MINISTRY OF EDUCATION AND SCIENCE OF UKRAINE
KYIV NATIONAL UNIVERSITY OF TECHNOLOGIES AND DESIGN
Faculty of Chemical and Biopharmaceutical Technologies
Department of Biotechnology, Leather and Fur

QUALIFICATION THESIS

on the topic **Physiological differences of *Populus wilsonii* tissue cells based on gene expression**

First (Bachelor's) level of higher education
Specialty 162 "Biotechnology and Bioengineering"
Educational and professional program "Biotechnology"

Completed: student of group BEBT-20
Chuanyang JING

Scientific supervisor
Liubov ZELENA, Ph.D., As. prof.

Reviewer
Olga ANDREYEVA, Dr. Sc., Prof.

Kyiv 2024

KYIV NATIONAL UNIVERSITY OF TECHNOLOGIES AND DESIGN

Faculty: Chemical and Biopharmaceutical Technologies
Department: Biotechnology, Leather and Fur
First (Bachelor's) level of higher education
Specialty: 162 Biotechnology and Bioengineering
Educational and professional program Biotechnology

APPROVE
Head of Department of Biotechnology, Leather and Fur, Professor,
Doctor of Technical Science
Olena MOKROUSOVA

_____
«___»_____2024

ASSIGNMENTS
FOR THE QUALIFICATION THESIS
Chuanyang Jing

1. Thesis topic **Physiological differences of *Populus wilsonii* tissue cells based on gene expression**
scientific supervisor Liubov Zelena, PhD
approved by the order of KNUTD "__"_____20__, №_____

2. Initial data for work: assignments for qualification thesis, scientific literature on the topic of qualification thesis, materials of Pre-graduation practice

3. Content of the thesis (list of questions to be developed): literature review; object, purpose, and methods of the study; experimental part; conclusions

4. Date of issuance of the assignments_____

EXECUTION SCHEDULE

№	The name of the stages of the qualification thesis	Terms of performance of stage	Note on performance
1	Introduction	From 01 April 2024 to 11 April 2024	
2	Chapter 1. Literature review	From 06 April 2024 to 20 April 2024	
3	Chapter 2. Object, purpose, and methods of the study	From 21 April 2024 to 30 April 2024	
4	Chapter 3. Experimental part	From 01 May 2024 to 10 May 2024	
5	Conclusions	From 07 May 2024 to 12 May 2024	
6	Draw up a bachelor's thesis (final version)	From 20 May 2024 to 25 May 2024	
7	Submission of qualification work to the supervisor for feedback	From 25 May 2024 to 27 May 2024	
8	Submission of bachelor's thesis to the department for review (14 days before the defense)	27 May 2024	
9	Checking the bachelor's thesis for signs of plagiarism (10 days before the defense)	31 May 2024	
10	Submission of bachelor's thesis for approval by the head of the department (from 7 days before the defense)	03 June 2024	

I am familiar with the task:

Student _____Chuanyang JING

Scientific supervisor _____Liubov ZELENA

SUMMARY

Chuanyang Jing.Physiological differences of *Populus wilsonii* tissue cells based on gene expression – Manuscript

Qualification thesis on the specialty 162 «Biotechnology and Bioengineering». – Kyiv National University of Technologies and Design, Kyiv, 2024.

Populus wilsonii is a tall tree belonging to the family of *Salicaceae*, mainly distributed in Shaanxi, Gansu, Hubei, Sichuan, Yunnan, Tibet and other provinces in China. In this study, transcriptome analysis was performed on buds, leaves, roots, xylem and phloem of *Populus wilsonii* samples to reveal functional differences of different tissues. The results showed that the number of clean reads in bud, leaf, phloem, root and xylem ranged from 24,515,553-36,645,427, and the genome coverage ranged from 77.77% to 95.46%. By analyzing the specific expression genes of different tissues, there were 447, 477, 167, 963 and 474 specific expression genes in the five sites, respectively. The most differentially expressed genes existed between middle tissues and other tissues, followed by root tissues and other tissues. These genes were mainly different in functions such as metabolism and protein transport. This study revealed the functional differences of various tissues from the aspect of gene expression, and provided a theoretical basis for further understanding of the physiological functions of the *Populus wilsonii*..

Keywords: Populus wilsonii , different tissues, transcriptome, functional enrichment

TABLE OF CONTENTS

INTRODUCTION **6**

CHAPTER 1 LITERATURE REVIEW 6

1.1 Introduction to Populus wilsonii wilsonii 7

1.2 Research progress of Populus wilsonii wilsonii 7

1.3 Transcriptome sequencing technology and application 8

1.4 Purpose and significance of this study 10

CHAPTER 2 MATERIALS AND METHODS 11

2.1 Data Acquisition 11

2.2 Gene quantification 11

2.3 Analysis of specific expression genes 11

2.4 Screening of differentially expressed genes 11

2.5 KEGG enrichment analysis 11

2.6 Cluster analysis 12

CHAPTER 3 RESULTS AND DISCUSSION 13

3.1 Sequencing data and gene expression quantification of different tissues of poplar 13

3.2 Analysis of specific genes in different sites 14

3.2.1 KEGG pathway analysis of specific genes expressed in different tissues of Poplar 14

3.3 Results of differential expression gene analysis 19

3.3.1 Analysis of differentially expressed genes 19

3 3.2Analysis of differentially expressed genes in different tissues of poplar22

3.4 Cluster analysis of different tissues of poplar 33

CONCLUSIONS 347

LIST OF REFERENCES DOCUMENTATION 38

INTRODUCTION

Poplar belongs to the poplar family, is a tall tree, up to 25 meters, its diameter can reach 1.5 meters. The bark has shallow longitudinal cracks, showing the characteristics of flaking, and the dorsal hair of the leaves is a flat band, which is significantly different from the cylindrical hair of other species in the genus Lipomicum, and the color is dark grayish-brown. The flower disc is divided, the buds are hypertrophic and oval in shape. The leaves are varied in morphology, ranging from broadly ovate, nearly round to broadly ovate oblong [1].

In 2006, researchers used Sanger sequencing technology to complete the whole genome project of poplar [3]. Now, poplar has been widely accepted as a model tree for forest genome research. Firstly, the genome of poplar (480Mb) is small and the size of rice genome is the same. The small genome of poplar and its high conversion efficiency are conducive to the construction of gene traps, enhancer traps and activation tag libraries [4].

Purpose of the study –

（1） To determine the gene expression of different tissues of *Populus wilsonii*

（2） Identify the genes involved

（3）To reveal the molecular mechanism of the development of *Populus wilsonii*

（ 4 ） Increase the cognition of Populus sativa and construct the genetic information database of Populus wilsonii.

Object of study –*Populus wilsonii*

Subject of study –Transcriptome analysis of different tissues of *Populus wilsonii*

CHAPTER 1
LITERATURE REVIEW

1.1 Introduction to *Populus wilsonii wilsonii*

The species is mainly distributed in Shaanxi, Gansu, Hubei, Sichuan, Yunnan, Tibet and other regions of China. It is born in the hillside forest at an altitude of 1300-3300 meters, especially on both sides of rivers. This kind of material is loose and used for furniture, sheet material, etc.

1.2 Research progress of *Populus wilsonii wilsonii*

Although researchers have conducted in-depth studies on model plants Arabidopsis and rice, due to the wide variety of species in the plant kingdom, the study of plant genomics needs the help of other model plants [2]. At the same time, poplar has rapid growth, easy asexual reproduction, and mature interspecific hybridizable tissue culture and genetic transformation technology. Now, an efficient and stable genetic transformation system has been established for various poplar trees [4][5]. Annotation of gene models derived from poplar genome sequences will help in protein prediction of mass spectrometry data. At the same time, genetic engineering based on poplar genome has achieved initial results in forest genetic improvement. For example, Tian Yingchuan et al. transferred the special insecticidal protein gene of Bacillus thuringiensis into European black poplar, screened and cultured poplar containing the special insecticidal protein gene integrated into chromosomes [6]. In general, the study of poplar genome helps us to understand the

key regulatory factors and signaling pathways in the growth process of poplar, and provides theoretical basis for optimizing growth conditions and increasing yield.

