Establishment of dimethyl labelling-based quantitative acetylproteomics in Arabidopsis

Shichang Liu\textsuperscript{a,1}, Fengchao Yu\textsuperscript{b,c,1}, Zhu Yang\textsuperscript{a,d}, Tingliang Wang\textsuperscript{e}, Hairong Xiong\textsuperscript{f}, Caren Chang\textsuperscript{g}, Weichuan Yu\textsuperscript{b,c,2} and Ning Li\textsuperscript{a,d,2}

\textsuperscript{a} Division of Life Science, The Hong Kong University of Science and Technology, Hong Kong SAR, China.

\textsuperscript{b} Division of Biomedical Engineering, The Hong Kong University of Science and Technology, Hong Kong SAR, China.

\textsuperscript{c} Department of Electronic and Computer Engineering, The Hong Kong University of Science and Technology, Hong Kong SAR, China.

\textsuperscript{d} The Hong Kong University of Science and Technology, Shenzhen Research Institute, Shenzhen, Guangdong, 518057, China.

\textsuperscript{e} Tsinghua-Peking Joint Center for Life Sciences, Center for Structural Biology, School of Life Sciences and School of Medicine, Tsinghua University, Beijing 100084, China.

\textsuperscript{f} College of Life Science, South-central University for Nationalities, Wuhan, 430074, China.

\textsuperscript{g} Department of Cell Biology and Molecular Genetics, University of Maryland, Maryland, USA. 20742-5815

\textsuperscript{1} Co-first author

\textsuperscript{2} Co-Corresponding Authors: Ning Li and Weichuan Yu

Correspondence: boningli@ust.hk and eeyu@ust.hk
Address: Division of Life Science, The Hong Kong University of Science and Technology, Hong Kong, SAR, China.

Telephone number: +852-23587335

Email address: boningli@ust.hk

Email address of other authors:
Shichang Liu: sliuaf@connect.ust.hk
Fengchao Yu: fyuab@connect.ust.hk
Zhu Yang: zhuyang@ust.hk
Tingliang Wang: wangtingliang@mail.tsinghua.edu.cn
Hairong Xiong: xionghr@mail.scuec.edu.cn
Caren Chang: carenc@umd.edu

Running Title: Quantitative acetylproteomics in Arabidopsis
ABBREVIATIONS

PTMs : post-translational modifications

XIC : extracted ion chromatogram.

4C : 4 components

MS : mass spectrometry

SQUA-D: stable isotope-based quantitation-dimethyl labelling

UPAs : unique PTM peptide arrays

iTRAQ : isobaric tags for relative and absolute quantitation

TMT : tandem mass tags

Col-0 : Columbia-0

ein3/eil1 : ethylene insensitive 3/ethylene insensitive 3-like 1

AOA : aminooxyacetic acid

ACC : 1-aminocyclopropane-1-carboxylic acid

CDG : detergent-free cesium chloride (CsCl) density gradient

UEB : urea protein extraction buffer

t : top fraction

m : middle fraction

b : bottom fraction

MSPD : membrane-solubilizing and protein-denaturing buffer

F : forward mixing of peptide samples
R : reciprocal mixing of peptide samples

SCX : strong cation exchange

NCE : normalized collision energy

DDA : data-dependent acquisition

AGC : automatic gain control

FDR : false discovery rate

ANOVA : analysis of variance

SD : Standard deviation

GO : Gene Ontology

L : light dimethyl labelling

H : heavy dimethyl labelling

PSM : peptide spectra match

OAP : overly acetylated proteins

UPSPs : unique PTM site patterns

HSP : heat shock protein

HSF : heat stress transcription factor
Summary

Protein acetylation, one of many types of post-translational modifications (PTMs), is involved in a variety of biological and cellular processes. In the present study, we describe a dimethyl-labelling-based extracted ion chromatogram (XIC) quantitation protocol, which was applied to study dynamic acetylproteomic changes in Arabidopsis. The workflow integrates dimethyl labelling with chromatography-based acetylpeptide enrichment followed by mass spectrometry (MS) analysis. A software program named Stable isotope-based Quantitation-Dimethyl labelling (SQUA-D) was developed to measure dynamic changes in acetylpeptide levels. An example of hormone-regulated acetylation was confirmed by immunoblot analysis. Eventually, 7,456 unambiguous acetylation sites were found, from which 5,250 acetylation sites were identified more than once, and they locate on 2,638 different acetylproteins. Out of these repetitively discovered acetylation sites, 4,228 of them (i.e., 80.5%) were novel. These acetylproteins are exemplified by the histone superfamily, ribosomal and heat shock proteins, and those proteins related to stress/stimulus responses and energy metabolism, whereas those novel acetylproteins enriched by the cesium chloride (CsCl) density gradient centrifugation contain many cellular trafficking proteins, membrane-bound receptors and receptor-like kinases that are involved in brassinosteroid, light and gravity and development signalling pathways. In addition, we identified twelve highly
conserved acetylation site motifs within histones, P-glycoproteins, actin depolymerizing factors, ATPases, transcription factors and receptor-like kinases. SQUA-D was then used to quantify 33 hormone-enhanced and 31 hormone-suppressed unique PTM peptide arrays (UPAs). This dimethyl-labelling-based quantitative acetylproteomic protocol, in conjunction with SQUA-D, may be applied for the quantitation of any PTMs in model eukaryotes and agricultural crops and in tissue samples from animals and humans.

**Keywords:** Stable isotope dimethyl labelling, CsCl density gradient centrifugation, acetylproteomics, quantitative PTM proteomics, Arabidopsis, SQUA-D
Introduction

In vitro dimethyl labelling of peptides has emerged as one of the fastest and most dependable chemical-labelling strategies for quantitative proteomics since its introduction in 2003 (1). This approach has the advantages of high cost-effectiveness, high labelling efficiency and minimal side reactions (2-4). This unique chemical labelling method utilizes formaldehyde and sodium cyanoborohydride to react with the primary amines of the N-terminal residues of peptides as well as with lysine side chains to generate dimethylamines (1). In vitro light and heavy isotope-coded dimethyl labelling can be achieved once the total cellular proteome is proteolytically digested into smaller peptides (5). Dimethyl labelling together with label-free strategy or other types of chemical labelling, such as TMT (tandem mass tags) and iTRAQ (isobaric tags for relative and absolute quantitation), have become common choices for quantitative proteomics in cell lines, tissues and organisms that cannot be labelled via metabolic isotopic labelling (6, 7). In contrast to label-free quantitative proteomics, which normally requires a higher number of technical replicates both to ensure the consistency of sample preparation and to reduce the variability arising from LC-MS/MS analysis (3, 8, 9), dimethyl labelling may require fewer technical replicates as peptides are labelled and mixed at the beginning of peptide enrichment, which eliminates the variation derived from enrichment processes. In theory, there should be no limitations
for applying this labelling method to samples from any biological source (10).

Because of the unique advantages of dimethyl labelling in quantitative proteomics, it has been applied to study the proteomes of many unicellular and multicellular organisms, including bacteria (11, 12), human cells (1, 13), mouse and rat cells (14, 15) and plants (16-18). Dimethyl labelling approaches have further been applied to quantitatively identify phosphoproteins, the predominant class of post-translationally modified proteins, in rat and mouse cells (19, 20), zebrafish (21), and in various human cells (22, 23). Similar examples of dimethyl labelling include glycoproteomics studies in rat (24, 25) and human cells (26). The early work of Boersema et al. has combined dimethyl labelling-based quantitative proteomics and antibody-based enrichment to profile tyrosine phosphorylation in human Hela cells (27).

To further develop a workflow for high-throughput dimethyl-labelling-based quantitative PTM proteomic analysis, we took protein acetylation as a PTM proteomic study because it is widely recognized as one of the key PTMs involved in the regulatory mechanisms of many physiological processes (28-31). Many acetylation sites have been identified in histone proteins (31-34). The acetyl moieties of histones generally function as an epigenetic code that is recognized by ‘readers’, which leads to changes in chromatin structure to modulate gene transcription (35). Acetyl-CoA is believed to
be an important molecule for balancing gene transcription and metabolic pathways through protein acetylation (36). Given the importance of acetylation, thousands of acetylation sites have been previously found and the acetylation levels have been quantified through either isotopic labelling or label-free quantification strategies in bacteria (37, 38), mammalian cell lines (39-41), malaria parasites (42), mouse hearts (43), mouse livers (44-46) and human cells (47).

In plants, protein acetylation is viewed as a widespread type of PTM that mediates a diverse range of biological processes and metabolic pathways through the modification of targets such as chlorophyll binding proteins, Rubisco subunits and ATP synthase (48, 49). Histone acetylases and deacetylases are known to regulate the acetylation levels of histone proteins (50). Many studies on acetylation at the single protein level, especially those of histone acetylation in various dicots and monocots, support the importance of protein acetylation in plant growth and development (51-57). To achieve a global view of acetylproteins in plants, acetylproteomics has been performed in numerous plant species, revealing nearly 1,015 acetylproteins in rice (58-60), 245 in soybean (61), 353 in stiff brome (62), 277 in wheat (63), 684 in strawberry (64) and 1,383 in Arabidopsis (48, 65-67). These findings have greatly improved our understanding of plant acetylproteomes in general, and have provided useful information for elucidating the mechanisms underlying the regulation of biological processes by protein acetylation.
Many proteomic approaches have been proposed to quantify either peptides or proteins (68-71). Among them, the isotopic-labelling coupled with extracted ion chromatogram (XIC) has been widely used in many biological studies (68). Both SILAC (72) and dimethyl labelling-based quantitative proteomics (1) belong to this type of approach. In order to accurately analyze the data generated from these quantitative proteomics, various computer programs have also been constructed (73-75). Among them, there are several widely used software, such as MaxQuant (76), pQuant (77), XInteract (XPRESS, (78)), and Skyline (79). Their performances and robustness have been tested in many biological studies (80) except for a few of unique biological studies. Thus, more quantitative tools, such as DeMix-Q (81), PyQuant (82), moFF (83), and pyQms (84) were consequently developed. In the present study, we have also built an in-house PTM peptide quantitation software Stable isotope-based Quantitation-Dimethyl (SQUA-D) dedicated to the dimethyl labelling-based quantitative PTM proteomics. Several reasons are behind the development of SQUA-D software rather than using the common tools mentioned above directly. Firstly, these commonly used tools only quantify post-translationally modified peptides, whereas our aim was to quantify unique PTM peptide arrays (UPAs). Each UPA comprises a group of acetylpeptides with different missing cleavages but sharing the same unique PTM site pattern (UPSP). Analysing UPAs allowed us to quantify acetylpeptides in groups instead of analysing
each peptide individually. Secondly, SQUA-D introduces a batch effect adjustment function for quantification, as batch effects exist in almost all high-throughput experiments (85). Thirdly, MaxQuant does not support Mascot-based identification (86) for subsequent quantitation. Considering that Mascot is one of the most widely used proteomics tools, we aimed to integrate its identification results into our quantification workflow. Finally, SQUA-D was designed to pair unidentified yet highly confident peaks with previously identified ones given that each sample contains both light and heavy labelled ions. As a result, SQUA-D analysis revealed significantly more signal-regulated acetylation sites than previous studies. This entire workflow of dimethyl-labelling-based acetylproteomics can be summarized as 4-component (4C) quantitative PTM proteomics, comprising chemical labelling (i.e., dimethyl labelling), chromatographic enrichment followed by mass spectrometry analysis, computational analysis and confirmation of identified PTM sites.
Experimental Procedure

Experimental Design and Statistical Rationale:

Four-component (4C) quantitative PTM proteomics, comprising chemical labelling (i.e., dimethyl labelling, Figure 1A), chromatographic enrichment followed by mass spectrometry analysis, computational analysis and confirmation of identified PTM sites, were performed. The whole workflow is shown in Figure 1B as following: I. Plant preparation and treatment (including the control and hormone-treated plants); II. CsCl density gradient centrifugation fractionation of the total cellular proteins without being solubilized by detergents and peptides preparation; III. Dimethyl labelling and mixing of peptides; IV. Chromatographic separation and affinity purification of acetylpeptides, V. LC-MS/MS analysis; VI, Mascot searching-based dentification; and VII. SQUA-D-based quantification. Three robust biological replicates were performed to meet the minimum requirement for statistical analysis. In total, there were 54 LC-MS/MS injections (i.e., 3 biological replicates × 3 CsCl density gradient fractions × 2 types of mixing [F and R] × 3 SCX-HPLC fractions). We would like to find out if a given UPA was regulated by the treatment, which meant that its average log-ratio was significantly different from zero. Thus, we performed one sample t-test for the log-ratios from the same UPA. After the t-test, each UPA has a p-value. Then, we performed multiple
testing correction using the Benjamini and Hochberg procedure (87) and the output was false discovery rate (FDR).

**Plant materials, growth conditions and hormone treatment**

The wild-type Arabidopsis thaliana ecotype Columbia-0 (Col-0) seeds were purchased from the Arabidopsis Biological Resource Center (ABRC, Columbus, OH, USA). The *ein3/eil1* double mutants were gifts from Dr. Hongwei Guo of South University of Science and Technology of China. The Arabidopsis plants were cultivated on glass agar jars containing 100 μM aminooxyacetic acid (AOA, an inhibitor of 1-aminocyclopropane-1-carboxylic acid (ACC) synthase to block biosynthesis of endogenous ethylene) in accordance to previously developed methods (88). The Arabidopsis growth chamber was controlled at 22 °C ± 2 °C and had a 16-hour light/8-hour dark regime. Ethylene treatment was performed by growing Arabidopsis plants on the agar medium supplemented with 10 μM ACC. Plants harvested for protein extraction were grown for 21 days with a density of 15 plants per jar. The harvested plant tissues were frozen immediately in liquid nitrogen and stored at -140 °C.
Cesium chloride (CsCl) density gradient centrifugation-based protein fractionation

The total cellular protein was extracted from the control and treated plants, respectively, using detergent-free CsCl density gradient (CDG) buffer for protein fractionation, which contains 3 M CsCl and a modified urea protein extraction buffer (UEB, (88)). The frozen plant tissues were ground into fine powder and mixed with CDG buffer at a ratio of 1 : 4 (w/v). After centrifugation at 218,000 × g for 2 hours at 10 °C, the top (t), middle (m) and bottom (b) protein fractions were obtained. The highly membrane-concentrated top and bottom fractions were re-dissolved using 5 volumes of membrane-solubilizing and protein-denaturing buffer (MSPD), which contains 20 mM Tris-HCl, PH 7.8, 8 M urea, 10 mM EDTA (ethylene diamine tetraacetic acid), 10 mM EGTA (ethylene glycol tetraacetic acid), 50 mM NaF, 2% Glycerol, *1% Glycerol-2-Phosphate disodium salt hydrate, *1 mM PMSF (phenylmethylsulfonyl fluorid), 1% SDS (sodium dodecyl sulfate) and 1.2% Triton-X100. Proteins from each one of three fractions were precipitated and quantitated according to the previously described (89).

