

Original Research Article

<https://doi.org/10.20546/ijcmas.2021.1002.137>

Functional Characterization of the Two Genes Encoding 1-deoxy-d-xylulose 5-phosphate Synthase in *Coleus forskohlii*

Sandeep Singh Pagoch¹, Ramesh Kumar², Yashbir S. Bedi¹ and Suphla Gupta^{1,3*}

¹Plant Biotechnology Department; Indian Institute of Integrative Medicine, Jammu, Jammu & Kashmir, India

²ICAR-National Bureau of Plant Genetic Resources, New Delhi, India

³Academy of Scientific & Innovative Research (AcSIR), CSIR, Anusandhan Bhawan 2 Rafi Marg, New Delhi, India

*Corresponding author

ABSTRACT

Keywords

Carotenoid, Color complementation assay, 2-C-methyl-D-erythritol 4-phosphate (MEP) pathway, Isoprenoid biosynthesis, Modeling

Article Info

Accepted:

10 January 2021

Available Online:

10 February 2021

The MEP pathway provides the initial precursors for the biosynthesis of terpenoids and carotenoid. Herein, we are reporting for the first time isolation, characterization and functional validation of the two isoforms of *CfDXS* gene from *Coleus forskohlii*. The full-length cDNAs of *CfDXS1* and *CfDXS2* were 2371 bp and 2541 bp, respectively having an open reading frame of 2148 bp and 2169 bp. The ORFs of *CfDXS1* and *CfDXS2* encoded protein of 715 and 722 amino acids, respectively, sharing 62.25% identity. Phylogenetic tree analysis classified the two isoforms in their respective plant clades. The *CfDXS1* gene was constitutively expressed in all the tissues (root, stem and leaf) of plant whereas *CfDXS2* gene expression was found to be higher in roots. The amino acid residues of the active site essential for the catalysis with the substrate, glyceraldehyde 3-phosphate, were found to be conserved in the *CfDXS1* and *CfDXS2* proteins. Identification of the function of *CfDXS2* in the biosynthesis of carotenoid biosynthesis was confirmed using color complementation assay as it accelerated the biosynthesis of β -carotene, which was not as effective in *CfDXS1*. Therefore, our results suggested the differential role of *CfDXS1* and *CfDXS2* in the plant growth and development.

Introduction

Terpenoids constitute a structurally diverse group of compounds, which have been implicated in a variety of biological processes¹⁻³. In plants, these compounds act as phytohormones, pigments, membrane constituents and as defense molecules^{1,4}. Apart from these, several terpenoid plant

secondary metabolites, often being species specific, are of human interest due to their medicinal and aromatic properties⁵. Owing to their biological significance and the economic value, there has been growing interest to understand terpenoid biosynthesis at the molecular level and to engineer their pathways in plants and microorganisms⁶.

Isopentenyl diphosphate (IPP) and its isomer DMAPP are the universal precursors of isoprenoids and, therefore, these are also termed as isoprene building blocks^{7, 8}. Depending upon the number of isoprene units, present in a structure, the terpenoids can be further classified into mono, di, tri, tetra-terpenes, etc⁵. The isoprene building blocks are biosynthesized by two separate metabolic routes, namely mevalonate pathway and 2-C-methyl-D-erythritol 4-phosphate pathway, which are localized in cytosol and plastid, respectively⁹⁻¹¹. Although certain level of cross-talk between the two pathways is apparent, the MEP pathway has been suggested to be the primary contributor of isoprene precursors, required in the biosynthesis of pigments (chlorophylls and carotenoids), phytohormones (gibberellins and abscisic acid), and monoterpene and diterpene secondary metabolites etc. The first reaction of MEP pathway is essentially a transketolase-like decarboxylation involving substrates namely pyruvate and glyceraldehyde-3-phosphate and catalyzed by the 1-deoxy-D-xylulose 5-phosphate synthase (*DXS*) enzyme¹²⁻¹⁷. Following this reaction, there are a series of downstream enzymatic steps, which ultimately lead to the biosynthesis of isoprene precursors.

The *DXS* enzyme, catalyzing the very first step within the MEP pathway is the pivotal enzyme for isoprenoid biosynthesis. Various studies have documented that there is a close correlation between *DXS* transcript and isoprenoid content in plant tissues^{11,18-20}. Also, the enhanced expression of *DXS* was reported to intensify the terpenoid biosynthesis in plants²¹⁻²⁴. Several factors such as light, developmental cues, elicitor treatment have been found to have modulating effect over the expression level of *DXS* genes^{8,25-28}. These studies suggested that both transcriptional and posttranscriptional mechanism is operative in the control of *DXS* expression²⁸.

The *DXS* genes have been cloned and characterized from different plant species²⁹⁻³¹. They constitute a small gene family, which can be further resolved into three phylogenetically distinct groups³². The different isoforms of *DXS* has been reported to be expressing differentially in different plant parts as well as under different exogenous factors. Furthermore, based on the expression profile of *DXS* genes, it was concluded that various *DXS* isoforms might be contributing towards biosynthesis of different functional types of isoprenoid compounds. For example, the group I *DXS* genes are supposed to be primarily involved in the biosynthesis of isoprenoid compounds having primary function³³. On the other hand, the group II *DXS* genes have primarily been linked to the biosynthesis of terpenoid secondary metabolites³⁴.

Coleus forskohlii accumulates forskolin, a diterpene class of secondary metabolite, used as a cAMP-regulating and vasodilator drug³⁵. A detailed genetic makeup of the forskolin biosynthesis, in terms of genes and enzymes is desirable for the development of strategies to modulate its biosynthesis in its native plant system as well as to engineer the pathway in microorganisms. Recently, the exploration of transcriptome data has provided important insights into the genes involved in the committed steps leading to the forskolin biosynthesis³⁶. However, the MEP pathway genes, which play a key role in the biosynthesis of diterpene compounds, have not been studied in detail from this plant species so far. Considering important role of the MEP pathway in isoprenogenesis, we herein, report on the cloning and characterization of two *CfDXS* genes for *C. forskohlii*. Based on the *in silico* analysis, functional studies in an engineered bacterium and expression analysis, we conclude that these *CfDXS* genes might be involved in specific aspects of terpenoid biosynthesis in *C. forskohlii*.

Materials and Methods

Plant materials

A rich stock of *C. forskohlii* established at the experimental farm of IIIM (CSIR-Indian Institute of Integrative Medicine, Jammu, India, 32°44'N longitude, 74°55'E latitude; 305 m in altitude) was used as a source material.

