

Differences in the *CaMYBA* Genome Between Anthocyanin-pigmented Cultivars and Non-pigmented Cultivars in Pepper (*Capsicum annuum*)

Sho Ohno^{1*}, Maiko Ueno¹ and Motoaki Doi¹

¹Graduate School of Agriculture, Kyoto University, Kyoto 606-8502, Japan

Anthocyanin in pepper is beneficial as a food antioxidant compound and as a pigment for ornamentals, while unexpected anthocyanin accumulation in fruit, known as black spots, reduces the commercial quality of some cultivars. Previous studies demonstrated that the *Anthocyanin* (*A*) locus determines the anthocyanin accumulation in pepper fruits, and an MYB transcription factor, *CaMYBA*, was found to be located near the *A* locus. However, the causal gene sequence of the *A* locus has not yet been identified. With progress regarding genome information in pepper, two other homologous MYB genes were found to be located near *CaMYBA*, and they are also considered to be candidate genes for the *A* locus. In this study, we attempted to identify the causal gene sequence of the *A* locus by performing linkage analysis, genomic sequence analysis, and gene expression analysis of the three candidate MYB genes. A crossing experiment between pigmented ‘Peruvian Purple’ and non-pigmented cultivars confirmed that anthocyanin accumulation in the pigmented cultivar was controlled by a single locus. Gene expression analysis demonstrated that a basic helix-loop-helix transcription factor, *CaMYC*, and *CaMYBA* were expressed abundantly in pigmented cultivars, but the other two MYB genes were not. Genotyping of the F₂ population derived from the cross demonstrated that the anthocyanin accumulation phenotype was highly linked to *CaMYBA*, but not to *CaMYC*. The DNA sequence of *CaMYBA* in pigmented cultivars had an insertion of a 4.3 kb retrotransposable element *LINE-1* in the first intron, but that of non-pigmented cultivars did not. No pigmented cultivar-specific sequence was found in the promoter region of *CaMYBA*. Therefore, it was suggested that *CaMYBA*, but not the other two homologous MYB genes, is the *A* locus gene, and insertion of *LINE-1* in *CaMYBA* appeared to be important for the regulation of anthocyanin accumulation, although the mechanism by which the *LINE-1* insertion induces *CaMYBA* expression is unknown.

Key Words: delphinidin, LINE-1 retrotransposable element, MYB transcription factor.

Introduction

Peppers (*Capsicum*) belong to the *Solanaceae* family and originated in South and Central America. Generally, immature fruits of *C. annuum* are consumed as vegetables and mature fruits are consumed as spices all over the world. In addition, some *Capsicum* plants have purple flowers, foliage, and colourful fruits, and are used as ornamental plants.

The purple pigment of pepper is anthocyanin, a flavonoid compound, and its chemical structure was re-

ported as delphinidin-3-(4-p-coumaroyl)-rutinoside-5-glucoside (Lightbourn et al., 2008). Anthocyanin accumulation in pepper is beneficial in terms of antioxidant capacity when consumed or as a pigment for ornamentals. However, unexpected anthocyanin accumulation in fruit causes black spots and could be a problem by reducing the commercial quality of some cultivars.

Anthocyanin is one of the most studied secondary metabolites in plants. In the flavonoid biosynthetic pathway, anthocyanin is synthesized from two substrate compounds, malonyl-CoA, and *p*-coumaroyl-CoA, by chalcone synthase (CHS), chalcone isomerase (CHI), flavanone 3-hydroxylase (F3H), dihydroflavonol 4-reductase (DFR), anthocyanidin synthase (ANS), and flavonoid UDP-glucosyltransferase (Tanaka et al., 2008). Although the regulatory mechanisms depend on the species, the expression of some enzyme genes are

Received; April 10, 2019. Accepted; August 19, 2019.
First Published Online in J-STAGE on November 15, 2019.
Part of this study was presented at the 2018 Spring Meeting and 2019 Spring Meeting of the Japanese Society for Horticultural Science.
No conflicts of interest declared.
* Corresponding author (E-mail: sohno@kais.kyoto-u.ac.jp).

regulated by transcription factors such as MYB, a basic helix-loop-helix (bHLH), and WD repeats (WDR) (Koes et al., 2005). Anthocyanin biosynthesis in pepper is controlled by mainly one locus known as the *Anthocyanin* (*A*) locus, which is an incomplete dominant gene for anthocyanin color in the foliage, flowers and immature fruit (Wang and Bosland, 2006). This *A* locus was mapped to chromosome 10 (Chaim et al., 2003), and afterwards mapped near to an MYB transcription factor (*CaMYBA*) that is orthologous to *PhAN2* in petunia (*Petunia hybrida*) (Borovsky et al., 2004).

