



OPEN

Comparative transcriptome analysis of a taxol-producing endophytic fungus, *Aspergillus aculeatinus* Tax-6, and its mutant strain

Weichuan Qiao[✉], Tianhao Tang & Fei Ling

Taxol is a rare but extremely effective antitumor agent extracted from *Taxus yew* barks. *Taxus* plants are valuable and rare species, and the production of taxol from them is a complex process. Therefore, taxol-producing endophytic fungi seem to be a promising alternative because of their high practical value and convenient progress. In this study, the transcriptome of an endophytic fungus, *Aspergillus aculeatinus* Tax-6 was analyzed in order to understand the molecular mechanisms of producing fungal taxol. The results showed that genes involved in the mevalonate (MVA) pathway and non-mevalonate (MEP) pathway were expressed, including isopentenyl pyrophosphate transferase, geranyl pyrophosphate transferase, and geranylgeranyl pyrophosphate synthetase. However, those downstream genes involved in the conversion of taxa-4(5)-11(12)-diene from geranylgeranyl pyrophosphate were not expressed except for taxane 10- β -hydroxylase. Additionally, a mutant strain, *A. aculeatinus* BT-2 was obtained from the original strain, *A. aculeatinus* Tax-6, using fungicidin as the mutagenic agent. The taxol yield of BT-2 was 560 $\mu\text{g L}^{-1}$, which was higher than that of Tax-6. To identify the mechanism of the difference in taxol production, we compared the transcriptomes of the two fungi and explored the changes in the gene expression between them. When compared with the original strain, Tax-6, most genes related to the MVA pathway in the mutant strain BT-2 showed upregulation, including GGPPS. Moreover, most of the downstream genes were not expressed in the mutant fungi as well. Overall, the results revealed the pathway and mechanism of taxol synthesis in endophytic fungi and the potential for the construction of taxol-producing genetic engineering strains.

Taxol, one of the taxanes, can be used for curing tumors as an integration in multidrug regimens^{1,2}. Formerly, taxol could be obtained from *Taxus chinensis*, but since the trees are a rare plant species, the concentration of taxol in the extract is only 0.01–0.03%. Thus, chemical synthesis has been explored for taxol production^{4,5}. However, the technologies are inhibited for practical use owing to their complicated and expensive processes, which has led researchers to focus on biological methods^{6–8}. Endophytic fungi in yew trees can be used for biological fermentation of taxol production, and has been observed to be a promising method because the fungi are widely available and taxol is produced conveniently by fermentation⁹.

To date, numerous endophytic fungi have been found to produce taxol, the hosts of which are the *Taxus* spp., *Platyclusus orientalis* (L.), *Wollemlia pine*, *Torreya grandifolia*, *Terminalia arjuna*^{10–13}. However, the taxol concentration of most of the microbial fermentation solutions is lower than 1 mg L^{-1} , which makes it difficult to be used in industrial applications⁹. In view of this shortcoming, several methods have been proved to increase the produced amount of fermentation taxol¹⁴. For example, a mutant strain was obtained from taxol-producing fungi, which showed a production of 225.2 $\mu\text{g L}^{-1}$ vs. 20 $\mu\text{g L}^{-1}$ in the original strain¹⁵. Additionally, some chemicals can induce an *Aspergillus niger* strain to improve the taxol production by two folds¹⁶. However, the production

Department of Environmental Engineering, College of Biology and the Environment, Nanjing Forestry University, Nanjing 210037, Jiangsu Province, China. ✉email: hqgwc@njfu.edu.cn

concentration of taxol is still unsatisfactory; therefore, studies have been conducted to use molecular biotechnology to modify the expression of genes involved in taxol biosynthesis. Apart from these, genome shuffling has been conducted to increase taxol production¹⁷. Ajikumar et al. developed a novel method to promote taxol production via multivariate-modular engineering in the metabolic-pathway. The titers of taxadiene, the first committed taxol intermediate, increased to 1 g L⁻¹ in *Escherichia coli*¹⁸. Nevertheless, only very few successful studies have been reported, primarily because of the lack of molecular information regarding the biosynthesis of taxol in fungi¹⁹.

Next generation sequencing is an advanced technology for complex transcriptome analysis, which allows for a quick and comprehensive insight into gene structure, facilitating the recognition of target genes and the analysis of gene expression of an organism^{19,20}. De novo assembly performs well for short read sequence data and has been widely used to investigate the synthesis of taxol in *Taxus*. Qiao et al.²¹ analyzed the transcriptome of *Cephalotaxus hainanensis* and found a large number of gene numbers involved in taxol synthesis. Sun et al.²² conducted comparative transcriptome analysis between non-elicited *Taxus X Media* and methyl jasmonate-elicited cultures and found that some genes were involved in the synthesis of terpenoid backbone and the bioconversion steps from geranyl pyrophosphate to 10-deacetyl Baccatin III. However, to the best of our knowledge, investigations of the taxol biosynthesis pathway and taxol-related genes in taxol-producing endophytic fungi remain unclear to date²³.

In our previous study²⁴, we isolated a taxol-producing endophytic fungus, *Aspergillus aculeatinus* Tax-6, from *Taxus X Media*, and researched a method for enhancing taxol production. In this study, we obtained a mutant strain with a higher taxol production from Tax-6 to comprehensively understand the taxol synthesis pathway in endophytic fungi. By comparing the gene expression of the wild-type and mutant strains of Tax-6, more than 60 genes involved in taxol synthesis and transcriptome factors that may influence the taxol biosynthesis pathway were obtained. We also found that the mutant strain performed better under environmental stress.

Materials and methods

Taxol-producing strain. The taxol-producing endophytic fungus *A. aculeatinus* Tax-6 (CCTCC M 2,016,614) was isolated from the bark of *Taxus X Media*²⁴.

Extraction of fungal taxol. The extraction of fungal taxol followed the method reported by Wang, et al.¹⁶. First, the endophytic fungus was aerobically cultured in potato dextrose broth medium at 28 °C for eight days. The entire culture was passed through four layers of cheesecloth, after which the culture fluid was collected and extracted three times (about two hours each time), then further purified with 1/10 volume of hexane. All of the organic phase were collected and then dried at 45 °C by rotary evaporator (Buchi R-200, Switzerland). Finally, the residue was dissolved in 1 mL of methanol and filtered through a 0.2 µm polymeric filter.

Determination of fungal taxol. A C18 column (250 mm × 4.4 mm × 10 µm) was used for separation of taxol production on a Dionex Ultimate-3000 HPLC system. The mobile phase consisted of methanol: water (70:30) applied at a flow rate of 1 mL min⁻¹ and a column temperature of 40 °C. The registration of the peak and retention time was recorded by reading the UV absorbance at 227 nm. The fungal taxol was quantified by comparing the peak area to that of a standard sample (standard taxol was diluted to 0.2, 0.4, 0.6, 0.8 and 10 mg L⁻¹ by methanol). The HPLC chromatogram including the sample and the standard compound of taxol have been provided in our previous study²⁴.

Mutagenesis assay of taxol-producing strain. A chemical mutagenesis assay was applied to obtain mutant fungus with more taxol production using mycostatin as the mutagen. *A. aculeatinus* Tax-6 was cultured in potato dextrose agar (PDA) medium with a certain concentration of mycostatin (CAS:1400-61-9, Sigma-Aldrich) solution on a culture dish. After five days, the colonies were picked out from the dishes in which the fungi had high fatality rates (90–95%) were picked out from the dishes, and then introduced by the streak plate method into a fresh PDA medium to culture for five days again. The selected fungus was repeatedly purified (five times) through inoculation before the mutant strain was obtained. The mutant fungus with the highest taxol yield was then named BT-2.

Transcriptome analysis. *RNA extraction.* Total RNA of the two fungi was extracted by the CTAB method according to the protocol of the RNA extraction kit (Invitrogen 15596-026). The residue of DNA was removed by DNase (Sangon Biotech, China). Qubit2.0 (Q32866, Invitrogen, USA) was used to measure the concentration of the extracted RNA, and then gel electrophoresis (DYY-11, LiuYi Instrument, China) was used to test the genome contamination and integrity of the extracted RNA. Each experiment had three biological replicates for the transcriptome sequencing (RNA-seq).