Protein digestion and dimethyl-labelling of peptides

Protein samples were dissolved into preheated (37 °C) trypsin digestion buffer (40mM
Tris-HCl, PH 8.0) with protease trypsin (100:1, w:w). The final concentration of urea should be lower than 1 M. The in-solution digestion was performed for 12 hours at 37 °C. The digested peptides were desalted and enriched by C18 Sep-Pak cartridges (Waters Corporation, United Kingdom). The peptides from both control and treated plants were re-suspended by 100 mM sodium acetate (pH5.5) and then were divided into two parts equally. The two parts were labelled with light isotope-coded chemicals ($^{12}$CH$_2$O and NaBH$_3$CN) and heavy isotope-coded dimethyl chemicals ($^{13}$CD$_2$O and NaBH$_3$CN), respectively (2). The mixing of an equal amount of heavy isotope-labelled peptides from treated plants with light isotope-labelled peptides from control plants was defined as the Forward mixing experiment (F), and vice versa as the Reciprocal mixing experiment (R). Both F and R samples were considered to be independent experimental replicates. In total, six experimental replicates (F1, F2, F3, R1, R2 and R3) were performed from three biological replicates. Because CDG centrifugation also generated 3 fractions (t, m and b fraction) per experimental replicate, each biological replicate produces 6 experimental peptide mixtures ((F and R mixings) x (t, m and b fraction)), which generates an equivalent of 18 mixed peptide samples from three biological replicates (Figure 1B).
HPLC (high performance liquid chromatography) fractionation and acetyl-affinity enrichment of acetylated peptides

Each of 18 peptide samples was further fractionated into 3 sub-fractions on HPLC using a 200 × 9.4 mm strong cation exchange (SCX) column (PolySULFOETHYL ATM, 5 μm, 200 Å, 209SE0502, PolyLCINC, Columbia, MD, USA) at a flow rate of 2.5 mL/min. Ultraviolet absorption was measured at 214 nm to monitor the eluted peptides. A total of 54 acetylpeptide fractions were harvested. To enrich the acetylated peptides, all 54 peptide samples were suspended in a NETN buffer, which contains 100 mM NaCl, 1 mM EDTA, 50 mM Tris-HCl, 0.5% NP-40, pH 8.0) and incubated with the beads conjugated with anti-acetyllysine antibody (PTM Biolabs, Hangzhou, China) at 4 °C overnight with gentle shaking. The resulting mixture solutions were washed 4 times using NETN buffer and 2 times with ddH2O, followed by the final elution using 0.1% TFA. The eluates were desalted using C18 ZipTip (ZTC18S960, Millipore, MA, USA). Finally, 54 enriched acetylpeptide samples were prepared and enriched, which were then subjected to LC-MS/MS analysis.

Liquid chromatography-tandem mass spectrometry (LC-MS/MS) analysis

The LC-MS/MS analysis of acetylpeptides was performed on an EASY-nLC 1000
UPLC system (Thermo Fisher Scientific, Odense, Denmark) coupled with a Q Exactive™ hybrid quadrupole-Orbitrap mass spectrometer (Thermo Fisher Scientific, Germany). The resulting acetylpeptides were loaded onto an Acclaim PepMap 100 C18 pre-column (Dionex, Sunnyvale, CA, USA) and separated in an Acclaim PepMap RSLC C18 analytical column (Dionex, Sunnyvale, CA, USA) at a constant flow rate of 300 nL/min on a gradient of 0–24 min 6–22% B, 24–32 min 22–40% B, 32–37 min 40–80% B, 37–40 min 80% B, where B is (0.1% FA in 98% ACN). Resolution was set at 70,000 for intact peptides and at 17,500 for ion fragments under a normalized collision energy (NCE) of 30. The data-dependent acquisition (DDA) mode was adopted for the top 20 precursor ions, all of which exceed a threshold ion count of 5E3 with 15s dynamic exclusion. To generate MS/MS spectra, about 5E4 ions were accumulated while automatic gain control (AGC) was used to avoid overfilling the Orbitrap. The precursor m/z scan range was set between 350 to 1800.

Mass spectrometry data analysis

The raw MS data files generated by the Thermo Q Exactive Mass Spectrometer were converted into the mzXML format using ProteoWizard (version: 3.0.11133 b4-bit) (90). The false discovery rate (FDR) was estimated using the target-decoy strategy (91).
target database which was used was the TAIR10 (35,387 proteins, https://www.arabidopsis.org/download_files/Sequences/TAIR10_blastsets/TAIR10_pep_20101214_updated) while the decoy database was generated by random shuffling the target database. We applied Mascot (Version 2.5.1, Matrix Science) for peptide identification as described previously (92). Trypsin was used as the protease to digest acetylproteins. The mass tolerance was $\pm$10 ppm for MS$^1$ and 0.02 Da for MS$^2$. The maximum missed cleavage was six. The fix modification was carboxymethyl (57.021464 Da) on Cysteine. The variable modifications were oxidation (15.994915 Da) on Methionine, light dimethyl labeling ($^{13}$C$_2$H$_4$, 28.031300 Da) on Lysine and peptide N-terminal, heavy dimethyl labeling ($^{13}$C$_2$D$_4$, 34.063117) on Lysine and peptide N-terminal, acetylation ($^{15}$N$_2$O, 42.010565) on Lysine and protein N-terminal. MascotPercolator (Version 3.1) (93) was appended to estimate the FDR. MascotPercolator converts the original FDR into the $q$-value. The peptide-spectrum match (PSM) cut-off threshold was set at $q$-value $\leq$ 0.01.

**Quantification of acetylpeptides**

In-solution protease digestion of acetylproteins may produce the peptides with missed cleavages. In order to quantitate the acetylpeptides with the common acetylation site(s),
both completely digested and partially digested acetylpeptides were combined into one peptide group (or called array) and defined as the unified PTM peptide array (UPA). For example, “n(34.063)GGK(42.011)GLGK(42.011)GGAK(34.063)Rc” and “n(34.063)GGK(42.011)GLGK(42.011)GGAK(34.063)c”, where “n” stands for peptide N-terminal, “c” stands for peptide C-terminal, “34.063” stands for heavy dimethyl labelling, and “42.011” stands for acetylation, are combined together. We built a tool named SQUA-D (Stable isotope-based Quantitation-Dimethyl labelling, Version 1.0) to perform the quality control, quantification analysis, and statistical analysis. The quality control is based on the following criteria:

1) the number of the PSMs of the light dimethyl-coded acetylpeptide is larger than or equal to one;

2) the number of the PSMs of the heavy dimethyl-coded acetylpeptide is larger than or equal to one;

3) the number of the PSMs from different experimental replicates (F1, F2, F3, R1, R2, and R3) should be larger than and equal to five;

4) The number of the identified PSMs from the forward experiments divided by the total number of PSM is larger than or equal to 0.15;

5) The number of the identified PSMs from the reciprocal experiments divided by the total number of PSM is larger than or equal to 0.15;
After the quality control, SQUA-D extracts ion chromatograms for light- and heavy-labelled peptides; pairs light- and heavy-labelled ion chromatograms; and calculates log-ratios using the maximum intensities (76) of the smoothed ion chromatograms. In order to make the most use of the data, SQUA-D also pairs the unidentified but highly confident ion chromatogram with the identified one. For example, if SQUA-D detects that an identified light-labelled peptide does not pair a heavy-labelled peptide, it calculates the theoretical $m/z$ value and the isotopic pattern of the heavy-labelled peptide. Then, SQUA-D finds the highly possible ion satisfying three criteria: the $m/z$ value equals the theoretical $m/z$ value; the Pearson correlation of the theoretical isotopic pattern and the observed isotopic pattern is larger than or equal to 0.7 (94) and there is overlap between two retention time ranges from the light- and heavy-labelled ion chromatograms. Finally, SQUA-D pairs these two ion chromatograms and calculates a log-ratio. SQUA-D also uses half of the minimum value among all extracted intensities as a replacement for the zero intensities of certain peptides. After calculating the log-ratios for the peptides from the UPAs, each log-ratio is adjusted with the median value of all the log-ratios from the same replicate to circumvent the inevitable labelling and mixing errors were subtracted. Finally, SQUA-D adjusts the batch effects and performs the statistical testing.
Statistical analysis of acetylpeptides

In our PTM proteomic workflow, there were 6 experimental replicates (i.e., F1, F2, F3, R1, R2, and R3) and 18 mixing replicates (or batches, tF, tR, mF, mR, bF, and bR for one biological replicate), which integrate three experimental factors, i.e. plant harvesting repeats, CDG fractionation types, “Forward” and “Reciprocal”. Thus, three-way analysis of variance (ANOVA) to test if there were any batch effect was applied (95). If there were, we adjusted them based on an empirical Bayes method proposed by Johnson et al. (96).

In principle, the log-ratios were firstly modeled as

\[ Y_{ijg} = \alpha_g + X\beta_g + \gamma_{ig} + \delta_{ig}\varepsilon_{ijg}, \]  

Where \( i \) indicates different batches, \( j \) indicates different log-ratios in the same batch, \( g \) indicates different UPA, \( Y_{ijg} \) is the log-ratio, \( \alpha_g \) is the mean of all log-ratios from UPA \( g \), \( X \) is the experimental design matrix, \( \beta_g \) is the coefficient for experimental conditions, \( \gamma_{ig} \) is the additive batch effect, and \( \delta_{ig}\varepsilon_{ijg} \) is the multiplicated batch effect. In our data, \( X \) is an identity of matrix because we have taken care of two experimental conditions and used log-ratio to represent them. Thus, the \( X\beta_g \) term can be ignored in the Equation (1). With Equation (1), the batch effect adjustment is achieved by adjusting \( \gamma_{ig} \) term and \( \delta_{ig}\varepsilon_{ijg} \) term.

The data was first standardized with
\[ Z_{ijg} = \frac{y_{ijg} - \hat{a}_g}{\hat{\sigma}_g^2}, \]  
(2)

Where \[ \hat{\sigma}_g^2 = \frac{1}{N} \sum_{ij} (y_{ijg} - \hat{a}_g)^2. \] Therefore \( Z_{ijg} \) was modeled with Gaussian distribution:

\[ Z_{ijg} \sim N(y_{ijg}, \hat{\sigma}_g^2). \]  
(3)

The empirical Bayes approach was applied to estimate \( y_{ig} \) and \( \hat{\sigma}_g^2 \):

\[ y_{ig} \sim N(\psi_i, \tau_i^2), \]  
(4)

\[ \hat{\sigma}_g^2 \sim InvGamma(\lambda_i, \theta_i). \]  
(5)

After estimating the parameters, the log-ratios were adjusted with

\[ y_{ijg}^* = \frac{\hat{\sigma}_g}{\hat{\sigma}_g^2} (Z_{ijg} - \hat{y}_{ig}) + \hat{a}_g. \]  
(6)

At the end, one-sample \( t \)-test was employed to test if a given UPA’s log-ratios had a mean significantly different from zero. After \( t \)-test, multiple test correction based on the Benjamini–Hochberg procedure was performed and resulted in the outcome of BH-FDR (87). BH-FDR is different from the target-decoy-based FDR for selection of acceptable PSM. The BH-FDR threshold was set to 0.1. Standard deviation (SD) of all UPA’s log-ratio mean values were calculated afterwards. The UPA with log-ratio mean values outside \( 0.5 \times \text{SD} \) and BH-FDR smaller than or equal to 0.1, were selected as the significant findings (see Supplementary Table S3). The flowchart of computational programs of Mascot identification coupled with extracted ion chromatogram (XIC) based quantification is shown Supplementary Figure S1.
Bioinformatics analysis

The conserved acetylation site motif analysis was performed using the Motif-X software (97). Molecular functions enrichment, Cellular components enrichment and Biological processes enrichment were conducted online by the Database for Annotation, Visualization and Integrated Discovery (DAVID) v6.7 (https://david-d.ncifcrf.gov/tools.jsp) with the following thresholds: Gene count ≥ 5, P-value ≤ 0.01, FDR Q-value ≤ 0.1 and then was drawn by software Cytoscape 3.3.0. The databases for Gene Ontology (GO) analysis was obtained from the website of the Arabidopsis information resource (TAIR). Comparison of the GO analysis results between the acetylproteins and the leaf proteome of Arabidopsis was performed using the following equation (92):

\[
R_i = \log_2 \left( \frac{N_i}{N} \right) \frac{N_i'}{N'}
\]

where \( N \) and \( N' \) represents the total matching number of all categories of the identified acetylproteins and that of the proteins from leaf proteome of Arabidopsis (98), respectively, whereas \( N_i \) and \( N_i' \) is the matching number of \( i \)-th category of the identified acetylproteins and that of the proteins from the leaf proteome of Arabidopsis (98), respectively. Comparison of the GO analysis results between the acetylproteins
and phosphoproteins (Full datasets of PhosPhAt 4.0, (99)) was performed using the same equation.

