

# Structural organization of *TFL1*-like genes in representatives of the tribe Phaseoleae DC.

Ekaterina Krylova<sup>1,2</sup>, Ksenia Strygina<sup>1</sup>, and Elena Khlestkina<sup>1</sup>

<sup>1</sup>Federal Research Center N. I. Vavilov All-Russian Institute of Plant Genetic Resources, Bol'shaya Morskaya ul., 42–44, Saint Petersburg, 190000, Russian Federation

<sup>2</sup>Department of Genetics and Biotechnology, Faculty of Biology, Saint Petersburg State University, Universitetskaya nab., 7–9, Saint Petersburg, 199034, Russian Federation

Address correspondence and requests for materials to Ekaterina Krylova, e.krylova@vir.nw.ru

## Abstract

The type of stem growth is one of the key features in determining plant architectonics. Stem growth type is an economically important trait. It interconnects with stem length, flowering duration, yield, resistance to lodging, and suitability of mechanized cultivation. Mutations in the *TFL1* gene and its homologs have been demonstrated to change meristem indeterminacy across genera. The aim of this work was to characterize and compare the structural organization of *TFL1*-like genes in representatives of the tribe Phaseoleae (pigeonpea *Cajanus cajan*, soybean *Glycine max*, common bean *Phaseolus vulgaris*, adzuki bean *Vigna angularis*, mung bean *V. radiata*, and cowpea *V. unguiculata*) based on *in silico* analysis, including analysis of nucleotide sequences, predicted elements in promoter regions, predicted amino acid sequences, putative functional domains and 3D protein structures. We investigated *TFL1* (one gene for adzuki bean, four copies for soybean, two copies for other studied species), *ATC* (two copies for soybean, one gene for other investigated species), and *BFT* (two copies for soybean, one gene for other studied species) gene family members found in whole-genome sequences databases available for representatives of the tribe Phaseoleae. The presence of duplicated copies for all genes in soybean may be a result of the last genome duplication event during the evolution of this species. Duplication of *TFL1* gene to two copies in most of studied species of the tribe Phaseoleae is probably accompanied by the maintenance of the functional state of these genes. The exception is *VrTFL1.2* of *V. radiata*, which likely had lost its functionality. This work broadens the existing data about the number of gene copies, their structural divergence and evolution, and the expected functional differences. This information will be important for understanding the molecular genetic mechanisms underlying the maintenance of indeterminacy in the growth of the shoot apical meristem, as well as in the control of the transition to the reproductive phase of plant development.

**Keywords:** *ATC*, *BFT*, *Cajanus*, *Glycine*, Phaseoleae, *Phaseolus*, *TFL1*, *Vigna*

## Introduction

The transformation of the vegetative shoot apical meristem into the inflorescence meristem and the subsequent formation of the floral meristem and flower organs are important stages in the development of the flower. The initiation of the formation of the floral meristem is well studied in the model plant *Arabidopsis thaliana* (L.) Heynh. It is controlled by main floral meristem identity genes *LFY* (*LEAFY*), *API* (*APETALA1*), and *TFL1* (*TERMINAL FLOWER1*) (Benlloch, Berbel, Serrano-Mislata and Madueno, 2007). When *LFY* expression reaches the critical level, the meristematic function of the apical meristem cells changes, and the floral meristems are laid (Weigel et al., 1992; Weigel and Nilsson, 1995; Bla'zquez, Soowal, Lee and Weigel, 1997). *LFY* is one of the activators of the *API* gene, which is also involved in determining the identity of the floral meristem (Benlloch, Berbel, Serrano-Mislata and Madueno, 2007). The antago-

**Citation:** Krylova, E., Strygina, K., and Khlestkina, E. 2021. Structural organization of *TFL1*-like genes in representatives of the tribe Phaseoleae DC. *Bio. Comm.* 66(2): 85–108. <https://doi.org/10.21638/spbu03.2021.201>

**Authors' information:** Ekaterina Krylova, Researcher, [orcid.org/0000-0002-4917-6862](https://orcid.org/0000-0002-4917-6862); Ksenia Strygina, PhD, Senior Researcher, [orcid.org/0000-0001-6938-1348](https://orcid.org/0000-0001-6938-1348); Elena Khlestkina, Dr. of Sci. in Biology, Professor RAS, Director, [orcid.org/0000-0002-8470-8254](https://orcid.org/0000-0002-8470-8254)

**Manuscript Editor:** Anton Nizhnikov, Department of Genetics and Biotechnology, Faculty of Biology, Saint Petersburg State University, Saint Petersburg, Russia

**Received:** August 18, 2020;

**Revised:** December 29, 2020;

**Accepted:** January 3, 2021.

**Copyright:** © 2021 Krylova et al. This is an open-access article distributed under the terms of the License Agreement with Saint Petersburg State University, which permits to the authors unrestricted distribution, and self-archiving free of charge.

**Funding:** The part of the present work devoted to analysis of *VuTFL1.1* and *VuTFL1.2* genes was supported by a Saint Petersburg State University grant (Pure ID 62228741). The part of the work devoted to phylogenetic analysis, analysis of nucleotide sequences, predicted elements in promoter regions, predicted amino acid sequences, putative functional domains and 3D protein structures of *TFL1*, *ATC*, *BFT* was prepared within the framework of the state task "Genomic and postgenomic technologies for searching for new genetic markers of economically important traits and new allelic variants of agriculturally valuable genes in the gene pool of cultivated plants and their wild relatives" (VIR project No 0481-2019-0001).

**Ethics statement:** This paper does not contain any studies involving human participants or animals performed by any of the authors.

**Competing interests:** The authors have declared that no competing interests exist.

nist of the *LFY* gene is the *TFL1* gene, which maintains the indeterminacy of the activity of the shoot apical meristem, and, thereby, delays the plant transition to flowering (Benlloch, Berbel, Serrano-Mislata and Madueno, 2007; Moraes, Dornelas and Martinelli, 2019; Périlleux, Bouché, Randoux, and Orman-Ligeza, 2019). *TFL1* belongs to the family of phosphatidylethanolamine binding proteins (PEBPs). This family is an evolutionarily conservative group that is found in all taxa from bacteria to animals and plants (Mimida et al., 2001; Benlloch, Berbel, Serrano-Mislata and Madueno, 2007; Jin, Nasim, Susila, and Ahn, 2020). *TFL1* belongs to a small gene family *CENTRORADIALIS* / *TERMINAL FLOWER 1* / *SELF-PRUNING* (*CETS*). This family in *A. thaliana* consists of six genes: *MOTHER OF FT AND TFL1* (*MFT*), *FLOWERING LOCUS T* (*FT*), *TWIN SISTER OF FT* (*TSF*), *BROTHER OF FT AND TFL1* (*BFT*), *ARABIDOPSIS THALIANA CENTRORADIALIS HOMOLOG* (*ATC*), and *TFL1* (Goretti et al., 2020; Jin, Nasim, Susila, and Ahn, 2020). These genes are divided into three groups: (1) the *MFT* gene that acts during seed germination, (2) a group of *FT*-like genes (*FT* and *TSF*) — floral activators, and (3) a group of *TFL1*-like genes (*TFL1*, *ATC*, *BFT*) — floral inhibitors (Jin, Nasim, Susila and Ahn, 2020).

*FT*, *TFL1* and *TFL1*-like proteins interact with *FLOWERING LOCUS D* (*FD*) (Benlloch, Berbel, Serrano-Mislata and Madueno, 2007; Huang, Jane, Chen and Yu, 2012; Ryu et al., 2014). Under inductive conditions *FT* moves from the leaves through the phloem into the shoot apical meristem and interacts with *FD*; this complex activates the expression of floral meristem identity genes. Under the non-inductive conditions of a short day, *TFL1* forms a heterodimer with the *FD* protein, which leads to blocking of *FD* transcriptional activity. Under stressful conditions, *BFT* competes with *FT* for binding with *FD* protein (Huang, Jane, Chen and Yu, 2012). Thus, *FT* facilitates the plant's transition from vegetative development to the reproductive stage, initiating flowering, and *TFL1* acts as a repressor of flowering initiation.

The function of *ATC* is unclear. Presumably, *ATC* gene is able to functionally replace the *TFL1* in the case of disturbances of its structure (Huang, Jane, Chen and Yu, 2012). However, the phenotypes of *atc* and *tfl1*, as well as different levels and patterns of gene expression, suggest that, unlike *TFL1*, *ATC* is not involved in determining the identity of floral meristems (Mimida et al., 2001). According to the literature, *BFT* expression is stimulated only under abiotic stress conditions (Chung et al., 2010; Yoo et al., 2010; Ryu, Park and Seo, 2011; Ryu et al., 2014).

The tribe Phaseoleae DC. includes economically important species, many of which are the most important food legumes in many countries (Boukar et al., 2015). Species of the genus *Vigna* Savi cowpea (asparagus bean, *V. unguiculata* (L.) Walp.), mung bean (*V. radiata* (L.) R. Wilczek.), and adzuki bean (*V. angularis* (Willd.) Ohwi et Ohashi) are included in the human diet as basic dietetic

products for millions. Many species can provide a wide variety of food products throughout the growing season. According to the Food and Agriculture Organization (FAO), the annual world production of various *Vigna* species is approaching 12 million hectares in Asia, Africa, Southern Europe, Central and South America. The cowpea *V. unguiculata* is a multipurpose crop, the interest in which has greatly increased in recent years among breeders (<http://www.fao.org>; Boukar et al., 2015; Burlyaeva et al., 2019; Vishnyakova et al., 2019). The mung bean *V. radiata* has been consumed as a traditional food all over the world for over 3,500 years. Its seeds and sprouts have many compounds important for human health. The adzuki bean is used as food; besides, the mung bean and adzuki bean have long been used in traditional medicine in China.

Genomes of many representatives of the Phaseoleae have been sequenced, but the molecular genetic mechanisms of stem growth control and the role of *TFL1*-like genes in this process are not well understood (Schmutz et al., 2010, 2014; Varshney et al., 2012; Kang et al., 2014; Lonardi et al., 2019). At the present time information about *TFL*-like gene copies and their functionality is required for subsequent works. *In silico* analysis is the basis for future experimental researches of molecular genetic mechanisms in the control of the transition to flowering. The aim of this work was to characterize the structural organization of *TFL1*-like genes in representatives of the tribe Phaseoleae.

## Materials and methods

The search for homologous sequences of *TFL1* (GenBank: AT5G03840) was conducted using the BLASTN algorithm in databases EnsemblPlants (<http://plants.ensembl.org/index.html>), Phytozome v12.1 (<https://phytozome.jgi.doe.gov/pz/portal.html#>), *Vigna* Genome Server (<https://viggs.dna.affrc.go.jp/>), and LIS — Legume Information System (<https://legumeinfo.org>) in the genomes of the main representatives of the tribe Phaseoleae: pigeonpea *Cajanus cajan* (L.) Millsp., soybean *Glycine max* (L.) Merr., common bean *Phaseolus vulgaris* L., adzuki bean *V. angularis*, mung bean *V. radiata*, and cowpea *V. unguiculata* (accessed Jan. 20, 2020) (Gonzales et al., 2005; Goodstein et al., 2012; Kersey et al., 2014; Sakai et al., 2015, 2016). The multiple alignment of nucleotide and amino acid sequences was made using MULTALIN v5.4.1 (<http://multalin.toulouse.inra.fr/multalin/>) (Corpet, 1988). The cluster analysis was performed using MEGAX software and the Neighbour-joining algorithm with 1000 bootstrap replicates (Felsenstein, 1985; Saitou and Nei, 1987; Tamura, Nei and Kumar, 2004; Kumar et al., 2018). Genes *FT* (GenBank: AT1G65480), *TSF* (GenBank AT4G20370), and *MFT* (GenBank: AT1G18100) were used as the out-group. The prediction of element in promoter regions was carried out with New PLACE (<https://www.dna.affrc.go.jp/PLACE/?action=newplace>) (Higo, Ugawa, Iwamoto

and Korenaga, 1999) and PlantPAN (<http://plantpan.itps.ncku.edu.tw/>) (Chow et al., 2019). The annotation of the functional domains was carried out using InterPro (<http://www.ebi.ac.uk/interpro/>) (Finn et al., 2016). Modelling of the tertiary structure of the predicted amino acid sequences was made using SWISSMODEL (<https://swissmodel.expasy.org/interactive>) based on 1wko.1.A, 1qou.1, and 6igi.1 templates from PDB (Waterhouse et al., 2018). For multiple structural alignment of 3D protein structures, the PDBeFold service was used (<https://www.ebi.ac.uk/msd-srv/ssf/>). Protein similarity was determined with the LALIGN tool ([https://fasta.bioch.virginia.edu/fasta\\_www2/fasta\\_www.cgi?rm=lalign&pgm=lal](https://fasta.bioch.virginia.edu/fasta_www2/fasta_www.cgi?rm=lalign&pgm=lal)). The default settings were used for all software.

