South African Journal of Botany 119 (2018) 286–294

Contents lists available at ScienceDirect

South African Journal of Botany

journal homepage: www.elsevier.com/locate/sajb

Transcriptomic approach to address low germination rate in *Cyclobalnopsis gilva* seeds

M. Zaynab [a], D. Pan [a], M. Fatima [c], S. Chen [b,*], W. Chen [a,*]

[a] College of Life Sciences, Fujian Agriculture and Forestry University, Fuzhou 350002, PRChina
[b] College of Forestry, Fujian Agriculture and Forestry University, Fuzhou 350002, PRChina
[c] College of Crop Science, Fujian Agriculture and Forestry University, Fuzhou 350002, PRChina

ARTICLE INFO

Article history:
Received 20 June 2018
Received in revised form 2 September 2018
Accepted 19 September 2018
Available online xxxx

Edited by I Demir

Keywords:
Cyclobalnopsis gilva
Carbohydrate metabolism
Glycolysis and TCA cycle
Sucrose and starch metabolism

ABSTRACT

Cyclobalnopsis gilva is a woody plant with lower seed germination rate in its natural habitats as compared with other plants. The regulation at transcriptional level to of seed germination in *C.gilva* is yet to be discovered. Illumina HiSeq 2500 platform was performed to investigate the transcriptome of germinated and non-germinated seeds of *C.gilva*. A total of 40,782 differentially expressed unigenes (DEGs) were identified in germinated and non-germinated seeds and compared with the control. Carbohydrate metabolisms were enriched significantly through the analysis of KEGG pathways of DEGs. The physiological parameters and mRNA expression levels relevant to carbohydrate metabolisms including glycolysis and TCA cycle, and sucrose and starch metabolism were found to be poor in non-germinated seeds. The consistent results of the transcriptomic data, physiological analyses, and mRNA expression levels suggest that *C.gilva* seeds fail to germinate due to the lower activity of carbohydrate metabolism. Our findings will expand the knowledge on the transcriptional responses of respiratory pathways, sucrose, and starch metabolism related to lower seed germination rate of *C.gilva*.

1. Introduction

Seed germination is the first important stage in the plant life cycle (Gioria et al., 2016). In general, seed germination is described as a process including three phases: Phase I is a rapid initial phase, namely seed imbibition during which storage starch and protein in the embryo changes gradually, thus glycolysis and anaerobic respiration begin to occur; Phase II is a plateau phase during which the uptake of water results in an increased oxygen supply, hence mitochondria becomes activated; Phase III is the post-germination phase during which the embryonic axes elongate along with the radicle development (Gimeno-Gilles et al., 2009). Energy production plays vital roles in seed germination. Energy is mainly generated by anaerobic respiration in the beginning of seed germination and subsequently by aerobic respiration along with an intensification of water uptake (Pergo and Ishii-Iwamoto, 2011). Physiological processes during germination require considerable energy (Bao et al., 2017) and the main energy source for cellular metabolism is glucose, which is catabolized by subsequent processes—glycolysis and tricarboxylic acid (TCA) cycle, and finally oxidative phosphorylation to produce ATP. Energy production pathways like respiration are important in the whole-seed germination by providing required ATP.