Immunoblot analysis

Proteins were extracted from plant tissues using the UEB described previously (89). Immunoblot analysis was performed using the following antibodies: (1) Anti-Human acetyl-histone H4(K5) monoclonal antibody (PTM-163), which targets both the K5 acetylation site of SGRGacKGGKGLGK and the K6 acetylation site of Arabidopsis histone superfamily protein (AT1G07660); (2) the Anti-actin (plant) monoclonal antibodies (a0480, Sigma Aldrich, St Louis, MO, USA). The first antibody was purchased from the PTM Biolabs, Hangzhou, China. 50 μg of the resulting extracted proteins were loaded onto 12% SDS-PAGE gels before being immobilized onto a PVDF membrane (GE Healthcare) and then probed with the antibodies mentioned above. $t$-test was used to evaluate the significance between control and treatment.
Results

**LC-MS/MS and computational analysis of Arabidopsis acetylproteomes**

A comprehensive dimethyl-labelling-based PTM proteomics workflow was performed on Arabidopsis (Figure 1A, 1B, I – V, the first and second C procedures), consisting of plant treatment (I); fractionation of proteins by detergent-free CsCl density gradient centrifugation (II); trypsin preparation, dimethyl labelling and mixing of peptides (III); and chromatographic separation and affinity purification of acetylpeptides (IV) followed by LC-MS/MS analysis (V). In a given biological experiment, the total cell lysate isolated from control or the treated plants (approximately 60 g tissues, respectively), was separated into top (t), middle (m) and bottom (b) fractions, which fractionated approximately 120 mg, 360 mg and 40 mg of proteins, respectively, on a CsCl density gradient (Figure 1B). As a result, a total of six protein fractions, *i.e.* three fractions from each of control and the treated plants. Peptides were subsequently produced by tryptic digestion of the six cellular protein fractions, which produced approximately 80 mg, 270 mg, and 30 mg of peptides from t, m, and b fraction, respectively. From the generated peptides of each fraction, 80 mg, 90 and 30 mg of peptides from t, m and b fraction were used for dimethyl labelling and subsequent mixing. In the three forward (F) peptide sample mixes, three peptide samples from control and three from treated plants were chemically labelled with light (L) and heavy
(H) stable isotope-coded dimethyl chemicals (see Experimental Procedure for details), respectively. Subsequently, three reciprocal (R) mixes of isotope-labelled peptides were generated in a similar fashion. Next, each of the six technically mixed peptide samples (tF1, mF1, bF1, tR1, mR1, bR1, where “1” represents the first biological replicate) was further fractionated into three sub-fractions by SCX-HPLC, resulting in a total of 18 peptide fractions (2 to 15 mg peptides were collected from each fraction), all of which were considered to have come from a single biological replicate. With two additional biological replicates, a total of 54 L/H peptide samples were produced. Each of these 54 peptide samples (1 to 1.5 mg of peptides per fraction) was highly enriched for acetylpeptides firstly using the acetyl moiety-specific antibodies (see Experimental Procedure). The consequent acetylpeptides were subsequently subjected to LC-MS/MS analysis. Using an FDR cut-off threshold of 1%, a total of 16,503 non-redundant (unique) light (L) and heavy (H) stable isotope-coded acetylpeptides (unique PSMs) were identified from 105,958 redundant (L and H) acetylpeptides, corresponding to 4,480 acetylproteins (Supplementary Table S1a). The computational analysis (the third C) procedure began with Mascot-based database searching (Supplementary Figure S1). Using a Mascot delta score ≥ 10 to process the PSM data set (Supplementary Table S1b) as previously described (100), a total of 63,490 redundant acetylpeptides were identified from the mass spectrometry analysis (Supplementary Table S1b;
Supplementary Figure S1), out of which 60,077 repeatable acetylpeptides were detected (i.e., with ≥ 2 PSMs; Supplementary Table S1c). Of these repeatedly identified acetylpeptides, 5,246 (i.e., 8.3% of 60,077) were derived from histone proteins (Supplementary Table S1c), indicating that the histone superfamily is one of the most abundant acetylprotein families in Arabidopsis.

Further analysis of the 63,490 redundant acetylpeptides (Supplementary Table S1b) revealed 5,399 and 5,434 light and heavy isotope-coded non-redundant acetylpeptides, respectively (Figure 2A), with a ratio of 1.0065, suggesting a successful non-biased chemical labelling of cellular peptides between the control and treated plants. To summarize the number of unique acetylpeptides produced from each experimental replicate (either a forward or a reciprocal mix of L/H peptides from a single biological replicate), we combined all peptides identified from tF1, mF1, and bF1 peptide mixes into one experimental replicate named F1. Similarly, tR1, mR1, and bR1 were combined into experimental replicate R1. The number of unique acetylpeptides identified increased when all six experimental replicates were combined sequentially (Figure 2B). In total, 6,734 unique acetylpeptides (either L or H acetylpeptides ≥ 1) were identified from these 6 experimental replicates (F1, R1, F2, R2, F3 and R3, Figure 2B), among which 4,540 unique acetylpeptides were identified repeatedly (i.e., detected at least twice, L + H ≥ 2; L + L ≥ 2; or H + H ≥ 2, Supplementary Table S1d).
Bioinformatic analysis performed on these 4,540 unique and repeatable acetylpeptides revealed 5,250 acetylation sites (PTM sites, Supplementary Table S1d) and 12 highly conserved motifs surrounding these acetylation sites, each of which was associated with either a single or double acetylation site (Figure 2C and Supplementary Figure S2). The double acetylation motifs AEKKPAEK, GKGKKGL, and PKAGKKLP were identified in histone proteins and are shown in Figure 2C, indicating that the acetylation motifs of histone proteins are conserved. Moreover, we identified the acetylation motif (Q/N)LSGGQKQR(I/V) in P-glycoproteins, VAVKRL in a number of kinases, F(L/M)ELKA(T/K)KR in actin depolymerizing factor, (Q/E)AKRRAE(I/V) in ATPases, KRCKEK(W/F)EN in several transcription factors, along with other motifs not identified in special protein superfamilies, such as A(A/T/Y)EKAK, EK(I/F/M)KEK, EEK(Q/A/L/Y)KL and EE_K(K/R) (Supplementary Figure S2). The presence of acetylation motifs in several transcription factors (Supplementary Figure S2G) indicates that acetylation may affect their functions, which has been demonstrated previously (101). Particularly interesting are the acetylation motifs of certain kinases (Supplementary Figure S2B), which illustrate the crosstalk between acetylation and phosphorylation that has been previously reported (37). How acetylation and phosphorylation mutually affect each other on modification of PTM proteins in Arabidopsis remains to be elucidated.
The 4,540 unique and repeatable acetylpeptides contain a total of 5,250 acetylation sites (i.e., 17 from protein N-termini and others from lysine side chains) distributed on 2,638 acetylproteins (Supplementary Table S1d). Among these 5,250 acetylation sites, 4,228 lysine acetylation sites are novel (Figure 2D) according to previous publications (48, 65-67). Comparison of GO analysis between the proteins of novel acetylation sites and those of previously identified acetylation sites (Supplementary Figure S3) indicated that proteins of novel acetylation sites were enriched in cellular components ER, Golgi apparatus, and plasma membrane and molecular functions receptor binding or activity and kinase activity (Supplementary Figure S3). Fifty-nine plasma membrane proteins of novel acetylation sites have kinase activity (Supplementary Table S4), eight out of which also have receptor binding or activity (Supplementary Table S4), including Barely any meristem 1 (BAM1, AT5G65700), Barely any meristem 2 (BAM2, AT3G49670), BAK1-interacting receptor-like kinase (BIR1, AT5G48380), BRI1-associated receptor kinase (BAK1, AT4G33430), Phototropin 1 (PHOT1, AT3G45780), Phytochrome B (PHYB, AT2G18790), Somatic embryogenesis receptor-like kinase 1 (SERK1, AT1G71830), and Somatic embryogenesis receptor-like kinase 2 (SERK2, AT1G34210). It is surprising to find that all these receptor-like kinases BAM1, BAM2 (102), BIR1 (103), BAK1(104), PHOT1 (105, 106), PHYB (107, 108), and SERK1/2 (109) and other 3 kinases Brassinosteroid-signalling kinase 1 (BSK1, AT4G35230)
(110), BSK3 (AT4G00710) (111), and HERCULES receptor kinase 1 (HERK1, AT3G46290) (112) as well as several cellular trafficking proteins were found to be acetylated from the CsCl density gradient centrifugation method-based PTM proteomics. These proteins are well-known to be involved in hormone (e.g., brassinosteroid), light, gravity and development stage signalling pathways.

Moreover, over half (56.6%) of the 2,638 acetylproteins contain a single acetylation site, whereas 1.3% of them contain more than 8 acetylation sites (Figure 2E). These acetylproteins with multiple acetylation sites were termed overly acetylated proteins (OAPs). Both molecular function enrichment and cellular component enrichment of these OAPs showed that they are distributed in the cytosol, chloroplast, plasma membrane, nucleus, vacuole and ribosome, and many have binding activities (Supplementary Figure S4). Biological process enrichment also showed that the OAPs are involved in histone protein-based complex organization and assembly, translation, protein folding and responses to stress stimuli, such metal ion concentrations and temperature (Figure 2F). These observations suggest that the activities of these proteins within the above mentioned biological processes may be regulated by acetylation.
Quantitative analysis of the acetylproteomes of hormone-treated Arabidopsis

From the 4,540 unique and repeatable acetylpeptides, we selected 1,288 acetylpeptides for peptide quantitation according to criteria described previously (92); Supplementary Table S2a; see Experimental Procedure for selection criteria). These quantifiable acetylpeptides were then converted into 1,155 unique PTM peptide arrays (UPAs) for quantitation (Supplementary Table S2b). Because multiple protein isoforms of a gene family, or highly conserved protein domains within a family of proteins, may contain the identical unique PTM site pattern (UPSP), and because incomplete tryptic digestion may produce many unique acetylpeptides that share the same UPSP, all acetylpeptides with the same UPSP were combined to form a UPA for quantitation (Supplementary Table S2b).

The batch effect adjustment (96) was performed for the 18 mixed acetylpeptide samples (MS data collected from three SCX-HPLC sub-fractions were combined into a single batch of acetylpeptide samples) produced from three biological replicates (Figure 1B, each biological replicate produces 6 mixed-peptide technical replicates; for example: tF1, mF1, bF1, tR1, mR1, bR1, where “1” represents the first biological replicate). The batch effect adjustment aimed to eliminate technical errors caused by variations in plant harvesting, protein fractionation, peptide preparation, chemical labelling and mixing (Supplementary Figure S1; see Experimental Procedure for details). As a result, 33
hormone-enhanced and 31 hormone-suppressed UPAs were identified by our dimethyl-
labelling-based quantitative PTM proteomics method (Supplementary Table S3;
Supplementary Figure S5). These hormone-regulated UPAs each had fold changes of ≥
1.23 or ≤ -1.23 (where the negative sign indicates down-regulation) following a
multiple hypothesis test correction (HB-FDR ≤ 0.1; (87); Figure 3A and Supplementary
Table S3).

In this quantitative acetylproteomic analysis, we found that the histone protein
AT1G07660 was the most abundant acetylprotein, accounting for 2,606 redundant
acetylpeptides, or 49.7% of all histone acetylpeptides discovered. The acetylation UPA
AT1G07660-K6-K9-K13 (orange circle, Figure 3A) has three acetylation sites: K6, K9
and K13. Both its MS and MS/MS spectra, as shown in Figure 3B and 3C, demonstrate
that UPA AT1G07660-K6-K9-K13 is hormone suppressed, whereas another acetylation
UPA, AT1G07660-K6-K13-K17, which originated from the same protein and shares
the acetylation sites K6 and K13, is hormone enhanced (Supplementary Table S3 and
Supplementary Figure S6).

An HSP 70 family protein (AT3G09440) has a hormone-enhanced acetylation UPA,
K74, and two hormone regulation-independent acetylation UPAs, K252 and K530
(Figure 3D). The translational elongation factor EMB2726 (AT4G29060) has a
hormone-enhanced acetylation UPA, K357, and three hormone-independent acetylation
UPAs, K129, K198 and K889 (Figure 3E). A histone superfamily protein (AT2G28720)
has a hormone-suppressed double acetylation UPA, K39-K40, while the same protein
has another five acetylation UPAs, K12-K13-K17, K28-K33, K28-K33-K34, K70 and
K135, which are hormone independent (Figure 3F). Another heat shock protein, HSP
90-7 (AT4G24190), has a hormone-suppressed acetylation UPA, K559, and two
hormone-independent acetylation UPAs, K520 and K546 (Figure 3G). The ADP/ATP
carrier 1 (AAC1, AT3G08580) has two hormone-suppressed acetylation UPAs, K325
and K336, while its other three acetylation UPAs, K241, K323 and K341, are
independent of hormone regulation (Figure 3H). The most interesting finding was the
ATP synthase subunit beta (ATCG00480), which has a hormone-enhanced acetylation
UPA, K50; a hormone-suppressed acetylation UPA, K134; and two hormone-
independent acetylation UPAs, K86 and K495 (Figure 3I). These findings show that our
approach can quantitatively identify up- or down-regulated acetylation UPAs derived
from acetylproteins.
Immunoblot validation of a hormone-regulated unique acetylation site in histone proteins

To validate the results from XIC-based quantitative proteomics (the fourth C procedure, confirmation of hormone-regulated acetylation sites), immunoblot analysis was performed using AT1G07660-K6 acetylation-specific antibodies. The level of AT1G07660-K6 acetylation was reduced -2.46 and -3.64 fold upon long-term hormone treatment in Col-0 and in the ethylene-insensitive mutant ein3/eil1, respectively (Figures 3J). The acetylation site motif SGRGKGG (Figure 3K), shared by eight histone isoforms of the same superfamily, was predicted to be ethylene suppressed. As shown by gas chromatography (GC), the ethylene production rate of ACC-treated wild-type Arabidopsis was 19.5 times higher than that of the control plants (Supplementary Figure S7). These immunoblot results confirmed the results obtained from the dimethyl-labelling-based quantitative PTM proteomics method.

Bioinformatic analysis of acetylproteins

GO analysis, comparing the identified acetylproteins with the leaf proteome of Arabidopsis, showed that the acetylproteins were concentrated in the following categories: molecular function A1 - A3 (i.e., structural molecule activity, DNA or RNA
binding, and transporter activity), biological process A4 - A6 (i.e., electron transport or energy pathways, response to abiotic or biotic stimulus, and response to stress), and cellular component A7 - A9 (i.e., ribosome, cell wall, and plastid; Figure 4A and Supplementary Figure S8). These findings indicate that acetylation occurs on proteins with specialized molecular functions, involved in unique biological processes and located within specific cellular components.

It was interesting to find that 55.14% of the 2,628 repetitively identified acetylproteins (i.e., 1,449 acetylproteins) have at least one phosphorylation site according to the database PhosPhAt 4.0 (Supplementary Figure S9A), suggesting that other type of PTMs may be tightly associated with the acetylated proteins. How acetylation and phosphorylation PTMs influence each other is an intriguing question, which remains to be addressed in the future work. GO analysis of the identified acetylproteins and the phosphoproteins showed that the acetylproteins are concentrated in the following molecular functions: structural molecule activity, other enzyme activity, and transporter activity, biological processes: electron transport or energy pathways, response to abiotic or biotic stimulus, and response to stress, and cellular components: ribosome, cell wall, and plastid (Supplementary Figure S9). Relative to the acetylproteins, the phosphoproteins are concentrated in molecular functions: nucleic acid binding, kinase activity, and transcription factor activity, biological processes: signal transduction,
DNA or RNA metabolism, and DNA-dependent transcription, and cellular components: Golgi apparatus, plasma membrane, and nucleus (Supplementary Figure S9). The asymmetric distribution of specific PTMs among cellular proteins and different organelles may suggest that protein acetylation plays unique roles in cellular events.

