Nonsense Mutation of an MYB Transcription Factor Is Associated with Purple-Blue Flower Color in Soybean

RYOJI TAKAHASHI, EDUARDO R. BENITEZ, MAURICE E. OYOO, NISAR A. KHAN, AND SETSUKO KOMATSU

From the National Institute of Crop Science, Tsukuba, Ibaraki, 305-8518 Japan (Takahashi, Benitez, Khan, and Komatsu); and the Graduate School of Life and Environmental Sciences, University of Tsukuba, Tsukuba, Ibaraki, 305-8518 Japan (Takahashi and Oyoo). M.E. Oyoo is now at the Kenya Agricultural Research Institute-Mtwapa, PO Box 16, Mtwapa, 80109, Kenya. N.A. Khan is now at the Central Research Institute of Electric Power Industry, Abiko, Chiba, 270-1194 Japan.

Address correspondence to R. Takahashi at the address above, or e-mail: masako@affrc.go.jp.

Previous studies revealed that the recessive allele of the W2 locus generated purple-blue color and high vacuolar pH of flower petals in soybean. The location of W2 gene was reportedly close to simple sequence repeat marker Satt318 in molecular linkage group B2. We used information from the soybean genome to clone a candidate gene for W2. An MYB transcription factor gene belonging to G20 group was found in the vicinity of Satt318. Full-length cDNAs were cloned from purple-flowered cultivar Harosoy (W2 allele) and purple-blue flowered cultivars, Nezumisaya and w2-20 (w2 allele), by reverse transcription–PCR and designated as GmMYB-G20-1. Its open reading frame was 1083 bp long that encoded 361 amino acids in Harosoy. GmMYB-G20-1 had 53.7% similarity in amino acid sequence with the PH4 gene of petunia controlling blueness and vacuolar pH of flower petals. GmMYB-G20-1 of Nezumisaya and w2-20 had 3 base substitutions compared with that of Harosoy. The first substitution generated a stop codon in the MYB domain, resulting in truncated polypeptides. Cleaved amplified polymorphic sequence (CAPS) marker was developed to detect the base substitution. The polymorphic CAPS marker co-segregated with alleles at the W2 locus in the F2 population. These results suggest that GmMYB-G20-1 might correspond to the W2 gene.

Key words: blue flower, Glycine max, MYB transcription factor, soybean, W2 gene

Flower color of soybean (Glycine max (L.) Merr.) is primarily controlled by 6 genes, namely, W1, W2, W3, W4, Wm, and Wp (reviewed by Palmer et al. 2004; Takahashi et al. 2008). Among them, W1, W3, W4, Wm, and Wp encode structural genes for flavonoid biosynthesis: W1, flavonoid 3’5’-hydroxylase (Buzzell et al. 1987; Zabala and Vodkin 2005); W3 and W4, dihydroflavonol 4-reductase (Fasoula et al. 1995; Xu et al. 2010); Wm, flavonol synthase (Buzzell et al. 1977; Takahashi et al. 2006); Wp, flavanone 3-hydroxylase (Zabala and Vodkin 2005). The molecular function and genomic location of W2 have not yet been unequivocally determined thus far.

Nagai (1926) crossed a white-flowered cultivar, Shakujou, with a purple-flowered cultivar, Nakaide. The F2 plants segregated into 9 purple:3 purple-blue:4 white flowers. Based on these results, Matsuura (1933) assigned the gene symbol W2; w2 produces purple-blue flower in combination with W1. Takahashi et al. (2008) repeated the cross between Shakujou and Nakaide and observed that the 77 F2 plants segregated into 36 purple:20 purple-blue:21 white flower in agreement with the results (χ² = 3.54, 0.1 < P < 0.2) of Nagai (1926). A line fixed for blue-purple flower was developed from the cross and was designated as w2-20.

Takahashi et al. (2008) performed genetic analysis by crossing between a purple-blue flowered cultivar Nezumisaya and a purple-flowered cultivar Harosoy and revealed that the flower color of Nezumisaya is controlled by a single gene whose recessive allele is responsible for purple-blue flower. Complementation analysis performed by crossing with w2-20 revealed that flower color of Nezumisaya is controlled by the W2 locus. Linkage mapping revealed that W2 was located between 2 simple sequence repeat (SSR) markers, Satt318 and Satt020, at 1.1 cM apart from Satt318 in molecular linkage group B2 (Takahashi et al. 2008). Yang et al. (2010) developed a marker SL017 based on the sequence of soybean expressed sequence tags (ESTs) similar to the tomato R2R3 MYB gene ANT1 involved in anthocyanin biosynthesis. They mapped SL017 to a chromosomal location similar to W2.

