RESEARCH ARTICLE Legume Research- An International Journal, Volume 47 Issue 7: 1073-1082 (July 2024)

Identification and Analysis of Phosphatidyl Ethanolamine Binding Protein (PEBP) Family Gene in Alfalfa

Wenna Fan[1], Mengyao Zhang[1], Wenfeng Ma[1], Yixin Yang[1], Yaqi Shi[1]

10.18805/LRF-796

ABSTRACT

Background: Phosphatidyl ethanolamine binding protein (PEBP) plays an important role in the formation of flowers. To identify the PEBP gene family of alfalfa in blossom, the PEBP gene family was selected from the transcriptome database of alfalfa with bioinformatics analysis, for their protein properties, functional domains, evolutionary relationships and expression patterns.

Methods: A BLASTP search using Arabidopsis PEBP sequences as a query was performed against alfalfa proteome with a minimum identity of 75% and E-value of 1E-5. BLASTP search resulted in 17 protein candidates. Finally, we obtained a consensus list of ten proteins from the HMMER and BLASTP search. Redundant sequences were removed using CD-HIT which resulted in a total of nine PEBP candidates in alfalfa.

Result: The results showed that there was a total of nine PEBP candidates in alfalfa, which could be divided into three sub-gene families, such as TFL1-like, MFT-like and FT-like groups. Chromosomal location analyses showed that nine PEBP members are distributed on five of the eight alfalfa chromosomes with chromosome 7 having the largest number of PEBP members (four members) while only three alfalfa PEBP proteins were annotated as MsFTa1, MsFTa2 and MsFTc. In addition, the motifs of this gene family were conservative with few differences in the number and types of conservative motifs. MsFTa1 and MsFTa2 are two strong candidates for florigen in alfalfa, while MsG0480024000.01 is a candidate which functions as a floral repressory.

Key words: Alfalfa, Conserved motif, Expression analysis, PEBP gene family, Phylogenetic analysis.

INTRODUCTION

Phosphatidylethanolamine-binding protein (PEBP) is a unique type of phosphorus conservative protein in the domain of Acyl ethanolamine binding proteins, widely present in animals and plants, control flowering transition and vegetative growth in plants, such as florigen production, light response and abiotic stressors of reaction (Zhu *et al.*, 2013; Dong *et al.*, 2023). The plant PEBP protein mainly consists of three groups: MFT-like (Mother of FT and TFL1-Like) FT-like (Flowering Locus T-Like) and TFL1-like (Terminal Flower 1-Like). The PEBP family genes all have a single, highly conserved PEBP domain, accounting for 80% of the coding region. The PEBP family genes are closely related to flowering hormones, especially FT-like protein. The expression mechanism of PEBP family genes regulated by photoperiod varies are different in different plants. The study of PEBP family genes and their expression regulation in plant breeding not only has theoretical significance, but also has certain potential application prospects. In plant breeding, flowering and maturity stages are important plant trait indicators.

Alfalfa is an important legume forage and studying its flowering mechanism is of great significance. A recent study showed that alfalfa is highly related to *Medicago truncatula* and they were estimated to have diverged around 8 million years ago (Shen *et al.*, 2020). Furthermore, alfalfa has not experienced recent whole-genome duplication (Shen *et al.*, 2020). Based on the bioinformatics analysis of transcriptome sequencing data, we identified the PEBP gene family in alfalfa. Further understanding of the composition of PEBP gene family members in alfalfa and their flowering regulation mechanisms is important for strengthening the future research programmes in this crop.

[1]Animal Science and Technology College, Henan University of Science and Technology, Luoyang, Henan 471003, China.

Corresponding Author: Wenna Fan, Animal Science and Technology College, Henan University of Science and Technology, Luoyang, Henan 471003, China. Email: chou0516@163.com

How to cite this article: Fan, W., Zhang, M., Ma, W., Yang, Y. and Shi, Y. (2024). Identification and Analysis of Phosphatidyl Ethanolamine Binding Protein (PEBP) Family Gene in Alfalfa. Legume Research. 47(7): 1073-1082. doi: 10.18805/LRF-796.

Submitted: 22-01-2024 **Accepted:** 09-03-2024 **Online:** 13-04-2024

MATERIALS AND METHODS

Identification of alfalfa PEBP family

Alfalfa protein sequences were retrieved from Shen *et al.* (2020). PBP domain (PF01161) HMM profile retrieved from Pfam database (v33.1, May 2020) (El-Gebali *et al.*, 2019) was used to search for the PBP domain-containing sequences in alfalfa proteome with HMMER v3.1b2 (http://hmmer.org/). In addition, BLASTP (v2.10.1+) (Altschul *et al.*, 1997) was employed to find homologues using Arabidopsis PEBP proteins with sequence identity >75% and E-value < 1E-5. Redundant sequences were removed using CD-HIT with default parameters (Huang *et al.*, 2010). Pfam (El-Gebali *et al.*, 2019) was used to check the presence of PBP domain in candidate PEBP protein sequences.

Phylogenetic relationship and bioinformatics analysis

For phylogenetic tree construction, multiple sequence alignment was performed using CLUSTALW v2.1 (https://www.genome.jp/tools-bin/clustalw) (Thompson *et al.*, 1994) with the "SLOW/ACCURATE" option. TrimAl v1.4 (Capella-Gutiérrez *et al.*, 2009) was used to remove spurious residues with the "automated1" option. The phylogenetic tree was constructed using IQ-TREE (Trifinopoulos *et al.*, 2016) with the best-fit model "JTTDCMut+G4" and 1,000 ultrafast bootstrap (Hoang *et al.*, 2017) replicates. The tree was displayed using FigTree v1.4.4 (http://tree.bio.ed.ac.uk/software/figtree/) with Bootstrap support values shown at branches. Alignments of Arabidopsis and alfalfa PEBP protein and coding sequences were shaded using pyBoxshade v1.0 (https://github.com/mdbaron42 / pyBoxshade). Properties of PEBP proteins were calculated using ProtParam (Gasteiger *et al.*, 2005) and subcellular localization was predicted using ProtComp v9.0 (http://www.softberry.com/berry.phtml?topic=protcomppland group=programsand subgroup=proloc). Genomic coordinates of nine alfalfa PEBP family members were retrieved from the alfalfa General Feature Format (GFF) annotation file (Shen *et al.*, 2020) and MapGene2Chrom (Chao *et al.*, 2015) was used to display the locations on alfalfa chromosomes (http://mg2c.iask.in/mg2c_v2.1/).

