TÜBİTAK

Turkish Journal of Botany

http://journals.tubitak.gov.tr/botany/

Research Article

Turk J Bot
(2020) 44: 396-409
© TÜBİTAK
doi:10.3906/bot-2001-22

Identification and *in silico* analysis of *PHANTASTICA* gene in *Saintpaulia ionantha* H. Wendl (Gesneriaceae)

Mina KAZEMIAN[1], Elham MOHAJEL KAZEMI[1,*], Maryam KOLAHI[2], Valiollah GHASEMI OMRAN[3]

[1]Department of Plant Biology, Faculty of Natural Sciences, University of Tabriz, Tabriz, Iran
[2]Department of Biology, Faculty of Science, Shahid Chamran University of Ahvaz, Ahvaz, Iran
[3]Genetic and Agricultural Biotechnology Institute of Tabarestan, University of Agriculture Science and Natural Resources, Sari, Iran

Received: 15.01.2020 • **Accepted/Published Online:** 24.04.2020 • **Final Version:** 16.07.2020

Abstract: *Saintpaulia ionantha* H. Wendl (synonym *Streptocarpus ionanthus*) has been considered an ornamental plant. The *PHANTASTICA* (*PHAN*) gene has a significant role in the formation of plant organs adaxial–abaxial polarity. In the present study the *PHAN* gene was identified in *Saintpaulia* for the first time using PCR. To determine the expression pattern of the *PHAN* gene, gene expression was compared at 3 developmental stages using real time PCR. For in silico analyses, protein characteristics, secondary and 3D structures of protein, intracellular localization, and the phylogenetic tree were investigated using bioinformatics tools. The partial CDS of *PHAN* gene contained 756 nucleotides, a predicted protein encoded with 252 amino acids with a molecular weight of 28,474 Da. The presence of the HTH MYB-type motif in this sequence indicated that the PHAN protein belonged to the MYB family. The secondary structure of the PHAN protein was composed of 45% α-helix. The predicted PHAN protein 3D structure revealed that this protein can bind to nucleic acids. The binding site sequence included amino acids 10(R), 12(W), 13(H), 66(W), and 67(K). The intracellular localization of PHAN protein is believed to be in the nucleus. Phylogenetic analysis revealed that species belonging to each family were grouped in same clade. Gene expression revealed that in the later stages of petal development, expression was reduced. The results suggest that the expression pattern of the *PHAN* gene is that of adaxial–abaxial symmetry initiation during petal development as well as in the leaves of *Saintpaulia*. It appears that genetic engineering using biotechnology would be beneficial for improving the ornamental value of *Saintpaulia*.

Key words: Adaxial–abaxial, DNA binding site, HTH domain, MYB family, ornamental plant

1. Introduction

Flowers have been considered as a connection between humans and nature. The importance of variability in floral phenotypes can be seen among different plant species (Endress and Matthews, 2006). Over the past decade, due to increasing market demand, plant production conditions have changed (Aida et al., 2018). The establishment and development of well–equipped greenhouses for the purpose of ornamental plant cultivation have dramatically increased (Aida et al., 2018). It has been reported that the total global cultivation of ornamental plants was high (360,000 ha) with the global rate of ornamental flower consumption toping 200 billion US dollars, with a total of 22 billion dollars for the export of flowers in 2014 (Azadi et al., 2016). With the ever increasing commercial importance of ornamental plants in recent decades, there has been a significant increase in the number of studies performed on ornamental plants and flowers globally (Azadi et al., 2016). *Saintpaulia ionantha* H. Wendl (synonym *Streptocarpus ionanthus*), which is also known as African violet, is considered as an ornamental herbaceous plant originally from Tanzania and Kenya (Kolehmainen et al., 2010). African violet belongs to the Gesneriaceae family and Lamiales order (Haston and De Craene, 2007). The sympetalous petals are quinary, with various colors (purple, velvet, pale blue, red, pink, and white) in different shapes, with the possibility of flowering throughout the year (Nhut et al., 2018). Petals and leaves in most plants usually have 2 common characteristics: 1–They are formed as lateral organs (Hepworth and Lenhard, 2014); 2–During developmental stages they become flat and adaptable to the environment (Truskina and Vernoux, 2017). Flower formation is controlled by the gene regulatory network (GRN) that is involved in plant maturity, flowering time, and other related processes. The process of floral formation is characterized by the induction of the homeotic genes to specify the organ identities in the flower (Endress and Matthews, 2006). Many genes played significant roles in

* Correspondence: e.mohajelkazemi@tabrizu.ac.ir

This work is licensed under a Creative Commons Attribution 4.0 International License.

floral development. Sobral and Costa (2017) found that many proteins like GMM13, SQUA, AGL12, AGL17, AGL6, TM3, StMADS11, FLC, AGL15, AG, DEF/GLO, and AGL2 have roles in floral development. Floral organ differentiation is controlled by transcription factors, containing BELLRINGER (BLR), ETTIN (ETT), JAGGED (JAG), and REPRESSOR OF GIBBERELLIC ACID (RGA) (Roberts and Roalson, 2017).

Recent research has shown that shoot apical meristem (SAM) conducted signals to the cells on the adaxial surface of the lateral primordium, which was different from the signal produced for cells on the abaxial surface (Truskina and Vernoux, 2017). As a result, the transmission of signals to the new lateral petal primordium determined the orientation of adaxial–abaxial polarity and cell specification (Truskina and Vernoux, 2017). Cells on the abaxial surface of the lateral primordium do not receive every signal that is received by cells on the adaxial surface, thereby remaining in their default state. In this regard, the function of the *PHANTASTICA (PHAN)* gene, the homologue of ASYMMETRIC LEAVES 1 (AS1) in Arabidopsis and ROUGH SHEATH2 (RS2) in maize, is significant for adaxial–abaxial symmetry (Waites and Hudson, 1995; Timmermans et al., 1999; Byrne et al., 2002).

Adaxial–abaxial polarity specification involves antagonistic interaction of abaxial and adaxial genes (Fukushima and Hasebe, 2014; Machida et al., 2015). Abaxial factors KANADI1 (KAN1) and KAN2 repress the *AS2* gene (Wu et al., 2008). Adaxial factors AS1-AS2 directly inhibit the abaxial gene ETTIN /AUXIN RESPONSE FACTOR3 (ETT/ARF3) and indirectly inhibit ETT/ARF3 and ARF4 through tasiR-ARF small RNA (Iwasaki et al., 2013; Husbands et al., 2015). Moreover, the AS1–AS2 protein complex functions in the regulation of the proximal–distal formation by directly inhibiting *KNOX1* homeobox genes (BP, KNAT2) that are expressed in the shoot apical meristem (Ikezaki et al., 2009).

The *PHAN* gene was identified in *Antirrhinum majus* belonging to the order of Lamiales by Waites and Hudson (1995). It was afterward revealed that this gene was effective in the formation and establishment of adaxial–abaxial symmetry in leaves, bracteole, sepals, and petals (Bustamante et al., 2016). The *PHAN* gene encoded a transcription factor belonging to the MYB family, which was effective in the expression of genes involved in the development of these organs (Bustamante et al., 2016). PHAN expression resulted in the formation of the lateral organ adaxial surface and was also effective in determination of the proximal–distal axis, which occurred through suppression of the *KNOX1* genes (Byrne et al., 2002; Burko et al., 2013). Studies have highlighted that *AS1* and *2* genes were effective in the formation of petal and sepal boundaries, size determination, and the orientation of petals through floral development (Xu et al., 2008). Gubert et al. (2014) determined that the *AS1* gene was suppressed by *KNOX* genes associated with AS2. Additionally, through regulating the expression of the gene network in receptacles and peduncles, AS1 also resulted in the regulation of floral organ primordia (Gubert et al., 2014). Ge et al. (2014) reported that in *Medicago phan* mutants there was a lack of composite leaves while having short and ectopic petioles (Ge et al., 2014). Leaf primordium initiation in the peripheral zone of SAM required the *KNOX1* gene, AS1/PHAN, and auxin phytohormone activity (Burko et al., 2013). Ge et al. (2014) reported that the marginal zone formation of *Medicago* leaves in *phan* mutants depended on the pin10 protein and the activity of auxins as well. On the other hand, *PHAN* mutants were defected with a cauliflower–like inflorescence in *Medicago* (Ge et al., 2014). The data also revealed that PHAN had a low expression in shoot apical meristem, whereas its expression in the leaf primordium was significantly increased (Ge et al., 2014). Tattersall et al. (2005) revealed that *PHAN* mutants in *Pisum sativum* had leaves that were defective in adaxial–abaxial symmetry as well as ectopic petioles.

Waites and Hudson (1995) highlighted that *Antirrhinum majus PHAN* mutants had smaller leaves than the wild type. In fact, through the primordium initiation stage, no difference was evident between the leaves of mutants and wild types (Waites and Hudson, 1995). As the tissues expanded and flattened, however, differences became evident (Waites and Hudson, 1995). Other effects on the *Antirrhinum majus* PHAN mutants can be observed from changes in the corolla. The 3 lower petals had different orientations and the petal lobes were different in PHAN mutants. In fact, parts of the lobes were expanded while other parts had not grown (Waites and Hudson, 1995). In the *Arabidopsis* AS1 mutant, petal and sepal abscission was delayed and abscission zones displaced (Gubert et al., 2014). Moreover, in the senescence stage, changes in sepal chlorophyll content were evident as well. In this regard, the zones of the proximal-distal axis remained green while the chlorophyll content degenerated in other zones (Gubert et al., 2014).

