Plant Gene 5 (2016) 100–108

Contents lists available at ScienceDirect

Plant Gene

journal homepage: www.elsevier.com/locate/plantgene

Molecular cloning, gene expression profiling and in silico sequence analysis of vitamin E biosynthetic genes from the oil palm

Kong Sze Ling [a], Siti Nor Akmar Abdullah [a,b,*], Ho Chai Ling [c], Mohd Din Amiruddin [d]

[a] Laboratory of Plantation Crops, Institute of Tropical Agriculture, Universiti Putra Malaysia, 43400, UPM, Serdang, Selangor, Malaysia
[b] Department of Agriculture Technology, Faculty of Agriculture, Universiti Putra Malaysia, 43400, UPM, Serdang, Selangor, Malaysia
[c] Department of Cell and Molecular Biology, Faculty of Biotechnology and Biomolecular Sciences, University Putra Malaysia, 43400, UPM, Serdang, Selangor, Malaysia
[d] Advanced Biotechnology and Breeding Centre, Malaysian Palm Oil Board, No. 6, Persiaran Institusi, B. B. Bangi, 43000 Kajang, Selangor, Malaysia

ARTICLE INFO

Article history:
Received 30 July 2015
Received in revised form 6 January 2016
Accepted 18 January 2016
Available online 26 January 2016

Keywords:
Oil palm
Homogentisate phytyltransferase (HPT)
Homogentisate geranylgeranyl transferase (HGGT)
Degenerate primer
qPCR
Promoter analysis

ABSTRACT

Homogentisate geranylgeranyl transferase (HGGT) and homogentisate phytyltransferase (HPT) are the two key enzymes involved in condensation of homogentisic acid (HGA) with a prenyldiphosphate to produce tocotrienols and tocopherols in plants, respectively. The partial cDNAs encoding HGGT and HPT enzymes were successfully isolated from the two oil palm species, *Elaeis guineensis* and *Elaeis oleifera* by PCR amplification using degenerate primers. Subsequently, full length cDNA sequences were completed by rapid amplification of cDNA ends (RACE) and further annotated using various bioinformatics tools. The analysis revealed the presence of an UbiA prenyltransferase conserved domain in all four deduced amino acid sequences and suggested that oil palm HGGT and HPT are more evolutionarily related to their counterparts from other monocot plant species. Quantitative gene expression analysis was carried out to elucidate the transcript profiles of the oil palm HGGT and HPT in different oil palm tissues and at different developmental stages of the mesocarp. The HPT was constitutively expressed in all analyzed tissues except in 15 w.a.a kernel whereas oil palm HGGT showed preferential expression in mesocarp and kernel tissues. However, HPT was highly expressed at the fruit ripening stage of 17 w.a.a mesocarp when active oil deposition occurs. Genome-walking PCR successfully amplified the promoter regions of HGGT and HPT from *E. guineensis*. Computational analysis using PlantCare and PLACE databases revealed several *cis*-regulatory elements including phytohormone-responsive, light-responsive and abiotic factor-responsive elements which may be involved in coordinating expression of both genes. Taken together, this study provides useful information about important features of the cDNA and promoter sequences as well as an insight into the transcriptional regulation of these key vitamin E genes for future genetic improvement efforts.

1. Introduction

Tocochromanols, commonly known as vitamin E, which are exclusively synthesized by photosynthetic organisms, are essential fat soluble nutrients in the human diet (Dörmann, 2003). The eight structurally related forms of tocopherols and tocotrienols (α-, β-, γ-, and δ-tocopherols and α-, β-, γ-, and δ-tocotrienols) together constitute the content of vitamin E. Tocotrienols structurally differ from tocopherols by the presence of three *trans* double bonds in the hydrocarbon tail while the aliphatic tail of tocopherols is fully saturated (Kamal-Eldin and Appelqvist, 1996). Vitamin E is well accepted as nature's most effective chain-breaking antioxidant that prevents the degradation of polyunsaturated fatty acids in membrane by reactive oxygen species (Dörmann, 2003). It was suggested that tocotrienols possess powerful neuroprotective, antioxidant, anti-cancer and cholesterol lowering properties which are not often exhibited by tocopherols (Sen et al., 2006). The pharmaceutical properties demonstrated by tocotrienols would make tocotrienol-accumulating crops potentially very beneficial for human health.

