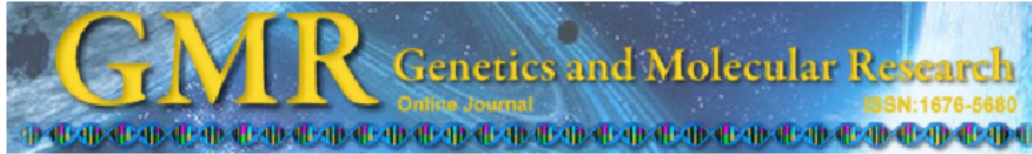

Molecular cloning and functional characterization of a glutathione *S*-transferase involved in both anthocyanin and proanthocyanidin accumulation in *Camelina sativa* (Brassicaceae)

Y. Wang[1]*, Y. Tang[1]*, M. Zhang[2], F. Cai[3], J. Qin[2], Q. Wang[1], C. Liu[1], G. Wang[1], L. Xu[1], L. Yang[1], J. Li[1], Z. Wang[1,4] and X. Li[1,2]

[1]China-Korea Joint Institute for Animal Science of Jilin Province, Changchun, China
[2]College of Plant Sciences, Jilin University, Changchun, Jilin, China
[3]Hospital of the Aeronautical University of China People Liberation Air Force, Changchun, China
[4]Food Science and Engineering College, Jilin Agricultural University, Changchun, China

*These authors contributed equally to this study.
Corresponding authors: Z. Wang / X. Li
E-mail: wanggui3038@163.com / lixi@mcmaster.ca

Genet. Mol. Res. 11 (4): 4711-4719 (2012)
Received March 29, 2012
Accepted June 8, 2012
Published September 25, 2012
DOI http://dx.doi.org/10.4238/2012.September.25.4

ABSTRACT. Recently, we found that the *Arabidopsis* TT19 protein, a glutathione *S*-transferase, has two functional domains that influence both anthocyanin and proanthocyanidin accumulation. To further understand the function of this protein in the other species, we cloned a cDNA encoding a glutathione *S*-transferase (namely *CMGSTF12*) from *Camelina sativa*, an oil crop that has received renewed interest due to its biofuel value and high omega-3 levels. Southern blot analysis demonstrated one copy of *CMGSTF12* in *C. sativa*. Transformation of the *Arabidopsis* loss-of-function *tt19-1*

 ©FUNPEC-RP www.funpecrp.com.br

mutant with *CMGSTF12* cDNA complemented accumulation of anthocyanin in vegetative tissues and resulted in the wild-type level of proanthocyanidin (both extractable and unextractable) in seeds. No obvious flavonoid accumulation changes were detected in the transgenic seeds, indicating that *CMGSTF12* may only involve the lower flavonoid pathway, further proving that the TT19 protein controls accumulation of unextractable proanthocyanidin.

Key words: *Camelina sativa*; Glutathione *S*-transferase; *CMGSTF12*; TT19

INTRODUCTION

Camelina sativa L. Cranz, a member of family Brassicaceae, is a relative of the genetic model organism *Arabidopsis thaliana*, which has gained increased popularity for its potential in biofuel applications (Fröhlich and Rice, 2005; Davis et al., 2011). *C. sativa* has been cultivated for oil production since prehistoric times and is widely distributed on marginal land (Imbrea et al., 2011). Phytochemical analysis has shown that the vegetative tissues of this plant contain flavonols and anthocyanins and its seeds are a source of health-promoting omega-3 fatty acids (Onyilagha et al., 2003; Ghamkhar et al., 2010). Flavonoid content and composition have been shown to influence the agronomically important traits seed dormancy, storability, and quality (Lepiniec et al., 2006). Anthocyanins and proanthocyanidins (PAs) are important plant pigments that are end-products of the flavonoid pathway. Flavonoid biosynthesis genes are downstream of the phenylpropanoid (monomeric lignin) pathway (Figure 1) (Marles et al., 2003). Although knowledge of this pathway has advanced significantly since the determination of anthocyanin reductase gene (*BANYULS*) function (Xie et al., 2003), the steps involved in the transportation of flavonoids from their presumed site of synthesis (endoplasmic reticulum) or uptake to their final destination in vacuoles are still poorly understood.

