Association genetics and transcriptome analysis reveal a gibberellin-responsive pathway involved in regulating photosynthesis

Jianbo Xie, Jiaxing Tian, Qingzhang Du, Jinhui Chen, Ying Li, Xiaohui Yang, Bailian Li and Deqiang Zhang

1 National Engineering Laboratory for Tree Breeding, College of Biological Sciences and Technology, Beijing Forestry University, No. 35, Qinghua East Road, Beijing 100083, China
2 Key Laboratory of Genetics and Breeding in Forest Trees and Ornamental Plants, Ministry of Education, College of Biological Sciences and Technology, Beijing Forestry University, No. 35, Qinghua East Road, Beijing 100083, China
3 Department of Forestry, North Carolina State University, Raleigh, NC 27695-8203, USA

* Correspondence: DeqiangZhang@bjfu.edu.cn
† These authors contributed equally to this work

Received 1 October 2015; Accepted 22 March 2016

Abstract

Gibberellins (GAs) regulate a wide range of important processes in plant growth and development, including photosynthesis. However, the mechanism by which GAs regulate photosynthesis remains to be understood. Here, we used multi-gene association to investigate the effect of genes in the GA-responsive pathway, as constructed by RNA sequencing, on photosynthesis, growth, and wood property traits, in a population of 435 Populus tomentosa. By analyzing changes in the transcriptome following GA treatment, we identified many key photosynthetic genes, in agreement with the observed increase in measurements of photosynthesis. Regulatory motif enrichment analysis revealed that 37 differentially expressed genes related to photosynthesis shared two essential GA-related cis-regulatory elements, the GA response element and the pyrimidine box. Thus, we constructed a GA-responsive pathway consisting of 47 genes involved in regulating photosynthesis, including GID1, RGA, GID2, MYBGa, and 37 photosynthetic differentially expressed genes. Single nucleotide polymorphism (SNP)-based association analysis showed that 142 SNPs, representing 40 candidate genes in this pathway, were significantly associated with photosynthesis, growth, and wood property traits. Epistasis analysis uncovered interactions between 310 SNP–SNP pairs from 37 genes in this pathway, revealing possible genetic interactions. Moreover, a structural gene–gene matrix based on a time-course of transcript abundances provided a better understanding of the multi-gene pathway affecting photosynthesis. The results imply a functional role for these genes in mediating photosynthesis, growth, and wood properties, demonstrating the potential of combining transcriptome-based regulatory pathway construction and genetic association approaches to detect the complex genetic networks underlying quantitative traits.

Key words: Association genetics, epistatic interactions, gibberellins, photosynthesis pathway, Populus tomentosa, transcriptome.
Introduction

Gibberellins (GAs) are important phytohormones that regulate a wide range of plant growth and developmental processes, such as photosynthesis (Biemelt et al., 2004; Swain and Singh, 2005). Recent work has demonstrated that GA signaling, particularly in the upstream signaling pathway, is initiated through GA binding to the GA INSENSITIVE DWARF1 (GID1) receptor. The Arabidopsis thaliana genome contains three GID1-like genes, GID1a, GID1b, and GID1c (Griffiths et al., 2006). Activation of GID1 enables the interaction between GID1 and negative regulators including, in rice (Oryza sativa), the DELLA protein SLENDER RICE 1 (SRL1), or, in Arabidopsis, five DELLA proteins, which are encoded by RGA, GAI, RGL1, RGL2, and RGL3 (Dill et al., 2001; Fu et al., 2002; Murase et al., 2008).

DELLA proteins, which belong to the GRAS family of plant transcriptional regulators, are named after their highly conserved N-terminal DELLA motif, and mediate GA-responsive gene expression (Boss and Thomas, 2002). GA regulates growth by promoting the degradation of these repressive DELLA proteins, which are ubiquitinated via the GID2-based complex and degraded through the 26S proteasome. The GA-specific F-box proteins GID2 and SLEEPY1 mediate this degradation (Dill et al., 2004). DELLA proteins modulate downstream gene expression by interacting with multiple regulatory proteins, such as the basic helix-loop-helix transcription factors PHYTOCHROME INTERACTING FACTORS (PIFs) (Harberd et al., 2009), and GRAS protein SCARECROW-LIKE 3 (SCL3) (Zhang et al., 2011). In addition, MYBGA (also called GAMYB), an R2R3 MYB transcription factor, could be induced by GA, which binds to the GA response element (GARE) and activates the transcription of downstream target genes. MYBGA is regarded as a major transcription factor in the GA signaling pathway (Tsui et al., 2006).

Cis-regulatory elements in the promoters of the target genes involved in the GA induction of α-amylase genes have been extensively studied using barley aleurone cells. GA-inducible expression of the high-pI α-amylase, Amy1/6–4, depended on the presence of three cis-regulatory elements, the amylase box, the GARE, and the pyrimidine box in the promoter region (Gubler and Jacobsen, 1992). One additional regulatory motif, the Opaque-2 binding sequence, was required for GA-dependent expression of the low-pI α-amylase gene Amy32b (Lanahan et al., 1992). Moreover, the GARE and the pyrimidine box were shown to be essential for GA-induced expression of the cysteine proteinase gene EPB1 (Cercós et al., 1999).

