ICBAA2017-8

GENOME MINING FOR PRODUCTION OF NUTRITIONALLY RICH PALM OIL

Siti Nor Akmar Abdullah[1], Sulaiman Rufai Babura[1], Kong Sze Ling[1], Umaiyal Munusamy[1], Marhalil Marjuni[2] and Mohd Din Amiruddin[2]

[1]Institute of Plantation Studies, Universiti Putra Malaysia, 43400 UPM, Serdang, Selangor, Malaysia;
[2]Malaysian Palm Oil Board, No. 6, Persiaran Institusi, Bandar Baru Bangi, 43000 Kajang, Selangor, Malaysia
snaa@upm.edu.my

Abstract: The vitamin E in palm oil contains high proportion of tocotrienol (70% tocotrienols and 30% tocopherols) while most other vegetable oils contain only tocopherols. Both tocopherols and tocotrienols are strong chain-breaking antioxidants. The interest in tocotrienols escalated when the other medicinal properties such as neuroprotective, anticancer and cholestrol-lowering were discovered. This report is based on studies carried out on the different vitamin E biosynthetic pathway genes in order to understand their molecular regulatory mechanism in the two important oil palm species, *E. guineensis* and *E. oleifera*. The expression profiles of the different genes correlate with oil accumulation period in the fruit mesocarp suggesting the pivotal role of vitamin E in maintaining oxidative stability of the oil. The presence of a set of common regulatory elements in the promoters of these genes indicate that the pathway is coordinately control by endogenous and environmental cues at the transcriptionl level. PCR-based SNP markers for routine screening of the oil palm breeding materials were developed using mismatch primers. Some of the innovative approaches used in our laboratory for functional studies to know the effects of sequence variants on the function of the vitamin E biosynthetic genes are described.

Keywords**:** Oil palm, vitamin E, homogentisate geranylgeranyl transferase, genetic variation, SNP markers

INTRODUCTION

Palm oil is a nutritious edible oil endowed with a variety of health promoting micronutrients including carotenoids, sterols and vitamin E. Tocopherol and tocotrienol commonly known as vitamin E consist of four different forms designated as α-, β-, γ- and δ-. These lipid soluble vitamins are widely used in animal feed, nutritional and dietary supplements, food, beverages and cosmetics. Vitamin E is nature's most effective chain-breaking antioxidant (Khanna *et al.* 2003). Tocotrienol in addition is valuable for its cholesterol lowering and anti-cancer properties while α-tocotrienol can prevent inducible neurodegeneration (Qureshi *et al.* 2002; Das *et al*. 2005). Thus, tocotrienol accumulating crops are potentially very beneficial for human health with oil palm offering the greatest potential (Babura et al., 2017). Oil palm is nature's richest source of vitamin E where the level in crude palm oil is around 600 ppm consisting of 70% tocotrienols and 30% tocopherols. Due to the narrow genetic base, there is high limitation in using current planting materials for improving tocotrienol content in oil palm. However, the genetic materials in the *E. guineensis* and *E. oleifera* germplasm collections of the Malaysian Palm Oil Board (MPOB) could potentially be the source of alleles for vitamin E improvement. For example, it was demonstrated that there is a high level of variation in the vitamin E content ranging from 300 – 1600 ppm in the *E. guineensis* germplasm collection from Tanzania and Angola. While the beneficial traits of *E. oleifera* include shorter height increment and high content of vitamin E up to 1500 ppm (Wahid et al. 2005).

Our interest is on charaterisation of genes encoding enzymes of the vitamin E biosynthetic pathway as well as other associated regulatory proteins from both *E. guineensis* and *E. oleifera* and to identify sequence variation for molecular marker development. Our work focuses on genes encoding homogentisate geranylgeranyl transferase (HGGT) and homogentisate phtyl transferase (HPT) that catalyses the first committed step of tocotrienol and tocopherol pathway, respectively. However, the genes that encode enzyme in subsequent steps shared by both tocopherols and tocotrienols involving a cyclase and methyltransferases for producing the different forms of tocopherols and tocotrienols were also studied to give a better understanding on the biochemical and molecular regulatory mechanisms of this biosynthetic pathway.

