Pol. J. Environ. Stud. Vol. 34, No. 4 (2025), 4927-4941
DOI: *10.15244/pjoes/190648* ONLINE PUBLICATION DATE: 2025-01-31

Original Research

Transcriptomic Analysis of Key Genes in Sex Differentiation of Male and Female Flower Buds during Physiological Differentiation of *Juglans mandshurica*

Mengmeng Zhang[1, 2], Jingze Li[1, 2], Baiting Qin[3], Chunping Liu[1, 2], Lijie Zhang[1, 2]*

[1]Shenyang Agricultural University, Shenyang, Liaoning, China
[2]Key Laboratory of Forest Tree Genetics, Breeding and Cultivation of Liaoning Province, Shenyang, Liaoning, China
[3]College of Life Engineering, Shenyang Institute of Technology, China

Received: 14 December 2023
Accepted: 30 June 2024

Abstract

By comparing the transcriptome data of male and female flower buds of *Juglans maudsurica*, we investigated the genetic regulation differences in male and female flower bud development and identified key genes. In this study, male and female flower buds of the male and female precocious type at the physiological differentiation stage were used as test materials. Through transcriptome sequencing data analysis, a total of 3,501 differentially expressed genes (DEGs) were identified, including 2,236 up-regulated genes and 2,184 down-regulated genes(Differentially expressed genes overlap between distinct comparison groups). Enrichment function and metabolic pathway analyses revealed that the sex differentiation process *J. mandshurica* flower buds involves complex biological processes and regulatory mechanisms. The phytohormone signal transduction pathway was notably enriched in all the control groups, indicating its crucial role in the process of sex differentiation. Furthermore, the number of up-regulated and down-regulated genes varied among the different controls, suggesting the presence of inhibitory effect genes during flower bud development. Using bidirectional clustering heat map analysis, we identified 11 key genes associated with sex differentiation in female-first flower buds and 10 key genes associated with sex differentiation in male-first flower buds. Overall, this study uncovers genetic regulatory differences between female-predominant and male-predominant flower buds during *J. mandshurica* flower bud development, and identifies several key genes and metabolic pathways associated with flower bud and sex differentiation. These findings contribute to a better understanding of the biological regulatory mechanisms underlying bud sex differentiation.

Keywords: *J. mandshurica*, sex differentiation, transcriptome, key genes, flower buds, differential gene expression, enrichment pathway, metabolic pathways

*e-mail: Zlj330@syau.edu.cn

Introduction

According to previous research, the *J. mandshurica* possesses both dioecious and heterozygous traits and exhibits two main types of reproduction: male-first-female and female-first-male [1]. Flower formation is a critical node in the plant life cycle, as it plays a vital role in maintaining plant yield and quality by regulating flower development [2]. The *J. mandshurica* reproductive characteristic of the huckleberry, which is dioecious and heterozygous, affects the flowering time between male and female flowers, leading to unstable pollination and fruiting rates [3]. Therefore, understanding the *J. mandshurica* of sex differentiation regulation of the huckleberry is important for improving fruit set and yield, and subsequently increasing its economic value.

Sexual differentiation is a crucial mechanism that allows plants to adapt to their surroundings and compete with other species [4]. This process is coordinately controlled by sex chromosomes [5-10], phytohormones [11-14], and environmental factors [15-18], which results in the emergence of distinct morphologies. Research has thoroughly investigated sex differentiation in various plant species, such as *Juniperus chinensis*, *Ginkgo biloba*, and *Populus*. The deletion of plant-related enzyme-encoding genes, rate-limiting step enzymes in ethylene (ETH) biosynthesis, gibberellins, oleoresin steroids, dehydration status of the plant, fungal infection, and key transcription factors have been found to play a crucial role in the control of sex differentiation [19-27]. Advances in biotechnology, genomics, and high-throughput sequencing have enabled the investigation of species such as *Citrus reticulata*, *Morus alba*, and *Camellia oleifera* using transcriptomic analysis, which provides an important reference to reveal the molecular regulation of plant sex differentiation [28-37]. Despite the existence of relevant literature, research specifically addressing the asynchrony between male and female flower bud development in Juglans regia is limited [38].

In this study, we performed transcriptome sequencing of *J. mandshurica* flower buds during the crucial physiological differentiation stage to unravel the underlying molecular mechanisms governing the development of male and female buds. Through comparative analysis of differentially expressed genes, our aim was to identify and characterize the pivotal candidate genes involved in this complex biological process. By delving into this intricate process, we can gain profound insights into the mechanisms driving the development of male and female buds. Moreover, it enables us to shed light on the broader aspects of reproductive biology in woody plants.This study lays a solid foundation for expanding the existing knowledge on the reproductive biology of woody plants. It also offers a theoretical basis and technical support for future endeavors aimed at enhancing fruit yield and quality, as well as selecting and breeding superior varieties for forestry production. The outcomes of this research have immense potential for advancements in the field and contribute to the sustainable development of the forestry industry.

Materials and Methods

Plant Material

The experimental materials were collected on April 2, 2021, during the physiological differentiation period of rowan bud development, from rowan trees growing in the natural forests of the experimental forestry field in the Dichegou area of Liaoning Province (latitude and longitude coordinates). Three robust female-preferred and male-preferred rowan plants were carefully selected. Male- and female-preferred flower buds were collected from these plants. The pre-pistillate female flower bud is designated as T1, with three replicates labeled as A1, A2, and A3; the pre-pistillate male flower bud is designated as T2, with three replicates labeled as A4, A5, and A6; the pre-male female flower bud is designated as M1, with three replicates labeled as B1, B2, and B3; the pre-male male flower bud is designated as M2, with three replicates labeled as B4, B5, and B6, totaling 12 samples (Table 1). Upon collection, the male and female flower buds were rapidly frozen in liquid nitrogen and subsequently stored at -80ºC. The samples were handled and preserved under consistent conditions to maintain their integrity for further analysis (Fig. 1).

RNA-Seq and Data Processing

A total of 12 RNAseq libraries were constructed (A1-6, B1-6), with each group comprising three libraries as triplicates. The collected samples were sent to Paisano Biotechnology Ltd (Shanghai, China) for high-throughput sequencing of the cDNA libraries. The data filtering procedure involved the utilization of Cutadapt to eliminate junctions at the 3' end and reads with an average mass fraction below Q20 [39-42]. This was done to obtain clean reads suitable for further analysis (Table 2).

Identification of Differentially Expressed Genes (DEGs) and Functional Annotation

The updated version of TopHat2, HISAT2 software, was employed to align the filtered reads to the reference genome. HTSeq was utilized to perform statistical comparisons of the read count values for each gene, representing the raw gene expression levels. To facilitate the comparability of gene expression levels across genes and samples, the expression amounts were normalized using FPKM (Fragments Per Kilobase of transcript per Million mapped reads) [43, 44]. FPKM is calculated as the number of fragments per kilobase of gene length per million mapped fragments, based on the alignment results. Differential expression analysis of genes was conducted using DESeq [45]. The criteria

Table 1. Statistics of the RNA-Seq results.

Sample	Read No.	Bases (bp)	Q30 (bp)	N (%)	Q20 (%)	Q30 (%)
T1-A1	45368780	6850685780	6341361430	0.009338	96.86	92.56
T1-A2	44598348	6734350548	6218875194	0.010669	96.78	92.34
T1-A3	44148488	6666421688	6181735385	0.011707	96.94	92.72
T2-A4	45558676	6879360076	6346050033	0.010119	96.72	92.24
T2-A5	44185528	6672014728	6045407008	0.004519	95.92	90.6
T2-A6	41358082	6245070382	5678672846	0.004431	96.1	90.93
M1-B1	41429644	6255876244	5681077290	0.003885	96.04	90.81
M1-B2	44856302	6773301602	6098898224	0.004014	95.64	90.04
M1-B3	43141164	6514315764	5930449431	0.004257	96.17	91.03
M2-B4	42024482	6345696782	5756959178	0.004058	96	90.72
M2-B5	46434224	7011567824	6374994728	0.004345	96.11	90.92
M2-B6	43188954	6521532054	5926553114	0.004346	96.08	90.87

Note: T1, Pre-pistillate female flower buds; T2, Pre-pistillate male flower buds; M1, Pre-male female flower buds; M2, Pre-male male flower buds; A1-A3, A4-A6, B1-B3, and B5-B6 represent three replicates for each of the four groups.

for identifying differentially expressed genes were a fold change |log2FoldChange|>1 and a significance level of P-value<0.05.