At present, many important achievements have been made in the study of physiological function of poplar by using poplar genome. For example, firstly, Wang Lijun et al. from Southwest University screened 20 R2R3-MYB transcription factors that may be involved in the biosynthesis regulation of phenylpropanoid compounds from poplar genome through bioinformatics analysis technology, and then further cloned the transcription factor MYB6, using genetic and biochemical experiments. The regulatory mechanism of MYB6 on flavonoid and lignin biosynthesis was analyzed in poplar [7]. Secondly, Zongdan et al. from Southwest Agricultural University resequenced the genomic DNA of 25 poplar species. Based on the chloroplast genome of Populus wilsonii pilocarpa as a reference genome, 25 complete chloroplast genomes were obtained to determine the chloroplast genome structure of Populus wilsonii, and completed a comparative analysis of chloroplast genome sequences of different poplar species. This study is of great significance for the determination and study of poplar chloroplast genome sequence, as well as for the correct understanding of chloroplast DNA structure, origin, species affinity, species diversity and chloroplast genetic engineering [8].

1.3 Transcriptome sequencing technology and application

The RNA-Seq technology can detect the overall transcriptional activity of a specific species at the level of mononuclear acid, so as to comprehensively and quickly obtain almost all transcript information of the species in a certain state [9].

RNA-Seq technology can help us to organize and analyze the specific genes and differentially expressed genes of different tissues of poplar.

At present, the use of RNA-Seq technology for scientific research has achieved good results in many aspects. For example, first, Liu Pengpeng et al. conducted metagenomic sequencing analysis on the types of endophytic bacteria in Astragalus that were not detected by traditional culture and non-culture methods through the second-generation high-throughput sequencing technology, explored the diversity of endophytic bacteria in Astragalus, and used sequencing technology to help the culture of Astragalus [9]. Second, based on high-throughput sequencing technology, Zhang Jefeng et al. analyzed the leaf transcriptomics and hormone levels of a certain variety of rice stem at different growth stages [10]. Third, Wang Rui et al. screened candidate genes related to low nitrogen stress in sorghum based on RNA-Seq technology, and completed analysis of transcriptional changes of sorghum in response to low nitrogen stress, laying a foundation for further exploration of excellent genes for improving nitrogen use efficiency [11]. Fourthly, by analyzing the extracted wheat anther total RNA, Wu Qian et al. verified the changes of related differentially expressed genes in anther development using high-throughput sequencing technology, laying a foundation for further revealing the mechanism of chemical androgenicide CH1 inducing wheat male sterlity [12].

These studies show that RNA-Seq technology is highly efficient and practical in plant transcriptome analysis, and the use of RNA-Seq technology can help us analyze the metabolism, genetics and other functions of poplar. In general, high-throughput sequencing technology has the characteristics of low cost, ultra-high throughput, simple process, high sensitivity and precision, which makes its application in the study of plant microbial diversity has advantages. Therefore, high-throughput sequencing technology is an effective technical means to recognize plants and microorganisms [13].

1.4 Purpose and significance of this study

Poplar has strong adaptability, drought resistance, cold resistance, saline-alkali resistance, and is often used to create fast-growing forest, farmland shelter forest, soil and water conservation forest, sand fixing forest, bank protection forest and urban and rural greening. Transcriptome sequencing and bioinformatic analysis were used to determine the gene expression in different tissues of poplar. By analyzing the gene expression profile of poplar, we can identify genes involved in key biological processes such as photosynthesis and material transport, providing important clues for understanding the development and function of poplar. Through analyzing the transcriptome data of different tissues of Populus wilsonii officinalis, the gene expression of different tissues of Populus wilsonii officinalis was discussed, so as to reveal the molecular mechanism of the development of Populus wilsonii officinalis.

In this experiment, through transcriptome sequencing of different tissues of Populus wilsonii officinalis, specific expression genes of different tissues, differentially expressed genes between different tissues and the expression trend of genes in different tissues were analyzed. Through functional enrichment analysis of corresponding gene sets, factors and signaling pathways with significantly different expression levels in different tissues of Populus wilsonii officinalis were expected to be found. It provides important clues for understanding the growth and development of Populus wilsonii officinalis, which is helpful to increase the cognition of Populus wilsonii officinalis and construct the genetic information database of Populus wilsonii officinalis.

CHAPTER 2
MATERIALS AND METHODS

2.1 Data Acquisition

The raw bioinformatics data for the PRJCA008004 project was obtained from the BIG Submission System (BIG Sub), a dataset containing detailed cleaning data for five different tissues: leaf, bud, phloem, xylem and root.

2.2 Gene quantification

First, we compared the cleaned RNA sequence with the genome sequence. On this basis, Cufflinks (v2.2.1) [20] software was used for gene quantification. Finally, the number of FPKM per transcript was used to calculate the expression of the gene.

2.3 Analysis of specific expression genes

After obtaining the expression levels of genes in different tissues, genes with expression level (FPKM) higher than 1 were taken as effective expression genes, and the specific genes expressed in different tissues were analyzed by Wayne diagram.

2.4 Screening of differentially expressed genes

Through ImageGP [21] (http://www.bic.ac.cn/BIC/#/) DE window analysis analysis, respectively load the chair Yang transcript expression in different tissue matrix, group information, Differentially expressed genes were screened with false discovery rate (FDR) less than 0.05 and Fold change (FC) greater than 2.

2.5 KEGG enrichment analysis

KEGG enrichment analysis was performed with the TBtools program [22], and q-value<0.05 was used as the criterion for significant enrichment.

2.6 Cluster analysis

In order to analyze the expression trend of genes in different tissues of Populus wilsonii officinalis, stem programs [23] were used to perform cluster analysis on buds, leaves, phloem, xylem and roots of Populus wilsonii officinalis, and functional analysis was performed on the gene sets with significant performance.

CHAPTER 3

RESULTS AND DISCUSSION

3.1 Sequencing data and gene expression quantification of different tissues of poplar

The samples for transcriptome analysis were bud site, leaf site, phloem site, root site and xylem site. Clean bases of 7,702,341,298, 7,312,556,460, 10,965,845,302, 7,387,490,240, 10,286,616,988 were obtained from buds, leaves, phloem, roots and xylem, respectively. The GC content of each part was between 44.32% and 46.51%, the Q30 was 0.9471%, 0.9416%, 0.9428%, 0.9425%, 0.9347%, and the coverage was between 77.77% and 95.46%. The sequencing statistics of various tissues of B. chinensis are shown in Table 3-1. The clean bases data were processed using the Baimai cloud platform, and the expression levels of each gene, that is, the FPKM values, were obtained, which provided a basis for further research.

Table 3.1 – **Summary of transcriptome data of different tissues of Populus wilsonii wilsonii**

（sample）	（Clean reads）	（ Clean bases）	（GC Content）	Q30(%)	（ Mapped Reads）	（Mappled Rate）
芽	25,782,466	7,702,341,298	44.7%	0.9471	48,095,328	93.27%
叶	24,515,553	7,312,556,460	44.32%	0.9416	46,805,581	95.46%
韧皮部	36,645,427	10,965,845,302	46.51%	0.9428	56, 998,075	77.77%
根	24,735,027	7,387,490,240	44.47%	0.942	44,133,181	89.21%
木质部	34,560,191	10,286,616,988	45.3%	0.9347	61,182,923	88.52%

3.2 Analysis of specific genes in different sites

High-throughput sequencing technology was used to analyze the differentially expressed key genes in various tissues of A. chinensis. The results showed that 447 genes were specifically expressed in bud, 477 in leaf, and 167 in phloem. There were 963 and 474 genes specifically expressed in root and xylem, respectively. As can be seen from the Wayne map, there are 15,707 unique genes shared by these five sites.

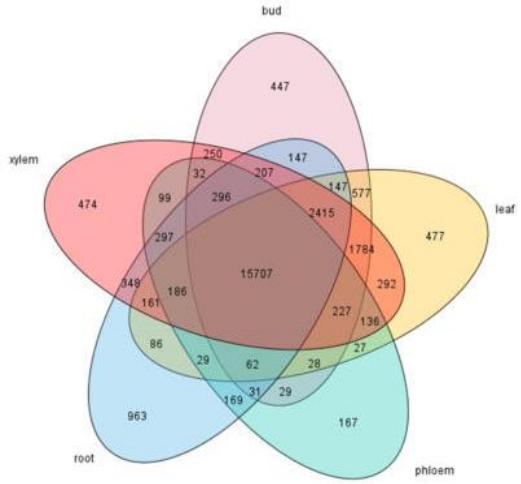


Figure 3.1 – Venn diagram of specific expressed genes

3.2.1 KEGG pathway analysis of specific genes expressed in different tissues of Poplar

The TBtools-II v2.089 program [22] was used to conduct KEGG pathway analysis on the specific genes of different tissues of the screened Populus wilsonii chinensis [1]. Get the following information.

Through the enrichment analysis of KEGG, Table 3-2 shows the KEGG pathway of some bud specific genes. The metabolic pathway with 98 genes was the most enriched pathway, followed by the transcription factor pathway and the transporter pathway with 32 and 31 genes, respectively.

