

# ESTs Analysis Reveals Putative Genes Involved in Symbiotic Seed Germination in *Dendrobium officinale*

Ming-Ming Zhao<sup>1,2</sup>, Gang Zhang<sup>1</sup>, Da-Wei Zhang<sup>1</sup>, Yu-Yun Hsiao<sup>3</sup>, Shun-Xing Guo<sup>1\*</sup>

**1** Institute of Medicinal Plant Development, Chinese Academy of Medical Sciences & Peking Union Medical College, Beijing, People's Republic of China, **2** Medical Laboratory College, Beihua University, Jilin, People's Republic of China, **3** Department of Life Sciences, Orchid Research Center, National Cheng Kung University, Tainan, Taiwan

## Abstract

*Dendrobium officinale* (Orchidaceae) is one of the world's most endangered plants with great medicinal value. In nature, *D. officinale* seeds must establish symbiotic relationships with fungi to germinate. However, the molecular events involved in the interaction between fungus and plant during this process are poorly understood. To isolate the genes involved in symbiotic germination, a suppression subtractive hybridization (SSH) cDNA library of symbiotically germinated *D. officinale* seeds was constructed. From this library, 1437 expressed sequence tags (ESTs) were clustered to 1074 Unigenes (including 902 singletons and 172 contigs), which were searched against the NCBI non-redundant (NR) protein database (E-value cutoff,  $e^{-5}$ ). Based on sequence similarity with known proteins, 579 differentially expressed genes in *D. officinale* were identified and classified into different functional categories by Gene Ontology (GO), Clusters of orthologous Groups of proteins (COGs) and Kyoto Encyclopedia of Genes and Genomes (KEGG) pathways. The expression levels of 15 selected genes emblematic of symbiotic germination were confirmed via real-time quantitative PCR. These genes were classified into various categories, including defense and stress response, metabolism, transcriptional regulation, transport process and signal transduction pathways. All transcripts were upregulated in the symbiotically germinated seeds (SGS). The functions of these genes in symbiotic germination were predicted. Furthermore, two fungus-induced calcium-dependent protein kinases (CDPKs), which were upregulated 6.76- and 26.69-fold in SGS compared with un-germinated seeds (UGS), were cloned from *D. officinale* and characterized for the first time. This study provides the first global overview of genes putatively involved in *D. officinale* symbiotic seed germination and provides a foundation for further functional research regarding symbiotic relationships in orchids.

**Citation:** Zhao M-M, Zhang G, Zhang D-W, Hsiao Y-Y, Guo S-X (2013) ESTs Analysis Reveals Putative Genes Involved in Symbiotic Seed Germination in *Dendrobium officinale*. PLoS ONE 8(8): e72705. doi:10.1371/journal.pone.0072705

**Editor:** David L. Roberts, University of Kent, United Kingdom

**Received:** January 4, 2013; **Accepted:** July 15, 2013; **Published:** August 13, 2013

**Copyright:** © 2013 Zhao et al. This is an open-access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

**Funding:** This work was supported by the National Nature Science Foundation of China (No. 31070300, 31170016). The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

**Competing interests:** The authors have declared that no competing interests exist.

\* E-mail: sxguo2006@yahoo.com.cn

## Introduction

Orchidaceae, which comprises an estimated 20000 to 35000 species, constitutes one of the most diverse families [1], and most species of this family have medicinal and ornamental properties [2]. Orchidaceae are among the most evolutionarily and ecologically significant plants and are known for a wide variety of epiphytic and terrestrial growth forms and hardiness; they successfully colonize almost every habitat on earth, including soil (terrestrial), rock surfaces (lithophytic) and other plants (epiphytic). Almost all orchids seeds are extremely small (0.3–14 µg), dust-like and contain notably few nutrient reserves [3]. Lacking nutrients for early seedling development, orchids are dependent on a supply of carbohydrate and other nutrients derived from mycorrhizal fungal hyphae during the achlorophyllous seedling stage [4]. In parallel, fungal hyphae

invade and colonize the orchid root cortex and subsequently form elaborate intracellular coils called pelotons, which are structures distinctive in both appearance and distribution patterns with orchid tissues [5]. Several species of fungi that engage in symbiosis and provide nutrients, such as nitrogen (N) and carbon (C), during orchid seed germination have been isolated and characterized. Most of these species have been classified as orchid-associated Rhizoctonia-type fungi [6,7], in which Tulasnellaceae, Sebacinaceae and Ceratobasidiaceae have been reported to support seed germination and protocorm development [6,8,9].

In recent years, research has been conducted on the determinants of symbiosis between orchids and mycorrhizal fungi. Studies on morphological features have revealed that along with the mycorrhizal fungi that colonize tissues, the nuclei of infected cells of a mycorrhizal protocorm are visibly

hypertrophied [10]. These cells display elevated DNA content [11,12], and the starch grains within the cell have vanished [13]. In parallel, with the formation of complex pelotons, a series of biochemical reactions, including signaling, ROS, and  $Ca^{2+}$ ... etc. are systemically induced [14], and polysaccharides, inorganic salts, trace elements and secondary metabolites subsequently accumulate [15]. However, the molecular events associated with these processes remain poorly understood. Studies on other types of mycorrhizal symbiosis at the molecular level have made some progress. In arbuscular mycorrhizal (AM) symbiosis, functions of many genes that are essential for the various stages of nodule development have been elucidated in two model legume species, *Medicago truncatula* and *Lotus japonicas* [16]. AM symbiosis has been shown to be associated with root nodule symbioses (RNS), and a common pathway in AM endosymbiosis in legumes has been established [17]. Seven known AM-associated genes (*CASTOR*, *SymRK*, *POLLUX*, *NUP133*, *NUP85*, *Cyclops* and *CCaMK*) were shown to be recruited during the evolution of RNS [18]. Comparatively, the molecular mechanisms of orchidaceae symbiotic seed germination are indistinct. The only report regarding orchid mycorrhizal (OM) symbiosis describes the characterization of gene expression in roots of *Cypripedium parviflorum* incubated with a mycorrhizal fungus. However, only two genes encoding trehalose-6-phosphate synthase phosphatase (*Tps*) and nucleotide binding protein (*NuBP*) were identified and cloned [19]. Therefore, there is an urgent need for a clearer picture of gene expression in OM symbiosis.

*D. officinale*, a representative species of orchidaceae with great medicinal value, was selected for further analysis in this study. *D. officinale*, which has been officially recorded in Chinese pharmacopoeia, contains a number of effective components, including sesquiterpene fluorine ketone, dibenzyl phenols, philippines phenols, alkaloids and polysaccharides. *D. officinale* has been used in traditional Chinese medicine as a Yin tonic to nourish the stomach, relieve throat inflammation and fatigue, promote the secretion of body fluids, reduce peripheral vascular obstruction, prevent the development of cataracts and enhance the immune system. However, limited natural resources and high demand threaten the survival of the species, which has been listed as an endangered species and catalogued in the Chinese Plant Red Book since 1987 [8]. The low germination rate under natural conditions is the main factor that limits *D. officinale* propagation and its protection [20]. Revealing the mechanisms of germination would be beneficial for both protecting germ plasm resources and improving germination rates in nature. The objective of this work was to assess the genetic basis of symbiotic germination in orchid seeds. To understand the mechanisms of this process better at the molecular level, an SSH cDNA library of symbiotically germinated *D. officinale* seeds was constructed. Differentially expressed genes were examined by sequencing and analyzing ESTs present in the library. Because symbiosis is the basis of seed germination, the functions of these genes were predicted, and the genes were classified as symbiosis-related genes or germination-associated genes. The collection of identified genes presented in the study provides an important genomic

resource for future molecular studies on symbiotic seed germination in *D. officinale* and other OM symbiosis.

