Plant Omics Journal POJ

POJ 7(4):244-252 (2014) ISSN:1836-3644

Comparison and analysis of transcriptome profiles of haploid and diploid *Populus* by digital gene expression

Xuliang Mao[1], Xuemei Xu[2], Haizhen Zhang[1], Jingli Yang[1,3*], Zhanchao Wang[1], Chenghao Li[1,*]

[1]State Key Laboratory of Forest Genetics and Tree Breeding, Northeast Forestry University, 26 Hexing Road, Harbin 150040, China
[2]Library of Northeast Forestry University, 26 Hexing Road, Harbin 150040, China

***Corresponding author:85831647@qq.com; chli0@163.com**

Abstract

Haploid (HP) plants have been used for practical and basic research for many years. The development of haploid plants of *Populus simonii Carr.* × *P. nigra* L. was inhibited compared with the wild-type (diploid). In order to understand the molecular mechanisms and the genes expression characteristics, we performed gene expression profiling of the shoot tips and leaves of haploid and diploid plants using high-throughput tag-sequencing (Tag-seq) analysis, based on an Illumina Digital Gene Expression (DGE) platform. Approximately, 3.0–3.4 million total clean sequence tags with 0.09–0.1 million distinct clean tag sequences were obtained for each library. The clean tags were mapped to the reference sequences for annotation of expressed genes. About 53.93–56.64% of the distinct clean tags were mapped to the reference genes. Differentially expressed genes between haploid shoot tip library and diploid shoot tip library and differentially expressed genes between haploid leaf library and diploid leaf library were identified (4464 and 3856, respectively). The differentially expressed genes were annotated using Gene Ontology and the Kyoto Encyclopedia of Genes and Genomes orthology identifiers, which revealed that the differentially expressed genes were mainly annotated to cell compartments, especially the plastid and chloroplast, and were enriched in binding and catalytic activity and metabolic process. Orthology analysis revealed that the differentially expressed genes were mainly involved in plant hormone signaling and biosynthesis, and were generally downregulated, such as the cytokinins isopentenyl transferase. Some genes related to the cell cycle were differentially expressed in the haploid. We concluded that the reduced level of hormones in the haploid plants mainly contribute to the inhibited development phenotype by changing the expression of genes involved in the cell cycle.

Keywords: Gene expression; haploid; leaf; Poplar; shoot tip.
Abbreviations: CK_cytokinin; DGE_Digital Gene Expression; DEGs_differentially expressed genes; DH_doubled haploid; GO_Gene Ontology; GA_gibberellin; HP_haploid; IAA_indole-3-acetic acid; KEGG_Kyoto Encyclopedia of Genes and Genomes.

Introduction

Poplar has become a widely planted forest tree because of its rapid growth, many natural and artificial interspecific hybrids, and its environmental and economic importance. With the completion of the *Populus trichocarpa* genome sequence, a wide range of genomic and genetic resources is now available for this species (Tuskan et al., 2006). Thus, poplar has been selected as a model tree for biochemical, genetic and genomic studies (Du et al., 2012; Jansson and Douglas, 2007). Economically, its development as a wood plant has become one of the most pivotal issues. Haploid (HP) plants have been used for practical and basic research for many years. HPs have been used in breeding, mutagenesis and transformation, and functional studies of genomic regions and genes related to microspore embryogenesis; thus, HPs have played roles in genetic and genomic research (Szarejko and Forster, 2007). There are several available methods to obtain HPs, of which anther or isolated microspore culture is the most effective and widely used *in vitro*. Many publications have reported advances in anther culture (Dunwell, 2010). Through anther culture, we obtained HPs, which were cultured under the same conditions as the diploids (DPs). Interestingly, the HP plants showed striking developmental alterations. The most noticeable changes in the shoot were a severely inhibited development, with shorter internodes, leading to a dwarfed growth habit. In contrast to the wild-type (DPs), leaf expansion was reduced and the surface size of leaves was smaller than DPs. The growth of plants depends on continuous function of the meristems. Shoot meristems are responsible for all the post-embryonic aerial organs, such as leaves, stems and flowers. Thus, in our study, shoot tips and leaves were obtained to analyze the dynamics of transcription. In recent years, next-generation sequencing (NGS) technologies, such as 454 and Illumina platforms, have been widely used in gene sequencing. Novel applications in biology and medicine are becoming a reality, extending beyond genomic sequencing (Ansorge, 2009). This technology can rapidly produce huge numbers of short sequencing reads, making it possible to analyze a complex sample containing a large amount of nucleic acids by simultaneously sequencing the contents of the entire sample. A number of studies have used next-generation sequencing technologies for genome-scale expression analyses in higher eukaryotes. Illumina's Digital Gene Expression (DGE) platform can generate, at its current capacity, 90 to 100 million reads per run of an eight-lane flow cell. It can reduce library saturation from abundant transcripts and enhance the capacity for rare transcript detection (Morrissy et al., 2009; Babbitt et al., 2010). This technique has been used in plant

Table 1. Statistics of DGE sequencing.

Category	Parameter	Value for *Populus* library	
		DP shoot tip libraries/ HP shoot tip libraries	DP leaf libraries/ HP leaf libraries
Raw tag	Total no. of tags	3475895/3535416	3477807/3252728
	No. of distinct tags	232489/239498	246055/219488
Clean tag	Total number	3346386/3413641	3306528/3056391
	Distinct Tag number	103785/118400	109580/95509
All Tag Mapping to Gene	Distinct Tag number	58780/63856	57648/53338
	Distinct Tag % of clean tag	56.64%/53.93%	52.61%/55.85%
All Tag-mapped Genes	Number	26528/26377	25119/25295
	% of ref. genes	57.95%/57.62%	54.87%/55.26%
Unambiguous Tag-mapped Genes	number	17898/17877	16846/17153
	% of ref genes	39.1%/39.05%	36.8%/37.47%

Raw tag: Sequence data transformed from sequencing-received raw image data.
Clean tag: Tags after filtering dirty tags from raw data