METHODS

Sequencing of the targeted vitamin E genes of interest ranging fom 6 – 8 kb from *E. guineensis* germplam materials with varying vitamin E content was carried out using PacBio long read sequencing. The expression profiling in the different oil palm tissues at the different developmental stages in the fruit mesoacarp was performed using quantitative real time PCR (qPCR) analysis normalized against at least three housekeeping genes. The sequences were mapped to the genome sequence at GenBank as reference and multiple sequence alignment was done for the identification of sequence variants in both the promoter and the different regions including 5' and 3'-UTR, exons and introns. Promoter motifs were determined using publicly accessible promoter motif databases. Screening of the different accessions in the MPOB germplasm collection was performed using mismatch PCR primers based on SNP positions enabling scoring the absence or presence of bands. Variants that may be functionally important discovered in the germplasm materials was introduced into the gene in the commercial oil palm variety by site directed mutagenesis and the function of the mutated gene was compared to the unmodified one in transgenic model plant, *Arabidopsis thaliana*. The expression profiling in the transgenic lines was carried out using qPCR analysis before performing HPLC analysis to see the effects on vitamin E content and composition.

RESULTS AND DISCUSSION

The size of the whole *HPT* and *HGGT* genes range between 6 – 8 kb as they are intron rich. The expression profiles of the different genes correlate with oil accumulation period in the fruit mesocarp suggesting the pivotal role of vitamin E in maintaining oxidative stability of the oil (Kong et al., 2016). The study focuses on the region from approximately 1 kb upstream of the translation start site harbouring the promoter until the 3'-UTR regions. The sequences in the oil palm genome sequence in GenBank was used as reference in carrying out the multiple sequence alignment. It was found that the majority of the sequence variants between genes from different accession with varying vitamin E content occured within the intron. This is about ten times more than the variablity found within the promoter and coding regions. This is illustrated in Figure 1 for the number of variants identified when comparing the *HGGT* from different *E. guineensis* germplasm materials (Figure 1).

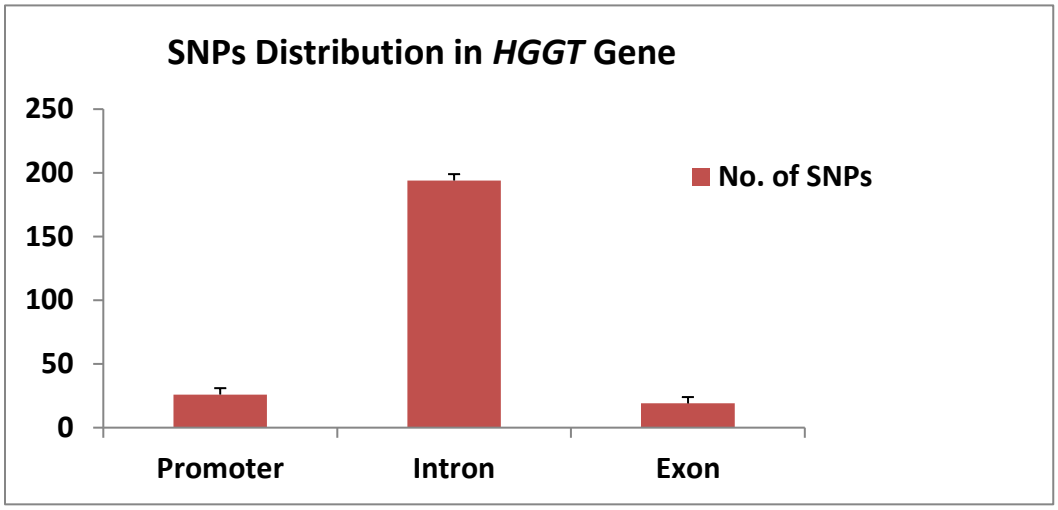

Figure 1: The number of variants identified when comparing the *HGGT* from different *E. guineensis* germplasm materials.

Comparison of the promoter sequence between *EgHPT* and *EgHGGT* showed the presence of common regulatory motifs that may be important in coordinating expresion of the genes involved in palm vitamin E biosynthesis. Examples include motif for response to dehydration, heat and the phytohormone gibberellin (Table 1).

Table 1: *Cis*-acting elements found in common in the promoters of *E. guineensis HPT* and *HGGT* genes.

