J Korean Soc Appl Biol Chem (2015) 58(4):555–562
DOI 10.1007/s13765-015-0077-z

Online ISSN 2234-344X
Print ISSN 1738-2203

REVIEW

Genetic and molecular regulation of flower pigmentation in soybean

Jagadeesh Sundaramoorthy · Gyu Tae Park · Jeong-Dong Lee · Jeong Hoe Kim · Hak Soo Seo · Jong Tae Song

Received: 11 April 2015 / Accepted: 21 April 2015 / Published online: 12 May 2015

Abstract Flower color is one of the key traits, which has been widely considered for genetic studies on soybean. A variety of flower colors, such as dark purple, purple, purple blue, purple throat, magenta, pink, near white, and white, has been identified in cultivated soybean (*Glycine max*). Out of the 19,649 soybean accessions deposited in the United States Department of Agriculture-Germplasm Resources Information Network database, 67 % have purple flowers, 32 % have white flowers, and merely 1 % have flowers with different colors. In contrast, almost all accessions of wild soybean (*Glycine soja*) have only purple flowers. Flavonoids, mainly anthocyanins, are the most common pigments contributing to flower coloration in soybean. In the recent decades, the flavonoid biosynthesis pathway for anthocyanins has been well established, and some of the genes controlling flower color in soybean have been identified and characterized. Flower pigmentation of soybean is mainly controlled by six independent loci (*W1*, *W2*, *W3*, *W4*, *Wm*, and *Wp*) along with the combination of various other factors such as anthocyanin structure, vacuolar pH, and co-pigments. In this review, we summarize the current status of genetic and molecular regulation of flower pigmentation in cultivated and wild varieties of soybean.

Keywords Anthocyanin · Flower color · *Glycine max* · *Glycine soja* · Soybean

J. Sundaramoorthy · G. T. Park · J.-D. Lee · J. T. Song (✉)
School of Applied Biosciences, Kyungpook National University, Daegu 702-701, Republic of Korea
e-mail: jtsong68@knu.ac.kr

J. H. Kim
Department of Biology, Kyungpook National University, Daegu 702-701, Republic of Korea

H. S. Seo
Department of Plant Bioscience, Seoul National University, Seoul 151-742, Republic of Korea

H. S. Seo
Bio-MAX Institute, Seoul National University, Seoul 151-818, Republic of Korea

Introduction

Flower color is one of the key traits considered for genetic studies since the rediscovery of Mendel's laws. Flower coloration is induced by the deposition of flavonoids (including anthocyanins), carotenoids, and betalains (Mol et al. 1998; Grotewold 2006; Tanaka and Brugliera 2013). Although flower pigmentation has been extensively studied in maize (*Zea mays*), petunia (*Petunia hybrida*), and *Arabidopsis* (Lepiniec et al. 2006), the knowledge is relatively limited for soybean [*Glycine max* (L.) Merr.], because flower color is used only as marker in breeding programs (Koes et al. 2005). However, flower color and patterning have recently become a topic of interest due to their wide use in applied research studies (Mol et al. 1998).

A variety of flower colors has been identified in soybean. Among them, purple and white are the most common colors (Fig. 1). Early genetic studies showed that purple and white-flower colors are controlled by a single pair of genes in which purple is dominant over white (Takahashi and Fukuyama 1919; Woodworth 1923; Nagai 1926). Out of the 19,649 soybean accessions deposited in the United States Department of Agriculture-Germplasm Resources Information Network (USDA-GRIN) germplasm collection, 13,133 accessions (67 %) have purple flowers, 6344 accessions (32 %) have white flowers, and the remaining

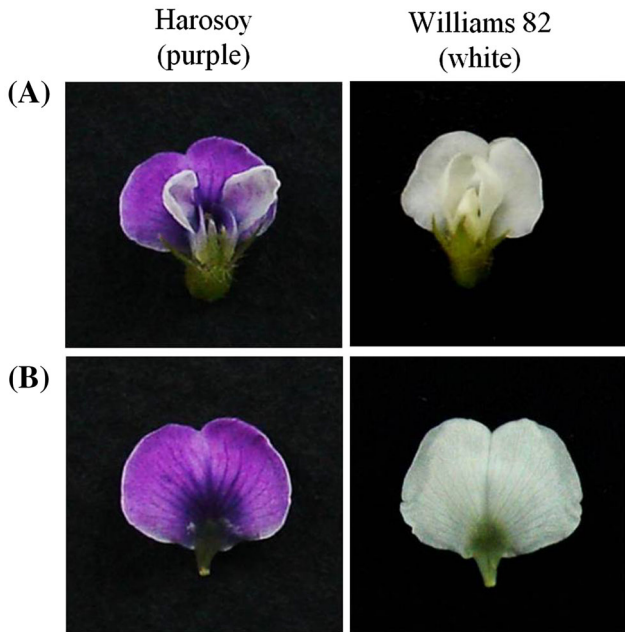

Fig. 1 Representation of the soybean genotype *W1W1w3w3W4W4* (Harosoy) with purple flowers and the soybean genotype *w1w1w3w3W4W4* (Williams 82) with white flowers. (**A**) Whole flower. (**B**) Standard petal

accessions (1 %) have different flower colors such as dark purple, purple blue, purple throat, magenta, pink, and near white (Table 1). In contrast, almost all the accessions of wild soybean [*Glycine soja* Seib. and Zucc.] have purple flowers (Chen and Nelson 2004). Flower pigmentation in soybean is mainly controlled by six independent loci, *W1*, *W2*, *W3*, *W4*, *Wm*, and *Wp* (reviewed by Palmer et al. 2004). In addition, factors, such as anthocyanin structure, pH of the vacuole (where the anthocyanins are localized) and co-pigmentation have been found to influence flower coloration (Mol et al. 1998; Yoshida et al. 2009).

