LjFT* gene was most likely involved in the regulation of flowering in *L. japonica*. This provides some reference for the future study of *LjFT* gene. In addition, previous studies have shown that ectopic expression of *LjMYB12* in *A. thaliana* can increase *PAL* activity and flavonoid content, and promote the transcription of a series of flavonoid biosynthesis genes [44]. That is, the transcription level of *LjMYB12* is directly proportional to the total flavonoid content of honeysuckle during flower development. Studies also shown that the expression pattern of luteoloside before flowering is higher than that after flowering [45]. Our platform analysis found that the expression patterns of *LjMYB12* and its co-expressed genes aligns with that of luteoloside (Fig. 6). These results validate the study of Qi et al., and provide a reference for further studies of *LjMYB12*. Therefore, our analysis can provide a valuable reference for the future utilization of the platform.

In addition, although the species in the current platform are mainly distributed in China, the platform stores genomic, transcriptomic, and other annotation data, providing researchers worldwide with the possibility of conducting cross-species comparative studies. However, while the LoniComp platform provides more genomic data, more transcriptome resources and more analytical tools than LjaFGD, it should be pointed out that the LoniComp platform also has some limitations that require further improvement. Firstly, only three genomes from two species are integrated into the platform. Whenever there is accessible genomic,

transcriptomic, or other data, including those of *Lonicera* species around the world, we will integrate them into the data platform in a timely manner. Secondly, the continuous accumulation of other omics data, such as epigenomics, metabolomics, and proteomics, will provide essential support for the ongoing updates and enrichment of our data platform. Lastly, with the advancement of artificial intelligence technologies, integrating AI into the database will play a crucial role in enhancing the construction and functionality of the platform in the future.

We believe that as ongoing advancements in sequencing technology, reduced costs, and sustained investment, multi-omics data will keep growing. In addition, as data resources continue to accumulate, we believe that the database platform established in this study can provide potential assistance to researchers around the world. Access the site for free at www.gzybioinformatics.cn/LoniComp.

Supplementary Information

The online version contains supplementary material available at https://doi.org/10.1186/s12864-025-11507-y.

Supplementary Material 1.

Supplementary Material 2.

Acknowledgements
We would like to thank all individuals and organizations who contributed to the development of this study.

Authors' contributions
JZ and JTY developed the platform, with JZ also drafting the manuscript. JZ, BP, JXY, and QX revised it, while QX and JY provided financial support. BP, JXY, PZ, MZ contributed to the platform's construction. QP supported server maintenance and database management. All authors approved the final version.

Funding
This work was supported by the National Natural Science Foundation of China (No. 32260140), Guizhou Provincial Science and Technology Planning Project (Qiankehe Jichu MS [2025] No. 152), the University Science and Technology Innovation Team of the Guizhou Provincial Department of Education ([2023]071), the Guizhou Provincial Science and Technology Projects (ZK[2022]505), and the National and Provincial Scientific and Technological Innovation Talent Team of the Guizhou University of Traditional Chinese Medicine (GZYTDHZ[2022]003, GZYTDHZ[2024]002).

Data availability
Genomic data and genome annotation of *L. japonica* was obtained from the National Genomics Data Center (NGDC) (Accession number: GWHAAZE00000000), *L. japonica* (sijhua) Genomic data was obtained from the GenBank database (Accession number: SAMN24662184), and the genome annotation file was obtained from the Figshare platform (https://figshare.com/articles/online_resource/honeysuckle_genome_final_gene_gff3/18092708/6). Transcriptomic data retrieved from the NCBI Sequence Read Archive (SRA), the China National Gene Bank database (CNGBdb, https://db.cngb.org/), and the National Genomics Data Center (NGDC) are list in supplementary table 1.

Declarations

Ethics approval and consent to participate
Not applicable.

Consent for publication
Not applicable.

Competing interests
The authors declare no competing interests.