In higher plants, the photosynthetic apparatus in chloroplasts is composed of nuclear- and chloroplast-encoded components (Pfannschmidt et al., 2001). Acclimation of the chloroplast to fluctuations in the environment involves changes in the level of expression of nuclear chloroplast genes and appears to be regulated primarily, but not exclusively, at the transcriptional level (Berry et al., 1988; Jung and Mockler, 2014; Richly et al., 2003). The transcription of photosynthetic genes can be regulated by light, cytokinin, photorespiration, iron availability, and biotic stresses (Carmona-Silva and Salvucci, 2013; Cortelleven and Schmülling, 2015; Feng et al., 2014; Lefebvre-Legendre et al., 2015). Identification of cis-binding motifs in nuclear genes coding for chloroplast proteins or transcription factor-binding promoter regions provides important clues about the regulatory pathway controlling chloroplast development and photosynthesis. A previous study demonstrated that the expression of ATPC, which encodes the gamma-subunit of the chloroplast ATP synthase, was induced by both cytokinin and light via the same CAAT-box cis-acting element located in its promoter (Kusnetsov et al., 1999). The promoters of the nuclear-encoded photosystem I genes PSAD, PSAF, and PETE respond to light-induced redox signals from the plastid (Pfannschmidt et al., 2001). Accumulating evidence suggests that GA plays a positive role in enhancing photosynthetic activity in plant species (Alvim, 1960; Arteca and Dong, 1981; Huerta et al., 2008). However, it is still not known how GA mediates regulation of photosynthesis, which supplies carbon skeletons and energy for growth and development.

Woody plants have distinct features that distinguish them from most herbaceous plants, including large size, long lifespans and perennial growth. Tree growth and wood formation are complex, dynamic processes that require coordinated regulation of diverse metabolic pathways in the whole plant (Persson et al., 2005). Wood formation, as an important secondary growth process, is tightly linked to photosynthesis, and the networks that regulate photosynthesis in trees are likely to be more complicated than those in herbaceous plants. While traditional molecular genetic studies with transgenic trees may potentially characterize the functions and interactions of genes in the GA response pathway, these approaches remain technically challenging in trees and may not provide a global view of the functions and interactions of these genes. As an alternative approach, association genetics, that is, characterizing polymorphisms in a gene and detecting associations with phenotypes or between loci, could provide clues to the functional and genetic interactions of genes in a pathway. Epistasis can also be used to define the functional relationships between genes in the same pathways, and to identify genetic regulatory networks and interactions of genes in response to GA (Roguev et al., 2008).

Poplars (Populus spp.) have long been used as a model system for studies of angiosperm tree physiology and genetics, and can be used to study the regulation of photosynthesis. A vast amount of genetic variation has arisen during the evolution of Populus, as is evident in natural populations (McKown et al., 2014); this variation provides a potential source of allelic variation for annotation of gene function and the identification of genetic regulatory networks of GA-responsive pathways.

In this study, we constructed a pathway of the regulation of photosynthesis in response to GA, using a transcriptome-based promoter-enrichment approach. To explore the functions and regulatory interactions of the genes involved in photosynthesis, growth, and wood formation, we used single nucleotide polymorphism (SNP) and multi-SNP association approaches to investigate the nature of genetic variation
(additive, dominant, and epistatic effects) for these genes and to identify associations with 14 quantitative traits in *P. tomentosa*. This multifaceted approach identified downstream target genes of MYBGa in trees, and serves as a model for interrogating complex signaling networks. Furthermore, this analysis provides a new strategy for examining the genetic architecture involved in the regulation of complex traits.

### Materials and methods

#### Plant material, growth conditions, and GA₃ treatment

In this study, an association population of 435 unrelated individuals, covering almost the entire climatic range and original provenance of *P. tomentosa*, was used for the SNP association studies (Du et al., 2013b). One-year-old clones of *P. tomentosa* LM50 were grown under 16 h light/8 h dark conditions. Seedlings were sprayed with either 100 μmol L⁻¹ GA₃ or water as a control. For each type of sample, we conducted three replicates for GA₃ treatment, trait measurement, and quantitative real-time PCR (qRT-PCR).

#### Measurement of physiological and biochemical characteristics

Leaves were harvested from 1-year-old *P. tomentosa* at 0, 3, 6, 12, and 24 h after GA₃ treatment for measurement of total protein concentration, sucrose phosphate synthase activity, peroxidase (POD) activity, and malondialdehyde content. Total protein concentration was measured by the method of Bradford (1976), using bovine serum albumin as the standard. Sucrose phosphate synthase activity was analyzed according to the protocol of a sucrose phosphate synthase assay kit (Nanjing Jiancheng Bioengineering Institute, Jiangsu Province, China). POD activity was measured with a plant POD assay kit (Nanjing Jiancheng Bioengineering Institute). Malondialdehyde content was determined according to the method of Heath and Packer (1968).

