Three R2R3 MYB transcription factor genes from *Capsicum annuum* showing differential expression during fruit ripening

Jian-Guo Li¹,², Hui-Liang Li¹ and Shi-Qing Peng¹*

¹Key Laboratory of Tropical Crop Biotechnology, Ministry of Agriculture, Institute of Tropical Biosciences and Biotechnology, Chinese Academy of Tropical Agricultural Sciences, Haikou 571101, China.

²Center of Inspection and Testing for Agricultural Products, Ministry of Agriculture, Analysis and Testing Center, Chinese Academy of Tropical Agricultural Sciences, Haikou 571101, China.

Accepted 20 May, 2011

Three R2R3-MYB genes, designated *CaMYB1*, *CaMYB2* and *CaMYB3*, were isolated from hot pepper (*Capsicum annuum*, L.). *CaMYB1*, *CaMYB2* and *CaMYB3* encode polypeptides consisting of 340, 262 and 345 amino acids respectively, containing R2R3 domain and the signature motif specific for the interaction between MYB and bHLH proteins in the R3 domain. Phylogenetic analysis based on the deduced amino acid sequences of these three R2R3 MYB transcription factor members revealed that *CaMYB1* and *CaMYB2* clustered together with the anthocyanin-related subgroup of R2R3 MYB proteins from other plants, while *CaMYB3* did not. *CaMYBs* transcripts accumulation was detected in all stages of fruit development and in flower and leaves. Three *CaMYBs* transcription factors showed differential expression during fruit ripening. Anthocyanin biosynthetic gene expression patterns were quite different in young leaves, flower, and the four stages of fruit development. *CaMYB1* and *CaMYB2* may regulate anthocyanin biosynthesis in hot pepper.

Key words: Anthocyanin, *Capsicum annuum*, gene expression, R2R3 MYB transcription factor.

INTRODUCTION

The MYB transcription factors family is one of the most abundant classes of transcription factors in plants (Rosinski and Atchley, 1998). MYB transcription factors are classified into three sub-families; MYB1R, MYBR2R3 and MYB3R factors; depending on the number of imperfect repeats (R1, R2 and R3) in the DNA-binding domain, each consisting of about 50 amino acids and including a helix-turn-helix structure (Jin and Martin, 1999).

Among these MYB transcription factors, R2R3-MYBs constitute the largest transcription factors gene family in plants, with 126 R2R3 MYB genes identified in Arabidopsis (Stracke et al., 2001). The R2R3-MYB genes have been extensively studied and members of the MYB family have been found to be involved in diverse physiological and biochemical processes including regulation of meristem formation, floral and seed development (Penfield et al., 2001; Schmitz et al., 2002; Shin et al., 2002; Steiner-Lange et al., 2003), the control of cell morphogenesis (Lee and Schiefelbein, 1999, 2001; Higginson et al., 2003), the control of the cell cycle (Ito et al., 2001; Araki et al., 2004), and regulation of secondary metabolism (Borevitz et al., 2000; Jin et al., 2000; Nesi et al., 2001; Baudry et al., 2004).

Some were also involved in various defense and stress responses (Vailleau et al., 2002; Abe et al., 2003; Denekamp and Smeekens, 2003; Nagaoka and Takano, 2003) and in light and hormone signaling pathways.
Plant anthocyanin pigmentation in leaves, flowers, and fruit imparts violet to black color. Anthocyanins are the end product of the flavonoid biosynthetic pathway. Anthocyanin pigments have key roles in plants for their function in attraction of pollinators and seed dispersers. These compounds also function as ultraviolet protectants, antimicrobial agents, and feeding deterrents, in signaling between plants and microbes, and in male fertility of some species (Stommel et al., 2009).

Hot pepper (Capsicum annuum L.) fruits are usually used as vegetable foods and as spice. Hot pepper fruit color is important for culinary product quality. Anthocyanin content affects fruit color of pepper during its ripening stage (Stommel et al., 2009). Anthocyanin structural gene transcription requires the expression of at least one of each of three distinct transcription factor families: MYC, MYB, and WD40 (Griesbach, 2005; Lightbourn et al. 2007). To study the MYB transcription factors controlling anthocyanin accumulation in hot pepper, three R2R3-MYB homologs were cloned and characterized from hot pepper. The expression profiles of CaMYBs and anthocyanin biosynthetic genes in the hot pepper were also investigated. This study perhaps contributes towards an understanding of the relation of CaMYBs and anthocyanin biosynthesis in the hot pepper.

