THE CROUCHER FOUNDATION

General Report Summary

Project Number: CAS10ST (CAS10SC01 and CAS10SC01-M1)
Project Title: *Identification of in vivo substrates for MAP-kinase cascades that play key roles in ethylene signaling using phosphoproteomics and interactomics*
(R6086-5101/5151)
Project Title in Chinese:
使用磷酸化蛋白组学和相互作用组学寻找植物细胞里乙烯信号传导过程中的 **MAPK** 激酶底物

Principal Investigator: LI, Ning/LIFS/The Hong Kong University of Science and Technology
Collaborator(s) outside Hong Kong : Li, Jia Yang/Genetic Institute /CAS

This grant was awarded in Oct-2010.

Publication Summary of this particular project (number only)

(a)	Already published in international refereed journals	1
(b)	Being prepared for publication in international refereed journals	2
(c)	Published conference papers	0
(d)	Conference papers already presented but not published	1
(e)	Chapter in books	1

Number of research student(s) registered (exclusively in association with this project, if any) 2

Research Teams

	CAS team	**HKUST team**
Name of Project Coordinator (with title)	Dr. **Jia Yang LI** (Academician of CAS)	Dr. **Ning Li**
Position	Vice President of CAS	Professor
Name of Co-Investigator	None	None
Other project team members(if any)	Wang ,Yong Hong	Wong, Wai Shing
Mailing Address	Institute of Genetics and Development 52, Sanlihe, Beijing 100864, China	Department of Biology, Clearwater Bay, Hong Kong SAR
Phone Fax E-mail	010-68597602 010-68527508 jyli@cas.cn	852-2358-7335 852-2358-1559 boningli@ust.hk

Research Activities Report

Objectives of this study

This proposed project includes 3 objectives:

1. **To large scale profile ethylene-regulated phosphopeptides in both *ctr1* mutant and wild type Arabidopsis using SDS-PAGE and SCX separation, IMAC and TiO_2 enrichment coupled with nanoLC-MS/MS.**
2. ***In vitro* validation of the ethylene differentially regulated phosphorylation sites and to measure the differentially regulated plant kinase activities related to ethylene treatment.**
3. **To investigate the *in vivo* biological functions of a number of selected phosphorylated proteins obtained from above phosphoproteomics approach.**

Research activities in progress

1) ^{15}N Stable isotope labeling of *ctr1* mutant (a MAPKKK-deficient mutant) and the wild type Arabidopsis

Recently, we have developed a quantitative phosphoprotoemic method, SILIA (Guo and Li, 2011, Phytochemistry), to label the total proteins *in planta* using stable isotope, heavy nitrogen. With the advent of SILIA, a differential phosphoproteomics is performed in between WT plants and *ctr1* mutants. To that end, *Arabidopsis thaliana* ecotype (Col-0) and its mutant (*ctr1-1*) were surface-sterilized and grown in glass jar containing ME2 medium supplemented with 18 mM of either **^{14}N-KNO3 or ^{15}N-KNO3**. Seedlings were also treated with or without 5 μM ACC in each ^{14}N- or ^{15}N- labeled experiment (as shown in **Fig 1**). Seedlings were grown with a 14 hours light cycle for 3 weeks before being harvested. These plants are ready to be analyzed by GE-LC-MS/MS to identify phosphoproteins unique to either WT or ctr1 under ethylene (or ACC) treatment (Guo and Li, 2011, phytochemistry)

2) Development of the SILIA-based relative and accurate measurement method for phosphorylated proteins

In the first period of study, we have successfully established an *in planta* quantitative proteomics method for measurement of phosphorylation and dephosphorylation status (Guo and Li, 2011). Using SILIA-based quantitative phosphorpoteomic method, we found that the ethylene-related transcription factor ERF110 originally discovered by bioinformatics (Li et al., 2009, Proteomics) was indeed down-regulated by ethylene *in planta* (**Fig 2**).

3) Development of the absolute quantitation of isoforms of post-translationally modified protein in transgenic Arabidopsis (AQUIP)

In combination with the *stable isotope labeling in Arabidopsis* (SILIA) method, we also developed a novel PTM proteomics protocol, AQUIP, to absolutely and accurately quantitate the phosphor-PTM of a ethylene response factor ERF110 at position 62aa in *Arabidopsis* and to measure changes in phosphosite abundance in both the treated or mutant plant samples as compared to the control plant (Li et al., 2012, Mol Cell Proteomics, **Fig 3**).

4) High-throughput profiling of ethylene-regulated phosphoproteins in ethylene response mutants

With the advent of the *stable isotope labeling in Arabidopsis* (SILIA, Guo and Li, 2011)-based quantitative proteomics, we have performed several high-throughput differential phosphoproteomics analysis on ethylene response mutants *ctr1, eer1 (rcn1), ein3eil1* as well as on ethylene-treated wild type plants. Hundreds and thousands of phosphopeptides have been identified from these plants. SILIA-based differential phosphoproteomics revealed that there are ctr1, eer1, ein3eil1 and ethylene-treated WT-specific phosphosites. Some of them are considered to be putative MAPKKK substrates as well as the putative substrates of RCN1 PP2A phosphatase. These two modifying and demodifying enzymes might function in an antagonistic fashion and regulate ethylene response in a coordinated manner (one manuscript is currently in preparation, **Fig 4**).

5) In vitro plant kinase assay validation of both MS- and bioinformatics-derived phosphopeptides

A BLAST-based phosphorylation site motif-mining approach was adopted to discover more ethylene-regulated phosphorylation sites, by which a short segment of polypeptide sequence in length of 13 - 21 amino acids, containing the MS-derived phosphopeptide and having its phosphosite located at the middle of this query polypeptide sequence, was usually searched against the entire Arabidopsis protein sequence database (*A. thaliana*, taxid: 3702) and target proteins with 55.5% homology or more to the query protein sequence were usually selected and further analyzed using amino acid sequence alignment program in ClustalW (http://www.ebi.ac.uk/clustalw/). The chosen proteins sharing the similar phosphorylation site were grouped into a motif. By this method, we have constructed dozens of conserved phosphorylation site motifs related to ethylene-regulated phsophosite motifs.

