Full Length Research Paper

Molecular cloning and functional analysis of the gene encoding geranylgeranyl diphosphate synthase from *Jatropha curcas*

Juan Lin*, YuanJie Jin, Xin Zhou and JingYa Wang

State Key Laboratory of Genetic Engineering, Institute of Plant Biology, School of Life Sciences, Fudan University, Shanghai, People's Republic of China.

Geranylgeranyl diphosphate (GGPP) synthase (GGPPS, EC: 2.5.1.29) catalyzes the condensation of isopentenyl diphosphate (IPP) with allylic diphosphates to yield (all-E)-GGPP. GGPP is one of the key precursors in the biosynthesis of biologically significant isoprenoid compounds. Here we report for the first time the cloning of a full-length cDNA encoding GGPPS (Jc-GGPPS) from *Jatropha curcas* L. The full-length cDNA was 1414 base pair (bp), with an 1110-bp open reading frame (ORF) encoding a 370-amino-acids polypeptide. Bioinformatic analysis revealed that Jc-GGPPS is a member of the polyprenyltransferases with two highly conserved aspartate-rich motifs, and showed high homology to other plant GGPPSs. Phylogenetic tree analysis indicated that plant GGPPSs could be classified into two groups, angiosperm and gymnosperm, while Jc-GGPPS had closer relationship with angiosperm plant GGPPSs. Functional analysis of Jc-GGPPS in GGPPS-deficient mutant plasmid pAC-BETA ($\geq$ crtE) demonstrated that Jc-GGPPS mediated the biosynthesis of carotenoid and provided the general precursor for diterpenes biosynthesis.

Key words: *Jatropha curcas* L., geranylgeranyl diphosphate synthase, rapid amplification of cDNA ends, genetic complementation.

INTRODUCTION

Geranylgeranyl diphosphate (GGPP) synthase (GGPPS, EC: 2.5.1.29), is a homodimeric, short chain trans-prenyltransferase responsible for the synthesis of GGPP and occurs nearly ubiquitously in plants, animals and bacteria (Ogura and Koyama, 1998). It is a branch point enzyme, which regulates coordination with the other prenyltransferases (GDP and FDP synthase respectively) of the precursor flux towards mono-, sesqui-, and diterpenoids production (Hefner et al., 1998; Laskaris et al., 2000). Consequently, this may affect the ratio of the produced mono-/ sesqui-/ di-terpenes. Thus, in view of the key role of this enzyme in diterpenes biosynthesis, the isolation and characterization of its cDNAs deserves further attention. The activity of GGPPS was first detected in plants, such as pumpkin (*Cucurbita moschata*) (Ogura et al., 1972), then in animal species, such as pig (*Miniature*

swine) (Sagami et al., 1981). Until now GGPPS cDNAs have been cloned and characterized from plant species, such as *Lycopericon esculentum* (Ament et al., 2006), *Antirrhinum majus* (Tholl et al., 2004), *Adonis aestivalis* (Cunningham and Gantt, 2007), *Hevea brasiliensis* (Takaya et al., 2003), *Chrysanthemum morifolium* (Kishimoto and Ohmiya, 2006), *Catharanthus roseus* (Bantignies et al., 1996), *Sinapis alba* (Kloer et al., 2006), *Picea abies* (Schmidt and Gershenzon, 2007), *Coleus forskohlii* (Engprasert et al., 2004), *Scoparia dulcis* and *Croton sublyratus* (Sitthithaworn et al., 2001), *Capsicum annuum* (Kuntz et al., 1992), *Lupinus albus* (Aitken et al., 1995). In spite of the extensive studies of the GGPPS genes in many plant species, only a few of the isolated cDNAs have been functionally characterized. Previous studies indicated that the GGPPS are an important branch point enzyme in terpenoid biosynthesis, and that the GGPPS gene is involved in the biosynthesis of diterpenes as well as in protein prenylation. Heterologous expression of GGPPS gene in *Escherichia coli* or

---

*Corresponding author. E-mail: linjuan@fudan.edu.cn.

Saccharomyces cerevisiae showed it is related to universal diterpenes biosynthesis, such as casbene (Dudley et al., 1986), oryzalexins (West et al., 1990), paclitaxel (Hefner et al., 1998) and rubber (Takaya et al., 2003). The expression of GGPPS gene from *P. abies* in transcription lever showed that this gene had an important role in the induced formation of oleoresin diterpene (Schmidt and Gershenzon, 2007). These results indicate that a large increase in GGPPS activity corresponds to diterpenes accumulation. Structural studies from many organisms, and comparison of the amino acid sequences, revealed there are two aspartate rich regions [DDxx(xx)D](D is aspartate, x any amino acid) in these GGPPSs, the FARM (First Aspartate Rich Motif) and the SARM (Second Aspartate Rich Motif) located in the enzymes' large, central, hydrophobic cavity responsible for substrate binding and catalysis. The two aspartate rich regions act as binding sites for allylic substrates and IPP (isopentenyl diphosphate). Substitution of any of the aspartate molecules in those motifs with different amino acids resulted in reduced enzymatic activity (Liang et al., 2002).

Jatropha curcas (Euphorbiaceae) occurs naturally in Central America, but had been cultivated in many tropical and subtropical areas of America, Africa and Asia (Lin et al., 2004). Because *J. curcas* is not eaten by most animals, including goats and sheep, this shrub has often been planted as a natural fence around gardens and fields. The plant produces large seeds which contain approximately 50% oil, which is being used as biodiesel in Central America (e.g. Nicaragua), Africa (e.g. West Africa and Zimbabwe) and Asia (e.g. China) (Horst et al., 2007). The aerial parts of the plant, especially the seed, contain phorbol esters which make the plant unpalatable and toxic to most vertebrates, insects and snails (Lin et al., 2004). However, the presence of high levels of phorbol esters prevents its use in animal feeding (Makkar et al., 1997). Various chemical and physical treatments have been evaluated to extract or inactivate phorbol esters so that seed meals rich in proteins can be used as feed resources. However, not much progress has been reported so far. Therefore, it is necessary to develop new methods for inactivating or decreasing phorbol esters, or bringing them to below threshold toxicity limits. Phorbol esters are the tetracyclic diterpenes generally known for their tumor promoting activity. They occur naturally in many plants of the family Euphorbiacaeae and Thymelaeaceae. GGPPS is suggested to be a key enzyme in the biosynthesis of phorbol esters. Therefore, we anticipate that the GGPPS gene will be a good model to study the regulatory mechanisms controlling phorbol esters accumulation at the transcriptional level. A prerequisite for such a study is to clone a functional GGPPS gene. It is important to explore in detail phorbol esters biosynthesis in genetics and enzymology for bioengineering of phorbol esters production through genetic manipulation of the isoprenoid biosynthetic path-

way. We are interested in studying *J. curcas* GGPPS with the expectation that the enzyme participates in the formation of a possible initiator molecule in phorbol esters biosynthesis.