There are many reports analysing anthocyanin biosynthetic genes in pepper (Aza-González et al., 2013; Li et al., 2011; Lightbourn et al., 2007), but the causal gene sequence for anthocyanin pigmentation has not yet been identified. Recent reports in petunia and potato (*Solanum tuberosum*) indicate that several paralogous MYB genes (*PhAN2*, *PhAN4*, *PhPHZ*, and *PhDPL* in petunia, *StAN1*, *StMYBA1*, and *StMYB113* in potato) can regulate anthocyanin biosynthesis (Albert et al., 2011; Liu et al., 2016). Genomic DNA sequences of several pepper cultivars, ‘Criollos de Morelos 334 (CM334)’ (Kim et al., 2014), ‘Zunla-1’, Chiltepin (*C. annuum* var. *glabriusculum*) (Qin et al., 2014), and UCD10X (Hulse-Kemp et al., 2018), suggested that three MYB genes, including *CaMYBA*, are located near each other on the same chromosome. For example, in ‘CM334’ (version 1.55), *CaMYBA*, *CaMYBB*, and *CaPHZ* were allocated as CA10g11650 (Chr10 182955325..182956589), CA10g11690 (Chr10 184012261..184013487), and CA10g11710 (Chr10 184898196..184899719), respectively. Therefore, not only *CaMYBA*, but also the two homologous MYB genes, are considered as candidate genes for the *A* locus. In this study, we performed linkage analysis, genomic DNA sequence analysis and gene expression analysis of the three candidate MYB genes to identify the *A* locus gene sequence in anthocyanin-pigmented cultivars and non-pigmented cultivars of pepper.

Materials and Methods

Plant materials

C. annuum anthocyanin-pigmented cultivars ‘Peruvian Purple’, ‘Murasaki’, and ‘Purple Flash’, and non-pigmented cultivars ‘Takanotsume’, ‘Nikko’, and ‘Sapporo-Onaga’ were used for the experiments (Fig. 1A–F). For genomic PCR, anthocyanin-pigmented cultivars ‘Conga’ and ‘Nishiki-sango’, and non-pigmented cultivars ‘Shishihomare’ and ‘Manganji’ were also used. All plants were grown in the field or greenhouses (heated to keep the minimum temperature > 5°C from November to April) of the experimental farm at Kyoto University (Kyoto, Japan). For pot cultivation, plants were grown in 24-cm pots filled with mixed soil (akadama soil:leaf mold = 7:3) and IB Kasei (JCAM AGRIC. CO., LTD., Tokyo, Japan) (N:P:K = 10:10:10) was used as a fertilizer.

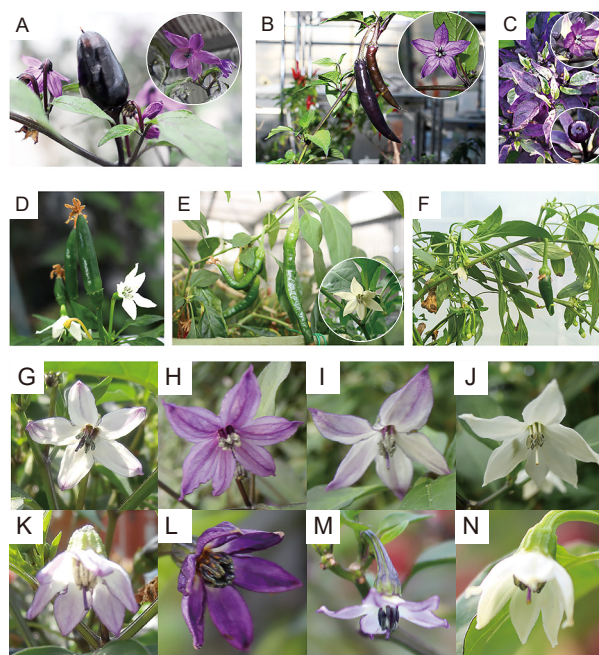


Fig. 1. Pepper cultivars used in this study. (A) ‘Peruvian Purple’, (B) ‘Murasaki’, (C) ‘Purple Flash’, (D) ‘Takanotsume’, (E) ‘Nikko’, (F) ‘Sapporo-Onaga’, (G) Bicolor flower of ‘Peruvian Purple’ crossed with ‘Takanotsume’ F₁, (H) Purple flower of ‘Peruvian Purple’ crossed with ‘Takanotsume’ F₂, (I) Bicolor flower of ‘Peruvian Purple’ crossed with ‘Takanotsume’ F₂, (J) White flower of ‘Peruvian Purple’ crossed with ‘Takanotsume’ F₂, (K) Bicolor flower of ‘Peruvian Purple’ crossed with ‘Nikko’ F₁, (L) Purple flower of ‘Peruvian Purple’ crossed with ‘Nikko’ F₂, (M) Bicolor flower of ‘Peruvian Purple’ crossed with ‘Nikko’ F₂ and (N) White flower of ‘Peruvian Purple’ crossed with ‘Nikko’ F₂.

HPLC analysis

The pigment compositions of flowers and fruits were measured by high-performance liquid chromatography (HPLC). The HPLC analysis process was according to Ohno et al. (2011). Delphinidin chloride (Nagara Science, Gifu, Japan) was used as a standard anthocyanin.

Crossing experiment

‘Peruvian Purple’ was crossed with ‘Takanotsume’ or ‘Nikko’, and then the F₁ and F₂ generations were produced. Segregation of pigmentation traits in flowers and fruits was recorded in both generations and they were used for genotyping.

Gene expression analysis

Total RNA was extracted from flowers and fruit peels using Sepasol RNA I Super G (Nacalai Tesque, Kyoto, Japan), and purified by precipitation with a High-salt solution (Takara Bio Inc., Shiga, Japan). It was reverse transcribed with ReverTra Ace (TOYOBO Co., Ltd., Osaka, Japan), and 1 µL of the resultant cDNA was diluted 10-fold and used as a template for semi-quantitative RT-PCR. Semi-quantitative RT-PCR

was performed with Blend Taq polymerase (TOYOBO) using the primers shown in Supplemental Table S1. The PCR program was set at 94°C for 2 min, followed by 30 cycles of 94°C for 30 s, 55°C for 30 s, and 72°C for 1.5–2.5 min.