#### Transcriptome analysis

Measurement of biochemical characteristics (as described above) showed that these characteristics changed significantly after GA₃ treatment for 6 h, implying that this time point might show a substantial change in gene expression (see Supplementary Fig. S1 at JXB online). Therefore, we used the 6 h GA₃ treatment for RNA sequencing (RNA-seq). Leaves for RNA extraction were collected from *P. tomentosa* with or without treatment with GA₃. Three independent biological replicates were sequenced and analyzed. Paired-end sequencing was performed on an Illumina HiSeq2005. Raw data are available for download at the NCBI Sequence Read Archive under accession number SRP060593.

Clean reads were mapped using TopHat (Trapnell et al., 2012) to generate read alignments for each sample. Genomic annotations were obtained from Phytozome (http://www.phytozome.net/). The transcript isoform level and gene level counts were calculated and differential transcript expression was then computed using Cuffdiff (Trapnell et al., 2012). Gene set enrichment analysis was performed using AgriGO analysis tools (http://bioinfo.cau.edu.cn/agriGO/).

#### SNP discovery and genotyping

The 435 unrelated natural individuals were re-sequenced to >15× coverage (raw data) using the Illumina GA II sequencing platform. Filtered reads were first mapped to the *Populus* reference genome sequence using the SOAPaligner (SOAP2, version 2.20) with default options (Li et al., 2009). The mapping rate in different accessions varied from 81% to 92%, and the effective mapping depth was ~11× for most individuals.

To obtain high-quality SNPs, we selected uniquely mapped paired-end reads to perform SNP calling. The genotype likelihood of the genomic site for each tree was calculated with SOAPsnp with default parameters (Li et al., 2009). To certify our SNP calling results, we randomly compared the results with our previous SNP data from 10 candidate genes in 120 trees from genome re-sequencing using PCR-Sanger sequencing (Du et al., 2015). The accuracy of SNP calling was 99.7%, indicating the high quality of the SNP-calling platform. The gene-derived biallelic SNPs within the full-length genes, including 2 kb upstream and 1 kb downstream, were extracted using VCFtools (Danecek et al., 2011) according to the annotation of the *Populus* genome (version 3.0) downloaded from Phytozome (http://www.phytozome.net/).

#### Determination of subcellular location and common cis-regulatory elements

The subcellular localization of each gene was determined by using ChloroP 1.1 (Emanuelsson et al., 1999). To determine the over-represented DNA motifs in GA-responsive photosynthetic differentially expressed genes (DEGs), the 2 kb promoter sequences preceding the 5’ end of each annotated transcription unit in the *Populus* annotation were extracted. Known plant promoter elements and their annotations were Pearson’s correlations between each pair of genes were determined for 13 selected genes, according to their transcript abundances over time, using the corrgram package in R version 3.2. Data were log-transformed before Pearson’s correlation. All primers used in the qRT-PCR analysis are listed in Supplementary Table S1.
Results

Gibberellin affects photosynthesis, growth, and wood properties

To investigate the roles of GA in regulating photosynthesis, growth, and wood formation, we examined nine characteristics in Populus plants that had been treated with GA (or sham treated with water as a control) for 1 month. We found that nearly all the characteristics exhibited a significant improvement under GA treatment (see Supplementary Fig. S2).

To further obtain a global view of gene functions in metabolism and photosynthesis, we used MapMan analysis, which groups DEG sets into hierarchical functional categories on the basis of putative biological function (Thimm et al., 2004) (Fig. 1C, D). These annotations provide a valuable source of data for investigation of the processes, functions, and pathways involved in photosynthesis, growth, and development following GA treatment.

GA-induced genes involved in photosynthesis and growth

To understand how GA promotes photosynthesis and growth, we examined the expression of 87 DEGs involved in cell wall, chloroplast division, cell division, cell expansion, and photosynthesis (Supplementary Table S5). Of these, 80.5% (n = 70) were induced (P < 0.05), as would be expected for actively dividing cells and chloroplasts in photosynthetic
In pathways related to photosynthesis, we found 26 DEGs involved in light reaction, six in carbon fixation, six in scavenging of reactive oxygen species, and two in photorespiration. Of these, 92.3% of the genes involved in light reaction were significantly up-regulated, by 2.00- to 7.38-fold, in response to GA treatment. However, most of the genes involved in carbon fixation, scavenging of reactive oxygen species, and photorespiration were repressed by GA treatment (P<0.05).

After GA treatment, the transcript abundance of genes involved in photosynthesis, cell expansion, chloroplast division, cell division, and cell wall exhibited large changes (Supplementary Table S5). Among these, nine members of the EXPANSIN family were significantly up-regulated, by 2.92- to 519.67-fold. Two genes (TUA6 and DRP4C) encoding important components of the chloroplast division ring were significantly induced, by 3.00- and 4.27-fold. Of genes involved in cell wall functions, only one (PME61) was repressed under GA treatment. The transcript abundance of cell-cycle-associated genes (cyclin D3, cyclin D6, cyclin P1, and cyclin P4) significantly increased in response to GA. In addition, 33 transcriptional regulators exhibited large changes in expression, including NAC, WRKY, YABBY, AUX/IAA, and R2R3-MYB, consistent with the idea that growth and development require large numbers of transcription factors. Therefore, change in expression of these transcripts suggests the involvement of these genes in photosynthesis- and growth-related processes.