Here we report the cloning and characterization of a cDNA from *J. curcas*, designated as *Jc-GGPPS*, which codes for a putative GGPPS likely to be involved in diterpenes biosynthesis. Sequence analysis showed that Jc-GGPPS shares sequence identity with GGPPS gene from other plants. Molecular analysis of Jc-GGPPS provides insight into the regulation of phorbol esters biosynthesis, and ultimately may help to decrease the toxin of *J. curcas* through the use of biotechnology.

MATERIALS AND METHODS

Plant materials

The seeds of *J. curcas* were collected from Panzhihua, Sichuan province, China. RNA was isolated using cetyl trimethyl ammonium bromide (CTAB)-acidic phenolic method (Jiang and Zhang, 2003). After isolation, total RNA was stored at -70°C before use in cDNA cloning.

Cloning of Jc-GGPPS full-length cDNA by rapid amplification of cDNA ends (RACE)

Total cDNA was reverse-transcribed from total RNA by superscript reverse-transcriptase (Invitrogen) and used as a template for the polymerase chain reaction (PCR) amplification of the conserved fragment of GGPPS gene from *J. curcas*. Two degenerate primers, ggppF (5'-GC HRTBGARATGATYCAYACHATGTC-3') (H = A/C/T, R = A/G, B = C/G/T, Y = C/T) and ggppR (5'-TTRCCMCGVCC TTCRATCCARTT-3')(R = A/G, M = A/C, V = A/C/G), were designed based on two highly conserved amino acid blocks (ALDMIHTMS and GKTAGKDL) of other GGPPSs and used for homology based isolation of the conserved fragment of Jc-GGPPS gene by standard gradient PCR amplification with an annealing temperature of 56°C. A 510 bp fragment was amplified and basic local alignment search tool (BLAST) analysis showed that this fragment was highly homologous to GGPPS from other plant species (data not shown). This fragment was, subsequently, used for designing gene-specific primers for the cloning of full-length cDNA of Jc-GGPPS by RACE.

For 3'-RACE of Jc-GGPPS, two gene-specific primers JcggppF1 (as 3'RACE first amplification primer) and JcggppF2 (as 3'RACE nest amplification primer), AP (as the RT primer) and AUAP (as the universal amplification primer) were used. The 3'-RACE was conducted in accordance with the protocol provided by the manufacturer (Clontech Laboratories Inc., USA). The amplified PCR product was purified and cloned into pMD 18-T vector (TaKaRa, China) and then sequenced.

The 5' RACE, including RT, dCtailing and PCR amplification, was conducted in accordance with the protocol provided by the manufacturer (Clontech Laboratories Inc., USA). Two gene-specific primers JcggppR1 (as 5'RACE first amplification primer) and JcggppR2 (as 5'RACE nest amplification primer), 5'-RACE CDS Primer (5'-(T)$_{25}$N-1N-3') and SMART II A oligonucleotide (as cDNA synthesize), UMP (as the first amplification primer), NUP (as the nest amplification primer) were used for 5'RACE. The amplified PCR product was purified and cloned into pMD 18-T vector (TaKaRa, China) and then sequenced. By aligning and assembling the products of the 3' RACE, 5' RACE and the conserved fragment,

Table 1. Primers for the cloning of *Jc-GGPP* gene.

Primer	Sequence
ggppF	5'-GC(A/T/C)(A/G)T(G/T/C)GA(A/G)ATGAT(T/C)CA(T/C)AC(A/T/C)ATGTC-3'
ggppR	5'-TT(G/A)CC(C/A)CG(G/C/A)CCTTC(G/A)ATCCA(G/A)TT-3'
JcggppF1	5'-TCGGCTGTTGGGTCAGAAGGACTTGTTG-3'
JcggppF2	5'-CAAGGTGTATAGGATTATTATTTCAAG-3'
JcggppR1	5'-CACTTGAAATAATAATCCTATACACCTT-3'
JcggppR2	5'-AAGTCCTTCTGACCCAACAGCCGATCCA-3'
JcGGPPfull-F	5'- ATGGCTTTCTCTGCAACAGCCCCTGC-3'
JcGGPPfull-R	5'-TTAATTTTGCCTACTAGCAATGTAATTAGCA-3'
5'-RACE CDS Primer	5'-(T)$_{25}$N-1N-3'
UPM	Long: 5'-CTAATACGACTCACTATAGGGCAAGCAGTGGTATCAACGCAGAGT-3'
	Short: 5'-CTAATACGACTCACTATAGGGC-3'
NUP	5'-AAGCAGTGGTATCAACGCAGAGT-3'
SMART II A Oligo	5'-AAGCAGTGGTATCAACGCAGAGTA-3'
AP	5′- GGCCACGCGTCGACTAGTACTTTTTTTTTTTTTTTTTT-3′
AUAP	5′- GGCCACGCGTCGACTAGTAC-3′

the full-length GGPPS sequence from *J. curcas* (Jc-GGPPS) was deduced and, subsequently amplified by proof-reading PCR amplification with primers JcGGPPfull-F and JcGGPPfull-R. The PCR procedure was conducted under the following condition: 3 min at 94°C, 30 cycles (30 s at 94°C, 30 s at 50°C, 90 s at 72°C and 10 min at 72°C). The amplified PCR produced was purified and cloned into pMD 18-T vector (TaKaRa, China) and then sequenced. All the primers used in PCR amplification were listed in Table 1 and synthesized by Shanghai Sangon Biotechnological Company, China. In total, three independent positive clones were selected and sent for sequencing to confirm the correct sequence of the gene, avoiding PCR errors.

Bioinformation analysis

Bioinformation analysis of Jc-GGPPS was performed online at www.ncbi.nlm.nih.gov and http://us.expasy.org/prosite. VNTI Suite 8.0 was used for multiple alignment analysis of the full-length plant GGPPS amino acid sequence. Phylogenetic tree analysis of GGPPS from other species was aligned using Clustal W (http://align. genome.jp/) program. The neighbor- joining method was used to construct the phylogenetic tree.