**Materials and Methods**

**Plant material**

Hot pepper (C. annuum L.) cv. Haijiao 4 (green-fruited) was planted at the Experimental Farm of the Chinese Academy of Tropical Agriculture (Hainan, China). Fruits of four development stages (stage 1, young fruit at 7 days post anthesis; stage 2, young fruit at 15 days post anthesis; stage 3, green-ripe fruit at 30 days post anthesis; stage 4, red-ripe fruit at 50 days post anthesis), or flower and young leaves of hot pepper were collected. Fruit flesh and other tissues were cut into small pieces, and immediately frozen in liquid nitrogen, and stored at -70°C or were used immediately.

**Isolation of RNA**

Total RNA was extracted according to the Chang method (Chang et al., 1993). The quality and concentration of the extracted RNA was checked by agarose gel electrophoresis and measured by spectrophotometer (DU-70, Beckman, Fullerton, CA).

**Internal conserved fragment cloning of CaMYBs**

The degenerated primers, P1 (5'- CGGAATTCTTDDATYTCT RTTRTCTGNT -3') and P2 (5'- CGGAATTCDSSARAGYTGGCGC G -3'), were designed according to the conserved regions of R2R3 MYB transcription factor in the GenBank, where D is G,A or T, R is A or G, Y is C or T, S is G or C and N is A,C,G or T. Total RNA (1 µg) was used in reverse transcription PCR (one-step RT-PCR kit, TAKARA, Dalian, China), under the following condition: 1 h reverse transcription at 50°C, 2 min denaturation at 94°C, followed by 32 cycles of amplification (94°C for 30 s, 50°C for 30 s, 72°C for 30 s). The amplified product was purified (Tiangen, Beijing, China) and cloned into the pGEM-T easy vector (Promega, Madison, WI), followed by sequencing.

**3'-RACE of CaMYBs**

3'-ready cDNA was synthesized by reverse transcribing 1 µg of the total RNA with 3'-CDS primer A (provided in the kit). 3'-rapid amplification of cDNA ends (RACE) primer were designed (Table 1) and synthesized based on the cloned internal conserved fragment. 3'-RACE was carried out in a total volume of 25 µl containing 1 µl 3'-ready cDNA, and performed for 35 cycles of amplification (94°C for 15 s, 68°C for 30 s, and 72°C for 1 min). The product was purified and cloned into pGEM-T easy vector (Promega), followed by sequencing.

**5'-RACE of CaMYBs**

An aliquot of 1 µg of the total RNA was reverse transcribed with 5'-CDS primer A and SMART II A oligonucleotide (provided in the kit) to obtain the 5'-ready cDNA, using the SMARTTM RACE cDNA Amplification Kit (Clontech, Palo Alto, CA). 5'-RACE primer were designed (Table 1) and synthesized based on the cloned internal conserved fragment. Primary 5'-RACE PCR was performed with 5P1 and UPM primer (provided in the kit) in a total volume of 25 µl containing cDNA, and was denatured at 94°C for 3 min, followed by 32 cycles of amplification (5 s denaturation at 94°C, 10 s annealing at 68°C, and 1 min extension at 72°C). The product of the primary PCR was diluted 50-fold, and was used as a template for nested PCR performed with primer 5P2 and NUP (provided in the kit). The product was purified and cloned into the pGEM-T easy vector (Promega), and sequenced.

**Cloning of the full-length cDNA of CaMYBs**

After alignment and assembly of the sequences of the internal conserved sequence and the 3'-RACE and 5'-RACE products, the full-length cDNA sequence of the CaMYBs gene was deduced and subsequently obtained by PCR. The open reading frame (ORF) of CaMYBs was amplified by RT-PCR with the TAKARA RNA PCR kit. An aliquot of 1 µg of the total RNA was reverse transcribed according to the manufacturer’s manual, and 2 µl of the cDNA was used in PCR in a total volume of 50 µl under the following condition: 94°C for 2 min, followed by 32 cycles of amplification (94°C for 15 s, 58°C for 30 s, and 72°C for 2 min). The PCR product was purified and cloned into the pGEM-T easy vector (Promega) followed by sequencing.

**Multiple alignments and bioinformatic analyses**

Comparative and bioinformatic analyses of CaMYBs were carried out online at http://www.ncbi.nlm.nih.gov. The nucleotide sequence, deduced amino acid sequence, and ORF encoded by CaMYBs were analysed, and sequence comparison was conducted through database searches using the BLAST program (http://www.ncbi.nlm.nih.gov). The phylogenetic analysis of CaMYBs and MYB transcription factors from other species was aligned with CLUSTAL W (1.82) using default parameters. Phylogenetic tree was constructed using MEGA version 2.1 from CLUSTAL W alignments (Kumar et al., 2001). The neighbor-joining method was used to construct the tree (Saitou and Nei, 1987).