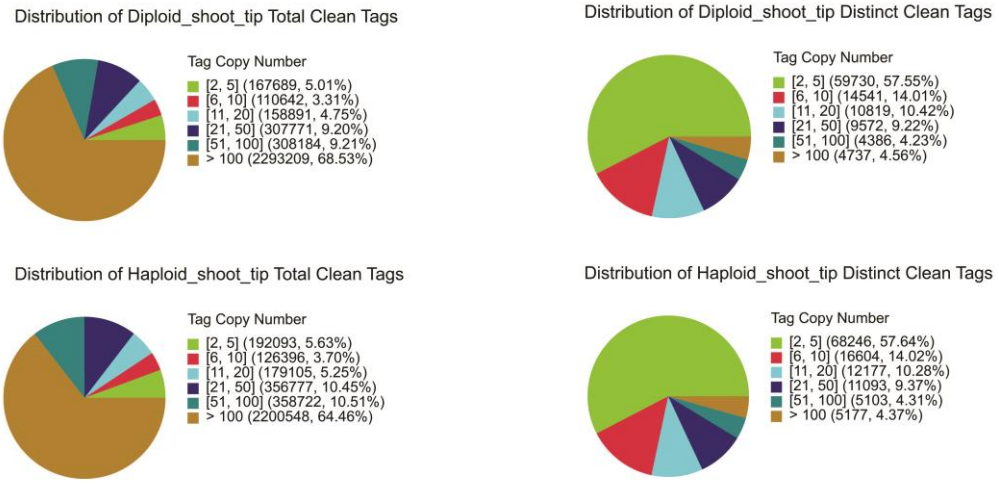


Fig 1. Distribution of clean tag copy number. Total clean tags represent the sum of all clean tag number. Distinct clean tags represent all types of clean tags. Numbers in square brackets indicate the range of copy numbers for a specific category of tags. Numbers in parentheses show the total tag copy number and percentage for all the tags in that category.

research. Transcriptome profiling of early developing cotton fiber by deep sequencing revealed significant differential gene expression in a fuzzless/lintless mutant (Wang et al., 2010). The analysis of short-read expression data of digital gene expression profiles in maize resolved specific expression signatures, which will help define mechanisms of action for the RA3 gene (Eveland et al., 2010). The present study represents the first genome-wide analysis of gene expression in *P. simonii Carr.* × *P. nigra* L. Using the Solexa sequencing system, the transcriptomes of shoot tips and leaves were compared between the haploid and the diploid. By investigating the expressions of genes related to plant development, differentially expressed genes involved in development in the shoot tips and leaves were identified. The purpose of the current study was to gain an insight into the molecular mechanisms underlying the differences in development between wild-type and HP poplars.

Results

Basic quantitative parameters of Digital Gene Expression library sequencing

We sequenced four DGE libraries: DP shoot tip *vs*. HP shoot tip and DP leaf *vs*. HP leaf, which generated approximately 3.2–3.5 million raw tags for each sample (Table 1). After removal of low-quality reads, the total number of tags per library ranged from 3.0 to 3.4 million. Heterogeneity and redundancy are two significant characteristics of mRNA expression. A small number of categories of mRNA have very high abundance, while the majority may have a low level of expression. As shown in Fig. 1, the distribution of total tags and distinct tags over different tag abundance categories showed similar patterns for both DGE libraries, suggesting that they had a similar distribution.

Screening of differentially expressed genes (DEGs)

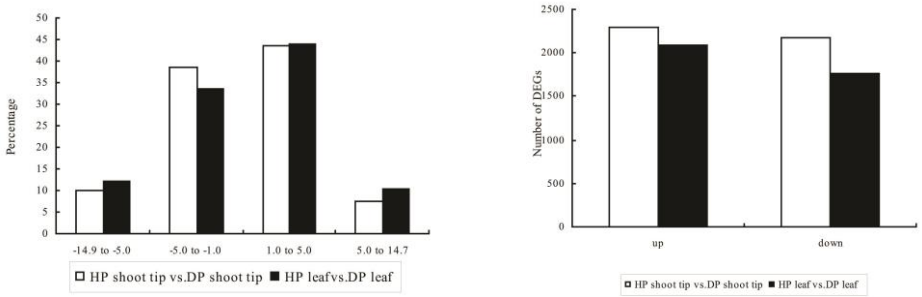
To analyze the altered phenotype, a rigorous algorithm (Audic and Claverie, 1997) was developed to identify DEGs between two samples. In the shoot tip libraries of HP and DP plants, we identified 4464 DEGs. Among them, 2294 genes were upregulated and 2170 were downregulated. The maximum value of $\log_2$ Ratio (HP shoot tip/DP shoot tip) was up to 14, and the minimum was 1; more than 80% of the genes (3680) were up- or downregulated between 1.0- and 5.0-fold (Fig. 2). The DEGs identified between the leaf libraries are listed in Fig. 2.

Gene Ontology categorization of DEGs

The results showed that DEGs in the shoot tip and leaf libraries had a similar Gene Ontology (GO) enrichment. In the ontology of the cellular component, the most enriched terms were cell, cell part, and membrane and organelle.

Table 2. List of pathway enriched differentially expressed genes in shoot tip libraries and leaf libraries of HP and DP plants.

Pathway	The number of DEGs	
	HP shoot tip *vs*. DH shoot tip	DH leaf *vs*. HP leaf
Metabolic pathways	680 (26.16%)	675 (29.48%)
Biosynthesis of secondary metabolites	425 (16.35%)	419 (18.3%)
Plant hormone signal transduction	171 (6.58%)	157 (6.86%)
Plant-pathogen interaction	160 (6.16%)	158 (6.9%)
Ribosome	98 (3.77%)	60 (2.62%)
Protein processing in endoplasmic reticulum	83 (3.19%)	71 (3.1%)
Phenylpropanoid biosynthesis	74 (2.85%)	63 (2.75%)
Splicedome	67 (2.58%)	48 (2.1%)
Starch and sucrose metabolism	64 (2.46%)	79 (3.45%)
Ubiquitin mediated proteolysis	13 (0.5%)	42 (1.83%)
RNA transport	58 (2.23%)	30 (1.31%)
Purine metabolism	55 (2.12%)	40 (1.75%)
Amino sugar and nucleotide sugar metabolism	50 (1.92%)	60 (2.62%)
Glycolysis/Gluconeogenesis	44 (1.69%)	55 (2.4%)
Pyrimidine metabolism	42 (1.62%)	29 (1.27%)
Stilbenoid, diarulheptanoid and gingerol biosynthesis	40 (1.54%)	39 (1.7%)
Oxidative phosphorylation	39 (1.5%)	36 (1.57%)
Endocytosis	38 (1.46%)	31 (1.35%)



(A) (B)

Fig 2. Analysis of differentially expressed genes in the shoot tip libraries and leaf libraries of HP and DP. (A) Summary of the number of differentially expressed genes in the HP shoot tip *vs.* DP shoot tip and HP leaf *vs*. DP leaf. The threshold was "FDR＜ 0.001 and the absolute value of $\log_2$ Ratio ≥1". (B) Fold change distribution of differentially expressed genes.

