



## Article

# The Analysis of Partial Sequences of the Flavonone 3 Hydroxylase Gene in *Lupinus mutabilis* Reveals Differential Expression of Two Paralogues Potentially Related to Seed Coat Colour

Norberto Guilengue <sup>1,2</sup>, Helena G. Azinheira <sup>1,3</sup>, Sofia Alves <sup>1,3</sup> , João Neves-Martins <sup>1</sup> , Pedro Talhinhos <sup>1,3,\*</sup> and Leonor Morais-Cecílio <sup>1,3,\*</sup>

- <sup>1</sup> Instituto Superior de Agronomia, Universidade de Lisboa, 1349-017 Lisbon, Portugal; guilenguen@gmail.com (N.G.); hmga@isa.ulisboa.pt (H.G.A.); salves@isa.ulisboa.pt (S.A.); nevesmartins@isa.ulisboa.pt (J.N.-M.)
- <sup>2</sup> Agricultural Engineering Course, Agricultural Faculty, Instituto Superior Politécnico de Gaza, Londe, Chokwe 1204, Mozambique
- <sup>3</sup> LEAF-Linking Landscape, Environment, Agriculture and Food, Associated Laboratory TERRA, Instituto Superior de Agronomia, Universidade de Lisboa, 1349-017 Lisbon, Portugal
- \* Correspondence: ptalhinhos@isa.ulisboa.pt (P.T.); lmorais@isa.ulisboa.pt (L.M.-C.)

**Abstract:** Flavonone 3 hydroxylases (EC 1.14.11.9) are key enzymes in the synthesis of anthocyanins and other flavonoids. Such compounds are involved in seed coat colour and stem pigmentation. *Lupinus mutabilis* (tarwi) is a legume crop domesticated in the Andean region, valued for the high protein and oil content of its seeds. Tarwi accessions are being selected for cultivation in Europe under defined breeding criteria. Seed coat colour patterns are relevant breeding traits in tarwi, and these are conditioned by anthocyanin content. We identified and isolated part of the tarwi flavonone 3-hydroxylase gene (*LmF3h*) from two accessions with distinct seed coat colour patterns. Two partial *LmF3h* paralogues, with predicted 20% amino-acid changes but little predicted tertiary structure alterations, were identified in the coloured seed genotype, while only one was present in the white seed genotype. Upon selection and validation of appropriate reference genes, a RT-qPCR analysis showed that these paralogues have different levels of expression during seed development in both genotypes, although they follow the same expression patterns. DNA and transcription analyses enabled to highlight potential *F3H* paralogues relatable to seed coat pigmentation in tarwi and, upon biochemical and genetic confirmation, prompt marker-assisted breeding for relevant phenotypic traits associated with flavonoid synthesis.

**Keywords:** *Lupinus mutabilis*; *F3H*; seed coat; gene expression; RT-qPCR reference genes



**Citation:** Guilengue, N.; Azinheira, H.G.; Alves, S.; Neves-Martins, J.; Talhinhos, P.; Morais-Cecílio, L. The Analysis of Partial Sequences of the Flavonone 3 Hydroxylase Gene in *Lupinus mutabilis* Reveals Differential Expression of Two Paralogues Potentially Related to Seed Coat Colour. *Agronomy* **2022**, *12*, 450. <https://doi.org/10.3390/agronomy12020450>

Academic Editors: Jose Maria Barrero and Matthew W. Blair

Received: 10 December 2021

Accepted: 9 February 2022

Published: 11 February 2022

**Publisher's Note:** MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



**Copyright:** © 2022 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

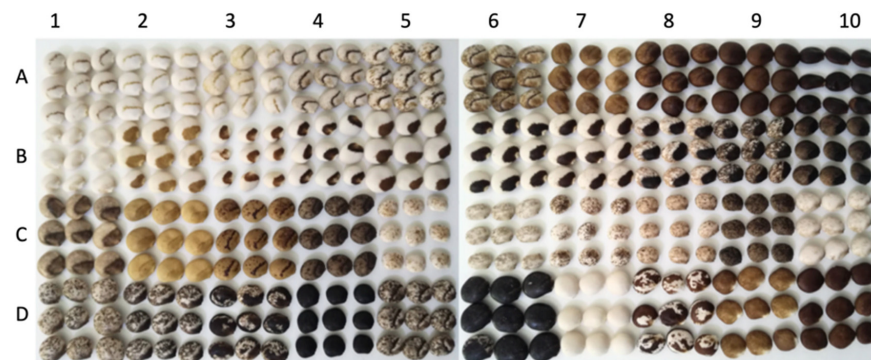
## 1. Introduction

Flavonoids are important secondary metabolites that include chalcones, flavones, flavonols, flavandiols, anthocyanins, and proanthocyanidins or condensed tannins. These metabolites are known for their multiple benefits for humans and plants. For humans, they are associated with positive effects, for instance as antioxidant, antibacterial, anti-cancer, anti-mutagenic, anti-atherosclerosis, and anti-inflammatory agents. In plants, they play important roles in protection against biotic and abiotic stresses, act as fertility regulators, and in auxin biosynthesis. They increase the agronomic and industrial values of products, conferring pigmentation to flowers, fruits, vegetative tissues and attracting pollinators [1–5]. Within the flavonoids, anthocyanins are the most extensively distributed pigments in higher plants. Their biological and genetic processes, as well as gene regulation, have been widely studied in different systems [3]. Anthocyanins are deposited in the tissues via two

biosynthesis pathways, which is well-described in several studies that point to the involvement of structural and regulatory genes [6]. Structural genes encode multiple enzymes: phenylalanine ammonia lyase (PAL); cinnamate 4-hydroxylase (C4H); chalcone synthase (CHS); flavanone 3-hydroxylase (F3H); dihydroflavonol 4-reductase (DFR); anthocyanidin synthase (ANS); and UDP-glucose: flavonoid 3-O-glucosyltransferase (UGT) [7,8]. Regulatory genes encode the transcription factors WIP-type Zn-Finger, AtMYB123, AtbHLH042, MADS AtAGL32, WD40 or WDR, and AtWRKY44 to control the temporal and spatial expression of the structural genes [1,9,10]. F3H (EC 1.14.11.9) is the central enzyme of the flavonoid biosynthetic pathway. This enzyme is associated with several functions, mainly on plant resistance to stresses caused by drought, salt, cold, damage caused by fungi, and ultraviolet-B radiation [11]. In several legume species such as *Glycine max*, *Medicago truncatula*, *M. sativa* and *Phaseolus vulgaris*, studies of characterization, expression, and function of the *F3H* gene have been carried out in different tissues to better understand the role of this gene [12–15], but in species of the genus *Lupinus* the information is still scarce. Most of these works have studied the expression of the *F3H* gene in roots, flowers, and stems and little is known about the role of *F3H* in the colouration of the seeds of legumes. However, in a study carried out in *Arabidopsis thaliana*, *F3H* was reported as being associated with spotted pale brown seed coat colour [1,10]. The seed coat acts as a protection layer after physiologic maturation, and contributes to the maintenance of the physiologic quality, seed longevity, disease resistance, control of seed development, and dormancy [16–18]. The brown pigmentation of the seed coat has been studied in other species such as *Glycine max* and *Phaseolus vulgaris* and has been associated with the presence of proanthocyanidins [19–21]. Beneficial effects have been associated with the consumption of coloured seeds due to its diverse phytochemical contents, such as flavonoids ([22] and references therein). The acceptability of grain seed by consumers is related with seed coat colour evaluation because many consumers are looking for foods with high nutritional value and health benefits [23]. Brown seed coat is a desirable agronomic trait, hence several studies in various species have been performed to elucidate seed coat pigmentation mechanisms.

Lupins (*Lupinus* spp.) are legume crops used for food and feed, appreciated for the high protein content of their seeds, for their adaptability to marginal cultivation conditions and for their high beneficial effect during crop rotation due to high nitrogen input to the soil. In parallel with the three lupin crops domesticated in the Mediterranean basin (white, yellow, and narrow-leaved lupin), tarwi (*L. mutabilis* Sweet) was domesticated in the Andean region and used as a legume crop by the Incas [24]. Tarwi is appreciated for the high protein and oil contents of its seeds [25,26] and is being selected and adapted to cultivation in Europe under diverse environmental conditions [27,28]. Breeding criteria include growth habit, seed characteristics and yield, nitrogen fixation, and disease response. Anthracnose, caused by *Colletotrichum lupini*, is one of the most important and widespread diseases of lupins [29], and tarwi is generally regarded as susceptible [30]. Nevertheless, stem anthocyanin-pigmentation was shown to be related to lesser disease susceptibility [31]. In several legumes, disease resistance has been associated with coloured seeds. Work performed in the 1970s referred to the importance of coloured seeds in resistance against root-rot disease in common beans [32], as well as resistance against fungi and seed vivipary in sorghum [33]. Stasz et al. [34], evaluating the time and site of infection by *Pythium ultimum* of resistant and susceptible germinating pea seeds, found that anthocyanin-coloured pea seeds had higher levels of resistance against this pathogen than non-coloured ones. Islam et al. [35] also came to the conclusion that coloured seeds have a beneficial effect on resistance against common bean diseases. On the other hand, the consumption of coloured seeds of certain legumes has been associated with multiple beneficial health effects. For instance, the consumption of coloured common bean and black soybean seeds has been identified as being fundamental in inhibiting the proliferation of different cancer cells in humans ([36] and references therein). A study of phenolic composition was carried out by Xu and Chang [37] in several legumes and found that black and red varieties of common beans and black soy have higher total phenolic content, total flavonoid content, and condensate

tannin content, that are said to be highly beneficial to health. Soybean seeds with a coloured coat are known to be rich in isoflavanones and proteins, which confer benefits to human health [38,39]. Segev et al. [40] stated that coloured chickpea seeds rich in antioxidants and with high total polyphenol content have great potential to be used as functional foods. Mineral content and bioavailability in common beans were correlated to seed coat colour [41] and human selection of bean genotypes regarding invisible micronutrient traits has been related to seed coat colour, cooking time, and palatability attributes [42]. Tarwi seed coat presents diverse colours and colour patterns that are more variable than in most other grain legumes and certainly more than in other lupin crops (Figure 1). Nevertheless, little is known about mineral content variation among tarwi genotypes [25]. Seed coat colours and colour patterns are perceived in diverse ways by consumers [43,44]. In tarwi, the white colour is reported as the most preferred by consumers [27], indicating little consumption of coloured grains. However, many countries encourage consumption of coloured foods to prevent chronic diseases [45]. Therefore, gaining knowledge about where and when these anthocyanins are expressed in legume seeds is important for the development of specific markers that can be used to establish breeding strategies regarding superior cultivars. Gene expression studies have been carried out in several legume species [46–49]. However, in species of the genus *Lupinus*, little is known about the genes that control seed pigmentation, an even about the expression profile of these genes. Knowing the expression profile of these genes in different seed development stages is essential to control the factors affecting colour expression in seeds and, on the other hand, to produce quality seeds. The main aim of this study is to characterize contrasting genotypes in the colour of seed coat (brown and white) based on partial conserved gene sequence and expression, with the purpose of supporting marker-assisted tarwi breeding regarding phenotypic traits related to anthocyanin pigmentation.



