# A rapid alkalinization factor-like peptide EaF82 impairs tapetum degeneration during pollen development

Chiu-Yueh Hung[1], Farooqahmed S. Kittur[1], Keely N. Wharton[1,a], Jianjun Chen[2], Makendra L. Umstead[1,b], D'Shawna B. Burwell[1], Martinique Thomas[1,c], Qi Qi[1], Jianhui Zhang[1], Carla E. Oldham[1], Kent O. Burkey[3] and Jiahua Xie[1,*]

[1] Department of Pharmaceutical Sciences, Biomanufacturing Research Institute & Technology Enterprise, North Carolina Central University, Durham, 27707, USA

[2] Environmental Horticulture Department and Mid-Florida Research and Education Center, Apopka, University of Florida, Apopka, 32703, USA

[3] USDA-ARS Plant Science Research Unit and Department of Crop and Soil Sciences, North Carolina State University, Raleigh, 27695, USA

[a] Present address: Yale University, School of Medicine, New Haven, CT 06520, USA

[b] Present address: EMD Serono Inc., Rockland, MA 02370, USA

[c] Present address: Rho Inc., Durham, NC 27713, USA

[*] Corresponding author: Jiahua Xie (jxie@nccu.edu); 1-919-530-6705; Department of Pharmaceutical Sciences, North Carolina Central University, 1801 Fayetteville Street, Durham, 27707, USA

**Abstract**

In plants, many small peptides have been discovered and are thought to function as growth regulators. However, the functions of most of them remain unknown. In this study, we systematically characterized EaF82, a novel cysteine-rich peptide isolated from shy-flowering 'Golden Pothos' plants. Our studies revealed that EaF82 is closely related to the rapid alkalinization factors (RALFs) and displays alkalinizing activity. Its heterologous expression in Arabidopsis impaired tapetum degeneration and reduced pollen production and seed yields. RNA-Seq, RT-qPCR and biochemical analyses showed that overexpressing *EaF82* down-regulated a group of genes involved in pH changes, cell wall modifications, tapetum degeneration and pollen maturation as well as endogenous Arabidopsis clade IV *AtRALFs*, and decreased proteasome activity and ATP levels. Yeast two-hybrid screening and co-immunoprecipitation identified AKIN10, an energy sensing kinase, to be its interacting partner. These results suggest that EaF82 action may be mediated through AKIN10 leading to alterations of transcriptome and energy metabolism, thereby causing ATP deficiency and impairing pollen development. Our study reveals a new regulatory role for RALF peptide in tapetum degeneration.

**Key words:** ATP deficiency; cysteine-rich peptide; *Epipremnum aureum*; pollen production; kinase AKIN10; rapid alkalinization factor; tapetum degeneration.

## Introduction

Polypeptide hormones have long been recognized as growth and development regulators in animals and yeast; but little is known about such polypeptides in plants until the discovery of systemin in 1991[1]. Since then, numerous plant peptide hormone sequences were identified[2] and the quest for understanding their functions and action modes is growing. One such peptide family is rapid alkalinization factors (RALFs), which are small cysteine-rich peptides (CRPs) and widely present in plant kingdom[3]. Many RALF sequences have been identified from genomic sequences[4,5]. A phylogenetic analysis of 795 RALFs from 51 species showed that they fall into four distinct clades[5]. However, only a handful of RALFs belonging to clades I, II and III have been functionally characterized to be involved in cell expansion, root and root hair development, pollen tube elongation/rupture and immunity, while the majority of them have not been studied[6]. The clade IV is the least investigated and distinct from other clades with greater evolutionary divergence[5,6].

Studies related to the function and mechanism of RALFs are attracting greater attention due to their importance in plant growth and development, and defense against biotic and abiotic stresses[7,8]. Recent studies on Arabidopsis AtRALF1/23/4/19 demonstrated that their actions are perceived through interactions with CrRLK1L (*Catharanthus roseus* Receptor Like Kinase 1 Like) receptors[6]. Binding of AtRALFs to CrRLK1L receptors triggers downstream signal transduction including down-regulation of $H^+$-ATPase resulting in extracellular pH increase. The increased extracellular pH was thought to strengthen the cell wall to prevent pathogen invasion and make alkaline apoplasts an unfavorable environment for pathogens[9]. The pH increase also counteracts the acidification-induced loosening of cell wall leading to inhibition of root growth and pollen tube elongation[10]. As for clade IV RALFs, many members are thought to be involved in pollen development based on their transcriptional patterns. Studies on male sterile Chinese cabbage lines found that three out of 14 identified differentially expressed *RALFs* belong to clade IV[11]. Studies on four pollenless Arabidopsis tapetum mutants also revealed that five out of seven differentially expressed *RALFs* are members of clade IV[12]. In Arabidopsis, six out of eight pollen-abundant *AtRALFs* (4/8/9/15/19/25/26/30) were found to belong to clade IV[13]. However, to date, their precise roles and underlying mechanism of action in pollen development are still unknown.

We previously isolated a novel gene *EaF82* from shy-flowering 'Golden Pothos' plant (*Epipremnum aureum*)[14,15] which encodes a small CRP[16]. Its expression was positively regulated by auxin[16], but its function was unknown because of the shy-flowering nature of 'Golden Pothos' plants, which makes its functional study challenging. In the present study, we show that EaF82 is closely related to clade IV RALFs and the synthetic EaF82 peptide can induce alkalinization of tobacco suspension cell culture medium as typical RALFs. Overexpressing *EaF82* in Arabidopsis and tobacco plants was found to dramatically reduce seed-setting. RNAseq analysis was used to identify those affected genes in Arabidopsis transgenic lines overexpressing *EaF82*. Furthermore, the interacting partners of EaF82 were screened by the yeast two-hybrid (Y2H) assay and verified, and the downstream effects of decreased ATP levels caused by accumulated EaF82 were investigated. Our studies reveal the inhibitory role of EaF82 in tapetum degeneration. Its potential action mechanism through binding to AKIN10, a catalytic α-subunit of energy sensor - sucrose non-fermenting related kinase 1 (SnRK1), is discussed.

## RESULTS

**EaF82 is a clade IV RALF-like peptide.** Our previous search for EaF82 homologs using protein BLAST only found an antimicrobial peptide MiAMP1 from *Macadamia integrifolia* as the closest match with E-value of 0.094 and 48% similarity[16]. In the current study, we further analyzed its sequence features and used a recent comprehensive peptide classification system based on peptide structural features and biological functions[8] to find its closely related peptide homologs. In the deduced 120-amino acid EaF82 peptide sequence from the cloned cDNA, a 30-amino acid signal peptide (Fig. 1a) was predicted by SignalP 5.0[17]. EaF82 contains four cysteines at positions 42, 54, 81 and 95, which potentially form two intramolecular disulfide bridges (Fig. 1a). Its primary structural features were found to be most similar to a group of CRPs classified as "nonfunctional precursor" by Tavormina et al.[8]. The numbers of cysteines and the amino acid patterns around disulfide bonds are known to be conserved for proteins having similar folding and function, and can be used as the basis for protein classification[18]. Thus, these criteria were used for further classification and revealed that the features of EaF82 were closest to RALFs among the listed CRPs, having two predicted intramolecular disulfide bridges and an *N*-terminal signal peptide[19].

The hallmark of RALFs is their ability to rapidly alkalinize tobacco cell culture media upon addition of exogenous peptide[7]. To investigate whether EaF82 also possesses a RALF-like alkalinizing activity, EaF82 peptide without signal peptide was synthesized (designated as EaF82-S) and its activity was measured, including kinetic parameters Km and Vmax. EaF82-S exhibited alkalinizing activity with a Vmax of delta pH ~0.4, which was about half of reported values[7], and Km of 1 nM (Fig. 1b) close to reported values for RALFs[7,10]. The pH increase peaked at 30 min and returned to baseline after 60 min (Fig. 1c). In contrast, the EaF82 alkalinizing activity was abolished when the peptide was reduced and alkylated to break disulfide bonds (a negative control) (Supplementary Fig. 1a). The results indicate that EaF82 has a RALF-like alkalinizing activity.

To examine the evolutionary relationship of EaF82 with RALFs, phylogenetic analysis was performed using the method and sequence information described by Campbell & Turner [5]. We found that EaF82 belonged to clade IV-C (Fig. 1d). Since the model plant Arabidopsis was to be used to heterologously express *EaF82* for functional studies, we further used 13 out of 14 Arabidopsis clade IV-C RALFs from previous publication[5] to align with EaF82 by MUSCLE[20] and analyze their evolutionary relationships using MEGA X[21]. The reported AtRALF17 (AT2G32890)[5] was excluded as it lacks common features of RALF family members and is very likely not a RALF peptide[22]. Phylogenetic analysis revealed that EaF82 was closely related to AtRALF8/9/15 (Fig. 1e), which are known to be abundantly expressed in pollen[13].

**EaF82 is expressed and accumulated in anthers but not in pollen.** To investigate its physiological function, a reporter gene *sGFP*(S65T) in-frame with *EaF82* driven by *EaF82* promoter (*EaF82p::EaF82-sGFP* designated as TA) or a constitutive *CaMV 35S* promoter (*35Sp::EaF82-sGFP* designated as TB) (Supplementary Fig. 2) was transformed into Arabidopsis. The former was used to determine the *EaF82* expression sites while both of them were used to investigate the effects of EaF82 peptide on Arabidopsis growth and development. When the *EaF82* transcriptional and translational expression levels in TA transgenic seedlings were analyzed by RT-PCR and immunoblotting, expected sizes of PCR and protein products were detected (Supplementary Fig. 3a). These validated TA transgenic seedlings showed strong GFP signals in roots, especially at the basal and apical meristems of primary and lateral roots (Supplementary Fig. 3b, c), where auxin is known to be accumulated[23]. The results are consistent

with our previous finding in which GUS was found to accumulate at the same sites when driven by *EaF82* promoter[16]. GFP signals were also observed in anthers and filaments of mature flowers (opened), but only occasionally in released pollen grains (Fig. 2a, b). Similarly, in our previously created *EaF82p::GUS* plants[16], less GUS activity was observed in mature flowers than in early stages of closed flower buds (Fig. 2c), where high auxin accumulation has been reported[24].

EaF82 is closely homologous to AtRALF8/9/15 (Fig. 1e), which are reported to be abundant in pollen[13]. However, the expression of both reporter genes driven by *EaF82* promoter was rarely detected in pollen grains (Fig. 2b, c). Thus, its protein level in pollen was further examined in detail. Seedlings with confirmed EaF82 accumulation from four independent TA lines (Supplementary Fig. 4a) were transferred to soil and grew into mature plants to collect pollen grains. These TA seedlings grew normally on MS medium (Supplementary Fig. 4b) and became mature plants on soil, indicating that EaF82 did not inhibit vegetative growth as reported for AtRALF1/8/23[6]. Collected pollen grains had little detectable EaF82 and GFP proteins as revealed by immunoblotting (Fig. 2d). Indeed, under the confocal microscope only a few pollen grains displayed GFP signals (Fig. 2e). These results indicate that EaF82 was not abundant in pollen. Thus, the observed GFP signals in anthers (Fig. 2a) did not originate from pollen but very likely from the surrounding tissues.

Interestingly, the above TA lines showed many unpollinated pistils and short siliques on the primary inflorescence stalks (Fig. 2f). This observation raised the question of whether EaF82 plays a role in seed-setting since a group of RALFs could affect the rate of seed-setting through regulating pollen tube elongation[10]. We examined pollen viability and pollen tube elongation. There were no differences in the general features of pollen grains compared to those of wild-type (WT) (Fig. 2g). Moreover, there were no obvious differences in pollen germination ability and *in vitro* pollen tube elongation compared to WT (Fig. 2h). Therefore, the effects of EaF82 on pollen viability are likely minimum, which is consistent with the observed result with no substantial EaF82 accumulation in produced pollen grains (Fig. 2d). These results led us to speculate that the unpollinated pistils and short siliques could be the results of high EaF82 accumulation in the surrounding tissues of pollen sacs leading to either no pollen or less pollen available at certain stages. Since the TA lines were driven by auxin responsible *EaF82* promoter[16], it is possible that

the accumulation of EaF82 along the inflorescence stalk was uneven due to the uneven auxin distribution affected by the environmental conditions.

The above speculation that high EaF82 accumulation in the surrounding tissues of pollen sacs might affect pollen development was underpinned by the observations from TB lines with the *EaF82* driven by a strong constitutive *CaMV 35S* promoter. First, although the same floral dipping procedure for creating TA lines was used, our transformation efforts with the TB cassette resulted in only two independent lines (TB-1 and TB-2), despite several attempts. The difficulty of TB transformation seemingly implies that the presence of high EaF82 might cause the developmental problem in transformed pollen and affect the production of transformed seeds. When two TB lines were grown on MS medium, their seedlings were normal as the WT (Supplementary Fig. 5a), but EaF82 and GFP were detected only in TB-2 line by immunoblotting (Supplementary Fig. 5b). TB-2 plants bore few pollen grains and produced only 1-2 small siliques per plant even though they produced many flowers (Fig. 3a, b). TB-2 line exhibited more severe seed abortions than TA lines (Fig. 2f). In order to obtain ~100 siliques, 72 independent plants were planted per subline of TB-2a and TB-2b. The average numbers of seeds per silique in TB-2 plants were reduced to only 4 compared to 49 in the case of WT (Fig. 3c).

Since the major difference between TA and TB lines was the promoters, we further tested promoter activities using multiple independent tobacco (*N. tabacum*) transgenic lines carrying genetic cassettes *35Sp::GUS* or *EaF82p::GUS* (Supplementary Fig. 2) to compare their *GUS* transcript levels. RT-qPCR results showed ~3.5-fold higher *GUS* expressions in transgenic *35Sp::GUS* plants than those of transgenic *EaF82p::GUS* (Supplementary Fig. 6), indicating that the activity of *CaMV 35S* promoter is stronger than that of *EaF82* promoter. Thus, the severe seed abortion phenotype in TB-2 probably was the result of higher expression of EaF82-sGFP than that of TA lines, supporting that the high expression levels of EaF82 may be responsible for the reduction of seed yield.

