Industrial Crops & Products 124 (2018) 510–518

Contents lists available at ScienceDirect

Industrial Crops & Products

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

Mapping podophyllotoxin biosynthesis and growth-related transcripts with high elevation in *Sinopodophyllum hexandrum*

Mengfei Li[a,b], Ping Sun[a], Tianlan Kang[c], Hua Xing[a], Delong Yang[a], Jinlin Zhang[a,d,**], Paul W. Paré[b,*]

[a] Gansu Provincial Key Lab of Arid Land Crop Science/College of Life Science and Technology, Gansu Agricultural University, Lanzhou 730070, PR China
[b] Department of Chemistry and Biochemistry, Texas Tech University, TX 79409, USA
[c] Institute of Industrial Crop and Promotion, Lanzhou 730000, PR China
[d] College of Pastoral Agriculture Science and Technology, Lanzhou University, Lanzhou 730000, PR China

ARTICLE INFO

Keywords:
Sinopodophyllum hexandrum
High elevation
Growth promotion
Podophyllotoxin biosynthesis
Transcriptomic analysis

ABSTRACT

Sinopodophyllum hexandrum, a perennial rhizomatous herb that produces the anti-cancer metabolite podophyllotoxin (PPT), is distributed in the Himalayan region. While greater plant growth and PPT accumulation is observed at higher compared with lower elevations, the mechanism responsible for elevation-dependent growth promotion and PPT biosynthesis has not been reported. Here, the *de novo* sequenced transcriptome of *S. hexandrum* is described. When plants were grown at 3300 versus 2300 m above sea level (asl), a total of 53,691 unigenes were generated and 736 differentially expressed genes (DEGs) were observed. DEGs (357) were identified with 234 characterized genes showing greater up than down regulation, 171 versus 63, respectively. Functional annotation classified 23 DEGs involved in PPT biosynthesis, including phenylpropanoid enzymes: phenylalanine ammonia layse (*PAL*), cinnamyl alcohol-dehydrogenase (*CAD*), dirigent protein (*DIR*), pinoresinol-lariciresinol reductase (*PLR*), CYP719A23, CYP71CU1, and 2-oxoglutarate/Fe(II)-dependent dioxygenase (*2-ODD*) which directly participate in PPT biosynthesis. Over 200 DEGs that participate in plant growth and development are categorized into: cell morphogenesis, bio-signaling, primary metabolism, photosynthesis/energy, transcription/polynucleotide metabolism, translation/protein synthesis, transport, and stress tolerance. This transcriptomic analysis provides insight into the mechanism that enhances plant growth and PPT accumulation in *S. hexandrum* under higher elevation conditions.

1. Introduction

Sinopodophyllum hexandrum (family Berberidaceae) is a perennial rhizomatous herb distributed in alpine Himalayan regions at elevations from 1300 to 4500 m asl (Alam and Naik, 2009; Wu et al., 2015; Li et al., 2018). Rhizomes are the major source of highly valued podophyllotoxin (PPT), which is used worldwide for the treatment of skin lesions and employed as a natural product precursor for the semi-synthetic production of anticancer chemotherapies including: etoposide, etopophos, and teniposide (Canela et al., 2000; Lv and Xu, 2011; Marques et al., 2013). It is currently an endangered species because of unregulated uprooting of wild plants to meet the ever-increasing demand for PPT extraction in the pharmaceutical industry (Choudhary et al., 1998; Gupta and Dutta, 2011).

Large scale farming is urgent and necessary to increase PPT production and in turn reduce unregulated harvesting of *S. hexandrum* plants (Kumari et al., 2016), since PPT content is greater in wild or field grown plants (Kitchlu et al., 2011; Liu et al., 2015; Li et al., 2018) than plants cultivated under glasshouse conditions (Pandey et al., 2007; Kushwaha et al., 2012; Seegers et al., 2017) or in an in vitro cell-culture system (Chattopadhyay et al., 2003; Rajesh et al., 2013, 2014). Plant growth and PPT accumulation are influenced by environmental conditions including: soil pH and nutrients, temperature, light (ultraviolet, duration, and irradiance), precipitation, and geographical factors

