

Transcript profiling of jasmonate-elicited *Taxus* cells reveals a β -phenylalanine-CoA ligase

Karla Ramírez-Estrada¹, Teresa Altabella^{1,2}, Miriam Onrubia³, Elisabeth Moyano³, Cedric Notredame^{3,4}, Lidia Osuna⁵, Robin Vanden Bossche^{6,7}, Alain Goossens^{6,7}, Rosa M. Cusido¹ and Javier Palazon^{1,*}

¹Secció de Fisiologia Vegetal, Facultat de Farmàcia, Universitat de Barcelona, Barcelona, Spain

²Center for Research in Agricultural Genomics (CRAG) (CSIC-IRTA-UAB-UB), Cerdanyola, Barcelona, Spain

³Departament de Ciències Experimentals i de Salut, Universitat Pompeu Fabra, Barcelona, Spain

⁴Comparative Bioinformatics, Centre for Genomic Regulation (CRG), Barcelona, Spain

⁵Centro de Investigación Biomédica del Sur, Instituto Mexicano del Seguro Social (IMSS), Xochitepec, Mexico

⁶Department of Plant Systems Biology, VIB, Ghent, Belgium

⁷Department of Plant Biotechnology and Bioinformatics, Ghent University, Ghent, Belgium

Received 20 November 2014;

revised 28 January 2015;

accepted 16 February 2015.

*Correspondence (Tel +34 934020267;

fax +34 93 4029043 9 3313809;

email javierpalazon@ub.edu)

Accession number: The sequences reported in this study have been deposited in the GenBank database: TB768: KM593667; TB506: KP178208; TB331: KP178206; TB574: KP178209; TB224: KP178204; TB328: KP178205; TB596: KP178210; TB812: KP178213; TB047: KP178200; TB739: KP178212; TB212: KP178203; TB645: KP178211; TB493: KP178207; TB195: KP178202; TB115: KP178201.

Keywords: cDNA-AFLP, CoA ligase, methyl jasmonate, β -phenylalanine, Taxol, *Taxus baccata*, cell cultures.

Summary

Plant cell cultures constitute eco-friendly biotechnological platforms for the production of plant secondary metabolites with pharmacological activities, as well as a suitable system for extending our knowledge of secondary metabolism. Despite the high added value of taxol and the importance of taxanes as anticancer compounds, several aspects of their biosynthesis remain unknown. In this work, a genomewide expression analysis of jasmonate-elicited *Taxus baccata* cell cultures by complementary DNA-amplified fragment length polymorphism (cDNA-AFLP) indicated a correlation between an extensive elicitor-induced genetic reprogramming and increased taxane production in the targeted cultures. Subsequent *in silico* analysis allowed us to identify 15 genes with a jasmonate-induced differential expression as putative candidates for genes encoding enzymes involved in five unknown steps of taxane biosynthesis. Among them, the *TB768* gene showed a strong homology, including a very similar predicted 3D structure, with other genes previously reported to encode acyl-CoA ligases, thus suggesting a role in the formation of the taxol lateral chain. Functional analysis confirmed that the *TB768* gene encodes an acyl-CoA ligase that localizes to the cytoplasm and is able to convert β -phenylalanine, as well as coumaric acid, into their respective derivative CoA esters. β -phenylalanyl-CoA is attached to baccatin III in one of the last steps of the taxol biosynthetic pathway. The identification of this gene will contribute to the establishment of sustainable taxol production systems through metabolic engineering or synthetic biology approaches.

Introduction

Taxol is the registered trade name of Bristol-Myers Squibb for the generic drug paclitaxel, which is currently considered as one of the most effective anticancer drugs ever developed. The production of taxol in *Taxus* spp. cell cultures has become a major challenge for plant biotechnology due to the growing worldwide demand and its scarcity in nature (Onrubia *et al.*, 2013a). The natural source of taxol is the inner bark of several *Taxus* spp., where it accumulates at a very low concentration. For these reasons, several research laboratories and biotechnological companies are involved in designing new strategies based on elicited *Taxus* cell cultures (ETCCs) to optimize the biotechnological production of this anticancer compound. ETCCs also represent an excellent platform to carry out basic studies of taxane biosynthesis aimed at deepening our understanding of this metabolic pathway and its regulation (Cusido *et al.*, 2014).

In the highly complex taxol biosynthetic pathway, the first product bearing the taxane ring is taxa-4(5),11(12)-diene, catalysed from (E,E,E)-geranylgeranyl diphosphate (GGPP) by the plastidic enzyme taxadiene synthase (TXS) (Hezari *et al.*, 1997; Koepp *et al.*, 1995). A succession of steps, involving hydroxylases, acyl transferases and other enzymes, leads to the production of 10-deacetyl baccatin III (10-DABIII) and baccatin III (BIII), both taxoids now widely used as advanced intermediates for the semi-synthesis of taxol and its analogues (Croteau *et al.*, 2006; Malik *et al.*, 2011; Onrubia *et al.*, 2013a). Subsequently, a β -phenylalanine-derived side chain is attached to BIII, and after two more biosynthetic steps, the target molecule is assembled (Figure 1). Croteau and co-workers have identified, cloned and heterologously expressed the genes encoding twelve enzymes involved in taxol biosynthesis: GGPP synthase (GGPPS), TXS, five hydroxylases, five acyl transferases and phenylalanine aminomutase (PAM) (see reviews: Croteau *et al.*, 2006; Guo *et al.*, 2006; Lange and

Please cite this article as: Ramírez-Estrada, K., Altabella, T., Onrubia, M., Moyano, E., Notredame, C., Osuna, L., Vanden Bossche, R., Goossens, A., Cusido, R.M. and Palazon, J. (2015) Transcript profiling of jasmonate-elicited *Taxus* cells reveals a β -phenylalanine-CoA ligase. *Plant Biotechnol. J.*, doi: 10.1111/pbi.12359

Ahkami, 2013; Nims *et al.*, 2006; Walker and Croteau, 2001). However, the genes encoding three other hydroxylases, an oxidase, epoxidase, oxomutase, and β -phenylalanine-CoA ligase have remained unidentified, with seven steps of the pathway yet to be defined (Croteau *et al.*, 2006; Nims *et al.*, 2006) (Figure 1). The side chain attached to the C13 position of the taxane core is known to be derived from β -phenylalanine, but some doubts remain regarding its formation. Although phenylisoserine has also been postulated as the side chain (Floss and Mocek, 1995), Croteau *et al.* (2006) reported that β -phenylalanine was incorporated into BIII three times more efficiently than phenylisoserine. The activation of β -phenylalanine to β -phenylalanyl-CoA by HSCoA is catalysed by β -phenylalanine-CoA ligase, but the structure of this enzyme and the identity of the gene encoding it have remained undetermined to date.

Plant cells generally respond to elicitor action by increasing the production of secondary metabolites involved in the plant

defence system. This effect has been repeatedly described in *Taxus* spp. cell cultures treated with biotic and abiotic elicitors (Cusido *et al.*, 2002; Bentebibel *et al.*, 2005; Li *et al.*, 2012; Onrubia *et al.*, 2013b). The enhanced production usually results from a higher expression of the genes controlling the metabolic pathways (De Geyter *et al.*, 2012; Expósito *et al.*, 2009; Frense, 2007; Malik *et al.*, 2011; Vongpaseuth and Roberts, 2007). Hence, comparing the transcriptome of elicited cultures of a selected cell line with an unelicited control is an effective strategy to acquire new insights into the genetic response to elicitation and, ultimately, to launch successful gene discovery programs in plant metabolism, particularly in species with no genome sequences available, such as *Taxus* spp. (Cusido *et al.*, 2014; Goossens, 2015; Nims *et al.*, 2006 and references therein).

Previously, we have used cDNA-amplified fragment length polymorphism (AFLP) (Colling *et al.*, 2013; Goossens *et al.*, 2003) to profile a selected *T. baccata* cell line treated with methyl

