Comprehensive analysis of the lysine succinylome and protein co-modifications in developing rice seeds

Xiaoxi Meng¹, Hana Mujahid¹, Yadong Zhang¹,², Xiaojun Peng³, Edilberto D. Redoña⁴, Cailin Wang², *, Zhaohua Peng¹, *

¹ Department of Biochemistry, Molecular Biology, Entomology and Plant Pathology, Mississippi State University, Starkville MS, 39762, USA
² Institute of Food Crops, Jiangsu Academy of Agricultural Sciences, Jiangsu High Quality Rice Research and Development Center, Nanjing Branch of China National Center for Rice Improvement, Nanjing 210014, China
³ Department of Bioinformatics, Jingjie PTM Biolab Co. Ltd, Hangzhou 310018, China
⁴ Delta Research and Extension Center, Stoneville MS, 38776, USA

* Corresponding author.
E-mail: zp7@BCH.msstate.edu (ZP), clwang@jaas.ac.cn (CW)

Running title: Lysine Succinylome of Developing Rice Seeds
ABBREVIATIONS:
Ksu: lysine succinylation. PTM: post-translational modification. nano-LC-MS/MS: nanoscale liquid chromatography coupled to tandem mass spectrometry. Kac: lysine acetylation. Kmal: lysine malonylation. Kcr: lysine crotonylation. Khib: lysine 2-hydroxyisobutyrylation. dpa: days post-anthesis. NAA: naphthylacetic acid. DTT: dithiothreitol. IAA: iodoacetamide. TFA: trifluoroacetic acid. FA: formic acid. ACN: acetonitrile. NSI: nano electrospray ionization. NCE: normalized collision energy. AGC: Automatic gain control. FDR: false discovery rate. GO: Gene Ontology. KEGG: Kyoto Encyclopedia of Genes and Genomes. PPIs: protein-protein interactions. KGDHC: E2 subunit of α-ketoglutarate dehydrogenase.

ABSTRACT:
Lysine succinylation has been recognized as a post-translational modification (PTM) in recent years. It is plausible that succinylation may have a vaster functional impact than acetylation due to bulkier structural changes and more significant charge differences on the modified lysine residue. Currently, however, the quantity and identity of succinylated proteins and their corresponding functions in cereal plants remain largely unknown. In this study, we estimated the native succinylation occupancy on lysine was between 2% to 10% in developing rice seeds. Eight hundred fifty-four lysine succinylation sites on 347 proteins have been identified by a thorough investigation in developing rice seeds. Six motifs were revealed as preferred amino-acid sequence arrangements for succinylation sites, and a noteworthy motif preference was identified in proteins associated with different biological processes, molecular functions, pathways, and domains.
Remarkably, heavy succinylation was detected on major seed storage proteins, in conjunction with critical enzymes involved in central carbon metabolism and starch biosynthetic pathways for rice seed development. Meanwhile, our results showed that the modification pattern of in vitro nonenzymatically succinylated proteins was different from those of the proteins isolated from cells in Western blots, suggesting that succinylation is not generated via nonenzymatic reaction in the cells, at least not completely. Using the acylation data obtained from the same rice tissue, we mapped many sites harboring lysine succinylation, acetylation, malonylation, crotonylation, and 2-hydroxisobutyrylation in rice seed proteins. A striking number of proteins with multiple modifications were shown to be involved in critical metabolic events. Given that these modification moieties are intermediate products of multiple cellular metabolic pathways, these targeted lysine residues may mediate the crosstalk between different metabolic pathways via modifications by different moieties. Our study exhibits a platform for extensive investigation of molecular networks administrating cereal seed development and metabolism via PTMs.