Table 3.2 **KEGG pathway of some specific genes at bud site**

（Term Name）	(GeneHisInSelectedSet)	p-value
Metabolism	98	0.048591197
Transcription factors	32	0.0000201
Transporters	31	0.008297702
Lipid metabolism	26	0.0000737
Biosynthesis of other secondary metabolites	18	0.0100942
Glycerolipid metabolism	8	0.005402558
Glycerophospholipid metabolism	7	0.012287158
Sesquiterpenoid and triterpenoid biosynthesis	6	0.0000586
Homologous recombination	6	0.013533589
Fatty acid elongation	5	0.000678
Stilbenoid, diarylheptanoid and gingerol biosynthesis	5	0.003733415
alpha-Linolenic acid metabolism	5	0.006959966
Flavonoid biosynthesis	5	0.024459922
Tryptophan metabolism	4	0.046033464
Tropane, piperidine and pyridine alkaloid biosynthesis	3	0.043654863
Glycosaminoglycan binding proteins	3	0.048645932

According to the enrichment analysis of KEGG, the pathway with the most enriched genes in leaves is the protein family: metabolism pathway, with 85 genes

enriched, followed by protein kinase pathway and signal transduction pathway, with 44 and 39 genes enriched, respectively.

Table 3.3 **KEGG pathway of some unique genes in leaf parts**

Term Name	GeneHisInSelectedSet	p-value
Protein families: metabolism	85	0.0390233
Protein kinases	44	3.21E-08
Signal transduction	39	1.18E-04
Environmental Information Processing	39	7.25E-04
Organismal Systems	36	1.67E-04
Environmental adaptation	36	1.67E-04
Plant-pathogen interaction	35	3.19E-05
Unclassified: metabolism	31	0.015102539
Enzymes with EC numbers	27	0.029527112
Plant hormone signal transduction	23	0.002838051
MAPK signaling pathway - plant	22	4.12E-04
Lipid metabolism	21	0.008253881
DNA repair and recombination proteins	19	0.023017379
Transcription machinery	16	0.022856889
Starch and sucrose metabolism	15	0.009226001
Homologous recombination	13	1.49E-07
Replication and repair	13	5.76E-04
Metabolism of terpenoids and polyketides	13	0.006658216
Phenylpropanoid biosynthesis	11	0.018210146

The results of KEGG enrichment showed that there were two pathways with the most abundant genes in phloem: the pathway of environmental adaptation and the

pathway of organic system, both with 13 genes enriched, followed by the pathway of plant and pathogen interaction with 11 genes enriched.

Table 3.4 **KEGG pathway of some specific genes in phloem region**

Term Name	GeneHisInSelectedSet	p-value
Environmental adaptation	13	0.003107067
Organismal Systems	13	0.003107067
Plant-pathogen interaction	11	0.008884614
Transcription machinery	7	0.017139573
Metabolism of other amino acids	6	0.010204854
Metabolism of terpenoids and polyketides	5	0.028175219
Cytochrome P450	4	0.00360131
Cyanoamino acid metabolism	4	0.045300791
Lipid biosynthesis proteins	3	0.022010388
Carotenoid biosynthesis	2	0.04744652
Polyketide biosynthesis proteins	1	0.04511686

The results of KEGG enrichment showed that the pathway with the highest concentration of genes in roots was the metabolic pathway with 238 genes, followed by the protein family: metabolic pathway and the protein family: signal transduction and cell process pathway with 176 and 134 genes respectively.

Table 3.5 **KEGG pathway of some specific genes at the root position**

Term Name)	GeneHisInSelectedSet	p-value
Metabolism	238	2.09E-05
Protein families: metabolism	176	0.00345731
Protein families: signaling and cellular processes	134	4.32E-05
Transporters	92	1.44E-11
Environmental Information Processing	80	7.28E-06
Signal transduction	63	0.003055561
Biosynthesis of other secondary metabolites	54	1.28E-09
Transcription factors	54	1.34E-04
Organismal Systems	53	0.025419893
Environmental adaptation	53	0.025419893
Plant-pathogen interaction	52	0.004195674
Metabolism of terpenoids and polyketides	31	4.24E-06
Plant hormone signal transduction	37	0.020179132
Metabolism of other amino acids	36	1.13E-07
Phenylpropanoid biosynthesis	34	1.41E-08
Membrane transport	31	1.39E-06
ABC transporters	17	1.39E-06
Selenocompound metabolism	17	1.84E-10
Galactose metabolism	16	9.11E-04

The results of KEGG enrichment showed that there were three pathways with the highest concentration of genes in the xylem region, namely transcription factor pathway, organic system pathway and environmental adaptation pathway, with 36 genes enriched in each. The next is the plant-pathogen interaction pathway, with 35 genes enriched.

Table 3.6 **KEGG pathway of some xylem specific genes**

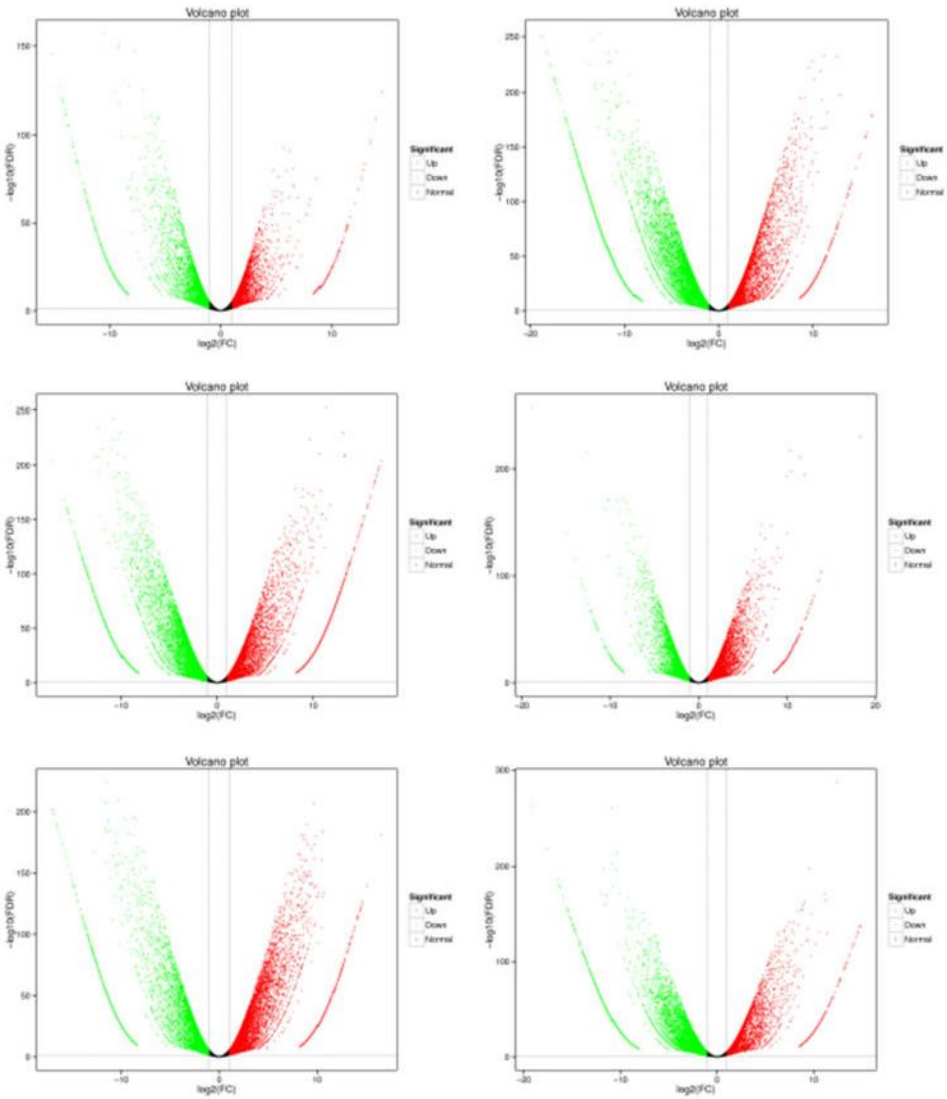
Term Name	GeneHisInSelectedSet	p-value
Transcription factors	36	3.43E-07
Organismal Systems	36	5.58E-05
Environmental adaptation	36	5.58E-05
Plant-pathogen interaction	35	1.01E-05
Enzymes with EC numbers	32	6.88E-04
Transporters	29	0.022020404
Starch and sucrose metabolism	19	1.16E-04
MAPK signaling pathway - plant	18	0.006284151
Chaperones and folding catalysts	14	0.045955046
Glycosylphosphatidylinositol (GPI)-anchored proteins	4	0.047610392