To validate these bioinformatics-predicted individual phosphosites of phosphorylation motifs and those authentic phosphoistes derived from MS study of ethylene response mutants (ctr1, eer1, ein3eil1 and ethylene treated WT plants), 23-aa long polyhistidine-tagged peptides were synthesized chemically. Out of 19 MS-derived phosphosites, 13 of them were in fact phosphorylated by the *in vitro* plant kinase assay (**Table 1**). Once more, this "Blast and Align"-based simple phosphosite prediction method (Li et al. 2009) is proven to be useful in identification of novel phosphosites and phosphorylation motifs.

6) *In planta* validation of both MS- and bioinformatics-derived phosphosites

Ethylene is able to regulate protein phosphorylation via both EIN2-dependent and –independent pathways (Li et al., 2009). High throughput proteomic profiling on ethylene response mutants strongly support this conclusion. To provide *in planta* genetic evidence to this conclusion, a bioinformatics-predicted ethylene response factor 110 has been introduced into *ein2-5, etr1-1* and wild type plants. A gain-of-function of bushy phenotype has been obtained from ERF110/ein2-5 and ERF110/wild type transgenic plants. However, S62A point mutation abolished this bushy phenotype in both ein2-5 and wt background. These results suggest that S62 phosphorylation/dephosphorylation occurs in an EIN2-independent pathway yet production of ERF11o protein is EIN2-dependent (**Fig 5**, Li et al., Mol Cell Proteomics, 2012 in press and Zhu et al., 2012, in preparation).

To provide in planta validation results on these hundreds and thousands of phosphosites found from ethylene response mutants, candidate phosphorylated proteins (or putative substrates of MAPK cascades), we have developed a binary vector containing a histidine$_8$-biotin-histidine$_8$ (HBH) tag for target protein fusion, this new tandem tag for two-step affinity purification under fully denaturing conditions from *Arabidopsis* cell lysate has been linked to several candidate ethylene regulated phosphoproteins and been transformed into ein3eil1 Arabidopsis. The transgenic plants over-expressing these phosphoproteins will be studied using mass spectrometers to validate the ethylene inducibility (**Fig 6**).

7) Development of nanowires for high-throughput proteomic analysis of kinase and phosphatase

Provided with success of identification of large number of phosphoproteins from ethylene response mutants, the next important research focus is to identify kinases (MAPKKK, MAPKK and MAPK) and phosphatases that catalyze the these ethylene-regulated phosphosites. Eventually, we intend to establish PTM networks of ethylene signaling. This work should provide the first example to study plant cell signaling using functional PTM proteomics, interactomics and molecular systems biology. To this end, we have developed nanowire technology to fix polypeptides for a high-throughput isolation of kinase and phosphatases. Given the large surface area of nanowires and nano particles, we have fabricated 3D-core radiating α-Fe_2O_3 nanowires, which have been modified with Ni-NTA for the conjugation of his-tagged substrate proteins or peptides (from this study) to nanowires. The Ni-NTA-functionalized α-Fe_2O_3 nanowires present better purification efficiency, whose protein-binding capacity was about 30ug/ mg. After stored them as dry powder for months, these modified nanowires still showed a stability in protein or peptide purification. The functionalized silicon nanowire-chips were prepared for immobilizing the MAPK substrate peptides through hexohistidine tag (**Fig 7**) to purify protein kinases like MAPKs (Hu et al., 2012, in preparation)

Appendix. Figures 1 to 7 and table 1 are provided in appendix.

Publication list

1. Li, YJ[R], Shu, YW[S], Peng, CC[R], Zhu, L[S], Guo, GG[S] and Li, Ning (**2012**) AQUIP: Absolute Quantitation of Isoforms of Post-translationally modified proteins in transgenic organism. **Mol Cell Proteomics**. **In press.**
2. Yang et al., others and Li, JY and Ning LI. Higher resolution mapping of ethylene signaling components via quantitative and differential phosphoproteome analysis of ethylene response mutants. **In preparation**.
3. Zhu et al., others and Li, JY and Ning LI. EIN2-independnet phosphorylation of Ser62 on ethylene response factor 110 is required to maintain normal flowering time of Arabidopsis. **In preparation.**