Gene structure analysis and conserved motifs of alfalfa and arabidopsis *PEBP* genes

Genomic and coding sequences of alfalfa and Arabidopsis were obtained and used to illustrate the exon-intron gene structures with the Gene Structure Display Server (GSDS) 2.0 (Hu *et al.*, 2015). Conserved motifs discovery was performed using MEME (Bailey and Elkan, 1994) with a maximum number of 10 motifs to be found. *cis*-regulatory elements were explored with 2000 bp promoter sequences upstream of the start codon using PlantCARE (Lescot *et al.*, 2002).

Multiple sequence alignment of alfalfa and arabidopsis PEBP protein sequences expression

Multiple protein sequence alignment was performed using CLUSTALW (https://www.genome.jp/tools-bin/clustalw). Alignment was shaded using pyBoxshade v1.0 (https://github.com/ mdbaron42/pyBoxshade).

***Cis*-regulatory elements in alfalfa *PEBP* promoters**

Cis-regulatory elements play critical roles in regulating gene expression. To determine the *cis*-regulatory elements in alfalfa *PEBP* promoters, we inspected the 2000 bp promoter sequences upstream of the start codon (ATG) using PlantCARE (Lescot *et al.*, 2002).

Expression profiles of alfalfa *PEBP* relative genes from RNA-seq data

RNA-seq dataset from tissues of the elongating stem, post-elongating stem, flower, leaf, nodule and root were retrieved from O'Rourke *et al.* (2015). Normalized transcripts per million (TPM) values were obtained using TPMCalculator v0.0.4 (Roberto *et al.*, 2018) after mapping RNA-seq reads to the alfalfa reference genome (Shen *et al.*, 2020) using HISAT v2.2.1 (Kim *et al.*, 2019). Expression heatmap was drawn using ClustVis (Metsalu and Vilo, 2015) and expression values were shown as ln(TPM+1).

RNA-seq dataset from early and late flowering genotypes of alfalfa was obtained from Ma *et al.* (2021). Normalized transcripts per million (TPM) values were obtained using TPMCalculator v0.0.4 (Roberto *et al.*, 2018) after mapping RNA-seq reads to the alfalfa reference genome (Shen *et al.*, 2020) using HISAT v2.2.1 (Kim *et al.*, 2019). Expression heatmap was drawn using ClustVis (Metsalu and Vilo, 2015) and expression values were shown as ln(TPM+1).

RNA-seq dataset from our previous study (Zhang *et al.*, 2015) which used leaf tissues of one fall dormant cultivar (Maverick, fall dormancy class of 1) and one non-fall dormant cultivar (CUF101, fall dormancy class of 9) under two-time points (May and September in China: 34° N, 113° E). Normalized transcripts per million (TPM) values were obtained using TPMCalculator v0.0.4 (Roberto *et al.*, 2018) after mapping RNA-seq reads to the alfalfa reference genome (Shen *et al.*, 2020) using HISAT v2.2.1 (Kim *et al.*, 2019). Expression heatmap was drawn using ClustVis (Metsalu and Vilo, 2015) and expression values were shown as ln(TPM+1).

Accession numbers

Accession numbers of Arabidopsis PEBP genes used in this study: *AtFT* (*AT1G65480*), *AtTSF* (*AT4G20370*), *AtMFT* (*AT1G18100*), *AtTFL1* (*AT5G03840*), *AtATC* (*AT2G27550*) and *AtBFT* (*AT5G62040*).

RESULTS AND DISCUSSION

The PEBP gene family plays an important role in floral transformation, floral morphogenesis and species development in plants. There is a significant difference in the number of members of this gene family between monocotyledonous and dicotyledonous plants, with a relatively higher number of PEBP genes in monocotyledonous plants (Dong *et al.*, 2023). This study searched for 9 members of the PEBP gene family in alfalfa through the local database blast comparison, which resulted in more members of dicotyledonous plants such as *Vitis vinifera* L., *Cucumis sativus* and *Populus trichocarpa* and much less of monocotyledonous plants such as *Musa acuminata*, *Zea mays* L. and *Sorghum bicolor* (Liu *et al.*, 2016).

Identification of PEBP gene family in alfalfa

To identify all PEBP gene family members in alfalfa, we conducted a profile search with the PBP domain (PF01161) against the alfalfa proteome by using HMMER. A total of 12 protein candidate sequences were obtained with an E-value threshold of 1E-3. In the meantime, we retrieved all six PEBP

protein sequences, FT, TSF, TFL1, ATC, BFT and MFT, from Arabidopsis. A BLASTP search using Arabidopsis PEBP sequences as a query was performed against alfalfa proteome with a minimum identity of 75% and E-value of 1E-5. BLASTP search resulted in 17 protein candidates. Finally, we obtained a consensus list of ten proteins from the HMMER and BLASTP search. Redundant sequences were removed using CD-HIT which resulted in a total of nine PEBP candidates in alfalfa. All nine candidate proteins harboured PBP domains and were retained for further analyses (Table 1). Since three out of the nine alfalfa PEBP proteins were annotated as MsFTa1 (MsG07800401 29.01.T01), MsFTa2 (MsG078004013 0.01.T01) and MsFTc (MsG07800 40133.01.T01) (Shen *et al.*, 2020), here we kept the same annotation for these three PEBP members.

We then analyzed the physical and chemical properties including length, molecular weight, isoelectric point, instability index and grand average of hydropathicity (GRAVY) value of alfalfa PEBP proteins along with Arabidopsis PEBP members using ProtParam (Gasteiger *et al.*, 2005). Alfalfa PEBP proteins had a length range of 106 to 210 amino acids (aa) compared to around 175 aa in Arabidopsis PEBP proteins while molecular weight ranged from 12.15 to 24.15 kD compared to 19-20 kD in arabidopsis (Table 1). The larger molecular weight range in alfalfa might be due to the length of two PEBP proteins, MsG0780035 995.01.T01 (106 aa) and MsG018000336 1.01.T01 (210 aa). Isoelectric point ranged from 7.69 to 9.64 with one exception of MsG0680031714.01. T01 (4.76). Instability analysis indicated that three out of nine alfalfa PEBP proteins are stable (instability index less than 40) while there is only one stable protein in Arabidopsis (AtATC). GRAVY values showed that all PEBP proteins are hydrophilic (with negative values) except MsG0480024000.01.T01 which is hydrophobic (with positive value). Furthermore, protein subcellular localization showed that all PEBP members are located in the cytoplasm.

