Research paper

Genome-scale transcriptome analysis of the desert poplar, *Populus euphratica*

Qiang Qiu¹, Tao Ma¹, Quanjun Hu, Bingbing Liu, Yuxia Wu, Haihong Zhou, Qian Wang, Juan Wang and Jianquan Liu²

Molecular Ecology Group, Key Laboratory of Arid and Grassland Ecology, Lanzhou University, Lanzhou 730000, China; ¹These authors contributed equally to this work; ²Corresponding author (liujq@nwipb.ac.cn)

Received October 6, 2010; accepted February 10, 2011; published online March 22, 2011; handling Editor Ron Sederoff

*Populus euphratica* is well-adapted to extreme desert environments and is an important model species for studying the effects of abiotic stresses on trees. Here we present the first deep transcriptomic analysis of this species. To maximize representation of conditional transcripts, mRNA was obtained from living tissues of desert-grown trees and two types of callus (salt-stressed and unstressed). *De novo* assembly generated 86,777 Unigenes using Solexa sequence data. These sequences covered 92% of previously reported *P. euphratica* expressed sequence tags (ESTs) and 90% of the TIGR poplar ESTs, and a total of 58,499 high-quality unique sequences were annotated by BLAST similarity searches against public databases. We found that 27% of the total Unigenes were differentially expressed (up- or down-regulated) in response to salt stress in *P. euphratica* callus. These differentially expressed genes are mainly involved in transport, transcription, cellular communication and metabolism. In addition, we found that numerous putative genes involved in ABA regulation and biosynthesis were also differentially regulated. This study represents the deepest transcriptomic and gene-annotation analysis of *P. euphratica* to date. The genetic knowledge acquired should be very useful for future studies of the molecular adaptation of this tree species to abiotic stress and facilitate genetic manipulation of other poplar species.

**Keywords:** abiotic stress, Illumina/Solexa sequencing, *Populus euphratica*, transcriptome.

Introduction

Salinity and drought are two major environmental stressors that inhibit tree growth in semiarid and arid areas (Boyer 1982). To survive in such areas, plants require complex responses to these abiotic stressors (Bohnert et al. 1995), involving signal transduction, regulation of gene expression, ionic homeostasis, scavenging of reactive oxygen species, accumulation of compatible solutes and growth regulation (e.g. Hasegawa et al. 2000, Seki et al. 2001, Bartels and Sunkar 2005, Yamaguchi-Shinozaki and Shinozaki 2006, Munns and Tester 2008). These responses have been partly elucidated in both trees and herbs through the identification and characterization of genes that are activated in response to abiotic stress (Ingram and Bartels 1996, Seki et al. 2002, Xiong et al. 2002, Zhu 2002, Shinozaki et al. 2003).

*Populus euphratica* Oliv. (Salicaceae) is naturally distributed in western China and adjacent Middle-Eastern countries (Browicz 1977). It plays an important role in maintaining local arid ecosystems (Ma et al. 1997) and can grow in deserts with extremely hot and dry summers, while other congeners are known to be drought sensitive (Wang et al. 2007, 2008, Ding et al. 2010). In addition, *in vitro* experiments have indicated that this species can tolerate salt concentrations up to 450 mM and mannitol concentrations up to 400 mM (Watanabe et al. 2000, Gu et al. 2004b). Thus, *P. euphratica* has been widely considered as a model species for elucidating abiotic resistance mechanisms of trees, e.g., responses to salinity or drought stress (Chen et al. 1997, 2009, Gu et al. 2004a, Ottow et al. 2005, Wu et al. 2007, Zhang et al. 2007, Sun et al. 2009, 2010, Ye et al. 2009, Zeng et al. 2009). In a recent
transcriptomal effort, Brosché et al. (2005) recovered 14,000 expressed sequence tags (ESTs) from normalized and sub-
tracted cDNA libraries of normal and stress-exposed samples of P. euphratica trees using Sanger EST sequencing. This study 
yielded valuable information, but the known scale of the P. trichocarpa genome (Tuskan et al. 2006), and findings that 
EST sequences generated from Sanger sequencing do not 
cover the entire transcriptional profile (Hale et al. 2009) indicate that a minority of the P. euphratica genes were recovered.

