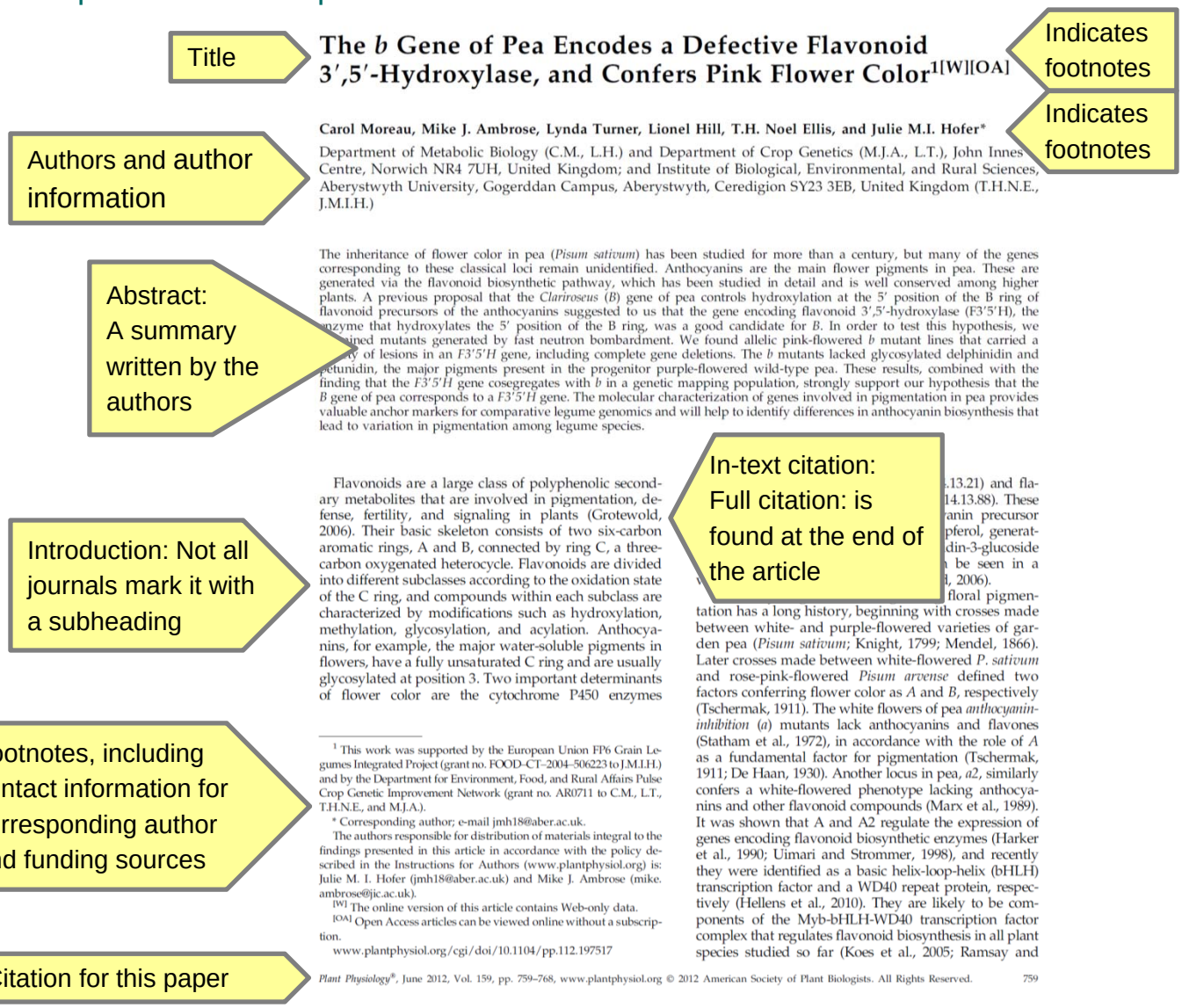
Case study: Reading a Primary Research Article from *Plant Physiology*

This case study examines a recent article published in the journal *Plant Physiology*. The full article is appended to this PDF. Because of space constraints, only the major points from the paper are covered in the case study, and the biochemical pathway is presented in simplified form.



Title

The *b* Gene of Pea Encodes a Defective Flavonoid 3′,5′-Hydroxylase, and Confers Pink Flower Color[1][W][OA]

Indicates footnotes

Indicates footnotes

Authors and author information

Carol Moreau, Mike J. Ambrose, Lynda Turner, Lionel Hill, T.H. Noel Ellis, and Julie M.I. Hofer*

Department of Metabolic Biology (C.M., L.H.) and Department of Crop Genetics (M.J.A., L.T.), John Innes Centre, Norwich NR4 7UH, United Kingdom; and Institute of Biological, Environmental, and Rural Sciences, Aberystwyth University, Gogerddan Campus, Aberystwyth, Ceredigion SY23 3EB, United Kingdom (T.H.N.E., J.M.I.H.)

Abstract: A summary written by the authors

The inheritance of flower color in pea (*Pisum sativum*) has been studied for more than a century, but many of the genes corresponding to these classical loci remain unidentified. Anthocyanins are the main flower pigments in pea. These are generated via the flavonoid biosynthetic pathway, which has been studied in detail and is well conserved among higher plants. A previous proposal that the *Clariroseus* (*B*) gene of pea controls hydroxylation at the 5′ position of the B ring of flavonoid precursors of the anthocyanins suggested to us that the gene encoding flavonoid 3′,5′-hydroxylase (F3′5′H), the enzyme that hydroxylates the 5′ position of the B ring, was a good candidate for *B*. In order to test this hypothesis, we obtained mutants generated by fast neutron bombardment. We found allelic pink-flowered *b* mutant lines that carried a variety of lesions in an *F3′5′H* gene, including complete gene deletions. The *b* mutants lacked glycosylated delphinidin and petunidin, the major pigments present in the progenitor purple-flowered wild-type pea. These results, combined with the finding that the *F3′5′H* gene cosegregates with *b* in a genetic mapping population, strongly support our hypothesis that the *B* gene of pea corresponds to a *F3′5′H* gene. The molecular characterization of genes involved in pigmentation in pea provides valuable anchor markers for comparative legume genomics and will help to identify differences in anthocyanin biosynthesis that lead to variation in pigmentation among legume species.

In-text citation: Full citation: is found at the end of the article

Introduction: Not all journals mark it with a subheading

Flavonoids are a large class of polyphenolic secondary metabolites that are involved in pigmentation, defense, fertility, and signaling in plants (Grotewold, 2006). Their basic skeleton consists of two six-carbon aromatic rings, A and B, connected by ring C, a three-carbon oxygenated heterocycle. Flavonoids are divided into different subclasses according to the oxidation state of the C ring, and compounds within each subclass are characterized by modifications such as hydroxylation, methylation, glycosylation, and acylation. Anthocyanins, for example, the major water-soluble pigments in flowers, have a fully unsaturated C ring and are usually glycosylated at position 3. Two important determinants of flower color are the cytochrome P450 enzymes

.13.21) and fla-
14.13.88). These
anin precursor
pferol, generat-
din-3-glucoside
be seen in a
, 2006).
floral pigmentation has a long history, beginning with crosses made between white- and purple-flowered varieties of garden pea (*Pisum sativum*; Knight, 1799; Mendel, 1866). Later crosses made between white-flowered *P. sativum* and rose-pink-flowered *Pisum arvense* defined two factors conferring flower color as *A* and *B*, respectively (Tschermak, 1911). The white flowers of pea *anthocyanin-inhibition* (*a*) mutants lack anthocyanins and flavones (Statham et al., 1972), in accordance with the role of *A* as a fundamental factor for pigmentation (Tschermak, 1911; De Haan, 1930). Another locus in pea, *a2*, similarly confers a white-flowered phenotype lacking anthocyanins and other flavonoid compounds (Marx et al., 1989). It was shown that A and A2 regulate the expression of genes encoding flavonoid biosynthetic enzymes (Harker et al., 1990; Uimari and Strommer, 1998), and recently they were identified as a basic helix-loop-helix (bHLH) transcription factor and a WD40 repeat protein, respectively (Hellens et al., 2010). They are likely to be components of the Myb-bHLH-WD40 transcription factor complex that regulates flavonoid biosynthesis in all plant species studied so far (Koes et al., 2005; Ramsay and

Footnotes, including contact information for corresponding author and funding sources

[1] This work was supported by the European Union FP6 Grain Legumes Integrated Project (grant no. FOOD-CT–2004–506223 to J.M.I.H.) and by the Department for Environment, Food, and Rural Affairs Pulse Crop Genetic Improvement Network (grant no. AR0711 to C.M., L.T., T.H.N.E., and M.J.A.).

* Corresponding author; e-mail jmh18@aber.ac.uk.

The authors responsible for distribution of materials integral to the findings presented in this article in accordance with the policy described in the Instructions for Authors (www.plantphysiol.org) is: Julie M. I. Hofer (jmh18@aber.ac.uk) and Mike J. Ambrose (mike.ambrose@jic.ac.uk).

[W] The online version of this article contains Web-only data.

[OA] Open Access articles can be viewed online without a subscription.

www.plantphysiol.org/cgi/doi/10.1104/pp.112.197517

Citation for this paper

Plant Physiology®, June 2012, Vol. 159, pp. 759–768, www.plantphysiol.org © 2012 American Society of Plant Biologists. All Rights Reserved. 759

The first page of a typical article from *Plant Physiology*. (See text for more information about each section)

It's like learning to ride a bike

It takes some practice to learn to read a scientific paper, but with a little effort you should be able to navigate your way around with confidence. This is a guided walk through a recently published article to get you started.

The Title

The title of the paper needs to be factual, informative, and concise. Most journals have a strict character limit; *Plant Physiology*'s limit is 150 characters. You could think of the title as the abstract in a tweet.

Q1. *What is the title of this article?*

The Authors

Single-author papers are rare, especially in biology. Most papers report the efforts of a team and so have two or more authors. Each author must have made a significant contribution to the research and writing of the paper (for guidelines about authorship, see www.plantphysiol.org/site/misc/ifora.xhtml#Authorship). Minor contributors can be recognized in the Acknowledgments section at the end of the paper. The order in which the authors are listed is important. Typically, the first person listed (the "first author") conducted much of the research and gets the most credit. Sometimes two or more people are given co–first authorship, which is usually indicated in a footnote of the paper. First authors are typically graduate students or postdoctoral researchers carrying out their research in the lab of a more senior scientist, who is typically listed in the last position. If the research involved a collaboration among more than one lab group, the senior authors are typically all listed at the end. In between the first-author position and the senior-author position, others who made significant contributions to the paper are listed. Some papers list the contributions of each author at the end of the paper. The corresponding author is indicated in the footnotes and is the contact person for the paper; usually the first or last author is the corresponding author. The institutional affiliations of each author are listed in this section.