Received: 9 January 2025 Accepted: 19 March 2025
Published online: 01 April 2025

References

1. Ge L, Xie Q, Jiang Y, Xiao L, Wan H, Zhou B, Wu S, Tian J, Zeng X. Genus Lonicera: New drug discovery from traditional usage to modern chemical and pharmacological research. Phytomedicine : international journal of phytotherapy and phytopharmacology. 2022;96:153889.
2. Srivastav M, Clement WL, Landrein S, Zhang J, Howarth DG, Donoghue MJ. A phylogenomic analysis of Lonicera and its bearing on the evolution of organ fusion. Am J Bot. 2023;110(4):e16143.
3. Minami M, Makino T. Effects of Lonicera japonica Flower Bud Extract on Citrobacter rodentium-Induced Digestive Tract Infection. Medicines (Basel). 2020;7(9):52.
4. Bandyopadhyay S, Abiodun OA, Ogboo BC, Kola-Mustapha AT, Attah EI, Edemhanria L, Kumari A, Jaganathan R, Adelakun NS. Polypharmacology of some medicinal plant metabolites against SARS-CoV-2 and host targets: Molecular dynamics evaluation of NSP9 RNA binding protein. J Biomol Struct Dyn. 2022;40(22):11467–83.
5. Rai A, Kamochi H, Suzuki H, Nakamura M, Takahashi H, Hatada T, Saito K, Yamazaki M. De novo transcriptome assembly and characterization of nine tissues of Lonicera japonica to identify potential candidate genes involved in chlorogenic acid, luteolosides, and secoiridoid biosynthesis pathways. J Nat Med. 2017;71(1):1–15.
6. Fang Z, Li J, Yang R, Fang L, Zhang Y. A Review: The Triterpenoid Saponins and Biological Activities of Lonicera Linn. Mole (Basel, Switzerland). 2020;25(17):3773.
7. Commission CP. Pharmacopoeia of the People's Republic of China. Beijing: People's Medical Publishing House and Chemical Industry Press; 1963.
8. Commission CP. Pharmacopoeia of the People's Republic of China. Beijing: People's Medical Publishing House and Chemical Industry Press; 2005.
9. Zeng Q, Cheng Z, Li L, Yang Y, Peng Y, Zhou X, Zhang D, Hu X, Liu C, Chen X. Quantitative analysis of the quality constituents of Lonicera japonica Thunberg based on Raman spectroscopy. Food Chem. 2024;443:138513.
10. Shang X, Pan H, Li M, Miao X, Ding H. Lonicera japonica Thunb.: ethnopharmacology, phytochemistry and pharmacology of an important traditional Chinese medicine. J Ethnopharmacol. 2011;138(1):1–21.
11. Wang T, Yang B, Guan Q, Chen X, Zhong Z, Huang W, Zhu W, Tian J. Transcriptional regulation of Lonicera japonica Thunb. during flower development as revealed by comprehensive analysis of transcription factors. BMC Plant Biol. 2019;19(1):198.
12. Yu H, Guo K, Lai K, Shah MA, Xu Z, Cui N, Wang H. Chromosome-scale genome assembly of an important medicinal plant honeysuckle. Scientific data. 2022;9(1):226.
13. Lyu H, Liu W, Bai B, Shan Y, Paetz C, Feng X, Chen Y. Prenyleudesmanes and A Hexanorlanostane from the Roots of Lonicera macranthoides. Mole (Basel, Switzerland). 2019;24(23):4276.
14. Lv LL, Li LY, Xiao LQ, Pi JH. Transcriptomic and targeted metabolomic analyses provide insights into the flavonoids biosynthesis in the flowers of Lonicera macranthoides. BMC Biotechnol. 2024;24(1):19.
15. Pu X, Li Z, Tian Y, Gao R, Hao L, Hu Y, He C, Sun W, Xu M, Peters RJ, et al. The honeysuckle genome provides insight into the molecular mechanism of carotenoid metabolism underlying dynamic flower coloration. New Phytol. 2020;227(3):930–43.