3.3 Results of differential expression gene analysis

3.3.1 Analysis of differentially expressed genes

Leaf, root, phloem of poplar sample; By comparing the gene expression profiles of leaves and different parts of leaves, we found that: (1) there were 8443 differentially expressed genes in leaves and leaf regions, among which 4214 and 4229 differentially expressed genes in leaf regions and leaf regions, respectively. Second, we found 16,276 gene expression changes in different regions of three-leaf leaves and leaves, among which 6465 genes had significant changes in the positions of leaves and leaves, and 6465 and 9811 genes in different positions of leaves and leaves, respectively. Thirdly, there were 14250 differentially expressed genes in leaf root and root, of which 6519 were up-regulated and 7731 were down-regulated. Fourthly, there were 9708 gene expression differences in the position of three-leaf leaves and leaves, among which 4946 and 4762 had significant changes in the position of three-leaf leaves and leaves. Fifth, there were a total of 14,443 gene expression differences at the root and root positions, of which 7610 and 6833 had significant changes at the root and root positions. Sixth, there were 11602 gene

expression differences between the root and root positions, among which 4211 and 7391 genes had significant changes at the root and root positions. Seventh, there are 11,574 genes in root and stem parts, of which 6795 are up-regulated and 4779 are down-regulated. Eighth, a total of 16937 genes were detected at two loci, of which 9893 were up-regulated and 7044 were down-regulated. Ninth, in the cleavage region and the cleavage region, a total of 14,773 genes had significant differences in the cleavage region and the cleavage region, of which 9239 genes had significant changes in the cleavage region and the cleavage region, and 5534 genes had significant changes in the cleavage region. The tenth is 10,909 genes with different expression in XY region and bud region, of which 5,055 are up-regulated and 5,854 are down-regulated.



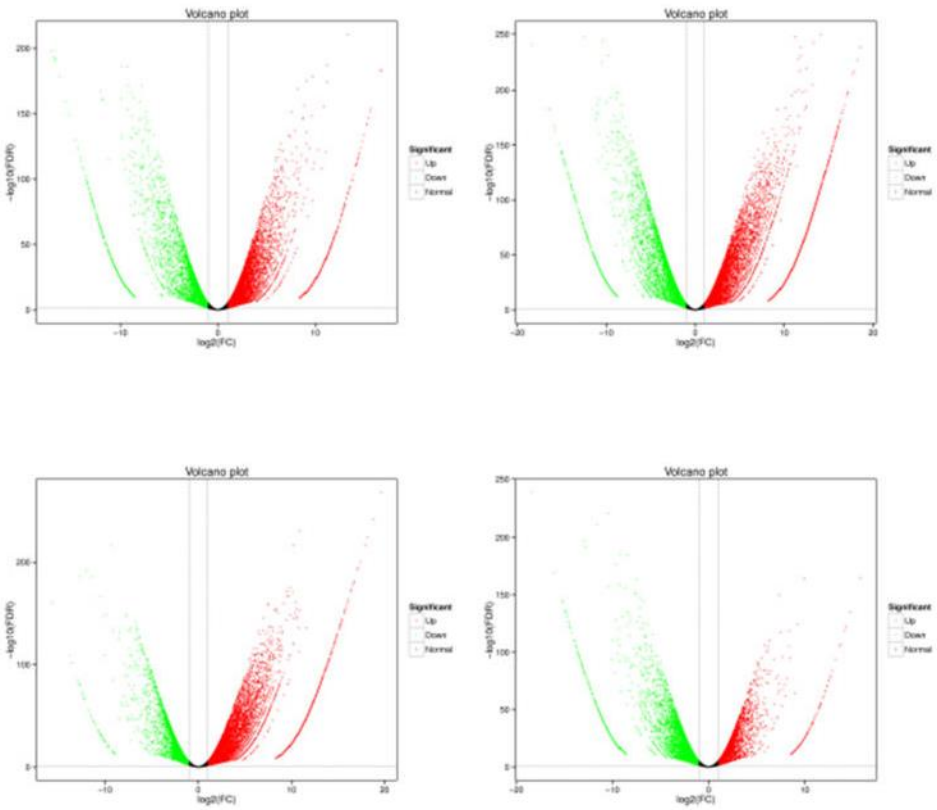


Figure 2 – Volcano map of differentially expressed genes in different parts of poplar

From left to right, from top to bottom, FIG. 1. Volcanoes between leaf and bud FIG. 2. Volcanoes between leaf and phloem FIG. 3. Volcanoes between leaf and root FIG. 4. Volcanoes between leaf and xylem FIG. 5. Volcanoes between root and bud FIG. 6. Volcanoes between root and phloem FIG. 7.root and xylem Volcanoes between phloem and bud Figure 8. Volcanoes between phloem and xylem Figure 9. Volcanoes between xylem and bud Figure 10. Volcanoes between Xylem and Bud

Note: Each point on the phenotypic volcano map represents a different gene, and the horizontal direction is a multiple of the difference in the expression of the same gene in the two samples. The vertical axis is a negative logarithm, which represents a genetically significant difference in the expression of a gene. The greater the absolute value on the horizontal coordinates, the greater the difference in gene expression multiples between the two samples. The larger the ordinate value is, the more obvious the difference is and the more reliable the selected variant genes are.

Among them, the green dots are low-level differentially expressed genes, the red dots are high-level differentially expressed genes, and the dark dots are non-significant genes [16][17].

3.3.2 Analysis of differentially expressed genes in different tissues of poplar

The differentially expressed genes in bud homophyll, phloem, root and xylem of Poplar were mapped (FIG. 3). As shown in the figure, there were 1529 differentially expressed genes between bud and leaf, 1915 differentially expressed genes between bud and phloem, 1681 differentially expressed genes between bud and root, and 1166 differentially expressed genes between bud and xylem. There were 596 differentially expressed genes between bud and different tissues.

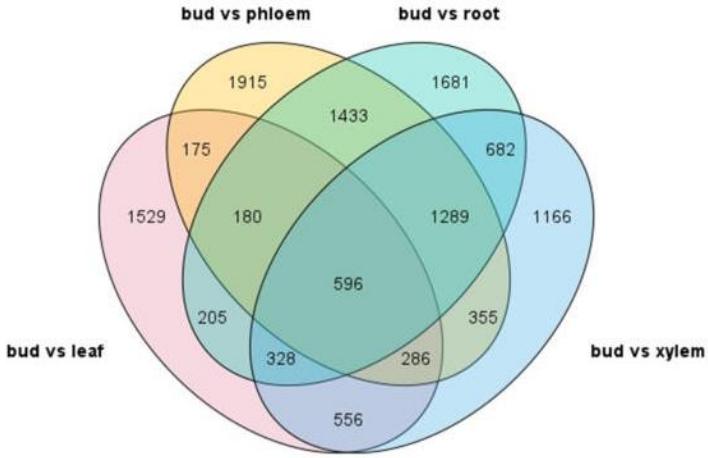


Figure 3 – Venn diagram of differentially expressed genes between bud parts and other parts of Populus wilsonii chinensis