Although morphological studies have been effective in understanding floral organogenesis, during the last decade these studies have been complemented with molecular analysis (Roberts and Roalson, 2017). Flower development is carried out with adaptability to pollinators and breeding systems. In this regard, a change in the flower pattern during development is considered a plant evolution (Lowenstein et al., 2019). Among the various types of flower symmetry, the actinomorphic and zygomorphic symmetry are significant. Moreover, zygomorphic flowers

are considered as the factors of progression in interaction between the plant and pollinators (Lowenstein et al., 2019). In the Gesneriaceae family, evolution and alteration of zygomorphic flowers to actinomorphic was observed in some cases, indicating the importance of molecular and structural studies of flower organ formation in this family (Weber et al., 2013). Undoubtedly, the economic value of African violet as an ornamental plant is considerable to everyone who commercially cultivates this plant. Not only are the flowers of African violet attractive, but it also has the ability to produce flowers throughout the year, making this plant important commercially. During the last decade, interest in gene identification as well as bioinformatics analysis for the purpose of discovering biological phenomena has increased dramatically (Kalpana et al., 2018). Such information can be useful in identifying and understanding how to improve developmental processes as well as predicting certain genetic abnormality (Kalpana et al., 2018). The aim of the present research was to identify the *PHAN* gene as well as evaluating its expression in order to determine gene expression pattern in floral developmental stages, phylogenetic analysis, predicting protein structural characteristics, and PHAN protein localization in the African violet for the first time. In present research we have focused on PHAN protein properties because there was insufficient information about this protein in African violet.

2. Materials and methods

2.1. Sample preparation

African violet samples were collected in 3 independent biological replicates from the greenhouse of the Genetics and Biotechnology Institute of Tabarestan, under photoperiod conditions of 16 h (light)/8 h (darkness) at 25 °C. Leaf and petal samples were immediately frozen with liquid nitrogen for molecular evaluation.

2.2. Reverse transcription

To design the primer, partial *PHAN* gene sequences in the Gesneriaceae family with the accession numbers EU330390.1, EU330392.1, and EU330394.1 were extracted from the NCBI database. The sequences were aligned using MEGA 7 software; the conserved sequences were considered in the primer design pattern using Primer3 plus software and an Oligo analyzer. Total RNA was extracted from the samples using the RNX plus isolation reagent (SINACLON Co.). Afterwards, the quality and quantity of the extracted RNA was determined respectively using 1.5% agarose gel and NanoDrop 2000 (USA) at 260 nm (OD 1.8). In order to eliminate the possible contamination of the extracted solution with DNA, DNase I enzyme was used (Thermo Fisher Scientific, USA). cDNA synthesis was performed based on the instructions provided with the Thermo Fisher Scientific Kit (USA). The PCR (T100 Thermal Cycler, Bio-Rad Co., USA) thermal condition was 95 °C (5 min), 95 °C (30 s), 58 °C (1 min), 72 °C (1 min) and 72° C (7 min) in 35 cycles (Peykari and Zamani, 2019). The PCR product was visualized utilizing 1.5% agarose gel electrophoresis. The PCR product (756 bp cDNA) was obtained from 1 pair primer (Table 1). Sequencing was performed by Macrogen Co., Korea based on the Capillary Electrophoresis Sequencing (CES) automation system. The cDNA sequences were read in 4 replicates. The best quality (with no errors) has been chosen for submitting in NCBI.

2.3. Quantitative Real Time-PCR

In order to evaluate gene expression, young leaves (control) and petals were collected in 3 different developmental stages (lateral floral bud, 5 mm buds, and open flowers). Quantitative Real Time-PCR (qRT-PCR) was performed using the CFX96 Real-Time System (Bio-Rad Co., USA) and The SYBR GREEN MASTERMIX Kit (Thermo Fisher Scientific, USA). The *ACTIN* gene (AB596843) was used as the house keeping gene so that the expression of the *PHAN* gene could be normalized (Stephan et al. 2019). The total volume of the samples (10 µL) included forward and reverse primers (0.3 µL), master mix (5 µL), cDNA (2 µL) and water DNase free (2.7 µL). The qRT-PCR thermal condition was 95 °C (10 min), 95 °C (15 s), 60 °C (1 min) and 72 °C (15 s) in 40 cycles. A negative control (without cDNA) was considered for every reaction. The data was

Table 1. Sequences for primers used in PCR and real-time PCR.

Reaction	Target gene Primer	Sequence 5'–3'
PCR	PHAN-F	TGGCGACCTGAAGAGGAT
	PHAN-R	TTCTCCCTCCTTCGGCAT
Real-time PCR	ACTIN-F	TTGATTCTGGTGACGGGGTG
	ACTIN-R	AGCAAGATCCAACCGCAGAA
Real-time PCR	PHAN-F	TCGGGAACAGAAGGAGAACA
	PHAN-R	GGGTGCTGGTGCATCTGTAT

analyzed using the GeneX software according to the ΔΔCt method. The difference between various means (Tukey test) was performed using SPSS 20 software (IBM Corp., Armonk, NY, USA) and the significance level was set at $P \leq 0.05$ (Stephan et al., 2019).

2.4. Bioinformatics analysis

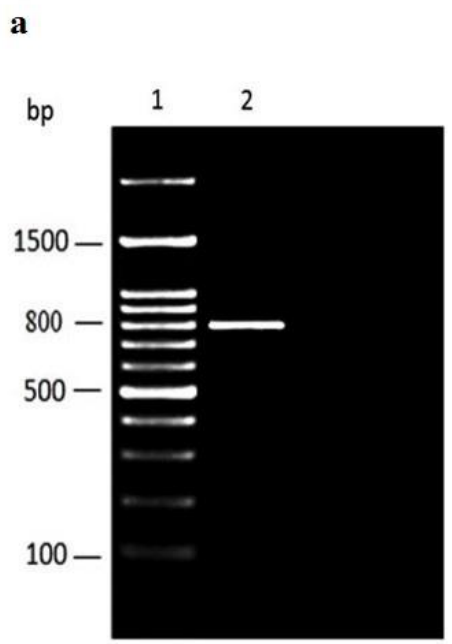
The partial cDNA sequence was translated into protein using Mega7 software and NCBI-Blast was used to confirm the PHAN sequence. In order to undertake the phylogenetic analysis, the NCBI database was searched using PHANTASTICA and ASSYMETRIC LEAVES1 as keywords for other genera. Afterwards, the phylogenetic tree was constructed using Mega7 software (Tamura et al. 2013). In order to predict protein localization, the LOCtree3 and DeepLoc1.0 online software were used (Almagro Armenteros et al. 2017; Goldberg et al. 2014). Since the structure did not exist in databases such as RCSB, the secondary and 3D structure of the PHAN protein were obtained using I-TASSER, COACH, and Phyre2 software (Kelley et al. 2015; Yang et al. 2015). Prediction of the physiochemical characteristics and motif sequences were respectively performed using Pfam, ProtParam, and PROSITE software (Kumar et al. 2008; Gasteiger et al. 2003; Finn et al. 2013).

3. Results

3.1. *PHAN* gene identification and expression pattern

Confirmation of *PHAN* gene amplification in *Saintpaulia ionantha* was obtained by loading the PCR product on 1.5% agarose gel and 756 nucleotide bands visualized (Figure 1a). Sequencing of the partial CDS of the *PHAN* gene in *Saintpaulia* revealed that this gene contained 756 bp (Figure 1b) and the protein was encoded with 252 amino acids. After primary evaluations, the gene was submitted to the NCBI database (MK54207.1 accession number).

To reveal the expression pattern of the *PHAN* gene, gene expression was evaluated during different developmental stages (Figures 2a–2d). The *PHAN* gene expression level in lateral buds at the petal initiation stage was significantly different from control samples ($P \leq 0.05$). At this stage the expression level was 4.5 times higher than that of the control sample. *PHAN* gene expression at stage 2 also showed a significant difference compared to the control sample ($P \leq 0.05$). PHAN gene expression in stage 2 was 2 times higher than that of open flowers (stage3). In contrast to stage 1 and 2, there was low expression of the *PHAN* gene in open flowers. However, there was no significant difference as compared to the control sample ($P \geq 0.05$).