Previous studies revealed that alleles at W1, W3, W4, Wm, and Wp loci affect the amount and/or structure of flavonoids (Iwashina et al. 2007, 2008; Takahashi et al. 2010). In contrast, flavonoids in purple-blue flowers were similar to those in purple flowers, suggesting that quantitative or structural differences of anthocyanins or copigmentation...
with other flavonoids may not be responsible (Iwashina et al. 2008). Blue flower color generally depends on the production of appropriate anthocyanin pigments like delphinidin-derived anthocyanins, self-association, intra- or intermolecular copigmentation, association with metal ions, and the vacuolar pH (Davies and Schwinn 1997; Mol et al. 1998). In cornflower, blue color is produced by a tetramer complex consisting of 6 molecules each of cyanidin-type anthocyanin and flavone, with one ferric iron, one magnesium, and 2 calcium ions (Shiono et al. 2005). In morning glory, vacuolar pH of flower petals is relatively low in flower buds, but upon further maturation, the vacuolar pH increases and the petal acquire a strong blue color (Yoshida et al. 1995). This color change and the increase of vacuolar pH require a Na+/H+ exchanger encoded by PURPLE gene (Fukada-Tanaka et al. 2000).

Flower petals of petunia normally have a lower pH than morning glory flowers, and the color of wild-type flowers stays on the reddish (low pH) side of the color spectrum (de Vlaming et al. 1983). Takahashi et al. (2008) measured pH of sap obtained from soybean banner petals of purple and purple-blue flowers. Purple flowers had a pH value of 5.73–5.77, whereas purple-blue flowers had a value of 6.07–6.10. It may be possible to obtain more distinct varietal differences in vacuolar pH similar to that obtained for morning glory (Yoshida et al. 1995; Yamaguchi et al. 2001) if separation and pH measurement of epidermal cells where anthocyanins are deposited is possible in soybean. These results suggested that W2 is responsible for vacuolar acidification of flower petals and that vacuolar pH of soybean flowers was largely similar to that of petunia. In petunia, 7 genes, PH1 to PH7, have been identified that, when mutated, cause a bluer flower color and an increase in the pH of crude petal extracts, suggesting that these genes are required for acidification of the vacuole (de Vlaming et al. 1983; Chuck et al. 1993; van Houwelingen et al. 1998). Among them, PH4 and PH6 have been cloned and determined to encode transcription factors, an R2R3 MYB protein and a basic helix-loop-helix (bHLH) protein, respectively (Spelt et al. 2002; Quattrochio et al. 2006). This study was conducted to clone a candidate gene for W2 using information from the soybean genome.

Materials and Methods

Plant Materials

The Canadian cultivar Harosoy, the US cultivar Clark with purple flowers (W1W1 W2W2 w3w3 W4W4 WmWm WpWp), the Japanese landrace Nezumisaya, and the experimental line w2-20 with purple-blue flowers (W1W1 w2w2 w3w3 W4W4 WmWm WpWp) were used in this study. In addition, an F2 population derived from a cross between Nezumisaya and Harosoy (Takahashi et al. 2008) was used for linkage mapping. Alleles at W2 locus in each F2 plant was determined by flower color segregation of the respective F3 families.

RNA Extraction and cDNA Cloning

Total RNA was extracted from 200 mg of banner petals from Harosoy, Clark, Nezumisaya, and w2-20 using the TRIZOL Reagent (Invitrogen) according to the manufacturer’s instructions. cDNA was synthesized by reverse transcription of 5 μg of total RNA from flower petals using the Superscript III First-Strand Synthesis System (Invitrogen) and an oligo(dT) primer according to the manufacturer’s instructions. The full-length cDNA was cloned by end-to-end PCR from Harosoy, Nezumisaya, and w2-20, using a pair of PCR primers (5'-TCCCTTAAATTCAAGAAGAGAATGG-3' and 5'-GATGAAGCTGTGAAACCAATA-3') designed based on the genome sequence of US cultivar Williams 82 deposited in the soybean genome database (Phytozome, http://www.phytozome.net/soybean.php). The PCR mixture contained 0.5 μg of cDNA, 10 pmol of each primer, 5 pmol of nucleotides, and 1 unit of ExTaq in 1 x ExTaq Buffer supplied by the manufacturer (Takara Bio) in a total volume of 25 μl. A 5 min denaturation at 94 °C was followed by 30 cycles of 30-s denaturation at 94 °C, 1 min annealing at 53 °C, and 1 min extension at 72 °C. A final 7 min extension at 72 °C completed the program. The PCR was performed in an Applied Biosystems 9700 thermal cycler. The ~1 kbp PCR product was cloned into pCR 2.1 vector (Invitrogen) and sequenced.