Appendix

Fig 1

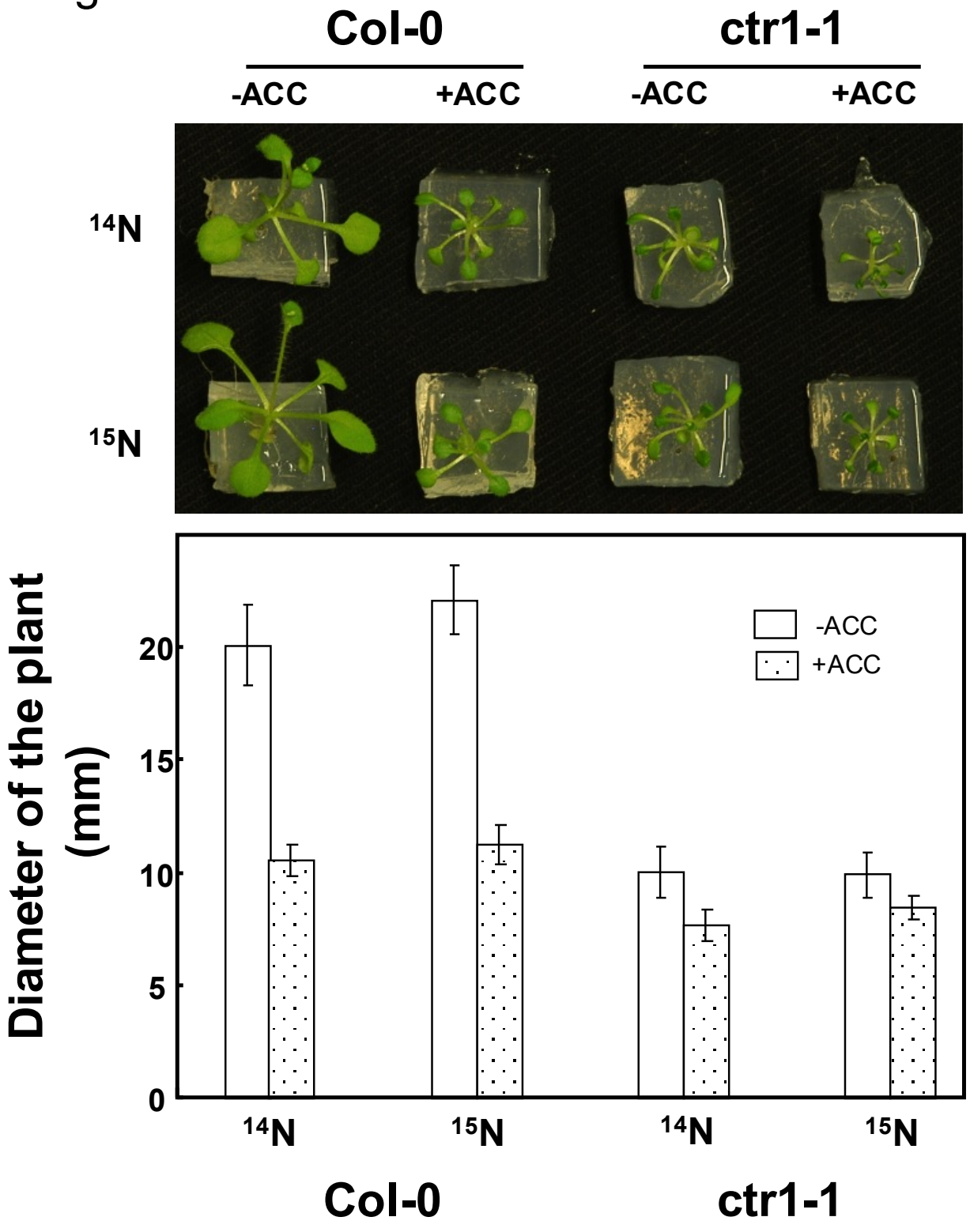


Figure 2

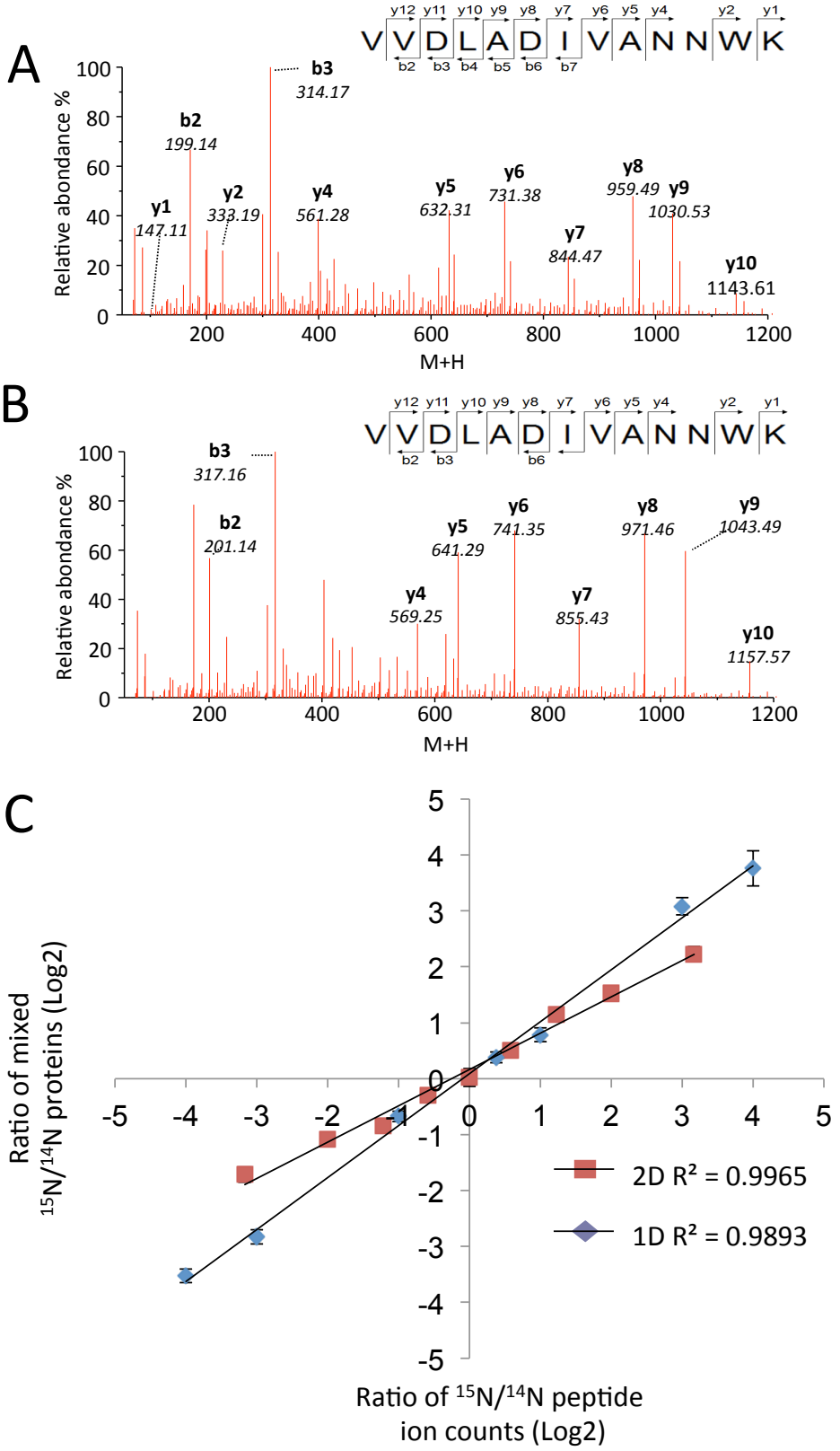

A
y12 y11 y10 y9 y8 y7 y6 y5 y4 y2 y1
V V D L A D I V A N N W K
b2 b3 b4 b5 b6 b7
Relative abondance %
b3 314.17
b2 199.14
y1 147.11
y2 333.19
y4 561.28
y5 632.31
y6 731.38
y7 844.47
y8 959.49
y9 1030.53
y10 1143.61
M+H
B
y12 y11 y10 y9 y8 y7 y6 y5 y4 y2 y1
V V D L A D I V A N N W K
