

Supplementary material

Exon Junction Complex (EJC) Core Genes Play Multiple Developmental Roles in *Physalis floridana*

Pichang Gong^{1,#}, Jing Li^{1,2,#}, Chaoying He^{1,2,*}

¹ State Key Laboratory of Systematic and Evolutionary Botany, Institute of Botany, Chinese Academy of Sciences, Beijing 100093 China;

² University of Chinese Academy of Sciences, Beijing 100049 China.

These authors contributed equally to this work.

* For correspondence: E-mail: chaoying@ibcas.ac.cn; Tel: +86-10-62836085

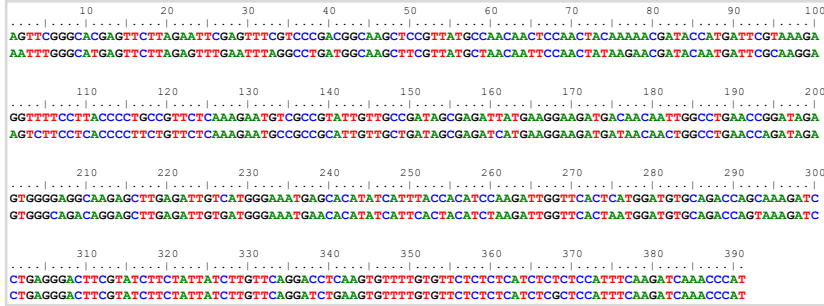
Supplementary material contains:

Supplementary Figures S1-S13

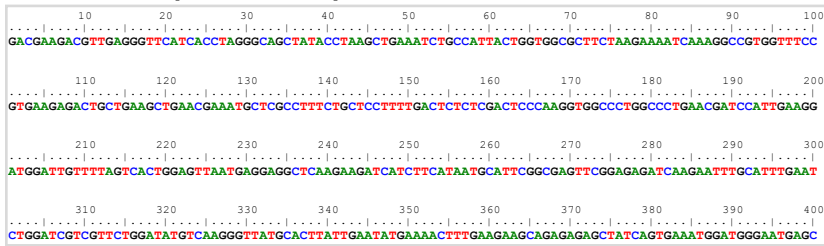
Supplementary Tables S1-S3

PFMAGO-VIGS probe; 391 bp

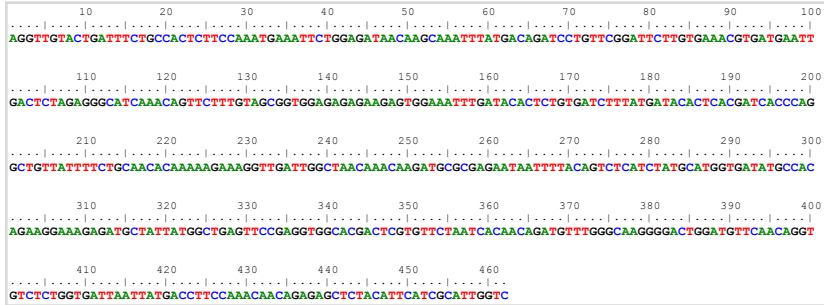
(Sequence identity between *PFMAGO1* and *PFMAGO2* was 89.0%)



PFY14-VIGS probe; 400 bp



PFelF4AIII-VIGS probe; 462 bp



PFBTZ-VIGS probe; 370 bp

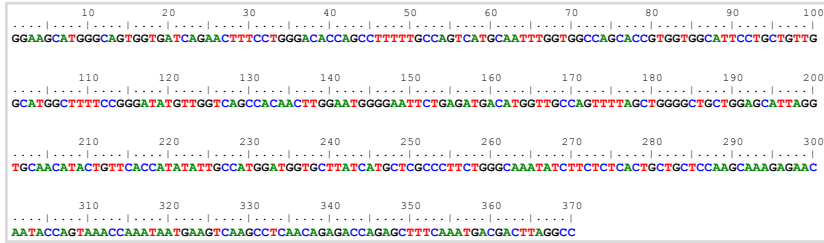


Fig. S1 Sequences used to trigger gene silencing in VIGS. From top to bottom, the trigger sequence of *PFMAGO*-, *PFY14*-, *PFelF4AIII*- and *PFBTZ*-, and each length is given.