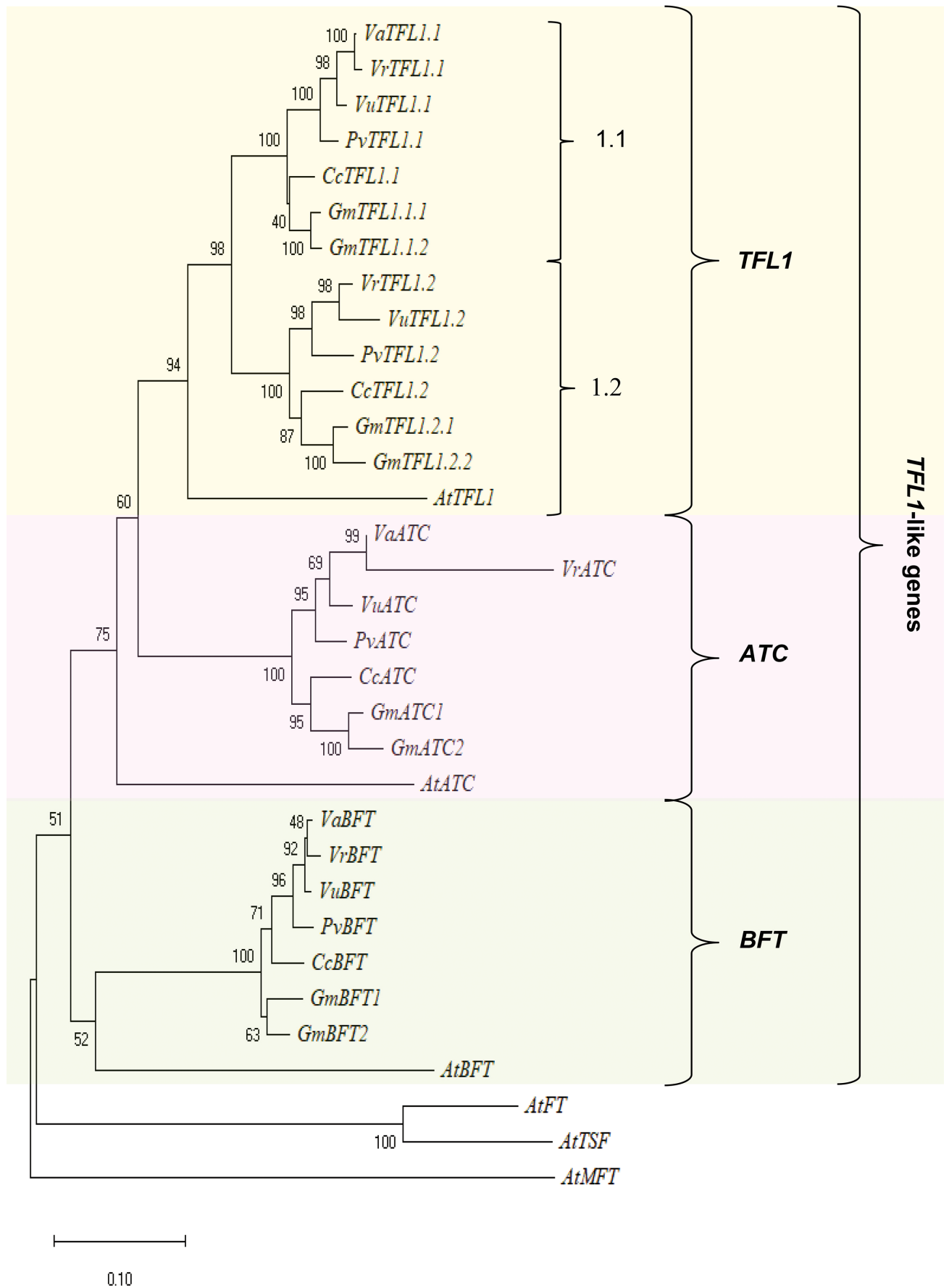
## Results

### Identification and phylogenetic analysis of *TFL1*-like genes of Phaseoleae species

The search for homologous sequences of the *TFL1* gene of *A. thaliana* was carried out in the genomes of the main representatives of the tribe Phaseoleae. In total, four sequences of *TFL1*-like genes were identified in the genomes of *C. cajan*, *P. vulgaris*, *V. radiata*, and *V. unguiculata*.

**Table 1. Identified homologs of *TFL1*-like genes in the genomes of studied members of the tribe Phaseoleae. \* — in concordance with NCBI**

Gene of <i>A. thaliana</i>	Species	Ortholog	Sequence	Chromosome	
TFL1	<i>C. cajan</i>	<i>CcTFL1.1</i>	C. cajan10074	3	
	<i>G. max</i>	<i>GmDt1 / GmTFL1b / GmTFL1.1.1</i>	AB511820* / Glyma19g194300	19	
		<i>GmTFL1.1.2</i>	Glyma03g194700	3	
	<i>P. vulgaris</i>	<i>PvTFL1y / PvTFL1.1</i>	JN418231* / Phvul.001G189200	1	
	<i>V. angularis</i>	<i>VaTFL1.1</i>	Vigan04g345000	4	
	<i>V. radiata</i>	<i>VrDet1 / VrTFL1.1</i>	Vigrad03g04510	3	
	<i>V. unguiculata</i>	<i>VuTFL1 / VuTFL1.1</i>	KJ569523* / Vigun01g173000	1	
	<i>C. cajan</i>	<i>CcTFL1.2</i>	C. cajan36529	Scaffold137665:134,613..135,929	
		<i>G. max</i>	<i>GmTFL1.2.1</i>	Glyma10g071400	10
			<i>GmTFL1.2.2</i>	Glyma11g209500	11
		<i>P. vulgaris</i>	<i>PvTFL1.2</i>	Phvul007g229300	7
		<i>V. radiata</i>	<i>VrTFL1.2</i>	Vigrad08g05490	8
		<i>V. unguiculata</i>	<i>VuTFL1.2</i>	Vigun07g059700	7
ATC	<i>C. cajan</i>	<i>CcATC</i>	C. cajan31760	Scaffold132593:212,899..214,119	
	<i>G. max</i>	<i>GmATC1</i>	Glyma12g184000	12	
		<i>GmATC2</i>	Glyma13g317100	13	
	<i>P. vulgaris</i>	<i>PvATC</i>	Phvul005g124600	5	
	<i>V. angularis</i>	<i>VaATC</i>	Vigan07g072900	7	
	<i>V. radiata</i>	<i>VrATC</i>	Vigrad04g03610	4	
	<i>V. unguiculata</i>	<i>VuATC</i>	Vigun05g236900	5	
BFT	<i>C. cajan</i>	<i>CcBFT</i>	C. cajan16048	8	
	<i>G. max</i>	<i>GmBFT1</i>	Glyma09g143500	9	
		<i>GmBFT2</i>	Glyma16g196300	16	
	<i>P. vulgaris</i>	<i>PvBFT</i>	Phvul04g119700	4	
	<i>V. angularis</i>	<i>VaBFT</i>	Vigan11g190300	11	
	<i>V. radiata</i>	<i>VrBFT</i>	Vigrad01g14660	1	
<i>V. unguiculata</i>	<i>VuBFT</i>	Vigun04g159500	4		



**Fig. 1.** Analysis of phylogenetic similarity of *TFL1*-like genes (CDS) using Neighbor-Joining method with program MEGAX. Genes *FT* (GenBank: AT1G65480), *TSF* (GenBank AT4G20370), *MFT* (GenBank: AT1G18100) are as outgroup.

*lata*, three genes in the genome of *V. angularis*, and eight genes in the genome of *G. max*. All found sequences are listed in Table 1.

As shown in Figure 1, all identified sequences are divided into three clades. The first clade includes sequences highly homologous to the *TFL1* gene of *Arabidopsis* (Fig. 1, yellow). All studied species have two copies of *TFL1* in their genome, which we designated as *TFL1.1* and *TFL1.2*. The exceptions are *V. angularis* and *G. max*, in the genomes of which there are one and four *TFL1* genes, respectively (Table 1).

At the same time, the *VaTFL1.1* sequence of *V. angularis*, according to the *Vigna* Genome Server database, is located in chromosome 4. However, in accordance with the EnsemblPlants database, this gene is located in chromosome 3, and, according to the LIS database, in chromosome 6. In our study, we present according to the *Vigna* Genome Server (Table 1).

The other two clades include genes homologous to genes *ATC* (GenBank: AT2G27550) and *BFT* (GenBank: AT5G62040) of *Arabidopsis* (Fig. 1, pink and green clades, respectively). It should be noted that the *ATC* clade is paraphyletic. In all studied species, except for *G. max*, one sequence of the *ATC* and *BFT* genes was found. In the genome of *G. max* two copies of these genes were identified (Table 1).

### Structural organization of identified genes

The exon-intron structure of most identified *TFL1*-like genes is the same in all species, with the exception of *G. max* (see below). Most of the identified genes of *G. max*, *C. cajan*, *P. vulgaris*, as well as *V. radiata*, *V. angularis* and *V. unguiculata*, consist of four exons and three introns (Fig. 2). An exception is the *VrATC* gene, which consists of two exons and one intron. In addition, the *VaATC* gene, according to the genomic *Vigna* Genome Server, consists of four exons and three introns, but according to the EnsemblPlants database, it has five exons and four introns. In our study, we present data in accordance with the *Vigna* Genome Server.

The second and third exons have highly conserved sequences: in all identified genes the lengths were 62 and 41 bp, respectively (Fig. 2). The first exon was the most variable in length. In contrast to exons, the intron lengths varied greatly in all identified genes; long introns were noted for *TFL1*-like genes of *C. cajan*.

Analysis of the identified *G. max* genes showed that their exon-intron structures are not the same. Among *TFL1*-like genes, according to the genome assembly Glyma.Wm82.a2.v1, *GmTFL1.1.2* (*Glyma03g194700*) consists of five exons and four introns. In the assembly Glyma.Wm82.a1.v1.1, it corresponds to the sequence *Glyma03g35250*, which consists of four exons and three

introns. In addition, a similar structure was noted for the identified *GmATC* genes—these genes also have five exons and four introns.

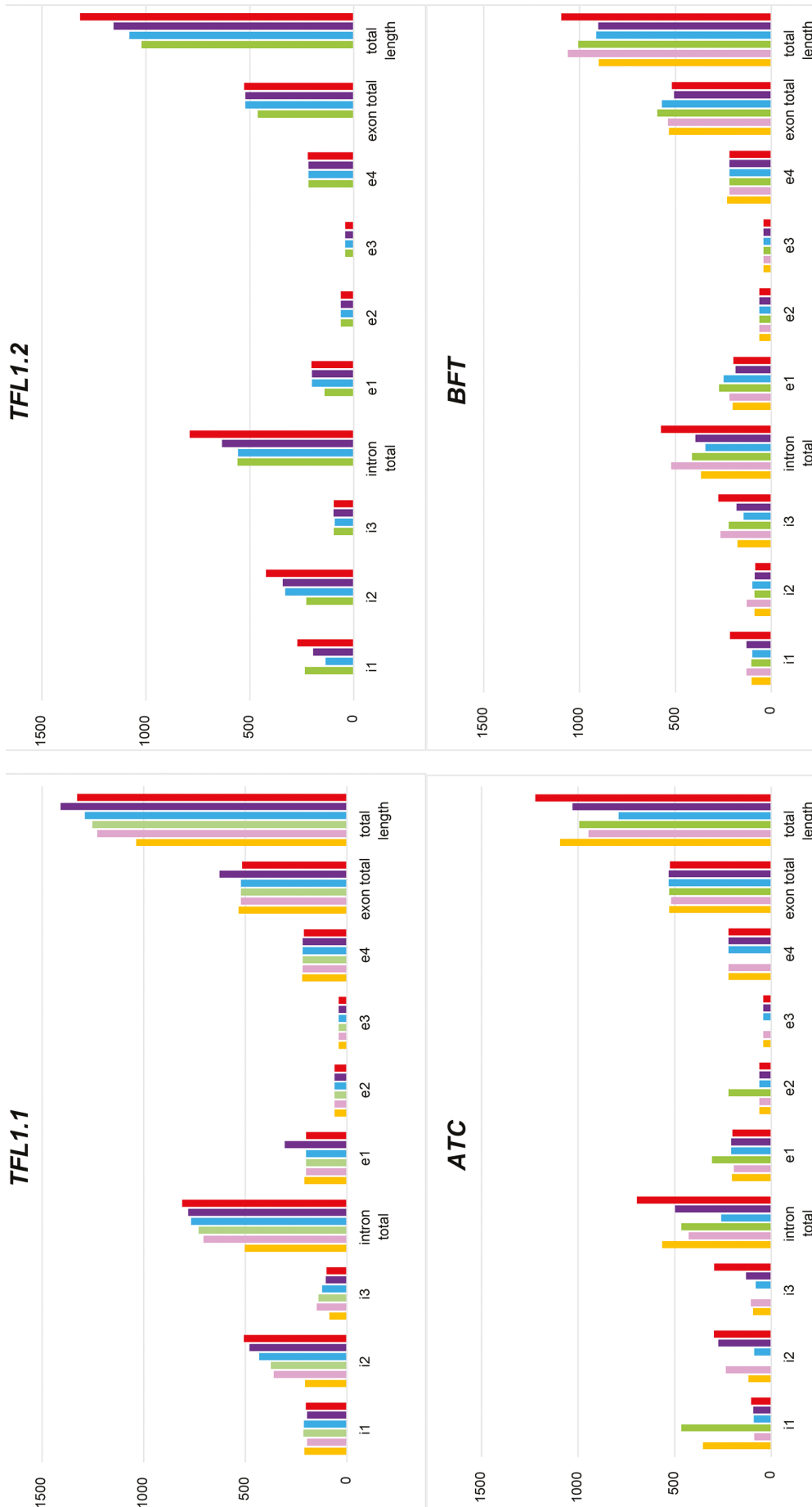
### Analysis of promoter regions of *TFL1*-like genes of *V. radiata* and *V. unguiculata*

Due to the fact that the data of *TFL1*-like genes of *V. angularis* are contradictory, in further work we will carry out more detailed analysis of *TFL1*-like genes of two species of the genus *Vigna* — *V. radiata* and *V. unguiculata*. The search for *cis*-acting regulatory DNA elements of identified *V. radiata* and *V. unguiculata* *TFL1*-like genes was in the region ~1000 bp upstream to the ATG start translation codon (Fig. 3, Supplementary file 1). All sequences had strongly conserved motifs, such as CAAT-box and TATA-box. Analyses of promoter elements of all identified sequences revealed many motifs for binding of transcription factors (TFs), light-induced, and hormone-responsible promoter elements. We identified binding sites for proteins that are involved in regulation of genes that are responsive to water stress (Supplementary file 1).