b2 b3 b6
Relative abondance %
b3 317.16
b2 201.14
y4 569.25
y5 641.29
y6 741.35
y7 855.43
y8 971.46
y9 1043.49
y10 1157.57
M+H
C
Ratio of mixed 15N/14N proteins (Log2)
2D R² = 0.9965
1D R² = 0.9893
Ratio of 15N/14N peptide ion counts (Log2)

Figure 3

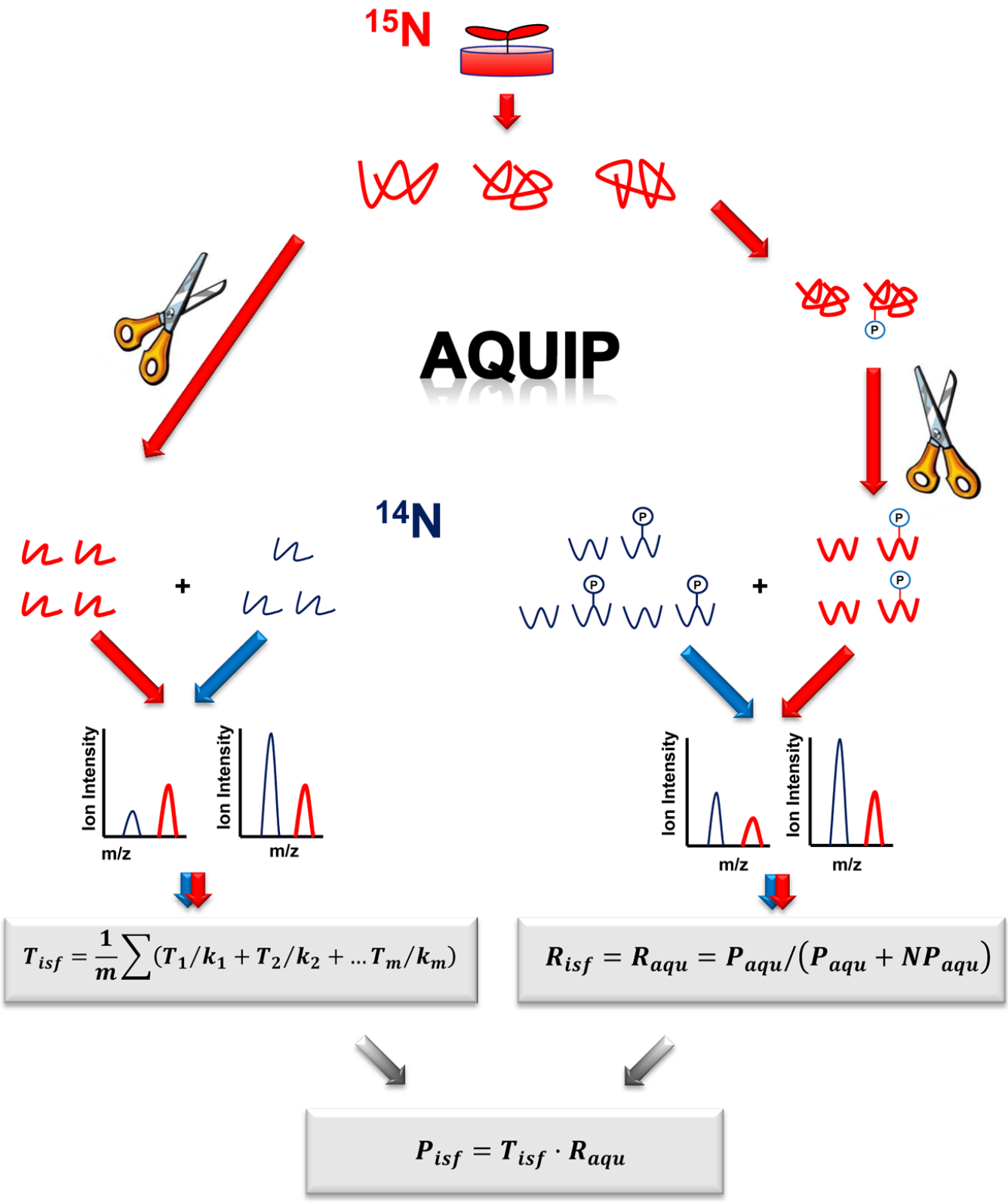
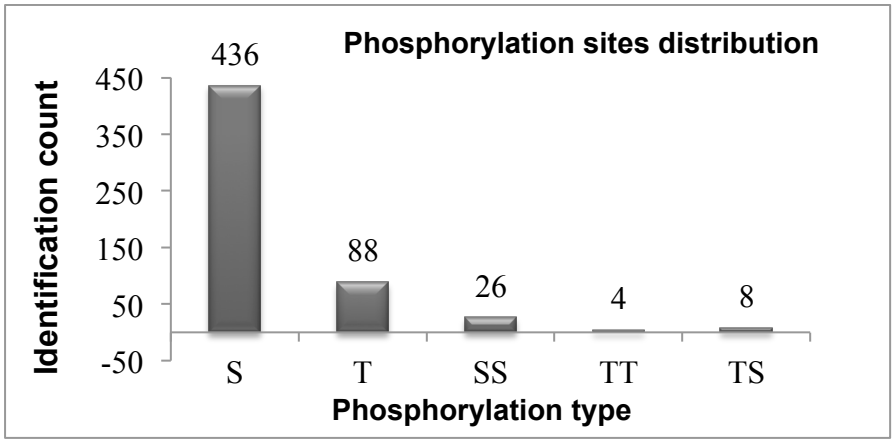
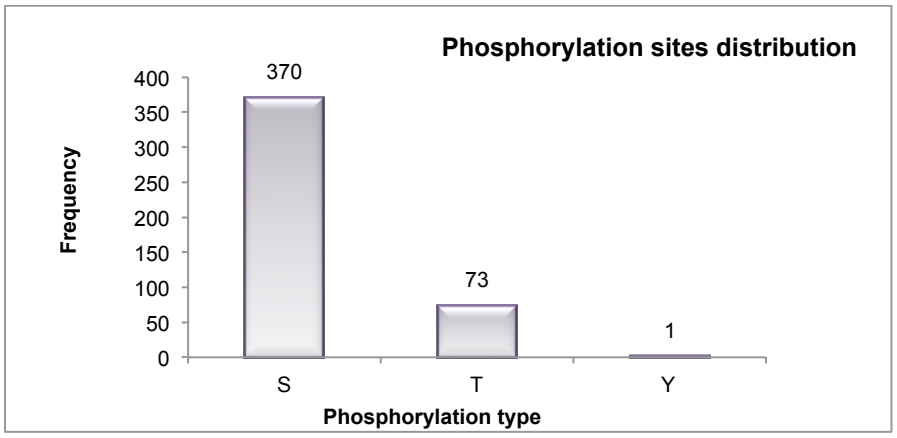