Several research studies have been carried out in soybean to identify the molecular mechanism of flower pigmentation. However, flower pigmentation and regulation mechanisms remain unclear. This review summarizes the recent advancements in molecular and regulatory mechanisms that are responsible for the flower color development in soybean.

Table 1 Distribution of values for flower color of soybean in USDA-GRIN (2015)

Definition	Number of accessions	Ratio (%)
Purple	13,133	66.838
White	6344	32.287
Dark purple	122	0.621
Light purple	15	0.076
Blue	12	0.061
Purple throat	10	0.051
Near white	6	0.031
Magenta	5	0.025
Mutable purple	2	0.010
Total	19,649	100

Anthocyanin and flavonol biosynthesis pathway and flower pigmentation

Anthocyanins are the most prominent pigments in soybean, which are responsible for purple, blue, and pink color in flowers. In addition, they protect plants from different stresses such as UV irradiation and reactive oxygen species (Nagata et al. 2003; Gould 2004). Three derivatives of anthocyanins, namely delphinidin, pelargonidin, and cyanidin, are the most commonly found in plants (Schwinn and Davies 2004). Delphinidin falls into the blue-purple range, pelargonidin into the pink range, and cyanidin into the red range (Harborne 1967). In soybean, most of the accessions have purple flowers, which show that delphinidin branch of anthocyanin biosynthesis is predominantly active in the determination of flower color (Iwashina et al. 2007).

Anthocyanins are water-soluble pigments and derived from a branch of the flavonoid pathway. In anthocyanin biosynthesis, the enzyme chalcone synthase (CHS) catalyzes the initial step by condensing one molecule of 4-coumaroyl-CoA and three molecules of malonyl-CoA for the synthesis of naringenin chalcones (Fig. 2). Naringenin chalcones act as the precursors for the synthesis of all the classes of flavonoids, isoflavonoids, and anthocyanins. The stereospecific isomerization of chalcone isomerase (CHI) converts naringenin chalcones to colorless naringenin, and then flavanone 3-hydroxylase (F3H) converts naringenin to dihydrokaempferol. Subsequently, the hydroxylation of dihydrokaempferol to dihydroquercetin and dihydromyricetin is catalyzed by flavonoid 3′-hydroxylase (F3′H) and flavonoid 3′5′-hydroxylase (F3′5′H), respectively. Dihydroflavonol 4-reductase (DFR) catalyzes the reduction of dihydroflavonols (dihydrokaempferol, dihydroquercetin, and dihydromyricetin) to colorless leucoanthocyanidins (Holton and Cornish 1995).

Furthermore, anthocyanidin synthase (ANS) oxidizes leucoanthocyanidins to anthocyanidins with the help of ferrous ion (Sparvoli et al. 1994; Gollop et al. 2001). In nature, anthocyanidins are unstable, but glycosylation increases their stability and hydrophilicity (He et al. 2010). Uridine diphosphate-flavonoid glucosyltransferase (UFGT) catalyzes the *O*-glycosylation of anthocyanidins and synthesizes more stable anthocyanins (delphinidin 3,5-di-*O*-glucoside, pelargonidin 3,5-di-*O*-glucoside, and cyanidin 3,5-di-*O*-glucoside) (Hostel 1981; Springob et al. 2003). Subsequent methylation of delphinidin 3,5-di-*O*-glucoside

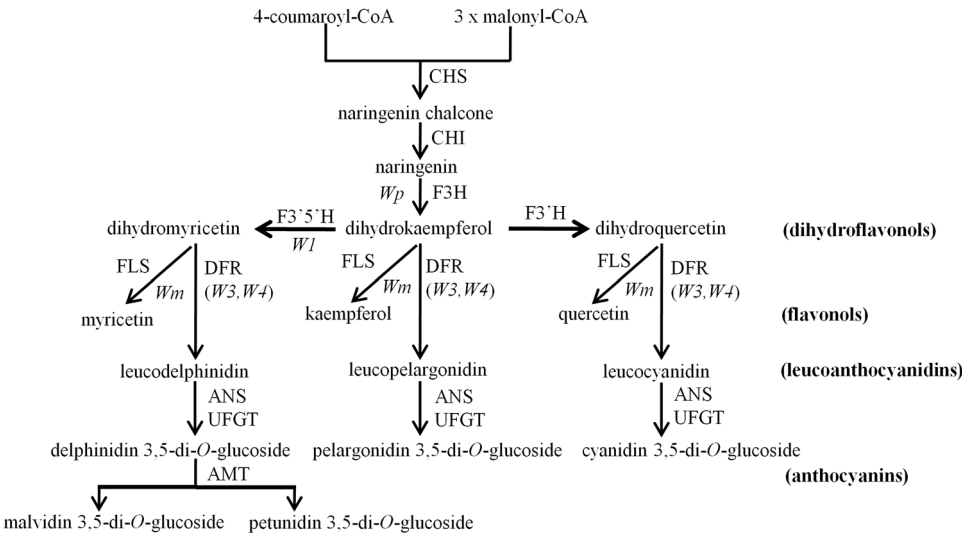

Fig. 2 Schematic representation of anthocyanin and flavonol biosynthesis pathway involved in flower development. *AMT* anthocyanin methyltransferase; *ANS* anthocyanidin synthase; *CHI* chalcone isomerase; *CHS* chalcone synthase; *DFR* dihydroflavonol-4-reductase; *F3H* flavanone 3-hydroxylase; *F3′H* flavonoid 3′-hydroxylase; *F3′5′H* flavonoid 3′5′-hydroxylase; *FLS* flavonol synthase; *UFGT* UDP-flavonoid glucosyltransferase

to petunidin 3,5-di-*O*-glucoside and malvidin 3,5-di-*O*-glucoside is catalyzed by anthocyanin methyltransferase (AMT) (Tanaka et al. 2005).