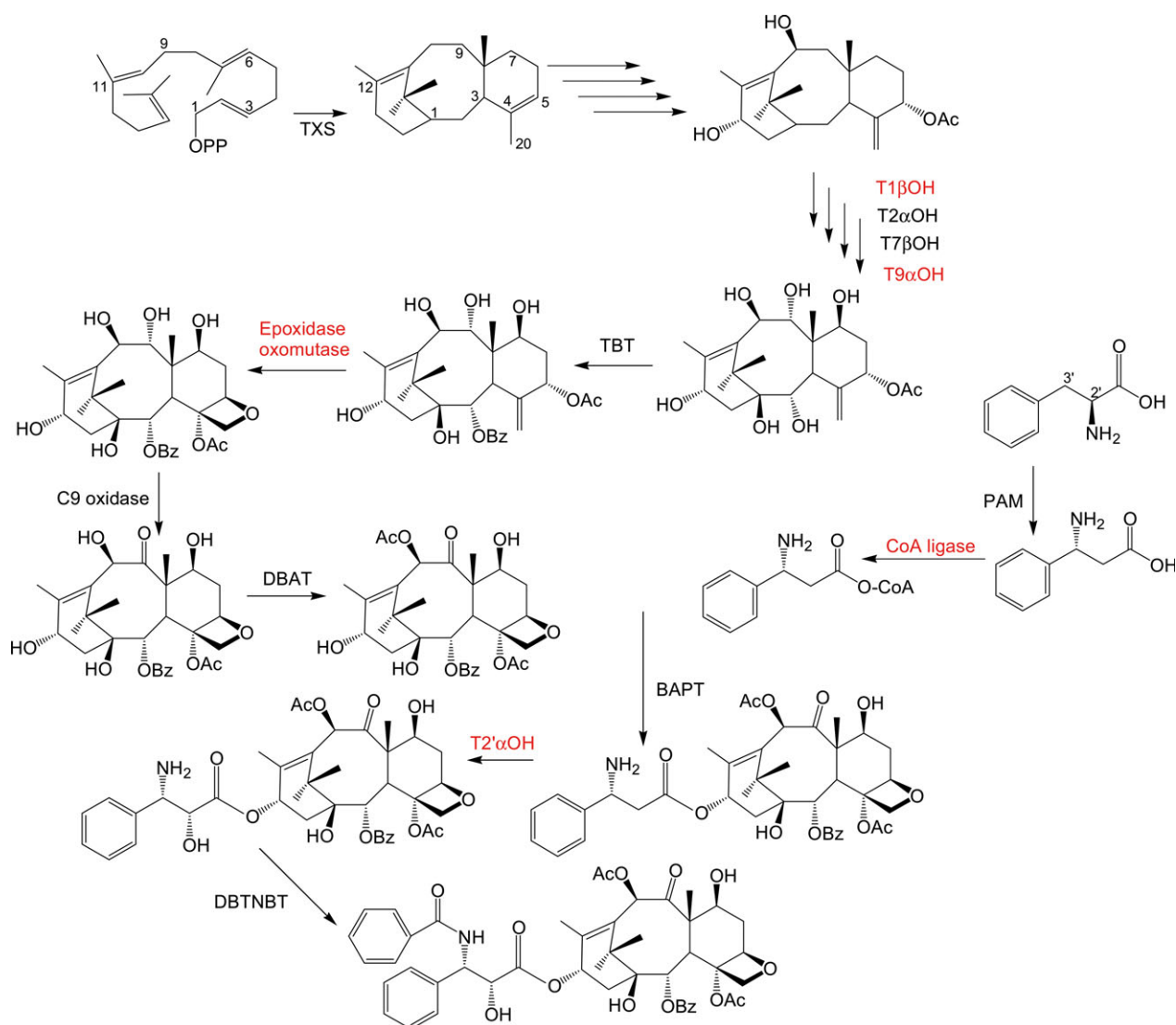


Figure 1 Postulated taxol biosynthesis from geranylgeranyl diphosphate. The multiple arrows indicate multiple steps. The unknown genes and corresponding proteins are in red. TXS, Taxadiene synthase; T1βOH, Taxane 1β-hydroxylase; T2αOH, Taxane 2α-hydroxylase; T7βOH, Taxane 7β-hydroxylase; T9αOH, Taxane 9α-hydroxylase; TBT, Taxane-2α-O-benzoyl transferase; DBAT, 10-Deacetylbaicatin III-10-O-Acetyltransferase; PAM, Phenylalanine aminomutase; BAPT, C-13-Phenylpropanoyl-CoA transferase; T2'αOH, Taxane 2'α-hydroxylase; DBTNBT, Debenzoyl taxol N-benzoyl transferase.

jasmonate (MeJA), which has led to the identification of a novel regulator of taxol biosynthesis, the small signalling peptide Taximin (Onrubia *et al.*, 2014). Here, we report on the characterization, sequencing and cloning of a gene selected through cDNA-AFLP and *in silico* analysis, which encodes β -phenylalanine-CoA ligase, an enzyme involved in the formation of the taxol side chain, one of the unresolved steps of the taxol biosynthetic pathway. The functionality of this enzyme and its subcellular localization have been determined.

Results

Transcript profiling reveals potential candidates for five unknown steps of taxane biosynthesis

In this work, cDNA-AFLP was applied for genomewide transcript profiling of *T. baccata* cells cultured in a two-stage system (Palazon *et al.*, 2003). Cell suspensions were first cultured for 7 days in the growth medium (GM), and then, they were maintained for 24 days in the production medium (PM), either with or without (control) the elicitor MeJA. As expected, based on the previous studies (Cusido *et al.*, 2002, 2014), cell metabolism in the PM culture was reprogrammed in favour of secondary metabolism, and taxane production was significantly increased in the elicited cell cultures (Figure S1).

For transcript profiling, samples were taken at 1, 2, 4, 8 and 16 h, 1, 2 and 4 days after their subculture in PM. The 7-d GM sample was selected as the negative control for secondary metabolism gene expression patterns. Using 128 BstYI–MseI primer combinations that yield 'one-gene-one-tag' (Breyne and Zabeau, 2001; Colling *et al.*, 2013; Goossens *et al.*, 2003), the quantitative temporal accumulation patterns of 8,192 transcript tags were determined and analysed. Tags of interest were those differentially expressed in the control and elicited cells (Figure 2a). Notably, a subset of these genes was more highly expressed in both control and elicited PM conditions relative to the 7-days GM condition, thus representing genes that are activated merely by the switch from GM to PM. Of the 858 differentially expressed transcript tags isolated (Figure 2b), 667 fragments (77.7%) gave good-quality sequences by direct sequencing of the PCR products, indicating they might represent unique gene tags. Homology searches with sequences from the putative unique gene tags revealed that 369 (55.3%) were similar to known genes. In Figure 2, the selected tags are classified according to the FunCat.

The similarity threshold maintained for BLAST searches was established at a value of $10e^{-2}$; however, due to the small size of some tags, in a few cases a lower *e*-value was accepted when unambiguous matches were observed. Of the sequenced tags, 309 (46.3%) displayed similarity with genes of known functions and 60 (9.0%) with genes without allocated functions (unclassified proteins); no homology with a known sequence was found for 298 of the sequenced tags (44.7% with no hits).

The major functional category was 'Metabolism and Energy' (Figure 2b), with tags homologous to genes from primary and secondary metabolism, followed by 'Cell Organization and Defence'. As we stimulated the cultures with jasmonates, which are involved in plant defence against a wide range of stimuli (Pauwels *et al.*, 2009; Wasternack, 2007), our study has produced a good list of possible candidate genes for the seven remaining unknown steps of the taxol biosynthetic pathway, which involve hydroxylases, an oxomutase, epoxidase, oxidase and CoA ligase. Tags classified in 'Metabolism and Energy' showing a clear MeJA induction pattern and homology to genes

with functions similar to those expected for the unknown steps of taxol biosynthesis were selected to study their possible involvement in this process (Figure 2a). From these gene tags, which are listed in Table 1, full-length (FL) cDNA sequences were obtained by rapid amplification of cDNA ends (RACE).

Also in the 'Metabolism and Energy' category, we found up-regulated tags corresponding to known genes involved in taxol biosynthesis (*GGPPS*, *Geranylgeranyl diphosphate synthase*; *TXS*, *Taxadiene synthase*; *TAT*, *Taxadiene-5 α -ol-O-acetyl transferase*; *T13OH*, *Taxane 13 α -hydroxylase*; *T2 α OH*, *Taxane 2 α -hydroxylase*; *T7 β OH*, *Taxane 7 β -hydroxylase*; *TBT*, *Taxane-2 α -O-benzoyl transferase*; *DBAT*, *10-deacetylbaicatin III-10-O-acetyl-transferase*; *PAM*, *Phenylalanine aminomutase*; *BAPT*, *C-13-phenylpropanoyl-CoA transferase*; *DBTNBT*, *Debenzoyl taxol N-benzoyl transferase*). qRT-PCR expression profiling showed a clear induction of these genes after 1 h of MeJA treatment up to 4 days, while no or very low and sporadic expression was observed in the nonelicited cultures (Figure S2).

Additionally, numerous novel genes, either with or without existing homologues, and with or without known functions, were identified as jasmonate-responsive genes. These included two genes both encoding a putative abietadiene synthase, an enzyme catalysing a step towards other diterpenes that are well known to occur in the resin of gymnosperms (Keeling and Bohlmann, 2006) and showing coregulation with the taxol synthesis genes, thus suggesting a coordinated activation of diterpene synthesis by elicitor treatment in general (Figure 2a). Furthermore, the jasmonate biosynthesis genes themselves were MeJA-induced (Figure S3a), confirming that the positive amplification loop in the jasmonate signal transduction cascade that is well known in angiosperms also occurs in gymnosperm species (Li *et al.*, 2012; Pauwels *et al.*, 2009). Finally, a set of genes putatively encoding enzymes involved in the biosynthesis of phenolic compounds was found to be differentially regulated by MeJA as well (Figure S3b). Although the pattern of the latter genes is less uniform than that of the taxane/diterpene pathway genes, the majority of them showed nonetheless a higher expression level after elicitor treatment, which is in agreement with the reported stimulatory effects of MeJA on the phenolic metabolism in *T. baccata* and *T. chinensis* cell cultures (Jalalpour *et al.*, 2014; Li *et al.*, 2012). Tags whose expression pattern changed or increased after MeJA elicitation and presented homology to genes encoding transcription factors or regulatory proteins were previously studied, resulting in the identification of *TB595*, a gene encoding the regulatory signalling peptide Taximin (Onrubia *et al.*, 2014).