15N
AQUIP
14N
Ion Intensity
m/z
Ion Intensity
m/z
Ion Intensity
m/z
Ion Intensity
m/z
$T_{isf} = \frac{1}{m}\sum(T_1/k_1 + T_2/k_2 + \ldots T_m/k_m)$
$R_{isf} = R_{aqu} = P_{aqu}/(P_{aqu} + NP_{aqu})$
$P_{isf} = T_{isf} \cdot R_{aqu}$

Figure 4

Phosphoproteomics of *ein3eil1*



Phosphoproteomics of *eer1*



Phosphoproteomics of *ctr1*

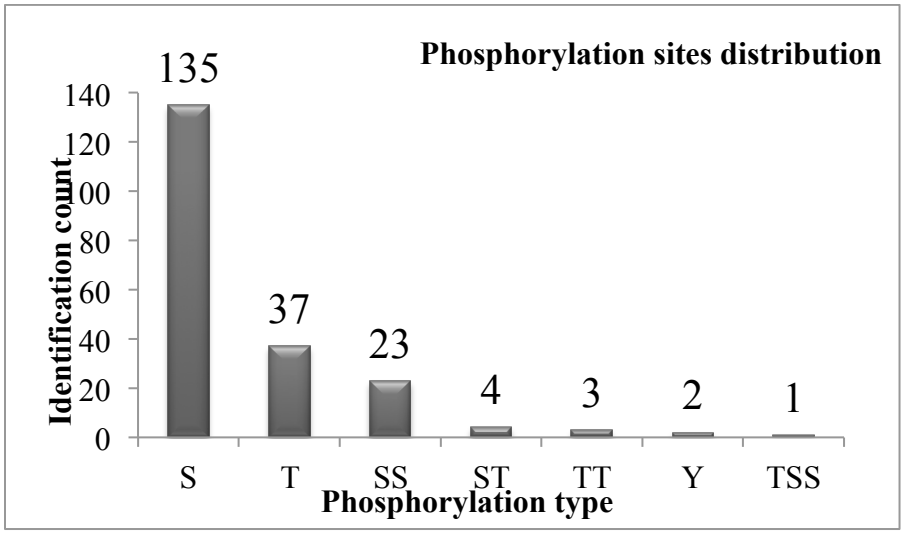


Table 1. *In vitro* plant kinase assay using *plant* kinase extract and synthetic peptides