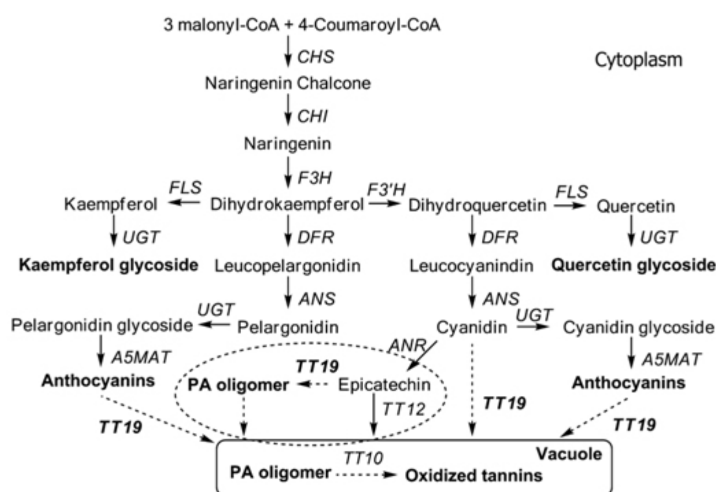

Figure 1. Proanthocyanidin (PA) biosynthetic pathway. Dotted lines indicate regions of the pathway, which are still not fully defined.

In *Arabidopsis*, three genes involved in vacuolar transportation of flavonoids have been identified to date. *AHA10* encodes an H^+-ATPase, and the transparent testa (TT) 12 gene encodes a multidrug and toxic compound extrusion membrane-spanning transporter-like protein that acts as a vacuolar flavonoid/H^+-antiporter in PA accumulation in the seed coat (Debeaujon et al., 2001; Baxter et al., 2005). Our recent research has suggested that a third enzyme, the *Arabidopsis* glutathione *S*-transferase (GST) TT19 protein, has two functional domains, one that affects anthocyanin accumulation through a domain containing a glutathione-binding site and another that affects PA accumulation through a 3'-conserved domain (Li et al., 2011). Until now, the molecular features of flavonoid biosynthesis in *C. sativa* have not been characterized.

In the present study, we identified a complementary DNA (cDNA) encoding a GST (homologue of the *Arabidopsis TT19* gene) - namely *CMGSTF12*. Transformation of the *Arabidopsis* loss-of-function *tt19-1* mutant with *CMGSTF12* cDNA increased the accumulation of anthocyanin in vegetative tissues and resulted in wild-type levels of PA (both extractable and unextractable) in seeds. The results from this study indicated that *CMGSTF12* is a functional GST involved in both anthocyanin and PA accumulation in *C. sativa*.

MATERIAL AND METHODS

Plant material, chemicals, *Escherichia coli* strains, and plasmids

C. sativa, *A. thaliana* wild-type Col-0, and transgenic plants were grown in soil. Seeds were cold-treated at 4°C for 2 days and then allowed to germinate and grow in a greenhouse at 22°C with a 16-h light/8-h dark cycle supplemented with halogen lamps. Alternatively, seeds were germinated on Murashige and Skoog medium containing 3% sucrose in a controlled growth cabinet (16-h light/8-h dark cycle) at 22°C. All solvents used in this study were of high-performance liquid chromatography (HPLC) grade. We purchased 3'- and 5'-rapid amplification of cDNA ends (RACE) kits from TaKaRa (Dalian, China). *E. coli* strains DH10B (Invitrogen, Carlsbad, CA, USA) and the sequence vector PGEM-T Easy (Promega, Madison, WI, USA) were used to clone the *CMGSTF12* gene.