One interesting observation from the biological process enrichment is that ethylene-regulated acetylation was significantly enriched in the processes of histone protein-based complex organization and assembly, photosynthesis, and responses to various stimuli and stresses (e.g., metal ions, radiation, light, osmotic stress, hormone and bacteria, Figure 4B). Through analysis of protein-protein interactions, we found that the predominant ethylene-regulated acetylproteins were histone proteins, ribosomal proteins, and heat shock proteins, which may collectively control transcription, translation, protein folding, and other biological processes, including responses to stimuli and stresses, photosynthesis, and ATP synthesis and transport (Supplementary Figure S10).
Discussion

The results of our comprehensive and quantitative PTM proteomics workflow have substantially expanded the database of Arabidopsis protein acetylation sites, revealing 4,245 acetylation sites not described in previous publications (48, 65-67). To compare the acetylproteins found from our CsCl density gradient fractions (i.e., t, m, and b fractions) with those found by Hartl et al. (67), we performed cellular component analysis (Supplementary Figure S11). Acetylproteins of the ER, Golgi apparatus, plasma membrane and mitochondria are found to be concentrated in the fraction t and b, whereas acetylproteins of the cytosol, nucleus and ribosome are concentrated in the fraction m (Supplementary Figure S11). The cellular locations of acetylproteins from the fraction m were similar to those identified by Hartl et al., especially in the Golgi apparatus, mitochondria, nucleus, and ribosome (Supplementary Figure S11). The ratio of proteins from the t, m, and b fractions was approximately 3 : 9 :1. It is likely that the abundant acetylpeptides from the fraction m would have masked the acetylpeptides from the other two fractions in LC-MS/MS analysis if we would not have separated acetylproteins from the fraction t and b. Thus, the enrichment of membrane proteins using the CsCl density gradient centrifugation may be the reason why we have identified more acetylation sites. Comparison of GO analysis between the proteins of novel acetylation sites and those proteins of previously identified acetylation sites
(Supplementary Figure S3) also supports our conclusion. The acetylproteins of ER, Golgi apparatus, and plasma membrane were mostly enriched for their novel acetylation sites (Supplementary Figure S3), which is consistent with the concentrated acetylproteins present in the fraction t and b of the CsCl density gradient (Supplementary Figure S11). Another reason why we have successfully identified a significantly larger number of acetylpeptides may result from the use of urea-based protein-denaturing buffer (UEB) during the initial cell lysis, which immediately inactivates proteases and deacetylases or any other protein modification enzymes to prevent in vitro degradation or deacetylation or modification of acetylproteins (88).

To find relatively low-abundance post-translationally modified peptides from peptide digests of cell or tissue lysates, enrichment by charge properties, chemical reactions, antibody recognition or other affinity-based binding methods are frequently included in proteomics workflows (113). For example, both charge property-based immobilized metal affinity chromatography (IMAC) and titanium dioxide (TiO$_2$) have been successfully applied to phosphoproteomics (114). Other strategies, including lectin affinity binding, boronic acid chemistry and hydrazide chemistry, have been used to purify glucopeptides (115). Antibody-based immunoprecipitation has been used extensively to identify PTM sites for ubiquitination (116), phosphorylation (117), methylation (118), acetylation (45), β-hydroxybutyrylation (119), glutarylation (120),
In the present study, we further improved the steps of protein isolation and peptide enrichment (Figure 1B) and applied a four-dimensional chromatography-based workflow, comprising CsCl density gradient centrifugation, SCX-HPLC, antibody affinity purification and C18-HPLC before MS/MS analysis. This new workflow may allow us to further enrich acetylpeptides derived from cellular membrane systems.

Based on Mascot search results, the proportion of acetylpeptides in our six experimental replicates, F1, R1, F2, R2, F3, and R3, was 92.71%, 96.64%, 86.90%, 88.60%, 95.28% and 96.46%, respectively. This high purification efficiency makes the results of our multi-dimensional acetylpeptide enrichment workflow highly persuasive. This novel procedure may be further applied to the study of other PTM proteomes.

To establish a versatile quantitative proteomics approach for field crops, animals or even human tissues, rather than just medium-grown Arabidopsis or cell lines, chemical labelling and label-free quantification methods are preferred over metabolic labelling. However, label-free-based quantification relies on highly consistent sample preparation and analysis (3, 8, 9). It is difficult to achieve high consistency in high-throughput PTM proteomics since multiple chromatography steps are usually involved to enrich for post-translationally modified peptides. Using chemical labelling for proteomics involves MS$^1$-based quantification (i.e., dimethyl labelling) and MS$^2$-based quantification (i.e., TMT and iTRAQ). Unlike MS$^1$-based quantification using XIC, MS$^2$-based
quantification measures a single MS² spectrum (2). For quantitative proteomics based
on chemical labels, especially quantitative PTM proteomics, it is better to perform the
labelling and mixing early in the protocol to reduce variance. Relative to dimethyl
labelling, TMT and iTRAQ cost much more if used in early steps (2). Ratio adjustment
of the ion intensities between light and heavy isotope-coded acetylpeptides can solve
the problem of incomplete dimethyl labelling and biased mixing. Moreover, although
different experimental replicates may be performed under well-controlled conditions,
there are still effects that are difficult to eliminate, such as variation in room
temperatures and various experimental biases. These effects are artificially introduced
into the data and may bias any conclusions. Batch effects exist in most high-throughput
data (85). Unfortunately, few studies have adjusted for batch effects when analysing
proteomics data. In our data, we found three types of effects (see Experimental
Procedure). Thus, we used three-way ANOVA to show that there were indeed batch
effects in our data (p-value = 2.08×10⁻⁷ for the effect of the CDG fractionation; p-value
= 0.00246 for the effect of the plant harvesting replicates). We adjusted for these effects
using an empirical Bayesian approach (96). Such an adjustment approach has been
widely used in previous microarray data. Before the adjustment, we identified only
eight UPAs with BH-FDR ≤ 0.1. After the adjustment, we found a total of 64 UPAs
with BH-FDR ≤ 0.1. The follow-up validation showed that those findings were true
positives. Thus, in this paper, we have demonstrated that there are batch effects in proteomics data, and adjusting for those effects will result in more findings.

Previous studies have found that various organisms utilize specialized PTMs for ribosomal proteins (121-123). These PTMs help guide nuclear events, expand molecular structures and facilitate activity regulation (124-129). Yang et al. (127) found that mitochondrial protein synthesis is enhanced by the reversible acetylation of mitochondrial ribosomal protein L10. N-terminal acetylation of ribosomal proteins was previously demonstrated to be necessary for the maintenance of protein synthesis (128). An important finding of the present quantitative acetylproteomics study was the identification of four hormone-regulated UPAs in ribosomal proteins. Both the assembly of ribosomes and translation activity may be controlled via acetylation events at these sites. We thus hypothesize that acetylation of ribosomal proteins may be an additional molecular mechanism that regulates protein translation, regardless of its positive or negative impact.

In addition to ribosomal proteins, HSPs, known as stress proteins, are responsible for protein folding, translocation, disaggregation and degradation of damaged proteins (130-132). In plants, the transcription of HSPs is regulated by at least 21 heat stress transcription factors (HSFs) (133). Previous studies have found that PTMs influence
the ATP binding and chaperone activity of HSP 90 (134, 135). Point mutation K294Q in the HSP 90α protein was used to mimic constitutive acetylation, reducing protein-protein interactions with some of its target proteins. Additionally, the K294R isoform of HSP 90α was used to mimic constitutive deacetylation, which resulted in stronger interactions with some of its target proteins (136). Our quantitative acetylproteomics results show that the acetylation of an HSP 70 family protein was hormone suppressed while that of an HSP 90 family protein was hormone enhanced (Supplementary Table S3). These results indicate that the plant hormone ethylene, or other signals, may regulate HSP-dependent protein-folding and stress responses by regulating protein acetylation. This finding may represent a major discovery of the prevalence of protein acetylation in regulating ethylene-dependent stress responses.

In addition, our acetylproteomics approach identified 44 acetylated histone proteins, and 8 histone UPAs were regulated by ethylene. Quantitative PTM proteomics revealed that the acetylation UPA AT1G07660-K6-K9-K13 is ethylene suppressed while the acetylation UPA AT1G07660-K6-K13-K17, from the same protein, is ethylene enhanced (Supplementary Table S3). The ion intensities of acetylpeptide AT1G07660-K6-K9-K13 were much more abundant than those of AT1G07660-K6-K13-K17 (Supplementary Figure S6), indicating that acetylpeptide AT1G07660-K6-K9-K13 may contribute more to the overall acetylation level of K6 and K13, since K6 and K13 are
common acetylation sites between these two acetylation UPAs. Immunoblotting analysis was performed on the K6 acetylation site of histone protein AT1G07660 (Figure 3J), showing that this acetylation site is ethylene suppressed in both Col-0 and ein3/eil1. This immunoblot-validated, ethylene-suppressed acetylation site of histone protein AT1G07660 is homologous to the K5 site of human and yeast histone H4. Given that histone H3 of Arabidopsis and yeast are homologous, Zhang et al (137) used acetylation site-specific antibodies from yeast to study the acetylation of K14 and K23 of Arabidopsis H3. They found that the acetylation of these sites was enhanced by ethylene in etiolated Col-0 seedlings after 4 hours of ethylene treatment. Assuming that the acetylation of histone proteins facilitates transcription by acting as transcription factor recognition sites and by providing space for the transcription process (50), these new findings suggest that the ethylene-regulated acetylation of histone proteins may contribute to ethylene-induced gene transcription.

GO analysis of the identified acetylproteins and the phosphoproteins revealed that phosphoproteins were concentrated in kinase-dependent signal transduction and transcription factor-dependent transcription in the plasma membrane and nucleus (Supplementary Figure S9). This bioinformatic result is consistent with the involvement of protein phosphorylation and dephosphorylation in ethylene signalling (138-142). Changes in protein phosphorylation have been demonstrated to directly regulate the
activities of transcription factors (142, 143). In contrast, the identified acetylproteins are largely involved in the regulation of enzyme activities (Supplementary Figure S9), thus affecting energy metabolism (144-147). Acetylation of both histone and non-histone proteins is mainly involved in stress responses (Supplementary Figure S9). For example, histone deacetylase 6 (HDA6) regulates responses to salt stress (148). Drought stress-induced deacetylation of protein N-termini is regulated by the plant hormone abscisic acid (149). In addition, both protein acetylation and phosphorylation may mutually affect each other (Supplementary Figure S9A), as previously reported (150, 151) even though the cause and effect relationship is not known at this time.

In summary, the combination of detergent-free CsCl density gradient centrifugation fractionation of the total cellular protein with the dimethyl labelling, XIC-based PTM peptide quantitation and the in-house built SQUA-D software analysis allowed us to identify 5,250 acetylation sites and 64 hormone-regulated acetylpeptide UPAs, (this number of UPAs may increase if SQUA-D is later improved for its efficiency and versatility in handling the results from other PSM identification computational programs like SEQUEST, MaxQuant, Skyline, pFind, Comet, Crux, and so on), which share the identical UPSP. These hormone-regulated changes in the acetylation UPAs may result from either alteration of the acetylation level or from alteration of the protein level itself (141). To differentiate these two possibilities, we have also analysed the
total cellular protein peptides (Supplementary Table S5) and found that at least 7 hormone-regulated acetylation UPAs (7 out of 15 regulated acetylation UPAs with quantified unmodified peptides, supplemental Table S5) have the same level of unmodified peptides in between both control and hormone-treated protein samples or opposite regulation trends in between acetylated and unmodified peptides, suggesting that it is the acetylation level changes in response to hormone treatment instead of protein level alteration in plant cells. For example, in the case of the histone superfamily protein AT1G07660, the K6 acetylation level of this protein is ethylene-suppressed according to both quantitative PTM proteomics (Figure 3A and 3B) and immunoblot assay analysis (Figure 3J). However, its protein level is not significantly changed by ethylene treatment (Supplementary Table S5).
Author Contributions

S. Liu performed growth and treatment of plants, conducted protein isolation and peptide preparation and proteomics-related experiments; F. Yu, Z. Yang, T. Wang and H. Xiong performed preliminary MS/MS analysis; F. Yu, N. Li and W. Yu performed extensive MS/MS data analysis; S. Liu performed immunoblot analysis; N. Li, S. Liu, F. Yu, H. Xiong and C. Chang wrote the manuscript; N. Li was involved in project planning, experimental design, project execution, supervision of experiments, and communication with collaborators and responsible for distribution of materials integral to the findings.
Acknowledgments

This research was supported by grants 31370315, 31570187 (National Science Foundation of China), 661613, 16101114, 16103615, 16103817, AoE/M-403/16 (RGC of Hong Kong), SRFI1I1EG17AG-A, SRFI1I1EG17-A (the HKUST internal support from the Energy Institute of HKUST) and SBI09/10.EG01-A (the Croucher Foundation CAS-HKUST Joint Laboratory matching fund). We thank Mr. Shuaijian Dai and Miss Nan Yang for ethylene measurement. We thank Mr. Yung Sheung Tang for English writing improvement.
Data Availability

The mass spectrometry proteomics data has been deposited to the ProteomeXchange (152) Consortium via the PRIDE (153) partner repository with the dataset identifier PXD006177 (Reviewer account details: Username: reviewer83752@ebi.ac.uk Password: BqBce0sA). The matched spectra have been uploaded onto MS-Viewer (http://msviewer.ucsf.edu/prospector/cgi-bin/mssearch.cgi?report_title=MS-Viewer&search_key=2e58f6g9sh&search_name=msviewer). In the MS-Viewer Report, there are six columns including “scanNum” (i.e., the scan number of the corresponding spectrum), “charge” (i.e., precursor charge), “fraction” (i.e., fraction name of the corresponding mass spectrometry data file), “peptide” (i.e., peptide sequence), “mod” (i.e., modification information), and “Mascot delta score (i.e., PTM delta score)” (i.e., Mascot delta score that equals the top Mascot score minus the second-best Mascot score from the same peptide sequence).
Reference

1. Hsu, J.-L., Huang, S.-Y., Chow, N.-H., and Chen, S.-H. (2003) Stable-isotope dimethyl labeling for quantitative proteomics. *Analytical chemistry* 75, 6843-6852

2. Boersema, P. J., Raijmakers, R., Lemeer, S., Mohammed, S., and Heck, A. J. (2009) Multiplex peptide stable isotope dimethyl labeling for quantitative proteomics. *Nature protocols* 4, 484-494