Similar to anthocyanins, flavonols and their derivatives are also synthesized from dihydroflavonols by flavonol synthase (FLS) followed by glycosylation, methylation, and acylation (To and Wang 2006). In purple flowers, a high amount of malvidin 3,5-di-*O*-glucoside and small traces of delphinidin 3,5-di-*O*-glucoside, petunidin 3,5-di-*O*-glucoside, and delphinidin 3-*O*-glucoside have been detected among anthocyanins. In addition, eight flavonol derivatives were also found, and kaempferol 3-*O*-gentiobioside was detected as the primary flavonol (Iwashina et al. 2007, 2008). In soybean, F3′H and F3′5′H are the two essential enzymes, which are mainly involved in the pigmentation of seed coat and flowers, respectively (Zabala and Vodkin 2003; Han et al. 2010; Moreau et al. 2012).

Among the six flower-color-controlling loci (*W1*, *W2*, *W3*, *W4*, *Wm*, and *Wp*) of soybean, all loci encode enzymes involved in flavonoid biosynthesis, except for *W2*. The loci *W1, W3, W4*, *Wm*, and *Wp* encode F3′5′H, DFR1, DFR2, FLS, and F3H1, respectively, whereas *W2* encodes an MYB transcription factor, which is involved in the vacuolar acidification of flower petals (Takahashi et al. 2008, 2011).

The role of structural genes of anthocyanin biosynthesis pathway in flower pigmentation

F3′5′H gene at the W1 locus

In anthocyanin biosynthesis, F3′5′H (a cytochrome P450 enzyme) hydroxylates dihydrokaempferol to produce dihydromyricetin from which a delphinidin class of anthocyanins (purple color) is synthesized. Buzzell et al. (1987) found that *W1* is responsible for the synthesis of delphinidin-3-glucosides and the production of purple flowers in soybean. *W1* was mapped between Satt348 and Satt160 on MLG F (chromosome 13) in soybean linkage map (Song et al. 2004).

Zabala and Vodkin (2007a) identified the *F3′5′H* gene as a candidate gene for *W1* and provided molecular evidence for the association with *W1* locus. In Williams 82 (*w1* allele), a small 65-bp insertion with tandem repeats and a 12-bp deletion in the third exon cause premature termination of translation and make recessive *w1* allele non-functional and produced white flowers. Further analysis of *F3′5′H* showed very low expression levels in both purple- and white-flower lines, implying that the low amount of the F3′5′H enzyme is sufficient to synthesize delphinidins and determines the flower color in soybean (Zabala and Vodkin 2007a). However, anthocyanins were not detected in white flowers of the *w1* mutant which strongly suggests that the *W1* locus encodes the F3′5′H enzyme (Iwashina et al. 2007). In addition, Yang et al. (2010) developed an indel-based marker, SL019, from the third exon of the *F3′5′H* gene (where a 65-bp indel was present), which perfectly detects the polymorphism between the *W1* and *w1* alleles of the *W1* locus from all the purple- and white-flower lines, respectively.

Recently, Park et al. (2014) analyzed the *W1* locus from 99 landraces with white flowers and showed that all the white-flower landraces have the *w1* allele, which is identical in sequence with the *w1* allele of Williams 82 (studied by Zabala and Vodkin 2007a). In addition, a phylogenetic analysis showed the possible origin of the *w1* recessive allele from a group of purple-flower *G. max* accessions, which were found to cluster together with white-flower *G. max* accessions in the constructed phylogenetic tree.

In contrast, almost all the accessions of wild soybean (*G. soja*) have purple flowers. The absence of flower color variations in wild soybean remains unclear. However, a few *G. soja* accessions have been identified with color variations. In 1998, a white-flower line, PI 424008C, was found among the progenies of the *G. soja* accession PI 424008A that has purple flowers. Genetic analysis showed that the white color in PI 424008C was caused by a mutation in the *W1* locus similar to that in the white-flower *G. max* accessions (Chen and Nelson 2004). Takahashi et al. (2010) identified an accession of *G. soja* (B09121) with light purple flowers and assigned it as the *w1-lp* mutant. They also showed that the *w1-lp* allele is dominant over the *w1* allele (*W1* > *w1-lp* > *w1*). In the *w1-lp* mutants, cDNA sequence analysis of the *F3′5′H* gene indicated a unique single-base substitution in the nucleotide position 653, which results in a notable amino acid change from valine to methionine (position 210). However, not much difference was detected in the transcription level between the *w1-lp* and *W1* alleles. Flavonoid analysis of the flower petals in the *w1-lp* line showed a very little amount of all the four major anthocyanins that are common in purple flowers (malvidin 3,5-di-*O*-glucoside, petunidin 3,5-di-*O*-glucoside, delphinidin 3,5-di-*O*-glucoside, and delphinidin 3-*O*-glucoside) and also small traces of 5′-unsubstituted derivatives of these anthocyanins, which suggests that a mutation in the *F3′5′H* gene may lead to a reduction of the F3′5′H enzymatic activity and an increase of the F3′H enzymatic activity in the *w1-lp* mutant line (Takahashi et al. 2010).

Recently, Takahashi et al. (2012) found a single plant with purple and white variegated flowers in a *G. soja* accession (B00146). This single plant with variegated flowers (B00146-*m*) developed progenies with white flowers (B00146-*w*) as well as with purple flowers (B00146-*r*). The *w1-m* allele of variegated flowers is allelic to the *W1* locus, and sequence analysis of the *F3′5′H* gene in B00146-*m* showed the insertion of the *Tgs1* transposon (CACTA family) in the first exon. Excision of this active transposon led to the development of revertant (purple) lines and mutant (white) lines. In B00146-*w,* CA insertion in the *F3′5′H* gene as a footprint of *Tgs1* at the transposon insertion site led to a truncated polypeptide with 59 amino acids, which might cause the complete loss of its function. As in *G. max*, *G. soja* also displayed white flower when *W1* locus was dysfunctional, which revealed that *W1* is the most important locus for flower color production in soybean. In addition, *W1* locus has a pleiotropic effect on flower and hypocotyl colors. Cultivars with the dominant *W1* allele produce purple flowers and purple hypocotyls, whereas cultivars with the recessive *w1* allele produce white flowers and green hypocotyls (Takahashi and Fukuyama 1919).