Interestingly, there were few DEGs annotated to the mitochondria, Golgi apparatus and endoplasmic reticulum categories, while relatively large numbers of DEGs were enriched in plastid and chloroplast categories (Supplemental Table 1). In particular, in the leaf libraries, all the DEGs involved in the photosystem were downregulated (data not shown). In the ontology of molecular function, most of the DEGs were enriched in binding and catalytic activity categories. In the ontology of biological process, large numbers of DEGs were enriched in cellular component organization or biogenesis, metabolic process and response to stimulus categories (Fig. 3).

Pathway analysis

There were 2599 DEGs in the shoot tip libraries and 2290 DEGs in the leaf libraries that had pathway annotation. Pathway analysis revealed that they were involved in 125 pathways and 124 pathways, respectively. Of the annotated DEGs in the shoot tip libraries and leaf libraries, 16.35% and 18.3% were involved in the biosynthesis of secondary metabolites, and 26.16% and 29.48% were involved in metabolic pathways, and the plant hormone signal transduction pathways were represented by 6.58% and 6.86%, respectively. Table 2 shows the pathways involving large numbers of DEGs.

DEGs related to hormone synthesis

We identified genes whose encoded proteins are involved in hormone synthesis pathways in the KEGG database. Genes in the auxin synthesis pathway were mainly downregulated, especially in the leaf libraries. In the cytokinin synthesis pathways, such as the zeatin biosynthesis pathway, the gene encoding isopentenyl transferase (IPT), which is the rating-limiting enzyme, was downregulated. Genes related to gibberellins (GAs) synthesis were generally downregulated, especially in the leaves (Table 3).

Plant hormone signaling

As shown in Table 2, genes involved in hormone signaling were among the most abundant DEGs. In the shoot tip libraries, the genes involved in auxin signaling were all differentially expressed. These genes included *AUXIN-RESISTANT1* (*AUX1*), which is a type of auxin influx carrier; *TIR1*, which is an F box protein that functions as an auxin receptor; AUX/IAA, which is a rapidly induced auxin response gene; *auxin response factors* (*ARF*); *Gretchen Hagen3* (*GH3*); and *SMALL AUXIN UP RNA* (*SAUR*). Of these genes, only three were upregulated; the others were significantly downregulated, including *ARF* and *SAUR* (Fig.

4A; Table 4). The cytokinin signaling pathway comprises a

Table 3. Differentially expressed genes involved in hormone biosynthesis.

Hormones	DEGs of HP shoot tip *vs*. DP shoot tip	DEGs of HP leaf *vs*. DP leaf	*Annotation*
Auxin	POPTR_0006s05920.1 (-3.3)	POPTR_0006s26000.1 (-9.8)	*FMO*
		POPTR_0001s36280.1 (-4.5)	
	POPTR_0002s02690.1 (1.8)	POPTR_0002s02690.1 (-4.1)	*CYP71E1*
		POPTR_0002s02800.1 (-3.7)	
	POPTR_0017s13470.1 (-9.1)	POPTR_0011s04810.1 (9.9)	*F4ST_FLACH*
	POPTR_0011s04810.1 (1.9)		
	POPTR_0007s13650.1 (-4.5)	POPTR_0007s13650.1 (-2.8)	*ALDH(NAD+)*
	POPTR_0015s08540.1 (-1.8)	POPTR_0015s08540.1 (-2.5)	
	POPTR_0012s08010.1 (-1.1)	POPTR_0012s08010.1 (-2.1)	
	POPTR_0001s16730.1 (3.9)		
	POPTR_0002s08230.1 (1.5)		
		POPTR_0001s33470.1 (-1.0)	*nitrilase*
Zeatin	POPTR_0008s20650.1 (-2.4)	POPTR_0008s20650.1 (-3.8)	*IPT*
	POPTR_0010s14940.1 (-9.1)	POPTR_0003s19150.1 (-1.1)	*CYP735A*
	POPTR_0002s26640.1 (2.0)	POPTR_0008s02630.1 (8.8)	
	POPTR_0014s04250.1 (8.6)	POPTR_0019s09260.1 (1.6)	
GA	POPTR_0004s15320.1 (-9.4)	POPTR_0015s14030.1 (-10.7)	*GA 20-ox*
	POPTR_0012s00660.1 (-1.2)	POPTR_0004s15320.1 (-9.4)	
	POPTR_0015s00510.1 (1.1)	POPTR_0015s00510.1 (-1.4)	
	POPTR_0001s28480.1 (2.5)	POPTR_0001s45960.1 (1.0)	*giberrellin-20*
	POPTR_0001s45960.1 (2.5)		

The value in the parentheses presents the value of $\log_2$ Ratio (HP_shoot_tip/DP_shoot_tip) and the value of $\log_2$ Ratio(HP_leaf/DP_leaf), respectively.

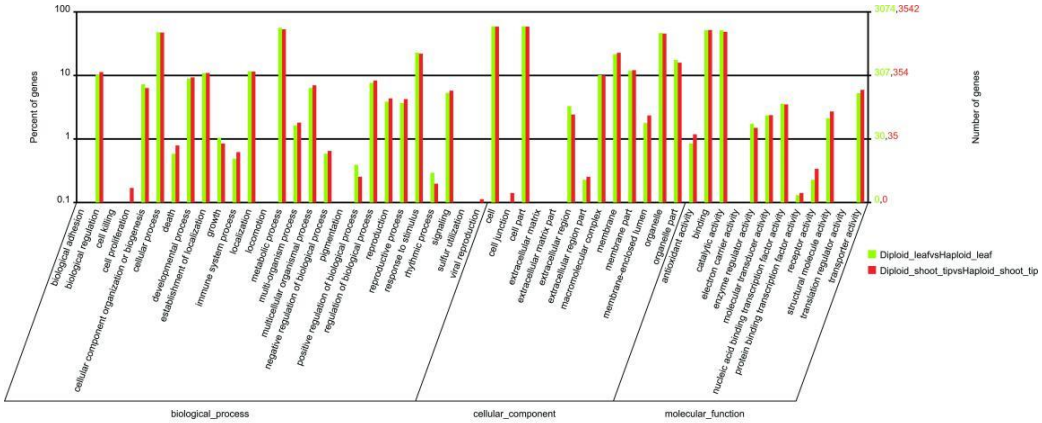


Fig 3. Comparison of differentially expressed genes with GO annotation. Annotated genes were classified into three ontologies (molecular function, cellular component and biological process) and 52 subgroups.