Genotyping analysis of CaMYBA and CaMYC in the 'Peruvian Purple' × 'Takanotsume' populations

To genotype the *CaMYBA* allele, PCR-RFLP was performed. The *CaMYBA* of 'Takanotsume' has a *HaeIII* restriction enzyme digestion site in the fourth exon (GGCC), but 'Peruvian Purple' does not (GACC). PCR was performed using Blend Taq polymerase (TOYOBO) using the primers Forward: 5'-GAAGCTAATAACTGCTCCTCATCGA-3' and Reverse: 5'-CTTACATTGAAGATGCGTGGA-3'. Then, the PCR product was digested with *HaeIII* (Takara Bio) for 1 hr at 37°C and electrophoresed. As a result, the band of the *CaMYBA* allele from 'Peruvian Purple' was 615 bp, while the bands of the *CaMYBA* allele from 'Takanotsume' were 426 bp and 189 bp. To genotype the *CaMYC* allele, a PCR length polymorphism in the promoter region was used. 'Peruvian Purple' has two repeats of 119 bp sequence, while 'Takanotsume' has only one 116 bp sequence. PCR was performed using Blend Taq polymerase (TOYOBO) using the primers (Forward: 5'-CGTAGTAAGGAGTGTAACCTTTGAT-3' and Reverse: 5'-AATAGTCATCAACATGAACCTTCAT-3') and electrophoresed. The PCR program for genotyping was set at 94°C for 2 min, followed by 30 cycles of 94°C for 30 s, 55°C for 30 s, and 72°C for 1.5–2 min.

Genomic analysis of CaMYBA

Genomic DNA was extracted from leaves using MagExtractor Plant Genome (TOYOBO). To clone the full length *CaMYBA* genome, the PCR product was amplified with Primer-Forward: 5'-GATATCATGAATACTGCTATT-3' and Primer-Reverse: 5'-CTTACATTGAAGATGCGTGGA-3' using Blend Taq polymerase (TOYOBO) and cloned into a pTAC-1 vector (BioDynamics Laboratory Inc., Tokyo, Japan). The PCR program was set at 94°C for 2 min, followed by 30 cycles of 94°C for 30 s, 55°C for 30 s, and 72°C for 3 min. To clone the insertion sequence in the *CaMYBA* genome, the PCR product amplified with Forward: 5'-A AAAAGCAGGCTCGGAAGGAGGCTAGTCAAATC AAAT-3' and Reverse: 5'-AGAAAGCTGGGTAATAG CAGTATTCATGATATC-3' using a KOD FX Neo polymerase (TOYOBO) was cloned into a pDONR221 vector using Gateway™ BP Clonase™ II Enzyme mix (Invitrogen, Carlsbad, CA, USA). The PCR program was set at 94°C for 2 min, followed by 35 cycles of 98°C for 10 s, 55°C for 30 s, and 68°C for 8 min. To clone the promoter region, two primer sets (Forward: 5'-CACATGTAGGATGATAGCTTCATGA-3' and Reverse: 5'-GGGACGTGGATGGTTAGCAATCACT-3',

Forward: 5'-AATAACCAACCGTTCATCATATCT-3' and Reverse: 5'-CTTTTAAGAGCTGTGTCCCTTGC T-3') were designed to amplify overlapping regions based on the genome sequence of 'CM334'. Primers used for sequencing are listed in Supplemental Table S2. For genomic PCR of *CaMYBA* among different cultivars, PCR was performed with the above conditions using a KOD FX Neo polymerase (TOYOBO) with primers (Forward: 5'-CGGAAGGAGGCTAGTCA AATCAAAT-3' and Reverse: 5'-AATAGCAGTATTCA TGATATC-3'). For genomic PCR of *CaMYBA* among 'Peruvian Purple' × 'Nikko' F₂ populations, PCR was performed with the above conditions except the extension time was changed to 3 min using KOD FX Neo polymerase (TOYOBO) with primers (Forward: 5'-AC CTTACAAATATATACAAAATTT-3' and Reverse: 5'-AATAGCAGTATTTCATGATATC-3'). The accession number of the 'Peruvian Purple' *CaMYBA* genome, including the promoter region, is deposited as LC473089.

Results

Pigmentation locus of 'Peruvian Purple'

From HPLC analysis, pigmented cultivars accumulated delphinidin-based anthocyanin in their flowers and fruits, whereas non-pigmented cultivars did not accumulate any anthocyanins (Table 1).

In order to confirm that the pigmentation locus was identical to the *A* locus, 'Peruvian Purple', which has purple foliage, purple flowers and purple fruits was crossed with 'Takanotsume' or 'Nikko'. As for flower color, all F₁ plants of 'Peruvian Purple' × 'Takanotsume' had bicolor flowers (Fig. 1G), suggesting incomplete dominance of the pigmentation gene. Pigmentation in the F₂ generation of 'Peruvian Purple' × 'Takanotsume' was found to be Purple:Bicolor: White = 38:87:47 ($\chi^2 = 0.97$, $P > 0.05$) (Fig. 1H–J; Table 2). The segregation ratio was not significantly different from 1:2:1, suggesting one locus is associated with pigmentation of these parent lines. Similar results were obtained from a crossing between 'Peruvian Purple' and 'Nikko'. All F₁ plants had bicolor flowers (Fig. 1K), and pigmentation in the F₂ generation was Purple:Bicolor:White = 35:51:20 ($\chi^2 = 4.40$, $P > 0.05$) (Fig. 1L–N; Table 2). As for fruit color, all F₁ plants had purple pigmentation and F₂ plants had Purple pig-

Table 1. Anthocyanin accumulation in flowers and fruits.