Library construction and Illumina sequencing. The preparation of the transcriptome library of the two fungi was conducted by Sango Biotechnology (Shanghai, China). A VAHTS™ mRNA-seq V2 Library Prep Kit for Illumina® (Vazyme Biotech Co., Ltd, China) was used to construct an RNA-seq library according to the manufacturer's protocol. Briefly, an amount of 200 ng of the total RNA sample was bound to oligo-dT attached magnetic beads to select poly-A mRNA. The selected mRNAs were fragmented to 150–500 bp at high temperature in the presence of magnesium ions. The cleaved mRNA fragments were then used to synthesize first strand cDNA using reverse transcriptase and random hexamer primers. Second strand cDNA synthesis followed, using DNA Polymerase I and RNase H. The cDNA products were purified with VAHTS DNA Clean Beads, end-repaired and A-tailed using End Rep Mix and dA-Tailing Buffer Mix, and then ligated with RNA Adapters. The library was amplified by PCR, and its quality was checked on Agilent 2,100 Bioanalyzer using Agilent DNA 1,000 kit

(Agilent, Cat. No. 5067-1504). The bioanalyzer results should reveal a peak of products at the size expected based on the average fragment size and length of the adaptors. Each single cDNA library was sequenced using Illumina HiSeq 2000 (Illumina, San Diego, CA, USA) and a v2.5 TruSeq Paired End HiSeq flow cell/cluster kit (Illumina, San Diego, CA, USA) according to the manufacturer's instructions²⁵. The length of paired end was 2×150 bp.

De novo assembly. Prinseq v 0.19.5 (<https://prinseq.sourceforge.net/>) was used to control the quality of the raw reads. The sequencing connectors were removed first by Cutadapt (<https://pypi.python.org/pypi/cutadapt/1.2.1>), and the bases were removed when the quality of their read tails was lower than 20, and then the N-containing sequences (< 35 bp) were cut before the assessment of sequences contamination. De novo assembly was employed to construct the transcripts because of the absence of a reference genomic sequence. The raw reads of the samples after quality control were combined, and then were assembled de novo using Trinity v r20140717 (<https://trinityrnaseq.sourceforge.net/>)²⁶. The overlapping reads with specific length (> 200 bp) were first assembled into contigs, after which the reads were mapped back onto the contigs using RSeQC v 2.6.1 (<https://rseqc.sourceforge.net/>) with the paired-end method and gap-filling in silico. Chimeric reads were subsequently excluded from the missing assembly using three or more read pairs as the criterion to define order and distance between two contigs. This step was able to detect the contigs from the same transcript and could be used to calculate the distance between two contigs. The contigs were then assembled to obtain longer sequences, but this step ended when constructed sequences could not be extended any more by the Trinity software and they were considered unique transcripts. Finally, the unique genes were obtained using the sequence-splicing redundancy removal routine to process the unique transcripts.

EST-SSR detection and primer design. The unique genes and transcripts were screened for SSR markers using MISA (<https://pgrc.ipk-gatersleben.de/misa/>). The parameters were adjusted for identification of perfect mono-nucleotide, di-nucleotide, tri-nucleotide, tetra-nucleotide, pentanucleotide, and hexa-nucleotide motifs with a minimum of 10, 6, 5, 5, 5, and 5 repeats, respectively^{27,28}. The primers were designed by primer3 using default parameters (<https://primer3.sourceforge.net/>).

Transcriptome sequencing data analysis. The unigenes gene sequences were blasted using NCBI blast 2.2.28 + software with the NR (NCBI Non-redundant Protein Sequences), NT (NCBI Nucleotide Sequences), KOG (EuKaryotic Ortholog Groups), CDD (Conserved Domain Database), PFAM (Protein Family), Swissprot (a manually annotated and reviewed protein sequence database), TrEMBL, GO (Gene Ontology), and KEGG (Kyoto Encyclopedia of Genes and Genomes) libraries²⁹⁻³¹, respectively. All the annotation details with a similarity > 30% and $E < 1 \times 10^{-5}$ were obtained through the combination of the unigenes. To classify the functions of genes obtained, BLASTX searches of the NR database, the SwissProt database and TrEMBL were carried out using an E-value cutoff of 1×10^{-5} .

Based on the annotation in the NR database, the Blast2Go program (<https://www.blast2go.com/b2gHOME>) was employed to classify the gene functions into molecular functions, biological processes and cellular components³². After retrieving the associated GO terms, the possible functions of unigenes were predicted against KOG database. In parallel, the KEGG database (<https://www.genome.jp/kegg>) was used to conduct the pathway assignments of the unique genes using BLASTX²⁹.

To investigate the response of chemical mutation on original fungus Tax-6, Fragments Per Kilobase of transcript per Million fragments mapped (FPKM) was used as an index to measure the gene expression level using the RSEM software, which can quantify transcript abundance from RNA-seq data without a reference genome³³, after which we conducted false discovery rate (FDR) calculation and statistical analysis by EdgeR³⁴. We defined differential expressed gene (DEG) as a gene with twofold expression changes between *A. aculeatinus* Tax-6 and BT-2 samples and FDR of less than 0.05.

qRT-PCR analysis. The qRT-PCR experiments were used for detecting the expression levels of six genes involved in taxol biosynthesis³⁵, which were performed using gene-specific primers with a QuantiFast SYBR Green PCR Kit (Qiagen) on a Bio-Rad T100™ Thermal Cycler Real-Time PCR Detection System (Bio-Rad)³⁶. The primer sequences were designed on two consecutive exons and the amplification efficiencies are shown in Table S1. The real-time melting curve and agarose gel electrophoresis were used to check the specificity of the primers³⁶. qRT-PCR reactions were performed in 20 μ L volumes containing 2 μ L of diluted cDNA, 0.4 μ L of 10 μ M forward primer, 0.4 μ L of 10 μ M reverse primer, and 10 μ L of 2 \times SybrGreen qPCR Master Mix (Applied Biosystems). The thermal conditions of the qRT-PCR reactions were 95 °C for 3 min, and 45 cycles of 95 °C for 15 s, 57 °C for 20 s, and 72 °C for 30 s³⁷. The relative expression levels were calculated using the $2^{-\Delta\Delta C_t}$ method³⁸. GAPDH was used as the housekeeping gene. Each experiment was performed with three biological and technical replicates²⁴. Analysis of variance (ANOVA) was used to examine differences in mean values. The 95% confidence interval ($p < 0.05$) was set as the significance threshold, which was used to evaluate the significance of the difference in gene expression between strains Tax-6 and BT-2³⁶.

Results

Growth and taxol yield of the mutant strain BT-2. In this study, mycostatin was used to induce the genetic mutation of the strain Tax-6. As shown in Fig. S1, mycostatin can inhibit the growth of the original strain Tax-6, and the strain can't grow when mycostatin concentration in the medium was 15 mg L⁻¹. The mutant strain BT-2 was selected after the genetic mutation of mycostatin to Tax-6. The highest taxol yield of BT-2 was 560 μ g L⁻¹ on day 8, which was higher than that of Tax-6²⁴ (Fig. 1).

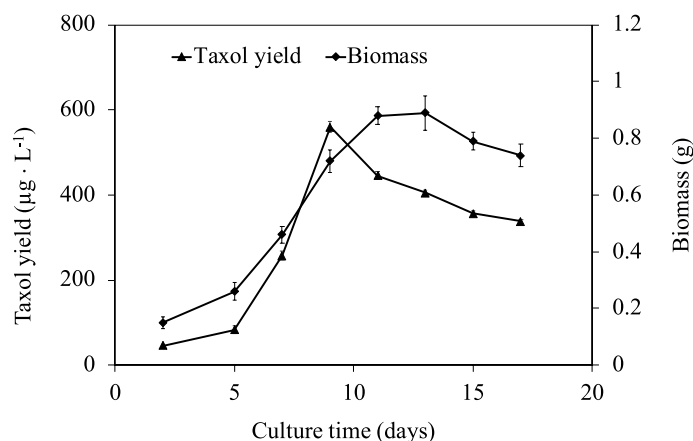


Figure 1. Taxol yield and biomass of the mutant strain BT-2 cultured in PBD medium for different culture time.

Strains	Before QC		≥ Q20 (%)	After QC	
	Total reads	Base numbers (bp)		Total reads	Base numbers (bp)
Tax-6	55,976,870	8,396,530,500	99.86	55,898,665	7,985,573,998
BT-2	68,241,536	10,236,230,400	99.72	68,048,645	9,871,570,321

Table 1. The result of removing miscellaneous from the original data.

