GMR Genetics and Molecular Research
Online Journal
ISSN:1676-5680

Cloning and polymorphism analysis of the 2-methyl-6-phytyl-1,4-benzoquinol methyltransferase gene (*VTE3*) in *Arachis hypogaea*, *A. duranensis*, and *A. ipaënsis*

F.Z. Liu, A.Q. Guo and Y.S. Wan

State Key Laboratory of Crop Biology,
Shandong Key Laboratory of Crop Biology, College of Agronomy,
Shandong Agricultural University, Taian, Shandong, China

Corresponding author: Y.S. Wan
E-mail: liufz@sdau.edu.cn / yswan@sdau.edu.cn

Genet. Mol. Res. 12 (2): 1859-1871 (2013)
Received February 7, 2012
Accepted August 29, 2012
Published January 4, 2013
DOI http://dx.doi.org/10.4238/2013.January.4.5

ABSTRACT. One of the critical enzymes involved in vitamin E biosynthesis in plants is 2-methyl-6-phytyl-1,4-benzoquinol methyltransferase (MPBQ MT). The full-length *VTE3* cDNA (designated *rVTE3-1* and *-2*) encoding MPBQ MT and the full-length DNA of *VTE3* (designated *gVTE3-1* and *-2*) were isolated from cultivated peanuts (*Arachis hypogaea*). The full-length DNA of *VTE3* (designated *gVTE3-A* and *-B*) was isolated from the wild groundnut species *A. duranensis* (A-genome) and *A. ipaënsis* (B-genome), and polymorphism analysis of *VTE3* was performed. The results demonstrated that *rVTE3-1* and *-2* both have a DNA sequence that is 1059 bp long and encodes 351 amino acids; the homology of the 2 amino acid sequences was 98.6%. The *gVTE3-1* of cultivated Fenghua 2 peanut samples was 2710 bp long, with 3 introns located at 44-163, 772-1295, and 1603-2437 bp, and the Fenghua 2 *gVTE3-2* was 2706 bp long, with 3 introns located at

44-169, 778-1291, and 1599-2433 bp. The homology for *gVTE3-1* and *-2* across 13 cultivated peanut samples was 99.9 and 100%, respectively. *gVTE3-1* and *-2* were from A- and B-genome, respectively, with 96.6% homology between the 2 sequences. The present study demonstrated that abundant polymorphisms were present in the VTE3 genes from different genomes. Additionally, polymorphisms were observed in the *gVTE3-1* alleles of the 13 cultivars and wild *A. duranensis* species but not in the *gVTE3-2* alleles of the 13 cultivars and wild *A. ipaënsis* species.

Key words: Gene cloning; *Arachis* sections; Vitamin E; 2-Methyl-6-phytyl-1,4-benzoquinone methyltransferase

INTRODUCTION

The antioxidant vitamin E plays an important physiological role in plants, animals, and human beings. There are 8 different forms of vitamin E (4 tocopherols: α, β, γ, and δ; 4 tocotrienols: α, β, γ, and δ), and it occurs mainly in the form of tocopherols in dicotyledonous plants (Rimbach et al., 2002; DellaPenna and Pogson, 2006). Different types of tocopherols have dissimilar activities. Among the 4 tocopherols, α-tocopherol has the highest activity, whereas β-, γ-, and δ-tocopherols achieve only 50, 10, and 3% of the activity level of α-tocopherol, respectively. Given that humans and animals both possess significant levels of proteins with a high affinity for α-tocopherol, α-tocopherol absorption and utilization is more significant than that of β-, γ-, and δ-tocopherols in these organisms (Brigelius-Flohe and Traber, 1999).

Natural vitamin E can only be synthesized by photosynthetic bacteria and green plants, and humans and animals depend solely on their dietary intake for vitamin E (Cheng et al., 2003). Plant tocopherols exist primarily in non-green plant tissues, such as the oil of many crop seeds; however, high-activity α-tocopherol has a much lower proportion of tocopherols when compared to low-activity tocopherols (Grusak and DellaPenna, 1999). As one of the critical enzymes involved in plant vitamin E biosynthesis, 2-methyl-6-phytyl-1,4-benzoquinol methyltransferase (MPBQ MT) can catalyze MPBQ to generate 2,3-dimethyl-6-phytyl-1,4-benzoquinone (DMPBQ). MPBQ and DMPBQ, however, are cyclized by tocopherol cyclase to yield δ- and γ-tocopherol, respectively, which are subsequently further converted to β- and α-tocopherol by γ-tocopherol methyltransferase (γ-TMT) (Shintani and DellaPenna, 1998; Porfirova et al., 2002; Shintani et al., 2002).

The *Arabidopsis* genes *VTE3* and *VTE4*, which encode MPBQ MT and γ-TMT, were transferred to soybeans together, and a significant increase of 85% was observed in α-tocopherol expression in soybean seeds (Van Eenennaam et al., 2003). The *Arabidopsis* genes hydroxyphenylpyruvate dioxygenase and *VTE3* were transferred to maize together, and a 3-fold increase in vitamin E expression was observed in corn grain (Naqvi et al., 2011). Peanuts are one of the most important commercial oil crops. Some key enzymes (such as MPBQ MT, tocopherol cyclase, and γ-TMT) can affect vitamin E biosynthesis; thus, they can be used to increase the vitamin E content and proportion of α-tocopherol in peanut-based oil crops by improving expression of the enzyme gene through the transference to peanut.