DFR genes at the W3 and W4 loci

The DFR enzyme is known to catalyze the production of leucoanthocyanidins from dihydroflavonols. DFR is an essential enzyme, to step forward, the synthesis of anthocyanins rather than flavonols in the biosynthetic pathway. In soybean, the genes encoding DFR enzymes co-segregate with the two loci, *W3* and *W4,* and they function (or interact) epistatically to each other in *W1* background (Fasoula et al. 1995; Xu and Palmer 2005). Soybean accessions show different flower colors depending on the allelic variations of *W3* and *W4*. For example, the genotype *W1W3W4* produces dark purple flowers, *W1w3w4* genotype produces near white flowers, *W1w3W4* genotype produces purple flowers, and *W1W3w4* produces purple throat flowers (Hartwig and Hinson 1962; Yan et al. 2014).

To identify the gene that is related to the *W3* locus, Fasoula et al. (1995) carried out a restriction fragment length polymorphism analysis with 23 restriction enzymes using a DFR probe, which contained a complete partial sequence of *DFR1* (Wang et al. 1994). Out of them, the enzyme *Hae*III showed polymorphism in the *DFR1* gene between the purple throat (*W1W3w4*) and the near white (*W1w3w4*) flowers. Furthermore, recombination analysis of a population developed by a cross between a purple-throat-flower line (*W3*) and a near-white-flower line (*w3*) showed a complete co-segregation and the absence of recombination between the *W3* and *w3* alleles. Fasoula et al. (1995) suggested that the phenotype of purple throat flowers in soybean is possibly due to the undetermined excision of some transposable elements, which cause changes in the expression pattern of the *DFR1* gene.

A recent sequence homology search and microsatellite marker analysis also showed that *DFR1* is between Sat_287 and Satt467 on MLG B2 (chromosome 14) indicating that *DFR1*, among the homologous *DFR* genes, is most likely associated with the *W3* locus (Yang et al. 2010). However, the *W3* locus has not been fully characterized yet.

The *W4* locus is also known to encode DFR in the anthocyanin biosynthesis pathway and was mapped between Satt386 and Sct_137 on MLG D2 (chromosome 17). The *DFR2* gene was identified as a candidate gene for the *W4* locus (Xu and Palmer 2005). Mutations in the *W4* locus with *W1w3* background results in different types of pigment accumulation and color patterning in flower petals. Five mutant alleles, *w4*, *w4-m*, *w4-dp*, *w4-p*, and *w4-lp*, have been identified in the *W4* locus that produce near white, variegated, dilute purple, pale purple, and light purple flower color, respectively (Palmer et al. 1989; Groose and Palmer 1991; Xu and Palmer 2005; Xu et al. 2010; Yan et al. 2014).

The complete loss of function of the *W4* locus (*w4* recessive allele) in different lines (Clark-*w4*, 222-A-3, and kw4) develops only near white flowers, irrespective of the

mutation type (Yan et al. 2014). Single-base substitution at the 5′ splice site of the fourth intron of *DFR2* in Clark-*w4* led to the retention of the fourth intron in the transcript. A single-base deletion in 222-A-3 resulted in a truncated polypeptide of only 24 amino acids that lacked the NADPH binding domain, which is important for the DFR2 function. In the wild soybean mutant kw4, no expression of *DFR2* in flower petals resulted in substantial reduction of anthocyanins, possibly due to the deletion of a 367-bp fragment in the third intron of *DFR2* or a nucleotide polymorphism in the promoter region (Yan et al. 2014). Further experiments are required to clarify this mechanism.

Analysis of the *DFR2* gene from another mutant *w4-m* showed the insertion of an active transposon (*Tgm9*) in the second intron (Xu et al. 2010). Excision of this active transposon in some somatic cells during flower development resulted in the production of variegated flowers, and this kind of transposon excision also caused stable revertants. Two stable revertants, *w4-dp* (*W1w3w4-dp*) and *w4-p* (*W1w3w4-p*), were developed from the *w4-m* line, and they produced dilute purple and pale purple flowers, respectively (Palmer and Groose 1993; Xu and Palmer 2005). In the *w4-dp* and *w4-p* mutants, excision of the *Tgm9* transposon leaves 4- and 0-bp footprints, respectively. Moreover, the excised fragment of *Tgm9* incorporates into the promoter region at 1043 and 1034 bp of the upstream transcription start site in the *w4-dp* and *w4-p* alleles, respectively (Xu et al. 2010). Expression analysis showed a very low expression level of *DFR2* in the *w4* and *w4-dp* mutants, but a higher expression level in the *w4-p* mutant than in the *W4* dominant line. Flower petals also showed the presence of unique dihydroflavonols in the *w4-p* mutant (Yan et al. 2014).

Recently, analysis of an EMS-induced mutant (*W1w3w4-lp*) with light purple flowers generated from a purple-flower soybean line showed that the *w4-lp* was responsible for the light purple flowers (Yan et al. 2014). In the *w4-lp* mutant, a single-base substitution in *DFR2* resulted in an amino acid change from arginine to histidine (position 39) and led to low expression levels of *DFR2* (Yan et al. 2014). Further experiments may be required to clarify the reasons of low transcript abundance and study the functional activity of DFR2 in the *w4-lp* mutant.

It is clear that mutations of *DFR2* with *w3* background leads to unique flavonoid compositions and different shades of purple, such as dilute purple, pale purple, and light purple, depending on the expression level of *DFR2* or the activity of the enzyme. However, a comprehensive study of the promoter is necessary to identify the cis-regulated expression of *DFR2*.