## Materials and Methods

### Plant material preparation

The capsules of *D. officinale* were collected from Menghai experimental base of Yunnan branch, Institute of Medicinal Plant Development, Chinese Academy of Medical Sciences & Peking Union Medical College Xishuanbanna, Yunnan, China. Our plant materials gathered from experimental base, and did not require specific permission. The corresponding author will be responsible for it. *D. officinale* is an endangered species; our purpose of this experiment is to reveal the mechanism of seed germination in nature conditions and to better protect the endangered species.

Nine capsules from eight maternal *D. officinale* growing on *Quercus* were collected (November, 2011) and divided into two groups. One group of seeds was cultured with *Sebacina* sp., which was isolated from *D. officinale* symbiotic seeds [8] (Genbank number EU910926 and public culture collection accession number CGMCC No. 3398) on oatmeal agar (OMA) medium [21] was germinated to the third stage, which is characterized by the appearance of protomeristem [22]. These seeds were defined as SGS. The other group defined as UGS was obtained by culturing *D. officinale* seeds on OMA medium without fungi. Seeds were sown according to the procedures described by Stewart et al. (2003) [23] and incubated in a high-humidity (75±5%) chamber at 25±2°C for a 16 h light (1500-2000 Lx) period and 8 h dark period for 5 weeks. Tissue-cultured seedlings (stored at the National Engineering Laboratory for Breeding of Endangered Medicinal Materials, Institute of Medicinal Plant Development, Chinese Academy of Medical Sciences & Peking Union Medical College) were carefully transferred to the 20 cm diameter pots and cultured in growth media (bark: pebble: coarse humus=3:1:1) in a high-humidity (75±5%) chamber at 25±2°C for a 16 h photoperiod with a light density of 400-1000 Lx for two months for aseptic tissue collection. All plant materials, including intact root, leaf, stem, UGS and SGS were collected, immediately frozen in liquid nitrogen and stored at -80° C prior to RNA extraction. A fraction of SGS and UGS were fixed in 2.5% glutaraldehyde (final concentration: 0.1M) for morphological observation.

### Total RNA extraction and SSH library construction

100 mg tissues of each sample including SGS, UGS, root, stem and leaf were gathered to extract total RNA using the EASYspin Plus Kit (Aidlab, Beijing, China) according to the manufacturer's recommendations. Genomic DNA was removed with RQ<sub>1</sub> RNase-free DNase (Promega, Madison, WI, USA). The quality and quantity of total RNA were assessed using a 1.2% formamide denaturing gel using an RNA ladder (Invitrogen, Carlsbad, CA, USA) and a NanoDrop™ 2000 spectrophotometer (Thermo, Fisher Scientific, Waltham, MA, USA), respectively. cDNA was synthesized via reverse transcription using M-MLV Reverse Transcriptase kit (Promega, Madison, WI, USA) and stored at -20° C. A "forward" subtraction library was constructed using the PCR-

Select™ cDNA Subtraction Kit (Clontech, Palo Alto, CA, USA). The tester and driver cDNA populations were generated from total RNA extracted from SGS and UGS, respectively. The final PCR products were purified using a QIAquick PCR purification Kit (Qiagen, Hilden, Germany) and directly cloned into the pGEM-T Easy vector (Promega, Madison, WI, USA), which was then transformed into competent *Escherichia coli* DH5α cells (Trans, Beijing, China). Transformed cells were grown on standard LB/ampicillin/X-gal/IPTG plates at 37°C for blue/white colony screening.

### EST assembly, annotation and functional classification

Positive recombinant plasmids were sequenced by Beijing Genewiz, Inc. The resulting sequences were trimmed to eliminate vectors and adaptor sequences with Codoncode Aligner 4.0.4 (CondonCode Corporation, Dedham, MA) and were subsequently assembled into unique sequences (containing contigs and singletons) with the CAP3 Sequence Assembly Program [24]. The insert sequences were manually assessed for similarities against the NCBI non-redundant protein database (NR) using BLASTX after translating DNA sequence into the respective amino acid sequence in six frames. ESTs with significant database matches ( $E$ -value  $\leq 1e-5$ ) were classified into functional categories according to the Gene Ontology database (<http://www.geneontology.org>), and gene expression was analyzed via EST sampling using COG database (<http://www.ncbi.nlm.nih.gov/COG/>). Pathway assignments were mapped according to the Kyoto Encyclopedia of Genes and Genomes (KEGG) database (<http://www.genome.ad.jp/kegg/kegg2.html>) (version KEGG 50) [25].

### Real-time qPCR analysis

To verify the quality of the SSH library and analyze the characterization of gene expression in fungi-infected plant tissues, 15 Unigenes were selected from different functional classes and their expression level were determined in SGS and UGS. Real-time qPCR technique were also used to determine the tissue expression pattern of two fungus-induced *CDPK* genes in non-infected plant tissues including UGS, root, stem and leaf, and their possible physiological functions were estimated. Using 2 µg of total RNA isolated from each sample as a template, reverse transcription reaction was performed (20 µL total volume) according to the Moloney Leukemia Virus Reverse Transcriptase (M-MLV RT) (Promega, Madison, WI, USA) kit protocol. The cDNAs were diluted 1:40 in ddH<sub>2</sub>O before performing the qPCR assay. Amplification curves were generated for all gene-specific primers designed using primer 5 (Table S1). Real-time qPCR was performed using a real-time SYBR green kit (Takara, Dalian, China) and the ABI 7500 real-time PCR System (Applied Biosystems, Foster City, CA, USA). The assay mix (final volume 25 µL) contained 12.5 µL 2× SYBR® Premix Ex Taq™ Master Mix (Takara Bio, Dalian, China), 0.5 µL each primer (10 µM), 0.5 µL ROX Reference Dye, 2 µL cDNA template, and 9 µL ddH<sub>2</sub>O. The thermal cycling conditions were as follows: 95° C for 30 s, followed by 40 cycles of 95° C 15 s and 60° C 40 s. The relative expression was calculated as the ratio of target gene expression

normalized to *GAPDH* expression [26]. All standards were amplified in duplicate, and samples were amplified in triplicate. All qPCR products were subjected to melting curve analysis and verified via gel electrophoresis. Cycle threshold (Ct) values were generated using the ABI PRISM 7500 Software Tool (Applied Biosystems, Foster City, CA, USA), and the relative expression ratios were calculated using the comparative 2<sup>-ΔΔCT</sup> method of relative gene quantification [27].