In silico-based selection of the best candidate genes for taxane biosynthesis

To prioritize the 15 putative FL sequences of the targets indicated above (Table 1), an *in silico* procedure was used to build *ad hoc* position-specific similarity metrics. Target sequences were blasted against the SwissProt database to identify the closest homologous gene with a known function, thereby yielding a list of candidate functions and associated genes. Functional templates were then identified in functional databases (MACiE, SwissProt), looking for homologues known to interact with substrates similar to taxol intermediates. Template residues in contact with the substrates were then identified and extrapolated to the candidate genes, which were finally prioritized according to their similarity with the template on the considered substrate interacting positions. The aim of this procedure was to focus the template/candidate

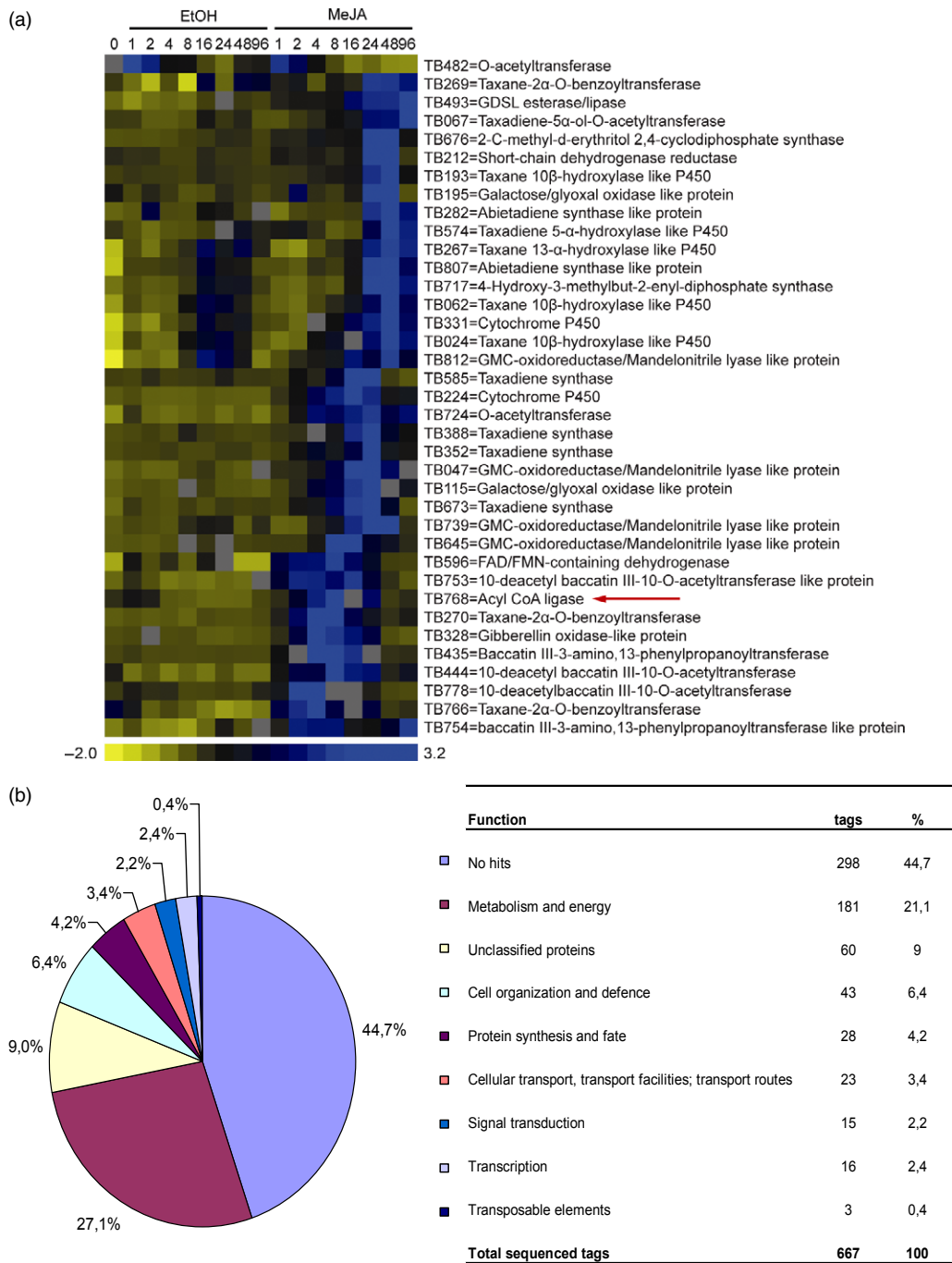


Figure 2 cDNA-AFLP-based transcript profiling of *Taxus baccata* cells. (a) Average linkage hierarchical clustering of differentially expressed gene tags identified by cDNA-AFLP. Treatments and time points (in hours/days) are indicated at the top. Blue and yellow boxes reflect transcriptional activation and repression, respectively, relative to the average expression level. Grey boxes correspond to missing time points. Shown are subclusters of the *T. baccata* genes encoding known and candidate enzymes catalysing taxol and other diterpene biosynthesis. The TB768 tag is marked by the arrowhead on the right. The numbers on the colour scale range indicate the normalized variance that is calculated as follows: the mean expression value over the time course was subtracted from each individual data point after which the obtained value was divided by the standard deviation (Goossens *et al.*, 2003). (b) Sequenced tags classified in functional categories according to the FunCat classification.

comparison on the residues most likely to determine specificity and whose evolutionary pattern is unlikely to be neutral. Using this approach, we mapped and determined the active site conservation of the 15 cDNA-AFLP candidates mentioned above, grouped in five functional categories (CoA ligase, C1 and C9 hydroxylases, C4-C20 epoxidase, C9 oxidase and oxomutase)

(Table 1). The best candidate was TB768, a putative CoA ligase, which showed the highest level of conservation on the substrate binding positions with an acyl-CoA ligase (B9GQ39) and other CoA ligases, all of them with the same e-value of 0.0 (Table 1). This protein was modelled structurally against two PDB templates (3a9v, a 4-coumarate-CoA ligase and 1ba3, a Luciferin

Table 1 Complete list of candidates grouped according to their tentative function in the taxol biosynthetic pathway (tentative function), tag name from the cDNA-AFLP analysis (Tag), the corresponding FL-cDNA sequence obtained after RACE and submitted to GenBank (GenBank Accession), the annotation and SwissProt accession number from the best hit (BLAST hit code) obtained by BLAST with the FL sequence, and the e-value with that hit

Tentative function	Tag	GenBank accession	Annotation	BLAST hit code	e-value
CoA ligase	TB768	KM593667	Acyl-CoA ligase	B9GQ39	0.0
C1 and C9 hydroxylases	TB506	KP178208	Cinnamate 4-hydroxylase	B2Z6P5	0.0
	TB331	KP178206	Putative cytochrome P450	D7T4D5	1.0E-118
	TB574	KP178209	Putative cytochrome P450	D7T4D5	1.0E-118
	TB224	KP178204	Putative cytochrome P450	D5A7Y6	1.0E-152
C4-C20 epoxidase	TB328	KP178205	Putative gibberellin oxidase	A9NW35	1.0E-45
C9 oxidase	TB596	KP178210	Unknown (FAD/FMN-containing dehydrogenase)	B8LR72	8.0E-87
	TB812	KP178213	Unknown (GMC oxidoreductase)	A9NXU5	1.0E-151
	TB047	KP178200	Unknown (GMC oxidoreductase)	A9NXU5	1.0E-109
	TB739	KP178212	Unknown (GMC oxidoreductase)	A9NXU5	1.0E-148
	TB212	KP178203	Unknown (short-chain dehydrogenase reductase)	A9NU86	6.0E-91
	TB645	KP178211	Unknown (GMC oxidoreductase)	A9NXU5	1.0E-158
	TB493	KP178207	Unknown	A9SVY3	1.0E-44
	TB195	KP178202	Putative GDSL esterase	F6GTQ6	1.0E-164
Oxomutase	TB115	KP178201	Unknown (galactose/glyoxal oxidase related)	D8RR64	1.0E-163
			Unknown (galactose/glyoxal oxidase related)		

4-monooxygenase). The Phyre2 model was deemed 100% confident (Figure 3). We eventually aligned all the considered sequences with the *Populus tomentosa* 4-coumarate-CoA ligase (4CL), the closest experimentally characterized plant protein (Q941M3). The similarity between TB768 and Q941M3 was 41% (as estimated from the T-Coffee MSA), but this increased to 81.8% when only the 25 residues most likely to interact with the substrate were considered (Figure 4).

The deduced amino acid sequence of the *TB768* gene (Figure S4) was aligned with those of other known plant 4CLs, starting with Met (Figure 4), showing an amino acid similarity ranging between 40% and 65%. 4CLs are dependent on AMP for their highest activity. Amino acid residues shown to be important for the catalytic activity of plant 4CLs (Hu *et al.*, 2010) are conserved in the *TB768* gene. These include the postulated AMP-binding site (QGYGLTE), the specific residues for AMP binding (Gly306, Ala307, Tyr330), the residues with catalytic functions (Lys438, Gln443, Lys523) and the specific residue for CoA binding (Gly441). These data corroborate the potential 4CL activity of the protein encoded by the *TB768* gene.

As an additional verification, we studied the temporal expression kinetics of this gene in the elicited *T. baccata* cell cultures also by qRT-PCR (Figure S5). *TB768* expression increased slightly when the cell cultures were transferred from the GM to the PM, following the same pattern as most of the genes involved in the taxol biosynthetic pathway, with maximum expression on the first day of elicitation (Figures S2 and S5). It is worth noting that the accumulation of the *TB768* transcript was not as high as observed for several of the hydroxylases studied and was more similar to that of the genes encoding the last transferases active in the taxol pathway. This suggests that the biosynthetic step controlled by the product of the *TB768* gene could be another limiting step in taxol formation, although further studies are needed to confirm this assumption.