phosphorelay mechanism that is initiated by binding of cytokinin to histidine kinase receptors. The genes encoding the hybrid histidine kinase receptor, *CRE1*, which binds to cytokinin and induces it to autophosphorylate, were upregulated and downregulated. Two genes encoding *histidine phosphotransfer proteins* (*AHPs*) were slightly upregulated and downregulated, respectively. In the nucleus, both type-B and type-A response regulators (ARRs) were differentially expressed. Genes encoding type-B ARRs were up- and downregulated; the gene was more upregulated than it was downregulated. However, the genes encoding type-A ARRs were all downregulated (Table 5). By contrast, DEGs in the leaf libraries involved in auxin signaling were generally upregulated (Fig. 4B; Table 4). The DEGs associated with cytokinins showed the same change trend in the leaf libraries as they did in the shoot tip libraries (Table 5). GA levels are related directly to stem growth and are required for stem elongation (Xu et al., 1997). In the GA signaling pathway, genes encoding the GA receptor *GIBBERELLIN-INSENSITIVE DWARF1* (*GID1*) were mainly downregulated. No significant difference in expression of the gene encoding GID2 was observed. Many genes encoding DELLA were differentially expressed, and the fold-changes of the upregulated genes were larger in both the shoot tip libraries and in the leaf libraries (Table 6). For DEGs involved in the cell cycle, the genes in the shoot tip libraries were generally downregulated and those in the leaf libraries were generally upregulated (Table 7).

Discussion

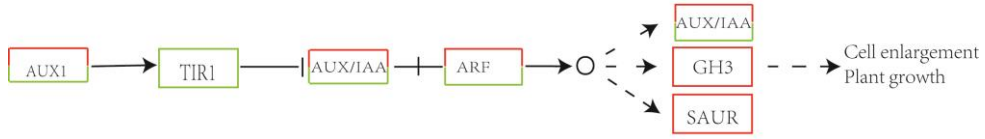
For many years, trees in natural forests have shown long generation times and heterogeneous genetics. Since then, the influence of HP induction on forest breeding and the theoretical study of genetics and development have become significant. Although HPs have been used in breeding for many years, there have been no reports of DEGs between HPs and DPs. In the present study, we compared the gene expression between HP and DP shoot tips of *P. simoniiv Carr.* × *P. nigra* L. using DGE. We identified DEGs involved in developmental processes and provided explanations for the observed phenotypic differences and provided some useful

Table 4. Differentially expressed genes involved in signal transduction of IAA.

Gene name	DEGs of HP shoot tip *vs.* DP shoot tip	DEGs of HP leaf *vs.* DP leaf
AUX1	POPTR_0006s09940.1 (-2.4)	POPTR_0016s12100.1 (-3.2)
	POPTR_0016s12100.1 (-3.0)	POPTR_0009s15200.1 (-1.5)
	POPTR_0001s10430.1 (3.6)	POPTR_0010s19840.1 (2.9)
TIR1	POPTR_0005s17580.1 (-2.3)	POPTR_0002s10310.1 (-2.3)
	POPTR_0002s10310.1 (-2.2)	POPTR_0017s08860.1 (-1.4)
AUX/IAA	POPTR_0005s05560.1 (-2.7)	POPTR_0010s08880.1 (-8.7)
	POPTR_0008s17220.1 (-2.1)	POPTR_0002s25830.1 (-1.3)
	POPTR_0006s25250.1 (-1.2)	POPTR_0001s17770.1 (2.2)
	POPTR_0002s25830.1 (-1.2)	POPTR_0013s03860.1 (1.1)
	POPTR_0013s03860.1 (-1.7)	
ARF	POPTR_0005s25800.1 (1.5)	POPTR_0005s25800.1 (2.1)
	POPTR_0017s01870.1 (-8.7)	POPTR_0001s13110.1 (1.3)
	POPTR_0012s10790.1 (-1.6)	POPTR_0006s12930.1 (-1.5)
	POPTR_0010s23000.1 (-1.7)	
	POPTR_0003s00890.1 (-2.0)	
GH3	POPTR_0007s10350.1 (3.8)	POPTR_0002s20790.1 (9.3)
	POPTR_0001s43990.1 (-3.3)	POPTR_0007s10350.1 (5.8)
SAUR	POPTR_0004s17180.1 (-9.5)	POPTR_0006s13920.1 (10.6)
	POPTR_0001s46220.1 (-4.1)	POPTR_0002s02590.1 (1.3)
	POPTR_0019s01610.1 (-4.1)	
	POPTR_0009s12890.1 (-1.4)	
	POPTR_0015s00920.1 (-1.6)	
	POPTR_0002s17700.1 (-1.6)	

The value in the parentheses presents the value of $\log_2$ Ratio (HP_shoot_tip/DP_shoot_tip) and the value of $\log_2$ Ratio(HP_leaf/DP_leaf), respectively.

(A)



(B)

Fig 4. DEGs involved in auxin signaling. (A) DEGs in shoot tip libraries between HP and DP. (B) DEGs in leaf libraries between HP and DP. Upregulated genes were presented in red boxes and down-regulated genes were in green boxes.

insights into poplar development.

DEGs related to hormone synthesis

In HP plants, the general downregulation of plant hormones mainly contributed to the altered phenotype. First, the endogenous concentrations of cytokinins were reduced, mainly by downregulation of adenosine phosphate-isopentenyl transferase (IPT). The expression of *AtIPT8* (also known as *PGA22*) in *Arabidopsis* resulted in elevated levels of cytokinins (Sun et al., 2003). The results of our study were consistent with mutants lacking *AtIPT3*, *5* and *7*, which were severely inhibited during shoot growth (Miyawaki et al., 2006). These phenotypes resembled those reported for cytokinin oxidase overexpressors, and agreed with previous experimental evidence that cytokinins positively regulate shoot elongation (Werner et al., 2003). In addition, the downregulation of genes involved in auxin biosynthesis was consistent with the reduced level of IAA in *Arabidopsis* mutants (Werner et al., 2003). The genes encoding GA biosynthetic enzymes were mainly downregulated, one of which (POPTR_0004s15320.1) was highly homologous with the gene encoding GA 20-oxidase, which can increase GA levels and promote xylem fiber length (Eriksson et al., 2000).