Abbreviations: asl, above sea level; DEGs, differentially expressed genes; DR, down-regulated; GO, gene ontology; KEGG, Kyoto Encyclopedia of Genes and Genomes; KOG, euKaryotic orthologous groups of proteins; NCBI, National Center for Biotechnology Information; NR, non-redundant protein database; Pfam, protein families; PPT, podophyllotoxin; qRT-PCR, quantitative reverse transcription polymerase chain reaction; RPKM, reads per kb per million; UR, up-regulated

* Corresponding author.
** Corresponding author.
E-mail addresses: zlzhang@lzu.edu.cn (J. Zhang), paul.pare@ttu.edu (P.W. Paré).

https://doi.org/10.1016/j.indcrop.2018.08.007
Received 19 May 2018; Received in revised form 6 July 2018; Accepted 5 August 2018

(slope, slope direction, and elevation) (Alam and Naik, 2009; Kushwaha et al., 2012; Li et al., 2013; Kumari et al., 2014; Seegers et al., 2017; Liu et al., 2015; Yang et al., 2016; Li et al., 2018). To date, elevation has been shown to be a determinate factor in PPT accumulation (Li et al., 2018). Based on either a dry weight or per plant basis, plants grown at higher elevations between 2800 to 3600 m asl contained approximate a 2-fold greater PPT content than plants at 1300 to 2300 m asl (Alam and Naik, 2009; Alam et al., 2009; Li et al., 2013; Pandey et al., 2015; Li et al., 2018). Elevation-dependent changes responsible for enhancing growth and PPT accumulation have not been identified.

To understand the temperature mediated molecular responses including those associated with PPT biosynthesis, an examination of transcriptional changes at two temperatures has been reported (Kumari et al., 2014). In this study, plants were grown in a green house in which the temperature was held constant at either 15 or 25 °C and light levels were supplemental with a photosynthetic photon flux density of 200 μmol/m^2/s for 16 h per day. While this study has provided an increased understand of the role of temperature in regulating PPT biosynthesis in *S. hexandrum*, the role of elevation in enhancing plant growth and PPT accumulation has not been directly addressed. To more closely match natural environmental conditions that can be introduced with differing elevations such as precipitation, light duration/quality, and temperature fluctuation/extremes, this study examines transcriptional alterations for plants grown at two different elevations. The complexity of how different elevations impact gene expression and influence growth and PPT accumulation can now be probed.

2. Materials and methods

2.1. Plant material

Seeds of *S. hexandrum* were collected from plants grown in Gannan Tibetan Autonomous Prefecture in September 2012 and sown at the same site of 3300 m asl in May 2013 as previously reported (Li et al., 2018). Some seedlings were transplanted to Weiyuan of Gansu province at 2300 m asl site in May 2014 and allowed to adapt for 2 years. Leaves from two experimental sites (Fig. S1) were collected (n = 60 total plants) in July 2016, pooled and rapidly frozen in liquid nitrogen for transcriptional analysis.

2.2. Total RNA isolation and Illumina sequencing

Total RNA samples were isolated using Trizol reagent (Invitrogen, CA, USA) according to the manufacturer's protocol and then purified using RNase-free DNase I (TakaRa, Dalian, China). Samples were quantified using both Bioanalyzer Chip RNA 7500 series II (Agilent Technologies, USA) and Nanodrop 1000 (NanoDrop Technologies, USA) instrumentation. Poly-A mRNA was enriched and then used to prepare paired-end cDNA library (2 × 126 nt) according to the manufacturer specifications (Illumina, San Diego, USA). To remove poor quality reads that can distort the sequence analysis, sequences were trimmed with 5′ and 3′ prime ends no more than 25% and until the first 25 bases contain less than 3 ambiguities. Paired-end sequencing was performed using an Illumina Hiseq 2000 platform.