Of all the identified binding sites, the most common were sites for transcription factors containing bHLH domain (basic helix-loop-helix), bZIP TFs, and AP2-like TFs, one is the ERF protein, which regulates plant response to ethylene. In all promoter regions TF binding sites of RAV1 family were identified. Proteins of this family are involved in the regulation of various growth and development processes. In addition, all promoters had multiple binding sites for ARR proteins, which are regulators of gene expression for the primary response to cytokinins. Also, in most promoters (an exception is *VuTFL1.1*), regions corresponding to the W-box were found. These regions are involved in the response to salicylic acid. Among the identified hormone-sensitive motifs, sites for binding to a cytokinin-binding protein, sites characteristic of gibberellin-regulated genes, as well as sites characteristic of genes regulated by abscisic acid, were found. At the same time, in contrast to the *Arabidopsis TFL1* gene promoter (Kushwah, Ahmad and Ali, 2014), the analyzed promoter regions did not have binding sites for ARF transcription factors that act as activators or repressors of the transcription of auxin-regulated genes.

### Annotation of the functional domains

All predicted proteins belong to the PEPB protein family. Schemes of *TFL1*-like proteins of *V. unguiculata* (in comparison with *V. radiata*) are presented in Supplementary file 2, and all identified domains and critical amino acid residues are designated. Most of the proteins have several overlap domains: PEBP\_euk CDD



**Fig. 2.** Length polymorphisms of exons and introns of identified *TFL1*-like genes. Orange colour — *A. thaliana*, pink — *V. angularis*, green — *V. radiata*, blue — *V. unguiculata*, violet — *P. vulgaris*, red — *C. cajan*

(IPR035810)—specific for PEBP proteins of eukaryotes; domain PEBP (IPR008914)—specific domain for PEBP proteins; and conservative domain IPR001858.

All identified protein sequences have all critical amino acid residues (Asp71, His85, His87, Glu109, Phe120, Asp140) (Fig. 4, Supplementary file 2). Protein VrTFL1.2 has 20 amino acid deletion, but it has all critical amino acid residues. Also, for VrTFL1.2 protein the conservative domain IPR001858 was not found.

In all proteins of *V. radiata* and *V. unguiculata*, Phe137 was replaced by Thr137; Ser142 is presented. VuTFL1.1, VuTFL1.2, VrTFL1.1, and VrTFL1.2 contain non-Tyr and non-Trp at position 134 and 138 based on alignment with *Arabidopsis* FT. These proteins are probably flowering repressors (see discussion).

Presently, ATC is considered the most likely orthologue of *CEN* gene of snapdragon *Antirrhinum majus* L. (Mimida et al., 2001; Huang, Jane, Chen and Yu, 2012). In our work, we consider ATC in comparison with CEN (Supplementary file 2). All critical amino acid residues (His87-Asp143, Asp73, His89, Glu111, Phe122) are conserved in VuATC.

VrATC protein does not contain PEBP (IPR008914) or conservative (IPR001858) domains. VrATC contains non-tyrosine and non-tryptophan at position 134 and 138 based on alignment with AtFT. In the sequence of the VrATC protein, three important amino acids have been replaced: His86Ile, His88Leu, and Asp72Ser, and so the function of this protein could be lost (Fig. 4, Supplementary file 2).

Analysis of the amino acid sequences of the identified proteins VuBFT and VrBFT showed that the pair of conserved amino acids His88-Asp144 (according to AtTFL1), which determines the specificity of the protein function as a flowering repressor, is preserved. VuBFT and VrBFT contain non-tyrosine and non-tryptophan at position 134 and 138 based on alignment with AtFT. Thus, these proteins, like TFL1 and ATC, are probably flowering repressors (see discussion).

### The prediction of 3D structures of TFL1-like proteins

**TFL1.** The 1wko.1.A template, deposited in PDB, was used to predict 3D structures of all identified TFL1 proteins. It corresponds to *Arabidopsis* TFL1 protein with 100% identity. The applied template achieves about 73.6% identity with VuTFL1.1, VrTFL1.1 (VrDet1), VuTFL1.2, VrTFL1.2 proteins (Supplementary file 3).

Overall root-mean-square deviation (RMSD) of atomic positions between 3D structures of AtTFL1 and VuTFL1.1 is 0.19 Å; the same RMSD is found between 3D structures of AtTFL1 and VuTFL1.2 (Supplementary file 4). RMSD between AtTFL1 and VrDet1 is 0.168 Å; between VuTFL1.1 and VrDet1—0.122 Å. The

smallest RMSD (0.034 Å) is between VuTFL1.1 and VuTFL1.2, indicating their high homology of 3D structures. The highest RMSD is between 3D structures of VrTFL1.2 with AtTFL1, VuTFL1.1, VrDet1, VuTFL1.2. This indicates a high degree of divergence of these aligned 3D structures. Sequence identity of protein sequences is the highest (99.4%) between VuTFL1.1 and VrDet1 (Supplementary file 4). Identity between all TFL1-like proteins of *V. unguiculata* and *V. radiata* with AtTFL1 reaches more than 79%.

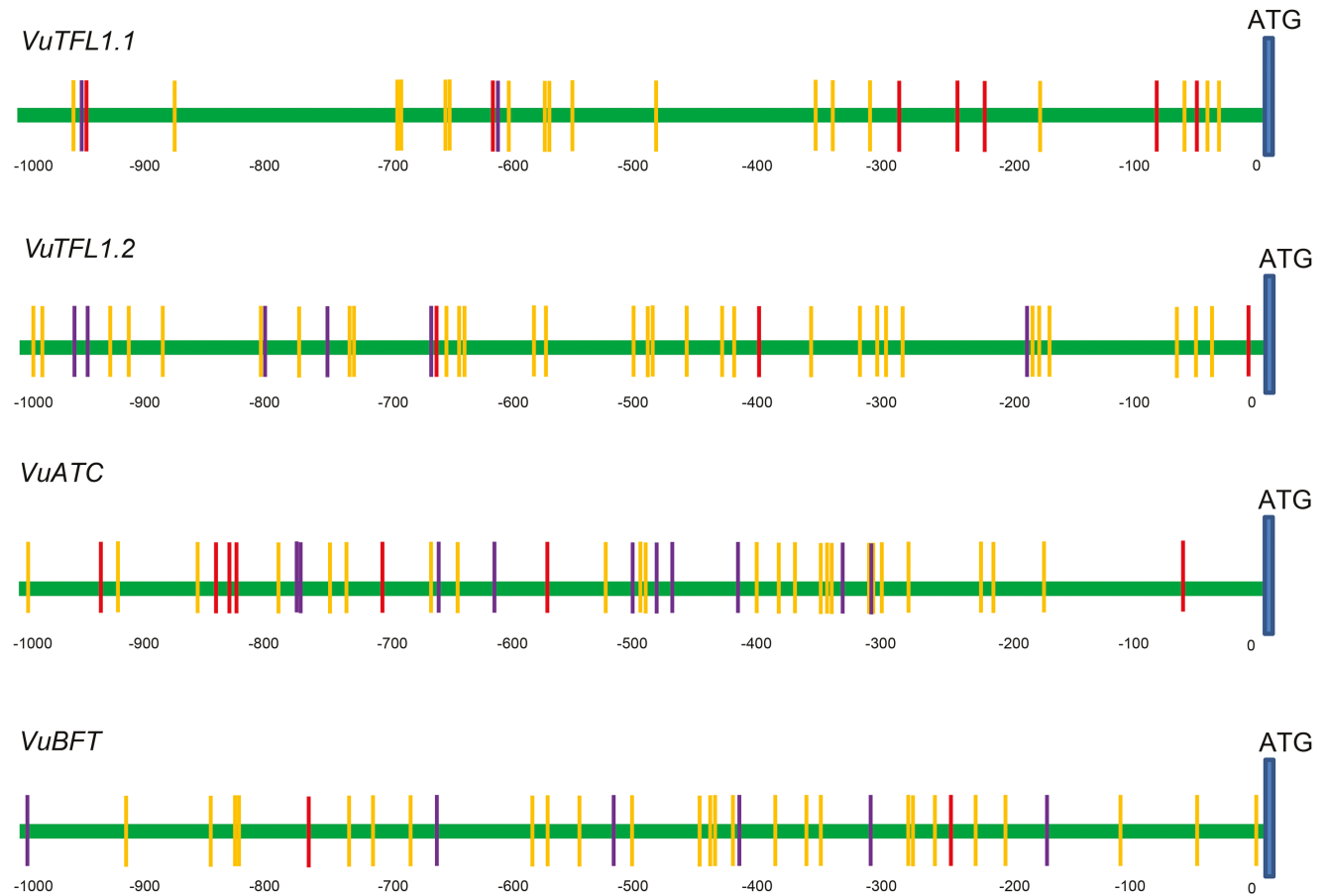
**ATC.** The 1qou.1 template was used to predict 3D structures of all identified ATC-like proteins. It corresponds to snapdragon CEN and it is presented as a dimer (Banfield and Brady, 2000). It is supposed that protein dimer is unfunctional (Banfield and Brady, 2000).

The applied template 1qou.1 constitutes about 78.61% identity with *Arabidopsis* ATC; identity with VuATC and VrATC is 76.61% and 63.69%, respectively (Supplementary file 3). The smallest RMSD is between CEN and VuATC (0.187 Å) (Supplementary file 4). RMSD between other identified proteins is high, which indicates the low homology of their 3D structures. Sequence identity of protein sequences is about 90% between VuATC and AtATC, CEN (Supplementary file 4). Identity between VrATC and AtATC, CEN is about 80%. Amino acid sequence identity of VuATC and VrATC is 86.6% (Supplementary file 4).

**BFT.** The 6igi.1 template was used to predict 3D structures of all identified BFT-like proteins. It corresponds to *Arabidopsis* FT protein. It is presented as monomer. The applied template 6igi.1 constitutes about 61.11% identity with *Arabidopsis* BFT, identity with VuBFT and VrBFT is about 60% (Supplementary file 3). High RMSD values were obtained by multiple structural alignments of 3D structures of all identified BFT-like proteins. It indicates a high degree of divergence of these aligned 3D structures (Supplementary file 4). Amino acid sequence identity of VuBFT and VrBFT is 98.9%. Sequence identity of these proteins with AtBFT is about 85%.

### Discussion

The group of *TFL1*-like genes includes *TFL1*, *ATC* and *BFT*. There are several studies of the role of the *ATC* and *BFT* genes (Mimida et al., 2001; Chung et al., 2010; Yoo et al., 2010; Ryu, Park and Seo, 2011; Huang, Jane, Chen and Yu, 2012; Ryu et al., 2014). It has been shown that the *ATC* gene is a flowering repressor, which plays a critical role in regulating transition to flowering under non-inductive photoperiodic conditions (Mimida et al., 2001; Huang, Jane, Chen and Yu, 2012). *BFT* expression is stimulated under certain abiotic stress conditions with a diurnal pattern (Chung et al., 2010; Yoo et al., 2010; Ryu, Park, and Seo, 2011; Ryu et al., 2014).



**Fig. 3.** The localization of promoter motifs on 5'-upstream region before ATG transcription start site. Orange — light-responsive site, red — TF binding motif, violet — hormone-responsive site.

The *TFL1* gene is one of the main regulators of the plant's transition to flowering. Plants with mutations in the *TFL1* gene show an early transition to flowering, while the overexpression of this gene showed a prolonged vegetative stage of development, a later transition to flowering with the formation of highly branched inflorescences (Alvarez, Guli, Yu and Smyth, 1992; Ratcliffe et al., 1998; Benlloch, Berbel, Serrano-Mislata and Madueno, 2007). *TFL1* homologs have been identified in many angiosperms. The *TFL1* genes in soybean *G. max* and common bean *P. vulgaris* are the best-studied among the tribe Phaseoleae (Kwak, Velasco and Gepts, 2008; Liu et al., 2010; Tian et al., 2010; Kwak, Toro, Debouck and Gepts, 2012; Repinski, Kwak and Gepts, 2012). *TFL1* genes were also found in the genomes of the mung bean *V. radiata*, the cowpea *V. unguiculata*, and also in the genome of the pigeonpea *C. cajan* (Gumber and Singh, 1997; Dhanasekar and Reddy, 2014; Mir et al., 2014; Li et al., 2018). In the present study, we analyzed the genomes of other members of the tribe Phaseoleae and identified highly homologous genes to *TFL1*—*ATC* and *BFT*.

### Identification of new *TFL1*-like genes

***TFL1*.** Members of the tribe Phaseoleae are ploidy. Among the considered species, cowpea *V. unguiculata*, mung bean *V. radiata*, adzuki *V. angularis*, common bean *P. vulgaris*, pigeonpea *C. cajan* are diploid ( $2n=22$ ). *G. max* is a partially diploidized tetraploid ( $2n=40$ ), whose chromosomes have undergone many rearrangements (Schmutz et al., 2010).

Four copies of *TFL1* in soybean *G. max* were known. It was shown that these paralogs formed from a whole-genome duplication that occurred about 50 million years ago. Within each pair, genes arose by duplication about 13 million years ago (Tian et al., 2010). In this study, these genes were designated as *GmTFL1.1.1* / *GmTFL1.1.2* and *GmTFL1.2.1* / *GmTFL1.2.2* (in chromosomes 19, 3, 10 and 11, respectively). Based on our results of the nucleotide sequences analysis, as well as the expression pattern of the *GmTFL1.1.1* and *GmTFL1.1.2* genes (Tian et al., 2010), it was assumed that these genes are maintained in a functional state.

Earlier, three *TFL1* genes, *PvTFL1x*, *PvTFL1y* and *PvTFL1z*, were identified in the genome of common



bean *P. vulgaris*. These genes were mapped in chromosomes 4, 1, and 7, respectively (Kwak, Velasco and Gepts, 2008). In our study, the *PvTFL1y* gene, which is a functional homolog of the *TFL1* of *Arabidopsis*, was designated as *PvTFL1.1* (Fig. 1). Based on the location (chromosome 7) of the *PvTFL1.2* identified in the present study, we assume that it corresponds to the previously isolated gene *PvTFL1z*. Based on our results, *PvTFL1x* most likely is the *PvBFT* gene (see below).