Sequencing and assembly. The transcriptomes of original strain *A. aculeatinus* Tax-6 and its mutant *A. aculeatinus* BT-2 were sequenced with Illumina sequencing technology to generate 55,976,870 and 68,241,536 raw reads, respectively (Table 1). The coverage of sequencing is shown in Fig. S2. The coverage was calculated by samtools (<https://samtools.sourceforge.net/>), which could reflect the proportion of fully detected and undetected genes in each sample, and could find whether there were more specific genes among samples. After strict quality control, 55,898,665 and 68,048,645 clean reads were obtained from *A. aculeatinus* Tax-6 and BT-2, and were combined before being assembled to 45,242 unigenes. The N50 and mean size of the combined unigenes were 2029 bp and 995 bp, respectively (Table S2). The length distribution of unigenes from 200 to more than 2000 bp is shown in Fig. S3. The distribution of GC content of the combined unigenes and transcripts is shown in Fig. S4. The combined transcripts data from *A. aculeatinus* Tax-6 and BT-2 are also shown in Table S2. 50 species were specified to assess the pollution of the sequences. The results showed that the first few species were very close to the source *A. aculeatinus* (Table S3), which indicated that the sample had no obvious pollution although a small read number of other species were identified.

Functional annotation. The proportion of annotation of each database is shown in Table S4. These annotations gave us a valuable resource for suggesting specific pathways and function genes in *A. aculeatinus* Tax-6. To classify the functions of genes obtained, BLASTX searches of the NR database were conducted using an E-value cutoff of 1.0×10^{-5} , accounting for 68.89% of all unigenes. Species distribution analysis revealed that *A. aculeatinus* Tax-6 had a number of homologous sequences with several species, and the genes from *Chaetomium globosum* had the highest similarity (30.23%), followed by *A. niger* (7.13%), *A. niger* CBS 513.88(7.08%), *Beauveria bassiana* D1-5 (6.28%), and *A. kawachii* IFO 4,308 (5.55%) (Fig. 2). In addition, 23,462 unique sequences (51.86%) were annotated into GO terms using Blast2Go, 5,555 (12.28%) of which were annotated with KEGG pathways using the Single-directional Best Hit (SBH) method³⁹. Additionally, 22,474 unigenes (49.68%), 22,794 unigenes (50.38%), 19,462 unigenes (43.02%), 20,874 unigenes (46.14%), 14,410 unigenes (31.85%), 31,026 unigenes (68.58%) and 3,350 unigenes (20.17%) were annotated in the databases of CDD, NT, PFAM, Swiss-Prot, KOG and TrEMBL, respectively (Table S4). In total, 34,742 unigenes (76.79%) could be annotated with at least one database, and 3,765 unigenes (8.32%) could be annotated in all databases, with relatively good defined functional annotations.

The GO classification data was combined from *A. aculeatinus* Tax-6 and BT-2. The obtained putative gene functions assigned by homology searches were classified into biological processes, cellular components and molecular function based on the results of Blast Uniprot (Fig. 3). The results showed that encoded proteins involved in biological processes ranked first, including metabolic processes, cellular processes and single-organism processes. For the cellular component domain, most of the unique sequences were located in the cell and cell parts. For the molecular function domain, most unigenes took part in catalytic activity and binding. The high number of genes involved in metabolic processes and cell parts indicated that the endophytic fungi could actively produce secondary metabolites like taxol.

NR species distribution

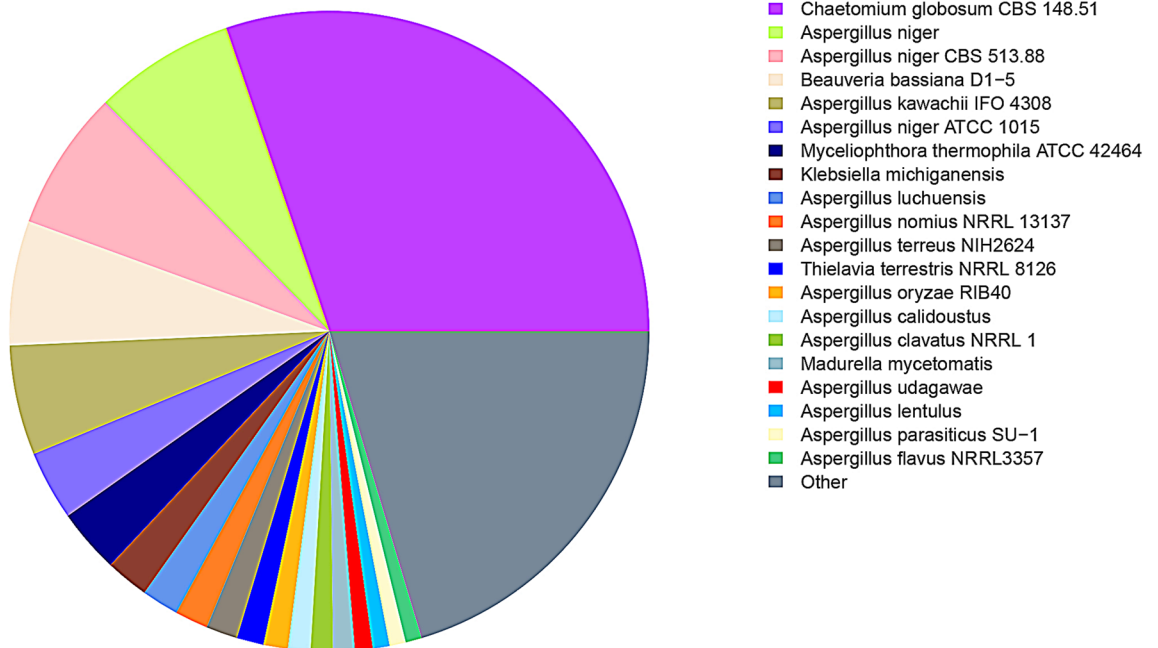


Figure 2. Percentage of top blast hits by species.

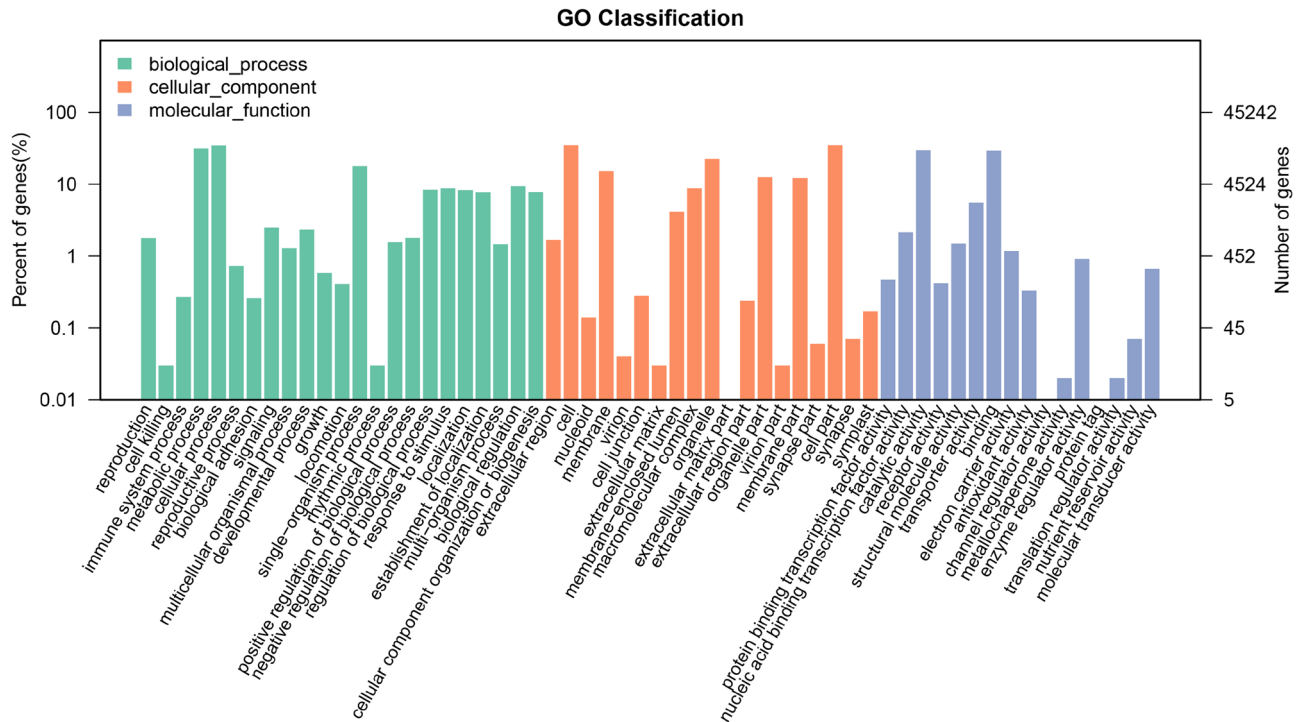


Figure 3. GO classification of Tax-6 and BT-2.