**Analysis of CaMYBs expression**

RT-PCR for the analysis of CaMYBs expression was performed
using total RNA from hot pepper tissues, and amplified with CaMYBs specific primers (Table 1). Specific primers were designed from the low homology regions of CaMYBs coding sequences and the 3’- untranslated region (UTR). The ACT gene was used as an internal control parallel in the reactions, amplified with ACT specific primers AF (5’-CAGTGGTCGACAACTGGTAT-3’) and AR (5’-ATCCCTCCAATCCAGACACTGTT-3’). PCR reaction was carried out in 22 cycles of programmed temperature control for 30 s at 95°C, 30 s at 50°C and 1 min at 72°C with a 5 min preheat at 95°C and a 10 min final extension at 72°C. The PCR products were analysed by agarose gel electrophoresis with ethidium bromide staining.

### Determination of anthocyanin content

Anthocyanin content was quantified by the pH differential spectrophotometry method (Niu et al., 2010). Three measurements for each biological replicate sample were performed.

<table>
<thead>
<tr>
<th>Primer Name</th>
<th>Sequence</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>3’ RACE-PCR primers</strong></td>
<td></td>
</tr>
<tr>
<td>3CAMYB11</td>
<td>GTTGCAGGCTGAGGTGGATAAATTACT</td>
</tr>
<tr>
<td>3CAMYB12</td>
<td>GTGGTTTAAGGAACAGAAAGAAAGAT</td>
</tr>
<tr>
<td>3CAMYB21</td>
<td>GCTGTAGATTGAGGTGGCTGAATTAT</td>
</tr>
<tr>
<td>3CAMYB22</td>
<td>GGCACGTATAAGAAGAGGTGACTTTGG</td>
</tr>
<tr>
<td>3CAMYB31</td>
<td>TCTTGCCGGTTACGATGGTGTAATTCG</td>
</tr>
<tr>
<td>3CAMYB32</td>
<td>TCTCCTCAAGGTAACATCGGCGCTTC</td>
</tr>
<tr>
<td><strong>5’ RACE-PCR primers</strong></td>
<td></td>
</tr>
<tr>
<td>5CAMYB11</td>
<td>TCTTCACTTATTACCAGTCTCTGTCGAG</td>
</tr>
<tr>
<td>5CAMYB12</td>
<td>ATTGGCTATTGGTGCCCATTGGTTC</td>
</tr>
<tr>
<td>5CAMYB21</td>
<td>CGAGTCCAGTAGTTTTCCACATCGT</td>
</tr>
<tr>
<td>5CAMYB22</td>
<td>GCAATCTCCAGCAATAAGTGACAC</td>
</tr>
<tr>
<td>5CAMYB31</td>
<td>CGGTGATTCTTTATTGCGTTATCGGTT</td>
</tr>
<tr>
<td>5CAMYB32</td>
<td>TCGCGCTTTAAGTAACTCGGCTATGT</td>
</tr>
<tr>
<td><strong>Semi-quantitative RT-PCR primers</strong></td>
<td></td>
</tr>
<tr>
<td>CaMYB11</td>
<td>GCAAGTCTTGCAGCATTGTCAC</td>
</tr>
<tr>
<td>CaMYB12</td>
<td>TTAACCATTTCGCCAGTTCGTC</td>
</tr>
<tr>
<td>CaMYB21</td>
<td>TTCTAGGACAGAGATGGTCAC</td>
</tr>
<tr>
<td>CaMYB22</td>
<td>CCCAGTAAGTACTGCTCTTC</td>
</tr>
<tr>
<td>CaMYB31</td>
<td>CCGTCAATGGATATATCTCATCTC</td>
</tr>
<tr>
<td>CaMYB32</td>
<td>ATCCCATGCGATTAGACGCCC</td>
</tr>
<tr>
<td>CaCHS1</td>
<td>CCTGCACCTCTAGTAACAGAC</td>
</tr>
<tr>
<td>CaCHS2</td>
<td>GTGACGCGATCCAGAAGATAGAG</td>
</tr>
<tr>
<td>CaCHI1</td>
<td>AGAGTGCCGTTCTTATCTCTGT</td>
</tr>
<tr>
<td>CaCHI2</td>
<td>TCCAGACCGCTTCTAGACAGTT</td>
</tr>
<tr>
<td>CaF3H1</td>
<td>GCACACTGATCGAGGACCAC</td>
</tr>
<tr>
<td>CaF3H2</td>
<td>TCTTCACTTGGTGTTACGTGCT</td>
</tr>
<tr>
<td>CaDFR1</td>
<td>CTCTTGCGCTTGGTCTCGAGAC</td>
</tr>
<tr>
<td>CaDFR2</td>
<td>ACGCTCGACGCTGCTCATT</td>
</tr>
<tr>
<td>CaDNS1</td>
<td>GCCTGTTCTTCCAGGTTCTTGTC</td>
</tr>
<tr>
<td>CaDNS2</td>
<td>CAATCAGTCTGTGCTCAGCGGT</td>
</tr>
<tr>
<td>CaUFGT1</td>
<td>ACAAGGCAATGAGCACCCTATT</td>
</tr>
<tr>
<td>CaUFGT2</td>
<td>TTGTGCGATGTCAGCTCTT</td>
</tr>
<tr>
<td><strong>ACT specific primers</strong></td>
<td></td>
</tr>
<tr>
<td>AF</td>
<td>CAGTGGTCGACAACTGGTAT</td>
</tr>
<tr>
<td>AR</td>
<td>TCCTCGACATTCCAGACACTGTT</td>
</tr>
</tbody>
</table>
Figure 1. The deduced amino acid sequences of CaMYBs are compared with anthocyanin-related MYB proteins of other species. Amino acid residues that are identical in all seven sequences are shaded darkly, while well-conserved residues are shaded in gray. The R2R3-binding domain is underlined. The box indicates specific residues that form the motif implicated in bHLH co-factor interaction. The accession numbers of these proteins, or translated products, in the GenBank database are as follows: AmMYB1 (Antirrhinum majus, ABB83826), VvMYBA1 (V. vinifera, AB242302), CaA (Capsicum annuum, AJ608992), ZmC1 (Zea mays, AAK81903) and FaMYB1 (Fragaria ananassa, AAK84064).