Plant hormone signaling

Hormone signaling plays diverse and critical roles during plant development. In particular, hormone interactions regulate meristem function, thereby controlling the formation of all organs in the plant. The inhibited development of HP plants could result from the generally repressed hormone signaling in the HP plants. Cytokinin signaling is a two-component signal transduction. One of the transductions is via ARR proteins. In *Arabidopsis*, ARR proteins are classified into three families: A-ARR, B-ARR and pseudo-ARR (Ha et al., 2012). Some of the type-B ARRs function as transcriptional activators of cytokinin-induced gene expression, whereas the type-A ARRs serve as repressors. A-type ARRs rapidly activate those genes that might be involved in feedback repression of CK responses

Table 5. Differentially expressed genes involved in signal transduction of CKs.

Gene name	DEGs involved in signal transduction of CKs	
	DEGs of HP shoot tip *vs*. DP shoot tip	DEGs of HP leaf *vs*. DP leaf
CRE1	POPTR_0011s09660.1 (3.1)	POPTR_0013s00510.1 (8.9)
	POPTR_0013s00510.1 (1.7)	POPTR_0011s09660.1 (4.4)
	POPTR_0008s05610.1 (1.1)	POPTR_0008s05610.1 (2.7)
	POPTR_0005s11380.1 (1.6)	POPTR_0014s15660.1 (1.8)
	POPTR_0006s01780.1 (-3.9)	
	POPTR_0010s16640.1 (-2.0)	
	POPTR_0001s24680.1 (-1.3)	
	POPTR_0008s09480.1 (-1.2)	
AHP	POPTR_0013s02950.1 (1.9)	POPTR_0013s02950.1 (3.7)
	POPTR_0010s02780.1 (-1.3)	POPTR_0008s20220.1 (1.2)
B-ARR	POPTR_0003s09940.1 (9.5)	POPTR_0006s20220.1 (9.5)
	POPTR_0016s04840.1 (9.1)	POPTR_0016s04830.1 (2.5)
	POPTR_0019s03750.1 (2.5)	POPTR_0453s00230.1 (1.4)
	POPTR_0013s04530.1 (1.1)	POPTR_0013s05670.1 (1.4)
	POPTR_0016s00310.1 (1.9)	POPTR_0006s27800.1 (-9.9)
	POPTR_0007s11470.1 (1.6)	POPTR_0007s01230.1 (-5.9)
	POPTR_0007s15110.1 (1.5)	POPTR_0016s12940.1 (-3.9)
	POPTR_0453s00230.1 (2.2)	POPTR_0009s10980.1 (-1.2)
	POPTR_0018s06410.1 (-2.1)	
	POPTR_0006s27800.1 (-3.2)	
A-ARR	POPTR_0008s19730.1 (2.2)	POPTR_0006s03950.1 (3.2)
	POPTR_0006s03950.1 (10.5)	POPTR_0008s19730.1 (1.8)

The value in the parentheses presents the value of $\log_2$ Ratio (HP_shoot_tip/DP_shoot_tip) and the value of $\log_2$ Ratio(HP_leaf/DP_leaf), respectively.

Table 6. Differentially expressed genes involved in GA signaling pathway.

Gene name	DEGs of HP shoot tip *vs*. DP shoot tip	DEGs of HP leaf *vs*. DP leaf
GID1	POPTR_0001s06970.1 (2.7)	POPTR_0009s10800.1 (11.6)
	POPTR_0002s22840.1 (1.7)	POPTR_0010s13770.1 (10.3)
	POPTR_1116s00200.1 (-9.5)	POPTR_0009s10790.1 (4.4)
	POPTR_0009s10780.1 (-3.6)	POPTR_0004s15050.1 (4.3)
	POPTR_0014s03220.1 (-3.4)	POPTR_0014s03230.1 (1.6)
	POPTR_0016s06510.1 (-2.3)	POPTR_0002s22840.1 (1.6)
	POPTR_0009s10750.1 (-1.8)	POPTR_0004s15020.1 (-12.2)
	POPTR_0009s10700.1 (-1.8)	POPTR_1116s00200.1 (-9.5)
	POPTR_0014s06890.1 (-1.6)	POPTR_0190s00210.1 (-9.2)
	POPTR_0005s04240.1 (-1.4)	POPTR_0009s10780.1 (-4.0)
	POPTR_0013s02980.1 (-1.1)	POPTR_0001s06970.1 (-2.2)
		POPTR_0013s02980.1 (-1.2)
GID2	POPTR_0002s12380.1 (1.8)	
DELLA	POPTR_0009s03800.1 (9.3)	POPTR_0001s24890.1 (8.9)
	POPTR_0009s03830.1 (3.8)	POPTR_0009s05330.1 (4.0)
	POPTR_0005s14540.1 (-4.1)	POPTR_0001s44090.1 (3.1)
	POPTR_0002s14550.1 (-2.2)	POPTR_0009s03830.1 (2.5)
	POPTR_0007s12480.1 (-2.1)	POPTR_0002s07430.1 (1.0)
	POPTR_0006s11500.1 (-1.8)	POPTR_0017s06900.1 (-3.7)
	POPTR_0017s01730.1 (-1.6)	POPTR_0005s14540.1 (-2.8)
	POPTR_0014s05910.1 (-1.5)	POPTR_0007s12480.1 (-2.7)
		POPTR_0014s05940.1 (-1.4)
		POPTR_0001s37270.3 (-1.4)

The value in the parentheses presents the value of $\log_2$ Ratio (HP_shoot_tip/DP_shoot_tip) and the value of $\log_2$ Ratio(HP_leaf/DP_leaf), respectively.

Table 7. Differentially expressed genes associated to cell cycle.