In the genome of *V. radiata*, two *TFL1* sequences were previously identified in chromosomes 3 and 8 (Li et al., 2018). It was found that the *VrDet1* gene performs functions similar to those of the soybean *Dt1* and common bean *PvTFL1y* (Li et al., 2018). In the present study, this copy was designated as *VrTFL1.1*. The gene identified in chromosome 8 was designated as *VrTFL1.2*. Analysis of protein sequence of *VrTFL1.2* showed 20 amino acid deletion, and the conservative domain was not identified (Fig. 4, Supplementary file 2). This protein is probably nonfunctional.

In the genome of cowpea *V. unguiculata*, only one *TFL1* gene, *VuTFL1*, was previously identified and described (Dhanasekar and Reddy, 2014). We identified an additional sequence homologous to the *VuTFL1* gene (in the present study designated as *VuTFL1.1*) in chromosome 7. This copy is designated as *VuTFL1.2*.

For the first time in the genome of adzuki *V. angularis* we identified one sequence, *VaTFL1.1* (in chromosome 4 according to the *Vigna* Genome Server), which is highly homologous to the *TFL1* gene of *Arabidopsis*. The genome version (Va3.0) of this species is available on Legume Information System (Kang et al., 2015). Besides that, genome sequencing of adzuki bean using single-molecule real-time (SMRT) sequencing technology is available on *Vigna* Genome Server (Sakai et al., 2015). According to different databases' information, the chromosomal localization of *VaTFL1.1* is different. And we did not identify a sequence homologous to the *VaTFL1.1* gene like for other studied members of the tribe Phaseoleae.

Thus, as a result of our study, an additional copy of the *TFL1* gene was revealed in the genomes of members of the tribe Phaseoleae (except adzuki) (Fig. 1, clade 1.2). The genes of this group (clade 1.2) are the result of the duplication of genes from clade 1.1 into other chromosomes (Table 1).

**ATC u BFT.** In the genomes of all the studied members of the tribe Phaseoleae, we identified one *ATC* and one *BFT* gene (Fig. 1, Table 1). The exception is soybean *G. max*, in the genome of which due to the polyploid nature, two copies of *ATC* were identified in chromosomes 12 and 13—*GmATC1* and *GmATC2*, respectively. In addition, as for other *TFL1*-like genes, two homologous *BFT* genes in chromosomes 9 and 16, *GmBFT1* and *GmBFT2*, were identified in the soybean genome (Fig. 1, Table 1). The genes are likely to be a homoeologous pair resulting from a whole genome duplication.

Based on the location of the *ATC* and *BFT* genes, we assume that previously only the *BFT* gene of the common bean *P. vulgaris* was identified. This gene was designated as *PvTFL1x* in the original work and was mapped in chromosome 4 in the *P. vulgaris* genome (Kwak, Velasco and Gepts, 2008). The activity of this gene in response to stress has not been analyzed. In our study this gene is in the *BFT* clade (Fig. 1, green clade), and we assume that the *PvTFL1x* gene is the *PvBFT* gene.

The results of our study are consistent with the literature data. The identified sequences of highly homologous *TFL1* are located in chromosomes 1 and 7 of the *V. unguiculata* and *P. vulgaris* (genes *VuTFL1.1*, *VuTFL1.2*, *PvTFL1.1* and *PvTFL1.2*). In addition, the *VuATC* / *PvATC* and *VuBFT* / *PvBFT* gene pairs are located in chromosomes 5 and 4, respectively. The homologous sequences of all genes of the two species *V. angularis* and *V. radiata* have been identified in other chromosomes.

### Genome synteny of members of tribe Phaseoleae

The genomes of the main members of the tribe Phaseoleae have been sequenced; work on genome sequencing of other species is ongoing (Schmutz et al. 2010, 2014; Varshney et al. 2012; Kang et al., 2014; Lonardi et al., 2019). Due to this, it became known that the genomes are distinguished by a high degree of collinearity (Lonardi et al., 2019). The highest level of macrosynteny was found between the genomes of the cowpea *V. unguiculata* and the common bean *P. vulgaris* (Vasconcelos et al., 2015; Munoz-Amatriain et al., 2017). Six of the eleven cowpea chromosomes are largely syntenic with six common bean chromosomes. Each of the remaining five cowpea chromosomes corresponds to parts of two *P. vulgaris* chromosomes. The exception is chromosomes Vu05 and Vu08 of *V. unguiculata* (in accordance with the previous numbering, chromosomes 1 and 5, respectively), which have an extended synteny with the *P. vulgaris* chromosome Pv08 (Lonardi et al., 2019).

Comparative mapping of genomes of three *Vigna* species (*V. unguiculata*, *V. angularis*, *V. radiata*) and common bean *P. vulgaris* showed that six chromosomes of *V. unguiculata* (Vu04, Vu06, Vu07, Vu09, Vu10 and Vu11) largely have synteny with single chromosomes of three other species (Lonardi et al., 2019). The presence of rearrangements in chromosomes Vu02, Vu03 and Vu08 suggests that they are characteristic of the divergence of *Vigna* from *Phaseolus*, while chromosomal rearrangements in Vu01 and Vu05 are characteristic of the *Vigna* species.

### *TFL1*-like genes — flowering repressors

**Structure of *TFL1*-like promoters.** The plant transition from the vegetative to reproductive phase is controlled by complex factors (Benlloch, Berbel, Serrano-Mislata

and Madueno, 2007). Photoperiod and temperature are the main exogenous factors. In the present study we reveal many light-induced promoter elements in all identified sequences (Fig. 3, Supplementary file 1). Endogenous factors such as phytohormones, circadian clock, and senescence are also important. At the stage of transition to flowering, changes of phytohormonal concentrations occur. In all identified promoter sequences we found many hormone-responsible elements (Fig. 3, Supplementary file 1). It should be noted that the pattern of distribution of identified promoter elements of VrATC was different from other studied sequences.

**Structure of TFL1-like and FT proteins.** The *TFL1* and *FT* genes encode small proteins with ~60% identity — 177 and 175 amino acids in length, respectively (Hanzawa, Money and Bradley, 2005; Ahn et al., 2006). A ligand-binding region similar to the structure of mammalian PEBP proteins was previously found in the TFL1 and FT proteins (Banfield and Brady, 2000; Ahn et al., 2006).

In PEBP proteins, the amino acids His86, Asp70, Glu110 and Tyr120 are known to play a critical role in ligand binding. A similar structure was shown for TFL1, FT and CEN plant proteins. Unlike mammalian PEBPs, TFL1 and CEN contain Phe123 and Phe125, respectively, instead of Tyr120, and FT contains Val120 (Ahn et al., 2006).

The amino acids Tyr85 in FT and His88 in TFL1 are located at the beginning of the ligand binding domain and form a hydrogen bond with Glu109 and Glu112, respectively. The study of the proteins' crystal structure, as well as comparison with the CEN protein, have shown that substitutions of His88 in TFL1 and Tyr85 in FT critically affect protein function. Substitutions of His88Tyr in TFL1 and Tyr85His in FT lead to a reversal of the functional significance of proteins — TFL1 becomes an activator of flowering, while the FT protein acquires the function of a repressor (Hanzawa, Money and Bradley, 2005). Earlier, according to the results of experiments with chimeric proteins, it was shown that the replacements in the fourth exon of the *TFL1* and *FT* genes significantly changed the functions of proteins (Ahn et al., 2006). The sequence was subdivided into four segments (A, B, C, D). It was shown that segments A and C did not differ significantly between TFL1 and FT proteins, while segments B and D were more variable. Segments B and C together are required for FT protein, and segment B or C is required for TFL1. It was shown that the B segment evolves rapidly in TFL1s, but practically does not change in FTs. (Ahn et al., 2006). In our work we will carry out more detailed analysis of TFL1-like proteins of two species, *V. radiata* and *V. unguiculata*. In the present study all identified protein sequences (VuTFL1.1, VuTFL1.2, VrTFL1.1, VrTFL1.2) have all critical amino acid residues (Asp71, His85, His87, Glu109, Phe120, Asp140)

(Fig. 4, Supplementary file 2). Besides that, all identified TFL1-like protein sequences of cowpea and mung bean have all four segments (Fig. 4).

Segment B, encoded by the fourth exon, forms an outer loop that determines the fundamental difference in the structure of the FT and TFL1 proteins (Ahn et al., 2006). The loop contains amino acids Gln140 in FT and Asp144 in TFL1. They interact with functionally important amino acids Tyr85/His88. At the same time, only His88 in TFL1 forms a second hydrogen bond with Asp144, while Tyr85 and Gln140 (which corresponds to Asp144) in FT does not form one (Ahn et al., 2006). Our analysis of TFL1-like protein sequences showed that all critical amino acid residues are preserved (Fig. 4, Supplementary file 2).

Further studies of the protein structures of FT and TFL1 identified additional key amino acids for the proper functioning of proteins. It was shown that flowering activators, FT-like proteins, have Tyr134, and flowering repressors, TFL1-like proteins, have an amino acid different from Tyr134. In addition, FT-like proteins have Trp138, and TFL1-like proteins have another amino acid in this position (Ho and Weigel, 2014; Wickland and Hanzawa, 2015). Changes in the amino acid composition in the loop may explain the functional diversification of FT homologues from flowering activators to repressors. Thus, in the beet *Beta vulgaris* L., two FTs with opposite functions were identified (*BvFT1* is a repressor, *BvFT2* is a flowering activator). Simultaneous replacement of three amino acids (Tyr134Asn, Gly137Gln and Trp138Gln) in the loop changes the function of the *BvFT1* protein from a flowering repressor to an activator (Pin et al., 2010). In the present study all identified proteins (VuTFL1.1, VuTFL1.2, VrTFL1.1, VrTFL1.2) contain non-tyrosine and non-tryptophan at position 134 and 138 based on alignment with AtFT (Fig. 4).

In addition, the surface charges of FT and TFL1 are crucial for determining the functions. The surface charge of Glu112 in TFL1 is positive, while the charge of Glu109 in FT is negative. The replacement of negatively charged Glu with positively charged Lys (Glu109Lys) in FT leads to a function change, while the replacement of Glu109Asp did not cause the effect (Ho and Weigel, 2014). The replacement of Gln140 with a negatively charged or neutral amino acid (Gln140Asp and Gln140Ala) did not lead to a change in the FT function. However, Gln140 replacement with a positively charged amino acid (Gln140Lys and Gln140Arg) leads to the functional change. The charge change within the ligand-binding domain (Asp71Asn) did not affect protein function (Ho and Weigel, 2014).

Thus, the ligand-binding region and the conformation-changing loop are the most important regions for determining the functional significance of proteins, and

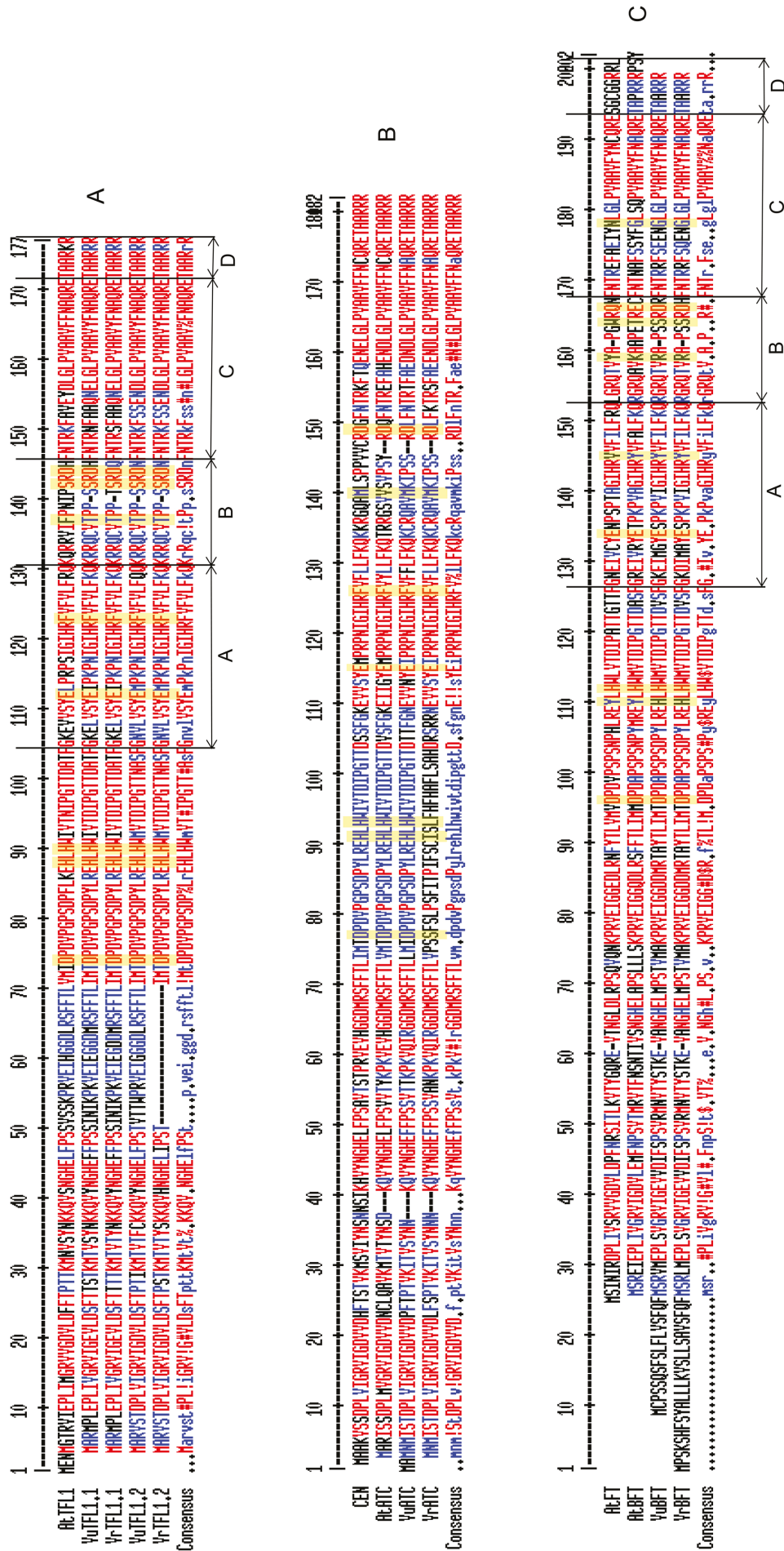


Fig. 4. Multiple amino acid sequences alignment. A — TFL1-like proteins, B — ATC-like proteins, C — BFT-like proteins. Red is high consensus, blue — low consensus, black — neutral consensus. Yellow designates critical amino acid residues.