**Overexpressing EaF82 impaired tapetum degeneration during pollen development in Arabidopsis.** The observed pollenless phenotype in TB-2 (Fig. 3a) implies that the pollen development was compromised. We wondered whether the reduced seed-setting was the results of defective pollen development. Considering the observed accumulation of EaF82 in anthers and closed flower buds (Fig. 2a-c) where pollen is under development, we first investigated at

which stage pollen development was impaired by histological analysis. In Arabidopsis, flower buds are clustered and the flower development can be divided into 12 developmental stages using a series of landmark events as described by Smyth et al.[25]. The anther development can be further divided into 14 stages based on the visualization of distinctive cellular events under microscope[26]. According to the key events of each stage, the pollenless TB-2 line was examined and found that stages of microspore mother cells undergoing meiosis and generating tetrads of haploid microspores, and microspores releasing from the tetrads (up to stage S8) were similar to those in WT (Fig. 3d). The abnormality was noticed at stages S11 and S12. At S11, the disappearance of tapetum occurred in WT but not in TB-2, and hence most of TB-2 pollen grains were aborted at S12 (Fig. 3d). These observations were further confirmed in another TB-2 flower cluster (Supplementary Fig. 7), indicating that overexpressing EaF82 impairs tapetum degeneration leading to defective pollen development.

**Overexpressing EaF82 also affected pollen development and seed-setting in tobacco plants.** To verify that the observed effects of EaF82 on pollen development and seed-setting were not limited to Arabidopsis, a genetic cassette *35Sp::EaF82* (designated as TC) (Supplementary Fig. 2) was expressed in *N. benthamiana* plants. Because tobacco leaf disc transformation method is efficient, this approach was used to generate 15 independent transgenic lines (T1-T15) with seven of them exhibiting a 3 to 1 segregation. Out of these seven lines, four (T2, T3, T8 and T11) had detectable EaF82 in leaves (Fig. 4a) and flowers (Fig. 4b) as revealed by immunoblotting. Mature anthers of these four transgenic lines were often shriveled in appearance with no or less pollen grains while those of the vector control lines were dehisced with many released pollen grains (Fig. 4c). Transgenic flowers were either not fertilized to develop seed pods or partially fertilized to develop small seed pods compared to those normal pods produced from the control lines (Fig. 4d). The seed numbers per pod of lines T2, T3, T8 and T11 were reduced by 42%, 22%, 50% and 58%, respectively, compared to the control lines (Fig. 4e). The results from these transgenic lines further support that EaF82 inhibits pollen development and seed setting in tobacco plants.

**Overexpressing EaF82 reduced the expression of genes involved in tapetum degeneration.** To elucidate the observed impaired tapetum degeneration during pollen development, RNAseq

analysis was performed to gain molecular insights of EaF82 associated transcriptional alterations. To ensure the pollen abortion is associated with overexpressed EaF82 but not sGFP, Arabidopsis transgenic lines with the TC (*35Sp::EaF82*) genetic cassette (Supplementary Fig. 2) were created. Similar to TB, transformation with TC genetic cassette was challenge and produced only two lines (TC-1 and TC-2) with detectable EaF82 levels (Fig. 5a, b). They grew normally during the vegetative stage compared to the WT but exhibited seed abortion (Fig. 5c, d) like TA and TB lines. Their unopened flowers had detectable EaF82 (Fig. 5e) and later could develop some viable pollen grains with normal pollen tube growth (Fig. 5f, g). However, many TC siliques were short in length compared to most WT siliques (Fig. 5h). A large-scale measurement using siliques from 10 primary stems per line showed that approximately 20.5% of TC-1 and 27.8% of TC-2 flowers did not develop siliques longer than 0.3 cm (Fig. 5h), which contained no seed at all (Supplementary Fig. 8). Only 16.6% of TC-1 and 18.8% of TC-2 siliques were longer than 1.0 cm while that of WT was 88.5% (Fig. 5h).

To gain molecular insights into transcriptional changes during pollen development, the gene expression profiles of the unopened flower bud clusters covering all anther developmental stages from TC-1 and TC-2 lines were analyzed by RNAseq along with a vector control line. Two TC lines were used to perform double verification of any observed differentially expressed genes (DEGs). The numbers of reads were between 18.3 to 20.0 million (Supplementary Table 1), and listed genes reported as FPKM (Fragments Per Kilobase Million) were 28,296 (Supplementary Data Set 1), covering ~90% of total nuclear, mitochondria and chloroplast genes (Supplementary Fig. 9a). The numbers of common DEGs in TC-1 and TC-2 either increased or decreased 2-fold with adjusted *p*-value for false discovery rate (FDR) < 0.05 compared to the vector control were 158 and 1197, respectively (Fig. 5i; Supplementary Table 2, 3). Hierarchical cluster analysis of 74 down-regulated DEGs with $\log_2$FC≤-1.5 (reduced ≥ 2.8-fold) plotted in heatmaps with R function showed a strong correlation between TC-1 and TC-2 lines (Supplementary Fig. 9b). The above results demonstrated that overexpressing *EaF82* induced ~5% nuclear gene expression changes ≥ 2-fold in the early development of flowers.

**GO analysis and RT-qPCR validation revealed a large group of down-regulated DEGs involved in cell wall modifications and pH changes.** To determine the affected pathways and related genes in overexpressing *EaF82* lines, identified 158 up- and 1197 down-regulated DEGs

were subjected to Gene Ontology (GO) analysis using PANTHER v16[27], and some selected DEGs were subjected to RT-qPCR validation. In the up-regulated DEGs, only regulation of developmental process (GO:0050793) in biological process and sequence-specific DNA binding (GO:0043565) in molecular function were significantly enriched with a similar group of genes (Supplementary Data Set 2), including flowering-related genes *AGL19, TFL1, AGL20 (SOC1), AGL24, AGL42, FD* and *SAP*[28,29]. All four AGAMOUS-LIKE DEGs *AGL19, AGL20, AGL24* and *AGL42* were selected for validation by RT-qPCR. The results show similar fold increases to those observed by RNAseq analysis (Table 1).

In the down-regulated DEGs, notably enriched GO terms related to the observed characteristics of transgenic plants were those involved in the regulation of pH, and cell wall modification (Fig. 5j). The overrepresented DEGs include genes encoding 29 pectin methylesterases (PMEs)/pectin methylesterase inhibitors (PMEIs), along with six $H^+$-ATPases and 16 cation/$H^+$ antiporters involved in regulating pH changes (Supplementary Data Set 3; Supplementary Table 4). Among them, twenty-eight were selected and validated by RT-qPCR. Their downregulations from RNAseq analysis were well confirmed by RT-qPCR results with similar fold decreases (Table 1, 2). These DEGs are known to play important roles in modulating the cell walls as an adaptation to stresses during plant development[30]. PME activity is regulated by pH changes through coordinating with $H^+$-ATPases and cation/$H^+$ antiporter[31]. Both PMEs and PMEIs are highly expressed in flower buds and anthers[32,33]. Among the down-regulated pollen specific *PMEs* (Supplementary Table 4), *PPME1* and *PME48* are regulated by RGA, a GA repressor DELLA[34]; while *PME5/VGD1*, *PME4/VGDH1* and *VGDH2* are regulated by MYB80, a transcription factor regulating both tapetum and pollen development[35]. These down-regulated DEGs together with reported roles of RGA and MYB80 at the late stage of tapetum degeneration and pollen development[34,35] suggest that cell wall modification was altered during the pollen development, supporting observed impaired tapetum degeneration and pollen abortion (Fig. 3d).

Moreover, seven DEGs (*AGP*5/6/11/14/23/24/40) encoding highly glycosylated arabinogalactan proteins (AGPs), which play key roles in pollen wall formation[36], were also in down-regulated DEGs (Supplementary Table 4). In Arabidopsis, pollen grains from *AGP6* and *AGP11* double mutants have been reported to exhibit pollen wall collapse[37], and *AGP6*, *AGP11*, *AGP23* and *AGP40* are known to be involved in nexine formation[36]. These down-regulated

309 *AGP*s present a strong correlation with the observed aborted pollen at stages S11 and S12 (Fig.
310 3d; Supplementary Fig. 7).
311
312 **DEGs involved in tapetum degeneration.** Pollen is developed inside the anther locules[38]. The
313 innermost layer of anther locules is the tapetum, which is in direct contact with the germ cells
314 and is the main tissue providing nutrition and enzymes for pollen development and pollen wall
315 formation[39,40]. Mounting evidence suggests that the pollen developmental process is tightly
316 linked to the development of tapetum[40,41] with the latter divided into three developmental stages:
317 tapetum differentiation, tapetum formation, and tapetum degeneration through program cell
318 death (PCD)[39]. Tapetal cells appear at the stage S5 and their degradation is initiated at the stage
319 S10[26]. Our histological results of pollen development (Fig. 3d; Supplementary Fig. 7) showed
320 that overexpressing EaF82 interrupted tapetum degeneration at S11, but not tapetum
321 differentiation and formation. This was supported by RNAseq data, which showed that many
322 known genes involved in early stages of tapetum differentiation and tapetum formation[40] were
323 detected but not differentially expressed (Supplementary Table 5), whereas expression of *CEP1*
324 involved in late stage of tapetum degeneration was found to be reduced ~16-fold (Supplementary
325 Tables 4, 5). *CEP1* encodes a papain-like cysteine protease and participates in tapetal cell wall
326 hydrolysis[42]. Previous studies have shown that overexpression *CEP1* advanced the tapetal cell
327 wall degeneration to early S7, while *cep1* mutant lacking functional CEP1 failed tapetum
328 degeneration[42]. Generally, tapetum degeneration accompanies with tapetum cell wall
329 degradation[42]. Therefore, ~16-fold down-regulated *CEP1* together with a large number of cell
330 wall modification associated genes (Supplementary Table 4) supports the observed impairment
331 of tapetal cell degradation (Fig. 3d; Supplementary Fig. 7).
332 In addition to the above genes, *AtRALF*4/8/9/19/25 down-regulated in four tapetum
333 mutants[12], were also found in our down DEGs (Table 2; Supplementary Table 4). Their down-
334 regulation was further confirmed by RT-qPCR showing similar fold reductions in their
335 expressions (Table 2). These *AtRALF*s might be responsible for cell-to-cell communication
336 between tapetal cells and pollen cells, an important process for pollen development. Their real
337 functions however, have not been investigated yet[12].
338

**Overexpressing EaF82 reduced proteasome activity and decreased ATP levels.** The tapetum degeneration is a process resembling apoptosis-like PCD[41], and is essential for pollen development[43]. Pollen abortion has been observed with delayed tapetum PCD[44]. Both proteases and proteasomes are critical to the progression of tapetum PCD[43,45]. Besides *CEP1*, there are seven additional downregulated DEGs encoding proteases (Supplementary Table 4), suggesting that PCD processes may be defective. Although no proteasome genes were found in our DEGs, we examined the translational abundance of three subunits (Rpn6, Rpn10 and Rpt5) of the proteasome by immunoblotting, and the proteasome activity in the early development of flowers. Among them, Rpt5a is one of the six AAA-ATPases of proteasome and essential for pollen development[46]. Immunoblotting showed no differences in their protein abundances (Fig. 6a). However, the proteasome activity in TC-1 and TC-2 lines was reduced by 33% and 49% compared to the WT, or by 48% and 60% compared to the vector control, respectively (Fig. 6b).

The results of reduced proteasome activity prompted us to further examine ATP levels because PCD is an energy dependent process and its initiation and execution are affected by ATP, $Ca^{2+}$ and NO[47,48]. Moreover, the proteasome assembly and activities are also ATP-dependent[49]. With the same tissues used for examining proteasome subunits and proteasome activity, the ATP levels were found to be indeed reduced by 38% and 57% compared to the WT, or 46% and 63% compared to the vector control, respectively (Fig. 6c). Their similar trends of reduction suggested that the reduced proteasome activity was likely the result of low ATP supply at the early developmental stages of flowers even though no DEGs directly linked to ATP production, such as genes coding for ATP synthases, were found. ATP serves not only as an intracellular energy molecule, but also as a signal molecule in the extracellular matrix of plant cells through coordinating with $Ca^{2+}$and ROS[50,51]. ATP defective mutants are male sterile[52,53], while elevated ATP increases seed yields[54]. ATP deficiency in TC lines is also consistent with many of observed down-regulated DEGs, which encode proteins directly or indirectly affected by ATP, such as six ATP-binding cassette (ABC) transporters and six $H^+$-ATPases, as well as a group of ATP-binding receptor-like protein kinases (Table 1; Supplementary Table 4). These results suggest that overexpressing EaF82 lowered ATP levels, and triggered down-regulation of genes encoding for ATP-binding proteins which in turn affected PCD activity and resulted in delaying tapetum degeneration. Additionally, tapetum degeneration is essential for releasing nutrients to support pollen grain development and maturation[39,40]. In the DEG, a large number of

genes encoding transporters and transmembrane proteins for shuttling sugars, amino acids, and ions were also found to be down-regulated (Supplementary Table 4), suggesting that the nutrient transport probably was limited to the developing pollen grains.

**Interacting partners of EaF82.** To understand how overexpressing EaF82 causes a decrease in ATP levels and down-regulation of so many genes, we conducted Y2H screening followed by 1-by-1 interaction validation assay with *EaF82-S* as the "bait" and a cDNA library made from Arabidopsis mitotic flower buds as the "prey" to identify the EaF82 interacting partners. Using the PBS (Predicted Biological Score) system, a total of 46 EaF82 interacting proteins with score A, B, C and D were obtained as good candidates (Supplementary Table 6). Based on their biological functions and cellular components from published studies, seven candidates ABCF4, ALATS, FKBP-like peptidyl-prolyl cis-trans isomerase family protein, PAPP2C, TCH4, AKIN10 and SYTA (Supplementary Table 7) were selected to perform 1-by-1 Y2H validation assay. The assay results were shown in Supplementary Fig. 10 and summarized in Supplementary Table 8. Among the seven candidates, we found three: ABCF4, PAPP2C and AKIN10, to have the strongest interactions (Supplementary Fig. 10; Supplementary Table 8). ABCF4 (AT3G54540, also named as AtGCN4) is an ATPase and regulates plasma membrane $H^+$-ATPase activity[55], PAPP2C (AT1G22280) is a protein phosphatase[56] while AKIN10 (AT3G01090) is a major cellular energy sensor in plants and orthologous to mammalian AMP-activated protein kinase (AMPK)[57].