Functional analysis of Jc-GGPPS in *E. coli*

pAC-BETA and pTrc-ATIPI plasmids, provided by Dr. Francis X. Cunningham, Department of Cell Biology and Molecular Genetics, University of Maryland, USA, were used to investigate the biological function of Jc-GGPPS. The plasmid, pAC-BETA, contains functional genes for geranylgeranyl pyrophosphate synthase (*crt*E), phytoene synthase (*crt*B), phytoene desaturase (*crt*L) and lycopene cyclase (*crt*Y). It also retains a chloramphenicol resistance gene. Cells of *E. coli* containing this plasmid produce and accumulate β-carotene, resulting in yellow colonies. The plasmid, pTrc-ATIPI, retains an ampicillin resistance gene and an IPI gene, the product of which can accelerate the accumulation of β-carotene (Cunningham and Gantt, 2000). After digesting the pMD 18-T vector containing Jc-GGPPS with *Bgl*II and *Not*I, the coding region of Jc-GGPPS was cloned into pTrc-ATIPI (The plasmid was digesting with *Bgl*II and *Not*I and releasing the ATIPI fragment from plasmid pTrc-ATIPI) and

the recombined plasmid (pTrc+Jc-GGPPS) was used to transform *E. coli* DH5α. The monoclonies containing pTrc+Jc-GGPPS could be screened on solidified Luria-Bertani (LB) medium containing 100 mg l^{-1} ampicillin. The plasmid, pAC-BETA (≥ *crt*E) (The plasmid pAC-BETA was digested with *Hind*III and *Sac*I and the released GGPPS (*crt*E) gene fragment from plasmid pAC-BETA) was used to transform *E. coli* DH5α. The monoclonies containing pAC-BETA (≥ *crt*E) could be screened on solidified LB medium containing 34 mg l^{-1} chloramphenicol. The pTrc+Jc-GGPPS and pAC-BETA (≥ *crt*E) were, subsequently, extracted from host *E. coli*, respectively, and used for co-transformation into the Top10F' strains. The Top10F' strains with pAC-BETA (≥ *crt*E)+pTrc or pAC-BETA(*crt*E), were used as control. The pTrc+Jc-GGPPS and pAC-BETA (≥ *crt*E) co-transformants were selected on LB medium containing 100 mg l^{-1} ampicillin and 34 mg l^{-1} chloramphenicol at 37°C for 48 h. The pAC-BETA (≥ *crt*E)+pTrc co-transformation were selected on LB medium containing 100 mg l^{-1} ampicillin and 34 mg l^{-1} chloramphenicol at 37°C for 48 h. The pAC-BETA (≥ *crt*E) transformants were selected on LB medium containing 100 mg l^{-1} ampicillin at 37°C for 48 h. The color of the transformants can be used as a visible marker to test if Jc-GGPPS can accelerate the accumulation of β-carotene.

RESULTS

Molecular cloning of the full-length cDNA of Jc-GGPPS

Total RNA isolated from the seeds of *J. curcas*, and the degenerate primers ggppF (sense) and ggppR (anti-sense) were used to specifically amplify a 510-bp product by homology-based reverse transcriptase (RT)-PCR (Figure 1A). A BLAST search showed that the PCR product was homologous to GGPPS genes from plant species (data not shown). cDNA ends of 726 and 523 bp were amplified by 5' RACE (Figure 1B) and 3' RACE (Figure 1C), respectively, based on the 510 bp conserved fragment. The core fragment, 3' and 5' end, was assembled
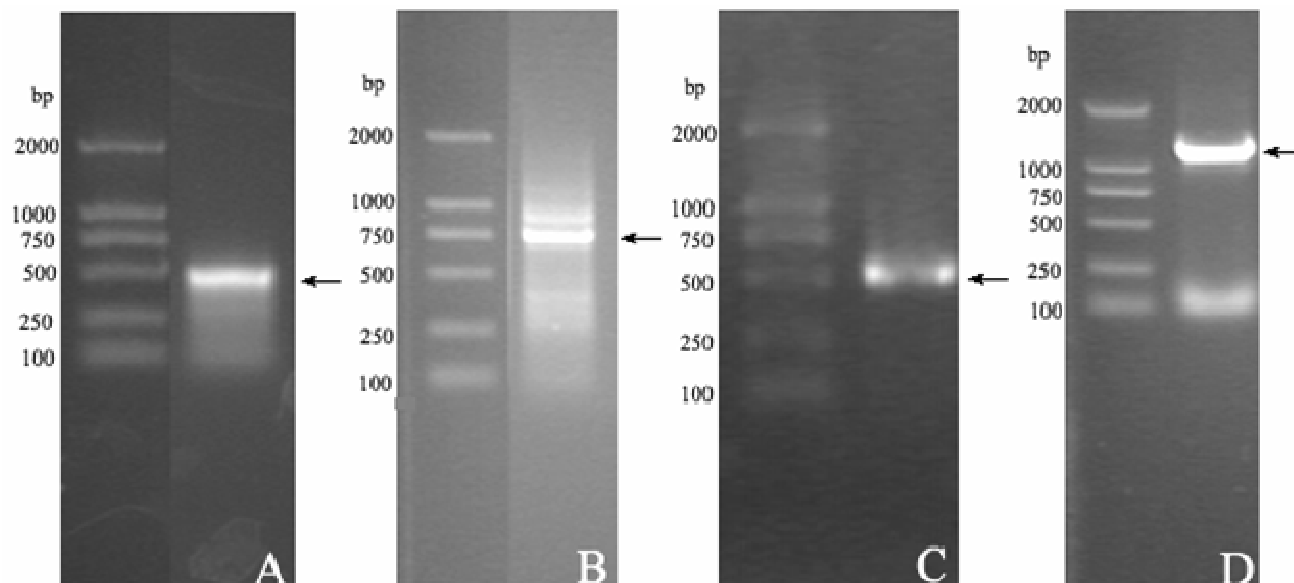
Figure 1. Agarose gel electrophoresis for the cloning of the Jc-GGPPS by RACE-PCR. A, B, C and D indicated the 510, 523 and 726 bp DNA bands of core fragment, 3′ end, 5′ end and the full-length cDNA amplifications, respectively.

with Vector NTI Suite 8.0 and the deduced full-length Jc-GGPPS cDNA sequence obtained was, subsequently, confirmed by sequencing. The full-length cDNA of Jc-GGPPS was 1414 bp with a 5' and 3' untranslated regions, poly A tail, and contained an 1113 bp ORF encoding a 370 amino-acid protein (Figures 1D and 2) (GenBank Accession Number: GU585938).