Tyr85-Gln140 in FT and His88-Asp144 in TFL1 play a key role in determining the protein function (Hanawa, Money and Bradley, 2005; Ahn et al., 2006; Ho and Weigel, 2014). However, despite the fact that in TFL1 of *Brassica L.* His88 is replaced by Arg (His88Arg), the function of the protein is preserved (Mimida, Sakamoto, Murata and Motoyoshi, 1999). In the present work, all critical amino acids are retained in the identified sequences of TFL1s of *V. unguiculata* and *V. radiata* (Fig. 4, Supplementary material 2). According to amino acid sequences these proteins are probably flowering repressors. Based on *in silico* analysis, including analysis of predicted elements in promoter regions and 3D protein structures, we noticed that the pattern of cis-promoter elements is similar for *VuTFL1.1* and *VuTFL1.2* (Fig. 3, Supplementary file 1). High homology of 3D structures of these proteins (the smallest RMSD) was identified (Supplementary file 4). Analysis of protein sequence of VrTFL1.2 showed 20 amino acid deletion, and the conservative domain was not identified (Fig. 4, Supplementary file 2). Besides that, the highest RMSD between 3D structures of VrTFL1.2 with other identified TFL1-like proteins (*VuTFL1.1*, VrDet1, *VuTFL1.2*) was noticed, and so for VrTFL1.2 a high degree of divergence of other aligned 3D protein structures was detected (Supplementary file 4). On the basis of our results we assume that protein VrTFL1.2 is probably nonfunctional. Additional research is required to clarify this supposition.

*In silico* analysis of identified protein sequence of VuATC showed that all critical amino acid residues are conserved and this protein contains non-tyrosine and non-tryptophan at position 134 and 138 based on alignment with AtFT (Fig. 4, Supplementary file 2). In the predicted sequence of the VrATC protein, three amino acids that are important for the protein functioning are replaced (His86Ile, His88Leu, Asp72Ser) (Supplementary materials 2). It is known that a charge change within the ligand-binding domain does not affect the function of FT (Ho and Weigel, 2014); therefore, replacing a negatively charged amino acid with an uncharged Ser in VrATC probably does not change its function. However, the 3D structure of VrATC compared to AtATC and VuATC was changed (Supplementary materials 3). Based on *in silico* analysis, including analysis of predicted elements in promoter regions, predicted amino acid sequences, putative functional domains and protein 3D structures we proposed pseudogenization of VrATC.

The role of BFT is still not fully understood. Based on the results of the first studies, it was suggested that BFT is more similar to FT than to TFL1 (Ahn et al., 2006). A comparison of the amino acid sequences of these proteins showed that in BFT Tyr is located at position 85 (corresponds to Tyr85 in FT). In addition, the sequence of the conserved B segment is more similar between BFT and FT than between BFT and TFL1 (Yoo et al., 2010). The

position homologous to Gln140 / Asp144 (FT / TFL1) in BFT is Glu141 (Ahn et al., 2006). Asp and Glu are negatively charged and have similar biochemical properties.

In the current research, two amino acids that are characteristic of the flowering activator proteins (134Tyr, 138Trp in FT) were not noted in the VuBFT and VrBFT (Supplementary material 2). The conservative amino acid Asn152, characteristic of FT proteins, has been replaced by glycine in BFT. A pair of amino acids critical for functioning as a flowering repressor, His88-Asp144 according to AtTFL1, is preserved in VuBFT and VrBFT.

Based on the obtained data, it can be assumed that most identified proteins could function as blockers of the transition to flowering. Additional studies are needed to establish the possible role of the VrTFL1.2 and VrATC proteins in the transition to the reproductive stage.

## Conclusion

Identification and analysis of genes responsible for the type of stem growth are required for the successful selection of modern varieties. The genes characterized in our study belong to the group of *TFL1*-like genes. In the genomes of all the studied members of the tribe Phaseoleae, we identified duplication of *TFL1* gene. *V. angularis* is an exception. The obtained results confirm the high evolutionary conservatism of genes involved in the molecular genetic control of the flowering initiation. Thus, our data based on *in silico* analysis — including analysis of nucleotide sequences, predicted elements in promoter regions, predicted amino acid sequences, putative functional domains and 3D protein structures — is necessary for the next step of studying the molecular genetic mechanisms underlying the maintenance of indeterminacy in the growth of the shoot apical meristem, as well as in the control of the transition to flowering.

## References

- Ahn, J. H., Miller, D., Winter, V. J., Banfield, M. J., Lee, J. H., Yoo, S. Y., Henz, S. R., Brady, R. L., and Weigel, D. 2006. A divergent external loop confers antagonistic activity on floral regulators FT and TFL1. *The EMBO Journal* 25:605–614. <https://doi.org/10.1038/sj.emboj.7600950>
- Alvarez, J., Guli, C. L., Yu, X. H., and Smyth, D. R. 1992. *Terminal flower*: a gene affecting inflorescence development in *Arabidopsis thaliana*. *The Plant Journal* 2(1):103–116. <https://doi.org/10.1111/j.1365-313X.1992.00103.x>
- Banfield, M. J. and Brady, R. L. 2000. The structure of *Antirrhinum* Centroradial Protein (CEN) suggests a role as a kinase regulator. *Journal of Molecular Biology* 297(5):1159–1170. <https://doi.org/10.1006/jmbi.2000.3619>
- Benlloch, R., Berbel, A., Serrano-Mislata, A., and Madueno, F. 2007. Floral initiation and inflorescence architecture: a comparative view. *Annals of Botany* 100(3):659–676. <https://doi.org/10.1093/aob/mcm146>
- Blázquez, M. A., Soowal, L. N., Lee, I., and Weigel, D. 1997. *LEAFY* expression and flower initiation in *Arabidopsis*. *Development* 124(19):3835–3844.

- Boukar, O., Fatokun, C. A., Roberts, P. A., Abberton, M., Huynh, B. L., Close, T. J., Kyei-Boahen, S., Higgins, T. J. V., and Ehlers J. D. 2015. Cowpea. In: Grain Legumes. Springer: 219–250. [https://doi.org/10.1007/978-1-4939-2797-5\\_7](https://doi.org/10.1007/978-1-4939-2797-5_7)
- Burlyayeva, M. O., Gurkina, M. V., Chebukin, P. A., Perchuk, I. N., and Miroschnichenko, E. V. 2019. New varieties of vegetable cowpea (*Vigna unguiculata* subsp. *sesquipedalis* (L.) Verdc.) and prospects of their cultivation in southern Russia. *Vegetable crops of Russia* 5:33–37. <https://doi.org/10.18619/2072-9146-2019-5-33-37> (In Russian)
- Chow, C. N., Lee, T. Y., Hung, Y. C., Li, G. Z., Tseng, K. C., Liu, Y. H., Kuo, P. L., Zheng, H. Q., and Chang, W. C. 2019. PlantPAN3.0: a new and updated resource for reconstructing transcriptional regulatory networks from ChIP-seq experiments in plants. *Nucleic Acids Research* 47(D1):D1155–D1163. <https://doi.org/10.1093/nar/gky1081>
- Corpet, F. 1988. Multiple sequence alignment with hierarchical clustering. *Nucleic Acids Research* 16(22):10881–10890. <https://doi.org/10.1093/nar/16.22.10881>
- Chung, K. S., Yoo, S. Y., Yoo, S. Y., Lee, J. S., and Ahn J. H. 2010. *BROTHER OF FT AND TFL1 (BFT)*, a member of the *FT/TFL1* family, shows distinct pattern of expression during the vegetative growth of *Arabidopsis*. *Plant Signaling and Behavior* 5(9):1102–1104. <https://doi.org/10.4161/psb.5.9.12415>
- Dhanasekar, P. and Reddy, K. S. 2015. A novel mutation in *TFL1* homolog affecting determinacy in cowpea (*Vigna unguiculata*). *Molecular Genetics and Genomics* 290(1):55–65. <https://doi.org/10.1007/s00438-014-0899-0>
- FAO Departments and Offices. FAO; 2020 [cited 2020 July]. Available from: <http://www.fao.org/faostat/ru/#data/QC>
- Felsenstein, J. 1985. Confidence limits on phylogenies: An approach using the bootstrap. *Evolution* 39:783–791. <https://doi.org/10.1111/j.1558-5646.1985.tb00420.x>
- Finn, R. D., Attwood, T. K., Babbitt, P. C., Bateman, A., Bork P., Bridge, A. J., Chang, H. Y., Dosztányi, Z., El-Gebali, S., Fraser, M., Gough, J., Haft, D., Holliday, G. L., Huang, H., Huang, X., Letunic, I., Lopez, R., Lu, S., Marchler-Bauer, A., Mi, H., Mistry, J., Natale, D. A., Necci, M., Nuka, G., Orengo, C. A., Park, Y., Pesseat, S., Piovesan, D., Potter, S. C., Rawlings, N. D., Redaschi, N., Richardson, L., Rivoire, C., Sangrador-Vegas, A., Sigrist, C., Sillitoe, I., Smithers, B., Squizzato, S., Sutton, G., Thanki, N., Thomas, P. D., Tosatto, S. C. E., Wu, C. H., Xenarios, I., Yeh, L. S., Young, S. Y., and Mitchell, A. L. 2017. InterPro in 2017-beyond protein family and domain annotations. *Nucleic Acids Research* 45(D1):D190–D199. <https://doi.org/10.1093/nar/gkw1107>
- Gonzales, M. D., Archuleta, E., Farmer, A., Gajendran, K., Grant, D., Shoemaker, R., Beavis, W. D., and Waugh, M. E. 2005. The Legume Information System (LIS): an integrated information resource for comparative legume biology. *Nucleic Acids Research* 33(Database issue):D660–D665. <https://doi.org/10.1093/nar/gki128>
- Goodstein, D. M., Shu, S., Howson, R., Neupane, R., Hayes, R. D., Fazo, J., Mitros, T., Dirks, W., Hellsten, U., Putnam, N., and Rokhsar, D. S. 2012. Phytosome: a comparative platform for green plant genomics. *Nucleic Acids Research* 40(D1):D1178–D1186. <https://doi.org/10.1093/nar/gkr944>
- Goretti, D., Silvestre, M., Collani, S., Langenecker, T., Méndez, C., Madueño, F., and Schmid, M. 2020. TERMINAL FLOWER1 functions as a mobile transcriptional cofactor in the shoot apical meristem. *Plant Physiology* 182(4):2081–2095. <https://doi.org/10.1104/pp.19.00867>
- Gumber, R. K. and Singh, S. 1997. Genetics of flowering patterns in pigeonpea: further evidence for two gene control. *Euphytica* 96:233–236. <https://doi.org/10.1023/A:1003086519511>
- Hanzawa, Y., Money, T., and Bradley, D. 2005. A single amino acid converts a repressor to an activator of flowering. *Proceedings of the National Academy of Sciences USA* 102(21):7748–7753. <https://doi.org/10.1073/pnas.0500932102>
- Higo, K., Ugawa, Y., Iwamoto, M., and Korenaga, T. 1999. Plant cis-acting regulatory DNA elements (PLACE) database. *Nucleic Acids Research* 27(1):297–300. <https://doi.org/10.1093/nar/27.1.297>
- Ho, W. W. H. and Weigel, D. 2014. Structural features determining flower-promoting activity of *Arabidopsis* FLOWERING LOCUS T. *The Plant Cell* 26(2):552–564. <https://doi.org/10.1105/tpc.113.115220>
- Huang, N. C., Jane, W. N., Chen, J., and Yu, T. S. 2012. *Arabidopsis thaliana* CENTRORADIALIS homologue (ATC) acts systemically to inhibit floral initiation in *Arabidopsis*. *The Plant Journal* 72(2):175–184. <https://doi.org/10.1111/j.1365-313X.2012.05076.x>
- Jin, S., Nasim, Z., Susila, H., and Ahn, J. H. 2021. Evolution and functional diversification of *FLOWERING LOCUS T/TERMINAL FLOWER1* family genes in plants. *Seminars in Cell and Developmental Biology* 109:20–30. <https://doi.org/10.1016/j.semcd.2020.05.007>
- Kang, Y. J., Kim, S. K., Kim, M. Y., Lestari, P., Kim, K. H., Ha, B. K., Jun, T. H., Hwang, W. J., Lee, T., Lee, J. Shim, S., Yoon, M. Y., Jang, Y. E., Han, K. S., Taeprayoon, P., Yoon, N., Somta, P., Tanya, P., Kim, K. S., Gwag, J. G., Moon, J. K., Lee, Y. H., Park, B. S., Bombarely, A., Doyle, J. J., Jackson, S. A., Schafleitner, R., Srinives, P., Varshney, R. K., and Lee, S. H. 2014. Genome sequence of mungbean and insights into evolution within *Vigna* species. *Nature Communications* 5:5443. <https://doi.org/10.1038/ncomms6443>
- Kang, Y. J., Satyawati, D., Shim, S., Lee, T., Lee, J., Hwang, W. J., Kim, S. K., Lestari, P., Laosatit, K., Kim, K. H., Ha, T. H., Chitkineni, A., Kim, M. Y., Ko, J.-M., Gwag, J.-G., Moon, J.-K., Lee, Y.-H., Park, B.-S., Varshney, R. K., and Lee, S. H. 2015. Draft genome sequence of adzuki bean, *Vigna angularis*. *Scientific Reports* 5:8069. <https://doi.org/10.1038/srep08069>
- Kersey, P. J., Allen, J. E., Christensen, M., Davis, P., Falin, L. J., Grabmueller, C., Hughes, D. S., Humphrey, J., Kerhornou, A., Khobova, J., Langridge, N., McDowall, M. D., Maheswari, U., Maslen, G., Nuhn, M., Ong, C. K., Paulini, M., Pedro, H., Toneva, I., Tuli, M. A., Walts, B., Williams, G., Wilson, D., Youens-Clark, K., Monaco, M. K., Stein, J., Wei, X., Ware, D., Bolser, D. M., Howe, K. L., Kulesha, E., Lawson, D., and Staines, D. M. 2014. Ensembl Genomes 2013: scaling up access to genome-wide data. *Nucleic Acids Research* 42(Database issue):D546–552. <https://doi.org/10.1093/nar/gkt979>
- Kushwah, S. N., Ahmad, I., and Ali, S. 2014. Characterization of promoter of *Terminal Flower1 (TFL1)* gene of *Arabidopsis*. *Research Journal of Biotechnology* 9(3):35–40.
- Kumar, S., Stecher, G., Li, M., Nknyaz, C., and Tamura, K. 2018. MEGA X: Molecular evolutionary genetics analysis across computing platforms. *Molecular Biology and Evolution* 35(6):1547–1549. <https://doi.org/10.1093/molbev/msy096>
- Kwak, M., Velasco, D., and Gepts, P. 2008. Mapping homologous sequences for determinacy and photoperiod sensitivity in common bean (*Phaseolus vulgaris*). *Journal of Heredity* 99(3):283–291. <https://doi.org/10.1093/jhered/esn005>
- Kwak, M., Toro, O., Debouck, D. G., and Gepts, P. 2012. Multiple origins of the determinate growth habit in domesticated common bean (*Phaseolus vulgaris*). *Annals of*