**Figure 1.** Common patterns of *Lupinus mutabilis* (tarwi) seed coat: eye-brow (A1–A10, C3 and D5); crescent (B1–C2 and C4); spotted (C5, C7 and C9); moustache (C6, C8, C10, D1 and D2); marbled (D3 and D8). Variants: A1 (white with brown eye-brow), A2 (white with cream brown eye-brow); A3 (white with brown spotted and cream brown eye-brow); A4–A6 (white with brown spotted and dark brown eye-brow); A7 (light brown with brown eye-brow); A8 and A9 (brown with dark brown eye-brow); A10 (dark brown with dark brown eye-brow); B1 (white with grey crescent); B2 (white with cream brown crescent), B3–B5 (white with brown crescent), B6 (white with black crescent); B7 (white with dark brown crescent); B8 and B9 (white with brown spotted and black crescent); B10 (grey with black eye-brow); C1 (light grey with dark grey crescent); C2 (cream with cream crescent); C3 (cream with spotted and brown eye-brow); C4 (grey with spotted and dark grey crescent); C5 and C7 (white with brown spotted); C9 (white with grey spotted); C6, C10 and D1 (white with grey spotted and grey moustache), D2 (white with grey spotted and black moustache); C8 (white with brown spotted and moustache), C9 (grey spotted); D3 (white with black marbled); D4 (black); D5 (white with grey spotted and black eye-brow); D6 (combination black marbled and black colour); D7 (white); D8 (white with brown marbled); D9 (light brown); D10 (brown) (according to IBPGR [50]).

## 2. Materials and Methods

### 2.1. Plant Material

Plant material used in this study were *L. mutabilis* accessions SBP (white seeds) and LM18 (brown seed coat with dark brown eyebrow), from the *Lupinus* germplasm collection at Instituto Superior de Agronomia, Universidade de Lisboa (Lisbon, Portugal). The seeds used in this experiment came from the 7th generation of plants multiplied under self-pollination conditions (using an insect net over the multiplication plot). LM18 is a promising breeding line, presenting high efficiency in converting vegetative growth to seed production [28] and exhibiting a moderate resistance response to anthracnose by developing anthocyanin pigmentation around infection areas [31]. SBP is a breeding line with homogeneous white seeds and high susceptibility to anthracnose (unpublished data). *Lupinus angustifolius* ('Illyarrie') leaves were also used for DNA analyses.

### 2.2. Gene Analysis

DNA was extracted from leaves of young plants of each of the three genotypes (narrow-leaved lupin and both tarwi accessions) using the DNeasy® Plant mini kit (Qiagen, Hilden, Germany) according to the manufacturer instructions and as previously described [28].

The soybean flavanone 3-hydroxylase gene (GLYMA\_02G048400) was used to interrogate the narrow-leaved lupin genome ([https://plants.ensembl.org/Lupinus\\_angustifolius/Info/Index](https://plants.ensembl.org/Lupinus_angustifolius/Info/Index); accessed 8 April 2020; [51]) using BLASTN and BLASTX tools, and protein sequence XP\_019420372 (*L. angustifolius* naringenin, 2-oxoglutarate 3-dioxygenase-like) was used for subsequent comparisons. PCR primers were drawn based on nucleotide and protein alignments between soybean and narrow-leaved lupin *F3H* genes along with *Vigna angularis* LOC108323664 (naringenin, 2-oxoglutarate 3-dioxygenase) (Figure S1). The alignment of the sequences for the identification of the conserved regions was performed with the ClustalW algorithm. The conserved region in the three species was used for the design of primers.

Primers 5' GATTGGAGAGAGATTGTGACATA 3' (forward) and 5' GTGATCAGCATTCTTGAACC 3' (reverse) were used to amplify *L. mutabilis* DNA. A PCR was performed under the following conditions: pre-denaturation 5 min at 94 °C, 40 cycles of 30 s at 94 °C, 45 s at 49 °C, and 1 min at 72 °C, and a final extension at 72 °C for 10 min. PCR reactions were carried out in a final volume of 10 µL containing 20 ng of DNA, 0.5 µM of each primer and 5 µL of dNTP + Taq DNA polymerase (NZYTech, Lisbon, Portugal). PCR products were visualised by electrophoresis at 2% agarose gels and, when necessary, PCR products were gel-excised using the GeneJET Gel Extraction Kit (Thermo Fisher Scientific, Waltham, MA, USA), and sequenced. The analysis of sequencing chromatograms was performed using the SeqMan module of DNASTar v5.05 (Lasergene, Madison, WI, USA).

Local DNA and polypeptide alignments were conducted with MegaX [52] using the Muscle algorithm. In comparison with the *L. angustifolius* (XP\_019420372), intronic and exonic regions were identified in *L. mutabilis*. The exonic region of *L. mutabilis* was used to deduce the polypeptide sequence and to construct a dendrogram using the Maximum Likelihood method and comparing similarities with *L. angustifolius* orthologs. The evolutionary relationship among sequences was inferred by using the Maximum Likelihood method based on the JTT matrix-based model. A phylogenetic tree was generated from 500 bootstrap datasets to provide statistical support for each node. We used iTOL (<https://itol.embl.de> accessed 8 April 2020) for displaying and annotating the generated phylogenetic trees.

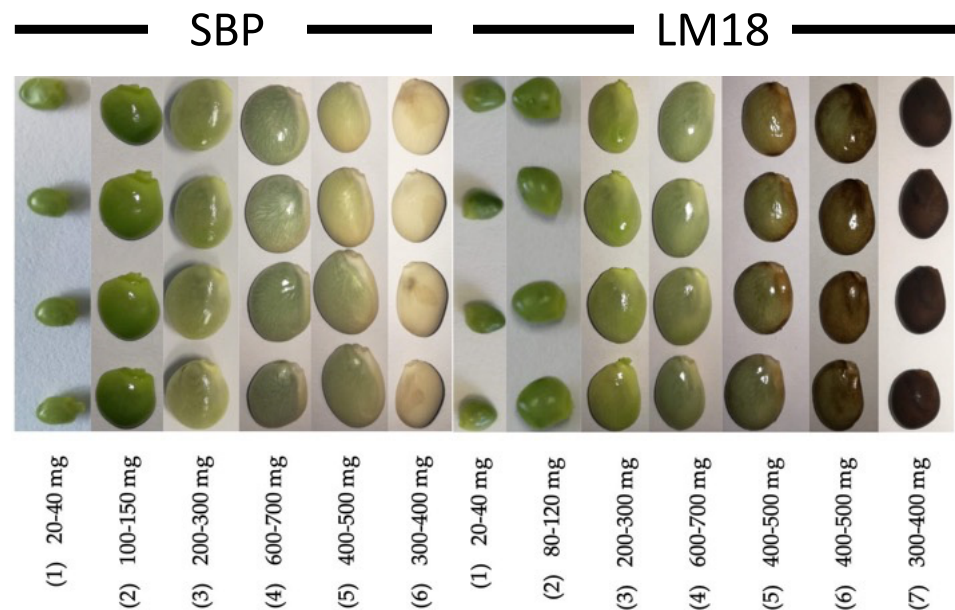
A protein analysis was performed to compare predicted three-dimensional and secondary structures of *L. mutabilis* using Phyre2 [53].

### 2.3. Transcription Analysis

Expression of the putative *F3H* gene along the development of tarwi seed coat in white and coloured genotypes was assessed through RT-qPCR. For this, RNA was extracted from seed coats of SBP and LM18 accessions along seed development (Figure 2). Following



Zabala and Vodkin [54], samples were divided in six or seven groups according to fresh seed weight (expressed in mg/seed) as exemplified in Figure 2.



**Figure 2.** Seed development stages (1–6, with the corresponding fresh weight expressed in mg/seed) of *Lupinus mutabilis* accessions SBP (white seed; left) and LM18 (brown seed with dark-brown eye-brow; right) used for transcriptomic analysis.

Three plants (biological replicates) were used for each stage of seed development and for each genotype. For reasons of material quantity, in the first two stages, the seeds collected in each plant were put together, making a single sample per each plant and stage. In the other developmental stages at least 10 seeds were collected to isolate seed coat. In the first two stages of seed development, in both genotypes, it was not possible to isolate the seed coat because the cotyledons were not formed and the inside of the seeds was watery, thus the whole seed was used. Seed coat was isolated from the third stage of seed development onwards. All seeds were immediately processed after being removed from the pods. In each seed development stage about 300–400 mg was collected and from these samples around 100 mg of material was used for dissection using sterile scalpels, immediately frozen in liquid nitrogen and stored at  $-80^{\circ}\text{C}$ . The total RNA was extracted using the Spectrum™ Plant Total RNA Kit (Sigma-Aldrich, St. Louis, MO, USA) and evaluated as previously described [55]. After evaluation of RNA quality and integrity, cDNA was synthesised from 1  $\mu\text{g}$  total RNA in a final volume of 20  $\mu\text{L}$  using the RevertAid H Minus Reverse transcriptase (Thermo Scientific, Waltham, MA, USA) according to the manufacturer's instructions. The cDNA products were stored at  $-20^{\circ}\text{C}$  for future use. For transcription analysis, three biological and two technical replicates were performed.

### 2.3.1. Reference Genes

Reference genes for RT-qPCR were selected considering the literature available for narrow-leaved lupin [56,57] and validated in tarwi. Five reference genes were tested (Table 1). To ensure the reliability of the potential reference gene, the expression profiles of two fragments were measured and normalised with the most stable reference gene determined using the RefFinder tool [58], which provided a classification based on four methods: Bestkeeper [59]; geNorm [60]; NormFinder [61,62]; and the  $\Delta\text{Ct}$  method [63].