Bioinformatic analysis of Jc-GGPPS

Jc-GGPPS had a calculated molecular weight of 40.4 kDa and an isoelectric point value of 6.03. Protein-protein BLAST analysis showed that the deduced Jc-GGPPS amino acid sequence had much wider and higher homologies to the existing GGPPS as indicated by nucleotide-nucleotide BLAST. The amino acid sequence of Jc-GGPPS showed 45 - 65% identity and 60 - 80% positives in local alignments to GGPPS from *Solanum lycopersicum* to *Synechococcus* sp. PCC. ChloroP1.1 Prediction (http://www.cbs.dtu.dk/services/ChloroP/) showed Jc-GGPPS had a 51 amino acid localization signal in their N-terminal regions that target them into specific subcellular compartments. The N-terminal region of Jc-GGPPS was predicted to be localized in chloroplasts. Two-dimensional structural prediction of Jc-GGPPS was performed using the SOPMA server (http://bip. weizmann.ac.il/ bio_tools/fag.html). Based on the Hierarchical Neural Network analysis, Jc-GGPPS protein was composed of 56.49% α-helix, 6.76% extended strand, 4.59% β-turn and 32.16% random coil.
Molecular homologous modeling of Jc-GGPPS using the Swiss-Model server (http:// cn.expasy.org/swissmod/

SWISS-MoDEL.html) yielded a spatial architecture of Jc-GGPPS (Figure 3) that was very similar to *E. coli* GGPPS. According to the obtained result, the fold consisted of a large central cavity formed by a bundle of 15 α-helices, and was the putative catalytic site. This finding is consistent with results obtained with *Taxus walliciana* GGPPS (Lan and Aun, 2006). The two aspartate-rich DDXX(XX)D sequences that are highly conserved in isoprenyl diphosphate synthase were located on opposite walls of this cavity, facing each other. All five of the identified conserved regions are clustered around this cavity.
The deduced Jc-GGPPS amino acid sequence showed homology with other plant GGPPS, including those from *S. lycopersicum* (with 79% positive and 64% identity), *Daucus carota* (with 77% positive and 64% identity), *C. sublyratus* (with 75% positive and 61% identity), *Ginkgo biloba* (with 72% positive and 61% identity), *H. brasiliensis* (with 74% positive and 60% identity), *A. majus* (with 74% positive and 60% identity), *A. aestivalis* (with 67% positive and 59% identity) and others (Figure 4). These results strongly suggest that Jc-GGPPS should be a functional protein catalyzing the production of mevalonate. Jc-GGPPS was the first GGPPS gene cloned from *J. curcas*, therefore, it was of interest to investigate its evolutionary position among the phylogenetic tree of various GGPPS. A phylogenetic tree of GGPPS was constructed from different organisms, including plants (seventeen angiosperms and six gymnosperms), four animals, one fungus and four archaebacteria using the Clustal W program. The phylogenetic tree showed that GGPPSs could be classified into plants groups, as well as animals, fungus, and archaebacteria GGPPS
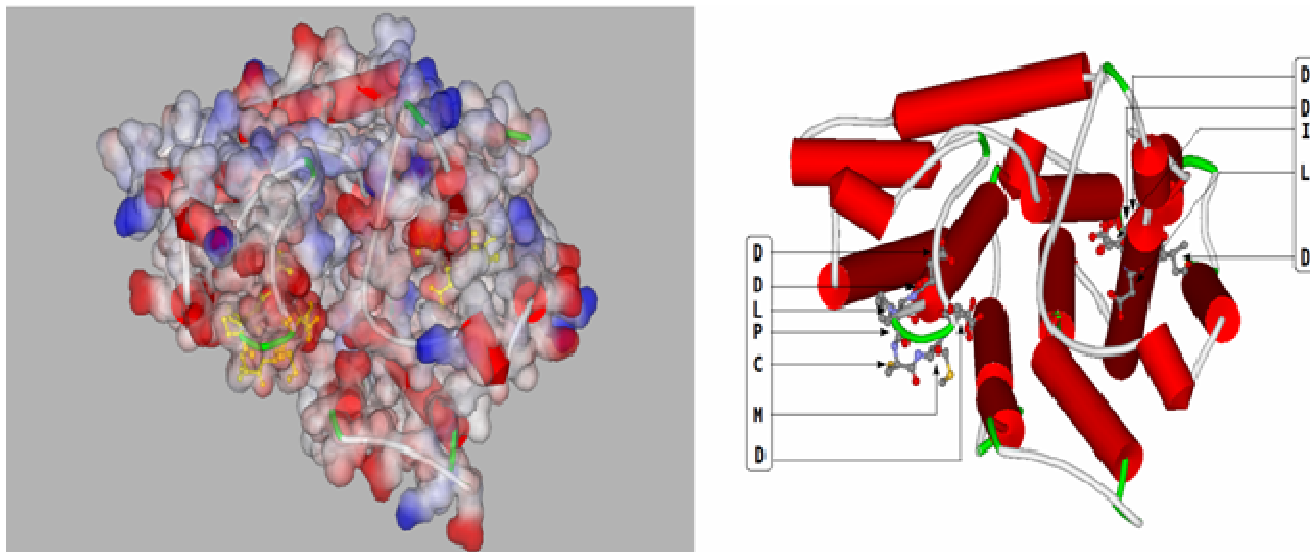
Figure 3. The 3-D structure of Jc-GGPPS constructed by the homology-based modeling A: the 3-D structure of Jc-GGPPS with surface filling; B: the peptide skeleton of Jc-GGPPS shoen in 3-D model, α-helix was shown in column shape, random coil in rope shape and sheet in arrowed-plate shape. The Asp-rich motifs (DDlpcmD and DDilD) were marked and shown.

groups (Figure 5). Jc-GGPPS bore a closer relationship to GGPPSs from angiosperms than those from gymnosperms.

Biological function analysis of Jc-GGPPS in *Escherichia coli*

In the color complementation assay, we used *E. coli* strain Top 10F' to test if Jc-GGPPS encoded the anti-cipated functional enzyme. The plastid methylerythritol 4-phosphate (MEP) pathway exists in *E. coli.* As a result, the Top 10F' containing plasmid pAC-BETA can manu-facture and accumulate β-carotene and form yellow colonies. The Top 10F' containing plasmid pAC-BETA(β crtE), in which the *crtE* encoding GGPPS had been deleted, cannot manufacture and accumulate β-carotene and form white colonies. When pTrc-GGPPS containing the coding region of Jc-GGPPS was cotransformed into *E. coli* Top 10F' with plasmid pAC-BETA(βcrtE), the color of the bacterial colonies turned from white to yellow, indicating that Jc-GGPPS could induce the accumulation of β-carotene (Figure 6). Therefore, the function of the Jc-GGPPS has been confirmed by the change.

DISCUSSION

GGPPS plays different roles in various organisms. GGPPS supplies the essential acyclic precursor GGPP for the biosynthesis of structurally diverse group of compounds including diterpenes in plants. Diterpenes are produced by condensation of three units of dimethylallyl diphosphate (DMADP) with isopentenyl diphosphate (IPP).

