Research on the allelic variance of FT-like PEBP proteins in *Solanum tuberosum*

Research Report

Li Shi

Reg. Nr: 941129759050

Oct 2018

MSc Thesis

Plant Sciences Group

Laboratory of Plant Breeding

Breeding for Growth and Development, Potato Group

WAGENINGEN UNIVERSITY

Supervisor:

dr. C.W.B. Bachem

Examiners:

dr. C.W.B. Bachem

dr.ir. AB (Guusje) Bonnema

Table of Contents

List of figures

Abstract

Potato (*Solanum tuberosum*) has an important role in the human diet. Potato originates from the Andean regions of South America and evolved short-day-dependent tuber formation as a reproductive strategy. Due to linking tuberization with the well-characterized photoperiodic control of flowering in *Arabidopsis*, the differentiation of this storage organ in potato has become better understood in recent decades. StSP6A, a member of the *FLOWERING LOCUS T* (*FT*) gene family in potato, has been discovered and confirmed to act as a major component of the tuber-inducing signal, and is regulated by a related circadian clock gene and photoperiodic components involved in regulating tuberization in potato. It has been discovered that a homologue of FT called SP3D acts as a floral inducer in potato. Another homologue of FT in potato called SP5G acts as a repressor of SP6A in tuberization regulation. Apart from flowering inducing FT, another FT homologue called *TERMINAL FLOWER1* (*TFL1*) in *Arabidopsis* works as a repressor of flowering and StTFL is known to be involved in both flowering and tuberization regulation. In this research, we focused on allelic variance of StSP3D, StSP6A, StSP5G and StTFL in 3 tetraploid potato varieties: Altus, Columba and NRQ-11022.

The phylogenetic analysis was done by using the obtained DNA sequences of *StSP6A, StSP3D, StSP5G* and *StTLF*. The translated amino acid sequences of these 4 gene were used for allelic variance analysis. In this study, we found that all 4 of these proteins are quite conserved in these 3 tetraploid potato varieties. However, StSP6A has comparably more non-conservative changes than others and these non-conservative changes are potentially linked with protein post-translational modification, such as phosphorylation. Moreover, by using the I-TASSER sequence replacement modeling, we compared the 3D protein crystal structures of StSP6A, StSP3D, StSP5G, StTLF with each other and their reference proteins. The results from this experiment provide possible directions to gaining a better understanding of the expansion of FT in potato.

Key words: Potato, Tuberization, FT, PEBP.

1.Introduction

Potato (*Solanum tuberosum*) and many other tuber-bearing plant species have an important role in the human diet. For potato, tuberization, formation of the storage organs, is an important reproductive strategy (Longman, & Wilson. 1993). The transition from vegetative to reproductive growth is controlled by photoperiodic information *via* the circadian rhythm (Abelenda, Navarro, & Prat. 2014). Over the past years, we have seen a remarkable advance in understanding the differentiation of this storage organ, due to linking it with the well-characterized photoperiodic control of flowering in *Arabidopsis*. In this decade, several studies discovered and confirmed that a member of the *FLOWERING LOCUS T* (*FT*) gene family acts as a major component of the tuber-inducing signal and discovered the related circadian clock gene and photoperiodic components involved in regulating this 'tuberigen' signal.

FT is a member of the CETS (CENTRORADIALIS(CEN), TERMINAL FLOWER 1 (TFL1) and SELF PRUNING(SP)) protein family, which shares homology with mammalian phosphatidylethanolamine-binding protein (PEBP) (Karlgren *et al.* 2011). It has been shown that the family of PEBP is an evolutionary conserved group of proteins that occur in all taxa (Karlgren *et al.* 2011) and these PEBP genes seem to be function as regulators in controlling various growth and differentiation signaling pathway (Yeung *et al.* 1999, Chautard *et al.* 2004). In *Arabidopsis thaliana*, PEBP family consist of 6 members (see figure 1); With further research that has been done, it has been revealed 15 PEBP homologs in potato and 13 in tomato (see figure 1).

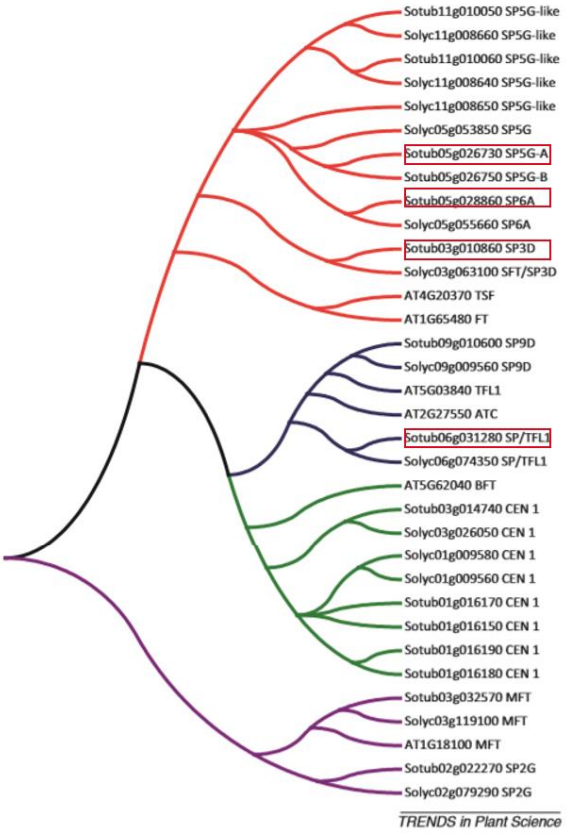

Figure 1. Phylogenetic tree showing sequence relationships among FT homologs (Abelenda, Navarro, & Prat. 2014). FT-like genes are grouped into three major clades: the FT-like, TFL (CEN)-like, and MFT-like clades. The interested gene in this study was boxed in red.

In *Arabidopsis*, flowering is initiated by the expression and transport of AtFT (Andrés, & Coupland, 2012). Orthologues of FT have been shown to regulate flowering across many different plant species, including *Hd3A* in rice (*Oryza sativa*) (Yano *et al.* 2000; Hayama *et al.* 2002; Kojima *et al.* 2002) and SP3D in tomato and potato. In potato, a homologue of *FT* (StSP6A) has been demonstrated to be the mobile 'tuberigen' signal that initiates tuberization and is regulated separately from the flowering pathway in potato (Navarro *et al.* 2011). Apart from flowering inducing FT, another FT homologue called *TERMINAL FLOWER1* (*TFL1*) in *Arabidopsis* works as a repressor of flowering and it maintains indeterminacy of the shoot apical meristem (Simon *et al.* 1996; Hanano & Goto. 2011). Notably, the repressor function of TFL1 can be convert into flowering promoting function by replacing a single amino acid at position 88 (His in TFL1, 85Tyr in FT) (Hanzawa, Money, & Bradley, 2005). In further studies, Ahn *et al.* (2006) have found that the most substantial differences between FT and TFL1 are in the external loop (residues 128–145), which corresponds to segment B, identified as critical for FT and TFL1 activity *in vivo*.