**Table 1.** *Lupinus* genes tested as RT-qPCR reference genes for gene expression analysis of tarwi *F3H* partial gene.

| Reference Gene                              | Primer Name and Direction | Primer Sequence                 | Tm <sup>1</sup> (°C) | Size (bp) | Eff <sup>2</sup> | Eff <sup>2</sup> (%) |
|---|---------------------------|---------------------------------|----------------------|-----------|------------------|----------------------|
| Ubiquitin C ( <i>Ubc</i> )                  | LangUBC-F                 | 5' CTGACAGCCCACTGAATTGTGA 3'    | 60.8                 | 108.0     | 1.952            | 0.930                |
|   | LangUBC-R                 | 5' TCTTGGGCATAGCAGCAAGC 3'      | 61.0                 |           |                  |                      |
| <i>Helicase (Hel)</i>                       | LangHEL-F                 | 5' TTGTACGAGGTCGGTGCTCT 3'      | 60.9                 | 127.0     | 1.947            | 0.927                |
|   | LangHEL-R                 | 5' ACAAGCAACCAAATATTGCACCATA 3' | 60.0                 |           |                  |                      |
| <i>Alcohol dehydrogenase class-3 (Adh3)</i> | LangADH3-F                | 5' AGCACACAGCGTAGGCATC 3'       | 58.0                 | 91.0      | 1.957            | 0.932                |
|   | LangADH3-R                | 5' AGTTGATGAGTACATAACCCACA 3'   | 58.0                 |           |                  |                      |
| <i>ATP synthase (ATPsyn)</i>                | LangATPsyn-F              | 5' AGTATGCTGTTCTGTTCTGTC 3'     | 59.0                 | 145       | –                | –                    |
|   | LangATPsyn-R              | 5' ATGGTGATCTTCTCCTTCTTAG 3'    | 59.0                 |           |                  |                      |
| <i>Alpha tubulin (αTub)</i>                 | LangTUBA-F                | 5' CGGGTTAGAAAGTTGGCGGA 3'      | 58.0                 | 101       | 1.944            | 0.926                |
|   | LangTUBA-R                | 5' CAACAAGAGAGATCCCAAACC 3'     | 58.0                 |           |                  |                      |

<sup>1</sup> Tm: melting temperature; <sup>2</sup> Eff: efficiency.

### 2.3.2. Quantitative Reverse-Transcription PCR

According to the results obtained in the genomics analysis, RT-qPCR primers were designed for tarwi *LmF3h\_a* and *LmF3h\_b* fragments (Table 2). RT-qPCR reactions were carried out using a CFX96 thermal cycler (Bio-Rad, Hercules, CA, USA) with the Bio-Rad CFX Manager software (Bio-Rad). For such, 5 µL of cDNA, 7.5 µL of SsoFast EvaGreen (Bio-Rad) and 0.45 µL of each primer (10 ng/µL) were used for each reaction. The amplification programme was performed at 95 °C for 3 min followed by 40 cycles of 10 s at 95 °C and 30 s at 60 °C. To produce melting curves, an additional denaturing cycle with temperature ranging between 65 °C and 95 °C in 0.5 °C increments was performed for all reactions after amplification. The specificity amplification of products was confirmed by a melting curve analysis and agarose gel electrophoresis. Each set of reactions included a negative control with no template. The efficiencies of RT-qPCR primer pairs were calculated:  $E = (10^{[-1/\text{slope}]} - 1) \times 100$  [64], using LinRegPCR [65]. The mRNA expression of the genes of interest (goi) was quantified using the comparative threshold cycle (Ct) values method [66]. The Ct value of the reference genes was subtracted from the Ct value of the genes of interest to obtain  $\Delta C_t$ . The first stage of seed development for each phenotype was used as the starting point for RT-qPCR data normalisation using reference genes. For such, the geometric mean of the multiple reference genes was calculated using  $2^{-\Delta\Delta C_t}$ , where  $\Delta\Delta C_t$  equals to  $\Delta C_{t\text{sample}} - \Delta C_{t\text{control}}$  [63].

**Table 2.** Primers used for RT-qPCR analysis along seed coat development in white and brown seeds of *Lupinus mutabilis*.

| Gene of Interest | Primer Name and Direction | Primer Sequence                | Tm <sup>1</sup> (°C) | Size (bp) | Eff <sup>2</sup> | Eff (%) |
|------------------|---------------------------|--------------------------------|----------------------|-----------|------------------|---------|
| <i>F3h_a</i>     | LmutF3h_a-F               | 5' GGAAGTATTATCAGAAGCAATGGG 3' | 60.0                 | 120       | 1.988            | 0.947   |
|                  | LmutF3h_a-R               | 5' AAGTGTAAGATCAGGTTGAGGG 3'   | 60.0                 |           |                  |         |
| <i>F3h_b</i>     | LmutF3h_b-F               | 5' TACCCTAAATGTCCACAACCTG 3'   | 60.0                 | 129       | 1.926            | 0.917   |
|                  | LmutF3h_b-R               | 5' TCCAAGTCTTTCCATTATCCCT 3'   | 60.0                 |           |                  |         |

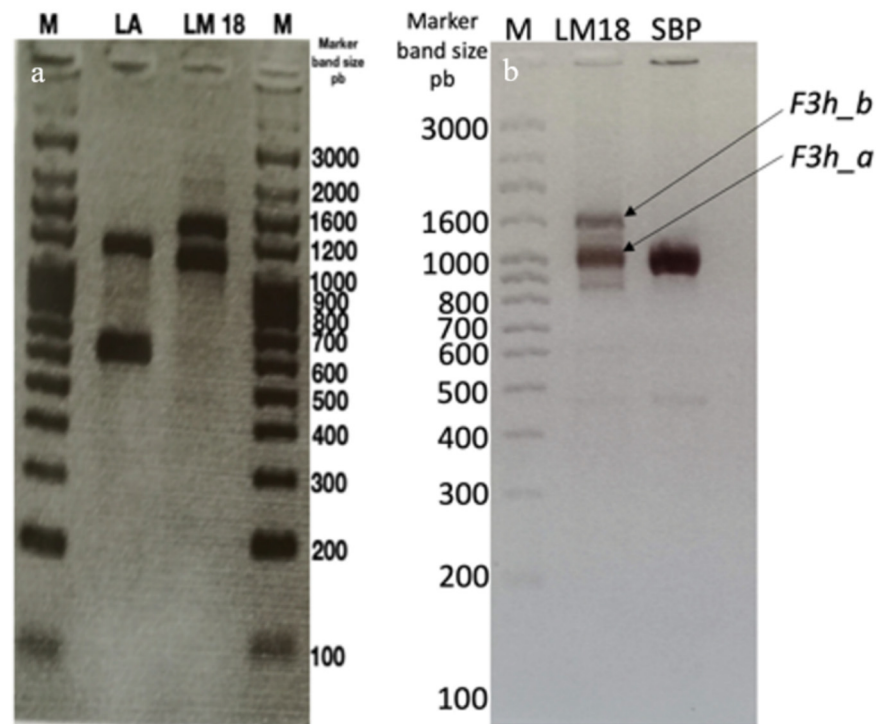
<sup>1</sup> Tm: melting temperature; <sup>2</sup> Eff: efficiency.

## 3. Results

### 3.1. DNA Sequences Analysis

Based on soybean and narrow-leaved lupin *F3H* gene sequences, we amplified an orthologous region in the tarwi genome. Two fragments were obtained in *L. angustifolius* (with ~700 pb and ~1200 pb) and two with ~1100 pb and ~1600 pb were obtained in *L. mutabilis* LM18 (brown seeds with dark-brown eyebrow) (Figure 3a). The amplification

of DNA from the *L. mutabilis* SBP genotype (white seeds) only yielded one fragment with ~1100 bp (Figure 3b).

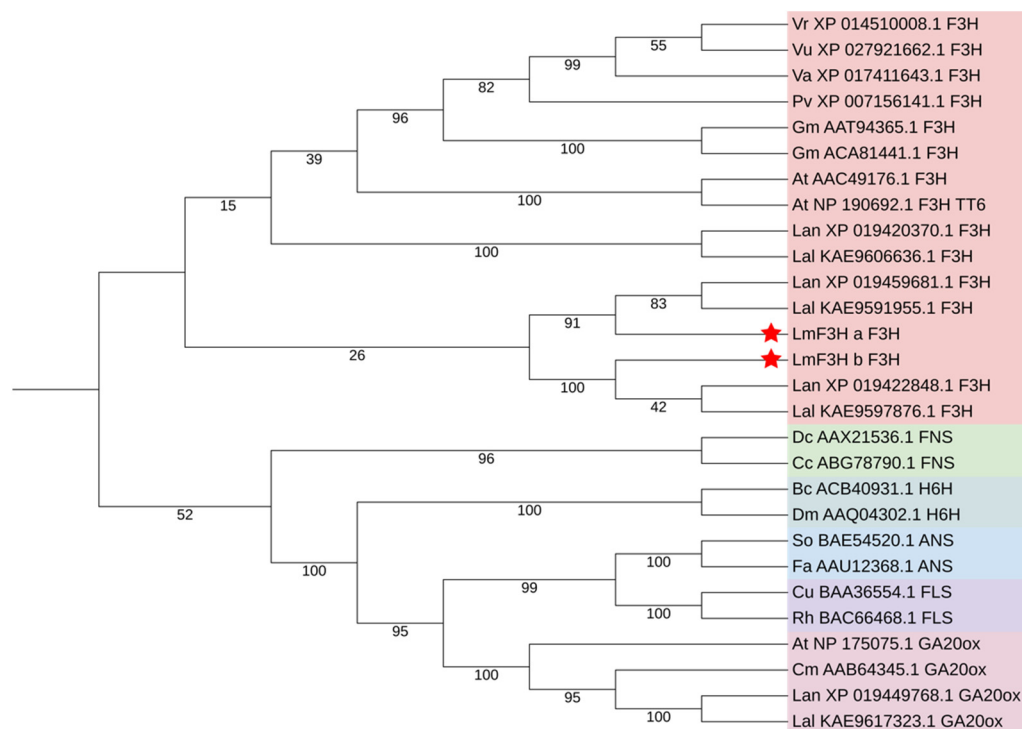


**Figure 3.** Amplification profiles obtained *Lupinus angustifolius* (LA; panel a) and *L. mutabilis* (LM18 and SBP; panels (a,b) using the *F3H* primers 5′GATTGGAGAGAGATTGTGACATA3′ (forward) and 5′GGTTCAAGAATGCTGATCAC3′ (reverse) separated in 2% agarose gel. M- NZYDNA Ladder V marker.