Following the completion of differential gene analysis, the differentially expressed genes were classified based on the annotation information of the genome. GO enrichment analysis was performed using topGO [46], which calculates gene lists and the number of genes associated with each GO term using the differentially expressed genes annotated with GO terms. It then determines the significance of enrichment using hypergeometric distribution (with the criteria for significant enrichment being a P-value<0.05). This analysis helps identify the significantly enriched GO terms associated with differential genes, providing insight into the main biological functions governed by these genes. KEGG (Kyoto Encyclopedia of Genes and Genomes) enrichment analysis was conducted using clusterprofiler [47]. It calculates the list of genes and the number of genes involved in each pathway using the differentially expressed genes annotated with KEGG pathways. A hypergeometric distribution is employed to determine the significance of enrichment (with the criteria for significant enrichment being a P-value<0.05). This analysis identifies the KEGG pathways significantly enriched with differential genes compared to the whole genomic background, revealing the major biological functions regulated by these genes.

Cluster Analysis and Trend Analysis

The combined sets and samples of differentially expressed genes from all comparison groups were subjected to two-way clustering using the R language

Fig. 1. Different types of floral buds comparison. a, gynoecious precocious female floral bud; b, gynoecious precocious male floral bud; c, androecious precocious female floral bud; d, androecious precocious male floral bud.

Table 2. Quantitative primer table.

Primer name	Primer sequence
ARF3-F	CTCTGACGGATGCTGCTA
ARF3-R	GCTCTTCAAGAATTTACGG
KNAT1-F	CTACCCAAAGATGCCAGAC
KNAT1-R	CCAGTTGATTCAGCCAAT
MED28-F	ATGGTGGAGCGACAAGCG
MED28-R	GGCAAGCACGGAAGCAAA
UGP1-F	GCAATGTACCTCGGTCTC
UGP1-R	ATAAAGCCATCCTCCAAA
NMD3-F	AGTTGATGATAGGAAGGGTA
NMD3-R	AGTGGAACAGAAGGCAAG
RPL27A3-F	AGGACAGCGTGCCTTTGG
RPL27A3-R	TCTTCTCGGCGATCTTGG
IAA6-F	TATTGGACGGAAGTGGTG
IAA6-R	TAAGAGCAGAAAGTTCAGAGCTCTT
EF1-F	AGGAGCTGGAGAAGGAGCCCA
EF1-R	AACAGCAACGGTCTGGCGCA

Pheatmap software package. This clustering was based on the expression levels of the same gene in different samples and the expression patterns of different genes in the same sample. The results of the two-way clustering heatmap analysis were utilized to further classify the genes into clusters based on the similarity of their expression patterns, defaulting to nine clusters. It is postulated that genes within each cluster belong to a distinct category and are more likely to exhibit similar functions.

The visualization of the expression trends of different gene types across samples is represented by the blue trend line, enabling a focused analysis on key genes. The distance calculation was performed using the Euclidean method, and hierarchical clustering employing the complete linkage method was used for clustering. This approach aids in identifying genes with similar expression patterns and provides insights into their potential functional similarities.

qRT-PCR Validation of RNA-seq Data

Seven differentially expressed key genes were randomly selected for validation of their expression in different samples using RT-qPCR, aiming to confirm the accuracy of the RNA-Seq data. Total RNA extraction in this experiment was performed using the TIANGEN kit, and the qPCR reaction system had a total volume of 20 μL, comprising 10 μL of 2 × ChamQ Universal SYBR qPCR Master Mix (Vazyme Biotechnology, Nanjing, China), each containing 0.4 μmol/L forward and reverse primers, 1 μL of tenfold diluted cDNA template, and 8.2 μL of ddH2O.

The PCR conditions were as follows: initial denaturation at 9ºC for 30 s, followed by 40 cycles of denaturation at 95ºC for 10 s, and annealing/extension at 60ºC for 30 s. These PCR conditions were employed to ensure accurate amplification of the specific gene targets and to maintain the integrity of the qPCR process.

Results

RNA-Seq Data Analysis

The Q30 base percentage of all 12 sequenced libraries exceeded 90%. Additionally, the percentage of high-quality sequence bases in each library after filtering surpassed 90% of the sequenced bases. The filtered reads were aligned to the reference genome, and the comparison results revealed a mapping ratio of over 70%, indicating the appropriate selection of the reference genome and the absence of relevant data contamination (Table 2). The gene expression profiles across sample groups exhibited a favorable abundance pattern, signifying a uniform distribution of gene expression levels (Fig. 2A). Violin plots were primarily employed to visualize the distribution of probability densities of the data samples (Fig. 2B).

PCA analysis facilitated the clustering of similar samples, with the proximity between samples indicating higher similarity. Based on the PCA analysis, the three biological replicates within each of the four subgroups clustered together, reflecting the consistency within each subgroup (Fig. 2C). Pearson's correlation coefficient was utilized to assess the correlation of gene expression levels among different samples. The resulting analysis portrayed the strength and direction of the linear relationship between variables, with correlation coefficients exceeding 0.95 (Fig. 2D), indicative of robust inter-sample correlation.

Comparative Transcriptomic Analysis of Different Types of Male and Female Flower Buds

To comprehensively elucidate the variability in gene expression patterns during the development of male and female flower buds, differential gene expression analysis was conducted using DESeq across four comparison groups (T1 VS T2, T1 VS M1, T2 VS M2, M1 VS M2). Genes with a fold change |log2FoldChange|>1 and a significant P-value<0.05 were considered for differential analysis. A total of 3,501 differentially expressed genes (DEGs) were identified, comprising 741 DEGs (315 up-regulated, 426 down-regulated) in the T1_vs_T2 comparison group, 552 DEGs (400 up-regulated, 152 down-regulated) in the T1_vs_M1 comparison group, 1124 DEGs (512 up-regulated, 612 down-regulated) in the M1_vs_M2 comparison group, and 2003 DEGs in the T2_vs_M2 comparison group (1009 up-regulated,

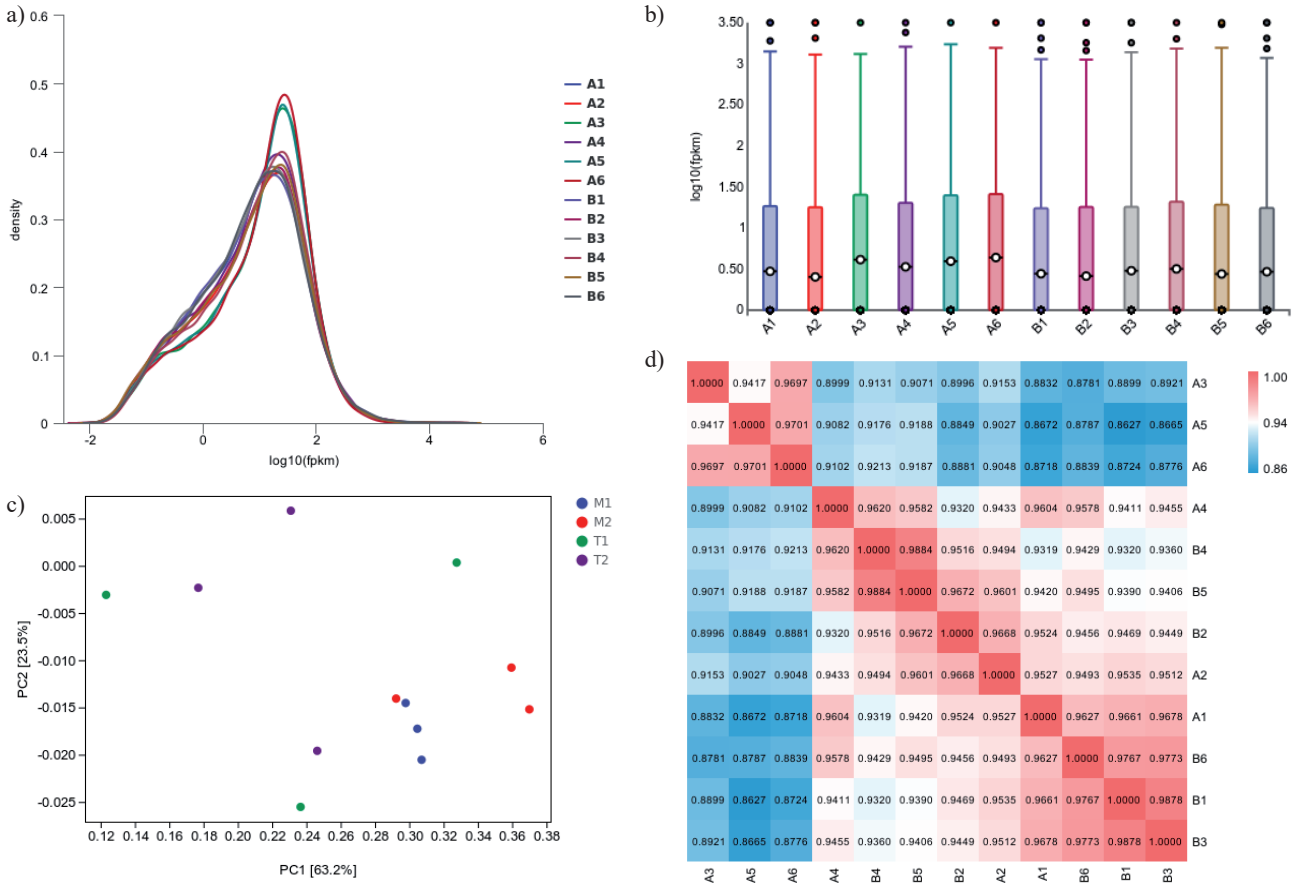

Fig. 2. RNA-Seq data analysis. a) density plot: the x-axis represents the log10 (FPKM) value of the gene, while the y-axis illustrates the density of the gene distribution corresponding to the level of expression. b) violin plot: the horizontal line within the central box denotes the median, with the upper and lower edges of the box representing the 75th and 25th percentiles, respectively. The wider areas adjacent to the box depict kernel density estimates. c) the x-axis corresponds to the first principal component, and the y-axis represents the second principal component. Different shapes within the plot denote distinct samples, while varying colors indicate different groupings. d) The left and top sides of the plot showcase the sample clustering patterns, while the right and bottom sides display the sample names. Different colored squares signify varying degrees of correlation between two samples.

