Molecular Cloning and Functional Analysis of the Gene Encoding 3-hydroxy-3-methylglutaryl Coenzyme A Reductase from Hazel (Corylus avellana L. Gasaway)

Ye Chun Wang1, Binhui Guo1, Fei Zhang1, Hongyan Yao1, Zhiqi Miao1 and Kexuan Tang1,2,*

1 Plant Biotechnology Research Center, School of Agriculture and Biology, School of Life Science and Technology, Fudan-SJTU-Nottingham Plant Biotechnology R&D Center, Shanghai Jiao Tong University, Shanghai 200030, PR China
2 State Key Laboratory of Genetic Engineering, School of Life Sciences, Morgan-Tan International Center for Life Sciences, Fudan-SJTU-Nottingham Plant Biotechnology R&D Center, Fudan University, Shanghai 200433, PR China

Received 1 May 2007, Accepted 9 June 2007

The enzyme 3-hydroxy-3-methylglutaryl-CoA reductase (HMGR; EC1.1.1.34) catalyzes the first committed step of isoprenoids biosynthesis in MVA pathway. Here we report for the first time the cloning and characterization of a full-length cDNA encoding HMGR (designated as CgHMGR, GenBank accession number EF206343) from hazel (Corylus avellana L. Gasaway), a taxol-producing plant species. The full-length cDNA of CgHMGR was 2064 bp containing a 1704-bp ORF encoding 567 amino acids. Bioinformatic analyses revealed that the deduced CgHMGR had extensive homology with other plant HMGRs and contained two transmembrane domains and a catalytic domain. The predicted 3-D model of CgHMGR had a typical spatial structure of HMGRs. Southern blot analysis indicated that CgHMGR belonged to a small gene family. Expression analysis revealed that CgHMGR expressed high in roots, and low in leaves and stems, and the expression of CgHMGR could be up-regulated by methyl jasmonate (MeJA). The functional color assay in Escherichia coli showed that CgHMGR could accelerate the biosynthesis of β-carotene, indicating that CgHMGR encoded a functional protein. The cloning, characterization and functional analysis of CgHMGR gene will enable us to further understand the role of CgHMGR involved in taxol biosynthetic pathway in C. avellana at molecular level.

Keywords: Hazel, 3-hydroxy-3-methylglutaryl-CoA reductase (HMGR), RACE, Taxol biosynthesis

Introduction

Taxol, one of the best selling anti-cancer drugs worldwide, was originally isolated with extremely low yield from the bark of Taxus brevifolia (Wani et al., 1971), which is one of the slowest growing trees in the world. It is evident that the harvesting of Taxus bark is not a viable long-term option for taxol production in industry. Therefore it is necessary and important to find alternative sources to produce taxol, one of the most expensive anti-cancer drugs. In 1998, an exciting discovery was announced that taxol could be extracted from the angiosperm Corylus avellana L. by Hofman and her colleagues, and later by other groups (Hofman et al., 1998; Boone et al., 2000). Very recently, it has been further confirmed that hazel cell cultures could produce taxol and taxanes under controlled conditions (Bestoso et al., 2006). Although the amount of taxol found in hazel was about one-tenth that of the yew, hazel was a widely grown species and grows much faster than yew. Therefore, hazel is a new potential source for taxol production in the future (Service, 2000). In addition, knowledge of the biosynthetic pathway is important for improving the production of this drug. Taxol synthesis can be regulated by overexpressing the key rate-limiting enzyme genes of taxol biosynthetic pathway (Eisenreich et al., 1996; Jennewein et al., 2004). Isolation and analysis of genes encoding key enzymes involved in the taxol biosynthetic pathway would be helpful to enhance taxol production through

Database Accession No: EF206343

Abbreviations: HMGR, 3-hydroxy-3-methylglutaryl coenzyme A reductase; MEP, 2C-methyl-D-erythritol 4-phosphate; MVA, mevalonic acid; MeJA, methyl-jasmonate; PCR, Polymerase chain reaction; RACE, rapid amplification of cDNA ends; RT-PCR, reverse transcriptase-polymerase chain reaction.