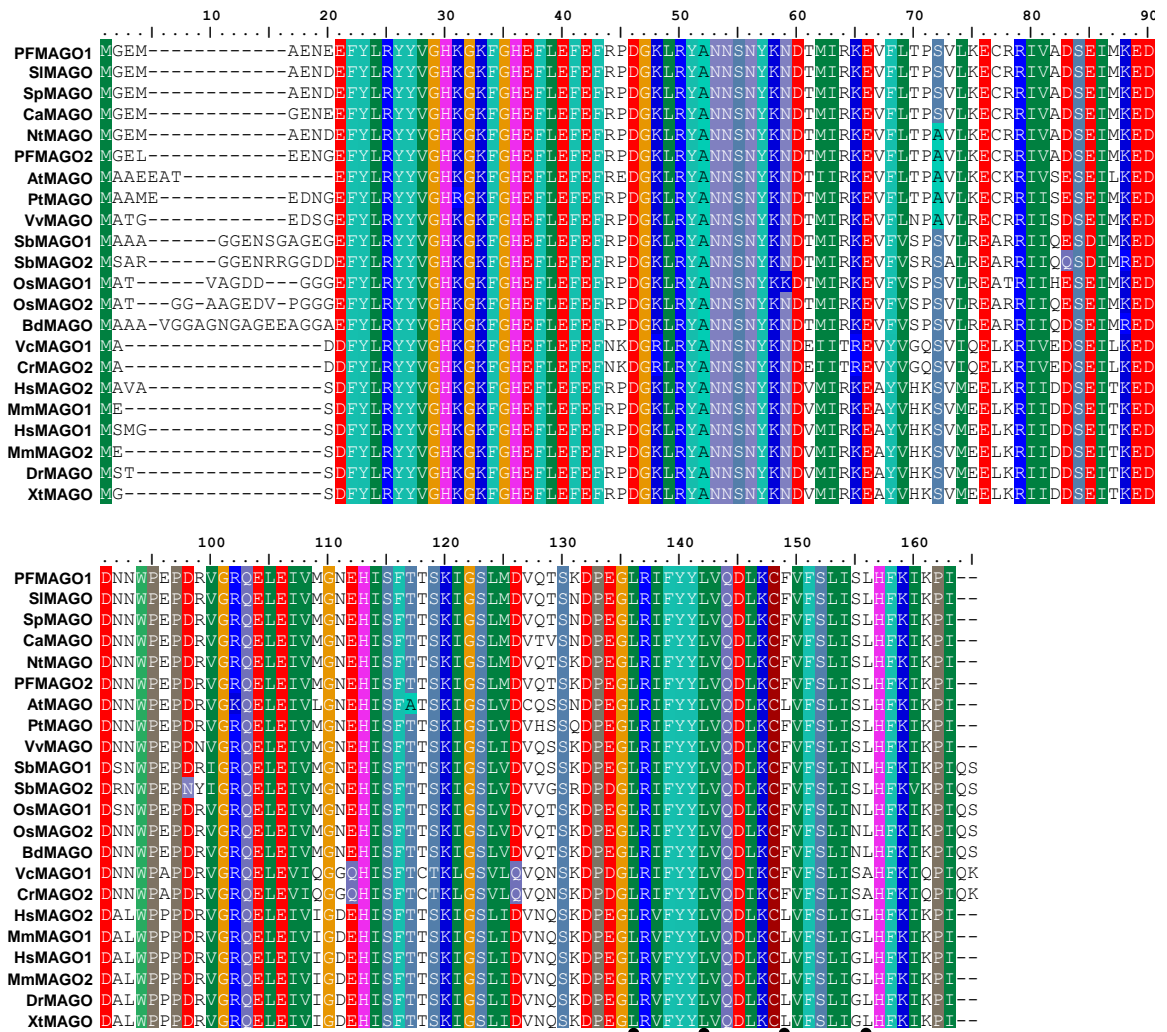


Fig. S2 Multiple sequence alignment of the MAGO protein family.

Shaded regions with color indicate 100% amino acid similarity. Four conserved leucine residues that constitute a potential leucine zipper in the C-terminus (Pozzoli et al. 2004; Chu et al. 2009) are labeled by black cycles.

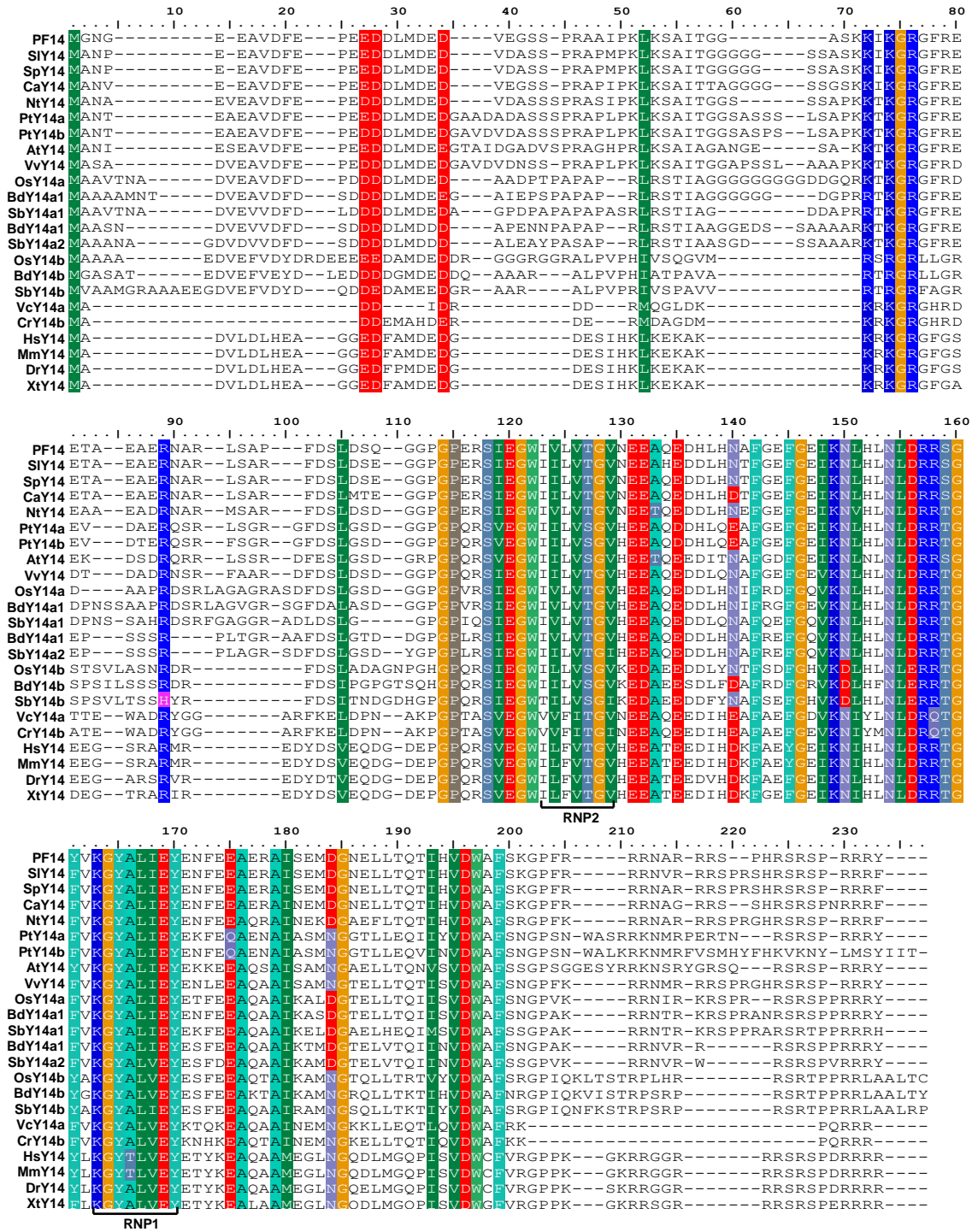


Fig. S3 Multiple sequence alignment of the Y14 protein family. Shaded regions with color indicate 100% amino acid similarity. Conserved RNA recognition motifs of RNP1 and RNP2 (Lau et al. 2003; Shi and Xu 2003; Chu et al. 2009) are labeled in the middle area.