3. Zhou, Y., Shan, Y., Zhang, L., and Zhang, Y. (2014) Recent advances in stable isotope labeling based techniques for proteome relative quantification. *Journal of chromatography A* 1365, 1-11

4. Lassowskat, I., Hartl, M., Hosp, F., Boersema, P. J., Mann, M., and Finkemeier, I. (2017) Dimethyl-Labeling-Based Quantification of the Lysine Acetylome and Proteome of Plants. *Photorespiration*, pp. 65-81, Springer

5. Kovanich, D., Cappadona, S., Raijmakers, R., Mohammed, S., Scholten, A., and Heck, A. J. (2012) Applications of stable isotope dimethyl labeling in quantitative proteomics. *Analytical and bioanalytical chemistry* 404, 991-1009

6. Thompson, A., Schäfer, J., Kuhn, K., Kienle, S., Schwarz, J., Schmidt, G., Neumann, T., and Hamon, C. (2003) Tandem mass tags: a novel quantification strategy for comparative analysis of complex protein mixtures by MS/MS. *Analytical chemistry* 75, 1895-1904

7. Ross, P. L., Huang, Y. N., Marchese, J. N., Williamson, B., Parker, K., Hattan, S., Khainovski, N., Pillai, S., Dey, S., and Daniels, S. (2004) Multiplexed protein quantitation in Saccharomyces cerevisiae using amine-reactive isobaric tagging reagents. *Molecular & cellular proteomics* 3, 1154-1169

8. Lau, H.-T., Suh, H. W., Golkowski, M., and Ong, S.-E. (2014) Comparing SILAC-and stable isotope dimethyl-labeling approaches for quantitative proteomics. *Journal of proteome research* 13, 4164-4174

9. Grossmann, J., Roschitzki, B., Panse, C., Fortes, C., Barkow-Oesterreicher, S., Rutishauser, D., and Schlapbach, R. (2010) Implementation and evaluation of relative and absolute quantification in shotgun proteomics with label-free methods. *Journal of proteome research* 73, 1740-1746

10. Sap, K. A., and Demmers, J. A. (2012) Labeling Methods in Mass Spectrometry Based Quantitative Proteomics. *Integrative Proteomics*, InTech

11. Tolonen, A. C., Haas, W., Chilaka, A. C., Aach, J., Gygi, S. P., and Church, G. M. (2011) Proteome-wide systems analysis of a cellulosic biofuel-producing microbe. *Molecular systems biology* 7, 461

12. Wijte, D., van Baar, B. L., Heck, A. J., and Altema, A. M. (2010) Probing the proteome response to toluene exposure in the solvent tolerant Pseudomonas putida S12. *Journal of proteome research* 10, 394-403

13. Tang, B., Li, Y., Zhao, L., Yuan, S., Wang, Z., Li, B., and Chen, Q. (2013) Stable isotope dimethyl labeling combined with LTQ mass spectrometric detection, a quantitative proteomics technology used in liver cancer research. *Biomedical reports* 1, 549-554

14. Huang, H.-J., Tsai, M.-L., Chen, Y.-W., and Chen, S.-H. (2011) Quantitative shot-gun proteomics and MS-based activity assay for revealing gender differences in enzyme contents for rat liver microsome. *Journal of proteomics* 74, 2734-2744

15. de Graaf, E. L., Vermeij, W. P., de Waard, M. C., Rijksen, Y., van der Pluijm, I., Hoogenraad, C. C.,
Hoeijmakers, J. H., Altelaar, A. M., and Heck, A. J. (2013) Spatio-temporal analysis of molecular determinants of neuronal degeneration in the aging mouse cerebellum. *Molecular & Cellular Proteomics* 12, 1350-1362

16. Ytterberg, A. J., Peltier, J.-B., and Van Wijk, K. J. (2006) Protein profiling of plastoglobules in chloroplasts and chromoplasts. A surprising site for differential accumulation of metabolic enzymes. *Plant Physiology* 140, 984-997

17. Murphy, J. P., Kong, F., Pinto, D. M., and Wang-Pruski, G. (2010) Relative quantitative proteomic analysis reveals wound response proteins correlated with after-cooking darkening. *Proteomics* 10, 4258-4269

18. Li, L., Song, J., Kalt, W., Forney, C., Tsao, R., Pinto, D., Chisholm, K., Campbell, L., Fillmore, S., and Li, X. (2013) Quantitative proteomic investigation employing stable isotope labeling by peptide dimethylation on proteins of strawberry fruit at different ripening stages. *Journal of proteomics* 94, 219-239

19. Huang, S. Y., Tsai, M. L., Wu, C. J., Hsu, J. L., Ho, S. H., and Chen, S. H. (2006) Quantitation of protein phosphorylation in pregnant rat uteri using stable isotope dimethyl labeling coupled with IMAC. *Proteomics* 6, 1722-1734

20. Corradini, E., Vallur, R., Raaijmakers, L. M., Feil, S., Feil, R., Heck, A. J., and Scholten, A. (2014) Alterations in the cerebellar (Phospho) proteome of a cyclic guanosine monophosphate (cGMP)-dependent protein kinase knockout mouse. *Molecular & Cellular Proteomics* 13, 2004-2016

21. Lemeer, S., Jopling, C., Gouw, J., Mohammed, S., Heck, A. J., Slijper, M., and den Hertog, J. (2008) Comparative phosphoproteomics of zebrafish Fyn/Yes morpholino knockdown embryos. *Molecular & Cellular Proteomics* 7, 2176-2187

22. Melo-Braga, M. N., Schulz, M., Liu, Q., Swistowski, A., Palmisano, G., Engholm-Keller, K., Jakobsen, L., Zeng, X., and Larsen, M. R. (2014) Comprehensive quantitative comparison of the membrane proteome, phosphoproteome, and sialiome of human embryonic and neural stem cells. *Molecular & Cellular Proteomics* 13, 311-328

23. Tsai, C.-F., Wang, Y.-T., Yen, H.-Y., Tsou, C.-C., Ku, W.-C., Lin, P.-Y., Chen, H.-Y., Nesvizhskii, A. I., Ishihama, Y., and Chen, Y.-J. (2015) Large-scale determination of absolute phosphorylation stoichiometries in human cells by motif-targeting quantitative proteomics. *Nature communications* 6, 6622

24. Khidekel, N., Ficarro, S. B., Clark, P. M., Bryan, M. C., Swaney, D. L., Rexach, J. E., Sun, Y. E., Coon, J. J., Peters, E. C., and Hsieh-Wilson, L. C. (2007) Probing the dynamics of O-GlcNAc glycosylation in the brain using quantitative proteomics. *Nature chemical biology* 3, 339-348

25. Parker, B. L., Palmisano, G., Edwards, A. V., White, M. Y., Engholm-Keller, K., Lee, A., Scott, N. E., Kolarich, D., Hambly, B. D., and Packer, N. H. (2011) Quantitative N-linked glycoproteomics of myocardial ischemia and reperfusion injury reveals early remodeling in the extracellular environment. *Molecular & Cellular Proteomics* 10, M110. 006833

26. Weng, Y., Qu, Y., Jiang, H., Wu, Q., Zhang, L., Yuan, H., Zhou, Y., Zhang, X., and Zhang, Y. (2014)
An integrated sample pretreatment platform for quantitative N-glycoproteome analysis with combination of on-line glycopeptide enrichment, deglycosylation and dimethyl labeling. *Analytica chimica acta* 833, 1-8

27. Boersema, P. J., Foong, L. Y., Ding, V. M., Lemeer, S., van Breukelen, B., Philp, R., Boekhorst, J., Snel, B., den Hertog, J., and Choo, A. B. (2010) In-depth qualitative and quantitative profiling of tyrosine phosphorylation using a combination of phosphopeptide immunoaffinity purification and stable isotope dimethyl labeling. *Molecular & Cellular Proteomics* 9, 84-99

28. Phillips, D. (1963) The presence of acetyl groups in histones. *Biochemical Journal* 87, 258

29. Kim, S. C., Sprung, R., Chen, Y., Xu, Y., Ball, H., Pei, J., Cheng, T., Kho, Y., Xiao, H., and Xiao, L. (2006) Substrate and functional diversity of lysine acetylation revealed by a proteomics survey. *Molecular cell* 23, 607-618

30. Choudhary, C., Weinert, B. T., Nishida, Y., Verdin, E., and Mann, M. (2014) The growing landscape of lysine acetylation links metabolism and cell signalling. *Nature reviews Molecular cell biology* 15, 536-550

31. Menzies, K. J., Zhang, H., Katsyuba, E., and Auwerx, J. (2016) Protein acetylation in metabolism [mdash] metabolites and cofactors. *Nature Reviews Endocrinology* 12, 43-60

32. Verdin, E., and Ott, M. (2015) 50 years of protein acetylation: from gene regulation to epigenetics, metabolism and beyond. *Nature reviews Molecular cell biology* 16, 258-264

33. Zhang, K., Williams, K. E., Huang, L., Yau, P., Siino, J. S., Bradbury, E. M., Jones, P. R., Minch, M. J., and Burlingame, A. L. (2002) Histone Acetylation and Deacetylation Identification of Acetylation and Methylation Sites of HeLa Histone H4 by Mass Spectrometry. *Molecular & Cellular Proteomics* 1, 500-508

34. Chu, F., Ren, X., Chasse, A., Hickman, T., Zhang, L., Yuh, J., Smith, M. T., and Burlingame, A. L. (2011) Quantitative mass spectrometry reveals the epigenome as a target of arsenic. *Chemico-biological interactions* 192, 113-117

35. Turner, B. M. (2000) Histone acetylation and an epigenetic code. *Bioessays* 22, 836-845

36. Xing, S., and Poirier, Y. (2012) The protein acetylome and the regulation of metabolism. *Trends in plant science* 17, 423-430

37. van Noort, V., Seebacher, J., Bader, S., Mohammed, S., Vonkova, I., Betts, M. J., Kühner, S., Kumar, R., Maier, T., and O’Flaherty, M. (2012) Cross-talk between phosphorylation and lysine acetylation in a genome-reduced bacterium. *Molecular systems biology* 8, 571

38. Weinert, B. T., Satpathy, S., Hansen, B. K., Lyon, D., Jensen, L. J., and Choudhary, C. (2017) Accurate quantification of site-specific acetylation stoichiometry reveals the impact of sirtuin deacetylase CobB on the E. coli acetylome. *Molecular & Cellular Proteomics* 16, 759-769

39. Choudhary, C., Kumar, C., Gnad, F., Nielsen, M. L., Rehman, M., Walther, T. C., Olsen, J. V., and Mann, M. (2009) Lysine acetylation targets protein complexes and co-regulates major cellular functions. *Science* 325, 834-840

40. Zhu, X., Liu, X., Cheng, Z., Zhu, J., Xu, L., Wang, F., Qi, W., Yan, J., Liu, N., and Sun, Z. (2016)
Quantitative analysis of global proteome and lysine acetylome reveal the differential impacts of Vpa and Saha on HL60 cells. *Scientific reports* 6

41. Zhang, C., Zhai, Z., Tang, M., Cheng, Z., Li, T., Wang, H., and Zhu, W. G. (2017) Quantitative proteome-based systematic identification of SIRT7 substrates. *Proteomics*

42. Cobbold, S. A., Santos, J. M., Ochoa, A., Perlman, D. H., and Llinàs, M. (2016) Proteome-wide analysis reveals widespread lysine acetylation of major protein complexes in the malaria parasite. *Scientific reports* 6

43. Fernandes, J., Weddle, A., Kinter, C. S., Humphries, K. M., Mather, T., Szweda, L. I., and Kinter, M. (2015) Lysine acetylation activates mitochondrial aconitase in the heart. *Biochemistry* 54, 4008-4018

44. Hebert, A. S., Dittenhafer-Reed, K. E., Yu, W., Bailey, D. J., Selen, E. S., Boersma, M. D., Carson, J. J., Tonelli, M., Balloon, A. J., and Higbee, A. J. (2013) Calorie restriction and SIRT3 trigger global reprogramming of the mitochondrial protein acetylome. *Molecular cell* 49, 186-199

45. Rardin, M. J., Newman, J. C., Held, J. M., Cusack, M. P., Sorensen, D. J., Li, B., Schilling, B., Mooney, S. D., Kahn, C. R., and Verdin, E. (2013) Label-free quantitative proteomics of the lysine acetylome in mitochondria identifies substrates of SIRT3 in metabolic pathways. *Proceedings of the National Academy of Sciences* 110, 6601-6606

46. Svinkina, T., Gu, H., Silva, J. C., Mertins, P., Qiao, J., Fereshtehian, S., Jaffe, J. D., Kuhn, E., Udeshi, N. D., and Carr, S. A. (2015) Deep, quantitative coverage of the lysine acetylome using novel anti-acetyl-lysine antibodies and an optimized proteomic workflow. *Molecular & Cellular Proteomics* 14, 2429-2440

47. Chen, S.-H., Chen, C.-R., Chen, S.-H., Li, D.-T., and Hsu, J.-L. (2013) Improved Nα-acetylated peptide enrichment following dimethyl labeling and SCX. *Journal of proteome research* 12, 3277-3287

48. Wu, X., Oh, M.-H., Schwarz, E. M., Larue, C. T., Sivaguru, M., Imai, B. S., Yau, P. M., Ort, D. R., and Huber, S. C. (2011) Lysine acetylation is a widespread protein modification for diverse proteins in Arabidopsis. *Plant physiology* 155, 1769-1778

49. Gao, X., Hong, H., Li, W.-C., Yang, L., Huang, J., Xiao, Y.-L., Chen, X.-Y., and Chen, G.-Y. (2016) Downregulation of Rubisco activity by non-enzymatic acetylation of RbcL. *Molecular plant* 9, 1018-1027

50. Shahbazian, M. D., and Grunstein, M. (2007) Functions of site-specific histone acetylation and deacetylation. *Annu. Rev. Biochem.* 76, 75-100

51. Li, D.-X., Chen, W.-Q., Xu, Z.-H., and Bai, S.-N. (2015) HISTONE DEACETYLASE6-defective mutants show increased expression and acetylation of ENHANCER OF TRIPTYCHON AND CAPRICE1 and GLABRA2 with small but significant effects on root epidermis cellular pattern. *Plant physiology* 168, 1448-1458

52. Li, H., Soriano, M., Cordewener, J., Muiño, J. M., Riksen, T., Fukuoka, H., Angenent, G. C., and Boutilier, K. (2014) The histone deacetylase inhibitor trichostatin a promotes totipotency in the male gametophyte. *The Plant Cell* 26, 195-209