The genus *Arachis* originated from South America and contains 80 documented species assembled into 9 taxonomic groups according to morphology, geographic distribution, and sexual compatibility (Moretzsohn et al., 2009). The *Arachis* section is the largest, comprising 29 wild diploid species and the tetraploid species *A. hypogaea* and *A. monticola* (Moretzsohn et al., 2009). The cultivated peanut (*A. hypogaea*) is an allotetraploid (2n = 4x = 40) with 2 genome types (A- and B-genome), and thus can be considered to have originated from 2 distinct wild diploid species of *Arachis* (Kochert et al., 1996). To date, most researchers support the hypothesis that the cultivated peanut originated via hybridization of *A. duranensis* (A-genome) and *A. ipaënsis* (B-genome), followed by a rare spontaneous duplication of chromosomes (Halward et al., 1991; Kochert et al., 1996; Fávero et al., 2006; Seijo et al., 2004, 2007). According to its characteristics, the cultivated peanut is divided into the subspecies *A. hypogaea* ssp *fastigiata* and *A. hypogaea* ssp *hypogaea*. The former comprises 2 variants (*A. hypogaea* ssp *fastigiata* var. *fastigiata* and *A. hypogaea* ssp *fastigiata* var. *vulgaris*), as does the latter (*A. hypogaea* ssp *hypogaea* var. *hypogaea* and *A. hypogaea* ssp *hypogaea* var. *hirsuta*). Peanut cultivars are classified into Runner, Virginia, Spanish, and Valencia market types (Hong et al., 2010).

Analysis of random amplified polymorphic DNA and inter-simple sequence repeat markers on peanut cultivars has revealed a low level of polymorphisms and a narrow genetic base, suggesting that variant cultivars share the same wild ancestors (Raina et al., 2001). Studies using amplified fragment length polymorphism (Milla et al., 2005) and fluorescence *in situ* hybridization (Seijo et al., 2004) methods confirmed that *A. hypogaea* variant cultivars share the same ancestral origin, and the present study concluded that *A. duranensis* and *A. ipaënsis* are likely to be the ancestors of contemporary cultivated species. Singh (1988) synthesized amphidiploid peanuts, hybridized them with different cultivars, and studied the cell morphology and chromosome synapsis. Analysis of isozymes in different cultivars (Lu and Pickersgill, 1993) demonstrated that differences exist between the 2 peanut subspecies. Both studies (Singh, 1988; Lu and Pickersgill, 1993) indicated that the 2 cultivated subspecies of peanuts possess distinct and independent evolutionary systems. While most researchers believe that different types of cultivars share similar ancestry, these findings confirm the suggestion put forth by several contemporary researchers that the 2 subspecies have independent evolutionary systems. By comparing homologous genes, the common ancestry of different peanut subspecies can be inferred at the molecular level.

To our knowledge, although *VTE3* plays an important role in vitamin E biosynthesis in peanuts, little is known about the conservation and diversity of *VTE3* genes in cultivars and wild species. In addition, and most importantly, the wild ancestry of various peanut cultivars is questionable, and different researchers have reached diverse conclusions regarding this issue. In the present study, we expect to reveal the molecular characteristics of *VTE3* and investigate peanut *VTE3* polymorphism by cloning *VTE3* genes from cultivated and wild peanut species. The present study not only provides new evidence for the determination of the ancestral origin of different peanut subspecies, but also lays the foundation for enhancing vitamin E production in plants, especially in commercial peanut crops, using genetic engineering.

MATERIAL AND METHODS

Material

The resources for *VTE3* collection and analysis were selected from different plant

material genotypes, comprising 13 distinct peanut cultivars (Table 1) and 2 wild species, i.e., *A. duranensis* and *A. ipaënsis*.

In silico cloning of peanut *VTE3*

To predict the full-length cDNA sequence of *VTE3* in peanuts, *Arabidopsis thaliana VTE3* gene sequences were obtained from the National Center for Biotechnology (NCBI) database (accession No. NM_116206.3) and used as an information probe against the peanut database for expressed sequence tags (dbEST) by the Basic Local Alignment Search Tool (BLAST) analysis to retrieve homologous ESTs, resulting in a single homologous sequence to be recorded and used as the seed sequence (accession No. ES764300.1). Next, the determined seed sequence was used in BLAST against the same dbEST, revealing partially overlapping ESTs, and these sequences were assembled into a longer new sequence. Subsequently, the above mentioned 2 steps were repeated until the new sequence could not be extended further. Finally, with an open reading frame (ORF) finder programming application, the final assembled sequence was established and its encoding amino acid sequence was determined. These results were compared with the corresponding *A. thaliana VTE3* nucleotide and amino acid sequences documented previously. Based on the computer-simulated cloned sequence of peanut *VTE3*, the primer sequences F1 (5'-TTCGTGCTCCTCTCTCATCC-3') and R1 (5'-TGACGTTGGGCAAACTGTAT-3') were designed. These primers were synthesized by Shanghai Biological Engineering Co. Ltd. (China).

DNA extraction of cultivated and wild peanut species

Young leaves were collected from the cultivated and wild peanut species, and subsequent total DNA extraction was performed using the hexadecylmethylammonium bromide method (Rogers and Bendich, 1985).