**KEYWORDS:** Post-translational modification; Lysine succinylation; Succinylome; Rice; Storage nutrient; Seeds

**INTRODUCTION:**

Post-translational modifications (PTMs) are covalent modifications that transpire during or after protein biosynthesis. Lysine succinylation (Ksu) is an evolutionarily-conserved PTM (1, 2) that attaches a succinyl group (-CO-CH₂-CH₂-CO-) to a protein lysine residue (2). The addition of a succinyl group induces a mass shift of +100.0186 Da and generates a negative charge on lysine
residue under physiological pH (2, 3). Since succinylation results in a bulkier structural change and more significant charge difference on lysine, one could postulate that it generates a grander impact on the substrate protein’s structures and functions in comparison to well-studied lysine acetylation and methylation (1, 2).

Succinyl-CoA levels and E2 subunit of α-ketoglutarate dehydrogenase (KGDHC) have emerged as the primary regulators of protein succinylation through non-enzymatic and enzymatic ways (4-6), respectively. The succinylation effectiveness of KGDHC is superior in comparison to succinyl-CoA alone (6). SIRT5 and SIRT7 are mammalian sirtuins of class III family histone deacetylases, and they were identified as the key enzymes for lysine desuccinylation in cells and tissues (7-10). Apart from desuccinylation, SIRT5 exhibits broader activities for demalonylation and deglutarylation, but it demonstrates low activity for deacetylation (8, 9, 11, 12). Moreover, CobB, a known Sir2-like prokaryotic deacetylase, can catalyze both deacetylation and desuccinylation in *E.coli* (13).

Prior research has suggested that lysine succinylation is a pervasive modifier among histone and non-histone proteins (2, 4, 7, 13-25). Recently, the advancements in the mass spectrometry technology and succinyl-peptides enrichment methods facilitated the identification of hundreds to thousands of succinylation sites in both prokaryotes and eukaryotes (4, 7, 13-15, 17, 19-25). The proteome analysis of lysine succinylation has been reported in *E.coli* (2, 4, 13), *B.subtilis* (14), *V.parahaemolyticus* (19), *M.tuberculosis* (17, 20), *H.sapiens* (4, 11), *M.musculus* (4, 7, 8, 11), *S.cerevisiae* (4), and *T.gondii* (15). This vital aspect has substantially extended our understanding of protein succinylation. Plant succinylomes have also been reported in tomato seedlings (21), rice
germinating embryos (22), *B. distachyon* seedling leaves (23), common wheat (24), and hybrid *Taxus* species (25). However, possible protein lysine co-modification by various modifications has not been explored in plants.

Rice is one of the most significant cereals as it serves as the staple food for over half of the world’s population (26). In rice grain, the majority of nutrients are stored in the form of starch, lipid, and protein, which extensively contribute to grain nutritional value, milling properties, appearance, and cooking quality (27). The content and composition of storage starch and protein are directly associated with seed development. Recently, PTMs of lysine acetylation (28), malonylation (29), and 2-hydroxyisobutyrylation (30) have been reported in developing rice seeds. In this report, we successfully identified 854 lysine succinylation sites across 347 proteins with a false discovery rate (FDR) of ≤ 1% in developing rice seeds. Our results indicate that lysine succinylation is a highly conserved modification. It frequently occurs in the rice proteome with a preference on carbon metabolic pathways, starch biosynthetic pathways, and the major seed storage proteins. Furthermore, a large number of proteins involved in crucial metabolic processes were revealed to embrace various modifications in lysine residues. This analysis provides a comprehensive view of lysine modifications in developing rice seeds.

**EXPERIMENTAL PROCEDURES**

*Plant materials and growth conditions*

Rice (*Oryza sativa* L. *japonica* cv. Nipponbare) leaves and roots were sampled from 20-day-old seedlings grown in an incubator at 28 °C (16-h-day/8-h-night). The flowers, pollen, 7, 15, and 21
days post-anthesis (dpa) developing rice seeds and mature rice seeds were collected from rice plants grown in a greenhouse of the Department of Biochemistry and Molecular Biology, Mississippi State University, MS, USA. The cultured cells are rice (*Oryza sativa* L. *japonica* cv. Nipponbare) NB2P suspension cell cultures, which were maintained as reported (31).