b

```
1   CGT GCT TAC GTT AAG CAA TAT GGC CCT AGA GAT TGG CAC CTC TTG
46  CCA GAG CGC ATG AAC AAA CCT CTG AAC AGG GAT CCA AAA TCC TGC
91  TTA GAA AGG TGG AAG AAC TAT CTC AAA CCC GGA ATC AAA AAA GGT
136 TCC CTA ACT GAA GAG GAG CAA CAT CTT GTT ATC CAT CTT CAA TCC
181 AAG CAC GGC AAC AAA TGG AAA AAA ATA GCG GCT GAA GTT CCG GGG
226 CGA ACA GCC AAG AGA CTT GGA AAG TGG TGG GAA GTT TTC AAA GAG
271 AAA CAA CAT CGG GAA CAG AAG GAG AAC ACC AAG ATT ATC GAA CCG
316 GTT GAA GAG GGT AAA TAC GAC CAT ATA CTG GAG ACC TTT GCA GAA
361 AAG CTT GTA AAA GAA CGA GCC ACT GCC ACC CCA GGG ATT AAT ATG
406 GCA ACT TCA AAT GGT GGT TTT CTG CAT ACA GAT GCA CCA GCA CCC
451 TCT GCA GCA TTG CCG CCG TGG CTT GCT AGT TCT AGC ATG AAT ACG
496 TCT ACG AGG CCA CCA TCT CCA TCT GTG ACT TTA AGT CTC TCC CCC
541 TCA ACT ATT ACT CCC ACA CCA CCA ATA CCA TGG TTG CAA ACT GAT
586 AGA TTA GTT GAT AAC AAT ACC CCC ACA CAC GGT TTA AGC AAT TTG
631 CCC TGT CTT GGG ATG GCT CCT CTA TCT GGA GAG AAT CCA TTA CTT
676 GTG GAA CTC ATG GAA TGC TGC AGG GAT TTG GAA GAA GGG CAT CAT
721 GCT TGG GCC ACG CAT AAG AAA GAA GCA GCC TGG AGA
```

Figure 1. Identification of partial CDS of *PHAN* gene in *Saintpaulia ionantha*. a, PCR product of partial CDS of PHAN using relative species primers, 756 bp was determined. 1: Ladder 100, 2: PCR product. b, The sequence of the *PHAN* gene.

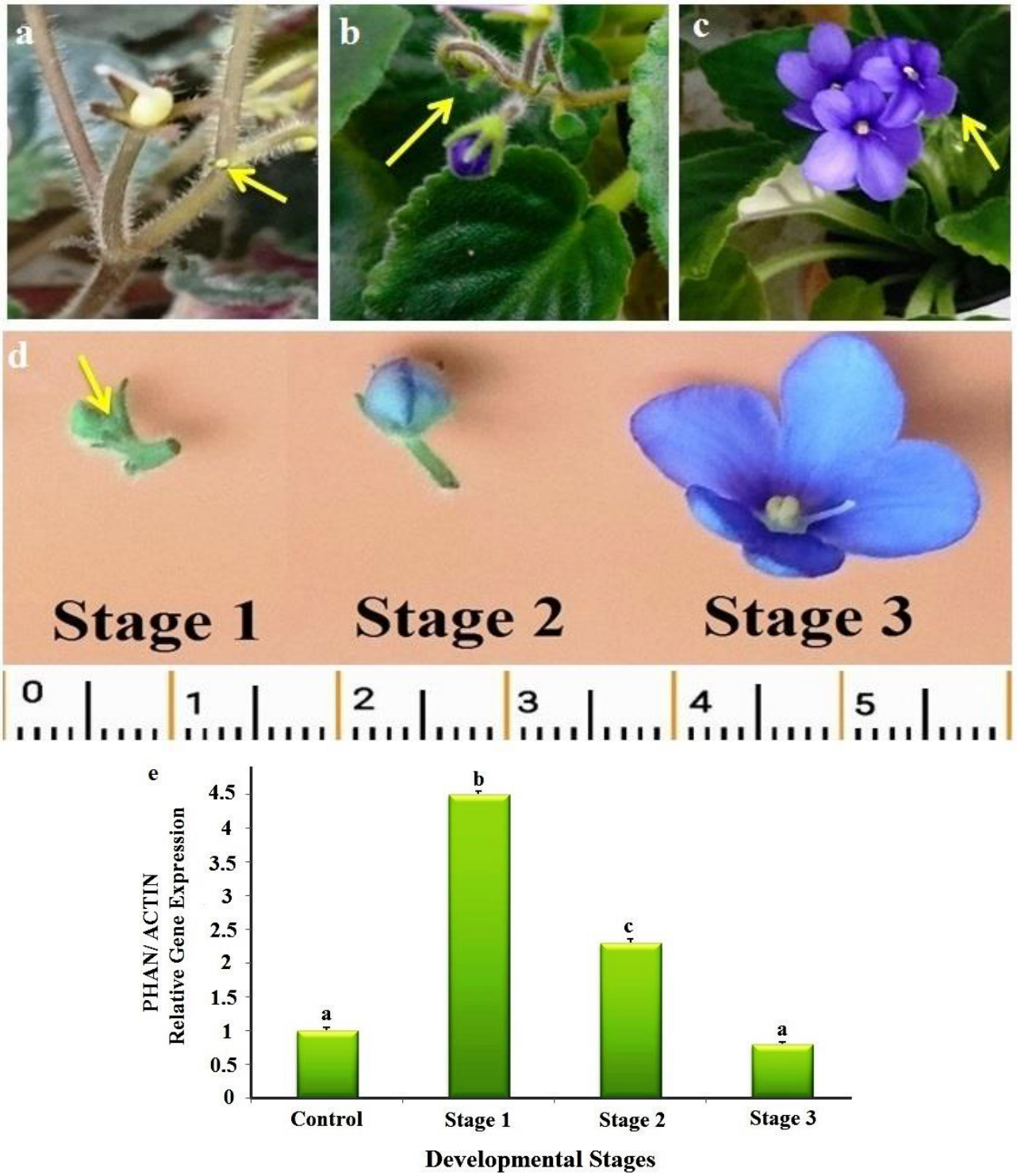


Figure 2. Changes in *PHAN* gene expression level at different petal developmental stages. a, Stage 1, b, Stage 2, c Stage 3, d, lateral bud, 5 mm flower and open flower. The yellow arrows represent the areas sampled. e, Evaluation of *PHAN* gene expression at 3 developmental stages in *Saintpaulia ionantha*. Different letters in each column indicate significant difference at $P \leq 0.05$ level. Error bars represent standard deviation (SD). Data are the mean ± SD of 3 replicates. *PHAN* gene expression level was high in the lateral buds, but lower in the mature flower. Young leaves were used as control samples.

The results indicate high levels of *PHAN* gene expression at stage 1 with lowest levels at stage 3 (Figure 2e).

3.2. Characterization of the primary structure of the PHAN protein

Protein blast results revealed that *Streptocarpus glandulosissimus* had more than 95% identity with the predicted PHAN protein in the Gesneriaceae family (Figure 3). The ProtParam software indicated that the partial CDS of *PHAN* gene sequence in *Saintpaulia ionantha* encoded 252 amino acids with a molecular weight of 28474 Da. PHAN protein sequence alignment indicated highly conserved amino acids and identity in the Gesneriaceae family (Figure 3a). Tryptophan (W) was highly conserved at the DNA binding site domain in the N-terminal region of the protein sequence (Figure 3a). Moreover, the Pfam software predicted a domain belonging to the MYB protein family (Figure 3c). Also, the presence of a Helix-Loop-Helix (HTH) MYB-type motif in the predicted protein sequence illustrated that this motif existed at the N-terminal of the PHAN protein, allowing the protein to bind the DNA (Figure 3d). The theoretical isoelectric point (PI) of the PHAN protein was estimated as 9.09. The most common amino acids were Leu (10%), Glu and Lys (8.7%), while Phe (1.2%), Tyr (1.6%), and Cys had the lowest occurrence among regular amino acids. The total number of negatively and positively charged residues were 29 and 35, respectively. PHAN protein characteristics indicated that the half-life (in vivo) of this protein in the bacteria and yeast evaluated was approximately 2 min. The instability, aliphatic index, and grand average of hydropathy (GRAVY) were 51.9, 71.6, and 0.75, respectively (ProtParam online software).

3.3. Phylogenetic analysis

To conduct molecular phylogenetic analysis, 50 plant species belonging to different families were selected from the NCBI database. *Selaginella* was considered as the out group. Two main roots were evident in the tree. Phylogenetic analysis revealed that members of different plant families were separated and grouped well. The species related to Gesneriaceae, Bignoniaceae, Rosaceae, Brassicaceae, Fabaceae, Oleaceae, Vitaceae and Solanaceae families were grouped in 1 clade. Indeed, these results revealed the presence of conserved sequences in the PHAN protein for each family (Figure 4). The Pedaliaceae, Bignoniaceae, Phyrmaceae and Gesneriaceae family belong to the order of Lamiales and were located close to each other; however, the Oleaceae family was grouped in the lower branch of the tree compared to families in the same order. The *Saintpaulia ionantha* PHAN protein sequence along with the cofamily species was well–separated in the Gesneriaceae family, indicating the phylogenetic value of this protein. This classification indicated the high similarity of the PHAN protein to relative plants. Indeed, the Pedaliaceae and Bignoniaceae belonging to the Lamiales order can be considered as the closest families to the Gesneriaceae family (Figure 4).