53. Singh, P., Yekondi, S., Chen, P.-W., Tsai, C.-H., Yu, C.-W., Wu, K., and Zimmerli, L. (2014) Environmental History Modulates Arabidopsis Pattern-Triggered Immunity in a HISTONE ACETYLTRANSFERASE1–Dependent Manner. *The Plant Cell* 26, 2676-2688

54. Weiste, C., and Dröge-Laser, W. (2014) The Arabidopsis transcription factor bZIP11 activates
auxin-mediated transcription by recruiting the histone acetylation machinery. *Nature Communications* 5, 3883

55. Liu, C., Li, L-C., Chen, W-Q., Chen, X., Xu, Z-H., and Bai, S-N. (2013) HDA18 affects cell fate in Arabidopsis root epidermis via histone acetylation at four kinase genes. *The Plant Cell* 25, 257-269

56. Liu, X., Chen, C-Y., Wang, K-C., Luo, M., Tai, R., Yuan, L., Zhao, M., Yang, S., Tian, G., and Cui, Y. (2013) PHYTOCHROME INTERACTING FACTOR3 associates with the histone deacetylase HDA15 in repression of chlorophyll biosynthesis and photosynthesis in etiolated Arabidopsis seedlings. *The Plant Cell* 25, 1258-1273

57. Zhou, Y., Tan, B., Luo, M., Li, Y., Liu, C., Chen, C., Yu, C-W., Yang, S., Dong, S., and Ruan, J. (2013) HISTONE DEACETYLASE19 interacts with HSL1 and participates in the repression of seed maturation genes in Arabidopsis seedlings. *The Plant Cell* 25, 134-148

58. Nallamilli, B. R. R., Edelmann, M. J., Zhong, X., Tan, F., Mujahid, H., Zhang, J., Nanduri, B., and Peng, Z. (2014) Global analysis of lysine acetylation suggests the involvement of protein acetylation in diverse biological processes in rice (*Oryza sativa*). *PLoS One* 9, e89283

59. He, D., Wang, Q., Li, M., Damaris, R. N., Yi, X., Cheng, Z., and Yang, P. (2016) Global proteome analyses of lysine acetylation and succinylation reveal the widespread involvement of both modification in metabolism in the embryo of germinating rice seed. *Journal of proteome research* 15, 879-890

60. Xiong, Y., Peng, X., Cheng, Z., Liu, W., and Wang, G-L. (2016) A comprehensive catalog of the lysine-acetylation targets in rice (*Oryza sativa*) based on proteomic analyses. *Journal of proteomics* 138, 20-29

61. Smith-Hammond, C. L., Swatek, K. N., Johnston, M. L., Thelen, J. J., and Miernyk, J. A. (2014) Initial description of the developing soybean seed protein Lys-N ε-acetylome. *Journal of proteomics* 96, 56-66

62. Zhen, S., Deng, X., Wang, J., Zhu, G., Cao, H., Yuan, L., and Yan, Y. (2016) First Comprehensive Proteome Analyses of Lysine Acetylation and Succinylation in Seedling Leaves of Brachypodium distachyon L. *Scientific reports* 6, 31576

63. Zhang, Y., Song, L., Liang, W., Mu, P., Wang, S., and Lin, Q. (2016) Comprehensive profiling of lysine acetylproteome analysis reveals diverse functions of lysine acetylation in common wheat. *Scientific reports* 6, 21069

64. Fang, X., Chen, W., Zhao, Y., Ruan, S., Zhang, H., Yan, C., Jin, L., Cao, L., Zhu, J., and Ma, H. (2015) Global analysis of lysine acetylation in strawberry leaves. *Frontiers in plant science* 6

65. Finkemeier, I., Laxa, M., Miguez, L., Howden, A. J., and Sweetlove, L. J. (2011) Proteins of diverse function and subcellular location are lysine acetylated in Arabidopsis. *Plant physiology* 155, 1779-1790

66. König, A.-C., Hartl, M., Boersema, P. J., Mann, M., and Finkemeier, I. (2014) The mitochondrial lysine acetylome of Arabidopsis. *Mitochondrion* 19, 252-260

67. Hartl, M., Füßl, M., Boersema, P. J., Jost, J. O., Kramer, K., Bakirbas, A., Sindlinger, J., Plöchinger, M., Leister, D., and Uhrig, G. (2017) Lysine acetylome profiling uncovers novel histone deacetylase substrate proteins in Arabidopsis. *Molecular systems biology* 13, 949
68. Ong, S.-E., and Mann, M. (2005) Mass spectrometry–based proteomics turns quantitative. 
  *Nature chemical biology* 1

69. Pan, S., Aebersold, R., Chen, R., Rush, J., Goodlett, D. R., McIntosh, M. W., Zhang, J., and 
  Brentnall, T. A. (2008) Mass spectrometry based targeted protein quantification: methods and 
  applications. *Journal of proteome research* 8, 787-797

70. Domon, B., and Aebersold, R. (2010) Options and considerations when selecting a quantitative 
  proteomics strategy. *Nature biotechnology* 28, 710-721

71. Bakalarski, C. E., and Kirkpatrick, D. S. (2016) A biologist's field guide to multiplexed 
  quantitative proteomics. *Molecular & Cellular Proteomics* 15, 1489-1497

72. Ong, S.-E., Blagoev, B., Kratchmarova, I., Kristensen, D. B., Steen, H., Pandey, A., and Mann, M. 
  (2002) Stable isotope labeling by amino acids in cell culture, SILAC, as a simple and accurate approach 
  to expression proteomics. *Molecular & cellular proteomics* 1, 376-386

73. Mueller, L. N., Brusniak, M.-Y., Mani, D., and Aebersold, R. (2008) An assessment of software 
  solutions for the analysis of mass spectrometry based quantitative proteomics data. *Journal of 
  proteome research* 7, 51-61

74. Cappadona, S., Baker, P. R., Cutillas, P. R., Heck, A. J., and van Breukelen, B. (2012) Current 
  challenges in software solutions for mass spectrometry-based quantitative proteomics. *Amino acids* 43, 
  1087-1108

75. Rauniyar, N., and Yates III, J. R. (2014) Isobaric labeling-based relative quantification in shotgun 
  proteomics. *Journal of proteome research* 13, 5293-5309

76. Cox, J., and Mann, M. (2008) MaxQuant enables high peptide identification rates, 
  individualized ppb-range mass accuracies and proteome-wide protein quantification. *Nature 
  biotechnology* 26, 1367-1372

77. Liu, C., Song, C.-Q., Yuan, Z.-F., Fu, Y., Chi, H., Wang, L.-H., Fan, S.-B., Zhang, K., Zeng, W.-F., and 
  He, S.-M. (2014) pQuan improves quantitation by keeping out interfering signals and evaluating the 
  accuracy of calculated ratios. *Analytical chemistry* 86, 5286-5294

78. Han, D. K., Eng, J., Zhou, H., and Aebersold, R. (2001) Quantitative profiling of differentiation- 
  induced microsomal proteins using isotope-coded affinity tags and mass spectrometry. *Nature 
  biotechnology* 19, 946-951

79. MacLean, B., Tomazela, D. M., Shulman, N., Chambers, M., Finney, G. L., Frewen, B., Kern, R., 
  Tabb, D. L., Liebler, D. C., and MacCoss, M. J. (2010) Skyline: an open source document editor for creating 
  and analyzing targeted proteomics experiments. *Bioinformatics* 26, 966-968

80. Schubert, O. T., Röst, H. L., Collins, B. C., Rosenberger, G., and Aebersold, R. (2017) Quantitative 
  proteomics: challenges and opportunities in basic and applied research. *Nature Protocols* 12, 1289-1294

81. Zhang, B., Käll, L., and Zubarev, R. A. (2016) DeMix-Q: quantification-centered data processing 
  workflow. *Molecular & Cellular Proteomics* 15, 1467-1478

82. Mitchell, C. J., Kim, M.-S., Na, C. H., and Pandey, A. (2016) PyQuant: a versatile framework for 
  analysis of quantitative mass spectrometry data. *Molecular & Cellular Proteomics* 15, 2829-2838
83. Argentini, A., Goeminne, L. J., Verheggen, K., Hulstaert, N., Staes, A., Clement, L., and Martens, L. (2016) moFF: a robust and automated approach to extract peptide ion intensities. Nature methods 13, 964-966

84. Leufken, J., Niehues, A., Sarin, L. P., Wessel, F., Hippler, M., Leidel, S. A., and Fufezan, C. (2017) pyQms enables universal and accurate quantification of mass spectrometry data. Molecular & Cellular Proteomics 16, 1736-1745

85. Leek, J. T., Scharpf, R. B., Bravo, H. C., Simcha, D., Langmead, B., Johnson, W. E., Geman, D., Baggerly, K., and Irizarry, R. A. (2010) Tackling the widespread and critical impact of batch effects in high-throughput data. Nature Reviews Genetics 11, 733-739

86. Cottrell, J. S., and London, U. (1999) Probability-based protein identification by searching sequence databases using mass spectrometry data. Electrophoresis 20, 3551-3567

87. Benjamini, Y., and Hochberg, Y. (1995) Controlling the false discovery rate: a practical and powerful approach to multiple testing. Journal of the royal statistical society. Series B (Methodological), 289-300

88. Guo, G., and Li, N. (2011) Relative and accurate measurement of protein abundance using 15 N stable isotope labeling in Arabidopsis [SILIA]. Phytochemistry 72, 1028-1039

89. Qing, D., Yang, Z., Li, M., Wong, W. S., Guo, G., Liu, S., Guo, H., and Li, N. (2016) Quantitative and functional phosphoproteomic analysis reveals that ethylene regulates water transport via the C-terminal phosphorylation of aquaporin PIP2; 1 in Arabidopsis. Molecular plant 9, 158-174

90. Chambers, M. C., Maclean, B., Burke, R., Amodei, D., Ruderman, D. L., Neumann, S., Gatto, L., Fischer, B., Pratt, B., and Egertson, J. (2012) A cross-platform toolkit for mass spectrometry and proteomics. Nature biotechnology 30, 918-920

91. Elias, J. E., and Gygi, S. P. (2007) Target-decoy search strategy for increased confidence in large-scale protein identifications by mass spectrometry. Nature methods 4, 207-214

92. Yang, Z., Guo, G., Zhang, M., Liu, C. Y., Hu, Q., Lam, H., Cheng, H., Xue, Y., Li, J., and Li, N. (2013) Stable isotope metabolic labeling-based quantitative phosphoproteomic analysis of Arabidopsis mutants reveals ethylene-regulated time-dependent phosphoproteins and putative substrates of constitutive triple response 1 kinase. Molecular & Cellular Proteomics 12, 3559-3582

93. Käll, L., Canterbury, J. D., Weston, J., Noble, W. S., and MacCoss, M. J. (2007) Semi-supervised learning for peptide identification from shotgun proteomics datasets. Nature methods 4, 923-925

94. Park, J., Piehowski, P. D., Wilkins, C., Zhou, M., Mendoza, J., Fujimoto, G. M., Gibbons, B. C., Shaw, J. B., Shen, Y., and Shukla, A. K. (2017) Informed-Proteomics: open-source software package for top-down proteomics. Nature Methods 14, 909-914

95. Pavelka, N., Fournier, M. L., Swanson, S. K., Pelizzola, M., Ricciardi-Castagnoli, P., Florens, L., and Washburn, M. P. (2008) Statistical similarities between transcriptomics and quantitative shotgun proteomics data. Molecular & Cellular Proteomics 7, 631-644

96. Johnson, W. E., Li, C., and Rabinovic, A. (2007) Adjusting batch effects in microarray expression data using empirical Bayes methods. Biostatistics 8, 118-127
97. Chou, M. F., and Schwartz, D. (2011) Biological sequence motif discovery using motif-x. *Current protocols in bioinformatics*, 13.15. 11-13.15. 24
98. Baerenfasser, K., Hirsch-Hoffmann, M., Svozil, J., Hull, R., Russenberger, D., Bischof, S., Lu, Q., Gruissem, W., and Baginsky, S. (2011) pep2pro: a new tool for comprehensive proteome data analysis to reveal information about organ-specific proteomes in Arabidopsis thaliana. *Integrative Biology* 3, 225-237
99. Durek, P., Schmidt, R., Heazlewood, J. L., Jones, A., MacLean, D., Nagel, A., Kersten, B., and Schulze, W. X. (2009) PhosphAt: the Arabidopsis thaliana phosphorylation site database. An update. *Nucleic acids research* 38, D828-D834
100. Savitski, M. M., Lemeer, S., Boesche, M., Lang, M., Mathieson, T., Bantscheff, M., and Kuster, B. (2011) Confident phosphorylation site localization using the Mascot Delta Score. *Molecular & cellular proteomics* 10, M110. 003830
101. Sterner, D. E., and Berger, S. L. (2000) Acetylation of histones and transcription-related factors. *Microbiology and Molecular Biology Reviews* 64, 435-459
102. Hord, C. L., Chen, C., DeYoung, B. J., Clark, S. E., and Ma, H. (2006) The BAM1/BAM2 receptor-like kinases are important regulators of Arabidopsis early anther development. *The Plant Cell* 18, 1667-1680
103. He, Z., Wang, Z.-y., Li, J., Zhu, Q., Lamb, C., Ronald, P., and Chory, J. (2000) Perception of Brassinosteroids by the Extracellular Domain of the Receptor Kinase Bri1. *Science* 288, 2360-2363
104. Li, J., Wen, J., Lease, K. A., Doke, J. T., Tax, F. E., and Walker, J. C. (2002) BAK1, an Arabidopsis LRR receptor-like protein kinase, interacts with BRI1 and modulates brassinosteroid signaling. *Cell* 110, 213-222
105. Sun, Y., Fan, X.-Y., Cao, D.-M., Tang, W., He, K., Zhu, J.-Y., He, J.-X., Bai, M.-Y., Zhu, S., and Oh, E. (2010) Integration of brassinosteroid signal transduction with the transcription network for plant growth regulation in Arabidopsis. *Developmental cell* 19, 765-777
106. Christie, J. M. (2007) Phototropin blue-light receptors. *Annu. Rev. Plant Biol.* 58, 21-45
107. Kim, B., Jeong, Y. J., Corvalán, C., Fujioka, S., Cho, S., Park, T., and Choe, S. (2014) Darkness and gulliver2/phyB mutation decrease the abundance of phosphorylated BZR1 to activate brassinosteroid signaling in Arabidopsis. *The Plant Journal* 77, 737-747
108. Feng, S., Martinez, C., Gusmaroli, G., Wang, Y., Zhou, J., Wang, F., Chen, L., Yu, L., Iglesias-Pedraz, J. M., and Kircher, S. (2008) Coordinated regulation of Arabidopsis thaliana development by light and gibberellins. *Nature* 451, 475-479
109. Gou, X., Yin, H., He, K., Du, J., Yi, J., Xu, S., Lin, H., Clouse, S. D., and Li, J. (2012) Genetic evidence for an indispensable role of somatic embryogenesis receptor kinases in brassinosteroid signaling. *PLoS genetics* 8, e1002452
110. Kim, T.-W., Guan, S., Sun, Y., Deng, Z., Tang, W., Shang, J.-X., Sun, Y., Burlingame, A. L., and Wang, Z.-Y. (2009) Brassinosteroid signal transduction from cell-surface receptor kinases to nuclear transcription factors. *Nature cell biology* 11, 1254-1260
111. Tang, W., Kim, T.-W., Oses-Prieto, J. A., Sun, Y., Deng, Z., Zhu, S., Wang, R., Burlingame, A. L., and Wang, Z.-Y. (2008) BSKs mediate signal transduction from the receptor kinase BRI1 in Arabidopsis. *Science* 321, 557-560