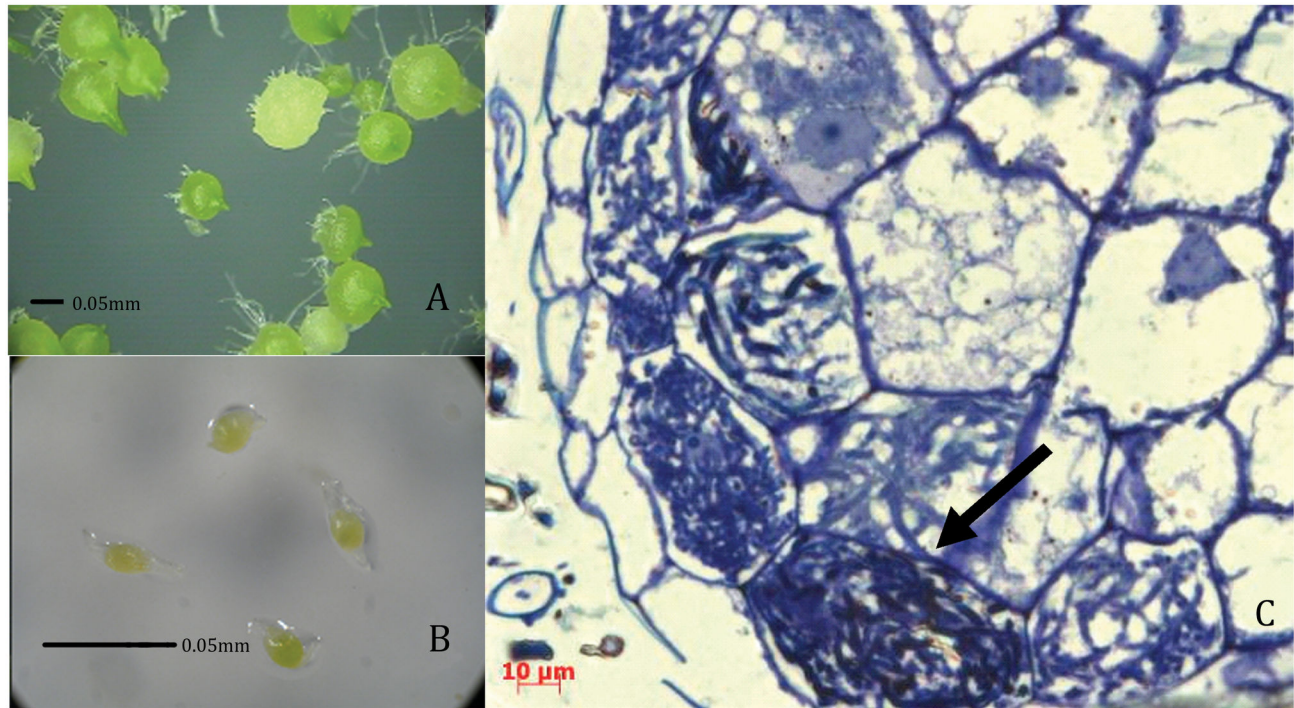
### Cloning and characterization of two full-length CDPK genes

Isolation of the 3'/5'-end of the two *CDPK* genes was performed using a SMART™ RACE cDNA Amplification Kit (Clontech, Japan) according to the manufacturer's instructions. Based on the sequences of the SSH cDNA Library, the gene-specific primers (GSP) DoCDPK1- R/F and DoCDPK32-like-R/F (Table S1) were designed. A thermal profile was obtained on a ABI-Veriti™ 96-well thermo cycler (Applied Biosystems, Foster City, CA, USA), and the assay mix (25 µL reaction volumes) contained 2.5 µL 10× Advantage® 2 PCR buffer, 0.5 µL dNTPs (10 mmol·L<sup>-1</sup>), 0.5 µL DoCDPK1-R/F or DoCDPK32-like-R/F 10 µmol·L<sup>-1</sup>, 0.5 µL UPM10 µmol·L<sup>-1</sup>, 1.0 µL 3'/5'-RACE ready cDNA, 0.5 µL 50× Advatange® 2 Polymerase Mix5 U·µL<sup>-1</sup> and 19.5 µL ddH<sub>2</sub>O. The PCR amplification program consisted of a 2 min activation step followed by 40 cycles of denaturation for 20 s at 95° C, annealing for 30 s at 68° C, and extension for 1.5 min step at 72° C and was terminated with an extension step for 7 min at 72° C. The amplified products were purified, cloned into a pGEM-T Easy vector (Promega, Madison, WI, USA) and transformed into competent DH5α cells. Positive recombinant plasmids were screened with AMP/IPTG/X-gal media and sequenced by Beijing Genewiz, Inc. After clustering and splicing with the core sequence, the ORFs of each sequence were obtained using BLASTX and ORF finder. Primers (Table S1) targeting the ORF determined in RACE screens were used to sequence and confirm the first reconstituted RACE sequences. The nucleotide sequences of *DoCDPK1* and *DoCDPK32-like* reported in the current study are available in the GenBank database under the accession numbers JX002662 and JQ927225, respectively. Sequences were aligned using MegAlign in the Lasergene software package of DNASTar 6.0 (Madison, WI, USA). A neighbor-joining phylogenetic tree was generated using MEGA 4.0 software package (version 4.0, <http://www.megasoftware.net>, Tempe, AZ, USA) with default parameters.

## Results

### Morphological examination

After cultivation for 5 weeks, seeds were examined using a dissecting microscope and were collected for further morphological assessment and RNA extraction. As shown in Figure 1A, almost all seeds cultivated with fungi germinated to the third stage. In contrast, all seeds cultivated without fungi remained un-germinated (Figure 1B). The germinated seeds infected with fungi were examined using semi-thin sections [28]. In the third-stage-germinated seeds, hyphae and mycelia knots convolved to form pelotons in plant cells, thereby



**Figure 1. Observation of *D. officinale* Seeds cultured for 5 weeks (A and B) and invasion in SGS (C).** Panels A and B show the pattern of third-stage SGS and UGS of *D. officinale*, respectively, using dissecting microscope. Scale bar (bold line), 0.05mm. Panel C: a semi-thin section was subjected to morphological examination using a Zeiss Axio Imager A1 microscope. The arrow indicates the pelotons of *Sebacina* sp. in *D. officinale* cell. Scale bar, 10  $\mu$ m.

doi: 10.1371/journal.pone.0072705.g001

indicating that a symbiotic relationship had been established (Figure 1C).