Accession No.	Putative function	Sequence of synthetic peptides	Kinase assay[§]
AT2G19810.1	Zinc finger (CCCH-type) family protein	HHHHHHLLSR***S*****LGSNLGNDVVASLR	n/p
AT2G46180.1	GC4 | GC4 (golgin candidate 4)	HHHHHHRKLFPK***S*****TEDLSR	n/p
AT1G19870.1	IQD32 calmodulin binding	HHHHHHKHSLKRKV***S*****NPSFIAAQSK	n/p
AT4G35785.1	Nucleic acid binding / nucleotide binding	HHHHHHRRKRPRTP***T*****PGHYLGLK	n/p
AT1G20696.1	HIGH MOBILITY GROUP / DNA bindin	HHHHHHKWKSL***S*****DSEKAPYVAK	n/p
AT1G74910.1	ADP-glucose pyrophosphorylase family protein	HHHHHHRATLKRV***S*****SFEALQPATR	n/p
AT2G28000.1	CHAPERONIN-60ALPHA / ATP binding / protein binding	HHHHHHGRNVVLDEFG***S*****PKVVNDG	n/p
AT5G52040.1	ATRSP41 RNA binding / nucleic acid binding / nucleotide binding	HHHHHHRKGRGESR***S*****PPPYEKR	n/p
AT5G27030.1	**TOPLESS-RELATED 3**	**HHHHHHKRPRTPPA*T**PGIVDYK**	**n/p**
AT1G18840.1	**IQD30; calmodulin binding**	**HHHHHHKDPDDSL*S*ESEKIQQEIA**	**n**
AT3G63280.1	**NIMA-RELATED KINASE / ATP binding / protein kinase**	**HHHHHHKHELMKV*S*NPTERRRR**	**n**
AT4G00900.1	**ER-TYPE CA2+-ATPASE 2) / calcium-transporting ATPase**	**HHHHHHPLKKKLDEFG*S*RLTTAIC**	**n**
AT5G35670.1	IQ-domain 33 / calmodulin binding	HHHHHHRSYLCGDEFN***S*****VRAVNDS	n/p
AT5G49120.1	Senescence-associated protein-related	HHHHHHKNPLSEGVLISPKVVNK	n
AT3G27580.1	ATPK7; kinase/ protein serine/threonine kinase	HHHHHHTSARSM***S*****FVGTHEYLAPEIIK	n/p
AT1G62630.1	Disease resistance protein (CC-NBS-LRR class), putative	HHHHHHRLLQIFGSNIF***S*****PDRQL	n/p
AT3G13224.2	RNA recognition motif (RRM)-containing protein	HHHHHHKKSLNR***S*****PPSYGSHPR	n/p
AT2G20950.1	CONTAINS InterPro DOMAIN/s: Phospholipase-like	HHHHHHKSASKSNPS***S*****PPHLAEGR	n/p
AT5G52040.1	ATRSP41 RNA binding / nucleic acid binding / nucleotide binding	HHHHHHSPSPYKRARL***S*****PDYKRDDR	n
AT1G20760.1	Calcium-binding EF hand family protein	HHHHHHPSTPLSRFGN***S*****PPRFSDASAR	n
AT5G23060.1	CaS | CaS (Calcium sensing receptor)	HHHHHHPSRIIPAASR***S*****FGTRSGTK	n
AT2G17410.1	ARID/BRIGHT DNA-binding domain-containing protein	HHHHHHPESKLSEDTG***S*****PHHHADILMVR	n
AT5G10360.1	EMB3010 | EMB3010 (embryo defective 3010)	HHHHHHRSESLAKKR***S*****RLSSAPAKPV	n

[§]MS-identification of substrate peptides from *in vitro* plant kinase assay:
n stands for non-phosphorylated peptide; *p* stands for phosphorylated peptide.
*Stands for the MS-detected phosphosites either from *in vitro* or *in vivo* experiments.
Bold, the bioinformatics-predicted putative phosphopeptide substrates are bolded

Figure 5

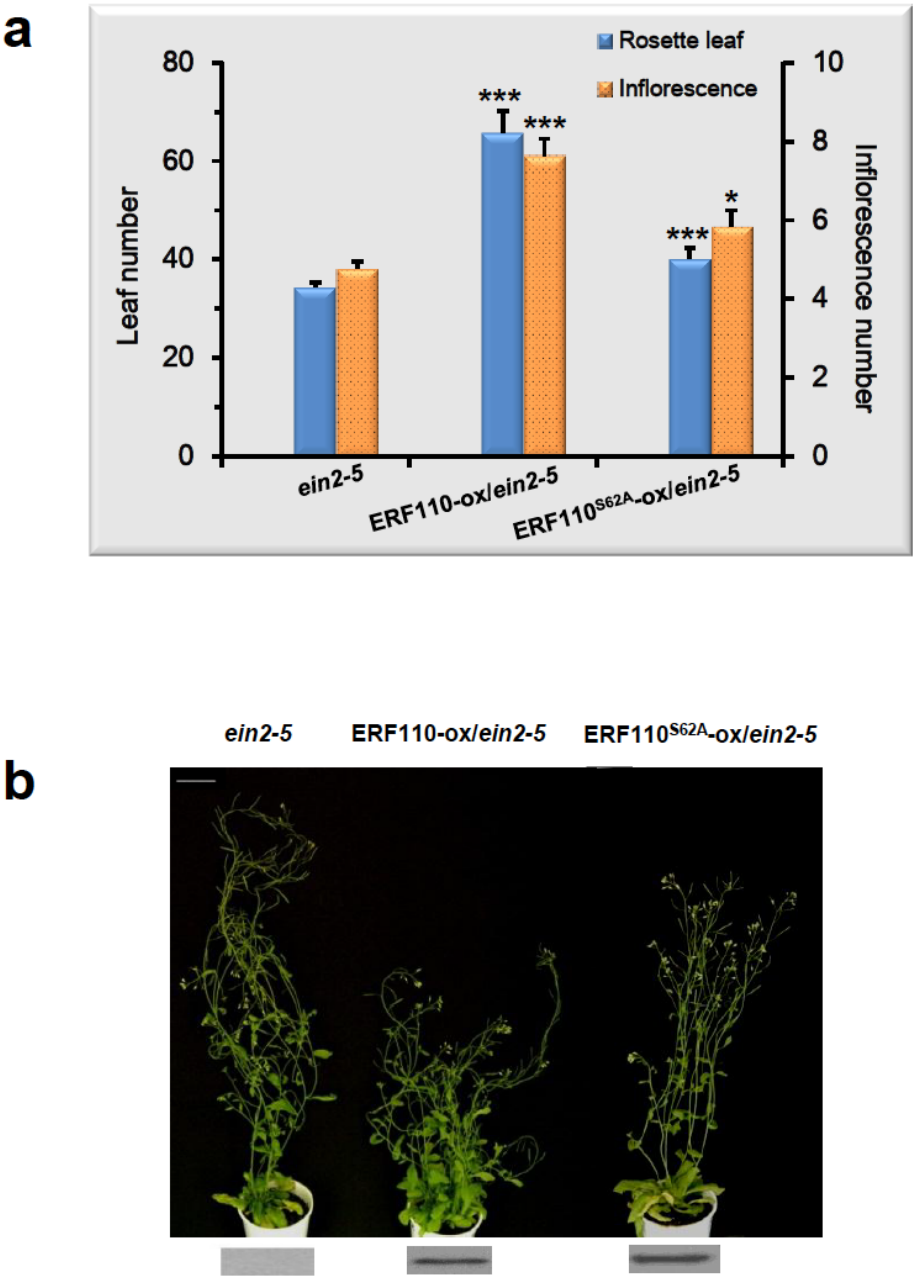
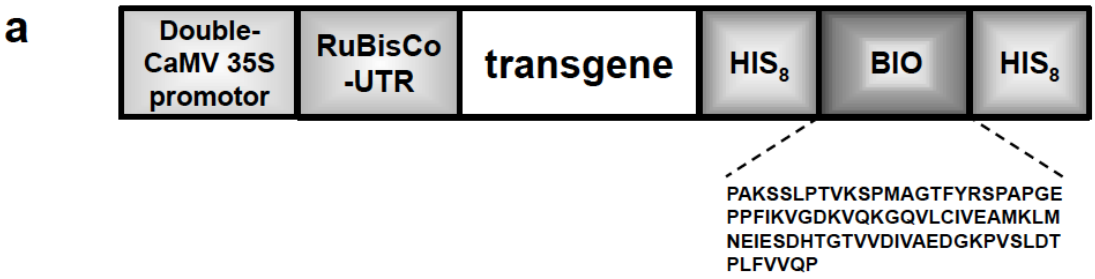
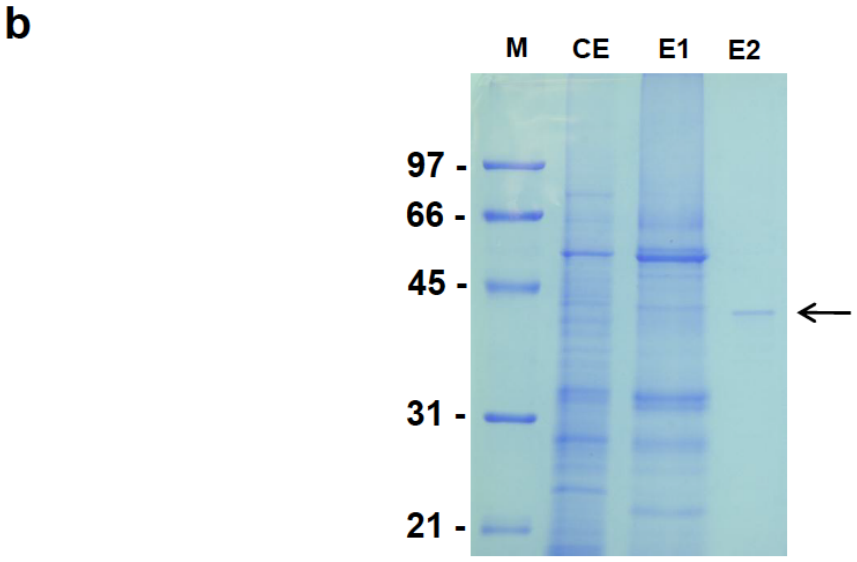