Elements	***Cis*-element**	**Sequences**	**Position from TSS in p*EgHPT***	**Position from TSS in p*EgHGGT***
Gibberellins Response	Pyrimidine Box	CCTTTT	44(+), 874(+)	19 (-)
	GAREAT	TAACAAR	924(+)	654(+)
Dehydration Response	MYBCORE	CNGTTR	13(+), 854(+), 459(-), 689(-), 770(-)	154(+), 151(-)
	MYCCONSENSUS	CANNTG	7(+), 13(+), 35(+), 501(+), 675(+), 7(-), 13(-), 35(-), 501(-), 675(-)	48(+), 220(+), 308(+), 855(+), 48(-), 220(-), 308(-), 855(-)
Heat Stress	HSE	AAAAAATTTC	887(-)	688(+), 719(+)

Initially, species- specific sequence variant in the promoter of *HGGT* was used for molecular marker development to avoid the difficulty of designing primers in the coding region as the genes have many introns. The motif, ACGT, is an important phytohormone responsive motif (Mehrotra and Mehrotra 2010) Using SNP that is located within this promoter motif, a mismatch primer was produced. Application of this primer on different populations of *E. guineensis* and *E. oleifera* was able to differentiate the two species with 100% accuracy based on simple visualisation of the presence or absence of a band. The same technique is currently being used to screen the *E. guineensis* germplasm materials based on vitamin E content.

The identified potentially functionally important sequnce variant such as that occuring within promoter motif and those resulting in the non-conservative substitution of amino acid in the coding region was introduced into the sequence of the commercial variety through site-directed mutagenesis to see its effects on gene functionality. The functional studies were performed in transgenic Arabidopsis harbouring normal and mutated versions of the genes. Fig. 2 shows the selection for the transgenic lines bearing the modified genes to compare with lines harbouring the unmodified gene from the commercial variety. Recent results showed that the different versions of the gene were successfully expressed based on qPCR analysis using cDNA from leaf tissues of the different T3 homozygous transgenic lines. HPLC analysis to determine the changes in vitamin E content and composition is in progress.

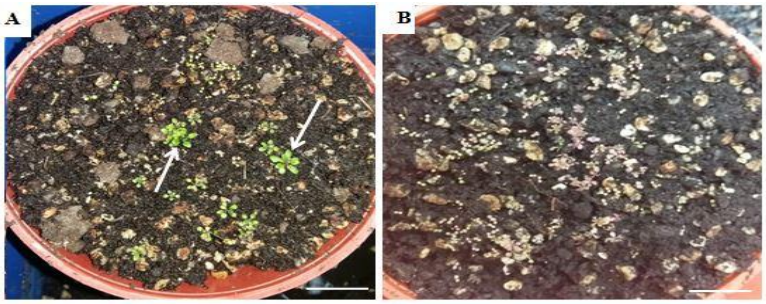

Figure 2: Selection for the transgenic lines bearing the mutated *E. guineensis* vitamin E biosynthetic gene to see the effects of sequence variant on vitamin E composition and content.

CONCLUSIONS

This paper reports on the strategy used to determine functionally important sequence variant from the sequences of vitamin E biosynthetic genes of different accessions in MPOB germplasm materials. The variant form the basis for developing SNP markers for screening the different population for development of a reliable marker to screen for high vitamin E oil palm planting materials.

REFERENCES

Babura, S.R., Abdullah, S.N.A., Khazaai, H. (2017). Advances in Genetic Improvement for Tocotrienol Production: A Review. Journal of Nutritional Science and Vitaminology 63 (4), 215-221.

Das, S., Powell, S.R., Wang, P., Divald, A., Nesaretnam, K., Tosaki, A., Das, D.K. (2005). Cardioprotection with palm tocotrienol: antioxidant activity of tocotrienol is linked with its ability to stabilize proteasomes. Am J Physiol 289: 361-367.

Khanna, S., Roy, S., Ryu, H., Bahadduri, P., Swaan, P.W., Ratan, R.R, Sen, C.K. (2003) Molecular basis of Vitamin E action. Journal of Biological Chemistry 278 (44):43508-43515.

Kong, S.L., Abdullah, S.N.A., Ho, C.l., Amiruddin, M.D. (2016). Molecular cloning, gene expression profiling and in silico sequence analysis of vitamin E biosynthetic genes from the oil palm**.** Plant Gene 5, 100-108.

Mehrotra, R., Mehrotra, S. (2010) Promoter activation by ACGT in response to salicylic and abscisic 280 acids is differentially regulated by the spacing between two copies of the motif. Journal of Plant Physiology 167 (14):1214-1218.

Qureshi, A.A., Sami, S.A., Salser, W.A., Khan, F.A. (2002) Dose-dependent suppression of serum cholesterol by tocotrienol-rich fraction (TRF25) of rice bran in hypercholesterolemic humans. Atherosclerosis 161 (1):199-207.

Wahid, M.B., Abdullah, S.N.A., Henson, I.E. (2005). Oil Palm. Plant Production Science 8 (3), 288-297.