Full-length cloning of the *CfDXS* gene

Total RNA was isolated from young leaf, stem and root of *C. forskohlii* plant using Trizol (Invitrogen, US) as per the manufacturer's instructions. RNase-free DNase (Fermentas, Burlington, Canada) treated RNA (5 µg) was subjected to the first strand cDNA synthesis using RevertAid cDNA synthesis kit (Fermentas, Burlington, Canada) as per manufacturer's instructions.

Design of degenerate primers

The forward and reverse primers (Table 1) were designed based on the conserved amino acid domains retrieved from the GenBank database at National Centre of Biotechnology Information (NCBI) of the DXS protein assembled from different plant species. The Reverse Transcriptase-Polymerase chain reaction (RT-PCR) was performed using 20 µl of PCR reaction product containing 1U of Taq polymerase in a 20 µl reaction on Mastercycler® Pro S (Eppendorf, Germany) using the thermal conditions as described in Table 1. The amplicon obtained was gel extracted, purified (Qiagen, Netherland) and cloned in pTZ57R/T vector (InsTAclone™ PCR Cloning kit; Fermentas, Thermo Fisher Scientific, USA), before transformation in *E. coli* DH5α cells (New England Biolabs, Ipswich, MA, USA). The positive clones obtained were selected on the basis of blue/white screening and were further

confirmed by sequencing using a BigDye terminator cycle sequencing kit (Applied Biosystems, Foster City, CA, USA) with an ABI PRISM® 3130XL genetic analyzer (Applied Biosystems, Foster City, CA, USA). The sequenced cloned fragment was analyzed using BLASTX. The resulted sequence was used for designing the gene-specific primers (GSP) for 3'- and 5'-RACE.

3'- and 5'-RACE

The 3'- and 5'-rapid amplification of cDNA ends (RACE) was carried out using the First Choice® RLM RACE kit (Ambion®, Life Technologies, USA) as per manufacturer's instructions. 3'- and 5'-RACE PCR was performed using their respective specific primers (GSP) and adapters (provided in the kit) (Table 1). This was followed by nested PCR of the 3' and 5' template with respective gene-specific inner primers (GSP) and adapters (provided in the kit). Both the reactions were performed in thermo cycler with conditions given in Table 1. The 3' and 5'-RACE amplicons obtained were cloned in pTZ57R/T vector and sequenced. The 5'-RACE amplicon, core fragment and 3'-RACE amplicon were aligned using NCBI BLAST align software (<http://www.ncbi.nlm.nih.gov/>) at its default settings.

The open reading frame (ORF) was located in the full-length sequence of the *CfDXS* genes using the ORF finder. The gene specific full-length primers (Table 1), designed from the ORF regions of *CfDXS*, were used for the amplification of ORF region using high fidelity proof-reading DNA polymerase (Phusion, Fermentas) under the PCR thermal conditions as given in Table 1. The resultant confirmed full-length amplicons were sequenced, ligated in pJET vector and subsequently introduced and maintained in *E. coli* DH5α strain.

Sequence and bioinformatic analyses

The amino acid sequence similarities of the CfDXS proteins were analyzed using the BLAST tool of NCBI database. Primer3 Input software (Primer3_www.cgi v.0.2; <http://frodo.wi.mit.edu/>) was used for designing the primers. The ExpASy translate tool (<http://web.expasy.org/cgi-bin/translate/dna>) was used for translating the nucleotide sequence of CfDXS1 and CfDXS2. Prot Param (<http://www.expasy.ch/tools/protparam.html>) was used to estimate the properties of deduced amino acid sequence of CfDXS1 and CfDXS2. The conserved sequence of the CfDXS1 and CfDXS2, their active site residues and the residues of the side chain involved in the interaction with the substrate required for the biosynthesis of the product were identified by comparing the CfDXS sequences with the already published *E. coli* and *D. radiodurans* sequence³⁷.

The secondary structure was predicted by using the PBIL based SOPMA³⁸. TargetP (<http://www.cbs.dtu.dk/services/TargetP/>) and PSORT (<http://psort.hgc.jp/form.html>) were used for cellular localization. The glycosylation sites of the protein were identified using the NetNGlyc 1.0 Server (<http://www.cbs.dtu.dk/services/NetNGlyc/>) whereas NetPhos 2.0 (<http://www.cbs.dtu.dk/services/NetPhos/>) was used to predict the post-translational modifications. Sequences of CfDXS along with their homologs from different plant species retrieved from the NCBI database were aligned using the ClustalW (<http://www.ebi.ac.uk>) using default settings. The phylogenetic tree was constructed by Neighbour-joining method using MEGA 6.0 software³⁹. The bootstrapping analysis was performed with 1000 replicates which defined the measure of the confidence in the subtree rooted at the node. The three-dimensional structures of

CfDXS1 and CfDXS2 protein were predicted by Phyre2⁴⁰ using the crystal structure of *Deinococcus radiodurans* DXS as template.

Tissue-specific semi-quantitative expression analysis

Tissue-specific expression of CfDXS1 and CfDXS2 gene was analyzed by isolating total RNA from leaves, stem and roots of *C. forskohlii*. DNase-treated RNA (5 µg) was used for first strand cDNA synthesis using RevertAid cDNA synthesis kit (Fermentas, Burlington, Canada) as per manufacturer's instructions. Semi-quantitative expression analysis was done using gene specific primer and actin primer (Table 1). Actin was taken as an internal control to normalize the amount of cDNA. PCR was performed at the following optimized condition: one cycle of 94°C for 1 min, 40 cycles of 94°C for 10 sec, 60°C for 20 sec and 72°C for 25 sec.