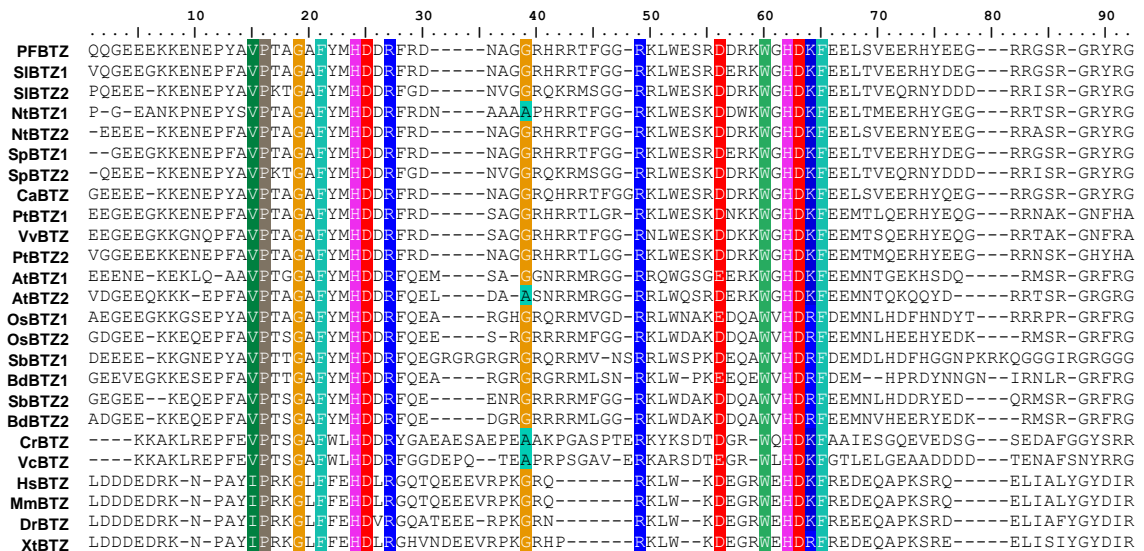


Fig. S5 Multiple sequence alignment of the SELOR domain in the BTZ protein family.

Only the SELOR domain (Degot et al. 2004), part of the whole MSA, is shown. Shaded regions with color indicate 100% amino acid similarity.

A	PFMAGO1	PFMAGO2	PFY14	PFelF4AIII	PFBTZ
Residue range	9-151	9-151	77-170	5-391	144-217
Tertiary structural modeling	Nt- β 1- β 2- β 3- β 4- α 1- β 5- β 6- α 2- α 3-Ct	Nt- β 1- β 2- β 3- β 4- α 1- β 5- β 6- α 2- α 3-Ct	Nt- β 1- α 1- β 2- β 3- α 2- β 4- β 5-Ct	Nt- α 1- α 2- β 1- α 3- β 2- α 4- β 3- α 5- β 4- β 5- α 6- α 7- β 6- α 8- β 7- β 8- α 9- β 9- α 10- β 10- α 11- β 11- α 12- β 13- α 13- β 13- α 14- α 15-Ct	No α -helix nor β -sheet
Template (PDB ID)	X-Ray, 2.30 Å, PDB ID: 3EX7				
Identity (%)	82.52	81.25	54.29	78.97	27.63
GMQE	0.89	0.88	0.43	0.92	0.03
QMEAN4	-1.05	-0.95	-0.22	-0.46	-5.24
Organism	<i>Homo sapiens</i>				

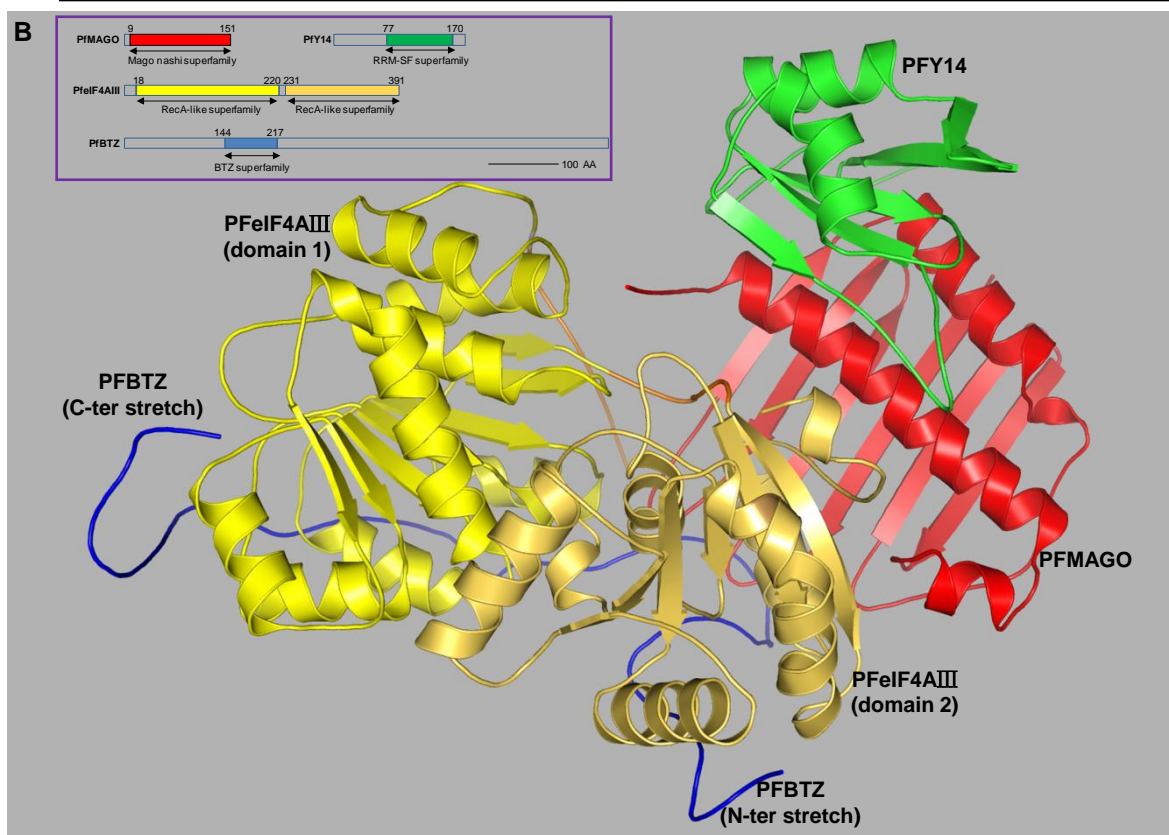


Fig. S6 Tetrameric structure modeling of EJC core components in *P. floridana*.

a Quality factors of *Physalis* EJC protein structure modeling. Nt, protein N-terminal; Ct, protein C-terminal; α , α -helix; β , β -sheet. **b** The tetrameric structure reconstruction of the *Physalis* EJC core based on human EJC (PDB ID: 3EX7). Conserved protein domains of EJC core components were predicted using the Pfam program and are presented in the pink box. The position of each conserved region or motif is given. Bar, 100 amino acids (AA).