Analysis of anthocyanin biosynthetic genes expression

RT-PCR for the analysis of CaMYBs expression was performed using total RNA from hot pepper tissues, and amplified with anthocyanin biosynthetic genes specific primers (Table 1). The ACT gene was used as an internal control parallel in the reactions, amplified with ACT specific primers AF(5'-CAGTGGTCGA CAACTTGAT-3') and AR (5'-ATCCTCCAATCCAGACACTT-3'). PCR reaction was carried out in 25 cycles of programmed temperature control for 30 s at 95°C, 30 s at 50°C and 1 min at 72°C with a 5 min preheat at 95°C and a 10 min final extension at 72°C. The PCR products were analysed by agarose gel electrophoresis with ethidium bromide staining.

RESULTS

Cloning and characterization of CaMYBs

About 180-bp PCR products were amplified with degenerated primers P1 and P2, and then PCR product was purified and cloned into the pGEM-T Easy vector following by sequencing. Based on sequences from 20 individual recombinant plasmids, three candidate fragments were obtained, which showed similarity with other plant R2R3 MYB transcription factor genes as revealed by a BlastX search. After sequence extension by 3'-RACE and 5'-RACE, sequences of 1093, 1086 and 1045 bp in length were obtained for CaMYB1, CaMYB2 and CaMYB3, respectively.

CaMYB1, CaMYB2 and CaMYB3 encode polypeptides consisting of 340, 262 and 345 amino acids respectively; containing R2R3 domain and the signature motif specific for the interaction between MYB and bHLH proteins in the R3 domain (Figure 1). Blast analysis revealed that deduced protein sequences are already deposited in database at NCBI (CaMYB1 - AAO5796.1; CaMYB2 - CAE7545.1; CaMYB3 - ABO11121.1). Phylogenetic analysis based on the deduced amino acid sequences of these three R2R3 MYB transcription factor members revealed that CaMYB1 and CaMYB2 clustered together with the anthocyanin-related subgroup of R2R3 MYB proteins from other plants, while CaMYB3 did not (Figure 2). CaMYB1 and CaMYB2 maybe regulate anthocyanin biosynthesis in hot pepper.

Expression of CaMYBs transcription factors during fruit ripening

In order to study the CaMYBs transcription pattern, total RNA was isolated from young leaves, flower, and the four
Figure 2. Phylogenetic analysis of CaMYBs and other species by MEGA version 2.1 from CLUSTAL W alignments. The neighbor-joining method was used to construct the tree. The anthocyanin-related MYB proteins of other species used in the evolutionary analysis are retrieved from Genbank including InMYB2 (Ipomoea nil, BAE94709), MaMYB (Mimulus aurantiacus ACA04006), OgMYB1 (Oncidium Gower Ramsey, ABS58501), GhMYB10 (Gerbera hybrid, CAD87010), AmROSEA1 (Antirrhinum majus, ABB83826), AmROSEA2 (Antirrhinum majus, ABB83827), MdMYB8 (Malus domestica, ABB84756.1), MdMYB10 (Malus domestica, ABB84753.1), VvMYBA1 (V. vinifera, AB242302), VvMYBA2 (V. vinifera, AB097924), ZmC1 (Zea mays, AAK81903), FaMYB1 (Fragaria ananassa, AAK84064), and CaA (Capsicum annuum, AJ608992).