RNA extraction and cDNA synthesis

Thirty-five days after being pegged into the soil, young peanut pods were collected and preserved by freezing in liquid nitrogen (-70°C) for subsequent RNA extraction. The extraction method used was the TRIzol Reagent Kit (Beijing TransGen Biotech Co. Ltd., China). Using the EasyScript First-Strand cDNA Synthesis Kit (Beijing TransGen Biotech Co. Ltd.), RNA sequences were reverse-transcribed to generate cDNA.

PCR amplification of *VTE3* from cultivated and wild peanut species

VTE3 cDNA transcribed from RNA sequences collected from the Fenghua 2, Lanna 1, and Lipudahuasheng cultivars were PCR-amplified using the F1 and R1 primers. The amplification process was as follows: 94°C for 5 min; 94°C for 30 s; 55°C for 30 s; 72°C for 1.5 min; 35 cycles; and 72°C for 10 min. *VTE3* DNA from the 13 cultivars and 2 wild species genomes were PCR-amplified using the same F1 and R1 primers. The amplification protocol was as follows: 94°C for 5 min; 94°C for 30 s; 55°C for 30 s; 72°C for 2.5 min; 35 cycles; and 72°C for 10 min.

PCR product cloning and sequencing

A TIANgel™ Gel Extraction Kit (Beijing TransGen Biotech Co. Ltd.) was used to recycle PCR products and link to the cloning vector pEASY-T1 (Beijing TransGen Biotech Co. Ltd.). The complex was then transferred into *Escherichia coli* DH5α via overnight culture on the plates and identified by colony PCR. Finally, target fragments were sent to the Beijing Genomics Institute for sequencing. To guarantee experimental reproducibility, between 6 and 10 clones from each species were selected for sequencing.

Analysis of *VTE3* DNA and amino acid sequences in cultivated and wild peanut species

Sequence homology was analyzed by BLAST search and the DNAMAN software. ClustalX and ClustalW were used to align the DNA sequences as well as the amino acid sequences, and the PHYLIP software was used to construct a *VTE3* phylogenetic tree based on this alignment. To investigate *VTE3* in a comprehensive manner, bioinformatic analyses were performed as follows: various physicochemical properties of *VTE3*, including molecular weight and protein theoretical isoelectric point (pI), were assessed by ProtParam (http://www.expasy.ch/tools/protparam.html); transmembrane helices in VTE3 proteins were predicted by TMHMM (http://www.cbs.dtu.dk/services/TMHMM/); the VTE3 protein secondary structure was analyzed by the SOPMA software (http://npsa-pbil.ibcp.fr/cgi-bin/npsa_automat.pl?page=/NPSA/npsa_sopma.html); the subcellular localization of VTE3 proteins was predicted by WoLF PSORT (http://wolfpsort.org/); and finally, conserved domains in VTE3 proteins were searched by the CDD software (http://www.ncbi.nlm.nih.gov/Structure/cdd/wrpsb.cgi).

RESULTS

Cloning of cultivated and wild peanut species *VTE3*

The full-length of the assembled EST (1396 bp) in peanuts has an ORF (1056 bp) and 71.8% homology with the *A. thaliana VTE3* coding region, and the homology of the corresponding amino acid sequences between the 2 species was approximately 77%. The results of this study therefore suggest that the assembled sequence is highly likely to be peanut *VTE3*.

We obtained 1.2-kb fragments by PCR amplification of reverse-transcribed cDNA with the F1 and R1 primers; the fragments were linked to pEASY-T1 vectors for sequencing. The sequencing results revealed 2 different sequences from the products of 3 cultivars, which were designated as *rVTE3-1* and *-2* (Table 1). PCR amplification of extracted DNA using primers F1 and R1 produced 2.7-kb fragments that were linked to pEASY-T1 vectors for sequencing. The sequencing results revealed 2 different sequences from the DNA amplification products of 13 cultivars; the sequences were designated as *gVTE3-1* and *-2* (Table 1). PCR amplification of extracted DNA using primers F1 and R1 revealed that the other 2 sequences were produced from *A. duranensis* and *A. ipaënsis*, respectively, and the sequences were designated as *gVTE3-A* and *-B*.

Table 1. Germplasm accessions of cultivated peanuts and the symbols of DNA/cDNA sequences.

Subspecies	Varieties	Accessions	DNA sequence symbol	cDNA sequence symbol
fastigiata	var. *vulgaris*	Fenghua 2	*FhgVTE3-1*, *FhgVTE3-2*	*FhrVTE3-1*, *FhrVTE3-2*
fastigiata	var. *vulgaris*	Shanhua 8	*Sh8gVTE3-1*, *Sh8gVTE3-2*	-
fastigiata	var. *vulgaris*	Yuanza 9102	*YzgVTE3-1*, *YzgVTE3-2*	-
fastigiata	var. *fastigiata*	Krapt.st.16	*KtgVTE3-1*, *KtgVTE3-2*	-
fastigiata	var. *fastigiata*	ICG6848	*IggVTE3-1*, *IggVTE3-2*	-
fastigiata	irregular type	Shanhua 7	*Sh7gVTE3-1*, *Sh7gVTE3-2*	-
fastigiata	irregular type	Shanhua 9	*Sh9gVTE3-1*, *Sh9gVTE3-2*	-
hypogaea	var. *hypogaea*	Lanna 1	*LngVTE3-1*, *LngVTE3-2*	*LnrVTE3-1*, *LnrVTE3-2*
hypogaea	var. *hypogaea*	Nongda 818	*NdgVTE3-1*, *NdgVTE3-2*	-
hypogaea	var. *hypogaea*	ICGV86699	*IvgVTE3-1*, *IvgVTE3-2*	-
hypogaea	var. *hypogaea*	E1	*E1gVTE3-1*, *E1gVTE3-2*	-
hypogaea	var. *hirsuta*	Lipudahuasheng	*LpgVTE3-1*, *LpgVTE3-2*	*LprVTE3-1*, *LprVTE3-2*
hypogaea	var. *hirsuta*	Rugaoxiyangsheng	*RggVTE3-1*, *RggVTE3-2*	-

