Non-Targeted Metabolomics Reveals Patterns of Metabolic Changes during Poplar Seed Germination

Chunpu Qu 1,2, Jinyuan Chen 3, Lina Cao 2, Xiangjin Teng 1,4, Jinbo Li 2, Chengjun Yang 2, Xiuli Zhang 3, Yuhong Zhang 5, Guanjun Liu 1,2, and Zhiru Xu 1,3,*

1 State Key Laboratory of Tree Genetics and Breeding (Northeast Forestry University), School of Forestry, Northeast Forestry University, Harbin 150040, China
2 School of Forestry, Northeast Forestry University, Harbin 150040, China
3 College of Life Science, Northeast Forestry University, Harbin 150040, China
4 College of Life Science and Technology, Harbin Normal University, Harbin 150025, China
5 Key Laboratory of Forestry Plant Ecology of Ministry of Education, Northeast Forestry University, Harbin 150040, China
* Correspondence: xuzhiru2003@nefu.edu.cn; Tel.: +86-130-1900-7421

Received: 12 July 2019; Accepted: 2 August 2019; Published: 6 August 2019

Abstract: Research Highlights: This study was the first to use metabolomics techniques to investigate seed germination in poplar, a model woody plant. Our results lay a foundation for uncovering changes in metabolite levels during woody plant seed germination and for understanding the underlying mechanism. Background and Objectives: Poplar is a model woody plant. Because poplar can be easily propagated asexually, the molecular mechanism of poplar seed germination has not been well studied. However, long-term asexual reproduction of poplar leads to seedlings with weak resistance, high vulnerability to degradation, and reduced growth potential. Materials and Methods: The non-targeted metabolomics technique was used to analyze changing trends in metabolite contents during the poplar seed germination process. Results: We found that the number of differential metabolites increased as seed germination progressed. Metabolic pathway analysis of differential metabolites revealed that galactose metabolism and alanine, aspartate, and glutamate metabolism were significantly enriched during all germination periods. MapMan-based visual analysis of metabolic pathways of differential metabolites indicated that glutamine, glutamic acid, phenylalanine, arginine, and asparagine contents increased with germination time, with most metabolites related to glucose metabolism following similar trends. Contents of most metabolites related to the tricarboxylic acid cycle exhibited a fluctuating pattern. Conclusions: This study has revealed the major changes taking place in primary metabolite contents during poplar seed germination and has laid the foundation for elucidation of the molecular mechanism of poplar seed germination.

Keywords: poplar; seed germination; metabolite; non-targeted metabolomics

1. Introduction

Seed germination, the starting point of the higher plant life cycle, is a complex process involving changes at transcripational, translational, physiological, and biochemical levels that is highly susceptible to biotic and abiotic conditions [1–4]. As defined by Bewley et al. [5], seed germination covers the period from imbibition until the penetration of the seed coat by the hypocotyl. This process is divided into three phases based on the fresh weight of the seed after water absorption: Rapid water absorption, slow water absorption, and hypocotyl elongation. A broader definition of seed germination includes the seed post-germination period as well as the time from hypocotyl penetration of the seed coat until the cotyledon is fully unfolded [5–7].
Previous studies of seed germination have focused on physiological and biochemical changes rather than molecular mechanisms. More recent investigations based on the emerging field of omics have laid the foundation for revealing molecular mechanisms underlying the complex germination process. For instance, Girke et al. [8] performed an in-depth study of seed germination in the model plant *Arabidopsis* and detected approximately 2600 genes having seed-specific expression. Howell et al. [9] identified over 2000 transcripts with a unique, low-level expression pattern in dry seeds and observed a peak in expression levels at 1 or 3 h after imbibition. A wide range of studies have also been carried out in species such as soybean [10,11], rice [12,13], and broccoli [14].

Poplar is the common name for members of the genus *Populus*. Because poplar reproduces vegetatively quite readily, little research has been carried out on its sexual reproduction [15,16]. From the perspective of species evolution and environmental adaptation, long-term asexual reproduction is extremely unfavorable. In recent years, research has once again focused on the molecular basis of seed germination, the key process of sexual reproduction. Zhang et al. [17,18] used proteomics methods to determine that energy dependence, protein synthesis and degradation, cell defense, and rescue-related pathways are significantly correlated with poplar seed vigor. Qu et al. [19] revealed patterns of transcriptional and metabolic changes during poplar seed germination and identified some genes closely related to primary metabolic changes through targeted network correlation analysis. Zhang et al. [20] performed an integrative transcriptome analysis of three seed germination phases from *Populus euphratica* Oliv. and *Populus pruinosa* Schrenk., and identified the specifically expressed genes in each phase. No reports have appeared, however, concerning metabolic changes during poplar seed germination.