Genome-scale transcript analysis aims to capture an unbi-
ased view of the complete RNA transcript profile of a species (Wilhelm et al. 2008), allowing the transcriptional level of each 
gene in a given tissue at a given point in its life cycle to be 
monitored. The development of next-generation, massively 
parallel sequencing technologies (e.g., Illumina/Solexa-based 
RNA-Seq technology) has significantly improved the possible 
depth of transcriptome sequencing, with significantly reduced 
costs (Wang et al. 2009). Using this high-throughput RNA 
sequencing, it is possible to recover transcripts that are 
expressed at extremely low levels, and to isolate substantial 
numbers of novel transcripts that previous large-scale Sanger 
sequencing procedures could not identify (Morozova and Marra 2008, Wang et al. 2009). In addition, such sequencing-
based methods detect absolute expression, rather than relative 
gene expression changes, and thus overcome many of the 
inherent limitations of microarray-based systems (Hoen et al. 
method is now being widely applied to non-model species for 
which genomic information is not available, and many ESTs and 
numerous novel transcripts are being recovered (Trick et al. 

Here, we present a de novo assembly of the P. euphratica 
transcriptome using Solexa data. Our data were collected by 
sequencing cDNA libraries of living tissues from mature trees 
growing in the Talim Basin desert, which had a long period to 
adapt to the local conditions before salt-stressed callus and 
unstressed callus were sampled. We specially examined the 
gene expression dynamics of this species in response to salt 
stress and identified a core set of stress-related transcripts. 
The acquired information should facilitate attempts to elucidate 
response mechanisms of this species to abiotic stress and to 
develop stress-tolerant poplar trees through genetic manipula-
tion of wild genetic resources.

**Materials and methods**

**Plant material**

Three sets of samples representing desert-grown trees, 
control-callus samples and salt-stressed callus samples were 
examined in this study. We collected three replicate samples 
(all of the same fresh weight) of roots, leaves, flower buds, 
flowers, xylem and phloem from two mature male P. euphratica 
trees and one mature female tree in the Talim Basin desert in 
Xinjiang. In addition, we cultivated calli using the method 
described by Zhang et al. (2004a), replaced the growth 
medium of one set with fresh, unamended medium and that of 
another set with fresh medium supplemented with 100 mM 
NaCl (to impose salt stress), and then harvested both sets 24 h 
later. All samples were rapidly stored at ~80 °C until required 
for RNA extraction.

**RNA extraction and quality determination**

Total RNA was extracted three times from each of the sample 
sets, using a CTAB procedure (see Chang et al. 1993). A260/A280 ratios of the RNA samples dissolved in 10 mm Tris (pH 
7.6) ranged from 1.9 to 2.1. The integrity of the RNA samples 
was examined with an Agilent 2100 Bioanalyzer and their RIN 
(RNA integrity number) values ranged from 8.6 to 10.0, with 
no sign of degradation. RNA from each replicate was pooled 
(in equal volumes) to obtain a single RNA sample for cDNA 
preparation and RNA-Seq, and equal amounts of mRNA from 
different tissues of the desert-grown trees were pooled to 
make single samples.

**Illumina cDNA library preparation and sequencing**

For cDNA synthesis and Solexa sequencing, 20 µg of total 
RNA was used, at a concentration of ≥400 ng/µl. Poly(A) 
mRNA was first purified using beads with oligo(dT). Then, the 
mRNA was fragmented into small pieces using divalent cations 
at an elevated temperature. Based on these cleaved RNA frag-
ments, we used random hexamer-primer and reverse tran-
scriptase (Invitrogen) to synthesize first-strand cDNA. 
Second-strand cDNA was synthesized using RNase H 
(Invitrogen) and DNA polymerase I (New England BioLabs). 
We constructed three paired-end cDNA libraries with insert 
sizes of 200 bp, and then sequenced the cDNA using an 
Illumina (San Diego, CA, USA) Genome Analyzer according to 
the manufacturer’s protocols with a read length of 75 bp.

**De novo assembly and assessment**

Reads from each library were assembled separately. Adapter 
sequences and reads containing too many (>8) unknown 
basics or low-quality bases (>50% of the bases with a quality 
score ≤5) were filtered using in-house Perl scripts. The aver-
age proportion of clean reads for each library was ~96.5%. 
The transcriptome sequence was assembled into distinct con-
tigs with short reads by SOAPdenovo software (Li et al. 2010) 
(http://soap.genomics.org.cn), which applies the de Bruijn 
graph data structure to construct contigs. The reads were then 
realigned to the contig sequences, and the paired-end rela-
ationships between the reads were used to construct scaffolds 
between contigs. To fill the intra-scaffold gaps, we then used 
the paired-end information to retrieve read pairs that had one 
read well-aligned on the contigs and another read located in
the gap region, and then locally assembled the collected reads. After gap closure, we constructed a non-redundant Unigene set from all three assembled datasets using the EST assembly program TGICL (Pertea et al. 2003).