Given the key role of AKIN10 as the energy sensor in plants and its orthologous to mammalian AMPK, the interaction between AKIN10 and EaF82 (Fig. 7a) could be critical and is likely to be responsible for observed ATP deficiency, and aberrant tapetum degradation and pollen development. ABCF4 could be also involved in the EaF82-induced intracellular pH increase leading to cell wall modification. To elucidate the observed induced ATP deficiency (Fig. 6b), therefore, we focused on AKIN10 and performed co-IP analysis to validate its interaction with EaF82.

An initial attempt to directly detect EaF82 in the pull-downed AKIN10 and EaF82 complex from extracts of TC transgenic floral tissues was unsuccessful despite both AKIN10 and EaF82 were detectable in the total protein extracts. Since the *EaF82-S* without signal peptide was used for alkalinizing activity assay and in Y2H analysis, both the vector control (C) and TC transgenic

line extracts were spiked with EaF82-S peptide and then co-IP performed. The idea was that if AKIN10 is indeed an interacting partner of EaF82, the added EaF82-S peptide should form a complex with endogenous AKIN10, and by co-IP with anti-AKIN10 antibody, the complex should be able to pull-down. Consistent with this idea, we were able to pull-down AKIN10-EaF82-S complex from both the vector control and TC transgenic line extracts (Fig. 7b; lanes 4 and 7). Under the same incubation conditions without the anti-AKIN10 antibody, neither AKIN10 nor EaF82 was detected (Fig. 7b; lanes 3 and 6) while EaF82 was not detected when EaF82-S was incubating with protein A/G magnetic bead suspension (Fig. 7b; lane 8), indicating that no adsorption of AKIN/10 or EaF82-S to protein A/G magnetic beads occurred. The detected EaF82 (Fig. 7b; lane 4 and 7) could only be pull-downed by the anti-AKIN10 bound protein A/G magnetic beads when it was present in the complex with AKIN10. These results indicate that AKIN10 can interact with EaF82, supporting the Y2H finding.

**Elevated levels of AKIN10 in the ATP-deficient *EaF82* transgenic flowers.** In mammals, the level of AMPK is increased when the AMP/ATP ratio is high, resulting in phosphorylation of multiple downstream targets to increase ATP production and decrease ATP consumption[58]. In plants, the level of AMPK homolog AKIN10 is also increased as the AMP/ATP ratio increases, while its gene expression level remains constant[59]. In addition, it has been reported that overexpression of *AKIN10* resulted in late flowering and defective silique development[60,61]. For the identified EaF82 interacting partner AKIN10 (Fig. 7a, b), its gene *AKIN10* and another member *AKIN11* were detected in our RANseq analysis but not differentially expressed (Supplementary Data Set 1). AKIN10 has also been reported to be degraded in a proteasome-dependent manner[62] and induced by low energy conditions such as dark growth and hypoxia[59,63]. To determine whether reduced ATP level and proteasome activity cause elevation of AKIN10 levels, we used immunoblotting with anti-AKIN10 antibody and discovered that TC-1 and TC-2 had an average of ~4-fold higher levels of AKIN10 and ~2-fold higher levels of AKIN11 compared to WT and vector control (Fig. 7c). These results are in agreement with the status of low ATP and reduced proteasome activity in transgenic lines (Fig. 6b, c).

## DISCUSSION

*EaF82* is a novel gene differentially expressed between green and yellow/white sectors of variegated 'Golden Pothos' leaves with its transcript and protein accumulation elevated in color defective sectors in comparison with their green counterparts[16]. *EaF82* contains no intron[16], and encodes a 120 amino acid long peptide with a 30 amino acid signal peptide at the *N*-terminus and four cysteine residues (Fig. 1a), sharing the typical features with numerous plant RALFs[19]. In present study, EaF82 identity as a member of RALF family was verified by its rapid alkalinization capacity (Fig. 1b, c), a hallmark of most RALFs[7], and its overexpression to affect cellular pH regulation and cell wall modification related genes (Table 1; Supplementary Table 4). Phylogenetic analysis indicated that it belongs to clade IV (Fig. 1d), the least characterized group of RALFs. Clade IV members are mainly expressed in reproductive tissues[5], but under certain conditions they may be induced in other tissues. For example, AtRALF8 reported to be abundant in pollen[13] while it was also found to be induced in roots during drought and nematode infection[64]. In 'Golden Pothos', a shy-flowering plant[15], *EaF82* expression correlates with IAA distribution in vegetative parts with no known function[16]. Here, we demonstrated that overexpression of EaF82 in Arabidopsis impaired tapetum degeneration resulting in low pollen production and seed-setting, but appeared not to inhibit growth in any vegetative parts (Fig. 5c), including seedlings (Supplementary Fig. 5a).

In flowering plants, pollen development occurs in anther as a complex process from initial microsporogenesis to the production of mature pollen grains[26,65]. During this development, the sporophytic anther tissues, in particular the tapetum cell layer, play an essential role with the involvement of many genes to regulate developmental process, and provide materials for pollen wall formation[12,65]. Our histological analysis showed that the tapetum degeneration was impaired in TB lines (Fig. 3d; Supplementary Fig. 7). The impairment of tapetum degeneration by EaF82 was further supported by our RNAseq and RT-qPCR results wherein we found that a group of *RALFs* reported in Arabidopsis tapetum mutants[12] were also down-regulated in our transgenic lines. In addition, *CEP1* and seven additional protease genes identified in our down-regulated DEGs are involved in tapetum PCD, suggesting that they might be involved in impairing tapetum degeneration and causing reduction in nutrient supply to timely support pollen wall formation. Moreover, the down-regulation of seven *AGPs* that are important for pollen wall formation[36] and many DEGs involved in nutrient transports (Supplementary Table 4) support our observed pollen formation defect (Fig. 3d; Supplementary Fig. 7). As for pollen development,

462 many genes encoding PME and PMEI for cell wall modification[30] and a group of auxin-
463 responsive genes essential for pollen development[66] were in down-regulated DEGs
464 (Supplementary Table 4). Moreover, 21 transcription factors, including some known to regulate
465 pollen formation and be pollen-specific, were also in down-regulated DEGs (Supplementary
466 Table 4). All of these down-regulated genes lead us to conclude that EaF82 compromised pollen
467 formation via the impairment of tapetum degeneration.

468 In addition, our RNAseq and RT-qPCR data showed that the overexpression of *EaF82*
469 resulted in the repression of seven *AtRALF* genes (*AtRALF4/8/9/15/19/25* and *At4g14020*) (Table
470 2). This intriguing phenomenon could be due to a feedback mechanism for regulating the
471 expression of endogenous *AtRALFs*. Once these seven *AtRALF* propeptides undergo proteolytic
472 processing to become mature peptides, they all have calculated molecular weights between 6.6 to
473 8.7 kD and pIs (isoelectric points) greater than 7 (9.3 - 10.6), indicating that they are basic small
474 peptides (Table 2). Mature EaF82 (EaF82-S) also has a calculated pI of 8.7. It is possible that the
475 overexpressed EaF82 peptide with higher accumulation levels and positive charge might
476 compete with these AtRALFs to bind to their interacting partners and affect regulatory pathways
477 leading to their transcriptional repressions. Nevertheless, the regulatory mechanism of *EaF82* on
478 the down-regulation of these *AtRALFs* needs future study.

479 Concerning the action mode of EaF82, the Y2H assay identified AKIN10 as an interacting
480 partner (Fig. 7a). AKIN10 is a catalytic α-subunit of SnRK1 complex that comprises of an *N*-
481 terminal Ser/Thr kinase domain, an adjacently linked ubiquitin-associated (UBA) domain, and a
482 large *C*-terminal regulatory domain involved in the interaction with the regulatory (β and γ)
483 subunits and upstream phosphatases[63]. *AKIN10* is expressed in flowers, anthers, and pollen[67,68].
484 Under normal conditions, AKIN10 is localized mainly in cytosol[69]. Our co-IP assay validated the
485 interaction between EaF82 and AKIN10 (Fig. 7b). In EaF82 overexpressed lines, it is reasonable
486 to assume that EaF82 binds to AKIN10 and modulates SnRK1 kinase activity and
487 phosphorylation of downstream targets leading to transcriptomic changes. Although we did not
488 examine any phosphorylation changes of proteins downstream of SnRK1, we did observe ~5%
489 transcriptomic changes at the early developmental stages of transgenic flowers overexpressing
490 *EaF82*. Our Y2H results showed that EaF82 interacted with AKIN10 through its *C*-terminal end,
491 which contains kinase associated domain 1 (KA1, position 492-533) (Supplementary Fig. 11).
492 The KA1 is a conserved domain of AMPKs from yeast to humans involved in autoinhibition,

tethering of acidic phospholipids and binding of peptide ligands[70]. Whether EaF82 binds to the KA1 domain to affect SnRK1 activity needs further investigation as the function of KA1 domain in AKIN10 is still unclear[63].

## Materials and Methods

**Construction of genetic cassettes.** Three pBI121 vector-based new genetic cassettes *EaF82p::EaF82-sGFP*, *35Sp::EaF82-sGFP* and *35Sp::EaF82* (designated as TA, TB and TC) were created in this study (Supplementary Fig. 2). The genetic cassette *EaF82p::GUS* was created in previous study[16]. The vector pBI121 carrying the *CaMV35S* promoter (*35Sp*) driving bacterial *uidA* (*GUS*) (*35Sp::GUS*) was used as a control (designated as C). To create these genetic cassettes, the intronless gene, *EaF82*, previously isolated from *E. aureum* containing a 1.2 kb *EaF82* promoter and a 0.8 kb full length *EaF82* including *EaF82* terminator[16] was used. The two previously constructed plasmid DNAs, CEJ826 and CEJ937 (CEJ # is a nomenclature for plasmid DNA in Xie's laboratory), and the pBI121 vector were used to construct these genetic cassettes. The plasmid DNA CEJ826 contains the *EaF82* promoter (*EaF82p*) and full length *EaF82* including *EaF82* terminator in pCR®-blunt 3.5 kb vector (Invitrogen). The plasmid DNA CEJ937 contains *35Sp* and 0.357 kb *EaF82* coding region without TGA stop codon in-frame fused with *sGFP* in pUC19 (Invitrogen). To construct the TC genetic cassette, the *EaF82* fragment isolated from CEJ826 by EcoRI-blunted and EcoRV digestion was ligated to vector pBI121 pre-digested by SacI-blunted and SmaI. The resultant contained *35Sp* driving full length *EaF82* including *EaF82* terminator. To construct the TB genetic cassette, the *35Sp::EaF82-sGFP-NosT* fragment isolated from CEJ937 by HindIII and EcoRI digestion was ligated to pBI121 pre-digested by HindIII and EcoRI. The resultant plasmid DNA was named CEJ964. To construct the TA genetic cassette, the *EaF82* promoter fragment isolated from CEJ826 by XbaI and EcoRV digestion was ligated to pre-digested CEJ964 by HindIII-blunted and XbaI. The resultant construct TA had *EaF82-sGFP* driven by *EaF82p*, whereas TB had the same *EaF82-sGFP* but driven by *35Sp*. The *Agrobacterium* strain LBA4404 was used to harbor these constructs.

**Created transgenic plants.** Transgenic plants were generated with *Arabidopsis thaliana* Col-0, *Nicotiana benthamiana* tobacco using previously described transformation methods[16,71] except

the kanamycin concentration for selection in *N. benthamiana* was 300 mg/L. The Arabidopsis *EaF82p::GUS* transgenic plants were created previously[16]. All experiments were performed using homozygous lines.

**Alkalinization assay.** The alkalinization assay was performed with tobacco (*Nicotiana tabacum*, W38) suspension-cultured cells (Supplementary Fig. 1b) derived from leaf calli following the method described by Pearce et al.[7]. The EaF82-S peptide (Fig. 1a) was synthesized by Lifetein (Somerset, NJ, USA) with a purity of 94.55%. Its molecular weight was 10 kD as confirmed by an Electrospray Ionization (ESI) Mass Spectrometry. Alkylated EaF82-S peptide was prepared as previously reported[7] and used as a negative control. Initial 70 ml of suspension cells ($OD_{600}$=0.2) in liquid medium, pH 5.6 containing MS basal salts and vitamins (Research Product International), were grown in a 300 ml flask under dark at 25°C with shaking at 160 rpm till the cell density reached 5-8 x $10^4$ cluster cells/ml after 3-4 days of subculture. For the assay, each 10 ml cell suspension was aliquoted into a 6-well plate and acclimated with shaking for 2-3 h. The peptide was reconstituted in $ddH_2O$ to give a concentration of 100 µM as a stock solution. Peptide concentrations of 1, 5, 10, 50 and 100 nM were used for assay. The pH was measured by a Mettler Toledo™ InLab 413 pH Electrodes (Mettler-Toledo).

**RT-PCR and RT-qPCR.** The procedures for RNA isolation, RT-PCR and RT-qPCR were the same as described in Hung et al.[14]. Data from three sets of biological samples were averaged. The primers are listed in Supplementary Table 9. The QuantumRNA™ Universal 18S Internal Standard (Invitrogen) was used as an internal control.

**RNAseq and DEG analyses.** The Illumina sequencing was performed by DNA-Link (San Diego, CA, USA). The sequencing reactions were run on the Illumina NextSeq 500 with single-end 76 bp read. The Consensus Assessment of Sequence and Variation (CASAVA) software version 1.8.2 (Illumina) was used to remove adaptor sequences, nucleotide library indexes and generate fastq files. The RNAseq reads were mapped to the *A. thaliana* reference genome TAIR10 by TopHat[72] to produce aligned reads and FPKM[73]. The gene annotations were based on NCBI database. The DEG analysis was conducted by Cufflinks and Cuffdiff[72]. The PANTHER classification system was used for functional pathway analysis[27].