Fenghua 2 *VTE3* nucleotide sequence analysis

The *FhrVTE3-1* and *-2* sequences shared 99.0% homology, and the 2 sequences shared 98.7 and 97.7% homology, respectively, with the assembled sequences produced by the *in silico* cloning. Both sequences had ORFs of the same length encoding 351 amino acids; the start and stop codons were located at 59 and 1114 bp, respectively. The sequence near the ATG start codon adhered to the Kozak rule (AXXATGG), and a polyadenylation signal (AATAAA) was located 25 bp downstream from the stop codon (Figure 1). *FhrVTE3-1* and *-2* both had an encoding length of 1059 bp, the homology of the 2 sequences was 97.8%, and 8 single nucleotide polymorphism (SNP) sites existed within the sequences. The protein sequences shared 98.6% homology with 5 amino acid differences.

The *FhgVTE3-1* and *-2* sequences shared 96.6% homology. The DNA sequences of *FhgVTE3-1* were 2710 bp long, having 3 introns located at 44-163, 772-1295, and 1603-2437 bp with lengths of 120, 524, and 835 bp, respectively. The DNA sequences of *FhgVTE3-2* were 2706 bp long, having 3 introns located at 44-169, 778-1291, and 1599-2433 bp with lengths of 126, 514, and 835 bp, respectively. In addition, the intron-splicing pattern adhered to the GA-AG rule (Figure 1).

Analysis of the *FhgVTE3-1* and *-2* sequences demonstrated that the 3 introns possessed 3, 15, and 18 SNPs (Figure 1), and the full-length sequences had 56 variable sites, of which 44 were SNPs, accounting for 79% of introns.

Analysis of restriction enzyme sites in *FhgVTE3-1* and *-2* by DNAMAN indicated that *FhgVTE3-2* possessed 3 restriction enzyme recognition sites (*Pac*I, *Afl*II, and *Bgl*II). In contrast, none of the 3 restriction sites was found in *FhgVTE3-1* (Figure 2).

VTE3 nucleotide sequence analysis in different cultivars

Sequence alignment analysis revealed 100% homology of the *FhrVTE3-1*, *LnrVTE3-1*, and *LprVTE3-1* sequences (Table 1), as well as 100% homology of the *FhrVTE3-2*, *LnrVTE3-2*, and *LprVTE3-2* sequences (Table 1).

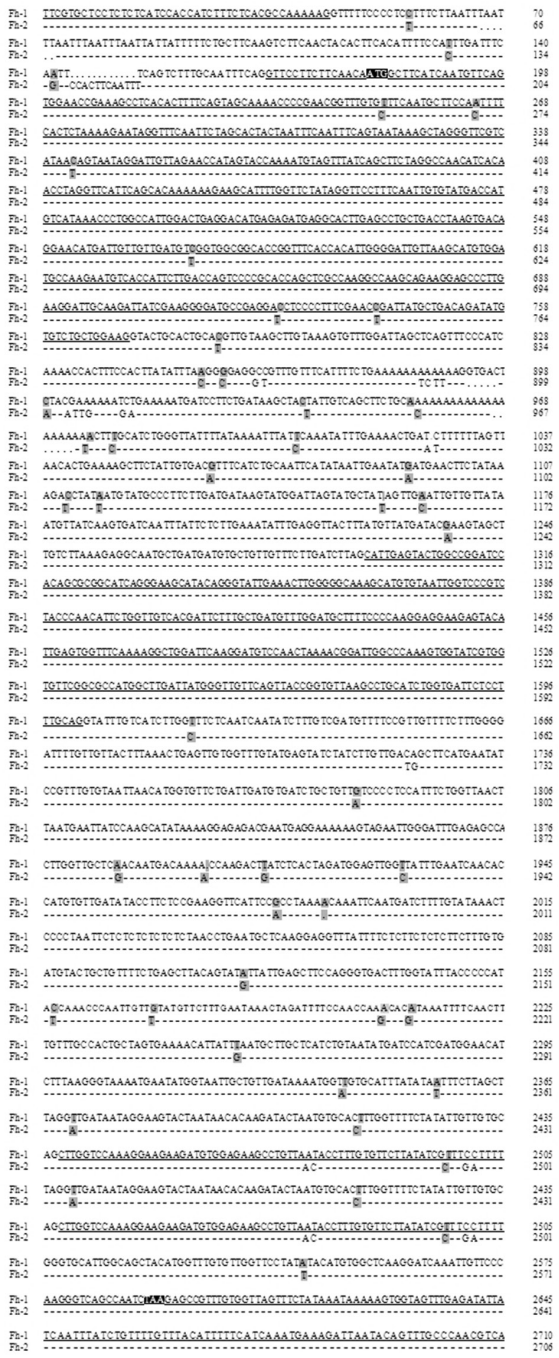

Figure 1. Alignment of 2 genomic nucleotide sequences of *VTE3* from *Arachis hypogaea*. Fh-1 = *FhgVTE3-1*; Fh-2 = *FhgVTE3-2*; exons are underlined; (·) = nucleotide deletions; the initiator codon ATG and the stop codon TAA are shaded by dark blocks and the sites of SNP variations by pale gray.