Cultivar	Flower	Fruit
Peruvian Purple	+	+
Murasaki	+	+
Purple Flash	+	+
Takanotsume	–	–
Nikko	–	–
Sapporo-Onaga	–	–

+: Abundant, –: Not detected

mentation:Green = 118:54 ($\chi^2 = 3.75$, $P > 0.05$) in ‘Peruvian Purple’ × ‘Takanotsume’, and Purple pigmentation:Green = 81:25 ($\chi^2 = 0.11$, $P > 0.05$) in ‘Peruvian Purple’ × ‘Nikko’ (Table 3). The ratio was not significantly different from Purple pigmentation:Green = 3:1. Therefore, it was suggested that one locus regulates anthocyanin pigmentation in ‘Peruvian Purple’.

Flavonoid biosynthetic gene expression analysis

Semi-quantitative RT-PCR was performed for gene expression analysis. In flowers, as for enzymatic genes compared to pigmented cultivars, *CaF3'5'H*, *CaDFR*, and *CaANS* were down-regulated in non-pigmented cultivars (Fig. 2A). Downregulation of multiple structural enzyme genes suggested the involvement of transcription factors. We analyzed three MYB genes (*CaMYBA*, *CaPHZ*, and *CaMYBB*), two bHLH transcription factors (*CaMYC* and *CaJAF13*) and one WDR transcription factor (*CaWDR*), and only *CaMYBA* and *CaMYC* were down-regulated in non-pigmented cultivars (Fig. 2A). Similar data was obtained from fruits (Fig. 2B). These data suggested that *CaPHZ* and *CaMYBB* were not important for pigmentation of the cultivars used in this experiment, and that *CaMYBA* or *CaMYC* could be the anthocyanin determinant locus.

Genomic segregation analysis of *CaMYBA* and *CaMYC*

The genomic location of *CaMYBA* and *CaMYC* was analyzed using the pepper genome database. This indicated that *CaMYBA* is located on chromosome 10,

while *CaMYC* is located on chromosome 9. Using ‘Peruvian Purple’ and ‘Takanotsume’ F₂ populations exhibiting different pigmentation, PCR segregation analysis was performed. For *CaMYC*, PCR was performed using a sequence length polymorphism found in the promoter region between ‘Peruvian Purple’ and ‘Takanotsume’. When PCR was performed, however, no correlation was found between the genotype of *CaMYC* and anthocyanin pigmentation (Fig. 3), suggesting that *CaMYC* is not the determinant factor for pigmentation. For *CaMYBA*, PCR-RFLP was performed using a single sequence polymorphism found in the fourth exon between ‘Peruvian Purple’ and ‘Takanotsume’. When PCR-RFLP was performed, the genotype of *CaMYBA* and pigmentation phenotype were totally co-segregated (Fig. 3). Considering the results of semi-quantitative RT-PCR and this segregation analysis, *CaMYBA* is most likely to be the *A* locus, as reported previously (Borovsky et al., 2004).

Difference in the *CaMYBA* genome between pigmented cultivars and non-pigmented cultivars

According to the ‘CM334’ genome sequence from the NCBI database, the cDNA of *CaMYBA* (AJ608992) was 1096 bp long and composed of four exons and three introns (Fig. 4A). When we performed genomic PCR using primers amplifying the full length coding region, we unexpectedly could only amplify full length genome fragments from non-pigmented cultivars, but not from pigmented ones. Since several PCR analyses indicated that an insertion sequence is likely to be in the

Table 2. Phenotypic segregation of anthocyanin pigmentation in flowers.

Parental cultivar and cross combination	Population size	Number of plants			Expected ratio (Purple:Bicolor:White)	Chi-square (P-value)
		Purple	Bicolor	White		
P1: Peruvian Purple	10	10	0	0		
P2: Takanotsume	10	0	0	10		
P3: Nikko	10	0	0	10		
F1: (P1×P2)	8	0	8	0	0:1:0	
F1: (P1×P3)	8	0	8	0	0:1:0	
F2: (P1×P2)	172	38	87	47	1:2:1	$\chi^2 = 0.97$ ($P = 0.62$)
F2: (P1×P3)	106	35	51	20	1:2:1	$\chi^2 = 4.40$ ($P = 0.15$)

Table 3. Phenotypic segregation of anthocyanin pigmentation in fruits.

Parental cultivar and cross combination	Population size	Number of plants		Expected ratio (Purple:Green)	Chi-square (P-value)
		Purple	Green		
P1: Peruvian Purple	10	10	0		
P2: Takanotsume	10	0	10		
P3: Nikko	10	0	10		
F1: (P1×P2)	8	8	0	1:0	
F1: (P1×P3)	8	8	0	1:0	
F2: (P1×P2)	172	118	54	3:1	$\chi^2 = 3.75$ ($P = 0.053$)
F2: (P1×P3)	106	81	25	3:1	$\chi^2 = 0.11$ ($P = 0.74$)