Promoters of GA-responsive photosynthetic genes share regulatory motifs

To explore the putative regulatory mechanisms of the photosynthetic DEGs, we explored the cellular location of the gene products and identified the cis-regulatory elements enriched in the promoter regions of the DEGs. Scanning for chloroplast transit peptides revealed that 32 (80%) of the 40 GA-responsive photosynthetic genes had these peptides, indicating that their gene products were likely targeted to the chloroplast, consistent with their functions.

To identify the enriched regulatory elements acting in this pathway, the motifs present in the promoters of the annotated genes in the Populus genome were used as the background, and hypergeometric tests were used to detect the enriched motifs in the promoters of the 40 photosynthetic genes. Known plant elements and their annotations were predicted using the latest version of PlantPAN2.0. Finally, several motifs were specifically enriched in the promoters of GA-responsive photosynthetic genes, and were present in 80–100% of promoters (Table 1). Two light-responsive elements (SOLRIP1AT and INRNTPSADB) were over-represented relative to the entire genome. In addition, two well-characterized ABA-responsive...
elements (PROXBBNNAPA and ABRELATERD1) were enriched in the promoters; however, the ABRE binding factors, belonging to group A of the bZIP family (Jakoby et al., 2002), were not induced by GA treatment (Supplementary Figure S3). The stress-responsive cis-element MYB1AT-related transcription factor gene MYB80 (Phan et al., 2011) also did not show a large change in expression. Examination of all the distribution of these motifs revealed that most transcription start sites, with some weak peaks scattered throughout the promoters. As shown in Figure 2, 15–82% of motifs were enriched within 0.5 kb of the TSS. A total of 150–191 motifs were identified in these promoter sequences; all the regulatory elements are shown in Supplementary Table S6. Of these, typical CAAT and TATA boxes, which are critical for transcription initiation, occurred in all the promoter regions.

Notably, two essential motifs, PYRIMIDINEBOXHVEPB1 (TTTTTTCC) and GARE2OSREP1 (TAACGTA) (Chen et al., 2006; Martinez-Andujar et al., 2012; Zhang et al., 2004), which are associated with GA responses, were significantly enriched in the promoters (P<0.05) (Table 1). A scan of the promoters revealed that more than half of the photosynthetic DEGs have multiple copies of these motifs in their promoter regions (Supplementary Table S6). Three additional GA-related cis-elements were over-represented in the promoter regions of these genes (TATCCAOSAMY, TATCCAYMOTIFOSRAMY3D, and TATCCACHVAL21), supporting the differential expression of these genes under GA treatment. A poplar R2R3MYB gene, MYBGa (Potri.003G189700), whose expression is induced by GA, is postulated to be part of the GA-responsive pathway leading to GA-inducible gene expression. Because most (37 out of 40) of the photosynthetic DEGs have the pyrimidine-box and GARE motifs in their promoters, and these two elements are important components of the GA response complex (GARC), which interacts with MYBGa, these 37 photosynthetic genes likely represent the direct targets of MYBGa (Supplementary Table S7). Therefore, we elucidated some downstream targets of MYBGa in regulating photosynthesis and constructed a GA-responsive pathway consisting of five groups (A–E) of genes involved in regulating photosynthesis in *Populus* (Fig. 3A).

To provide insights into gene interactions and functions in this GA-responsive pathway, we selected 12 candidate genes and used qRT-PCR analysis to examine their expression over a time-course of 24 h following GA treatment. These 12 genes exhibited distinct but partially overlapping patterns of expression (Fig. 3B). Upstream genes in groups A–D and downstream targets of MYBGa (group E) exhibited two distinct transcriptional signatures; the abundance of upstream genes in groups A–D changed more significantly than the downstream genes in group E across the five time points, suggesting complex regulatory interactions among upstream genes in groups A–D. Next, we constructed a gene–gene correlation matrix using the expression patterns of these 12 genes and identified 26 positive or negative correlations (P≤0.05, Fig. 3C). Of these gene pairs, *Pt-MYBGa* and *Pt-GID1B.1* (P≤0.05, R=0.94) showed highly significant correlations.

### Table 1. Regulatory motifs enriched in promoters of 40 GA-responsive genes involved in photosynthesis (P<0.05)

<table>
<thead>
<tr>
<th>Consensus sequence</th>
<th>Cis-regulatory element</th>
<th>Associated type of transcription factor</th>
<th>%a</th>
<th>P valueb</th>
</tr>
</thead>
<tbody>
<tr>
<td>TATCCAY</td>
<td>TATCCAYMOTIFOSRAMY3D</td>
<td>MYBS3, MCB1, and MCB2</td>
<td>0.95</td>
<td>0.005927</td>
</tr>
<tr>
<td>GARE2OSREP1</td>
<td></td>
<td>MYBGA</td>
<td>0.95</td>
<td>0.041736</td>
</tr>
</tbody>
</table>

aPercentage of genes containing that particular cis-element.
bHypergeometric test P<0.05.

downstream targets of MYBGa in regulating photosynthesis and constructed a GA-responsive pathway consisting of five groups (A–E) of genes involved in regulating photosynthesis in *Populus* (Fig. 3A).