**Protein isolation, SDS-PAGE and immunoblotting.** For extracting total proteins from leaves and seedlings, Plant Total Protein Extraction Kit (Sigma-Aldrich) was used. For extracting pollen proteins, the method was adopted from Chang & Huang[74]. In brief, hundreds of opened flowers were harvested in an ice-cold tube. About 5x volume of cold extraction buffer (HEPES potassium solution) was added. After brief mixing, the mixture was centrifuged at 350 *g* for 1 min to collect pollen. The purity of collected pollen was examined under a microscope (Supplementary Fig. 1c) before disruption by a blue pestle on ice. After 1 h incubation on ice, extracted pollen proteins in the supernatant were collected following centrifugation at 18,000 *g* for 30 min at 4°C. For extracting proteins from unopened flower clusters, tissues ground in liquid nitrogen were mixed in 1:2.5 ratio with extraction buffer containing 50 mM HEPES pH 7.8, 2 mM EDTA, 1 mM DTT and 1x Halt™ Protease Inhibitor Cocktail, EDTA-free (Thermo Scientific). After centrifugation at 18,000 *g* for 20 min at 4°C, the protein extracts were collected.

SDS-PAGE, immunoblotting and the subsequent detection of chemiluminescent signals along with staining of blots were performed as previously described[16]. To detect EaF82, custom-made anti-EaF82 antibody was used[16]. To detect GFP, the blots were probed with 1:200 diluted mouse monoclonal anti-GFP (sc-9996, Santa Cruz Biotechnology). After three washes with TBST, blots were incubated with 1:10,000 diluted HRP-conjugated anti-mouse IgGκ (sc-516102, Santa Cruz Biotechnology). For detecting AKIN10, the blots were probed with 1:500 diluted anti-AKIN10 (AS10919, Agrisera) in PBST containing 1% dry-milk, followed by the 1:20,000 diluted HRP-conjugated anti-rabbit IgG (H+L) (AS014, ABclonal). For quantification of band intensity, the same blot was probed with 1:15,000 diluted plant actin mouse mAb (AC009, ABclonal) as an internal control, followed by the 1:20,000 diluted HRP-conjugated anti-mouse IgG (Jackson ImmunoResearch). Three independent experiments were performed and scanned images of band intensities on X-ray films were quantified using Image J (https://imagej.nih.gov/ij/). For examining the proteasome subunits, the same procedure used for detecting AKIN10 was also used except for the incubation conditions for each primary antibody. Anti-RPN6 (AS15 2832A, Agrisera) was 1:1,000 diluted in TBST containing 3% dry-milk and incubated at 23°C for 1 h. Anti-RPN10 (PHY0102S, PhytoAB) was 1:2,000 diluted in TBST and

incubated at 23°C for 1 h. Anti-Rpt5a/b (PHY1747A, PhytoAB) was 1:1,000 diluted in PBST and incubated at 4°C for 16 h.

To determine the size of EaF82, the Novex™ tricine gel system (Invitrogen) for revealing low molecular weight proteins was used. In brief, extracted proteins were first mixed with equal volume of tricine SDS sample buffer containing 50 mM DTT. Before loading onto 16% tricine gel, the sample mixtures were denatured at 85°C for 2 min. The Spectra™ Multicolor Low Range Protein Ladder, a mixture of six proteins ranging from 1.7 to 40 kD, was used as size standards. Proteins were separated in tricine SDS running buffer under constant 50 volts for 4 h and later increasing to 100 volts for 1 h. They were then transferred to 0.2 μm PVDF membrane in Novex Tris-Glycine transfer buffer with 20% methanol under constant 20 volts for 90 min. Immunoblotting was performed as described for detecting EaF82.

**Histological analysis and seed counting**

GUS assay was performed as previously described[16]. To detect GFP, tissues were observed directly under a fluorescence stereo microscope Nikon SMZ1000 (Nikon) equipped with ET-Narrow Band EGFP to minimize autofluorescence ex 480 nm/20 and em 510 nm/20 (49020, Chroma Technology). The images were taken by Nikon Digital Sight DS-Fi1 and analyzed by software NIS-ELEMENTS BR 3.0 (Nikon). To observe pollen, pollen grains were released by gently crushing the anthers on slides. For observing pollen germination, the method described by Krishnakumar & Oppenheimer[75] was used to prepare slides containing germinated pollen grains. The pollen images were observed under a Zeiss LSM 710 microscope with ZEN software (Zeiss). For pollen viability assay, the stamens were tapped on a drop of iodine/potassium iodide TS 1 solution (RICCA chemical company) placed on a glass slide. After 5 min in dark, the pollen grains were imaged under a Keyence BZ-X700 microscope (Keyence). To observe pollen development, the whole unopened flower cluster was fixed in FAA solution (4% [v/v] formaldehyde, 5% [v/v] acetic acid and 50% [v/v] ethanol) with gentle vacuum for 5 min then kept at 4°C for 24 h. The dehydration, paraffin embedding, and sectioning steps were the same as described by Hung et al.[15]. After immobilizing on slides, the deparaffinized specimens were stained for 2-3 min with fresh 0.05% (w/v) Toluidine blue O solution in 0.1 M citrate phosphate buffer, pH 6.8 and then rinsed in running water. After mounting on Fluoromount (Sigma-

Aldrich), the images were taken by a Keyence BZ-X700. For counting seeds, the images of seeds were counted by VisionWork®LS software (UVP).

**ATP measurement.** The method of ATP measurement was adopted from Napolitano & Shain[76]. In brief, ATP was extracted by mixing 1 mg of tissue powder with 40 µl of 50 mM HEPES buffer, pH 7.4 containing 33 mAU/ml Novagen® proteinase K (Millipore). The mixture was incubated at 50°C for 15 min then at 80°C for 5 min. After centrifugation at 20,000 *g* for 10 min at 4°C, the supernatant was collected and used for ATP assay using an adenosine 5-triphosphate (ATP) bioluminescent assay kit (Sigma-Aldrich). The generated bioluminescence signal was measured with a SpectraMax M5 plate reader (Molecular Devices).

**Proteasomal activity assay.** Proteasomal activity was assayed as described by Vallentine et al.[77] except that 10 mM ATP and 5% (v/v) glycerol were included in the extraction buffer but omitted in the reaction buffer. Each assay had equal amounts of extracted proteins (3 µg). The generated signal was measured using a microplate reader PHERAstar (BMG Labtech).

**Yeast two-hybrid.** The Y2H screening and 1-by-1 direct interaction assays were conducted by Hybrigenics Services SAS (Paris, France). The coding sequence for EaF82 (amino acid residues 30-120) was PCR-amplified from previously constructed CEJ982 containing *EaF82*, and cloned in frame with the LexA DNA binding domain (DBD) into pB27 vector as a *C*-terminal fusion to LexA (N-LexA-EaF82-C). Hybrigenics' reference for this construct was hgx4998v1_pB27. The entire insert sequence in the construct was confirmed by sequencing and then used as a bait to screen a random-primed *Arabidopsis thaliana* meiotic buds cDNA library constructed into pP6 vector. Cloning vectors pB27 and pP6 were derived from the original pBTM116[78] and pGADGH[79] plasmids, respectively.

The N-*LexA-EaF82*-C bait was tested in yeast and found neither toxic nor autoactivating by itself. Therefore, it was used for the ULTImate Y2H™ screening. A total of 115 million clones (11-fold the complexity of the library) were screened using a mating approach with YHGX13 (Y187 ade2-101::loxP-kanMX-loxP, mata) and L40ΔGal4 (mata) yeast strains as previously described[80]. A total of 271 His+ colonies were selected on a medium lacking tryptophan, leucine and histidine. The prey fragments of positive clones were amplified by PCR and sequenced at

their 5' and 3' junctions. The obtained sequences were subjected to search the corresponding interacting protein in the GenBank database (NCBI) using a fully automated procedure. A confidence PBS score, which relies on both local and global scores, was assigned to each interaction as described by Formstecher et al.[81]. Briefly, the local score was analyzed by considering the redundancy and independency of prey fragments, as well as the distribution of reading frames and stop codons in overlapping fragments. Then, the global score was analyzed by taking into consideration of the interactions that were found in all the screens performed at Hybrigenics using the same library. The global score indicates the probability of a nonspecific interaction. The PBS scores were divided into six categories (A to E) for practical use purpose. Category A represents the highest confidence while D stands for the lowest confidence. Moreover, the category E specifically flags interactions involving highly connected prey domains previously found several times in screens performed on libraries derived from the same organism, while F stands for a false-positive with several of these highly connected domains that have been confirmed as false-positives of the technique. The PBS scores have been reported to correlate with the biological significance of interactions well[82,83]. Interacting proteins with PBS scores A, B, C and D all have been confirmed to have biological relevance independently[82]. Therefore, all interacting proteins with PBS scores A, B, C and D could be considered as good candidates.

Using PBS score system, seven candidates were selected from the obtained EaF82 interacting proteins with score A, B, C and D (marked as P in Supplementary Table 7) to further perform 1-by-1 Y2H assay for validation. This direct 1-by-1 interaction assay was performed. Seven fragments were extracted from the ULTImate Y2H™ screening and cloned in frame with the Gal4 activation domain (AD) into plasmid pP7. The AD construct was checked by sequencing the 5' and 3' ends of the inserts. Hybrigenics' references for these seven preys are listed in Supplementary Table 7.

To perform 1-by-1 pairwise Y2H interaction assays, the bait and prey constructs were transformed into the yeast haploid cells L40ΔGal4 (mata) and YHGX13 (Y187 ade2-101::loxP-kanMX-loxP, matα), respectively. These assays were based on the HIS3 reporter gene (growth assay without histidine). As negative controls, the bait plasmid was tested in the presence of empty prey vector (pP7) and all prey plasmids were tested with the empty bait vector (pB27). The interaction between SMAD and SMURF was used as the positive control[84]. Interaction pairs

were tested in duplicate as two independent clones (clone 1 and clone 2) for the growth assay. For each interaction, undiluted and $10^{-1}$, $10^{-2}$, $10^{-3}$ dilutions of the diploid yeast cells (culture normalized at $5\times10^{7}$ cells/ml) expressing both bait and prey constructs were spotted on several selective media. The DO-2 selective medium lacking tryptophan and leucine was used as a growth control to verify the presence of both the bait and prey plasmids. The different dilutions were also spotted on a selective medium without tryptophan, leucine and histidine (DO-3). Four different concentrations (1, 5, 10 and 50 mM) of 3-AT, an inhibitor of the HIS3 gene product, were added to the DO-3 plates to increase stringency and reduce possible autoactivation by the bait and prey constructs. The "DomSight" (Hybrigenics Services, SAS) displaying the comparison of the bait fragment and the Selected Interacting Domain (SID) of the prey proteins with the functional and structural domains (databases of protein domains: PFAM, SMART, TMHMM, SignalP, Coil algorithms) on these proteins were used for data visualization.

**Co-immunoprecipitation (co-IP) analysis**. To validate the interaction between EaF82 and AKIN10, co-IP analysis was performed. Each 333 mg ground floral tissues from Arabidopsis vector control or TC transgenic line was resuspended in 1 ml of extraction buffer (50 mM Tris-HCl, pH 7.5 containing 150 mM NaCl, 0.05% Triton-X100, 10% glycerol, and both protease and phosphatase inhibitor cocktail) and kept on ice for 30 min for soaking. The samples were then homogenized using a BeadBug 6 microtube homogenizer (Benchmark Scientific) and centrifuged at 15,000 *g* for 10 min. The clear extracts were transferred to new tubes and centrifuged again to remove particulates. The clear extracts were transferred to new tubes.

To prepare the input samples, each 100 µl clear extracts was mixed with 37 µl of 4X LDS sample buffer (NuPAGE™, Invitrogen, USA) and 15 µl 10X reducing agent. The mixture was then heated at 70˚C for 10 min and stored at -20˚C until analysis. The remaining extracts were used for co-IP. First, both control and TC extracts were split into two 250 µl fractions. They then were spiked with or without 20 µg of EaF82-S peptide and divided into two tubes. After that, one tube was added 10 µg of anti-AKIN10 antibody (Cedarlane, USA) while the other tube received none (as a negative control). Another control was prepared by adding EaF82-S peptide alone to protein A+G magnetic beads. All samples were gently shaken overnight at 4˚C for binding. After that, 50 µl of protein A/G magnetic beads pre-equilibrated with extraction buffer were added and incubated further for 3 h. After binding, all samples were centrifuged at 800 *g* for 1 min to

separate beads. The supernatant was removed and the beads were first washed thrice with 400 µl of extraction buffer, followed by two washes with 0.1M Tris-HCl buffer, pH 7.5. The bound complex was released from beads by adding 150 µl of preheated (70˚C) 4X LDS sample buffer and centrifuged at 10,000 *g* for 5 min. The eluates were then transferred to new tubes, mixed with 15 µl of 10X reducing agent (NuPAGE™, Invitrogen), and heated at 70˚C for 10 min. Immunoblotting was then performed to detect EaF82 and AKIN10 in the eluate. Briefly, 20 µl of input and 30 µl of co-IP samples were loaded onto a 12% NuPAGE™ Bis-Tris gel (Invitrogen) and separated at 200 V for 40 min. Following separation, the proteins were transferred onto a PVDF membrane overnight at 4˚C. The membrane was then blocked with fat-free milk powder in PBST (1%) for 1 h at room temperature, and then was first incubated with rabbit anti-AKIN10 antibody (1:500 dilution) followed by incubation with Veriblot IP detection reagent (Abcam, 1:500 in blocking solution). Protein bands were detected with Supersignal™ west pico chemiluminescent substrate (Thermofisher) and images were captured with the iBright™ CL1500 imaging system (Thermofisher). To detect EaF82, the above blot was first stripped with Restore™ stripping buffer (Thermofisher), washed thrice for 10 min each with 1X PBST, and then blocked as described above. The remainder of the procedure to detect EaF82 was the same as described earlier.