Organ	DEGs	$\log_2$ Ratio(HP/DP)	Annotation
Shoot tip	POPTR_0010s18690.1	3.5	*CAF-1*
	POPTR_0008s04320.1	3.4	*ARGININE-RICH CYCLIN 1*
	POPTR_0010s10240.2	2.9	*RHF1a*
	POPTR_0015s03450.1	2.7	*t-SNARE SNAP25*
	POPTR_0018s09330.1	2.6	*CDC27B*
	POPTR_0002s22430.1	2.4	*CAF1*
	POPTR_0004s22710.1	1.4	*CDK-subunit 2*
	POPTR_0001s33730.1	1.2	*ribonucleotide reductase (RNR)*
	POPTR_0007s01200.1	1.0	*RPN subunits*
	POPTR_0006s03370.1	-8.7	*Cyclin B*
	POPTR_0014s04930.1	-3.5	*CYCP4*
	POPTR_0002s14440.2	-3.5	*CYCP3;2*
	POPTR_0002s12350.1	-2.2	*WPP2*
	POPTR_0002s09700.1	-1.9	*Trm112p-like protein*
Leaf	POPTR_0296s00210.1	9.8	*cyclin p4;1*
	POPTR_0014s04930.1	8.9	*cyclin p4;1*
	POPTR_0014s15400.1	8.7	*condensin complex subunit*
	POPTR_0006s11390.1	3.7	*CDK5*
	POPTR_0006s17730.1	2.4	*ANAPC3*
	POPTR_0007s06510.1	2.2	*E3 ubiquitin ligase*
	POPTR_0012s11610.1	1.8	*cyclin p4;1 (CYCP4;1)*
	POPTR_0007s02640.1	-3.6	*cyclin p2;1*
	POPTR_0002s22430.1	-1.3	*CAF1*
	POPTR_0014s05510.1	-1.3	*E3 ubiquitin ligase*

$\log_2$ Ratio(HP/DP)= $\log_2$ Ratio (HP_shoot_tip/DP_shoot_tip or HP_leaf/DP_leaf).

(Hwang and Sheen, 2001). By contrast, the upregulated A-ARR in HP plants likely leads to a low level of cytokinin signaling, which then contributed to the smaller shoot apical meristem. It was reported that mutations in type-A ARRs could lead to a significantly larger shoot apical meristem and an altered phyllotaxy (Giulini et al., 2004). In the cellular component category, there were fewer genes annotated to the mitochondrial part, Golgi apparatus and endoplasmic reticulum categories. In contrast, a large number of genes were annotated to plastid or chloroplast categories. Cytokinins promote the expression of key genes of chloroplast division, thus promoting chloroplast division (Okazaki et al., 2009). We can infer that the downregulated cytokinin synthesis and signaling in the HP plants would affect plastids and chloroplasts. In particular, all the DEGs involved in the photosystem were downregulated in leaves.

Auxin promotes cell elongation and is the major factor that promotes stem elongation. In our study, the genes involved in the auxin signaling pathways were mainly downregulated in the shoot tip of HP plants. In particular, ARF, GH3 and SAUR were significantly downregulated. The F box protein TIR1 directly interacts with AUX/IAA proteins and promotes their degradation (Dharmasiri et al., 2003). Auxin acts by promoting the degradation of Aux/IAA proteins. The expressions of genes encoding ARFs are regulated and they in turn regulate the transcription of downstream genes. The three major classes of auxin response genes (Aux/IAAs, SAURs and GH3s) were generally downregulated. However, in the leaves of HP plants, the DEGs involved in auxin signaling were generally upregulated. It was reported that a significant reduction of the leaf area was caused primarily by decreased cell division during leaf development, which was compensated for only partially by increased cell size (Werner et al., 2003). It was inferred that upregulated auxin signaling led to enlarged leaf cell sizes in the haploid leaves. Auxin signaling can regulate the transcription of GA metabolism genes and the gibberellin-auxin interaction plays a role in stem growth (Frigerio et al., 2006). In the GA signaling pathway, it is now clear that the DELLA proteins are negative regulators of GA function, and GA downregulates DELLA proteins through protein degradation, likely via the proteasome pathway. The genes encoding GA receptors (e.g. GID1) were mainly downregulated in the HP plants. Mutants of GID1 displayed reduced stem height and a dwarf phenotype, which is a similar phenotype to the HP plants (Griffiths et al., 2006). The influence of hormones, especially cytokinins, on morphogenesis is primarily achieved through cell cycle regulation. Several cell cycle genes are regulated by cytokinins, including cdc2, CycD3 and different B-type cyclins. In the shoot tip of HP plants, a gene associated with M phase of the mitotic cell cycle, which is highly similar to the gene encoding cyclin b, was significantly downregulated. In addition, one of the downregulated genes was annotated as WPP2, whose reduced expression in *Arabidopsis* causes decreased mitotic activity in roots (Patel et al., 2004). However, in the leaves of HP plants, the DEGs involved in the cell cycle were generally upregulated. Several DEGs were annotated as P-type cyclins (CYCPs), which exert distinct functions in the plant cell cycle, although this has still to be proven experimentally. CYCPs are mostly expressed in proliferating cells, and in differentiating and mature tissues. If the cell size of HP leaves were partially enlarged, as stated in a previous report (Werner et al., 2003), CYCPs may play a role in coordination between the cell cycle and cell enlargement under the control of auxin and cytokinin signaling.

Materials and Methods

Plant material and culture conditions

HP plants of *P. simoniiv Carr.* × *P. nigra* L. were obtained as previously described by Qu et al. (2007). HP and DP plants were maintained in half-strength MS (Murashige and Skoog, 1962) medium with a photoperiod of 16 h light/8 h darkness condition at 24°C. To make sure that the haploid and diploid

plants were at the same growth stage, the plants were incubated in adventitious buds medium (1/2MS containing 0.5 mg/L BA and 0.05 mg/L NAA) for 1 month. They were then transferred into rooting medium (1/2MS containing 0.2 mg/L IBA) for 1 month. Shoot tips and the fourth leaves from the apex were harvested and frozen immediately in liquid nitrogen for subsequent RNA extraction.

Total RNA extraction

An E.Z.N.A. Plant RNA Kit (Omega Biotek, USA) was used to isolate total RNA, according to the manufacturer's protocol.