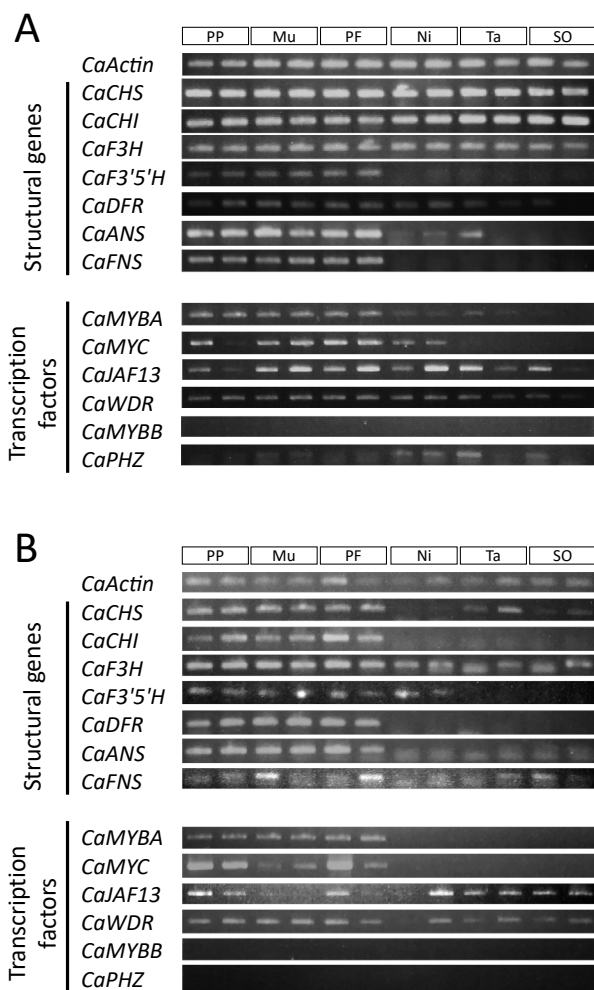


Fig. 2. Semi-quantitative RT-PCR for flavonoid biosynthetic genes. (A) Flower, (B) Fruit. Two biological replicates for each cultivar were performed. Abbreviations for cultivar names: PP, Peruvian Purple; Mu, Murasaki; PF, Purple Flash; Ni, Nikko; Ta, Takanotsume, and SO, Sapporo-Onaga.

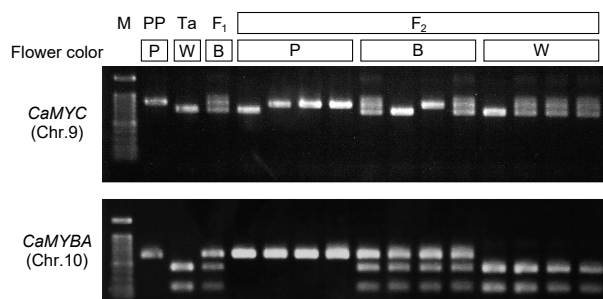


Fig. 3. Genotyping of *CaMYBA* and *CaMYC* in the 'Peruvian Purple' crossed with 'Takanotsume' population. Abbreviations for cultivar names: PP, Peruvian Purple and Ta, Takanotsume. Abbreviations for flower colors: P, purple, B, bicolor, and W, white. M indicates 100 bp marker.

first intron, long-PCR was performed. We could amplify an approximately 4.3 kb longer sequence in pigmented cultivars than non-pigmentation cultivars (Fig. 4B), and this insertion sequence of 'Peruvian

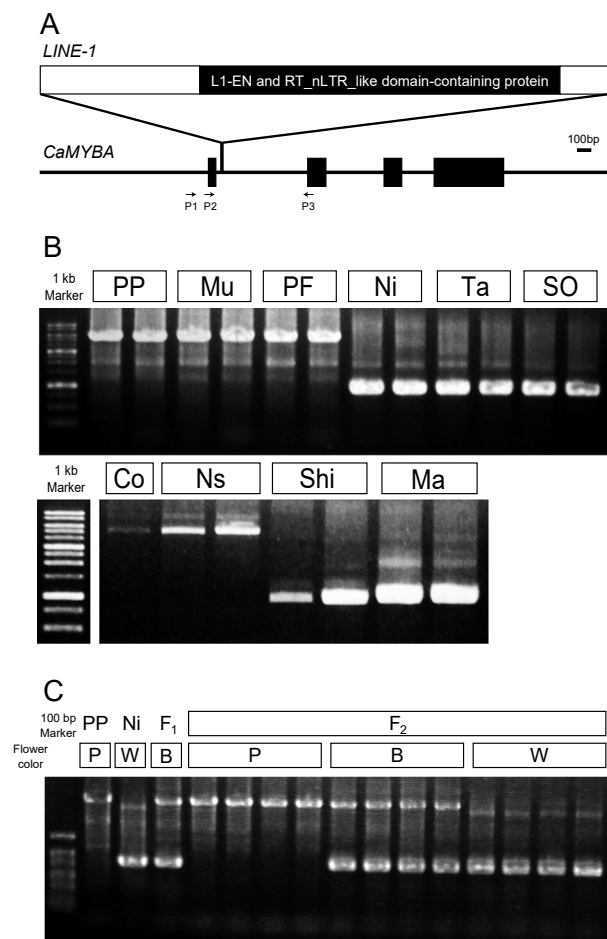


Fig. 4. Structure of the *CaMYBA* gene. (A) Structure of the *CaMYBA* gene of 'Peruvian Purple'. Black rectangles indicate exons. Arrows in P1 and P3 indicate the position of primers used in (B) and Arrows in P2 and P3 indicate the position of primers used in (C). (B) Genomic PCR among pigmented cultivars and non-pigmented cultivars. Two biological replicates for each cultivar were performed. (C) Genomic PCR among 'Peruvian Purple' x 'Nikko' F₂ populations. Abbreviations for cultivar names: PP, Peruvian Purple; Mu, Murasaki; PF, Purple Flash; Ni, Nikko; Ta, Takanotsume; SO, Sapporo-Onaga; Co, 'Conga'; Ns, 'Nishiki-sango'; Shi, 'Shishihomare' and Ma, 'Manganji'. Abbreviations for flower colors: P, purple, B, bicolor, and W, white.