To obtain high-quality sequences for further annotation and analysis, we excluded Unigene sequences that might represent non-coding RNAs, identified by comparing our sequences with known non-coding RNAs deposited in the Rfam database (http://www.sanger.ac.uk/Software/Rfam/, release 10.0). Unigene sequences assigned to microbial (MBGD: http://mbgd.genome.ad.jp/), fungal and virus (based on data downloaded from the NCBI database) sources were also filtered out. In addition, sequences for which >50% of the bases aligned with sequences in UTRdb (http://utrdb.ba.itb.cnr.it/) and/or contained <200 non-UTR bases were excluded.

In order to assess the sequence coverage of the transcriptome assemblies, we searched our Unigene sequences by BLAST (Altschul et al. 1997) against (i) 13,845 unassembled P. euphratica ESTs from GenBank; (ii) 172,068 assembled ESTs from 14 poplar species in the TIGR Plant Transcript Assemblies database (Childs et al. 2007) (accessed August 2009; http://planta.tigr.org/index.shtml); (iii) P. trichocarpa protein-coding transcript and genome sequences (JGI release 2.0; http://www.phytozome.net/poplar.php); and (iv) all available expressed sequences identified from 254 plant species in the TIGR Plant Transcript Assemblies database.

Functional annotation
We annotated sequences based on a set of sequential BLAST searches (Altschul et al. 1997) designed to find the most descriptive annotation for each sequence. The assembled unique transcripts were compared with sequences in GenBank’s non-redundant database using the BLASTN algorithm, the GI accessions of best hits were retrieved, and the GO accessions were mapped to GO terms according to molecular function, biological process and cellular component ontologies (http://www.geneontology.org/). The remaining sequences that putatively encoded proteins were searched against the Swiss-Prot protein database (http://www.expasy.ch/sprot), the KEGG pathways database (Kanehisa et al. 2008) and the COG database (http://www.ncbi.nlm.nih.gov/COG/), applying a typical E-value cutoff level of <1E−5.

Analysis of differential EST expression
Gene expression levels were measured in the RNA-Seq analyses as numbers of reads per kilobase of exon region in a given gene per million mapped reads (RPKM) (Mortazavi et al. 2008). To identify genes regulated by salt stress, we determined the number of reads for each coding region in the control and salt-stress callus libraries (after normalization to account for the difference in number of total reads), and then calculated the ratio of reads in the two libraries. The statistical significance of the differential expression value for each gene was determined using the method described by Audic and Claverie (1997), and the results of all statistical tests were corrected for multiple testing with the Benjamini–Hochberg false discovery rate (FDR). Sequences were deemed to be significantly differentially expressed if the adjusted P value obtained by this method was <0.001 and there was at least a twofold change (>1 or <1 in log 2 ratio value) in sequence count between two libraries.

To ascertain the biological significance of the detected differences in gene expression profiles, ESTs with significant expression differences under salt stress were assigned to functional classifications according to corresponding Arabidopsis locus identifiers by BLASTX to Arabidopsis gene models (TAIR ver. 9, http://www.arabidopsis.org/), applying an E-value cutoff level of E−10. The data were then entered into a functional catalogue according to the Munich Information Center for Protein Sequence (MIPS) classification scheme (http://www.helmholtz-muenchen.de/en/ibis; see Table S4 available as Supplementary Data at Tree Physiology Online for a complete list).