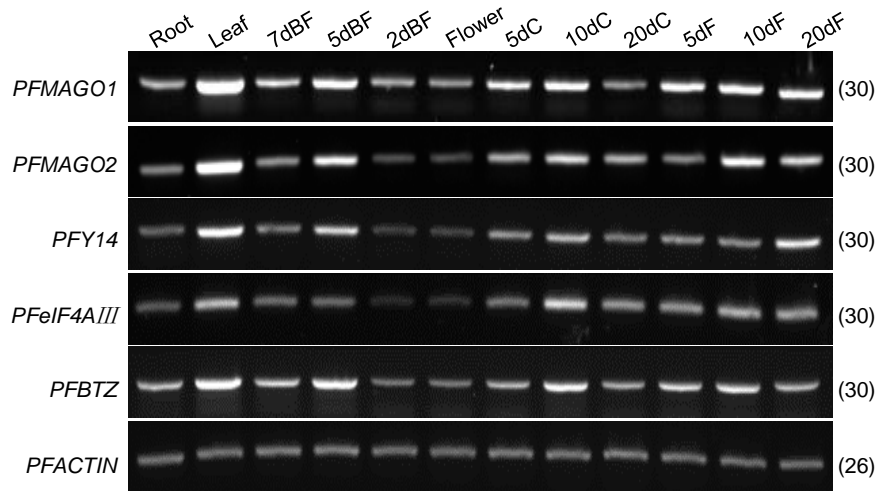


Fig. S7 Expression of *Physalis* EJC core genes by semi-quantitative RT-PCR. The total RNAs from roots, leaves, flower buds (7dBF, 5dBF, and 2dBF abbreviated for the flower buds of 7-, 5-, and 2-d before flowering), blooming flowers, fruiting calyx (5dC, 10dC and 20dC is short of the calyx of 5-, 10- and 20-d after fertilization) and developing berries (5dF, 10dF and 20dF represent fruits of 5-, 10- and 20-d after fertilization) were subjected to RT-PCR. The PCR products of 30 cycles for each EJC genes were demonstrated; 26 cycles of *PFACTIN* gene were shown as a loading control.

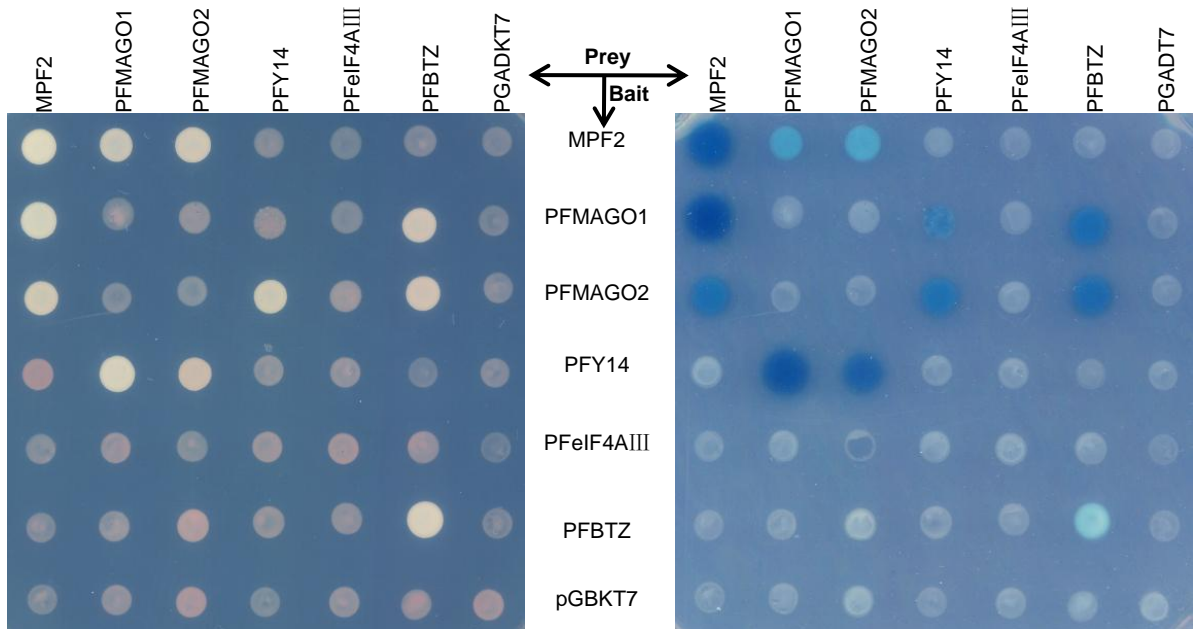


Fig. S8 PPIs among *Physalis* EJC core proteins revealed by yeast two-hybrid analysis.

Yeast cell growth on a high-stringency selective plate (SD/-Leu-Trp-His-Ade) (Left panel) and nonlethal β -galactosidase tests were used to confirm the interactions (right panel). The empty pGADT7 and pGBKT7 constructs were used as the negative controls.

A

Gene/VIGS	Infected seedling number	Mutated seedling number	Gene expression of mutants at 14 DAI compared to NC	Survival number (ratio)	Pollen viability	Fruit setting rate by self / cross-pollination
WT	-	-	1.00±0.05 (<i>PFMAGO1</i>) 1.06±0.16 (<i>PFMAGO2</i>) 0.99±0.21 (<i>PFY14</i>) 0.95±0.11 (<i>PFelF4AIII</i>) 1.03±0.18 (<i>PFBTZ</i>)	30 (100%)	96.1%	95.5%/94.3%
<i>TRV2</i> (NC)	30	-	1.0 (for each gene)	30 (100%)	95.8%	95.3%/94.6%
<i>PFMAGO</i>	120	115	0.26±0.13 (<i>PFMAGO1</i>) 0.32±0.12 (<i>PFMAGO2</i>)	25 (21.7%)	45.2%	35.3%/89.5%
<i>PFY14</i>	120	112	0.22±0.15	0 (0)	35.3%	0/0
<i>PFelF4AIII</i>	120	113	0.35±0.21	44 (38.9%)	90.9%	95.2%/94.1%
<i>PFBTZ</i>	120	115	0.33±0.18	115 (100%)	50.1%	67.5%/92.3%