Transcriptomics uncovers physiological indicators of possible changes in an organism, whereas metabolomics reveals the terminal products present in the signaling pathway and thus reflects the physiological state of an organism at that time. Compared with other omics, metabolomics is more suitable for the isolation and study of specific markers at various stages of biological development because it is closer to the phenotype. No relevant research has been conducted on the patterns of metabolic change during poplar seed germination. In this study, we used metabolomics techniques to investigate these changes, with the goal of identifying biological pathways associated with changes in primary metabolites. We also aimed to provide an integrated view of these metabolic changes to reveal the underlying molecular mechanisms.

2. Materials and Methods

Seed germination metabolomic raw data were obtained from the State Key Laboratory of Forest Genetics and Breeding (Harbin, Heilongjiang province, China). The sampling schedule and methodology have been described previously [19], in short, seeds produced in the same year from superior poplar trees were selected from the greenhouse of Northeast Forestry University. The seeds were placed in a petri dish with filter paper and cultured at a constant temperature of 24 °C in dark. The seed germination was divided into five periods according to the difference in fresh weight after water absorption, periods of rapid and slow water absorption were defined as stages 2 (0.75 h) and 3 (6 h), respectively, while the hypocotyl extension period was defined as stage 4 (24 h), the dry seed phase was defined as stage 1 (0 h), and cotyledon unfolding was defined as stage 5 (48 h), and true-leaf unfolding was defined as stage 6 (144 h). In the collection of samples, seeds were blotted with absorbent paper to remove surface moisture, quickly wrapped in tin foil, frozen in liquid nitrogen, and stored at −80 °C.

First, metabolomic data were obtained from six quality control and 36 experimental samples, from which 3790 features were extracted and detected. After the elimination of outliers identified by the interquartile range method, missing values in the metabolic raw data were replaced by numbers corresponding to half of the minimum value. In addition, the overall normalization method was used in the data analysis [21]. The resulting three-dimensional data, including peak number, sample name, and normalized peak area were subjected to principal component analysis (PCA) and orthogonal
projections to latent structures-discriminate analysis (OPLS-DA) as implemented in SIMCA14.1 v14.1 (MKS Data Analytics Solutions, Umeå, Sweden) [22]. PCA was used to reveal the distribution of the original data, while OPLS-DA was applied to obtain a higher level of group separation and a better understanding of variables responsible for the classification. The OPLS-DA results were used to construct a loading plot showing the contribution of each variable to differences between groups. The loading plot, which also showed important variables situated far from the origin, was complex because of the large number of variables. To refine this analysis, the first principal component of variable importance in the projection (VIP) was obtained. VIP values exceeding 1.0 were first selected as changed metabolites; the remaining variables were assessed by Student’s t-test ($p > 0.05$), and those showing no significant difference between the two comparison groups were discarded [23]. In addition, metabolic pathway searches were carried out using KEGG (http://www.genome.jp/kegg/) [24] and MetaboAnalyst (http://www.metaboanalyst.ca/) commercial databases [25]. Screening and classification of primary metabolites was performed in MapMan according to the software instructions [26].

3. Results

3.1. Metabolic Data Analysis

To reveal physiological responses during poplar seed germination, we evaluated changes in metabolites at different stages of poplar seed germination by liquid chromatography-hybrid quadrupole time-of-flight mass spectrometry (LC-QTOF-MS). Significant chromatographic differences in seed germination were observed at different times. The raw data were manipulated as follows. First, missing values in the original data were recoded based on the minimum one-half method of numerical simulation. Next, the data were normalized using the total ion current of each sample [21]. Finally, we identified 285 metabolites (Supplementary Table S1). After logarithmic transformation and centralization formatting in SIMCA v14.1, the data were subjected to PCA and OPLS-DA. In the resulting PCA score plots (Figure 1), the raw data and different samples were well separated. As shown in Figure 2, large differences between samples were also uncovered by the application of the multivariate analysis method OPLS-DA.