To explore the functions of genes in groups A–E of the GA-responsive pathway, we identified SNPs in the genes, including the regions 2 kb upstream and 1 kb downstream of each gene. In total, we detected 4446 common SNPs (frequency ≥0.05) within 47 full-length genes, approximately 19 SNPs per kb (Supplementary Tables S8 and S9). Analysis with a mixed linear model in TASSEL 2.1, which takes into account the kinship matrix and genetic structure (K+Q), identified 152 significant associations (Q<0.01) representing 142 SNPs in 40 candidate genes and phenotypes, including four photosynthetic, three growth, and seven wood property traits (Fig. 4A and Supplementary Table S10). The individual SNPs explained between 0.01% (*Pt-PSAE-2-SNP21*) and
Gibberellin-responsive pathway in regulating photosynthesis

20.89% (Pt-GID2.3-SNP17) of the phenotypic variation. These findings suggest that these genes play important roles in tree photosynthesis, growth, and wood properties.

For SNPs in group A, three SNPs (Pt-GID1B.1-SNP82, Pt-GID1B.1-SNP73, and Pt-GID1B.2-SNP45) were significantly associated with V, two SNPs (Pt-GID1B.1-SNP58 and Pt-GID1B.1-SNP96) with Ci, and one (Pt-GID1B.1-SNP66), Ci, with H. For SNPs in group B, 15 were significantly associated with 10 traits; for example, Pt-GID2.1-SNP3 was associated with HEMC and Cond. Interestingly, eight SNPs in Pt-GID2.2 were significantly associated with FW, V, FL, and LC, providing new insights into the potential role of Pt-GID2.2 in plant growth and wood formation. In group C, a total of 17, 53, and 42 SNPs were detected in Pt-RGA.1, Pt-RGA.2, and Pt-RGA.3, respectively. Three SNPs in Pt-RGA.1 (SNP22, SNP45, and SNP5) were associated with V, LC, and Ci, respectively. Two SNPs in Pt-RGA.1 (SNP29 and SNP46), and three SNPs in Pt-RGA.3 (SNP36, SNP12, and SNP13), were simultaneously associated with V. Pt-RGA.2-SNP49 was significantly associated with Ci. For SNPs in group D, nine SNPs in Pt-MYBGa were significantly associated with six traits. Notably, nine SNPs in Pt-MYBGa were significantly associated with two photosynthetic traits (Ci and Trmmol), two wood property traits (MFA and HEMC), and two growth traits (V and H). Among the SNPs in group E, as expected, 100 SNPs were significantly associated with photosynthetic traits, including Pt-PSBQ.1-SNP82 (A>T, \( Q=2.50E-07 \)) and Pt-PSBQ.2-SNP131 (G>T, \( Q=0.003 \)) associated with Pn (Supplementary Table S10). Notably, Pt-GID2.3-SNP17 (T>C, \( Q=0.0029 \)) was significantly associated with FL, and explained 20.89% of the phenotypic variation; this was the highest contribution to phenotype. The significant marker–trait associations indicated that these genes in the GA-responsive pathway play important roles in photosynthesis, growth, and wood formation.

We next used the epiSNP package (Ma et al., 2008), emphasizing multi-SNP additive and dominant effects for each quantitative trait, to perform multi-SNP association with four photosynthetic traits (Pn, Cond, Ci, and Trmmol) and 10 growth and wood quality traits (H, D, HC, LC, V, MFA, AC, FW, and HEMC). Under the additive and dominant effects models, we detected 497 significant associations for 157 SNPs in 36 genes associated with all traits (Supplementary Table S11) (\( Q<0.01 \)). The total numbers of
identified SNP–trait associations varied across trait categories, with 66 associations for photosynthesis, 313 associations for wood properties, and 118 associations for growth traits. All 157 SNP markers exhibited significant associations with at least one trait, and the number of significant SNPs ranged from eight to 39 for each trait (Supplementary Table S11). Correspondingly, each trait was associated with variation in at least eight candidate genes. For photosynthetic traits, SNP–trait associations showed that 28 genes associated with four traits, including genes from all five groups A–E. For wood properties, additive models identified 24 genes with SNPs associated with variation in wood quality traits; of these, Pt-MYBGa SNPs associated with five traits. For growth traits, we detected 36 unique SNPs significantly associated with all three traits. Under the dominant effect model, we detected 214 associations with positive dominance values and 83 with negative values, across all trait categories (Supplementary Table S11). Many genes associated with multiple traits within/across all trait categories, and the same SNPs from the same gene associated with different traits. The number of SNP associations with negative versus positive effects across all traits categories was, respectively, 36 and 26 for photosynthetic traits, 132 and 181 for wood property traits, and 37 and 81 for plant growth traits. We also found that different SNPs from the same gene had positive and negative dominant effects on different traits (Supplementary Table S11). For example, Pt-RGA.2-SNP88 had a negative dominant effect on AC, but Pt-RGA.2-SNP48 had a positive effect. In summary, SNPs in all five groups associated with at least five traits in common, reflecting the possible genetic interaction between the different groups in the GA-responsive pathway, and indicating that they may function in the same pathway.