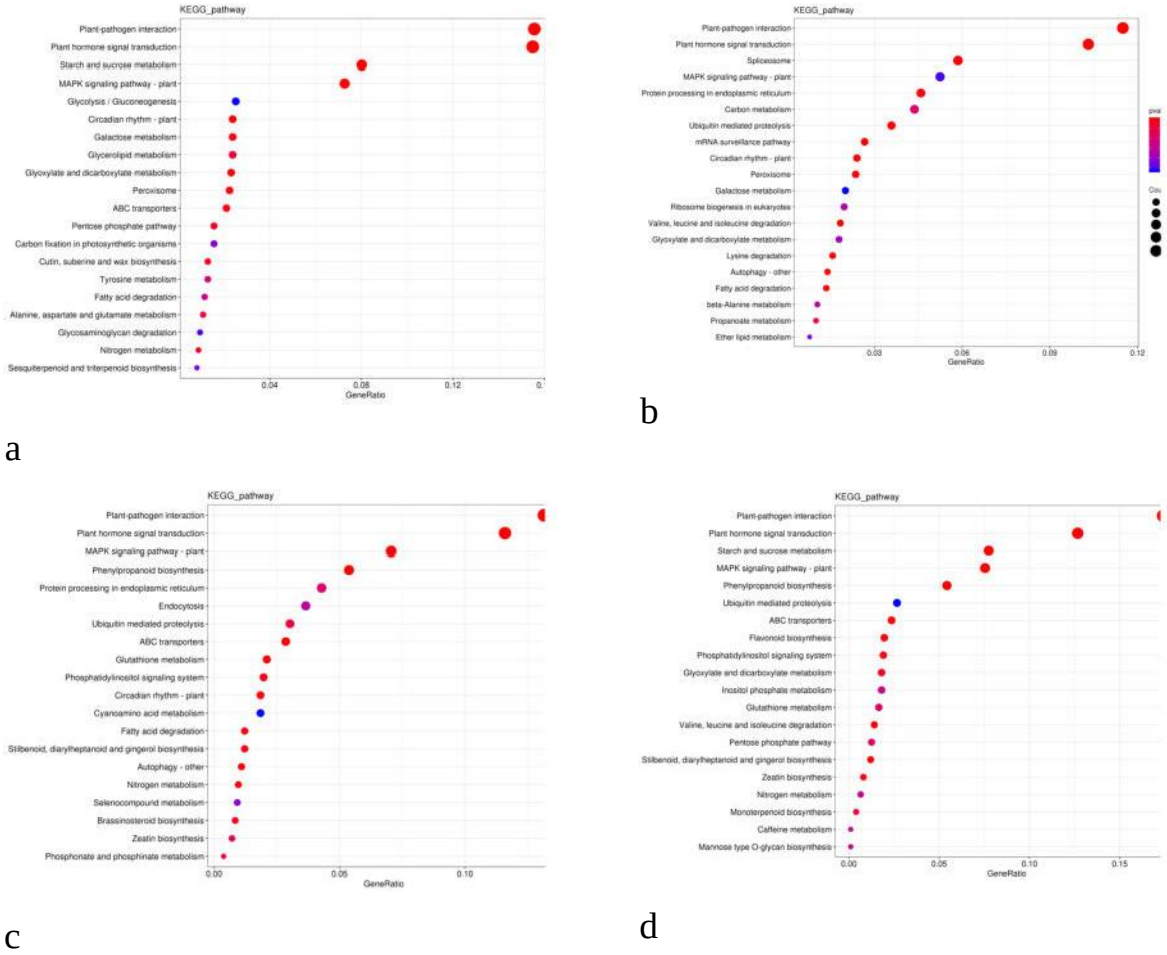


Figure 4 – KEGG scatter plots of differense-expressed genes in poplar buds and different tissues（a is bud and leaf, b is bud and phloem, c is bud and root, d is bud and xylem）

KEGG enrichment analysis was performed on the obtained common differentially expressed genes. According to the KEGG results (Table 3-7), among the differentially expressed genes of all buds and different tissues, a total of 101 genes were involved in the protein family: metabolic pathway, followed by organic system pathway and environmental adaptation pathway, with 63 genes enriched.

Table 3.7 **KEGG pathway of differentially expressed genes in poplar buds and different tissues**

Term Name	GeneHisInSelectedSet)	p-value
Protein families: metabolism	101	0.03268680 3
Organismal Systems	63	8.77E-13
Environmental adaptation	63	8.77E-13
Plant-pathogen interaction	56	1.54E-11
Environmental Information Processing	44	0.00244266 4
Signal transduction	40	0.00404350 4
Plant hormone signal transduction	29	6.95E-04
Peptidases and inhibitors	24	0.01008403 4
Starch and sucrose metabolism	17	0.01366976 2
Cytochrome P450	13	0.00548243 4
Galactose metabolism	10	0.00507265 1
Circadian rhythm - plant	7	0.02012407 5
Cysteine and methionine metabolism	7	0.04515811 2
CD molecules	6	0.01280034 5
Isoquinoline alkaloid biosynthesis	4	0.03940940 8
Ubiquinone and other terpenoid-quinone biosynthesis	4	0.04237779 1
Other types of O-glycan biosynthesis	3	0.02924087 1
Nitrogen metabolism	3	0.04684871 1
Transcription	2	0.01862341 9

The differentially expressed genes among leaves and buds, phloem, roots and xylem of poplar were mapped (FIG. 5). As shown in the figure, there were 793

differentially expressed genes between leaves and buds, 2424 differentially expressed genes between leaves and phloem, 485 differentially expressed genes between leaves and roots, and 268 differentially expressed genes between leaves and xylem. There were 1376 differentially expressed genes between leaves and different tissues.

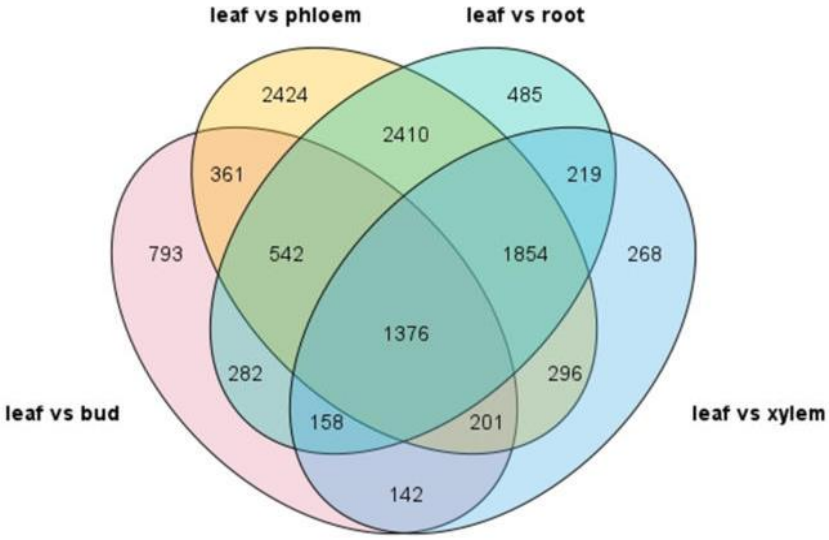
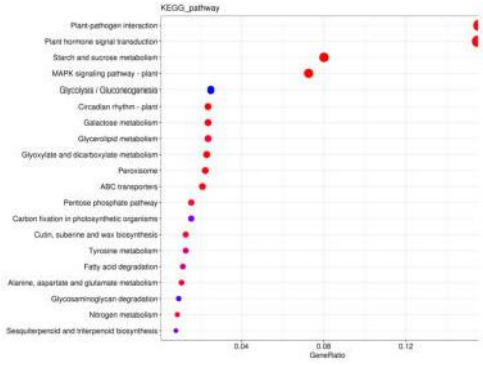
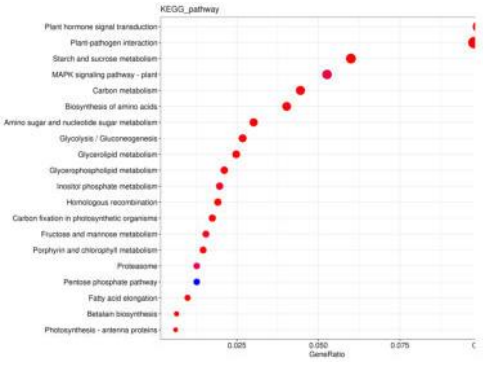


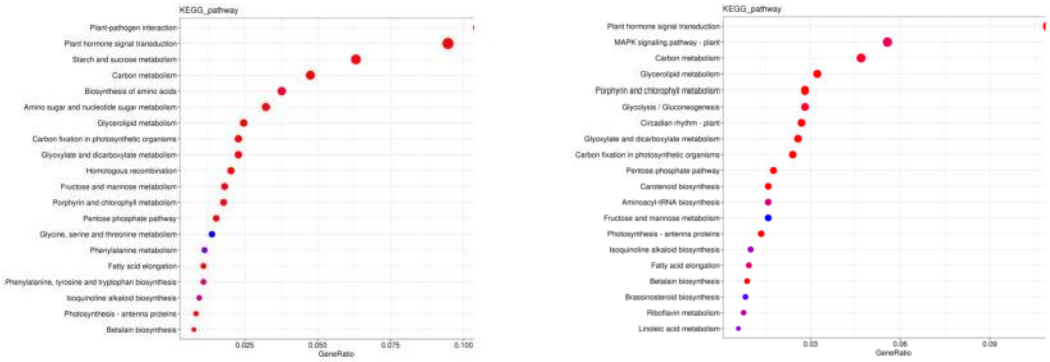
Figure 5 – Venn diagram of differentially expressed genes in leaf parts and other parts of poplar



a



b



c d

Figure 6 – KEGG scatter plots of differense-expressed genes in poplar leaves and different tissues（a is leaf and bud, b is leaf and phloem, c is leaf and root, and d is leaf and xylem.）

KEGG enrichment analysis was performed on the obtained common differense-expressed genes. According to the KEGG results (Table 3-8), among the differense-expressed genes in all leaves and different tissues, 357 genes were involved in metabolic pathways, followed by protein family: metabolism, carbohydrate metabolism and environmental information processing pathways, with 258 enriched genes respectively. 143. 133.