***TB768* encodes a phenylalanine-CoA ligase involved in the formation of the taxol side chain**

The CoA activation of β -phenylalanine occurs towards the end of the taxol biosynthetic pathway, before the attachment of the phenylalanyl chain to the taxane backbone (Croteau *et al.*, 2006)

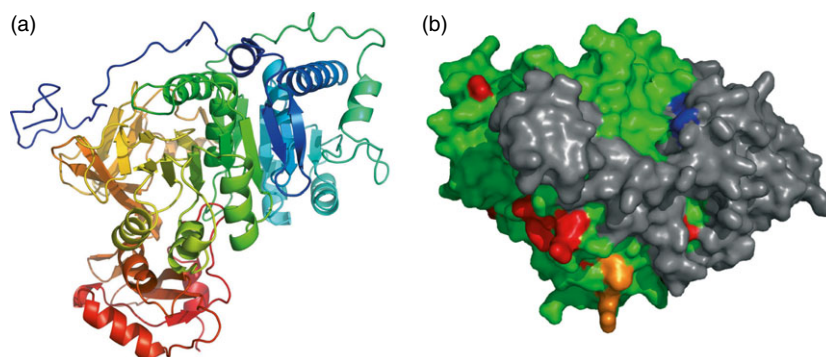


Figure 3 Phyre2 modelling and predicted conserved domains. (a) The predicted tertiary structure of the *TB768* protein. (b) Predicted CoA ligase catalytic domains of the *TB768* protein. The active site is indicated in green, the AMP-binding site in red, the putative CoA binding site in orange and the acyl activating site in blue.

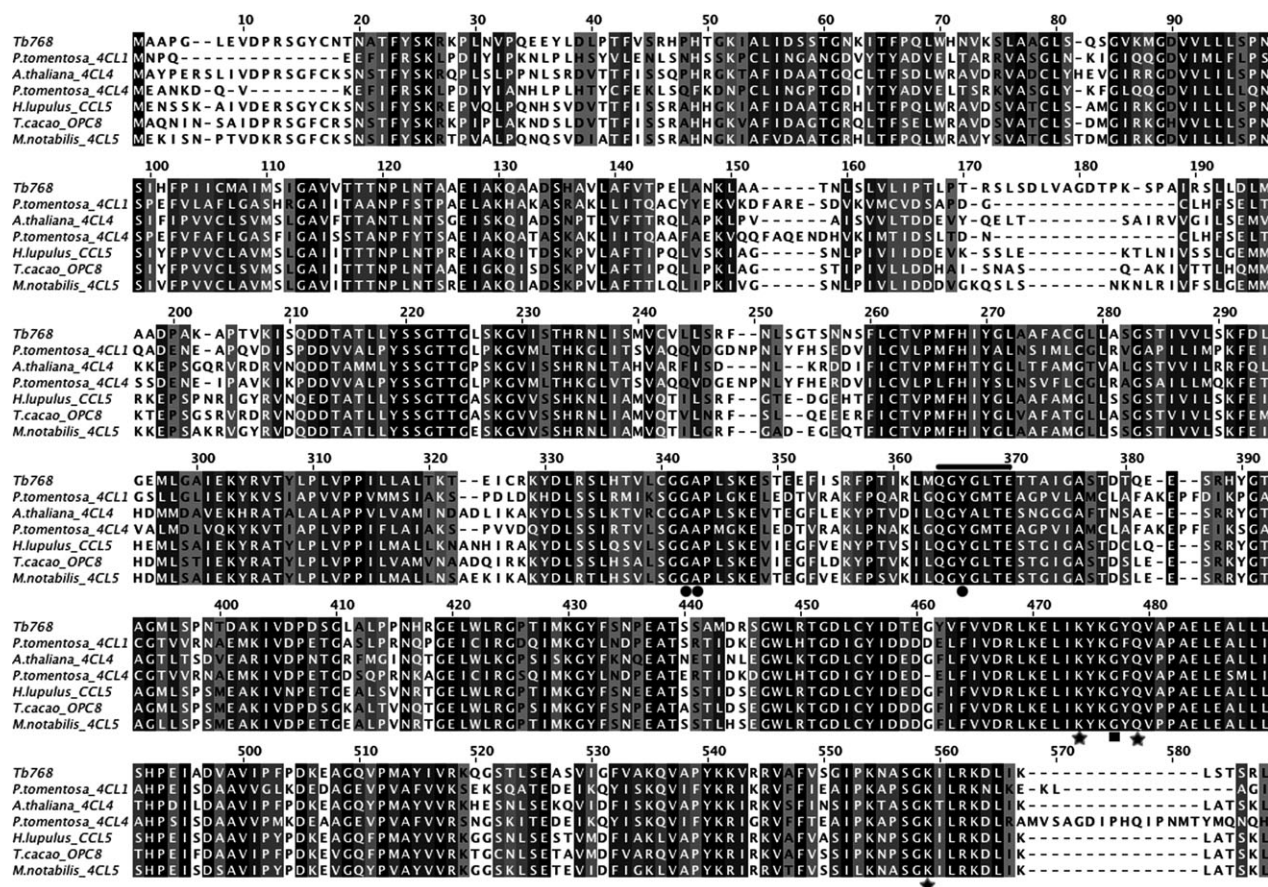


Figure 4 Multiple sequence alignment of several CoA ligases with the following sequence identity with respect to TB768: *Populus tomentosa* 4CL1: 40%; *Arabidopsis thaliana* 4CL4: 55%; *P. tomentosa* 4CL4: 43%; *Humulus lupulus* CCL5: 64%; *Teobroma cacao* OPC8: 63%; and *Morus notabilis* 4CL5: 65%. White letters on black background indicate completely conserved residues (100% SID), less conserved residues are indicated by shading. Residues (Lys438, Gln443, Lys523) with catalytic functions are labelled with stars; the specific residue (Gly441) for CoA binding is indicated with a triangle. Inside the AMP-binding surface (QGYGLTE) the specific residues for AMP binding, Gly306, Ala307, and Tyr330, are indicated by circles.

(Figure 1), and is mediated by an unknown enzyme. The selected candidate gene *Tb768* was transiently expressed as an N-terminal myc-tag fusion protein in *Nicotiana benthamiana* plants. The soluble ectopically expressed protein was purified with anti-myc-tag beads and eluted in neutral pH to maintain the putative activity of the enzyme. The isolated 59-kDa enzyme was concentrated four times and visualized by immunoblot analysis (Figure S6). For the *in vitro* enzyme activity assay, β -phenylalanine and 4-coumaric acid were added separately to the reaction mixtures containing the purified TB768 protein, ATP, $MgCl_2$ and CoA. Reaction products were detected by HPLC-tandem mass spectrometry (MS/MS). The purified TB768 enzyme was found to catalyse the CoA activation of both β -phenylalanine and 4-coumaric acid, and the respective catalytic products were identified by their retention time and the MS/MS fragmentation pattern in comparison with the standards (Figures S7 and S8). The β -phenylalanyl-CoA (m/z 916.2 eluting at 3.93 min) and the 4-coumaroyl-CoA (m/z 914.3 eluting at 3.95 min) molecules were confirmed by multiple reaction monitoring (MRM) (Figure 5), a technique based on the selection of standard predefined mass fragmentations. The precursor and product ions with the highest intensity were chosen for the MRM analysis. In the control assays without the enzyme, no ions corresponding to the acyl-CoA products were detected (Figure 5). These results confirm

that the *Tb768* gene encodes an acyl-CoA ligase that transforms β -phenylalanine into its CoA ester and may be involved in taxol biosynthesis. The enzyme was also found to be nonspecific and able to accept other substrates, such as 4-coumarate. The specific activity of the purified protein was calculated for each substrate, being 102.309 $\mu\text{mol}/\text{min}/\text{mg}$ for β -phenylalanine, and 617.76 $\mu\text{mol}/\text{min}/\text{mg}$ for 4-coumaric acid (Figure 5d). We can therefore consider that the *Tb768* gene encodes a β -phenylalanine- and 4-coumarate-CoA ligase from *T. baccata* (TBPCCL).

TBPCCL is a cytosolic enzyme

Analysis with SignalP v4.0 (Petersen *et al.*, 2011) predicted that there is no signal peptide for the TBPCCL protein (Figure 6a), suggesting that the enzyme is located in the cytosol. To investigate the subcellular localization of the TBPCCL, the *Tb768* gene was C-terminal fused to the yellow fluorescent protein (YFP) in the pEarlyGate101 vector. The resulting construct was transiently expressed in *N. benthamiana* leaves. Confocal microscopy analysis showed a YFP signal as a thick band at the cell periphery, indicating that the TBPCCL could be located in the plasma membrane or the cytosol. To discern the exact subcellular localization of the TBPCCL-YFP fusion protein, the cell walls of the *N. benthamiana* transfected leaves were stained with propidium iodide (PI) prior to confocal analysis. We were able to see that the