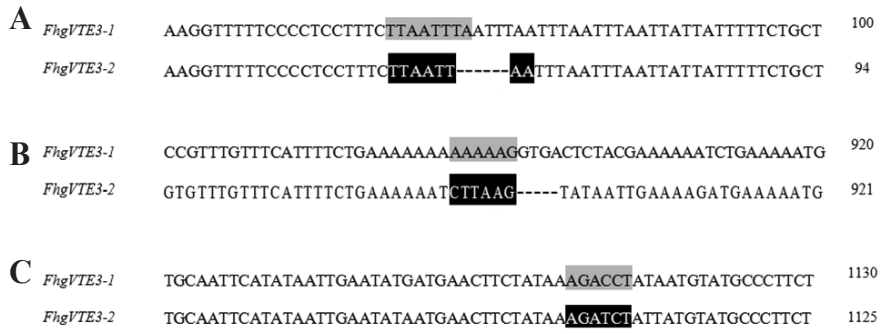

Figure 2. Alignment of partial genomic nucleotide sequences of *VTE3* from *Arachis hypogaea*. (-) = nucleotide deletions. **A.** The endonuclease recognition site of *Pac*I is shaded by dark blocks and the unrecognitive nucleotides are shaded by pale gray; **B.** the endonuclease recognition site of *Afl*II is shaded by dark blocks and the unrecognitive nucleotides are shaded by pale gray; **C.** the endonuclease recognition site of *Bgl*II is shaded by dark blocks and the unrecognitive nucleotides by pale gray.

The homology of *gVTE3-1* nucleotide sequences in the 13 cultivars was 99.9%, and polymorphisms were found to exist within the homologous sites. According to the different numbers of A-bases (A's) found within the *gVTE3-1* nucleotide sequences, the 13 cultivars could be classified into 4 categories: category 1 comprised Fenghua 2, Krapt. st. 16, and ICG6848 with 4 A's at the diversity site; category 2 was made up of Yuanza 9102, E1, and Rugaoxiyangsheng with 3 A's; category 3 included Shanhua 7, 8 and 9, Lanna 1, Nongda 818, and Lipudahuasheng, which shared 2 A's; category 4 comprised only ICGV86699, which contains 4 A's less than the cultivars in category 1 (Figure 3). The former 3 categories comprised different peanut varieties, and no consistency was found between the varieties. The *gVTE3-2* nucleotide sequences in the 13 cultivated varieties were found to be totally conserved with 100% homology.

Figure 3. Alignment of partial genomic nucleotide sequences of *VTE3* from 13 accessions (as shown in Table 1) of *Arachis hypogaea*.

Phylogenetic analysis of *VTE3* sequences from cultivars and wild species

The *FhgVTE3-1* (Fenghua 2) and -*2* (Fenghua 2) and *gVTE3-A* (*A. duranensis*) and -*B* (*A. ipaënsis*) sequences were aligned with ClustalX, and a phylogenetic tree (Figure 4) was constructed by PHYLIP based on the maximum parsimony method. In addition, bootstrap (1000 replicates) was used to assess the tree. The tree revealed that the *FhgVTE3-1* and *gVTE3-A* clusters formed a group and the *FhgVTE3-2* and *gVTE3-B* clusters formed another group. Sequence alignment analysis revealed that the homology of *FhgVTE3-1* and *gVTE3-A* was 98.7% and that of *FhgVTE3-2* and *gVTE3-B* was 100%. Currently, the majority of researchers believe that *A. duranensis* and *A. ipaënsis* are the donors of A-and B-genome in the cultivated peanut. Therefore, the present study inferred that *gVTE3-1* and -*2* in cultivated peanuts are from A- and B-genome, respectively.

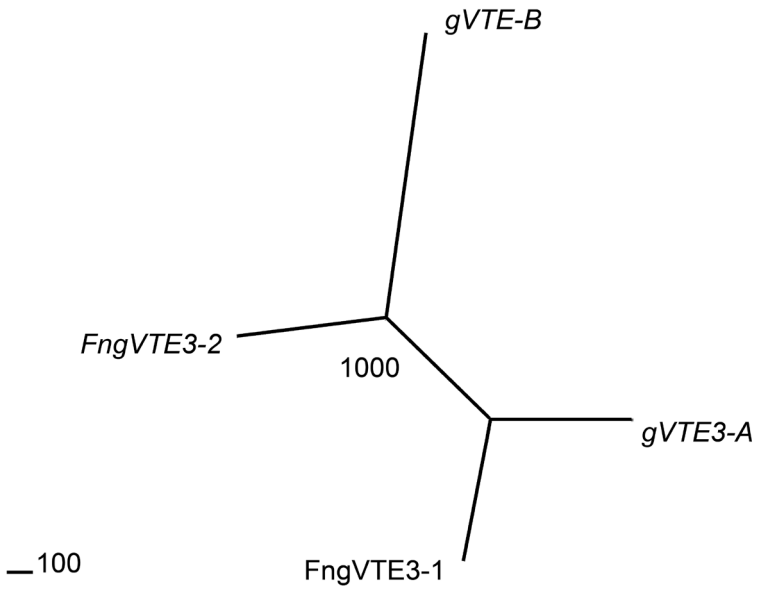

Figure 4. Phylogenetic tree based on the genomic nucleotide sequences of *FhgVTE3-1* and -*2*, *gVTE3-A* and -*B*.