Purple' was cloned and sequenced. The full length of this insertion sequence was 4321 bp, and it contained an open reading frame for sense orientation and encoding a 907 amino acid polypeptide with a non-LTR (long terminal repeat) retrotransposon *LINE-1* endonuclease domain (cd09076) and a non-LTR retrotransposon reverse transcriptase domain (cd01650) (Fig. 4A), which are typical of *LINE L1* retrotransposons (Wicker et al., 2007). This polypeptide sequence shared high identity (53%) with *Zea mays* Retrovirus-related Pol polyprotein *LINE-1* (ONM52836). We could detect an 18 bp (TGTTGTTTAAATTTTGGA) target site duplication sequence, but no terminal inverted repeat-like sequence was found. To confirm the relationship between the

LINE-1 insertion in *CaMYBA* and anthocyanin pigmentation, genomic PCR in the ‘Peruvian Purple’ × ‘Nikko’ F_2 population was performed. We could detect a longer band from ‘Peruvian Purple’ and F_2 populations that had purple flowers, a short band from ‘Nikko’ and F_2 populations that had white flowers, and both bands from F_1 and F_2 populations that had bicolor flowers (Fig. 4C).

To investigate whether there is a polymorphism in the *CaMYBA* promoter region specific to pigmented cultivars, we isolated and compared a region of approximately 1300 bp upstream from the putative transcription start site of the *CaMYBA* gene among six cultivars; however, there were no sequence polymorphisms specific to pigmented cultivars (Supplemental Fig. S1). Therefore, it was suggested that this *LINE-1* insertion is important for the pigmentation allele of *CaMYBA*.

Discussion

Anthocyanin accumulation in pepper was reported to be determined by the *A* locus, and this locus was tightly mapped to *CaMYBA* (Borovsky et al., 2004). However, it remained unclear whether the causal gene of anthocyanin pigmentation in pepper was *CaMYBA* because the draft genome of pepper ‘CM334’ indicates that there are at least two homologous MYB genes (*CaMYBB* and *CaPHZ*) that are located near *CaMYBA* on chromosome 10. It was reported that four homologous MYB genes (*PhAN2*, *PhAN4*, *PhDPL*, and *PhPHZ*) in petunia, and three homologous MYB genes (*StAN1*, *StMYBA1*, and *StMYB113*) in potato regulate anthocyanin biosynthesis (Albert et al., 2011; Liu et al., 2016). A sequence comparison indicated that *CaMYBA*, *CaMYBB*, and *CaPHZ* are orthologous to *StMYB113*, *StAN1*, and *StMYBA1*, respectively (Supplemental Fig. S2). Therefore, it was considered that not only *CaMYBA*, but also *CaMYBB* and *CaPHZ*, were candidate genes for the *A* locus. From semi-quantitative RT-PCR, expressions of *CaMYBB* and *CaPHZ* were not detected from pigmented cultivars, while abundant expression of *CaMYBA* was detected from pigmented cultivars (Fig. 2). Therefore, it was suggested that *CaMYBB* and *CaPHZ* were not the determinant factors of anthocyanin biosynthesis, and that *CaMYBA* is a strong candidate for the *A* locus.

Previously, virus-induced gene silencing (VIGS) of the *CaMYBA* gene induced down-regulation of several enzyme genes (Aguilar-Barragán and Ochoa-Alejo, 2014; Zhang et al., 2015) and *CaMYC* (Zhang et al., 2015). However, VIGS of *CaMYC* reduced the expression levels of several enzyme genes, but not *CaMYBA* (Lu et al., 2019). This suggested that *CaMYBA* regulated expression of not only flavonoid biosynthetic enzyme genes, but also *CaMYC*, similar to the overexpression of the potato MYB gene *StAN1* in tobacco, which induces endogenous bHLH *NtAN1a* and *NtAN1b* (Liu et al., 2016). In pigmented cultivars, ex-

pression levels of both *CaMYBA* and *CaMYC* were up-regulated in flowers (Fig. 2A) and fruit peels (Fig. 2B); however, only the *CaMYBA* genotype was linked to the pigmentation (Fig. 3), suggesting that *CaMYBA* functions upstream of *CaMYC*.