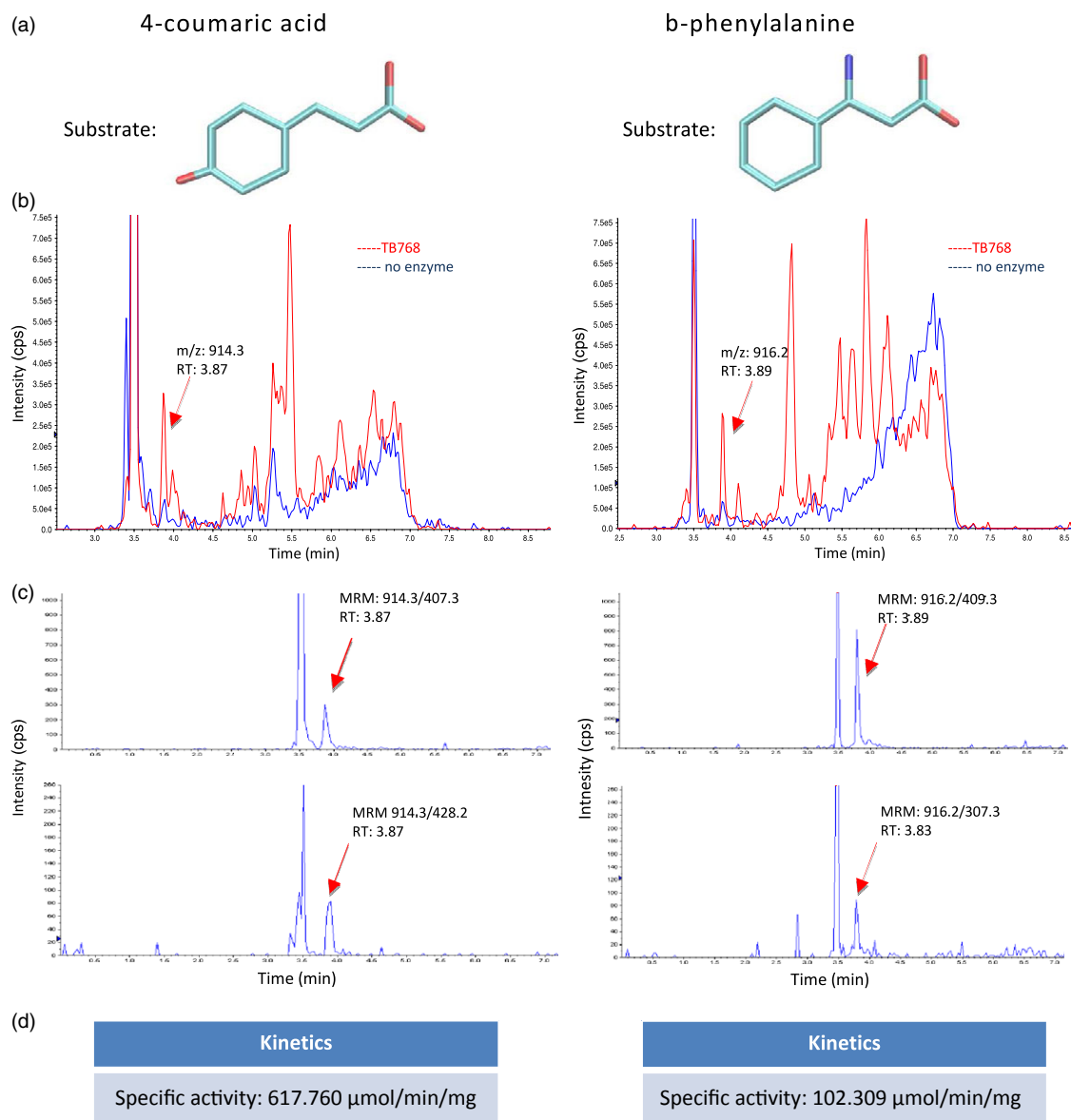


Figure 5 HPLC-MS/MS analysis of the *in vitro* enzymatic activity assays of the TB768 protein. (a) Molecular structure of TB768 substrates. (b) Extract ion chromatogram of the 4-coumaric acid and β -phenylalanine assay. The red lines correspond to the assay with the TB768 enzyme. The arrows indicate the peak of the molecular ion of each product at the corresponding retention time (3.87 min for p-coumaroyl-CoA and 3.89 min for β -phenylalanyl-CoA). The blue lines correspond to the same assay without the TB768 enzyme. (c) MRM analysis of 4-coumaric acid and β -phenylalanine-CoA ligase activity. The arrows indicate the peak at the corresponding retention time of the two specific ion fragments analysed for each molecule used to confirm their identity (914.3/407.3 and 914.3/428.2 for 4-coumaroyl-CoA with a RT: 3.87; 916.2/409.3 and 916.2/307.3 with a RT: 3.89 for β -phenylalanyl-CoA). (d) Specific activity.

PI signal was separated from the TBPCCL-YFP signal, indicating that the TBPCCL-YFP fusion protein has a cytosolic expression pattern (Figure 6b). To confirm the cytosolic location of the enzyme, we performed a fluorescence recovery after a photobleaching (FRAP) analysis, which is an effective technique that facilitates the estimation of protein mobility inside the cells. Exposure of a tagged fluorescent fusion protein to a high intensity laser light for a few seconds produces an irreversible fluorophore photobleaching in the region of interest (ROI). Afterwards, the nonphotobleached fusion proteins can migrate into the photobleached area and the fluorescence is recovered. Therefore, FRAP allowed a direct estimation of the protein mobility in the cell. Free

cytosolic proteins have high fluorescence recovery rates, as they have more mobility than those immersed in membranes (Held *et al.*, 2008; Wu *et al.*, 2006). Close to complete instant fluorescence recovery was observed in cells expressing the TBPCCL-YFP fusion protein, to almost the same extent as the free GFP-expressing cells (Figure 6d,e), thereby supporting a cytosolic location.

Discussion

In previous studies, we have substantially improved taxol production in two-stage *Taxus* spp. cell cultures by the addition

of elicitors (Palazon *et al.*, 2003; Cusido *et al.*, 2002; Bentebibel *et al.* 2005, Onrubia *et al.*, 2013b; Sabater-Jara *et al.*, 2014). However, any further enhancement in production of this anticancer agent depends on an improved understanding of its biosynthesis. This prompted us to undertake an in-depth study of the molecular mechanisms that regulate secondary metabolism in targeted *Taxus* spp. cell cultures. Accordingly, variations in the transcriptomic profile of a selected cell line were analysed to explore the relationship between the metabolic and transcriptomic response to elicitation. This approach allowed us to identify new genes potentially involved, directly or indirectly, in taxol metabolism.

Our transcriptome study showed that all genes known to be involved in taxol biosynthesis were more highly expressed after MeJA elicitation, although to a variable extent (Figures S2 and S5). Taxane production also clearly increased when MeJA was added to the medium at the beginning of the second stage of the culture, achieving the highest levels in the last days of the experiment, as repeatedly observed before (Bentebibel *et al.*, 2005; Onrubia *et al.*, 2010) (Figure S1). The predominant taxane was BIII, whose levels peaked at day 20, decreasing thereafter. At this peak, the BIII and taxol content in the elicited culture was 20.5 and 19 times higher, respectively, than that in the mock-treated culture. It can be inferred from these results that the metabolic step responsible for the formation of BIII from its precursor DABIII was not rate-limiting in the elicited *T. baccata* cell cultures studied, as BIII accumulation was high despite the very low levels of DABIII, indicating that after its formation, this precursor is probably immediately transformed into the product BIII. The lower taxol production, as compared to that of BIII, may

be due to the low expression of the two transferases responsible for connecting the side chain to BIII and/or involved in the final step in taxol biosynthesis. This fact was previously observed by Sabater-Jara *et al.* (2014) in *T. media* cell cultures.

This study allowed us to identify not only the genes encoding known key enzymes of the taxol biosynthetic pathway, whose expression was modified by the elicitor, but also possible candidate genes encoding enzymes for the unknown steps, including hydroxylases, an oxidase, oxomutase, epoxidase and CoA ligase. Bioinformatics evidence suggested that one of these candidate genes, *TB768*, might encode a CoA ligase, specifically the one controlling the activation of β -phenylalanine to form the side chain at the C13 position of BIII (Figure 1). This gene was therefore selected for further functionality studies.

The enzymes that form CoA esters of cinnamic acids are known as 4-coumarate-CoA ligases (4CLs; E.C 6.2.1.12) (Allina *et al.*, 1998). In both angiosperm and gymnosperm plants, the genes encoding these enzymes form a multigene family of differentially regulated members with possible functions in different specific biological processes, although all activate a range of very similar substrates (Hamberger and Hahlbrock, 2004). The fact that the product of the *TB768* gene activated both β -phenylalanine and 4-coumarate, which share a very similar chemical structure, suggested that this enzyme belongs to the 4CL family (Figure 4). Docimo *et al.* (2013) reported that in several plant species, the 4CL family members show a variable expression pattern and substrate specificity. In rice, the expression profile of the five known 4CL genes differs according to the tissues, developmental stage and stress response, and all the encoded enzymes show distinct kinetic properties in function of