### Epistatic interactions among genes in the GA-responsive pathway

Potential epistatic interactions of the SNPs in genes of the GA-responsive pathway for photosynthetic, growth, and wood quality traits were examined using the epiSNP package, which can identify SNP–SNP interactions in a population. In total, 310 SNP–SNP pairs from groups A–E were significantly associated with 14 traits (Supplementary Table S12). After statistical significance analysis, we identified 164 SNP–SNP pairs (Q<0.01) associated with all 14 traits in genes from group D (MYBGa) and group E (MYBGa target genes). Multi-way SNP–SNP interaction analysis showed that MYBGa SNPs have epistatic interactions with 20 genes in group E (Fig. 4B). For Pt-MYBGa and its target genes, 49 SNP–SNP pairs were identified as being associated with 13 traits. Thus, a total of 20 MYBGa target genes in group E have epistatic interactions with MYBGa (Supplementary Table S13), supporting the existence of epistatic interactions between MYBGa and the target photosynthetic genes. Of the 37 genes in group E, 67.5% were found to have epistatic interactions with genes in group A, 59.5% with genes in group B, and 67.5% with genes in group C. In addition, we identified 124 associations (Q<0.01) with 14 traits from different groups of A–D. Of these, 63.7% represented mRNA–mRNA (from different groups) interactions (Supplementary Table S12), supporting the potential genetic interactions of upstream genes in groups A–D in the GA-responsive pathway.
The genotype combinations between SNP markers in *Pt-MYBGa* and its putative target genes associated with high and low values are shown in Fig. 5A for *Pn* and Fig. 5B for *Trmmol*. The patterns of high- and low-value groups clearly differ across each of the different multi-locus dimensions that were considered to be evidence of epistasis, or gene–gene interactions among *Pt-MYBGa* and its putative target genes (Fig. 5A, B). In addition, the dendograms of interactions among phenotype, *Pt-MYBGa*, and its six putative target genes showed a strong joint effect of *Pt-MYBGa*-SNP1 and *Pt-AGT*-SNP129, as well as *Pt-MYBGa*-SNP1 and *Pt-LHCb4.3*-SNP1 (Fig. 5C, D). To further visualize the two-way interaction for photosynthetic and wood property traits, we created an entropy-based interaction graph for *Trmmol* and *FW*; this also pointed toward interactions between SNPs in *Pt-MYBGa* and its putative target genes (*Pt-PSAK*, *Pt-PSAF*, *Pt-FBA2*, *Pt-PSBP1*, *Pt-AGT*, and *Pt-PSBQ.2*) (Fig. 5E, F). All six variants explained 0.14–3.72% of phenotypic variation by themselves. Notably, we found 164 two-locus-genotype combinations between *MYBGa* and its putative target genes (Supplementary Table S12). Taken together, the results of epistasis analysis (Fig. 4B) confirm the GA-responsive pathway and support the existence of epistatic interactions among genes in this pathway, indicating that genes in groups A–E may contribute to the same pathway in promoting photosynthesis, wood formation, and tree growth.

**Discussion**

**A model for the GA-responsive pathway in regulating photosynthesis in trees**

GAs regulate many aspects of plant development. Previous studies have demonstrated extensive changes in the transcriptome following GA application; GA also induced fruit coloring and seed abortion, and promoted flower opening (Cheng et al., 2015). Using RNA-seq, Mutasa-Göttgens et al. (2012) uncovered global transcriptional responses to exogenous GA in the shoot apex. However, little attention has been paid to changes in gene expression in photosynthetic tissues from woody plants. In the present study, we used RNA-seq to examine the global gene expression profiles in leaves following GA treatment and found that large numbers of genes involved in photosynthesis, cell wall, chloroplast division, and cell division exhibited large transcriptional changes, indicating a requirement for expression of these genes during GA-induced growth (Supplementary Table S5). These results also provide the potential to study GA-regulated genes involved in a broad range of physiological functions, especially in photosynthesis and growth.

In recent decades, significant progress has been made in the identification of upstream GA signaling components and *trans* and *cis*-acting factors that regulate downstream targets of MYBGa in higher plants (Sun and Gubler, 2004). Early-GA-regulated genes encode a number of positive and negative transcriptional regulators that likely act as a transcriptional cascade, activating the expression of downstream targets of MYBGa involved in cell elongation and other processes (Sun and Gubler, 2004). For the upstream regulatory cascade, GA appears to de-repress its signaling pathway by inducing proteolysis of its negative regulator DELLa via the GID2 complex, to release transcription factors (Sun and Gubler, 2004). In the present study, we found that GID1B1, GID1B2, RGA.2, and MYBGa were differentially regulated by GA, supporting the idea that upstream receptors and signaling components may respond at different times after GA treatment. Consistent with our findings, time-course studies showed a lag of approximately 1 h between GA-induced SLN1 (DELLa protein in barley) degradation and increased expression of early-response genes such as the transcription factor gene *MYBGa* (Fu et al., 2002).
GA-regulated genes act in many processes, such as cell elongation and germination (Sun and Gubler, 2004), and their respective promoters have putative motifs related to the GA response (Aftab et al., 2010; Chen et al., 2006). Our computational analysis to detect binding motifs enriched in the promoter regions found several elements that were enriched in the promoters of 40 photosynthetic DEGs (Table 1). Some light-responsive elements, such as SORLIP1AT (Hudson and Quail, 2003) and INRNTPSADB (Zheng et al., 2014), were enriched in the promoter regions; it may be a coincidence that genes responsive to GA contain these regulatory elements, as many photosynthetic genes are light regulated (Tyagi and Gaur, 2003). Notably, our results revealed that two essential cis-regulatory motifs, PYRIMIDINEBOXHVEPB1 and GARE2OSREP1 (Chen et al., 2006; Martínez-Andújar et al., 2012; Zhang et al., 2004), related to the GA response,