Q2. *How many authors does this paper have? How many institutions and departments are represented? Who is the corresponding author?*

Q3. *Which organization provided the funds used to carry out the research, and which author was awarded the funds?*

The Abstract

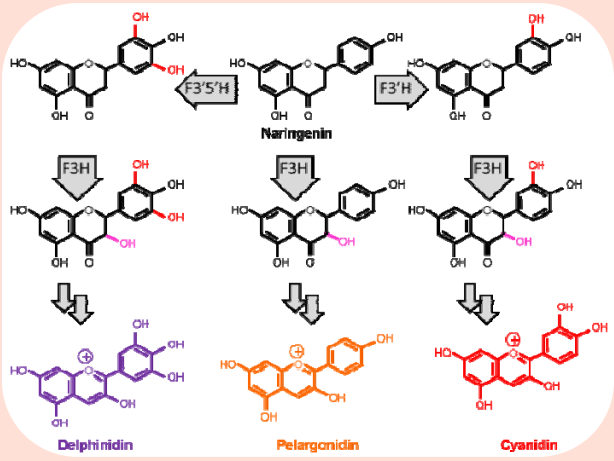
The abstract is a summary of the entire study. The authors highlight the question that they addressed, the methods they used, the hypotheses they tested, and the results of their experiments, and explain what their results mean, all in one paragraph (for *Plant Physiology*, a paragraph of no more than 250 words).

Q4. *After reading the abstract, what do you understand about this paper? What is the experimental organism being studied? In your own words, what question is being addressed in this paper?*

The Introduction

In some journals, including *Plant Physiology*, the heading "Introduction" is not used, but the introductory information is always the first part of the article after the abstract. The introduction provides the background and justification for the experiment. The introduction describes why the study was carried out and the question being investigated or hypotheses being tested. Statements of facts should be supported with a citation to another published article.

Moreau et al. (2012) introduces the pigment molecules found in flowers called anthocyanins, which are types of a category of chemicals called flavonoids. The article states that the biochemistry and genetics of anthocyanin production have been studied in peas and other plants. The first paper cited in the Introduction is a review article by Grotewold, published in 2006, called "The Genetics and Biochemistry of Floral Pigments," published in the *Annual Review of Plant Biology*. This review article gathers together a lot of information from many articles. It describes the biochemical pathway for anthocyanin synthesis and the reactions that are catalyzed by the enzymes discussed in Moreau et al. (2012) and so provides important background information. A simplified version of the anthocyanin biosynthetic pathway is shown below.

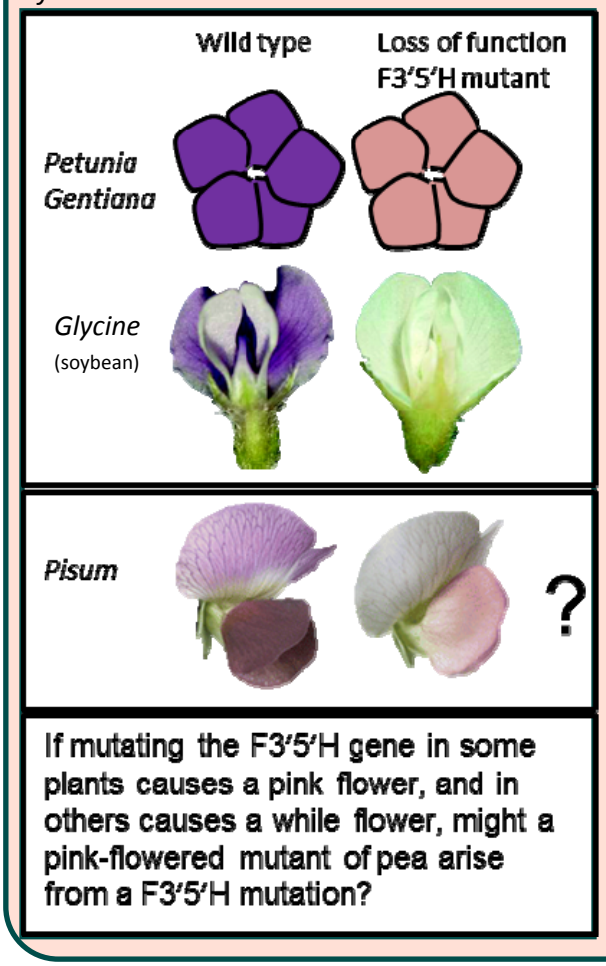


As shown in the figure above, a colorless precursor, Naringenin, can be hydroxylated at the 3′ position by the enzyme F3′H, or at the 3′ and 5′ positions by the enzyme F3′5′H. These compounds, as well as the unmodified precursor, are subsequently converted into pigments. The presence or absence of the hydroxyl groups affects the color of the pigments. Delphinidin (and related compound petunidin, not shown), are purple and are hydroxylated at both the 3′ and 5′ positions.

Moreau et al. (2012) states that the *b* mutant of pea, which has pink flowers rather than purple flowers, resembles some purple-turned-pink mutants in other plants (*Petunia* and *Gentiana*) that resulted from mutations in the *F3'5'H* gene. However, in *Glycine max* (soybean) plants mutated in the *F3'5'H* gene, the flowers are white, not pink like the flowers of the corresponding mutants of *Petunia* and *Gentiana*.

Moreau et al. (2012) wonder if the *b* mutant of pea, which has pink flowers, might arise from a loss-of-function mutation in the *F3'5'H gene*. Another possibility is that pea, which is a

legume related to soybean, might produce white flowers when the *F3'5'H* gene is mutated. To help clear up this apparent discrepancy, the authors decided to investigate the *b* mutant in pea. They set out to determine which gene is mutated in the *b* mutant and to address the role of the mutated gene on flower pigment synthesis.



Q5. *How many references are cited in the Introduction? How many of them would you want to look up to fully understand the study?*
Q6. *Do the authors clearly state what questions they addressed in this study?*

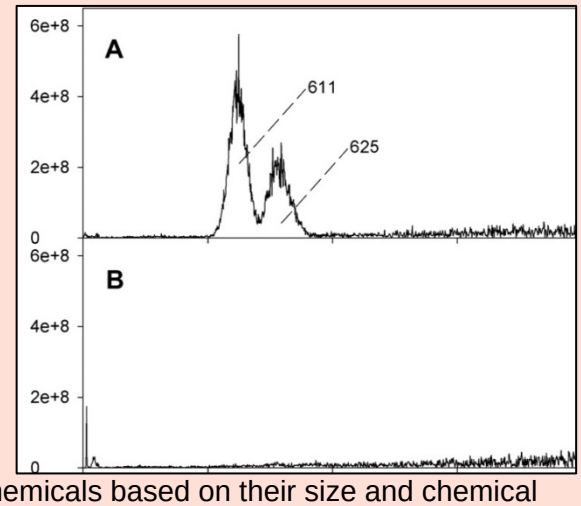
The Results
The results section includes a description of the research conducted and the results obtained. Results can be presented as tables, large datasets, and figures, which can include graphs, videos, diagrams, and photographs.

In the article by Moreau et al. (2012), the results are presented in four figures.

Figure 1 shows photographs of the experimental and control organisms to demonstrate the phenotypic effect of the *b* mutation. Figure 1A shows the pigmentation pattern of a wild-type pea. Figure 1B shows a flower from a *b* mutant of pea that is less pigmented, and Figure 1C shows an unstable *b*

mutant in which the gene is active in some parts of the petal (the darker parts) and inactive in others (the lighter parts).

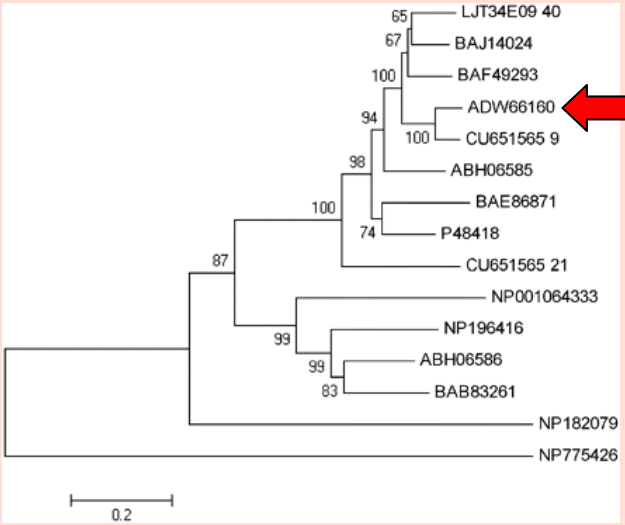
Figure 2 shows a chromatographic separation of pigments extracted from wild-type (A) and *b* mutant (B) petals. This instrument detects



chemicals based on their size and chemical properties. The peaks labeled 611 and 635 represent delphinidin and petunidin, which are present in the wild-type sample and absent in

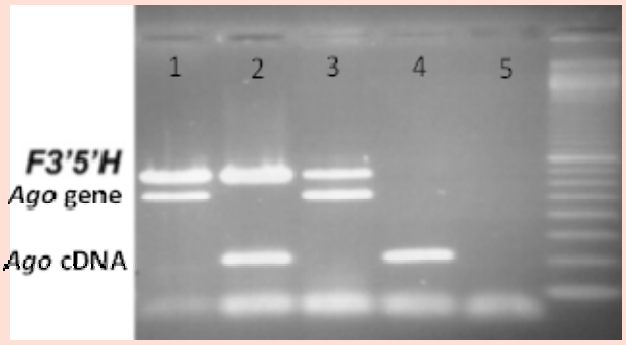
the *b* mutant sample. (Panels C and D, and supplemental Figure 1, show additional assays that identify the pigment profiles of wild-type and mutant petals).

Using a gene cloning method, the authors isolated the *F3'5'H* gene from wild-type peas based on homology to the gene from related plants. The sequence of the pea protein was compared to those of other plants. The sequence alignments are shown in Supplemental Figure 2, and a phylogenic representation is shown in **Figure 3**.



The pea enzyme (indicated by a red arrow) is most closely related to the enzyme found in another legume, *Medicago truncatula*, indicated as CU651565 9.