The nucleotide sequence analysis of the three DNA fragments (Figure 3b) revealed no differences between the 1100 bp fragments of SBP and LM18. This fragment was termed *LmF3h\_a* (Genbank reference MW387156). The 1600 bp fragment, only identified in brown seeds, was termed *LmF3h\_b* (Genbank reference MW387157). Upon comparison with the *L. angustifolius* naringenin, 2-oxoglutarate 3-dioxygenase-like protein sequence (XP\_019420372), an intronic and an exonic region were identified in each of *LmF3h\_a* and *LmF3h\_b* sequences. The alignment of *LmF3h\_a* and *LmF3h\_b* sequences revealed that, in the intron, 254 out of 674 nucleotides (37.7%) have mutations. In the exon, of the 373 nucleotides, there are 73 with mutations (19.6%). Of these 73 mutations, 29 are synonymous and 44 are non-synonymous, that is, 60.3% of exon mutations are non-synonymous, which translates to 26 amino acids different in an exon with 123 amino acids (20.3%) (Figure S2), half of which represent changes in the amino acid family. Even though the rate of exon mutation is lower than that of the intron (19.6% and 37.7%, respectively), most mutations in the exon are non-synonymous. Although non-synonymous mutations represented 11.8% of the nucleotides of the exon, they translate into 20.3% differences in polypeptides.

The variability observed between *LmF3h\_a* and *LmF3h\_b* is too great for them to be considered alleles of the same gene (~80%). This finding is supported by the relative position of *LmF3H\_a* and *LmF3H\_b* in the dendrogram (Figure 4) with three similar proteins of *L. angustifolius* (Supplementary material Table S1). *LmF3H\_a* and *LmF3H\_b* grouped in different clusters suggesting that they represent paralogous genes rather than alleles of the same gene (Supplementary material Table S1 and Figure 4). Moreover, the *F3H* proteins including the tarwi fragments group in the same clade with other *F3H* proteins of several legume species and the well-characterized *Arabidopsis* TT6, while the other 2-oxoglutarate

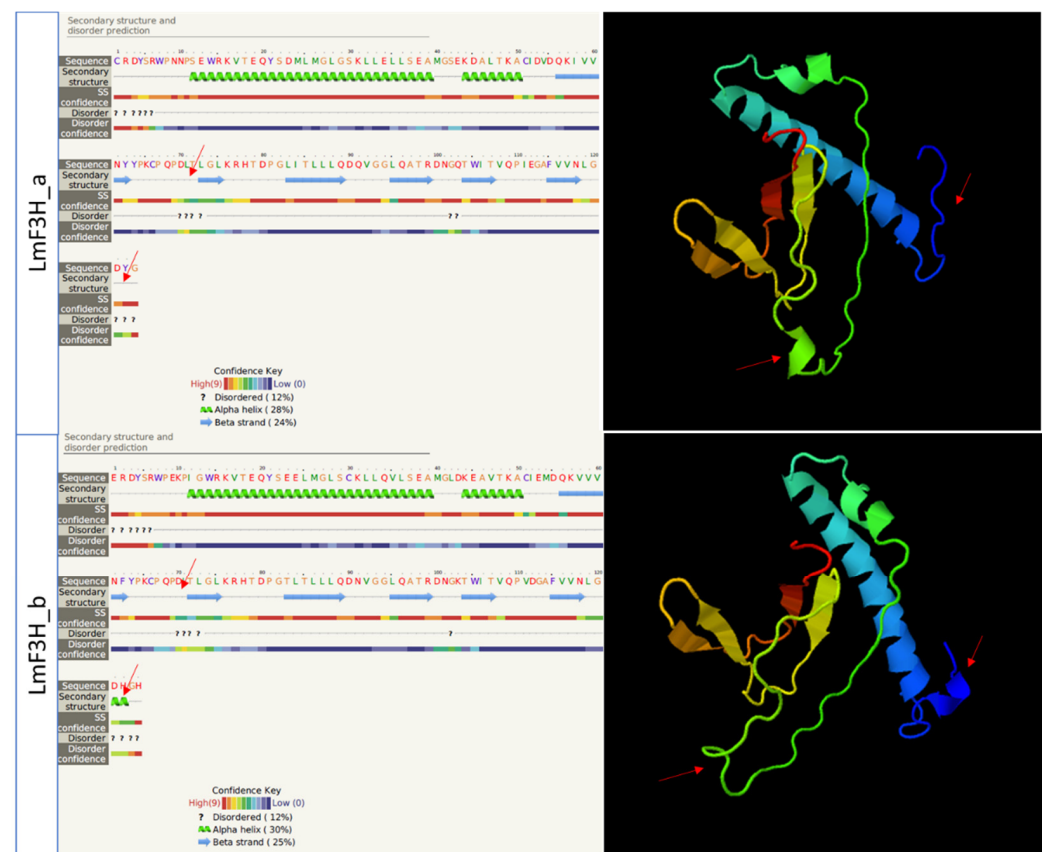
dioxygenases bearing proteins (FNS, ANS, FLS, H6H and G20ox) constitute an independent group, which is in accordance with the percent of identity between F3H proteins and the other 2-oxoglutarate dioxygenases bearing proteins (Supplementary material Table S1).



**Figure 4.** Phylogenetic tree illustrating similarity between *Lupinus mutabilis* polypeptide sequences LmF3h\_a and LmF3h\_b (red stars) and other F3H proteins from other *Lupinus* species, other F3H protein sequences from several Fabaceae species and from *Arabidopsis thaliana* including TT6, as well as to flavone synthase (FNS), anthocyanidin synthase (ANS), flavonol synthase (FLS), hyoscyamine 6 $\beta$ -hydroxylase (H6H) and gibberellin 20 oxidase (G20ox) of other different species. At—*Arabidopsis thaliana*; Bc—*Brugmansia candida*; Cu—*Citrus unshiu*; Cm—*Cucurbita maxima*; Cc—*Cuminum cyminum*; Dm—*Datura metel*; Dc—*Daucus carota*; Fa—*Fragaria × ananassa*; Gm—*Glycine max*; Lal—*Lupinus albus*; Lan—*L. angustifolius*; Lm—*L. mutabilis*; Pv—*Phaseolus vulgaris*; Rh—*Rosa hybrida*; So—*Spinacia oleracea*; Va—*Vigna angularis*; Vr—*V. radiata*; Vu—*V. unguiculata*.

Differences reported in the primary structure of the LmF3H\_a and LmF3H\_b predicted that polypeptides generate little differences in the predicted secondary and tertiary structures, as depicted in Figure 5. The analysis of secondary structures of the polypeptide fragments and the prediction of the disorder revealed the same percentage of disorder estimated at 12% for the two proteins, differing only in the alpha helix and beta strand with about 28 and 23% for the LmF3H\_a and around 30 and 25% for the LmF3H\_b, respectively. The differences observed for the alpha helix and beta strand in the two sequences were only 2%, and did not produce major changes in the predicted structures of the protein fragments. LmF3H\_a and LmF3H\_b fragments possess three out of five F3H highly conserved motifs (2, 3 and 4) which were detected in legumes including lupins as well as in the *Arabidopsis thaliana* TT6 gene with proved function as F3H [67]. These motifs contain two histidines (H220 and H278) and one aspartic acid (D222) of the ferrous iron binding sites, as well as the three strictly conserved prolines (P148, P204, P207), indicated by arrows (Figure 6). These were suggested to have an important role on the folding process of the protein [68]. Tarwi fragments also show the specific domains of the F3H proteins (Supplementary material Figure S4).





**Figure 5.** Predicted secondary (left) and tertiary (right) structures of LmF3H\_a (top) and LmF3H\_b (bottom) partial polypeptides in *Lupinus mutabilis*. Arrows indicate the region responsible Figure 3. H\_a protein the beta strand starts in aa. 73 and ends in 74 position while, in LmF3H\_b protein, this region is longer, starting at aa. 72 and ending in position 75. The mutation changing amino acid Y (LmF3H\_a) to H (LmF3H\_b) at position 122 resulted in the production of the short alpha helix. In the three-dimensional representations, images are coloured by rainbow N→C terminus.

### 3.2. Gene Expression Analysis

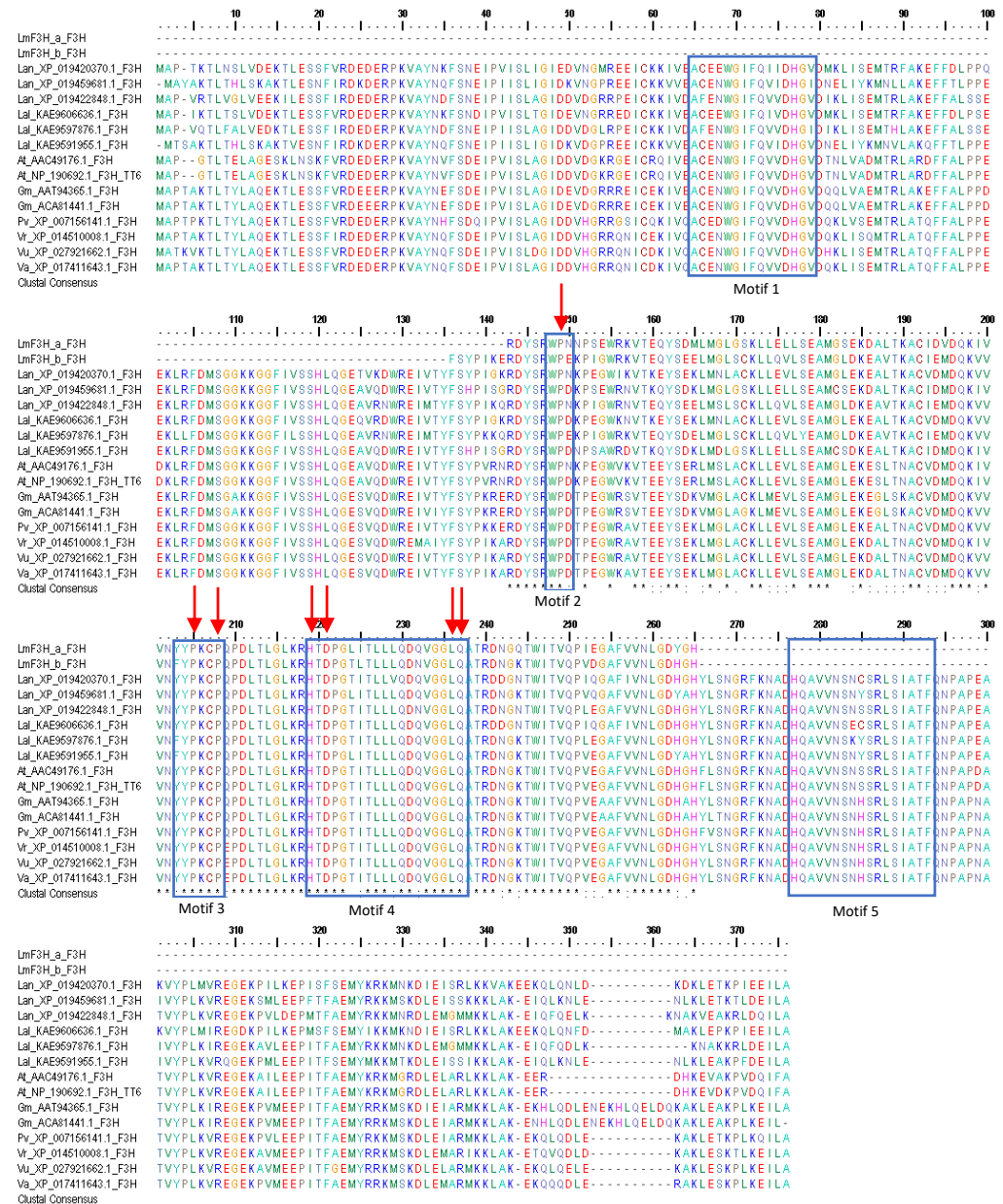
#### 3.2.1. Selection and Validation of Reference Genes