Sequencing Analysis

Nucleotide sequences of both strands were determined with the BigDye terminator cycle method using an ABI3100 Genetic Analyzer (Applied Biosystems). Primers for sequencing the internal regions were 5'-ACCTCCTCCAGTGATCATCA-3’ and 5’-AGGAACGGGCAATAGATGG-TG-3’ . Nucleotide sequences and the putative amino acid translations were analyzed with the BLAST program (Altschul et al. 1997). Gene structure was estimated based on the comparison between cDNA sequences and the corresponding genome sequences of Williams 82 deposited in the soybean genome database. Sequence alignment was performed with ClustalW (http://clustalw.ddbj.nig.ac.jp/top-j.html) using default settings.

CAPS Analysis and Linkage Mapping

Genomic DNA of Harosoy, Clark, Nezumisaya, w2-20, and 92 F2 plants was isolated from trifoliolate leaves by hexadecyl-trimethylammonium bromide (Murray and Thompson 1980). CAPS PCR primers (5’-CTCCATGCTGCAACAAATGTGT-3’, 5’-ATGAGATCTTCTCGTCGG-3’) were designed to detect a single-base substitution found in Nezumisaya. The base substitution would result in presence/absence of the FokI restriction site in the amplified product to generate a polymorphism (Figure 1). The PCR mixture contained 30 ng of genomic DNA, 25 pmol of each primer, 5 pmol of nucleotides, and 1 unit of ExTaq in 1 x ExTaq Buffer supplied by the manufacturer (Takara Bio) in a total volume of 25 μl. After an initial 30-s denaturation at 94 °C, there were 30 cycles of 30-s denaturation at 94 °C, 1 min annealing at 59 °C, and 1 min
extension at 72 °C. A final 7 min extension at 72 °C completed the program. The amplified products were digested with FokI, and the digests were separated on an 8% nondenaturing polyacrylamide gel in 1/2×TBE buffer (90 mM Tris-borate, 2 mM ethylenediaminetetraacetic acid, pH 8.0). After electrophoresis, the gel was stained with ethidium bromide and the DNA fragments were visualized under UV light. A linkage map was constructed by combining the CAPS marker genotypes and SSR marker genotypes previously examined in the F2 population (Takahashi et al. 2008) using MAPMAKER/EXP. ver. 3.0 (Lander et al. 1987) with the threshold logarithm of odds score of 3.0.

**Semiquantitative RT-PCR Analysis**

To test the transcription level of the GmMYB-G20-1 gene, PCR reactions were carried out in a volume of 25 μl, using 125 ng of cDNA. A soybean actin gene (GenBank accession number: J01298) was used as a control. The number of cycles was optimized and chosen not to exceed the mid-log phase of product yield (22 cycles for actin and 30 cycles for GmMYB-G20-1). For GmMYB-G20-1, the initial 30-s denaturation at 94 °C was followed by 22 cycles of 30-s denaturation at 94 °C, 1 min annealing at 56 °C and 1 min extension at 72 °C. A final 7 min extension at 72 °C completed the program. The primers were 5’-GGCGAAGCTGGCCTCCGCTTGATAGAAGCTAGCCAAGGCCTCCTCGTC and 5’-GGCGAAGCTGGCCTCCGCTTGATAGAAGGCCTCCTCGTC. For actin, the initial 30-s denaturation at 94 °C was followed by 22 cycles of 30-s denaturation at 94 °C, 1 min annealing at 56 °C and 1 min extension at 72 °C. A final 7 min extension at 72 °C completed the program. The primers were 5’-GACGCTGAGGATATTCAACC and 5’-AGAAATCTGTGAGGTCACGA. PCR products were loaded on a 2% agarose gel, stained by ethidium bromide and visualized under UV light.

**Accession Numbers**

Sequence data from this article have been deposited with the DDBJ Data Libraries under accession numbers AB597932 (Harosoy) and AB597933 (Nezumisaya).

**Results and Discussion**

**cDNA Cloning**

A survey of the compiled genome sequence of the US cultivar Williams 82 from Satt318 (Gm14:28809850) to Satt020 (Gm14:42022366) revealed the existence of 3 MYB transcription factors (Glyma14g24290.1, Glyma14g24500.1, and Glyma14g27260.1). Sequence alignment suggested that Glyma14g24500.1 (Gm14:29,396,832 to Gm14:29,398,217), located in the vicinity of Satt318, was similar to the petunia PH4 gene. The putative amino acid sequences of the other 2 MYB genes and the bHLH gene were not similar to those of...
either the petunia \(PH4\) or \(PH6\) genes. A putative full-length cDNA of Glyma14g24500.1, about 1 kb long, was obtained by reverse transcription–PCR (RT–PCR) and end-to-end PCR using a pair of primers specific to the 5' and 3' ends. The cDNA was designated as \(GmMYB-G20-1\) based on the MYB subgroup name as described below. Sequence analysis revealed that its open reading frame contained 1086 nucleotides in Harosoy, Nezumisaya, and \(w_2\)-20. The putative polypeptide (GmMYB-G20-1) consisted of 361 amino acids was deduced from the cDNA sequence of Harosoy (Figure 1). The apparent molecular mass and isoelectric point were 40,500 and 6.60, respectively.