Phylogenetic and chromosomal location analyses

PEBP family members are classified into three groups, namely FT-like, MFT-like and TFL1-like (Carmona *et al.*, 2007; Hedman *et al.*, 2009; Wang *et al.*, 2015). All nine alfalfa PEBP proteins along with Arabidopsis PEBP homologues were aligned and used to construct a phylogenetic tree. Phylogenetic analysis showed that alfalfa PEBP proteins were classified into TFL1-like, MFT-like and FT-like groups (Fig 1). The largest group is FT-like which contains five alfalfa PEBP proteins, followed by three proteins in the TFL1-like group, while the MFT-like group contains only one alfalfa PEBP protein (Fig 1).

To determine the locations of alfalfa PEBP proteins on the alfalfa chromosomes, gene location coordinates were obtained and used to pin the PEBP members onto alfalfa chromosomes. It showed that nine PEBP members are distributed on five of the eight alfalfa chromosomes (Fig 2). Chromosome 7 has the largest number of PEBP members (four members), followed by Chromosome 6 which has two members, while Chromosomes 1, 2 and 4 have one member each (Fig 2). Gene labels with different colors represent different PEBP groups. Chromosome 7 has a cluster of three FT homologues as reported by Shen *et al.* (2020). Moreover, this region also echoes the syntenic region in its closest relative *Medicago truncatula* which also harbors three FT homologues (Shen *et al.*, 2020).

Gene structure analysis and conserved motifs of alfalfa and arabidopsis *PEBP* genes

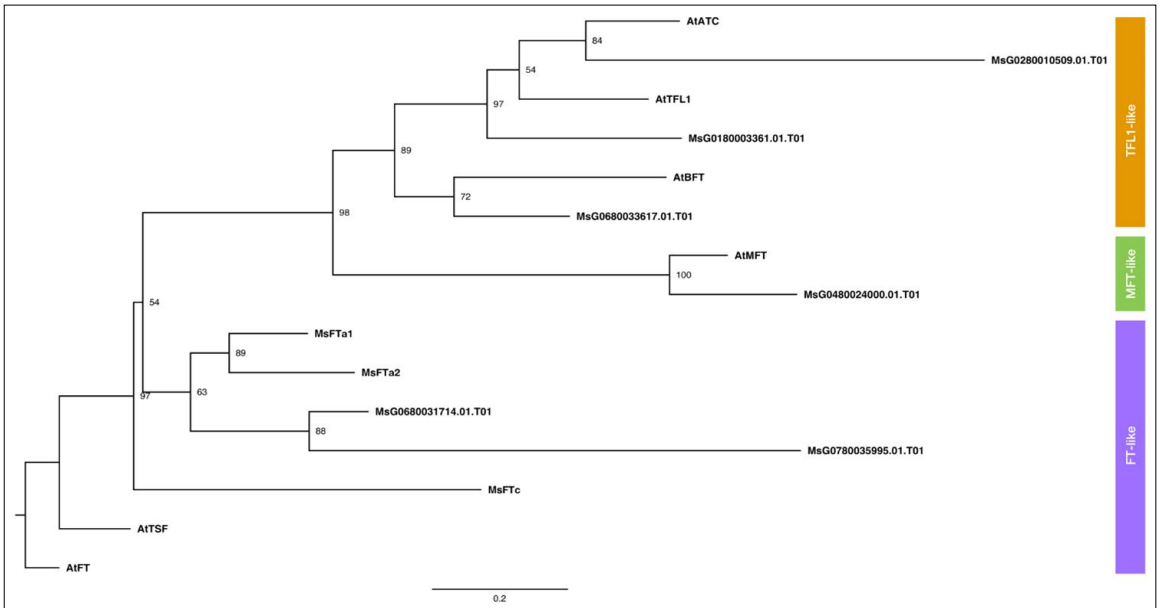
Genomic and coding sequences of alfalfa and Arabidopsis were obtained and used to illustrate the exon-intron gene structures with the Gene Structure Display Server (GSDS)

Table 1: Properties of alfalfa and arabidopsis PEBP proteins.

Protein ID	Length (amino acids)	Molecular weight (kD)	pI	Instability index	GRAVY value	Subcellular localization
MsG0180003361.01.T01	210	24.15	9.44	48.00	-0.310	Cytoplasmic
MsG0280010509.01.T01	120	14.20	9.64	49.41	-0.453	Cytoplasmic
MsG0480024000.01.T01	193	21.31	8.91	36.72	0.164	Cytoplasmic
MsG0680031714.01.T01	162	17.84	4.76	48.33	-0.183	Cytoplasmic
MsG0680033617.01.T01	172	19.61	9.44	56.29	-0.512	Cytoplasmic
MsG0780035995.01.T01	106	12.15	8.93	43.20	-0.147	Cytoplasmic
MsG0780040129.01.T01 (MsFTa1)	176	19.70	8.52	36.42	-0.432	Cytoplasmic
MsG0780040130.01.T01 (MsFTa2)	177	19.85	8.64	36.84	-0.271	Cytoplasmic
MsG0780040133.01.T01 (MsFTc)	171	19.06	7.69	40.53	-0.106	Cytoplasmic
AtFT	175	19.81	7.75	48.81	-0.259	Cytoplasmic
AtTSF	175	19.69	7.76	41.54	-0.283	Cytoplasmic
AtMFT	173	19.13	7.93	51.44	-0.179	Cytoplasmic
AtTFL1	177	20.16	9.69	44.52	-0.224	Cytoplasmic
AtATC	175	19.90	7.01	33.33	-0.275	Cytoplasmic
AtBFT	177	20.01	9.16	48.12	-0.271	Cytoplasmic