Bioinformatic analyses of protein sequences predicted by peanut *VTE3* nucleotide sequences

The amino acid sequences predicted by analysis of the *FhrVTE3-1* and -*2* coding sequences were designated Ah1 and Ah2. Bioinformatics analyses (see Material and Methods) were performed on the 2 proteins. For Ah1, the molecular weight, pI, and grand average of hydropathicity were 39.90 kDa, 9.27, and -0.288, respectively. The molecular weight, pI, and grand average of hydropathicity of Ah2 were 39.86 kDa, 9.33, and -0.279, respectively. The results indicate that the 2 proteins are hydrophobic.

Conserved domain analysis showed that Ah1 and Ah2 shared the same conserved domain of *S*-adenosylmethionine-dependent methyltransferases (Figure 5). Secondary structure analysis revealed that Ah1 comprises 31.34% α-helices, 22.51% extended strands, 5.13% β-turns, and 41.03% random coils; Ah2 comprises 28.49% α-helices, 23.36% extended strands, 5.13% β-turns, and 43.02% random coils. Transmembrane structure analysis using TMHMM determined that the transmembrane region was found at positions 318-340 in both the Ah1 and Ah2 proteins. In addition, analysis of subcellular localization by WoLF PSORT revealed that Ah1 and Ah2 are located in the chloroplast.

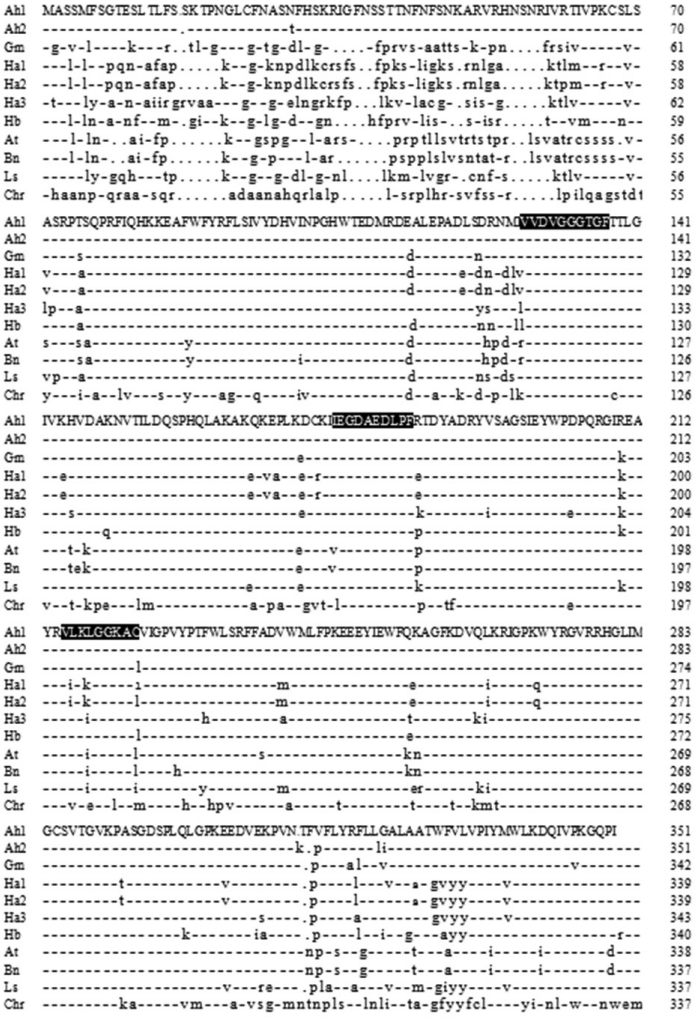

Figure 5. Alignment of amino acid sequences of 2-methyl-6-phytyl-1,4-benzoquinol methyltransferase in peanuts compared with other plants. Ah = *Arachis hypogaea*; Gm = *Glycine max*; Ha = *Helianthus annuus*; Hb = *Hevea brasiliensis*; At = *Arabidopsis thaliana*; Bn = *Brassica napus*; Ls = *Lactuca sativa*; Chr = *Chlamydomonas reinhardtii*. (-) = same amino acid residues. Conserved domains are shaded by dark blocks.

Homology analysis of *VTE3* amino acid sequence in peanuts compared with other plants

Comparison of the homologous amino acid sequences of peanut Ah1 and Ah2 with homologous proteins from other plants demonstrated that Ah1 shares 84.9, 81.2, 78.1, 78.1, 77.3, 76.7, and 69.7% homology with MPBQ MT from soybeans, rubber, lettuce, sunflower, *A. thaliana*, *Brassica napus*, and *Chlamydomonas*, respectively. It also revealed that Ah2 shares 84.6, 81.2, 77.8, 77.6, 76.4, 75.9, and 70.4% homology with the MPBQ MT from the species listed above.