F3H gene at the Wp locus

The F3H enzymes catalyze an early and very important step in anthocyanin metabolism, in which flavonones are converted to dihydroflavonol (dihydrokaempferol). Genetic analysis showed that the *Wp* locus corresponds to F3H, and the homozygous recessive *wp* allele in *W1* background develops pink flowers in soybean (Stephens and Nickell 1992). The *Wp* locus was mapped between SL007 and Satt216 on MLG D1b (chromosome 2) (Yang et al. 2010), and the *F3H1* gene was identified as a candidate gene for the *Wp* locus (Zabala and Vodkin 2005). A mutant (*wp* allele) with pink flowers was derived from a mutable line with a high rate of instability in flower color (variegated flowers) (Johnson et al. 1998; Zabala and Vodkin 2005).

Analysis of the *F3H1* gene from the *wp* line showed the insertion of an active transposon (*Tgm-Express1*) in the second intron, and expression analysis showed very low expression levels in the *wp* mutants compared to the *Wp* dominant lines. This active transposon also created a variety of chimerical transcripts with varying open-reading frames in the *wp* mutants. About twelve distinct chimeric transcripts and one non-chimeric transcript were identified in the *wp* mutants by Zabala and Vodkin (2007b). The non-chimeric transcript of the *wp* mutant was identical to the normal transcript of the *Wp* allele, even though the *F3H1* transcripts with typical sizes were found in very low amounts in the *wp* mutant, suggesting that they are still sufficient for synthesizing anthocyanin pigments to develop pink flowers (Zabala and Vodkin 2007b).

Other anthocyanin biosynthesis pathway enzymes

The enzyme CHS is the first key enzyme in anthocyanin biosynthesis pathway, which catalyzes the production of chalcone that acts as the precursor of flavonoid biosynthesis. In petunia and tobacco, the *CHS* genes have been isolated and characterized for flower color and pigmentation (Koes et al. 1986; Wang et al. 2006). In soybean, the *CHS* repeats and *I* locus were mapped between A454.p2 and GMNOD2B on MLG A2 (chromosome 8) (Yang et al. 2010). The *CHS* gene family has been studied extensively for seed coat color (Tuteja et al. 2004, 2009).

The F3′H enzyme is involved in the conversion of dihydrokaempferol to dihydroquercetin. Genetic analysis showed that the *F3′H* gene was identified as the *T* locus (Toda et al. 2002; Zabala and Vodkin 2003). The *T* locus was mapped between Satt286 and Satt365 on MLG C2 (chromosome 6) (Yang et al. 2010) and controls seed coat/hilum color in soybean (Toda et al. 2002).

The enzyme ANS is essential for the oxidation of leucoanthocyanidins to anthocyanins, and there is a high sequence similarity between the genes encoding these enzymes. Among the *ANS* genes (*ANS1*, *ANS2*, and *ANS3*) in soybean, *ANS2* and *ANS3* have identical sequences, and they are likely from the same gene. *ANS1* and *ANS2* were mapped between BE806308 and Sat_272 on MLG B1

(chromosome 11) and between Sat_414 and Satt129 on MLG D1a (chromosome 1), respectively (Yang et al. 2010).

UFGT catalyzes the glycosylation of anthocyanin to colored anthocyanin 3-*O*-glucosides by transferring glucose moiety from uridine diphosphate-glucose to C-3 hydroxyl group of anthocyanin (Buzzell et al. 1987; Todd and Vodkin 1993). The putative genes *UGT78K1* and *UGT78K2* were identified by homology search using BLASTn and characterized as the *UFGT* genes that control seed coat pigment (Kovinich et al. 2010, 2011). However, *UFGT* has not been precisely mapped in soybean linkage map.

In addition, the transcript level of the *UFGT* and *ANS* genes were regulated by the *R* locus, which encodes the R2R3 MYB transcription factor. The *R* locus was mapped between BARCSOYSSR_09_1489 and BARCSOYSSR_09_1506 on MLG K (chromosome 9) (Gillman et al. 2011). When the *R2R3 MYB* gene is overexpressed, the *UFGT* (*UGT78K1* and *UGT78K2*) and *ANS* (*ANS2*/*ANS3*) genes are also up-regulated and produce black seed coat/hilum (Gillman et al. 2011; Kovinich et al. 2011). In contrast, loss of function of the *R2R3 MYB* gene causes a down-regulation in both the *UFGT* and *ANS* genes and produces brown seed coat/hilum. It remains to be answered whether the MYB transcriptional factor directly promotes the transcriptional activation of the *UFGT* and *ANS* genes.

In soybean, all four enzymes discussed in this review have been studied extensively for their involvement in seed coat/hilum color, but not in flower pigmentation. However, genetic regulation of these enzymes with flower color are well established in other plant species, such as in petunia and *Arabidopsis* (To and Wang 2006); hence it is likely that these enzymes play an important role in soybean flower pigmentation too. These enzymes and their corresponding and regulating loci need to be addressed for flower color in soybean.

Additional factors affecting flower pigmentation

Co-pigmentation by the FLS gene at the Wm locus

Flavonols (myricetin, kaempferol, and quercetin) are yellow-color components, and they act as either pigments or co-pigments to anthocyanins (Harborne 1967); however, in soybean, flavonols act only as co-pigments (Takahashi et al. 2007). In flavonoid biosynthesis, both anthocyanin and flavonols are derived from dihydroflavonols by the establishment of a double bond between C-2 and C-3 positions by the action of FLS (Forkmann 1991). Analysis of a mutant with magenta flowers from Harosoy (Harosoy-*wm*) showed that the *Wm* locus controls magenta color (Buzzell et al. 1977). The recessive *wm* allele at the *Wm* locus was found to be associated with low levels of flavonol glucoside synthesis in flowers and leaves, which suggested that *Wm* is responsible for the production of flavonol and probably encodes FLS. The *Wm* locus was mapped between Satt252 and Satt425 on MLG F (chromosome 13) (Takahashi et al. 2007).