To understand the metabolic pathways in Tax-6 and BT-2, a BLASTX search against the KEGG database was used to map unique sequences to specific pathways as shown in Fig. 4. A total of 5,555 unique genes were annotated with the KEGG database, and the pathways were classified into cellular processes, environmental information processes, genetic information processes, metabolism and organismal systems. The largest unique pathways were carbohydrate metabolism, translation, signal transduction and amino acid metabolism.

KEGG Classification

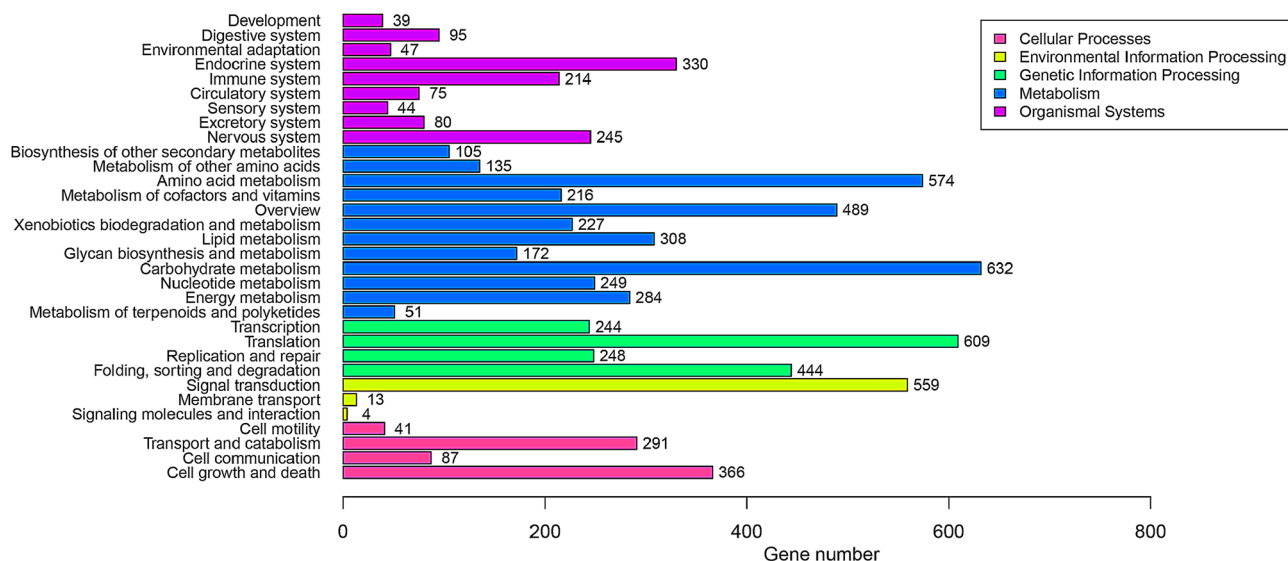


Figure 4. KEGG pathway classification^{29–31}.

Genes involved in taxol synthesis in *A. aculeatinus* Tax-6. The biosynthesis pathway of taxol in *Taxus* plants has been revealed, which includes the synthesis of the precursor of terpenoid— isopentenyl pyrophosphate (IPP), carbon ring backbone—Baccatin III and side chains of taxol. However, the biosynthesis pathway of taxol originated from endophytic fungi was not clear. According to the KEGG annotation in this study, over 20 unique sequences were involved in terpenoid backbone biosynthesis (Table S5), including 1-deoxyxylulose 5-phosphate reductoisomerase (DXR), 4-hydroxy-3-methylbut-2-enyl diphosphate reductase (HDR), 3-hydroxy-3-methylglutaryl-coenzyme A synthase (HMGS) and 4-hydroxy-3-methylbut-2-enyl diphosphate reductase (HMDR). These included genes were related to IPP and dimethylallyl diphosphate (DMAPP) biosynthesis in the mevalonate (MVA) pathway and the non-mevalonate (2-C-Methyl-D-erythritol-4-phosphate, MEP) pathway^{40,41}. In addition, genes responsible for geranyl pyrophosphate (GPP) and geranylgeranyl pyrophosphate (GGPP) synthesis were also found⁴², which are the important taxol precursor and can promote the production of taxol. However, genes involved in the conversion of taxa-4(5)-11(12)-diene from GGPP, the key step in taxol production, were not found (Table 1). Taxa-4(5)-11(12)-diene is a key carbon ring backbone of the taxol structure. The unique sequence was found to be homologous to taxane 10- β -hydroxylase (T10 β H) involved in taxol synthesis by *Ozonium* sp. BT2, which is used in taxol biosynthesis and is known as P450 hydroxylase⁴³.

Comparison of the expression levels of genes between *A. aculeatinus* Tax-6 and BT-2. As shown in Fig. 5, most of the genes involved in metabolic pathways in the mutant fungus, *A. aculeatinus* BT-2, changed when compared with that in Tax-6; moreover, there were 11,279 up-regulated genes and 8,097 down-regulated genes in BT-2 (Fig. S5). About 15.69% (3,041) of the unique genes showed changes of more than two folds in 312 pathways. The increase in pathways included terpenoid backbone biosynthesis, the cell cycle and the citrate cycle. Pathways that showed decreased gene abundance were related to phenylalanine metabolism, metabolism of xenobiotics by cytochrome P450 and base excision repair. As shown in Fig. S6, in the terpenoid backbone biosynthesis pathway, the MVA pathway and MEP pathway were both found^{29–31}. When compared with the original fungus Tax-6, most genes showed up-regulated expression in this pathway in BT-2, including GPPS, which transferred the product of the dimethylallyl-PP pathway to GPP, the backbone of taxol. As shown in Fig. S7, when compared with Tax-6, the gene expression of phenylalanine biosynthesis changed in the phenylalanine metabolism pathway of BT-2^{29–31}. According to Fleming et al.⁴⁴, phenylalanine is converted to β -phenylalanine and then further to phenylisoserine, after which phenylisoserine is combined with the backbone of taxol (like Baccatin III) and forms the side chain of taxol. Fig. S8 shows that the expression of genes related to glycine, serine and threonine metabolism changed in BT-2 compared with Tax-6^{29–31}. Most of these genes were up-regulated, which showed that doses of glycine, serine and threonine could promote taxol production. The result is consistent with the previous researches^{24,45}.

DXR, 1-hydroxy-3-methylglutaryl enzyme A reductase (HMGR), IPPS, GPPS, and geranylgeranyl diphosphate synthase (GGPPS) are the key enzymes involved in the MVA pathway and MEP pathway, and T10 β H is used to catalyze taxadiene to Baccatin III. To find out effects of important enzymes on taxol biosynthesis, we compare the expression levels of DXR, HMGR, IPPS, GPPS, GGPPS and T10 β H between Tax-6 and BT-2 by qRT-PCR. As shown in Fig. 6, compared to Tax-6, the expressions of genes involved in the six enzymes in BT-2 were all upregulated. However, compared with Tax-6, the expression of GGPPS in BT-2 was up-regulated more significantly than that of other enzymes ($p < 0.01$), whereas the change in T10 β H expression was not obvious ($p > 0.05$).