A screening test was performed for the five reference genes selected using DNA and cDNA, and the result revealed that only the ATPsyn gene primers did not amplify tarwi DNA and cDNA (Figure S3a,b), therefore ATPsyn gene was discarded. The analysis for selection reference genes were carried out for the remaining four genes. Although all genes can be used, we recommend the use of the *Ubc* and *Hel* genes regardless of the colour of the seed coat because they were the most specific (Supplementary material Figure S5). Based on the geNorm, NormFinder and  $\Delta C_t$  algorithms, we conclude that all genes are stable (Supplementary material Table S2 and Figure S6). This finding is supported by results of the variation in pairs, V2/3 and V3/4 were 0.024 and 0.015, respectively (Supplementary material Figure S7), less than 0.15, suggesting that these genes can be used for normalization of qPCR data in tarwi.

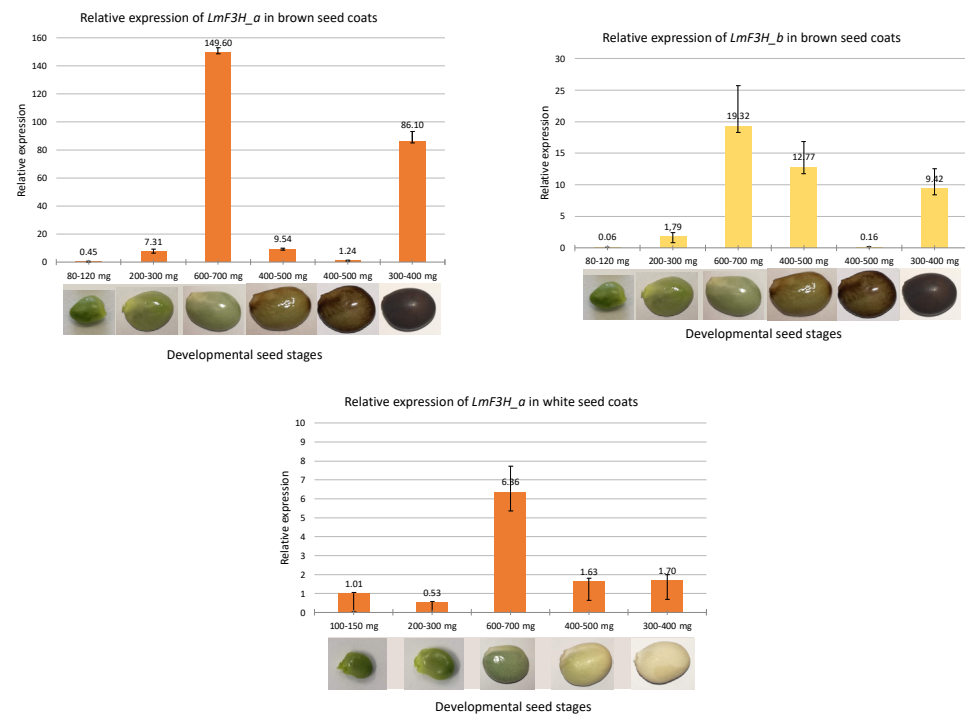
#### 3.2.2. Expression Profiles of *LmF3h\_a* and *LmF3h\_b*

Different stages of seed development in brown and white phenotypes were compared with each other in terms of the expression of *LmF3h\_a* and *LmF3h\_b* (Figure 7). A transcript analysis in different stages of development of brown and white seeds showed different levels of expression of *LmF3h\_a* and *LmF3h\_b* gene fragments. For brown seed, transcripts of both fragments present an increasing expression in the initial stages of seed development, unlike white seeds where there was a decrease in *LmF3h\_a* expression in the third stage of

seed development. We noted that in both phenotypes the pattern of expression profiles is the same for *LmF3h\_a* similar to the teeth of a saw, with the greatest expression in stage four, decreasing at stage six and raising again in stage seven. Furthermore, in both paralogues (*LmF3h\_a* and *LmF3h\_b*) and phenotypes (brown and white seeds) the maximum expression of transcripts was observed in the transition stage (600–700 mg) and the highest level of expression of *LmF3h\_a* paralogue was recorded in brown seeds, rather than in white seeds.



**Figure 6.** Alignment of the *Lupinus mutabilis* LmF3H\_a and LmF3H\_b protein fragments with other F3H proteins showing the conserved nature of the tarwi sequences. The five motifs specific of these proteins are outlined. Arrows indicate the strictly conserved amino acids present in the tarwi fragments. At—*Arabidopsis thaliana*; Gm—*Glycine max*; Lal—*Lupinus albus*; Lan—*L. angustifolius*; Pv—*Phaseolus vulgaris*; Vr—*Vigna radiata*; Vu—*V. unguiculata*.



**Figure 7.** Expression levels (and standard error) of the *LmF3h\_a* and *LmF3h\_b* paralogs in different stage of *Lupinus mutabilis* seed development, normalised (using initial stage of seed development as control for LM18 and SBP genotype) with *Ubc*, *Hel*, *Adh3* and *atub* reference genes, in different experimental samples. *LmF3h\_b* paralogue is only present in LM18 genotype.

#### 4. Discussion

In this study we identified and sequenced a portion of a putative *F3H* gene in *L. mutabilis*. In spite of sequence homologies, further studies will be needed in order to obtain the full sequence of the gene and to validate its biochemical role.

*LmF3h\_a* and *LmF3h\_b* paralogous fragments were detected in the genome of tarwi accessions with brown seed coat, while in accessions with white seed coat only the *LmF3h\_a* paralogous fragment was detected. This result suggests that the *LmF3h\_b* paralogue may be associated with the brown seed pigmentation process, although an analysis of different tarwi genotypes is needed to thoroughly fundament this relationship. A similar result was reported by Himi et al. [69], that observed the differential expression of the *F3H-A1*, *F3H-B1*, and *F3H-D1* genes in wheat grains and coleoptiles and found that the transcripts of the three genes were all detected in red seeds, while in white seeds only the *F3H-A1* gene was detected. It is interesting to note that these results corroborate those reported in the present study, highlighting the fact that both *LmF3h* paralogous fragments are expressed in the brown-coated seeds. The non-detection of the *LmF3h\_b* paralogue in white seeds can be a result of a deletion or of a large insertion into the amplified region that did not allow its amplification by PCR. Insertions can be related with transposition events, and some genes of the flavonoid biosynthetic pathway are known to be able to house transposable elements that can prevent their transcription [70]. In fact, a long terminal repeat type retroelement was reported in the *F3H* gene in *Torenia* [71] which is responsible for white flowers. The similarity between our polypeptides (*LMF3H\_a* and *LMF3H\_b*) and those of *L. angustifolius* (XP\_019459681, XP\_019420370 and XP\_019422848) suggests that we are in presence of two paralogs, and that, taking into account the likely synteny of these two lupins, the paralogs *LmF3h\_a* and *LmF3h\_b* may be located on different chromosomes. The comparison of the *LmF3H\_a* and *LmF3H\_b* nucleotide sequences to those of *F3H* genes from other Fabaceae species and to those of *Arabidopsis thaliana* evidence a considerable level of genetic diversity that partially translates to divergence in the respective proteins, but which contrasts with a highly conserved secondary structure, suggesting a conserved

function among them. The inference that *LmF3H\_a* and *LmF3H\_b* are paralogues lays chiefly in the comparison to *L. angustifolius* sequences. It is important to bear in mind the phylogenetic placement of *Lupinus* [72] in the genistoid clade of the Papilionoideae subfamily, (mostly composed of woody plant species, including *Cytisus*, *Genista*, *Spartium*, *Teline*, and *Ulex*), for which little or no information is available regarding the *F3H* gene or others genes related to flavonoid biosynthesis, apart from the major grain and forage legumes that are placed in the *indigoferoid/millettioid* (e.g., *Glycine*, *Phaseolus*, *Pueraria*, and *Vigna*) or in the *Hologalegina* clades (e.g., *Cicer*, *Lathyrus*, *Lens*, *Lotus*, *Medicago*, *Melilotus*, *Pisum*, *Trifolium*, and *Vicia*). The estimated 56 million year diverge between the genistoids and the remaining Papilionoideae clades [73] may explain the divergence found in the nucleotide and protein sequence between the putative *F3H* partial sequences from lupins and *F3H* from other grain legumes.