Takahashi et al. (2007) identified a candidate gene, *gmfls1*, in the *Wm* locus, which encodes a 334-amino-acid-long polypeptide, GmFLS1, consisting of conserved dioxygenase domains (A and B), and this gene showed homology with previously reported *FLS* genes from other plant species. Bacterial heterologous expression assay showed that GmFLS1 of Harosoy (*Wm* allele) has the activity of FLS, whereas a single-base deletion in the *wm* allele resulted in a truncated polypeptide and devoid of FLS activity. Thus, co-pigmentation between anthocyanin and flavonol glucosides may contribute to purple flowers in soybeans along with the *Wm* alleles, whereas the recessive *wm* allele substantially reduces the content of flavonol glucosides, and it may inhibit co-pigmentation, resulting in magenta flowers (Takahashi et al. 2007).

MYB transcription factor gene at the W2 locus and regulation of vacuolar pH

MYB transcription factors represent a group of proteins that consist of a conserved MYB DNA-binding domain. In plants, an MYB-protein subfamily is illustrated as the R2R3-type MYB domain and the *R2R3-type MYB* genes control many paths of plant secondary metabolism (Stracke et al. 2001).

The purple-blue-flower landrace, Nezumisaya, was identified, and a complementation test showed that purple blue color was controlled by the *W2* locus, which was mapped between Satt318 and Satt020 on MLG B2 (chromosome 14) (Takahashi et al. 2008). Flavonoid analysis of flower petals showed that the alleles of *W1*, *W3*, *W4*, *Wm*, and *Wp* loci affect the structure and/or amount of flavonoids (Iwashina et al. 2007, 2008; Takahashi et al. 2010). In contrast, flavonoids in the purple blue flowers were similar to that of purple flowers, which suggest that quantitative or structural differences of anthocyanins or co-pigmentation were not responsible for the purple blue flowers (Iwashina et al. 2008).

To identify the factor affecting the color change in the purple blue flowers, Takahashi et al. (2008) investigated the physiological basis of flower color. An increase in pH has a bluing effect in flower petals, while a decrease in pH causes a reddening effect (To and Wang 2006). In accordance, the sap extract of standard petal from purple flowers showed pH values from 5.73 to 5.77, whereas purple blue flowers had pH values from 6.07 to 6.10, which suggests that the recessive allele *w2* may be responsible for the acidification of flower petals and produces purple blue flowers (Takahashi et al. 2008). Similarly, in petunia, when

mutation occurs in the *PH* genes (*PH1 to PH7*), plants produce blue flowers and show increased pH in petal sap extract (deVlaming et al. 1983; Chuck et al. 1993).

To ascertain how the transcription factor controls the vacuolar pH of flower petals, Takahashi et al. (2011) identified a candidate gene, *GmMYB-G20-1*, which encodes an MYB transcription factor. This gene has a 53.7 % amino-acid sequence similarity with the *PH4* gene of petunia, which controls the vacuolar pH and blue color. *GmMYB-G20-1* of purple flower (*W2* allele) encodes a 361-amino-acid-long polypeptide, whereas *GmMYB-G20-1* of purple blue flower mutants (*w2* allele) has a nonsense mutation in the MYB domain, resulting in a truncated polypeptide (Takahashi et al. 2011). Furthermore, virus-induced gene silencing in Harosoy (purple flowers) revealed that the silencing of *GmMYB-G20-1* changes the flower color from purple to gray or blue. These results showed that *GmMYB-G20-1* corresponds to the *W2* locus and controls the proteins involved in pH regulation of petals as a transcription factor (Takahashi et al. 2013).

Conclusion

Flower color and pigmentation have been recognized as a potential tool for elucidating the basis of genetics and biochemistry. Flower color is developed by the accumulation of flavonoids (including anthocyanins), carotenoids, and betalains. Anthocyanin pigments act as major components in flower color development. The structural genes, which control anthocyanin and flavonol biosynthesis pathway, have been studied to some extent. Mutations in these genes create variations in soybean flower color. In addition, other factors, such as environment, regulatory genes, co-pigments, and vacuolar pH also play an important role in flower coloration. However, the factors influencing biosynthesis and regulation of anthocyanins are still unclear in soybean. Therefore, detailed studies are necessary to understand the whole mechanism of flower coloration. Furthermore, the available knowledge of active transposable elements, such as *Tgs1*, *Tgm9*, and *Tgm-Express1*, may provide a path for the utilization of these active transposons as tagging tools in soybean.

Acknowledgments This work was carried out with the support of the 'Next-Generation BioGreen21 Program for Agriculture & Technology Development' (Project No. PJ01109202), Rural Development Administration, Republic of Korea.

References

Buzzell RI, Buttery BR, Bernard RL (1977) Inheritance and linkage of magenta flower gene in soybeans. Can J Genet Cytol 19:749–751

Buzzell RI, Buttery BR, MacTavish DC (1987) Biochemical genetics of black pigmentation of soybean seed. J Hered 78:53–54

Chen Y, Nelson RL (2004) Identification and characterization of a white-flowered wild soybean plant. Crop Sci 44:339–342

Chuck G, Robbins T, Nijjar C, Ralston E, Courtney-Gutterson N, Dooner HK (1993) Tagging and cloning of a petunia flower color gene with the maize transposable element *Activator*. Plant Cell 5:371–378

deVlaming P, Schram AW, Wiering H (1983) Genes affecting flower color and pH of flower limb homogenates in *Petunia hybrida*. Theor Appl Genet 66:271–278

Fasoula DA, Stephens PA, Nickell CD, Vodkin LO (1995) Cosegregation of purple-throat flower color with dihydroflavonol reductase polymorphism in soybean. Crop Sci 35:1028–1031

Forkmann G (1991) Flavonoids as flower pigments: the formation of the natural spectrum and its extension by genetic engineering. Plant Breed 106:1–26