Table 3.8 **KEGG pathway of differentially expressed genes shared by leaf and different tissues**

Term Name	GeneHisInSelectedSet	p-value
Metabolism	357	7.08E-10
Protein families: metabolism	258	1.02E-04
Carbohydrate metabolism	143	1.07E-11
Environmental Information Processing	133	2.83E-13
Signal transduction	122	2.14E-12

Protein kinases	107		1.56E-11
Organismal Systems	101		2.90E-08
Environmental adaptation	101		2.90E-08
Plant-pathogen interaction	90		1.21E-07
Unclassified: metabolism	90		3.64E-04
Plant hormone signal transduction	81		4.50E-11
Enzymes with EC numbers	72	5	0.01748233
Lipid metabolism	63		1.37E-05
MAPK signaling pathway - plant	61		1.48E-07
Glycosyltransferases	43	4	0.02480617
Starch and sucrose metabolism	42		1.39E-04
Biosynthesis of other secondary metabolites	42	6	0.03408196
Metabolism of cofactors and vitamins	34	8	0.00261033
Metabolism of terpenoids and polyketides	32	8	0.00125774

The differentially expressed genes in phloem, bud, leaf, root, and xylem were mapped by Wayne diagram (FIG. 7). As shown in the figure, there were 1971 differentially expressed genes between phloem and bud, 4947 differentially expressed genes between phloem and leaf, 2242 differentially expressed genes between phloem and root, and 825 differentially expressed genes between phloem and xylem. There were no differentially expressed genes between phloem and different tissues.

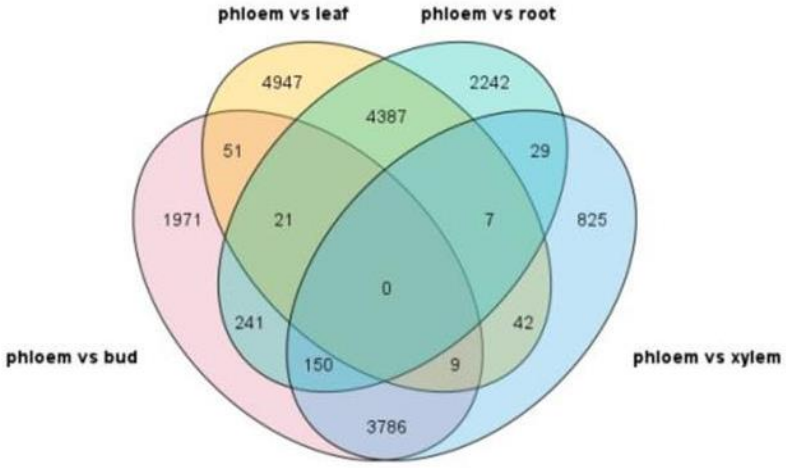


Figure 7 - Venn diagram of differentially expressed genes between phloem and other parts of poplar

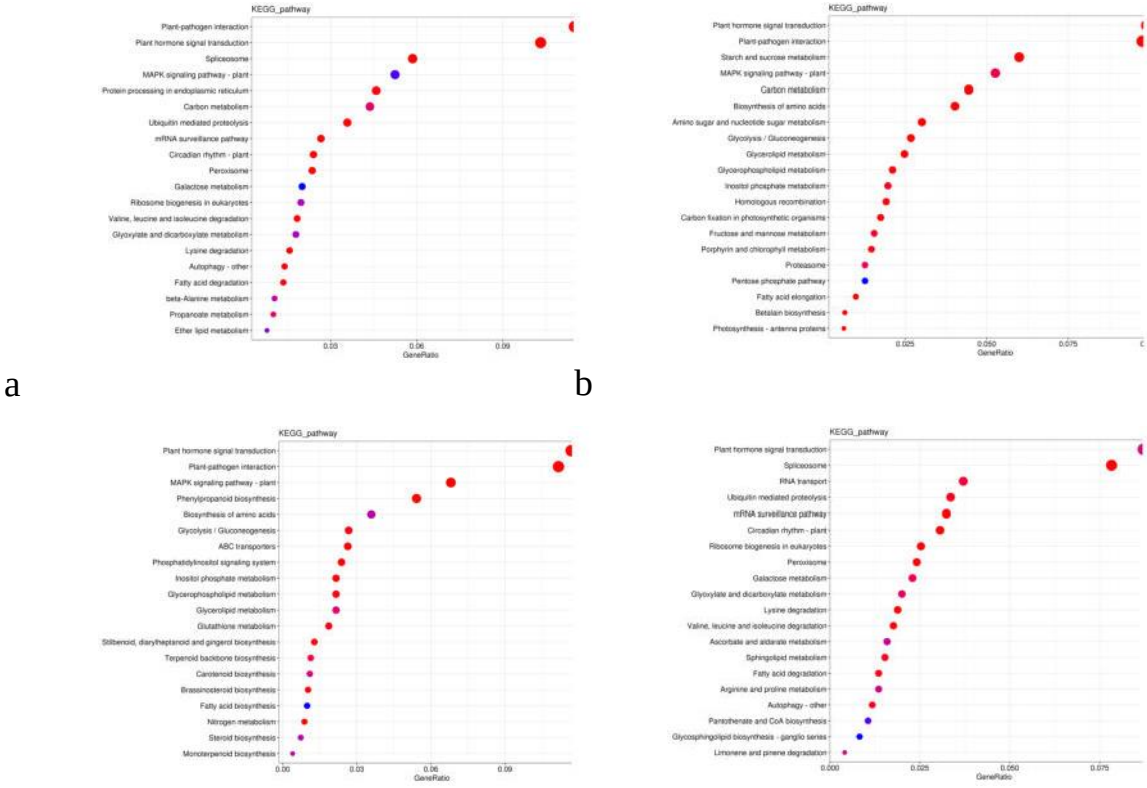


Figure 8 - KEGG scatterplots of differentially expressed genes in phloem and different tissues（a is phloem and bud, b is phloem and leaf, c is phloem and root, and d is phloem and xylem.）

The differentially expressed genes in roots, buds, leaves, phloem, and xylem of

Poplar were mapped (FIG. 9). As shown in the figure, there were 2079 differentially expressed genes between root and bud, 5351 differentially expressed genes between root and leaf, 1930 differentially expressed genes between root and phloem, and 506 differentially expressed genes between root and xylem. There were 21 differentially expressed genes between root and different tissues.

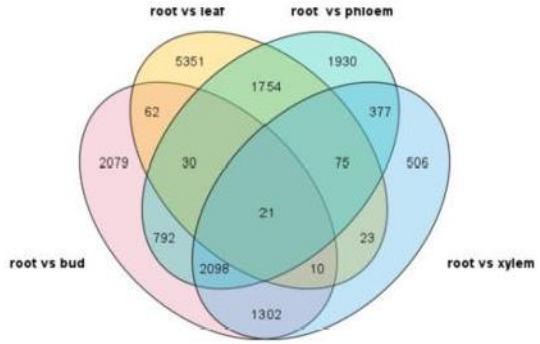
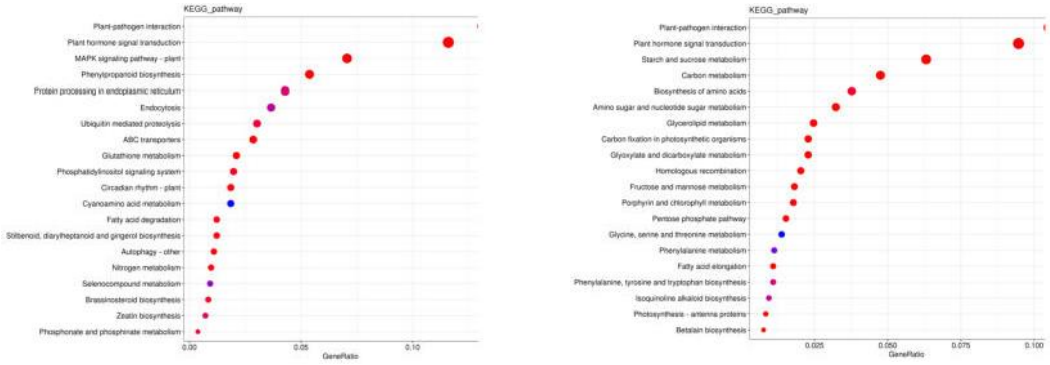
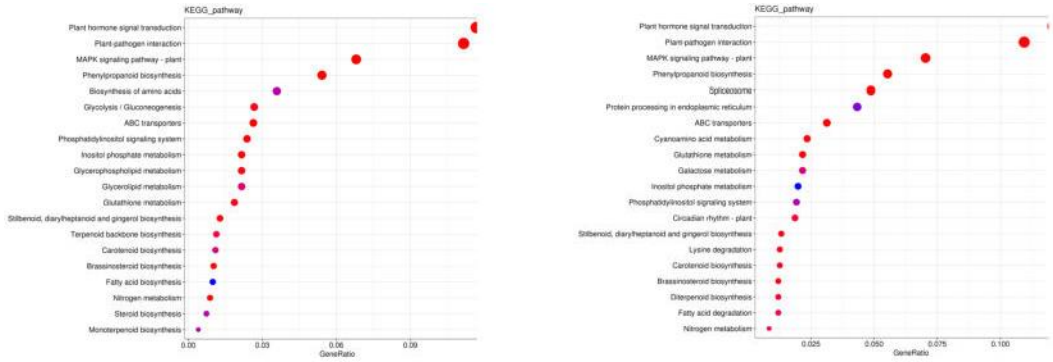