### Sequence assembly and EST annotation

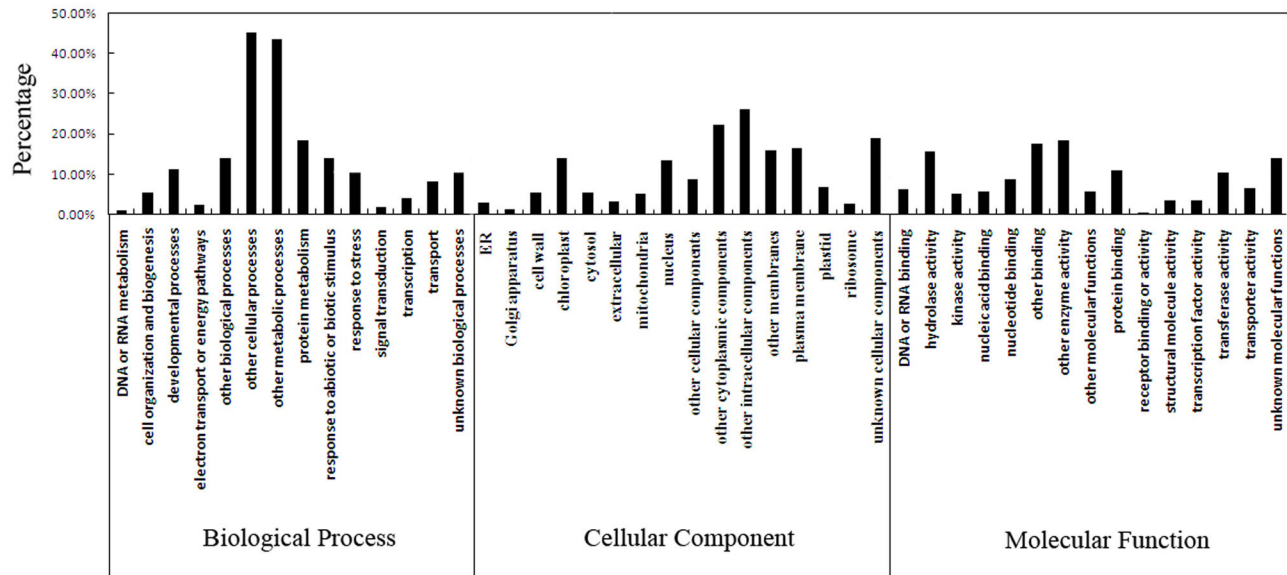
In total, 1565 positive clones were sequenced, and ESTs containing no insert sequence or inserts shorter than 100 bp were removed prior to cluster analysis. A total of 1437 ESTs were obtained, with an average length of 518 bp. The remaining ESTs were aligned and assembled into 172 contigs (average length of 569 bp; range: 179-2202 bp) and 902 singletons (average length of 509 bp; range: 110-1096 bp). The ESTs clusters were divided into seven classes, including singles with one EST. Accordingly, 121 contigs containing 2 ESTs comprised the greatest proportion of the total accepted ESTs with the expression of singletons. We found 50 contigs containing ESTs ranging from 3 to 20 and one contig containing 68 ESTs, thereby suggesting a high level of transcription associated with the corresponding genes.

Genes for which the expression was altered at stage 3 symbiotic germination were identified. A total of 1074 Unigenes were subjected to BLASTX analysis for a homologous search against the NR database in GenBank. Of these, 777 Unigenes (70.08%) displayed at least one significant alignment with an existing gene in the database (E-value cutoff,  $e^{-5}$ ), and the remaining 297 Unigenes (29.92%) did not match any known sequence. These matched ESTs were named according to

homologous sequences in the NCBI database and divided into two sets: *Dendrobium* and *Sebacina*. According to the results, 579 differentially expressed genes were identified in *Dendrobium*, and of these, 416 Unigenes had a known function. Based on comparison with genes expressed in UGS, a portion of genes were induced by OM symbiosis, including those in response to OM fungal signals and plant pathogens. The other genes were related to seed germination, including many that have been associated with carbon and nitrogen metabolism. A total of 198 Unigenes were identified in *Sebacina*; of these, 155 matched a known function. Most of these genes were involved in membrane transport and mineral nutrition metabolism.

### Functional classification of *D. officinale* subtractive ESTs

GO annotation was used to provide descriptions of gene products associated with molecular functions, cellular components and biological processes [29]. To determine the possible functions, we used the GO classification system for plants developed at TAIR, which is based on the *Arabidopsis thaliana* genome ([http://www.arabidopsis.org/help/helppages/go\\_slim\\_help.jsp](http://www.arabidopsis.org/help/helppages/go_slim_help.jsp)). The annotations of unique sequences for *D. officinale*, for which GO categories could be assigned, were most frequently associated with biological processes followed by cellular components and molecular functions. Figure 2



**Figure 2. Representation of gene ontology assignments for plant homologous ESTs derived from SSH library.** The GO slim Classification for Plants Developed at TAIR was used to functionally characterize the ESTs.

doi: 10.1371/journal.pone.0072705.g002

shows the percentages of transcript models of subtractive ESTs for *D. officinale* in the 3 GO categories. These ESTs were classified into 14 categories of biological processes, primarily protein metabolism (18.42%), response to abiotic or biotic stimuli (14.11%), response to stress (13.43%) and unclassified processes. Plasma membrane (16.18%) and chloroplast (13.6%) were the most frequently represented categories in cellular components other than unclassified components. The largest proportion of molecular function ESTs were categorized as hydrolase activity (15.49%), protein binding (10.67%) and unclassified molecular functions.

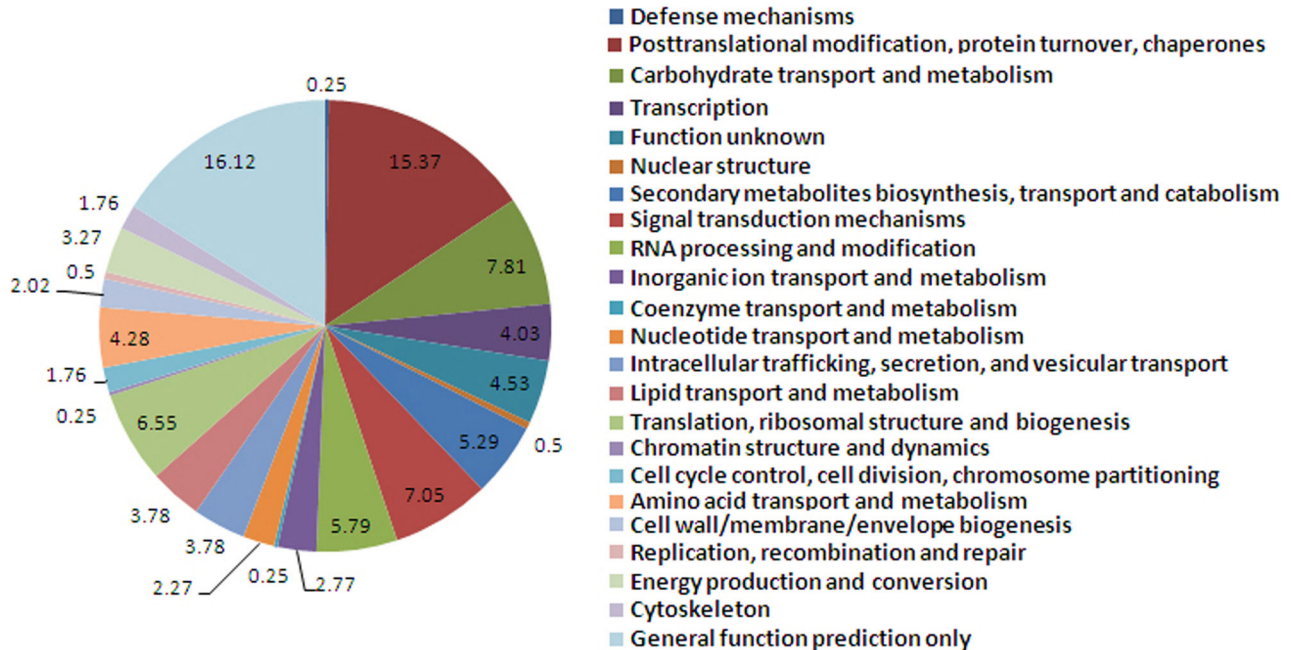
Clusters of Orthologous Groups of proteins (COGs) were delineated by comparing protein sequences encoded in complete genomes, representing major phylogenetic lineages. COGs can predict individual protein features and the whole-genome protein function. The tagged proteins can be clustered into COGs comparing with all proteins in the COGs. To classify the possible functions of deduced proteins annotated by the SWISS-PROT database, we searched the COGs database for plant homologous proteins. Among all deduced plant homologous proteins, 369 displayed significant hit with COGs database, including 23 functional groups (Figure 3). These results show that the largest proportion of functions were protein metabolisms (15.37%), carbohydrate transport and metabolism (7.81%) and signal transduction (7.05%). Such categories as secondary metabolite biosynthesis, transport and catabolism (5.29%) and lipid transport and metabolism (3.78%) also represented a higher proportion than other classifications, thereby reflecting the high level of activity possibly linked to seed germination.