Cloning of a full-length cDNA sequence from *C. sativa*

A highly conserved region was identified through comparison of the *TT19* cDNA sequences from *A. thaliana* (accession No. NM_121728) and *Thellungiella halophila* (BAJ34302). Based on this region, primers P1 and P2 (Table 1) were designed to amplify an approximately 358-bp DNA fragment using reverse transcriptase polymerase chain reaction (RT-PCR) with the cDNA from 24-day-old *C. sativa* seedlings as the template. Based on the partial sequences obtained, 5'- and 3'-RACE were performed to amplify the remaining cDNA regions of *CMGSTF12* according to manufacturer instructions (TaKaRa). The open reading frame (ORF) of *CMGSTF12* was amplified using gene-specific primers P3 and P4 (Table 1) and high-fidelity DNA polymerase Pfu Phusion (Invitrogen). The PCR product was cloned into pGEM-T Easy vector, yielding the plasmid pGEM-CMGSTF12.

Table 1. Primers used in this study.

Name	Sequences
P1	5'-AGCTTTTTGAATCACGAG-3'
P2	5'-ATATCGGTTATACTCATC-3'
P3	5'-CGCGGATCCATGGTTGTGAAACTATAC-3'
P4	5'-CGCGAGCTCTCAGTGACCAGCCATCTC-3'
P5	5'-ACGAGCCATTGCGAGATACT-3'
P6	5'- GGTTAACCACAATGGGATGG-3'

The sequences underlined indicate restriction enzyme sites.

Sequence alignment and phylogenetic analysis of CMGSTF12

Amino acid sequences were aligned using the AlignX program, part of the Vector NTI suite (Invitrogen) with default parameter settings (gap opening penalty, 10; gap extension penalty, 0.05; gap separation penalty range, 8; identity for alignment delay, 40%). A phylogenetic tree was constructed using the neighbor-joining algorithm in Molecular Evolutionary Genetics Analysis version 5.0 with 1000 bootstrap trials (Li et al., 2011).

Southern blot

For Southern blot analysis, genomic DNA was isolated from 24-day-old seedlings using the cetyltrimethylammonium bromide method (Saghai-Maroof et al., 1984), digested with *Bam*HI, *Hin*dIII, and *Eco*RI restriction enzymes, and resolved on 1.2% agarose gels. Southern blot analysis was performed on nylon membranes according to Southern (1975). The *CMGSTF12* coding region was radiolabeled with α-^{32}P deoxycytidine triphosphate using Ready-to-Go DNA labeling beads (Amersham Pharmacia Biotech, Piscataway, USA) and used as a probe. Hybridization was carried out at 55°C for 2 h with 0.8 M NaCl, and blots were washed at high stringency (0.2X saline-sodium citrate, 0.1% sodium dodecyl sulfate).

Real-time PCR

RNAs extracted from various growth stages using a commercial RNAEasy Mini Kit (Qiagen, Valencia, CA, USA) were used in quantitative real-time PCRs (qPCR) with gene-specific primers P5 and P6 (Table 1). qRT-PCRs were conducted using SuperScript III First-Strand Synthesis SuperMix, a Platinum SYBR Green qPCR SuperMix-UDG kit (Invitrogen), and a StepOnePlus Real-Time PCR System (Applied Biosystems, Carlsbad, USA) following manufacturer instructions. qRT-PCR primers (Table 1) were designed using the online primer design tool OligoPerfect^{TM} Designer (Invitrogen). The qRT-PCR mixture contained 8 μL diluted cDNA, 10 μL 2X SYBR Green qPCR Master Mix (Invitrogen), and 200 nM of each gene-specific primer in a final volume of 20 μL. Control PCR without cDNA templates were also performed for each primer pair.

qRT-PCR was performed using a StepOnePlus Real-Time PCR System (Applied Biosystems) as previously described (Gao et al., 2009) under the following conditions: 2 min at 50°C followed by 2 min at 95°C, 40 cycles for 15 s at 95°C, and for 30 s at 60°C on a 96-well optical reaction plate (Bio-Rad Laboratories, Hercules, USA). The specificity of amplicons was verified through melting curve analysis (60°-95°C) after 40 cycles. Mean expression values were calculated from assays in triplicate RNA preparations from independent plant samples. PCR efficiency was between 88 and 92%.