16. Xiao Q, Li Z, Qu M, Xu W, Su Z, Yang J. LjaFGD: Lonicera japonica functional genomics database. J Integr Plant Biol. 2021;63(8):1422–36.
17. Yin X, Xiang Y, Huang FQ, Chen Y, Ding H, Du J, Chen X, Wang X, Wei X, Cai YY, et al. Comparative genomics of the medicinal plants Lonicera macranthoides and L. japonica provides insight into genus genome evolution and hederagenin-based saponin biosynthesis. Plant Biotechnol J. 2023;21(11):2209–23.
18. Reiser L, Subramaniam S, Li D, Huala E. Using the Arabidopsis Information Resource (TAIR) to Find Information About Arabidopsis Genes. Curr Protoc Bioinform. 2017;60:1.11.11-11.11.45.
19. Jones P, Binns D, Chang HY, Fraser M, Li W, McAnulla C, McWilliam H, Maslen J, Mitchell A, Nuka G, et al. InterProScan 5: genome-scale protein function classification. Bioinformatics (Oxford, England). 2014;30(9):1236–40.
20. Koren S, Walenz BP, Berlin K, Miller JR, Bergman NH, Phillippy AM. Canu: scalable and accurate long-read assembly via adaptive k-mer weighting and repeat separation. Genome Res. 2017;27(5):722–36.
21. Kim D, Paggi JM, Park C, Bennett C, Salzberg SL. Graph-based genome alignment and genotyping with HISAT2 and HISAT-genotype. Nat Biotechnol. 2019;37(8):907–15.
22. Shumate A, Wong B, Pertea G, Pertea M. Improved transcriptome assembly using a hybrid of long and short reads with StringTie. PLoS Comput Biol. 2022;18(6): e1009730.
23. Li H. Protein-to-genome alignment with miniprot. Bioinformatics. 2023;39(1):btad014.
24. Hoff KJ, Stanke M. Predicting Genes in Single Genomes with AUGUSTUS. Curr Protoc Bioinformatics. 2019;65(1): e57.
25. Haas BJ, Salzberg SL, Zhu W, Pertea M, Allen JE, Orvis J, White O, Buell CR, Wortman JR. Automated eukaryotic gene structure annotation using EVidenceModeler and the Program to Assemble Spliced Alignments. Genome Biol. 2008;9(1):R7.
26. Ontology G. C: Gene Ontology Consortium: going forward. Nucleic Acids Res. 2015;43(Database issue):1049–56.
27. Kanehisa M, Sato Y, Morishima K. BlastKOALA and GhostKOALA: KEGG Tools for Functional Characterization of Genome and Metagenome Sequences. J Mol Biol. 2016;428(4):726–31.
28. Zhou J, Xu Y, Lin S, Guo Y, Deng W, Zhang Y, Guo A, Xue Y. iUUCD 2.0: an update with rich annotations for ubiquitin and ubiquitin-like conjugations. Nucleic Acids Res. 2018;46(D1):D447-d453.
29. Emms DM, Kelly S. OrthoFinder: phylogenetic orthology inference for comparative genomics. Genome Biol. 2019;20(1):238.
30. Elbourne LD, Tetu SG, Hassan KA, Paulsen IT. TransportDB 2.0: a database for exploring membrane transporters in sequenced genomes from all domains of life. Nucleic Acids Research. 2017;45(D1):D320-d324.
31. Zheng Y, Jiao C, Sun H, Rosli HG, Pombo MA, Zhang P, Banf M, Dai X, Martin GB, Giovannoni JJ, et al. iTAK: A Program for Genome-wide Prediction and Classification of Plant Transcription Factors, Transcriptional Regulators, and Protein Kinases. Mol Plant. 2016;9(12):1667–70.
32. Wu P, Peng M, Li Z, Yuan N, Hu Q, Foster CE, Saski C, Wu G, Sun D, Luo H. DRMY1, a Myb-Like Protein, Regulates Cell Expansion and Seed Production in Arabidopsis thaliana. Plant Cell Physiol. 2019;60(2):285–302.
33. Cao Y, Yan H, Sheng M, Liu Y, Yu X, Li Z, Xu W, Su Z. KAKU4 regulates leaf senescence through modulation of H3K27me3 deposition in the Arabidopsis genome. BMC Plant Biol. 2024;24(1):177.