994 down-regulated). The substantial number of differentially expressed genes observed across the various comparison groups highlights the divergent genetic regulation between the two types of flower buds during male and female flower bud development. Further analyses will be pivotal in identifying key developmental genes and enhancing our understanding of the mechanisms underlying plant male and female floral organ development (Fig. 3A-F).

GO and KEGG Enrichment Analysis of Common DEGs in Four Comparison Groups

The GO enrichment analysis of the four comparison groups revealed a significant enrichment of biological process functions. Specifically, five biological process functions — bud phylogeny, germinal bud phylogeny, floral development, floral organ development, and template transcriptional regulation — were enriched in both the T1VsT2 and M1VsM2 groups, with bud phylogeny showing the highest level of enrichment. These findings indicate that the gene regulation of several biological process functions is involved in the sex differentiation of male and female flower buds. Notably, shoot phylogeny was significantly enriched in the sex differentiation process, suggesting its substantial influence on the sex determination of flower buds (Fig. 4A,C,E,G).

The analysis of KEGG pathways in the four comparison groups identified a total of 212 differentially expressed genes (91 up-regulated and 121 down-regulated) mapped to 70 pathways in the T1 vs T2 group. Enrichment analysis revealed significant enrichment of these 212 DEGs in essential biological processes such as phytohormone signaling, phenylpropane biosynthesis, photosynthesis antenna proteins, riboflavin metabolism, biosynthesis of neomycin/kanamycin/gentamycin, nitrogen metabolism, starch and sucrose metabolism, and galactose metabolism. Notably, the higher number of down-regulated genes indicates potential inhibition of certain biological processes during bud differentiation, warranting further investigation (Fig. 4B).

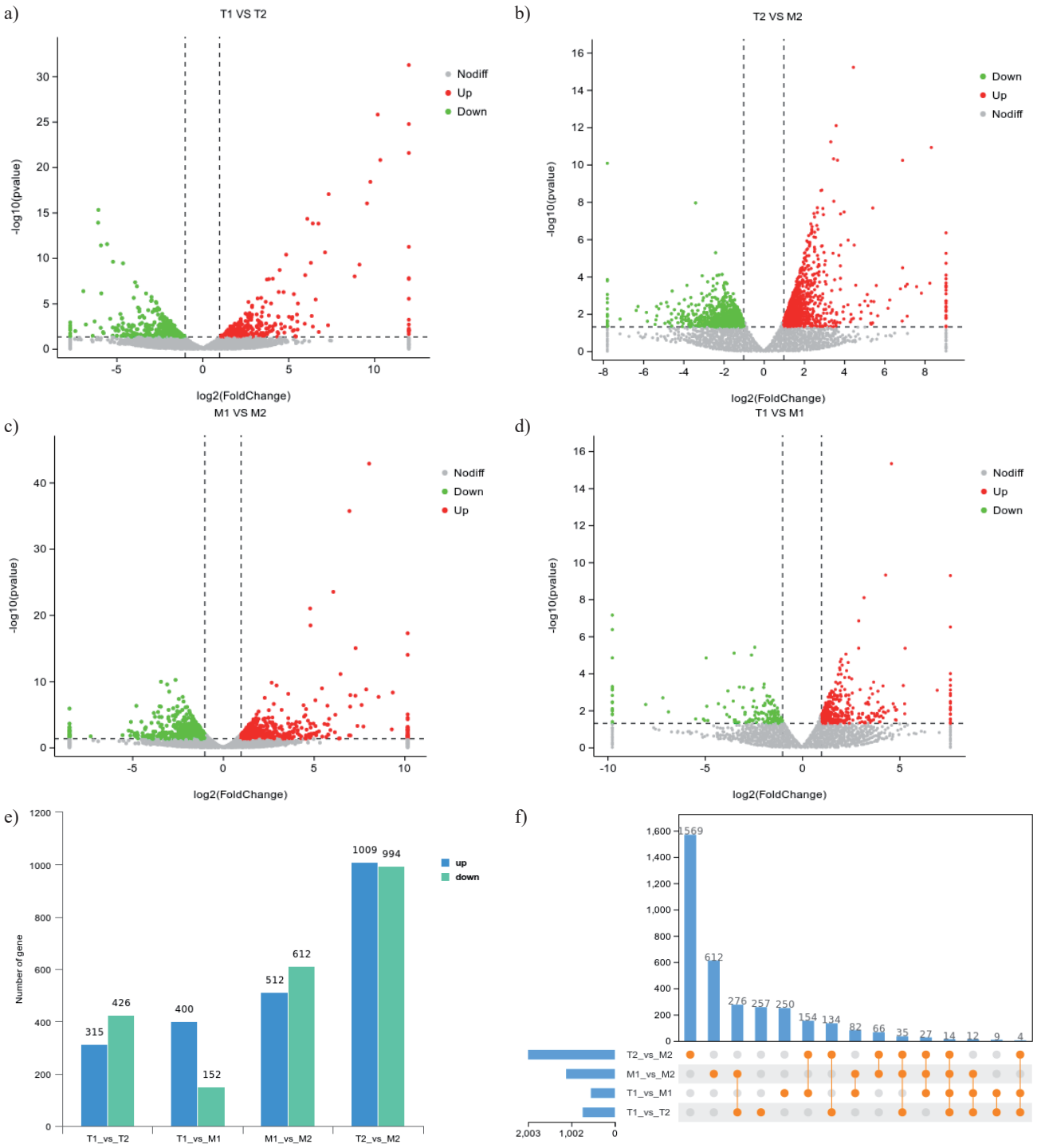

Fig. 3. Comparative transcriptomic analysis of different types of male and female flower buds. A-D, volcano plots depicting log2FoldChange on the horizontal axis and significance level as negative logarithms to the base 10 on the vertical axis. Dashed lines represent thresholds for differential expression, with colors indicating up-regulated, down-regulated, or non-significantly differentially expressed genes. E, statistical summary of differential expression results: Comparison groups are represented on the horizontal axis, while the vertical axis illustrates the number of differentially expressed genes, with colors denoting up- or down-regulation. F, Upset plots displaying the number of differential genes in each comparison group, shared differential genes across multiple comparison groups, and the unique set of differential genes in each comparison group. Points on the horizontal axis represent unique differential genes, while lines connect points corresponding to shared differential genes across multiple comparison groups.

In the M1 vs M2 group, a total of 383 differentially expressed genes (80 up-regulated and 203 down-regulated) were mapped to 97 enriched pathways. In addition to the pathways identified in the T1 vs T2 group, specific enriched pathways such as MAPK signaling, arginine and proline metabolism, and glycosaminoglycan degradation were also identified. These pathways may play crucial roles in bud