The unique sequences were assigned to KEGG annotations, which provide an alternative functional annotation of genes

according to their associated biochemical pathways [25] based on sequence similarity searches against the KEGG database (<http://www.genome.jp/kegg/pathway.html>). Unigenes of *D. officinale* mapped to the KEGG pathways are shown in Table 1. Among these, 100 Unigenes were related to global maps, 170 corresponded to metabolism, 41 mapped to genetic information processing, 10 belonged to cellular processes and 8 were classified as organismal systems. Unigenes mapped to metabolism were the most abundant, which were associated with carbohydrate metabolism, amino acid metabolism, lipid metabolism, energy metabolism and nucleotide metabolism, all of which were related to germination. Eight genes were involved in environmental adaptation, and three of these, mitogen-activated protein kinase (MPK), *CDPK* and transcription factor *WRKY*, participated in plant-pathogen interaction pathways and are assumed to be relevant for orchid symbiotic interactions [30,31].

### Real-time qPCR analysis of symbiotic germination-related genes

Real-time qPCR was employed to experimentally verify the expression patterns of 15 differentially expressed genes based on their putative functions and the results of the SSH library. Expression analysis was used to verify the validity of stage 3 symbiotic germination-related ESTs and to more precisely determine the timing of expression of the filtered genes of interest. As shown in Figure S1, all selected genes, including those encoding UDP-glucosyltransferase (17.5-fold),  $\beta$ -glucosidase (57.20-fold), immediate-early fungal elicitor (9.28-fold), chitinase (33.94-fold), leucine-rich repeat receptor kinase (LRR-RK) (29.37-fold),  $\beta$ -1, 3-glucanase (23.67-fold), cysteine protease (23.67-fold), NAC transcription factor (83.34-fold),



**Figure 3. The functional classification of proteins deduced from plant homologous Unigenes in the SSH library using COGs.**

doi: 10.1371/journal.pone.0072705.g003

CDPK32-like (26.69-fold), agglutinin (4.43-fold), auxin-responsive protein (3.75-fold), catalase (3.24-fold), early nodulin putative (5.39-fold), cation exchanger (8.49-fold) and CDPK1 (6.76-fold), were verified to be upregulated in SGS.

### Characterization and sequence analysis of full-length DoCDPK1 and DoCDPK32-like

CDPKs are the largest group of complex calcium sensors, and they play a key role in a variety of biological processes, including plant growth and development, abiotic and biotic stress adaptations, and seed germination, via calcium signaling pathways [32]. To gain further insight into the role of CDPKs in *D. officinale* symbiotic germination, the corresponding full-length cDNAs of two CDPKs were isolated via RACE assay. *DoCDPK1* (*D. officinale* CDPK1) and *DoCDPK32-like* (*D. officinale* CDPK32-like), according to the nomenclature of reference [33], were obtained via 3'RACE and 5'RACE, and positive clones were sequenced. Trimmed out of vectors using VecScreen followed by re-assembling, two full-length cDNA (2236 bp and 2124 bp) were obtained. To confirm the full-length cDNA sequences, PCR was performed with primers (DoCDPK1-orfF/R; DoCDPK32-like-orfF/R) (Table S1) spanning ORFs. Single bands of 1.9 kb and 1.85 kb for *DoCDPK1* and *DoCDPK32-like*, respectively, were observed via agarose gel electrophoresis. After cloning and sequencing, the products were verified to be consistent with the spliced sequences described above. The two sequences were deposited into GenBank under the accession numbers JX002662 (*DoCDPK1*) and JQ927225 (*DoCDPK32-like*), respectively. *DoCDPK1* contained a 1578 bp ORF with a 5'-

UTR (un-translated region) of 266 bp and a 3'-UTR of 392 bp. The ORF of *DoCDPK32-like* was 1407 bp in size and comprised of a 5'-UTR of 442 bp and a 3'-UTR of 275 bp. Each sequence contained a eukaryotes polyA tail and the typical Kozak consensus sequence (A/GNNATGG) located at the beginning of the coding region. Using Compute PI/MW, the predicted molecular weight of DoCDPK1 and DoCDPK32-like were 58.34 KDa and 51.48 KDa, respectively, and the prediction of *pI* values was 5.94 and 6.11, respectively. Both deduced proteins contain a conserved serine/threonine protein kinase catalytic domain. DoCDPK1 contains the typical four EF-hand motifs. In contrast, this motif was not found in DoCDPK32-like.

### Multiple Sequence Alignment, Phylogenetic relationships and tissue expression pattern analysis of DoCDPK1 and DoCDPK32-like

DoCDPK1 and DoCDPK32-like were aligned to several other plant CDPKs using MegAlign in DNASTar 6.0 (Figure S2). DoCDPK1 had the highest level of homology (86%), with PgCDPK1 (ACY78680) of *Panax Ginseng* followed by *Zea mays* ZmCDPK2 NP\_001105542 and *Hevea Brasiliensis* HbCDPK1 (ACB71246) with 83% identity. DoCDPK32-like displayed the highest level of homology with *Vitis vinifera* VvCDPK25-like (47%), while the homology was 42% with the other three proteins. BLAST searches and subsequent domain analysis also revealed that each of these CDPKs displayed similar domain compositions. The protein kinase ATP-binding signature and S-TKc active-site signature of all sequences were well conserved. Regarding the protein kinase ATP-

**Table 1.** Unigenes mapped in KEGG pathways.

KEGG Pathways	Sub-pathways of KEGG Pathways	Number of Unigenes
Global Map		100
	Metabolic Pathways	68
	Biosynthesis of Secondary Metabolites	32
Metabolism		170
	Carbohydrate Metabolism	62
	Energy Metabolism	18
	Lipid Metabolism	21
	Nucleotide Metabolism	11
	Amino Acid Metabolism	26
	Metabolism of Other Amino Acids	10
	Glycan Biosynthesis and Metabolism	10
	Metabolism of Cofactors and Vitamins	2
	Metabolism of Terpenoids and Polyketides	2
	Biosynthesis of Other Secondary Metabolites	8
Genetic Information Processing		41
	Transcription	7
	Translation	9
	Folding, Sorting and Degradation	24
	Replication and Repair	1
Cellular Processes		10
	Transport and Catabolism	10
Organismal Systems		8
	Environmental Adaptation	8

binding signature, each sequence started with Leu–Gly and ended with Lys. Similarly with respect to the S-TKc active-site signature, the active-site was Asp for all sequences.