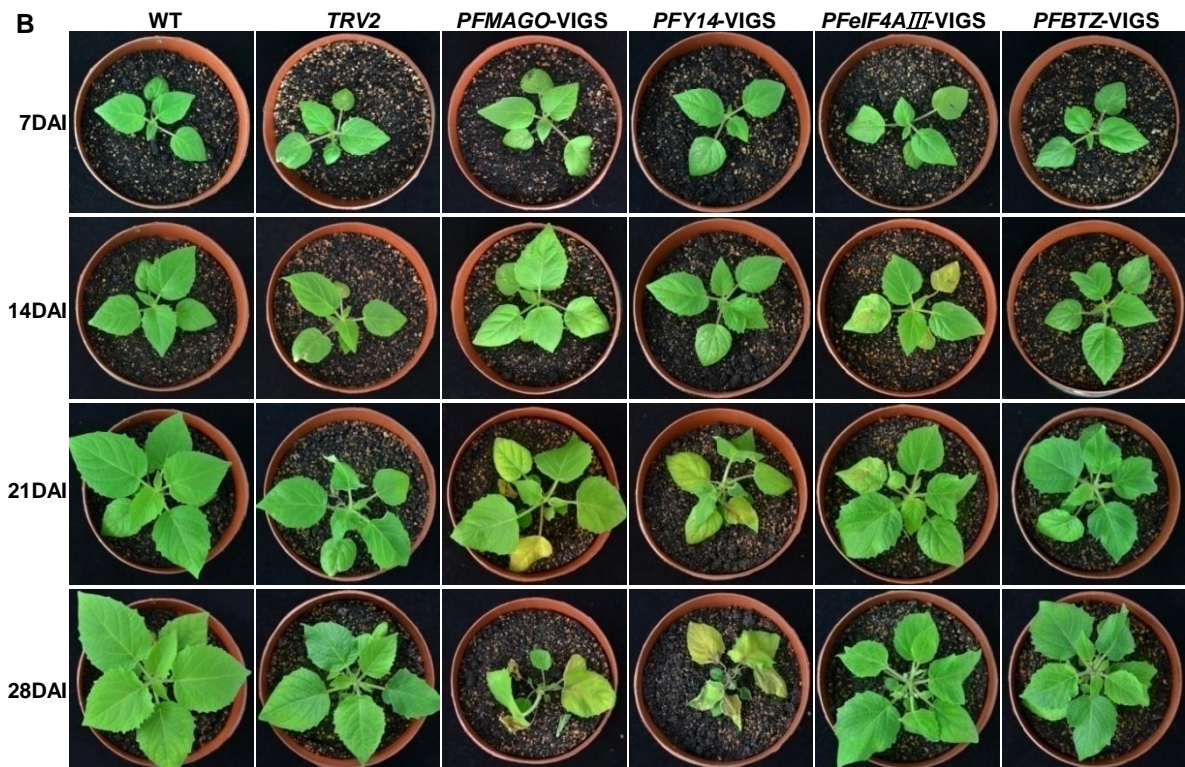


Fig. S9 Overall analysis of EJC core gene-VIGS in *P. floridana*.

a Evaluated parameters are as indicated. **b** Phenotypic variation of VIGS plants at vegetative stages compared to WT and NC. 7DAI, 14DAI, 21DAI and 24DAI indicated 7-, 14-, 21- and 24-d after infection, respectively.

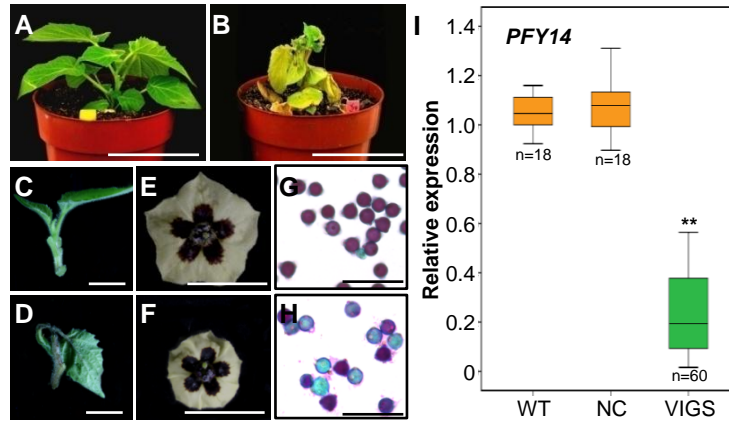


Fig. S10 *PFY14*-VIGS analysis in *P. floridana*.

a, b Plant growth of NC and *PFY14*-VIGS plants at 28DAI stage. **c, d** Shoot apex of NC and *PFY14*-VIGS plants. **e, f** Flower morphology of NC and *PFY14*-VIGS plants. **g, h** Pollen viability of NC and *PFY14*-VIGS plants, which was evaluated by I₂-KI staining. **i** The *PFY14* expression in flowers of plants as indicated. WT, wild type; NC, negative control; VIGS, *PFY14*-VIGS flowers showing severe pollen abortion. The flower number (n) investigated in each case was indicated. Bars, 10 cm in **a, b**, 1 cm in **c-f**, 100 μ m in **g, h**. Significance relative to NC was evaluated by a two-tailed Student's *t* test, and ** indicates significance at $P < 0.01$.