*To whom correspondence should be addressed. Tel: 86-21-62932002; Fax: 86-21-62824073
E-mail: kxtang1@yahoo.com or kxtang1@163.com
taxol bioengineering in the near future.

In plants, it is well known that there are two distinct isoprenoids biosynthesis pathways, the plastidic 2C-methyl-D-erythritol 4-phosphate (MEP) pathway and the cytosolic mevalonic acid (MVA) pathway (Rohmer et al., 1993). But the two pathways are not separated absolutely. In some extens, there are some forms of Crosstalk between them (Laule et al., 2003). As diterpenes, taxol is formed in the isoprenoid pathway. It has been demonstrated that the precursor of taxol is biologically synthesized from both MVA pathway and MEP pathway (Lansing et al., 1991; Zamir et al., 1991; Eisenreich et al., 1996; Palazon et al., 2003). The 3-hydroxy-3-methylglutaryl-CoA reductase (HMGR; EC1.1.1.34) catalyzes the first committed step in MVA pathway for biosynthesis of isoprenoids (Chappell, 1995) and earlier studies suggested that Taxus produced high rates of radioactively-labelled taxol when supplied with labeled mevalonate, demonstrating that taxol was produced from mevalonate, which was catalyzed by HMGR (Lansing et al., 1991; Zamir et al., 1992). Therefore, HMGR probably acts as one of the key regulatory enzymes for taxol biosynthesis in MVA pathway. Up to now, there has been no report on the cloning of HMGR gene from hazel.

In this paper, we report for the first time the molecular cloning and characterization of CgHMGR gene from hazel by RACE technique. The expression profiles of the CgHMGR and its biological function in E. coli were also investigated.

Materials and Methods

Plant materials and growth conditions. The taxol-producing hazel (Corylus avellana L. Gasaway), kindly provided by Drs. Shawn A. Mehlenbacher and Bruce R. Bartlett at Oregon State University, USA, was used as the research material. For surface sterilization, hazel seeds were soaked in 75% (v/v) ethanol for 1 min and then rinsed 3 times with sterile distilled water. Hazel embryos were put out and then germinated on solid Murashige and Skoog (1962) medium (pH 5.8) with 3% (w/v) sucrose and grown in darkness at 25°C for two weeks. After two weeks, the germinated seedlings were transferred to soil and grown in a greenhouse at 27 ± 1°C with 16 h light and 8 h dark photoperiod.

Plant treatments. Leaves from four-week-old seedlings were sprayed with 2 mM MeJA (dissolved in 0.1% ethanol and water). Meanwhile, leaves were also sprayed with only 0.1% ethanol without any elicitors as the control. The treated and control plant leaves were harvested respectively at 0, 6, 16, 24, 36, 48, 72, 96, and 108 h after treatment for RNA extraction. All experiments were carried out in triplicates.

DNA and RNA isolation. The four-leaf stage seedlings were used for genomic DNA and RNA isolation in the experiments. Genomic DNA and total RNA of hazel were extracted using the modified CTAB method (Paterson et al., 1993; Izakola et al., 2001). The quality and concentration of the extracted DNA and RNA were checked by agarose gel electrophoresis and monitored by Nucleic Acid Protein Analyzer (DU-640, Beckman). The DNA and RNA samples were stored at −20°C and −70°C respectively prior to use.