In 2011, Taoka *et al* constructed a model suggesting that a rice FT homolog, OsHd3a, interacts with the rice FD homolog, OsFD1, *via* 14-3-3 proteins (see figure 1A). This tri-protein complex (florigen activation complex, FAC) will activate a number of MADS box genes, including *APETALA1*, *FRUITFULL* and *SUPPRESSOR OF OVEREXPRESSION OF CONSTANS 1* to stimulate floral organ initiation (Wigge *et al.* 2005). However, *RICE CENTRORADIALIS* (*RCN*), rice TFL1-like proteins, competes with OsHd3a for 14-3-3 binding to repress flowering (Kaneko-Suzuki, *et al.* 2018) (see figure 1A). This model has been proven in temperate grasses (*Brachypodium distachyon*) by correctly predicting that inhibiting the binding to 14-3-3 protein would have a strong effect on reproduction timing (Qin *et al.* 2017). In potato, the StSP3D and StSP6A have been shown to promote flowering and tuberization, respectively. For flowering, we expect the StSP3D will form a complex analogous to the florigen activation complex (FAC) as described in rice to inducing flowering.

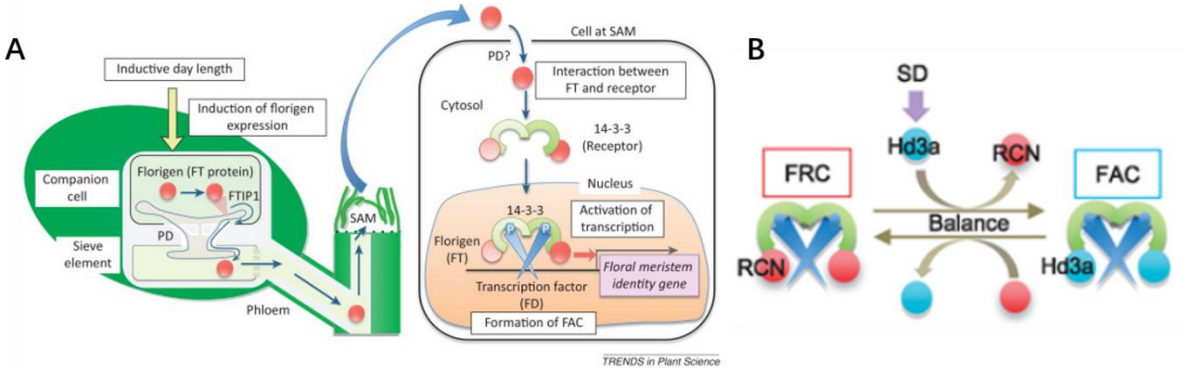

Figure 2. Model for the regulation of photoperiodic flowering by florigen. A. Under inductive day lengths, florigen (FT protein) is produced in the companion cells of vascular tissue. Then, FT moves to the shoot apical meristem (SAM) through the phloem. In SAM cells, FT interacts with 14-3-3 protein in the cytoplasm, next the FT–14-3-3 complex enters the nucleus to interact with FD. The resultant FT–14-3-3–FD complex (florigen activation complex, FAC) binds to the promoter regions of floral meristem identity genes (e.g., AP1), thereby activating their gene expression to promote flowering (Taoka et al. 2011; Taoka et al. 2013). B.in the cells of the SAM, RCN competes with Hd3a for 14-3-3 binding and represses florigenic activity. RCN forms the florigen repression complex (FRC) with 14-3-3 and OsFD1, whereas Hd3a forms the FAC. The balance between the FRC and FAC is regulated depending on the ratio of Hd3a to RCN in the cell (Kaneko-Suzuki, et al. 2018).

For tuberization, we have known for some time that environmental signals together with the inner clock regulate expression of the CONSTANS (CO) factor. Then, StCO controls activation of the StSP6A gene, *via* StSP5G, which has a repressor function (Abelenda, Navarro, & Prat. 2014). It is very interesting to note that the closely related FT homologs in potato function in a very different way. The knowledge of the difference between StSP6A, StSP5G, and StSP3D is limited. Recently, researchers have found that a FAC-like complex, the tuberigen activation complex (TAC), comprised of StSP6A, St14-3-3s and StFDL1, regulate potato tuber formation (Teo *et al.* 2016). The repressor function of StSP5G may result from a competition with StSP6A to bind with 14-3-3 proteins. Therefore, knowing the allelic variance at the 14-3-3 protein binding position of StSP6A, StSP5G, and StSP3D from 3 different tetraploid varieties with different earliness can provide more insight in tuber formation.

Apart from the critical binding sites with 14-3-3 proteins, segment B in the fourth exon of FT-like protein has a critical role in conferring biological specificity on FT and TFL1 (Ahn *et al.* 2006). Interestingly, they have also found segment B are highly variable in TFL1 orthologs but almost invariant in FT orthologs. It seems like this external loop, which is encoded by segment B, has important effect on FT function. The role of the FT fourth exon has not been fully described. The allelic variance at the fourth exon is worth extra attention.

Based on the background information, the formulated research questions are:

1. Are there variations in the coding region that will produce amino acid changes?
2. Are these changes conserved and are the amino acid changes likely to affect the protein function?
3. If there are non-conserved changes how does this affect the protein 3D structure?

2.Materials and methods

In this study, two experiments will be performed. First, blasting the target gene sequence with 3 tetraploid potato's full *de-novo* genome sequences to obtain scaffolds. Then, aligning the scaffolds with the target genes sequence to observe allelic variance of *StSP6A, StSP3D, StSP5G* and *StTLF*. In the second experiment, the obtained DNA sequence will be translated into amino acid sequence and searched for any change which may affect the protein function. Moreover, we will also build a 3D protein model for each allele of these three genes to have a better view of any differences.