For phylogenetic analysis, DoCDPK1, DoCDPK32-like and sixty-four homologous CDPKs were aligned using ClustalW in Mega 4.0 (Figure S3). The results indicate that DoCDPK1 belongs to group II-a (13), it also clustered with *Oryza sativa* OsCDPK19 in the same branch and was phylogenetically closest to the AtCDPK9 and AtCDPK33 branches. DoCDPK32-like grouped with GmCDPK32-like and was distant from other plant CDPKs groups.

Real-time qPCR was used to profile *DoCDPK1* and *DoCDPK32-like* (Figure S4). The results show that both genes are constitutively expressed. The highest level of the *DoCDPK1* transcript level was found in leaf samples, with 8.54-fold expression, whereas *DoCDPK32-like* was expressed 17.34-fold higher in root samples.

## Discussion

Although widespread investigations of orchid seedling germination have been performed for well over a century, the mechanism of this germination at the molecular level remains unclear. Much less is known regarding the genetically

controlled checkpoints that facilitate or restrict symbiotic development and the plant and fungal genes that are involved in the process. This work reflects a preliminary characterization and provides the first investigation of orchid symbiotic germination related genes, especially focusing on plant response genes. Compared with *Sebacina*, the sequence in *Dendrobium* is well known and the plant genotype can affect the extent of the fungi colonization and the response. A variety of plant homologous ESTs were classified and most were assigned as germination-related and symbiosis-related. These Unigenes were identified, annotated and found to participate in many biological processes, such as transcription, transduction, transport and metabolism. By analyzing the annotation of Unigenes mapped onto KEGG database metabolism pathways, the most abundant subset belonged to carbohydrate metabolism followed by amino acid and lipid metabolism (Table 1). Two carbohydrate metabolism Unigenes encoding UDP-glucosyltransferase and  $\beta$ -glucosidase were significantly upregulated by 17.35- and 57.20-fold in SGS, respectively. In the current report, various genes involved in transport processes, including carbohydrate transporter, cation transporters, ATP/ADP-transporters and nucleoside transporters, were determined to be upregulated, thereby indicating that a soluble or insoluble source of carbohydrate, nitrogen and phosphorus translocated from the fungus to the plant. This observation suggests that a symbiotic relationship was established and nutrients accumulated in the plant for the next germination stage.

In this work, we paid close attention to genes putatively involved in signaling, and several genes associated with signal transduction were identified during the course of OM symbiosis. The transcript encoding LRR-RK, which has been reported to be involved in plant defense response and shoot development [34], displayed a 29.37-fold increase in SGS. CDPKs and MPKs are two common protein kinases involved in diverse signal regulation pathways. MPKs, which mediate intracellular transmission and amplification of extracellular stimuli, have been demonstrated to be essential for plant-microbe symbiosis [35]. CDPKs have also been reported to be involved in symbiotic systems and seed germination. *MtCDPK3* was confirmed to be upregulated during nodulation in *M. truncatula* [36]. *OsCDPK18* was shown to be transcriptionally activated in response to inoculation with *Glomus intraradices* [37]. ABA-stimulated CDPK, *ACDPK1*, derived from grapes, promotes plant growth and enhances ABA sensitivity in seed germination [38]. Moreover, *AtCDPK32* has been shown to be involved in ABA-regulated seed germination [39]. Real-time qPCR analyses indicate that the two CDPKs cloned in the current study display a tissue-specific expression pattern and are both upregulated in SGS. This result suggests that these genes may be induced by fungi and may play a vital role in the regulation of *D. officinale* symbiosis.

Symbiosis in which fungus infects and colonizes the plant would trigger a defense response. The production of hydrolytic enzymes by the plant may be important for lysis of the fungal cell wall. Chitinase and  $\beta$ -glucanase are two common enzymes often implicated in the response of plants to parasite infection [6], and their expression can be induced by fungi [40,41].

Fungal elicitors are proteins that induce necrosis in plants, thereby causing the plant to become resistant to pathogens [42]. Genes encoding immediate-early fungal elicitor proteins, chitinase and beta-1, 3-glucanase, were selected and their expression levels were elevated 2-fold in SGS compared with UGS. This situation is similar to AM symbiotic relationships [6]. A phytohormone-related gene encoding auxin response protein was also discovered and seems to be upregulated in SGS. The auxin response protein had been reported to function in defense response in *Pinus pinaster* roots against fungal pathogen attacks [43].

Two pathways in which genes may participate in the regulation of the plant–fungus interaction are shown in Figure S5. The polyamines spermidine (Spd), spermine (Spm), and their precursor, putrescine (Put), are low-molecular-weight aliphatic amines produced ubiquitously in many physiological processes in both eukaryotes and prokaryotes, especially higher plants [44]. Kytöviita et al. [45] have reported that polyamines are involved in establishing and maintaining mycorrhizal symbiosis. The concentration of free polyamines in roots is altered by inoculating with ectomycorrhizal (ECM) fungi, and the supply of specific polyamines enhances the formation of both the ECM [46] and AM [47]. In our work, three major enzymes involved in polyamine synthesis pathways were identified among the differentially expressed genes, including arginine decarboxylase (ADC), S-adenosylmethionine decarboxylase (SAMDC) and spermidine synthase (SPDS) (Figure S5A). These genes may collectively function in polyamine circulation during OM symbiosis. Another pathway associated with plant–pathogen interactions is activated by  $Ca^{2+}$  according to KEGG ath04626, including *CDPK*, *MPK* and *WRKY* for instance (Figure S5B).

Compared with *Dendrobium* derived ESTs, relatively few genes were identified from *Sebacina*. A majority of these ESTs are involved in membrane transport and mineral nutrition metabolism, including  $H^+$ -ATPases, sucrose transporter, protein transporter, inorganic phosphate transporter and lipase.  $H^+$ -ATPases are reported to play a key role in establishing the electrochemical gradient required for the transfer of nutrients across the plasma membrane in AM symbiosis [48]. As the expression of these genes is altered during the interaction, they should be essential for fungus to establish a compatible relationship with the plant. The understanding of gene expression regulation in OM fungi is much more limited; therefore, more studies are required to test this hypothesis and determine these genes' functions in OM symbiosis.

In this study, we successfully used a combination of SSH library and real-time qPCR techniques to determine gene expression patterns during orchid symbiotic germination. These genes and their putative functions provide insight into *Dendrobium* symbiotic seed germination and provide candidate genes for further investigations of the molecular events that control OM symbiosis.