3.4. Prediction of PHAN protein secondary and 3D structures and subcellular localization

In order to predict the PHAN protein secondary and 3D structures, I-TASSER, and Phyre2 software were utilized. The protein secondary structure in *Saintpaulia ionantha* basically consisted of α-helix (45%) and disordered parts (20%). The predicted secondary structure was 78% accurate (Phyre2 software). The B factor (I-TASSER software) was derived from a comparison of the proteins existing on the PDB database, showing the inherent thermal mobility of residues in the protein (Figures 5a and 5b). The results indicated the vast expanse of the α-helix and random twists in the predicted PHAN protein structure. Ten templates with the highest probabilities were derived from the PDB library using the I-TASSER software as targets (Table 2). Templates with Z-Scores above 1 were used, which represented the highest significant values. The final predicted model for PHAN protein obtained through the application of I-TASSER and Phyre2 software is shown in Figure 6. The most probable organism and protein family were *Trichomonas vaginalis* and the MYB3 protein family. Moreover, ligand binding sites were predicted using COACH software. Based on COACH results, the predicted PHAN protein included 4 templates based on the C-Score (0–1). The C-Score indicated the confidence score of the prediction. The results highlight that PHAN protein has the ability to bind to nucleic acids. The binding site of protein with higher C-Scores included the amino acids of 10(R), 11(D), 12(W), 13(H), 31(L), 35(K), 64(N), 65(K), 66(W), 67(K), and 82(G) (Table 3).

The DeepLoc1.0 and LOCtree3 software was used for PHAN protein intracellular localization. Results showed that the PHAN protein was located in the nucleus and nonsecretory pathways (Figure 7).

4. Discussion

4.1. *PHAN* gene expression during lateral organ development

In the present study, the expression pattern of the *PHANTASTICA* gene during the petal development was investigated and in silico analysis of the *PHAN* gene in *Saintpaulia ionantha* conducted. Identification of the *PHAN* gene and its effects on the adaxial–abaxial symmetry of petals in African violets were identified for the first time. Highest similarity (over 95%) was found between *Saintpaulia* and *Streptocarpus glandulosissimus* belonging to the Gesneriaceae family. Thepartial CDS of the *PHAN* gene encodes 252 amino acids in *Saintpaulia*. Sequence alignment in related plants showed similarity and high conservation of amino acids. On the other hand, this gene

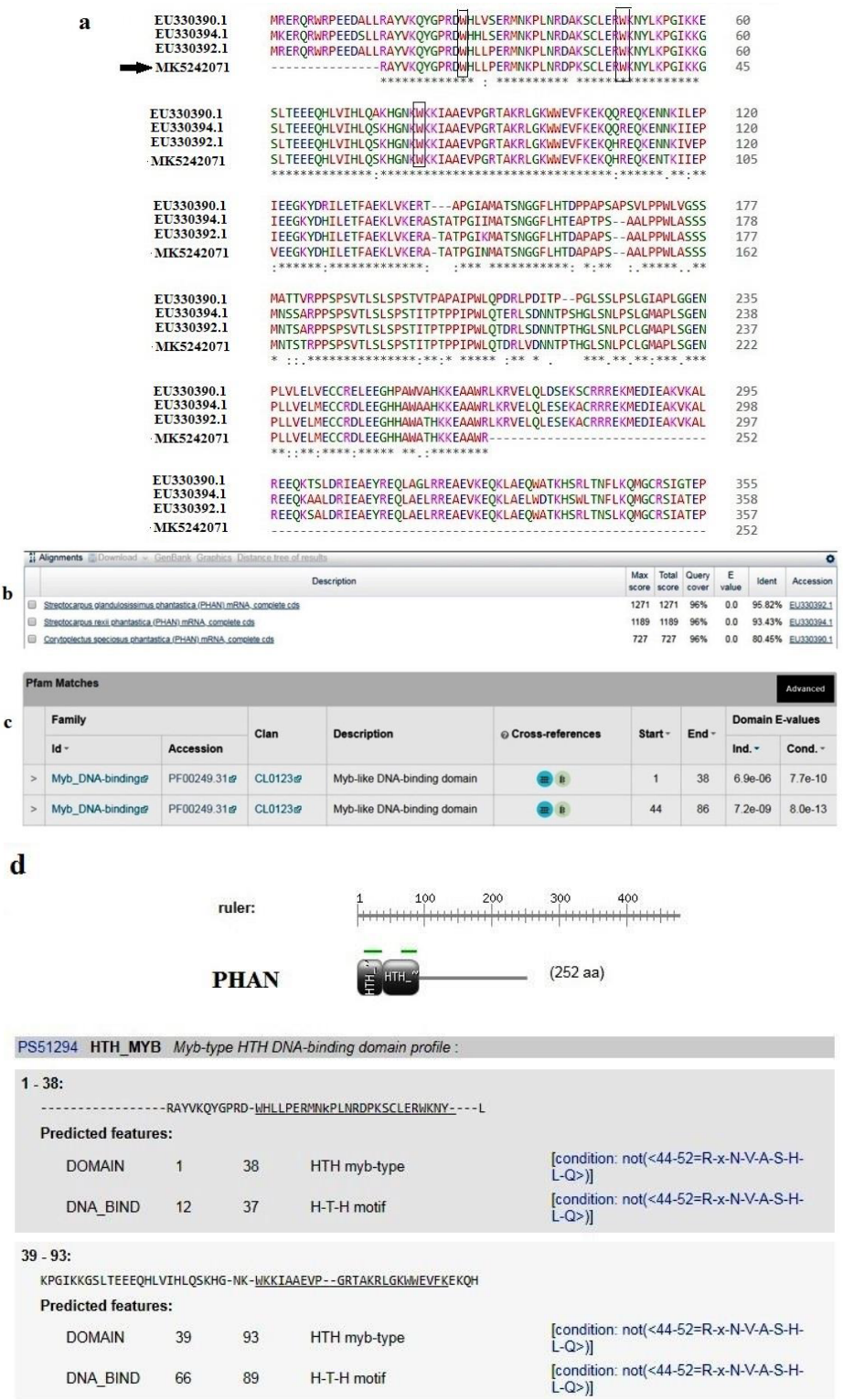


Figure 3. Alignment and comparison of predicted PHAN protein sequence. a, Highly conserved amino acids were determined using Clustal W software. Tryptophan (W) is marked with a black rectangle, and is a highly conserved amino acid in the HTH motif at the N-terminal of the protein. (*) Conserved amino acids, (:) Strongly similar properties, (.) Weakly similar amino acids. *Sainpaulia* is represented with an arrow. b, PHAN protein blast in the Gesneriaceae family indicated an almost 80–95% similarity between the relative genera. c, Pfam software illustrated that the PHAN protein belongs to the MYB family. d, Presence of HTH-MYB-Type domain and DNA binding capability in the N terminal of the PHAN protein sequence.