Several reports have mentioned the role of retrotransposable elements in anthocyanin biosynthesis. In grape (*Vitis vinifera*), a Ty3-gypsy-type retrotransposon *Gret1* inserted in the 5'-flanking region of *VvmybA* blocks *VvmybA* expression (Kobayashi et al., 2004). In Sicilian blood orange ‘Tarocco’ (*Citrus sinensis*), a Copia-like retrotransposon *Tcs1* inserted adjacent to a gene encoding Ruby, an MYB transcriptional activator of anthocyanin production, controls fruit-specific and cold-inducible *Ruby* expression (Butelli et al., 2012). In this case, the start of *Ruby* transcription lies within the LTR of *Tcs1*, indicating that the LTR sequence functions as a promoter and transcription start site. In apple (*Malus domestica*), a gypsy-type retrotransposon insertion upstream of *MdMYB1*, a core transcriptional activator of anthocyanin biosynthesis, is associated with a red-skinned phenotype (Zhang et al., 2019). Here, all pigmented cultivars and purple flower plants of the ‘Peruvian Purple’ × ‘Nikko’ F_2 population had a *LINE-1* insertion in the first intron of *CaMYBA*, while all non-pigmented cultivars and white flower plants of the ‘Peruvian Purple’ × ‘Nikko’ F_2 population did not (Fig. 4B, C). The sequence of the *CaMYBA* promoter region was almost the same among cultivars tested in this study (Supplemental Fig. S1), indicating the difference in expression level of *CaMYBA* is due to a *LINE-1* insertion in the first intron. Thus, two possibilities can be considered for the *A* locus: one is that *CaMYBA* is identical to the *A* locus, and the insertion of *LINE-1* into the first intron induces *CaMYBA* expression. The other is that *LINE-1* insertion is not related to the induction of *CaMYBA* expression and the *A* locus encodes another transcription factor which regulates *CaMYBA* expression. In the latter case, the gene must be located close to *CaMYBA* because the *CaMYBA* genotype perfectly coincided with the pigmentation phenotype (Figs. 3 and 4; Borovsky et al., 2004). Although, the latter possibility cannot be excluded, the former hypothesis is the more likely. As for the former possibility, the question arises as to how the *LINE-1* inserted allele can express more *CaMYBA* than the non-inserted allele, despite the fact that both alleles contain the full coding region. The possibility that a *LINE-1* insertion created a novel promoter is excluded because expression of the first exon of *CaMYBA* was detected (Supplemental Fig. S3). It is known that intron-mediated enhancement of gene expression exists (Gallegos and Rose, 2015). As a result, it is possible that the insertion promoted *CaMYBA* expression. The mechanism by which *LINE-1* insertion in the first intron affects *CaMYBA* expression needs to be elucidated.