The homologous MPBQ MT proteins from the species mentioned above (i.e., peanuts, soybeans, rubber, lettuce, sunflower, *A. thaliana*, *B. napus*, and *Chlamydomonas*) were aligned (Figure 5) by PHYLIP in combination with ClustalW and the subsequent phylogenetic tree construction (Figure 6). Figure 5 illustrates the high homology of MPBQ MT across different species, even in distant species such as *Chlamydomonas*, indicating strong conservation of genetic material in MPBQ MT. Figure 6 shows that, among the 11 amino acid sequences from 8 sampled species, *A. thaliana*, *B. napus*, and *Chlamydomonas* are clustered on 1 branch; sunflower and lettuce are clustered on another branch; peanut and soybean share 1 branch; and rubber is the sole occupant of the final branch. The first branch can be further classified into 2 sub-branches consisting of *A. thaliana* and *B. napus* on one sub-branch and *Chlamydomonas* on the other. This demonstrates that *Chlamydomonas* is much more distantly related to other plants used in this study. The protein molecular tree is consistent with previous studies utilizing classical taxonomy, specifically in the case of peanuts and soybeans belonging to legume groups, sunflowers and lettuces belonging to Asteraceae, and *A. thaliana* and *B. napus* belonging to Cruciferous.

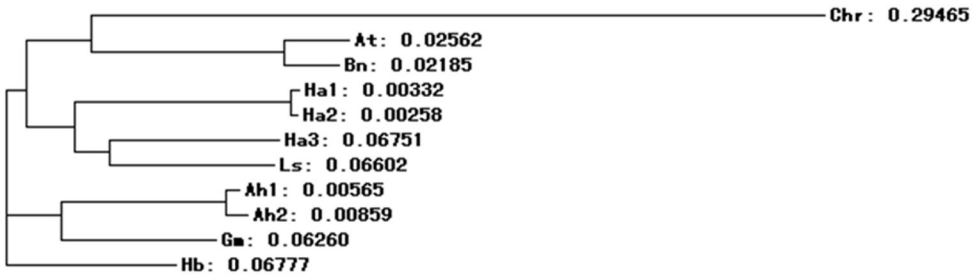

Figure 6. Phylogenetic tree of 2-methyl-6-phytyl-1,4-benzoquinol methyltransferase for related species. Chr = *Chlamydomonas reinhardtii*; At = *Arabidopsis thaliana*; Bn = *Brassica napus*; Ha = *Helianthus annuus*; Ls = *Lactuca sativa*; Ah = *Arachis hypogaea*; Gm = *Glycine max*; Hb = *Hevea brasiliensis*.

DISCUSSION

This study utilized EST assembly and RT-PCR techniques to obtain 2 peanut *VTE3* sequences, demonstrating that the coding nucleotide sequences share more than 70% homology with *A. thaliana VTE3*. The corresponding predicted amino acid sequences share more than 76% homology with the homologous MPBQ MT from any variety of *Arabidopsis* species, soybeans, rubbers, lettuces, sunflowers, and *B. napus*. All of the proteins had an S-adenosyl-methionine-dependent methyltransferase conserved domain in common.

Bioinformatic analyses demonstrated that peanut MPBQ MT is primarily hydrophobic, is located within the chloroplast, and possesses a distinct transmembrane region. These results are consistent with previous findings studying the related enzyme-producing genes involved in vitamin E synthesis (Arango and Heise, 1998). Therefore, the sequence we obtained in this study is actually the peanut *VTE3* sequence, and the corresponding MPBQ MT proteins are highly conserved across different species.

Phylogenetic analysis completed in the present study demonstrated that *FhgVTE3-1* and *-2* in Fenghua 2 are from the A- and B-genome, respectively. The 2 sequences were shown to contain 56 variable sites, and these sites were used to design primers that allowed *VTE3* genomes from different sources to be distinguished. Analysis of restriction enzyme recognition sites showed that *FhgVTE3-2* has 3 restriction sites corresponding to *Pac*I, *Afl*II, and *Bgl*II, whereas *FhgVTE3-1* cannot be recognized by any of the 3 restriction enzymes. This finding suggests that these restriction sites are likely to have genomic specificity. Differentiation methods utilizing *Pac*I, *Afl*II, and *Bgl*II recognition sites on the *VTE3* gene sequences from B-genome sources have thus been shown to give researchers the ability to distinguish between the 2 *VTE3* genes in peanut cultivars.

VTE3 sequences were cloned from 2 subspecies of peanut cultivars comprising 13 distinct varieties, producing 2 homologous *VTE3* DNA sequences, *gVTE3-1* and *-2*, cloned from each of the sampled varieties. Alignment showed that *gVTE3-2* is totally conserved with 100% homology across the 13 varieties, and *gVTE3-1* has 99.9% homology. The 0.1% difference can be accounted for by subtle differences in the intron repeat sequence between the various samples. Different molecular markers were used by Raina et al. (2001) and Milla et al. (2005), confirming the conclusions that the genetic base of cultivated peanuts is narrow and that polymorphisms occur only at low levels. Different peanut *VTE3* cultivars have been demonstrated to be highly homologous, with the 2 subspecies *A. hypogaea* ssp *fastigiata* and *A. hypogaea* ssp *hypogaea* being the most likely to share the same diploid wild ancestral species, although validation through further study, including larger cohorts and examination of additional molecular markers is needed to validate these results.

ACKNOWLEDGMENTS

We appreciate the assistance and dedication of Professor Huifang Jiang at the Oil Crops Research Institute of the Chinese Academy of Agricultural Sciences for supplying the wild species *Arachis duranensis* and *A. ipaënsis*. Research supported by the Shandong Province Peanut Seed Industry Project and financed by the Earmarked Fund for China Agriculture Research System (#CARS-14).

Conflicts of interest

The authors declare that they have no conflict of interest.