**Protein extraction**

Proteins were isolated using a phenol extraction method (28-30, 32, 33). The ground plant organ/tissue was mixed with an extraction buffer (0.9 M sucrose, 0.5 M Tris-HCl pH 8.7, 0.05 M EDTA, 0.1 M KCl, and 2% β-mercaptoethanol), combined with the subsequent addition of an equal volume of saturated phenol (pH 8.0), and homogenization for 30 minutes at 4°C. The phenol phase was recovered from homogenate by centrifugation at 5,000 g for 15 minutes at 4°C. The phenol extraction procedure was repeated three times. The final collection of phenol was mixed with five volumes of precipitation buffer (methanol with 0.1 M ammonium acetate and 1% β-mercaptoethanol) and resided overnight at −80°C for precipitation. The crude protein was obtained by centrifugation at 15,000 g for 15 minutes at 4 °C. Afterward, the protein pellet was washed three times with cold precipitation buffer, accompanied by three-time additional washes with ice-cold 70% ethanol. The pellet was lyophilized in a speed vacuum and stored at -80 °C before use.

**Nonenzymatic succinylation by succinyl-CoA in vitro**

Proteins isolated from 15 days post-anthesis (15 dpa) developing rice seeds were prepared in PBS buffer (pH 7.2) at 10mg/ml. A total of 100μl (1mg) protein was mixed with freshly prepared succinyl-CoA sodium salt (S1129, Sigma) in ddH2O to a final concentration of 0.5mM or 1mM. Reactions were incubated at 28 °C for 12 h with occasional shaking. Sequentially, the proteins
were purified by Sep-Pak C18 Plus Short Cartridge (WAT020515, Waters), and were lyophilized in a speed vacuum and stored at -80 °C before proper usage. All succinyl-CoA sodium salt treatments were performed with three technical replicates.

**Western blot analysis**

Protein samples extracted from different rice organs/tissue were separated by 12% SDS-PAGE and transferred onto PVDF membrane (EMD Millipore) for Western blot. Succinyalted proteins were detected by utilizing rabbit-derived pan anti-succinyl lysine antibody (PTM-401, PTM Biolabs) in a 1:1000 (v/v) dilution overnight at 4°C, accompanied by gentle shaking in reference to the supplier’s instruction.

**Trypsin digestion**

The lyophilized 15 dpa rice seed protein was re-dissolved in a buffer (8 M urea, 100 mM NH₄HCO₃, pH 8.0), and the protein content was determined with 2-D Quant kit (GE Healthcare). Before digestion, the protein was reduced with 10 mM DTT (Dithiothreitol) for 1 hour at 37 °C, and subsequently alkylated with 20 mM IAA (Iodoacetamide) for 45 minutes in dark setting at room temperature. For trypsin digestion, the protein sample was diluted with an addition of 100 mM NH₄HCO₃ for urea reduction to the desired concentration of less than 2M. Sequencing-grade trypsin (V5111, Promega Corporation) was supplemented at a 1:50 (w/w) enzyme-to-substrate mass ratio for an overnight digestion period and likewise at a 1:100 (w/w) enzyme-to-substrate mass ratio for an additional 4 hour digestion period at 37 °C.

**HPLC fractionation and affinity enrichment of lysine succinyalted peptides**

The peptides were fractionated into 12 fractions by high pH reverse-phase HPLC with Agilent 300
Extend C18 column (5 μm particles, 4.6 mm ID, 250 mm length, Agilent) after digestion. Briefly, peptides were initially separated into 80 fractions in a gradient of 2% to 60% acetonitrile in 10 mM ammonium bicarbonate (pH 10). Then, the peptides were further combined into 12 fractions in a noncontiguous manner as reported (34). For succinylated peptides enrichment, fractionated peptides were dissolved once more in NETN buffer (100 mM NaCl, 1 mM EDTA, 50 mM Tris-HCl, 0.5% NP-40, pH 8.0) and incubated with pre-washed anti-succinyl-lysine agarose beads (PTM-402, PTM Biolabs) at 4°C overnight, accompanied by gentle oscillation. Afterward, the beads were rinsed four times with NETN buffer and twice with ice-cold ddH2O. Enriched peptides were eluted by 0.1% TFA (Trifluoroacetic acid) from the beads and purified with C18 ZipTips column (EMD Millipore).