3.2. Screening of Differential Metabolites and KEGG Analysis

Significant differences in metabolites between time periods were assessed using Student’s t-test ($p < 0.05$) and the first principal component of the OPLS-DA model of VIP > 1. The pathways associated with differential metabolites, which were identified using the Kyoto Encyclopedia of Genes and Genomes (KEGG) Pathway database (www.kegg.jp/kegg/pathway.html), are shown in Supplementary Table S2. Most of these pathways were related to primary metabolites. In most pathways, the number of differential metabolites was not consistent over time (Figure 3). For example, the number of metabolites associated with starch and sucrose metabolism, the citrate cycle (TCA cycle), and glycolysis/gluconeogenesis slowly increased as seed germination progressed, whereas the number related to the pentose phosphate pathway increased until 24 h and then remained unchanged. The number of lipid-related metabolites, such as those involved in glycerolipid metabolism and glycerophospholipid metabolism, fluctuated and exhibited a decrease at approximately 24 h. Nucleic acid-related metabolites increased as germination progressed, while the number of metabolites associated with pyrimidine metabolism decreased after 48 h. The number of metabolites associated with amino acid pathways, which was significantly higher than that of other pathways, slowly increased over the course of seed germination; central amino acid metabolism-related compounds followed a similar trend. Metabolites related to zeatin biosynthesis were not detected until 6 h and then remained constant.
Figure 1. Principal component analysis (PCA) score plots of metabolic profiles during different seed germination times. (A–E) PCA score plots of seed germination at 0 and 0.75 h (A), 0 and 6 h (B), 0 and 24 h (C), 0 and 48 h (D), and 0 and 144 h (E).

3.3. Pathway Analysis of Differential Metabolites

Metabolites undergoing changes in abundance over the course of seed germination (as assessed by five different timepoint comparisons) were mapped to 38 KEGG biological pathways. Pathways significantly enriched in differential metabolites were identified using the criteria of $p < 0.05$. The results of the analysis are shown in Figure 4 and Supplementary Table S3. At 0.75 h, pathways related to galactose metabolism, amino sugar and nucleotide sugar metabolism, glycerolipid metabolism, and alanine, aspartate, and glutamate metabolism were significantly enriched with the monoterpenoid biosynthesis pathway having the highest impact value. Pathways associated with glyoxylate and dicarboxylate metabolism, galactose metabolism, amino sugar and nucleotide sugar metabolism, and alanine, aspartate, and glutamate metabolism were significantly enriched 6 h after seed germination, and glyoxylate and dicarboxylate metabolic pathways had the highest impact values. At 24 h, the aminoacyl-tRNA biosynthesis pathway was also significantly enriched, while the alanine, aspartate, and glutamate metabolic pathway had the highest impact value. At 48 h, pathways related to
phenylalanine, tyrosine, and tryptophan biosynthesis and carbon fixation in photosynthetic organisms were also enriched, and the pathway with the highest impact value was still the alanine, aspartate, and glutamate metabolic pathway. At 144 h, the phenylalanine metabolic pathway had a highest impact value, and pathways associated with galactose metabolism, alanine, aspartate, and glutamate metabolism, and carbon fixation in photosynthetic organisms were still significantly enriched. In general, the metabolic pathways enriched in differential metabolites at different poplar seed germination stages were relatively conserved.

Figure 2. Orthogonal partial least squares-discriminate analysis (OPLS-DA) score plots of metabolic profiles during different seed germination times. (A–E) OPLS-DA score plots of seed germination at 0 and 0.75 h (A), 0 and 6 h (B), 0 and 24 h (C), 0 and 48 h (D), and 0 and 144 h (E).
followed a similar trend. Metabolites related to zeatin biosynthesis were not detected until 6 h and then remained constant.

Figure 3. Changes in the number of differential metabolites in different pathways during seed germination. The abscissa represents the number of metabolites.

3.3. Pathway Analysis of Differential Metabolites

Metabolites undergoing changes in abundance over the course of seed germination (as assessed by five different timepoint comparisons) were mapped to 38 KEGG biological pathways. Pathways significantly enriched in differential metabolites were identified using the criteria of \( p < 0.05 \). The results of the analysis are shown in Figure 4 and Supplementary Table S3. At 0.75 h, pathways related to galactose metabolism, amino sugar and nucleotide sugar metabolism, glycerolipid metabolism, and alanine, aspartate, and glutamate metabolism were significantly enriched with the monoterpenoid biosynthesis pathway having the highest impact value. Pathways associated with glyoxylate and dicarboxylate metabolism, galactose metabolism, amino sugar and nucleotide sugar metabolism, and alanine, aspartate, and glutamate metabolism were significantly enriched 6 h after seed germination, and glyoxylate and dicarboxylate metabolic pathways had the highest impact values. At 24 h, the aminoacyl-tRNA biosynthesis pathway was also significantly enriched, while the alanine, aspartate, and glutamate metabolic pathway had the highest impact value. At 48 h, pathways related to phenylalanine, tyrosine, and tryptophan biosynthesis and carbon fixation in photosynthetic organisms were also enriched, and the pathway with the highest impact value was still the alanine, aspartate, and glutamate metabolic pathway. At 144 h, the phenylalanine metabolic pathway had the highest impact value, and pathways associated with galactose metabolism, alanine, aspartate, and glutamate metabolism, and carbon fixation in photosynthetic organisms were still significantly enriched. In general, the metabolic pathways enriched in differential metabolites at different poplar seed germination stages were relatively conserved.