MYB transcription factors are classified in 3 subfamilies depending on the numbers (1, 2, or 3) of adjacent repeats of helix-helix-turn-helix motifs that potentially bind with DNA. \(GmMYB-G20-1\), like the majority of plant MYB proteins, contains only the R2 and R3 repeats (Figure 1). MYBs are further subdivided into 42 subgroups based on similarities in the encoded proteins and intron–exon structures (Jiang et al. 2004). \(GmMYB-G20-1\) consisted of 361 amino acids was deduced from the cDNA sequence of Harosoy (Figure 1). The apparent molecular mass and isoelectric point were 40,500 and 6.60, respectively.

CAPS Analysis and Linkage Mapping

PCR using CAPS primers produced a 220-bp fragment in all 4 cultivars, Harosoy, Clark, Nezumisaya, and \(w_2\)-20 (Figure 3). After \(FokI\) digestion, DNA amplified from purple-flowered cultivars Harosoy and Clark were digested to produce a band with fragment length of about 170 bp. In contrast, amplified DNA of purple-blue flowered cultivars Nezumisaya and \(w_2\)-20 remained undigested. CAPS analysis generated distinct fragments in 89 of the 92 \(F_2\) plants. \(FokI\) restriction patterns of the \(F_2\) population co-segregated with flower color: \(F_2\) plants homozygous for \(W_2\) (purple flower) had only the short (170 bp) fragment; \(F_2\) plants fixed for \(w_2\) (purple-blue flower) had the undigested 220-bp fragment; heterozygous \(F_2\) plants (purple flower) had both bands (Figure 3). Linkage mapping suggested that \(GmMYB-G20-1\) was located at the same position with the \(W_2\) gene. BLAST analysis indicated that the ESTs (BM093788 and BM092559) that Yang et al.
(2010) used to develop the marker SL017 were located on chromosome 13. However, a sequence in GmMYB-G20-1 was identical to the forward primer of SL017, suggesting a possibility that they actually mapped GmMYB-G20-1.

Semiquantitative RT–PCR Analysis

Results of semiquantitative RT–PCR analysis are presented in Figure 4. Specific primers for GmMYB-G20-1 and the actin control produced bands of 504 and 552 bp, respectively. Transcript levels in flower petals of purple-flowered cultivars, Harosoy, and Clark were higher than those of purple-blue-flowered cultivars, Nezumisaya, and w2-20. The lower transcript levels in purple-blue flowers may be due to degradation by a cellular mRNA surveillance mechanism known as nonsense-mediated decay that functions to detect nonsense mutations and prevent the expression of truncated or erroneous proteins (Chang et al. 2007). A similar mechanism was also suggested for mutants of the flavonoid 3'-hydroxylase and flavonol synthase genes in soybean (Toda et al. 2002; Takahashi et al. 2006).

High percent similarity with petunia PH4 gene, genomic location, and association between dominance relationship and nonsense mutation suggest that GmMYB-G20-1 corresponds to W2. Survey of the genome sequence of Williams 82 suggested the existence of another gene with comparable amino acid similarity with PH4 (52.0%, Glyma13g09980.1) on chromosome 13. Functional differences and expression specificity of the 2 soybean MYB genes remain to be investigated. Transgenic experiments should be conducted to ascertain the function of GmMYB-G20-1. Mutation of PH4 has little or no effect on synthesis and modification of anthocyanins (Quattrocchio et al. 2006) similar to the W2. Grotewold et al. (2000) reported that 4 amino acid residues (L77, R80, R83, and L84) in the R3 region of the maize R2R3 MYB transcription factor C1 are necessary for interaction with a bHLH protein R. These amino acids are conserved in both PH4 and GmMYB-G20-1 of Harosoy (Figure 2). GmMYB-G20-1 protein possibly interacts with other transcription factor(s) and eventually affects the downstream protein(s) involved in proton metabolism. Protein analysis including yeast 2-hybrid assays may be necessary to investigate the functional mechanisms of W2 gene leading to high vacuolar pH of flower petals.

Funding

The Japanese Government (MEXT) Scholarship (062155 to M.E.O.)

Acknowledgments

The authors are grateful to the NIAS Genebank for providing the seeds of Nezumisaya and Dr. Joseph G. Dubouzet (SCION Research, New Zealand) for critical reading of the manuscript.

References


Matsuura H. 1933. A bibliographical monograph on plant genetics. 2nd ed. Sapporo (Japan): Hokkaido Imperial University.


Received January 10, 2011; Revised March 7, 2011; Accepted March 18, 2011

Corresponding Editor: Reid Palmer