34. Zhao A, Cui C, Li F, Li C, Naveed S, Dong J, Gao X, Rustgi S, Wen S, Yang M. Heterologous expression of the TaPI-PLC1-2B gene enhanced drought and salt tolerance in transgenic rice seedlings. Heredity. 2022;129(6):336–45.
35. Yan H, Zhang X, Li X, Wang X, Li H, Zhao Q, Yin P, Guo R, Pei X, Hu X, et al. Integrated Transcriptome and Metabolome Analyses Reveal the Anthocyanin Biosynthesis Pathway in Am Rosea1 Overexpression 84K Poplar. Frontiers in bioengineering and biotechnology. 2022;10: 911701.
36. De Jesus Martinez T, Hershberg EA, Guo E, Stevens GJ, Diesh C, Xie P, Bridge C, Cain S, Haw R, Buels RM, et al. JBrowse Jupyter: a Python interface to JBrowse 2. Bioinformatics (Oxford, England). 2023;39(1):btad032.
37. Deng W, Nickle DC, Learn GH, Maust B, Mullins JI. ViroBLAST: a standalone BLAST web server for flexible queries of multiple databases and user's datasets. Bioinformatics (Oxford, England). 2007;23(17):2334–6.
38. Yu G, Wang LG, Han Y, He QY. clusterProfiler: an R package for comparing biological themes among gene clusters. OMICS. 2012;16(5):284–7.
39. Morel P, Chambrier P, Boltz V, Chamot S, Rozier F, Rodrigues Bento S, Trehin C, Monniaux M, Zethof J, Vandenbussche M. Divergent Functional Diversification Patterns in the SEP/AGL6/AP1 MADS-Box Transcription Factor Superclade. Plant Cell. 2019;31(12):3033–56.
40. Heijmans K, Ament K, Rijpkema AS, Zethof J, Wolters-Arts M, Gerats T, Vandenbussche M. Redefining C and D in the petunia ABC. Plant Cell. 2012;24(6):2305–17.
41. Panikashvili D, Shi JX, Schreiber L, Aharoni A. The Arabidopsis ABCG13 transporter is required for flower cuticle secretion and patterning of the petal epidermis. New Phytol. 2011;190(1):113–24.
42. Zhang D, Yu Z, Zeng B, Liu X. Genome-wide analysis of the ABC gene family in almond and functional predictions during flower development, freezing stress, and salt stress. BMC Plant Biol. 2024;24(1):12.
43. Nguyen VNT, Usman B, Kim EJ, Shim SH, Jeon JS, Jung KH. An ATP-binding cassette transporter, OsABCB24, is involved in female gametophyte development and early seed growth in rice. Physiol Plant. 2024;176(3): e14354.
44. Qi X, Fang H, Chen Z, Liu Z, Yu X, Liang C. Ectopic Expression of a R2R3-MYB Transcription Factor Gene LjaMYB12 from Lonicera japonica Increases Flavonoid Accumulation in Arabidopsis thaliana. Int J Mol Sci. 2019;20(18):4494.
45. Kong DX, Li YQ, Bai M, He HJ. Liang GX. Wu HJIC, Products: Correlation between the dynamic accumulation of the main effective components and their associated regulatory enzyme activities at different growth stages in Lonicera japonica Thunb. 2017;96:16–22.
46. Freytes SN, Canelo M, Cerdán PD. Regulation of Flowering Time: When and Where? Curr Opin Plant Biol. 2021;63: 102049.
47. Chen Y, Zhang L, Zhang H, Chen L, Yu D. ERF1 delays flowering through direct inhibition of FLOWERING LOCUS T expression in Arabidopsis. J Integr Plant Biol. 2021;63(10):1712–23.
48. Song GQ, Walworth A, Zhao D, Jiang N, Hancock JF. The Vaccinium corymbosum FLOWERING LOCUS T-like gene (VcFT): a flowering activator reverses photoperiodic and chilling requirements in blueberry. Plant Cell Rep. 2013;32(11):1759–69.

Publisher's Note