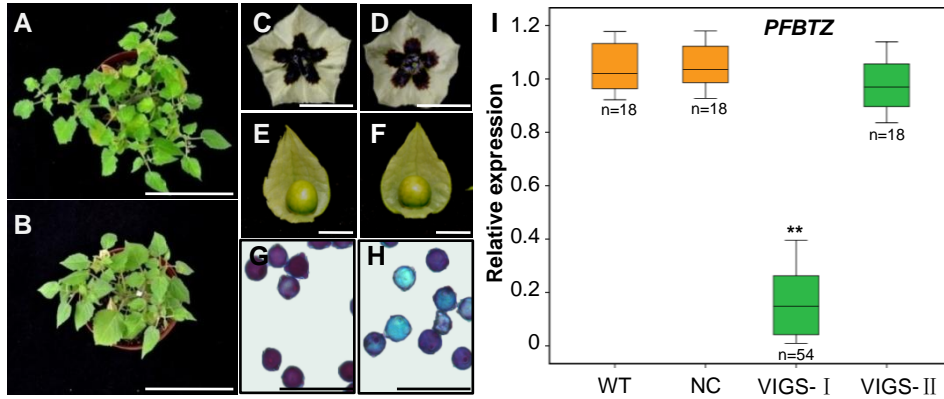


Fig. S11 *PFBTZ*-VIGS analysis in *P. floridana*.

a, b Phenotype of NC and *PFBTZ*-VIGS plants at the fruiting stage. **c, d** Flower morphology of NC and *PFBTZ*-VIGS plants. **e, f** Fruit morphology of NC and *PFBTZ*-VIGS plants. **g, h** Pollen viability of NC and *PFBTZ*-VIGS flowers, which was evaluated by I₂-KI staining. **i** The *PFBTZ* expression in flowers of plants as indicated. WT, wild type; NC, negative control; VIGS-I, *PFBTZ*-VIGS flowers showing pollen abortion; VIGS-II, *PFBTZ*-VIGS flowers having WT-like pollen development. The flower number (n) investigated in each case was indicated. Bars, 10 cm in **a, b**, 1 cm in **c-f**, 100 μm in **g, h**. Significance relative to NC was evaluated by a two-tailed Student's *t* test, and ** indicates significance at $P < 0.01$.

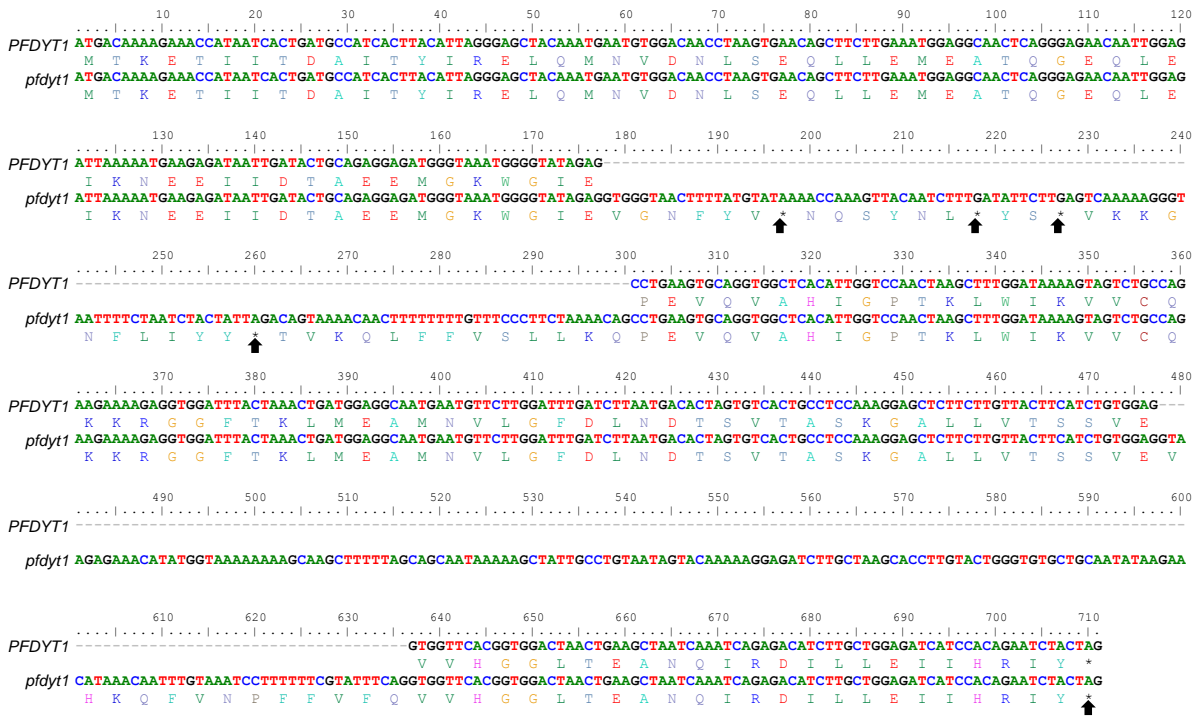


Fig. S12 Sequence comparison of different *PFDYT1* transcripts.

The normal cDNA of *PFDYT1* was 429 bp long, and putatively encoded a peptide of 142 amino acids (aa). Aberrant transcript (*pfdyt1*) was 711 bp, and maximally encoded a putative 65-aa peptide. Stop codon (*) was highlighted by black arrows. The - indicates the nucleotide acid gap.