2.3. Sequence assembly and functional annotation

De novo assembly was carried out using Trinity software (version 2.0.6) (Grabherr et al., 2008), with further filtering by quality score < 30 and length < 60 bp. Clean reads were mapped back onto the corresponding contigs and then all contigs were assembled to produce unigenes with no extension on either end. Unigenes were searched against the NCBI non-redundant protein database (NR), Swiss-Prot, TrEMBL, protein families (Pfam), and euKaryotic orthologous groups of proteins (KOG) using a BLASTx procedure with an e-value ≤ 10^{-5}. For unigene annotation, gene ontology (GO) was obtained based on NR database annotations and proteins were categorized according to their biological functions by BLAST algorithm-based searching against Kyoto Encyclopedia of Genes and Genomes (KEGG) database.

2.4. Differentially expressed genes (DEGs) identification

The expression level of each transcript was estimated according to a previously described protocol (Martazavi et al., 2008). Briefly, the number of mapped reads for each transcript was normalized into a reads per kb per million reads (RPKM) value to calculate the level of differential expression for each transcript. In analysis, a criterion of $|\log_2(\text{fold-change})| \geq 1$ and P value ≤ 0.05 between the two elevations was used to identify DEG. Functional gene classification was performed using UniProtKB/Swiss-Prot database.

2.5. qRT-PCR validation

Total RNA samples were isolated using an RNA simple total RNA kit (Tiangen, Beijing, China) according to the manufacturer's instructions. *Actin* was used as an internal standard to calculate relative gene expression amounts (Willems et al., 2008), and primer premier 6.0 was used to design the primer sequences (Table 1). cDNA synthesis and qRT-PCR analysis were performed using a one-step SYBR Prime Script PLUS RT-PCR kit (TaKaRa, Dalian, China). PCR amplification was performed in a 96-well platform (FTC-3000, Canada) according to the following protocol: one cycle at 42 °C for 5 min and 95 °C for 10 s; followed by 50 cycles at 95 °C for 5 s and 60 °C for 31 s. Melting curve analysis was performed after an incubation at 60 °C for 34 s.

3. Results

3.1. Global gene analysis

To identify molecular mechanisms responsible for greater plant growth and PPT accumulation in *S. hexandrum* with increased elevation, comparison of gene transcription for plants grown at 2300 and 3300 m asl was performed. A robust data set was collected; after data processing, 60.2 and 57.9 million high-quality reads were obtained at 2300 and 3300 m asl, respectively and a total of 53,691 unigenes were generated (Table 2; Fig. S2).

The percentage of genes that could be mapped onto a given protein database is shown in Table 3. Using the NR database, 26,748 unigenes were mapped into 10 identified plant species (59.5%) as well as "other" plant species (40.5%) (Fig. S3). Using the KOG database, proteins deduced from the identified nucleotide sequences were matched with genes, resulting in 21,421 unigenes identified; these genes were grouped into 25 functional categories (Fig. S4). Using the GO database,

Table 1
Sequences of primer employed in qRT-PCR analysis.

Gene name	Primer sequences (5' to 3')	Accession
ACT	Forward: GCAGGGATCCACGAGACCACC Reverse: CCCACCACTGAGCACATTGTTCC	AJ245731.1
PAL	Forward: ACGCTCTACGAACTTCTCCAC Reverse: CCGAATTGATTTCACGCTCA	KJ595569.1
CAD	Forward: ACCTGTTCTAGTGCCCCAAA Reverse: TGGAAGAGCAGGAAGAGGTG	KJ595572.1
DIR	Forward: AATGCAACAGCATCCATCGTA Reverse: CCAACTGGTGGGGAGTGAAA	KJ595571.1
PLR	Forward: GTTGAGGCCGTGAAATTGGT Reverse: TTTCCAGCTTCCTCGATTGC	EU240218.2
CYP719A23	Forward: AGCTTGCTAGAGTTCACGGT Reverse: CGAGCAGCATAGTCAGAGGA	KC110997.1
CYP71CU1	Forward: CTACACCCACCAGCTCCTTT Reverse: TCGGCCTTGTCCCATAACTT	KT390172.1
2-ODD	Forward: CATAGTCCTCTTCACCGGGT Reverse: ATGCTCTTCATCCACCAGCT	KT390173.1