Gillman JD, Tetlow A, Lee JD, Shannon JG, Bilyeu K (2011) Loss-of-function mutations affecting a specific *Glycine max* R2R3 MYB transcription factor result in brown hilum and brown seed coats. BMC Plant Biol 11:155

Gollop R, Farhi S, Perl A (2001) Regulation of the *leucoanthocyanidins dioxygenase* gene expression in *Vitis vinifera*. Plant Sci 161:579–588

Gould KS (2004) Nature's Swiss army knife: the diverse of protective roles of anthocyanins in leaves. J Biomed Biotechnol 5:314–320

Groose RW, Palmer RG (1991) Gene action governing anthocyanin pigmentation in soybean. J Hered 82:498–501

Grotewold E (2006) The genetics and biochemistry of floral pigments. Annu Rev Plant Biol 57:761–780

Han Y, Vimolmangkang S, Soria-Guerra RE, Rosales-Mendoza S, Zheng D, Lygin AV, Korban SS (2010) Ectopic expression of apple *F3′H* genes contributes to anthocyanin accumulation in the Arabidopsis *tt7* mutant grown under nitrogen stress. Plant Physiol 153:806–820

Harborne JB (1967) Comparative biochemistry of the flavonoids. Academic Press, London

Hartwig EE, Hinson K (1962) Inheritance of flower color of soybeans. Crop Sci 2:152–153

He F, Mu L, Yan GL, Liang NN, Pan QH, Wang J, Reeves MJ, Duan CQ (2010) Biosynthesis of anthocyanins and their regulation in colored grapes. Molecules 15:9057–9091

Holton TA, Cornish EC (1995) Genetics and biochemistry of anthocyanin biosynthesis. Plant Cell 7:1071–1083

Hostel W (1981) In the biochemistry of plants. In: Stumpf W, Conn PM (eds) Secondary plants products, vol 7. Academic press, New York, pp 725–753

Iwashina T, Githiri SM, Benitez ER, Takemura T, Kitajima J, Takahashi R (2007) Analysis of flavonoids in flower petals of soybean near-isogenic lines for flower and pubescence color genes. J Hered 98:250–257

Iwashina T, Oyoo ME, Khan NA, Matsumura H, Takahashi R (2008) Analysis of flavonoids in flower petals of soybean flower color variants. Crop Sci 48:1918–1924

Johnson EOC, Stephens PA, Fasoula DA, Nickell CD, Vodkin LO (1998) Instability of a novel multicolored flower trait in inbred and outcrossed soybean lines. J Hered 89:508–515

Koes RE, Spelt CE, Reif HJ, van den Elzen PJ, Veltkamp E, Mol JN (1986) Floral tissue of *Petunia hybrida* (V30) expresses only one member of the chalcone synthase multigene family. Nucleic Acids Res 14:5229–5239

Koes R, Verweij W, Quattrocchio F (2005) Flavonoids: a colorful model for the regulation and evolution of biochemical pathways. Trends Plant Sci 10:236–242

Kovinich N, Saleem A, Arnason JT, Miki B (2010) Functional characterization of a UDP-glucose: flavonoid 3-*O*-