Results and discussion
De novo assembly and quantitative assessment of the Illumina ESTs
After removing low-quality sequences and trimming adapter sequences, totals of 28 million, 55 million and 57 million clean reads of 75 bp were generated from the desert-grown tree, control-callus and salt-stressed callus cDNA libraries, respectively, in the Illumina GA runs (Table 1). The total length of the reads was >10.6 gigabases (Gb), equivalent to ~25-fold coverage of a genome of P. trichocarpa size. De novo assembly was carried out by SOAPdenovo, a genome assembly program developed specifically for next-generation short-read sequences (Li et al. 2010). The average contig size exceeded 170 bp in all three libraries. After using paired-end information to join the contigs into scaffolds and local assembly, we generated 44,593 scaffolds for desert-grown trees, 71,876 scaffolds for control-callus samples and 63,655 scaffolds for salt-stressed callus samples, with average lengths of 452, 591 and 589 nt, respectively. These scaffolds were assembled into 94,196 Unigenes by TGICL clustering tools. The sequencing throughout ranged from 1 to 24,406× (average 96×), and the average length of the Unigenes was 671 bp (Table 1). The size distribution and gap ratio of the assembly is shown in Figures S1 and S2 available as Supplementary Data at Tree Physiology Online.

We excluded possible non-coding RNA, untranslated region, microbial, fungal and virus sequences identified by comparing our Unigene sequences against entries in databases listed in the Materials and methods section. Finally, we identified a total of 86,777 high-quality Unigene sequences when 7419 possibly polluted sequences were excluded. All Unigene sequences...
obtained in this study can be accessed through File 1 available as Supplementary Data at Tree Physiology Online.

A BLAST comparison revealed that the assembly contains 92% of the NCBI <i>P. euphratica</i> ESTs and 90% of the TIGR poplar ESTs. In addition, we compared the recovered <i>P. euphratica</i> transcripts against <i>P. trichocarpa</i> transcripts from the genome assembly (Tuskan et al. 2006). Homologues of 40,249 sequences (~88% of the <i>P. trichocarpa</i> transcripts) were represented in our assembled transcriptome (Table 2). These results suggest that our assembly covered >90% of all <i>P. euphratica</i> transcripts.

### Functional annotation and characterization of transcripts

The entire Unigene sets were then annotated on the basis of similarities to known or putative sequences in the public databases. Using the best hits found by BLAST, an inferred putative function was assigned to the sequences and they were sorted into major functional categories (Figure 1). Among the 86,777 high-quality unique sequences, 58,499 (67.4%) had at least one significant match to an existing gene model in BLASTX searches (see Table S1 available as Supplementary Data at Tree Physiology Online). The remaining 28,278 (32.6%) unannotated Unigene sequences were further compared against the <i>P. trichocarpa</i> gene and genome sequences, and all available ESTs identified from 254 plant species. Using an <i>E</i>-value cutoff of <i>E</i>−10, at least 65 sequences did not show any significant match to known genes longer than 400 bp (see Table S2 available as Supplementary Data at Tree Physiology Online) and may be <i>P. euphratica</i>-specific novel ESTs.

The 100 most abundant transcripts distinctly differed among the three samples. In control callus, genes involved in auxin signaling, cell division and biogenesis were highly expressed, while stress-responsive genes (such as several antiporter, stress-induced transcription factor and diverse protease encoding genes) were strongly expressed in the salt-stressed callus. By contrast, the desert-growing trees strongly expressed photosynthesis-related genes and stress-related proteins, for example, aquaporin and glutathione transferase (see Table S3 available as Supplementary Data at Tree Physiology Online). Together, the recovered expression profiles largely support the phenotypic and physiological characterization of our three sample types.

### Transcript differences between control and salt-stressed callus

According to the applied criteria (twofold or more change and <i>P</i> < 0.001), 23,512 ESTs were identified as differentially expressed between the 24-h salt-stressed callus and control-callus samples: 7109 up-regulated and 16,403 down-regulated (Figure 2). This 24-h salt-stress treatment had apparently modified the expression of almost 27% of the total Unigenes, in accordance with previous observations in <i>Arabidopsis</i> (Kreps et al. 2002). Therefore, it is obvious that <i>P. euphratica</i> responds to abiotic stress by moderate reprogramming of its genome-scale transcriptome analysis.
transcriptome. Based on the MIPS functional catalogue, the 5775 differentially regulated ESTs with well-functional annotations and classifications were grouped into 18 main functional categories (Figure 3 and Table S5 available as Supplementary Data at Tree Physiology Online), the most heavily represented being cellular transport (10%), transcription (8%), cellular communication (10%) and metabolism (23%). Four important expression characteristics were further identified from the following detailed analyses.