Isolation of CgHMGR core cDNA fragment. Two degenerate oligonucleotide primers, HMGRF (5'-GGGCTC/TGATGGGAT/C/TGATGGGACA/TCTG-3') and HMGRR (5'-ACA/T/C/TGATGGGAT/C/TGATGGGACA/TCTG-3'), were designed and synthesized based on the conserved amino acid sequence regions of several plant HMGRs. RT-PCR was performed to amplify the core fragment of CgHMGR gene with the two degenerate primers according to the protocol of One Step RNA PCR Kit (TaKaRa). The PCR program was carried out under the following condition: 50 min at 50°C, 94°C for 2 min followed by 30 cycles of amplification (30 s of denaturation at 94°C, 30 s of annealing at 58°C, 1 min of extension at 72°C). After the final cycle, the amplification was extended for 7 min at 72°C. The amplified products were subcloned into pMD18-T vector (TaKaRa, Japan) and then sequenced by ABI 3730 Sequencer (Perkin-Elmer, USA). Subsequently, the core fragment was used to design and synthesize gene-specific primers to clone the 3' and 5' fragments of CgHMGR by RACE.

3' and 5' RACE and the full-length cDNA cloning of CgHMGR. A SMART™ RACE cDNA Amplification Kit (Clontech) was used to isolate the 3'-ends of CgHMGR cDNA. The 3'-ready and 5'-ready cDNA libraries were synthesized by reverse transcription 1 μg of total RNA with 3'-CD5 primer A and 5'-CD5 primer, respectively. For 3' RACE of CgHMGR, the first round PCR was performed to amplify the 3'-end cDNA with the 3'-gene-specific primer CgHMGR 3-1 (5'-GTTGGCGAAGGCTCTTCTCATC-3') and UPM (Universal Primer A Mix, provided by the kit), and 3'-RACE ready cDNA was used as template. The PCR product diluted 20-fold was used as template for the nested PCR with CgHMGR 3-2 (5'-CACATGCCATGCATCTC-3') and NUP (provided in the kit). For 3' RACE of CgHMGR, two gene-specific primers, CgHMGR 5-1 (5'-CCTATAGGAGCTCCTCCAACACAGC-3') and CgHMGR 5-2 (5'-GGCGTTTCTCGAGCCTTGTGGC-3') were designed and synthesized according to the sequence of the obtained core fragment. The first round PCR was performed with primer CgHMGR 5-1 and UPM, and 5'-RACE ready cDNA was used as template. The PCR product diluted 20-fold was used as template for the nested PCR with CgHMGR 5-2 and NUP. An Advantage 2 PCR Kit (Clontech) was used in the first round and nested PCR amplification of 5'-end and 3'-end. The PCR program was used according to the protocol (SMART™ RACE cDNA Amplification Kit, user manual Clontech) with 25 cycles of amplification (30 s at 94°C, 30 s at 68°C, 1 min at 72°C). Subsequently, the products were ligated into pMD18-T vector followed by sequencing. After aligning and assembling the sequences of the core fragment, 3' RACE and 5' RACE products, the full-length cDNA sequence of CgHMGR gene was deduced and subsequently obtained by PCR using PCR primers CgHMGRF (5'-GAACCAAAAAACCCCAATCCACACACACA-3') and CgHMGRR (5'-TACAAAAAGATACATACAAACAAACCAGGGGT-3') with 5'-RACE-ready cDNA as template under the following condition: 3 min at 94°C followed by 35 cycles of amplification (30 s at 94°C, 30 s at 60°C, 140 s min at 72°C), with an additional 7 min at 72°C. The PCR products were purified, cloned into pMD18-T vector and sequenced.
Bioinformatic analyses. The sequence analysis of CgHMGR was carried out online at the websites (http://www.ncbi.nlm.nih.gov and http://www.expasy.org). Vector NTI Suite 8 was used for sequence multi-alignment analysis. Homology-based structural modeling was accomplished by Swiss-Model (Schwede et al., 2003), and WebLab ViewerLite 4.0 was used for 3-D model display.