**Peptide sequence analysis and phylogenetic tree construction**. The SignalP 5.0 (http://www.cbs.dtu.dk/services/SignalP/)[17] was used to predict the signal peptide. The Swiss Institute of Bionformatics Expasy server (https://web.expasy.org/compute_pi/)[85] was used to predict pI and molecular weight. The construction of two phylogenetic trees was carried out following the information and procedures provided in Campbell & Turner[5] except skipping a manual optimization; and MEGA X[21] was used to perform evolutionary analysis of Arabidopsis clade IV-C RALFs with EaF82.

**Statistical analysis.** Statistical analysis was performed with one-way ANOVA and Fisher's least significant difference (LSD) test. The GraphPad Prism 7 (GraphPad Software) was used for calculating Michaelis–Menten constant (Km) and Vmax as well as statistic best fit value of R square and standard deviation of estimation (Sy.x).

**Data availability.** Arabidopsis Genome Initiative locus identifiers for each gene mentioned in this study are listed in Tables 1 and 2; Supplementary Tables 2, 3, and 4; Supplementary Data Sets 2 and 3. The accession number of *EaF82* is FJ666044. Raw data obtained from RNAseq analysis have been deposited into NCBI's Gene Expression Omnibus under the accession codes GSE171459 (https://www.ncbi.nlm.nih.gov/geo/).

## References


1. Pearce, G., Strydom, D., Johnson, S. & Ryan, C. A. A polypeptide from tomato leaves induces wound-inducible proteinase inhibitor proteins. *Science* **253**, 895-7 (1991).
2. Olsson, V. et al. Look closely, the beautiful may be small: precursor-derived peptides in plants. *Annu. Rev. Plant Biol.* **70**, 153-186 (2019).
3. Murphy, E., Smith, S. & De Smet, I. Small signaling peptides in Arabidopsis development: how cells communicate over a short distance. *Plant Cell* **24**, 3198-3217 (2012).
4. Sharma, A. et al. Comprehensive analysis of plant rapid alkalization factor (RALF) genes. *Plant Physiol. Biochem.* **106**, 82-90 (2016).
5. Campbell, L. & Turner, S. R. A comprehensive analysis of RALF proteins in green plants suggests there are two distinct functional groups. *Front. Plant Sci.* **8**, 37 (2017).
6. Blackburn, M. R., Haruta, M. & Moura, D. S. Twenty years of progress in physiological and biochemical investigation of RALF peptides. *Plant Physiol.* **182**, 1657-1666 (2020).
7. Pearce, G., Moura, D. S., Stratmann, J. & Ryan, C. A. RALF, a 5-kDa ubiquitous polypeptide in plants, arrests root growth and development. *Proc. Natl Acad. Sci. USA* **98**, 12843-12847 (2001).
8. Tavormina, P., De Coninck, B., Nikonorova, N., De Smet, I. & Cammue, B. P. The plant peptidome: an expanding repertoire of structural features and biological functions. *Plant Cell* **27**, 2095-2118 (2015).
9. Haruta, M. et al. Molecular characterization of mutant Arabidopsis plants with reduced plasma membrane proton pump activity. *J. Biol. Chem.* **285**, 17918-17929 (2010).
10. Covey, P. A. et al. A pollen-specific RALF from tomato that regulates pollen tube elongation. *Plant Physiol.* **153**, 703-715 (2010).

11. Shi, F., Zhou, X., Liu, Z. & Feng, H. Rapid alkalinization factor (*RALF*) genes are related to genic male sterility in Chinese cabbage (*Brassica rapa* L. ssp. *pekinensis*). *Sci. Hortic.* **225**, 480-489 (2017).
12. Li, D. D., Xue, J. S., Zhu, J. & Yang, Z. N. Gene regulatory network for tapetum development in *Arabidopsis thaliana*. *Front. Plant Sci.* **8**, 1559 (2017).
13. Loraine, A. E., McCormick, S., Estrada, A., Patel, K. & Qin, P. RNA-seq of Arabidopsis pollen uncovers novel transcription and alternative splicing. *Plant Physiol.* **162**, 1092-1109 (2013).
14. Hung, C. Y. et al. Identification of a Mg-protoporphyrin IX monomethyl ester cyclase homologue, EaZIP, differentially expressed in variegated *Epipremnum aureum* 'Golden Pothos' is achieved through a unique method of comparative study using tissue regenerated plants. *J. Exp. Bot.* **61**, 1483-1493 (2010).
15. Hung, C. Y. et al. Gibberellin deficiency is responsible for shy-flowering nature of *Epipremnum aureum*. *Sci. Rep.* **6**, 1-11 (2016).
16. Hung, C. Y. et al. Differential expression of a novel gene *EaF82a* in green and yellow sectors of variegated *Epipremnum aureum* leaves is related to uneven distribution of auxin. *Physiol. Plant.* **152**, 749-762 (2014).
17. Armenteros, J. J. et al. SignalP 5.0 improves signal peptide predictions using deep neural networks. *Nat. Biotechnol.* **37**, 420-423 (2019).
18. Marques, J. R., Fonseca, R. R. D., Drury, B. & Melo, A. Amino acid patterns around disulfide bonds. *Int. J. Mol. Sci.* **11**, 4673-4686 (2010).
19. Marshall, E., Costa, L. M. & Gutierrez-Marcos, J. Cysteine-rich peptides (CRPs) mediate diverse aspects of cell–cell communication in plant reproduction and development. *J. Exp. Bot.* **62**, 1677-1686 (2011).
20. Edgar, R. C. MUSCLE: a multiple sequence alignment method with reduced time and space complexity. *BMC Bioinform.* **5**, 1-19 (2004).
21. Kumar, S., Stecher, G., Li, M., Knyaz, C. & Tamura, K. MEGA X: molecular evolutionary genetics analysis across computing platforms. *Mol. Biol. Evol.* **35**, 1547-1549 (2018).
22. Abarca, A., Franck, C. M. & Zipfel, C. Family-wide evaluation of RALF peptides in *Arabidopsis thaliana*. *Plant Physiol.* **187**, 996-1010 (2021).

800 23. De Smet, I. & Jürgens, G. Patterning the axis in plants–auxin in control. *Curr. Opin.*
801 *Genet. Dev.* **17**, 337-343 (2007).
802 24. Yao, X. et al. Auxin production in diploid microsporocytes is necessary and sufficient for
803 early stages of pollen development. *PLoS Genet.* **14**, e1007397 (2018).
804 25. Smyth, D. R., Bowman, J. L. & Meyerowitz, E. M. Early flower development in
805 Arabidopsis. *Plant Cell* **2**, 755-767 (1990).
806 26. Sanders, P. M. et al. Anther developmental defects in *Arabidopsis thaliana* male-sterile
807 mutants. Sex. *Plant Reprod.* **11**, 297-322 (1999).
808 27. Mi, H. er al. PANTHER version 16: a revised family classification, tree-based
809 classification tool, enhancer regions and extensive API. *Nucleic Acids Res.* **49**, D394-
810 D403 (2021).
811 28. Parenicová, L. et al. Molecular and phylogenetic analyses of the complete MADS-box
812 transcription factor family in Arabidopsis: new openings to the MADS world. *Plant Cell*
813 **15**, 1538-1551 (2003).
814 29. Lee, J. & Lee, I. Regulation and function of SOC1, a flowering pathway integrator. *J.*
815 *Exp. Bot.* **61**, 2247-2254 (2010).
816 30. Barnes, W. J. & Anderson, C. T. Release, recycle, rebuild: cell-wall remodeling,
817 autodegradation, and sugar salvage for new wall biosynthesis during plant development.
818 *Mol. Plant* **11**, 31-46 (2018).
819 31. Pelloux, J., Rusterucci, C. & Mellerowicz, E. J. New insights into pectin methylesterase
820 structure and function. *Trends Plant Sci.* **12**, 267-277 (2007).
821 32. Pina, C., Pinto, F., Feijó, J. A. & Becker, J. D. Gene family analysis of the Arabidopsis
822 pollen transcriptome reveals biological implications for cell growth, division control, and
823 gene expression regulation. *Plant Physiol.* **138**, 744-756 (2005).
824 33. Pinzón-Latorre, D. & Deyholos, M. K. Characterization and transcript profiling of the
825 pectin methylesterase (PME) and pectin methylesterase inhibitor (PMEI) gene families in
826 flax (*Linum usitatissimum*). *BMC Genom.* **14**, 742 (2013).
827 34. Tian, G. W., Chen, M. H., Zaltsman, A. & Citovsky, V. Pollen-specific pectin
828 methylesterase involved in pollen tube growth. *Dev. Biol.* **294**, 83-91 (2006).

35. Phan, H. A., Iacuone, S., Li, S. F. & Parish, R. W. The MYB80 transcription factor is required for pollen development and the regulation of tapetal programmed cell death in *Arabidopsis thaliana*. *Plant Cell* **23**, 2209-2224 (2011).
36. Jia, Q. S. et al. Arabidopsis AT-hook protein TEK positively regulates the expression of arabinogalactan proteins for Nexine formation. *Mol. Plant* **8**, 251-260 (2015).
37. Coimbra, S., Costa, M., Jones, B., Mendes, M. A. & Pereira, L. G. Pollen grain development is compromised in Arabidopsis agp6 agp11 null mutants. *J. Exp. Bot.* **60**, 3133-3142 (2009).
38. Goldberg, R. B., Beals, T. P. & Sanders, P. M. Anther development: basic principles and practical applications. *Plant Cell* **5**, 1217 (1993).
39. Scott, R. J., Spielman, M. & Dickinson, H. G. Stamen structure and function. *Plant Cell* **16**, S46-S60 (2004).
40. Lei, X. & Liu, B. Tapetum-dependent male meiosis progression in plants: Increasing evidence emerges. *Front. Plant Sci.* **10**, 1667 (2019).
41. Parish, R. W. & Li, S. F. Death of a tapetum: a programme of developmental altruism. *Plant Sci.* **178**, 73-89 (2010).
42. Zhang, D. et al. The cysteine protease CEP1, a key executor involved in tapetal programmed cell death, regulates pollen development in Arabidopsis. *Plant Cell* **26**, 2939-2961 (2014).
43. Zhang, D. S. et al. Tapetum degeneration retardation is critical for aliphatic metabolism and gene regulation during rice pollen development. *Mol. Plant* **1**, 599-610 (2008).
44. Kawanabe, T., Ariizumi, T., Kawai-Yamada, M., Uchimiya, H. & Toriyama, K. Abolition of the tapetum suicide program ruins microsporogenesis. *Plant Cell Physiol.* **47**, 784-787 (2006).
45. Buono, R. A., Hudecek, R. & Nowack, M. K. Plant proteases during developmental programmed cell death. *J. Exp. Bot.* **70**, 2097-2112 (2019).
46. Gallois, J. L. et al. The Arabidopsis proteasome RPT5 subunits are essential for gametophyte development and show accession-dependent redundancy. *Plant Cell* **21**, 442-459 (2009).
47. Sun, J. et al. An ATP signalling pathway in plant cells: extracellular ATP triggers programmed cell death in *Populus euphratica*. *Plant Cell Environ.* **35**, 893-916 (2012).

48. Van Aken, O. & Van Breusegem, F. Licensed to kill: mitochondria, chloroplasts, and cell death. *Trends Plant Sci.* **20**, 754-766 (2015).
49. Marshall, R. S. & Vierstra, R. D. Dynamic regulation of the 26S proteasome: from synthesis to degradation. *Front. Mol. Biosci.* **6**, 40 (2019).
50. Brookes, P. S., Yoon, Y., Robotham, J. L., Anders, M. W. & Sheu, S. S. Calcium, ATP, and ROS: a mitochondrial love-hate triangle. *Am. J. Physiol. Cell Physiol.* **287**, C817-C833 (2004).
51. Tanaka, K., Gilroy, S., Jones, A. M. & Stacey, G. Extracellular ATP signaling in plants. *Trends Cell Biol.* **20**, 601-608 (2010).
52. Berthomé, R. et al. pur4 mutations are lethal to the male, but not the female, gametophyte and affect sporophyte development in Arabidopsis. *Plant Physiol.* **147**, 650-660 (2008).
53. Li, W. Q., Zhang, X. Q., Xia, C., Deng, Y. & Ye, D. MALE GAMETOPHYTE DEFECTIVE 1, encoding the FAd subunit of mitochondrial F1F0-ATP synthase, is essential for pollen formation in *Arabidopsis thaliana*. *Plant Cell Physiol.* **51**, 923-935 (2010).
54. Liang, C. et al. Impacts of high ATP supply from chloroplasts and mitochondria on the leaf metabolism of *Arabidopsis thaliana*. *Front. Plant Sci.* **6**, 922 (2015).
55. Kaundal, A. et al. GENERAL CONTROL NONREPRESSIBLE4 degrades 14-3-3 and the RIN4 complex to regulate stomatal aperture with implications on nonhost disease resistance and drought tolerance. *Plant Cell* **29**, 2233-2248 (2017).
56. Phee, B. K. et al. A novel protein phosphatase indirectly regulates phytochrome-interacting factor 3 via phytochrome. *Biochem. J.* **415**, 247-255 (2008).
57. Nukarinen, E. et al. Quantitative phosphoproteomics reveals the role of the AMPK plant ortholog SnRK1 as a metabolic master regulator under energy deprivation. *Sci. Rep.* **6**, 31697 (2016).
58. Hardie, D. G., Schaffer, B. E. & Brunet, A. AMPK: an energy-sensing pathway with multiple inputs and outputs. *Trends Cell Biol.* **26**, 190-201 (2016).
59. Cho, H. Y., Wen, T. N., Wang, Y. T. & Shih, M. C. Quantitative phosphoproteomics of protein kinase SnRK1 regulated protein phosphorylation in Arabidopsis under submergence. *J. Exp. Bot.* **67**, 2745-2760 (2016).