The acyclic molecule (C20) produced, geranylgeranyl diphosphate (GGPP), undergoes a range of cyclisations to produce the parent backbone of all diterpenoids. In the pathway of plant diterpenes biosynthesis, there are four enzymes, and they are the synthase (FPS), the GGPPS and the diterpene synthase which catalyze the last four steps of diterpenes biosynthesis, respectively. This is followed by a variety of modifications of the parent backbone to produce the many thousands of different plant-origin diterpenes (Trapp and Croteau, 2001). We isolated a full-length Jc-GGPPS cDNA from *J. curcas.* Sequence alignment of Jc-GGPPS with GGPPS from other plant species revealed that Jc-GGPPS contains two aspartate-rich sequences [DDxx(xx)D]. DDlpcmD and DDilD, which are thought to be responsible for substrate binding, are found in all isoprenyl diphosphate synthase (IDS) described to date. Jc-GGPPS shares less than 65% identity with other reported GGPPS amino acid sequence. These conserved regions, as well as the active site suggested that Jc-GGPPS are likely to have the same function as other reported GGPPS. Highly conserved residues were designated as domains I-VII. These plant GGPPSs used in the multiple alignment analysis were selected out and the result are shown in Figure 3. Domains II and V contained aspartate-rich motifs (ARM), which are proposed to be diphosphate-binding sites (Chen et al., 1994). Domain II represent a longer region that contained highly conserved DD (Asp) and RR (Arg) dipeptides, DD(X)$_9$RR (Figure 2). The DDXXD motif in Domain V was the most conserved region and was proposed to be an allyl-isoprenoid binding site (Ashby and Edwards, 1990). Both DD(X)$_9$RR and DDXXD motifs were important for the catalytic activity of GGPPS (Sandmann, 1994). Further-
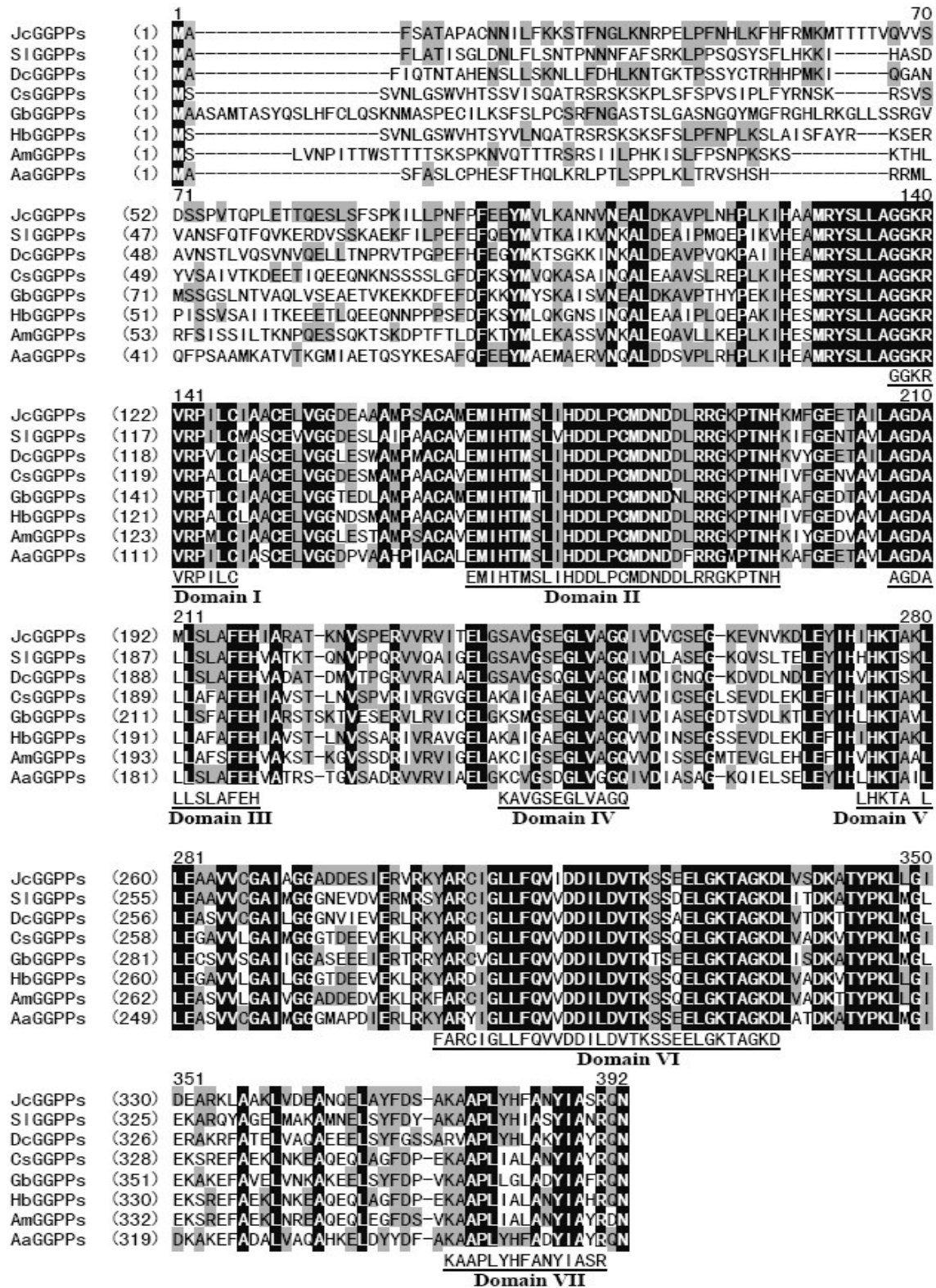
Figure 4. Alignment of Jc-GGPPS with other plant GGPPS proteins. The alignment was performed with Vector NTI Suite 8.0 by using the published sequences of plant GGPPS proteins including SlGGPPs (S. lycopersicum) (ABB82554), DcGGPPs (D. carota) (BAA78047), CsGGPPs (C. sublyratus)(BAA86284), GbGGPPs (Ginkgo biloba) (AAQ72786), HbGGPPs (H. brasiliensis)(BAB60678), AmGGPPs (A. majus) (AAS82860), AaGGPPs (A. aestivalis) (AAV74397). Conserved regions are underlined with a bold solid line. Conserved amino acid residues in all the sequences used in this alignment are black boxed, while similar amino acids are in gray boxes.