Southern blot analysis. Aliquots of hazel genomic DNA (25 µg/sample) were digested overnight at 37°C with BglII, EcoRI and HindIII respectively, which did not cut within the Hybond N+ region, electrophoresed in 0.8% agarose gel and transferred to the expression profiles of RT-PCR analysis. The membrane (Amersham Biosciences) according to the standard procedure. The 682-bp probe was generated by PCR using the full-length sequence of CgHMGR as template with primers HMGRSF (5'-CCAGAGCAGACA TCAGAGACAG-3') and HMGRSR (5'-GCGA TCGACC-3') and then labeled by biotin using the random primer method (the Amersham Multiprime Labeling system). Southern blot analysis was carried out according to the manufacturer’s protocols (Amersham Pharmacia). The hybridized signals were visualized by exposure to Fuji medical X-ray film at room temperature for 2-3 h.

RT-PCR analysis. Semi-quantitative RT-PCR was used to investigate the expression profiles of CgHMGR in different tissues including roots, stems and leaves, as well as leaves under different treatment with 2 mM MedA as mentioned earlier. Total RNA was extracted from roots, stems and treated leaves respectively, followed by incubation with RNase-free DNase I at 37°C for 30 min according to the manufacturer’s protocols (TaKaRa, Japan). Aliquots of total RNA (1 µg/sample) were used as templates in one-step RT-PCR with primers HMGRSF and HMGRSR. Meanwhile, two primers, 18SF (5'-AAGACCTACACCAAGCCCA-3') and 18SR (5'-AA GTGAGCCCACACTTACCA-3'), were also used to amplify the house-keeping gene (18S rRNA gene) as an internal control. The expression profiles of CgHMGR were cloned into the empty vector pTrc to generate the plasmid pTrc-CgHMGR which was verified by sequencing. The pTrc-AIPI was digested by PstI and ligated by T4 DNA ligase as control. The vectors pTcre-CgHMGR and pAC-BETA were co-transformed into the E. coli strain TOP10 F' as controls. Transformants were cultured on solid LB medium containing ampicillin (150 mg/L) and chloramphenicol (50 mg/L) at 37°C for 2 d. The color of the transformants was as a visible marker to test the function of CgHMGR.

Results and Discussion

Generation of the full-length cDNA of CgHMGR. The genes encoding HMGR have been cloned and characterized
from some higher plant species such as *Cucumis melo* (Kato-Emori et al., 2001), *Taxus x media* (Liao et al., 2004), *Ginkgo biloba* (Shen et al., 2006) and *Eucalyptus ulmoides* (Jiang et al., 2006) because of its importance for primary and secondary metabolite biosynthesis. Total RNA extracted from hazel leaves, and the degenerate primers (HMGRF and HMGRR) were used to amplify a 458-bp product by RT-PCR. A BLASTn search revealed that the obtained fragment had high
homology with HMGR genes from other plant species. Two pairs of gene-specific primers were then designed and synthesized for the 3' RACE and 5' RACE based on the obtained fragment. By using the above mentioned methods, the full-length cDNA sequence of CgHMGR was deduced, amplified and subsequently confirmed by sequencing. The full-length cDNA was 2064 bp with 5' and 3' untranslated regions (UTR) and a polyA tail. The ORF search result showed that CgHMGR contained a 1704 bp-ORF encoding a 567-amino-acid protein.

Characterization of the deduced CgHMGR protein. The predicted CgHMGR protein had a molecular mass of 60.83 kDa and a theoretical pI value of 6.43 (http://cn.expasy.org/tools/protparam.html), which was very similar to the previously reported *G. biloba* HMGR (Shen et al., 2006). Protein-Protein BLAST and multiple alignment analysis showed the deduced CgHMGR amino acid sequence had high homology with HMGR sequences from other plant species, such as *Hevea brasiliensis* (80%), *Camptotheca acuminate* (77%), *Nicotiana tabacum* (73%), *Lycopersicon esculentum* (72%), *Ginkgo biloba* (70%), *Taxus x media* (68%), respectively, suggesting that CgHMGR belonged to the HMGR family.