60. Tsai, A. Y. L. & Gazzarrini, S. AKIN10 and FUSCA3 interact to control lateral organ development and phase transitions in Arabidopsis. *Plant J.* **69**, 809-821 (2012).
61. Jeong, E. Y., Seo, P. J., Woo, J. C. & Park, C. M. AKIN10 delays flowering by inactivating IDD8 transcription factor through protein phosphorylation in Arabidopsis. *BMC Plant Bio.* **15**, 1-3 (2015).
62. Farrás, R. et al. SKP1–SnRK protein kinase interactions mediate proteasomal binding of a plant SCF ubiquitin ligase. *EMBO J.* **20**, 2742-56 (2001).
63. Broeckx, T., Hulsmans, S. & Rolland, F. The plant energy sensor: evolutionary conservation and divergence of SnRK1 structure, regulation, and function. *J. Exp. Bot.* **67**, 6215-6252 (2016).
64. Atkinson, N. J., Lilley, C. J. & Urwin, P. E. Identification of genes involved in the response of Arabidopsis to simultaneous biotic and abiotic stresses. *Plant Physiol.* **162**, 2028-2041 (2013).
65. Gómez, J. F., Talle, B. & Wilson, Z. A. Anther and pollen development: a conserved developmental pathway. *J. Integr. Plant Biol.* **57**, 876-891 (2015).
66. Cardarelli, M. & Cecchetti, V. Auxin polar transport in stamen formation and development: how many actors? *Front. Plant Sci.* **5**, 333 (2014).
67. Bitrián, M., Roodbarkelari, F., Horváth, M. & Koncz, C. BAC-recombineering for studying plant gene regulation: developmental control and cellular localization of SnRK1 kinase subunits. *Plant J.* **65**, 829-842 (2011).
68. Williams, S. P., Rangarajan, P., Donahue, J. L., Hess, J. E. & Gillaspy, G. E. Regulation of sucrose non-fermenting related kinase 1 genes in *Arabidopsis thaliana*. *Front. Plant Sci.* **5**, 324 (2014).
69. Ramon, M. et al. Default activation and nuclear translocation of the plant cellular energy sensor SnRK1 regulate metabolic stress responses and development. *Plant Cell* **31**,1614-1632 (2019).
70. Emptage, R. P., Lemmon, M. A. & Ferguson, K. M. Molecular determinants of KA1 domain-mediated autoinhibition and phospholipid activation of MARK1 kinase. *Biochem. J.* **474**, 385-398 (2017).

71. Musa, T. A., Hung, C. Y., Darlington, D. E., Sane, D. C. & Xie, J. Overexpression of human erythropoietin in tobacco does not affect plant fertility or morphology. *Plant Biotechnol. Rep.* **3**, 157-165 (2009).
72. Trapnell, C. et al. Differential gene and transcript expression analysis of RNA-seq experiments with TopHat and Cufflinks. *Nat. Protoc.* **7**, 562-578 (2012).
73. Kim, D. et al. TopHat2: accurate alignment of transcriptomes in the presence of insertions, deletions and gene fusions. *Genome Biol.* **14**, R36 (2013).
74. Chang, M. & Huang, S. Arabidopsis ACT 11 modifies actin turnover to promote pollen germination and maintain the normal rate of tube growth. *Plant J.* **83**, 515-527 (2015).
75. Krishnakumar, S. & Oppenheimer, D. G. Extragenic suppressors of the Arabidopsis zwi-3 mutation identify new genes that function in trichome branch formation and pollen tube growth. *Development* **126**, 3079-3088 (1999).
76. Napolitano, M. J. & Shain, D. H. Distinctions in adenylate metabolism among organisms inhabiting temperature extremes. *Extremophiles* **9**, 93-98 (2005).
77. Vallentine, P., Hung, C. Y., Xie, J. & Van Hoewyk, D. The ubiquitin–proteasome pathway protects *Chlamydomonas reinhardtii* against selenite toxicity, but is impaired as reactive oxygen species accumulate. *AoB Plants* **6**, plu062 (2014).
78. Vojtek, A. & Hollenberg, S. M. Ras-Raf interaction: two-hybrid analysis. *Methods Enzymol.* **255**, 331-342 (1995).
79. Bartel, P. L., Chein, C. T., Sternglanz, R. & Fields, S. in Cellular interactions in development: A practical approach. ed. Hartley, D.A. (Oxford University Press, Oxford) pp. 153-179 (1993).
80. Fromont-Racine M, Rain JC, Legrain P. 1997. Toward a functional analysis of the yeast genome through exhaustive two-hybrid screens. *Nat. Genet.* **16**, 277-282.
81. Formstecher, E. et al. Protein interaction mapping: a Drosophila case study. *Genome Res.* **15**, 376-384 (2005).
82. Rain, J. C. et al. The protein–protein interaction map of Helicobacter pylori. *Nature* **409**, 211-215 (2001).
83. Wojcik, J., Boneca, I. G. & Legrain, P. Prediction, assessment and validation of protein interaction maps in bacteria. *J. Mol. Biol.* **323**, 763-770 (2002).

84. Colland, F. et al. 2004. Functional proteomics mapping of a human signaling pathway. *Genome Res* **14**, 1324-32 (2004).

85. Gasteiger, E. et al. Protein identification and analysis tools on the ExPASy server. In The proteomics protocols handbook pp. 571-607. (Humana press, 2005).

**Acknowledgements**

We thank Qingping He and Maotao He for their assistance on histological analysis, Eva Johannes for the help on confocal microscopy, and Mythili Saravanan for the help on statistical analysis.

**Authors' contributions**

C.Y.H. and J.X. conceived and designed the experiments. C.Y.H., K.N.W., M.L.U., J.C., D.B.B., M.T., Q.Q., and J.Z. performed the experiments. C.Y.H., F.S.K., C.E.O., J.C., K.O.B. and J.X. analyzed the data. C.Y.H., F.S.K., and J.X. wrote the article with contributions of all the authors.

**Funding**

This study was supported by the National Science Foundation grant (HRD-1400946 to J.X. and C.E.O.), National Institute of General Medical Sciences grant (SC1GM111178 to J. X.) and a Startup Fund of Golden LEAF Foundation to BRITE.

**Competing interest statement**

The authors declare no competing financial interests.

**Materials & Correspondence**

Correspondence and requests for materials should be addressed to Jiahua Xie.

**Additional information**

**1. Supplementary Table 1, 5-9**

Table S1. The summary of read counts of RNAseq analysis in current study.

Table S5. A subset of FPKM related to tapetum genes.

Table S6. The EaF82 interacting proteins.

Table S7. The Hybrigenics' references of seven selected clones for 1-by-1 assays.

Table S8. Summary of interaction matrix and results.

Table S9. List of primer sequences used in current study.

**2. Supplementary Figures 1-11**

Fig. S1 Alkalinization assays of inactive EaF82-S as well as a representative of collected pollen grains and tobacco suspension cells.

Fig. S2. Genetic cassettes used in this study.

Fig. S3. Tissue specific expression of *EaF82* promoter in Arabidopsis transgenic *EaF82p::EaF82-sGFP* (TA) lines.

Fig. S4. Arabidopsis transgenic *EaF82p::EaF82-sGFP* lines (TA-1, -3, -4, and -5).

Fig. S5. Arabidopsis transgenic *35Sp::EaF82-sGFP* (TB) lines.

Fig. S6. RT-qPCR of *GUS* expression levels.

Fig. S7. Male gametophyte development of Arabidopsis transgenic *35Sp::EaF82-sGFP* (TB) line.

Fig. S8. The numbers of seeds per silique with different length (cm) in Arabidopsis transgenic 35Sp::EaF82 (TC) line.

Fig. S9. RNAseq analysis of early developmental flower buds of two independent *35Sp::EaF82* (TC1 and TC2) and vector control (C) transgenic lines.

Fig. S10. Solid growth tests on +/- Histidine and +/- 3-AT plates.

Fig. S11. A "DomSight" of AKIN10 (AT3G01090) displays the information of bait and prey structural, functional and interaction domains.

**Supplementary Tables 2-4 in Excel spreadsheets**

Table S2. Common DEGs that both TC-1 and TC-2 increase two-fold compared to the vector control.

Table S3. Common DEGs that both TC-1 and TC-2 decrease two-fold compared to the vector control.

Table S4. A subset of upregulated and downregulated DEGs with FDR<0.05.

**Supplementary Data Sets 1-3 in Excel spreadsheets**

Data Set S1. The FPKM values for normalized read counts of RNAseq analysis of Arabidopsis transgenic *35Sp::EaF82* (TC) and vector control (C) lines.

Data Set S2. Gene Ontology enrichment analysis of upregulated DEGs.

Data Set S3. Gene Ontology enrichment analysis of downregulated DEGs.

Table 1. RT-qPCR of a subset of selected DEGs

| | | TC-1 (n=3) | | TC-2 (n=3) | |
|---|---|---|---|---|---|
| **Gene_ID**[a] | **Gene description**[b] | **RNAseq** (FC) | **RT-qPCR** (FC ± SD) | **RNAseq** (FC) | **RT-qPCR** (FC ± SD) |
| ***Flowering and pollen development related*** | | | | | |
| AT2G45660 | ***SOC1/AGL20*** (AGAMOUS-like) | 2.52 | 2.15±0.27 | 2.55 | 2.57±0.13 |
| AT4G24540 | ***AGL24*** (AGAMOUS-like) | 3.02 | 2.96±0.35 | 3.27 | 3.80±0.86 |
| AT5G62165 | ***AGL42*** (AGAMOUS-like) | 2.10 | 2.13±0.05 | 2.10 | 2.35±0.18 |
| AT4G35900 | ***FD*** (Basic-leucine zipper transcription factor) | 3.26 | 1.80±0.16 | 3.84 | 2.24±0.07 |
| AT1G19890 | ***MGH3*** (male-gamete-specific histone H3) | -2.93 | -4.34±1.32 | -4.91 | -4.34±1.65 |
| AT1G19960 | transcription factor | -5.10 | -6.60±0.87 | -7.50 | -8.55±2.86 |
| AT1G21000 | PLATZ family transcription factor | -4.16 | -4.13±1.52 | -4.42 | -3.86±1.08 |
| AT1G35490 | bZIP family transcription factor | -5.35 | -7.04±3.20 | -7.11 | -7.31±2.50 |
| AT2G36080 | ***ABS2/NGAL1*** (AP2/B3-like transcription factor) | -4.40 | -5.75±1.67 | -4.58 | -5.69±1.69 |
| AT1G24520 | ***BCP1*** (*Brassica campestris* homolog pollen protein 1) | -6.65 | -3.15±1.65 | -14.91 | -98.37±5.98 |
| AT5G17480 | ***PC1/APC1/CML29*** (pollen calcium-binding protein 1) | -2.44 | -4.17±1.85 | -2.66 | -3.99±0.64 |
| AT4G10603 | ***SLR1-BP*** (S locus-related glycoprotein 1 binding pollen coat protein) | -2.29 | -3.15±0.55 | -2.64 | -2.57±0.57 |
| AT1G29140 | Pollen Ole e 1 allergen and extensin family protein | -3.75 | -5.85±1.94 | -4.76 | -6.05±3.63 |
| AT5G45880 | Pollen Ole e 1 allergen and extensin family protein | -4.09 | -6.26±2.00 | -6.70 | -5.81±0.82 |
| ***$H^+$-ATPase*** | | | | | |
| AT5G57350 | ***HA3*** | -2.12 | -2.83±0.74 | -2.28 | -2.16±0.46 |
| AT2G07560 | ***HA6*** | -2.95 | -4.31±1.35 | -3.48 | -3.87±1.00 |
| AT3G42640 | ***HA8*** | -2.19 | -2.50±0.95 | -2.74 | -2.54±1.49 |
| AT1G80660 | ***HA9*** | -3.84 | -4.92±1.00 | -4.52 | -5.08±1.54 |
| AT3G08560 | ***VHA-E2*** | -2.50 | -3.41±0.21 | -2.90 | -2.94±1.14 |
| AT4G25950 | ***VATG*** | -2.45 | N.A. | -2.95 | N.A. |
| ***Protein kinases*** | | | | | |
| AT2G07040 | ***PRK2A*** (Leucine-rich repeat protein kinase) | -5.07 | -8.55±4.81 | -7.81 | -8.47±3.81 |
| AT2G18470 | ***PERK4*** (Proline-rich protein kinase) | -3.24 | -4.52±1.24 | -4.07 | -4.18±0.51 |
| AT2G21480 | ***BUPS2*** (Malectin/receptor-like) | -3.96 | -6.52±4.87 | -5.45 | -7.75±2.38 |
| AT4G39110 | ***BUPS1*** (Malectin/receptor-like) | -3.14 | -1.99±0.49 | -4.65 | -2.10±0.84 |
| AT5G28680 | ***ANX2*** (Malectin/receptor-like) | -5.86 | -8.16±6.36 | -12.79 | -25.23±13.51 |

[a]: All DEGs selected have FDR<0.05 and fold changes (FC) great than two-fold; [b]: Gene names are italicized and bolded. N.A.: not analyzed.

Table 2. RT-qPCR of seven RALF genes and their predicted pI and molecular weight (MW)

| | | TC-1 (n=3) | | TC-2 (n=3) | | Predicted pI and MW[c] | |
|---|---|---|---|---|---|---|---|
| **Gene_ID**[a] | **Gene description (clade)**[b] | **RNAseq** (FC) | **RT-qPCR** (FC ± SD) | **RNAseq** (FC) | **RT-qPCR** (FC ± SD) | **pI** | **MW** (kD) |
| AT1G28270 | ***RALF4*** (III-B) | -4.27 | -7.00±2.79 | -5.96 | -8.37±4.59 | 9.76 | 8.7 |
| AT1G61563 | ***RALF8*** (IV-C) | -3.67 | -5.26±1.60 | -5.35 | -5.66±1.00 | 9.30 | 6.6 |
| AT1G61566 | ***RALF9*** (IV-C) | -3.51 | -4.77±1.26 | -5.27 | -5.74±1.38 | 9.27 | 6.6 |
| AT2G22055 | ***RALF15*** (IV-C) | -5.73 | -2.48±0.39 | -23.41 | -16.92±2.40 | 9.76 | 6.5 |
| AT2G33775 | ***RALF19*** (III-B) | -5.26 | -9.10±5.00 | -9.69 | -8.14±3.43 | 10.57 | 7.5 |
| AT3G25165 | ***RALF25*** (IV-B) | -8.24 | -2.48±0.67 | -23.55 | -82.52±18.38 | 10.08 | 7.0 |
| AT4G14020 | RALF family protein (IV-A) | -3.39 | -4.29±1.30 | -3.78 | -4.02±1.27 | 10.17 | 6.8 |

[a]: All DEGs selected have FDR<0.05 and fold changes (FC) great than two-fold; [b]: Gene names are italicized and bolded; [c]: Amino acid sequences were retrieved from TAIR (https://www.arabidopsis.org/). N.A.: not analyzed.