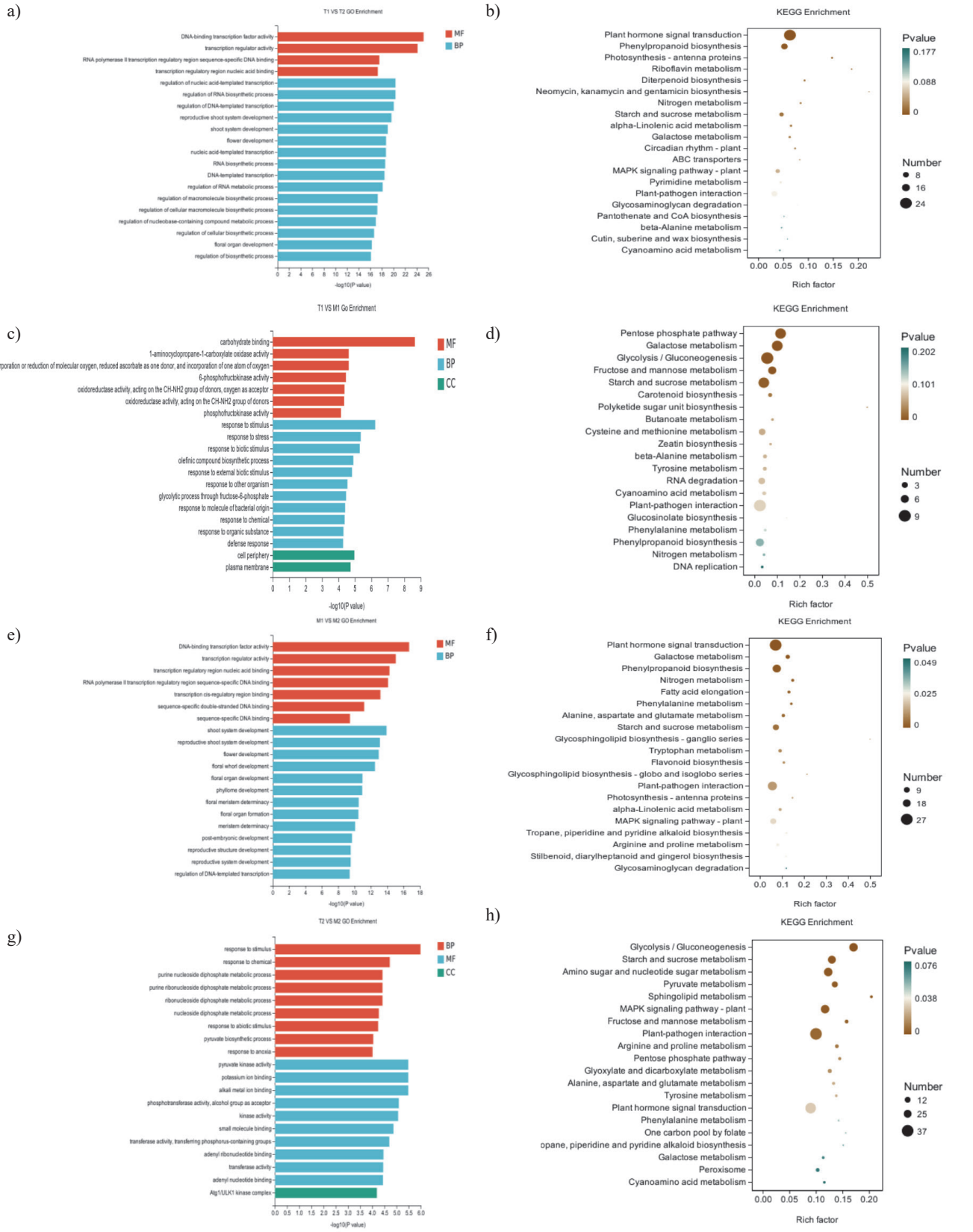

Fig. 4. GO and KEGG enrichment analysis of common DEGs in four comparison groups. **(A/C/E/G)**, GO Enrichment Histograms. The horizontal axis represents the specific GO Term, while the vertical axis depicts the -log10 (p-value) of GO Term enrichment. (B/D/F/H), KEGG Enrichment Factor Plots.The horizontal axis displays the rich factor (number of differentially expressed genes annotated to the Pathway divided by the total number of genes annotated to the Pathway), while the vertical axis indicates the specific Pathway. The size of the dots in the plots corresponds to the number of differentially expressed genes (up-regulated or down-regulated) annotated to the respective Pathway, and the darkness of the color represents the number of genes annotated to the Pathway. Additionally, the color shade reflects the level of significance for the enrichment (up- or down-regulated, depending on the set of genes selected for analysis).

development and growth. Furthermore, similar to the T1 vs T2 group, the higher proportion of down-regulated genes suggests the existence of inhibitory effects during flower bud development. Further research is warranted to explore the potential roles of these enriched pathways in flower bud development and growth given their significant enrichment and the observed higher number of down-regulated genes (Fig. 4F).

This study presents a comparative analysis of gene expression and pathway enrichment during flower bud differentiation. The T2 vs M2 and T1 vs M1 comparison groups were examined, revealing distinct patterns of differentially expressed genes (DEGs) and enriched pathways associated with cell biological processes such as energy supply, carbohydrate metabolism, and signaling. A total of 710 DEGs (346 up-regulated, 364 down-regulated) were identified in the T2 vs M2 comparison group, mapping to 117 pathways. In comparison, the T1 vs M1 group exhibited 159 DEGs (116 up-regulated, 43 down-regulated) mapped to 62 pathways. Notably, the T1 vs M1 comparison group demonstrated a higher proportion of up-regulated genes, indicating enhanced metabolic activities during flower bud differentiation. Specifically, enriched pathways including the pentose phosphate pathway, galactose metabolism, and glycolysis/gluconeogenesis were found to play pivotal roles in metabolism, with potential synergistic effects on bud differentiation. These findings shed light on the metabolic regulatory mechanisms underlying flower bud differentiation, serving as a valuable reference for future studies in this area (Fig. 4D,H).

Cluster Analysis of Differentially Expressed Genes

The transcriptome data underwent clustering and were subsequently partitioned into distinct clusters (default set at 9) based on bi-directional clustering heatmap analysis, reflecting the similarity of gene expression patterns. Each cluster signifies a group of genes likely to perform similar functions, enabling a focused analysis on key genes and the identification of two distinct cluster types (Fig. 5A, D).

The first type consists of Cluster 1, Cluster 2, Cluster 6, and Cluster 7, representing genes exhibiting higher variability in both types of buds. Within these clusters, gene expression was observed to be higher in female buds compared to male buds in Clusters 1, 2, and 7, and lower in Cluster 6 compared to male buds in both types. These findings suggest the potential involvement of genes within these clusters in bud differentiation in R. hirsutus and imply their role in sex development within Juglans mandshurica flower buds. The differential expression patterns of female and male buds across the clusters indicate that genes within Clusters 1, 2, and 7 may be associated with female-predominant sex differentiation, while genes in Cluster 6 may be linked to male-predominant sex differentiation (Fig. 5D).

The second type consisted of clusters 4, 5, 8, and 9. In clusters 4 and 5, the expression of male-predominant genes was found to be higher than that of female-predominant genes in both female and male flower buds. Conversely, clusters 8 and 9 demonstrated the opposite pattern, with higher expression of female-predominant genes compared to male-predominant genes. These findings suggest that certain genes within clusters 4 and 5 may be involved in female-predominant sex differentiation, while genes within clusters 8 and 9 may be associated with male-predominant sex differentiation. Interestingly, no significant difference in expression was observed between female and male flower buds within these four clusters. This indicates that these genes play a role in regulating sex differentiation in both female-predominant and male-predominant phenotypes. In other words, these genes regulate the sex differentiation of flower buds regardless of the predominant sex, resulting in no discernible expression difference between female and male flower buds within these four clusters (Fig. 5D)

A pie chart comparing the transcription factors involved in genes within the two types of clusters revealed that 672 transcription factors were associated with gynoecious-related clusters, while 326 transcription factors were implicated in androecious-related clusters. Notably, the number of transcription factor families related to androecious-related genes was higher, with bHLH having 31, ERF having 31, and C2H2 having 29 representatives. On the other hand, genes related to female-predominant traits displayed a higher number of transcription factor families, including bHLH with 69, ERF with 59, and NAC with 47 representatives. Interestingly, both male-predominant and female-predominant genes exhibited a high presence of bHLH and ERF families. These findings suggest that bHLH and ERF family transcription factors may play crucial roles in regulating sex expression in Nutria species. Additionally, the NAC family was also highly represented in female-predominant-related genes, indicating that NAC transcription factors likely play a significant role in the differential expression of female-predominant and male-predominant sex differentiation. Furthermore, other families such as MYB_related, WRKY, MIKC_MADS, FAR1, C3H, etc., also showed a certain representation in female-predominant-related genes, suggesting their potential regulatory roles in female development and sex expression in huckleberry (Fig. 5B,C)

Based on the aforementioned analysis, our focus encompassed genes within clusters 1, 2, 4, 5, and 7, with the aim of identifying key genes involved in the sex differentiation of female-predominant flower buds. Additionally, attention was directed towards genes within clusters 6, 8, and 9, to conduct a thorough analysis of the key genes implicated in the sex differentiation of male-predominant flower buds.