Figure 9 –Venn diagram of differentially expressed genes between the root and other parts of poplar



a b



c d

Figure 10 – KEGG scatterplots of differentexpression genes in poplar root and different tissues (a is the root and bud, b is the root and leaf, c is the root and phloem, d is the root and xylem.）

KEGG enrichment analysis was performed on the obtained common differentially expressed genes. According to the KEGG results (Table 3-9), among the common differentially expressed genes in all roots and different tissues, a total of 9 genes were involved in the metabolic pathway, followed by the carbohydrate metabolic pathway with 5 genes enriched.

Table 3-9 **KEGG pathway of differentially expressed genes in poplar root and different tissues**

Term Name	GeneHisInSelectedSet		p-value
Carbohydrate metabolism	5	9	0.02405904
Metabolism	9	2	0.00999317
Phenylpropanoid biosynthesis	2	5	0.03100559
Zeatin biosynthesis	1	3	0.03759894
Sesquiterpenoid and triterpenoid biosynthesis	1	9	0.03111844
Metabolism of terpenoids and polyketides	2	1	0.03673373

The differentially expressed genes in xylem, bud, leaf, phloem, and root of Poplar were mapped by Wynn diagram (FIG. 11). As shown in the figure, there were 3498 differentially expressed genes between xylem and bud, 3310 differentially expressed genes between xylem and leaf, 2152 differentially expressed genes between xylem and phloem, and 1737 differentially expressed genes between xylem and root. There were 16 differentially expressed genes in xylem and different tissues.

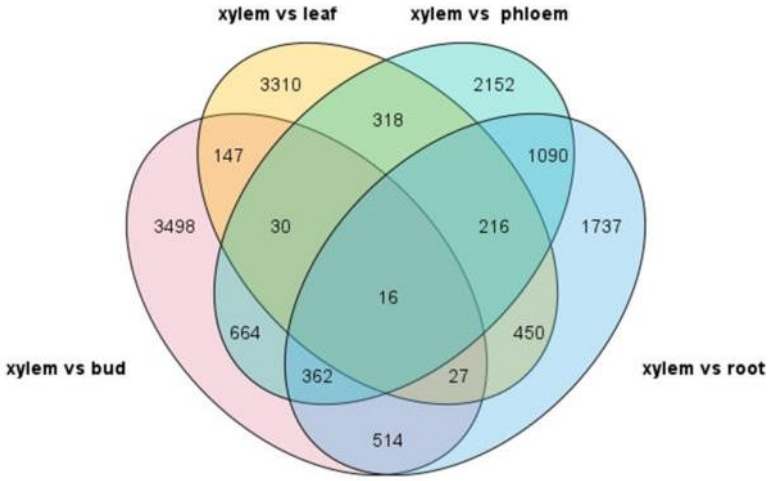
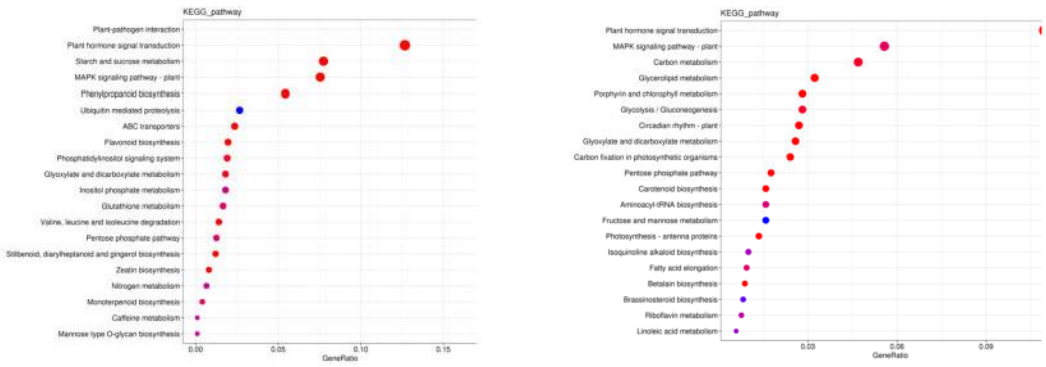
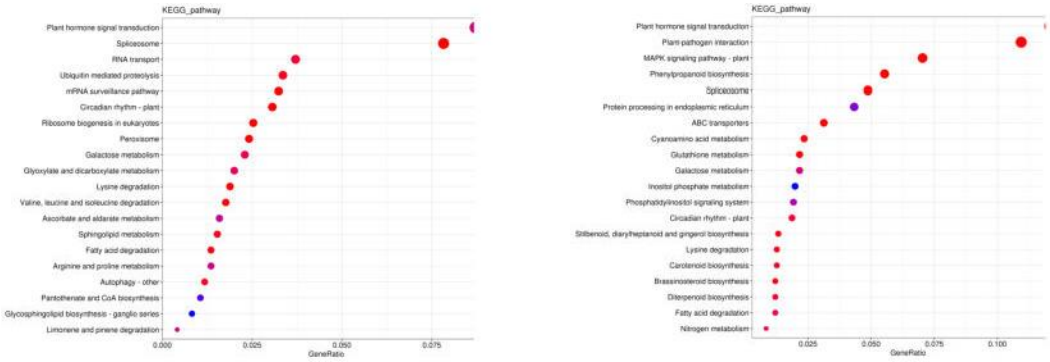


Figure 11– Venn diagram of differentially expressed genes in xylem and other parts of poplar



a b



c d

Figure 12 – KEGG scatter plots of differentexpression genes in xylem and different tissues of poplar（a is xylem and bud, b is xylem and leaf, c is xylem and phloem, and d is xylem and root.）

KEGG enrichment analysis was performed on the obtained common differentially expressed genes. According to the KEGG results (Table 3-10), among all the differentially expressed genes in xylem and different tissues, the protein family: metabolic pathway was the most enriched gene, reaching 903. Followed by environmental information processing, organic systems, and environmental adaptation pathways, 455, 440, and 440 genes were enriched, respectively.

Table 3.10 **KEGG pathway of differentially expressed genes in xylem and different tissues of poplar**

Term Name	GeneHisInSelectedSet	p-value
Protein families: metabolism	903	2.28E-07
Environmental Information Processing	455	2.22E-16
Organismal Systems	440	5.55E-16
Environmental adaptation	440	5.55E-16

Carbohydrate metabolism	413		3.42E-12
Protein kinases	319		8.88E-16
Unclassified: metabolism	304		7.23E-07
Transporters	286		2.05E-04
Enzymes with EC numbers	269		6.45E-06
Ubiquitin system	248		2.72E-05
Transcription factors	239		3.29E-08
MAPK signaling pathway - plant	198		8.88E-16
Biosynthesis of other secondary metabolites	187		2.69E-11
Glycosyltransferases	183		4.17E-10
Starch and sucrose metabolism	150		2.38E-12
Phenylpropanoid biosynthesis	108		6.90E-10
Cytochrome P450	98		4.55E-10
Metabolism of other amino acids	87	3	0.03621689
Metabolism of terpenoids and polyketides	84	1	0.01726530
Amino sugar and nucleotide sugar metabolism	52	3	0.03899162

3.4 Cluster analysis of different tissues of poplar

Through cluster analysis by stem program [23], we obtained 14 gene sets with significant characteristics (Figure 13). According to the change trend, the group with the same change trend is divided into eight categories. The red group contains 5,858 genes, the green group contains 2,769 genes, the purple group contains 2,557 genes, the yellow group contains 1,635 genes, the orange group contains 1,514 genes, the

pink group contains 1,058 genes, and the brown group contains 968 genes. The gray-blue group contains 588 genes.