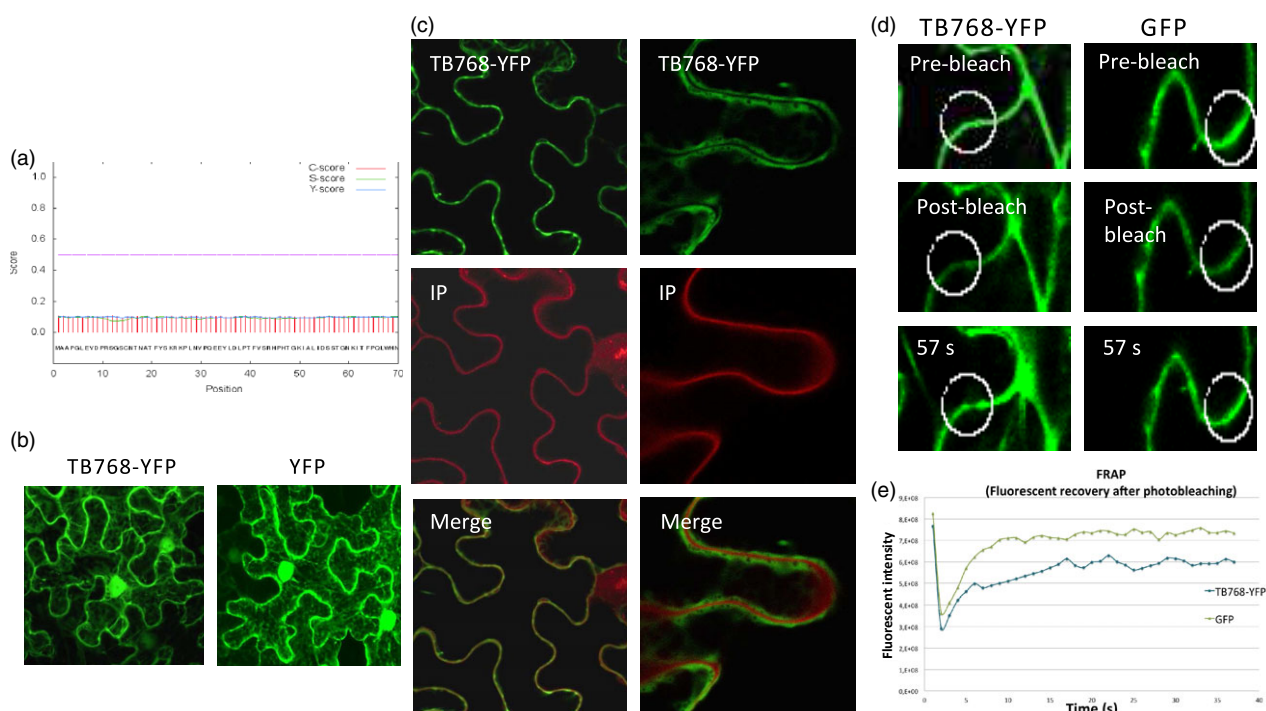


Figure 6 The CoA-ligase TBPCCL is a cytosolic enzyme. (a) Screenshot from signal peptide 4.0 showing no signal peptide for the TBPCCL enzyme. The flat graph indicates no cell compartment signalling for this protein, suggesting a predicted cytoplasmic localization. (b–e) Confocal microscopy analysis. (b) TB768-YFP and free GFP in *N. benthamiana* leaves. (c) TB768-YFP expression in *N. benthamiana* leaves stained with propidium iodide (PI). (d) FRAP analysis over 40 s; (e) Fluorescence graph showing the fast recovery of the TB768-YFP and free GFP after photobleaching.

the substrate. In *Populus tremuloides*, two isoforms have been determined, one mainly involved in lignin formation and the other in phenylpropanoid biosynthesis (Hu *et al.*, 2010). In most of the studied plants, the activity of the known 4CL isoforms is associated with both flavonoid and lignin biosynthesis, although they can use different substrates with variable efficiency. Lindermayr *et al.* (2002) determined the presence of four 4CL isoenzymes in *Glycine max*, each encoded by a single gene with different substrate specificity.

4CL-encoding genes have not been previously characterized in *Taxus* spp. The *TB768* gene identified in the present work encodes a protein that can form CoA esters with β -phenylalanine and also with 4-coumarate, and is referred to as TBPCCL, because this is the first acyl-CoA ligase characterized in *Taxus baccata*. The newly identified protein could control flavonoid/lignin formation, as well as activate β -phenylalanine for taxol biosynthesis. As other plant species (see above), also *Taxus* spp. likely possess more than one 4CL-encoding gene with variable expression pattern and substrate specificity. This is also supported by our transcriptome study, in which a second putative 4CL-encoding gene (*TB154*) was identified (Figure S3b) that displayed different expression kinetics compared to *TB768* and the taxane synthesis genes, and hence, because of a lack of coregulation, was not considered for further analysis.

Another aspect of taxol biosynthesis that needs further investigation is the subcellular localization of some enzymatic steps. The taxadiene ring is formed in the plastids, because the diterpene structure of this anticancer compound is assembled through the methylerythritol 4-phosphate (MEP) pathway. The hydroxylation steps that need cytochrome P450 enzymes usually occur in the microsomal fraction, but to date, the subcellular compartment in which β -phenylalanine is activated by HS-CoA has not been determined. To shed more light on this, TBPCCL was transiently expressed as a YFP fusion protein in *N. benthamiana*, which demonstrated that it is a cytosolic protein.

Like most 4CL enzymes, TBPCCL can act on several substrates, but most importantly, it binds to β -phenylalanine, transforming it into β -phenylalanyl-CoA, which constitutes the C13 side chain of the taxol ring system, essential for the anticancer activity of taxol (Kingston, 2000). After the side chain is attached to baccatin III, mediated by the BAPT enzyme, only two more steps, a hydroxylation and benzylation, are needed to obtain taxol. We have previously shown that the most limiting steps of taxol biosynthesis are at the end of the pathway, subsequent to the formation of 10-DATBIII (Cusido *et al.*, 2014). The identification of the *TBPCCL* gene, together with the application of metabolic engineering techniques, may therefore contribute to the establishment of highly productive taxol and related taxane cell cultures. Increased in-depth knowledge of the taxol biosynthetic pathway may ultimately facilitate at least its partial reconstitution through synthetic biology in heterologous systems.

Experimental procedures

Taxus cell cultures and elicitor experiment

Taxus baccata cell suspension cultures were established from stem-derived calli, as described previously (Cusido *et al.*, 2002). Cells were cultured following a two-stage system elicited with 100 μ M MeJA at the beginning of the second stage (Onrubia *et al.*, 2013b). Samples of both elicited and control cultures were harvested at the following time points: 0 h, 30 min, 1, 2, 4, 8 and 16 h, 1, 2 and 4 days after elicitation.

Transcriptome analysis by cDNA-AFLP

Total RNA from *T. baccata* cells was isolated with the RNeasy (Qiagen, Hilden, Germany) kit. cDNA-AFLP-based transcript profiling of *T. baccata* samples was performed as we described previously (Onrubia *et al.*, 2014). For qRT-PCR, cDNA was prepared from total RNA with SuperScript II reverse transcriptase (Invitrogen, Carlsbad, CA). qRT-PCR was performed using the SYBR Green PCR Master Mix (Roche, Sant Cugat del Valles, Spain) in a 384-well platform system (LightCycler[®] 480 Instrument; Roche). Gene-specific primers were designed with Primer3 software version 0.4.0 (Table S1). Expression levels were normalized to those of the 18S from *T. baccata* (Onrubia *et al.*, 2013b).

Cloning of *T. baccata* sequences

The candidate gene FL-cDNA sequences (Table 1) were obtained by RACE (GibcoBRL-Invitrogen, Carlsbad, CA). The *TB768* FL ORF (Figure S4) was PCR-amplified between Gateway *attB* sites and Gateway recombined into the pEarlyGate203 vector (Earley *et al.*, 2006). The resulting expression vector pEarlyGate203-TB768 fused the myc-tag sequence to the N-terminal of the TB768 protein and was used for transient overexpression in *N. benthamiana* plants. For subcellular localization, the *TB768* gene was cloned in pENTR/D-TOPO using the TOPO cloning kit (Invitrogen), and the sequence was verified and transferred to the pEarlyGate101 vector (Earley *et al.*, 2006) by Gateway recombination. In the resulting vector, the yellow fluorescent protein (YFP) was C-terminally fused to the TB768 protein.

In silico studies of the candidate genes

Candidate genes were blasted (PMID: 2231712, default parameters) against the PDB database to identify the best structural templates. They were then modelled in 3D using Phyre2 (Kelley and Sternberg, 2009). Resulting models with a confidence higher than 99% were displayed using PyMOL (PyMOL Molecular Graphics System, version 1.5.0.4 Schrödinger, LLC) and used for further identification of substrate-contacting residues. Candidate cDNA-AFLP sequences were also Blasted against the UNIPROT SwissProt section (PMID: 24253303) to identify the most closely related functionally characterized sequence, as well as the closest homologue featured in MACiE (24319146), the main reference database of enzyme proteins with annotated functional sites. Catalytic domains were also identified using the Conserved Domains Database (CDD) from the NCBI (<http://www.ncbi.nlm.nih.gov/Structure/cdd/cdd.shtml>) (Marchler-Bauer *et al.*, 2009, 2011). We used the SignalP v4.0 (Petersen *et al.*, 2011) to predict the existence of a signal peptide.

Expression of the *Tb768* gene in *N. benthamiana* plants

The expression vector pEarlyGate203-TB768 was electroporated in the *Agrobacterium tumefaciens* strain C58C1 carrying the helper pGv2260 (Earley *et al.*, 2006) to perform the agroinfiltration in *N. benthamiana* plants. The *Agrobacterium* strain carrying the pEarlyGate203-TB768 vector was co-infiltrated in a 1:1 ratio with an *Agrobacterium* strain containing the HC-Pro silencing suppressor (Goytia *et al.*, 2006) in leaves of 3- to 5-week-old *N. benthamiana* plants using the syringe method (Voinnet *et al.*, 2003); the plants were grown in long-day greenhouse conditions at 25 °C. Six days after agroinfiltration, the leaves were collected and a total protein extraction was performed using a modified buffer described by Dickerson *et al.* (1984) (0.1 M Tris-HCl pH7.8, 0.5%, Triton X-100, 15 mM β -mercaptoethanol and 15% of

glycerol). The crude extract was diluted with 0.1 M Tris buffer pH 7.5, and the soluble TB768 protein was purified following the protocol of the c-myc-tagged protein mild purification kit from MBL. The purified enzyme was concentrated and the final concentration was determined by a Bradford assay (Ernst and Zor, 2010). To determine the TB768 protein integrity, 1 µL of purified TB768 was analysed by Western blot, using an anti-Myc pAb (Genescript) (Figure S8).