Complementation of *Arabidopsis tt19-1* with *CMGSTF12*

The ORF of *CMGSTF12* from the vector pGEM-CMGSTF12 was digested with *Bam*HI and *Sac*I and ligated into binary vector pBI121 under the control of the 35*S* promoter, which led to construct pBI121-CMGSTF12. Wild-type *Arabidopsis* and *tt19-1* mutants that lacked TT19-GST activity were transformed with the *Agrobacterium tumefaciens* GV3101 containing the binary construct pBI121-CMGSTF12 via floral dipping (Clough and Bent, 1998) followed by selection with 50 µg/mL kanamycin and used for PA and flavonoid HPLC analysis.

Determination of anthocyanin and extractable and unextractable PA

Before extraction and quantification of anthocyanin, seedlings of Col-0, *tt19-1*, and complementary lines were grown on Murashige and Skoog agar plates containing 3% sucrose for 10 days. Anthocyanins were extracted from seedlings with 70% methanol and directly injected into an Agilent 1200 HPLC system and a Zorbax C_{18} column (Mississauga, Canada) as described by Li et al. (2011). Total anthocyanins were calculated at 525 nm using a standard curve of cyanidin chloride (Sigma-Aldrich, St. Louis, MO, USA). Extractable seed PAs were quantified after grinding 300 mg seed in 30 mL (10 mg/mL) acetone:H_2O (70:30, v/v). The extracts were dried; the volume was made up to 1 mL and hydrolyzed to anthocyanidins for 60 min at 95°C using 2 mL *n*-BuOH/HCl solution (95:5, v/v) and 0.1 mL $FeSO_4$ in 2 M HCl. After cooling and centrifugation, the supernatant was chromatographed on the Zorbax C_{18} column. The absorbance peak area of red anthocyanidins was recorded at 525 nm, and extractable PAs were calculated from a similarly treated PA B_2 standard curve. Unextractable PA was measured as the supernatant arising after boiling the solid residue (20 mg) remaining after acetone extraction in 1 mL *n*-BuOH/HCl solution (95:5, v/v).

RESULTS

Cloning and molecular characterization of *CMGSTF12*

With the known conserved features of *A. thaliana* and *T. halophila* TT19 genes, we used 24-day seedling cDNA as a template for PCR with primers P1 and P2 to isolate a partial DNA fragment of the GST gene from *C. sativa*. A 358-bp DNA fragment was obtained. Then, the complete sequence of this gene was determined using 5'- and 3'-RACE based on the sequence of the partial DNA, and the gene was named *CMGSTF12*. The *CMGSTF12* cDNA contains an ORF encoding a protein of 214 amino acids with a predicted molecular mass of 24.6 kDa. We performed multiple sequence alignment of CMGSTF12 with other previously reported TT19 genes. As shown in Figure 2A, phylogenetic analysis demonstrated that CMGSTF12 is closely grouped with *Arabidopsis* TT19. The highly homologous similarities found between these two sequences (Figure 2D) suggested that CMGSTF12 may function in the same way that *Arabidopsis* TT19 does.

When *Hin*dIII-, *Bam*HI-, and *Eco*RI-digested DNA samples were probed with *CMGSTF12*, the number of hybridization signals indicated that only one copy of *CMGSTF12* existed in the *C. sativa* genome (Figure 2B). Temporal expression of *CMGSTF12* in various tissues was detected using qPCR. The *CMGSTF12* transcript accumulated at the highest level in immature seeds (Figure 2C).