RT-qPCR is an extremely powerful and widely used tool for precise quantification of gene expression. In this study we reported a higher level of expression of the *LmF3h\_a* in brown seed coat than in white-coated seeds. The high levels of expression reported in our study corroborate Shirley et al. [74], Koornneef [75], Wisman et al. [67], and Lepiniec et al. [1] studies that associated brown seed coat with *F3H* gene expression. Similarly, high levels of expression of the *F3H-A*, *F3H-B*, and *F3H-C* alleles were detected in brown seed of *Ipomoea nil* [76]. Wan et al. [77], studying peanut pigmentation using transcriptomic approaches, found high levels of expression of *F3H*, *F3'H*, *DFR*, and *ANR* in brown peanuts, suggesting that these four genes represent the key for understanding the pigmentation of the peanut tegument. Himi and Noda [78] studied the expression of the *CHS*, *CHI*, *F3H*, and *DFR* genes in white and red wheat seeds. These authors noted a high expression of these transcripts in lines with red skin and low expression in white lines. Shao et al. [79] found that there were higher levels of flavonoids in black and red grain rice than in white phenotype lines, as well as in common bean seeds with brown, red and black seed coat compared with white seeds [80,81]. It was interesting to note that although there are differences in the levels of expression in the paralogues during seed development in both phenotypes, the pattern of expression is similar in the two genotypes with the maximum expression for *LmF3h\_a* and *LmF3h\_b* registered in an intermediate stage of seed development, specifically in the 600–700 mg stage. This is not in accordance with the results of Zabala and Vodkin [54], where they found *F3H* expressed at higher level early in black seed development (100–200 mg), which may reflect differences in seed development between lupins and soybean, namely in anthocyanin accumulation. Moreover, our results point to a redundancy of *LmF3h\_a* and *LmF3h\_b* roles in flavonols accumulation during seed development.

## 5. Conclusions

In our study it was possible to identify two putative paralogues of the *LmF3H* gene (*LmF3h\_a* and *LmF3h\_b*) in brown-coated *L. mutabilis* seeds and only one in white seed coats (*LmF3h\_a*), with *LmF3h\_a* being overexpressed in brown seeds. These results represent an important basis for developing molecular markers for early detection of plants generating brown coat seeds and to be used in tarwi breeding programmes, as coloured seeds are valued by consumers and seed coat colour in tarwi is not an easily genetically tractable trait. Moreover, it also provides a guideline for future studies of tarwi genome-wide association.

The results of the present study provide ground for further exploration and scientific understanding of the enormous phenotypic variability observed in tarwi seeds. On the other hand, they provide fundamental bases for the development of specific molecular markers for the early identification of the brown seed phenotype in the initial stages of plant development. They also serve as a basis for the study of chemical compounds synthesized in coloured-seed genotypes and their importance as a functional food, allowing, in the near future, breeding programmes to be directed towards the intensification of the brown pigmentation in tarwi, including potential health benefits regarding the bioavailability of minerals relatable to seed coat colour variability.



**Supplementary Materials:** The following supporting information can be downloaded at: <https://www.mdpi.com/article/10.3390/agronomy12020450/s1>, Supplementary material, Figure S1: Multiple sequence alignment of partial F3H nucleotide sequence from *Glycine max*, *Vigna angularis* and *Lupinus angustifolius* orthologs, illustrating primer location 5′GATTGGAGAGAGATTGTGACATA3′ (forward), 5′GTGATCAGCATTCTTGAACC3′ (reverse); Figure S2: Alignment of the F3H predicted polypeptide sequence of paralogues LmF3h\_a and LmF3h\_b of *Lupinus mutabilis*; Figure S3: PCR amplification patterns of reference genes using cDNA and genomic DNA as templates. (a) PCR from cDNA and (b) from DNA. M- NZYDNA Ladder VII marker; Figure S4: Conserved domains detected in LmF3H\_a, LmF3H\_b, and other F3H proteins evidencing the putative function of the *Lupinus mutabilis* proteins. LmF3H although incomplete contain the 2-OG-Fe(ii) conserved domain, which is specific of this superfamily; Figure S5: Melting curve analyses of: (a) *Adh3* with detection temperature around 78.5 °C; (b) *Hel* (80.5 °C); (c) *αTub* (80 °C), and (d) *Ubc* (79.5 °C) candidate reference genes for quantitative Reverse-Transcription PCR (RT-qPCR) of *Lupinus mutabilis*. In the four reference genes, pure and single amplicons were obtained; Figure S6: Analysis of the stability of the reference genes (under validation for subsequent analyses of *Lupinus mutabilis* seed coat development) based on the comprehensive classification of RefFinder. The most stables are *Ubc* and *Hel* and the least stables are *αTub* and *Adh3*; Figure S7: Variation in pairs to identify the optimal number of genes for normalisation; Table S1: Similarity matrix between *Lupinus mutabilis* Lm\_F3H\_a and Lm\_F3H\_b predicted amino acid sequences, F3H orthologs (with reference to the *L. angustifolius* chromosome location of the respective genes) and FNS, ANS, FLS, H6H and GA20ox proteins; Table S2: Stability values, classification by algorithm and general for the four reference genes under validation for subsequent analyses of *Lupinus mutabilis* seed coat development.

**Author Contributions:** Conceptualization, N.G., P.T. and L.M.-C.; methodology, N.G., H.G.A., S.A., P.T. and L.M.-C.; formal analysis, N.G.; investigation, N.G., H.G.A., S.A.; resources, H.G.A., J.N.-M., P.T. and L.M.-C.; writing—original draft preparation, N.G.; writing—review and editing, P.T. and L.M.-C.; supervision, J.N.-M., P.T. and L.M.-C.; project administration, J.N.-M.; funding acquisition, J.N.-M. All authors have read and agreed to the published version of the manuscript.

**Funding:** This research was funded by the Bio-based Industries Joint Undertaking under the European Union's Horizon 2020 research and innovation programme, the LIBBIO project, under grant agreement No. 720726 and the APC was also funded under the framework of the LIBBIO project, under grant agreement No 720726. This research was also financially supported by the R&D unit LEAF (UIDB/04129/2020 and UIDP/04129/2020; Fundação para a Ciência e a Tecnologia, Portugal).

**Institutional Review Board Statement:** Not applicable.

**Informed Consent Statement:** Not applicable.

**Data Availability Statement:** Nucleotide sequences obtained in the scope of this study were deposited in GenBank under accessions MW387156 (*LmF3h\_a*) and MW387157 (*LmF3h\_b*).

**Conflicts of Interest:** The authors declare no conflict of interest.

## References

1. Lepiniec, L.; Debeaujon, I.; Routaboul, J.M.; Baudry, A.; Pourcel, L.; Nesi, N.; Caboche, M. Genetics and biochemistry of seed flavonoids. *Annu. Rev. Plant Biol.* **2006**, *57*, 405–430. [[CrossRef](#)] [[PubMed](#)]
2. Brunetti, C.; Di Ferdinando, M.; Fini, A.; Pollastri, S.; Tattini, M. Flavonoids as antioxidants and developmental regulators: Relative significance in plants and humans. *Int. J. Mol. Sci.* **2013**, *14*, 3540–3555. [[CrossRef](#)] [[PubMed](#)]
3. Ahmed, N.U.; Park, J.; Jung, H.J.; Yoonkang Hur, Y.; Nou, S.-I. Anthocyanin biosynthesis for cold and freezing stress tolerance and desirable colour in *Brassica rapa*. *Funct. Integr. Genom.* **2015**, *15*, 383–394. [[CrossRef](#)] [[PubMed](#)]
4. Panche, A.; Diwan, A.; Chandra, S. Flavonoids: An overview. *J. Nutr. Sci.* **2016**, *5*, e47. [[CrossRef](#)]
5. Jarial, R.; Thakur, S.; Sakinah, M.; Zularisam, A.W.; Sharad, A.; Kanwar, S.S.; Singh, L. Potent anticancer antioxidant and antibacterial activities of isolated flavonoids from *Asplenium nidus*. *J. King Saud Univ.-Sci.* **2018**, *30*, 185–192. [[CrossRef](#)]
6. Jaakola, L.; Määtä, K.; Pirttilä, A.M.; Törrönen, R.; Kärenlampi, S.; Hohtola, A. Expression of genes involved in anthocyanin biosynthesis in relation to anthocyanin, proanthocyanidin, and flavonol levels during bilberry fruit development. *Plant Physiol.* **2002**, *130*, 729–739. [[CrossRef](#)]
7. Guo, N.; Cheng, F.; Wu, J.; Liu, B.; Zheng, S.N.; Liang, J.L.; Wang, X.W. Anthocyanin biosynthetic genes in *Brassica rapa*. *BMC Genom.* **2014**, *15*, 426. [[CrossRef](#)]
8. Grotewold, E. The genetics and biochemistry of floral pigments. *Annu. Rev. Plant Biol.* **2006**, *57*, 761–780. [[CrossRef](#)]