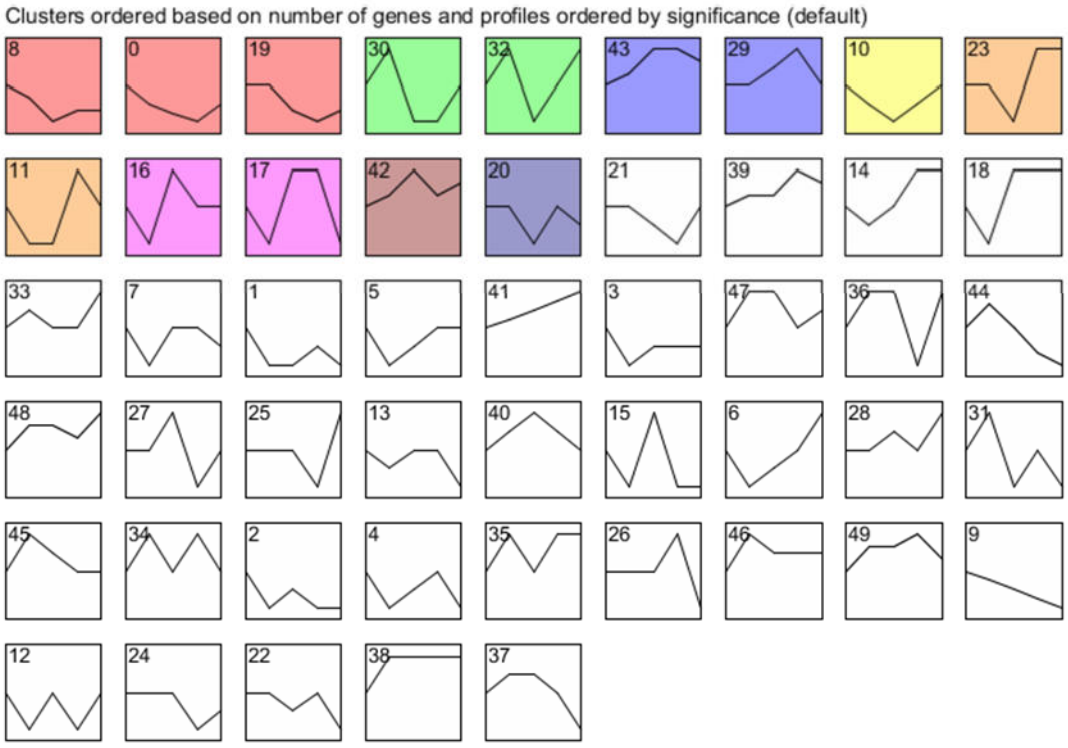


Figure 13. Genome trend of Populus wilsonii Populus wilsonii by cluster analysis

As a native tree species in China, it is very necessary to further study the gene expression of different parts of Populus wilsonii officinalis to reveal the functional differences through transcriptome analysis. The metabolism of poplar is closely related to photosynthesis and respiration in the leaves, and in the roots and buds, it is closely related to the rapid cell division and renewal metabolism in the roots and buds.

CONCLUSIONS

1. It was shown that the specific genes expressed in five different parts of Poplar
2. KEGG results showed that the pathway with the most abundant genes in bud was the metabolic pathway, the pathway with the most abundant genes in leaf was the protein family: metabolic pathway, and the pathway with the most abundant genes in root was also the metabolic pathway.
3. In phloem, the most abundant genes are pathway environmental adaptation pathway and plant pathogen interaction pathway, which are closely related to the protective effect of phloem. All these indicate that the expression of phloem gene is related to the location of phloem.
4. There were 596 differentially expressed genes between bud and other parts, 1376 differentially expressed genes between leaves and different tissues, 21 differentially expressed genes between roots and different tissues, and 16 differentially expressed genes between xylem and different tissues.
5. It was detected with KEGG analysis that 98 differentially expressed genes are involved in the metabolic pathway among the genes shared by buds and different tissues, and 357 genes are involved in the metabolic pathway among the differentially expressed genes between leaves and different tissues, followed by the protein family: metabolic pathway, which further proves that the expression of diosin is related to the location.

LIST OF REFERENCES

1. Shang Ce. Revision of the classification of Populus wilsonii in Pan-Himalayan region [D]. Beijing Forestry University,2017.
2. Li Ziyin, Chen Shouyi. Advances in functional genomics of plants [J]. Genetics,2000,(01).
3. Shi Jisen, Wang Zhanjun, Chen Jinhui. Advances in whole genome sequencing of woody plants [J]. Genetics,2012,34(02):145-156. (in Chinese)
4. Xu Meng, Pan Huixin, Zhang Bo, et al. Research progress of molecular biology in genetic improvement of forest trees [J]. Scientia Silviculina,2009,45(01):136-143. (in Chinese)
5. Cheng Qiang, Pan Huixin, Xu Lian, et al. Research progress of poplar genome project and molecular biology [J]. Journal of Nanjing Forestry University (Natural Science Edition),2009,33(01):131-135.
6. Breeding of transgenic European black poplar with insect resistance [J]. Chinese Journal of Bioengineering,1993,9(04).
7. Wang Lijun. Study on the mechanism of regulation of flavonoid and lignin biosynthesis by transcription factor MYB6 in poplar [D]. Southwest University,2018...
8. Zong Dan. Phylogenetic study of Populus wilsonii based on nucleoplasmic genome [D]. Southwest Forestry University.
9. Han Xia, Fang Jiali, Meng Shenkun 'er, et al. Application of high-throughput sequencing technology in the study of microbial diversity in medicinal plants [J]. Zhejiang Academy of Agricultural Sciences.

10. Zhang Jefeng, Li Fang, Zhong Haiying, et al. Water stem "Zhenong No. 7"; Analysis of leaf transcriptomics and hormone levels at different growth stages [J]. Journal of Agricultural Biotechnology, 2019,32(02).
11. Wang Rui, Zhang Fuyao, Zhan Pengjie, et al. Screening of candidate genes related to low nitrogen stress in sorghum based on RNA-Seq [J]. Acta Agricultura Sinica, 2019,50(03):669-685. (in Chinese)
12. Wu Qian, Guo Haoyu, Xi Tiantian, et al. Transcriptomics and related gene expression analysis of wheat male sterile anthers induced by chemical androgenicide CH1 [J]. Journal of Wheat Crops, 2018,44(03):289-298.
13. Cui Kai, Wu Weiwei, Diao Qiyu. Advances in research and application of transcriptome sequencing technology [J]. Bulletin of Biotechnology,2019,35(07).
14. Zhang Shengkui. Comprehensive evaluation of cassava germplasm resources and genome-wide association analysis of main agronomic traits [D]. Huazhong Agricultural University,2018.
15. Zhang Yong, Zhang Shouong, Qi Liwang, et al. Poplar: a model species for forest genomics research [J]. Bulletin of Botany,2006,(03):286-293.
16. Hao Guixia, Zhu Zhen, Zhu Zhiti. Progress in genetic engineering of poplar [J]. Advances in Bioengineering,2000,(04).
17. Rao Hong-yu, Huang Min-Ren. Current situation and prospect of genetic engineering of poplar [J]. Forestry Science and Technology Development,1999,(04).
18. Xu Nong, Huang Minren, Chen Daoming. Yang Shujiin Engineering Research Progress [J]. Journal of Nanjing Forestry University (Natural Science Edition),1993,(04).

19. Shi Wenjun, Wang Pan, Hu Runhao, et al. Analysis on the trend of differentially expressed genes in gill tissues of Alopecia alopecta under hypoxic-reoxygenation stress [J]. Journal of Southern Agricultural Sciences, 2019,53(03):735-747.
20. Xiaolong Cao, Haobo Jiang,Integrated modeling of protein-coding genes in the Manduca sexta genome using RNA-Seq data from the biochemical model insect,Insect Biochemistry and Molecular Biology,Volume 62,2015.
21. Haoting Ye, Qiuxiao Lin, Qinghua Mei, Qiuqiong Liu, Siwei Cao,Study on mechanism of transdermal administration of eugenol for pain treatment by network pharmacology and Molecular docking technology, Heliyon, Volume 10, 8202 4 Issue.
22. Chen Chengjie, Xia Rui. Domestic biological software in the era of TBtools Big data [J]. Scientific Observation,202,17(06).
23. Xiang Guolong. Study on preparation of black bone sheep induced pluripotent stem cells by reproducing 10 transcription factors [D]. Inner Mongolia Agricultural University,2023.