were significantly over-represented in these 40 photosynthetic DEGs (P<0.05) (Table 1). Promoter analysis indicated that the GA-responsive photosynthetic genes may be regulated by MYB genes through the GA-responsive elements. By focusing on identifying common regulatory elements of GA-responsive photosynthetic genes, we found new downstream targets of MYB genes and built a picture of the GA-responsive pathway in regulating photosynthesis (Fig. 3A). Supporting this, the time-course analysis of 12 candidate genes in this pathway showed positive or negative correlations among these genes (Fig. 3B, C). Thus, our study proposes downstream regulatory targets of the known GA-responsive pathway.

Functional interpretation of genetic associations

SNP-based association mapping provides a suitable approach for annotation of gene function and identification of genetic regulatory networks; thus, it may be useful for exploring the functions and interactions of genes in the GA pathway. Here, we used multi-gene association approaches, combining multiple full-length candidate genes from a shared biological pathway to identify numerous loci underlying variation in photosynthesis, growth, and wood property traits. All candidate genes were expressed in GA-treated leaves, and some of these genes were differentially regulated by GA (Supplementary Table S2), indicating that they are good candidates for participating in photosynthesis. Considering that different family members in the GA signaling cascade are often functional homologs (Claeys et al., 2014) and occur as large gene families in Populus, further work is needed to estimate the genetic effect of other family members. Previous studies suggest that various growth processes, such as cell division and cell expansion, could be regulated by a core GA signaling cascade (Jiang et al., 2012). This agrees with our observation that one upstream gene in groups A–D could affect between 1 and 11 traits (Supplementary Table S10). Taking GIDI as an example, five unique SNPs from Pt-GID1B.1 were associated with H, Ci, and V, and one SNP from Pt-GID1B.2 was associated with V. This finding was supported by a previous functional analysis, which showed that GIDI could affect tree height and lower male fertility (Griffiths et al., 2006). A similar situation was also observed for the DELLA protein gene RGA, which encodes an essential component of the GA signaling cascade. For RGA, 13 SNPs from group C (Pt-RGA.1, Pt-RGA.2, and Pt-RGA.3) were significantly associated with six traits, suggesting that DELLA proteins have roles in photosynthesis, tree growth, and wood formation. Furthermore, a previous study found that non-synonymous mutations in the conserved DELLA domain render the protein insensitive to degradation and constitutively block the GA response (Gou et al., 2010). Thus, SNPs within DELLA protein genes may change the stability of DELLA proteins, influence the release of transcription factors, and affect phenotypic variation (Gou et al., 2010). In addition, we found that nine SNPs in MYBGa were associated with six traits, suggesting that they have roles in photosynthesis, growth, and wood formation. Supporting this idea, a previous study showed that several MYB family members implicated in the regulation of genes encoding lignin biosynthetic enzymes potentially affect lignin content in poplar (Wilkins et al., 2009). Thus, these SNPs in MYBGa genes are potentially interesting subjects for future research.

Photosynthesis, as an essential primary biological process, is strongly associated with growth and wood formation. We identified several downstream photosynthetic genes in group E as being associated with 14 photosynthesis, growth, and wood property traits. Pt-PSBQ.1 can affect the activity of photosystem II (Thornton et al., 2004), and its absence could lead to decreased oxygen evolution (Kashino et al., 2006). As expected, 17 unique SNPs in Pt-PSBQ.1 were associated with photosynthetic, wood property, and growth traits (Supplementary Table S10). Another photosynthetic gene encoding ferredoxin (FDX1) was up-regulated in response to GA (log2FC=1.32). FDX1 also participates in other reactions in the chloroplast, including nitrogen and sulfur assimilation, amino acid and fatty acid synthesis, and redox regulation (Knaff and Hirasawa, 1991), and photosynthetic and non-photosynthetic tissues have different FDX isoforms (Green et al., 1991; Hanke et al., 2004). Our results showed that six SNPs in Pt-FDX1 were associated with H, Ci, and V (Supplementary Table S10), improving our understanding of Pt-FDX1 functions.

In this study, the repeated occurrence of multiple traits associated with one gene was notable. This finding can be explained by two hypotheses. On the one hand, significant genetic correlations frequently occurred among multiple, related traits, indicating trait co-selection. On the other hand, this may be an indication of pleiotropy in the broad sense. The potential pleiotropic loci may act as upstream receptors and signaling components that affect multiple traits, and the developmental integration of these different traits might have led to their genetic integration (McKown et al., 2014). This is consistent with our finding that Pt-GID1, Pt-GID2, Pt-RGA, and Pt-MYBGa are all associated with multiple traits (Supplementary Table S10). Future work will require the application of biological pathway-based association (Wang et al., 2010) to examine many more candidate genes involved in the diverse pathways of trees.