9. Quattrocchio, F.; Wing, J.F.; Leppen, H.T.C.; Mol, J.N.M.; Koes, R.E. Regulatory genes controlling anthocyanin pigmentation are functionally conserved among plant species and have distinct sets of target genes. *Plant Cell* **1993**, *5*, 1497–1512. [\[CrossRef\]](#)
10. Shirley, B. Flavonoid biosynthesis. A colorful model for genetics biochemistry cell biology and biotechnology. *Plant Physiol.* **2001**, *126*, 485–493. [\[CrossRef\]](#)
11. Han, Y.; Huang, K.; Liu, Y.; Jiao, T.; Ma, G.; Qian, Y.; Wang, P.; Dai, X.; Gao, L.; Xia, T. Functional Analysis of Two Flavanone-3-Hydroxylase Genes from *Camellia sinensis*: A Critical Role in Flavonoid Accumulation. *Genes* **2017**, *8*, 300. [\[CrossRef\]](#) [\[PubMed\]](#)
12. Charrier, B.; Coronado, C.; Kondorosi, A.; Ratet, P. Molecular characterization and expression of alfalfa (*Medicago sativa* L.) flavanone-3-hydroxylase and dihydroflavonol-4-reductase encoding genes. *Plant Mol. Biol.* **1995**, *29*, 773–786. [\[CrossRef\]](#) [\[PubMed\]](#)
13. Cotting, R.M.; Hosfield, G.L. Flavanone 3-Hydroxylase: A Candidate Gene Product for the P Color Gene. *USDA Pub.* **2005**, *48*, 38–39.
14. Shen, X.; Martens, S.; Chen, M.; Li, D.; Dong, J.; Tao, W. Cloning and characterization of a functional flavanone-3-hydroxylase gene from *Medicago truncatula*. *Mol. Biol. Rep.* **2009**, *37*, 3283–3289. [\[CrossRef\]](#)
15. Cheng, H.; Wang, J.; Chu, S.; Yan, H.; Yu, D. Diversifying Selection on Flavanone 3-Hydroxylase and Isoflavone Synthase Genes in Cultivated Soybean and Its Wild Progenitors. *PLoS ONE* **2013**, *8*, e54154. [\[CrossRef\]](#)
16. McClean, P.E.; Lee, R.K.; Otto, C.; Gepts, P.; Bassett, M.J. Molecular and phenotypic mapping of genes controlling seed coat pattern and colour in common bean (*Phaseolus vulgaris* L.). *J. Hered.* **2002**, *93*, 148–152. [\[CrossRef\]](#)
17. Weber, H.; Borisjuk, L.; Wobus, U. Molecular physiology of legume seed development. *Ann. Rev. Plant Biol.* **2005**, *56*, 253–279. [\[CrossRef\]](#)
18. Bellaloui, N. Soybean seed phenol lignin and isoflavones partitioning as affected by seed node position and genotype differences. *Food Nutr. Sci.* **2012**, *3*, 447–454. [\[CrossRef\]](#)
19. Todd, J.J.; Vodkin, L.O. Pigmented soybean (*Glycine max*) seed coats accumulate proanthocyanidins during development. *Plant Physiol.* **1993**, *102*, 663–670. [\[CrossRef\]](#)
20. Devic, M.; Guillemot, J.; Debeaujon, I.; Bechtold, N.; Bensaud, E.; Koornneef, M.; Pelletier, G.; Delseny, M. The *BANYULS* gene encodes a DFR-like protein and is a marker of early seed coat development. *Plant J.* **1999**, *19*, 387–398. [\[CrossRef\]](#)
21. Debeaujon, I.; Peeters, A.J.M.; Léon-Kloosterziel, K.M.; Koornneef, M. The *TRANSPARENT TESTA12* gene of *Arabidopsis* encodes a multidrug secondary transporter-like protein required for flavonoid sequestration in vacuoles of the seed coat endothelium. *Plant Cell* **2001**, *13*, 853–872. [\[CrossRef\]](#) [\[PubMed\]](#)
22. Gupta, R.; Min, C.W.; Kim, S.W.; Wang, Y.; Agrawal, G.K.; Rakwal, R.; Kim, G.S.; Won Lee, W.B.; Ko, J.M.; Baek, I.Y.; et al. Comparative investigation of seed coats of brown- versus yellow-colored soybean seeds using an integrated proteomics and metabolomics approach. *Proteomics* **2014**, *15*, 1706–1716. [\[CrossRef\]](#) [\[PubMed\]](#)
23. Dwivedi, S.L.; Upadhyaya, H.D.; Chung, M.I.; Vita, P.; García-Lara, L.S.; Guajardo-Flores, D.; Gutiérrez-Urbe, J.A.; Serna-Saldívar, S.O.; Rajakumar, G.; Sahrawat, K.L.; et al. Exploiting Phenylpropanoid Derivatives to Enhance the Nutraceutical Values of Cereals and Legumes. *Front. Plant Sci.* **2016**, *7*, 763. [\[CrossRef\]](#) [\[PubMed\]](#)
24. Atchinson, G.W.; Nevado, B.; Eastwood, R.J.; Contreras-Ortiz, N.; Reynel, C.; Madriñán, S.; Filatov, D.A.; Hughes, C.E. Lost crops of the Incas: Origins of domestication of the Andean pulse crop tarwi *Lupinus mutabilis*. *Am. J. Bot.* **2016**, *103*, 1592–1606. [\[CrossRef\]](#)
25. Carvajal-Larenas, F.E.; Linnemann, A.R.; Nout, M.J.R.; Koziol, M.; Van Boekel, M.A.J.S. *Lupinus mutabilis*: Composition uses toxicology and debittering. *Crit. Rev. Food Sci. Nutr.* **2016**, *56*, 1454–1487. [\[CrossRef\]](#)
26. Martins, J.M.N.; Talhinas, P.; Sousa, R.B. Yield and seed chemical composition of *Lupinus mutabilis* in Portugal. *Rev. Ciênc. Agrár.* **2016**, *39*, 518–525. [\[CrossRef\]](#)
27. Gulisano, A.; Alves, S.; Neves-Martins, J.; Trindade, L. Genetics and breeding of *Lupinus mutabilis*: An emerging protein crop. *Front. Plant Sci.* **2019**, *10*, 1385. [\[CrossRef\]](#)
28. Guilengue, N.; Alves, S.; Talhinas, P.; Neves-Martins, J. Genetic and genomic diversity in a tarwi (*Lupinus mutabilis* Sweet) germplasm collection and adaptability to Mediterranean climate conditions. *Agronomy* **2020**, *10*, 21. [\[CrossRef\]](#)
29. Talhinas, P.; Baroncelli, R.; Le Floch, G. Anthracnose of lupins caused by *Colletotrichum lupini*: A recent disease and a successful worldwide pathogen. *J. Plant Pathol.* **2016**, *98*, 5–14. [\[CrossRef\]](#)
30. Talhinas, P.; Neves-Martins, J.; Oliveira, H. Evaluation of anthracnose resistance in *Lupinus* spp. germplasm. *Rev. Ciênc. Agrár.* **2016**, *39*, 550–570. [\[CrossRef\]](#)
31. Guilengue, N.; Neves-Martins, J.; Talhinas, P. Response to anthracnose in a tarwi (*Lupinus mutabilis*) collection is influenced by anthocyanin pigmentation. *Plants* **2020**, *9*, 583. [\[CrossRef\]](#) [\[PubMed\]](#)
32. Statler, G.D. Resistance of bean plants to *Fusarium solani* f. sp. *phaseoli*. *Plant Dis. Rep.* **1970**, *54*, 698–699.
33. Harris, H.B.; Burns, R.E. Relationship between tannin content of sorghum grain and preharvest seed molding. *Agron. J.* **1973**, *65*, 957–959. [\[CrossRef\]](#)
34. Stasz, T.E.; Harman, G.E.; Marx, G.A. Time and Site of Infection of Resistant and Susceptible Germinating Pea Seeds by *Pythium ultimum*. *Phytopathology* **1980**, *70*, 730–733. [\[CrossRef\]](#)
35. Islam, F.M.A.; Rengifo, J.; Redden, R.J.; Basford, K.E.; Beebe, S.E. Association Between Seed Coat Polyphenolics (Tannins) and Disease Resistance in Common Bean. *Plant Foods Hum. Nutr.* **2003**, *58*, 285–297. [\[CrossRef\]](#)

36. Lima, A.; Oliveira, J.; Saúde, F.; Mota, J.; Ferreira, R.B. Proteins in Soy Might Have a Higher Role in Cancer Prevention than Previously Expected: Soybean Protein Fractions Are More Effective MMP-9 Inhibitors Than Non-Protein Fractions, Even in Cooked Seeds. *Nutrients* **2017**, *9*, 201. [\[CrossRef\]](#) [\[PubMed\]](#)
37. Xu, B.J.; Yuan, S.H.; Chang, S.K. Comparative analyses of phenolic composition, antioxidant capacity, and color of cool season legumes and other selected food legumes. *Food Sci.* **2007**, *72*, S167–S177. [\[CrossRef\]](#)
38. Xiao, C.W. Health Effects of Soy Protein and Isoflavones in Humans. *J. Nutr.* **2008**, *138*, 1244S–1249S. [\[CrossRef\]](#)
39. Zhang, R.F.; Zhang, F.X.; Zang, M.W.; Wei, Z.C.; Yang, C.Y.; Zhang, Y.; Tang, X.J.; Deng, Y.Y.; Chi, J.W. Phenolic Composition and Antioxidant Activity in Seed Coats of 60 Chinese Black Soybean (*Glycine max* L. Merr.) Varieties. *J. Agric. Food Chem.* **2011**, *59*, 5935–5944. [\[CrossRef\]](#)
40. Segev, A.; Badani, H.; Kapulnik, Y.; Shomerand, I.; Oren-Shamir, M.; Galili, S. Determination of phenolic compounds, flavonoids, and antioxidant activity colored chickpea (*Cicer arietinum* L.). *J. Food Sci.* **2010**, *75*, S115–S119. [\[CrossRef\]](#)
41. Ariza-Nieto, M.; Blair, M.W.; Welch, R.M.; Glahn, R.P. Screening of bioavailability patterns in eight bean (*Phaseolus vulgaris* L.) genotypes using the Caco-2 cell in vitro model. *J. Agric. Food Chem.* **2007**, *55*, 7950–7956. [\[CrossRef\]](#) [\[PubMed\]](#)
42. Blair, M.W.; Izquierdo, P.; Astudillo, C.; Grusak, M.A. A legume biofortification quandary: Variability and genetic control of seed coat micronutrient accumulation in common beans. *Front. Plant Sci.* **2013**, *4*, 275. [\[CrossRef\]](#) [\[PubMed\]](#)
43. Mishili, F.J.; Fulton, J.; Shehu, M.; Kushwaha, S.; Marfo, K.; Jamal, M.; Kergna, A.; De Borer, J.L. Consumer preferences for quality characteristics along the cowpea value chain in Nigeria Ghana and Mali. *Agribusiness* **2009**, *25*, 16–35. [\[CrossRef\]](#)
44. Kläsener, G.R.; Ribeiro, N.D.; Casagrande, C.R.; Arns, F.D. Consumer preference and the technological and nutritional quality of different bean colours. *Acta Sci. Agron.* **2020**, *42*, e43689. [\[CrossRef\]](#)
45. Mohan, V.; Ruchi, V.; Gayathri, R.; Bai, M.R.; Shobana, S.; Anjana, R.M.; Unnikrishnan, R.; Sudha, V. Hurdles in Brown Rice Consumption. In *Brown Rice*; Manickavasagan, A., Santhakumar, C., Venkatachalapathy, N., Eds.; Springer: Cham, Switzerland; New York, NY, USA, 2017; pp. 255–269. [\[CrossRef\]](#)
46. Bassett, M.J. Genetics of Seed Coat Color and Pattern in Common Bean. In *Plant Breeding Reviews*; Janick, J., Ed.; Wiley: New York, NY, USA, 2007; Volume 28, pp. 239–315. [\[CrossRef\]](#)
47. Gillman, J.D.; Tetlow, A.; Lee, J.D.; Shannon, J.G.; Bilyeu, K. Loss-of-function mutations affecting a specific *Glycine max* R2R3 MYB transcription factor result in brown hilum and brown seed coats. *BMC Plant Biol.* **2011**, *11*, 155. [\[CrossRef\]](#) [\[PubMed\]](#)
48. Yang, K.; Jeong, N.; Moon, J.K.; Lee, Y.H.; Lee, S.H.; Kim, H.M.; Hwang, C.H.; Back, K.; Palmer, R.G.; Jeong, S.C. Genetic analysis of genes controlling natural variation of seed coat and flower colors in soybean. *J. Hered.* **2010**, *101*, 757–768. [\[CrossRef\]](#)
49. Ubayasena, L.; Vijayan, P.; Bett, K.E.; Gray, G.R.; Küster, H.; Warkentin, T.D. Gene expression profiles of seed coats and biochemical properties of seed coats and cotyledons of two field pea (*Pisum sativum*) cultivars contrasting in green cotyledon bleaching resistance. *Euphytica* **2015**, *193*, 49–65. [\[CrossRef\]](#)
50. IBPGR. *Lupin Descriptors*; FAO: Rome, Italy, 1981.
51. Hane, J.K.; Ming, Y.; Kamphuis, L.G.; Nelson, M.N.; Carg, G.; Atkins, C.A.; Bayer, P.E.; Bravo, A.; Bringans, S.; Cannon, S.; et al. Comprehensive draft genome sequence for lupin (*Lupinus angustifolius*) an emerging health food: Insights into plant–microbe interactions and legume evolution. *Plant Biotechnol. J.* **2016**, *15*, 318–330. [\[CrossRef\]](#)
52. Kumar, S.; Stecher, G.; Li, M.; Knyaz, C.; Tamura, K. MEGA X: Molecular Evolutionary Genetics Analysis across computing platforms. *Mol. Biol. Evol.* **2018**, *35*, 1547–1549. [\[CrossRef\]](#)
53. Kelley, L.; Mezulis, S.; Yates, C.; Wass, M.; Sternberg, M. The Phyre2 web portal for protein modeling, prediction and analysis. *Nat. Protoc.* **2015**, *10*, 845–858. [\[CrossRef\]](#)
54. Zabala, G.; Vodkin, L.O. Methylation Affects Transposition and Splicing of a Large CACTA Transposon from a MYB Transcription Factor Regulating Anthocyanin Synthase Genes in Soybean Seed coats. *PLoS ONE* **2014**, *9*, e111959. [\[CrossRef\]](#) [\[PubMed\]](#)
55. Vieira, A.; Cabral, A.; Fino, J.; Azinheira, H.G.; Loureiro, A.; Talhinhas, P.; Pires, A.S.; Várzea, V.; Moncada, P.; Oliveira, H.; et al. Comparative validation of conventional and RNA-seq data-derived reference genes for qPCR expression studies of *Colletotrichum kahawae*. *PLoS ONE* **2016**, *11*, e0150651. [\[CrossRef\]](#)
56. Taylor, C.M.; Jost, R.; Erskine, W.; Nelson, M.N. Identifying Stable Reference Genes for RT-qPCR Normalisation in Gene Expression Studies of Narrow-Leafed Lupin (*Lupinus angustifolius* L.). *PLoS ONE* **2016**, *11*, e0148300. [\[CrossRef\]](#) [\[PubMed\]](#)
57. Kroc, M.; Koczyk, G.; Kamel, K.A.; Czepiel, K.; Fedorowicz-Strońska, O.; Krajewski, P.; Kosińska, J.; Podkowiński, J.; Wilczura, P.; Święcicki, W. Transcriptome-derived investigation of biosynthesis of quinolizidine alkaloids in narrow-leafed lupin (*Lupinus angustifolius* L.) highlights candidate genes linked to *iucundus* locus. *Sci. Rep.* **2019**, *9*, 2231. [\[CrossRef\]](#)
58. Xie, F.; Xiao, P.; Chen, D.; Xu, L.; Zhang, B. miRDeepFinder: A miRNA analysis tool for deep sequencing of plant small RNAs. *Plant Mol. Biol.* **2012**, *80*, 75–84. [\[CrossRef\]](#)
59. Pfaffl, M.W.; Tichopad, A.; Prgomet, C.; Neuvians, T.P. Determination of stable housekeeping genes differentially regulated target genes and sample integrity: BestKeeper-excel-based tool using pair-wise correlations. *Biotechnol. Lett.* **2004**, *26*, 509–515. [\[CrossRef\]](#)
60. Vandesompele, J.; De Preter, K.; Pattyn, F.; Poppe, B.; Van Roy, N.; De Paepe, A.; Speleman, F. Accurate normalization of real-time quantitative RT-PCR data by geometric averaging of multiple internal control genes. *Genome Biol.* **2002**, *3*, research00341. [\[CrossRef\]](#)