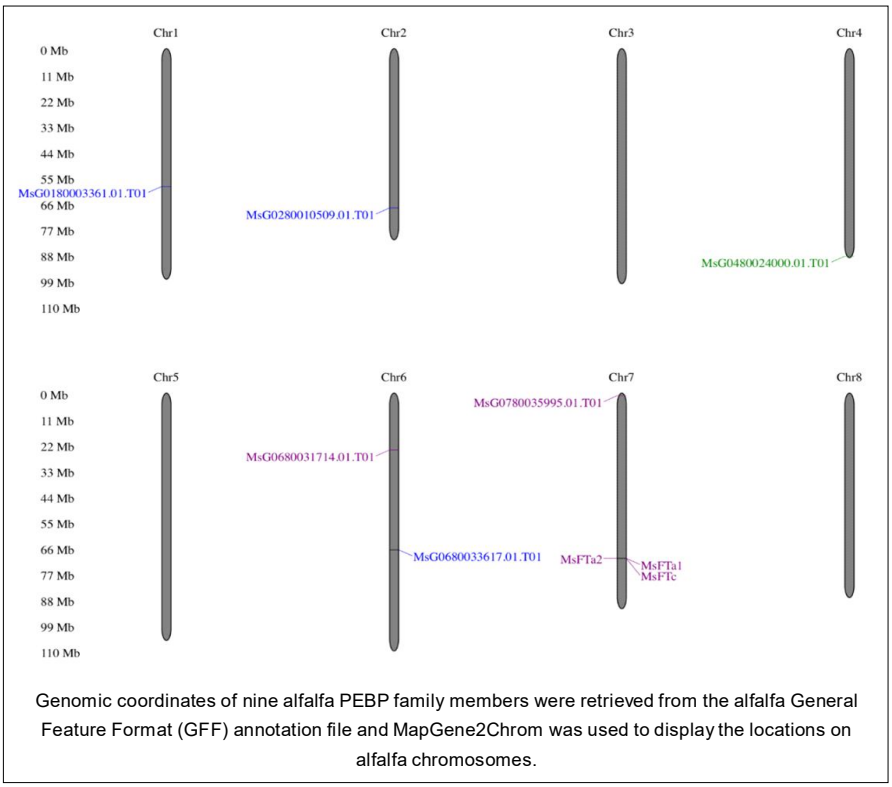
2.0 (Hu *et al*., 2015). Exons, introns and, upstream/ downstream regions were highlighted (Fig 3).

A total of 15 alfalfa and Arabidopsis PEBP proteins were submitted to the MEME suite for identification of conserved motifs. A maximum number of 10 motifs were allowed to be discovered (Fig 4). All other parameters were in default. Then we obtained a consensus list of ten proteins from HMMER and BLASTP search. Redundant sequences were removed using CD-HIT resulted in nine PEBP candidates in alfalfa.



Alfalfa PEBP proteins were grouped into TFL1-like, MFT-like and FT-like categories as labelled on the right.

Fig 1: Phylogenetic tree of alfalfa and Arabidopsis PEBP proteins.



Genomic coordinates of nine alfalfa PEBP family members were retrieved from the alfalfa General Feature Format (GFF) annotation file and MapGene2Chrom was used to display the locations on alfalfa chromosomes.

Fig 2: Chromosomal locations of alfalfa PEBP family members.

Multiple sequence alignment of alfalfa and arabidopsis PEBP protein sequences expression

The same residues were shaded in red while similar residues were shaded in Dodger blue with lowercases. Residues identical to consensus were shaded in Chatham blue. Exon boundaries are indicated using arrows (Fig 5). Two critical residues (Ahn *et al.*, 2014) distinguish FT and TFL1 are indicated with purple dots. Four segments within the fourth exon are underlined and labelled as A, B, C and D.

Cis-regulatory elements in alfalfa *PEBP* promoters

Cis-regulatory elements were classified into five categories, namely Light (light responsiveness), Circadian, Hormone, Stress and defense and Meristem expression (Table 2). Among the five categories, the category Light has the largest number of *cis*-regulatory elements (122 elements), followed

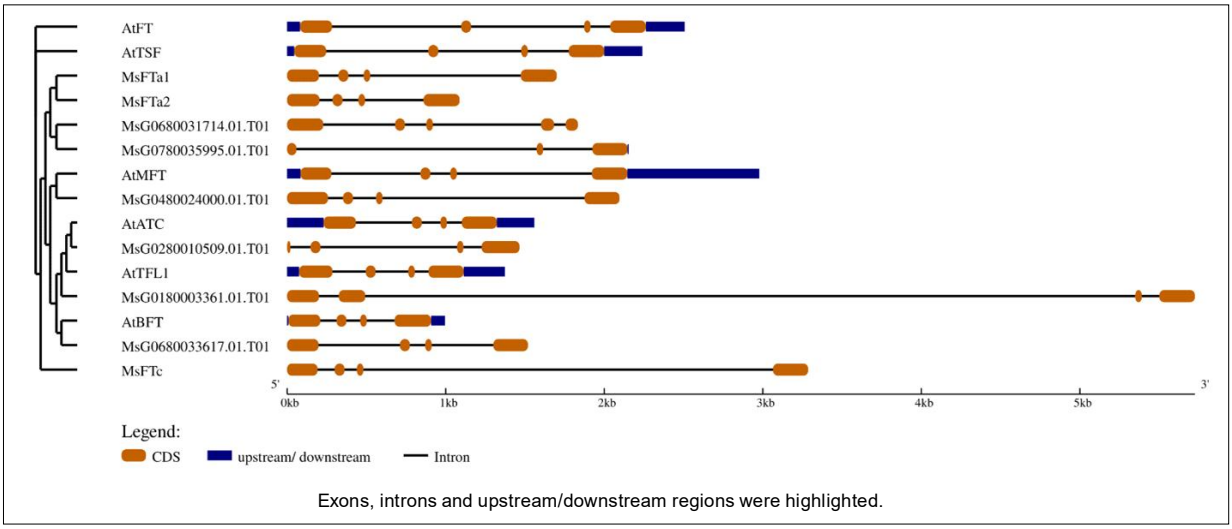


Fig 3: Gene structure analysis of alfalfa and arabidopsis *PEBP* genes.