- Botany* 110(8):1573–1580. <https://doi.org/10.1093/aob/mcs207>
- Li, S., Ding, Y., Zhang, D., Wang, X., Tang, X., Dai, D., Jin, H., Lee, S. H., Cai, C., and Ma, J. 2018. Parallel domestication with a broad mutational spectrum of determinate stem growth habit in leguminous crops. *The Plant Journal* 96(4):761–771. <https://doi.org/10.1111/tpj.14066>
- Liu, B., Watanabe, S., Uchiyama, T., Kong F., Kanazawa, A., Xia, Z., Nagamatsu, A., Arai, M., Yamada T., Kitamura K., Masuta, C., Harada, K., and Abe, J. 2010. The soybean stem growth habit gene *Dt1* is an ortholog of *Arabidopsis* *TERMINAL FLOWER1*. *Plant Physiology* 153(1):198–210. <https://doi.org/10.1104/pp.109.150607>
- Lonardi, S., Munoz-Amatrian, M., Liang, Q., Shu, S., Wanamaker, S. I., Lo, S., Tanskanen, J., Schulman, A. H., Zhu, T., Luo, M. C., Alhakami, H., Ounit, R., Hasan, A. Md., Verdier, J., Roberts, P. A., Santos, J. R. P., Ndeve, A., Dolezel, J., Vrana, J., Hokin, S. A., Farmer, A. D., Cannon, S. B., and Close, T. J. 2019. The genome of cowpea (*Vigna unguiculata* [L.] Walp.). *The Plant Journal* 98(5):767–782. <https://doi.org/10.1111/tpj.14349>
- Mimida, N., Sakamoto, W., Murata, M., and Motoyoshi, F. 1999. *TERMINAL FLOWER 1*-like genes in *Brassica* species. *Plant Science* 142(2):155–162. [https://doi.org/10.1016/s0168-9452\(99\)00020-5](https://doi.org/10.1016/s0168-9452(99)00020-5)
- Mimida N., Goto K., Kobayashi Y., Araki, T., Ahn, J. H., Weigel, D., Murata, M., Motoyoshi, F., and Sakamoto, W. 2001. Functional divergence of the *TFL1*-like gene family in *Arabidopsis* revealed by characterization of a novel homologue. *Genes to Cells* 6(4):327–336. <https://doi.org/10.1046/j.1365-2443.2001.00425.x>
- Mir, R. R., Kudapa, H., Srikanth, S., Saxena, R. K., Sharma, A., Azam S., Saxena, K., Penmetsa, R. V., and Varshney, R. K. 2014. Candidate gene analysis for determinacy in pigeonpea (*Cajanus* spp.). *Theoretical and Applied Genetics* 127:2663–2678. <https://doi.org/10.1007/s00122-014-2406-8>
- Moraes, T. S., Dornelas, M. C., and Martinelli, A. P. 2019. *FT/TFL1*: calibrating plant architecture. *Frontiers in Plant Science* 10:1–6. <https://doi.org/10.3389/fpls.2019.00097>
- Munoz-Amatrian, M., Mirebrahim H., Xu P., Wanamaker, S. I., Luo, M. C., Alhakami, H., Alpert, M., Atokple, I., Batiño, B. J., Boukar, O., Bozdog, S., Cisse, N., Drabo, I., Ehlers, J. D., Farmer, A., Fatokun, C., Gu, Y. Q., Guo, Y. N., Huynh, B. L., Jackson, S. A., Kusi, F., Lawley, C. T., Lucas, M. R., Ma, Y., Timko, M. P., Wu, J., You, F., Barkley, N. A., Roberts, P. A., Lonardi, S., and Close, T. J. 2017. Genome resources for climate-resilient cowpea, an essential crop for food security. *The Plant Journal* 89(5):1042–1054. <https://doi.org/10.1111/tpj.13404>
- Périlleux, C., Bouché, F., Randoux, M., and Orman-Ligeza, B. 2019. Turning meristems into fortresses. *Trends in Plant Science*. 24(5):431–442. <https://doi.org/10.1016/j.tplants.2019.02.004>
- Pin, P. A., Benlloch, R., Bonnet, D., Wremerth-Weich, E., Kraft, T., Gielen, J. J., and Nilsson, O. 2010. An antagonistic pair of FT homologs mediates the control of flowering time in sugar beet. *Science* 330(6009):1397–1400. <https://doi.org/10.1126/science.1197004>
- Ratcliffe, O. J., Amaya, I., Vincent, C. A., Rothstein, S., Carpenter, R., Coen, E. S., and Bradley, D. J. 1998. A common mechanism controls the life cycle and architecture of plants. *Development* 125(9):1609–1615.
- Repinski, S. L., Kwak, M., and Gepts, P. 2012. The common bean growth habit gene *PvTFL1y* is a functional homolog of *Arabidopsis* *TFL1*. *Theoretical and Applied Genetics* 124(8):1539–1547. <https://doi.org/10.1007/s00122-012-1808-8>
- Ryu, J. Y., Park, C. M., and Seo P. J. 2011. The floral repressor *BROTHER OF FT* and *TFL1* (*BFT*) modulates flowering initiation under high salinity in *Arabidopsis*. *Molecular Cells* 32(3):295–303. <https://doi.org/10.1007/s10059-011-0112-9>
- Ryu, J. Y., Lee, H. J., Seo, P. J., Jung, J. H., Ahn, J. H., and Park, C. M. 2014. The *Arabidopsis* floral repressor *BFT* delays flowering by competing with *FT* for *FD* binding under high salinity. *Molecular Plant* 7(2):377–387. <https://doi.org/10.1093/mp/sst114>
- Saitou, N. and Nei, M. 1987. The neighbor-joining method: A new method for reconstructing phylogenetic trees. *Molecular Biology and Evolution* 4(4):406–425. <https://doi.org/10.1093/oxfordjournals.molbev.a040454>
- Sakai, H., Naito, K., Ogiso-Tanaka, E., Takahashi, Y., Iseki, K., Muto, C., Satou, K., Teruya, K., Shiroma, A., Shimoji, M., Hirano, T., Itoh, T., Kaga, A., and Tomooka, N. 2015. The power of single molecule real-time sequencing technology in the de novo assembly of a eukaryotic genome. *Scientific Reports* 5:16780. <https://doi.org/10.1038/srep16780>
- Sakai, H., Naito, K., Takahashi, Y., Sato, T., Yamamoto, T., Muto, I., Itoh, T., and Tomooka, N. 2016. The *Vigna* Genome Server, 'VigGS': A genomic knowledge base of the genus *Vigna* based on high-quality, annotated genome sequence of the azuki bean, *Vigna angularis* (Willd.) Ohwi & Ohashi. *Plant Cell Physiology* 57(1):e2. <https://doi.org/10.1093/pcp/pcv189>
- Schmutz, J., Cannon, S. B., Schlueter, J., Ma, J., Mitros, T., Nelson, W., Hyten, D. L., Song, Q., Thelen, J. J., Cheng, J., Xu, D., Hellsten, U., May, G. D., Yu, Y., Sakurai, T., Umezawa, T., Bhattacharyya, M. K., Sandhu, D., Valliyodan, B., Lindquist, E., Peto, M., Grant, D., Shu, S., Goodstein, D., Barry, K., Futrell-Griggs, M., Abernathy, B., Du, J., Tian, Z., Zhu, L., Gill, N., Joshi, T., Libault, M., Sethuraman, A., Zhang, X.-C., Shinozaki, K., Nguyen, H. T., Wing, R. A., Cregan, P., Specht, J., Grimwood, J., Rokhsar, D., Stacey, G., Shoemaker, R. C., and Jackson S. A. 2010. Genome sequence of the palaeopolyploid soybean. *Nature* 463:178–183. <https://doi.org/10.1038/nature08670>
- Schmutz, J., McClean, P. E., Mamidi, S. G., Wu, A., Cannon, S. B., Grimwood, J., Jenkins, J., Shu, S., Song, Q., Chavarro, C., Torres-Torres, M., Geffroy, V., Moghaddam S. M., Gao, D., Abernathy, B., Barry, K., Blair, M., Brick, M. A., Chovatia, M., Gepts, P., Goodstein, D. M., Gonzales, M., Hellsten U., Hyten, D. L., Jia, G., Kelly, J. D., Kudrna, D., Lee, R., Richard, M. M. S., Miklas, P. N., Osorno, J. N. Rodrigues, J., Thareau, V., Urrea, C. A., Wang, M., Yu, Y., Zhang, M., Wing, R. A., Cregan, P. B., Rokhsar, D. S., and Jackson, S. A. 2014. A reference genome for common bean and genome-wide analysis of dual domestications. *Nature Genetics* 46:707–713. <https://doi.org/10.1038/ng.3008>
- Tamura, K., Nei, M., and Kumar, S. 2004. Prospects for inferring very large phylogenies by using the neighbor-joining method. *Proceedings of the National Academy of Sciences USA* 101(30):11030–11035. <https://doi.org/10.1073/pnas.0404206101>
- Tian, Z., Wang, X., Lee, R., Li, Y., Specht, J. E., Nelson, R. L., McClean, P. E., Qiu, L., and Ma, J. 2010. Artificial selection for determinate growth habit in soybean. *Proceedings of the National Academy of Sciences USA* 107(19):8563–8568. <https://doi.org/10.1073/pnas.1000088107>
- Yoo, S. J., Chung, K. S., Jung, S. H., Yoo, S. Y., Lee, J. S., and Ahn, J. H. 2010. *BROTHER OF FT AND TFL1* (*BFT*) has *TFL1*-like activity and functions redundantly with *TFL1* in inflorescence meristem development in *Arabidopsis*. *The Plant Journal* 63(2):241–253. <https://doi.org/10.1111/j.1365-313X.2010.04234.x>

- Varshney, R. K., Chen, W., Li, Y., Bharti, A. K., Saxena, R. K., Schlueter, J. A., Donoghue, M. T., Azam, S., Fan, G. and Whaley, A. M. 2012. Draft genome sequence of pigeonpea (*Cajanus cajan*), an orphan legume crop of resource-poor farmers. *Nature Biotechnology* 30:83. <https://doi.org/10.1038/nbt.2022>
- Vasconcelos, E. V., De Andrade Fonseca, A. F., Pedrosa-Harand, A., De Andrade Bortoleti, K. C., Benko-Iseppon, A. M., Da Costa, A. F., and Brasileiro-Vidal, A. C. 2015. Intra- and interchromosomal rearrangements between cowpea [*Vigna unguiculata* (L.) Walp.] and common bean (*Phaseolus vulgaris* L.) revealed by BAC-FISH. *Chromosome Research* 23(2):253–266. <https://doi.org/10.1007/s10577-014-9464-2>
- Vishnyakova, M. A., Aleksandrova, T. G., Buravtseva, T. V., Burlyayeva, M. O., Egorova, G. P., Semenova, E. V., Seferova, I. V., and Suvorova, G. N. 2019. Species diversity of the VIR collection of grain legume genetic resources and its use in domestic breeding. *Proceedings on Applied Botany, Genetics and Breeding* 180(2):109–123. <https://doi.org/10.30901/2227-8834-2019-2-109-123> (In Russian)
- Waterhouse, A. Bertonni, M., Bienert, S., Studer, G., Tauriello, G., Gumienny, R., Heer, F. T, de Beer, T. A P, Rempfer, C., Bordoli, L., Lepore, R., and Schwede, T. 2018. SWISS-MODEL: homology modelling of protein structures and complexes. *Nucleic Acids Research* 46(W1):W296–W303. <https://doi.org/10.1093/nar/gky427>
- Wickland, D. P. and Hanzawa, Y. 2015. The *FLOWERING LOCUS T / TERMINAL FLOWER 1* gene family: functional evolution and molecular mechanisms. *Molecular Plant* 8:983–997. <https://doi.org/10.1016/j.molp.2015.01.007>
- Weigel, D., Alvarez, J., Smyth, D. R., Yanofsky, M. F., and Meyerowitz, E. M. 1992. *LEAFY* controls floral meristem identity in *Arabidopsis*. *Cell* 69(5):843–859. [https://doi.org/10.1016/0092-8674\(92\)90295-n](https://doi.org/10.1016/0092-8674(92)90295-n)
- Weigel, D. and Nilsson, O. A. 1995. Developmental switch sufficient for flower initiation in diverse plants. *Nature* 377:495–500. <https://doi.org/10.1038/377495a0>