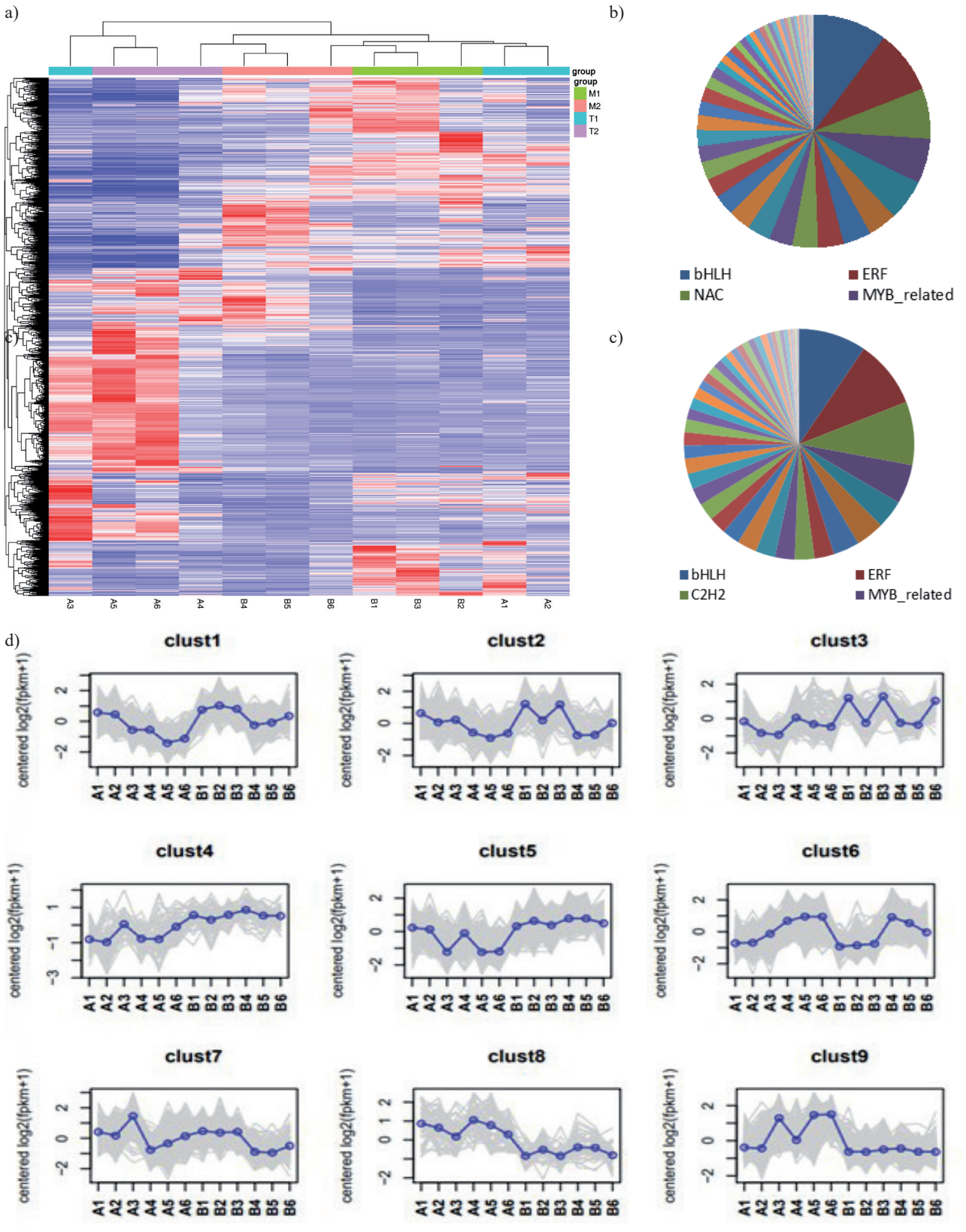

Fig. 5. Cluster analysis of differentially expressed genes. A, clustering diagram of differentially expressed genes, where genes are represented horizontally and each column represents a sample. The color red denotes highly expressed genes, while green indicates genes with low expression levels (According to the similarity of gene expression patterns, they were classified into 9 distinct clusters.) B, pie chart detailing the top transcription factor families with the highest number of differentially expressed genes associated with female-predisposed traits. C, pie chart depicting the transcription factor families predominantly present among differentially expressed genes associated with the female-predisposing phenotype. D, Scatter plot. The gray lines in the figure show the expression patterns of individual genes within each cluster, while the blue lines represent the average expression values of all genes in the cluster across samples.

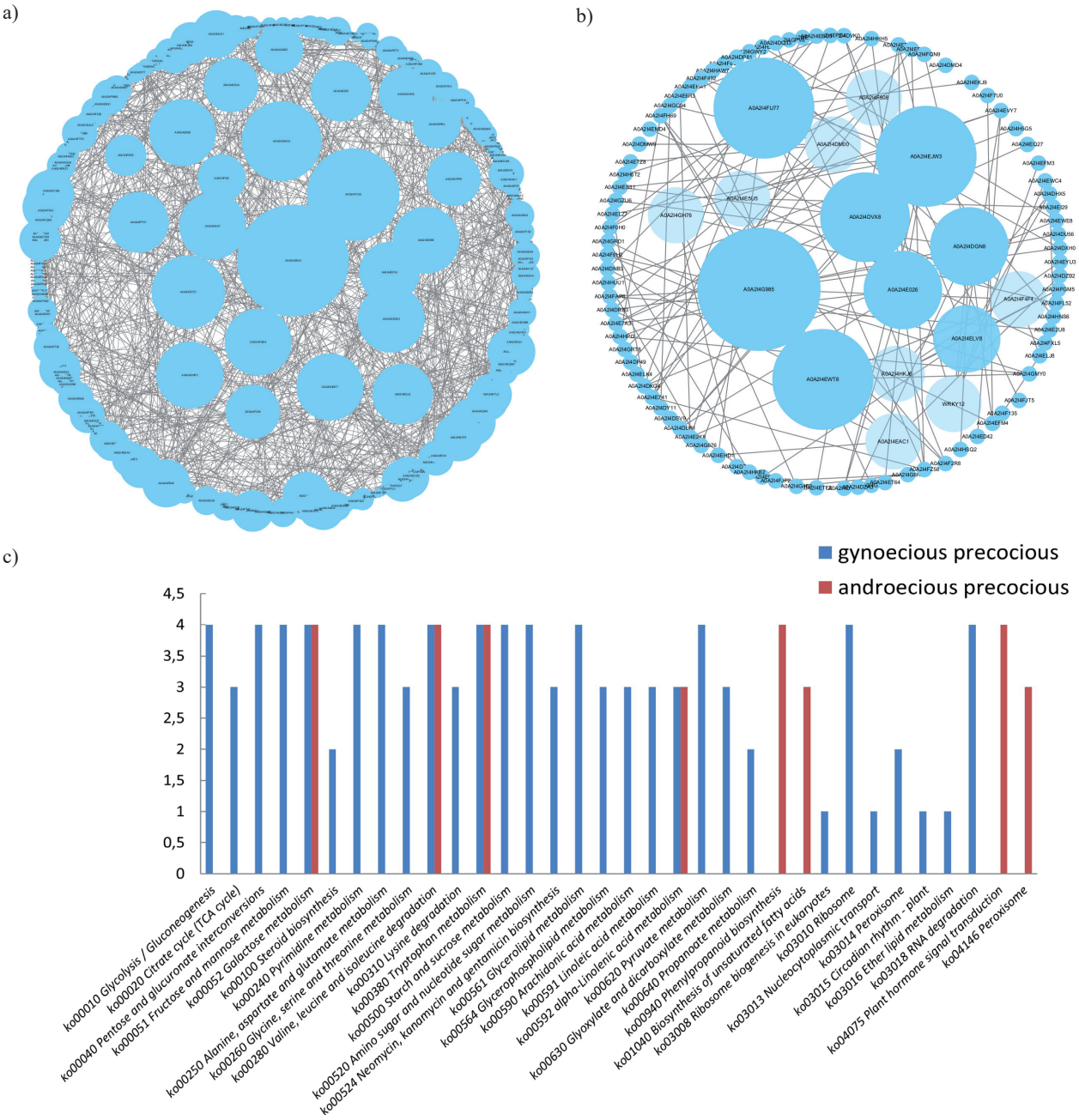

Fig. 6. Illustrates the protein-protein interaction (PPI) network diagrams for female-first male and female flower buds a) and male-first magnetic flower buds b). Additionally, c) presents a comparative analysis of the key gene-related pathways in male and female-first Nutmeg buds.

Screening the Key Genes

A protein-protein interaction (PPI) network graph was constructed using the genes identified in the key clusters as input, integrating known protein interaction databases. Within the PPI network, proteins with the highest number of associations were identified. Eleven key genes associated with female-predominant sex differentiation were screened: SUS3, HXK1, UGP1, RPL27AC, NMD3, CRY1, SGAT, DLS2, SDP1, CPS, and ENO1 (Fig. 6A). Additionally, ten key genes related to male-predominant sex differentiation were identified: IAA6, At3g47200, ARF6, AUX1, ARF3, KNAT1, OMT1, SRP43, KAT2, and MED28 (Fig. 6B).