2.1 Available sequence data

Table 1 Available *de-novo* genome sequences

Potato genotypes	Ploidy	Maturity
Altus	**Tetraploid**	**Late**
Columba	**Tetraploid**	**Early**
NRQ-11022	**Tetraploid**	**Unknown**

Table 2 Target genes

Gene name	Gene ID	Source
StSP6A	**Sotub05g028860.1.1**	**http://solanaceae.plantbiology.msu.edu/cgi-bin/gbrowse/potato/**
StSP3D	**Sotub03g010860.1.1**	
StSP5G(A)	**Sotub05g026730.1.1**	
SP/StTLF1	**Sotub06g031280.1.1**	

2.2 Primary protein structure analyzes

With the help of Richard Finkers and Danny Esselink, we have access to 3 tetraploid potato genome sequences. The scaffolds were obtained by using *StSP6A, StSP3D, StSP5G* and *StTLF* as query sequences in TBLASTX searches against the *de-novo* genome sequences. The obtained scaffolds will be aligned with FT-like PEBP sequences by using Megalign v7.1 version. The complete alignment of FT-like PEBP sequences will be manually edited using BioEdit 5.0.9 version (Hall 1991). Sequences will be translated in order to delimit the 5' and 3' noncoding sequences. The parts located upstream of the ATG and downstream of the stop codon was discarded. Introns was removed from the genomic sequences. Next, the translated amino acid sequences were aligned against each other. The alignment searched for amino acid changes, and the changes will be used to predict whether they are likely to affect the protein function.

2.3 Molecular model building

The amino acid sequences of StSP6A, StSP3D, StSP5G and StTLF were retrieved from the Spud DB database (Gene ID can be found in table 2). The structures of StSP6A, StSP3D, StSP5G and StTFL were homology modeled using the I-TASSER server (Zhang, 2008) using the crystal structure of 1WKP from Protein Data Bank as a template (Ahn *et al.* 2006).

2.4 Motif prediction

The PDB file of StSP6A, StSP3D, StSP5G and StTLF were retrieved from the I-TASSER server (Zhang, 2008). To screen the possible motifs, we used PDBeMotif website (http://www.ebi.ac.uk/pdbe-

site/pdbemotif/barChartPattern?pattern1=PS00006). The PDBeMotif relies on the information mainly from the PROSITE (Bairoch and Bucher, 1994)

2.4 Phylogenetic analysis of FT-like PEBP genes in 3 tetraploid varieties

The phylogenetic analysis was done by using MEGA X (https://www.megasoftware.net/home). Maximum Likelihood was the statistical method we used in this study.

3.Result

3.1 Allelic variance of FT-like PEBP proteins in potato

From 3 tetraploid potato genomes, we obtained 10 alleles of *StSP6A*, 7 alleles of *StSP3D*, 12 alleles of *StSP5G* and 8 alleles of *StTFL1*. All these DNA sequences have been aligned with their individual reference genes and then translated to amino acid sequence for further comparison. As can been seen from figure 3, StSP6A and StSP5G show more allelic variance in protein sequence than others. Comparing with their individual reference protein sequences, the StSP3D and StTFL alleles are mostly conserved and no obvious structure changes have been found.

In StSP6A alleles, there are non-conservative changes at 5 different positions. Based on protein structure modeling using the program I-TASSER (Zhang, 2008), these non-conservative changes do not seem likely to cause significant change in protein structure. Interestingly, by cross-comparing this result with the motif prediction results, we have found that four out of 5 nonconservative changes occurred on, or near potential phosphorylation sites of the protein (figure 1). All these 4 non-conservative changes have a direct effect on the predicted phosphorylation at or near the changed site. For example, most alleles of StSP6A have Ser at the 126 amino acid position, while in some others it was replaced with Leu (S126L). The former can be phosphorylated by Casein kinase II (CK-2) or protein kinase C. However, the Leu at position 126 can no longer be phosphorylated.

Moreover, a 13 amino acid (aa) deletion has been found in one StSP6A allele called Altus7076-2. These 13 aa deletions eliminate one of the critical binding sites for 14-3-3 protein and caused a significant change in protein structure. In addition, the 13 aa deletion also abolishes the anion-binding site which is considered as the signature of phosphatidylethanolamine-binding protein family. Therefore, this allelic variant protein may not be functional as a normal StSP6A protein.

Another Interesting finding is that StSP5G has significantly less predicted phosphorylation sites than other FT-like PEBP proteins. Based on the prediction, StSP5G only has 1 residue which has the potential to be phosphorylated. However, both StSP6A and StSP3D have 3 potential phosphorylated residues. The allelic variance of StSP5G is considerable higher since the nonconservative changes happens more often in StSP5G alleles. But the frequency of nonconservative changes happens in one column is low. Moreover, mutations happened at protein position 129 in StSP5G, which are considered to be non-conservative changes. These mutations might affect the StSP5G protein function.