## SUPPLEMENTS

## Supplementary file 1

*cis*-acting regulatory DNA elements identified in *TFL1*-like gene promoters of *Vigna unguiculata* and *V. radiata*. The localization of promoter motifs on 5'-upstream region before ATG transcription start site. Analysis was performed using New PLACE database. “+” — coding chain, “-” — noncoding chain

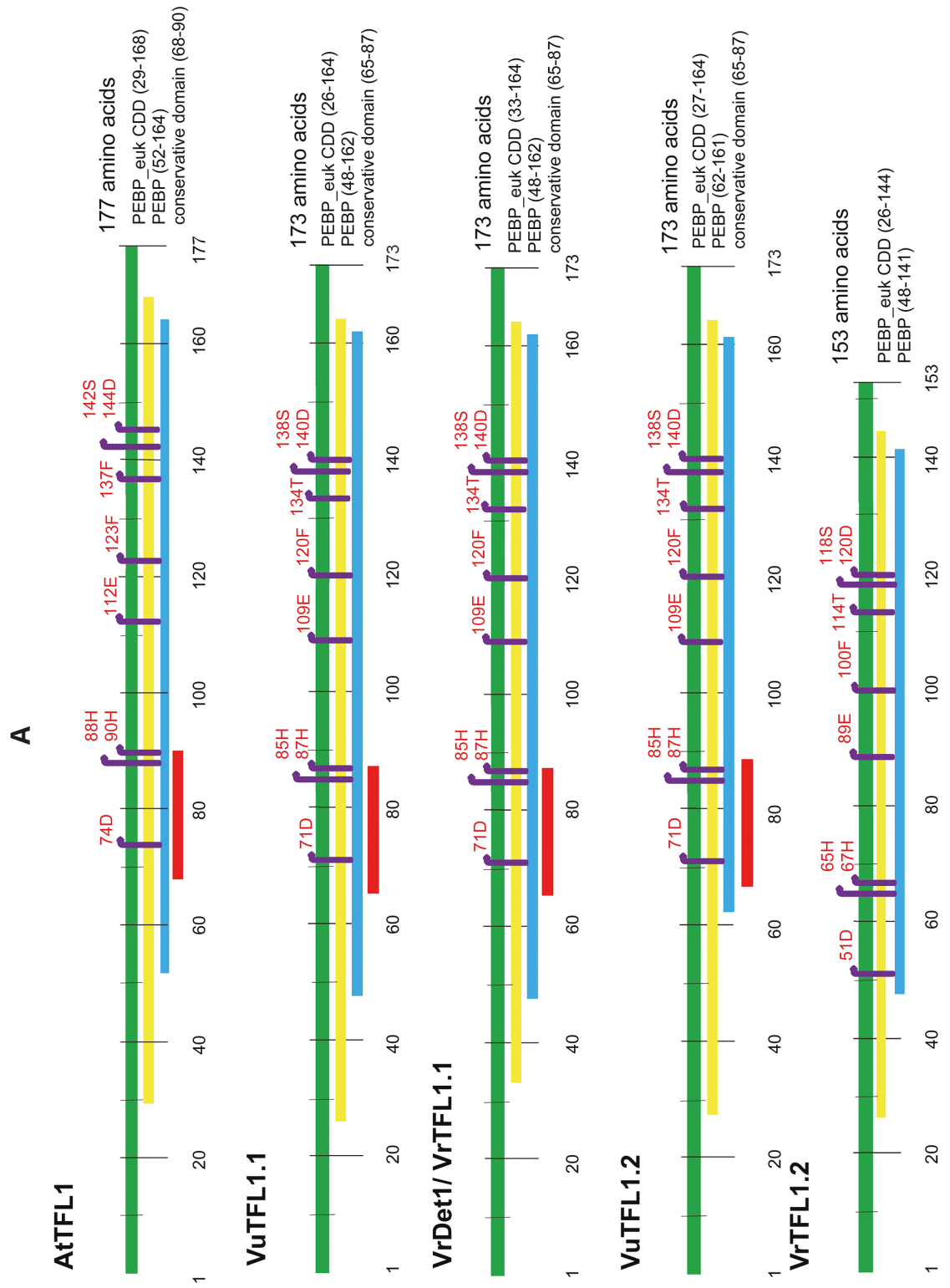
	Motif	Sequence	Gene	Position	Description
light-responsive sites	GATABOX	GATA	<i>TFL1.1</i>	67 (-) 347 (+) 654 (-) 956 (+)	S000039
			<i>TFL1.2</i>	70 (-) 289 (-) 366 (+) 437 (+) 439 (+) 505 (+) 576 (+) 588 (+) 646 (-) 648 (+) 658 (+) 731 (+) 807 (+) 929 (+) 990 (+)	
			<i>ATC</i>	219 (+) 283 (-) 306 (-) 314 (-) 316 (+) 378 (-) 498 (-) 500 (+) 666 (+) 739 (+) 750 (+) 859 (-) 920 (+)	
			<i>BFT</i>	115 (+) 231 (+) 265 (+) 282 (-) 357 (-) 369 (+) 429 (-) 434 (+) 454 (+) 576 (-) 588 (-) 688 (-) 715 (+) 733 (+) 826 (-) 911 (+)	
	GT1CONSENSUS	GRWAAW	<i>TFL1.1</i>	179 (-) 180 (-) 317 (-) 347 (+) 360 (+) 489 (+) 559 (-) 578 (+) 606 (-) 696 (-) 697 (-) 874 (+)	S000198
			<i>TFL1.2</i>	42 (-) 43 (-) 59 (-) 174 (+) 184 (+) 291 (-) 306 (+) 325 (+) 366 (+) 463 (+) 492 (-) 578 (-) 658 (+) 779 (-) 807 (+) 912 (+) 981 (-) 990 (+)	
			<i>ATC</i>	174 (-) 228 (-) 347 (-) 353 (-) 406 (-) 407 (+) 527 (+) 649 (-) 791 (-) 920 (+) 992 (+)	
			<i>BFT</i>	56 (-) 208 (-) 231 (+) 284 (-) 396 (+) 429 (+) 443 (-) 508 (-) 735 (-) 828 (-)	
	IBOX	GATAAG	<i>TFL1.1</i>	656 (-)	S000124
			<i>BFT</i>	115 (+) 454 (+)	
	IBOXCORE	GATAA	<i>TFL1.1</i>	347 (+) 655 (-)	S000199
			<i>TFL1.2</i>	71 (-) 290 (-) 366 (+) 506 (-) 577 (-) 658 (+) 733 (-) 990 (-)	
			<i>ATC</i>	307 (-) 920 (+)	
			<i>BFT</i>	115 (+) 231 (+) 283 (-) 429 (+) 454 (+) 734 (-) 827 (-)	
	IBOXCORENT	GATAAGR	<i>BFT</i>	454 (+)	S000424
	INRNTPSADB	YTCANTYY	<i>TFL1.1</i>	37 (+) 44 (+)	S000395
			<i>TFL1.2</i>	189 (-) 311 (-) 429 (-) 497 (-) 886 (+)	
			<i>ATC</i>	350(+) 390 (-)	
			<i>BFT</i>	9 (+) 446 (+) 550 (+) 847 (-)	
SORLREP3AT	TGTATATAT	<i>BFT</i>	264 (-)	S000488	

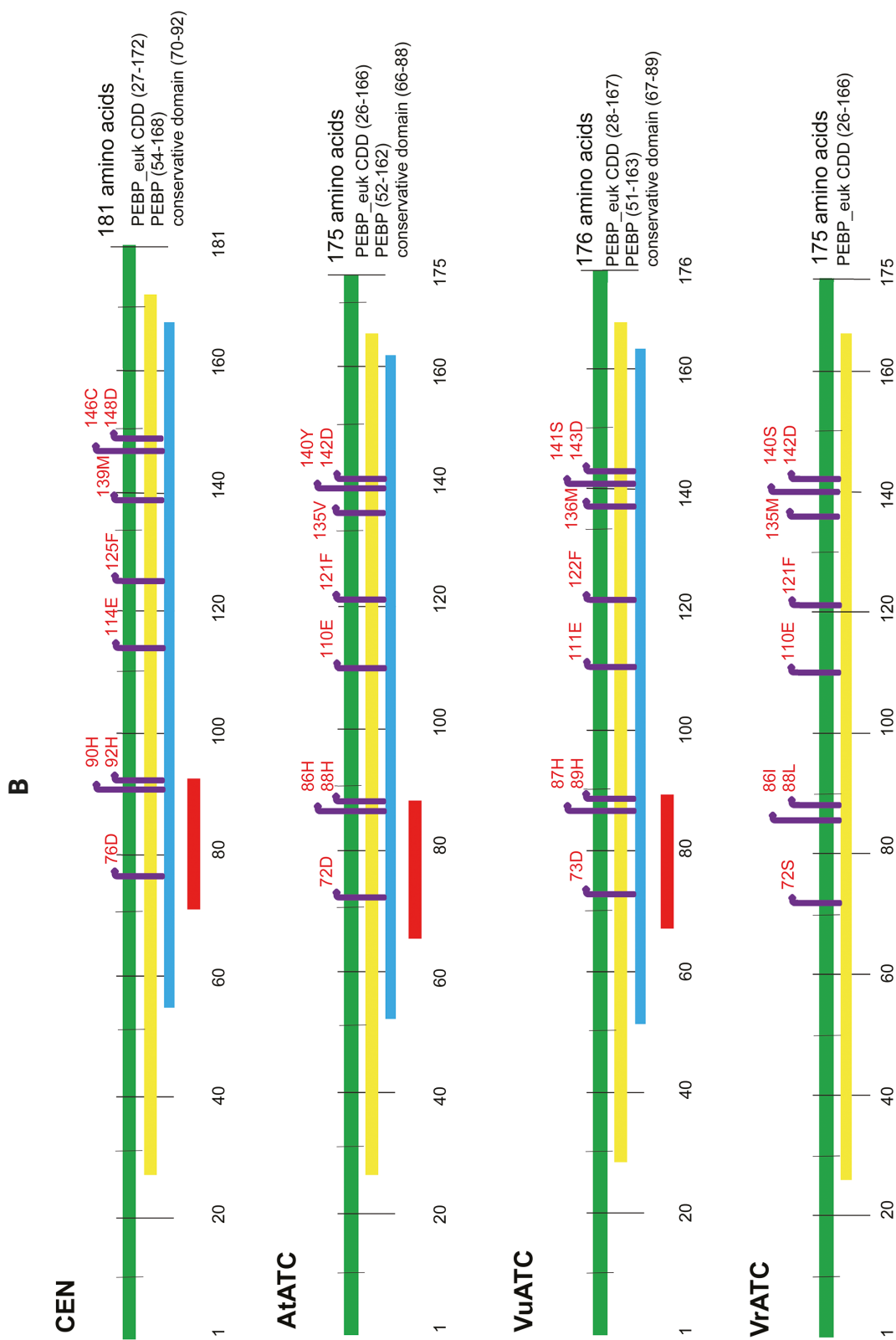


	Motif	Sequence	Gene	Position	Description
hormone-responsive sites	ASF1MOTIFCAMV	TGACG	ATC	777 (+)	S000024
	CPBCSPOR	TATTAG	TFL1.2	946 (-) 958 (-)	S000491
			ATC	617 (-)	
			BFT	521 (+)	
	ERELEE4	AWTTCAAA	ATC	338 (+)	S000037
	GAREAT	TAACAAR	TFL1.2	193 (+) 801 (-)	S000439
			ATC	311 (-) 473 (-) 489 (-)	
	GAREZOSREP1	TAACGTA	BFT	314 (-)	S000420
	MYCATRD22	CACATG	TFL1	616 (+) 950 (-)	S000174
			ATC	662 (+)	
BFT			422 (-)		
PROXBBNAPA	CAAACACC	TFL1.2	752 (+)	S000263	
WBOXATNPR1	TTGAC	TFL1.2	670 (-)	S000390	
		ATC	421 (-) 505 (-) 778 (+)		
		BFT	173 (+) 663 (+) 992 (-)		
TF binding motifs	ARR1AT	NGATT	TFL1.1	169 (-) 219 (+) 320 (+) 443 (-) 622 (+) 862 (-) 892 (-)	S000454
			TFL1.2	489 (-) 595 (-) 743 (-) 787 (-) 834 (+) 867 (+)	
			ATC	8 (-) 141 (+) 259 (+) 331 (-) 409 (+) 540 (+) 572 (+) 582 (+) 621 (-) 793 (+) 837 (+) 848 (+) 936 (-)	
			BFT	12 (-) 100 (-) 295 (+) 425 (-) 450 (+) 466 (-) 498 (+) 696 (-) 789 (-) 798 (-)	
	ATHB1ATCONSENSUS	CAATWATTG	TFL1.2	668 (-)	S000317
	ATHB5ATCORE	CAATNATTG	TFL1.2	668 (-)	S000371
	CRTDREHVCBF2	GTCGAC	TFL1.1	296 (-)	S000411
	DPBFCOREDCDC3	ACACNNG	TFL1.1	53 (-) 89 (+) 617 (+) 950 (-)	S000292
	HDZIP2ATATHB2	TAATMATTA	ATC	932 (-)	S000373
			BFT	769 (+)	
RAV1AAT	CAACA	TFL1.1	223 (+) 246 (-)	S000314	
		TFL1.2	12 (-) 408 (+)		
		ATC	66 (+) 577 (-) 703 (+) 825(-) 830 (-) 840 (-)		
		BFT	252 (+)		