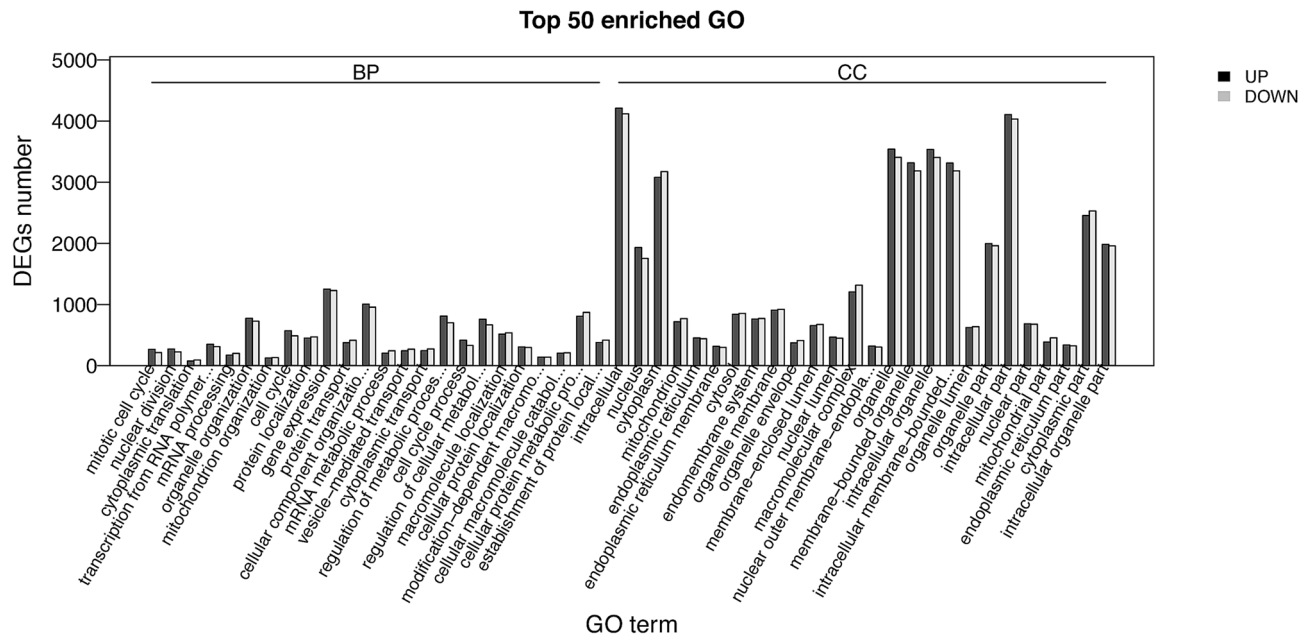


Figure 5. The different expression gene set (DEGs) number that were significantly different ($p < 0.05$) when comparing *A. aculeatinus* Tax-6 with BT-2.

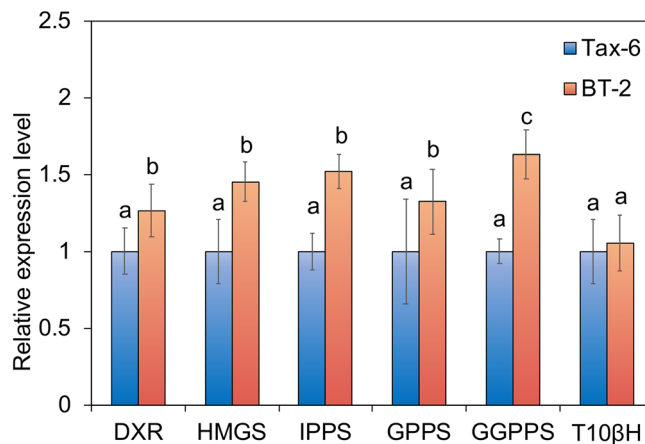


Figure 6. Comparison of transcriptome expression levels of genes involved in key enzymes involved in the biosynthetic pathway of paclitaxel between *A. aculeatinus* Tax-6 and BT-2. Values with different letters (a–c) differ significantly ($p < 0.05$). Error bars represent standard deviations ($n = 3$).

Discussion

The isoprenoids, which are the important precursors of taxol biosynthesis, can convert to IPP and DMAPP in higher plants^{46,47}. The biosynthesis of IPP and DMAPP includes the MVA and MEP pathways (Fig. 7)⁴⁸. In this study, the genes involved in the MVA pathway were found, such as HMGS, HMGR and HMDR⁴⁹, which catalyze the condensation of acetyl CoA and acetoacetyl CoA to form HMG CoA⁴⁹.

In the plastidial MEP pathway with 1-deoxyxylulose 5-phosphate (DOXP), pyruvate converts into glyceraldehyde-3-phosphate^{18,47,48}, and DXR and HDR are the key enzymes^{46,50,51}. The expressions of genes coding these two key enzymes were observed through the transcriptome analysis in this study. Moreover, the expressions of these two genes in BT-2 were up-regulated when compared with those of Tax-6.

GPPS, farnesyl pyrophosphate synthetase (FPPS) and GGPPS are the important enzymes for the synthesis of GGPP from GPP and IPP. GGPPS plays an important role in regulating carbon flow^{42,52}. In this study, the expressions of the gene coding GPP and GGPPS were found in Tax-6, which showed that the endophytic fungus had a similar isoprenoids synthesis pathway as *Taxus sp.* Although we did not observe the expression of the gene coding FPPS, this also provided the molecular evidence of which the endophytic fungus in *Taxus sp.*, which can produce taxol.

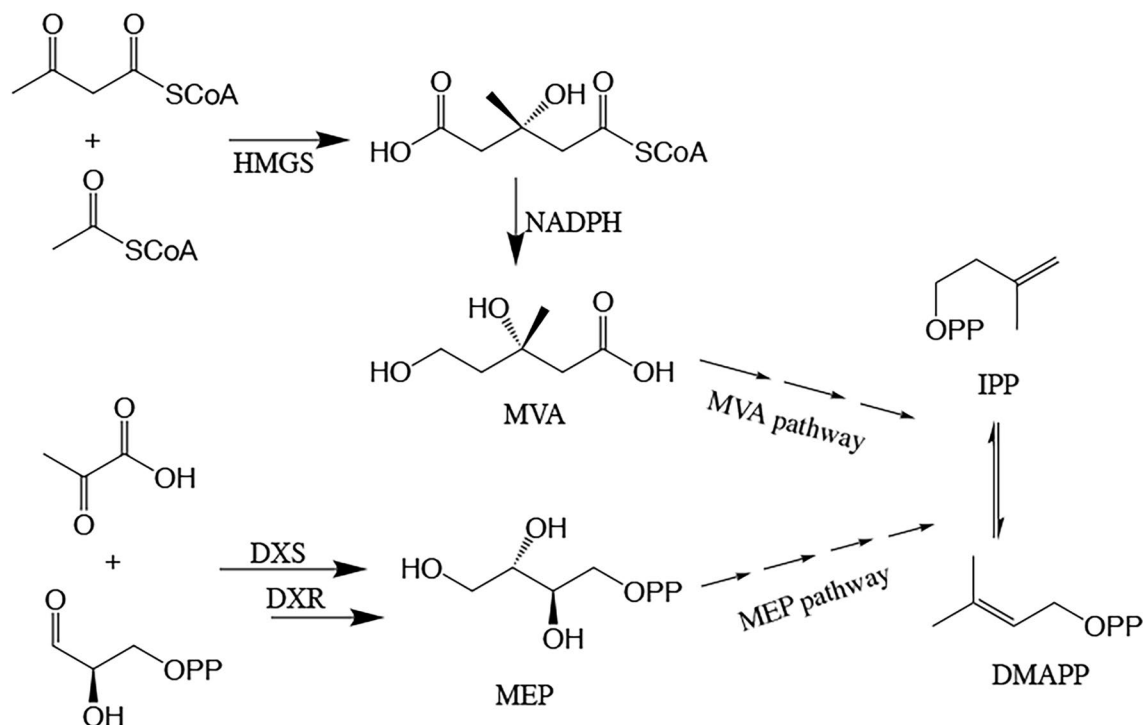


Figure 7. The MVA and MEP pathway of taxol biosynthesis. (Numbers of arrows represent the number of steps. MVA: mevalonate; DOXP: 1-deoxyxylulose 5-phosphate; IPP: isopentenyl pyrophosphate; DMAPP: dimethylallyl diphosphate; HMGS: 3-hydroxy-3-methylglutaryl-coenzyme A synthase; DXR: 1-deoxyxylulose 5-phosphate reductoisomerase; DXS: 1-deoxyxylulose 5-phosphate synthase; NADPH: triphosphopyridine nucleotide; MVA: mevalonate; MEP: 2-C-methyl-D-erythritol-4-phosphate).