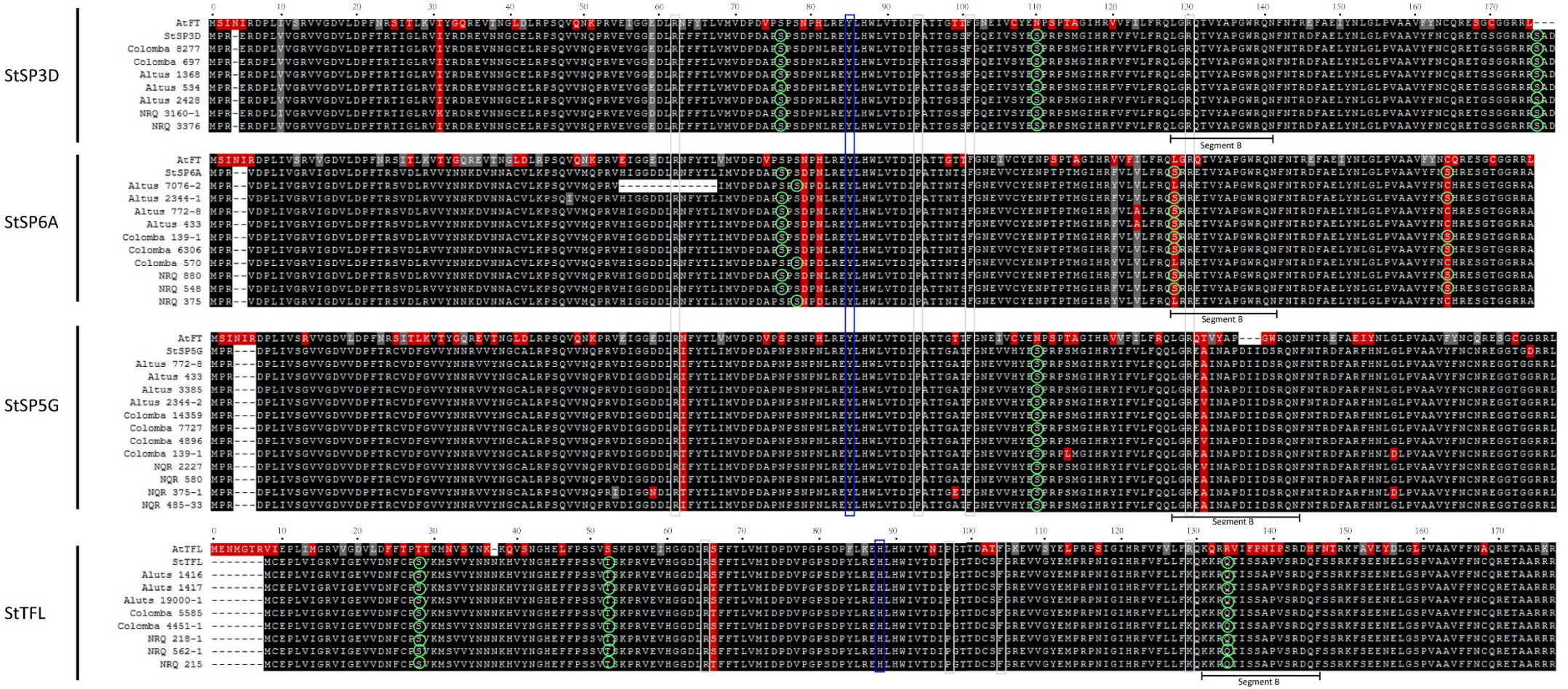

Figure 3. Allelic variance of FT-like PEBP proteins. Amino-acid sequences of the parental StSP3D, StSP6A, StSP5G and StTFL proteins. The white characters on black background indicate identical residues; The white characters on grey background indicate the residual has similar biochemical properties with other residuals in the same column (conservative changes); The white characters on red background indicate the nonconservative changes. The whole column background will be colored in grey or red when the percentage of sequences that must agree for identity or similarity lower than 80%. And dashes gaps introduced to optimize the sequence alignment. The green circles indicate the boxed residuals' potential to be phosphorylated (the possible phosphorylate site in AtFT and AtTFL not show in this figure). Tyr85/His88 is considered to be the most critical residue for distinguishing FT and TFL1 activity and it was boxed with blue. The four predicted most critical 14-3-3 binding sites have been boxed with light grey.

3.2 Comparison of protein structure of FT-like PEBP proteins

Ahn *et al.* (2006) discovered that the fourth exon of AtFT and AtTFL has a critical role in differentiating the function of these two closely related proteins. In their research and many other latter studies, segment B of the fourth exon was highlighted (Ahn *et al.* 2006; Taoka *et al.* 2013). In potato, the homologues of FT were shown to have functions beyond flowering. StSP3D and StSP6A have been known for inducing flowering and tuberization, respectively. However, another FT homologue called StSP5G works as a repressor for StSP6A in tuberization. Moreover, StTFL is known to be involved in both flowering and tuberization regulation. To have better understanding of how FT and TFL homologues developed diverse functions in potato, we compared primary protein structure to tertiary protein structure of StSP3D, StSP6A, StSP5G and StTFL with AtFT, OsHd3a and AtTFL.

As can been seen in figure 4A-F, all 3 potato FT homologues in potato have very similar primary and tertiary protein structures to AtFT. The differences in primary protein structure between species are more significant in TFL proteins (see figure 5B). The most notable differences that can be found between species are at the COOH-terminal ends (figure 4A-E and 5A). All StSP6A, StSP3D, StSP5G and StTFL proteins have an alpha helix (α-helix) tail at the C-terminus. Interestingly, OsHd3a also has an extra tail at C-terminal end but it did not form into the alpha helix structure (see figure 4A). However, this alpha helix tail did not appear in the AtFT protein C-terminal end, nor in AtTFL. Thus, the structure at the C-terminal end might have evolved differently in different species and the alpha helical tail might be involved in potato specific regulation.

Knowing the difference between StSP6A and StSP3D is important for understanding how they function in different ways. However, comparing these to protein structures with each other or with AtFT and OsHd3a, no significant structure difference could be found. The most notable structural difference in these FT homologues is the length and structure of alpha helix tail at C-terminal end. However, the 3 aa insertion in StSP5G at Segment B results in a significant shape change in the external loop (figure 4C).

In previous studies, both FT and TFL proteins have been known for binding with 14-3-3 and then FD (FAC) to promote flowering or repress flowering. In this experiment, we examined the most critical residues for binding with 14-3-3 proteins in all 4 FT-like PEBP proteins and used sequence-replaced modelling to test affinity. The results show that the most critical residues for binding with 14-3-3 proteins are quite conserved in StSP6A, StSP3D, StSP5G and StTFL. Therefore, all StSP6A, StSP3D, StSP5G and StTFL should be able to bind with the 14-3-3 protein. However, further experiments are required to confirm these interactions.