## Supporting Information

**Figure S1. Expression patterns of fifteen subtractive plant homologous genes from SSH library using qPCR analyses.**

Dark grey columns represent the RQ in SGS; light grey columns indicate the RQ in UGS. All standards were run in duplicate and samples were run in triplicate. 1: NAC transcription factor; 2: cation exchanger; 3: auxin-responsive protein; 4: LRR; 5: cysteine protease; 6: chitinase; 7:  $\beta$ -1,3-glucanase; 8: catalase; 9: UDPG; 10: CDPK1; 11: immediate-early fungal elicitor; 12: agglutinin; 13: early nodulin putative; 14:  $\beta$ -glucosidase; 15: CDPK32-like (TIF)

**Figure S2. Multiple sequence alignments of DoCDPK1 (A), DoCDPK32-like (B) and CDPKs proteins from other plants.**

Thick lines indicate the S-TKc domain ; Dot lines indicate the four  $Ca^{2+}$  binding EF-hand motifs. At: *Arabidopsis thaliana*; Do: *Dendrobium officinale*; Gm: *Glycine max*; Hb: *Hevea brasiliensis*; Os: *Oryza sativa*; Pg: *Peanax ginseng*; St: *Solanum tuberosum*; Vv: *Vitis vinifera*; Zm: *Zea mays*. (TIF)

**Figure S3. Phylogenetic tree of DoCDPK1 and DoCDPK32-like with CDPKs genes from other plants.**

At: *Arabidopsis thaliana*; Do: *Dendrobium officinale*; Gm: *Glycine max*; Os: *Oryza sativa* (TIF)

**Figure S4. Tissue-specific expression patterns of DoCDPK1 and DoCDPK32-like genes using real-time qPCR analyses (A and B, respectively).**

(TIF)

**Figure S5. Genes involved in polyamine synthesis pathways (A) and plant–pathogen interaction pathways (B).**

Shadow genes represent those obtained from the SSH library. SAM: S-adenosylmethionine; dcSAM: decarboxylated S-adenosylmethionine; SAMDC: S-adenosylmethionine decarboxylase; ADC: argininedecarboxylase; SPDS: spermidine synthase; SPMS: spermine synthase; CDPK: calcium-dependent protein kinase; Rboh: respiratory burst oxidase homo-logs; ROS: reactive oxygen species; MPK: mitogen-activated protein kinases (TIF)

**Table S1. Gene specific primers for real-time qPCR.**

(DOC)

## Author Contributions

Conceived and designed the experiments: MMZ GZ SXG. Performed the experiments: MMZ. Analyzed the data: MMZ YYH. Contributed reagents/materials/analysis tools: MMZ GZ. Wrote the manuscript: MMZ DWZ.