61. Andersen, C.L.; Jensen, J.L.; Orntoft, T.F. Normalization of real-time quantitative reverse transcription-PCR data: A model-based variance estimation approach to identify genes suited for normalization applied to bladder and colon cancer data sets. *Cancer Res.* **2004**, *64*, 5245–5250. [[CrossRef](#)]
62. Hellemans, J.; Mortier, G.; De Paepe, A.; Speleman, F.; Vandesompele, J. qBase relative quantification framework and software for management and automated analysis of real-time quantitative PCR data. *Genome Biol.* **2007**, *8*, R19. [[CrossRef](#)]
63. Silver, N.; Best, S.; Jiang, J.; Thein, S.L. Selection of housekeeping genes for genes expression studies in human reticulocytes using real time PCR. *BMC Mol. Biol.* **2006**, *7*, 33. [[CrossRef](#)]
64. Pfaffl, M.W. A new mathematical model for relative quantification in real-time RT-PCR. *Nucleic Acids Res.* **2001**, *29*, e45. [[CrossRef](#)] [[PubMed](#)]
65. Ruijter, J.M.; Ramakers, C.; Hoogaars, W.M.; Karlen, Y.; Bakker, O.; Van den Hoff, M.J.; Moorman, A.F. Amplification efficiency: Linking baseline and bias in the analysis of quantitative PCR data. *Nucleic Acids Res.* **2009**, *37*, e45. [[CrossRef](#)] [[PubMed](#)]
66. Livak, K.J.; Schmittgen, T.D. Analysis of relative gene expression data using real-time quantitative PCR and the 2(-Delta Delta C(T)) Method. *Methods* **2001**, *25*, 402–408. [[CrossRef](#)] [[PubMed](#)]
67. Wisman, E.; Hartmann, U.; Sagasser, M.; Baumann, E.; Palme, K.; Hahlbrock, K.; Saedler, H.; Weisshaar, B. Knock-out mutants from an En-1 mutagenized *Arabidopsis thaliana* population generate phenylpropanoid biosynthesis phenotypes. *Proc. Natl. Acad. Sci. USA* **1998**, *95*, 12432–12437. [[CrossRef](#)]
68. Xiong, S.; Tian, N.; Long, J.; Chen, Y.; Qin, Y.; Feng, J.; Xiao, W.; Liu, S. Molecular cloning and characterization of a flavanone 3-Hydroxylase gene from *Artemisia annua* L. *Plant Physiol. Biochem.* **2016**, *105*, 29–36. [[CrossRef](#)]
69. Himi, E.; Maekawa, M.; Noda, K. Differential Expression of Three Flavanone 3-Hydroxylase Genes in Grains and Coleoptiles of Wheat. *Intern. J. Plant Genom.* **2011**, *2011*, 369460. [[CrossRef](#)]
70. Itoh, Y.; Higeta, D.; Suzuki, A.; Yoshida, H.; Ozeki, Y. Excision of Transposable Elements from the Chalcone Isomerase and Dihydroflavonol 4-Reductase Genes May Contribute to the Variegation of the Yellow-Flowered Carnation (*Dianthus caryophyllus*). *Plant Cell Physiol.* **2002**, *43*, 578–585. [[CrossRef](#)]
71. Nishihara, M.; Yamada, E.; Saito, M.; Fujita, K.; Takahashi, H.; Nakatsuka, T. Molecular characterization of mutations in white-flowered torenia plants. *BMC Plant Biol.* **2014**, *14*, 86. [[CrossRef](#)]
72. LPWG. Legume phylogeny and classification in the 21st century: Progress, prospects and lessons for other species-rich clades. *Taxon* **2013**, *62*, 217–248. [[CrossRef](#)]
73. Lavin, M.; Herendeen, P.S.; Wojciechowski, M.F. Evolutionary rates analysis of Leguminosae implicates a rapid diversification of lineages during the tertiary. *Syst. Biol.* **2005**, *54*, 575–594. [[CrossRef](#)]
74. Shirley, B.W.; Kubasek, W.L.; Storz, G.; Bruggemann, E.; Koornneef, M.; Ausubel, F.M.; Goodman, H.M. Analysis of Arabidopsis mutants deficient in flavonoid biosynthesis. *Plant J.* **1995**, *8*, 659–671. [[CrossRef](#)] [[PubMed](#)]
75. Koornneef, M. Mutations affecting the testa colour in Arabidopsis. *Inf. Serv.* **1990**, *27*, 1–4.
76. Park, K.I.; Nitasaka, E.; Hoshino, A. Anthocyanin mutants of Japanese and common morning glories exhibit normal proanthocyanidin accumulation in seed coats. *Plant Biotechnol.* **2018**, *35*, 259–266. [[CrossRef](#)] [[PubMed](#)]
77. Wan, L.; Li, B.; Pandey, M.K.; Wu, Y.; Lei, Y.; Yan, L.; Dai, X.; Jiang, H.; Zhang, J.; Wei, G.; et al. Transcriptome Analysis of a New Peanut Seed Coat Mutant for the Physiological Regulatory Mechanism Involved in Seed Coat Cracking and Pigmentation. *Front. Plant Sci.* **2016**, *7*, 1491. [[CrossRef](#)] [[PubMed](#)]
78. Himi, E.; Noda, K. Red grain colour gene (R) of wheat is a Myb-type transcription factor. *Euphytica* **2005**, *143*, 239–242. [[CrossRef](#)]
79. Shao, Y.; Tang, F.; Huang, Y.; Xu, F.; Chen, Y.; Tong, C.; Chen, H.; Bao, J. Analysis of Genotype × Environment interactions for polyphenols and antioxidant capacity of rice by association mapping. *J. Agric. Food Chem.* **2014**, *62*, 5361–5368. [[CrossRef](#)] [[PubMed](#)]
80. Akond, A.S.M.G.M.; Khandaker, L.; Berthold, J.; Gates, L.; Peters, K.; Delong, H.; Hossain, K. Anthocyanin total polyphenols and antioxidant activity of common bean. *Am. J. Food Technol.* **2011**, *6*, 385–394. [[CrossRef](#)]
81. Agostini-Costa, T.S.; Teodoro, A.F.P.; Alves, R.B.N.; Braga, L.R.; Ribeiro, I.F.; Silva, J.P.; Quintana, L.G.; Burle, M.L. Total phenolics flavonoids tannins and antioxidant activity of lima beans conserved in a Brazilian genebank. *Ciênc. Rural* **2015**, *45*, 335–341. [[CrossRef](#)]