a
Rosette leaf
Inflorescence
Leaf number
Inflorescence number
80
60
40
20
0
10
8
6
4
2
0
***
***
***
*
ein2-5
ERF110-ox/ein2-5
ERF110S62A-ox/ein2-5
b
ein2-5
ERF110-ox/ein2-5
ERF110S62A-ox/ein2-5

Figure 6





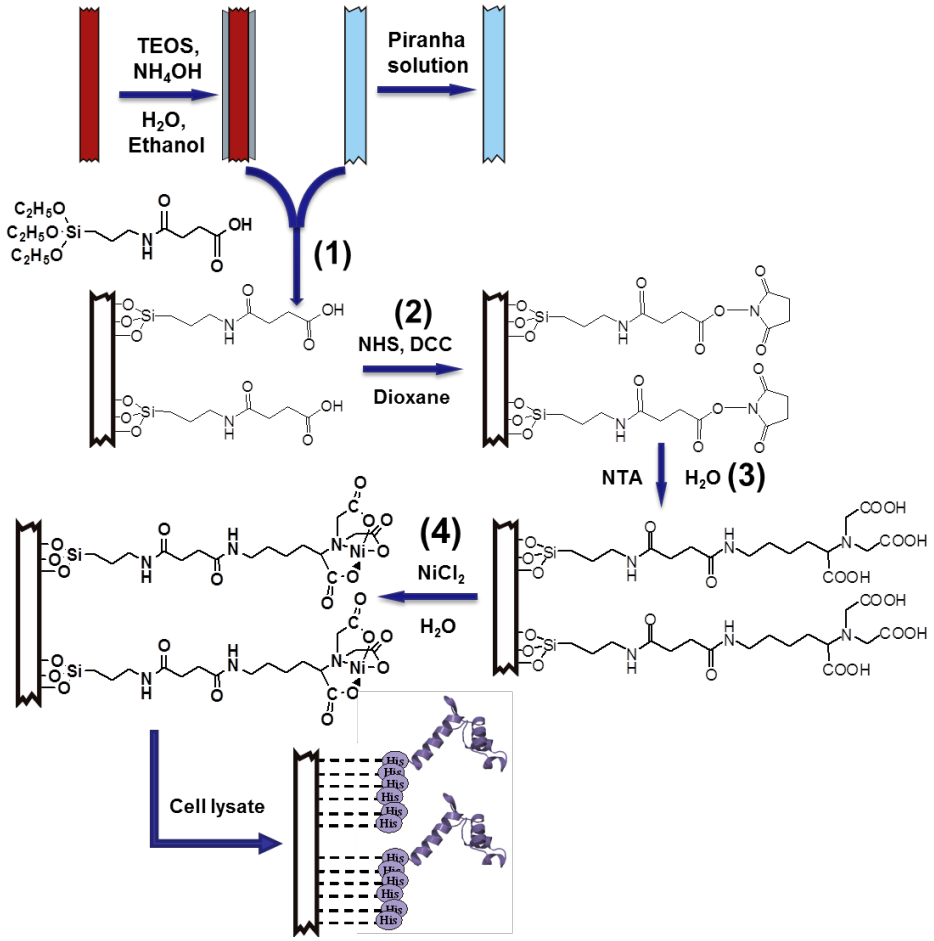
A construct of ERF110-HBH and tandem affinity purification of recombinant ERF110.