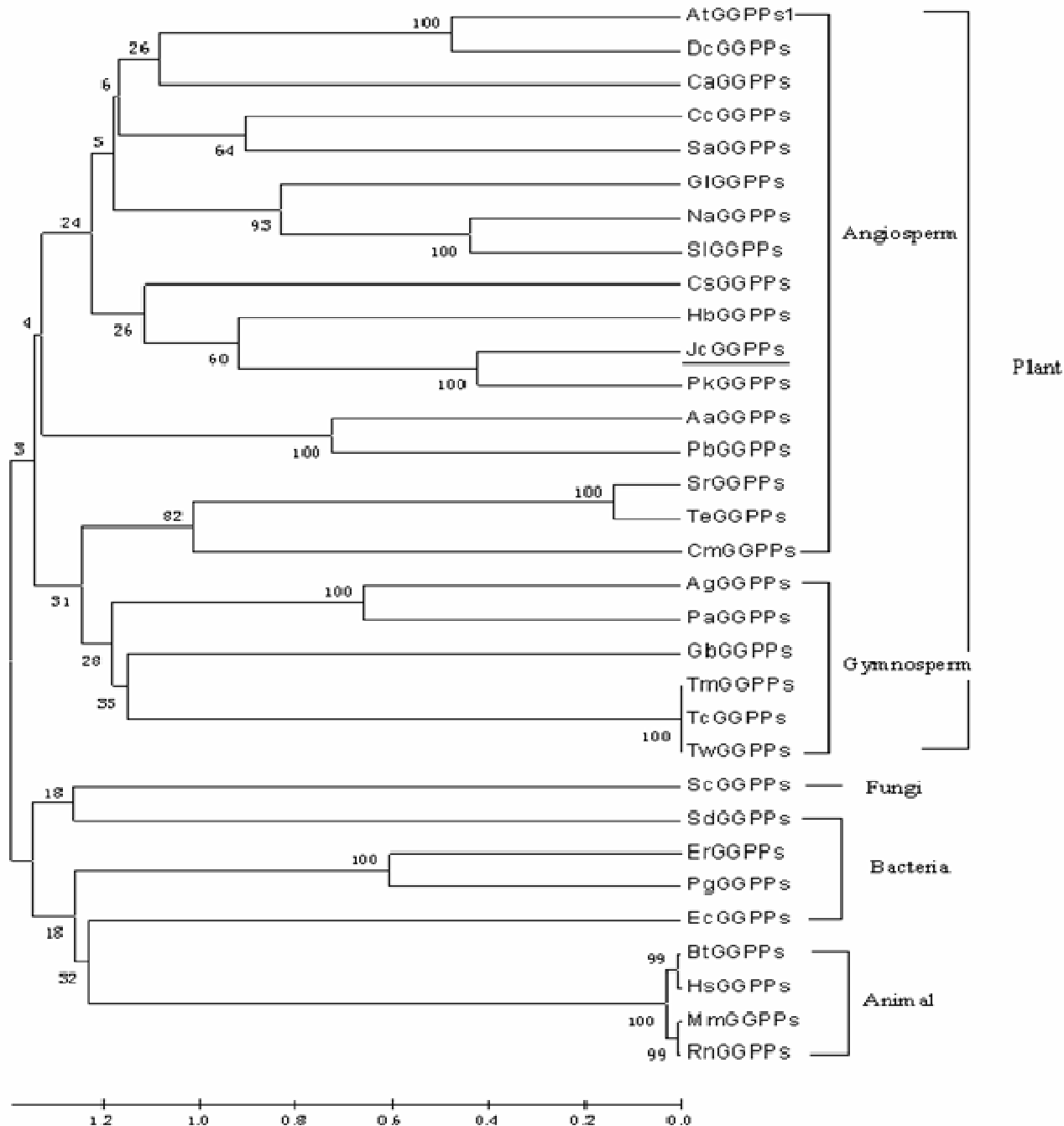
Figure 5. Phylogenetic relationships of GGPPS from plants, bacteria, fungi and mammals based on their amino acid sequences. These sequences were downloaded from NCBI, and species and the accession numbers are below: AtGGPPs (*A. thaliana*)(L25813), DcGGPPs (*D. carota*)(BAA78047), CaGGPPs (*C. annuum*)(CAA56554), CcPPGGs (*Cistus creticus*) (AAM21638), SaGGPPs (*S. alba*)(CAA67330), GlGGPPs (*G. lutea*) (BAB82463), NaGGPPs (*Nicotiana attenuata*) (ABQ53935), SlGGPPs (*S. lycopersicum*) (ABB82554), CsGGPPs (*C. sublyratus*)(BAA86284), HbGGPPs (*H. brasiliensis*)(BAB60678), PkGGPPs (*P. kurrooa*) (AAW66658), AaGGPPs (*A. aestivalis*) (AAV74397), PbGGPPs (*P. barbatus*) (AAR99082), SrGGPPs (*S. rebaudiana*) (ABD92926), TeGGPPs (*Tagetes erecta*)(AAG10424), CmGGPPs (*C. x morifolium*)(BAE79550), AgGGPPs (*A. grandis*)(AAL17614), PaGGPPs (*P. abies*) (ACA21461), GbGGPPs (*G. biloba*)(AAQ72786), Tm GGPPs (*T. x media*)(AAS49033), TcGGPPs (*Taxus canadensis*)(AAD16018), TwGGPPs (*Taxus wallichiana*) (ABC88389), ScGGPPs (*S. cerevisiae*)(SCU31632), SdGGPPs (*S. dulcis*)(BAA86285), ErGGPPs (*Erythrobacter* sp. NAP1)(ZP_01041671), PgGGPPs (marine gamma proteobacterium HTCC2080)(EAW41198), EcGGPPs (*E. coli*)(NP_414955), BtGGPPs (*B. taurus*)(NP_001073269), HsGGPPs (*Homo sapiens*)(AB016043), MmGGPPS (*M. musculus*) (NP_034412), RnGGPPS (*Rattus norvegicus*)(NP_001007627). The phylogenetic tree was constructed with the Neighbor-Joining method. The numbers indicated are the bootstrap values.
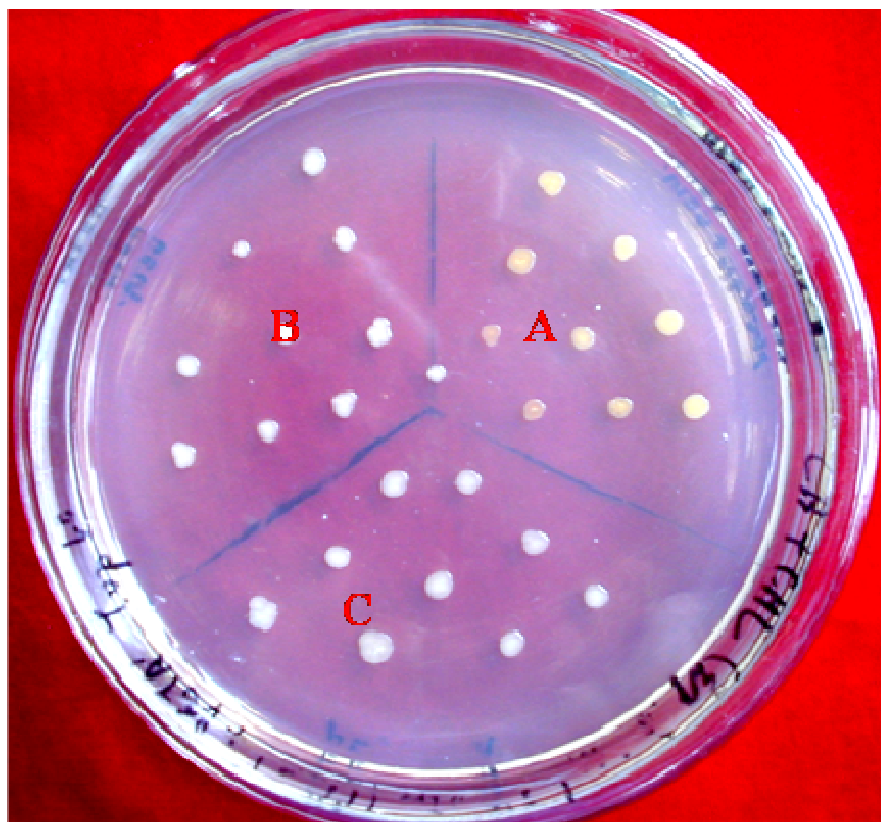
Figure 6. Functional demonstrations of Jc-GGPPS in Escherichia coli. The Escherichia coli strain TOP10 F' was respectively transformed with pAC-BETA (≥ crtE), and (A) pTrc-Jc-GGPPS harboring the Jc-GGPPS gene, (B) pAC-BETA (≥ crtE) and pTrc and (C) pAC-BETA (≥ crtE).