## References

- Cribb PJ, Kell SP, Dixon KW, Barrett RL (2003) Orchid conservation: a global perspective. Orchid conservation. Kota Kinabalu: Natural History Publications. pp. 1-24.
- Hsiao YY, Pan ZJ, Hsu CC, Yang YP, Hsu YC et al. (2011) Research on orchid biology and biotechnology. *Plant Cell Physiol* 52: 1467-1486. doi:10.1093/pcp/pcr100. PubMed: 21791545.
- Arditti J, Ghani AKA (2000) Numerical and physical properties of orchid seeds and their biological implications. *New Phytol* 146: 569-569. doi: 10.1046/j.1469-8137.2000.00675.x.
- Rasmussen HN, Rasmussen FN (2009) Orchid mycorrhiza: implications of a mycophagous life style. *Oikos* 118: 334-345. doi: 10.1111/j.1600-0706.2008.17116.x.
- Kristiansen KA, Freudenstein JV, Rasmussen FN, Rasmussen HN (2004) Molecular identification of mycorrhizal fungi in *Neuwiedia veratrifolia* (Orchidaceae). *Mol Phylogenet Evol* 33: 251-258. doi: 10.1016/j.ympev.2004.05.015. PubMed: 15336661.
- Smith SE, Read DJ (2008) Mycorrhizal symbiosis. New York: Academic Press. 787pp.
- Merckx VS (2013) Mycoheterotrophy: An Introduction. Mycoheterotrophy. Springer Verlag. pp. 1-17.
- Wang H, Fang H, Wang Y, Duan L, Guo S (2011) In situ seed baiting techniques in *Dendrobium officinale* Kimura et Migo and *Dendrobium nobile* Lindl: the endangered Chinese endemic *Dendrobium* (Orchidaceae). *World J Microbiol Biotechnol* 27: 2051-2059. doi: 10.1007/s11274-011-0667-9.
- Chutima R, Dell B, Lumyong S (2011) Effects of mycorrhizal fungi on symbiotic seed germination of *Pecteilis susanna* (L.) Rafin (Orchidaceae), a terrestrial orchid in Thailand. *Symbiosis* 53: 149-156. doi:10.1007/s13199-011-0120-8.
- Burgeff H (1932) Saprophytismus und symbiose. studien an tropischen orchideen. Jena G Fischer: 249.
- Alvarez MR (1968) Quantitative changes in nuclear DNA accompanying postgermination embryonic development in *Vanda* (Orchidaceae). *Am J Bot* 55: 1036-1041. doi:10.2307/2440469.
- Williamson B (1970) Induced DNA synthesis in orchid mycorrhiza. *Planta* 92: 347-354. doi:10.1007/BF00385100.
- Beyrie H, Smith S, Franco C, Peterson R (1995) Colonization of *Orchis morio* protocorms by a mycorrhizal fungus: effects of nitrogen nutrition and glyphosate in modifying the responses. *Can J Bot* 73: 1128-1140. doi:10.1139/b95-123.
- Hou XQ, Guo SX (2009) Interaction between a dark septate endophytic isolate from *Dendrobium* sp. and roots of *D. nobile* seedlings. *J Integr Plant Biol* 51: 374-381. doi:10.1111/j.1744-7909.2008.00777.x. PubMed: 21452589.
- Leake JR (2005) Plants parasitic on fungi: unearthing the fungi in myco-heterotrophs and debunking the 'saprophytic' plant myth. *Mycologist* 19: 113-122. doi:10.1017/S0269915X05003046.
- Oldroyd GED, Harrison MJ, Paszkowski U (2009) Reprogramming plant cells for endosymbiosis. *Science* 324: 753-754. doi:10.1126/science.1171644. PubMed: 19423817.
- Capoen W, Den Herder J, Sun J, Verplancke C, De Keyser A et al. (2009) Calcium spiking patterns and the role of the calcium/calmodulin-dependent kinase CcAMK in lateral root base nodulation of *Sesbania rostrata*. *Plant Cell* 21: 1526-1540. doi:10.1105/tpc.109.066233. PubMed: 19470588.
- Reinhardt D (2007) Programming good relations-development of the arbuscular mycorrhizal symbiosis. *Curr Opin Plant Biol* 10: 98-105. doi: 10.1016/j.pbi.2006.11.001. PubMed: 17127091.
- Watkinson JI, Welbaum GE (2003) Characterization of gene expression in roots of *Cypripedium parviflorum* var. *pubescens* incubated with a mycorrhizal fungus. *Acta Horticult*: 463-470.
- Cai X, Feng Z, Zhang X, Xu W, Hou B et al. (2011) Genetic diversity and population structure of an endangered Orchid (*Dendrobium loddigesii* Rolfe) from China revealed by SRAP markers. *Sci Horticul* 129: 877-881. doi:10.1016/j.scienta.2011.06.001.
- Warcup J (1981) The mycorrhizal relationships of Australian orchids. *New Phytol* 87: 371-381. doi:10.1111/j.1469-8137.1981.tb03208.x.
- Stewart SL, Zettler LW (2002) Symbiotic germination of three semi-aquatic rein orchids (*Habenaria repens*, *H. quinquiseta*, *H. macroceratitis*) from Florida. *Aquat Bot* 72: 25-35. doi:10.1016/S0304-3770(01)00214-5.
- Stewart S, Zettler L, Minso J, Brown P (2003) Symbiotic germination and reintroduction of *Spiranthes brevibrabis* Lindley, an endangered orchid native to Florida. *Selbyana* 24: 64-70.
- Huang X, Madan A (1999) CAP3: A DNA sequence assembly program. *Genome Res* 9: 868-877. doi:10.1101/gr.9.9.868. PubMed: 10508846.
- Ogata H, Goto S, Sato K, Fujibuchi W, Bono H et al. (1999) KEGG: Kyoto encyclopedia of genes and genomes. *Nucleic Acids Res* 27: 29-34. doi:10.1093/nar/27.20.e29. PubMed: 9847135.
- Zhang G, Song C, Zhao MM, Li B, Guo SX (2012) Characterization of an A-type cyclin-dependent kinase gene from *Dendrobium candidum*. *Biologica* 67: 360-368. doi:10.2478/s11756-012-0016-y.
- Pfaffl MW (2001) A new mathematical model for relative quantification in real-time RT-PCR. *Nucleic Acids Res* 29: e45. doi:10.1093/nar/29.9.e45. PubMed: 11328886.
- Aparicio S, Marsden P (2011) A rapid methylene blue-basic fuchsin stain for semi-thin sections of peripheral nerve and other tissues. *J Microsc* 89: 139-141.
- Berardini TZ, Mundodi S, Reiser L, Huala E, Garcia-Hernandez M et al. (2004) Functional annotation of the Arabidopsis genome using controlled vocabularies. *Plant Physiol* 135: 745-755. doi:10.1104/pp.104.040071. PubMed: 15173566.
- Romeis T (2001) Protein kinases in the plant defence response. *Curr Opin Plant Biol* 4: 407-414. doi:10.1016/S1369-5266(00)00193-X. PubMed: 11597498.
- Eulgem T, Somssich IE (2007) Networks of WRKY transcription factors in defense signaling. *Curr Opin Plant Biol* 10: 366-371. doi:10.1016/j.pbi.2007.04.020. PubMed: 17644023.
- Asano T, Tanaka N, Yang G, Hayashi N, Komatsu S (2005) Genome-wide identification of the rice calcium-dependent protein kinase and its closely related kinase gene families: comprehensive analysis of the CDPKs gene family in rice. *Plant Cell Physiol* 46: 356-366. doi: 10.1093/pcp/pci035. PubMed: 15695435.
- Hrabak EM, Dickmann LJ, Satterlee JS, Sussman MR (1996) Characterization of eight new members of the calmodulin-like domain protein kinase gene family from *Arabidopsis thaliana*. *Plant Mol Biol* 31: 405-412. doi:10.1007/BF00021802. PubMed: 8756605.
- Spaink HP (2002) Plant-microbe interactions: A receptor in symbiotic dialogue. *Nature* 417: 910-911. doi:10.1038/417910a. PubMed: 12087390.
- Lee H, Kim J, Im JH, Kim HB, Oh CJ et al. (2008) Mitogen-activated protein kinase is involved in the symbiotic interaction between *Bradyrhizobium japonicum* USDA110 and soybean. *J Plant Biol* 51: 291-296. doi:10.1007/BF03036129.
- Gargantini PR, Gonzalez-Rizzo S, Chinchilla D, Raices M, Giammaria V et al. (2006) A CDPK isoform participates in the regulation of nodule number in *Medicago truncatula*. *Plant J* 48: 843-856. doi:10.1111/j.1365-3113.2006.02910.x. PubMed: 17132148.
- Campos-Soriano L, Gómez-Ariza J, Bonfante P, San Segundo B (2011) A rice calcium-dependent protein kinase is expressed in cortical root cells during the presymbiotic phase of the arbuscular mycorrhizal symbiosis. *BMC Plant Biol* 11: 90. doi:10.1186/1471-2229-11-90. PubMed: 21595879.
- Yu XC, Zhu SY, Gao GF, Wang XJ, Zhao R et al. (2007) Expression of a grape calcium-dependent protein kinase ACDPK1 in *Arabidopsis thaliana* promotes plant growth and confers abscisic acid-hypersensitivity in germination, postgermination growth, and stomatal movement. *Plant Mol Biol* 64: 531-538. doi:10.1007/s11103-007-9172-9. PubMed: 17476573.
- Choi HI, Park HJ, Park JH, Kim S, Im MY et al. (2005) Arabidopsis calcium-dependent protein kinase AtCDPK32 interacts with ABF4, a transcriptional regulator of abscisic acid-responsive gene expression, and modulates its activity. *Plant Physiol* 139: 1750-1761. doi: 10.1104/pp.105.069757. PubMed: 16299177.
- Liu B, Xue X, Cui S, Zhang X, Han Q et al. (2010) Cloning and characterization of a wheat  $\beta$ -1, 3-glucanase gene induced by the stripe rust pathogen *Puccinia striiformis* f. sp. tritici. *Mol Biol Rep* 37: 1045-1052. doi:10.1007/s11033-009-9823-9. PubMed: 19757158.
- Amian AA, Papenbrock J, Jacobsen HJ, Hassan F (2011) Enhancing transgenic pea (*Pisum sativum* L.) resistance against fungal diseases through stacking of two antifungal genes (Chitinase and Glucanase). *GM Crops* 2: 104-109. doi:10.4161/gmcr.2.2.16125. PubMed: 21971070.
- Mikes V, Milat ML, Ponchet M, Ricci P, Blein JP (1997) The fungal elicitor cryptogein is a sterol carrier protein. *FEBS Lett* 416: 190-192. doi:10.1016/S0014-5793(97)01193-9. PubMed: 9369212.
- Reddy S, Pandey A, Melayah D, Marmeisse R, Gay G (2003) The auxin responsive gene Pp-C61 is up-regulated in *Pinus pinaster* roots following inoculation with ectomycorrhizal fungi. *Plants Cell Environ* 26: 681-691. doi:10.1046/j.1365-3040.2003.01003.x.
- Efrose RC, Fletmetakis E, Sfichi L, Stedel C, Kouri ED et al. (2008) Characterization of spermidine and spermine synthases in *Lotus japonicus*: induction and spatial organization of polyamine biosynthesis

- in nitrogen fixing nodules. *Planta* 228: 37-49. doi:10.1007/s00425-008-0717-1. PubMed: 18320213.
45. Kytöviita MM, Sarjala T (1997) Effects of defoliation and symbiosis on polyamine levels in pine and birch. *Mycorrhiza* 7: 107-111. doi:10.1007/s005720050170.
46. Niemi K, Häggman H, Sarjala T (2002) Effects of exogenous diamines on the interaction between ectomycorrhizal fungi and adventitious root formation in Scots pine in vitro. *Tree Physiol* 22: 373-381. doi:10.1093/treephys/22.6.373. PubMed: 11960762.
47. El Ghachtouli N, Martin-Tanguy J, Paynot M, Gianinazzi S (1996) First-report of the inhibition of arbuscular mycorrhizal infection of *Pisum sativum* by specific and irreversible inhibition of polyamine biosynthesis or by gibberellic acid treatment. *FEBS Lett* 385: 189-192. doi:10.1016/0014-5793(96)00379-1. PubMed: 8647248.
48. Requena N, Breuninger M, Franken P, Ocón A (2003) Symbiotic status, phosphate, and sucrose regulate the expression of two plasma membrane H<sup>+</sup>-ATPase genes from the mycorrhizal fungus *Glomus mosseae*. *Plant Physiol* 132: 1540-1549. doi:10.1104/pp.102.019042. PubMed: 12857834.