Heterologous expression of CfDXS2 in *E. coli*

Full-length CDS region of CfDXS2 gene was amplified after truncating the signal sequence. The forward and reverse primers were modified by adding the restriction sites of *XhoI* and *NotI*, respectively, upstream to start codon in the forward primer and downstream to stop codon in the reverse primer. The purified PCR amplified product was ligated with pJET vector which was further digested with the *XhoI/NotI* restriction enzymes. Digested PCR product was cloned in the pre-digested and purified pGEX4T-2 vector (viz. *XhoI/NotI*). The generated pGEX-CfDXS2 cassette was transformed in the BL21 (DE3) *E. coli* strain. The positive colonies grown on LB agar containing ampicillin (100 µg/ml) were selected and confirmed by sequencing. Transformant harboring pGEX-CfDXS2 cassette was inoculated in LB medium (100 ml) containing ampicillin (100 µg/ml) and

incubated overnight at 37°C. Primary culture (1%) was inoculated in LB broth (500 ml) containing respective antibiotic at the defined concentration. It was grown at 37°C until its optical density at 600 nm reached approximately at 0.5. Protein expression was induced by adding Iso-propyl β-D-1-thiogalactopyranoside (IPTG; Fermentas, Berligton, Canada) in culture at the concentration of 1 mM and it was incubated at 16°C/200 rpm for 12 h. Sorbitol, as an additive, was added (600 mM) in the culture after induction to increase the solubility of CfDXS2 protein. The bacterial cells were harvested by centrifugation and resuspended in 5x sodium dodecyl sulphate-polyacrylamide gel electrophoresis sample buffer (SDS-PAGE; 0.25 M Tris pH 6.8, 15% SDS, 50% glycerol, 25% β-mercaptoethanol, 0.01% bromophenol blue). The expression of the protein was analyzed on 10 % SDS-PAGE.

Functional analysis of CfDXS2 in *E. coli*

Biological function of CfDXS2 was examined using the plasmids pAC-BETA and pTrc-AtIPI which carries functional gene for the synthesis of β-carotene. The functional genes in the plasmid, pAC-BETA, included *crtE* (Geranylgeranyl pyrophosphate synthase), *crtB* (phytoene synthase), *crtI* (phytoene desaturase) and *crtY* (lycopene cyclase) genes⁴¹. It also contained a chloramphenicol resistance gene. The plasmid, pTrc-AtIPI, retained an ampicillin resistance gene and *AtIPI* gene, the product of which could enhance the accumulation of lycopene, the immediate precursor of β-carotene⁴². The PCR product obtained was digested with *Clai* and *NotI* and then ligated into the *Clai* and *NotI* sites of the pTrc-AtIPI plasmid to create a recombinant expression plasmid pTrc-CfDXS2. The putative positive plasmids were confirmed by restriction analysis and further verified by sequencing. The recombinant

plasmid pTrc-CfDXS2 having the target gene was transformed into bacterial host strain *E. coli* DH5α strain. The construct pTrc-CfDXS2 was co-transformed with pAC-BETA in the *E. coli* TOP10F' strain. The *E. coli* TOP10F' strains, co-transformed with pAC-BETA and pTrc-AtIPI and pAC-BETA alone, were used as control. The putative transformants were grown on selection medium, containing ampicillin (150 mg/ml) and chloramphenicol (50 mg/ml) and incubated at 37 °C for 2 days. Growth was observed in the selection media having transformants and the controls harboring both the plasmids without the target gene. No growth was seen on the plates having pAC-BETA only. Color of the transformants was used as a visible marker to test whether CfDXS2 can accelerate the accumulation of β-carotene.

Results and Discussion

Identification and cloning of two CfDXS cDNAs from *C. forskohlii*

A homology-based PCR approach was adapted for the cloning of CfDXS genes from *C. forskohlii*. Following PCR with a set of degenerate primers, a core fragment corresponding to the cDNA of DXS gene was amplified. Based on its sequence information, the 3' and 5' RACE were carried out to establish the corresponding cDNA of 2541 bps, containing an open reading frame of 2169 bps, with 111 and 257 bps of 3' and 5' UTRs, respectively [Fig. 1(A)]. Our BLASTX search against NCBI database revealed that this cDNA is homologous to that of DXS2 from other plant species and, therefore, we named it as CfDXS2. The DXS genes in plants have been reported to constitute a small gene family, which can be differentiated into three subfamilies, namely DXS1, DXS2 and DXS3^{30,32}. We attempted to identify other DXS gene family members from *C. forskohlii*.

To this end, we explored an assembled transcriptome data of *C. forskohlii* (available at Medicinal Plant RNA Seq Database, <http://www.medplantrnaseq.org/>) and carried out BLASTN search using *CfDXS2* cDNA sequence as a query. The analysis led us to identify a cDNA of 2371 bps, with an open reading frame of 2148 bps, which displayed homology to *DXS1* genes of other plant species [Fig. 1(B)]. The full-length cDNA sequence of two isoforms of *CfDXS* were submitted to the GenBank database of NCBI with *CfDXS1* having accession number KU178948 and *CfDXS2* having accession number KU178949. We cloned the corresponding CDS and named it as *CfDXS1*. The CDS regions of *CfDXS1* and *CfDXS2* cDNAs were found to have 62.2% of sequence identity.

***In silico* characterization and phylogenetic analyses**

The two cDNAs, *CfDXS1* and *CfDXS2* were found to encode putative proteins, consisting of 715 and 722 amino acids, respectively. The predicted isoelectric points of *CfDXS1* and *CfDXS2* proteins were observed to be 6.58 and 6.49, respectively with their respective calculated molecular weights being 76.76 kDa and 77.79 kDa. The BLASTP searches against NCBI database suggested that the *CfDXS1* and *CfDXS2* are highly homologous to known DXS proteins from other plant species.

The secondary structure of *CfDXS2* predicted by PBIL based on the SOPMA, identified the composition of 36.43 % α -helices, 20.91 % β -extended strands, 10.66 % β -turns and 31.99% random coils whereas *CfDXS1* was identified to contain the 36.50 % α -helices, 18.18 % β -extended strands, 10.63 % β -turns and 34.69% random coils. It is well established that the enzymes of MEP pathway are localized into the plastids. An N-terminal

signal peptide has been proposed to direct the targeting of the corresponding gene products to the plastids⁴³. By using TargetP program, we attempted to predict the sub-cellular localization of *CfDXS1* and *CfDXS2* enzymes. Our analysis suggested that *CfDXS2* contains a putative signal peptide for the chloroplast targeting. However, interestingly, *CfDXS1* could not be predicted to have a canonical plastid targeting signal. Similar to our observation, (Cordoba *et al.*, 2011)³⁰ reported that one gene product corresponding to a DXS2 protein in maize did not contain canonical signal peptide for the plastid targeting and, therefore, was observed to be localized in cytosol instead of the chloroplast. Further *in vivo* studies, using C-terminal GFP-fusion would be helpful to affirm sub-cellular localization of *CfDXS1* protein. The deduced amino acids of two isoforms of *CfDXS* subjected to NetNGlyc 1.0 were predicted to have one glycosylation site in *CfDXS2* (89th amino acid residue) and two glycosylation sites in *CfDXS1* (11th amino acid and 103rd amino acid residue). The NetPhos 2.0 used for finding the post-translational modifications predicted the phosphorylated sites of 17 serine, 7 threonine and 6 tyrosine residues in *CfDXS2* whereas *CfDXS1* was shown to contain 14 serine, 9 threonine and 9 tyrosine residues, all having the score higher than 0.5, identifying each as a potential hit.