## Figure legends

**Fig. 1** EaF82 peptide and alkalinization assay. **a** Amino acid sequence. The predicted signal peptide is highlighted in gray. Four cysteines (C) are marked in red with predicted potential intramolecular disulfide bridges indicated with blue brackets. Sequences for making antibody is underlined. **b** Alkalinization assays of EaF82-S (EaF82 without signal peptide). Six different concentrations (0, 1, 5, 10, 50, and 100 nM) were tested and the pH changes (Δ pH) were measured after 10, 20 and 30 min. Data plotted were the average of five independent experiments ± SD. The Km and Vmax are listed below. Data marked with the same letter are not significantly different by the LSD test at 5% level of significance. **c** Alkalinization activity measured for 60 min. Data represent the average of five independent experiments ± SD. **d** Phylogenetic tree of EaF82 among 795 RALFs from 51 plant species. Clade I, II, III and IV as well as their subgroups are categorized following Campbell and Turner[5]. **e** Phylogenetic tree of EaF82 with clade IV-C of AtRALFs.

**Fig. 2** The functional characterization of Arabidopsis transgenic lines carrying EaF82 promoter driving *EaF82-sGFP* (TA) or *GUS*. **a** The newly opened flower (left) shows GFP signal in anthers (right). A: anther. **b** Confocal microscopy shows GFP (left) and DIC (differential interference contrast, right) in stamen (S) and some pollen grains (PG). Bar = 100 μm. **c** GUS staining of the flowers of *EaF82p::GUS*. White arrow indicates GUS activity at the early

developmental stages of flowers. Bar = 1.5 mm. (**d**) Immunoblot of pollen proteins against anti-EaF82 and anti-GFP. Stained blots show protein loading. TA-1, -3, -4 and -5: four independent lines. TB-2: Proteins isolated from flowers of Arabidopsis transgenic *35Sp::EaF82-sGFP*. B: blank. C: vector control. WT: wild-type. M: protein size marker. **e** Confocal microscopy detecting GFP in pollen of Arabidopsis transgenic *EaF82p::EaF82-sGFP* (TA). Wild-type (WT) is a negative control. Bar = 100 µm. **f** A 10-week old TA plant (right) bears aborted siliques in primary inflorescence stalk (yellow bracket) compared to normal WT (left). Bar = 1 cm. **g** Pollen from fully opened flowers stained with iodine-potassium iodide. **h** Germinated pollen under germination medium. Bar = 100 µm.

**Fig. 3** Male gametophyte and pollen development of Arabidopsis transgenic *35Sp::EaF82-sGFP* lines (TB) comparing to wild-type (WT). **a** Fully opened flower (left) and stained anther with iodine-potassium iodide (right) of WT and TB line, respectively. Bar = 100 µm. **b** TB plants (right) produce no siliques with occasionally observed small silique (white arrow) compared to WT (left). Bar = 1 cm. **c** TB silique carries less seeds than that of WT (left). The numbers of seeds per silique are plotted as the mean ± SD (right). n: numbers of siliques. **d** Male gametophyte development of TB line compared to WT. At the stage S11, the undegenerated tapetum (red arrow) was observed in TB line. At the stages S11 and S12, underdeveloped pollen was stained in light green, while mature pollen is in dark blue that occasionally was observed in TB line (black arrow). Bar = 50 µm.

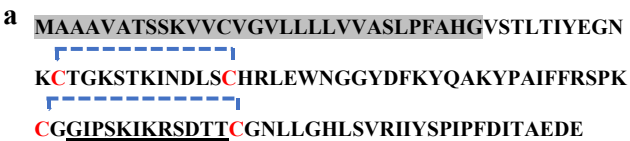
**Fig. 4** The functional characterization of tobacco transgenic *35Sp::EaF82* lines (T2, 3, 8 and 11) comparing to two independent vector control lines (C1 and C2). **a** Immunoblot of leaf proteins against anti-EaF82. Stained blot shows protein loading. **b** T8 (right) and C1 (left) opened flowers. **c** Immunoblot of unopened flower proteins against anti-EaF82. The unopened flowers are as indicated in (d) with orange star. **d** Transgenic plants with normal (red arrow) and aborted (white arrow) seed pods. Enlarged seed pods are shown below. **e** Seed counts per pod. Data plotted are the average from six independent plants per line using 10 seed pods per independent plant ± SD. Data marked with the same letter are not significantly different by the LSD test at 5% level of significance. Bar = 200 µm.

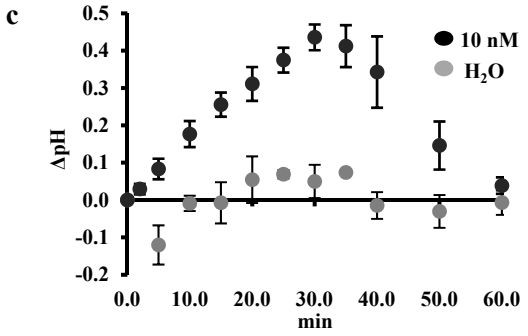
**Fig. 5** The functional characterization and RNAseq analysis of Arabidopsis transgenic *35Sp::EaF82* lines (TC-1 and -2) compared to wild-type (WT) and vector control (C). **a** RT-PCR of leaf tissues using primer pair specific to *EaF82* and *18S rRNA*. +: plasmid DNA. **b** Immunoblot of leaf proteins against anti-EaF82. Stained blot shows protein loading. M: protein size marker. **c** Normal growth of TC and WT plants. **d** Aborted siliques in TC-1 and TC-2. **e** Immunoblot of proteins from unopened whole flowers against anti-EaF82. **f** Stained pollen and **g** germinated pollen. The descriptions are the same as in Fig. 2. Bar=100 μm. **h** Histogram of the silique lengths. Data plotted are the percentages of total siliques. n: numbers of siliques. **i** The number of up- and down-regulated DEGs (≥ 2-fold) in TC-1 and -2 lines. **j** Enriched GO terms from down-regulated DEGs related to pH regulation and cell wall modification.

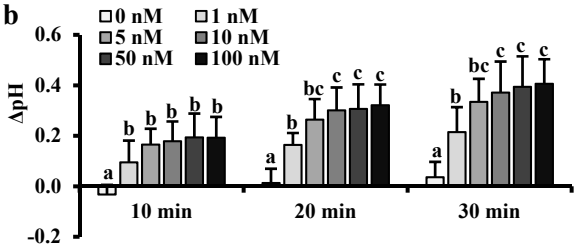
**Fig. 6** Immunoblots of three subunits of the proteasome, proteasome activity and ATP content in early developmental flowers of TC lines. **a** Anti-Rpn6, anti-Rpn10 and anti-Rpt5a showed no significant difference among all samples. **b** Proteasome activity. Data plotted are the average of three biological replicates ± SD. RFU: relative fluorescence units. **c** ATP content. Data plotted are the average of four biological replicates ± SD. FW: fresh weight.

**Fig. 7** Identification and characterization of an EaF82 interacting partner. **a** Yeast growth tests of 1-by-1 Y2H assay on selective medium without (DO-2) or with (DO-3) histidine and 3-aminotriazole (3AT). Supporting information is detailed in Supplementary Fig. 10. +: positive interaction; pB27ø and pP7ø: empty vectors. **b** Co-immunoprecipitation (co-IP) analysis to validate the Y2H results of EaF82 and AKIN10 interaction. Protein extracts from floral tissues of vector control (C) and TC transgenic lines were spiked with (+) or without (-) EaF82-S. The complex was pull-downed with (+) or without (-) anti-AKIN10 antibody. The immobilized anti-AKIN10 on protein A/G magnetic beads could pull-down EaF82 and AKIN10 complex from both C and TC lines (lanes 4 and 7). EaF82 was undetectable without immobilized anti-AKIN10 antibody (lanes 3 and 6). Lane 1: protein size marker. Lanes 2 and 5: extracted proteins from C and TC lines. Lane 8: EaF82-S alone bound to protein A/G magnetic beads. **c** A representative immunoblot of proteins from unopended flowers against anti-AKIN10 that detects both AKIN10 (61 kD) and AKIN11 (58 kD). The average of band intensity is plotted as three independent

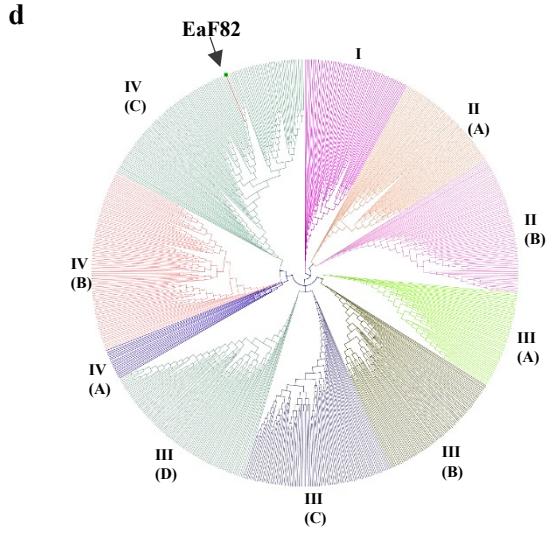
1150 experiments ± SD (below). Data marked with the same letter are not significantly different by the
1151 LSD test at 5% level of significance.




| | 10 min | 20 min | 30 min |
|---|---|---|---|
| Vmax (ΔpH) | 0.20±0.02 | 0.32±0.02 | 0.40±0.03 |
| Km (nM) | 1.10±0.72 | 0.96±0.41 | 0.91±0.41 |
| R square | 0.54 | 0.71 | 0.67 |
| Sy.x | 0.072 | 0.073 | 0.095 |