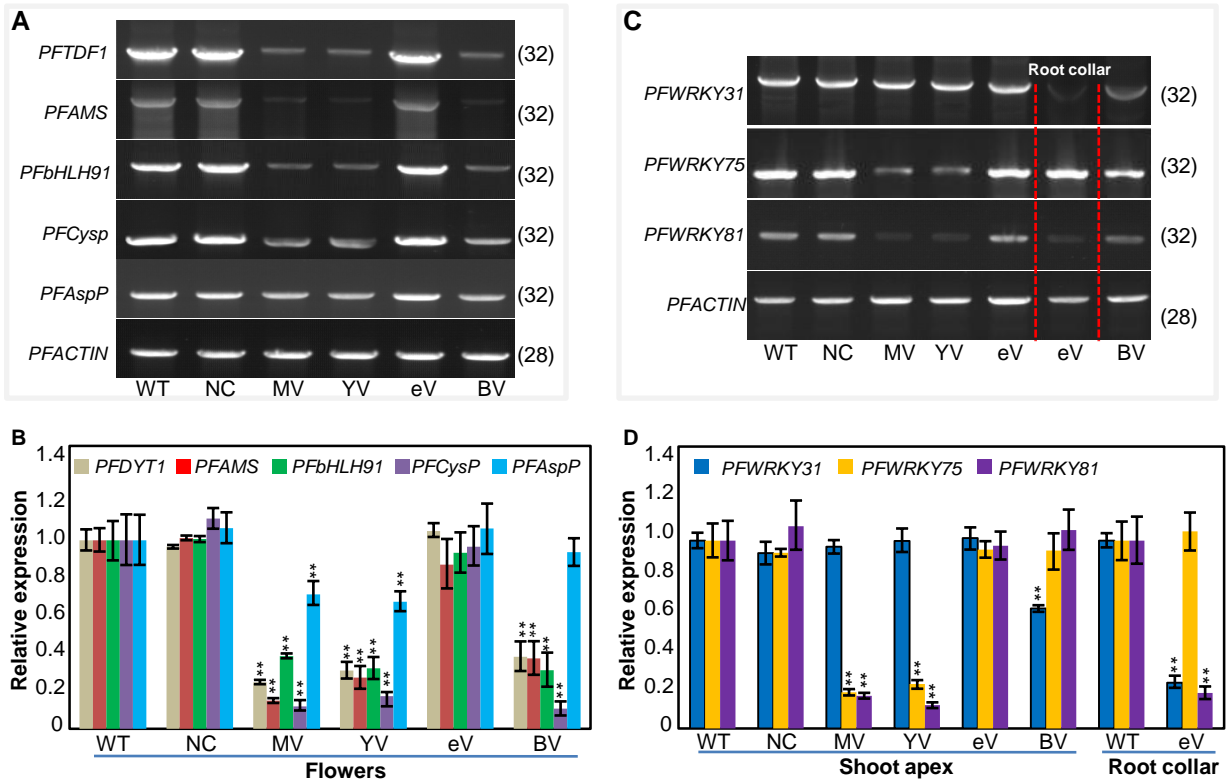


Fig. S13 Downregulations of EJC core genes affect gene expression.

a Full-length transcripts of the indicated genes in flower organs. WT, NC, MV, YV, eV and BV respectively represent wild type, negative control (*TRV2*), and *PFMAGO*-, *PFY14*-, *PfeIF4AIII*- and *PFBTZ-VIGS* *P. floridana* plants. The cycle number is given in the bracket. **b** The gene expression in 7DBF floral buds of the indicated plants. **c** Full-length transcripts of *PFWRKY* family genes in shoot apex or root collar. WT, NC, MV, YV, eV and BV respectively represent wild type, negative control (*TRV2*), and *PFMAGO*-, *PFY14*-, *PfeIF4AIII*- and *PFBTZ-VIGS* *P. floridana* plants. The sample between the two dashed lines is root collar, and others are the shoot apex. The cycle number is given in the bracket. **d** The gene expression of *PFWRKY* family in the shoot apex or root collar of the indicated plants. Significance relative to WT was evaluated by two-tailed Student's *t* test, and ** indicates significance at $P < 0.01$.

Table S1 Primers used in the present study.

Usage	Primer name	Forward sequence	Reverse sequence
qRT-PCR	<i>PFMAGO1</i>	AGAATTCGAGTTTCGTCGCCGAC	TGACAATCTCAAGCTCTTGCCTC
	<i>PFMAGO2</i>	AGTTTACCTGAGATATTACGTG	AAGGGGTGAGGAAGACTTCCCTTG
	<i>PFY14</i>	AATCAAAGGCCGTGTTCCG	AGCCTCCTCATTAACTCCAGTG
	<i>PfeIF4AIII</i>	TATTATGGCTGAGTCCGAGGTG	CGACCAATGCGATGAATGTAG
	<i>PFBTZ</i>	GTGGTGGCATTCTGCTGTTG	TTCTGACGCTGGCCTAAGTCG
	<i>PFDYT1 (E1F-E3R)</i>	ATGACAAAAGAAACCATAATCACTG	CTAGTAGATCCGTGGATGATCTC
	<i>PFDYT1 (E1F-I1R)</i>		GTTGTTTTACTGCTAATAG
	<i>PFDYT1 (I1F-E2R)</i>	AATCTTTGATATTCTTGAGTC	TCCAAGAACATTCTTGCCTC
	<i>PFDYT1 (I2F-E3R)</i>	GCTAAGCACCTTGTACTGGG	CTAGTAGATCCGTGGATGATCTC
	<i>PFDYT1 (E3F-E3R)</i>	CCTGAAGTGCAGGTGGCTCAC	
	<i>PFTDF1</i>	ACTAGCTGGGCAACTAGCTGGG	CTTTGTTGACCCATCTTGAGCACC
	<i>PFAMS</i>	GGATTGCTCGAGTTGTTCAGTACG	TCCGAATTTGATCTCCGTTATC
	<i>PFBHLH</i>	TACCAAGTTTAGGTCTCGAGAAG	GAGGTATGGAGCTTCAGGCATTGG
	<i>PFA_{sp}P</i>	GCTACCCTAGGAGTTGGACAACC	GCAAACTGAGTCACTAGGTCCG
	<i>PFC_{ys}P</i>	GTTGTGGAAGCTGTTGGGCGTTCTC	GCCACTTATTGAACTGATGCTG
	<i>PFWRKY31</i>	GTCAGATGATGGATACAATTGGAG	GATATCGATCGCTTGATGTTG
	<i>PFWRKY75</i>	TCTTGGGCTAATGGCAAGCATGG	GGAAGCATAGATTGTCCTGACTC
<i>PFWRKY81</i>	TAATCGTGGTCGGGAGTTTACC	GTCTTGAATTTGTCACCTGTTTTG	
<i>PFACTIN</i>	AACCGAATGGTGAAGCTGG	CCATATCATCCCAATTGCTGAC	
RT-PCR	<i>PFDYT1</i>	ATGACAAAAGAAACCATAATCACTG	CTAGTAGATCCGTGGATGATCTC
	<i>PFTDF1</i>	ATGGGAGGCCTCCTGTTGC	CTAGTAAAGGGATCCTCAGAG
	<i>PFAMS</i>	ATGGAGAGGCTTAGCCCTATTATG	GTGATGGGATGTCGGTGAAGTG
	<i>PFBHLH</i>	AGCATGAAGGAGTAGCTGAAGATG	GAGCTCTTGCTATTGCACTTGC
	<i>PFA_{sp}P</i>	GAACCTCAGACCCCTGTCACG	ATCAGGACTAAATCCAACAAGG
	<i>PFC_{ys}P</i>	ATGAAGCCATCTTATATTA	TCAATAATCCCTTAAGAGG
	<i>PFWRKY31</i>	ATGGCTGCTCAAGTTTCTCTTTTC	TCAGCAAAGAAATGACTCCATAAAC
	<i>PFWRKY75</i>	TATTCGCCAATATGTCCATCG	TTAGAAGGAAGCATAGATTG
<i>PFWRKY81</i>	ATGGGATCTTATTGCCGGAA	GTCTTGAATTTGTCACCTGTTTTG	
<i>PFACTIN</i>	ATGGCTGATGGTGAAGATATCC	GAAGGACCAGACTCATCACTC	
Y2H	<i>PFMAGO1</i>	GGCCATGGGGAAATTGGAAGAGAATG	ATGGATCCGGTTTGATCTTGAATGGAG
	<i>PFMAGO2</i>	GGCCATGGGGAGATGGCAGAGAAC	ATGGATCCGGTTTGATCTTGAATGGAG
	<i>PFY14</i>	ATCCATGGGGAACGGGAGGAGGCAG	ATGGATCCTTAGTATCTCCTCTGGGACTTC
	<i>PfeIF4AIII</i>	TGCCATGGAAGAAGACCCGTTAGTTTTTCG	AAGGATCCTATCAATTCAGCCACGTTCAATTGG
	<i>PFBTZ</i>	TGACATATGGCGAGTGCTTGGAAGAAGCAG	AAGGATCCTCACTGATCAAACCTCATCTCCG
BIFC	<i>PFMAGO1</i>	CACGGGGACTCTAGAATGGGGAAATGGAAGAGAATG	CGATGGATCCCACTAGTAAATGGTTGATCTTGAATG
	<i>PFMAGO2</i>	CACGGGGACTCTAGAATGGGGAGATGGCAGAGAAC	CGATGGATCCCACTAGTAAATGGTTGATCTTGAATG
	<i>PFY14</i>	CACGGGGACTCTAGAATGGGAACGGGAGGAGGC	CGATGGATCCCACTAGTATCTCCTCTGGGACTTC
	<i>PfeIF4AIII</i>	CACGGGGACTCTAGAATGGAAGAAGACCCGTTAGTTTTTC	CGATGGATCCCACTAGTATCAATTCAGCCACGTTCAATTGG
<i>PFBTZ</i>	CACGGGGACTCTAGAATGGCGAGTGTGGAGAAGCAG	CGATGGATCCCACTAGTCTGATCAAACCTCATCTCCGT	
Subcellular localization	<i>PFMAGO1</i>	AACCATGGGGAAATTGGAAGAGAATG	ATACTAGTAAATGGTTTGATCTTGAATGG
	<i>PFMAGO2</i>	ATCCATGGGGAGATGGCAGAGAAC	ATACTAGTAAATGGTTTGATCTTGAATGG
	<i>PFY14-1302</i>	AACCATGGGAACGGGAGGAGGCAG	CCACTAGTGTATCTGCTCTGGGACTTC
	<i>PfeIF4AIII</i>	CACCATGGAAGAAGACCCGTTAGTTTTTC	CCACTAGTATCAATTCAGCCACGTTCAATTGG
	<i>PFBTZ-1300</i>	AATCTAGATGGCGAGTGCTTGGAAGAAGCAG	AAGGATCCCTGATCAAACCTCATCTCCGATATC
VIGS	<i>PFMAGO</i>	ACGGATCCAGTTCGGCACGAGTTCT	TGCTCGAGATGGTTGATCTTGA
	<i>PFY14</i>	ACGGATCCGACGAAGACGTTGAGGG	TGCTCGAGGCTCATTCCATCCATT
	<i>PfeIF4AIII</i>	ACGGATCCAGTTGTTACTGATTT	TGCTCGAGCGACCAATGCGATGAA
	<i>PFBTZ</i>	ACGGATCCGGAAGCATGGCGAGT	TGCTCGAGGGCCTAAGTCGCTATTG