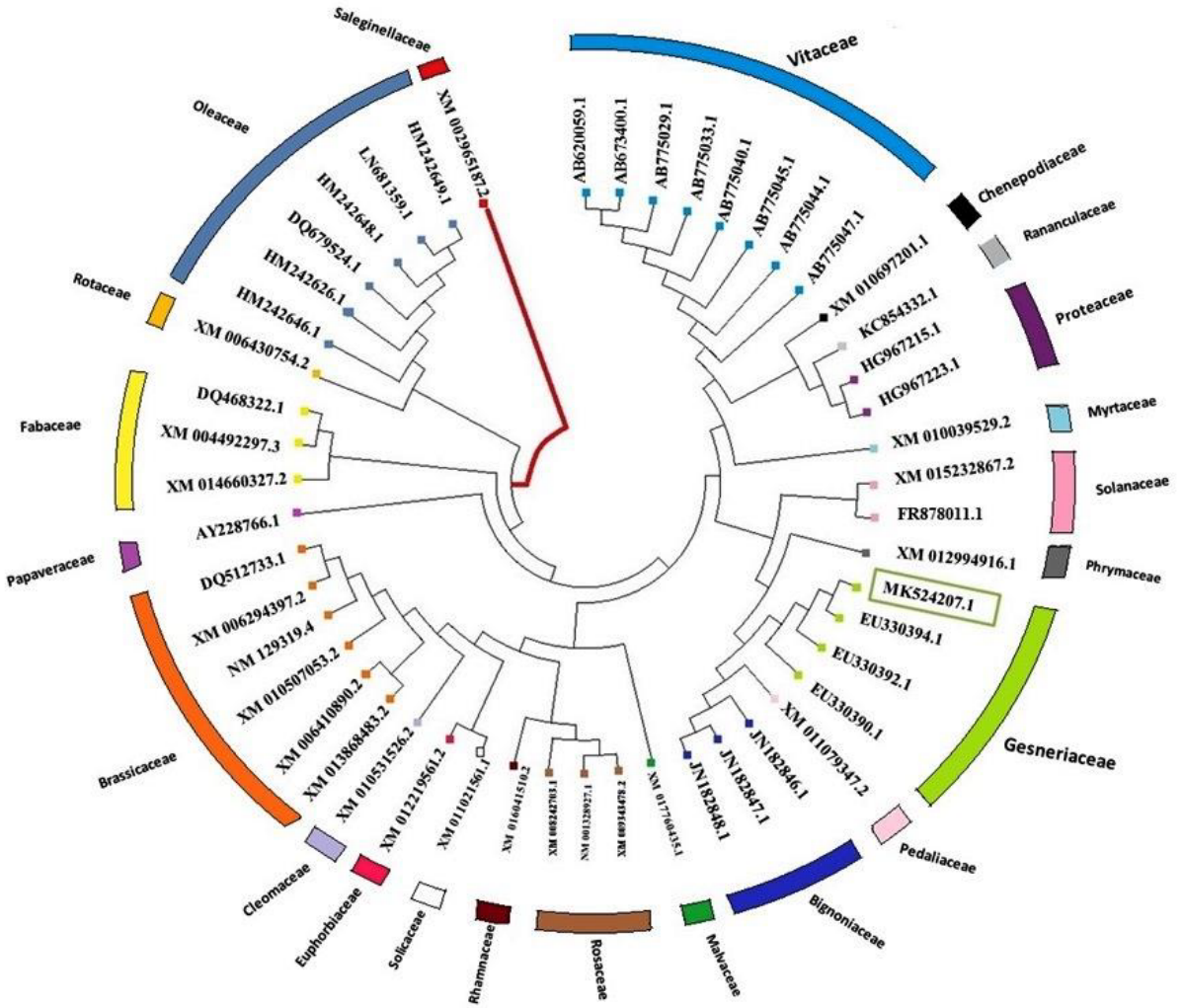


Figure 4. PHAN protein phylogenetic analyses in different families. The tree was constructed from the Neighbor Joining method using the Mega7 software based on homology between the sequences of 50 species. The 2 main roots in this tree have been highlighted in red which respectively included the primary and higher plants. The green square shows *Saintpaulia ionantha*. The species relating to each family have been marked with a similar color. The genera were listed from the upper branch to the lower one, respectively: *Cayratia trifolia*, *Cayratia japonica*, *Tetrastigma*, *Cyphostemma*, *Cissus*, *Vitis*, *Parthenocissus*, *Leea*, *Beta*, *Aquilegia*, *Leucadendron stellare*, *Leucadendron xanthoconus*, *Eucalyptus*, *Solanum*, *Nicotiana*, *Erythranthe*, *Streptocarpus inonanthus, Streptocarpus rexii*, *Streptocarpus glandulosissimus*, *Corytoplectus*, *Sesamum*, *Distictis*, *Bignonia*, *Dolichandra*, *Gossypium*, *Pyrus*, *Malus*, *Prunus*, *Ziziphus*, *Populus*, *Jatropha*, *Tarenaya*, *Brassica*, *Eutrema*, *Camelina*, *Arabidopsis*, *Capsella*, *Cardamine*, *Eschscholzia*, *Vigna*, *Cicer*, *Medicago*, *Citrus*, *Fraxinus xanthoxyloides*, *Fraxinus uhdei*, *Fraxinus excelsior*, *Ligustrum*, *Hesperelaea*, *Olea*, *Selaginella*.

encodes almost 360 and 357 amino acids in *Streptocarpus rexii* and *Antirrhinum*, respectively (Waites and Hudson, 1995). The full length of the PHAN homologue gene (AS1) in *Arabidopsis* and *Brassica* was respectively 1104 and 1080 bp which encoded 367 and 357 amino acids (Gubert et al., 2014). Studies have determined that adaxial–abaxial and proximal-distal symmetries were strongly effective in the formation of plant axis and morphogenesis (Tian et al., 2018). On the other hand, in *Arabidopsis*, *AS1/PHAN* and *AS2* genes – as a complex – were involved in the formation of the adaxial surface of leaves (Yang et al., 2018). From the present research it was determined that during petal initiation and flower maturation, gene expression levels were high and low respectively. As the petal adaxial–abaxial symmetry was completed, expression of the *PHAN* gene was significantly reduced compared to previous stages ($P \leq 0.05$). Moreover, the results with regards to the role of the *PHAN* gene in adaxial–abaxial polarity formation indicated that this gene was also effective in *Saintpaulia* petals development. Bustamante et al. (2016) highlighted that *PHAN* gene expression increased in the margin zones of leaves and petals primordia whereas in the mature stage

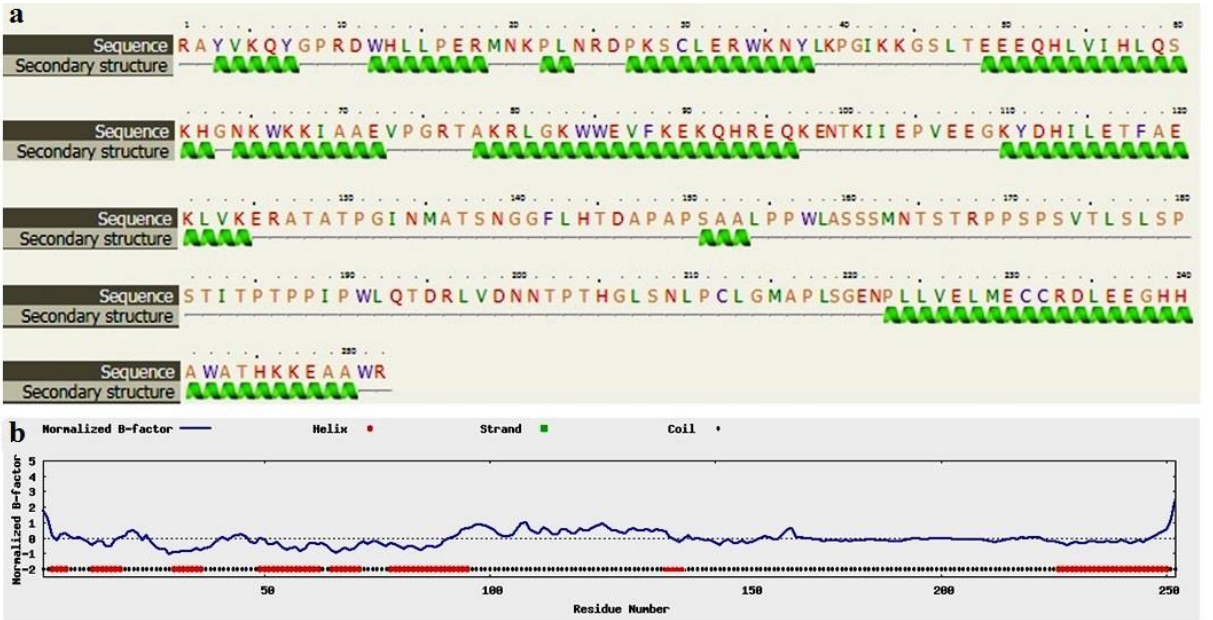


Figure 5. The secondary structure of predicted PHAN protein. a, The secondary structure was predicted using the Phyre2 software. The green color is indicative of the α-helix. b, Normalized B factor was predicted using I-TASSER software. The B factor indicated the inherent thermal mobility of atoms/residues in proteins, which was derived by comparing proteins which exist in the PDB database. The red lines are reflective of the α-helix.

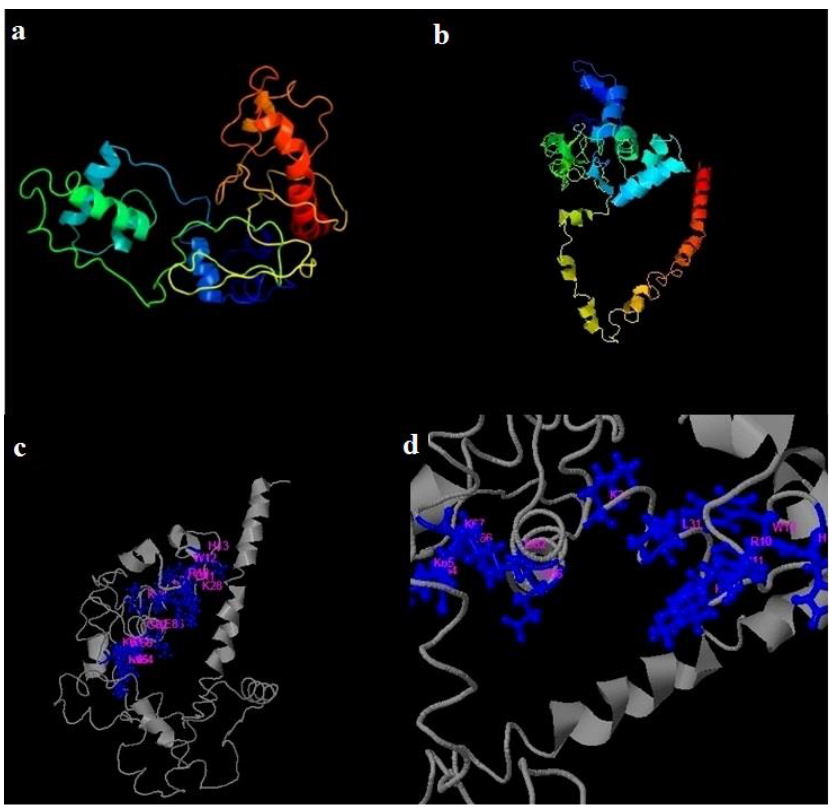


Figure 6. The predicted 3D structure of PHAN protein. a, The overall model of the 3D structure predicted by Phyre2, using the MYB protein family as model. b, An overview of the PHAN protein 3D structure constructed by I-TASSER software using the MYB3 protein as template. c-d, The ligand binding site was identified from the COACH software. Ligands were labeled in blue and the binding sites are highlighted in pink (The binding site amino acids are listed in Table 3).

Table 2. Top 10 threading templates used by I-TASSER software for PHAN protein.