Finally, the authors show that the *b* mutant they identified and characterized has a defect in the expression of the *F3'5'H* gene. **Figure 4** shows a characterization of this gene and a cDNA copy of the mRNA transcript of the gene. The authors were able to amplify the gene from wild type and *b* mutant DNA (lanes 1 and 3). They were able to amplify the cDNA from wild-type plants (lane 2) but not *b* mutant plants (lane 4). The lower bands show the amplification product of a control gene (Ago gene, lanes 1 and 3) and cDNA (Ago cDNA, lanes 2 and 4). The presence of the AGO cDNA band in lane 4 (which lacks the *b* shows that there is nothing



wrong with this sample; the lack of a band is specifically because of the lack of the *b* cDNA, not a problem with the cDNA in the sample in general. Lane 5 is another control, which used no DNA or cDNA; the smudge at the bottom of the gel comes from residual PCR primers used in the assay. The lane on the right shows size standards.

Q7. *How straightforward was it to understand the data presented in the figures?*
Q8. *What information do the figure legends provide?*
Q9. *Where do you find information about the experimental methods used?*
Q10. *What information is presented in the Supplemental Materials? Why do you think some information is put in this supplementary section?*

The Discussion

This section summarizes the finding of this study and interprets how the new information integrates with previous knowledge.

Here is the key finding of Moreau et al. (2012): "In this paper, we have presented genetic and biochemical evidence to show that *b* mutants lack a functional *F3'5'H* gene that results in a rose-pink flower color due to the presence of cyanidin- and peonidin-based anthocyanins."

The first part of the discussion analyzes the type of mutations the authors identified in the *b* mutants, including the nature of the unstable *b* mutant shown in Figure 1C.

The next part compares the *F3'5'H* genes in legumes. An interesting observation the authors make is that although the pea gene is most closely related to the gene from *Medicago truncatula,* this plant makes yellow flowers! This observation points to the complexity of the biochemical pathway of anthocyanins, as well as the possibility that a single amino acid mutation in the *Medicago* gene might make the enzyme it encodes non-functional (something interesting to follow up on!).

The last part of the discussion talks about flower color in soybean, starting with a discussion of the F3H gene and mutations of it. The third-to-last paragraph of the discussion observes, "However, it is not clear why a *w1* encoding a defective *F3'5'H* gene would condition white flower color in soybean, when the pea *b* mutant and other *F3'5'H* mutants derived from purple-flowered wild-type plants have pink flowers." The authors point out that although the soybean study showed that the *w1* mutation lies very close to the *F3'5'H* gene, its identity has not been proven to be the *F3'5'H,* leaving open the possibility that the *w1* mutation is in a different gene. If *w1* is not a mutation of the *F3'5'H* gene, then we no longer have the puzzling issue of different phenotypes for *F3'5'H* mutations in different species (something interesting to follow up on!). Thus, by characterizing the *F3'5'H* gene in pea, these authors have contributed some clarity to a set of puzzling observations and provided a new hypothesis to follow up on.

Q11. *In your own words, how does the discussion section differ from the results section?*
Q12. *Find the places in the discussion that specifically refer to the results and the data presented in the supplemental materials. Is each result discussed similarly, or are different types of data discussed differently?*
Q13. *What information is conveyed by the final paragraph of the Discussion?*

The Materials and Methods

In this section, the authors describe the sources of the biological materials they used and the conditions of their growth and the experimental procedures that they followed. Additional information about their methods, including the sequence of the DNA primers they used for sequencing, can be found in the Supplemental Materials section of the paper.

Q14. *In what way is the font of this section different from the remainder of the paper? Why do you think this information is presented differently?*

The Acknowledgments

People who helped out with photography and plant care and technical assistance are given thanks here.

Q15. *Where do you find guidelines that specify what kinds of contributions are necessary for authorship versus those that are recognized in this section?*

The References

This section lists articles that were cited in the text. Some journals list them in alphabetic order, and others list them in the order in which they appear in the text.

Q16. *How many references are listed? How many include one or more of the authors who contributed to this paper?*
Q17. *What kinds of articles are listed in the references section, and what kinds of sources are not included?*
Q18. *How would you go about finding these references? How does reading the HTML online version of the article facilitate accessing references?*

For more information

For more information about the format and style of a scientific paper, you might find the "Instructions for Authors" guidelines interesting. Every journal provides specific guidelines about the content and format of the papers it will publish, and reading the instructions for authors will help to familiarize you with this format. The complete Instructions for Authors for *Plant Physiology* is found at http://www.plantphysiol.org/site/misc/ifora.xhtml .

See also:

American Society of Plant Biologists. (2013). How to read a scientific paper http://journalaccess.aspb.org/ReadaSciPaper.

Carpi, A., Egger, A.E., and Kuldell, N.H. (2008). Scientific communication: Understanding scientific journals and articles," *Visionlearning* Vol. POS-1 (9).http://www.visionlearning.com/library/module_viewer.php?mid=158

Pechenik, J. (2013). *A Short Guide to Writing about Biology.* Prentice Hall, New Jersey.

Written by Mary E. Williams (2013) for the American Society of Plant Biologists.
www.aspb.org

The *b* Gene of Pea Encodes a Defective Flavonoid 3′,5′-Hydroxylase, and Confers Pink Flower Color[1][W][OA]

Carol Moreau, Mike J. Ambrose, Lynda Turner, Lionel Hill, T.H. Noel Ellis, and Julie M.I. Hofer*

Department of Metabolic Biology (C.M., L.H.) and Department of Crop Genetics (M.J.A., L.T.), John Innes Centre, Norwich NR4 7UH, United Kingdom; and Institute of Biological, Environmental, and Rural Sciences, Aberystwyth University, Gogerddan Campus, Aberystwyth, Ceredigion SY23 3EB, United Kingdom (T.H.N.E., J.M.I.H.)

The inheritance of flower color in pea (*Pisum sativum*) has been studied for more than a century, but many of the genes corresponding to these classical loci remain unidentified. Anthocyanins are the main flower pigments in pea. These are generated via the flavonoid biosynthetic pathway, which has been studied in detail and is well conserved among higher plants. A previous proposal that the *Clariroseus* (*B*) gene of pea controls hydroxylation at the 5′ position of the B ring of flavonoid precursors of the anthocyanins suggested to us that the gene encoding flavonoid 3′,5′-hydroxylase (F3′5′H), the enzyme that hydroxylates the 5′ position of the B ring, was a good candidate for *B*. In order to test this hypothesis, we examined mutants generated by fast neutron bombardment. We found allelic pink-flowered *b* mutant lines that carried a variety of lesions in an *F3′5′H* gene, including complete gene deletions. The *b* mutants lacked glycosylated delphinidin and petunidin, the major pigments present in the progenitor purple-flowered wild-type pea. These results, combined with the finding that the *F3′5′H* gene cosegregates with *b* in a genetic mapping population, strongly support our hypothesis that the *B* gene of pea corresponds to a *F3′5′H* gene. The molecular characterization of genes involved in pigmentation in pea provides valuable anchor markers for comparative legume genomics and will help to identify differences in anthocyanin biosynthesis that lead to variation in pigmentation among legume species.

Flavonoids are a large class of polyphenolic secondary metabolites that are involved in pigmentation, defense, fertility, and signaling in plants (Grotewold, 2006). Their basic skeleton consists of two six-carbon aromatic rings, A and B, connected by ring C, a three-carbon oxygenated heterocycle. Flavonoids are divided into different subclasses according to the oxidation state of the C ring, and compounds within each subclass are characterized by modifications such as hydroxylation, methylation, glycosylation, and acylation. Anthocyanins, for example, the major water-soluble pigments in flowers, have a fully unsaturated C ring and are usually glycosylated at position 3. Two important determinants of flower color are the cytochrome P450 enzymes flavonoid 3′-hydroxylase (F3′H; EC 1.14.13.21) and flavonoid 3′,5′-hydroxylase (F3′5′H; EC 1.14.13.88). These hydroxylate the B ring of the anthocyanin precursor molecules naringenin and dihydrokaempferol, generating substrates for the production of cyanidin-3-glucoside and delphinidin-3-glucoside, which can be seen in a variety of pigmented flowers (Grotewold, 2006).

The study of genetic loci regulating floral pigmentation has a long history, beginning with crosses made between white- and purple-flowered varieties of garden pea (*Pisum sativum*; Knight, 1799; Mendel, 1866). Later crosses made between white-flowered *P. sativum* and rose-pink-flowered *Pisum arvense* defined two factors conferring flower color as *A* and *B*, respectively (Tschermak, 1911). The white flowers of pea *anthocyanin-inhibition* (*a*) mutants lack anthocyanins and flavones (Statham et al., 1972), in accordance with the role of *A* as a fundamental factor for pigmentation (Tschermak, 1911; De Haan, 1930). Another locus in pea, *a2*, similarly confers a white-flowered phenotype lacking anthocyanins and other flavonoid compounds (Marx et al., 1989). It was shown that A and A2 regulate the expression of genes encoding flavonoid biosynthetic enzymes (Harker et al., 1990; Uimari and Strommer, 1998), and recently they were identified as a basic helix-loop-helix (bHLH) transcription factor and a WD40 repeat protein, respectively (Hellens et al., 2010). They are likely to be components of the Myb-bHLH-WD40 transcription factor complex that regulates flavonoid biosynthesis in all plant species studied so far (Koes et al., 2005; Ramsay and

[1] This work was supported by the European Union FP6 Grain Legumes Integrated Project (grant no. FOOD–CT–2004–506223 to J.M.I.H.) and by the Department for Environment, Food, and Rural Affairs Pulse Crop Genetic Improvement Network (grant no. AR0711 to C.M., L.T., T.H.N.E., and M.J.A.).

* Corresponding author; e-mail jmh18@aber.ac.uk.

The authors responsible for distribution of materials integral to the findings presented in this article in accordance with the policy described in the Instructions for Authors (www.plantphysiol.org) is: Julie M. I. Hofer (jmh18@aber.ac.uk) and Mike J. Ambrose (mike.ambrose@jic.ac.uk).

[W] The online version of this article contains Web-only data.

[OA] Open Access articles can be viewed online without a subscription.

www.plantphysiol.org/cgi/doi/10.1104/pp.112.197517

Glover, 2005). The gene encoding the Myb component of this complex in pea, as well as genes at other loci involved in pigment production, such as *Clariroseus* (*B*), *Roseus* (*Ce*), and *Fuscopurpureus* (*Cr;* Statham et al., 1972), remain to be identified.