Seed germination is highly correlated with seedling survival rate as well as subsequent seedling growth and development, thus directly affecting the quality of the seedlings. Studies on seed germination have mainly focused on aspects of seed physiology and biochemistry (Noman et al., 2015). It is reported that major physiological changes and reactivation of metabolic processes take place during seed germination (Rajjou et al., 2006; Holdsworth et al., 2008; Zaynab et al., 2017). ATPases are synthesized to support normal cellular function during the late phase of germination (Chen and Bradford, 2000; Mei and Song, 2010). Transcriptomic analysis indicated that the enzymes in TCA cycle are reported to be up-regulated after imbibition (Yao et al., 2016). We have a basic understanding of energy provision in seed germination, but how those gene expressions are related to energy production pathways regulating seed germination is largely unknown (He and Yang, 2013; Mangrauthia et al., 2016). High-throughputRNA-sequencing (RNA-seq) method is opening new horizons in the field of transcripts participated in specific biological metabolic processes (Mangrauthia et al., 2016). Over the years, model plant species such as *Oryza sativa*, barley, maize, and *Arabidopsis thaliana* have been analyzed through transcriptomic studies with respect to seed germination (Rajjou et al., 2004; Nakabayashi et al., 2005; Radchuk et al., 2007; Kimura and Nambara, 2010; Okamoto et al., 2010). A daunting challenge is to provide detailed transcriptomic profiling of germination process in woody plants. *Cyclobalnopsis gilva,* belongs to *Fagaceae*, a woody plant. It is native to East Asia and abundantly found in Japan and China.

* Corresponding authors.
E-mail addresses: fjcsp@126.com (S. Chen), weichen@fafu.edu.cn (W. Chen).

https://doi.org/10.1016/j.sajb.2018.09.024
0254-6299/

The furniture industry in the region prefers its wood due to its hardness. Seeds from *C.gilva* have a long quiescence, so sand-burying can break the dormancy of the quiescent seeds to promote germination. As compared with other plants in the same habitat and region, the germination rate of *C.gilva* seed is very low (less than 50%) affecting the seedling survival (Zaynab et al., 2017). Transcriptomic study of *C.gilva* will give new insights into the genes involved in energy metabolism. The current study was carried out to understand transcriptional responses to low germination rate in *C.gilva* seeds. This study will serve as a public information and resource for future functional genomic and genetic studies on the seed germination of *C.gilva*.

2. Materials and methods

2.1. Plant materials

Cyclobalnopsis gilva seeds were collected from Fujian province, China. Seeds of *C.gilva* were put in a container of water. Seeds sank in water were considered as quality seeds and were collected to do germination experiments, while floating seeds were removed. Five hundred seeds were selected for the experiment. Seeds were sown in wet sand (2 m length × 1 m width × 0.5 m height) in an open field with white plastic film roof (4 m length × 2 m width × 1 m height) for 60 d with an average temperature of 16 °C and a humidity of 60%, which is the favorable condition for *C.gilva* to grow from December 2016 to February 2017. The investigation of seed germination rate was calculated after 60-d wet sand treatment. Three biological replicates were conducted for this study. Each replicate (30 seeds), including control seeds (seeds not sown in wet sand), germinated seeds (namely SG), and non-germinated ones (namely NG), was collected and immediately immersed in liquid N_2 and stored at −80 °C for future experiments.

2.2. RNA extraction and construction of cDNA libraries for transcriptome sequencing

Total RNA from *C.gilva* seed tissue was extracted through a modified CTAB method (Chan et al., 2007). Agilent 2100 bioanalyzer was employed to assess the integrity of the total extracted RNAs. Two cDNA libraries were constructed, and sequencing of transcriptomes was carried out by Biomarker Technologies Co., Ltd. (Beijing, China). For purification and enrichment of mRNAs from the total RNAs of each sample, oligo (dT) magnetic beads were used according to instructions of Illumina manufacturer. The enriched mRNAs were short fragmented, and these fragments were reverse transcribed for the first and the second strands cDNA synthesis. Afterward, adopters were ligated with these double-stranded fragments and further appropriate DNA fragments were used as templates for PCR amplification.