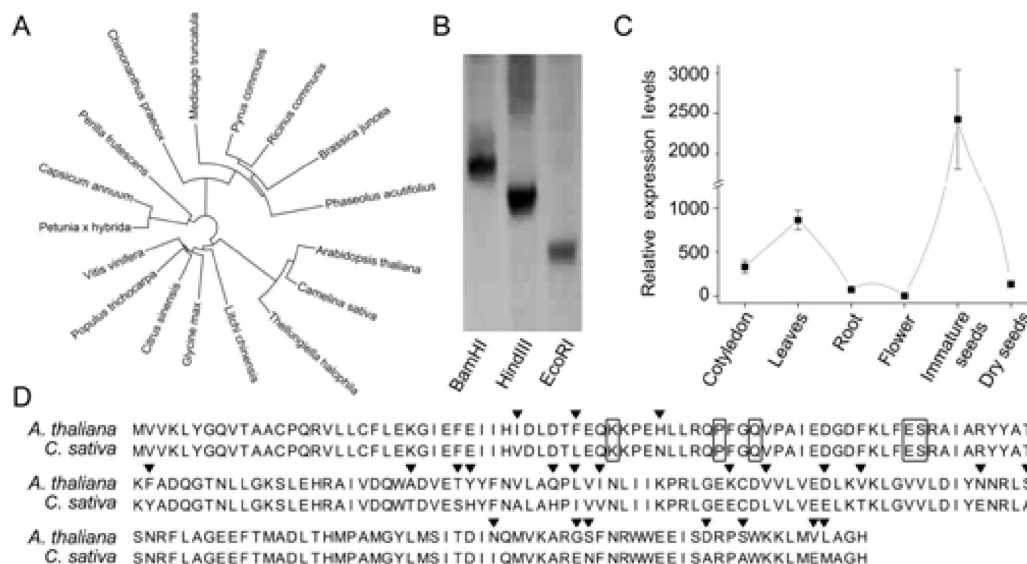

Figure 2. Sequence information of CMGSTF12. **A.** Phylogenetic analysis of CMGSTF12. **B.** Southern blot; Genomic DNA was digested with *Bam*HI, *Hin*dIII, and *Eco*RI restriction enzymes, and resolved on 1.2% agarose gels, then probed by α-^{32}P dCTP-labeled *FGCOMT1*. **C.** *CMGSTF12* expression levels in different developing stages. Dry seed mRNA level was designated as one and all results were shown as means of at least three independent RNA extractions with corresponding standard errors (SE). **D.** Alignment of CMGSTF12 (bottom) with *Arabidopsis* TT19 (top). Amino acid differences were highlighted with triangle and GSH binding sites were marked with boxes.

Complementation of *Arabidopsis tt19-1* with *CMGSTF12*

To verify the function of *CMGSTF12 in planta*, we overexpressed it under the CaMV *35S* promoter in an *Arabidopsis tt19-1* mutant (Figure 3A). RT-PCR analysis revealed that the transgene was highly expressed (Figure 3B). Vegetative tissues of *tt19-1* mutant lines showed no red pigment (anthocyanin) accumulation (Figure 3D). In contrast, all T_3 homozygous-complemented *tt19-1* expression lines exhibited anthocyanin pigmentation levels similar to those in the wild type (data not shown) when seedlings were grown on sucrose-supplemented media (Figure 3E and F).