The major anthocyanins found in wild-type pea lines that contribute to their purple flower color are delphinidin-, petunidin-, and malvidin-3-rhamnoside-5-glucosides (Statham et al., 1972). Rose-pink *b* mutants (Blixt, 1972) produce a different range of anthocyanins (pelargonidin-, cyanidin-, and peonidin-3-rhamnoside-5-glucosides), suggesting that the *B* gene controls hydroxylation of the anthocyanin B ring (Statham et al., 1972) and encodes a hydroxylase. Pink-flowered mutants identified in species that are typically purple flowered, such as *Petunia* × *hybrida* (Snowden and Napoli, 1998; Matsubara et al., 2005) and *Gentiana scabra* (Nakatsuka et al., 2006), were found to have resulted from the insertion of transposable elements into the gene encoding F3′5′H. If anthocyanin biosynthesis in pea were to conform to the enzymatic steps elucidated in other plant species (Grotewold, 2006), then the activity missing in *b* mutants would be predicted to correspond to that of a F3′5′H.

In soybean (*Glycine max*), however, the *wp* locus, which conditions a change in flower color from purple to pink (Stephens and Nickell, 1992), was reported to encode a flavanone 3-hydroxylase (F3H; EC 1.14.11.9; Zabala and Vodkin, 2005). Furthermore, an insertion/deletion mutation in a gene encoding a F3′5′H was associated with the white-flowered phenotype of the soybean *w1* mutant (Zabala and Vodkin, 2007). These results suggested that anthocyanin biosynthesis in legumes, or at least in soybean, may differ from that in other plant species studied, where *F3′5′H* mutations result in pink flowers (Snowden and Napoli, 1998; Matsubara et al., 2005; Nakatsuka et al., 2006) and *F3H* mutations result in white flowers (Martin et al., 1991; Britsch et al., 1992). More recently, a *Glycine soja* accession carrying a *w1-lp* allele was described as having pale pink banner petals and a flower color designated as light purple (Takahashi et al., 2010). Our analysis here of the *b* mutant of pea, which is also a legume, addresses the complexity of these findings in soybean.

Transposon-tagged mutations have facilitated the isolation of genes involved in anthocyanin biosynthesis in numerous plant species, and transposon tagging is a useful technology for gene identification that remains particularly relevant for species without sequenced genomes, such as pea. Endogenous retrotransposons and DNA transposons have been identified in pea, but the transposition rate of those studied to date has been too low to be exploited for gene tagging (Shirsat, 1988; Vershinin et al., 2003; Macas et al., 2007). The identification of active DNA transposons usually occurs when sectors are found on pigmented flowers or seeds. Because most cultivated pea crop varieties have white flowers, any chance identification of sectored flowers in the field is extremely limited. A secondary purpose of this study was to carry out a screen for sectors on purple-flowered peas with the aim of identifying an active transposon.

We generated pink-flowered fast neutron (FN) deletion mutants and used these to identify the gene corresponding to *B*. Among the pigmentation mutants we obtained were several new *b* alleles, including pink-sectored mutants, which we characterized further. Stable pink *b* mutants were shown to carry a variety of lesions in an *F3′5′H* gene, including complete gene deletions. Analysis of one of these deletion lines showed that it lacked delphinidin and petunidin, the major anthocyanins of the progenitor wild-type pea variety. These results, combined with the finding that the *F3′5′H* gene cosegregates with *b* in a genetic mapping population, strongly support our hypothesis that the pea gene *b* corresponds to a *F3′5′H*.

RESULTS

Generation of New *b* Mutant Alleles

We used FN mutagenesis to generate pigmentation mutants in line JI 2822, which is wild type at the flower color loci *A*, *A2*, *Albicans* (*Am*), *B*, *Ce*, and *Cr*. The fully open petals of JI 2822 flowers are nonuniformly pigmented (Fig. 1A); the adaxial standard petal is pale purple, the two wing petals are dark purple, and the two fused abaxial keel petals are very lightly pigmented. The standard and wing petals fade to a blue purple. The JI 2822 flower is described here as purple to conform with previous naming conventions (De Haan, 1930).

M2 and M3 progeny from the mutagenized population were screened for flower color variants that differed from the wild type. Six FN lines were identified with pale pink standards, rose-pink wing petals, and lightly pigmented keel petals (Fig. 1B). Backcrosses to JI 2822 showed that four of these lines, FN 1076/6, FN 2160/1, FN 2255/1, and FN 2438/2, carried stable recessive mutations that determined the pink flower trait. These lines yielded rose-pink F1 progeny when crossed to the *b* mutant type line, JI 118, confirming that they carried allelic mutations. Two further lines, FN 2271/3/pink and FN 3398/2164, were stable rose-pink and allelic to *b*; however, sibling individuals carried flowers with pink sectors on a purple background (Fig. 1C), suggesting they were unstable at the *b* locus.

The *b* mutation is also known to confer paler stem axil pigmentation than the wild type and paler pod color in genotypes carrying the purple-podded *Pur* allele (De Haan, 1930; Statham et al., 1972). All six FN *b* alleles likewise differed from JI 2822 in having paler axillary rings. No effect on pod color was observed in the FN alleles, because JI 2822 is a green-podded genotype (*pur*). The FN *b* mutants are described here as rose pink to incorporate previous conventions (Tschermak, 1911; De Haan, 1930) yet distinguish them from cerise-pink *ce* and crimson-pink *cr* mutants.

Figure 1. Pea *b* mutant phenotypes. A, Purple-flowered wild-type line JI 2822. B, Rose-pink-flowered *b* mutant line FN 2271/3/pink. C, Unstable *b* mutant line FN 2271/3/flecked with rose-pink sectors on a purple background.

The *b* Mutant Lacks Delphinidin and Petunidin

Methanol-HCl extracts of anthocyanins from the wing petals of line JI 2822 and a stable pink M3 plant, FN 2271/3/pink, were analyzed using liquid chromatography (LC) coupled with mass spectroscopy (MS). Chromatograms with two major peaks showed that JI 2822 contained two major anthocyanins (Fig. 2A; 611 and 625 atomic mass units [amu]). MS data averaged across the peaks indicated that these were anthocyanins isomeric to delphinidin and petunidin glycosylated with deoxyhexose and hexose sugars (Supplemental Fig. S1). Fragmentation of the sugar moieties as mass losses of 146 and 162 amu were consistent with Rha and Glc, respectively. Fragmentation consistent with the loss of both monosaccharide moieties individually was observed, which suggested that the anthocyanidins delphinidin (303 amu) and petunidin (317 amu) were monoglycosylated at two different positions (Supplemental Fig. S1). These results agree with earlier studies that identified delphinidin-3-rhamnoside-5-glucoside and petunidin-3-rhamnoside-5-glucoside among the anthocyanins present in wild-type pea (Statham et al., 1972).

The peaks indicating glycosylated delphinidin and petunidin were absent from FN 2271/3/pink samples (Fig. 2B). A range of ions consistent with glycosylated cyanidin and peonidin were present in FN 2271/3/pink and absent from JI 2822 (Fig. 2, C and D). These were isomeric to cyanidin glycosylated with deoxyhexose and hexose sugars (595 amu), peonidin glycosylated with deoxyhexose and hexose sugars (609 amu), and cyanidin glycosylated with a pentose and two hexose sugars (743 amu; Fig. 2C). Fragmentation of the sugars attached to cyanidin (287 amu) as mass losses of 162 , 294, and 456 amu was consistent with a pentose moiety buried beneath a Glc moiety (Supplemental Fig. S1). No single loss of 132 amu, expected of an exposed pentose, was observed. These results confirmed earlier studies that identified cyanidin-3-sambubioside-5-glucoside among the anthocyanins present in *b* mutants (Statham et al., 1972). Fragmentation of the sugars attached to cyanidin and peonidin (301 amu) as mass losses of 146 and 162 amu was consistent with cyanidin-3-rhamnoside-5-glucoside and peonidin-3-rhamnoside-5-glucoside, also previously identified in *b* mutants (Statham et al., 1972).

The conversion of cyanidin and peonidin to delphinidin and petunidin requires hydroxylation at the 5′ position of the B ring of the precursor flavonoids. Because the products of this conversion were not observed in *b* mutants, it was presumed that the *B* gene controls the hydroxylation of the anthocyanin B ring (Statham et al., 1972). Our studies confirmed this conclusion and suggested to us that the gene encoding F3′5′H was a good candidate for *B*.

Isolation of a Pea *F3′5′H* Gene from a Purple-Flowered Wild-Type Plant

We performed PCR on cDNA derived from JI 2822 wing petals using primers based on aligned *Medicago truncatula* and soybean *F3′5′H* sequences. This yielded a product encoding a partial open reading frame (ORF) with extensive sequence similarity to *F3′5′H*. We used primers based on this new pea sequence together with primers based on the *Medicago* sequence for adaptor-ligation PCR (Spertini et al., 1999), which enabled us to isolate genomic DNA sequences and a larger cDNA product including a TAG stop codon. Amplification and sequencing of a single PCR product, using primers at the 5′ and 3′ ends of the surmised contig, confirmed that a 1,548-bp cDNA encoded a cytochrome P450 monooxygenase 515 amino acids long.

A BLASTP search of *Medicago* genome pseudomolecules (version 3.5) using the chromosome visualization tool CViT (http://www.medicagohapmap.org) identified CU651565_9 on bacterial artificial chromosome (BAC) CU651565, a F3′5′H 515 amino acids in length, as the most similar sequence, with 89% identity. The predicted pea protein sequence is 79%, 78%, and 75% identical to predicted full-length F3′5′H sequences from lotus (*Lotus japonicus*; LjT34E09.40), soybean (AAM51564, ABQ96218, and BAJ14024), and butterfly pea (*Clitoria ternatea*; BAF49293), respectively. The soybean sequences are classified as CYP75A17 cytochrome P450s (Nelson, 2009). The Arabidopsis (*Arabidopsis thaliana*) sequence most closely related to the pea F3′5′H (48% identity) is the cytochrome P450 monooxygenase CYP75B1, encoded by *TRANSPARENT TESTA7* (At5g07990; GenBank accession no. NP196416). This 513-amino acid protein has been demonstrated to have F3′H activity (Schoenbohm et al., 2000), and it lies within a separate clade when compared with other plant F3′5′H sequences (Fig. 3).