Tocochromanols also play important roles in plants, such as protection of chloroplasts from photooxidative damage (Munné-Bosch and Alegre, 2002). As a strong antioxidant, vitamin E is beneficial in

Abbreviations: BLAST, basic local alignment search tool; bp, base pair; cDNA, complementary deoxyribonucleic acid; Ct, threshold cycle; CTAB, hexadecyl (or cetyl) trimethyl ammonium bromide; Da, Dalton; DNA, deoxyribonucleic acid; DNase I, deoxyribonuclease I; dNTP, deoxynucleoside triphosphate; GA, gibberellin; GGDP, geranylgeranyldiphosphate; GSP, gene-specific primer; HGA, homogentisic acid; HGGT, homogentisate geranylgeranyl transferase; HPT, homogentisate phytyltransferase; kb, kilobase; min, minutes; mRNA, messenger ribonucleic acid; NCBI, National Center for Biotechnology Information; ORF, open reading frame; PCR, polymerase chain reaction; PDP, phytyldiphosphate; pI, isoelectric point; PrDP, prenyldiphosphate; RACE, Rapid Amplification of cDNA End; RNA, ribonucleic acid; RT-PCR, reverse transcription PCR; sec, seconds; TSS, transcription start site; UTR, untranslated region; w.a.a, week after anthesis.

* Corresponding author at: Laboratory of Plantation Crops, Institute of Tropical Agriculture, Universiti Putra Malaysia, 43400, UPM, Serdang, Selangor, Malaysia.
E-mail address: snaa@upm.edu.my (S.N.A. Abdullah).

http://dx.doi.org/10.1016/j.plgene.2016.01.003

maintaining oxidative stability of vegetable oils and also in enhancing nutritional value of crop plants for human diets and livestock feeds (Hunter and Cahoon, 2007). Studies have shown that tocopherols occur ubiquitously in various plant tissues especially in leaves and seeds of most dicots and are also found in photosynthetic microbes such as *Synechocystis*. Tocotrienols, in contrast, are rarely found in vegetative tissues of plants but are the exclusive form of vitamin E present in the seed endosperm of most monocots, including agronomically important cereal grains such as wheat, rice and barley (Kamal-Eldin and Appelqvist, 1996). Soybean oil and palm oil are the main sources for commercially produced tocopherols and tocotrienols, respectively. These lipid-soluble products are found as vegetable oils components (Hunter and Cahoon, 2007). Palm oil is unique among the vegetable oils because it contains high level of tocotrienols (α, γ, and δ), which together account for about 70% of the total vitamin E content. Tocochromanols are found in the crude palm oil at levels of 600–1000 ppm (Choo et al., 2004). Palm vitamin E has been extensively studied for its nutritional and health properties, attributed largely to its high tocotrienol content (Sen et al., 2006).

The biosynthesis pathway of tocochromanols can be divided into four stages starting from the formation of the homogentisic acid head group. HGA is the first substrate needed in the biosynthesis of vitamin E by supplying the aromatic ring of the chromanol head group and it is derived from the shikimate pathway. The second required substrate is a prenyldiphosphate (PrDP), either phytyldiphosphate (PDP) or geranylgeranyldiphosphate (GGDP). These prenyl groups are supplied by the non-mevalonate pathway. The second stage is the transfer of a prenyl group to HGA which is the branch point for commitment to either tocopherols or tocotrienols and is a potential key regulatory step. Homogentisate geranylgeranyl transferase (HGGT) is generally believed to be the key enzyme that catalyzes the condensation of HGA and GGDP to generate tocotrienols due to the structural similarity of the tocotrienols side chain and GGDP. Meanwhile the committed step in the formation of tocopherols is the condensation of HGA with PDP through an enzyme designated as homogentisate phytyltransferase (HPT). The last two stages are the subsequent ring cyclization and methylation reactions to generate all eight forms of tocochromanol (Hunter and Cahoon, 2007; Collakova and DellaPenna, 2001; Bramley et al., 2000).

HPT was first identified from *Arabidopsis* and *Synechocystis* sp. PCC 6803 and it was further demonstrated that the level of seed tocopherols could be elevated up to 2-fold by seed-specific expression of the *Arabidopsis HPT1* in *Arabidopsis* (Collakova and DellaPenna, 2001; Savidge et al., 2002). In 2003, Cahoon and coworkers isolated the gene encoding HGGT from three monocot species (*Oryza* sp., *Hordeum* sp. and *Triticum* sp.) and over-expressed it in corn seeds driven by a strong embryo-specific promoter. This was resulted in the increment of tocotrienol content by 20-fold and also enhanced the vitamin E content by 6-fold. Oil palm may offer great advantages compared to other plants for genetic manipulation of vitamin E. However, the knowledge on biosynthesis pathway of vitamin E in oil palm as one of the basic prerequisites for genetic manipulation is quite limited. This will definitely become an impediment to improve oil palm vitamin E content through genetic engineering, development of molecular markers for cross species breeding as well as the other biotechnological approaches. The present study reports on characterization of the cDNA and promoter sequences of the HGGT and HPT in two important oil palm species. Quantitative gene expression analysis in various oil palm tissues was also performed to obtain an insight into the transcriptional regulation of these key vitamin E genes.