glucosyltransferase from the seed coat of black soybean (*Glycine max* (L.) Merr.). Phytochemistry 71:1253–1263
Kovinich N, Saleem A, Arnason JT, Miki B (2011) Combined analysis of transcriptome and metabolite data reveals extensive differences between black and brown nearly-isogenic soybean (*Glycine max*) seed coats enabling the identification of pigment isogenes. BMC Genom 12:381
Lepiniec L, Debeaujon I, Routaboul JM, Baudry A, Pourcel L, Nesi N, Caboche M (2006) Genetics and biochemistry of seed flavonoids. Annu Rev Plant Biol 57:405–430
Mol J, Grotewold E, Koes R (1998) How genes paints flowers and seeds. Trends Plant Sci 3:212–217
Moreau C, Ambrose MJ, Turner L, Hill L, Ellis TH, Hofer JM (2012) The *b* gene of pea encodes a defective flavonoid 3′,5′-hydroxylase, and confers pink flower color. Plant Physiol 159:759–768
Nagai I (1926) On the genetics of soybean (Japanese). Agric Hortic 1(1–14):107–108
Nagata T, Todoriki S, Masumizu T, Suda I, Furuta S, Du Z, Kikuchi S (2003) Levels of active oxygen species are controlled by ascorbic acid and anthocyanin in Arabidopsis. J Agric Food Chem 51:2992–2999
Palmer RG, Groose RW (1993) A new allele at the *w4m* locus derived from the *w4-m* mutable allele in soybean. J Hered 84:217–300
Palmer RG, Hedges BR, Benavente RS, Groose RW (1989) *w4*-mutable line in soybean. Dev Genet 10:542–551
Palmer RG, Pfeiffer TW, Buss GR, Kilen TC (2004) Qualitative genetics. In: Specht JE, Boerma HR (eds) Soybean: improvement, production, and uses. Agronomy monograph 16, 3rd edn. American Society of Agronomy, Madison Inc., Madison, pp 137–234
Park GT, Sundaramoorthy J, Park JB, Lee JD, Choi KS, Kim JH, Seo HS, Park SK, Song JT (2014) Diversity of the *W1* gene encoding flavonoid 3′,5′-hydroxylase in white- and purple-flowered soybeans. Plant Genet Resour. doi:10.1017/S1479262114000938
Schwinn KE, Davies KM (2004) Flavonoids. In: Davies KM (ed) Plant pigment and their manipulation. Annual plant reviews, vol 14. Blackwell Publishing, Oxford, pp 92–149
Song QJ, Marek LF, Shoemaker RC, Lark KG, Concibido VC, Delannay X, Specht JE, Cregan PB (2004) A new integrated genetic linkage map of the soybean. Theor Appl Genet 109:122–128
Sparvoli F, Martin C, Scienza A, Gavazzi G, Tonelli C (1994) Cloning and molecular analysis of structural genes involved in flavonoid and stilbene biosynthesis in grape (*Vitis vinifera* L.). Plant Mol Biol 24:743–755
Springob K, Nakajima J, Yamazaki M, Saito K (2003) Recent advances in the biosynthesis and accumulation of anthocyanins. Nat Prod Rep 20:288–303
Stephens PA, Nickell CD (1992) Inheritance of pink flower in soybean. Crop Sci 32:1131–1132
Stracke R, Werber M, Weisshaar B (2001) The R2R3-MYB gene family in *Arabidopsis thaliana*. Curr Opin Plant Biol 4:447–456
Takahashi Y, Fukuyama J (1919) Morphological and genetic studies on the soybean (Japanese). Hokkaido Agriculture Experiment Station Report 10
Takahashi R, Githiri SM, Hatayama K, Dubouzet EG, Shimada N, Aoki T, Ayabe S, Iwashina T, Toda K, Matsumura H (2007) A single-base deletion in soybean *flavonol synthase* gene is associated with magenta flower color. Plant Mol Biol 63:125–135
Takahashi R, Matsumura H, Oyoo ME, Khan NA (2008) Genetic and linkage analysis of purple-blue flower in soybean. J Hered 99:593–597
Takahashi R, Dubouzet JG, Matsumura H, Yasuda K, Iwashina T (2010) A new allele of flower color gene *W1* encoding flavonoid 3′5′-hydroxylase is responsible for light purple flowers in wild soybean *Glycine soja*. BMC Plant Biol 10:155
Takahashi R, Benitez ER, Oyoo ME, Khan NA, Komatsu S (2011) Nonsense mutation of an MYB transcription factor is associated with purple-blue flower color in soybean. J Hered 102:458–463
Takahashi R, Morita Y, Nakayama M, Kanazawa A, Abe J (2012) An active CACTA-family transposable element is responsible for flower variegation in wild soybean *Glycine soja*. Plant Genome 5:62–70
Takahashi R, Yamagishi N, Yoshikawa N (2013) A MYB transcription factor controls flower color in soybean. J Hered 104:149–153
Tanaka Y, Brugliera F (2013) Flower colour and cytochromes P450. Philos Trans R Soc B 368:20120432
Tanaka Y, Katsumoto Y, Brugliera F, Mason J (2005) Genetic engineering in floriculture. Plant Cell Tissue Organ Cult 80:1–24
To KY, Wang CK (2006) Molecular breeding of flower color. In: da Silva JAT (ed) Floriculture, ornamental, and plant biotechnology: advances and topical issues, vol 1. Global Science Books, Middlesex, pp 300–310
Toda K, Yang D, Yamanaka N, Watanabes S, Harada K, Takahashi R (2002) A single-base deletion in soybean *flavonoid 3′-hydroxylase* gene is associated with gray pubescence color. Plant Mol Biol 50:187–196
Todd JJ, Vodkin LO (1993) Pigmented soybean (*Glycine max*) seed coats accumulate proanthocyanidins during development. Plant Physiol 102:663–670
Tuteja JH, Clough SJ, Chan WC, Vodkin LO (2004) Tissue-specific gene silencing mediated by a naturally occurring *chalcone synthase* gene cluster in *Glycine max*. Plant Cell 16:819–835
Tuteja JH, Zabala G, Varala K, Hudson M, Vodkin LO (2009) Endogenous, tissue-specific short interfering RNAs silence the *chalcone synthase* gene family in *Glycine max* seed coats. Plant Cell 21:3063–3077
Wang CS, Todd JJ, Vodkin LO (1994) Chalcone synthase mRNA and activity are reduced in yellow seed coats with dominant *I* alleles. Plant Physiol 105:739–748
Wang CK, Chen PY, Wang HM, To KY (2006) Cosuppression of tobacco chalcone synthase using petunia chalcone synthase construct results in white flowers. Bot Stud 47:71–82
Woodworth CM (1923) Inheritance of growth habit, pod color and flower color in soybeans. J Am Soc Agron 15:481–495
Xu M, Palmer RG (2005) Genetic analysis and molecular mapping of a pale flower allele at the *W4* locus in soybean (*Glycine max* (L.) Merr.). Genome 48:334–340
Xu M, Brar HK, Grosic S, Palmer RG, Bhattacharyya MK (2010) Excision of an active CACTA-like transposable element from *DFR2* causes variegated flowers in soybean (*Glycine max* (L.) Merr.). Genetics 184:53–63
Yan F, Di S, Rodas FR, Torrico TR, Murai Y, Iwashina T, Anai T, Takahashi R (2014) Allelic variation of soybean flower color gene *W4* encoding dihydroflavonol 4-reductase 2. BMC Plant Biol 14:58
Yang K, Jeong N, Moon JK, Lee YH, Lee SH, Kim HM, Hwang CH, Back K, Palmer RG, Jeong SC (2010) Genetic analysis of genes controlling natural variation of seed coat and flower colors in soybean. J Hered 101:757–768
Yoshida K, Mori M, Kondo T (2009) Blue flower color development by anthocyanins: from chemical structure to cell physiology. Nat Prod Rep 26:884–915
Zabala G, Vodkin LO (2003) Cloning of the pleiotropic *T* locus in soybean and two recessive alleles that differentially affect structure and expression of the encoded flavonoid 3′ hydroxylase. Genetics 163:295–309
Zabala G, Vodkin LO (2005) The *wp* mutation of *Glycine max* carries a gene-fragment-rich transposon of the CACTA superfamily. Plant Cell 17:2619–2632
Zabala G, Vodkin LO (2007a) A rearrangement resulting in small tandem repeats in the *F3′5′H* gene of white flower genotypes is associated with the soybean *W1* locus. Crop Sci 47:113–124
Zabala G, Vodkin LO (2007b) Novel exon combinations generated by alternative splicing of gene fragments mobilized by a CACTA transposon in *Glycine max*. BMC Plant Biol 7:38