A phylogenetic tree of *CfDXS1* and *CfDXS2* and other proteins from different plant species was constructed. The topology of the tree resolved three distinct clades, namely DXS1, DXS2 and DXS3 (Fig. 2). The *CfDXS1* and *CfDXS2* were reported to be grouped within their respective clades. Furthermore, within each clade, DXS proteins from taxonomically related groups tended to group together. These observations suggested that different DXS gene family members, after having been diversified in a common ancestor, might have

evolved independently in different taxonomic groups. Similar observation pertaining to the phylogeny of DXS proteins has been made in earlier reports^{30,33,44}. More divergent *DXS3* genes have been reported from different plant species, including from the members of family *lamiaceae*⁴⁵. Therefore, we speculate that *C. forskohlii* genome might also encode for, yet to be identified *DXS3* gene(s).

Conserved regions in CfDXS protein

In order to study conserved residues among DXS proteins, we carried out a multiple sequence alignment (Fig. 3). The analysis further supported the conservation of DXS proteins including the CfDXS1 and CfDXS2. The consensus sequence of thiamine pyrophosphate (TPP)-binding motif of -G-D-G- and a conserved asparagine residue separated by 28 amino acid residues were found to be present in both the proteins CfDXS1 and CfDXS2²⁸. The residues for active site of DXS in *E. coli* were also identified in CfDXS2 (CfDXS1). The side chains of these residues in CfDXS2 (CfDXS1) such as His116 (His110), His380 (His374), Tyr473 (Tyr467), Arg501 (Arg495) and Asp508 (Asp502) were supposed to be involved in the interaction with the glyceraldehyde 3-phosphate (GAP) molecule³⁷ (Fig. 4 and Fig. S1). The amino acid residues reported to be critical for the catalysis in *E. coli* (*D. radiodurans*) viz. Glutamate370 (Glutamate373), Arginine398 (Arginine401) and Arginine478 (Arginine480) were also found to be conserved in the CfDXS2 (CfDXS1) viz. Glutamate451 (Glutamate445), Arginine479 (Arginine473) and Arginine559 (Arginine553). The possible interaction between Glutamate451 and Arginine479 and with the TPP-binding motif in case of CfDXS2 (Glutamate445 and Arginine473 and with the TPP-binding motif in CfDXS1) could be one of the possible reasons for the

activation of coenzyme for catalysis⁴⁶ and Arginine559 in CfDXS2 (Arginine553 in CfDXS1) could help recognize the GAP substrate³⁷. The conservation of characteristic residues of DXS proteins suggests that CfDXS1 and CfDXS2 genes might encode functional proteins *in planta*.

Homology-based modeling was used for designing the three-dimensional structural models of CfDXS1 and CfDXS2 amino acid sequence using Phyre2 with the >90% confidence. The distantly known template used to model the structure of CfDXS1 and CfDXS2 was that of *D. radiodurans* DXS with PDB: 2O1X. 586 amino acid residues (82%) in CfDXS1 and 584 amino acid residues (81%) in CfDXS2 were modeled at >90 % accuracy. Two templates with PDB: 2O1X and 2O1S in CfDXS1 protein and in CfDXS2 protein were selected to model in heuristics to maximize the sequence confidence, percentage identity and alignment coverage. 129 amino acid residues in CfDXS1 and 134 amino acid residues in CfDXS2 were modeled by *ab initio* which was highly unreliable. The homology modeling suggested that the CfDXS1 and CfDXS2 proteins have 3D structures comparable to that of other known DXS proteins. The location and identity of the active side residues were also reported to be highly conserved (Table S1 and Table S2).

Functional characterization of CfDXS2 gene

As evident through our *in silico* analysis, the putative CfDXS1 and CfDXS2 proteins contain characteristic conserved residues and structural features. These results suggested that both the proteins might be functional *in planta*. The open reading frame corresponding to CfDXS2 fused with that of GST was expressed in *E. coli* strain BL12 (DE3) using expression vector pGEX4T-2.

Table.1 Primer sequences and PCR conditions used in the present work for the amplification of genes from *C. forskohlii*

Name of the gene	Primer sequence (forward primer, F and reverse primer, R) (5'-3')	Used in the study	PCR Conditions	
<i>CfDXS(2)</i>	F: TGGGAYGTYGGYCAYCAG R: GGMGCCATNACMACCATGTT	Core Fragment	95 °C, 30 s; 55 °C, 30 s; 72 °C, 1:30 min (35 cycles).	
	GSP1:ATCCACGCGGCCATGGGAGG	3' End Amplification	95 °C, 30 s; 60 °C, 30 s; 72 °C, 1:30 min (35 cycles).	
	GSP2: GTC CCT CCC CAC CGC CAT CCC TA	5' End Amplification	95 °C, 30 s; 55 °C, 30 s; 72 °C, 1 min (35 cycles).	
	F:TTCTCGAGCAATGGCGTCTTGTGGAGC TATC R:TTGCGGCCGCCATGTTGATCAAATGA AGACTG	Full Length and Heterologous Expression in <i>E. coli</i>	95 °C, 30 s; 50 °C, 30 s; 72 °C, 3 min (35 cycles).	
	F:CGAGGAGCTCGGGATTTATTAC R:CCGTTGTTGGATCAAACCTTCAC	Real-Time analysis	95 °C, 30 s; 55 °C, 30 s; 72 °C, 20 sec (30 cycles).	
	F:TTTATCGATATGGCGTCTTGTGGAGCT ATC R:TTTGC GGCCGCTTACATGTTGATCAA TGAAGACTG	Functional analysis in <i>E. coli</i>	95 °C, 30 s; 55 °C, 30 sec; 72 °C, 3 min (35 cycles).	
	<i>CfDXS1</i>	R:GYCMTYCCY MAYCCYGMAGMGAT	5' End Amplification	95 °C, 30 s; 55 °C, 30 s; 72 °C, 1 min (35 cycles).
	F:CCTCTCCACTAAGGAACTGAAAC R:GTCTGTCTTAAACTCGGCATCT	Real-Time analysis	95 °C, 30 s; 55 °C, 30 s; 72 °C, 20 sec (30 cycles).	
	Actin	F: GAGAGTTTTGATGTCCCTGCCATG R: CAACGTCGCATTTTCATGATGGAGT	Real-Time analysis	95 °C, 30 s; 55 °C, 30 s; 72 °C, 20 sec (30 cycles)