It was well known that HMGRs were classified into two distinct classes: eukaryotic HMGRs (class I) and prokaryotic HMGRs (class II) (Bochar et al., 1999). The membrane domains of plant HMGRs (class I) contained two membrane-spanning helices (Istvan et al., 2000). Two transmembrane regions of CgHMGR were predicted by TMHMM 2.0 (http://www.cbs.dtu.dk/services/TMHMM-2.0/), revealing that one region was located between Pro34 (P) and Leu56 (L) and the other was located between Ile77 (I) and Val99 (V) along the polypeptide chain. The amino acid multiple alignment analysis showed that the N-terminal end of CgHMGR was quite diverse in both length and composition, and the C-terminal catalytic domain of CgHMGR had high similarity when compared with other plant HMGRs (Fig. 2). Plant HMGR comprised two HMG-CoA binding motifs (EMPIGYVQIP and TTEGCLVAV) and two NADPH-binding motifs (DAMGMNM and GSVGGTG) (Kato-Emori et al., 2001; Liao et al., 2004; Jiang et al., 2006; Shen et al., 2006). The highest homology regions appeared around the substrate-
binding sites (Fig. 2). The sequence analysis revealed that the functional motifs of CgHMGR were very similar to those of other plant HMGRs.

The result of a ramachandran plot was shown in Fig. 3. The ramachandran plot showed that all the main chain conformational angles fell into core and allowed regions. Therefore, the structure has good main chain stereochemistry. The 3-D structure of CgHMGR was analyzed by Swiss-Modeling on the basis of the Homo sapiens HMGR crystal structure. Comparison of CgHMGR protein with human HMGR revealed 56.4% identity in 426 aa overlap, which was good for homology-based modeling. The result showed that the catalytic domain of CgHMGR consisted of three domains including the small helical N-terminal domain, the large L-domain containing two HMG-CoA-binding motifs and a NADP(H)-binding motif and the smallest S-domain harboring a NADP(H)-binding motif (Fig. 4). The 3-D structure of CgHMGR strongly resembled human HMGR, indicating that they had potential catalytic similarities.

Southern blot analysis. In plants, HMGR is encoded by two genes in Arabidopsis thaliana (Enjuto et al., 1994), three genes in potato (Korth et al., 1997) and Hevea brasiliensis (Chye et al., 1992), and more genes in tomato (Denbow et al., 1996), wheat (Aoyagi et al., 1993) and mulberry (Jain et al., 2000). To investigate the genomic organization of CgHMGR in hazel, Southern blot analysis was carried out by digesting the hazel genomic DNA with three different restriction enzymes (BglII, EcoRI, and HindIII) respectively, followed by hybridization with the 682-bp probe generated by PCR using the coding sequence of CgHMGR as template. The result showed that at least three hybridizing bands were present in each lane under high stringency conditions (Fig. 5), indicating that CgHMGR belonged to a small gene family, like the counterparts from some other tree species such as Tavax x media and G. biloba (Liao et al., 2004; Shen et al., 2006).
Expression profile of CgHMGR in different tissues of *C. avellana*. To investigate the expression pattern of CgHMGR in different tissues, total RNA was isolated from roots, stems and leaves respectively, and subjected to RT-PCR analysis. The result showed that CgHMGR was constitutively expressed in all tested tissues but at different levels with the highest expression in roots, followed by leaves and stems (Fig. 6), similar to that reported in *Ginkgo biloba* (Shen et al., 2006). It was earlier reported that the content of taxol was the highest in roots (Liu et al., 2001), which was consistent with the expression pattern of CgHMGR gene.