112. Guo, H., Li, L., Ye, H., Yu, X., Algreen, A., and Yin, Y. (2009) Three related receptor-like kinases are required for optimal cell elongation in Arabidopsis thaliana. *Proceedings of the National Academy of Sciences* 106, 7648-7653

113. Doll, S., and Burlingame, A. L. (2014) Mass spectrometry-based detection and assignment of protein posttranslational modifications. *ACS chemical biology* 10, 63-71

114. Thingholm, T. E., Jensen, O. N., Robinson, P. J., and Larsen, M. R. (2008) SIMAC (sequential elution from IMAC), a phosphoproteomics strategy for the rapid separation of monophosphorylated from multiply phosphorylated peptides. *Molecular & cellular proteomics* 7, 661-671

115. Ongay, S., Boichenko, A., Govorukhina, N., and Bischoff, R. (2012) Glycopeptide enrichment and separation for protein glycosylation analysis. *Journal of separation science* 35, 2341-2372

116. Bustos, D., Bakalarski, C. E., Yang, Y., Peng, J., and Kirkpatrick, D. S. (2012) Characterizing ubiquitination sites by peptide-based immunoaffinity enrichment. *Molecular & Cellular Proteomics* 11, 1529-1540

117. Rush, J., Moritz, A., Lee, K. A., Guo, A., Goss, V. L., Spek, E. J., Zhang, H., Zha, X.-M., Polakiewicz, R. D., and Comb, M. J. (2005) Immunoaffinity profiling of tyrosine phosphorylation in cancer cells. *Nature biotechnology* 23

118. Afjehi-Sadat, L., and Garcia, B. A. (2013) Comprehending dynamic protein methylation with mass spectrometry. *Current opinion in chemical biology* 17, 12-19

119. Xie, Z., Zhang, D., Chung, D., Tang, Z., Huang, H., Dai, L., Qi, S., Li, J., Colak, G., and Chen, Y. (2016) Metabolic regulation of gene expression by histone lysine β-hydroxybutyrylation. *Molecular cell* 62, 194-206

120. Tan, M., Peng, C., Anderson, K. A., Chhoy, P., Xie, Z., Dai, L., Park, J., Chen, Y., Huang, H., and Zhang, Y. (2014) Lysine glutarylation is a protein posttranslational modification regulated by SIRT5. *Cell metabolism* 19, 605-617

121. Byrne, M. E. (2009) A role for the ribosome in development. *Trends in plant science* 14, 512-519

122. Carroll, A. J., Heazlewood, J. L., Ito, J., and Millar, A. H. (2008) Analysis of the Arabidopsis cytosolic ribosome proteome provides detailed insights into its components and their post-translational modification. *Molecular & cellular proteomics* 7, 347-369

123. Xue, S., and Barna, M. (2012) Specialized ribosomes: a new frontier in gene regulation and organismal biology. *Nature reviews Molecular cell biology* 13, 355-369

124. Polevoda, B., and Sherman, F. (2007) Methylation of proteins involved in translation. *Molecular microbiology* 65, 590-606

125. Mazumder, B., Sampath, P., Seshadri, V., Maitra, R. K., DiCorleto, P. E., and Fox, P. L. (2003) Regulated release of L13a from the 60S ribosomal subunit as a mechanism of transcript-specific
translational control. *Cell* 115, 187-198

126. Spence, J., Gali, R. R., Dittmar, G., Sherman, F., Karin, M., and Finley, D. (2000) Cell cycle–regulated modification of the ribosome by a variant multiubiquitin chain. *Cell* 102, 67-76

127. Yang, Y., Cimen, H., Han, M.-J., Shi, T., Deng, J.-H., Koc, H., Palacios, O. M., Montier, L., Bai, Y., and Tong, Q. (2010) NAD+-dependent deacetylase SIRT3 regulates mitochondrial protein synthesis by deacetylation of the ribosomal protein MRPL10. *Journal of Biological Chemistry* 285, 7417-7429

128. Kamita, M., Kimura, Y., Ino, Y., Kamp, R. M., Polevoda, B., Sherman, F., and Hirano, H. (2011) Nα-Acetylation of yeast ribosomal proteins and its effect on protein synthesis. *Journal of proteomics* 74, 431-441

129. Biever, A., Valjent, E., and Puighermanal, E. (2015) Ribosomal protein S6 phosphorylation in the nervous system: from regulation to function. *Frontiers in molecular neuroscience* 8

130. Tutar, L., and Tutar, Y. (2010) Heat shock proteins; an overview. *Current Pharmaceutical Biotechnology* 11, 216-222

131. Vabulas, R. M., Raychaudhuri, S., Hayer-Hartl, M., and Hartl, F. U. (2010) Protein folding in the cytoplasm and the heat shock response. *Cold Spring Harbor perspectives in biology* 2, a004390

132. Saibil, H. (2013) Chaperone machines for protein folding, unfolding and disaggregation. *Nature reviews Molecular cell biology* 14, 630-642

133. Al-Whaibi, M. H. (2011) Plant heat-shock proteins: a mini review. *Journal of King Saud University-Science* 23, 139-150

134. Li, J., Soroka, J., and Buchner, J. (2012) The Hsp90 chaperone machinery: conformational dynamics and regulation by co-chaperones. *Biochimica et Biophysica Acta (BBA)-Molecular Cell Research* 1823, 624-635

135. Mollapour, M., and Neckers, L. (2012) Post-translational modifications of Hsp90 and their contributions to chaperone regulation. *Biochimica et Biophysica Acta (BBA)-Molecular Cell Research* 1823, 648-655

136. Scroggins, B. T., Robzyk, K., Wang, D., Marcu, M. G., Tsutsumi, S., Beebe, K., Cotter, R. J., Felts, S., Toft, D., and Karnitz, L. (2007) An acetylation site in the middle domain of Hsp90 regulates chaperone function. *Molecular cell* 25, 151-159

137. Zhang, F., Qi, B., Wang, L., Zhao, B., Rode, S., Riggin, N. D., Ecker, J. R., and Qiao, H. (2016) EIN2-dependent regulation of acetylation of histone H3K14 and non-canonical histone H3K23 in ethylene signalling. *Nature communications* 7

138. Guo, H., and Ecker, J. R. (2004) The ethylene signaling pathway: new insights. *Current opinion in plant biology* 7, 40-49

139. Merchante, C., Alonso, J. M., and Stepanova, A. N. (2013) Ethylene signaling: simple ligand, complex regulation. *Current opinion in plant biology* 16, 554-560

140. Li, H., Wong, W. S., Zhu, L., Guo, H. W., Ecker, J., and Li, N. (2009) Phosphoproteomic analysis of ethylene-regulated protein phosphorylation in etiolated seedlings of Arabidopsis mutant ein2 using two-dimensional separations coupled with a hybrid quadrupole time-of-flight mass spectrometer.
141. Li, Y., Shu, Y., Peng, C., Zhu, L., Guo, G., and Li, N. (2012) Absolute quantitation of isoforms of post-translationally modified proteins in transgenic organism. *Molecular & Cellular Proteomics* 11, 272-285

142. Zhu, L., Liu, D., Li, Y., and Li, N. (2013) Functional phosphoproteomic analysis reveals that a serine-62-phosphorylated isoform of ethylene response factor110 is involved in Arabidopsis bolting. *Plant physiology* 161, 904-917

143. Whitmarsh, A., and Davis, R. (2000) Regulation of transcription factor function by phosphorylation. *Cellular and Molecular Life Sciences CMLS* 57, 1172-1183

144. Schwer, B., Bunkenborg, J., Verdin, R. O., Andersen, J. S., and Verdin, E. (2006) Reversible lysine acetylation controls the activity of the mitochondrial enzyme acetyl-CoA synthetase 2. *Proceedings of the National Academy of Sciences* 103, 10224-10229

145. Hirschey, M. D., Shimazu, T., Goetzman, E., Jing, E., Schwer, B., Lombard, D. B., Grueter, C. A., Harris, C., Biddinger, S., and Ilkayeva, O. R. (2010) SIRT3 regulates mitochondrial fatty-acid oxidation by reversible enzyme deacetylation. *Nature* 464, 121-125

146. Guan, K.-L., and Xiong, Y. (2011) Regulation of intermediary metabolism by protein acetylation. *Trends in biochemical sciences* 36, 108-116

147. Lin, H., Su, X., and He, B. (2012) Protein lysine acylation and cysteine succination by intermediates of energy metabolism. *ACS chemical biology* 7, 947-960

148. Chen, L.-T., Luo, M., Wang, Y.-Y., and Wu, K. (2010) Involvement of Arabidopsis histone deacetylase HDA6 in ABA and salt stress response. *Journal of experimental botany* 61, 3345-3353

149. Linster, E., Stephan, I., Bienvenut, W. V., Maple-Grødem, J., Myklebust, L. M., Huber, M., Reichelt, M., Sticht, C., Møller, S. G., and Meinnel, T. (2015) Downregulation of N-terminal acetylation triggers ABA-mediated drought responses in Arabidopsis. *Nature communications* 6

150. Yang, X.-J., and Seto, E. (2008) Lysine acetylation: codified crosstalk with other posttranslational modifications. *Molecular cell* 31, 449-461

151. Latham, J. A., and Dent, S. Y. (2007) Cross-regulation of histone modifications. *Nature structural & molecular biology* 14, 1017-1024

152. Deutsch, E. W., Csordas, A., Sun, Z., Jarnuczak, A., Perez-Riverol, Y., Ternent, T., Campbell, D. S., Bernal-Linares, M., Okuda, S., and Kawano, S. (2016) The ProteomeXchange consortium in 2017: supporting the cultural change in proteomics public data deposition. *Nucleic acids research*, gkw936

153. Vizcaíno, J. A., Csordas, A., Del-Toro, N., Dianes, J. A., Griss, J., Lavidas, I., Mayer, G., Perez-Riverol, Y., Reisinger, F., and Ternent, T. (2015) 2016 update of the PRIDE database and its related tools. *Nucleic acids research* 44, D447-D456
Figure legends

Figure 1

**Dimethyl labelling of the Arabidopsis acetylome, chromatographic enrichment for LC-MS/MS analysis of acetylpeptides, and quantitative identification of acetylation unique PTM peptide arrays (UPAs)**

(A) Chemical reactions for light- and heavy-isotope-coded dimethyl labelling. The N-terminus and lysine side chains of acetylpeptides (Ac-p) can be labelled. (B) Workflow for protein fractionation by CsCl density gradient centrifugation, peptide preparation, dimethyl labelling of the whole-plant acetylome, chromatographic enrichment, LC-MS/MS analysis, Mascot-based database searching, and quantification of acetylation UPAs with SQUA-D (*i.e.*, first 3C procedures of quantitative PTM proteomics). All extracted cellular proteins were separated into three fractions: top (t), middle (m), and bottom (b). The peptides generated from the t, m, and b fractions from control (C) and treated (T) Arabidopsis were labelled with light- (28 Da) or heavy- (34 Da) isotope-coded dimethyl, respectively (*i.e.*, the first C procedure). The mixes of light and heavy isotope-coded dimethyl-labelled peptides from the control and the treated plants, respectively, were defined as forward (F) replicates. Conversely, the mixes of heavy and light isotope-coded dimethyl-labelled peptides from the control and the treated plants, respectively, were defined as reciprocal (R) replicates. Both F (*i.e.*, tF, mF, and bF) and R (*i.e.*, tR, mR, and bR) replicates were produced from a single biological experiment. The peptide mixtures were further fractionated by SCX-HPLC, enriched with anti-acetyllysine antibodies, and analysed by LC-MS/MS (*i.e.*, the second C procedure). The resulting MS data were further analysed by Mascot and SQUA-D.
software for sequential identification and quantification (i.e., the third C procedure).
The Mascot-identified acetylpeptides were evaluated based on the following criteria: FDR ≤ 0.01, where FDR is the false discovery rate obtained by the target-decoy strategy. One biological replicate generates 2 experimental replicates (i.e., F and R). In total, there were 6 experimental replicates (i.e., F1, R1, F2, R2, F3, and R3) from 3 biological replicates, which produced 18 batches of peptide samples (i.e., 3 biological replicates × 3 CDG fractions [t, m, and b] × 2 labelling/mixing types [F and R]). Batch effect adjustment was applied to the batches that have batch effect. The measurement of log-ratios of XIC was followed by statistical evaluation which included both the t-test and multiple test correction (i.e., BH-FDR). The significant cut-off criteria were BH-FDR ≤0.1 and |log-ratio| ≥ 0.5×SD, where SD represents the standard deviation of all log-ratios. The detailed identification and quantification workflow (dry laboratory, VI-VII) are shown in Supplementary Figure S1. XIC denotes extracted ion chromatogram.

Figure 2

Analysis of Arabidopsis acetylpeptides

(A) Venn diagram showing the identified non-redundant light- and heavy-isotope-coded acetylpeptides (green and brown circles, respectively). (B) All identified non-redundant acetylpeptides from six experimental replicates (i.e., F1, R1, F2, R2, F3, and R3). (C) Conserved acetylation motifs of histone proteins from repeatedly identified acetylpeptides. Asterisks (*) indicate acetylation sites. (D) Venn diagram comparing previously reported acetylation sites and the acetylation sites identified in this study. Blue circles represent previously identified acetylation sites on lysine side chains. Red circles represent the identified lysine side chain acetylation sites from this study. (E)
Distribution of acetylation sites per protein. (F) Biological process enrichment for the overly acetylated proteins (OAPs) (with > 8 acetylation sites). The thickness of each line represents the number of overlapping proteins between two processes, where thicker lines indicate more proteins.