^a R=A/G; M=A/C; W=A/T; Y=C/T; S=C/G; K=G/T; D=A/G/T; H=A/C/T; B=G/C/T; N=A/T/G/C. ^bGSP1, gene specific primer for 3'-RACE
GSP2, gene specific primer for 5'-RACE

Table.S1 Twelve amino acid residues in the TPP binding site of CfDXS 1 and CfDXS 2. Conserved amino acid

<i>E. coli</i> DXS	<i>D. radiodurans</i> DXS	CfDXS1	CfDXS2
Ser52	Ser54	Ser113	Ser119
His80	His82	His141	His147
Gly121	Gly123	Gly182	Gly188
Ser123	Ala125	Ser184	Ser190
Asp152	Asp154	Asp213	Asp219
Asn181	Asn183	Asn242	Asn248
Met183	Met185	Gln244	Gln250
Lys284	Lys289	Lys359	Lys365
Ile368	Ile371	Ile443	Ile449
Glu370	Glu373	Glu445	Glu451
Phe395	Phe398	Phe470	Phe476
Arg398	Arg401	Arg473	Arg479

Fig.5

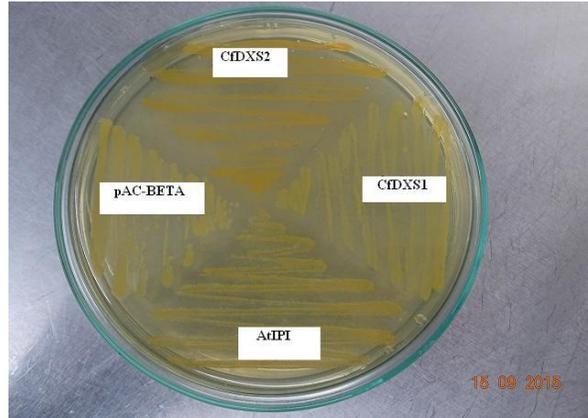


Fig.S1

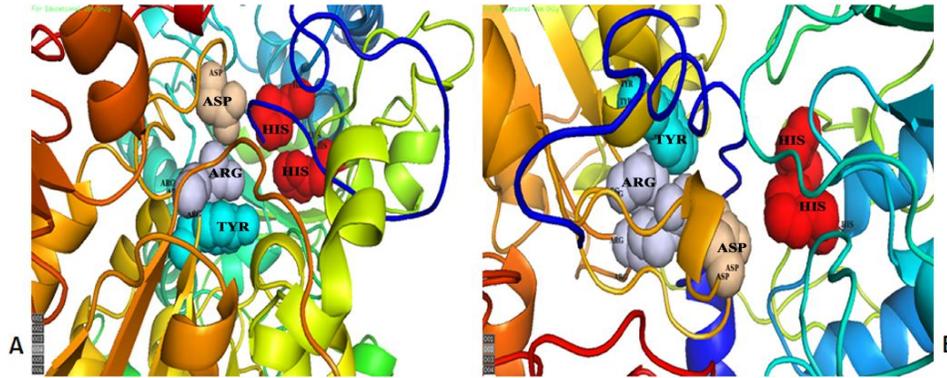


Fig.S2

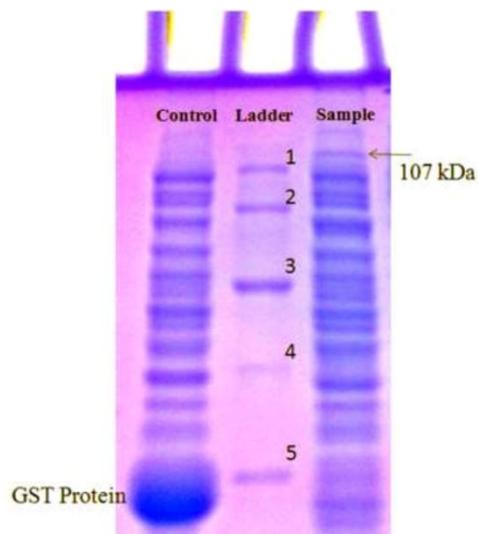
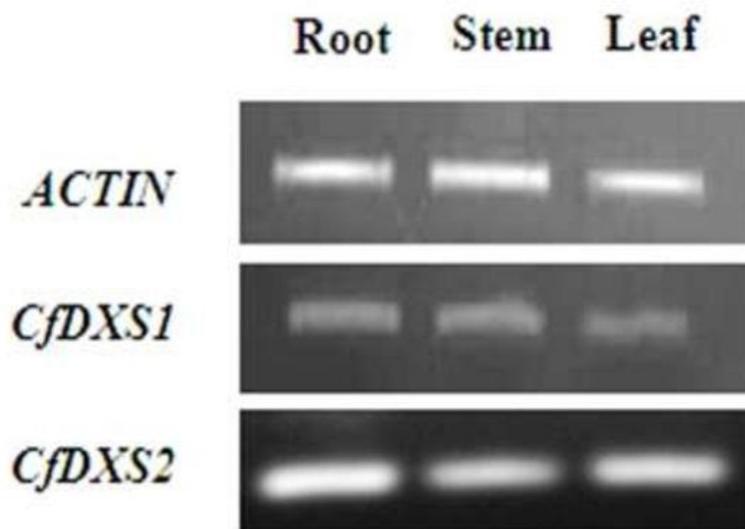