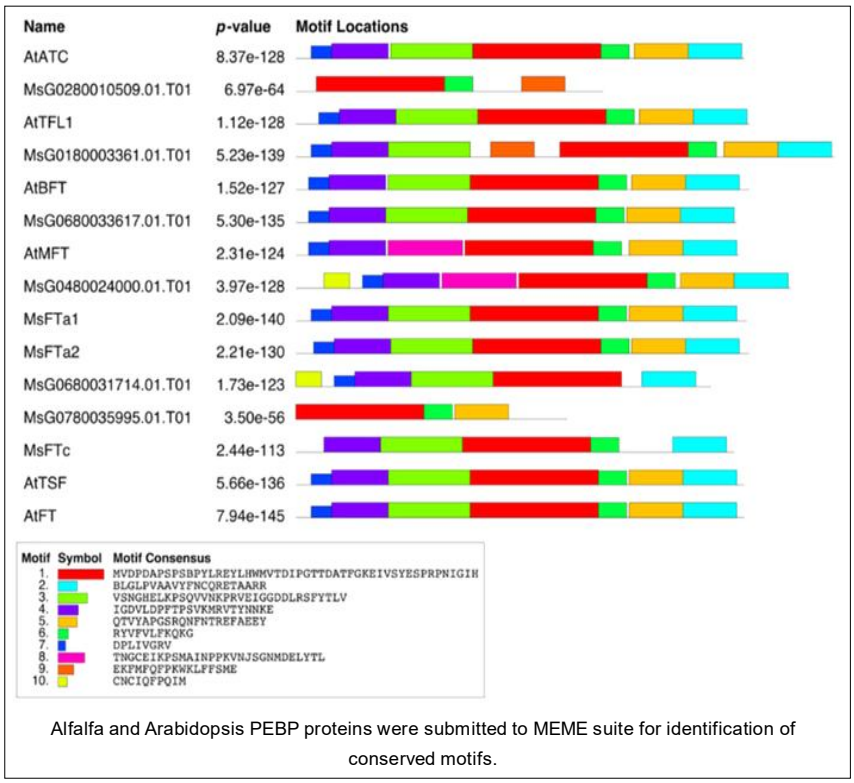


Fig 4: Conserved motifs in alfalfa and Arabidopsis PEBP proteins.

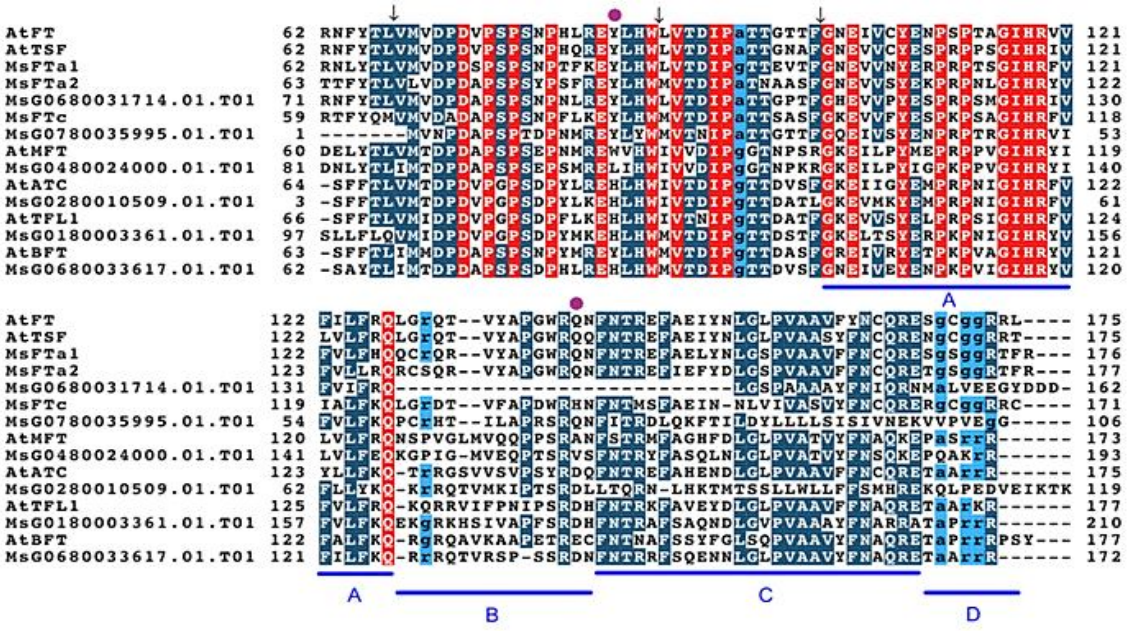
Table 2: *Cis*-regulatory elements in the promoters of alfalfa *PEBP* family members.

Category	Elements	*MsFTa1*	*MsFTa2*	*MsFTc*	*MsG0680031714.01*	*MsG0780035995.01*	*MsG0480024000.01*	*MsG0180003361.01*	*MsG0280010509.01*	*MsG0680033617.01*
Light	AAAC-motif	1	0	0	0	0	0	0	1	1
	ACE	0	2	1	0	0	0	0	0	0
	AE-box	0	0	1	0	0	1	1	1	0
	AT1-motif	0	0	0	0	0	1	0	0	0
	ATCT-motif	0	1	0	0	0	0	0	0	1
	Box 4	3	1	0	8	6	4	6	5	3
	chs-CMA1a	0	0	0	3	1	0	0	0	0
	G-Box	1	3	7	3	2	10	0	0	1
	GA-motif	0	0	2	0	1	1	0	0	0
	Gap-box	0	1	0	1	1	0	0	0	0
	GATA-motif	0	1	2	1	0	1	0	1	1
	GT1-motif	0	2	1	0	1	1	1	2	2
	LAMP-element	0	0	0	2	0	0	0	0	0
	MRE	0	0	0	0	1	0	1	1	1
	Sp1	0	0	2	0	0	0	0	0	0
	TCCC-motif	0	0	0	0	1	1	0	0	1
	TCT-motif	1	1	1	0	0	2	1	0	1
	Subtotal	**6**	**12**	**17**	**18**	**14**	**22**	**10**	**11**	**12**
Circadian	circadian	0	0	0	0	0	0	0	1	0
Hormone	ABRE(ABA)	2	3	7	3	1	8	0	0	1
	AuxRR-core(Auxin)	2	0	0	0	0	0	0	0	0
	TGA-element(Auxin)	1	0	2	0	0	0	0	0	0
	GARE-motif(GA)	0	0	0	0	1	0	0	0	0
	P-box(GA)	0	0	2	0	0	0	1	0	0
	TATC-box(GA)	0	0	0	0	0	1	0	0	0
	TCA-element	0	0	1	1	0	0	0	1	0
	Subtotal	**5**	**3**	**12**	**4**	**2**	**9**	**1**	**1**	**1**
Stress and defense	ARE (Anaerobic induction)	5	2	3	1	0	2	2	3	4
	CCAAT-box	1	0	2	0	1	3	0	0	1
	TC-rich repeats	1	0	0	0	1	1	0	2	0
	LTR(Low-temperature)	0	1	4	0	2	0	0	2	1
	MBS(Drought)	0	0	2	0	1	0	0	0	0
	CGTCA-motif(MeJA)	0	0	3	2	0	1	2	0	1
	TGACG-motif (MeJA)	0	0	3	2	0	1	2	0	1
	Subtotal	**7**	**3**	**17**	**5**	**5**	**8**	**6**	**7**	**8**
Meristem expression	CAT-box	1	1	0	0	0	0	0	1	0