2. Materials and methods

2.1. Plant materials

The mesocarp and kernel tissues of two oil palm species, *Elaeis guineensis* and *Elaeis oleifera* were harvested at various weeks after anthesis (7-, 10-, 12-, 15-, 17- and 19-w.a.a) and sliced into small pieces, snap frozen in liquid nitrogen and stored at −80 °C. All fruit bunches were obtained from tagged fruits in the same field. Unfolded spear leaves and young roots from *E. guineensis* seedlings were harvested and frozen at −80 °C.

2.2. Total RNA isolation and cDNA synthesis

Total RNA extraction was carried out using the modified method of Prescott and Martin (1987) and all of the extracted RNAs were treated with DNaseI (Fermentas, USA). The amount of total RNA was quantified using a Nanodrop spectrophotometer (ND-1000, Thermo Scientific, USA). The integrity of the total RNA was checked using 1% (*w*/*v*) agarose gel electrophoresis and separated at 75 V in 1× TAE buffer. Mesocarp Poly(A)$^+$ RNAs of both oil palm species at 17 w.a.a were purified using the Absolutely mRNA™ Purification Kit (Stratagene, CA, USA) and used as the template to synthesize the cDNA pool for gene isolation. First-strand cDNA was synthesized using Superscript III First-Strand Synthesis Kit (Invitrogen, USA) and 150 ng of each mRNA template in a final volume of 20 μl.

2.3. Isolation of the full length HGGT and HPT cDNAs

RT-PCR amplification was conducted using Advantage cDNA polymerase mix (Clontech, USA) and a pair of sense and antisense degenerate oligonucleotides which designed based on the conserved nucleotide region of known monocot HGGT and HPT protein sequences (Table 1). A similar forward degenerate primer (F5-2) was used to amplify both genes in oil palm. The PCR reaction was performed in a T-Professional Basic Thermal Cycler (Gradient) (Biometra, Germany) using an initial denaturation step at 94 °C for 5 min, followed by 35 cycles at 94 °C for 30 s, 58 °C for 30 s and 72 °C for 1 min and final extension at 72 °C for 5 min. The desired PCR fragment was cloned into pGEM-T Easy vector (Promega, USA) and sequenced. A homology search was conducted through NCBI database using BLASTX and BLASTP algorithms (http://www.ncbi.nlm.nih.gov/). Gene-specific primers (GSPs) (Table 1) were designed based on the isolated partial cDNAs for amplification of 5′- and 3′-ends of the oil palm HGGT and HPT genes by RACE PCR. The 5′- and 3′-RACE ready cDNA synthesis was carried out using SMARTer™ RACE cDNA Amplification Kit (Clontech, USA). Both 5′- and 3′-sequences were combined to produce the full length gene sequences. For long-distance PCR (LD-PCR), new gene specific primers were designed based on the identified 5′- and 3′-end sequences covering the open reading frame (ORF) of the oil palm HGGT and HPT genes. The primers were designated as follows: (1) *E. guineensis*; HGGT forward primer EEF1 (5′-GGTTGCAACGCATCATTAAGATATTG-3′), HGGT reverse primer EER1 (5′-GGAGAATTTTCAAAAGTCTATGATGATG-3′). HPT forward primer EGHPTF (5′-GGTTCTCTTACCAGGTCCCTGCAAGATCG-3′), HPT reverse primer EOHPTR (5′-CATAATGGATGATCGAAGTGGCACAG-3′); (2) *E. oleifera*; HGGT forward primer EEF2 (5′-GCTTCAAAATGACC TGCTTTTTATGG-3′), HGGT reverse primer EER1, HPT forward primer EOHPTF (5′-GGTCTCTTACCAGATCCCTGCAAGATCG-3′) and HPT reverse primer as EOHPTR.

2.4. Full length cDNA sequence analysis

Initially, ORFs for all four cDNAs were predicted using the ORF Finder program (http://www.ncbi.nlm.nih.gov) and subjected to BLASTX and BLASTP analyses. The sequence alignment and characterization of the deduced proteins were carried out using different tools available in SDSC Biology Workbench (http://seqtool.sdsc.edu). Multiple sequence alignment was done using ClustalW pairwise alignment algorithm. To find out general features of the cDNA sequence including amino acid composition and isoelectric point (pI), the analysis was performed using ExPASy Proteomics tools (http://cn.expasy.org/tools/protscale.html). Physical and chemical characteristics of all deduced amino acid