Comparative analyses of the pathways associated with these key genes across the four comparative groups revealed common pathways for female-predominant and male-predominant key genes, including galactose metabolism (ko00052), valine, leucine and isoleucine degradation (ko00280), tryptophan metabolism (ko00380), and α-linolenic acid metabolism (ko00592). Male-specific pathways encompassed primary amine biosynthesis (ko00940), unsaturated fatty acid biosynthesis (ko01040), phytohormone signaling (ko04075), peroxisome (ko04146), and female-specific

pathways such as glycolysis/gluconeogenesis (ko00010), the citric acid cycle (TCA cycle) (ko00020), α-linolenic acid metabolism (ko00592), pentose and glucuronate interconversions (ko00040), and additional metabolic and biosynthesis pathways.

Comparison further revealed differences in pathway composition between male-predominant and female-predominant key genes, indicative of biological disparities and differential expression of significant functions within Juglans mandshurica individuals. Detailed studies and analyses are requisite to comprehensively comprehend these differences and their impact on sex differentiation between female-predominant and male-predominant individuals in Juglans mandshurica (Fig. 6C).

Quantitative Real-Time PCR (qRT-PCR) Validation of Differentially Expressed Genes (DEGs)

To validate the accuracy of the transcriptome data, seven DEGs were randomly selected from the differentially expressed genes screened in the four comparison groups for real-time quantitative PCR (qRT-PCR) experiments. Three biological and technical replicates were performed for each organ and gene, respectively. The results showed that the expression regulation patterns were consistent with the results obtained by transcriptome sequencing (Fig. 7). The expression of all seven genes showed the same trend between the qPCR and RNA-Seq results, indicating that the RNA-Seq data are highly reliable for further analysis.

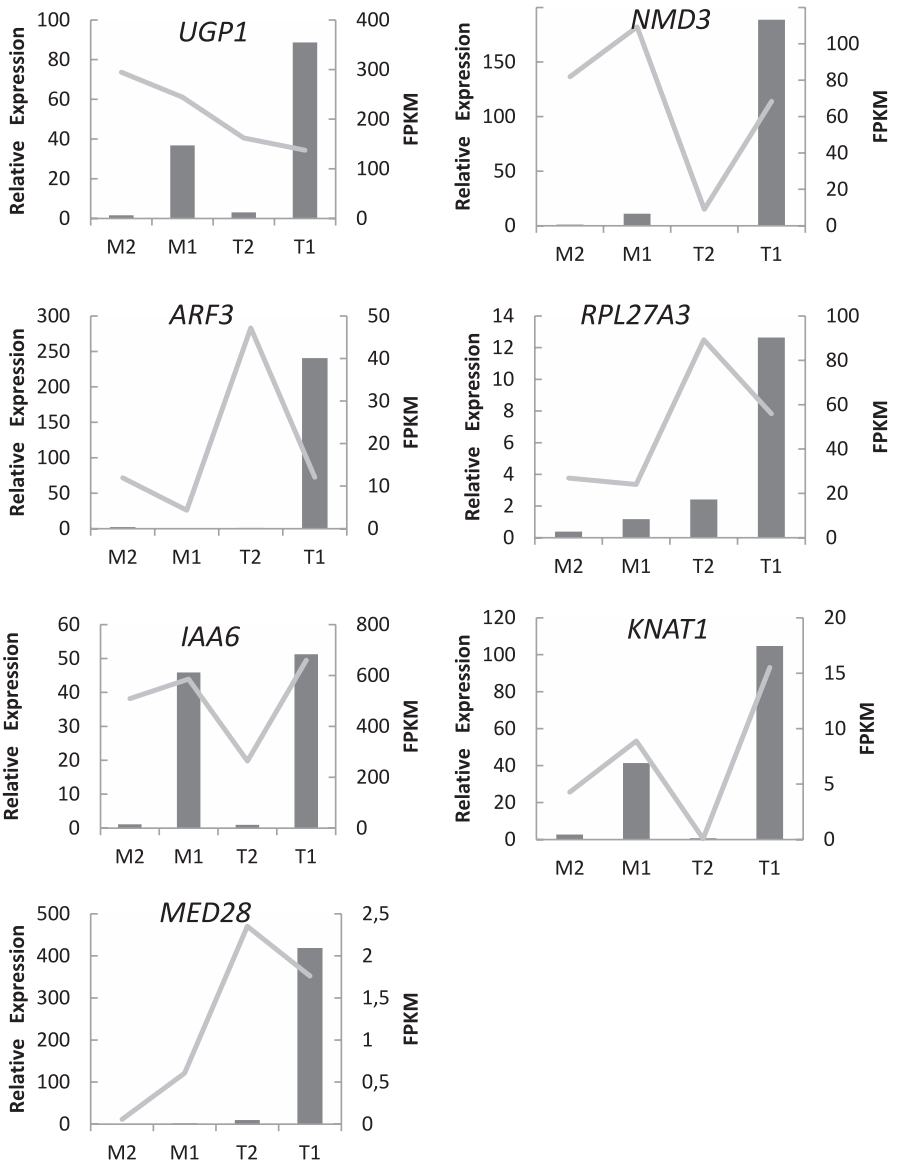

Fig. 7. Presents the results of the qRT-PCR analysis for the seven randomly selected key genes. The X-axis represents the different types of flower buds, while the left Y-axis represents the relative expression levels. The right Y-axis represents the FPKM values. The bars represent the qPCR while the line represents the results of RNA-seq.

Discussion

Understanding the molecular mechanisms that govern gene expression patterns in organisms represents a significant challenge in modern biology. This understanding is essential for obtaining deeper insights into the molecular underpinnings of organism development, morphology, and function. RNA-Seq is a widely adopted technique for exploring differentially expressed genes (DEGs) between female and male flowers, and has been extensively employed in various plant species [48, 49]. Leveraging the advantages of RNA-Seq technology holds remarkable potential for further elucidating the mechanisms of sex differentiation in *C. hirsutus*. In this study, RNA-Seq technology facilitated the identification of a substantial number of differentially expressed genes. A comparison of DEGs in male and female buds across different sets of comparative groups revealed a higher number of differentially expressed genes in male and female buds within the same type of comparison group. This observation suggests that genes play a pivotal role in influencing sex differentiation during bud development, irrespective of the prioritization of female or male buds. The varying counts of up-regulated and down-regulated differentially expressed genes in different control groups imply the presence of potential inhibitory effects during flower bud development. Consequently, further investigations could be directed towards exploring these differentially expressed genes to gain insights into abnormal flower bud development.

Through the analysis of enrichment and metabolic pathways associated with the identified differentially expressed genes (DEGs), we observed that the phytohormone signal transduction pathway exhibited the highest level of enrichment across all control groups. This finding aligns with a previously published article [50], which further validates the crucial role of phytohormones in the process of floral bud differentiation. Moreover, it provides a promising avenue for further exploration and enhancement of our understanding of the molecular mechanisms underlying bud sex differentiation. Notably, we observed variations in the enriched pathways across different control groups, predominantly involving diverse biological processes such as sugar metabolism, nitrogen metabolism, and nucleotide metabolism. These pathways and metabolic processes are likely involved in key biological events during bud differentiation, including bud growth and metabolism. Therefore, a comprehensive investigation into the roles and mechanisms of these pathways and metabolic processes in bud differentiation is warranted to uncover the intricate biological regulatory mechanisms governing the sex differentiation of male and female flower buds during physiological development.

Through bidirectional clustering heat map analysis and PPI network diagram examination, we identified 21 key genes. Among the male-predominant key genes, IAA6, ARF6, AUX1, and ARF3 were found to be associated with growth hormone-related genes, consistent with previous experimental verifications of the effect of hormones on inflorescence development [51]. Additionally, the transcription factor KNAT1, known for its homologous heterodimeric structural domain, plays a crucial role in inflorescence development, as evidenced by its significant function in several species [52, 53]. OMT1, primarily linked to phytocaffeic acid-O-methyltransferase, has been implicated in the methylation alteration pathway, possibly involved in the demethylation pathway in sex differentiation in Nutmeg buds [54]. Moreover, KAT2, known for its role in peroxisomal β-oxidation, has demonstrated effects on inflorescence meristem phenotypes [55]. Furthermore, SRP43 and MED28 were identified as components involved in RNA recognition and RNA polymerase II binding and transcriptional control of class II genes, respectively.

Conversely, among the female-predominant key genes, SUS3 and UGP1 were found to be important in sugar metabolism [56, 57], while HXK1 has been shown to play a significant role in the Arabidopsis photoperiodic pathway and phytohormone signaling pathway [58]. Similarly, CRY1 has been suggested to act as a photoreceptor regulating plant morphogenesis and GA signaling [59], while RPL27A3 and NMD3 encode ribosomal proteins. Furthermore, enzymes such as SGAT, DLD2, SDP1, CPS, and ENO1 were found to be crucial in flower bud differentiation, primarily responsible for plant photorespiration and sugar metabolism.

The observed differences between the female- and male-predominant key genes, with the former mainly involved in enzymes, sugar metabolism, and photoperiodic pathways, and the latter being primarily related to growth hormones, RNAs, and specific transcription factors, may be closely linked to the dioecious maturation mechanism of rowan. This finding holds significant theoretical and practical value, providing a scientific basis for further exploring the mechanism of rowan sex differentiation.