more, by analyzing the protein domain organization of plant GGPPSs, we found that all the GGPPSs had the same form of domain organization: the transit peptide, the linker peptide region between the transit peptide and the polyprenyltransferase domain, the functional region of GGPPSs (polyprenyltransferase) containing the five GGPPS domains (Sandmann, 1994), and the two ARM (Sitthithaworn et al., 2001) were organized in order along the peptide chain from N- to C-terminal.

The deduced amino acid sequence of the Jc-GGPPS signifies a putative presence of a transit peptide sequence to chloroplasts, which is predicted from ChloroP programs. It is likely that the GGPPS has similar functions as that of *Arabidopsis thaliana* GGPS-1, which is localize in the chloroplast and engages in the biosynthesis of biologically important isoprenoids, such as carotenoids, chlorophylls and gibberellins (Okada et al., 2000). Like the *A. thaliana GGPS1* gene, Jc-GGPPS did not contain intron sequence.

Analysis of the *in vivo* levels of GGPPSs in plants suggests that GGPPSs plays a regulatory role in diter-penoids biosynthesis. Altering the levels of expression of this enzyme has helped to define more clearly the influence of GGPPS on regulation of diterpenes biosynthesis in plants. For example, in tobacco, suppression of the expression of the GGPPS gene (*Naggpps*) results in dramatic reduction of levels of 17-hydroxygeranyllinalool diterpenoid glycosides (HGL-DTG) (Jassbi et al., 2008). These results indicated that the GGPPSs have a universal role in diterpene biosynthesis.

Functional expression of Jc-GGPPS was confirmed by complementation of the gene cluster for carotenoid biosynthesis of *Erwinia uredovora* lacking GGPPS. This system has been used to demonstrate the function of putative GGPPS from Arabidopsis (Zhu et al., 1997), sunflower (*Helianthus annuus*) (Oh et al., 2000) makandi (*C. forskohlii*) (Engprasert et al., 2004), human and mice (*Mus musculus*) (Kainou et al., 1999).

J. curcas have mainly been investigated as a source of oil. The seed kernel of the plant contains about 60% oil that can be converted into biodiesel and used as a substitute for diesel fuel. The seed cake remaining after oil extraction is an excellent source of plant nutrients (Goel et al., 2007). The use of genetic engineering to reduce toxin levels of *J. curcas* will be the first method developed to reduce expression levels of these toxins.

The cloning of Jc-GGPPS laid the molecular basis at a

genetic level to improved the materials. Further charac-terization of the Jc-GGPPS gene is needed to understand the regulation of phorbol esters biosynthesis in *J. curcas*, and to provide insights on how to decrease toxin content using biotechnology.

Abbreviations

GGPP, Geranylgeranyl diphosphate; **GGPPS,** geranylgeranyl diphosphate synthase; **IPP,** isopentenyl diphosphate; **ORF,** open reading frame; **FARM,** first aspartate rich motif; **SARM,** second aspartate rich motif; **D,** aspartate; **x,** any amino acid; **CTAB,** cetyl trimethyl ammonium bromide; **RACE,** rapid amplification of cDNA ends; **PCR,** polymerase chain reaction; LB, Luria-Bertani; **MEP,** methylerythritol 4-phosphate; **DMADP,** dimethylallyl diphosphate; **FPS,** farnesyl diphosphate synthase; **IDS,** isoprenyl diphosphate synthase; **ARM,** aspartate-rich motifs; **RR,** arginine dipeptides.

REFERENCES

Aitken SM, Attucci S, Ibrahim RK, Gulick PJ (1995). A cDNA encoding geranylgeranyl pyrophosphate synthase from *White lupin*. Plant Physiol. 108: 837-838.

Ament K, Van Schie CC, Bouwmeester HJ, Haring MA and Schuurink RC (2006). Induction of a leaf specific geranylgeranyl pyrophosphate synthase and emission of (e,e)-4,8,12-trimethyltrideca- 1,3,7,11-tetraene in tomato are dependent on both jasmonic acid and salicylic acid signaling pathways. Planta, 224: 1197-1208.

Ashby MN. Edwards PA (1990). Elucidation of the deficiency in two yeast coenzyme Q mutants. Characterization of the structural gene encoding hexaprenyl pyrophosphate synthetase. J. Biol. Chem. 265: 13157-13164.

Bantignies B, Liboz T, Ambid C (1996). Nucleotide sequence of a *Catharanthus roseus* geranylgeranyl pyrophosphate synthase gene. Plant Physiol. 110: p. 336.

Chen AJ, Kroon PA, Poulter CD (1994). Isoprenyl diphosphate synthases protein-sequence comparisons, a phylogenetic tree and redictions of secondary structure. Protein Sci. 3: 600-607.

Cunningham FX, Gantt E (2000). Identification of multigene families encoding isopentenyl diphosphate isomerase in plants by heterologous complementation in *Escherichia coli*. Plant Cell Physiol. 41: 119-123.