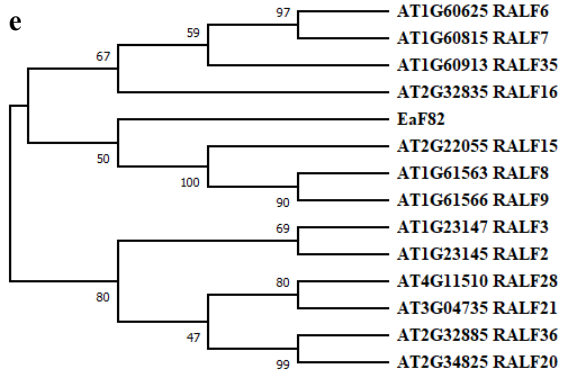




**Fig. 1**

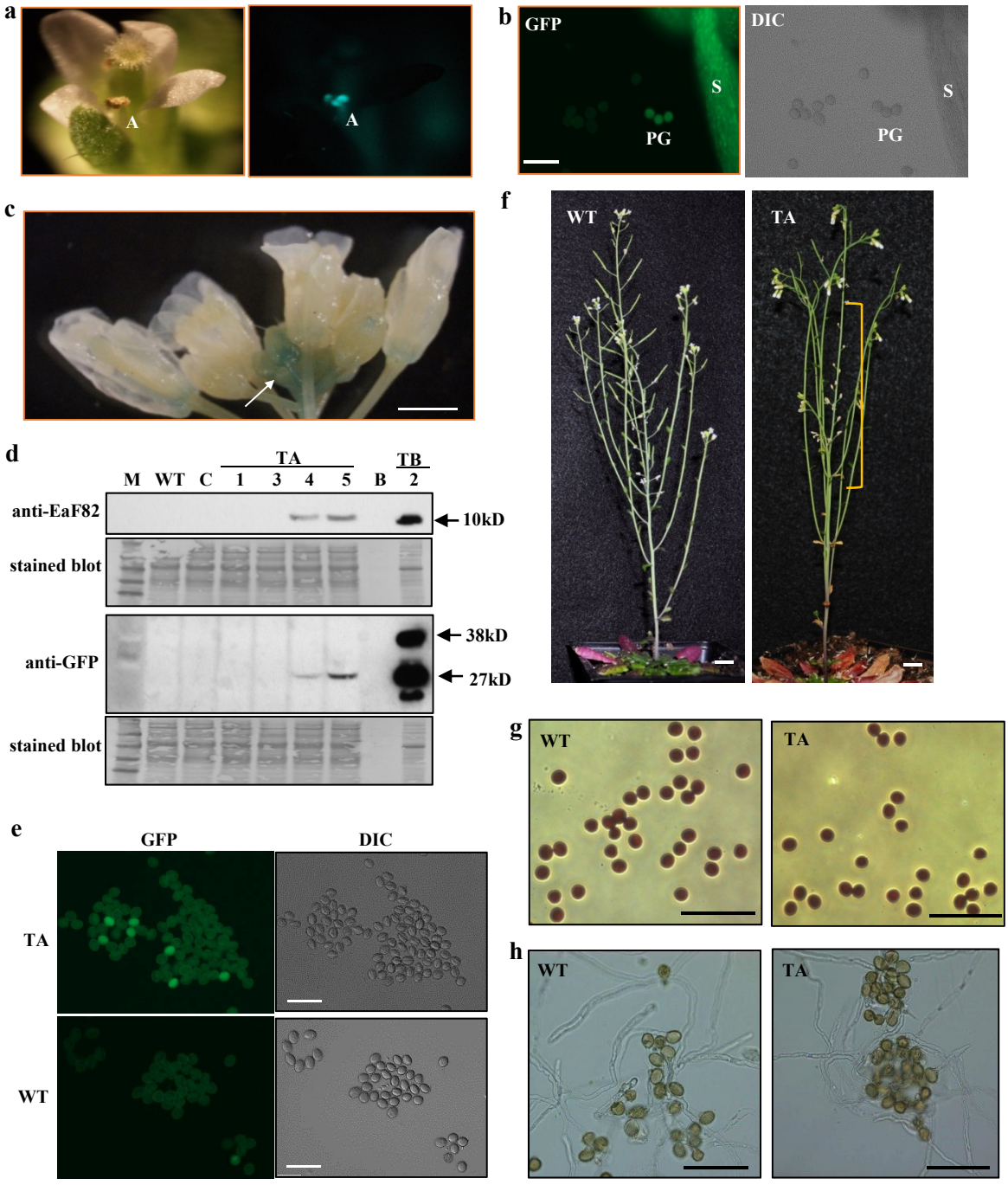



**Fig. 2**

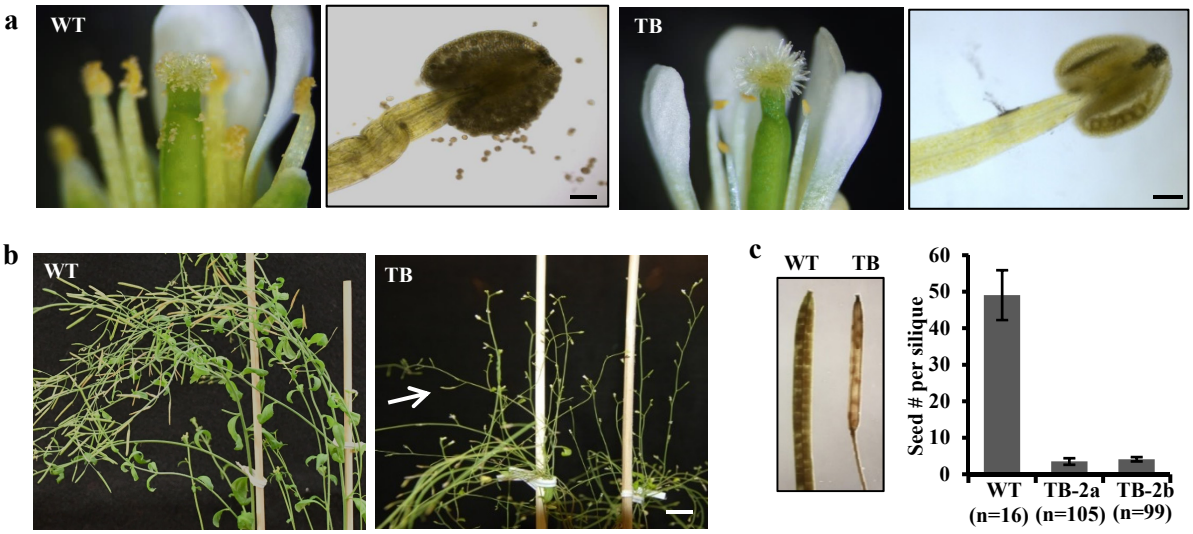



**Fig. 3**

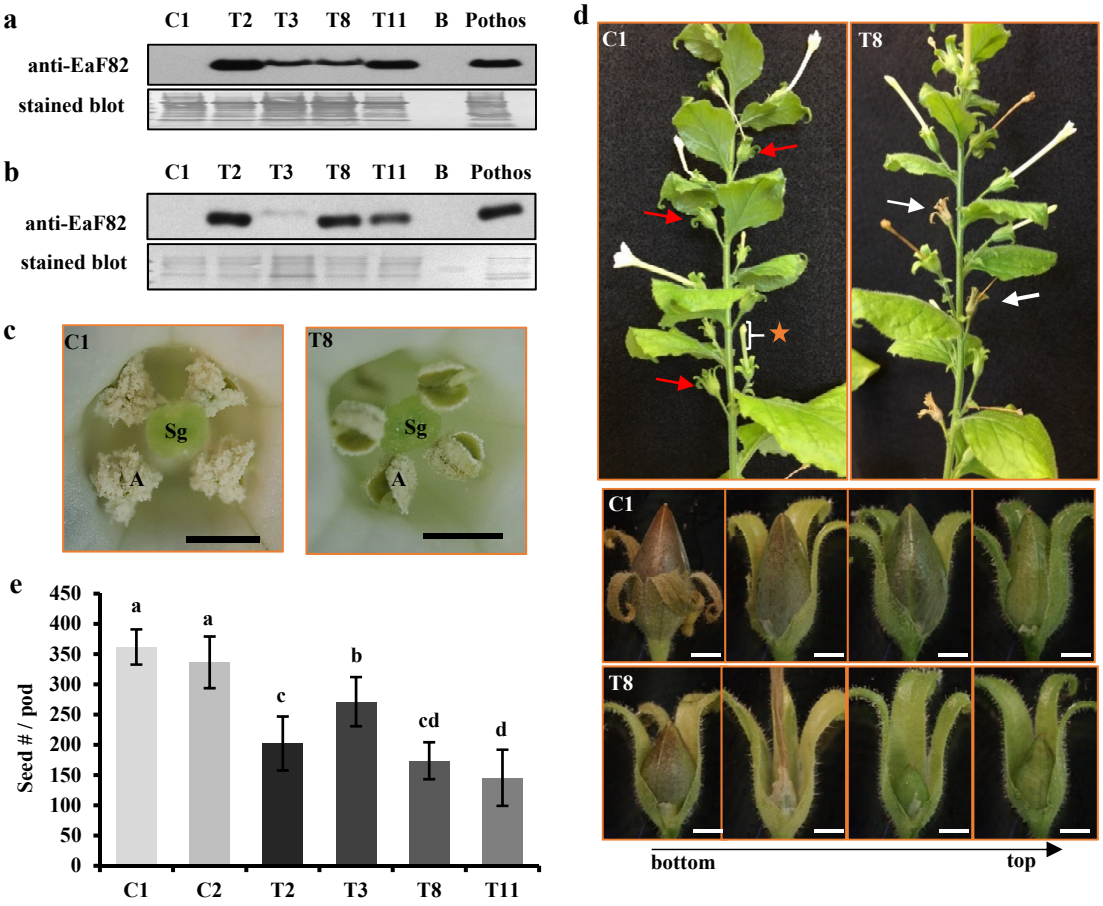



**Fig. 4**

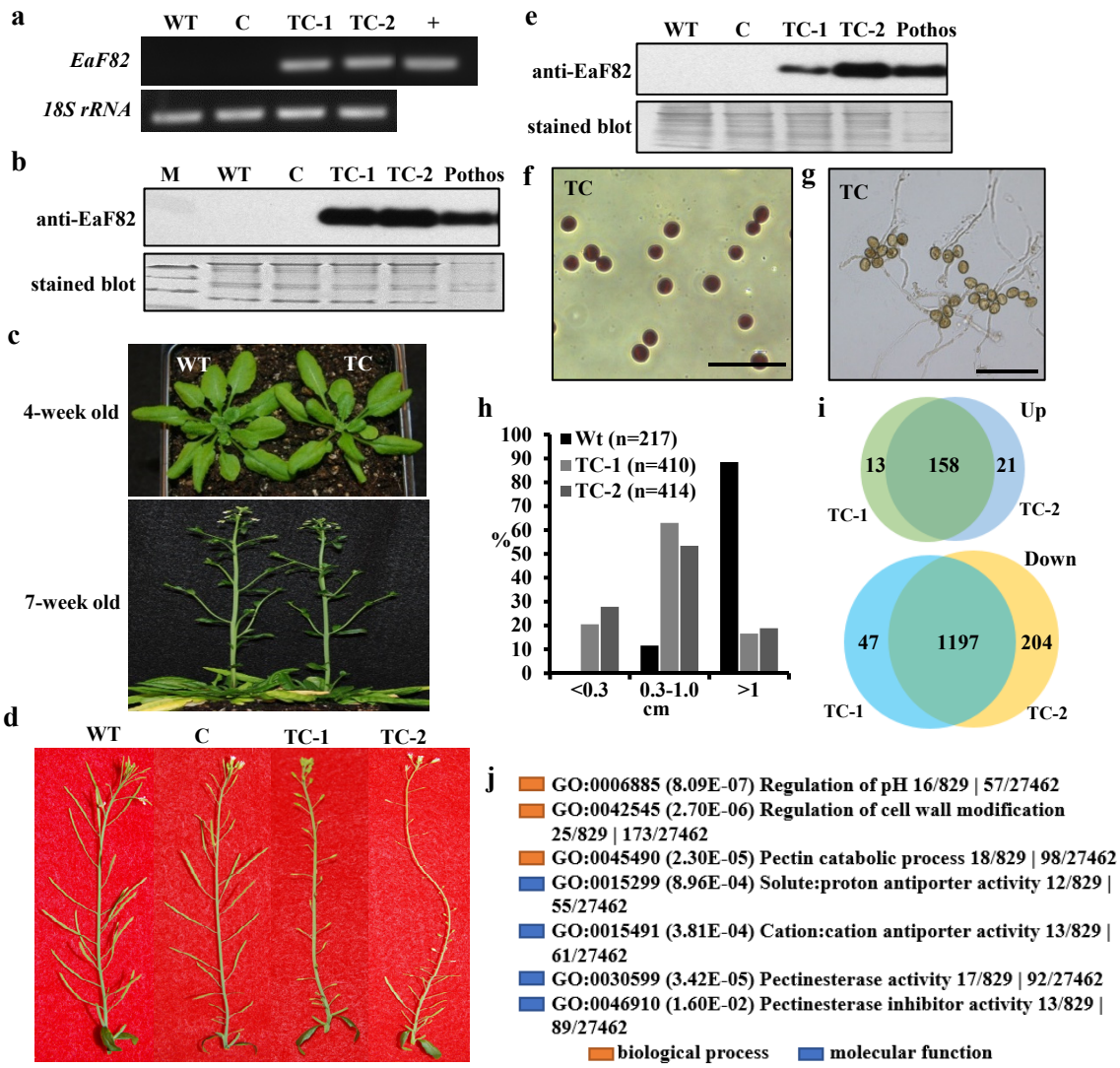



**Fig. 5**

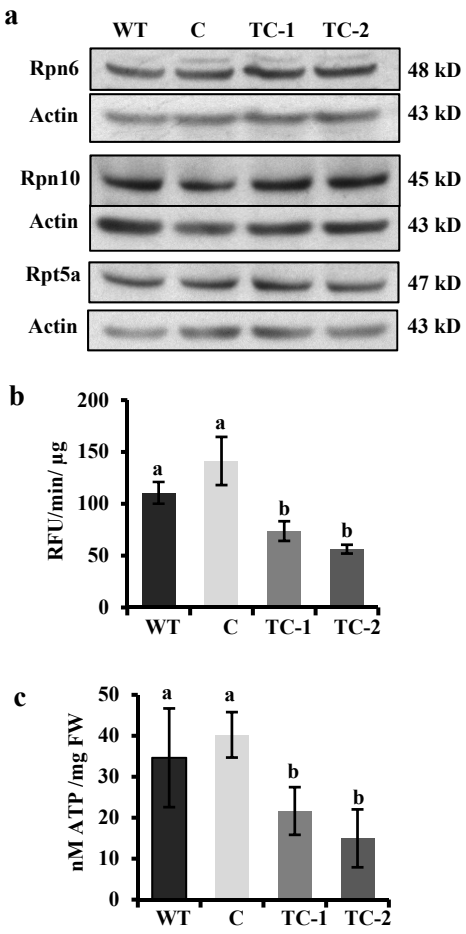
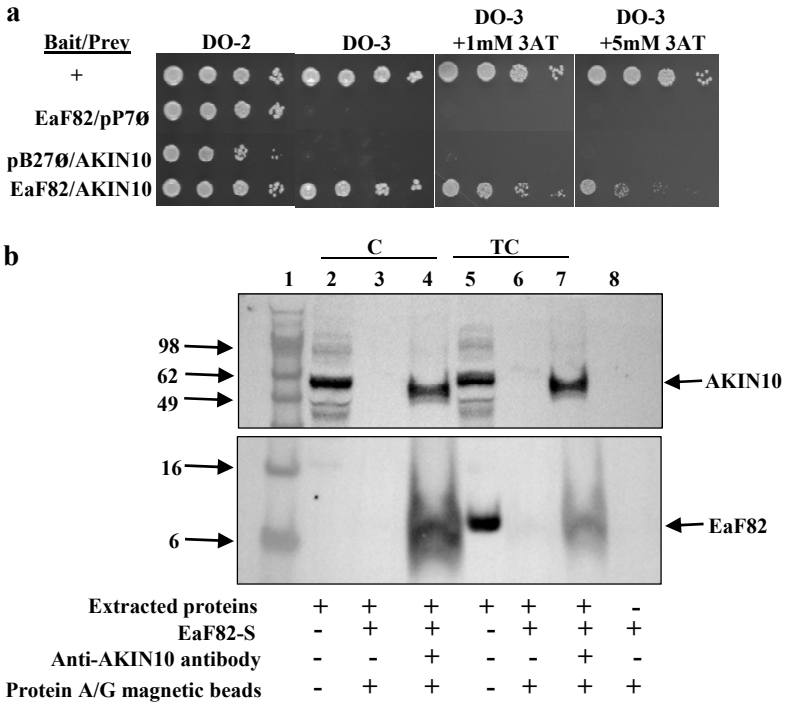



**Fig. 6**



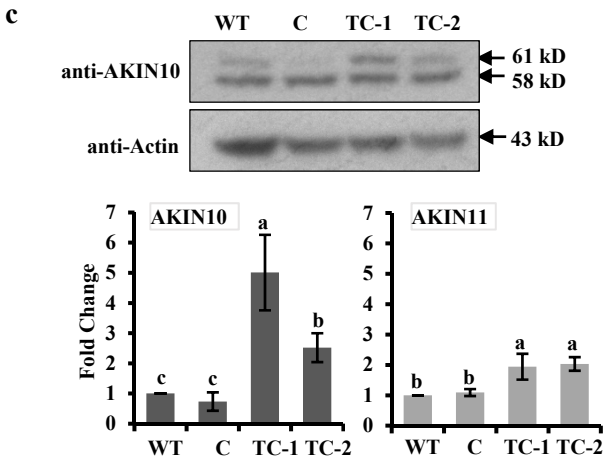




**Fig. 7**