Rank	PDB Hit	Organism	Iden1	Iden2	Cov	Z-score
1	3zqcA	*Trichomonas vaginalis* MYB3	0.12	0.22	0.98	1.59
2	1h89C	Homo sapiens	0.21	0.28	0.79	1.93
3	3zqcA	*Trichomonas vaginalis* MYB3	0.34	0.17	0.41	1.87
4	1h89C	Homo sapiens	0.12	0.28	0.96	1.65
5	3zqcA	*Trichomonas vaginalis* MYB3	0.34	0.17	0.42	2.63
6	5gmkC	*Saccharomyces cerevisiae* RNA binding protein	0.15	0.23	0.98	3.00
7	1h89C	Homo sapiens	0.30	0.28	0.35	2.19
8	6bk8S	*Saccharomyces cerevisiae* RNA binding protein	0.19	0.17	0.69	1.46
9	3zqcA	*Trichomonas vaginalis* MYB3	0.34	0.17	0.42	1.70
10	3zqcA	*Trichomonas vaginalis* MYB3	0.12	0.22	0.98	2.42

Ident1: The percentage sequence identity of the templates in the threading aligned region with the query sequence. Ident2: The percentage sequence identity of the whole template chain with query sequence. Cov: The coverage of the threading alignment which is equal to the number of aligned residues divided by the length of the query protein. Z-score > 1 good alignment.

Table 3. Top 4 threading templates used in the COACH software for PHAN protein.

Rank	C-score	Cluster size	PDB Hit	Organism	Ligand Name	Consensus binding residues
1	0.41	24	3zqcA	*Trichomonas vaginalis* MYB3	Nuc.Acid	10,11,12,13,31,35,64,65,66,67,82
2	0.31	18	1h89C	*Homo sapiens* MYB	Nuc.Acid	10,11,12,13,28,31,35,64,65,66,67,82,86
3	0.14	9	1h89C	*Homo sapiens* MYB	Nuc.Acid	28,29,32,33,37,44,79,82,83,87,88
4	0.01	1	3zqcA	*Trichomonas vaginalis* MYB3	Nuc.Acid	24,25,26,29,32,33,79,83,84,87,90

Cluster size: The total number of templates in a cluster. PDB Hit: Templates. C-score: Confidence score of the prediction.

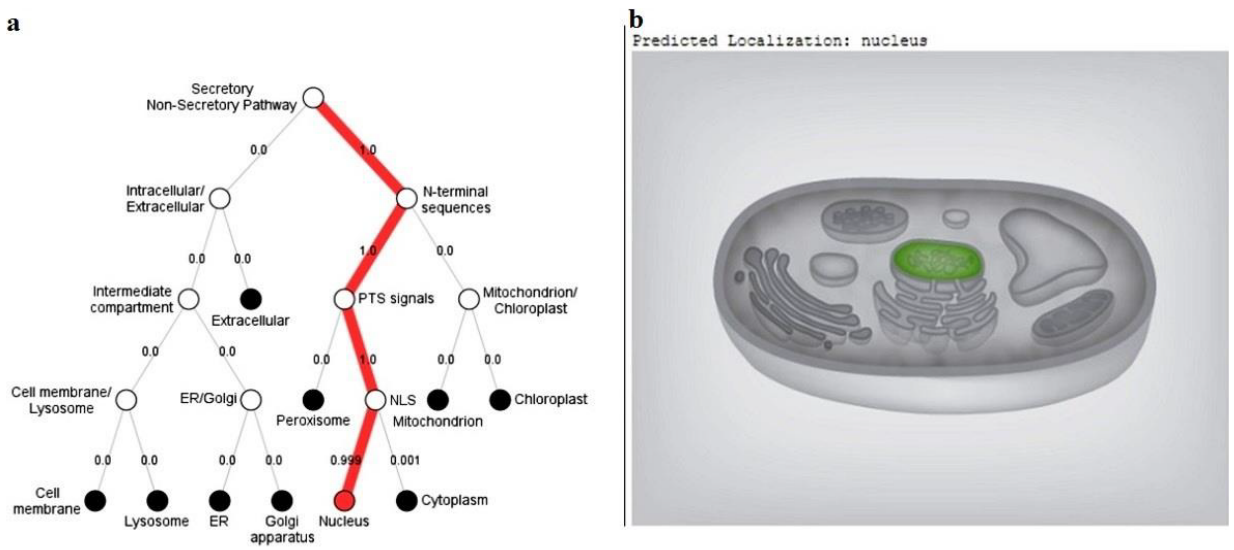


Figure 7. The predicted intracellular localization of PHAN protein by DeepLoc1 (a) and LocTree3 (b) software. The results show that PHAN protein was located in the nucleus.

expression was reduced. The results of present study are consistent with these findings.

Petal primordium formation requires processes such as division, expansion and cellular differentiation which are controlled by the gene network (Huang and Irish, 2016). Hugouvieux et al. (2018) reported that the formation of lateral organs adaxial–abaxial symmetry during floral meristem initiation, a group of cells existing on the peripheral zones are involved in lateral organ formation and make them distant from the axis. Expansion of sepals and petals became evident through certain divisions (Hugouvieux et al., 2018). The adaxial surface, formed at the same time, created a change in the polarity of the division pattern (Byrne et al., 2002). However, the adaxial surface of petal primordium had limited growth compared to the abaxial surface. Studies have indicated a difference between division patterns of adaxial and abaxial surfaces in the petal primordium stage (Tian et al., 2018). Consequently, the cells differentiate and form the second whorl (Tian et al., 2018). Yang et al. (2018) determined that the *AS1* gene suppressed auxin phytohormone activity genes and improved the stability of the adaxial–abaxial symmetry (Yang et al., 2018). Binding to the promoter region of these genes and activating miR390, resulted in gene silencing which consequently involved DNA methylation (Yang et al., 2018).

PHAN gene expression in *Antirrhinum majus* revealed that PHAN mRNA exists in the early floral meristem initiation stage which is involved in lateral organ identity (Waites and Hudson, 1995). Waites and Hudson (1995) reported that the adaxial surface cells conducted signals which changed lateral organ division. Ko et al. (2008) highlighted that the salt bridge was not constructed between R2 and R3 MYB domain inan *AS1* mutant, which inhibited the DNA binding ability. Consequently, AS1 expression was absent in leaves and the adaxial–abaxial symmetry were not initiated (Ko et al., 2008). Sun et al. (2002) showed that the *AS1/PHAN* gene plays a role in cell division and differentiation. *AS1/PHAN* gene overexpression was detected in leaves proliferated domain (Sun et al., 2002). Immunofluorescence analysis showed that in *AS1* mutants, KNOX protein accumulated ectopically. Moreover, *AS101* and *AS144* mutants exhibited an interruption in vascular tissue formation and with early opening flowers being smaller in size (Sun et al., 2002). Petal elongation was delayed in *PHAN* mutants, showing that the *AS1/PHAN* gene plays a role in many developmental processes (Sun et al., 2002). Ori et al. (2000) showed that mutation in the AS1 locus not only affected leaf elongation but also influenced petal flattening. *Antirrhinum PHAN* mutant had flat petals (Huang and Irish, 2016). Ikezaki et al. (2009) revealed that the *Arabidopsis AS1* mutant had small sepals, petals, and immature flowers. Overall, the results show that petal shape and size can be selected for the improvement of the value of ornamental flowers (Azadi et al., 2016).

4.2. Analysis of PHAN protein characteristics

Based on the present study, the PHAN protein belongs to the MYB family. Additionally, the PHAN protein secondary structure was predicted and consists of a vast expansion of α-helix as well as random twists in *Sainpaulia inonantha*. The MYB protein family exists in plants, animals and even viruses. As the largest family of transcription factors it is involved in many plant developmental processes (Lai et al., 2013). Lou et al. (2009) analyzed the MYB1 protein structure and proved that the beta strand was absent in this sequence. On the other hand, studies conducted by Wei et al. (2012) regarding the MYB3 protein structure in *Trichomonas* indicated that a unique-Beta existed at the end of the protein secondary structure. Moreover, Ogata et al. (1994) reported that only 1 beta-strand existed in the *Mus musculus* MYB protein secondary structure.

In the MYB protein family, 2 distinct conserved domains of N-terminal (DNA binding ability) and C-terminal (regulatory domain) have been identified (Huang et al., 2013). The N-terminal region was divided into 4 protein classes: R1MYB, R1R2R3 MYB, R2R3 MYB, and 4R MYB. Among the classes, the largest class related to R2R3 has about 100 members in the gymnosperms (Huang et al., 2013). The function of the genes in this class has been determined in *Petunia*, *Antirrhinum*, *Arabidopsis*, *Oryza sativa*, *Gossypium*, and *Zea mays* (Ambawat et al., 2013). Based on the results of this study, the presence of the DNA binding domain and HTH motif in the N-terminal region of the *Saintpaulia inonantha* PHAN protein was identified as 1 of the members of the MYB protein family, indicating that the findings are consistent with that of Ambawat et al. 2013. Indeed, MYB transcription factors are classified based on DNA binding domain conservation (having HLH, HTH, Zinc finger, etc. domains) (Feng et al., 2018). The MYB family has 3 conserved repeats (R); with the HTH motif in R2 and R3 (Feng et al., 2018). Each R contained 50–53 amino acids and α-helix, which provided DNA binding ability (Xu et al., 2017). Xu et al. (2017) also highlighted that 3 amino acids of tryptophan (w) were present in the HTH motif domains, which conserved HTH motif stability (Feng et al., 2018). Moreover, the C-terminal functional domain of MYB protein family contained conserved amino acids, which divided the family into 22 subgroups (Xu et al., 2017). The findings of the present study confirmed the presence of the ligand

binding site, DNA binding domain and HTH motifs in *Saintpaulia inonantha* PHAN protein. The conserved tryptophan amino acid in the *Saintpaulia* PHAN protein sequence was determined using bioinformatics analysis. The results confirmed the PHAN protein nucleic acid binding ability in *Saintpaulia*. The MYB protein family contained a DNA binding site, nuclear localization signal (NLS), and an activation domain that controlled the target gene expression as a regulator (Feng et al., 2018). In African violets, the probable nucleus location of the PHAN protein was predicted which was consistent with previous results.