The cyclization condensation of GGPP is catalyzed to form Taxa-4,11-diene by Taxadiene synthetase (TS) using GGPPS as a raw material⁵³. The following synthesis steps of taxol are seven hydroxylation and five acetyl-transferase reactions of taxa-4,11-diene, and the benzoyl acylation of side chains on C13 of the taxol backbone (Fig. 8). However, we did not observe an obvious expression of these genes through our transcriptome analysis except for that of T10βH. We suggest that this is the important reason for low taxol production from *A. aculeatinus*.

As a cytochrome P450, T10βH can hydroxylate taxadiene to taxadien-5α-ol and can catalyze taxadien-5α-yl acetate to 5α,10β-taxadien-diol monoacetate. The cytochrome P450-dependent hydroxylation reactions involved in T10βH are the second and fourth specific steps of taxol biosynthesis diverging from primary metabolism (i.e., from GGPP)⁵⁴.

From the transcriptome results, we found the expression of the most upstream genes. However, the genes involved in another key step for taxol biosynthesis, the conversion of GGPP to taxadiene, were not found, nor were the key downstream genes, except for T10βH. When compared with other transcriptome analyses of *Taxus*, our study found relatively fewer enzymes involved in taxol biosynthesis^{22,55}. This was likely because transcriptome did not match the genome perfectly, the expression of genes related to taxol production was subjected to environmental changes, and some genes may not be expressed or may only be expressed at low levels until some elicitors are added or are subjected to certain environment stress⁵⁶. We inferred that the absence of those functional genes was possibly because they were not expressed, or because taxol biosynthesis in endophytic fungi is different from that in plants⁸. Additionally, genes involved in the cell cycle were more active in BT-2, which could lead to the production of more biomass in the same culture time, which may also contribute to the production of taxol⁵⁷. The upregulation of genes related to glycine metabolism may also promote taxol production because glycine has been reported to increase taxol production⁵⁸. Finally, we identified NADH dehydrogenase and transcription factors RfeF(AFUA_4G10200)⁵⁹, basic leucine zipper (bZip)⁶⁰, and Specificity protein 1 (SFP1)⁶¹, which may also contribute to the biosynthesis of taxol, but the mechanism through which this occurs is still unclear.

Conclusion

We obtained a taxol-producing endophytic fungus, BT-2, by chemical mutagenesis, with a relatively higher taxol yield, conducted transcriptome analysis between two taxol-producing fungi with different taxol production and identified the potential mechanism for the changes in production. The results showed that the expressions of genes involved in the MVA and MEP pathway were almost observed. Moreover, these genes in the mutant strain were up regulated. However, the key downstream genes related to taxol biosynthesis, i.e., TS were hardly found, were far less there were far in fungi than that in *Taxus* plant. In our view, our transcriptome data not only revealed the molecular mechanism by which the endophytic fungus can produce taxol⁸, but also explained the reason

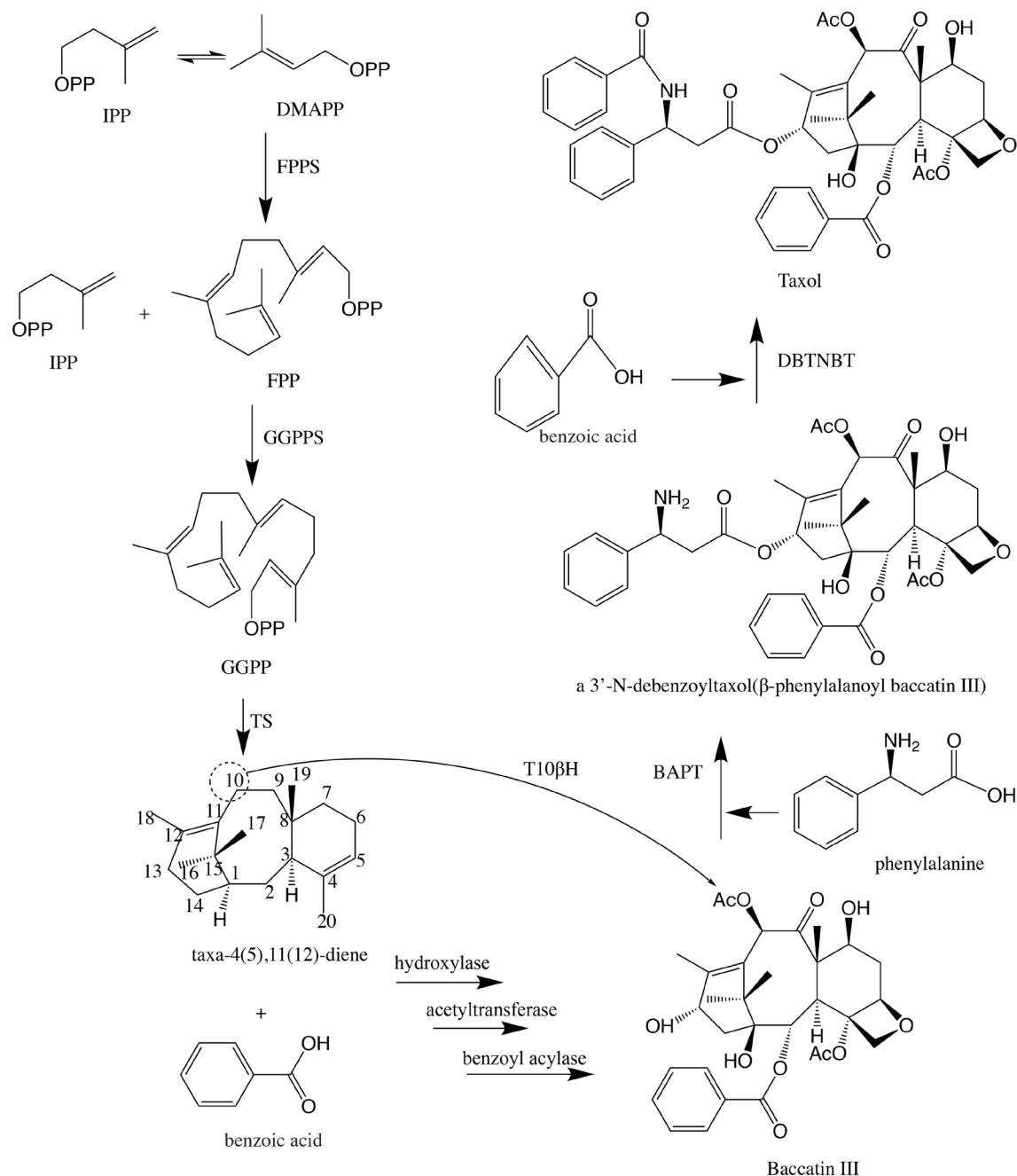


Figure 8. The downstream pathway of taxol biosynthesis. (Numbers of arrows represent the number of steps. DBAT: 10-deacetyl Baccatin III-O-acetyltransferase; BAPT: Baccatin III 13-O-(3-amino-3-phenylpropanoyl) transferase; DBTNT: 3'-N-debenzoyl-2'-deoxytaxol-N-benzoyltransferase; IPP: isopentenyl pyrophosphate; DMAPP: dimethylallyl diphosphate; FPPS: farnesyl pyrophosphate synthetase; GGPPS: geranylgeranyl pyrophosphate synthase; GGPP: geranylgeranyl pyrophosphate).

why the fungal taxol yields were still low. This study is the first to analyze the transcriptome of taxol-producing fungi and facilitate the promotion of taxol production by *A. aculeatinus*. However, additional studies to clarify the entire pathway of taxol biosynthesis in the endophytic fungi should be performed.

Received: 17 May 2019; Accepted: 11 June 2020

Published online: 29 June 2020

References

1. Yuan, J. *et al.* Stable gene silencing of cyclin B1 in tumor cells increases susceptibility to taxol and leads to growth arrest *in vivo*. *Oncogene* 25, 1753–1762. <https://doi.org/10.1038/sj.onc.1209202> (2006).