Many studies revealed that taxane production could be induced by various elicitors, such as methyl jasmonate (MeJA) in *Taxus* and hazel (Yukimune et al., 1996; Baebler et al., 2002; Bestoso et al., 2006). However, little was known about the expression profiles of genes involved in taxol biosynthetic pathway under elicitor treatments. Therefore, it is worthwhile studying expression patterns of genes in taxol biosynthesis under various treatments including MeJA, which will be helpful to uncover molecular induction mechanism for further improving taxol biosynthesis in hazel. Semi-quantitative RT-PCR was performed to study the expression of CgHMGR in response to MeJA. When treated with 2 mM MeJA, the expression of CgHMGR increased gradually within 48 h, reached the highest level after 72 h and maintained the level thereafter (Fig. 7), demonstrating that the expression of CgHMGR could be induced by MeJA. This result suggests that MeJA treatment might be an effective approach to induce higher production of taxol from hazel.

Confirmation of CgHMGR function in *E. coli*. E. coli could not synthesize carotenoids because of lacking of carotenogenic genes. However, *E. coli* introduced with foreign carotenogenic genes clusters had the ability to produce carotenoids, and transformed bacteria, owing to accumulating carotenoid pigments, could act as a visible marker for providing an easily screenable phenotype (Matthews and Wurzel, 2000).

To test if CgHMGR encoded the anticipated functional protein, the plasmids pAC-BETA and pTrek-CgHMGR harboring CgHMGR gene (A), pAC-BETA and pTrek (B), and pAC-BETA (C).

![Fig. 6. Expression profile of CgHMGR in different tissues of *C. avellana*. Total RNA was extracted from roots, stems and leaves respectively, and subjected to RT-PCR analysis. The 18S rRNA gene was used as the control for the normalization of RNA loading amount in RT-PCR reaction. The experiment was repeated three times.](image1)

![Fig. 7. Expression profile of CgHMGR under 2 mM methyl jasmonate (MeJA) induction treatment. RT-PCR analysis is performed by using total RNA isolated from treated leaves at different time points (0, 6, 16, 24, 36, 48, 72, 96, and 108 h). The 18S rRNA gene was used as the loading control. The experiment was repeated three times.](image2)

![Fig. 8. Functional demonstration of CgHMGR in *E. coli*. The *E. coli* strain TOP10 F' was respectively transformed with pAC-BETA and pTrek-CgHMGR harboring CgHMGR gene (A), pAC-BETA and pTrek (B), and pAC-BETA (C).](image3)
(Fig. 8), demonstrating that CgHMGR could accelerate the accumulation of β-carotene. The E. coli TOP10 F’ harboring single vector pAC-BETA could not grow on LB medium containing ampicillin and chloramphenicol due to the lack of ampicillin resistance gene. Since the precursors of β-carotene (tetraterpenes) and taxol (diterpenes) are biosynthesized through the same pathway (Eisenreich et al., 1998), it can be expected that CgHMGR helps to accelerate the accumulation of taxol in hazel. In fact, it has been reported that amorphadiene production was improved by 50% through transferring HMGR gene into engineered yeast (Ro et al., 2006).

In summary, we have successfully cloned and characterized a functional gene encoding 3-hydroxy-3-methylglutaryl-CoA reductase (HMGR) involved in the biosynthesis of taxol from C. avellana. Multiple alignments showed that the deduced CgHMGR had high identity to other plant HMGRs, and contained all conserved substrate-binding motifs of HMGRs. The three dimensional model of CgHMGR represented a typical spatial structure of HMGRs. Color complementation assay further confirmed that CgHMGR encoded a functional protein and played an important role in promoting β-carotene pathway flux. The cloning and characterization of CgHMGR will be helpful to understand the taxol biosynthesis in C. avellana at molecular level. Intensive studies, such as genetic transformation of CgHMGR gene into hazel seedlings, is expected that CgHMGR helps to accelerate the accumulation of taxol in hazel.

Acknowledgments The authors would like to acknowledge Drs. Shawn A. Mehlenbacher and Bruce R. Bartlett (Oregon State University, USA) for providing the hazel seeds and seedlings. This work was financially supported by the China National High-tech “863” Program and Shanghai Science and Technology Committee.

References