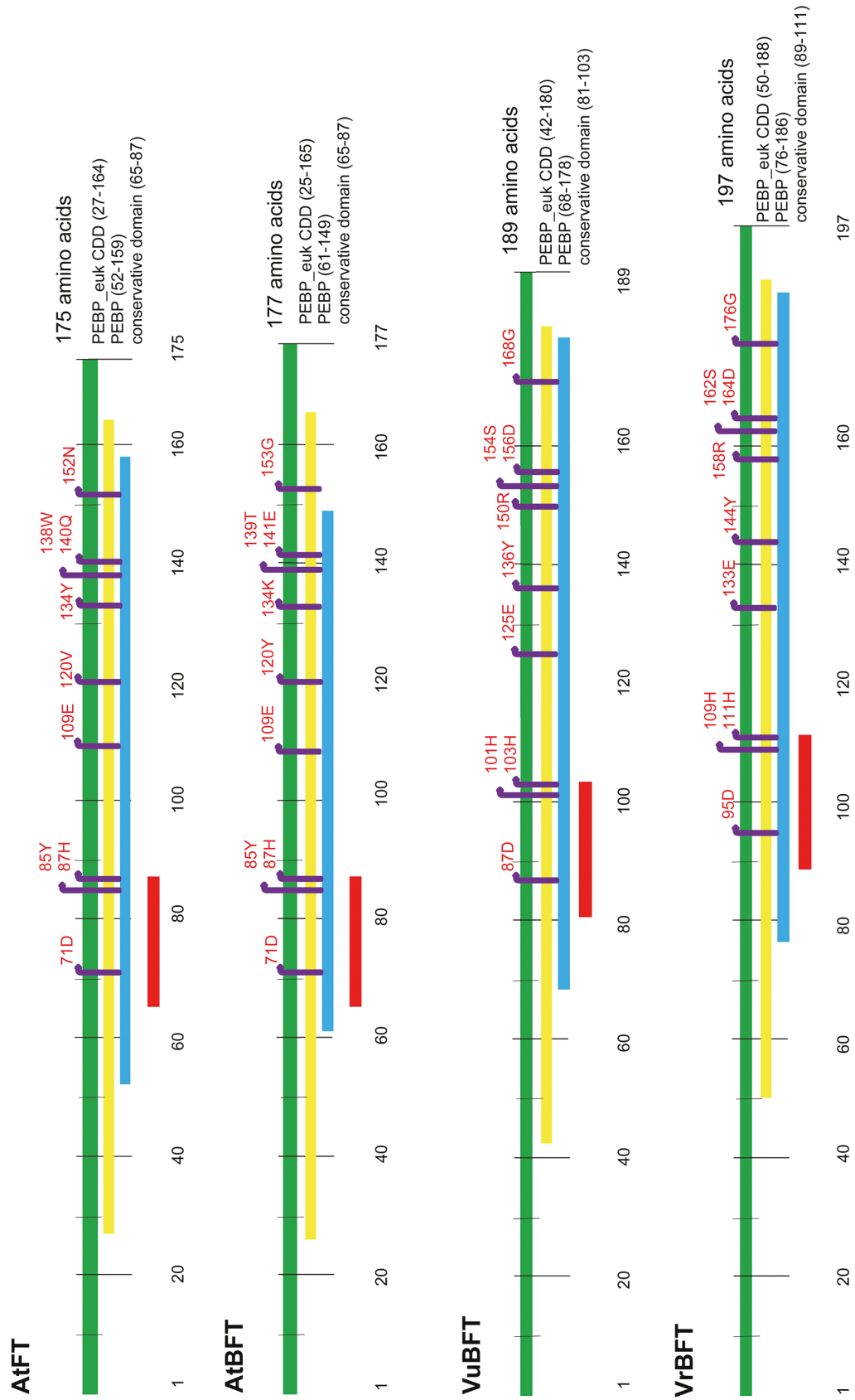
## Supplementary file 2

Annotation of functional domains (A) TFL1-like proteins, (B) ATC-like proteins, (C) BFT-like proteins





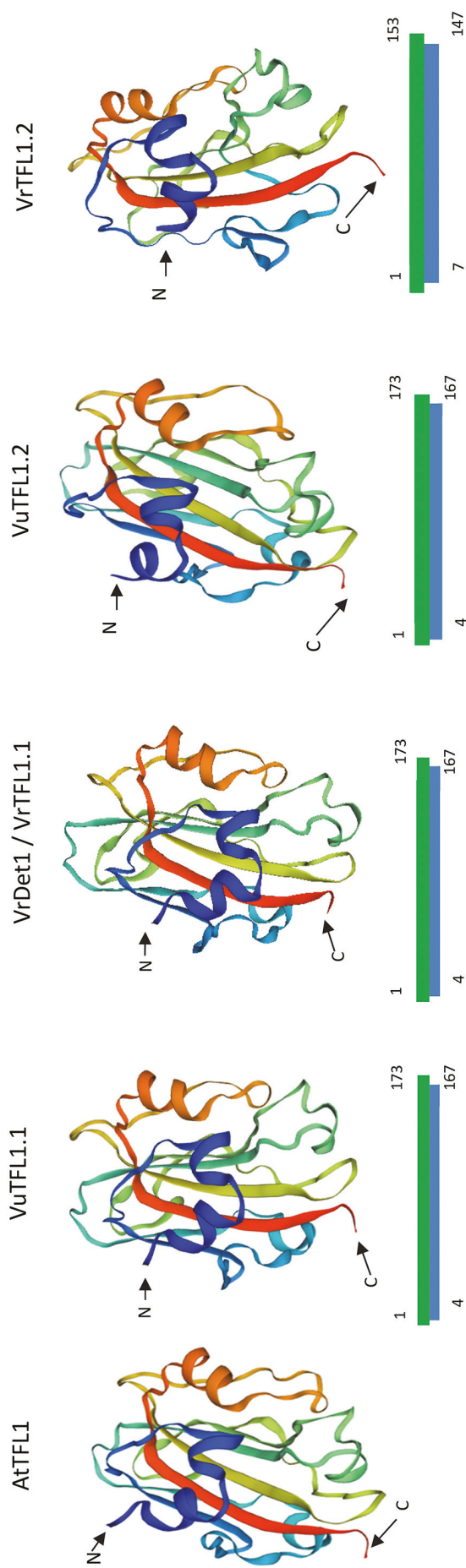
## C



**Supplementary file 3**

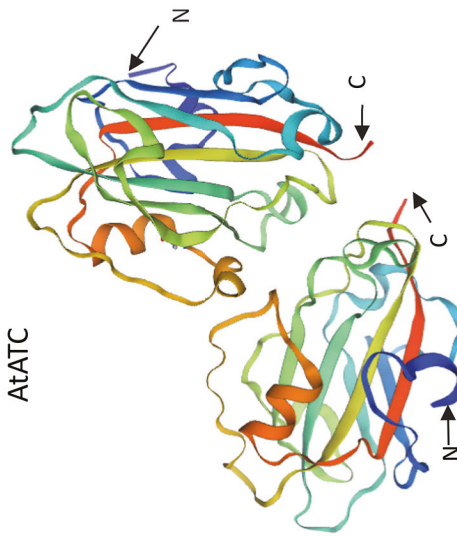
(A) Predicted 3D structures of *TFL1* gene products of *V. unguiculata* and *V. radiata*. 1wko.1.A template from PDB. (B) Predicted 3D structures of *ATC* gene products of *V. unguiculata* and *V. radiata*. 1qou.1 template from PDB. (C) Predicted 3D structures of *BFT* products of *V. unguiculata* and *V. radiata*. 6igi.1 template from PDB. Green line shows the length of investigating protein sequence as well as blue it's coverage by applied template

**A**

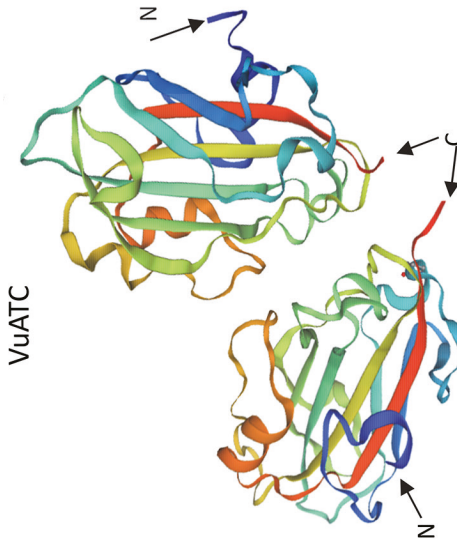


**B**

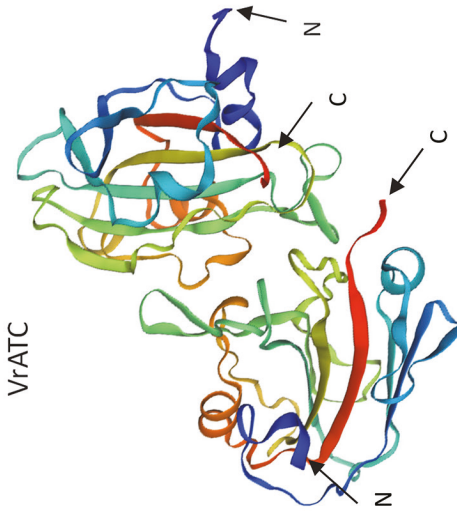
AtATC



VuATC

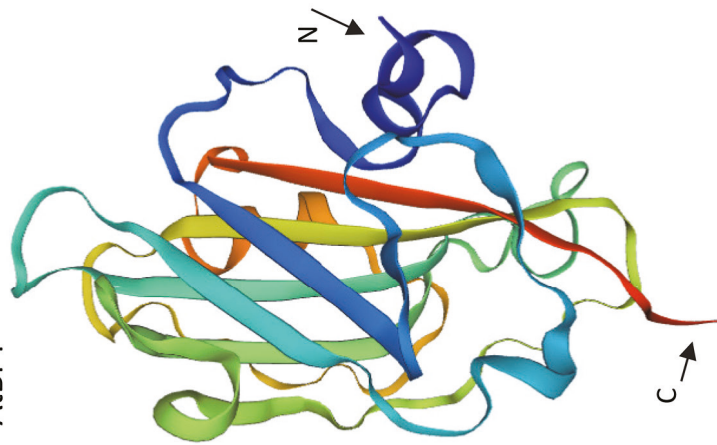


VrATC

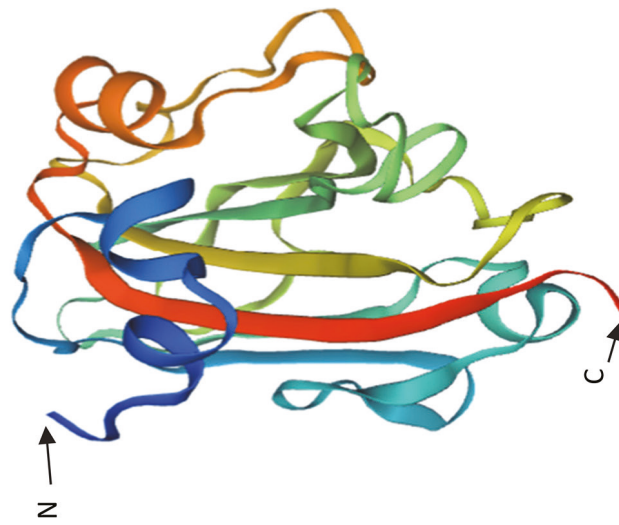


C

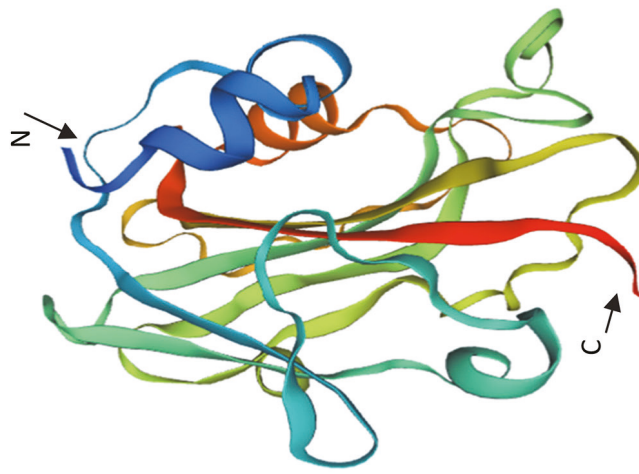
AtBFT



VuBFT



VrBFT



## Supplementary file 4

Protein sequence identity and overall root-mean-square deviation (RMSD) (A) TFL1-like proteins, (B) ATC-like proteins, (C) BFT-like proteins

### A

	AtTFL1	VuTFL1.1	VrDet1 / VrTFL1.1	VuTFL1.2	VrTFL1.2
AtTFL1	—	92/0.190	91.4/0.168	89.1/0.191	79.9/0.741
VuTFL1.1	92/0.19	—	<b>99.4</b> /0.122	94.8/ <b>0.034</b>	86.1/0.716
VrDet1	91.4/0.168	<b>99.4</b> /0.122	—	94.2/0.128	86.1/0.726
VuTFL1.2	89.1/0.191	94.8/ <b>0.034</b>	94.2/0.128	—	87.3/0.714
VrTFL1.2	79.9/0.741	86.1/0.716	86.1/0.726	87.3/0.714	—

### B

	AtATC	VuATC	VrATC	CEN
AtATC	—	<b>90.3</b> /0.433	79.9/0.713	96.5/0.408
VuATC	<b>90.3</b> /0.433	—	86.6/0.655	89.8/ <b>0.187</b>
VrATC	79.9/0.713	86.6/0.655	—	80.1/0.651
CEN	96.5/0.408	89.8/ <b>0.187</b>	80.1/0.651	—

### C

	AtBFT	VuBFT	VrBFT	FT
AtBFT	—	86.8/0.639	87.4/0.953	83.5/0.586
VuBFT	86.8/0.639	—	<b>98.9</b> /0.861	86.6/0.534
VrBFT	87.4/0.953	<b>98.9</b> /0.861	—	87.2/1.013
FT	83.5/0.586	86.6/0.534	87.2/1.013	—