CoA ligase *in vitro* enzyme activity assay

The CoA ligase reaction was carried out as described previously (Koetsier *et al.*, 2009) in a final volume of 400 µL, with 20 µg/mL of the concentrated TB768 protein, 10 mM β-phenylalanine or 1 mM 4-coumaric acid and 1.5 mM CoA as substrates, dissolved in 50 mM Tris-HCl, at pH 8 and pH 8.5, respectively. One control of each substrate without the enzyme was carried out in parallel. The reactions were incubated at 30 °C for 45 min, stopped by boiling for 10 min and lyophilized to dryness. The resulting samples were separately dissolved in aqueous 0.01 M HCl 50 µL (Muchiri and Walker, 2012) and the reaction products were analysed by HPLC-MS/MS as described in the Supporting Information experimental procedures. Enzymatic activity analyses were performed under the same conditions using a previously described spectrophotometric assay (Ehlting *et al.*, 1999). 4-coumaroyl-CoA and β-phenylalanyl-CoA formation was monitored at 336 and 315 nm, respectively, for 60 min using a UV/Vis spectrophotometer SpectraMax M3 Multi-Mode Microplate (Molecular Devices, Workingham Berkshire, UK). The product concentration was calculated with the extinction coefficient of each compound (β-phenylalanyl-CoA ε 315 nm: 3421 mM/cm; 4-coumaroyl-CoA ε 336 nm: 3885 mM/cm).

Subcellular location of the protein encoded by the TB768 gene

The construct pEarlyGate101-TB768 was transformed in *Agrobacterium tumefaciens* strain C58C1 carrying the helper pGv2260 (Deblaere *et al.*, 1985). The resulting strain was infiltrated in *N. benthamiana* plants as described above. After 3 days of normal growing conditions, the infiltrated leaves were harvested and analysed by confocal laser scanning microscopy (CLSM). Leaf segments were cut and abaxial leaf sides were scanned. Cell walls were stained by infiltration of the tissue with a saturated solution of propidium iodide (PI) 5 mg/mL. The images were acquired with an Olympus FV1000 confocal laser scanning microscope (Tokyo, Japan), using the 60× water-immersion NA: 1.20 objective. An argon laser at 514 nm was used to excite the YFP and PI. For visualization, the emission windows were set at 500–545 nm and 590–670, respectively. For fluorescence recovery after photobleaching (FRAP) analysis, a region of interest (ROI) of 10–12 µm was defined and photobleached with full laser power (100%) for 6 s. Recovery of the fluorescence was monitored by scanning the whole focused cell area at a low laser power (20%). The prebleaching photograph was taken with the same laser power.

Acknowledgements

Karla Ramírez is a predoctoral fellow of Mexican CONACYT. Miriam Onrubia is indebted to the European Union-Human Resources and Mobility for an Early Stage Training Grant (MEST-CT-2004-514632). This research has been supported by two grants from the Spanish MINECO (BIO2011-29856-C02-01 and

AGL2013-43522-R) and a grant from Generalitat de Catalunya (2014SGR1434). All authors gratefully acknowledge the support of Cost Action FA1006-Plant Engine.

Conflicts of interest

No conflicts of interest are to be declared.

References

- Allina, S.M., Pri-Hadash, A., Theilmann, D.A., Ellis, B.E. and Douglas, C.J. (1998) 4-Coumarate: coenzyme A ligase in hybrid poplar. *Plant Physiol.* **116**, 743–775.
- Bentebibel, S., Moyano, E., Palazon, J., Cusidó, R.M., Bonfill, M., Eibl, R. and Piñol, M.T. (2005) Effects of immobilization by entrapment in alginate and scale-up on paclitaxel and baccatin III production in cell suspension cultures of *Taxus baccata*. *Biotechnol. Bioeng.* **89**, 647–655.
- Breyne, P. and Zabeau, M. (2001) Genome-wide expression analysis of plant cell cycle modulated genes. *Curr. Opin. Plant Biol.* **4**, 136–142.
- Colling, J., Pollier, J., Makunga, N.P. and Goossens, A. (2013) cDNA-AFLP-based transcript profiling for genome-wide expression analysis of jasmonate-treated plants and plant cultures. *Methods Mol. Biol.* **1011**, 287–303.
- Croteau, R., Ketchum, R.E.B., Long, R.B., Kaspera, R. and Wildung, M.R. (2006) Taxol biosynthesis and molecular genetics. *Phytochem. Rev.* **5**, 75–97.
- Cusidó, R.M., Palazón, J., Bonfill, M., Navia-Osorio, A., Morales, C. and Piñol, M.T. (2002) Improved paclitaxel and baccatin III production in suspension cultures of *Taxus media*. *Biotechnol. Prog.* **18**, 418–423.
- Cusidó, R.M., Onrubia, M., Sabater-Jara, A.B., Moyano, E., Bonfill, M., Goossens, A., Pedreño, M.A. and Palazon, J. (2014) A rational approach to improving the biotechnological production of taxanes in plant cell cultures of *Taxus* spp. *Biotechnol. Adv.* **32**, 1157–1167.
- De Geyter, N., Gholami, A., Goormachtig, S. and Goossens, A. (2012) Transcriptional machineries in jasmonate-elicited plant secondary metabolism. *Trends Plant Sci.* **17**, 349–359.
- Deblaere, R., Byteler, B., DeGreve, H., Deboeck, F., Schell, J., Van Montagu, M. and Leemans, J. (1985) Efficient octopine Ti plasmid-derived vectors for *Agrobacterium*-mediated gene transfer to plants. *Nucleic Acids Res.* **13**, 4777–4788.
- Dickerson, D.P., Pascholati, S.F., Hagerman, A.E., Butler, L.G. and Nicholson, R.L. (1984) Phenylalanine ammonia-lyase and hydroxycinnamate:CoA ligase in maize mesocotyls inoculated with *Helminthosporium maydis* or *Helminthosporium carbonum*. *Physiol. Plant Pathol.* **25**, 111–123.
- Docimo, T., Consonni, R., Coraggio, I. and Mattana, M. (2013) Early phenylpropanoid biosynthetic steps in *Cannabis sativa*: link between genes and metabolites. *Int. J. Mol. Sci.* **14**, 13626–13644.
- Earley, K.W., Haag, J.R., Pontes, O., Opper, K., Juehne, T., Song, K. and Pikaard, C.S. (2006) Gateway-compatible vectors for plant functional genomics and proteomics. *Plant J.* **45**, 616–629.
- Ehlting, J., Büttner, D., Wang, Q., Douglas, C.J., Somssich, I.E. and Kombrink, E. (1999) Three 4-coumarate: coenzyme A ligases in *Arabidopsis thaliana* represent two evolutionarily divergent classes in angiosperms. *Plant J.* **19**, 9–20.
- Ernst, O. and Zor, T. (2010) Linearization of the Bradford protein assay. *J. Vis. Exp.* **38**, 1918.
- Expósito, O., Bonfill, M., Moyano, E., Onrubia, M., Mirjalili, H., Cusidó, R.M. and Palazon, J. (2009) Biotechnological production of taxol and related taxoids: current state and prospects. *Anticancer Agents Med. Chem.* **9**, 109–121.
- Floss, H.G. and Mocek, U. (1995) Biosynthesis of taxol. In *Taxol: Science and Applications* (Suffness, M., ed.), pp. 191–208. Boca Raton: CRC Press.
- Frense, D. (2007) Taxanes: perspectives for biotechnological production. *Appl. Microbiol. Biotechnol.* **73**, 1233–1240.
- Goossens, A. (2015) It is easy to get huge candidate gene lists for plant metabolism now, but how to get beyond? *Mol. Plant.* **8**, 2–5.
- Goossens, A., Hakkinen, S.T., Laakso, I., Seppanen-Laakso, T., Biondi, S., De Sutter, V., Lammertyn, F., Nuutila, A.M., Söderlund, H., Zabeau, M., Inze, D. and Oksman-Caldentey, K.M. (2003) A functional genomics approach toward the understanding of secondary metabolism in plant cells. *Proc. Natl Acad. Sci. USA*, **100**, 8595–8600.