The molecular phylogenetic analysis indicated that *Saintpaulia* was grouped in the Gesneriaceae family along with other relative species. This information may be used to determine common ancestry in this family. Consequently, family groups can be used to determine the conservation of the PHAN protein sequence among different plant species. Although Gesneriaceae, Phrymaceae, Bignoniaceae, and Pedaliaceae families related to the Lamiales order were grouped close to each other, and the Oleaceae family was grouped distant from the coorder families. These results suggest that there is a distinction of the *PHAN* gene sequence or possibly the presence of different PHAN protein isoforms in the Oleaceae family (Du et al., 2009). During *PHAN* gene evolution, it has been determined that many of the MYB protein family subgroups have been formed by losing 1 of the R domains while their ancestor had all 3 R domains (Du et al., 2009). These domains originated from triplication of MYB protein sequence from primary eukaryotes (Wilkins et al., 2009). In contrast, some researchers have identified the phylogenic origin of this family with only 1 DNA binding domain (Du et al., 2009). Du et al. (2009) highlighted that the *Arabidopsis* MYB protein family consisted of 125 members.

The molecular phylogenetic analysis revealed that all eukaryotes contained at least 1 member of the MYB protein family (Wilkins et al., 2009). The first member, known as V-MYB, was identified in virus (Wilkins et al., 2009). It was suggested that MYB genes originated from vertebrate genes (Du et al., 2009). On the other hand, the first MYB gene in plants was identified as the *Zea mays* C1 gene (Wilkins et al., 2009). According to the *Populus* genome molecular study, the MYB family genes had 192 members that contained all 3 R-MYB domains. Moreover, 252 genes have been identified in *Glycine max* containing 3RMYB and 4R MYB (Ambawat et al., 2013). Ambawat et al. (2013) suggested that the splicing mechanism of the MYB protein family produced different transcripts. The MYB/AS1 transcription factor in *Oryza sativa* has created 3 transcripts while the homologous gene in *Arabidopsis* produced 4 transcripts (Ambawat et al., 2013). The R2R3 MYB domain sequence was highly conserved in plants and seemed to play a role in land plants ontogeny (Xu et al., 2015). However, the MYB DNA binding domain was different in various species (Ko et al., 2008). Ko et al. (2008) indicated that in the *Arabidopsis* AS1 protein, a salt bridge was formed between acidic-basic amino acids of the adjacent helixes (R2R3 MYB domain). Studies have identified that some amino acids involved in salt bridge formation (including arginine and glutamic acid) were highly conserved, which was consistent with the results of this study (Ko et al., 2008). In the AS1 mutant, by altering the W7 in the R2R3 domain, DNA binding ability was inhibited. Indeed, the conserved tryptophan amino acid played a significant role in MYB DNA binding ability at the N-terminal of AS1/PHAN protein (Gubert et al., 2014). Identification of the *Antirrhinum majus* PHAN protein indicated the presence of 56 conserved amino acid residues at the N-terminal of the MYB domain protein (Waites and Hudson, 1995).

5. Conclusion

The various shapes of African violet petals can be used to determine the importance of studying genetic control processes of flower development in this species. The results highlighted that the *Saintpaulia ionantha* partial CDS of *PHAN* gene contains 756 bp and 252 aa. Moreover, the *PHAN* gene expression level indicated that as the flower approached the late developmental stages (maturity), gene expression decreased significantly compared to the previous stages ($P \leq 0.05$). Bioinformatics analysis revealed that the tryptophan amino acid (W) can be considered as a conserved amino acid in the N-terminal DNA binding domain. Moreover, the presence of the HTH-MYB type motif in the N-terminal of the protein sequence highlighted that PHAN protein contained the ability of DNA binding. Prediction of the PHAN protein intracellular localization showed that the PHAN protein was soluble and located in the nucleus (in the nonsecretory pathway of the cell). The present study suggested that the *PHAN* gene may play a significant role in adaxial–abaxial symmetry formation at the petal initiation stage. Therefore, studying the *Saintpaulia PHAN* gene can help us to better understand the different forms of leaves and petals in ornamental plants. Finally, it is suggested that changes in *PHAN* gene expression and genetic engineering of ornamental plants are effective in producing new and various petal phenotypes.

Conflict of Interest

The authors report no conflict of interest and they are responsible for the content and writing of the article.

References

Aida R, Ohmiya A, Onozaki T (2018). Current researches in ornamental plant breeding. Breeding Science 68 (1): 1-2. doi: 10.1270/jsbbs.68.1

Almagro Armenteros JJ, Sonderby CK, Sonderby SK, Nielsen H, Winther O (2017). DeepLoc: prediction of protein subcellular localization using deep learning. Bioinformatics 33 (21): 3387-3395. doi: 10.1093/bioinformatics/btx431

Ambawat S, Sharma P, Yadav N, Yadav R (2013). MYB transcription factor genes as regulators for plant responses: an overview. Physiology and Molecular Biology of Plants 19 (3): 307-321. doi: 10.1007/s12298-013-0179-1

Azadi P, Bagheri H, Nalousi AM, Nazari F, Chandler SF (2016). Current status and biotechnological advances in genetic engineering of ornamental plants. Biotechnology Advances 34 (6): 1073-1090.

Burko Y, Shleizer-Burko S, Yanai O, Shwartz I, Zelnik ID et al.(2013). A role for APETALA1/fruitful transcription factors in tomato leaf development. Plant Cell 25 (1): 2070-2083.

Bustamante M, Matus J, Riechmann JL (2016). Genome-wide analyses for dissecting gene regulatory networks in the shoot apical meristem. Journal of Experimental Botany 67 (6): 1639-1648. doi: 10.1093/jxb/erw058

Byrne M, Simorowski J, Martienssen R (2002). ASYMMETRIC LEAVES1 reveals Knox gene redundancy in *Arabidopsis*. Development 129: 1957-1965.

Du H, Zhang L, Liu L, Tang X, Yang W et al. (2009). Biochemical and molecular characterization of plant MYB transcription factor family. Biochemistry 74 (1): 1-11. doi: 10.1134/s0006297909010015

Endress PK, Matthews ML (2006). First steps toward a floral structural characterization of the major Rosid subclades. Plant Systematics and Evolution 260 (1): 223-51.

Feng K, Xu ZS, Que F, Liu JX, Wang F et al. (2018). An R2R3.MYB transcription factor, OjMYB1, functions in anthocyanin biosynthesis in *Oenanthe javanica*. Planta 247 (2): 301-315.

Finn RD, Bateman A, Clements J, Coggill P, Eberhardt RY et al. (2013). Pfam: the protein families' database. Nucleic Acids Research 42 (1): 222-230.

Fukushima K, Hasebe M (2014). Adaxial–abaxial polarity: the developmental basis of leaf shape diversity. Genesis 52 (1): 1-18.

Gasteiger E, Gattiker A, Hoogland C, Ivanyi I, Appel RD et al. (2003). ExPASy: the proteomics server for in-depth protein knowledge and analysis. Nucleic Acids Research 31 (13): 3784-3788.

Ge L, Peng J, Berbel A, Madueno F, Chen R (2014). Regulation of compound leaf development by PHANTASTICA in *Medicago truncatula*. Plant Physiology 164 (1): 216-228.

Goldberg T, Hecht M, Hamp T, Karl T, Yachdav G et al. (2014). LocTree3 prediction of localization. Nucleic Acids Research 42 (1): 350-355.

Gubert C, Christy M, Ward D, Groner W, Liljegren S (2014). ASYMMETRIC LEAVES1 regulates abscission zone placement in *Arabidopsis* flowers. BMC Plant Biology 14 (2): 195-202.

Haston E, De Craene LP (2007). Inflorescence and floral development in *Streptocarpus* and *Saintpaulia* (Gesneriaceae) with particular reference to the impact of bracteole suppression. Plant Systematics and Evolution 265 (2): 13-25.

Hepworth J, Lenhard M (2014). Regulation of plant lateral-organ growth by modulating cell number and size. Current Opinion in Plant Biology 17 (1): 36-42.

Huang T, Irish VF (2016). Gene networks controlling petal organogenesis. Journal of Experimental Botany 67 (1): 61-68.

Huang W, Sun W, Lv H, Xiao G, Zeng S et al. (2013). Isolation and molecular characterization of thirteen R2R3-MYB transcription factors from *Epimedium sagittatum*. International Journal of Molecular Sciences 14 (2): 594-610.