**Protein identification by nano-HPLC/MS/MS**

The lyophilized peptides were dissolved in 0.1% formic acid (FA) and 2% acetonitrile (ACN). Peptide analysis was conducted by nano-HPLC/MS/MS, utilizing an EASY-nLC 1000 UPLC system (Thermo Fisher Scientific) attached to a Q-Exactive Plus Hybrid Quadrupole-Orbitrap mass spectrometer (Thermo Fisher Scientific). Briefly, the peptides were eluted onto a reversed-phase analytical column (2 μm particles, 50μm ID, 15cm length, Acclaim PepMap RSLC, Thermo Fisher Scientific). The utilization of a consistent flow rate of 280 nl/min with a 16-minute linear gradient from 7% to 18% of the solvent buffer of 0.1% FA in 98% ACN was the initial response. A concentration increase of 18% to 22% for an 8 minute period and a concentration increase of 22% to 35% for another 8 minutes was achieved. The ascending concentration grasped 80% in another 5 minutes and immobilized at this value of the solvent buffer for the remaining 3 minutes.
The resulting peptides were subjected to a NanoSpray Ionization (NSI) source, succeeded by tandem mass spectrometry in Q Exactive™ Plus (Thermo Fisher Scientific) connected online to the UPLC. Intact peptides were detected at a resolution of 70,000 in the Orbitrap, and peptides were selected for MS/MS with the NCE setting to the preferred quantity of 30. Ion fragments were detected at a resolution of 17,500. A data-dependent procedure that alternated between one MS scan and 20 MS/MS scans was applied for the highest 20 precursor ions exceeding a threshold ion count of $1.0 \times 10^4$ in the MS survey scan with 15.0s dynamic exclusion. The operated electrospray voltage was 2.0 kV, and automatic gain control (AGC) was applied to avoid ion trap overcapacity. Ions with charge states of 2 to 5 were permitted, and $5 \times 10^4$ ions were collected for the generation of MS/MS spectra. The m/z scan range was approximately 350 to 1800 for MS scans.

The protein and succinylation sites were identified by MaxQuant software (http://www.maxquant.org/) coupled with Andromeda search engine (v.1.4.1.2) (35). The mass spectra of the raw data were searched against UniProt _Oryza sativa japonica_ database (63,195 sequences, released July 2014) concatenated with reverse decoy database. Trypsin/P was specified as a cleavage enzyme, allowing a maximum of 4 missing cleavages, 5 modifications per peptide, and 5 charges. Carbamidomethylation on cysteine was distinguished as a fixed modification. Oxidation on methionine, succinylation on both lysine and protein N-terminal were specified as variable modifications. The mass error was programmed to 10 ppm for precursor ions and 0.02 Da for fragment ions. The false discovery rate (FDR) thresholds for protein, peptide, and modification site were adjusted to a quantity of 1% (36), and the minimum peptide length was changed to a quantity of 7. Lysine succinylation sites identified with a localization probability of $< 0.75$ were
eliminated. The remaining parameters in MaxQuant were set to default values.

**Bioinformatics analysis of lysine succinylated peptides and proteins**

Gene Ontology (GO) annotation was obtained from the UniProt-GOA database (www. http://www.ebi.ac.uk/GOA/). If some identified proteins were not annotated by UniProt-GOA database, the InterProScan software was used based on protein sequence alignment method. Kyoto Encyclopedia of Genes and Genomes (KEGG) database was used to annotate protein pathway. Domain annotation was performed by using InterProScan based on protein sequence alignment on the InterPro domain database via Web-based interfaces and services. The functional annotation tool, DAVID bioinformatics resources 6.7, against the background of rice (*Oryza Sativa*) was employed for GO, KEGG pathway, and domain enrichment analyses. A two-tailed Fisher’s exact test was used to examine the identified succinylated proteins. Multiple testing correction was performed using the Benjamini–Hochberg false discovery rate (FDR) control method. The GO term, KEGG pathway term and domain categories with a corrected p-value < 0.05 and enrichment fold ≥ 2 were determined as significantly enriched.