- Goytia, E., Fernández-Clavino, L., Martínez-García, B., López-Abella, D. and López-Moya, J.J. (2006) Production of plum pox virus HC-Pro functionally active for aphid transmission in a transient expression system. *J. Gen. Virol.* **87**, 3413–3423.
- Guo, B.H., Kai, G.Y., Jin, H.B. and Tang, K.X. (2006) Taxol synthesis. *A. J. Biotechnol.* **5**, 15–20.
- Hamberger, B. and Hahlbrock, K. (2004) The 4-coumarate:CoA ligase gene family in *Arabidopsis thaliana* comprises one rare, sinapate-activating and three commonly occurring isoenzymes. *Proc. Natl Acad. Sci. USA*, **101**, 2209–2214.
- Held, M.A., Boulaflous, A. and Brandizzi, F. (2008) Advances in fluorescent protein-based imaging for the analysis of plant endomembranes. *Plant Physiol.* **147**, 1469–1481.
- Hezari, M., Ketchum, R.E.B., Gibson, D.M. and Croteau, R. (1997) Taxol production and taxadiene synthase activity in *Taxus canadensis* cell suspension cultures. *Arch. Biochem. Biophys.* **337**, 185–190.
- Hu, Y., Gai, Y., Yin, L., Wang, X., Feng, C., Feng, L., Li, D., Jiang, X. and N. Wang, D.C. (2010) Crystal structures of *Populus tomentosa* 4-Coumarate:CoA ligase shed light on its enzymatic mechanisms. *Plant Cell*, **22**, 3093–3140.
- Jalalpour, Z., Shabani, L., Afghani, L., Sharifi-Tehrani, M. and Amini, S.A. (2014) Stimulatory effect of methyl jasmonate and squalenol on phenolic metabolism through induction of LOX activity in cell suspension culture of yew. *Turk. J. Biol.* **38**, 76–82.
- Keeling, C.I. and Bohlmann, J. (2006) Diterpene resin acids in conifers. *Phytochemistry*, **67**, 2415–2423.
- Kelley, L.A. and Sternberg, M.J.E. (2009) Protein structure prediction on the web: a case study using the Phyre server. *Nat. Protoc.* **4**, 363–371.
- Kingston, D.G.I. (2000) Recent advances in the chemistry of taxol. *J. Nat. Prod.* **63**, 726–734.
- Koepp, A.E., Hezari, M., Zajicek, J., Stofer, V.B., LaFever, R.E., Lewis, N.G. and Croteau, R.B. (1995) Cyclization of geranylgeranyl diphosphate to taxadiene is the committed step of Taxol biosynthesis in Pacific yew. *J. Biol. Chem.* **270**, 8686–8690.
- Koetsier, M.J., Jekel, P.A., Vandenberg, M.A., Bovenberg, R.A. and Jansen, D.B. (2009) Characterization of a phenylacetate–CoA ligase from *Penicillium chrysogenum*. *Biochem. J.* **417**, 467–476.
- Lange, B.M. and Ahkami, A. (2013) Metabolic engineering of plant monoterpenes, sesquiterpenes and diterpenes – current status and future opportunities. *Plant Biotechnol. J.* **11**, 169–196.
- Li, S.T., Zhang, P., Zhang, M., Fu, Ch.H., Zhao, Ch.F., Dong, Y.S., Guo, A.Y. and Yu, L.J. (2012) Transcriptional profile of *Taxus chinensis* cells in response to methyl jasmonate. *BMC Genom.* **13**, 295.
- Lindermayr, Ch., Möllers, B., Fliegmann, J., Uhlmann, A., Lottspeich, F., Meimberg, H. and Ebel, J. (2002) Divergent members of a soybean (*Glycine max* L.) 4-coumarate:coenzyme A ligase gene family. Primary structures, catalytic properties and differential expression. *Eur. J. Biochem.* **269**, 1304–1315.
- Malik, S., Cusido, R.M., Mirjalili, M.H., Moyano, E., Palazon, J. and Bonfill, M. (2011) Production of the anticancer drug taxol in *Taxus baccata* suspension cultures: a review. *Proc. Biochem.* **46**, 23–34.
- Marchler-Bauer, A., Anderson, J.B., Chitsaz, F., Derbyshire, M.K., DeWeese-Scott, C., Fong, J.H., Geer, L.Y., Geer, R.C., Gonzales, N.R., Gwadz, M., He, S., Hurwitz, D.I., Jackson, J.D., Ke, Z., Lanczycki, C.J., Liebert, C.A., Liu, C., Lu, F., Lu, S., Marchler, G.H., Mullokandov, M., Song, J.S., Tasneem, A., Thanki, N., Yamashita, R.A., Zhang, D., Zhang, N. and Bryant, S.H. (2009) CDD: specific functional annotation with the conserved domain database. *Nucleic Acids Res.* **37**, D205–D210.
- Marchler-Bauer, A., Lu, S., Anderson, J.B., Chitsaz, F., Derbyshire, M.K., DeWeese-Scott, C., Fong, J.H., Geer, L.Y., Geer, R.C., Gonzales, N.R., Gwadz, M., Hurwitz, D.I., Jackson, J.D., Ke, Z., Lanczycki, C.J., Lu, F., Marchle, R.G.H., Mullokandov, M., Omelchenko, M.V., Robertson, C.L., Song, J.S., Thanki, N., Yamashita, R.A., Zhang, D., Zhang, N., Zheng, C. and Bryant, S.H. (2011) CDD: a conserved domain database for the functional annotation of proteins. *Nucleic Acids Res.* **39**, D225–D229.
- Muchiri, R. and Walker, K.D. (2012) Taxol biosynthesis: tyrocidine synthetase A catalyzes the production of phenylisoserinyl CoA and other amino phenylpropanoyl thioesters. *Chem. Biol.* **19**, 679–685.
- Nims, E., Dubois, C.P., Roberts, S.C. and Walker, E.L. (2006) Expression profiling of genes involved in paclitaxel biosynthesis for targeted metabolic engineering. *Metab. Eng.* **8**, 385–394.
- Onrubia, M., Moyano, E., Bonfill, M., Expósito, O., Palazón, J. and Cusido, R.M. (2010) An approach to the molecular mechanism of methyl jasmonate and vanadyl sulphate elicitation in *Taxus baccata* cell cultures: the role of *txs* and *bapt* gene expression. *Biochem. Eng. J.* **53**, 104–111.
- Onrubia, M., Cusido, R.M., Ramirez, K., Hernández-Vázquez, L., Moyano, E., Bonfill, M. and Palazon, J. (2013a) Bioprocessing of plant *in vitro* systems for the mass production of pharmaceutically important metabolites: Paclitaxel and its derivatives. *Curr. Med. Chem.* **20**, 880–891.
- Onrubia, M., Moyano, E., Bonfill, M., Cusido, R.M., Goossens, A. and Palazón, J. (2013b) Coronatine, a more powerful elicitor for inducing taxane biosynthesis in *Taxus media* cell cultures than methyl jasmonate. *J. Plant Physiol.* **170**, 211–219.
- Onrubia, M., Polle, R.J., Vanden Bossche, R., Goethals, M., Gevaert, K., Moyano, E., Vidal-Limon, H., Cusido, R.M., Palazón, J. and Goossens, A. (2014) Taximin, a conserved plant-specific peptide modulates secondary metabolism in *Taxus baccata* and *Nicotiana tabacum*. *Plant Biotechnol. J.* **12**, 971–983.
- Palazon, J., Cusido, R.M., Bonfill, M., Morales, C. and Piñol, M.T. (2003) Inhibition of paclitaxel and baccatin III accumulation by mevinolin and fosmidomycin in suspension cultures of *Taxus baccata*. *J. Biotechnol.* **101**, 157–163.
- Pauwels, L., Inzé, D. and Goossens, A. (2009) Jasmonate-inducible gene: what does it mean? *Trends Plant Sci.* **14**, 87–91.
- Petersen, T.N., Brunak, S., von Heijne, G. and Nielsen, H. (2011) SignalP 4.0: discriminating signal peptides from transmembrane regions. *Nat. Methods*, **8**, 785–786.
- Sabater-Jara, A.B., Onrubia, M., Moyano, E., Bonfill, M., Palazón, J., Pedreño, M.A. and Cusido, R.M. (2014) Synergistic effect of cyclodextrins and methyl jasmonate on taxane production in *Taxus x media* cell cultures. *Plant Biotechnol. J.* **12**, 1075–1084.
- Voinnet, O., Rivas, S., Mestre, P. and Baulcombe, D. (2003) An enhanced transient expression system in plants based on suppression of gene silencing by the p19 protein of tomato bushy stunt virus. *Plant J.* **33**, 949–956.
- Vongpaseuth, K. and Roberts, S.C. (2007) Advancements in the understanding of Paclitaxel metabolism in tissue culture. *Curr. Pharm. Biotechnol.* **8**, 219–236.
- Walker, K. and Croteau, R. (2001) Taxol biosynthetic genes. *Phytochemistry*, **58**, 1–7.
- Wasternack, C. (2007) Jasmonates: an update on biosynthesis, signal transduction and action in plant stress response, growth and development. *Ann. Bot.* **100**, 681–697.
- Wu, Y., Masison, D.C., Eisenberg, E. and Green, L.E. (2006) Application of photobleaching for measuring diffusion of prion proteins in cytosol of yeast cells. *Methods*, **39**, 43–49.

Supporting information

Additional Supporting information may be found in the online version of this article:

Figure S1 10-deacetylbaccatin III (DABIII), baccatin III, 10-deacetyl taxol (DAT), taxol and cephalomannine (Ceph) content (cell associated + extracellular) in *T. baccata* cell cultures grown for 24 days in the production medium in the absence (light gray) or presence (dark gray) of methyl jasmonate (MeJA, 100 μ M).

Figure S2 qRT-PCR analysis of the transcript levels of the GGPS, TXS, TAT, T13OH, T7OH, T2OH, TBT, DBAT, PAM, BAPT and DBTNBT genes relative to time zero of the cultures studied.

Figure S3 Average linkage hierarchical clustering of differentially expressed gene tags identified by cDNA-AFLP and corresponding to enzymes involved in metabolic pathways.

Figure S4 Serial Cloner vs. 2.6 display.

Figure S5 qRT-PCR analysis of the *TB768* gene in *Taxus baccata* cells cultured in the two-media system.

Figure S6 Immunoblot image obtained by chemiluminescence of purified and concentrated TB768 protein.

Figure S7 Extract ion chromatogram of 4-coumaroyl-CoA (m/z 914) and β -phenylalanyl-CoA (m/z 916) standards showing the retention time of each molecule.

Figure S8 Fragmentation pattern of the β -phenylalanyl-CoA and 4-coumaroyl-CoA.

Table S1 Sequences of the primers used to amplify the genes by qRT-PCR.

Data S1 Experimental procedures.