Fig.S3



The present study results demonstrated that the presence of sorbitol in the cell lysate increased the solubility of CfDXS2 protein. The heterologously expressed fusion protein obtained was of expected molecular weight (~107 kDa) (Fig. S2). To further confirm the functionality of *CfDXS2* gene, we explored functional complementation studies using a recombinant *E. coli* strain with carotenogenic gene cluster. The *E. coli* strain TOP10F⁷ harboring pAC-BETA plasmid is able to synthesize trace amount of β -carotene^{47,48}. The plasmid pAC-*AtIPI* contains *IPI* gene from *Arabidopsis*, which is responsible for isoprenogenesis. The *CfDXS2* ORF was cloned in pAC-*AtIPI* plasmid by disrupting *AtIPI* ORF. The recombinant plasmid, pAC-*CfDXS2*, thus developed was transformed in the TOP10F⁷ strain harboring pAC-BETA. The recombinant strain harboring both the plasmid displayed intense yellow coloration suggesting the boosting of isoprenoid pathway by *CfDXS2* expression (Fig. 5). These results indicated that *CfDXS2* might encode a functional protein responsible for isoprenogenesis. Similar color complementation system has been utilized for

functional studies of various genes of isoprenoid biosynthesis from other plant species⁴⁹⁻⁵². The expression of *CfDXS1* in the recombinant strain did not result into any significant alteration in the colors of colonies. Provided that like *CfDXS2*, *CfDXS1* contained all the conserved motifs, characteristics of DXS enzymes, this result was surprising. Further, the enzyme activity assays using substrates of DXS would properly explain these issues.

Expression analyses of DXS genes

In order to gain insights into the regulation of *CfDXS* genes in different tissues, we carried out expression analysis using semi-quantitative RT-PCR. The analysis suggested that *CfDXS1* and *CfDXS2* were expressed in different tissues of *C. forskohlii*. The *CfDXS1* expression level was reported to be roughly same in the different tissues tested, while *CfDXS2* was highly upregulated in roots as compared to other tissues (Fig. S3). The forskolin biosynthesis is primarily restricted to the roots of *C. forskohlii*⁵³. Therefore, it can be speculated that *CfDXS2* might be

responsible for driving MEP pathway in favour of forskolin biosynthesis, while CfDXS1 could be involved in isoprenogenesis to feed primary pathways. Similar conclusions pertaining to the division of function among members of *DXS* gene family have been documented by earlier works^{26,34,54}.

Taken together, we have identified and characterized two *CfDXS* genes from *C. forskohlii*. Our data suggests that these genes could be involved in MEP pathway of *C. forskohlii*. Further functional validation using overexpression and suppression approaches in *Coleus* should affirm their role in terpenoid biosynthesis in general and forskolin biosynthesis in particular.

In this study, we have characterized two isoforms of 1-deoxy-D-xylulose 5-phosphate synthase viz. *CfDXS1* and *CfDXS2* from *C. forskohlii*, an important medicinal plant. Multiple sequence alignment with other reported DXS showed high sequence similarity with CfDXS1 and CfDXS2 proteins. The TPP-binding motifs and other amino acid residues essential for the recognition of substrate and required for the catalysis were mostly found to be conserved in both the proteins. Based on the results of heterologous protein expression in *E. coli*, carotenoid pathway intensification in an engineered carotenogenic bacterium and the expression pattern, it was concluded that CfDXS2 might be involved in isoprenogenesis, required for the biosynthesis of diterpenoid class of secondary metabolites in *C. forskohlii*.

Author contributions

SSP carried out all the experimental work and wrote the manuscript. YSB provided vital inputs for the study. SG designed the study and analyzed the results.

Acknowledgments

Acknowledgment is due to Dr. Ram Ashrey Vishwakarma, Director, for providing the necessary facilities and encouragement. This work was financially supported by the network project grant (BSC106) of Council of Scientific and Industrial Research (CSIR). Authors wish to express their gratitude to Dr. Francis X. Cunningham for providing the vector used in the present study. SSP acknowledge Council of Scientific and Industrial Research (CSIR) for the Senior Research Fellowship. Acknowledgment is also due to Dr. Satiander Rana, Dr. Sumit G. Gandhi and Dr. Prashant Misra.

References

1. Pichersky, E. and Gershenzon, J. 2002. The formation and function of plant volatiles: perfumes for pollinator attraction and defense. *Curr Opin Plant Biol*, 5: 237-243.
2. Tholl, D. 2006. Terpene synthases and the regulation, diversity and biological roles of terpene metabolism. *Curr Opin Plant Biol*, 9: 297-304.
3. Gershenzon, J. and Dudareva, N. 2007. The function of terpene natural products in the natural world. *Nat Chem*, 3: 408-414.
4. Mazid, M., Khan, T.A. and Mohammad, F. 2011. Role of secondary metabolites in defense mechanisms of plants. *Biology and medicine*, 3: 232-249.
5. Croteau, R., Kutchan, T. and Lewis, N. 2000. Natural Products (Secondary Metabolites). *Biochemistry and Molecular Biology of Plants*, 1250–1268.
6. Ikram, N.K.B, Zhan, X, Pan X.W., King, B.C. and Simonsen, H.T. 2015. Stable heterologous expression of biologically active terpenoids in green plant cells. *Front Plant Sci*, 6
7. Eisenreich, W., Schwarz, M., Cartayrade,