Conserved amino-acid sequence motifs of succyl-21-mers (ten amino acids upstream and downstream of the succinylation site) were analyzed by motif-x (http://www.motif-x.med.harvard.edu). The entire protein sequence of *Oryza sativa* L. *japonica* was utilized as the background database. For motif logo-based clustering analysis, the −log10 (p-value) were z-transformed for each category (p-value < 0.05) and gathered by one-way hierarchical clustering in Genesis. Cluster membership was visualized by heat map through the “heatmap.2” function, located in the “gplots” R-package.
A BLAST analysis was accomplished by operating the two-directional BlastP software with BLOSUM62 matrix, with the objective of evaluating succinylated protein conservation across selected succinylomes of diverse organisms. BLAST outcomes that encompassed a corresponding e value < 1E-10, a score > 200, and an identity > 30 were taken as the best-scoring homologous proteins.

Protein-protein interaction networks were constructed against the STRING database (version 10.5). Evidence for protein-protein interaction was obtained from seven sources: experiments, databases, text mining, co-expression, neighborhood, co-occurrence, and gene fusion. Identified interactions with confidence scores ≥ 0.7 (high confidence) were obtained. Interaction networks from STRING were visualized with Cytoscape software.

**Experimental design and statistical rationale**

Proteome-wide succinylation of rice (Oryza sativa L. japonica cv. Nipponbare) was investigated by using 15 dpa developing seeds selected from 5 individual plants. Succinylated peptides were enriched using immunoaffinity enrichment strategies from 12 HPLC fractions and analyzed by high accuracy nano-HPLC/MS/MS. Since the current research aims to detect the modification sites instead of performing quantitative analysis among different tissues or treatments, no analyses with replicates were involved. Instead, we used high stringent screening criteria for mass spectrometry. The mass error was programmed to 10 ppm for precursor ions and 0.02 Da for fragment ions. FDR thresholds for proteins, peptides, and modification sites were set at 1% using MaxQuant software (36). Minimum peptide length was set at 7. Lysine succinylation sites identified with a localization probability of < 0.75 were eliminated. Succinylated proteins were annotated by GO, KEGG, and
domain enrichment analyses. A two-tailed Fisher’s exact test was used to examine the identified succinylated proteins. Multiple testing correction was performed using the Benjamini–Hochberg FDR control method. The GO term, KEGG pathway term and domain categories with a corrected p-value < 0.05 and enrichment fold ≥ 2 were determined as significantly enriched. For motif logo-based clustering analysis, the −log10 (p-value) were z-transformed for each category (p-value < 0.05) and gathered by one-way hierarchical clustering in Genesis. Protein in vitro nonenzymatic succinylation was achieved by succinyl-CoA sodium salt treatments. Quantitative comparative analysis of succinylation level was revealed by Western blot with three technical replicates.

RESULTS AND DISCUSSION

Protein lysine succinylation in different rice tissue/organs

To detect the status of lysine succinylation in rice proteomes, proteins from nine different tissue/organs were examined by Western blot. These distinctive tissue/organs included cultured cells, roots, leaves, flowers, pollens, 7 dpa seeds, 15 dpa seeds, 21 dpa seeds, and mature dry seeds. Specific antibodies for succinylated lysine residue were utilized under equal protein loading conditions (Fig. 1 and Supplementary Fig. S1). Multiple protein bands possessing different sizes and smears were detected from all tested samples (Fig. 1b). It was evident that different tissue/organs possessed different patterns of protein succinylation modification. These discoveries demonstrated that lysine succinylation is a highly specific modification system and widely presented in the rice plant.