Table S2 Molecular information of the isolated *Physalis* EJC core genes.

Gene	Accession number	ORF length (bp)	Protein length (Amino acid)	Grand average of hydropathicity
<i>PFMAGO1</i>	EF205415	456	151	-0.523
<i>PFMAGO2</i>	EF205416	456	151	-0.508
<i>PFY14</i>	KF051015	567	188	-0.973
<i>PFelF4AIII</i>	MH319840	1176	391	-0.171
<i>PFBTZ</i>	MH319841	2067	688	-1.024

Table S3 Sequences used in the present study.

Species	Abbreviation	Accession number			
		MAGO	Y14	eIF4A3	BTZ
<i>Physalis floridana</i>	PF	EF205415	KF051015	MH319840	MH319841
		EF205416			
<i>Solanum lycopersicum</i>	Sl	NM_001291909	NM_00129108	XM_004250609	XP_010318560
					XP_004252328
<i>Solanum pennellii</i>	Sp	XP_015067364	XP_01508009	XP_015058616	XP_015070832
					XP_015060990
<i>Nicotiana tomentosiformis</i>	Nt	XP_009593407	XP_009586749	XP_009616974	XP_009586607
					XP_009595299
<i>Capsicum annuum</i>	Ca	XP_016558805.1	XP_016574886	XP_016554349	PHT88780
<i>Populus trichocarpa</i>	Pt	XM_006375098	XM_002299753	XM_002306354	XM_006369105
			XM_002314049	XM_002310627	XM_002303242
<i>Arabidopsis thaliana</i>	At	NM_100094	NM_104029	NM_112866	NM_202458
					NM_101397
<i>Vitis vinifera</i>	Vv	XP_002281294	XP_002281228	XP_002275011	XP_010654695
<i>Oryza sativa</i>	Os	KF051011	KF051016	NM_001050208	XP_015650091
		KF051008	KF051017	NM_001057041	XP_015640548
<i>sorghum bicolor</i>	Sb	02g027660	XP_002439249	04g038330	XP_002455883
		07g027660	XP_002457845		XP_021303601
<i>Brachypodium distachyon</i>	Bd	XP_003573269	XP_003568971	XP_003570724	XP_003569296
			XP_003567720		XP_003565997
<i>Volvox carteri nagariensis</i>	Vc	XM_001770356	XP_002953417	XP_002958883	XP_002951521
<i>Chlamydomonas reinhardtii</i>	Cr	XM_001763749	XP_001696992	XP_001699375	XP_001698374
<i>Homo sapiens</i>	Hs	NM_002370	NM_005105	NM_014740	XM_005257163
		NM_018048			
<i>Mus musculus</i>	Mm	NM_001282737	NM_001102407	NM_138669	NM_138660
		NM_025564			
<i>Danio rerio</i>	Dr	NP_001017700	NP_001013363	NP_957372	NM_205716
<i>Xenopus tropicalis</i>	Xt	XP_002931471	NP_001039147	NP_001084200	NM_001012652
				NP_001107349	