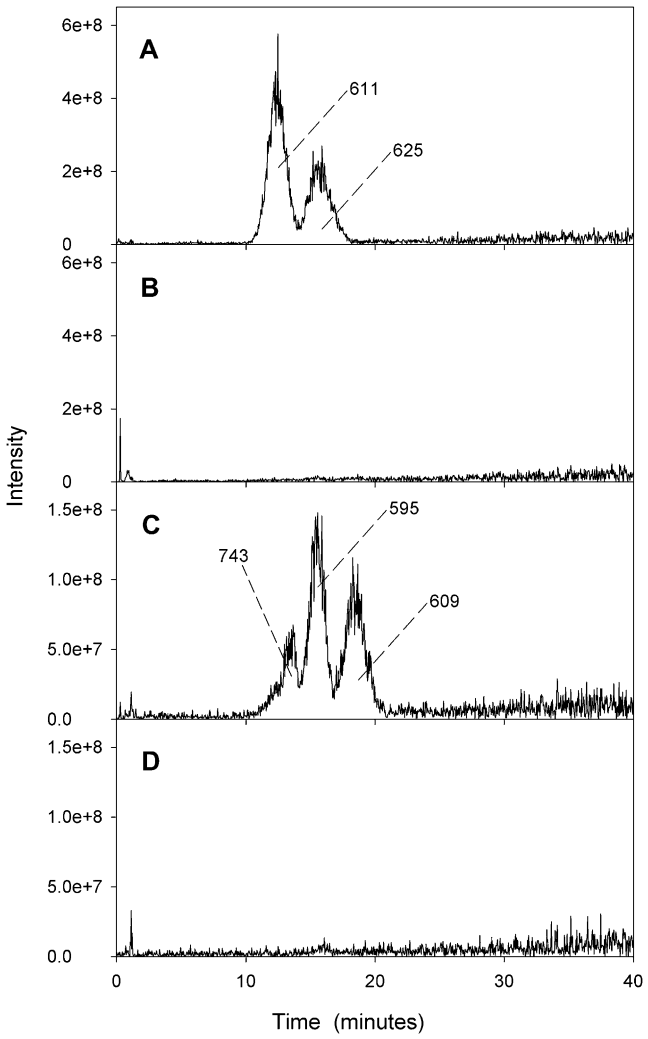


Figure 2. LC-MS analysis of anthocyanins present in the wild type and *b* mutant lines. A, Extracted ion chromatograms showing the summed intensities of ions with masses corresponding to delphinidin and petunidin, each glycosylated with Rha and Glc, present in JI 2822. These masses are *m/z* = 611 (delphinin) and *m/z* = 625 (petunin). B, Masses corresponding to delphinin and petunin absent from FN 2271/3/pink. A and B are plotted to the same scale. C, Extracted ion chromatograms showing the summed intensities of three alternative anthocyanin ions, with masses based on glycosylated cyanidin (*m/z* = 743, *m/z* = 595) and peonidin (*m/z* = 609), present in line FN 2271/3/pink. D, Masses corresponding to cyanin and peonin absent from JI 2822. C and D are plotted to the same scale. Chromatographic peaks are annotated with *m/z* of the mass responsible for the peak.

A 3,231-bp genomic DNA sequence was obtained from PCR products amplified from JI 2822 DNA using primers spanning the cDNA sequence and adaptor-ligation PCR products corresponding to the promoter and 3′ untranslated region (GenBank accession no. GU596479). The position of a single 530-bp intron, 915 bp downstream of the ATG start codon, was determined by alignment of the genomic DNA and cDNA sequences. A single intron is predicted in *Medicago* CU651565_9 at the same position, but in other legumes, such as soybean (Zabala and Vodkin, 2007) and lotus (LjT34E09.40), two introns are reported or annotated. In both these species, the position of the predicted second intron is coincident with the position of the pea intron. The first introns are predicted in different positions, 331 and 348 bp downstream of their ATG, for lotus and soybean, respectively.

Genetic Mapping of *F3′5′H* Reveals Cosegregation with *b*

A cleaved-amplified polymorphic sequence (CAPS) marker for *F3′5′H* that distinguished the JI 15 and JI 73 alleles was generated by *Taq*I cleavage of the PCR products amplified from genomic DNA. Cosegregation of the CAPS marker with *b* was tested directly in a JI 15 × JI 73 recombinant inbred population of 169 individuals, because JI 73 carries the recessive *b* allele. JI 73 also carries *k*, the homeotic conversion of wing petals to keel petals, and *d*, the absence of pigmentation in foliage axils, whereas JI 15 carries *ce*, an independent crimson-pink flower trait. The *b*, *ce* double mutant is almost white, so single and double mutants can be distinguished easily, except in a *k* mutant background, where only the pale standard petal gives a clue to flower color. The genotypes

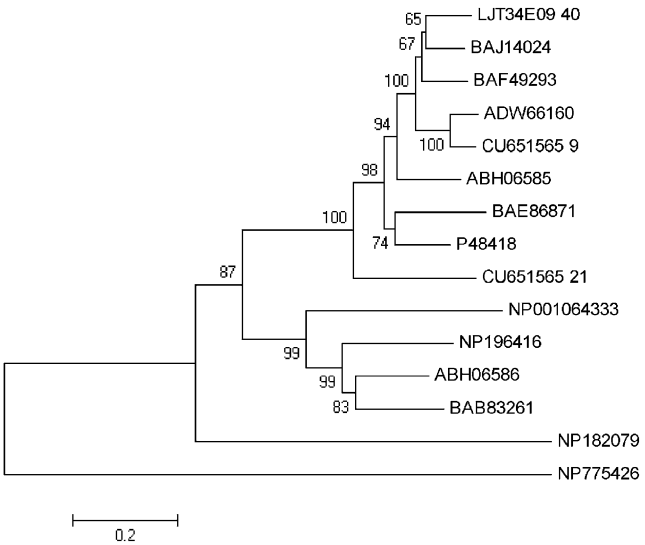


Figure 3. Phylogenetic analysis of cytochrome P450 sequences. The optimal neighbor-joining tree derived from the multiple sequence alignment in Supplemental Figure S2 is drawn to scale, with the sum of branch lengths = 4.7. The Jones-Taylor-Thornton amino acid substitution model was used in phylogeny construction, and the scale bar indicates the number of amino acid substitutions per site. Percentage support for 1,000 bootstrap replicates is shown at the branch points. Labeled lines show GenBank accession numbers as follows: LjT34E09_40, *L. japonicus*; BAJ14024, soybean; BAF49293, *C. ternatea*; ADW66160, *P. sativum*; CU651565_9, *M. truncatula*; ABH06585, *Vitis vinifera*; BAE86871, *G. scabra*; P48418, *Petunia* × *hybrida*; CU651565_21, *M. truncatula*; NP_001064333, *Oryza sativa*; NP196416, Arabidopsis; ABH06586, *V. vinifera*; BAB83261, soybean; NP182079, Arabidopsis; NP775426, *Rattus norvegicus*.

at *b* and *ce* are particularly difficult to distinguish in a *k, d* background, where axillary pigmentation is also absent. For these reasons, the cosegregation analysis was restricted to a subset of 160 of the 169 recombinant inbred lines. The *b* phenotype cosegregated exactly with the JI 73 *F3′5′H* CAPS marker lacking a *Taq*I restriction enzyme site (*b*:*B* = 71:89; χ^2 = 2.0, not significant), consistent with our hypothesis that this *F3′5′H* identifies a single gene that corresponds to *B*.

Identification of Lesions in *F3′5′H* Alleles from Pink-Flowered *b* Mutants

In order to provide further evidence of a correspondence between the pea gene encoding F3′5′H and *B*, we sequenced alleles from known mutants. The *b* mutant type line, JI 118, carries a single nucleotide polymorphism 332 bp downstream of the ATG. This G/A transition would result in a single amino acid change, G111E (Supplemental Figs. S2 and S3). Line JI 73, the *b* mapping parent used above, carries a 23-bp deletion in the ORF, 291 bp from the ATG start. This deletion would introduce a change in the reading frame at position 98, resulting in the inclusion of 29 residues unrelated to the wild type followed by a premature stop codon (Supplemental Fig. S3). PCR analysis using primers that spanned the *F3′5′H* gene showed that lines FN 2160/1, FN 2255/1, and FN 2438/2 as well as the stable pink line FN 2271/3/pink all carry complete gene deletions (Supplemental Fig. S4). FN 1076/6 contains a genomic rearrangement that is consistent with a reciprocal break and join between the *F3′5′H* gene and a predicted *Ogre* retroelement (Neumann et al., 2003). The 5′ segment of the *Ogre* element lies 1,330 bp downstream of the *F3′5′H* start codon, whereas the 3′ segment lies upstream of position 1,330 at the 3′ end of the *F3′5′H* gene (Supplemental Fig. S4).

Characterization of an Unstable Pink-Sectored *b* Mutant

Unstable *b* mutants occurred in the M3 families FN 2271/3/flecked (Fig. 1C) and FN 3398/2164. It was found that sectored pink M3 siblings gave rise to sectored or stable pink M4 progeny, whereas stable pink M3 plants gave rise to stable pink M4 progeny only. Wild-type purple M3 siblings gave rise to either stable wild type, or a mix of stable wild type and stable pink, or a mix of stable wild type, stable pink, and sectored pink M4 progeny. Sectored pink M4 progeny gave rise to sectored or stable pink M5 plants in the following generation. In order to study this instability further, PCR analysis was carried out on individual flowers and progeny plants of line FN 2271/3/flecked/8.

Primers 3′pinkS1 and 3′pinkS2comp amplified 693 bp of genomic DNA and reported on exon 1 and the intron of the *F3′5′H* gene. Primers 3′pinkS2 and 3′extR amplified 683 bp of genomic DNA or cDNA and reported on exon 2. Both pairs of primers were used in conjunction with control primers designed to a pea *Argonaute* gene, which verified that PCR amplification had occurred, even in the absence of a *F3′5′H* PCR product. Genomic DNA and cDNA were prepared from the purple petals of a JI 2822 wild-type flower and from the petals of an entirely pink flower on a FN 2271/3/flecked/8 plant that carried purple/pink-sectored flowers at other nodes. PCR using primers 3′pinkS2 and 3′extR showed the presence of the *F3′5′H* gene in JI 2822 and pink flower FN 2271/3/flecked/8 genomic DNA samples; however, cDNA amplification occurred in line JI 2822 only, suggesting that the *F3′5′H* gene was present but not expressed in the entirely pink FN 2271/3/flecked/8 flower (Fig. 4). Stable pink-flowered M4 progeny were grown from seed set on that entirely pink FN 2271/3/flecked/8 flower. When these were analyzed by PCR, exon 1 and exon 2 of *F3′5′H* failed to amplify from genomic DNA, suggesting that the gene was deleted in these progeny, as was observed previously in the stable pink-flowered line FN 2271/3/pink.