- A., Arigoni, D., Zenk M.H. and Bacher, A. 1998. The deoxyxylulose phosphate pathway of terpenoid biosynthesis in plants and microorganisms. *Chem Biol*, 5: 221-233.
8. Cordoba, E., Salmi, M. and Leon, P. 2009. Unravelling the regulatory mechanisms that modulate the MEP pathway in higher plants. *J Exp Bot*, 60: 2933-2943.
 9. Lichtenthaler, H.K., Rohmer, M. and Schwender, J. 1997. Two independent biochemical pathways for isopentenyl diphosphate and isoprenoid biosynthesis in higher plants. *Physiol Plant*, 101 (1997) 643–652.
 10. Hemmerlin A, Hoeffler JF, Meyer O, Tritsch D, Kagan IA, Grosdemange-Billiard C, Rohmer M and Bach TJ. 2003. Cross-talk between the cytosolic mevalonate and the plastidial methylerythritol phosphate pathways in tobacco bright yellow-2 cells. *J Biol Chem*, 278: 26666-26676.
 11. Dudareva N, Andersson S, Orlova I, Gatto N, Reichelt M, Rhodes D, Boland W and Gershenzon J. 2005. The nonmevalonate pathway supports both monoterpene and sesquiterpene formation in snapdragon flowers. *Proc Natl Acad Sci USA*, 102: 933–938.
 12. Sprenger GA, Schörken U, Wiegert T, Grolle S, De Graaf AA, Taylor SV, Begley TP, Bringer-Meyer S and Sahn H. 1997. Identification of a thiamin-dependent synthase in *Escherichia coli* required for the formation of the 1-deoxy-D-xylulose 5-phosphate precursor to isoprenoids, thiamin, and pyridoxol. *Proc Natl Acad Sci USA*, 94:12857-12862.
 13. Bouvier F, d'Harlingue A, Suire C, Backhaus RA and Camara B 1998. Dedicated roles of plastid transketolases during the early onset of isoprenoid biogenesis in pepper fruits. *Plant physiol*, 117:1423-1431.
 14. Lange BM, Wildung MR, McCaskill D and Croteau R 1998. A family of transketolases that directs isoprenoid biosynthesis via a mevalonate-independent pathway. *Proc Natl Acad Sci USA*, 95:s2100-2104.
 15. Lois LM, Campos N, Putra SR, Danielsen K, Rohmer M and Boronat A 1998. Cloning and characterization of a gene from *Escherichia coli* encoding a transketolase-like enzyme that catalyzes the synthesis of D-1-deoxyxylulose 5-phosphate, a common precursor for isoprenoid, thiamin, and pyridoxol biosynthesis. *Proc Natl Acad Sci USA*, 95:2105-2110.
 16. Miller B, Heuser T and Zimmer W 1999. A *Synechococcus leopoliensis* SAUG 1402-1 operon harboring the 1-deoxyxylulose 5-phosphate synthase gene and two additional open reading frames is functionally involved in the dimethylallyl diphosphate synthesis. *FEBS Lett*, 460:485-490.
 17. Kuzuyama T, Takagi M, Takahashi S and Seto H 2000. Cloning and characterization of 1-deoxy-d-xylulose 5-phosphate synthase from *Streptomyces* sp. Strain c1190, which uses both the mevalonate and nonmevalonate pathways for isopentenyl diphosphate biosynthesis. *J Bacteriol*, 182:891-897.
 18. Chahed K, Oudin A, Guivarc'h N, Hamdi S, Chénieux JC, Rideau M and Clastre M. 2000. 1-Deoxy- d-xylulose 5-phosphate synthase from periwinkle: cDNA identification and induced gene expression in terpenoid indole alkaloid-producing cells. *Plant Physiol Biochem*, 38:559–566.
 19. Totté N, Ende WVD, Van Damme EJ, Compennolle F, Baboeuf I and Geuns JM. 2003. Cloning and heterologous expression of early genes in gibberellin and steviol biosynthesis via the

- methylerythritol phosphate pathway in *Stevia rebaudiana*. *Can J Bot*, 81:517–522.
20. Paetzold H, Garms S, Bartram S, Wieczorek J, Urós-Gracia EM, Rodríguez-Concepción M, Boland W, Strack D, Hause B and Walter MH. 2010. The isogene 1-deoxy-D-xylulose 5-phosphate synthase2 controls isoprenoid profiles, precursor pathway allocation, and density of tomato trichomes. *Mol Plant*, 3:904–916.
 21. Estévez JM, Cantero A, Reindl A, Reichler S and León P. 2001. 1-Deoxy-D-xylulose-5-phosphate synthase, a limiting enzyme for plastidic isoprenoid biosynthesis in plants. *J Biol Chem*, 276:22901-22909.
 22. Khemvong S and Suvachittanont W. 2005. Molecular cloning and expression of a cDNA encoding 1-deoxy-D-xylulose-5-phosphate synthase from oil palm *Elaeis guineensis* Jacq. *Plant Sci*, 169:571-578.
 23. Gong YF, Liao ZH, Guo BH, Sun XF and Tang KX. 2006. Molecular cloning and expression profile analysis of *Ginkgo biloba* DXS gene encoding 1-deoxy-D-xylulose 5-phosphate synthase, the first committed enzyme of the 2-C-methyl-D-erythritol 4-phosphate pathway. *Planta Med*, 72:329.
 24. Muñoz-Bertomeu J, Arrillaga I, Ros R and Segura J. 2006. Up-regulation of 1-deoxy-D-xylulose-5-phosphate synthase enhances production of essential oils in transgenic spike lavender. *Plant physiol*, 142:890-900.
 25. Lois LM, Rodríguez-Concepción M, Gallego F, Campos N and Boronat A. 2000. Carotenoid biosynthesis during tomato fruit development: regulatory role of 1-deoxy-D-xylulose 5-phosphate synthase. *Plant J*, 22:503-513.
 26. Walter MH, Fester T and Strack D. 2000. Arbuscular mycorrhizal fungi induce the non-mevalonate methylerythritol phosphate pathway of isoprenoid biosynthesis correlated with accumulation of the ‘yellow pigment’ and other apocarotenoids. *Plant J*, 21:571-578.
 27. Okada A, Shimizu T, Okada K, Kuzuyama T, Koga J, Shibuya N, Nojiri H and Yamane H. 2007. Elicitor induced activation of the methylerythritol phosphate pathway toward phytoalexins biosynthesis in rice. *Plant Mol Biol*, 65:177-187.
 28. Han M, Heppel SC, Su T, Bogs J, Zu Y, An Z and Rausch T. 2013. Enzyme inhibitor studies reveal complex control of methyl-D-erythritol 4-phosphate (MEP) pathway enzyme expression in *Catharanthus roseus*. *Plos One*, 8.5:e62467.
 29. Zhang M, Li K, Zhang C, Gai J and Yu D. 2009. Identification and characterization of class 1 DXS gene encoding 1-deoxy-D-xylulose-5-phosphate synthase, the first committed enzyme of the MEP pathway from soybean. *Mol Biol Rep*, 36:879-887.
 30. Cordoba E, Porta H, Arroyo A, San Román C, Medina L. 2011. Rodríguez-Concepción M and León P, Functional characterization of the three genes encoding 1-deoxy-D-xylulose 5-phosphate synthase in maize. *J Exp Bot*, 62:2023–2038.
 31. Yang J, Adhikari MN, Liu H, Xu H, He G, Zhan R, Wei J and Chen W. 2012. Characterization and functional analysis of the genes encoding 1-deoxy-D-xylulose-5-phosphate reductoisomerase and 1-deoxy-D-xylulose-5-phosphate synthase, the two enzymes in the MEP pathway, from *Amomum villosum* Lour. *Mol Biol Rep*, 39:8287-8296.
 32. Kim BR, Kim SU and Chang YJ. 2005. Differential expression of three 1-deoxy-D-xylulose-5-phosphate synthase genes