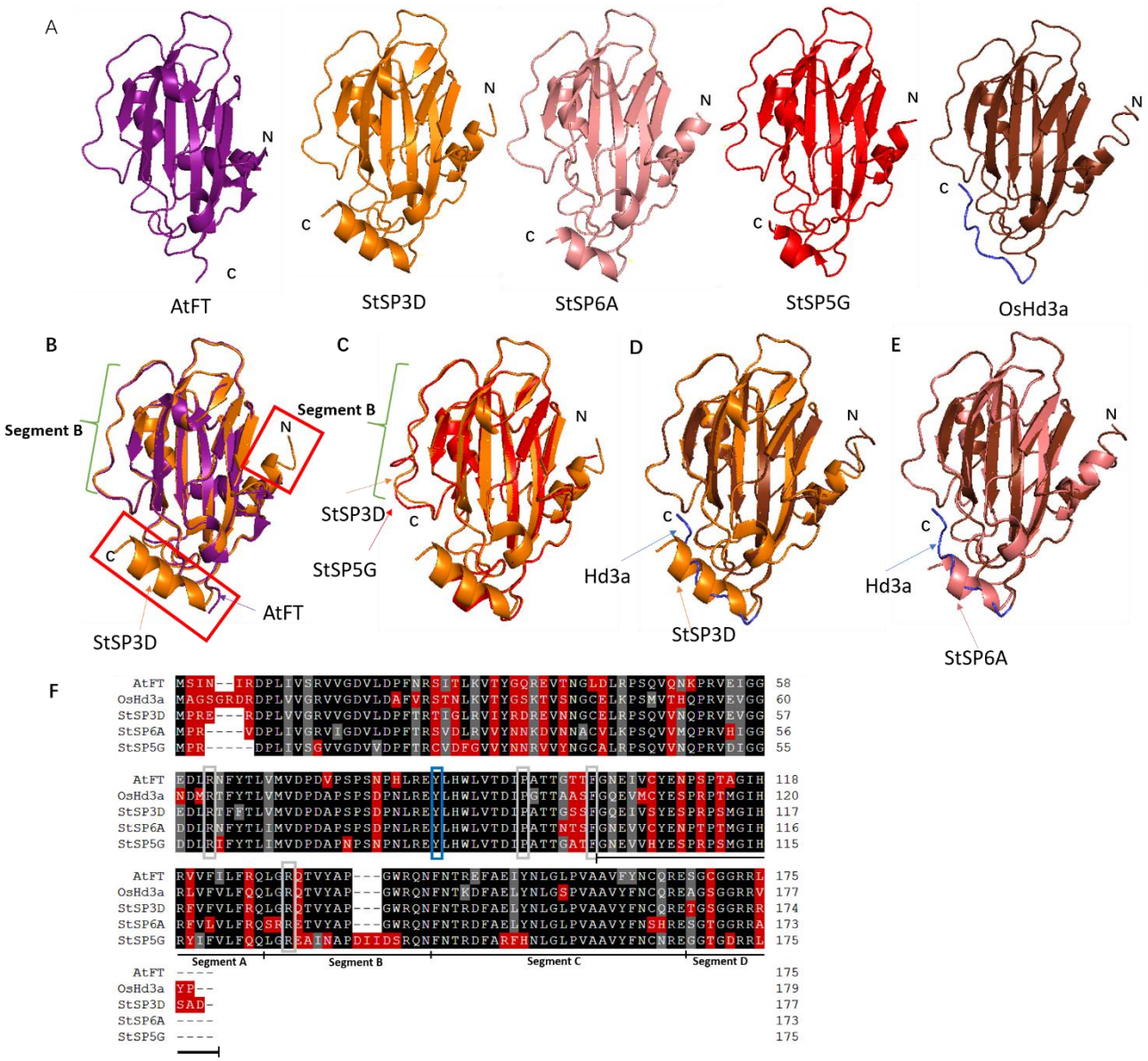

Figure 4. Protein sequence and tertiary structure. A. Cartoon diagrams of AtFT, StSP6A, StSP3D, StSP5G and OsHd3a. For better recognition, the C-terminal end of OsHd3a was colored in blue. B. Close-up showing an overlay of AtFT and StSP3D. C. Close-up showing an overlay of StSP3D and StSP5G. D. Close-up showing an overlay of OsHd3a and StSP3D. E. Close-up showing an overlay of OsHd3a and StSP6A. F. Amino-acid sequences of AtFT, StSP6A, StSP3D and StSP5G. white characters on a red background indicate identical residues; The black characters on yellow background indicate similar biochemical properties; And dashes gaps introduced to optimize the sequence alignment. The four segments of the fourth exon have been described in many other researches and they are shown as A, B, C and D. Tyr85, which is considered to be the most critical residue for distinguishing FT and TFL1 activity, is boxed with blue. The predicted most critical 14-3-3 binding sites have been boxed with grey

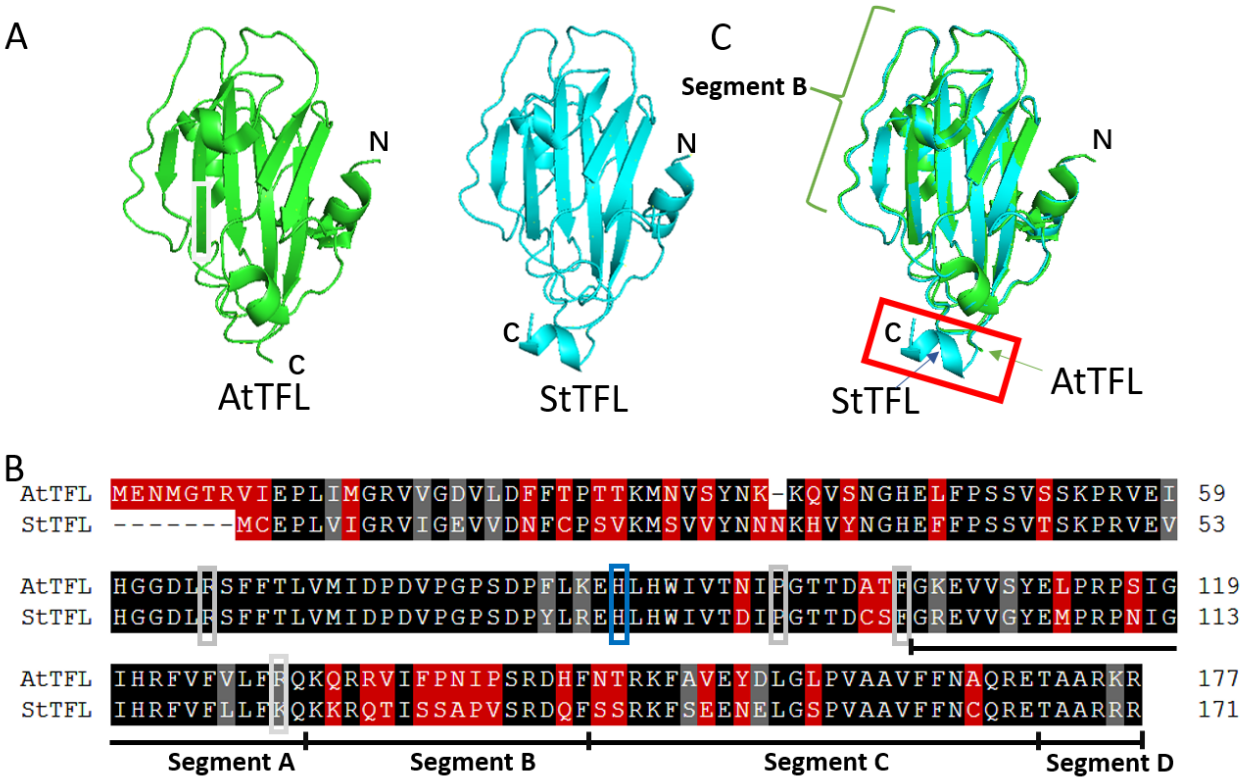

Figure 5. Protein sequence and secondary structure. A. Cartoon diagrams of AtTFL and StTFL. B. Amino-acid sequences of AtTFL and StTFL. The white characters on red background indicate identical residues; The black characters on yellow background indicate similar biochemical properties; And dashes gaps introduced to optimize the sequence alignment. The four segments of the fourth exon have been describe in many other researches and they are shown as A, B, C and D. 88His of AtTFL is considered to be the most critical residues for distinguishing FT and TFL1 activity and it was boxed with blue. The predicted most critical 14-3-3 binding sites have been boxed with grey. C. Close-up showing an overlay of AtTFL and StTFL.

3.3 Phylogenetic analysis of FT-like PEBP genes in 3 tetraploid varieties.

In order to have better a view of FT-like genes' differentiation, we did phylogenetic analysis. In total, 37 FT-like PEBP genes' sequences were identified for 4 genes. The reference sequences of *StSP6A, StSP3D, StSP5G* and *StTFL* were obtained from Spud DB dataset. Sequences of *AtTFL, AtFT* and *OsHd3a* have been used as reference.