(a) Double-CaMV 35S promotor represents double cauliflower mosaic virus (CaMV) 35S promotor. RuBisCO-UTR is the 5' untranslated region of RuBisCO small subunit 1A. Transgene is the genomic ERF110 DNA containing both introns and exons. HIS_8 is an octahistidine tag that binds to Ni^{2+} chelate resins and can be specifically detected with anti-His_6 antibody. BIO represents an 83-amino acid sequence long polypeptide derived from *Arabidopsis* biotin carboxyl carrier protein (BCCP) that is biotinylated *in vivo*. Octahistidine tag and BIO polypeptide fusion protein tag is called HBH tag.

(b) The recombinant ERF11 protein is a fusion protein in which ERF110 protein has a carboxyl-terminal HBH tag and over-expressed in an ethylene-insensitive *ein2-5 Arabidopsis*. The recombinant protein was purified sequentially using Ni^{2+} chelate affinity chromatography and streptavidin affinity magnetic beads as described in "Methods". **M**, Marker; **CE**, total crude protein extract (50 µg protein were loaded); **E1,** concentrated eluate (50 µg protein were loaded) from Ni^{2+}-NTA agarose; **E2,** eluate from streptavidin affinity magnetic beads. The arrow indicates the specific band of recombinant ERF110.

Figure 7

A


TEOS,
NH4OH
H2O,
Ethanol
Piranha
solution
(1)
(2)
NHS, DCC
Dioxane
NTA
H2O (3)
(4)
NiCl2
H2O
Cell lysate

B

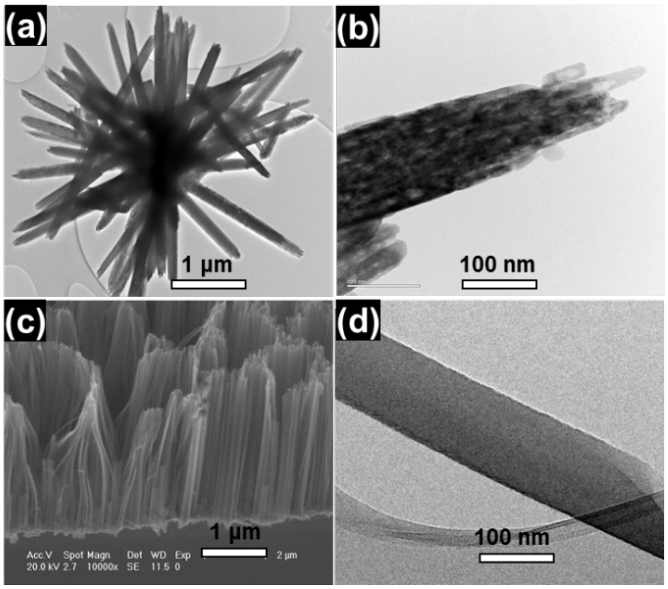
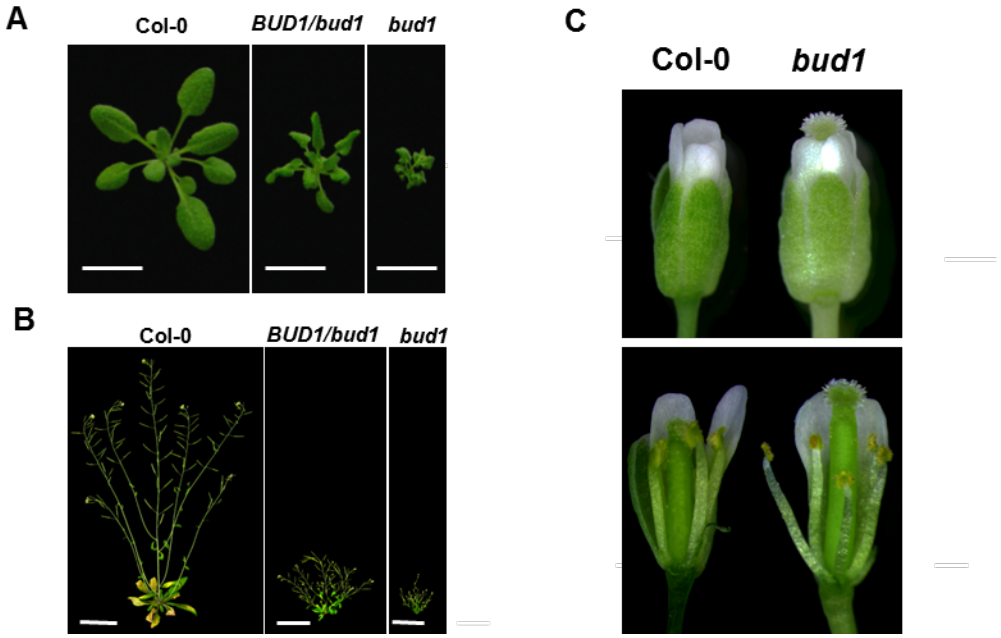

(a)
1 μm
(b)
100 nm
(c)
1 μm
(d)
100 nm

Figure 8



The phenotype of *bud1* mutant plants

A. Phenotype of 25 days plants. Bars=2 cm, B. Phenotype of 50 days plants, Bars=5 cm C. shorter stamens in *bud1* mutant