The *tt19-1* mutant had yellow seeds, which differed from the wild-type brown seed (Figure 3G and H), owing to a block in the flavonoid transportation pathway caused by the mutation in the TT19 gene (see Figure 1) (Li et al., 2011). Seeds from transgenic plants overexpressing *CMGSTF12* were brown (Figure 3I and J), indicating that *CMGSTF12* complemented the seed color phenotype of the *tt19-1* mutant. When extractable and unextractable PAs were profiled from our *Arabidopsis* lines using a butanol:HCl hydrolysis method, wild-type seeds had approximately 5-fold less unextractable seed PA and approximately 8-fold higher extractable seed PA levels than those in the *tt19-1* mutant (anthocyanidins released from extraction residue) (Figure 4A-D). By contrast, both complemented *tt19-1* lines had amounts of extractable and unextractable seed PA equivalent to those in wild-type seeds (Figure 4E-F). We analyzed the content of flavonol glycosides Q3R, QDR, KDR, and QHR in the wild type, *tt19-1*, and *CMGSTF12*-complemented *tt19-1* lines, and the flavonol derivatives were unchanged among all complemented lines (data not shown).

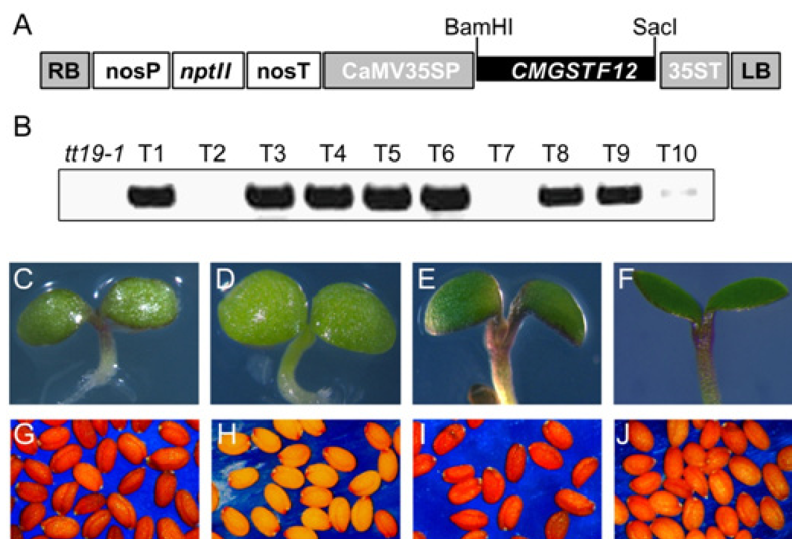

Figure 3. Complementation of *Arabidopsis tt19-1* by *CMGSTF12*. **A.** Constitutive expression vectors for transgenic analysis. **B.** PCR analysis in the transgenic (*lanes T1* to *T10*) and control *tt19-1*; C.-F. anthocyanin phenotype in cotyledon stage: **C.** wild-type Col-0; **D.** *tt19-1* mutant; **E.** and **F.** representative *CMGSTF12* complemented lines; G.-J. seed coat phenotype in mature seeds: **G.** wild-type Col-0; **H.** *tt19-1* mutant; **I.** and **J.** representative *CMGSTF12* complemented lines.