Literature Cited

- Aguilar-Barragán, A. and N. Ochoa-Alejo. 2014. Virus-induced silencing of *MYB* and *WD40* transcription factor genes affects the accumulation of anthocyanins in chilli pepper. *Biol. Plant.* 58: 567–574.
- Albert, N. W., D. H. Lewis, H. Zhang, K. E. Schwinn, P. E. Jameson and K. M. Davies. 2011. Members of an R2R3-MYB transcription factor family in *Petunia* are developmentally and environmentally regulated to control complex floral and vegetative pigmentation patterning. *Plant J.* 65: 771–784.
- Aza-González, C., L. Herrera-Isidró, H. G. Núñez-Palenius, O. M. De La Vega and N. Ochoa-Alejo. 2013. Anthocyanin accumulation and expression analysis of biosynthesis-related genes during chili pepper fruit development. *Biol. Plant.* 57: 49–55.
- Borovsky, Y., M. Oren-Shamir, R. Ovadia, W. D. Jong and I. Paran. 2004. The *A* locus that controls anthocyanin accumulation in pepper encodes a *MYB* transcription factor homologous to *Anthocyanin2* of *Petunia*. *Theor. Appl. Genet.* 109: 23–29.
- Butelli, E., C. Licciardello, Y. Zhang, J. Liu, S. Mackay, P. Bailey, G. Reforgiato-Recupero and C. Martin. 2012. Retrotransposons control fruit-specific, cold-dependent accumulation of anthocyanins in blood oranges. *Plant Cell* 24: 1242–1255.
- Chaim, A. B., Y. Borovsky, W. De Jong and I. Paran. 2003. Linkage of the *A* locus for the presence of anthocyanin and *fs10.1*, a major fruit-shape QTL in pepper. *Theor. Appl. Genet.* 106: 889–894.
- Gallegos, J. E. and A. B. Rose. 2015. The enduring mystery of intron-mediated enhancement. *Plant Sci.* 237: 8–15.
- Hulse-Kemp, A. M., S. Maheshwari, K. Stoffel, T. A. Hill, D. Jaffe, S. R. Williams, N. Weisenfeld, S. Ramakrishnan, V. Kumar, P. Shah, M. C. Schatz, D. M. Church and A. Van Deynze. 2018. Reference quality assembly of the 3.5-Gb genome of *Capsicum annuum* from a single linked-read library. *Hortic. Res.* 5: 4.
- Kim, S., M. Park, S. I. Yeom, Y. M. Kim, J. M. Lee, H. A. Lee, E. Seo, J. Choi, K. Cheong, K. T. Kim, K. Jung, G. W. Lee, S. K. Oh, C. Bae, S. B. Kim, H. Y. Lee, S. Y. Kim, M. S. Kim, B. C. Kang, Y. D. Jo, H. B. Yang, H. J. Jeong, W. H. Kang, J. K. Kwon, C. Shin, J. Y. Lim, J. H. Park, J. H. Huh, J. S. Kim, B. D. Kim, O. Cohen, I. Paran, M. C. Suh, S. B. Lee, Y. K. Kim, Y. Shin, S. J. Noh, J. Park, Y. S. Seo, S. Y. Kwon, H. A. Kim, J. M. Park, H. J. Kim, S. B. Choi, P. W. Bosland, G. Reeves, S. H. Jo, B. W. Lee, H. T. Cho, H. S. Choi, M. S. Lee, Y. Yu, Y. D. Choi, B. S. Park, A. Deynze, H. Ashrafi, T. Hill, W. T. Kim, H. S. Pai, H. K. Ahn, I. Yeom, J. J. Giovannoni, J. K. C. Rose, I. Sørensen, S. J. Lee, R. W. Kim, I. Y. Choi, B. S. Choi, J. S. Lim, Y. H. Lee and D. Choi. 2014. Genome sequence of the hot pepper provides insights into the evolution of pungency in *Capsicum* species. *Nat. Genet.* 46: 270–278.
- Kobayashi, S., N. Goto-Yamamoto and H. Hirochika. 2004. Retrotransposon-induced mutations in grape skin color. *Science* 304: 982.
- Koes, R., W. Verweij and F. Quattrocchio. 2005. Flavonoids: a colorful model for the regulation and evolution of biochemical pathways. *Trends Plant Sci.* 10: 236–242.
- Li, J. G., H. L. Li and S. Q. Peng. 2011. Three R2R3 MYB transcription factor genes from *Capsicum annuum* showing differential expression during fruit ripening. *Afr. J. Biotechnol.* 10: 8267–8274.
- Lightbourn, G. J., R. J. Griesbach, J. A. Novotny, B. A. Clevidence, D. D. Rao and J. R. Stommel. 2008. Effects of anthocyanin and carotenoid combinations on foliage and immature fruit color of *Capsicum annuum* L. *J. Heredity* 99: 105–111.
- Lightbourn, G. J., J. R. Stommel and R. J. Griesbach. 2007. Epistatic interactions influencing anthocyanin gene expression in *Capsicum annuum*. *Am. Soc. Hort. Sci.* 132: 824–829.
- Liu, Y., K. Lin-Wang, R. V. Espley, L. Wang, H. Yang, B. Yu, A. Dare, E. Varkonyi-Gasic, J. Wang, J. Zhang, D. Wang and A. C. Allan. 2016. Functional diversification of the potato R2R3 MYB anthocyanin activators AN1, MYBA1, and MYB113 and their interaction with basic helix-loop-helix cofactors. *J. Exp. Bot.* 67: 2159–2176.
- Lu, B. Y., G. X. Cheng, Z. Zhang, J. T. Sun, M. Ali, Q. L. Jia, L. De-Xu, Z. H. Gong and D. W. Li. 2019. CaMYC, a novel transcription factor, regulates anthocyanin biosynthesis in color-leaved pepper (*Capsicum annuum* L.). *J. Plant Growth Reg.* 38: 574–585.
- Ohno, S., M. Hosokawa, M. Kojima, Y. Kitamura, A. Hoshino, F. Tatsuzawa, M. Doi and S. Yazawa. 2011. Simultaneous post-transcriptional gene silencing of two different chalcone synthase genes resulting in pure white flowers in the octoploid dahlia. *Planta* 234: 945–958.
- Qin, C., C. Yu, Y. Shen, X. Fang, L. Chen, J. Min, J. Cheng, S. Zhao, M. Xu, Y. Luo, Y. Yang, Z. Wu, L. Mao, H. Wu, C. L. Hu, H. Zhou, H. Lin, S. González-Morales, D. L. Trejo-Saavedra, H. Tian, X. Tang, M. Zhao, Z. Huang, A. Zhou, X. Yao, J. Cui, W. Li, Z. Chen, Y. Feng, Y. Niu, S. Bi, X. Yang, W. Li, H. Cai, X. Luo, S. Montes-Hernández, M. A. Leyva-González, Z. Xiong, X. He, L. Bai, S. Tan, X. Tang, D. Liu, J. Liu, S. Zhang, M. Chen, L. Zhang, L. Zhang, Y. Zhang, W. Liao, Y. Zhang, M. Wang, X. Lv, B. Wen, H. Liu, H. Luan, Y. Zhang, S. Yang, X. Wang, J. Xu, X. Li, S. Li, J. Wang, A. Palloix, P. W. Bosland, Y. Li, A. Krogh, R. F. Rivera-Bustamante, L. Herrera-Estrella, Y. Yin, J. Yu, K. Hu and Z. Zhang. 2014. Whole-genome sequencing of cultivated and wild peppers provides insights into *Capsicum* domestication and specialization. *Proc. Nat. Amer. Soc. USA* 111: 5135–5140.
- Tanaka, Y., N. Sasaki and A. Ohmiya. 2008. Biosynthesis of plant pigments: anthocyanins, betalains and carotenoids. *Plant J.* 54: 733–749.
- Wang, D. and P. W. Bosland. 2006. The genes of *Capsicum*. *HortScience* 41: 1169–1187.
- Wicker, T., F. Sabot, A. Hua-Van, J. L. Bennetzen, P. Capy, B. Chalhoub, A. Flavell, P. Leroy, M. Morgante, O. Panaud, E. Paux, P. SanMiguel and A. H. Schulman. 2007. A unified classification system for eukaryotic transposable elements. *Nat. Rev. Genet.* 8: 973–982.
- Zhang, L., J. Hu, X. Han, J. Li, Y. Gao, C. M. Richards, C. Zhang, Y. Tian, G. Liu, H. Gul, D. Wang, Y. Tian, C. Yang, M. Meng, G. Yuan, G. Kang, Y. Wu, K. Wang, H. Zhang, D. Wang and P. Cong. 2019. A high-quality apple genome assembly reveals the association of a retrotransposon and red fruit colour. *Nat. Commun.* 10: 1494.
- Zhang, Z., D. W. Li, J. H. Jin, Y. X. Yin, H. X. Zhang, W. G. Chai and Z. H. Gong. 2015. VIGS approach reveals the modulation of anthocyanin biosynthetic genes by *CaMYB* in chili pepper leaves. *Front. Plant Sci.* 6: 500.