DISCUSSION

The early part of anthocyanin biosynthesis from chalcone to anthocyanidin is well conserved in higher plants and has been studied in detail (Grotewold, 2006). One of the key enzymes responsible for blue-purple coloration in flower petals is F3′5′H, which catalyzes hydroxylation at the 3′ and 5′ positions of the B ring of naringenin and dihydrokaempferol, yielding flavanone and dihydroflavonol precursors of the chromophore delphinidin (Grotewold, 2006; Yoshida et al., 2009). Flowers that lack this enzyme,

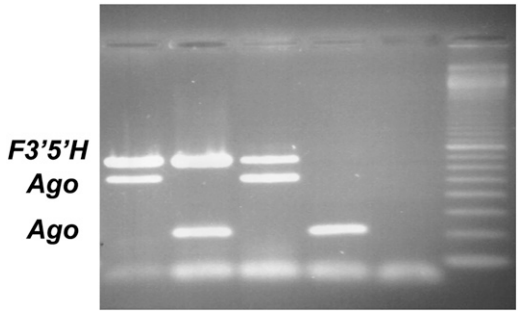


Figure 4. *F3′5′H* gene expression in unstable *b* mutant line FN 2271/3/flecked. PCR amplification of *F3′5′H* and *Argonaute* (*Ago*) genes from JI 2822 genomic DNA (lane 1), JI 2822 cDNA (lane 2), FN 2271/3/flecked/8 genomic DNA (lane 3), FN 2271/3/flecked/8 cDNA (lane 4), and no-DNA control (lane 5) is shown. Lane 6 shows 100-bp markers. The top *Ago* band represents PCR amplification products spanning two introns from genomic DNA, and the bottom *Ago* band represents PCR amplification products (without introns) from cDNA. The *F3′5′H* primers do not flank an intron; therefore, *F3′5′H* PCR products from genomic DNA and cDNA are the same size.

such as rose (*Rosa hybrid*) and carnation (*Dianthus caryophyllus*), contain only cyanidin and/or pelargonidin chromophores, so their natural coloration is restricted to yellow, pink, and red but not purple or blue. Flower color also can be affected by pH, the presence of copigments, and whether the anthocyanidin chromophores are polyacetylated or held in metal complexes (Yoshida et al., 2009). For example, hydrangea (*Hydrangea macrophylla*) sepals can be red, mauve, purple, violet, or blue, yet only one anthocyanin, delphinidin 3-glucoside, is present. It has been proposed that the anthocyanin and copigments in hydrangea sepals are held in a metal complex and that color depends on the concentrations of these components and the pH conditions (Kondo et al., 2005). In wild-type pea, the *F3′5′H* gene is intact and F3′5′H activity produces delphinidin-based anthocyanidins, which confer a purple flower color. In this paper, we have presented genetic and biochemical evidence to show that *b* mutants lack a functional *F3′5′H* gene that results in a rose-pink flower color due to the presence of cyanidin- and peonidin-based anthocyanins. The presence of these latter 3′-hydroxylated compounds in *b* mutants suggests that a F3′H exists in pea, contrary to previous conclusions (Statham et al., 1972).

Lesions Present in *F3′5′H* Alleles

Plant P450 monooxygenases have not been characterized structurally because they are extremely insoluble when purified; however, membrane-associated mammalian P450s have been studied by homology to the crystal structure of a soluble bacterial P450 (Ferrer et al., 2008). P450s have only three absolutely conserved residues: a Cys that serves as a ligand to the heme iron, and an EXXR motif that is thought to stabilize the core around the heme (Werck-Reichhart and Feyereisen, 2000). The Cys lies within the P450 consensus sequence FXXGXRXCXG in the heme-binding loop, corresponding to FGAGRRICAG in the pea F3′5′H (Supplemental Fig. S2). Another consensus sequence, A/GGXD/ETT/S, corresponds to a proton-transfer groove, and this corresponds to AGTDTS in the pea F3′5′H (Supplemental Fig. S2). The G111E mutation in the *b* type line, JI 118, does not occur in these conserved motifs, but the change in size and charge at this residue presumably affects protein function. Alignment of the pea F3′5′H sequence with homologous plant proteins (National Center for Biotechnology Information BLASTP) shows that substitutions occur at the G111 residue; however, none of the substitutes are charged residues, supporting our proposal that G111E is a detrimental change.

Line JI 73 carries a *b* allele with a spontaneous 26-bp deletion that is predicted to encode a truncated version of the F3′5′H protein. At the 3′ end of the 26-bp deleted sequence, there is a 10-bp motif, ATTTCTCAAA, that is repeated at the 5′ end of the deletion break point (Supplemental Fig. S3). This repeat pattern suggests that this stable *b* allele may have arisen from a spontaneous deletion event as a result of recombination and unequal crossing over. The same 26-bp deletion was found in lines JI 17, JI 132, and JI 2160 in the John Innes *Pisum* germplasm collection.

A genomic rearrangement consistent with a translocation event involving a retroelement was evident in line FN 1076/6. Here, sequencing showed that a break occurred in the *F3′5′H* gene, between nucleotides 1,329 and 1,330 downstream of the ATG, but we do not know whether the two fragmented portions of the *F3′5′H* gene remain on the same chromosome (Supplemental Fig. S4). The 5′ end of the genic disjunction was 95% identical to nucleotides 77,728 to 78,111 of a Ty3-gypsy *Ogre*-like retroelement (Neumann et al., 2003) identified in pea BAC clone JICPSV-297I9, whereas the sequence at the 3′ end of the disjunction was 95% identical to nucleotides 77,213 to 77,726 of the same retroelement. This indicates that a break occurred in the *Ogre* element between nucleotides 77,726 and 77,728 and that nucleotide 77,727 was missing from this copy of *Ogre* or was lost during the rearrangement. The presence of this retroelement does not necessarily implicate it in the mechanism of translocation but more likely reflects the abundance of the *Ogre* retroelement family. Data from 454 sequencing of cv Carerra estimated that copies of *Ogre* represent up to 33% of the pea genome (Macas and Neumann, 2007).

We gathered evidence of independent, recurring, spontaneous deletion events derived from unstable *b* alleles carried by lines FN 2271/3/flecked and FN 3398/2164. These sectored flowers carried an *F3′5′H* gene, presumably in nonepidermal tissue where it is not expressed, but repeatedly gave rise to stable pink deletion alleles in their progeny (Fig. 4). One possible explanation of these unstable *b* alleles is that FN 2271 M1 seed carried both a deletion of the *b* gene and a rearrangement of the chromosome carrying the wild-type *B* allele. This rearranged chromosome would be prone to the generation of acentric fragments that would fail to segregate properly at mitosis, generating sectors with a haploinsufficiency for many loci, including *b*. Individuals with the unstable phenotype would give rise to pink homozygous deletion progeny (with a wild-type karyotype). They would also generate progeny that are homozygous or heterozygous for the unstable chromosome, but the transmission of this unstable chromosome may be inefficient, or those that are transmitted efficiently may be selected for stability. In this scheme, the pink-flowered FN 2271 mutants derive from a simple deletion segregating in the population and the instability is not specifically associated with the *b* locus.

Alternatively, the unstable alleles at the *b* locus in the FN 2271 lineage may be prone to deletion, perhaps because of the action of a nearby transposon activated in the FN mutagenesis. Deletion of the *b* gene at one allele would be masked by the presence of the other, wild-type *B* allele, but the presence of such a deletion would reveal subsequent deletions of the *B* allele, which would be seen as pink sectors. In this scheme, deletion of *b* is not

generated directly by mutagenesis and the instability is associated specifically with the *b* gene. Pink flowers of this type could be indicators of a captured insertion element, but in no case did we find a stable pink mutant with the *F3′5′H* gene detectably present, even when these derived from seed set from an entirely pink flower on an unstable plant where the gene, but not the transcript, had been detected by PCR.

F3′5′H Homologs in Legumes

Cytochrome P450s are one of the largest enzyme families in plants. A search of annotated *Medicago* pseudomolecules (http://www.medicagohapmap.org) reveals 142 F3′5′H homologs (BLASTP, $P > 1e\text{-}40$), with approximately one-third of these located on chromosome 5. Gene clusters are found in many other organisms, and in *Medicago*, BACs containing five or more homologous ORFs occurred on chromosomes 2 (AC130800), 3 (AC145061), 5 (FP102223 and AC137079), and 6 (AC157489), although some of these may be pseudogenes. The soybean genome contains 712 cytochrome P450s, of which 380 are denoted pseudogenes (Nelson, 2009). *Medicago* BAC CU651565 carrying CU651565_9, the most similar intact ORF to pea F3′5′H, is unanchored in version 3.5 of the *Medicago* genome pseudomolecules; therefore, we were unable to gain any further evidence of orthology by analyzing collinearity with *b* gene-flanking markers. In the previous version of annotated *Medicago* pseudomolecules (version 3.0), BAC CU651565 was located on chromosome 3, which is syntenic with pea linkage group III, where *b* maps.

Another predicted *Medicago F3′5′H* gene, CU651565_21 (Fig. 3), lies only 52 kb from CU651565_9. The coding sequence of CU651565_21 corresponds to a protein 522 amino acids in length, which is anomalous compared with the lengths of related F3′5′H sequences (Supplemental Fig. S2). Multiple sequence alignment (Supplemental Fig. S2) suggests that CU651565_21 may in fact correspond to a 506-amino acid protein that would be 63% identical to CU651565_9 and 62% identical to the pea F3′5′H. An alternative intron-splicing model derived from ORFs annotated in *Medicago* pseudomolecule version 3.0 is presented (Supplemental Fig. S5).

It is not clear whether the closest related lotus and soybean sequences are orthologous to the pea *F3′5′H*, because they have two introns; therefore, they are structurally dissimilar to the pea and *Medicago* genes. The *Petunia* × *hybrida F3′5′H* also has two introns, whereas the *G. scabra F3′5′H* has one, indicating that intron number is a variable feature of these genes. Diversity of exon-intron structure has been noted among genes encoding P450 enzymes, with multiple gains and losses in their evolutionary history (Werck-Reichhart and Feyereisen, 2000).