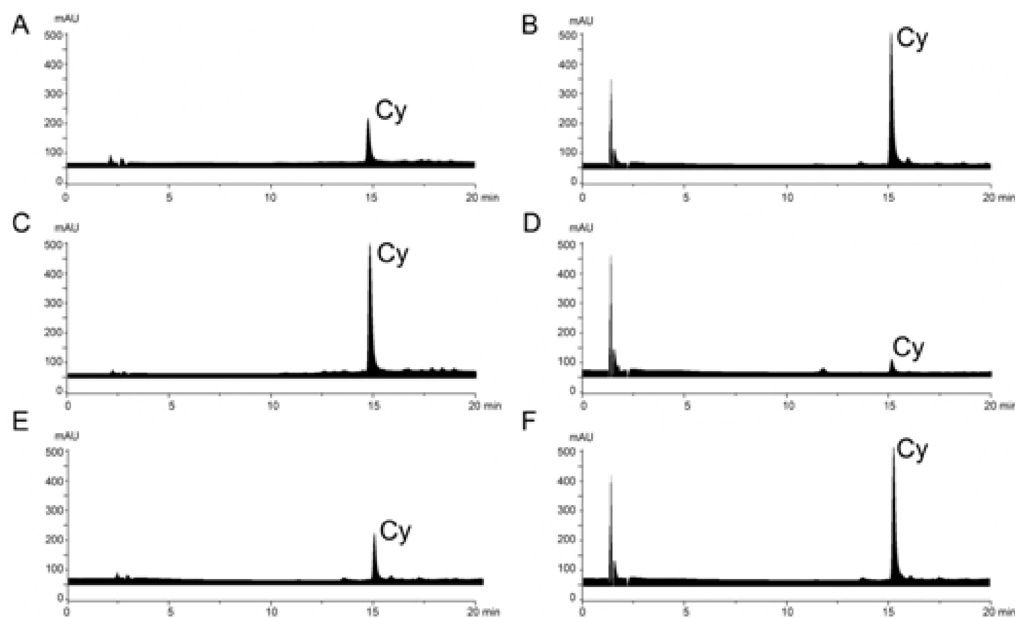

Figure 4. Representative HPLC-UV ($\lambda_{525\ nm}$) chromatogram profiles of hydrolyzed seed methanol extracts. **A.** Unextractable proanthocyanidin (PA) of the wild-type Col-0; **B.** extractable PA of the wild-type Col-0; **C.** unextractable PA of the *tt19-1* mutant; **D.** extractable PA of the *tt19-1* mutant; **E.** unextractable PA of complementation line; **F.** extractable PA of complementation line. Cy = cyanidin released after HCl hydrolysis.

DISCUSSION

The flavonoid pathway is responsible for red-blue pigments in flowers and yellow-brown pigments in seeds, in which the last steps in polymerization and transport are still poorly defined (Tian et al., 2008). We recently elucidated that the key *Arabidopsis* flavonoid pathway protein TT19-GST has two functional domains, one that influences both anthocyanin and PA accumulation in its 5'-GSH-binding domain and another that affects only PA accumulation through a 3'-substitution.

To understand the importance of the 3'-conserved domain function in the TT19 protein, we isolated a full-length cDNA encoding a GST (namely, *CMGSTF12*) from the agriculturally important plant *C. sativa.* The cDNA ORF is 645 bp and encodes a 214-amino acid peptide. Southern blot analysis results indicated that only one copy of *CMGSTF12* existed in the *C. sativa* genome. The *CMGSTF12* transcript accumulated at the highest level in immature seeds. The transformation of the *Arabidopsis* loss-of-function *tt19-1* mutant with *CMGSTF12* cDNA complemented the accumulation of anthocyanin in vegetative tissue and led to wild-type levels of PA (both extractable and unextractable) in seeds.

In our recently reported *Arabidopsis tt19-4* mutant, a G to T mutation occurred in a conserved 3'-domain in the *TT19* gene, which resulted in a missense mutation that converted Trp (amino acid 205) into Leu. Compared to the *tt19-1* mutant, the *tt19-4* mutant showed similar vegetative anthocyanin accumulation, but distinct PA profiles emphasized the importance of amino acid 205 (Trp). In the case of CMGSTF12, amino acid 205 remains as Trp, which, we proved, complements *tt19-1*. Sequence analysis data further suggested that the conserved 3'-domain of TT19 affected the dark seed coat of *Arabidopsis* and that seed coat color is related to extractable PA content. It also indicated that TT19 holds the key to the accumulation of unextractable PA. Overall, the results from this study showed that *CMGSTF12* is a functional GST involved in both anthocyanin and PA accumulations in *C. sativa*.

ACKNOWLEDGMENTS

Research supported by the Chinese Sika Deer Standardized System and Product Development of Key Technical Project (Q.K.W., #2011DFA32900) of the Ministry of Science and Technology of the People's Republic of China, the Jilin China-Korea Joint Institute for Animal Science, the Program for New Century Excellent Talents in University (X.L., #09-0423), and the National Natural Science Foundation of China (X.L., #31100219; J.C.Q., #31000149).