2.3. Illumina sequencing, assembly, and annotation

Sequencing of cDNA libraries was carried out using Illumina HiSeq™ 2500 and raw 100 nt paired end reads were generated. To filter out the reads, quality parameters including sequence duplication level, GC-contents, Q20, and Q30 were used and low-quality reads. The reads containing poly-N and those containing adapter were eliminated, and the high quality clean reads were obtained from the raw reads. De novo transcriptome assembly of clean reads was executed using the Trinity assembly program with default parameters (Grabherr et al., 2011). Trinity software contains three components: Inchworm, Chrysalis and Butterfly (http:// trinityrnaseq.sourceforge.net/). Initially, Inchworm created a k-mer dictionary by breaking all sequence reads ("k-mer," a sequence of a fixed length of k nucleotides, in practice, $k = 25$ bp). After removing error-containing, low-complexity and singleton k-mers, the most frequent k-mer was selected for a contig assembly. The contigs were obtained until the two sides of the sequence could not be extended with a k-1 overlap. Afterward, Chrysalis module was used to gather the linear contigs and construct de Bruijn graphs. Transcripts were assembled from all the generated contigs. Lastly, the butterfly module was used to analyze the de Bruijn graphs and generate transcript sequences. The main transcripts made up of more than 200 bp were selected as unigenes. BLASTX alignment (Altschul et al., 1997) was carried out between the generated unisequences and the public protein databases: non-redundant (Nr) (Deng et al., 2006), Swiss-Prot, Clusters of Orthologous Groups (COGs), Kyoto Encyclopedia of Genes and Genomes (KEGG) (Kanehisa et al., 2004), and Gene Ontology (GO) protein database (Ashburner et al., 2000).

2.4. Analysis of quantitative real-timePCR (qRT-PCR)

Total RNA from *C.gilva* seeds was extracted through a modified CTAB method (Chan et al., 2007). According to the transcriptomic data, primers were designed by using Primer Premier 5.0 software (Supplementary Table 1). RT-qPCR was conducted using SYBR Premix Ex Taq™ (Tli RNaseH Plus) kit (TaKaRa, Japan) at 95 °C for 30 s, followed by 40 cycles at 95 °C for 5 s, and 60 °C for 30 s. The whole thermocycling process was conducted in a BioRad CFX96 real-timePCR detection system. *Actin* gene was chosen as the internal standard to normalize gene expression. RT-qPCR was performed in three replicates of all samples. The $2^{-\Delta\Delta Ct}$ method was employed to compute the relative quantitative gene expression (Livak and Schmittgen, 2001).

2.5. Analysis of physiological parameters

The activities of α-amylase, β-amylase, PFK, PGK, PK, and NAD-MDH were evaluated using commercial reagent kits from Suzhou Comin Biotechnology Co., Ltd. (Suzhou, China) according to instructions from the manufacturer. Soluble sugar and starch contents were assayed by adopting the method described by Tollenaar and Daynard (1978). Each data point represents an average ± standard deviation (SD) of three repeated experiments.

2.6. Statistical analysis

Analysis of variance (ANOVA) was used to statistically analyze the data. The least significant difference (LSD) was applied to determine the significant differences among group means at $P < 0.05$.

3. Results

3.1. Sequence analysis and assembly from C.gilva seeds

To break dormancy, seeds of *C.gilva* were germinated in wet sand. The germination rate of *C.gilva* seeds was found to be as low as 37.68% in this study (Supplementary Table 2). Therefore, transcriptome analysis of *C.gilva* seeds was performed to understand the transcriptional changes in *C.gilva* during seed germination. After removing lower-quality reads, more than 6.77 billion clean reads with acceptable quality were generated in three samples (Table 1). After cleaning and quality checks, Q30 and GC percentages exceeded 94% and 43%, respectively (Table 1).

3.2. Analysis of differentially expressed unigenes (DEGs) by RNA-Seq

Compared with the control in the present study, a total of 40,782 unigenes were differentially expressed in the SG and non-germinated ones NG, among which 8146 unigenes were up-regulated while 8951 unigenes were down-regulated in SG samples, and 14,260 unigenes were up-regulated while 9425 unigenes were down-regulated in NG samples (Table 2).

After sequencing of transcriptome gene ontology was carried out on the basis of biological processes, molecular functions, and cellular components, GO analysis and annotation were done using three categories