by Stress and defense (66 elements) and Hormone (38 elements) categories. The circadian (involved in the circadian clock) and CAT-box (involved in meristem expression) are found in one and three PEBP members respectively (Table 2).

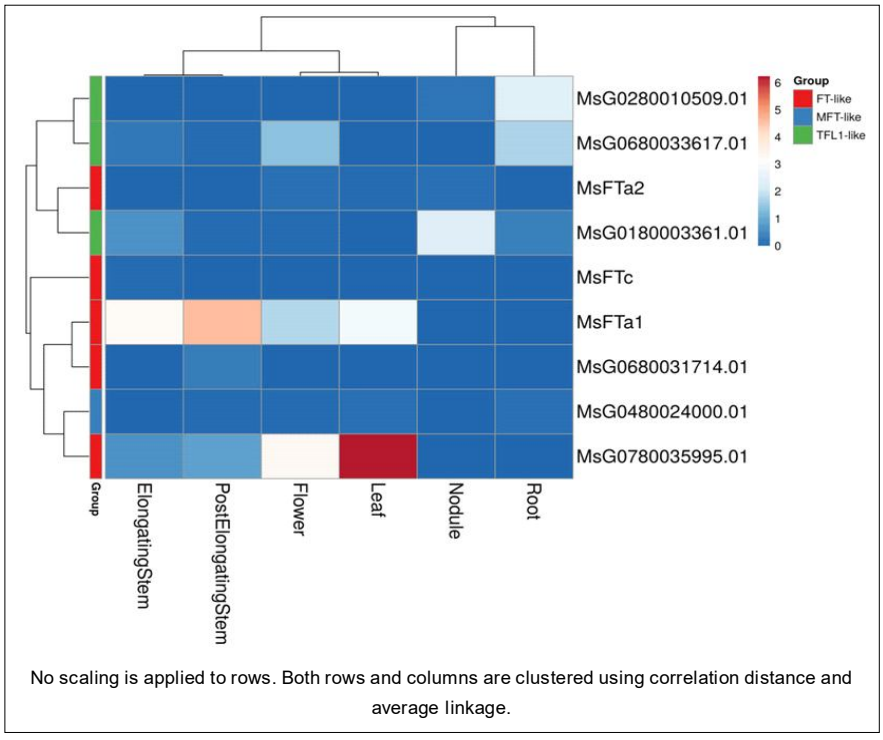
Gene expression profiles of alfalfa *PEBP* gene family from RNA-seq dataset

To check the expression patterns in various tissues, we used a public RNA-seq dataset (O'Rourke *et al.*, 2015) which includes tissues of the elongating stem, post-elongating



Exon boundaries are indicated using arrows. Two critical residues distinguish FT and TFL1 are indicated with purple dots. Four segments within the fourth exon are underlined and labelled as A, B, C and D.

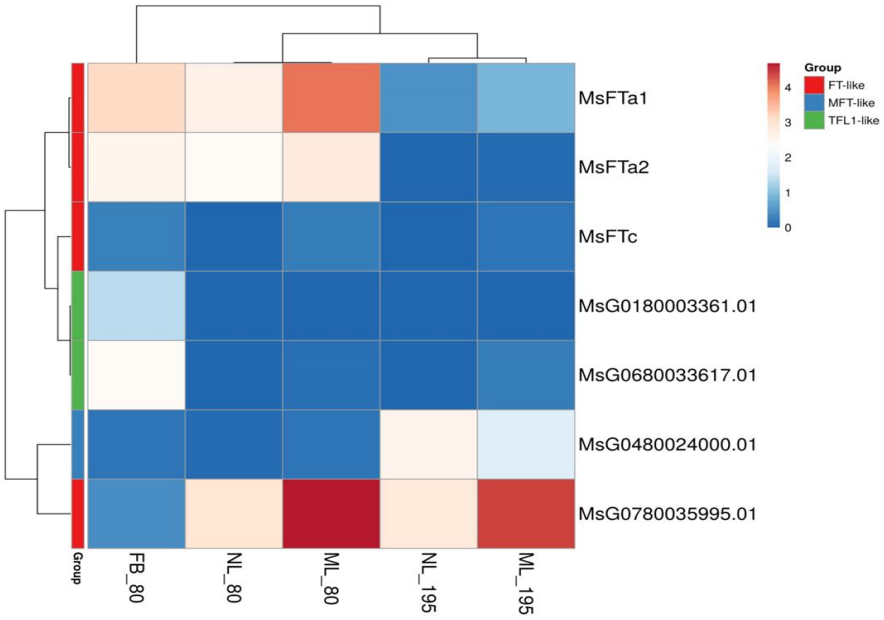
Fig 5: Multiple sequence alignment of alfalfa and Arabidopsis PEBP protein sequences.



No scaling is applied to rows. Both rows and columns are clustered using correlation distance and average linkage.