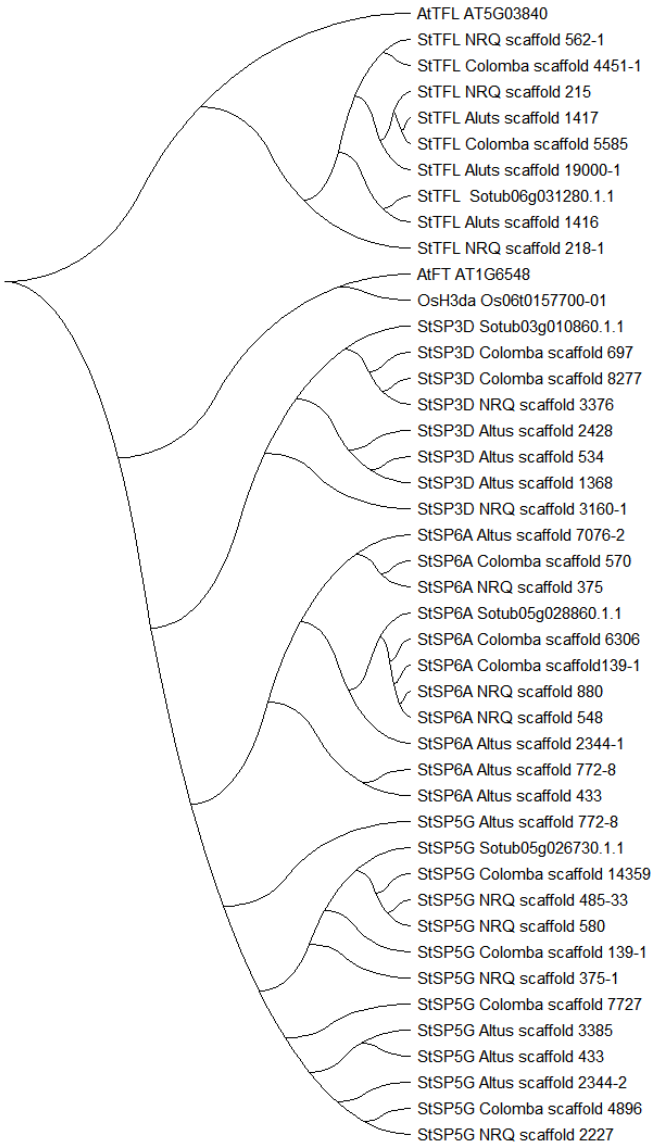

Figure 6. Phylogenetic tree of FT-like PEBP genes in 3 tetraploid varieties. FT-like PEBP genes are grouped into two clades: the FT like and the TFL like. The FT-like group includes members with a floral promotion function (AtFT, OsHd3a and StSP3D), tuberization inducer (StSP6A) and StSP5G. The TFL-like group includes StTFL and AtFT.

As can be seen in figure 6, FT-like PEBP genes are grouped into two clades: the FT like and the TFL like. The FT-like group includes members with a floral promotion function (AtFT, OsHd3a and StSP3D), tuberization inducer (StSP6A) and StSP5G. The TFL group includes AtTFL and StTFL genes. Moreover, compared to StSP3D, StSP6A and StSP5G, AtFT and OsHd3a show closer relationship to TFL group genes. This result contributes to explaining the diverse functions of FT like genes in potato. The different functions of StSP3D, StSP6A and StSP5G may benefit from having differences in their genetic code.

Notably, although StSP5G acts as a repressor in tuberization, it has more in common with FT in nucleotide sequence than with TFL. However, compared to StSP3D and StSP6A, StSP5G is more dissimilar with the reference genes of FT (AtFT and OsHd3a). In potato, StSP3D acts as a flowering inducer which showed higher similarity with two reference FT genes (AtFT, OsHd3a) than StSP6A. Moreover, in these three tetraploid varieties, StSP3D and TFL have less differentiation than StSP6A and StSP5G.

In conclusion, these five FT-like PEBP genes show a different variety level in the 3 examined tetraploid potatoes. In the FT-like clade, the sequences of *StSP6A, StSP3D* and *StSP5G* form separate clusters.

4.Discussion

4.1 the non-conservative changes in FT-like PEBP proteins

As a phosphatidylethanolamine-binding protein (PEBP) family member, FT protein crystal structure is very similar to mammalian PEBP (Ahn *et al.* 2006). Beside the highly conserved anion-binding pocket region, two other regions of the FT protein have been known for their critical impact on protein functions, segment B and Y85 (in *Arabidopsis*, H89 in AtTFL1) (Ahn *et al.* 2006; Hanzawa, Money, & Bradley, 2005). In this research, we screened 4 FT-like PEBP genes in 3 tetraploid potato varieties and the results show only 4 alleles of *StSP6A* contain non-conserved changes at 5 different sites. The changes that happened in *StSP3D, StSP5G* and *StTFL* alleles are mostly conservative changes. Interestingly, all these conservative and non-conservative changes did not seem like they would affect protein 3D structure or the binding ability with 14-3-3s. However, by cross-comparing the positions of non-conservative changes in StSP6A with the potential phosphorylation site, we have found that 4 non-conservative changes in StSP6A were on or near the potential phosphorylation sites of StSP6A.

Apart from the tertiary structure difference, the post-translational modification can also affect protein function, such as protein phosphorylation (Duan, & Walther, 2015). In mammals, PEBP has been known to work as an inhibitor for Raf kinase (Keller, Fu, & Brennan 2004). However, the phosphorylation of PEBP protein at Ser153 induced by protein kinase C (PKC) negatively affects the binding with Raf-1 kinase (Bian *et al.* 2015). In plants, it is possible that a similar post-translational modification could happened to FT-like PEBP proteins and affect their function.

Based on sequence-replaced modeling, all 4 FT-like PEBP proteins in potato are capable of binding 14-3-3s. Interestingly, most proteins interacting with plant 14-3-3s were phosphorylated on Ser or Thr residues present in a conserved binding motif (de Boer, van Kleeff, & Gao., 2013). Moreover, the phosphorylation of the target outside the 14-3-3 motif can also counteract the 14-3-3/target interaction (Duby *et al.*2009; de Boer, van Kleeff, & Gao., 2013). In StSP6A, the 126Ser in some alleles, which is two amino acids before the critical binding site (128Arg) for 14-3-3 protein, has the potential to be phosphorylated by Casein kinase II (CK-2) or PKC. It is possible that the phosphorylated 126Ser can inhibit the binding between StSP6A and 14-3-3. In Altus7076-2, Columba 507 and NRQ 375 alleles the Ser was replaced by Leu at position 126. Therefore, these 3 StSP6A alleles can be considered to induce “early” phenotype, since the binding with 14-3-3s might be easier and more stable. However, the 13aa deletion in Altus7076-2 deleted one of the other critical binding sites with 14-3-3, so the Altus variety does not have any allele which can induce “early” phenotype and it fits the late maturity phenotype of the Altus variety. Unfortunately, we only obtain 3 and 2 StSP6A alleles from Columba and NRQ genome sequence datasets, respectively. So far, both these two varieties have shown at least one StSP6A allele which can be considered to induce early maturity and we have already known that Columba has early maturity. Although we do not know how the NRQ variety would perform in tuberization, based on our results we presume the NRQ variety could have comparably early phenotype. This result provides a new line of research into plant propagation regulation. However, more experiments are required to test these predictions.