The amino acid sequence of CU651565_9, 89% identical to pea F3′5′H, is the closest match; however, the yellow (rather than purple/blue) pigmented flowers of *M. truncatula* suggest that there are differences in anthocyanin biosynthesis between these two species. All of the conserved P450 motifs are intact in CU651565_9, but a comparison with homologous sequences from other plant species shows differences that may be significant. For example, residue Phe-350, which is Leu or Val in aligned homologs (Supplemental Fig. S2), may disrupt F3′5′H function in *M. truncatula*. In support of this possibility, overexpression of the Myb transcription factor LAP1 in *M. truncatula* induced anthocyanin pigments, which were identified as glycosylated cyanidins and pelargonidins but not delphinidins (Peel et al., 2009). The absence of glycosylated delphinidins in these transgenic plants suggests a defect in F3′5′H activity, especially because glycosylated delphinidins were observed in white clover (*Trifolium repens*) overexpressing LAP1 (Peel et al., 2009).

Three soybean sequences (AAM51564, ABQ96218, and BAJ14024) are all 78% identical to pea F3′5′H; however, they are themselves nonidentical. ABQ96218 (Zabala and Vodkin, 2007) and AAM51564 (from cv Chin-Ren-Woo-Dou) are 99% identical and 509 and 508 amino acids long, respectively. They encode a CYP2 subfamily cytochrome P450, also classified as a CYP75A17 cytochrome P450 (Nelson, 2009), at locus Glyma13g04210 on linkage group F of soybean (http://soybase.org). ABQ96218, originating from cv Lee 68 and cloned from the Williams isoline L79-908, carries a G305D amino acid substitution (Zabala and Vodkin, 2007) in the conserved P450 proton-transfer groove motif that would likely render this allele nonfunctional (Supplemental Fig. S2). BAJ14024 (Takahashi et al., 2010) is a predicted F3′5′H from soybean cv Clark, 509 amino acids long, with invariant conserved motifs and 99% identical to both ABQ96218 and AAM51564.

Flower Pigmentation in Pea and Soybean

Soybean is believed to have been domesticated from purple-flowered *G. soja* (Takahashi et al., 2010). Studies of the standard (banner) petals of purple-flowered soybean cultivars show that these have a different sugar moiety at the 3 position of the C ring of their anthocyanidins compared with pea: the primary anthocyanins detected in soybean cv Clark (*W1W1 w3w3 W4W4 WmWm TT TdTd*) and cv Harosoy (*W1W1 w3w3 W4W4 WmWm tt TdTd*) were malvidin, delphinidin, and petunidin 3,5-di-*O*-glucoside and delphinidin 3-*O*-glucoside (Iwashina et al., 2008), whereas delphinidin and petunidin-3-rhamnoside-5-glucoside were the major anthocyanins found in the wing petals of pea line JI 2822 in this study, consistent with previous studies on line L 60 of pea (Statham et al., 1972). As the intensity of coloration in pea petals indicates (Fig. 1), the concentration of total anthocyanins in standard petals is less than in wing petals of pea at all stages of flower development (Statham and Crowden, 1974), whereas soybean flowers often have wing petals

that are less intensely pigmented than their standard petals.

The *Wp* gene of soybean lies on linkage group D1b, corresponding to chromosome 2 (http://soybase.org). The *wp* allele is reported to contain a 5,722-bp CACTA transposable element in intron 2 of a *F3H* gene, *F3H1*, with down-regulated expression (Zabala and Vodkin, 2005). A null mutation would result in a lack of the substrates dihydromyricetin, dihydrokaempferol, and dihydroquercetin required for conversion into anthocyanins (Grotewold, 2006; Iwashina et al., 2008); therefore, a null mutant would be expected to have white flowers and, indeed, white-flowered mutants have been observed in other plant species (Martin et al., 1991; Britsch et al., 1992). Analysis of a *wp* genotype obtained by back-crossing to soybean cv Loda showed that the *wp* line had a low flavonoid content: 9% of the total flavonol glycosides, no detectable kaempferol 3-*O*-glucoside, and 28% of dihydroflavonols compared with cv Clark (Iwashina et al., 2008). The presence of dihydroflavonols indicates that F3H activity occurs in the *wp* mutant, suggesting that it is not a null allele. Alternatively, if the CACTA element insertion does render *F3H1* null, a second *F3H* gene, *F3H2*, may be functional (Zabala and Vodkin, 2005).

Although the presence of anthocyanins in the *wp* mutant can be explained by the considerations above, the pale pink coloration (instead of pale purple) remains unexplained. Many factors such as copigments and vacuolar pH could influence soybean flower color, but the presence of an additional defective pigmentation gene, such as the ABQ96218 allele of *F3′5′H*, for example, would also cause pink flower color. A comparison of flower color and flavonoid content in available *Wp* and *wp* near-isogenic lines (Iwashina et al., 2008) and cosegregation analysis of *F3H1* and *wp* would help to confirm which structural genes were defective.

The soybean *w1* gene on chromosome 13 confers white flower color; accordingly, no HPLC peaks corresponding to anthocyanins were observed in a Clark-*w1* near-isogenic line (L63-2373, *w1w1*, *w3w3*, *W4W4*, *WmWm*, *TT*, *TdTd*; Iwashina et al., 2007). However, it is not clear why a *w1* encoding a defective *F3′5′H* gene would condition white flower color in soybean, when the pea *b* mutant and other *F3′5′H* mutants derived from purple-flowered wild-type plants (Snowden and Napoli, 1998; Matsubara et al., 2005; Nakatsuka et al., 2006) have pink flowers. Genetic linkage analysis of an F2 population segregating for *w1* showed that 12 white-flowered individuals out of 39 F2 progeny carried an *F3′5′H* allele containing a tandem repeat insertion that would result in premature termination of the protein (Zabala and Vodkin, 2007). This linkage evidence is consistent with *w1* being less than 1.1 centimorgan (Kosambi, 1944; Allard, 1956) from the tandem repeat-containing *F3′5′H* gene but with a high SE: the *F3′5′H* homozygotes in the purple flower class were not shown to be *W1* homozygotes by progeny testing, and the population size is small. Thus, it is not clear that a mutated *F3′5′H* gene conditions white flower color in soybean.

One possibility is that *w1* is a separate nonfunctional pigmentation locus, distinct from, but tightly linked to, the *F3′5′H* gene. This *w1* locus is predicted to be functional in a *G. soja* line carrying the *w1-lp* allele, which has pale pink banner petals (Takahashi et al., 2010), and nonfunctional in Clark-*w1*. A cross between these two lines produced purple-flowered F2 progeny at a frequency of 0.9% (Takahashi et al., 2010), which is consistent with recombination between a distinct *w1* gene and the *F3′5′H* gene. Soybean orthologs of genes encoding components of the Myb-bHLH-WD40 transcription factor complex that regulates anthocyanin biosynthesis (Koes et al., 2005; Ramsay and Glover, 2005), such as *a* and *a2* (Hellens et al., 2010), have not yet been identified. These are good candidates for the proposed *F3′5′H*-adjacent *w1* gene.

Pigmentation loci in pea, which have been studied in crosses for more than 100 years (Mendel, 1866; Tschermak, 1911), represent historic anchor markers that will aid comparative genomics between legume species as more physical maps are generated from sequenced genomes. Further biochemical studies, combined with genetic and genomic analyses, will help to elucidate the differences in anthocyanin biosynthesis that lead to variation in pigmentation among legume crop species such as soybean as well as important legume forage species such as alfalfa (*Medicago sativa*) and clover.

MATERIALS AND METHODS

Plant Material

The garden pea (*Pisum sativum*) type line for *b*, JI 118, also known as WBH 22 (Blixt, 1972), multiple marker line JI 73 (genotype *b*, also known as WBH 1238), multiple marker line JI 15 (genotype *B*, also known as WBH 1458), F13 recombinant inbred mapping population JI 15 × JI 73, and all FN mutant lines were obtained from the John Innes *Pisum* Germplasm collection. Plants were grown in 16-h daylength in John Innes No. 1 compost with 30% extra grit. DNA was prepared from leaves according to Vershinin et al. (2003), and RNA was prepared from flowers according to Hofer et al. (2009).

Mutagenesis

A total of 1,400 seeds of line JI 2822 were subjected to 20 Gray FN irradiation from a ^{252}Cf source at Oak Ridge National Laboratory. Irradiated M1 plants were self fertilized, and M2 families of up to four plants were screened for variant flower color phenotypes. Rose-pink mutants were backcrossed to JI 2822 to generate lines FN 1076/6, FN 2160/1, FN 2255/1, FN 2438/2, FN 2271/3/pink, and FN 3398/2164. These stable pink lines segregated purple: pink in a 3:1 ratio after backcrossing, indicating that the pigmentation mutations were recessive.

LC-MS

Purple (JI 2822) and pink (FN 2271/3/pink) wing petal tissue was harvested from 10 fully open flowers, ground in liquid N_2, and stored in methanol at −20°C. Sample aliquots of 10 μL containing 300 μg of tissue in methanol and 0.1 M HCl were analyzed by LC-MS using a Surveyor HPLC apparatus attached to a DecaXPplus ion-trap mass spectrometer (Thermo Fisher). Anthocyanins were separated on a

100- × 2-mm, 3-μm Luna C18(2) column (Phenomenex) using the following gradient of methanol (solvent B) versus 2 mM trifluoroacetic acid in water (solvent A), run at 230 μL min^{-1} and 30°C: 0 min, 2% B; 40 min, 70% B; 41 min, 2% B; 50 min, 2% B. Anthocyanins were detected by UV A_{520} and by positive electrospray ionization MS. Spray chamber conditions were 50 units of sheath gas, 5 units of auxiliary gas, 350°C capillary temperature, and 5.2-kV spray voltage. In order to investigate the structure of anthocyanins, data-dependent secondary fragmentation (MS2) spectra were collected at an isolation width of mass-to-charge ratio (*m/z*) = 4.0 and 35% collision energy.

Isolation of Pea F3′5′H cDNA and Genomic DNA

Total RNA was extracted from JI 2822 wing petals using the Qiagen RNeasy Plant Mini kit. DNA was removed from RNA samples by digestion with DNA-free DNaseI (Ambion) in buffers according to the manufacturer's protocol. Two micrograms of RNA was reverse transcribed with SuperScript reverse transcriptase (Invitrogen) from an oligo(T) primer in a 20-μL reaction. Amplification of a F3′5′H cDNA fragment from pea was achieved using 1 μL of 1:20 diluted first-strand cDNA in 20-μL PCRs containing 0.25 mM primers mtF35HF1 and mtF35HR2 (Supplemental Table S1) for 35 cycles with an annealing temperature of 62°C. Products were separated by electrophoresis on a 1% agarose gel in 1× Tris-borate/EDTA buffer. A 794-bp sequence obtained from this fragment was used to design additional primers for the amplification of 3,231-bp genomic DNA using successive rounds of adaptor ligation PCR (Spertini et al., 1999). The genomic DNA sequence was used to design primers pinkmtF1 and 3′extR for the amplification of a 1,595-bp cDNA clone, minus the ATG start codon and extending 50 bp beyond the TAG stop codon. This was cloned into a Topo4 vector (Invitrogen).