Firstly, a total of 579 differentially regulated ESTs were categorized as transporters, many of which (246) were strongly up-regulated in response to salt stress (see Table S5 available as Supplementary Data at Tree Physiology Online). For example, the expressions of two EST sequences (Unigene45498_All and Unigene24445_All), which are homologous to Arabidopsis potassium ion transmembrane transporter (HAK5) and Na⁺/H⁺ antiporter (NHX1) genes, were up-regulated 48-fold and 66-fold, respectively. The 100 most abundant transcripts under salt stress encompassed sequences encoding 17 transport-related proteins, including sodium and potassium ion transmembrane transporters, and chloride channel and ABC transporters (see Table S3 available as Supplementary Data at Tree Physiology Online), corroborating previous findings (Wang et al. 2008, Sun et al. 2009, Ye et al. 2009). These transporters, which are important for maintaining and re-establishing homeostasis of the cytoplasm, are induced and highly expressed in response to salinity stress (Hasegawa et al. 2000).

Secondly, there were also considerable changes in the expression of genes involved in transcription and cellular communication. We found 454 differentially regulated ESTs involved in transcription: 132 up-regulated and 322 down-regulated. Five hundred and ninety-six of the differentially expressed ESTs were involved in cellular communication: 166 up-regulated and 430 down-regulated (see Table S5 available as Supplementary Data at Tree Physiology Online). Several of the transcription factors have homologues, e.g. AP2-EREBP and bZIP, which are known to be stress-induced in model herb species (Arabidopsis and rice) (Novillo et al. 2004, Oh et al. 2005). In addition, two transcription factors (PeDREB2 and PeSCL7) have been isolated from P. euphratica, and their over-expression has been recently shown to increase drought- and salt-stress tolerance in transgenic
plants (Chen et al. 2009, Ma et al. 2010). However, there have been no functional analyses of homologues for a greater number of the transcription factors (e.g. AP2-EREBP11 and bZIP53, homologues of Unigene45134_All, Unigene3935_All, respectively) that were up-regulated >60-fold in our salt-stressed callus samples. Functional analyses of these transcription factors (Figure 4 and see Table S6 available as Supplementary Data at Tree Physiology Online) should provide more information on the complex regulatory networks involved in responses of trees to salt stress (Yamaguchi-Shinozaki and Shinozaki 2006).

Thirdly, to obtain a better understanding of the functional significance of the differentially regulated ESTs, we focused on...
those that play important roles in well-characterized metabolic pathways. Key aspects of the salt-stress responses of *P. euphratica* callus cultured in vitro include the accumulation of proline and sugars, which provide osmotic balance (Watanabe et al. 2000). We found that the expression of genes related to 40 metabolic pathways was significantly changed under salt stress (*P* < 0.05), including genes involved in carbohydrate, amino acid, energy, lipid, secondary metabolite, cofactor and vitamin, terpenoid and polyketide metabolism (Table 3). Therefore, *P. euphratica* may establish a new energetic and