4.2 The expansion of FT in potato

Having viable progenies at the right time of the year is very important for plants. FLOWERING LOCUS T has been known to play a key role in regulating flowering while ago. In *Arabidopsis,* the FT protein can be transported to the apex and induces the expression of flower meristem identity genes to initiate flowering (Wigge 2011). In this decade, the FT homologues have been known to be involved in regulating storage organ's formation (Navarro *et al.* 2011). It has been discovered that the homologue of FT called SP3D acts as floral inducer in potato and another homologue of FT called SP6A works as a mobile tuberization signal in potato. Another homologue of FT in potato called SP5G acts as a repressor of SP6A in tuberization regulation (Navarro *et al.* 2011). The FT-like PEBP protein developed various alleles with different functions in potato. However, the level in which changes took place resulting in these functional changes remains largely unknown.

The external loop encodes by segment B has been known for playing important role in conferring biological specificity on FT and TFL (Ahn *et al.* 2006). In this study, we have found that the 3aa insertion at segment B part of StSP5G result in significant change on the shape of external loop. Although it is still not clear how FT protein's external loop is involved in regulating plant reproduction, more and more results show even single amino acid changes on FT segment B will result in reverting its promoting function to repressing function (Ho, & Weigel, 2014). Their results specifically show W138S in AtFT significantly delayed flowering. Interestingly, unlike AtFT, StSP6A and StSP3D, StSP5G has Ser at protein position 138 and this change might play a key role in StSP5G repressing function. Similar example can also be found in one of two paralogs of FT in cultivated beets (Beta vulgaris ssp. vulgaris) (Pin *et al.* 2010). They suggest that BvFT1 was initially a flowering promoter. However, the mutations within the external loop of its protein caused a shift in function to repress flowering. Considering no 14-3-3 proteins binding sites appear in segment B, it is possible that the segment B structure is involved in other regulation, such as binding with SWEET11 (SUGARS WILL EVENTUALLY BE EXPORTED TRANSPORTERS 11) for sugar transportation (Abelenda *et al.* personal communication).

It is interesting to note that the expansion of PEBPs extends to other *Solanaceae*. The TERMINAL FLOWER, SPELF PRUNING and SINGLE FLOWER TRUSS nomenclature also used in potato, originates from tomato flowering phenotypes. It may be that the high degree of floral plasticity in this asterid clade underlies the expansion of this gene family. Furthermore, the evolution of *Solanaceae* in regions naturally exposed to high levels of abiotic stresses such as drought, heat and altitude, may have made regulators such as PEBPs ideal to accelerate development of traits for ensuring both sexual and vegetative reproductive efficiency such as development of runners (such as in *S. etuberosum*), underground storage organs (such as in *S. tuberosum*) and diverse floral morphologies (as in the *S. lycopersicum* species)(Christian Bachem Personal communication).

In this study, we have found that StSP3D and StSP6A have a very similar protein structure. Based on previous studies, it has been shown that both proteins function similarly in forming similar complexes (FAC for StSP3D and TAC for StSP6A), since both proteins are capable of binding 14-3-3 proteins (Teo *et al.* 2016). Moreover, both proteins are synthesized in leaves, then transported to different tissues (Abelenda, Navarro, & Prat. 2014). Knowing whether these two proteins can work in different tissue is important in studying their function. Interestingly, an experiment has shown that overexpressing StSP6A in non-tuber species *S. etuberosum* results in early flowering (Dr. Christian Bachem, personal communication). However,

overexpressing StSP6A in potato could only lead to early tuberization (Teo *et al.* 2016). It would appear that StSP6A is capable of inducing flowering in other species, but not in potato. Therefore, the different function of StSP3D and StSP6A might depend highly on being correctly recognized by the transportation system. Potato may have a special recognition process for transporting the right FT-like protein to the right position which did not exist in other species.

So far, not much is known about the transportation of FT proteins in potato. The fact that the alpha helix tail of StSP3D is slightly longer than StSP6A is the only notable difference in tertiary structure between StSP3D and StSP6A. This special alpha helix tail did not exist in AtFT or OsHd3a. It is possible that this potato specific alpha helix tail is involved in the potato specific recognition of StSP3D and StSP6A. Therefore, the alpha helix tail would be a good place to start for further study.

As mentioned in chapter 4.1, the post-translational modification might be involved in FT regulation in potato. The phosphorylation sites prediction shows both StSP3D and StSP6A have several possible phosphorylation sites, but they appear at very different positions. We have shown that one of the phosphorylation sites in StSP6A at position 126 may affect the binding with 14-3-3s. However, none of the predicted phosphorylation sites in StSP3D appear at the 14-3-3 binding motif or near the 14-3-3 binding motif. Additionally, there is a potential phosphorylation site at the alpha helix tail of StSP3D but no other FT-like PEBP proteins in potato. This post-translational modification might also be involved in the recognition of different FT homologues. Indeed, further experiments are required to have more insight. However, the phosphorylation motif prediction results provide a possible direction to gaining a better understanding of the key difference between StSP3D and StSP6A.

5.Ref:

Abelenda, J. A., Navarro, C., & Prat, S. (2014). Flowering and tuberization: a tale of two nightshades. Trends in plant science, 19(2), 115-122.

Ahn, J. H., Miller, D., Winter, V. J., Banfield, M. J., Lee, J. H., Yoo, S. Y., ... & Weigel, D. (2006). A divergent external loop confers antagonistic activity on floral regulators FT and TFL1. The EMBO journal, 25(3), 605-614.

Ahn, J. H., Miller, D., Winter, V. J., Banfield, M. J., Lee, J. H., Yoo, S. Y., ... & Weigel, D. (2006). A divergent external loop confers antagonistic activity on floral regulators FT and TFL1. The EMBO journal, 25(3), 605-614.

Andrés, F., & Coupland, G. (2012). The genetic basis of flowering responses to seasonal cues. *Nature Reviews Genetics*, *13*(9), 627.