Cunningham FX, Gantt E (2007). A portfolio of plasmids for identifi-cation and analysis of carotenoid pathway enzymes: *Adonis aestivalis* as a case study. Photosynth. Res. 92: 245-259.

Dudley MW, Green TR, West CA (1986). Biosynthesis of the macrocyclic diterpene casbene in castor bean (*Ricinus communis L.*) seedlings: The purification and properties of farnesyl transferase from elicited seedlings. Plant Physiol. 81: 343-348.

Engprasert S, Taura F, Kawamukai M, Shoyama Y (2004). Molecular cloning and functional expression of geranylgeranyl pyrophosphate synthase from *Coleus forskohlii* Briq. BMC Plant Biol. 4: p. 18.

Goel G, Makkar HP, Francis G, Becker K (2007). Phorbol esters: structure, biological activity, and toxicity in animals. Int. J. Toxicol. 26: 279-288.

Hefner J, Ketchum RE, Croteau R (1998). 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.

Horst W, Timm T, Su YF, Fredrich K (2007). Biofuels in China: An analysis of the opportunities and challenges of *Jatropha curcas*. China: World Agroforestry Centre-ICRAF, pp. 1-27.

Jassbi AR, Gase K, Hettenhausen C, Schmidt A, Baldwin IT (2008). Silencing geranylgeranyl diphosphate synthase in *Nicotiana attenuata* dramatically impairs resistance to tobacco hornworm. Plant Physiol. 46: 974-986.

Jiang JX, Zhang TZ (2003). Extraction of total RNA in cotton tissue with CTAB-acidic phenolic method. Cotton Sci. 15: 166-167.

Kainou T, Kawamura K, Tanaka K, Matsuda H, Kawamukai M (1999). Identification of the GGPS1 genes encoding geranylgeranyl diphosphate synthases from mouse and human. Biochim. Biophys. Acta, 1437: 333-340.

Kishimoto S, Ohmiya A (2006). Regulation of carotenoid biosynthesis in petals and leaves of chrysanthemum (*Chrysanthemum morifolium*). Physiol. Plant, 128: 436-447.

Kloer DP, Welsch R, Beyer P, Schulz GE (2006). Structure and reaction geometry of geranylgeranyl diphosphate synthase from *Sinapis alba*. Biochemistry, 45: 15197-15204.

Kuntz M, Römer S, Suire C, Hugueney P, Weil JH, Schantz R, Camara B (1992). Identification of a cDNA for the plastid-located geranylgeranyl pyrophosphate synthase from *Capsicum annuum*: correlative increase in enzyme activity and transcript level during fruit ripening. Plant J. 2: 25-34.

Lan XZ, Aun M (2006). Cloning and characterization of the geranylgeranyl diphosphate synthase gene in *Taxus walliciana*. J. Southwest Agric. Univ. 28: 537-543.

Laskaris G, Van Der Heijden R, Verpoorte R (2000). Purification and partial characterization of geranylgeranyl diphosphate synthase, from *Taxus baccata* cell cultures. An enzyme that regulates taxane biosynthesis. Plant Sci. 153: 97-105.

Liang PH, Ko TP, Wang AH (2002). Structure, mechanism and function of prenyltransferases. Eur. J. Biochem. 269: 3339-3354.

Lin J, Zhou XW, Tang KX, Chen F (2004). A Survey of the studies on the resources of *Jatropha curcas*. J. Trop. Subtrop. Bot. 12: 285-290.

Makkar H, Becher K, Sporer F (1997). Studies on nutritive potential and toxic constituents of different provenances of *Jatropha curcas*. J. Agric. Food Chem. 45: 3152-3157.

Ogura K, Koyama T (1998). Enzymatic aspects of isoprenoid chain elongation. Chem. Rev. 98: 1263-1276.

Ogura K, Shinka T, Seto S (1972). The purification and properties of geranylgeranyl pyrophosphate synthetase from pumpkin fruit. J. Biochem. 72: 1101-1108.

Oh SK, Kim IJ, Shin DH, Yang J, Kang H, Han HK (2000). Cloning, characterization and heterologous expression of a functional geranylgeranyl pyrophosphate synthase from sunflower (*Helianthus annuus* L.). J. Plant Physiol. 157: 535-542.

Okada K, Sato T, Nakagawa T, Kawamukai M, Kamiya Y (2000). Five geranylgeranyl diphosphate synthase expressed in different organs are localized into three subcellular compartments in Arabidopsis. Plant Physiol. 122: 1045-1056.

Sagami H, Ishi K, Ogura K (1981). Occurence and unusual properties of geranylgeranyl pyrophosphate synthetase of pig liver. Biochem. Int. 3: 669-675.

Sandmann G (1994) Carotenoid biosynthesis in microorganisms and plants. Eur. J. Biochem. 223: 7-24.

Schmidt A, Gershenzon J (2007). Cloning and characterization of isoprenyl diphosphate synthases with farnesyl diphosphate and geranylgeranyl diphosphate synthase activity from norway spruce (*Picea abies*) and their relation to induced oleoresin formation. Phytochem. 68: 2649-2659.

Sitthithaworn W, Kojima N, Viroonchatapan E, Suh DY, Iwanami N, Hayashi T, Noji M, Saito K, Niwa Y, Sankawa U (2001). Geranylgeranyl diphosphate synthase from *Scoparia dulcis* and *Croton sublyratus*. Plastid localization and conversion to a farnesyl diphosphate synthase by mutagenesis. Chem. Pharm. Bull. 49: 197-202.

Takaya A, Zhang YW, Asawatreratanakul K, Wititsuwannakul D, Wititsuwannakul R, Takahashi S, Koyama T (2003). Cloning, expression and characterization of a functional cDNA clone encoding geranylgeranyl diphosphate synthase of *Hevea brasiliensis*. Biochim. Biophys. Acta 1625: 214-220.

Tholl D, Kish CM, Orlova I, Sherman D, Gershenzon J, Pichersky E, Dudareva N (2004). Formation of monoterpenes in *Antirrhinum majus* and *Clarkia breweri* flowers involves heterodimeric geranyl diphosphate synthases. Plant Cell, 16: 977-992.

Trapp SC, Croteau RB (2001). Genomic organization of plant terpene synthases and molecular evolutionary implications. Genetic, 158: 811-832.

West CA, Lois AF, Wickham KA, Ren YY (1990). Diterpenoid phytoalexins: biosynthesis and regulation. Recent Adv. Phytochem. 24: 219-248.

Zhu XF, Suzuki K, Okada K, Tanaka K, Nakagawa T, Kawamukai M, Matsuda H (1997). Cloning and functional expression of a novel geranylgeranyl pyrophosphate synthase gene from *Arabidopsis thaliana* in *Escherichia coli*. Plant Cell Physiol. 38: 357-361.