2. Heinig, U., Scholz, S. & Jennewein, S. Getting to the bottom of taxol biosynthesis by fungi. *Fungal Divers.* **60**, 161–170. <https://doi.org/10.1007/s13225-013-0228-7> (2013).
3. Croom, E. M. Jr. Taxol and taxoids. *Sci. Appl.* **37**, 2701–2702 (1995).
4. Nicolaou, K. C. *et al.* Total synthesis of taxol. *Nature* **367**, 630–634. <https://doi.org/10.1038/367630a0> (1991).
5. Borman, S. Taxol to be made from renewable materials by organic semisynthesis. *Chem. Eng. News* **70**, 30–32. <https://doi.org/10.1021/cen-v070n041.p030> (1992).
6. Pezzuto, J. Taxol production in plant cell culture comes of age. *Nat. Biotechnol.* **14**, 1083. <https://doi.org/10.1038/nbt0996-1083> (1996).
7. Sah, B., Subban, K. & Chelliah, J. Cloning and sequence analysis of 10-deacetylbaconin III-10-O-acetyl transferase gene and WRKY1 transcription factor from taxol-producing endophytic fungus *Lasiodiplodia theobromae*. *FEMS Microbiol. Lett.* <https://doi.org/10.1093/femsle/fnx253> (2017).
8. Kusari, S., Singh, S. & Jayabaskaran, C. Rethinking production of Taxol(R) (paclitaxel) using endophyte biotechnology. *Trends Biotechnol.* **32**, 304–311. <https://doi.org/10.1016/j.tibtech.2014.03.011> (2014).
9. Gond, S. K., Kharwar, R. N. & White, J. F. Will fungi be the new source of the blockbuster drug taxol?. *Fungal Biol. Rev.* **28**, 77–84. <https://doi.org/10.1016/j.fbr.2014.10.001> (2014).
10. Strobel, G. A., Torczynski, R. & Bollon, A. *Acremonium* sp.-a leucinostatin A producing endophyte of European yew (*Taxus baccata*). *Plant Sci.* **128**, 97–108. [https://doi.org/10.1016/s0168-9452\(97\)00131-3](https://doi.org/10.1016/s0168-9452(97)00131-3) (1997).
11. Gangadevi, V. & Muthumary, J. A novel endophytic taxol-producing fungus *Chaetomella raphigera* isolated from a medicinal plant, *Terminalia arjuna*. *Appl. Biochem. Biotechnol.* **158**, 675–684. <https://doi.org/10.1007/s12010-009-8532-0> (2009).
12. Li, J. Y. *et al.* The induction of taxol production in the endophytic fungus—*Periconia* sp. from *Torreya grandifolia*. *J. Ind. Microbiol. Biotechnol.* **20**, 259–264. <https://doi.org/10.1038/sj.jim.2900521> (1998).
13. Uzma, F. *et al.* Endophytic fungi-alternative sources of cytotoxic compounds: a review. *Front. Pharmacol.* **9**, 309. <https://doi.org/10.3389/fphar.2018.00309> (2018).
14. Heinig, U. & Jennewein, S. Taxol: a complex diterpenoid natural product with an evolutionarily obscure origin. *Afr. J. Biotechnol.* **8**, 1370–1385. <https://doi.org/10.5897/AJB2009.000-9217> (2009).
15. Xu, F., Tao, W., Cheng, L. & Guo, L. Strain improvement and optimization of the media of taxol-producing fungus *Fusarium mairei*. *Biochem. Eng. J.* **31**, 67–73. <https://doi.org/10.1016/j.bej.2006.05.024> (2006).
16. Wang, Y. *et al.* Isolation and characterization of endophytic huperzine A-producing fungi from *Huperzia serrata*. *J. Ind. Microbiol. Biotechnol.* **38**, 1267–1278. <https://doi.org/10.1007/s10295-010-0905-4> (2011).
17. Zhao, K. *et al.* Screening and breeding of high taxol producing fungi by genome shuffling. *Sci. China Ser. C: Life Sci.* **51**, 222–231. <https://doi.org/10.1007/s11427-008-0037-5> (2008).
18. Ajikumar, P. K. *et al.* Isoprenoid pathway optimization for taxol precursor overproduction in *Escherichia coli*. *Science* **330**, 70–74. <https://doi.org/10.1126/science.1191652> (2010).
19. Shendure, J. & Ji, H. Next-generation DNA sequencing. *Nat. Biotechnol.* **26**, 1135–1145. <https://doi.org/10.1038/nbt1486> (2008).
20. Morozova, O. & Marra, M. A. Applications of next-generation sequencing technologies in functional genomics. *Genomics* **92**, 255–264. <https://doi.org/10.1016/j.ygeno.2008.07.001> (2008).
21. Qiao, F. *et al.* De Novo characterization of a *Cephalotaxus hainanensis* transcriptome and genes related to paclitaxel biosynthesis. *PLoS ONE* **9**, e106900. <https://doi.org/10.1371/journal.pone.0106900> (2014).
22. Sun, G. *et al.* Deep sequencing reveals transcriptome re-programming of *Taxus x media* cells to the elicitation with methyl jasmonate. *PLoS ONE* **8**, e62865. <https://doi.org/10.1371/journal.pone.0062865> (2013).
23. Miao, L.-Y. *et al.* Transcriptome analysis of a taxol-producing endophytic fungus *Cladosporium cladosporioides* MD2. *AMB Express* <https://doi.org/10.1186/s13568-018-0567-6> (2018).
24. Qiao, W., Ling, F., Yu, L., Huang, Y. & Wang, T. Enhancing taxol production in a novel endophytic fungus, *Aspergillus aculeatinus* Tax-6, isolated from *Taxus chinensis* var. *mairei*. *Fungal Biol.* **121**, 1037–1044. <https://doi.org/10.1016/j.funbio.2017.08.011> (2017).
25. Caporaso, J. G. *et al.* Ultra-high-throughput microbial community analysis on the Illumina HiSeq and MiSeq platforms. *ISME J* **6**, 1621–1624. <https://doi.org/10.1038/ismej.2012.8> (2012).
26. Grabherr, M. G. *et al.* Full-length transcriptome assembly from RNA-Seq data without a reference genome. *Nat. Biotechnol.* **29**, 644–652. <https://doi.org/10.1038/nbt.1883> (2011).
27. Wu, X., Wen, Y., Ueno, S. & Tsumura, Y. Development and characterization of EST-SSR markers for *Taxus mairei* (Taxaceae) and their transferability across species. *Silvae Genet.* **65**, 67–70. <https://doi.org/10.1515/sg-2016-0009> (2016).
28. Wang, L., Yu, C., Guo, L., Lin, H. & Meng, Z. In silico comparative transcriptome analysis of two color morphs of the common coral trout (*Plectropomus leopardus*). *PLoS ONE* **10**, e0145868. <https://doi.org/10.1371/journal.pone.0145868> (2016).
29. Kanehisa, M. & Goto, S. KEGG: Kyoto encyclopedia of genes and genomes. *Nucleic Acids Res.* **28**, 27–30. <https://doi.org/10.1093/nar/28.1.27> (2000).
30. Kanehisa, M., Sato, Y., Furumichi, M., Morishima, K. & Tanabe, M. New approach for understanding genome variations in KEGG. *Nucleic Acids Res.* **47**, D590–D595. <https://doi.org/10.1093/nar/gky962> (2019).
31. Kanehisa, M. Toward understanding the origin and evolution of cellular organisms. *Protein Sci.* <https://doi.org/10.1002/pro.3715> (2019).
32. Ana, C. *et al.* Blast2GO: a universal tool for annotation, visualization and analysis in functional genomics research. *Bioinformatics* **21**, 3674–3676. <https://doi.org/10.1093/bioinformatics/bti610> (2005).
33. Li, B. & Dewey, C. N. RSEM: accurate transcript quantification from RNA-Seq data with or without a reference genome. *BMC Bioinform.* **12**, 323 (2011).
34. Robinson, M. D., McCarthy, D. J. & Smyth, G. K. edgeR: a bioconductor package for differential expression analysis of digital gene expression data. *Bioinformatics* **26**, 139–140. <https://doi.org/10.1093/bioinformatics/btp616> (2010).
35. Liao, W. *et al.* Transcriptome assembly and systematic identification of novel cytochrome P450s in *Taxus chinensis*. *Front. Plant Sci.* <https://doi.org/10.3389/fpls.2017.01468> (2017).
36. Yu, D. J., Zhang, G. M., Chen, Z. L., Zhang, R. J. & Yin, W. Y. Rapid identification of *Bactrocera latifrons* (Dipt., Tephritidae) by real-time PCR using SYBR Green chemistry. *J. Appl. Entomol.* **128**, 670–676. <https://doi.org/10.1111/j.1439-0418.2004.00907.x> (2004).
37. Soliman, S. S. M., Trobacher, C. P., Tsao, R., Greenwood, J. S. & Raizada, M. N. A fungal endophyte induces transcription of genes encoding a redundant fungicide pathway in its host plant. *BMC Plant Biol.* <https://doi.org/10.1186/1471-2229-13-93> (2013).
38. Amini, S.-A., Shabani, L., Afghani, L., Jalalpour, Z. & Sharifi-Tehrani, M. Squalenol-induced production of taxol and baccatin in cell suspension culture of yew (*Taxus baccata* L.). *Turk. J. Biol.* **38**, 528–536. <https://doi.org/10.3906/biy-1401-47> (2014).
39. Kanehisa, M. *et al.* KEGG for linking genomes to life and the environment. *Nucleic Acids Res.* **36**, D480–484. <https://doi.org/10.1093/nar/gkm882> (2008).
40. Liu, T. & Khosla, C. A balancing act for taxol precursor pathways in *E. coli*. *Science* **330**, 44–45. <https://doi.org/10.1126/science.1195014> (2010).
41. Ma, H., Lu, Z., Liu, B., Qiu, Q. & Liu, J. Transcriptome analyses of a Chinese hazelnut species *Corylus mandshurica*. *BMC Plant Biol.* <https://doi.org/10.1186/1471-2229-13-152> (2013).
42. Hefner, J., Ketchum, R. E. & Croteau, R. Cloning and functional expression of a cDNA encoding geranylgeranyl diphosphate synthase from *Taxus canadensis* and assessment of the role of this prenyltransferase in cells induced for taxol production. *Arch. Biochem. Biophys.* **360**, 62–74. <https://doi.org/10.1006/abbi.1998.0926> (1998).
43. Yu, C. *et al.* Identification of potential genes that contributed to the variation in the taxoid contents between two *Taxus* species (*Taxus media* and *Taxus mairei*). *Tree Physiol.* **37**, 1659–1671. <https://doi.org/10.1093/treephys/tpx091> (2017).