REFERENCES

Baxter IR, Young JC, Armstrong G, Foster N, et al. (2005). A plasma membrane H+-ATPase is required for the formation of proanthocyanidins in the seed coat endothelium of *Arabidopsis thaliana*. *Proc. Natl. Acad. Sci. U. S. A.* 102: 2649-2654.

Clough SJ and Bent AF (1998). Floral dip: a simplified method for *Agrobacterium*-mediated transformation of *Arabidopsis thaliana*. *Plant J.* 16: 735-743.

Davis PB, Menalled FD, Peterson RKD and Maxwell BD (2011). Refinement of weed risk assessments for biofuels using *Camelina sativa* as a model species. *J. Appl. Ecol.* 48: 989-997.

Debeaujon I, Peeters AJ, Leon-Kloosterziel KM and Koornneef M (2001). The TRANSPARENT TESTA12 gene of *Arabidopsis* encodes a multidrug secondary transporter-like protein required for flavonoid sequestration in vacuoles of the seed coat endothelium. *Plant Cell* 13: 853-871.

Fröhlich A and Rice B (2005). Evaluation of *Camelina sativa* oil as a feedstock for biodiesel production. *Ind. Crops Prod.*

 ©FUNPEC-RP www.funpecrp.com.br

21: 25-31.
Gao MJ, Lydiate DJ, Li X, Lui H, et al. (2009). Repression of seed maturation genes by a trihelix transcriptional repressor in *Arabidopsis* seedlings. *Plant Cell* 21: 54-71.
Ghamkhar K, Croser J, Aryamanesh N, Campbell M, et al. (2010). *Camelina* (*Camelina sativa* (L.) Crantz) as an alternative oilseed: molecular and ecogeographic analyses. *Genome* 53: 558-567.
Imbrea F, Jurcoane S, Hălmăjan HV, Duda M, et al. (2011). *Camelina sativa*: a new source of vegetal oils. *Rom. Biotech. Lett.* 16: 6263-6270.
Lepiniec L, Debeaujon I, Routaboul JM, Baudry A, et al. (2006). Genetics and biochemistry of seed flavonoids. *Annu. Rev. Plant Biol.* 57: 405-430.
Li X, Gao P, Cui D, Wu L, et al. (2011). The *Arabidopsis tt19-4* mutant differentially accumulates proanthocyanidin and anthocyanin through a 3' amino acid substitution in glutathione S-transferase. *Plant Cell Environ.* 34: 374-388.
Marles MA, Ray H and Gruber MY (2003). New perspectives on proanthocyanidin biochemistry and molecular regulation. *Phytochemistry* 64: 367-383.
Onyilagha J, Bala A, Hallett R, Gruber M, et al. (2003). Leaf flavonoids of the cruciferous species, *Camelina sativa*, *Crambe* spp., *Thlaspi arvense* and several other genera of the family Brassicaceae. *Biochem. Syst. Ecol.* 31: 1309-1322.
Saghai-Maroof MA, Soliman KM, Jorgensen RA and Allard RW (1984). Ribosomal DNA spacer-length polymorphisms in barley: mendelian inheritance, chromosomal location, and population dynamics. *Proc. Natl. Acad. Sci. U. S. A.* 81: 8014-8018.
Southern EM (1975). Detection of specific sequences among DNA fragments separated by gel electrophoresis. *J. Mol. Biol.* 98: 503-517.
Tian L, Pang Y and Dixon RA (2008). Biosynthesis and genetic engineering of proanthocyanidins and (iso)flavonoids. *Phytochem. Rev.* 7: 445-465.
Xie DY, Sharma SB, Paiva NL, Ferreira D, et al. (2003). Role of anthocyanidin reductase, encoded by BANYULS in plant flavonoid biosynthesis. *Science* 299: 396-399.