Solexa/Illumina sequencing

BGI Shenzhen, China carried out the Solexa/Illumina sequencing. The main reagents and supplies were the Illumina Gene Expression Sample Prep kit and Solexa Sequencing Chip (flowcell). The main instruments were the Illumina Cluster Station and the Illumina HiSeq™ 2000 System. The summary experimental process was as follows: 6 μg total RNA was extracted and purified with Oligo (dT) magnetic beads. Oligo (dT) was then used as a primer to synthesize the first and second-strand cDNA. The 5′ ends of the tags can be generated by two types of endonuclease: NlaIII or DpnII. Usually, the bead-bound cDNA is subsequently digested with restriction enzyme NlaIII, which recognizes and removes the CATG sites. The fragments apart from the 3' cDNA fragments connected to Oligo(dT) beads were washed away and the Illumina adaptor 1 was ligated to the sticky 5′ end of the digested bead-bound cDNA fragments. The junction of Illumina adaptor 1 and the CATG site is the recognition site of MmeI, which cuts at 17 bp downstream of the CATG site, producing tags with adaptor 1. After removing 3′ fragments by magnetic beads precipitation, Illumina adaptor 2 was ligated to the 3′ ends of the tags, producing tags with different adaptors at each end to form the tag library. After 15 cycles of linear PCR amplification, 105-bp fragments were purified by 6% TBE PAGE Gel electrophoresis. After denaturation, the single-chain molecules were fixed onto the Illumina Sequencing Chip (Flowcell). Each molecule then grew into a single-molecule cluster sequencing template through in situ amplification. Four types of color-labeled nucleotides were then added, and sequencing was performed via the sequencing by synthesis (SBS) method. Each line of the flowcell tunnel generated millions of raw reads with sequencing lengths of 49 bp. All the samples were analyzed after three repetitions. The sequencing data from this study were deposited in the NCBI Sequence Read Archive (SRA, http://www.ncbi.nlm.nih.gov/Traces/sra/) under the accession number SRP030954.

Then raw sequences were transformed into clean tags after certain steps of data-processing. The summary data-processing steps were as follows: removed the 3' adaptor sequence and empty reads, removed the low quality tags which were with unknown sequences N', removed tags which were too long or too short, leaving tags of 21nt long, removed tags with a copy number of 1, then generated clean tags.

Gene expression annotation

All tags were annotated using the database provided by Illumina. Briefly, a preprocessed database of all possible CATG+17nt tag sequences was created, using the *Populus trichocarpa* genome and transcriptome. All clean tags were mapped to the reference sequences, allowing only a 1-bp mismatch. Clean tags mapped to reference sequences from multiple genes were filtered, and the remaining clean tags were designated as unambiguous clean tags. The number of unambiguous clean tags for each gene was calculated and then normalized to the number of transcripts per million clean tags (TPM) (Morrissy et al., 2009).

Screening of DEGs

A rigorous algorithm (Audic and Claverie, 1997) was used to identify differentially expressed genes between the four samples. The P-value corresponds to the differential gene expression test. The FDR was used to determine the threshold of P-value in multiple tests and analysis by manipulating the FDR value. We used an FDR ≤ 0.001 and the absolute value of $\log_2$Ratio ≥ 1 as the thresholds to judge the significance of gene expression difference (Benjamini and Yekutieli, 2001).

Gene ontology functional enrichment analysis

The gene ontology (GO) classification system was used to determine the possible functions of all genes tagged. In gene expression profiling analysis, GO enrichment analysis of functional significance applies a hypergeometric test to map all DEGs to terms in the GO database, looking for significantly enriched GO terms in the DEGs compared with the genome background. The formula used was:

$$P = 1 - \sum_{i=0}^{m-1} \frac{\binom{M}{i}\binom{N-M}{n-i}}{\binom{N}{n}}$$

where N is the number of all genes with GO annotation; n is the number of DEGs in N; M is the number of all genes that are annotated to certain GO terms; and m is the number of DEGs in M. GO terms with Bonferroni-corrected P-values ≤ 0.05 were defined as significantly enriched in DEGs.

Pathway enrichment analysis

The Kyoto encyclopedia of genes and genomes (KEGG) (Kanehisa and Goto, 2000) is a major public pathway-related database. Pathway enrichment analysis identifies significantly enriched metabolic pathways or signal transduction pathways in DEGs compared with the whole genome background. The formula used was the same as that used for GO analysis.

Conclusions

We found that a number of genes involved in plant hormone signaling pathways were regulated and genes involved in hormone biosynthesis were generally downregulated, especially those associated with cytokinins. In addition, in the shoot tip libraries, the genes involved in auxin signaling and the cell cycle were downregulated. However, in the leaf libraries, the genes involved in auxin signaling and the cell cycle were generally upregulated. These results indicated the different actions of auxin in the shoot tip cells and leaf cells. Thus, we concluded that the genes identified as being involved in hormone synthesis and signaling, especially cytokinin, were the fundamental factors that underlie the altered phenotype of HP plants. DGE analysis of different HP

plant organs and plants at different stages of development should be performed to obtain a deeper understanding of the molecular variation and to determine the networks involved.

Conflict of Interest

The authors declare no conflict of interest.

Acknowledgments

This research was supported by the Fundamental Research Funds for the Central Universities (DL13BAX14), and the National High Technology Research and Development Program of China (2013AA102704).

References

Ansorge WJ (2009) Next-generation DNA sequencing techniques. New Biotechnol. 25: 195-203.

Audic S, Claverie JM (1997) The significance of digital gene expression profiles. Genome Res. 7: 986–995.

Babbitt CC, Fedrigo O, Pfefferle AD, Boyle AP, Horvath JE, Furey TS, Wray GA (2010) Both noncoding and protein-coding RNAs contribute to gene expression evolution in the primate brain. Genome Biol Evol. 2: 67-79.

Benjamini Y, Yekutieli D (2001) The control of the false discovery rate in multiple testing under dependency. Ann Stat. 29: 1165-1188.

Dharmasiri N, Dharmasiri S, Jones AM, Estelle M (2003) Auxin action in a cell-free system. Curr Biol. 13: 1418–1422.

Du J, Yang JL, Li CH (2012) Advances in metallotionein studies in forest trees. Plant Omics. 5: 46–51.

Dunwell JM (2010) Haploids in flowering plants: origins and exploitation. Plant Biotechnol J. 8: 377–424.

Eriksson ME, Israelsson M, Olsson O, Moritz T (2000) Increased gibberellin biosynthesis in transgenic trees promotes growth, biomass production and xylem fiber length. Nature Biotechnol. 18: 784–788.