### Table 3. Statistical enrichment analysis for KEGG metabolic pathways.

<table>
<thead>
<tr>
<th>Metabolism pathway</th>
<th>Pathway ID</th>
<th>Number</th>
<th>Background</th>
<th>Up</th>
<th>Down</th>
<th>FDR P value</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Carbohydrate metabolism</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Starch and sucrose metabolism</td>
<td>ko:00500</td>
<td>724</td>
<td>92</td>
<td>275</td>
<td>5.35E–14</td>
<td></td>
</tr>
<tr>
<td>Pentose and glucuronate interconversions</td>
<td>ko:00040</td>
<td>204</td>
<td>30</td>
<td>98</td>
<td>5.35E–14</td>
<td></td>
</tr>
<tr>
<td>Ascorbate and aldarate metabolism</td>
<td>ko:00053</td>
<td>177</td>
<td>11</td>
<td>83</td>
<td>9.75E–08</td>
<td></td>
</tr>
<tr>
<td>Amino sugar and nucleotide sugar metabolism</td>
<td>ko:00520</td>
<td>295</td>
<td>33</td>
<td>88</td>
<td>5.49E–03</td>
<td></td>
</tr>
<tr>
<td><strong>Amino acid metabolism</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Phenylalanine metabolism</td>
<td>ko:00360</td>
<td>267</td>
<td>98</td>
<td>56</td>
<td>5.35E–14</td>
<td></td>
</tr>
<tr>
<td>Tyrosine metabolism</td>
<td>ko:00350</td>
<td>165</td>
<td>51</td>
<td>31</td>
<td>2.35E–05</td>
<td></td>
</tr>
<tr>
<td>Cysteine and methionine metabolism</td>
<td>ko:00270</td>
<td>330</td>
<td>69</td>
<td>77</td>
<td>2.79E–05</td>
<td></td>
</tr>
<tr>
<td>Alanine, aspartate and glutamate metabolism</td>
<td>ko:00250</td>
<td>185</td>
<td>57</td>
<td>31</td>
<td>7.97E–05</td>
<td></td>
</tr>
<tr>
<td><strong>Metabolism of other amino acids</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Cyanogenic acid metabolism</td>
<td>ko:00460</td>
<td>215</td>
<td>39</td>
<td>62</td>
<td>4.09E–05</td>
<td></td>
</tr>
<tr>
<td>Glutathione metabolism</td>
<td>ko:00480</td>
<td>199</td>
<td>52</td>
<td>31</td>
<td>1.66E–02</td>
<td></td>
</tr>
<tr>
<td><strong>Biosynthesis of other secondary metabolites</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Phenylpropanoid biosynthesis</td>
<td>ko:00940</td>
<td>567</td>
<td>157</td>
<td>132</td>
<td>5.35E–14</td>
<td></td>
</tr>
<tr>
<td>Flavonoid biosynthesis</td>
<td>ko:00941</td>
<td>309</td>
<td>79</td>
<td>70</td>
<td>5.19E–08</td>
<td></td>
</tr>
<tr>
<td>Stilbenoid, dihydropyrimidinoid and gingerol biosynthesis</td>
<td>ko:00945</td>
<td>282</td>
<td>83</td>
<td>55</td>
<td>5.53E–08</td>
<td></td>
</tr>
<tr>
<td>Isoquinoline alkaloid biosynthesis</td>
<td>ko:00950</td>
<td>61</td>
<td>27</td>
<td>8</td>
<td>3.53E–04</td>
<td></td>
</tr>
<tr>
<td>Flavonoid and flavanol biosynthesis</td>
<td>ko:00944</td>
<td>95</td>
<td>25</td>
<td>21</td>
<td>4.94E–03</td>
<td></td>
</tr>
<tr>
<td>Glucosinolate biosynthesis</td>
<td>ko:00966</td>
<td>63</td>
<td>14</td>
<td>19</td>
<td>5.06E–03</td>
<td></td>
</tr>
<tr>
<td>Tropane, piperidine and pyridine alkaloid biosynthesis</td>
<td>ko:00960</td>
<td>82</td>
<td>25</td>
<td>14</td>
<td>1.56E–02</td>
<td></td>
</tr>
<tr>
<td>Indole alkaloid biosynthesis</td>
<td>ko:00901</td>
<td>30</td>
<td>9</td>
<td>8</td>
<td>2.37E–02</td>
<td></td>
</tr>
<tr>
<td><strong>Energy metabolism</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Nitrogen metabolism</td>
<td>ko:00910</td>
<td>207</td>
<td>66</td>