**Figure 3**

**Quantitative analysis of the acetylome and immunoblot validation of an identified hormone-regulated acetylation site in Arabidopsis**

(A) Quantification results obtained via dimethyl labelling-based extracted ion chromatogram (XIC). In the upper panel, each circle represents a quantified UPA (Supplemental Table S2b). The circles from up regulated UPAs are marked with red and the circles from down regulated UPAs are marked with blue. The orange circle indicates the UPA (from histone superfamily protein AT1G07660) verified by immunoblot assay. The x-axis indicates the average of log-ratios from pairs of treatment and control ions. The y-axis indicates -10 times the log-transformed *p*-value. The significant threshold is *BH FDR* ≤ 0.1. The horizontal dashed line indicates the corresponding *p*-value threshold. Two vertical dashed lines indicate the boundary of ±0.5 × *SD* (standard deviation). The lower panel contains a histogram (green) and a fitted Gaussian distribution (red) of all quantified UPAs. The x-axis indicates the average log-ratio and the y-axis indicates the number of UPA in the corresponding histogram bin. Two vertical dashed lines indicate the boundary of ±0.5 × *SD* (standard deviation). (B-C) MS spectra (B) and MS/MS spectra (C) of the light and heavy isotope-coded dimethyl modified peptides of acetylation UPSP AT1G07660-K6-K9-K13 (orange circle in Figure 3A) which has been demonstrated to be ethylene-
suppressed by both XIC and immunoblot assay. (D-I) Profiles of acetylation level of quantified UPAs of Heat shock protein 70 family protein (AT3G09440) (D), translational elongation factor EMB2726 (AT4G29060) (E), Histone superfamily protein (AT2G28720) (F), Heat shock protein 90-7 (AT4G24190) (G), AAC1, ADP/ATP carrier 1 (AT3G08580) (H), and ATP synthase subunit beta (ATCG00480) (I). Green bars indicated that these acetylation UPAs passed the cutoff (BH FDR ≤ 0.1). (J) Immunoblot assay to detect acetylation on K6 of protein AT1G07660. Plants in experiment group were treated with ethylene precursor 1-aminocyclopropane-1-carboxylic acid (ACC) of 10 μM (orange bar) and untreated plants were used as control (green bar). Actin was also detected as loading control. Three biological replicates and two technic replicates for each biological replicate were performed for each genotype. Average values are shown with error bars (±SEM (standard error)). Statistical significance was determined using one-sample $t$-test. * represents $p < 0.05$, ** represents $p < 0.01$ and *** represents $p < 0.001$. (K) Conserved acetylation motifs that were extracted from the validated acetylation sites: K6 of protein AT1G07660. Asterisks (*) stands for acetylation sites.

**Figure 4**

**Bioinformatic analysis of repeatedly identified acetylproteins in Arabidopsis**

(A) Comparison between GO analysis of repeatedly identified acetylproteins and the Arabidopsis leaf proteome using the equation in the Experimental Procedures. Categories A1 - A9 represent the up-regulated categories among acetylproteins. Protein categories with the top three log2 ratios are shown. (B) Biological processes enrichment
for the hormone ethylene-regulated acetylproteins. The thickness of each line represents the number of overlapped proteins between two processes, where thicker lines indicate more proteins.
Light isotope-coded dimethyl labelling (L): \[ \text{Ac-p} \rightarrow \text{CH}_2\text{O} \rightarrow \text{Ac-p} \rightarrow \text{N} \]

Heavy isotope-coded dimethyl labelling (H): \[ \text{Ac-p} \rightarrow \text{NaBH}_3\text{CN} \rightarrow \text{Ac-p} \rightarrow ^{13}\text{N} \]

B

I. Plant preparation and treatment

Control (C)

Treated (T)

II. CsCl gradient density centrifugation and peptide preparation

CsCl gradient density fractionation

Urea-based protein extraction

In-solution trypsin digestion

III. Dimethyl labelling and mixing of peptides

Dimethyl labelling

Mixing of peptides

Heavy (H)

Light (L)

Reciprocal (R)

Forward (F)

IV. Chromatographic separation and affinity purification of acetylpeptides

Urea-based protein extraction

In-solution trypsin digestion

CsCl gradient density fractionation

In-solution trypsin digestion

V. LC-MS/MS analysis

VI. Identification

Mascot searching

SQUA-D -based quantification

VII. Quantification

XIC

Affinity purification

SCX fractionation

LC-MS/MS analysis

MS/MS spectra

MS spectra
**Figure 1**

*Dimethyl labelling of the Arabidopsis acetylome, chromatographic enrichment for LC-MS/MS analysis of acetylpeptides, and quantitative identification of acetylation unique PTM peptide arrays (UPAs)*

(A) Chemical reactions for light- and heavy-isotope-coded dimethyl labelling. The N-terminus and lysine side chains of acetylpeptides (Ac-p) can be labelled. (B) Workflow for protein fractionation by CsCl density gradient centrifugation, peptide preparation, dimethyl labelling of the whole-plant acetylome, chromatographic enrichment, LC-MS/MS analysis, Mascot-based database searching, and quantification of acetylation UPAs with SQUA-D (*i.e.*, first 3C procedures of quantitative PTM proteomics). All extracted cellular proteins were separated into three fractions: top (t), middle (m), and bottom (b). The peptides generated from the t, m, and b fractions from control (C) and treated (T) Arabidopsis were labelled with light- (28 Da) or heavy- (34 Da) isotope-coded dimethyl, respectively (*i.e.*, the first C procedure). The mixes of light and heavy isotope-coded dimethyl-labelled peptides from the control and the treated plants, respectively, were defined as forward (F) replicates. Conversely, the mixes of heavy and light isotope-coded dimethyl-labelled peptides from the control and the treated plants, respectively, were defined as reciprocal (R) replicates. Both F (*i.e.*, tF, mF, and bF) and R (*i.e.*, tR, mR, and bR) replicates were produced from a single biological experiment. The peptide mixtures were further fractionated by SCX-HPLC, enriched with anti-acetyllysine antibodies, and analysed by LC-MS/MS (*i.e.*, the second C procedure). The resulting MS data were further analysed by Mascot and SQUA-D software for sequential identification and quantification (*i.e.*, the third C procedure). The Mascot-identified acetylpeptides were evaluated based on the following criteria: FDR ≤ 0.01, where FDR is the false discovery rate obtained by the target-decoy strategy. One biological replicate generates 2 experimental replicates (*i.e.*, F and R). In total, there were 6 experimental replicates (*i.e.*, F1, R1, F2, R2, F3, and R3) from 3 biological replicates, which produced 18 batches of peptide samples (*i.e.*, 3 biological replicates × 3 CDG fractions [t, m, and b] × 2 labelling/mixing types [F and R]). Batch effect adjustment was applied to the batches that have batch effect. The measurement of log-ratios of XIC was followed by statistical evaluation which included both the *t*-test and multiple test correction (*i.e.*, BH-FDR). The significant cut-off criteria were BH-FDR ≤0.1 and |log-ratio| ≥ 0.5×SD, where SD represents the standard deviation of all log-ratios. The detailed identification and quantification workflow (dry laboratory, VI-VII) are shown in Supplementary Figure S1. XIC denotes extracted ion chromatogram.
Previously identified acetylation sites on lysine side chains (3122)

Our identified acetylation sites on lysine side chains (5233)

**C**

AT1G07790 **AEAK**<sup>A</sup> Histone superfamily protein
AT2G37470 **AEAK**<sup>A</sup> Histone superfamily protein
AT3G45980 **AEAK**<sup>A</sup> Histone superfamily protein
AT3G46030 **AEAK**<sup>A</sup> Histone superfamily protein
AT5G02570 **AEAK**<sup>A</sup> Histone superfamily protein
AT5G59910 **AEAK**<sup>A</sup> Histone superfamily protein

AT1G07660 **GKG**<sup>B</sup> Histone superfamily protein
AT2G38810 **GKG**<sup>B</sup> HTA8, Histone H2A 8
AT3G54560 **GKG**<sup>B</sup> HTA11, Histone H2A 11
AT5G59970 **GKG**<sup>B</sup> Histone superfamily protein

AT1G07790 **PKA**<sup>C</sup> Histone superfamily protein
AT2G28720 **PKA**<sup>C</sup> Histone superfamily protein
AT5G22880 **PKA**<sup>C</sup> HTB2, Histone B2
AT5G59910 **PKA**<sup>C</sup> Histone superfamily protein

**D**

2117 1005 4228

Previously identified acetylation sites on lysine side chains (3122)

Our identified acetylation sites on lysine side chains (5233)

**E**

Protein numbers

**F**

CHROMOSOME ORGANIZATION
MACROMOLECULAR COMPLEX
SUBUNIT ORGANIZATION
CELLULAR MACROMOLECULAR COMPLEX SUBUNIT
MACROMOLECULAR COMPLEX ORGANIZATION
NUCLEOSOME ORGANIZATION
CHROMATIN ORGANIZATION
CHROMATIN ASSEMBLY OR DISASSEMBLY
NUCLEOSOME ASSEMBLY
TRANSLATION
RESPONSE TO METAL ION
RESPONSE TO CADMIUM ION
RESPONSE TO INORGANIC SUBSTANCE
RESPONSE TO ABBIOTIC STIMULUS
RESPONSE TO HEAT
RESPONSE TO TEMPERATURE STIMULUS
RESPONSE TO PROTEIN FOLDING
Figure 2

Analysis of Arabidopsis acetylpeptides

(A) Venn diagram showing the identified non-redundant light- and heavy-isotope-coded acetylpeptides (green and brown circles, respectively). (B) All identified non-redundant acetylpeptides from six experimental replicates (i.e., F1, R1, F2, R2, F3, and R3). (C) Conserved acetylation motifs of histone proteins from repeatedly identified acetylpeptides. Asterisks (*) indicate acetylation sites. (D) Venn diagram comparing previously reported acetylation sites and the acetylation sites identified in this study. Blue circles represent previously identified acetylation sites on lysine side chains. Red circles represent the identified lysine side chain acetylation sites from this study. (E) Distribution of acetylation sites per protein. (F) Biological process enrichment for the overly acetylated proteins (OAPs) (with > 8 acetylation sites). The thickness of each line represents the number of overlapping proteins between two processes, where thicker lines indicate more proteins.
A. Scatter plot showing the distribution of acetylation events across different proteins.

B. Heat maps illustrating the relative ion intensity for different proteins.

C. Mass spectrometry data showing the relative ion intensity for different peptides.

D. Protein acetylation sites for HSP 70 family protein (AT3G09440).

E. Protein acetylation sites for EMB2726, translational elongation factor (AT4G29060).

F. Protein acetylation sites for Histone superfamily protein (AT2G28720).

G. Protein acetylation sites for Heat shock protein 90-7 (AT4G24190).

H. Protein acetylation sites for AAC1, ADP/ATP carrier 1 (AT3G08580).

I. Protein acetylation sites for ATPB, ATP synthase subunit beta (ATCG00480).

J. Relative intensity of acetylation for Col-0 and ein3/eil1 mutant lines.

K. Summary table of protein acetylation levels for various proteins.

** and *** denote statistical significance.
Figure 3

Quantitative analysis of the acetylome and immunoblot validation of an identified hormone-regulated acetylation site in Arabidopsis

(A) Quantification results obtained via dimethyl labelling-based extracted ion chromatogram (XIC). In the upper panel, each circle represents a quantified UPA (Supplemental Table S2b). The circles from up-regulated UPAs are marked with red and the circles from down-regulated UPAs are marked with blue. The orange circle indicates the UPA (from histone superfamily protein AT1G07660) verified by immunoblot assay. The x-axis indicates the average of log-ratios from pairs of treatment and control ions. The y-axis indicates -10 times the log-transformed p-value. The significant threshold is BH FDR $\leq 0.1$. The horizontal dashed line indicates the corresponding p-value threshold. Two vertical dashed lines indicate the boundary of $\pm 0.5 \times SD$(standard deviation). The lower panel contains a histogram (green) and a fitted Gaussian distribution (red) of all quantified UPAs. The x-axis indicates the average log-ratio and the y-axis indicates the number of UPA in the corresponding histogram bin. Two vertical dashed lines indicate the boundary of $\pm 0.5 \times SD$(standard deviation). (B-C) MS spectra (B) and MS/MS spectra (C) of the light and heavy isotope-coded dimethyl modified peptides of acetylation UPSP AT1G07660-K6-K9-K13 (orange circle in Figure 3A) which has been demonstrated to be ethylene-suppressed by both XIC and immunoblot assay. (D-I) Profiles of acetylation level of quantified UPAs of Heat shock protein 70 family protein (AT3G09440) (D), translational elongation factor EMB2726 (AT4G29060) (E), Histone superfamily protein (AT2G28720) (F), Heat shock protein 90-7 (AT4G24190) (G), AAC1, ADP/ATP carrier 1 (AT3G08580) (H), and ATP synthase subunit beta (ATCG00480) (I). Green bars indicated that these acetylation UPAs passed the cutoff ($t$ $\leq 0.1$). (J) Immunoblot assay to detect acetylation on K6 of protein AT1G07660. Plants in experiment group were treated with ethylene precursor 1-aminocyclopropane-1-carboxylic acid (ACC) of 10 $\mu$M (orange bar) and untreated plants were used as control (green bar). Actin was also detected as loading control. Three biological replicates and two technic replicates for each biological replicate were performed for each genotype. Average values are shown with error bars ($\pm$SEM (standard error)). Statistical significance was determined using one-sample t-test. * represents $p < 0.05$, ** represents $p < 0.01$ and *** represents $p < 0.001$. (K) Conserved acetylation motifs that were extracted from the validated acetylation sites: K6 of protein AT1G07660. Asterisks (*) stands for acetylation sites.
A. Molecular functions
- A1. structural molecule activity
- A2. DNA or RNA binding
- A3. transporter activity
- A4. electron transport or energy pathways
- A5. response to abiotic or biotic stimulus
- A6. response to stress
- A7. ribosome
- A8. cell wall
- A9. plastid

B. Log2 ratio of GO analysis between acetylproteins and the leaf proteome of Arabidopsis
Figure 4

Bioinformatic analysis of repeatedly identified acetylproteins in Arabidopsis

(A) Comparison between GO analysis of repeatedly identified acetylproteins and the Arabidopsis leaf proteome using the equation in the Experimental Procedures. Categories A1 - A9 represent the up-regulated categories among acetylproteins. Protein categories with the top three log₂ ratios are shown. (B) Biological processes enrichment for the hormone ethylene-regulated acetylproteins. The thickness of each line represents the number of overlapped proteins between two processes, where thicker lines indicate more proteins.