Eveland AL, Satoh-Nagasawa N, Goldshmidt A, Meyer S, Beatty M, Sakai H, Ware D, Jackson D (2010) Digital gene expression signatures for maize development. Plant Physiol. 154: 1024-1039.

Frigerio M, Alabadí D, Pérez-Gómez J, García-Cárcel L, Phillips AL, Hedden P, Blázquez MA (2006) Transcriptional regulation of gibberellin metabolism genes by auxin signaling in *Arabidopsis*. Plant Physiol. 142: 553–563.

Giulini A, Wang J, Jackson D (2004) Control of phyllotaxy by the cytokinin-inducible response regulator homologue ABPHYL1. Nature. 430: 1031–1034.

Griffiths J, Murase K, Rieu I, Zentella R, Zhang ZL, Powers SJ, Gong F, Phillip AL, Hedden P, Sun TP, Thomas SG (2006) Genetic characterization and functional analysis of the GID1 gibberellin receptors in *Arabidopsis*. Plant Cell. 18: 3399–3414.

Ha S, Vankova R, Yamaguchi-Shinozaki K, Shinozaki K, Tran L-SP (2012) Cytokinins: metabolism and function in plant adaptation to environmental stresses. Trends Plant Sci. 17: 172–179.

Hwang I, Sheen J (2001) Two-component circuitry in *Arabidopsis* cytokinin signal transduction. Nature. 413: 383–389.

Jansson S, Douglas CJ (2007) *Populus*: a model system for plant biology. Annu Rev Plant Biol. 58: 435–458.

Kanehisa M, Goto S (2000) KEGG: kyoto encyclopedia of genes and genomes. Nucleic Acids Res. 28: 27–30.

Miyawaki K, Tarkowski P, Matsumoto-Kitano M, Kato T, Sato S, Tarkowska D, Tabata S, Sandberg G, Kakimoto T (2006) Roles of *Arabidopsis* ATP/ADP isopentenyltransferases and tRNA isopentenyltransferases in cytokinin biosynthesis. Proc Natl Acad Sci USA. 103: 16598–16603.

Morrissy AS, Morin RD, Delaney A, Zeng T, McDonald H, Jones S, Zhao YJ, Hirst M, Marra MA (2009) Next-generation tag sequencing for cancer gene expression profiling. Genome Res. 19: 1825-1835.

Murashige T, Skoog F (1962) A revised medium for rapid growth and bioassays with tobacco tissue culture. Physiol Plantarum. 15: 473–497.

Okazaki K, Kabeya Y, Suzuki K, Mori T, Ichikawa T, Matsui M, Nakanishi H, Miyagishima SY (2009) The PLASTID DIVISION1 and 2 components of the chloroplast division machinery determine the rate of chloroplast division in land plant cell differentiation. Plant Cell. 21: 1769–1780.

Patel S, Rose A, Meulia T, Dixit R, Cyr RJ, Meier I (2004) *Arabidopsis* WPP-domain proteins are developmentally associated with the nuclear envelope and promote cell division. Plant Cell. 16: 3260–3273.

Qu GZ, Liu GF, Wang YC, Jiang J, Wang MH (2007) Efficient tissue culture and *Agrobacterium*-mediated transformation of haploid poplar derived from anthers. Russ J Plant Physiol. 54: 559–563.

Sun J, Niu QW, Tarkowski P, Zheng BL, Tarkowska S, Sandberg G, Chua NH, Zuo JR (2003) The *Arabidopsis AtIPT8/PGA22* gene encodes an isopentenyl transferase that is involved in de novo cytokinin biosynthesis. Plant Physiol. 131: 167–176.

Szarejko I, Forster BP (2007) Doubled haploidy and induced mutation. Euphytica. 158: 359–370.

Tuskan GA, Di Fazio S, Jansson S, Bohlmann J, Grigoriev I, Hellsten U, Putnam N, Ralph S, Rombauts S, Salamov A, Schein J, Sterck L, Aerts A, Bhalerao RR, Bhalerao RP, Blaudez D, Boerjan W, Brun A, Brunner A, Busov V, Campbell M, Carlson J, Chalot M, Chapman J, Chen G-L, Cooper D, Coutinho PM, Couturier J, Covert S, Cronk Q, Cunningham R, Davis J, Degroeve S, Déjardin A, de Pamphilis C, Detter J, Dirks B, Dubchak I, Duplessis S, Ehlting J, Ellis B, Gendler K, Goodstein D, Gribskov M, Grimwood J, Groover A, Gunter L, Hamberger B, Heinze B, Helariutta Y, Henrissat B, Holligan D, Holt R, Huang W, Islam-Faridi N, Jones S, Jones-Rhoades M, Jorgensen R, Joshi C, Kangasjärvi J, Karlsson J, Kelleher C, Kirkpatrick R, Kirst M, Kohler A, Kalluri U, Larimer F, Leebens-Mack J, Leplé JC, Locascio P, Lou Y, Lucas S, Martin F, Montanini B, Napoli C, Nelson DR, Nelson C, Nieminen K, Nilsson O, Pereda V, Peter G, Philippe R, Pilate G, Poliakov A, Razumovskaya J, Richardson P, Rinaldi C, Ritland K, Rouzé P, Ryaboy D, Schmutz J, Schrader J, Segerman B, Shin H, Siddiqui H, Sterky F, Terry A, Tsai CJ, Uberbacher E, Unneberg P, Vahala J, Wall K, Wessler S, Yang G, Yin T, Douglas C, Marra M, Sandberg G, Van de Peer Y, Rokhsar D (2006) The genome of black cottonwood, *Populus trichocarpa* (torr. & Gray). Science. 313: 1596–1604.

Wang QQ, Liu F, Chen XS, Ma XJ, Zeng HQ, Yang ZM (2010) Transcriptome profiling of early developing cotton fiber by deep-sequencing reveals significantly differential expression of genes in a fuzzless/lintless mutant. Genomics. 96: 369-376.

Werner T, Motyka V, Laucou V, Smets R, Van Onckelen H, Schmülling T (2003) Cytokinin-deficient transgenic Arabidopsis plants show multiple developmental alterations indicating opposite functions of cytokinins in the regulation of shoot and root meristem activity. Plant Cell. 15: 2532-2550.

Xu YL, Gage DA, Zeevaart JAD (1997) Gibberellins and stem growth in *Arabidopsis thaliana* (Effects of photoperiod on expression of the GA_4 and GA_5 loci). Plant Physiol. 114: 1471–1476.