3.4. Metabolite Network Visualization

The MapMan tool was used to visualize pathways of differential metabolites. Patterns of primary metabolite changes during seed germination are displayed in Figure 5. As shown in the figure, contents of metabolites related to glycolysis and upstream pathways, such as dihydroxyacetone phosphate and glucose-1-phosphate, were mostly elevated in seeds 48 h before germination, whereas sucrose and raffinose contents decreased. A similar situation was also observed in the case of ribulose-5-phosphate in the OPP. In contrast, the contents of metabolites related to the tricarboxylic acid cycle mostly fluctuated; except for ketoglutaric, citric, isocitric, aconitic, and succinic acids, their contents were low at approximately 6 h. In the urea cycle, metabolites other than arginine, ornithine, Argininosuccinic acid had higher contents before and after 0.75 h. Under amino acid metabolism, the contents of all amino acids in seeds increased during germination. The pattern of changes in contents of metabolites associated with lipid and nucleic acid metabolism appears to be more complicated. Fold changes in metabolite abundances are shown in Supplementary Table S4.
Figure 4. Pathway enrichment of differential metabolites. Pathway impact, which is an influence factor calculated from pathway topological analysis, is plotted on the abscissa, and $-\log(p)$, the negative logarithm of the $p$-value from pathway enrichment analysis, is plotted on the ordinate. Circle size is proportional to the degree of influence calculated from the topological analysis: The larger the circle, the larger the impact factor. The color of a circle reflects the negative logarithm of the $p$-value obtained from the enrichment analysis: The darker the color, the larger the absolute value of $-\log(p)$ and the more significant the enrichment. (A–E) Pathway enrichment during seed germination at 0.75 h (A), 6 h (B), 24 h (C), 48 h (D), and 144 h (E) vs. 0 h.
Figure 5. Changes in contents of metabolites related to primary pathways over the course of poplar seed germination. Different colors represent different metabolic pathways and polylines represent changes in metabolite abundance. 1–6 represents 0 h, 0.75 h, 6 h, 24 h, 48 h and 144 h, respectively.
4. Discussion

Untargeted metabolomics refers to the unbiased detection, using techniques such as LC-MS, of all small-molecule metabolites (relative molecular weights mainly <1000 Da) present before or after stimulation or disturbance of cells, tissues, organs, or organisms. Bioinformatics and pathway analyses of differential metabolites are used to reveal the physiological mechanisms underlying dynamic changes in endogenous small-molecule compounds. In this study, we used the non-targeted metabolomics technique to analyze trends in metabolite contents during poplar germination, which has not been previously investigated. Our results can serve as a foundation for revealing physiological changes in the seed germination process of woody plants.

Sucrose, one of the main storage carbohydrates in poplar seeds, is consumed throughout almost the entire period from imbibition to radicle protrusion in Arabidopsis seeds. Sucrose can be hydrolyzed to free hexoses, such as glucose and fructose, thereby affecting carbohydrate levels [27]. Previous studies have found a high positive correlation between changes in sucrose concentration and the expressions of sucrose-associated genes [28]. In our study, we observed that sucrose levels decreased rapidly after 0.75 h (Figure 5). Raffinose, another carbohydrate reserve whose role has recently been demonstrated [10,29,30], showed the same pattern of change. Another interesting phenomenon was the change in the content of galactose-1-phosphate associated with carbohydrate catabolism, which increased between 0.75 and 48 h, and that of dihydroxyacetone phosphate related to carbohydrate synthesis, which increased before 48 h (Figure 5). We believe that storage carbohydrates are initially mobilized before 48 h, consistent with previous studies in barley and rice [31,32]. Hexose phosphates, which are products of carbohydrate decomposition, pool for biosynthesis; these include glucose-1-phosphate, which had a higher content after 6 h. Noteworthily, ribulose-5-phosphate, AMP, GMP, succinic acid, and some lipids had higher contents before seed germination. Confirmation of whether these compounds act as storage substances requires further experimentation.