Epistasis of genes in the GA-responsive pathway

Non-additive interactions between segregating mutations (epistasis) in natural populations can significantly influence the rate and direction of evolutionary change (Xu et al., 2012). This process is also an important component of genetic variation and has great potential to improve predictions of long-term response to co-selection and inbreeding depression. Epistasis has also been used extensively to detect functional relationships between genes in the same biological and regulatory pathways, particularly in organisms with natural populations, in which large numbers of mutants can be generated, crossed, and phenotyped (Phillips, 2008), such as Populus.

We used epiSNP to identify SNP–SNP pairs associated with traits (Ma et al., 2008) and explore the interactions between genes in the GA-responsive pathway. The presence of allelic interaction could provide a clue to the underlying
genetic regulatory mechanisms (Gjuvsland et al., 2010). In our study, 164 significant associations ($Q<0.01$) from group D ($MYBGa$) and group E ($MYB$-target genes) were associated with all 14 traits examined (Fig. 4B and Supplementary Table S12), reflecting the possible genetic interaction of $MYBGa$ and potential target genes. Considering that $Pt-MYBGa$ and its potential target genes are closely related to photosynthesis, growth, and wood formation, we identified 20 $MYBGa$ target genes that have epistatic interactions with $MYBGa$ (Supplementary Table S13). For example, $Pt-MYBGa-SNP126$ and $Pt-PSBQ.1-SNP96$ were associated with seven traits. These observations were supported by the significant positive correlations in gene expression for the two gene pairs in our time-course transcript profiling (Fig. 3C). In addition, significant SNP–SNP interactions between group A, B, C, and E genes were also detected, suggesting their potential roles in the same pathway.

In contrast to other complex trait analyses or quantitative trait locus mapping, epistasis asks how a specific locus interacts with other loci and can identify a large fraction of the missed genetic information in the large genetic system (Phillips, 2008). In this study, we identified allelic variants from three unique genes with significant epistatic effects for Pn and Trmmol (Fig. 5A, B). These findings suggest the coselection of epistatic alleles at protein–protein interactions, although the mechanisms for their generation and maintenance are not clear. We demonstrated that the transcriptome-based regulatory pathway construction approach, along with the identification of common motifs and confirmation using a candidate gene association (single-SNP and multi-SNP) study, can be employed to identify naturally occurring allelic variation in genomic loci or genes associated with important photosynthesis, growth, and wood-quality traits, as well as their epistatic effects that contribute to the same pathway. However, the power of RNA-seq at one time point and qRT-PCR technology to dissect a full regulatory network is limited, so other methods such as yeast two-hybrid, chromatin immunoprecipitation, and time-course gene coexpression assays should be employed to build and verify a more comprehensive map of the full ‘interactome’ (Marchini et al., 2005). The incorporation of more genes (such as gene family members, kinases, and transporters) in shared biosynthetic pathways (Wang et al., 2010) would provide a more complete picture of the effects of genetic variation on photosynthesis, growth, and wood property traits.

**Supplementary data**

Supplementary data are available at JXB online.

**Table S1.** Real-time PCR primer sequences.

**Table S2.** The genes differentially expressed in response to GA treatment.

**Table S3.** The enriched GO terms of the differentially expressed genes in response to GA.

**Table S4.** Gene ontology enrichment analyses for up-regulated or down-regulated differentially expressed genes ($Q<0.05$).

**Table S5.** The 87 differentially expressed genes involved in photosynthesis, cell wall, chloroplast division, and cell expansion.

**Table S6.** All the regulatory elements present in the promoter regions.

**Table S7.** The 37 photosynthetic genes are likely the direct targets of MYBGa.

**Table S8.** The allelic variation in 47 candidate genes.

**Table S9.** The SNP genotype data of 47 genes.

**Table S10.** SNP markers significantly associated with photosynthetic, growth, and wood properties in the association population.

**Table S11.** SNP markers significantly associated with photosynthetic, growth, and wood properties under additive and dominant effects.

**Table S12.** Epistatic interactions between SNPs in genes in the GA-responsive pathway.

**Table S13.** The 20 MYBGa target photosynthetic genes.

**Figure S1.** Measurement of the biochemical characteristics following GA or water treatment.

**Figure S2.** Measurement of the photosynthesis, growth, and wood property characteristics following 1 month of GA or water treatment.

**Figure S3.** Phylogenetic analysis of the bZIP family.

**Acknowledgements**

This work was supported by the Fundamental Research Funds for Central Universities (No. 2015ZCQ-SW-01), the State Key Basic Research Program of China (No. 2012CB114506), and the Project of the National Natural Science Foundation of China (No. 31170622).

**References**


Artocia RN, Dong CN. 1981. Increased photosynthetic rates following gibberellic acid treatments to the roots of tomato plants. Photosynthesis Research 2, 243–249.

Berry JO, Carr JP, Klessig DF. 1988. mRNAs encoding ribulose-1, 5-bisphosphate carboxylase remain bound to polysomes but are not translated in amaranth seedlings transferred to darkness. Proceedings of the National Academy of Sciences of the United States of America 85, 4190–4194.


trans-acting elements involved in the co-ordinated gene expression regulated by gibberellins and abscisic acid. The Plant Journal 19, 107–118.