The high concordance between the qRT-PCR validation results and the transcriptome data attests to the accuracy and reliability of the transcriptome analysis. The qPCR results further support the differential expression patterns observed in the RNA-Seq data, indicating that the transcriptome data serve as a credible foundation for further analyses.

Sex differentiation in rowan is a complex and evolutionary process influenced by genes, environmental factors, hormones, and other regulatory factors. Despite advancements in genomics, molecular genetics, and gene mining technology, our understanding of the underlying mechanisms of sex differentiation in rowan remains limited. To further elucidate the regulatory network of sex differentiation, it is imperative to improve the existing framework by integrating multiple metabolic and pathway pathways. This comprehensive approach will facilitate a clearer understanding of the relationship

between sex differentiation and the distinct flower types in rowan. Future studies should focus on analyzing the functions and interactions of key genes and advancing the gene regulatory network of sex differentiation. Additionally, establishing an internal and external multifactorial interaction regulatory network will serve as a cornerstone for studying bud differentiation and contribute to the development of high-yield and stable cultivation practices and phenotypic genetic breeding of rowan.

Conclusions

Differential gene expression analysis of male and female flower buds in Juglans mandshurica revealed 3,501 differentially expressed genes (DEGs), including 2236 up-regulated genes and 2184 down-regulated genes. This suggests distinct genetic regulation differences between female-priority and male-priority flower buds during their development. Functional enrichment analysis and pathway annotation of *C. hirsutus* demonstrated the involvement of complex biological processes and regulatory mechanisms in flower bud differentiation. Notably, the phytohormone signal transduction pathway was identified as the most significantly enriched pathway, highlighting the important role of phytohormones in flower bud differentiation.

Bidirectional clustering heat map analysis was employed to identify key genes associated with sex differentiation in female-predominant and male-predominant flower buds. A total of 11 key genes involved in female-predominant sex differentiation were identified, including SUS3, HXK1, UGP1, RPL27AC, NMD3, CRY1, SGAT, DLS2, SDP1, CPS, and ENO1. Similarly, 10 key genes involved in male-predominant sex differentiation were identified, including IAA6, At3g47200, ARF6, AUX1, ARF3, KNAT1, OMT1, SRP43, KAT2, and MED28. The metabolic pathways associated with these key genes displayed significant variability, highlighting the importance of metabolic pathways in the study of sex differentiation.

Acknowledgment

We express our gratitude to Shanghai Personal Biotechnology Company for their valuable assistance with the sequencing analysis. We would also like to extend our thanks to the reviewers and editors for their insightful comments and suggestions.

Funding

This work has received support from the Applied Basic Research Project of the Liaoning Provincial Department of Science and Technology. (2022JH2/101300170.). The funders did not have any involvement in the study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Conflict of Interest

The authors declare no conflict of interest.

References

1. ZHANG L.J., GUO C., QIN B.T., LU X.J., YANG Y.C., QI Y.H., SHEN H.L. Phonological characteristics of flowering and pollen viability of *Juglans mandshurica*. Journal of Northeast Forestry University, **47**, 4, **2019**.
2. LI H., LI Y., ZHANG X., CAI K., LI Y., WANG Q., QU G., HAN R., ZHAO X.Y. Genome-wide identification and expression analysis of the MADS-box gene family during female and male flower development in *Juglans mandshurica*. Frontiers in Plant Science, **13**, 1020706, **2022**.
3. QIN B., LU X., SUN X., CUI J., DENG J., ZHANG L. Transcriptome-based analysis of the hormone regulation mechanism of gender differentiation in *Juglans mandshurica* Maxim. Peer J, **9**, e12328, **2021**.
4. DURAND R., DURAND B. Sexual differentiation in higher plants. Physiologia Plantarum, **60**, 267, **1984**.
5. YANG L., HUANG Y., FU Z., XU Q. Research Progress on the Epigenetic Mechanisms of Sex Determination in Horticultural Plants. Acta Horticulturae Sinica, **49**, 1602, **2022**.
6. XU Z., CHEN Y., GAO M., WU L., ZHAO Y., WANG Y. Research progress in sex differentiation in angiosperms. Scientia Silvae Sinicae, **55**, 157, **2019**.
7. LIU J., CHATHAM L., ARYAL R., YU Q., MING R. Differential methylation and expression of *HUA1* ortholog in three sex types of *Papaya*. Plant Science, **272**, 99, **2018**.
8. MASSONNET M., COCHETEL N., MINIO A., VONDROS A.M., LIN J., MUYLE A., CANTU D. The genetic basis of sex determination in grapes. Nature communications, **11**, 2902, **2020**.
9. LAI Y.S., SHEN D., ZHANG W., ZHANG X., QIU Y., WANG H., DOU X., LI S., WU Y., SONG J. Temperature and photoperiod changes affect cucumber sex expression by different epigenetic regulations. BMC plant biology, **18**, 268, **2018**.
10. MARKS R.A., SMITH J.J., CRONK Q., GRASSA C.J., MCLETCHIE D.N. Genome of the tropical plant *Marchantia inflexa*: implications for sex chromosome evolution and dehydration tolerance. Scientific reports, **9**, 8722, **2019**.
11. ZHENG Y.Q., LUO X.F., WANG X., MA J., JIANG Y.G., ZHAO J.J, AO Y. The role of phytohormones and their related miRNAs in sex differentiation of *Xanthoceras sorbifolium* Bunge. Scientia Horticulturae, **307**, 111498, **2023**.
12. LIU J., CHEN L., ZHOU P., LIAO Z., LIN H., YU Q., MING R. Sex biased expression of hormone related genes at early stage of sex differentiation in papaya flowers. Horticulture Research, **8**, 147, **2021**.
13. HUANG H., WANG H., HU X., ZHANG Z.Q. Identification of candidate genes associated with sex