- in rice. *Biotechnol Lett*, 27:997-1001.
33. Walter MH, Hans J and Strack D. 2002. Two distantly related genes encoding 1-deoxy-d-xylulose 5-phosphate synthases: differential regulation in shoots and apocarotenoid-accumulating mycorrhizal roots. *Plant J*, 31:243-254.
 34. Saladié M, Wright LP, Garcia-Mas J. 2014. Rodriguez-Concepcion M and Phillips MA, The 2-C-methylerythritol 4-phosphate pathway in melon is regulated by specialized isoforms for the first and last steps. *J Exp Bot* 275.
 35. Alasbahi RH and Melzig MF 2010. *Plectranthus barbatus*: a review of phytochemistry, ethnobotanical uses and pharmacology-part 2. *Planta Med*, 76:753-765.
 36. Zerbe P, Hamberger B, Yuen MM, Chiang A, Sandhu HK, Madilao LL, Nguyen A, Hamberger B, Bach SS and Bohlmann J. 2013. Gene discovery of modular diterpene metabolism in nonmodel systems. *Plant physiol*, 162:1073-1091.
 37. Xiang S, Usunow G, Lange G, Busch M and Tong L. 2007. Crystal structure of 1-deoxy-D-xylulose 5-phosphate synthase, a crucial enzyme for isoprenoids biosynthesis. *J Biol Chem*, 282:2676-2682.
 38. Geourjon C and Deleage G. 1995. SOPMA: significant improvements in protein secondary structure prediction by consensus prediction from multiple alignments. *Computer Applications in the Biosciences: CABIOS*, 11:681-684.
 39. Tamura K, Stecher G, Peterson D, Filipski A, Kumar S. 2013. MEGA6: molecular evolutionary genetics analysis version 6.0. *Mol Biol Evol*, 30:2725-2729.
 40. Kelley LA and Sternberg MJ. 2009, Protein structure prediction on the Web: a case study using the Phyre server. *Nat Protoc*, 4:363-371.
 41. Misawa N, Satomi Y, Kondo K, Yokoyama A, Kajiwara S, Saito T, Ohtani T and Miki W. 1995. Structure and functional analysis of a marine bacterial carotenoid biosynthesis gene cluster and astaxanthin biosynthetic pathway proposed at the gene level. *J Bacteriol*, 177:6575-6584.
 42. Cunningham FX and Gantt E. 2000. Identification of multi-gene families encoding isopentenyl diphosphate isomerase in plants by heterologous complementation in *Escherichia coli*. *Plant Cell Physiol*, 41:119-123.
 43. Carretero-Paulet L, Ahumada I, Cunillera N, Rodríguez-Concepción M, Ferrer A, Boronat A and Campos N. 2002. Expression and molecular analysis of the *Arabidopsis* DXR gene encoding 1-deoxy-d-xylulose 5-phosphate reductoisomerase, the first committed enzyme of the 2-c-methyl-d-erythritol 4-phosphate pathway. *Plant physiol*, 129:1581-1591.
 44. Phillips MA, Walter MH, Ralph SG, Dabrowska P, Luck K, Urós EM, Boland W, Strack D, Rodríguez-Concepción M, Bohlmann J and Gershenzon J. 2007. Functional identification and differential expression of 1-deoxy-D-xylulose 5-phosphate synthase in induced terpenoid resin formation of Norway spruce (*Picea abies*). *Plant Mol Biol*, 65:243-257.
 45. Ma Y, Yuan L, Wu B, Li XE, Chen S and Lu S. 2012, Genome-wide identification and characterization of novel genes involved in terpenoid biosynthesis in *Salvia miltiorrhiza*. *J Exp Bot*. err466.
 46. Schellenberger A. 1998. Sixty years of thiamin diphosphate biochemistry. *BBA-Protein Struct M*, 1385:177-186.
 47. Cunningham FX, Sun Z, Chamovitz D, Hirschberg J and Gantt E. 1994. Molecular structure and enzymatic function of lycopene cyclase from the *Cyanobacterium synechococcus* Sp Strain

- Pcc7942. *Plant Cell*, 6:1107-1121.
48. Cunningham FX, Pogson B, Sun Z, McDonald KA, DellaPenna D and Gantt E. 1996. Functional analysis of the beta and epsilon lycopene cyclase enzymes of *Arabidopsis* reveals a mechanism for control of cyclic carotenoid formation. *Plant Cell*, 8:1613-1626.
 49. Gao S, Lin J, Liu X, Deng Z, Li Y, Sun X and Tang K. 2006. Molecular cloning, characterization and functional analysis of a 2C-methyl-D-erythritol 2, 4-cyclodiphosphate synthase gene from *Ginkgo biloba*. *BMB Rep*, 39:502-510.
 50. Wang Y, Guo B, Zhang F, Yao H, Miao Z and Tang K. 2007. Molecular cloning and functional analysis of the gene encoding 3-hydroxy-3-methylglutaryl coenzyme A reductase from hazel (*Corylus avellana* L. Gasaway). *BMB Rep*, 40:861-869.
 51. Lin J, Jin Y, Zhou X and Wang J. 2010. Molecular cloning and functional analysis of the gene encoding geranylgeranyl diphosphate synthase from *Jatropha curcas*. *Afr J Biotechnol*, 9:3342-3351.
 52. Lan X. 2013. Molecular cloning and characterization of the gene encoding 2-C-methyl-D-erythritol 4-phosphate cytidyltransferase from hairy roots of *Rauvolfia verticillata*. *Biologia*, 68:91-98.
 53. Pateraki I, Andersen-Ranberg J, Hamberger B, Heskes AM, Martens HJ, Zerbe P, Bach SS, Møller BL, Bohlmann J and Hamberger B. 2014. Manoyl oxide (13R), the biosynthetic precursor of forskolin, is synthesized in specialized root cork cells in *Coleus forskohlii*. *Plant physiol*, 164:1222-1236.
 54. Vallabhaneni R and Wurtzel ET. 2009. Timing and biosynthetic potential for carotenoid accumulation in genetically diverse germplasm of maize. *Plant Physiol*, 150:562-572.

How to cite this article:

Sandeep Singh Pagoch, Ramesh Kumar, Yashbir S. Bedi and Suphla Gupta. 2021. Functional Characterization of the Two Genes Encoding 1-deoxy-d-xylulose 5-phosphate Synthase in *Coleus forskohlii*. *Int.J.Curr.Microbiol.App.Sci*. 10(02): 1158-1175.
doi: <https://doi.org/10.20546/ijcmas.2021.1002.137>