<td>61</td>
<td>5.35E–14</td>
<td></td>
</tr>
<tr>
<td>Methane metabolism</td>
<td>ko:00680</td>
<td>181</td>
<td>51</td>
<td>45</td>
<td>7.90E–08</td>
<td></td>
</tr>
<tr>
<td>Sulphur metabolism</td>
<td>ko:00920</td>
<td>93</td>
<td>34</td>
<td>8</td>
<td>2.78E–02</td>
<td></td>
</tr>
<tr>
<td><strong>Lipid metabolism</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Steroid biosynthesis</td>
<td>ko:00100</td>
<td>96</td>
<td>14</td>
<td>51</td>
<td>2.86E–11</td>
<td></td>
</tr>
<tr>
<td>Linoleic acid metabolism</td>
<td>ko:00591</td>
<td>94</td>
<td>20</td>
<td>23</td>
<td>2.11E–02</td>
<td></td>
</tr>
<tr>
<td>Synthesis and degradation of ketone bodies</td>
<td>ko:00072</td>
<td>13</td>
<td>9</td>
<td>0</td>
<td>3.16E–02</td>
<td></td>
</tr>
<tr>
<td><strong>Metabolism of cofactors and vitamins</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Retinol metabolism</td>
<td>ko:00830</td>
<td>81</td>
<td>23</td>
<td>25</td>
<td>5.58E–06</td>
<td></td>
</tr>
<tr>
<td>Ubiquinone and other terpenoid-quinone biosynthesis</td>
<td>ko:00130</td>
<td>117</td>
<td>42</td>
<td>13</td>
<td>4.23E–03</td>
<td></td>
</tr>
<tr>
<td><strong>Metabolism of terpenoids and polyketides</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Zeatin biosynthesis</td>
<td>ko:00908</td>
<td>65</td>
<td>20</td>
<td>23</td>
<td>2.93E–07</td>
<td></td>
</tr>
<tr>
<td>Brassinosteroid biosynthesis</td>
<td>ko:00905</td>
<td>55</td>
<td>21</td>
<td>13</td>
<td>5.64E–05</td>
<td></td>
</tr>
<tr>
<td>Limonene and pinene degradation</td>
<td>ko:00903</td>
<td>286</td>
<td>65</td>
<td>58</td>
<td>5.74E–04</td>
<td></td>
</tr>
<tr>
<td>Carotenoid biosynthesis</td>
<td>ko:00906</td>
<td>165</td>
<td>37</td>
<td>32</td>
<td>2.72E–02</td>
<td></td>
</tr>
<tr>
<td>Diterpenoid biosynthesis</td>
<td>ko:00904</td>
<td>80</td>
<td>20</td>
<td>16</td>
<td>4.72E–02</td>
<td></td>
</tr>
<tr>
<td><strong>Xenobiotics biodegradation and metabolism</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Metabolism of xenobiotics by cytochrome P450</td>
<td>ko:00980</td>
<td>129</td>
<td>50</td>
<td>29</td>
<td>2.72E–10</td>
<td></td>
</tr>
<tr>
<td>Drug metabolism—cytochrome P450</td>
<td>ko:00982</td>
<td>139</td>
<td>53</td>
<td>30</td>
<td>4.88E–10</td>
<td></td>
</tr>
<tr>
<td>Gamma-hexachlorocyclohexane degradation</td>
<td>ko:00361</td>
<td>245</td>
<td>68</td>
<td>50</td>
<td>1.54E–06</td>
<td></td>
</tr>
<tr>
<td>Naphthalene and anthracene degradation</td>
<td>ko:00626</td>
<td>226</td>
<td>62</td>
<td>43</td>
<td>4.75E–05</td>
<td></td>
</tr>
<tr>
<td>3-Chloroacrylic acid degradation</td>
<td>ko:00641</td>
<td>79</td>
<td>19</td>
<td>20</td>
<td>7.05E–03</td>
<td></td>
</tr>
<tr>
<td>Fluorobenzoate degradation</td>
<td>ko:00364</td>
<td>25</td>
<td>9</td>
<td>6</td>
<td>2.10E–02</td>
<td></td>
</tr>
<tr>
<td>Fluorene degradation</td>
<td>ko:00628</td>
<td>23</td>
<td>9</td>
<td>5</td>
<td>2.21E–02</td>
<td></td>
</tr>
<tr>
<td>1,1,1-Trichloro-2,2-bis(4-chlorophenyl)ethane (DDT) degradation</td>
<td>ko:00351</td>
<td>23</td>
<td>9</td>
<td>5</td>
<td>2.21E–02</td>
<td></td>
</tr>
<tr>
<td>1- and 2-Methylnaphthalene degradation</td>
<td>ko:00624</td>
<td>66</td>
<td>17</td>
<td>14</td>
<td>3.86E–02</td>
<td></td>
</tr>
</tbody>
</table>