differentiation and determination of gender diphasic plant *Lilium apertum* (Liliaceae). Scientia Horticulturae, **306**, 111431, **2022**.
14. HU X., HU F., FAN H., WANG X., HAN B., LIN Y. Effects of five plant growth regulators on blooming and fruit-setting of 'Feizixiao' litchi. Southwest China Journal of Agricultural Sciences, **29**, 915, **2016**.
15. LIU M., KORPELAINEN H., LI C. Sexual differences and sex ratios of dioecious plants under stressful environments. Journal of Plant Ecology, **14**, 920, **2021**.
16. GOODNOE T.T., HILL J.P., AHO K. Effects of variation in carbon, nitrogen, and phosphorus molarity and stoichiometry on sex determination in the fern *Ceratopteris richardii*. Botany, **94**, 249, **2016**.
17. LAI Y.S., ZHANG X., ZHANG W., SHEN D., WANG H., XIA Y., QIU Y., SONG J., WANG C., LI X. The association of changes in DNA methylation with temperature-dependent sex determination in cucumber. Journal of Experimental Botany, **68**, 2899, **2017**.
18. CHEN G., LI J., LIU Y., ZHANG Q., GAO Y., FANG K., XING Y. Roles of the GA-mediated *SPL* gene family and miR156 in the floral development of Chinese chestnut (*Castanea mollissima*). International Journal of Molecular Sciences, **20**, 1577, **2019**.
19. WANG Z., YADAV V., YAN X., CHENG D., WEI C., ZHANG X. Systematic genome-wide analysis of the ethylene-responsive *ACS* gene family: Contributions to sex form differentiation and development in melon and watermelon. Gene, **805**, 145910, **2021**.
20. LI J., SU X., GUO J., XU W., FENG L., WANG T., FU F., WANG G. Sex-related Ddifferences of *Ginkgo biloba* in growth traits and wood properties. Forests, **14**, 1809, **2023**.
21. LI Q., LI J., ZHANG L., PAN C., YANG N., SUN K., HE C. Gibberellins are required for dimorphic flower development in Viola philippica. Plant Science, **303**, 110749, **2021**.
22. CASTRO-CAMBA R., SÁMCHEZ C., VIDAL N., VIELBA J.M. Plant development and crop yield: The role of Gibberellins. Plants, **11**, 2650, **2022**.
23. MASUDA K., AKAGI T., ESUMI T., TAO R. Epigenetic flexibility underlies somaclonal sex conversions in hexaploid persimmon. Plant and Cell Physiology, **61**, 393, **2020**.
24. WANG L., HAN W., DIAO S., SUO Y., LI H., MAI Y., FU J. Study of sexual-linked genes (*OGI* and *MeGI*) on the performance of androecious persimmons (*Diospyros kaki* Thunb.). Plants, **10**, 390, **2021**.
25. ZANEWICH K.P., ROOD S.B. Limited sex differentiation in poplars: similar physiological responses to low temperature of males and females of three cottonwood taxa. Trees, **37**, 1217, **2023**.
26. MÜLLER N.A., KERSTEN B., LEITE MONTALVÃO A.P., MÄHLER N., BERNHARDSSON C., BRÄUTIGAM K., FLADUNG M. A single gene underlies the dynamic evolution of poplar sex determination. Nature plants, **6**, 630, **2020**.
27. CHEN X., WANG L., YAN X., TANG Z. Nutrition regulates sex expression in a gender diphasy plant, *Lilium concolor* var. *megalanthum*. Frontiers in Plant Science, **14**, 1252242, **2023**.
28. XU X., TAO J., XING A., WU Z., XU Y., SUN Y., WANG Y. Transcriptome analysis reveals the roles of phytohormone signaling in tea plant (*Camellia sinensis* L.) flower development. BMC Plant Biology, **22**, 471, **2022**.
29. WANG X., ZHAO F., WU Q., XING S., YU Y., QI S. Physiological and transcriptome analyses to infer regulatory networks in flowering transition of *Rosa rugosa*. Ornamental Plant Research, **3**, 1, **2023**.
30. LI Z., XIAO, W., CHEN H., ZHU G., LV F. Transcriptome analysis reveals endogenous hormone changes during spike development in *Phalaenopsis*. International Journal of Molecular Sciences, **23**, 10461, **2022**.
31. FENG X., ZHOU L., SHENG A., LIN L., LIU H. Comparative transcriptome analysis on drought stress-induced floral formation of *Curcuma kwangsiensis*. Plant Signaling & Behavior, **17**, 2114642, **2022**.
32. DICK C., ARENDT J., REZNICK D.N., HAYASHI C.Y. The developmental and genetic trajectory of coloration in the guppy (*Poecilia reticulata*). Evolution & development, **20**, 207, **2018**.
33. GIANI A.M., GALLO G.R., GIANFRANCESCHI L., FORMENTI G. Long Walk to genomics: History and current approaches to genome sequencing and assembly. Computational and Structural Biotechnology Journal, **18**, 9, **2020**.
34. LIU J., CHEN L., ZHOU P., LIAO Z., LIN H., YU Q., MING R. Sex biased expression of hormone related genes at early stage of sex differentiation in papaya flowers. Horticulture Research, **8**, 147, **2021**.
35. LIU Z., WANG H., XU Z., ZHANG H., LI G., WANG X., QIAN W. Transcriptome profiling of differentially expressed genes of male and female inflorescences in spinach (*Spinacia oleracea* L.). Genome, **64**, 777, **2021**.
36. LI N., MENG Z., TAO M., WANG Y., ZHANG Y., LI S., DENG C. Comparative transcriptome analysis of male and female flowers in *Spinacia oleracea* L. BMC Genomics, **21**, 1, **2020**.
37. YE Z., YU J., YAN W., ZHANG, J., YANG D., YAO G., LIU Z., WU Y., HOU X. Integrative iTRAQ-based proteomic and transcriptomic analysis reveals the accumulation patterns of key metabolites associated with oil quality during seed ripening of *Camellia oleifera*. Horticulture Research, **8**, 157, **2021**.
38. CAI J., JIA R., JIANG Y., FU J., DONG T., DENG J., ZHANG L. Functional verification of the *JmLFY* gene associated with the flowering of *Juglans mandshurica* Maxim. Peer J, **1**, e14938, **2023**.
39. MARTIN M. Cutadapt removes adapter sequences from high-throughput sequencing reads. EMBnet. Journal, **17**, 10, **2021**.
40. JIN W., YAN W., MA M., HASI A., CHE G. Genome-wide identification and expression analysis of the *JMJ-C* gene family in melon (*Cucumis melo* L.) reveals their potential role in fruit development. BMC Genomics, **24**, 771, **2023**.
41. WANG H., YE L., ZHOU L., YU J., PANG B., ZUO D., WANG H. Co-expression network analysis of the transcriptome identified hub genes and pathways responding to saline–alkaline stress in *Sorghum bicolor* L. International Journal of Molecular Sciences, **24**, 16831, **2023**.
42. ALAM B., LIU R., GONG J., LI J., YAN H., GE Q., GONG W. Hub genes in stable QTLs orchestrate the accumulation of cottonseed oil in upland cotton via catalyzing key steps of lipid-related pathways. International Journal of Molecular Sciences, **24**, 16595, **2023**.
43. MORTAZAVI A., WILLIAMS B.A., MCCUE K., SCHAEFFER L., WOLD B. Mapping and quantifying mammalian transcriptomes by RNA-Seq. Nature methods, **5**, 621, **2008**.
44. TRAPNELL C., WILLIAMS B.A., PERTEA G., MORTAZAVI A., KWAN G., VAN BAREN M.J., PACHTER L. Transcript assembly and quantification by

RNA-Seq reveals unannotated transcripts and isoform switching during cell differentiation. Nature biotechnology, **28**, 511, **2010**.
45. WEI L., DU Y., XIANG J., ZHENG T., CHENG J., WU J. Integrated mRNA and miRNA transcriptome analysis of grape in responses to salt stress. Frontiers in Plant Science, **14**, 1173857, **2023**.
46. CONESA A., GÖTZ S., GARCÍA-GÓMEZ J.M., TEROL J., TALÓN M., ROBLES M. Blast2GO: a universal tool for annotation, visualization and analysis in functional genomics research. Bioinformatics, **21**, 3674, **2005**.
47. CHEN Y., ZHANG X., FAN Y., SUI D., JIANG J., WANG L. The role of WRKY transcription factors in exogenous potassium (K+) response to NaCl stress in *Tamarix ramosissima*. Frontiers in Genetics, **14**, 1274288, **2023**.
48. LIU Z.X., GUO C.X., WU R., WANG J.J., ZHOU Y.P., YU X.L., ZHANG Y.X., ZHAO Z.H., LIE H., SUN S.S., HU M.K., QIN A.Z., LIU Y.M., YANG J.C., GEORGE B., SUN X.W. Identification of the regulators of epidermis development under drought-and salt-stressed conditions by single-cell RNA-seq. International journal of molecular sciences, **23**, 2759, **2022**.
49. ADAL A.M., DOSHI K., HOLBROOK L., MAHMOUD S.S. Comparative RNA-Seq analysis reveals genes associated with masculinization in female *Cannabis sativa*. Planta, **253**, 1, **2021**.
50. LI X., HAN R., CAI K., GUO R., PEI X., ZHAO X. Characterization of phytohormones and transcriptomic profiling of the female and male inflorescence development in Manchurian walnut (*Juglans mandshurica* Maxim.). International Journal of Molecular Sciences, **23**, 5433, **2022**.
51. QIN C., DU T., ZHANG R., WANG Q., LIU Y., WANG T., SU S. Integrated transcriptome, metabolome and phytohormone analysis reveals developmental differences between the first and secondary flowering in *Castanea mollissima*. Frontiers in Plant Science, **14**, 1145418, **2023**.
52. ZHAO Y., ZHOU H., WEI K., JIANG C., SONG X., LU M. Structure, expression and function analysis of Class I KNOX genes in *Populus*. Journal of Forest Research, **31**, 118, **2018**.
53. CHEN J.J., WANG W., QIN W.Q., MEN S.Z., LI H.L., MITSUDA N., WU A.M. Transcription factors KNAT3 and KNAT4 are essential for integument and ovule formation in *Arabidopsis*. Plant Physiology, **191**, 463, **2013**.
54. SONG Y., MA K., BO W., ZHANG Z., ZHANG D. Sex-specific DNA methylation and gene expression in andromonoecious poplar. Plant Cell Reports, **31**, 1393, **2012**.
55. YUAN D., ZHANG Y., WANG Z., QU C., ZHU D., WAN H., LIANG Y. BnKAT2 positively regulates the main inflorescence length and silique number in *Brassica napus* by regulating the auxin and cytokinin signaling pathways. Plants, **11**, 1679, **2022**.
56. FANG J., ZHU X., JIA H., WANG C. Research advances on physiological function of plant sucrose synthase. Journal of Nanjing Agricultural University, **40**, 759, **2017**.
57. ZHOU Y., BU Z., QIAN J., CHEN Y., QIAO L., YANG S., YANG Y. The UTP-glucose-1-phosphate uridylyltransferase of *Brucella melitensis* inhibits the activation of NF-κB via regulating the bacterial type IV secretion system. International Journal of Biological Macromolecules, **164**, 3098, **2020**.
58. MOORE B., ZHOU L., ROLLAND F., HALL Q., CHENG W.H., LIU Y.X., SHEEN J. Role of the *Arabidopsis* glucose sensor *HXK1* in nutrient, light, and hormonal signaling. Science, **300**, 332, **2003**.
59. ZHONG M., ZENG B., TANG D., YANG J., QU L., YAN J., ZHAO X. The blue light receptor CRY1 interacts with GID1 and DELLA proteins to repress GA signaling during photomorphogenesis in *Arabidopsis*. Molecular Plant, **14**, 1328, **2021**.