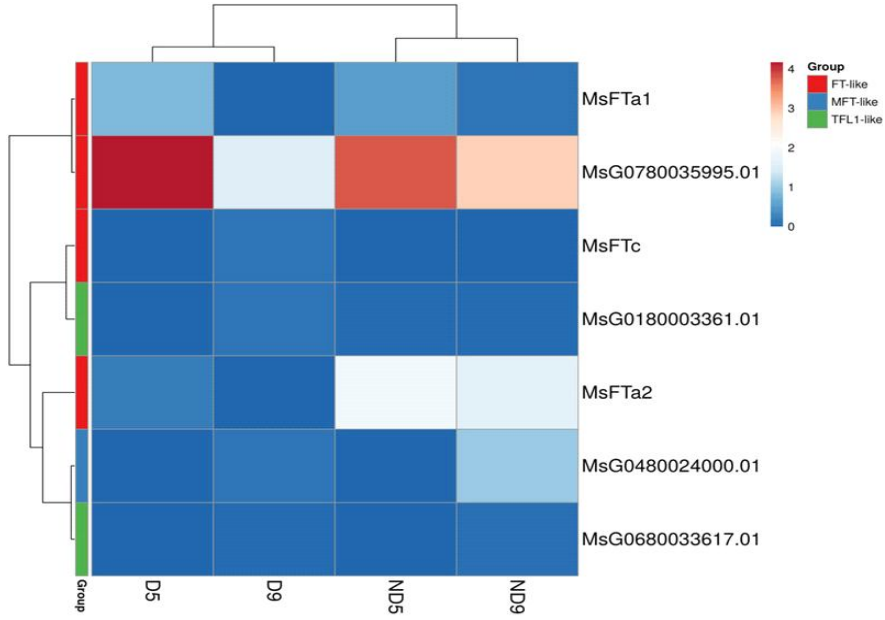
Fig 6: Tissue-specific expression of alfalfa *PEBP* genes.

stem, flower, leaf, nodule and root. It shows that for most alfalfa *PEBP* genes, gene expression is detected in all tissues studied (Fig 6). *MsFTa1* has the highest level in post elongating stem, followed by the leaf and elongating stem. Among all the alfalfa *PEBP* genes, *MsG0780035995*.01 has the highest expression in the leaf, followed by the flower (Fig 6). Furthermore, gene expression retrieved from a recent study (Ma *et al.*, 2021) of two genotypes, early



No scaling is applied to rows. Both rows and columns are clustered using correlation distance and average linkage. FB: Flower bud. NL: New leaf. ML: Mature leaf. 80, early flowering genetype 80. 195, late flowering genotype 195.

Fig 7: Expression profiles of alfalfa *PEBP* genes in early and late flowering genotypes.



No scaling is applied to rows. Both rows and columns are clustered using correlation distance and average linkage. D: Fall dormant cultivar Maverick. ND: Non-fall dormant cultivar CUF101. 5, May. 9, September. D5, fall dormant type in May.

Fig 8: Expression profiles of alfalfa *PEBP* genes in fall dormant and non-fall dormant cultivars.

flowering genotype 80 and late flowering genotype 195, showed that *MsFTa1* and *MsFTa2* are two strong candidates for florigen in alfalfa, while *MsG0480024000.01* is a candidate which functions as a floral repressor (Fig 7). In addition, *MsG0780035995*.01 shows similar expression profiles in both the early and late flowering genotypes and higher expression is detected in the mature leaf than in the new leaf (Fig 7).

Another RNA-seq dataset from our previous study (Zhang *et al*., 2015) which used leaf tissues of one fall dormant cultivar (Maverick, fall dormancy class of 1) and one non-fall dormant cultivar (CUF101, fall dormancy class of 9) under two-time points (May and September in China: 34° N, 113° E), was also retrieved to explore the expression of alfalfa *PEBP* genes. In our previous study, both cultivars showed similar phenotypes in May, while a fall dormancy-specific phenotype was observed in the fall dormant Maverick in September rather than in non-fall dormant CUF101. Gene expression heatmap shows that *MsFTa2* has higher expression in the non-fall dormant cultivar CUF101, compared to the fall dormant type Maverick (Fig 8). Unsurprisingly, *MsG0780035995*.01 has the highest expression among all *PEBP* genes. *MsG0780035995*.01 has higher expression in May compared to September. In September when fall dormancy-specific phenotypes were observed, lower expression of *MsG0780035995*.01 is shown in the fall dormant cultivar Maverick which might suggest that *MsG0780035995*.01 is playing a role in fall dormancy (Fig 8).

CONCLUSION

The identification and analysis of the PEBP family gene in alfalfa showed that there were nine PEBP candidates, which could be divided into three sub-gene families, such as TFL1-like, MFT-like and FT-like groups. These nine PEBP candidates are distributed on five of the eight alfalfa chromosomes, Chromosome 7 has the largest number of PEBP members. The Motif of this gene family had a certain conservative type and there were some differences in the number and types of conservative motifs. MsFTa1 and MsFTa2 are two strong candidates for florigen in alfalfa, while MsG0480024000.01 is a candidate which functions as a floral repressor.

ACKNOWLEDGEMENT

This work was supported by the National Natural Science Foundation of China (32102585).

Conflict of interest

All authors declared that there is no conflict of interest.

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. (2014). A divergent external loop confers antagonistic activity on floral regulators ft and tfl1. Embo Journal. 25(3): 605-614.

Altschul, S.F., Madden, T.L., Schäffer, A.A., Zhang, J., Zhang, Z., Miller, W. and Lipman, D.J. (1997). Gapped BLAST and PSI-BLAST: A new generation of protein database search programs. Nucleic Acids Research. 25(17): 3389-3402.

Bailey, T.L. and Elkan, C. (1994). Fitting a mixture model by expectation maximization to discover motifs in biopolymers. Proc. Int. Conf. Intell. Syst. Mol. Biol. 2: 28-36.

Capella-Gutiérrez, S., Silla-Martínez, J.M. and Gabaldón, T. (2009). trimAl: A tool for automated alignment trimming in large-scale phylogenetic analyses. Bioinformatics. 25(15): 1972-1973.