REFERENCES

Arango Y and Heise KP (1998). Tocopherol synthesis from homogentisate in *Capsicum anuum* L. (yellow pepper) chromoplast membranes: evidence for tocopherol cyclase. *Biochem. J.* 336: 531-533.

Brigelius-Flohe R and Traber MG (1999). Vitamin E: function and metabolism. *FASEB J.* 13: 1145-1155.

Cheng Z, Sattler S, Maeda H, Sakuragi Y, et al. (2003). Highly divergent methyltransferases catalyze a conserved reaction in tocopherol and plastoquinone synthesis in cyanobacteria and photosynthetic eukaryotes. *Plant Cell* 15: 2343-2356.

DellaPenna D and Pogson BJ (2006). Vitamin synthesis in plants: tocopherols and carotenoids. *Annu. Rev. Plant Biol.* 57: 711-738.

Fávero AP, Simpson CE, Valls JFM and Vello NA (2006). Study of the evolution of cultivated peanut through crossability studies among *Arachis ipaënsis*, *A. duranensis*, and *A. hypogaea*. *Crop Sci.* 46: 1546-1552.

Grusak MA and DellaPenna D (1999). Improving the nutrient composition of plants to enhance human nutrition and HEALTH1. *Annu. Rev. Plant Physiol. Plant Mol. Biol.* 50: 133-161.

Halward TM, Stalker HT, Larue EA and Kochert G (1991). Genetic variation detectable with molecular markers among unadapted germ-plasm resources of cultivated peanut and related wild species. *Genome* 34: 1013-1020.

Hong Y, Chen X, Liang X, Liu H, et al. (2010). A SSR-based composite genetic linkage map for the cultivated peanut (*Arachis hypogaea* L.) genome. *BMC Plant Biol.* 10: 17.

Kochert G, Stalker H, Gimenes M, Galgaro L, et al. (1996). RFLP and cytogenetic evidence on the origin and evolution of allotetraploid domesticated peanut, *Arachis hypogaea* (Leguminosae). *Am. J. Bot.* 83: 1282-1291.

Lu J and Pickersgill B (1993). Isozyme variation and species relationships in peanut and its wild relatives (*Arachis* L. - Leguminosae). *Theor. Appl. Genet.* 85: 550-560.

Milla SR, Isleib TG and Stalker HT (2005). Taxonomic relationships among *Arachis* sect. *Arachis* species as revealed by AFLP markers. *Genome* 48: 1-11.

Moretzsohn MC, Barbosa AV, Alves-Freitas DM, Teixeira C, et al. (2009). A linkage map for the B-genome of *Arachis* (Fabaceae) and its synteny to the A-genome. *BMC Plant Biol.* 9: 40.

Naqvi S, Farre G, Zhu C, Sandmann G, et al. (2011). Simultaneous expression of *Arabidopsis* rho-hydroxyphenylpyruvate dioxygenase and MPBQ methyltransferase in transgenic corn kernels triples the tocopherol content. *Transgenic Res.* 20: 177-181.

Porfirova S, Bergmuller E, Tropf S, Lemke R, et al. (2002). Isolation of an *Arabidopsis* mutant lacking vitamin E and identification of a cyclase essential for all tocopherol biosynthesis. *Proc. Natl. Acad. Sci. U. S. A.* 99: 12495-12500.

Raina SN, Rani V, Kojima T, Ogihara Y, et al. (2001). RAPD and ISSR fingerprints as useful genetic markers for analysis of genetic diversity, varietal identification, and phylogenetic relationships in peanut (*Arachis hypogaea*) cultivars and wild species. *Genome* 44: 763-772.

Rimbach G, Minihane AM, Majewicz J, Fischer A, et al. (2002). Regulation of cell signalling by vitamin E. *Proc. Nutr. Soc.* 61: 415-425.

Rogers SO and Bendich AJ (1985). Extraction of DNA from milligram amounts of fresh, herbarium and mummified plant tissues. *Plant Mol. Biol.* 5: 69-76.

Seijo JG, Lavia GI, Fernandez A, Krapovickas A, et al. (2004). Physical mapping of the 5S and 18S-25S rRNA genes by FISH as evidence that *Arachis duranensis* and *A. ipaensis* are the wild diploid progenitors of *A. hypogaea* (Leguminosae). *Am. J. Bot.* 91: 1294-1303.

Seijo G, Lavia GI, Fernandez A, Krapovickas A, et al. (2007). Genomic relationships between the cultivated peanut (*Arachis hypogaea*, Leguminosae) and its close relatives revealed by double GISH. *Am. J. Bot.* 94: 1963-1971.

Shintani D and DellaPenna D (1998). Elevating the vitamin E content of plants through metabolic engineering. *Science* 282: 2098-2100.

Shintani DK, Cheng Z and DellaPenna D (2002). The role of 2-methyl-6-phytylbenzoquinone methyltransferase in determining tocopherol composition in *Synechocystis* sp. PCC6803. *FEBS Lett.* 511: 1-5.

Singh AK (1988). Putative genome donors of *Arachis hypogaea* (*Fabaceae*), evidence from crosses with synthetic amphidiploids. *Plant Syst. Evol.* 160: 143-151.

Van Eenennaam AL, Lincoln K, Durrett TP, Valentin HE, et al. (2003). Engineering vitamin E content: from *Arabidopsis* mutant to soy oil. *Plant Cell* 15: 3007-3019.