Mutation Analysis

Genomic DNA from JI 2822 and FN mutant lines was analyzed using pairs of primers that spanned the *F3′5′H* gene sequence in order to determine the size of deletion alleles (Supplemental Table S1). Primers PsAGO1 and PsAGO2, flanking introns 19, 20, and 21 of a pea *Argonaute1* cDNA clone (accession no. EF108450), were included in the reactions as internal controls. For the analysis of unstable lines, wing petal cDNA and genomic DNA from JI 2822, plant FN 2271/3/flecked/8, and its progeny were analyzed. Touchdown PCR was performed using 250 nM primers 3′pinkS2 and 3′extR, 250 μM deoxyribonucleotide triphosphates, and 1 unit of Taq polymerase in a 10-μL volume of PCR buffer. Primers PsAGO1 and PsAGO2 were included in the reactions as internal controls. Components were denatured at 95°C for 180 s, before being subjected to one cycle of 94°C for 45 s, 62°C for 45 s, and 72°C for 90 s, followed by 10 further cycles with the annealing temperature 1°C lower at each cycle. Twenty-nine further cycles of 94°C for 45 s, 50°C for 45 s, and 72°C for 90 s were terminated at 72°C for 300 s. Reactions were held at 10°C for 300 s prior to analysis by agarose gel electrophoresis (Supplemental Fig. S4; Supplemental Table S1).

Genetic Mapping

A CAPS marker for F3′5′H was generated by *Taq*I cleavage of the 363- and 340-bp PCR products amplified from 100 ng of genomic DNA from parental lines JI 15 and JI 73, respectively, using primers pinkmtF1 and psf35hF2comp. Reactions contained 250 nM primers, 250 μM deoxyribonucleotide triphosphates, and 1 unit of Taq polymerase in a 20-μL volume of PCR buffer. Components were denatured at 94°C for 120 s, cycled through 94°C for 30 s, 55°C for 60 s, and 72°C for 120 s for 35 cycles, and finally incubated at 72°C for 5 min. Cleavage products of 293 bp from line JI 15 and 340 bp from line JI 73 were separated on a 2% agarose gel. Cosegregation of *b* with the 340-bp *F3′5′H* CAPS marker was tested for 160 lines out of 169 in total at the F13 generation of a recombinant inbred population derived from the cross JI 15 × JI 73. A total of 71 lines were *b/b* and carried the 340-bp marker, and 89 individuals were *B* and carried the 293-bp marker.

Sequencing

Sequencing was performed using the BigDye Terminator version 3.1 cycle sequencing kit (Applied Biosystems) at the John Innes Centre Genome Laboratory. Genomic DNA sequence was obtained from line JI 2822 using the primers listed in Supplemental Table S1. A 3,232-bp overlapping DNA sequence contig was generated using the program BioEdit (http://www.mbio.ncsu.edu/bioedit/bioedit.html). Overlapping DNA sequence contigs from *b* mutant lines JI 118, JI 73, and FN 1076/6 and cDNA sequences from lines JI 2822, JI 118, JI 73, and FN 1076/6 were obtained in the same way.

Sequence data from this article can be found in the GenBank/EMBL data libraries under the following accession numbers: JI 2822 F3′5′H cDNA sequence, GU596478; JI 2822 F3′5′H genomic DNA sequence, GU596479.

Supplemental Data

The following materials are available in the online version of this article.

Supplemental Figure S1. Ion fragmentation analysis of anthocyanins present in the wild type and *b* mutant lines.

Supplemental Figure S2. F3′5′H sequence analysis.

Supplemental Figure S3. Sequence characterization of mutant *b* alleles.

Supplemental Figure S4. Characterization of mutant *b* alleles by PCR.

Supplemental Figure S5. Proposed splicing model for *Medicago* gene CU651565_21.

Supplemental Table S1. Primers used for PCR and sequencing.

ACKNOWLEDGMENTS

We thank Andrew Davis for photography, Ruth Pothecary and Hilary Ford for plant care, and a Nuffield scholarship student, Priyanka Tharian, for technical assistance.

Received March 22, 2012; accepted April 3, 2012; published April 6, 2012.

LITERATURE CITED

Allard RW (1956) Formulas and tables to facilitate the calculation of recombination values in heredity. Hilgardia **24:** 235–278

Blixt S (1972) Mutation genetics in *Pisum*. Agri Hort Genet **31:** 1–293

Britsch L, Ruhnau-Brich B, Forkmann G (1992) Molecular cloning, sequence analysis, and *in vitro* expression of flavanone 3 beta-hydroxylase from *Petunia hybrida*. J Biol Chem **267:** 5380–5387

De Haan H (1930) Contributions to the genetics of *Pisum*. Genetica **12:** 321–439

Ferrer JL, Austin MB, Stewart C Jr, Noel JP (2008) Structure and function of enzymes involved in the biosynthesis of phenylpropanoids. Plant Physiol Biochem **46:** 356–370

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

Harker CL, Ellis THN, Coen ES (1990) Identification and genetic regulation of the chalcone synthase multigene family in pea. Plant Cell **2:** 185–194

Hellens RP, Moreau C, Lin-Wang K, Schwinn KE, Thomson SJ, Fiers MWEJ, Frew TJ, Murray SR, Hofer JMI, Jacobs JME, et al (2010) Identification of Mendel's white flower character. PLoS ONE **5:** e13230

Hofer J, Turner L, Moreau C, Ambrose M, Isaac P, Butcher S, Weller J, Dupin A, Dalmais M, Le Signor C, et al (2009) *Tendril-less* regulates tendril formation in pea leaves. Plant Cell **21:** 420–428

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

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

Knight T (1799) An account of some experiments on the fecundation of vegetables. Philos Trans R Soc Lond **89:** 195–204

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

Kondo T, Yuki TK, Yoshida K (2005) Essential structure of co-pigment for blue sepal-color development of hydrangea. Tetrahedron Lett **46:** 6645–6649

Kosambi DD (1944) The estimation of map distances from recombination values. Ann Eugen **12:** 172–175

Macas J, Neumann P (2007) Ogre elements: a distinct group of plant Ty3/gypsy-like retrotransposons. Gene **390:** 108–116

Macas J, Neumann P, Navrátilová A (2007) Repetitive DNA in the pea (*Pisum sativum* L.) genome: comprehensive characterization using 454 sequencing and comparison to soybean and *Medicago truncatula*. BMC Genomics **8:** 427

Martin C, Prescott A, Mackay S, Bartlett J, Vrijlandt E (1991) Control of anthocyanin biosynthesis in flowers of *Antirrhinum majus*. Plant J **1:** 37–49

Marx GA, Weeden NF, Muehlbauer FJ (1989) A new locus controlling anthocyanin production in *Pisum*. Pisum Newsletter **21:** 35–36

Matsubara K, Kodama H, Kokubun H, Watanabe H, Ando T (2005) Two novel transposable elements in a cytochrome P450 gene govern anthocyanin biosynthesis of commercial petunias. Gene **358:** 121–126

Mendel G (1866) Versuche über Pflanzen-Hybriden. Verhand Naturforsch Vereines Abhandlungen Brunn **4:** 3–47

Nakatsuka T, Nishihara M, Mishiba K, Hirano H, Yamamura S (2006) Two different transposable elements inserted in flavonoid 3′,5′-hydroxylase gene contribute to pink flower coloration in *Gentiana scabra*. Mol Genet Genomics **275:** 231–241

Nelson DR (2009) The cytochrome p450 homepage. Hum Genomics **4:** 59–65

Neumann P, Pozárková D, Macas J (2003) Highly abundant pea LTR retrotransposon Ogre is constitutively transcribed and partially spliced. Plant Mol Biol **53:** 399–410

Peel GJ, Pang Y, Modolo LV, Dixon RA (2009) The LAP1 MYB transcription factor orchestrates anthocyanidin biosynthesis and glycosylation in *Medicago*. Plant J **59:** 136–149

Ramsay NA, Glover BJ (2005) MYB-bHLH-WD40 protein complex and the evolution of cellular diversity. Trends Plant Sci **10:** 63–70

Schoenbohm C, Martens S, Eder C, Forkmann G, Weisshaar B (2000) Identification of the *Arabidopsis thaliana* flavonoid 3′-hydroxylase gene and functional expression of the encoded P450 enzyme. Biol Chem **381:** 749–753

Shirsat AH (1988) A transposon-like structure in the 5′ flanking sequence of a legumin gene from *Pisum sativum*. Mol Gen Genet **212:** 129–133

Snowden KC, Napoli CA (1998) *Psl*: a novel *Spm*-like transposable element from *Petunia hybrida*. Plant J **14:** 43–54

Spertini D, Béliveau C, Bellemare G (1999) Screening of transgenic plants by amplification of unknown genomic DNA flanking T-DNA. Biotechniques **27:** 308–314

Statham CM, Crowden RK (1974) Anthocyanin biosynthesis in *Pisum*: sequence studies in pigment production. Phytochemistry **13:** 1835–1840

Statham CM, Crowden RK, Harborne JB (1972) Biochemical genetics of pigmentation in *Pisum sativum*. Phytochemistry **11:** 1083–1088

Stephens PA, Nickell CD (1992) Inheritance of pink flower in soybean. Crop Sci **32:** 1131–1132

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

Tschermak E (1911) Examen de la théorie des facteurs par le recroisement méthodique des hybrides. Compt Rend Conf Int Genet Paris **IV:** 91–95

Uimari A, Strommer J (1998) Anthocyanin regulatory mutations in pea: effects on gene expression and complementation by *R*-like genes of maize. Mol Gen Genet **257:** 198–204

Vershinin AV, Allnutt TR, Knox MR, Ambrose MJ, Ellis THN (2003) Transposable elements reveal the impact of introgression, rather than transposition, in *Pisum* diversity, evolution, and domestication. Mol Biol Evol **20:** 2067–2075

Werck-Reichhart D, Feyereisen R (2000). Cytochromes P450: a success story. Genome Biol **1:** 3003.1–3003.9.

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 (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 (2007) 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:** S113–S124