44. Fleming, P. E., Mocek, U. & Floss, H. G. Biosynthesis of taxoids. mode of formation of the taxol side chain. *J. Am. Chem. Soc.* **115**, 805–807. <https://doi.org/10.5897/AJB11.170> (1993).
45. Fett-Neto, A. G., Melanson, S. J., Nicholson, S. A., Pennington, J. J. & Dicosmo, F. Improved taxol yield by aromatic carboxylic acid and amino acid feeding to cell cultures of *Taxus cuspidata*. *Biotechnol. Bioeng.* **44**, 967–971. <https://doi.org/10.1002/bit.260440813> (1994).
46. Lange, B. M., Rujan, T., Martin, W. & Croteau, R. Isoprenoid biosynthesis: the evolution of two ancient and distinct pathways across genomes. *Proc. Natl. Acad. Sci. USA* **97**, 13172–13177. <https://doi.org/10.1073/pnas.240454797> (2000).
47. Laule, O. *et al.* Crosstalk between cytosolic and plastidial pathways of isoprenoid biosynthesis in *Arabidopsis thaliana*. *Proc. Natl. Acad. Sci. USA* **100**, 6866–6871. <https://doi.org/10.1073/pnas.1031755100> (2003).
48. Bick, J. A. & Lange, B. M. Metabolic cross talk between cytosolic and plastidial pathways of isoprenoid biosynthesis: unidirectional transport of intermediates across the chloroplast envelope membrane. *Arch. Biochem. Biophys.* **415**, 146–154. [https://doi.org/10.1016/s0003-9861\(03\)00233-9](https://doi.org/10.1016/s0003-9861(03)00233-9) (2003).
49. Kai, G. *et al.* Molecular cloning and expression analyses of a new gene encoding 3-hydroxy-3-methylglutaryl-CoA synthase from *Taxus x media*. *Biol. Plant.* **50**, 359–366. <https://doi.org/10.1007/s10535-006-0050-0> (2006).
50. Takahashi, S., Kuzuyama, T., Watanabe, H. & Seto, H. A 1-deoxy-D-xylulose 5-phosphate reductoisomerase catalyzing the formation of 2-C-methyl-D-erythritol 4-phosphate in an alternative nonmevalonate pathway for terpenoid biosynthesis. *Proc. Natl. Acad. Sci. USA* **95**, 9879–9884. <https://doi.org/10.1073/pnas.95.17.9879> (1998).
51. Adam, P. *et al.* Biosynthesis of terpenes: studies on 1-hydroxy-2-methyl-2-(E)-butenyl 4-diphosphate reductase. *Proc. Natl. Acad. Sci. USA* **99**, 12108–12113. <https://doi.org/10.1073/pnas.182412599> (2002).
52. Aharoni, A. *et al.* Terpenoid metabolism in wild-type and transgenic *Arabidopsis* plants. *Plant Cell* **15**, 2866–2884. <https://doi.org/10.1105/tpc.016253> (2003).
53. Koepp, A. E. *et al.* Cyclization of geranylgeranyl diphosphate to taxa-4(5),11(12)-diene is the committed step of taxol biosynthesis in *Pacific yew*. *J. Biol. Chem.* **270**, 8686–8690. <https://doi.org/10.1074/jbc.270.15.8686> (1995).
54. Schoendorf, A., Rithner, C. D., Williams, R. M. & Croteau, R. B. Molecular cloning of a cytochrome P450 taxane 10 beta-hydroxylase cDNA from *Taxus* and functional expression in yeast. *Proc. Natl. Acad. Sci. USA* **98**, 1501–1506. <https://doi.org/10.1073/pnas.98.4.1501> (2001).
55. Prashant, S. & Balasankar, K. Comparative transcriptome analysis of non-model *Taxus* species for taxol biosynthesis. *J. Chem. Pharm. Res.* **7**, 800–805 (2015).
56. Matsuura, H. N. *et al.* Specialized plant metabolism characteristics and impact on target molecule biotechnological production. *Mol. Biotechnol.* **60**, 169–183. <https://doi.org/10.1007/s12033-017-0056-1> (2018).
57. Chen, J.-J. *et al.* Combinatorial mutation on the beta-glycosidase specific to 7-beta-xylosyltaxanes and increasing the mutated enzyme production by engineering the recombinant yeast. *Acta Pharm. Sin. B* **9**, 626–638. <https://doi.org/10.1016/j.apsb.2018.11.003> (2019).
58. Liu, S. P. *et al.* Heterologous pathway for the production of L-phenylglycine from glucose by *E. coli*. *J. Biotechnol.* **186**, 91–97. <https://doi.org/10.1016/j.jbiotec.2014.06.033> (2014).
59. Carberry, S., Neville, C. M., Kavanagh, K. A. & Doyle, S. Analysis of major intracellular proteins of *Aspergillus fumigatus* by MALDI mass spectrometry: identification and characterisation of an elongation factor 1B protein with glutathione transferase activity. *Biochem. Biophys. Res. Commun.* **341**, 1096–1104. <https://doi.org/10.1016/j.bbrc.2006.01.078> (2006).
60. Izawa, T., Foster, R. & Chua, N. H. Plant bZIP protein DNA binding specificity. *J. Mol. Biol.* **230**, 1131–1144. <https://doi.org/10.1006/jmbi.1993.1230> (1993).
61. Marion, R. M. *et al.* Sfp1 is a stress- and nutrient-sensitive regulator of ribosomal protein gene expression. *Proc. Natl. Acad. Sci. USA* **101**, 14315–14322. <https://doi.org/10.1073/pnas.0405353101> (2004).

Acknowledgements

This research was supported by the Priority Academic Program Development of Jiangsu Higher Education Institutions (PAPD).

Author contributions

W.Q. wrote the main manuscript text, T.T. prepared Figs. 1–8, and F.L. prepared the supplementary materials. All authors reviewed the manuscript.

Competing interests

The authors declare no competing interests.

Additional information

Supplementary information is available for this paper at <https://doi.org/10.1038/s41598-020-67614-1>.

Correspondence and requests for materials should be addressed to W.Q.

Reprints and permissions information is available at www.nature.com/reprints.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2020