Hugouvieux V, Silva CS, Jourdain A, Stigliani A, Charras Q et al. (2018). Tetramerization of MADS family transcription factors SEPALLATA3 and AGAMOUS is required for floral meristem determinacy in *Arabidopsis*. Nucleic Acids Research 46 (2): 4966-4977.

Husbands AY, Benkovics AH, Nogueira FT, Lodha M, Timmermans MC (2015). The ASYMMETRIC LEAVES complex employs multiple modes of regulation to affect adaxial–abaxial patterning and leaf complexity. Plant Cell 27 (2): 3321-3335.

Ikezaki M, Kojima M, Sakakibara H, Kojima S, Ueno Y et al. (2009) Genetic networks regulated by ASYMMETRIC LEAVES1 (AS1) and AS2 in leaf development in *Arabidopsis thaliana*: KNOX genes control five morphological events. Plant Journal 61 (1): 70-82.

Iwasaki M, Takahashi H, Iwakawa H, Nakagawa A, Ishikawa T et al. (2013). Dual regulation of ETTIN (ARF3) gene expression by AS1-AS2, which maintains the DNA methylation level, is involved in stabilization of leaf adaxial–abaxial partitioning in *Arabidopsis*. Development 140 (2): 1958-1969. doi: 10.1242/dev.085365

Kalpana K, Manjuvani S, Shoba K (2018). In silico comparative modeling of maturase K protein in *Cymbopogon martinii* plant. Research & Reviews: A Journal of Bioinformatics 5 (3): 30-36. doi: 10.1016/S1672-0229(08)60048-0

Kelley LA, Mezulis S, Yates CM, Wass MN, Sternberg MJ (2015). The Phyre2 web portal for protein modeling, prediction and analysis. Nature Protocols 10 (6): 845-855.

Ko ER, Ko D, Chen C, Lipsick JS (2008). A conserved acidic patch in the MYB domain is required for activation of an endogenous target gene and for chromatin binding. Molecular Cancer 7 (1): 77-79.

Kolehmainen J, Korpelainen H, Mutikainen P (2010). Inbreeding and inbreeding depression in a threathened endemic plant, the African violet (*Saintpaulia ionantha* ssp. grotei), of the East Usambara Mountains, Tanzania. African Journal of Ecology 48 (1): 576-587.

Kumar S, Nei M, Dudley J, Tamura K (2008). MEGA: a biologist-centric software for evolutionary analysis of DNA and protein sequences. Briefings in Bioinformatics 9 (1): 299-306.

Lai Y, Li H, Yamagishi M (2013). A review of target gene specificity of flavonoid R2R3-MYB transcription factors and a discussion of factors contributing to the target gene selectivity. Frontiers in Biology 8 (2): 577-598.

Lou YC, Wei SY, Rajasekaran M, Chou CC, Hsu HM et al. (2009). Solution NMR structure of the R2R3 DNA binding domain of Myb1 protein from protozoan parasite *Trichomonas vaginalis*. Nucleic Acids Research 37 (2): 2381-2394.

Lowenstein DM, Matteson KC, Minor ES (2019). Evaluating the dependence of urban pollinators on ornamental, non-native, and 'weedy' floral resources. Urban Ecosystems 22 (2): 293-302.

Machida C, Nakagawa A, Kojima S, Takahashi H, Machida Y (2015). The complex of ASYMMETRIC LEAVES (AS) proteins plays a central role in antagonistic interactions of genes for leaf polarity specification in *Arabidopsis*. Wiley Interdisciplinary Reviews-Developmental Biology 4 (1): 655-671.

Nhut DT, Trinh DB, Cuong D.M, Tung HT, Huy NP et al. (2018). Study on silver nanoparticles as a novel explant disinfectant for micropropagation of African violet (*Saintpaulia ionantha* H. Wendl.). Vietnam Journal of Biotechnology 16 (1): 1-10.

Ogata K, Morikawa S, Nakamura H, Sekikawa A, Inoue T et al. (1994). Solution structure of a specific DNA complex of the MYB DNA-binding domain with cooperative recognition helices. Cell 79 (2): 639-648. doi: 10.1016/0092-8674(94)90549-5

Ori N, Eshed Y, Chuck G, Bowman J, Hake S (2000). Mechanisms that control *knox* gene expression in the *Arabidopsis* shoot. Development 127 (2): 5523-5532.

Peykari N, Zamani K (2019). Cloning and characterization of a constitutive promoter of polyubiquitin gene from *Cicer ariethinum*. Journal of Crop Science and Biotechnology 9 (25): 35-45.

Roberts WR, Roalson EH (2017). Comparative transcriptome analyses of flower development in four species of *Achimenes* (Gesneriaceae). BMC Genomics 18 (1): 1-26. doi: 10.1186/s12864-017-3623-8

Sobral R, Costa MM (2017). Role of floral organ identity genes in the development of unisexual flowers of *Quercus suber* L. Scientific reports 7 (1): 10368. doi: 10.1038/s41598-017-10732-0

Stephan L, Tilmes V, Hulskamp M (2019). Selection and validation of reference genes for quantitative Real-Time PCR in *Arabis alpina*. PLoS ONE 14 (3): e0211172. doi: 10.1371/journal.pone.0211172

Sun Y, Zhou Q, Zhang W, Fu Y, Huang H (2002). ASYMMETRIC LEAVES1, an *Arabidopsis* gene that is involved in the control of cell differentiation in leaves. Planta 214 (1): 694-702. doi: 10.1007/s004250100673

Tamura K, Stecher G, Peterson D, Filipski A, Kumar S (2013). MEGA6: molecular evolutionary genetics analysis version 6.0. Molecular Biology and Evolution 30 (2): 2725-2729. doi: 10.1093/molbev/mst197

Tattersall AD, Turner L, Knox MR, Ambrose MJ, Ellis TH et al. (2005). The mutant *crispa* reveals multiple roles for PHANTASTICA in pea compound leaf development. Plant Cell 17 (3): 1046-1060. doi: 10.1105/tpc.104.029447

Tian X, Zou P, Miao M, Ning Z, Liao J (2018). RNA-Seq analysis reveals the distinctive adaxial–abaxial polarity in the asymmetric one-theca stamen of *Canna indica*. Molecular Genetics and Genomics 293 (4): 391-400. doi: 10.1007/s00438-017-1392-3

Timmermans MC, Hudson A, Becraft PW, Nelson T (1999). ROUGH SHEATH2: A Myb protein that represses knox homeobox genes in maize lateral organ primordia. Science 284 (1): 151-153.

Truskina J, Vernoux T (2017). The growth of a stable stationary structure: coordinating cell behavior and patterning at the shoot apical meristem. Current Opinion in Plant Biology 41 (1): 83-88. doi: 10.1016/j.pbi.2017.09.011

Waites R, Hudson A (1995). Phantastica: a gene required for dorsoventrality of leaves in *Antirrhinum majus*. Development 121 (3): 2143-2154.

Weber A, Clark JL, Moller M (2013). A new formal classification of Gesneriaceae. Scientific Journal of the Marie Selby Botanical Gardens 31 (2): 68-94.

Wei S, Lou Y, Tsai J, Ho M, Chou C et al. (2012). Structure of the *Trichomonas vaginalis* MYB3 DNA-binding domain bound to a promoter sequence reveals a unique C-terminal beta-hairpin conformation. Nucleic Acids Research 40 (1): 449-459.

Wilkins O, Nahal H, Foong J, Provart NJ, Campbell MM (2009). Expansion and diversification of the Populus R2R3–MYB family of transcription factors. Plant Physiology 149 (1): 981-993. doi: 10.1104/pp.108.132795

Wu G, Lin WC, Huang T, Poethig RS, Springer PS et al. (2008). KANADI1 regulates adaxial–abaxial polarity in Arabidopsis by directly repressing the transcription of ASYMMETRIC LEAVES2. Proceedings of the National Academy of Sciences USA 105 (1): 16392-16397.

Xu B, Li Z, Zhu Y, Wang H, Ma H et al. (2008). Arabidopsis genes AS1, AS2, and JAG negatively regulate boundary-specifying genes to promote sepal and petal development. Plant Physiology 146 (2): 566-575. doi: 10.1104/pp.107.113787

Xu W, Dubos C, Lepiniec L (2015). Transcriptional control of flavonoid biosynthesis by MYB–bHLH–WDR complexes. Trends in Plant Science 20 (3): 176-185. doi: 10.1016/j.tplants.2014.12.001

Xu ZS, Feng K, Que F, Wang F, Xiong AS (2017). A MYB Transcription Factor DcMYB6, is involved in regulating anthocyanin biosynthesis in purple carrot taproots. Scientific Reports 7 (1): 45-54.

Yang J, Yan R, Roy A, Xu D, Poisson J et al. (2015). The I-TASSER Suite: protein structure and function prediction. Nature Methods 12 (1): 7-15. doi: 10.1038/nmeth.3213

Yang T, Wang Y, Teotia S, Zhang Zh, Tang G (2018). The making of leaves: how small RNA networks modulate leaf development. Frontiers in Plant Science 9 (1): 1-7. doi: 10.3389/fpls.2018.00824