Downloaded from https://academic.oup.com/treephys/article-abstract/31/4/452/1644047 by guest on 23 February 2019
developmental equilibrium under salt stress (Bartels and Sunkar 2005, Ottow et al. 2005).

Finally, we found that numerous genes involved in abscisic acid (ABA) signalling regulation and biosynthesis were significantly up- or down-regulated in response to salt stress. This was consistent with expectations, since ABA plays crucial roles in drought and high-salinity stress responses, as well as diverse aspects of plant growth and development (Zhu 2002, Galvez-Valdivieso et al. 2009, Raghavendra et al. 2010). Furthermore, its concentration is known to increase during salt stress in *P. euphratica* (Chen et al. 2001, 2002). Three key ABA-binding receptors have been identified for pathways regulated by ABA: the flowering-time control protein (FCA), the Mg-chelatase H subunit and the G-protein-coupled receptor (GCR2) (Hirayama and Shinozaki 2007). Our results showed that homologues of genes encoding two of these receptors—FCA (Unigene484_All) and Mg-chelatase subunit H (Unigene2835_All)—were highly expressed in salt-stressed callus. Phosphatidic acid (PA), an important second messenger, is produced by phospholipase D (PLD) upon ABA treatment (Mishra et al. 2006), and two ESTs (Unigene7349_All and Unigene16355_All) that putatively encode homologues of *Arabidopsis* PLDα1 and PLDb1 were also significantly up-regulated in our salt-stressed callus sample. Furthermore, two protein-phosphatase 2C (PP2C)-like proteins encoded by ABA-insensitive loci AB11 and AB12 negatively regulate ABA responses in *Arabidopsis* (Zhang et al. 2004b), and we found that 46 ESTs homologous to *Arabidopsis* PP2C genes were differentially regulated: 20 were significantly up-regulated and 26 significantly down-regulated. Recent studies have also revealed connections between PP2C and SNF1-related kinases (SnRKs), particularly that SnRK2- and SnRK3-type kinases are important ABA regulators (Kim et al. 2003, Fujii et al. 2007). Accordingly, we found that one (Unigene38424_All) and two (Unigene14960_All, Unigene7553_All) ESTs that encode putative SnRK2.3- and SnRK3-type kinases, respectively, were up-regulated by salt stress.

The first step of the ABA-specific synthetic pathway is the conversion of zeaxanthin to violaxanthin, which is catalysed by zeaxanthin epoxidase (ZEP)—the first enzyme to be identified as an ABA biosynthetic enzyme (Seo and Koshiba 2002), and lack of which leads to ABA deficiency in mutants such as *Arabidopsis thaliana aba1* (Duckham et al. 1991). Our results showed that a possible AtZEP homologue (Unigene4188_All) was highly induced under salt stress. Another enzyme that plays a key role in ABA biosynthesis, 9-cis-epoxy-carotenoid dioxygenase (NCED) (Nambara and Marion-Poll 2005), has been found to improve drought and salt tolerance by increasing endogenous ABA levels (Iuchi et al. 2001). Our results clearly indicated that one EST (Unigene21682_All), putatively homologous to *AtNCED3*, was also up-regulated under salt stress. Hence, the differential expression profile indicates that key components of the ABA signalling and synthesis pathways are induced during salt stress, confirming the close relationship between ABA responses and salt stress in *P. euphratica*. Such a correlation is also consistent with the observed statistically significant enrichment of genes involved in KEGG metabolic pathways (Table 3), because carotenoid metabolism provides precursors for ABA biosynthesis (Seo and Koshiba 2002). However, the detailed mechanisms whereby components of ABA pathways participate in salt-stress responses in *P. euphratica* remain to be further elucidated.

**Conclusions and perspectives**

The development of new sequencing technologies in the past decade has provided opportunities for genome-wide transcriptomic analysis of non-model plants (Trick et al. 2009, Libault et al. 2010a, 2010b, Wu et al. 2010). Our results provide an extensive catalogue of the genes expressed in *P. euphratica*, and confirm that the applied sequence and assembly approach provides coverage of sufficient breadth and depth for potent transcriptome characterization and gene annotation, especially in non-model tree species. We have now increased the recorded number of EST sequences (13,845) sixfold for this species (Brosché et al. 2005). In addition, we have identified numerous potential stress-induced transporters, transcription factors and various components of ABA biosynthesis and signalling pathways that may be involved in the adaptation of this species to extreme environments (Chen et al. 2002, 2009, Umezawa et al. 2006, Wu et al. 2007). These genetic findings should be very useful for future studies of the molecular adaptation of this tree species to abiotic stress and genetic manipulation of other poplar species.

**Funding**

Financial support was provided by the National High-Tech Research and Development Program (863 Program) of China (grant no. 2009AA10Z101) and the Key Innovation Project of Ministry of Education of China (grant to L.J.Q).

**Supplementary data**

Supplementary data for this article are available at *Tree Physiology* Online.

**References**


Sun, J., M.J. Wang, M.Q. Ding et al. 2010. H(2)O(2) and cytosolic Ca(2+) signals triggered by the PM H(+)-coupled transport system mediate K(+) /Na(+) homeostasis in NaCl-stressed Populus euphratica cells. Plant Cell Environ. 33:943–958.