Figure 3. Expression of CaMYBs transcription factors in young leaves (1), flower (2), and the four stages of fruit development (3 to 6).

Anthocyanin content and expression of biosynthetic genes during fruit ripening

Anthocyanin content varied among the four stages of fruit development, and subjected to semi-quantitative RT-PCR analysis. CaMYBs transcripts accumulation was detected in all stages of fruit development and in flower and leaves. Three CaMYBs transcription factors showed differential expression during fruit ripening (Figure 3). The highest expression level of CaMYB1 during fruit development was detected in DAF30 fruit, the lowest expression level of CaMYB1 was detected in ripe fruit. A high level of expression of CaMYB1 was also detected in leaf tissue. The profile of CaMYB2 was consistent of CaMYB2 during fruit development, but the expression level of CaMYB1 was higher than those of CaMYB1 during fruit development. A high level of expression of CaMYB2 was detected in flower. The expression level of CaMYB3 decreased during fruit development.
development. As shown in Figure 4a, the content was low in ripe fruit and high in DAF30 fruit. Anthocyanin biosynthetic gene expression patterns were quite different in young leaves, flower, and the four stages of fruit development (Figure 4b).

**DISCUSSION**

R2R3 MYB genes have been cloned and characterized from horticulture crops, such as grape (Kobayashi et al., 2002), apple (Ban et al., 2007), pear (Feng et al., 2010), Chinese bayberry (Niu et al., 2010), mangosteen (Palapol et al., 2009), strawberry (Aharoni et al., 2001), and a few other species. The first characterized plant R2R3-MYB was C1 from maize, which regulates genes encoding enzymes of the anthocyanin biosynthetic pathway (Cone et al., 1993). Similar to C1, some MYB transcription factors have been proved to regulate anthocyanin biosynthesis (Ban et al., 2007; Feng et al., 2010; Niu et al.,...
In this study, we reported on the cloning and characterization of three genes encoding R2R3 MYB protein. The deduced amino acid sequence of CaMYBs showed extensive similarity to their counterparts in other species, containing R2R3 domain and the signature motif specific for the interaction between MYB and bHLH proteins in the R3 domain. Phylogenetic analysis based on the deduced amino acid sequences of these three R2R3 MYB transcription factor members revealed that CaMYB1 and CaMYB2 clustered together with the anthocyanin-related subgroup of R2R3 MYB proteins from other plants, while CaMYB3 did not. In addition, the deduced amino acid of CaMYB2 shared 96% identity with the CaA, controlling anthocyanin accumulation encode a MYB transcription factor in pepper (purple-fruited) (Borovsky et al., 2004). Based on sequence similarity, we hypothesize that CaMYB1 and CaMYB2, maybe regulate anthocyanin biosynthesis in hot pepper (green-fruited).

Anthocyanin content varied among the four stages of fruit development in hot pepper (C. annuum L.) cv. Haijiao 4 (green-fruited), the content was low in ripe fruit and high in DAF30 fruit. Only after about 10 days post anthesis does anthocyanin become visible, peaking 20 days post-anthesis in pepper S226 (purple-fruited). At ripening, anthocyanin disappears and the fruit turns orange due to the accumulation of carotenoid pigments (Borovsky et al., 2004). Plant anthocyanin biosynthetic genes have been identified or cloned in many plant species (Grotewold, 2004). Plant anthocyanin biosynthetic genes have been characterized of three genes encoding R2R3 MYB transcription factor gene. In this study, CaMYB1 and CaMYB2 expression, suggesting CaMYB1 and CaMYB2 may activate the expression of CaDFR.

In this study, the transcription of CaMYB1 and anthocyanin biosynthetic genes declined during late stages of fruit development, in agreement with the lower efficiency of anthocyanin synthesis and the degradation of anthocyanin may explain the fast decline in anthocyanin content. In summary, we have cloned three genes encoding R2R3 MYB transcription factors from hot pepper. Based on sequence similarity and expression profile, we hypothesize that CaMYB1 and CaMYB2 may regulate anthocyanin biosynthesis in hot pepper.

REFERENCES