Bairoch, A., & Bucher, P. (1994). PROSITE: recent developments. Nucleic Acids Research, 22(17), 3583.

Bian, J. M., Wu, N., Su, R. B., & Li, J. (2015). Phosphatidylethanolamine-binding protein is not involved in μ-opioid receptor-mediated regulation of extracellular signal-regulated kinase. Molecular medicine reports, 11(5), 3368-3374.

Carmel-Goren, L., Liu, Y. S., Lifschitz, E., & Zamir, D. (2003). The SELF-PRUNING gene family in tomato. Plant molecular biology, 52(6), 1215-1222.

Chautard H, Jacquet M, Schoentgen F, Bureaud N, Bénédetti H (2004) Tfs1p, a member of the PEBP family, inhibits the Ira2p but not the Ira1p Ras GTPase-activating protein in Saccharomyces cerevisiae. Eukaryotic Cell 3: 459-470

de Boer, A. H., van Kleeff, P. J., & Gao, J. (2013). Plant 14-3-3 proteins as spiders in a web of phosphorylation. Protoplasma, 250(2), 425-440.

Duan, G., & Walther, D. (2015). The roles of post-translational modifications in the context of protein interaction networks. PLoS computational biology, 11(2), e1004049.

Duby, G., Poreba, W., Piotrowiak, D., Bobik, K., Derua, R., Waelkens, E., & Boutry, M. (2009). Activation of plant plasma membrane H+-ATPase by 14-3-3 proteins is negatively controlled by two phosphorylation sites within the H+-ATPase C-terminal region. Journal of Biological Chemistry, 284(7), 4213-4221.

Hanano, S., & Goto, K. (2011). Arabidopsis TERMINAL FLOWER1 is involved in the regulation of flowering time and inflorescence development through transcriptional repression. *The Plant Cell*, tpc-111.

Hanzawa, Y., Money, T., & Bradley, D. (2005). A single amino acid converts a repressor to an activator of flowering. Proceedings of the National Academy of Sciences, 102(21), 7748-7753.

Hayama, R., Izawa, T., & Shimamoto, K. (2002). Isolation of rice genes possibly involved in the photoperiodic control of flowering by a fluorescent differential display method. *Plant and cell physiology*, *43*(5), 494-504.

Ho, W. W. H., & Weigel, D. (2014). Structural features determining flower-promoting activity of Arabidopsis FLOWERING LOCUS T. *The Plant Cell*, tpc-113.

Kaneko-Suzuki, M., Kurihara-Ishikawa, R., Okushita-Terakawa, C., Kojima, C., Nagano-Fujiwara, M., Ohki, I., ... & Taoka, K. I. (2018). TFL1-like proteins in rice antagonize rice FT-like protein in inflorescence development by competition for complex formation with 14-3-3 and FD. *Plant and Cell Physiology*, *59*(3), 458-468.

Karlgren, A., Gyllenstrand, N., Källman, T., Sundström, J. F., Moore, D., Lascoux, M., & Lagercrantz, U. (2011). Evolution of the PEBP gene family in plants: functional diversification in seed plant evolution. *Plant physiology*, pp-111.

Keller, E. T., Fu, Z., & Brennan, M. (2004). The role of Raf kinase inhibitor protein (RKIP) in health and disease. *Biochemical pharmacology*, *68*(6), 1049-1053.

Kojima, S., Takahashi, Y., Kobayashi, Y., Monna, L., Sasaki, T., Araki, T., & Yano, M. (2002). Hd3a, a rice ortholog of the Arabidopsis FT gene, promotes transition to flowering downstream of Hd1 under short-day

conditions. *Plant and Cell Physiology*, *43*(10), 1096-1105.
Longman, K. A., & Wilson, R. (1993). Rooting cuttingsof tropical trees. Cornmon wealth Science Council, London.
Navarro, C., Abelenda, J. A., Cruz-Oró, E., Cuéllar, C. A., Tamaki, S., Silva, J., ... & Prat, S. (2011). Control of flowering and storage organ formation in potato by FLOWERING LOCUS T. *Nature*, *478*(7367), 119.
Pin, P. A., Benlloch, R., Bonnet, D., Wremerth-Weich, E., Kraft, T., Gielen, J. J., & Nilsson, O. (2010). An antagonistic pair of FT homologs mediates the control of flowering time in sugar beet. Science, 330(6009), 1397-1400.
Qin, Z., Wu, J., Geng, S., Feng, N., Chen, F., Kong, X., ... & Wu, L. (2017). Regulation of FT splicing by an endogenous cue in temperate grasses. *Nature communications*, *8*, 14320.
Simon R., Igeño M.I., Coupland G. (1996). Activation of floral meristem identity genes in Arabidopsis. Nature 384: 59–62.
Taoka, K. I., Ohki, I., Tsuji, H., Furuita, K., Hayashi, K., Yanase, T., ... & Ogaki, Y. (2011). 14-3-3 proteins act as intracellular receptors for rice Hd3a florigen. *Nature*, *476*(7360), 332.
Taoka, K. I., Ohki, I., Tsuji, H., Kojima, C., & Shimamoto, K. (2013). Structure and function of florigen and the receptor complex. *Trends in plant science*, *18*(5), 287-294.
Teo, C. J., Takahashi, K., Shimizu, K., Shimamoto, K., & Taoka, K. I. (2016). Potato tuber induction is regulated by interactions between components of a tuberigen complex. Plant and Cell Physiology, 58(2), 365-374.
Tomato Genome Consortium. (2012). The tomato genome sequence provides insights into fleshy fruit evolution. Nature, 485(7400), 635.
Wigge, P. A. (2011). FT, a mobile developmental signal in plants. *Current Biology*, *21*(9), R374-R378.
Wigge, P. A., Kim, M. C., Jaeger, K. E., Busch, W., Schmid, M., Lohmann, J. U., & Weigel, D. (2005). Integration of spatial and temporal information during floral induction in Arabidopsis. Science, 309(5737), 1056-1059.
Yano, M., Katayose, Y., Ashikari, M., Yamanouchi, U., Monna, L., Fuse, T., ... & Sasaki, T. (2000). Hd1, a major photoperiod sensitivity quantitative trait locus in rice, is closely related to the Arabidopsis flowering time gene CONSTANS. *The Plant Cell*, *12*(12), 2473-2483.
Yeung K, Seitz T, Li S, Janosch P, McFerran B, Kaiser C, Fee F, Katsanakis K, Rose D, Mischak H, Sedivy J, Kolch W (1999) Suppression of Raf-1 kinase activity and MAP kinase signalling by RKIP. Nature 401: 173-177
Zhang, Y. (2008). I-TASSER server for protein 3D structure prediction. BMC bioinformatics, 9(1), 40.