Carmona, M.J., Calonje, M. and Martínez-Zapater, J.M. (2007). The FT/TFL1 gene family in grapevine. Plant Molecular Biology. 63(5): 637-650.

Chao, J., Kong, Y., Wang, Q., Sun, Y., Gong, D., Lv, J. and Liu, G. (2015). MapGene2Chrom, a tool to draw gene physical map based on Perl and SVG languages. Hereditas. 37(1): 91-97.

Dong, W., Xie, Q., Liu, Z., Han, Y., Wang, X., Xu, R. and Gao, C. (2023). Genome-wide identification and expression profiling of the bzip gene family in betula platyphylla and the functional characterization of BpChr04G00610 under low-temperature stress. Plant Physiology and Biochemistry. 198: 107676.

El-Gebali, S., Mistry, J., Bateman, A., Eddy, S.R., Luciani, A., Potte,r S.C., Qureshi, M., Richardson, L.J., Salazar, G.A., Smart, A., Sonnhammer, E.L.L., Hirsh, L., Paladin, L., Piovesan, D., Tosatto, S.C.E. and Finn, R.D. (2019). The Pfam protein family's database in 2019. Nucleic Acids Research. 47(D1): D427-D432.

Gasteiger, E., Hoogland, C., Gattiker, A., Duvaud, S.E., Wilkins, M.R., Appel, R.D. and Bairoch, A. (2005). Protein Identification and Analysis Tools on the ExPASy Server. In: The Proteomics Protocols Handbook. [Walker JM (ed)], Humana Press, Totowa, NJ, USA. pp 571-607.

Hedman, H., Källman, T. and Lagercrantz, U. (2009). Early evolution of the MFT-like gene family in plants. Plant Molecular Biology. 70(4): 359-369.

Hoang, D.T., Chernomor, O., von Haeseler, A., Minh, B.Q. and Vinh L.S. (2017). UFBoot2: Improving the Ultrafast Bootstrap Approximation. Molecular Biology and Evolution. 35: 518-522.

Hu, B., Jin, J., Guo, A.Y., Zhang, H., Luo, J. and Gao, G. (2015). GSDS 2.0: An upgraded gene feature visualization server. Bioinformatics. 31(8): 1296-1297.

Huang, Y., Niu, B., Gao, Y., Fu, L. and Li, W. (2010). CD-HIT Suite: A web server for clustering and comparing biological sequences. Bioinformatics. 26(5): 680-682.

Kim, D., Paggi, J.M., Park, C., Bennett, C., Salzberg, S.L. (2019). Graph-based genome alignment and genotyping with HISAT2 and HISAT-genotype. Nature Biotechnology. 37: 907-915.

Lescot, M., Déhais, P., Thijs, G., Marchal, K., Moreau, Y., Van de Peer, Y., Rouzé, P. and Rombauts, S. (2002). PlantCARE, a database of plant cis-acting regulatory elements and a portal to tools for in silico analysis of promoter sequences. Nucleic Acids Research. 30(1): 325-327.

Liu, Y.Y., Yang, K.Z., Wei, X.X. and Wang, X. (2016). Revisiting the phosphatidylethanolamine binding protein (PEBP) gene family reveals cryptic flowering locus t gene homologs in gymnosperms and sheds new light on functional evolution. New Phytologist. 212(3): 730-744.

Ma, D., Liu, B., Ge, L., Weng, Y., Cao, X., Liu, F., Mao, P. and Ma, X. (2021). Identification and characterization of regulatory pathways involved in early flowering in the new leaves of alfalfa (*Medicago sativa* L.) by transcriptome analysis. BMC Plant Biology. 21(1): 8.

Metsalu, T. and Vilo, J. (2015). ClustVis: A web tool for visualizing clustering of multivariate data using principal component analysis and heatmap. Nucleic Acids Research. 43(W1): W566-W570.

O'Rourke, J.A., Fu, F., Bucciarelli, B., Sam Yang, S., Samac, D.A., Lamb, J.F.S., Monteros, M.J., Graham, M.A., Gronwald, J.W., Krom, N., Li, J., Dai, X., Zhao, P.X. and Vance, C.P. (2015). The *Medicago sativa* gene index 1.2: A web-accessible gene expression atlas for investigating expression differences between *Medicago sativa* subspecies. BMC Genomics. 16: 502.

Roberto, V.A, Lorinc, S.P., Leonardo, M.R. and David, L. (2018). TPMCalculator: one-step software to quantify mRNA abundance of genomic features. Bioinformatics. 35(11): 1960-1962.

Shen, C., Du, H.L., Chen, Z., Lu, H.W., Zhu, F.G., Chen, H., Meng, X.Z., Liu, Q.W., Liu, P., Zheng, L.H., Li, X.X., Dong, J.L., Liang, C.Z. and Wang, T. (2020). The chromosome-level genome sequence of the autotetraploid alfalfa and resequencing of core germplasms provide genomic resources for alfalfa research. Molecular Plant. 13(9): 1250-1261.

Thompson, J.D., Higgins, D.G. and Gibson, T.J. (1994). CLUSTAL W: Improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. Nucleic Acids Research. 22(22): 4673-4680.

Trifinopoulos, J., Nguyen, L.T., von Haeseler, A. and Minh, B.Q. (2016). W-IQ-TREE: A fast online phylogenetic tool for maximum likelihood analysis. Nucleic Acids Research 44(W1): W232-W235.

Wang, Z., Zhou, Z., Liu, Y., Liu, T., Li, Q., Ji, Y., Li, C., Fang, C., Wang, M., Wu, M., Shen, Y., Tang, T., Ma, J. and Tian, Z. (2015). Functional evolution of phosphatidylethanolamine binding proteins in soybean and Arabidopsis. Plant Cell. 27(2): 323-336.

Zhang, S., Shi, Y., Cheng, N., Du, H., Fan, W. and Wang, C. (2015). De novo characterization of fall dormant and nondormant alfalfa (*Medicago sativa* L.) leaf transcriptome and identification of candidate genes related to fall dormancy. PLOS ONE. 10(3): e0122170.

Zhu, Y., Peng, Z.Y., Zhang, B. and Bi, Y.P. (2013). Advances in function research of PEBP family genes in plants. Shandong Agricultural Sciences. 45(2): 139-145.