Two pools feed plant carbohydrate metabolism: One composed of hexose phosphates and the other consisting of pentose and triose phosphates. The direction of flow through these pools depends on cellular requirements. The hexose phosphate pool comprises three metabolic intermediates: Glucose-6-phosphate, glucose-1-phosphate, and fructose-6-phosphate. The level of glucose-6-phosphate has been found to increase 10-fold in wild oat seeds 8 h after imbibition. Hexose or its derivatives, such as fructose-6-phosphate and glucose, also exhibit increasing trends [7,33]. As shown in Figure 5, glucose-1-phosphate levels increased after seed germination until 48 h in our study. Further downstream dihydroxyacetone phosphate in the glycolytic pathway also displayed similar changes, even in regards to the transcription levels of glycolytic-related genes [19]. Due to the net production of 4 mol ATP from each mol of glucose in the glycolytic pathway, the accumulation of glucose-1-phosphate in imbibed seeds may reflect a positive energetic status of germinating seeds [34].

Storage lipid reserve mobilization is commonly accepted to commence after germination [5,6,35], with storage lipid degradation-associated transcripts also increased at this stage [36]. Triacylglycerol, the major storage lipid, is broken down to fatty acids and glycerol during early germination in Arabidopsis seeds. Glycerol is used as an energy source via the glycerol-3-phosphate (G3P) shunt. As shown in Figure 5, G3P content was high at 0.75 h, and we presume that stored lipids had started to undergo catabolism, thereby providing raw materials and energy.

Although de novo protein synthesis is required for germination, levels of free amino acids in dry seeds are not sufficient for protein synthesis during germination [6]. In barley, most amino acid biosynthetic genes are expressed 48 h after imbibition [31]. In our study, levels of most amino acids increased from late germination to seedling stages. We thus hypothesize that free amino acids released by the degradation of storage proteins are mainly involved in new protein synthesis and then enter the amino acid pool for synthesis metabolism. The exceptions are ornithine and argininosuccinic acid, whose levels in our study were increased at 0.75 h and then decreased. Previous investigations have shown that pathways that shuttle alternative amino acids to the TCA for energy production are regulated by the availability of sugars [33,37,38]. In germinating yellow lupin seeds, for example,
arginine enters the TCA cycle via the urea recycling pathway, which is induced by sugar starvation [39]. When sugars are insufficient during germination, we believe that early protein degradation of the carbon skeleton may be related to the urea cycle and subsequently the TCA pathway, which synthesize substances and produce energy necessary for seed germination.

5. Conclusions

Seed germination in poplar is a very complex process involving significant changes in various metabolites. Our study has provided important insights into the dynamics of metabolites during seed germination in woody plants. Our findings should aid the development of new markers for seed quality improvement to enhance the environmental adaptation of forest tree seeds.

Supplementary Materials: The following are available online at http://www.mdpi.com/1999-4907/10/8/659/s1,
Table S1: Information on 285 metabolites annotated by qualitative matching analysis. Table S2: KEGG annotations of metabolites. Table S3: Pathways enriched at different seed germination timepoints. Table S4: Fold changes in metabolite abundances revealed by metabolite network visualization.

Author Contributions: C.Q. and Z.X. conceived and designed the study, C.Q. and J.C. performed most of the experiments, L.C., X.T., and J.L. conducted the sampling, C.Y. and X.Z. performed bioinformatics calculations, G.L. and Z.X. processed and analyzed the data, and Y.Z., C.Q., and G.L. wrote the manuscript.

Funding: This research was supported by the National Natural Science Foundation of China (31570648), Natural Science Foundation of Heilongjiang Province (C2018009). Special Fund for Basic Scientific research operation Fee of Central University (2572017EA05) and The 111 project (B16010).

Conflicts of Interest: The authors declare no competing financial interests.

References


32. He, D.; Han, C.; Yao, J.; Shen, S.; Yang, P. Constructing the metabolic and regulatory pathways in germinating rice seeds through proteomic approach. *Proteomics* 2011, 11, 2693–2713. [CrossRef]


37. Lehmann, T.; Ratajczak, L. The pivotal role of glutamate dehydrogenase (GDH) in the mobilization of N and C from storage material to asparagine in germinating seeds of yellow lupine. *J. Plant Physiol.* **2008**, *165*, 149–158. [CrossRef] [PubMed]