Cellular proteins displayed a distinct succinylation pattern compared to in vitro non-
enzymatically succinylated proteins in Western blots

Succinylation could occur via a chemical mechanism dependent on succinyl-CoA concentration and pH (4, 5). Using pan anti-succinyl lysine antibody, we observed that incubating proteins isolated from developing rice seeds with 0.5 mM or 1mM of succinyl-CoA sodium salt under high pH condition (pH7.2) caused a significant increase in lysine succinylation compared with native rice proteins without succinyl-CoA treatment (Fig. 2a, compare lanes 1 and 2 with lanes 3 and 4). We confirmed via Western blot that succinyl-CoA could non-enzymatically succinate rice proteins in vitro under certain pH. Western blot signals for the succinyl-CoA treated samples tended to be saturated (Fig. 2a, lane 3 and 4), and we speculate that all proteins were fully succinylated in the artificially treated samples.

To estimate global protein succinylation occupancy in developing rice seeds, the initial quantity of 1mM succinyl-CoA sodium salt-treated protein sample was reduced to 60%, 45%, 30%, 15%, 10%, 5%, 2%, and 1% for Western blot analysis (Fig. 2a, lane 5-12). The succinylation abundances were quantified based on the two distinguished bands (indicated in Fig. 2a) on the Western blot for each lane by using Image Studio Lite tool. After removal of the background, succinylation intensities of the upper band of native developing rice seed proteins dissolved in either SDS or PBS buffer (Fig. 2a, lane 1 and 2) were between the intensity presented in lane 9 and lane 10 (Fig. 2a), which was 10% and 5% amount of the protein sample saturated with succinylation (Fig. 2b), respectively. With the removal of the background, the succinylation intensities of the lower band of native developing rice seed proteins were slightly higher or close
to the intensity presented in lane 11 (Fig. 2a), which was 2% amount of the developing rice seeds protein saturated with succinylation (Fig. 2b). Three technical repeats were conducted and shown in Supplementary Fig. S2. This observation indicated that the native succinylation occupancy was between 2% to 10% in developing rice seeds.

Comparing the band patterns in lanes 1 and 2 with lanes 11 in the Western blot, it is clear that lane 11 is less intense than lanes 1 and 2 in the thick upper protein band (Fig. 2a and b). However, multiple weaker bands with larger protein size can be seen in lane 11 but not in lanes 1 or 2. Therefore, the results demonstrated that chemical mediated succinylation occurs more evenly to all protein bands, while native succinylation occurs more selectively - only occurring on specific protein bands. Understanding the global succinylation occupancy of lysine residues is essential for understanding succinylation functions in rice plant. Using SWATH-MS2 methodology, the majority of lysine succinylation modifications in wild-type E.coli proteins showed less than 2% occupancy (37). Our approximation of succinylation occupancy estimated in developing rice seeds was between 2% to 10%.

**Global profiling of succinylated peptides and proteins**

To examine the succinylated peptides on a global scale, the 15 dpa developing rice seeds were collected for protein extraction. After trypsin digestion, the proteins were fractioned into 12 fractions to reduce the complexity of the peptide samples. Then, the succinylated peptides were enriched with succinyl-lysine specific antibodies and analyzed by nano-HPLC/MS/MS.

Eight hundred fifty-four succinylation sites correlated with 347 proteins were identified
retaining an FDR of $\leq 1\%$ and recorded in Supplementary Table S1. The raw data and annotated MS spectra have been deposited to PRIDE with the accession number of PXD005582 (30) and PXD013664, respectively. Representative mass spectra of succinylated peptides were shown in Supplementary Fig. S3. The majority of ascertained peptides were determined to have fluctuating lengths between 8 and 20, which complements the distinct property of tryptic peptides, suggesting that the sample preparation achieved the standard quality (Supplementary Fig. S4). To evaluate the distribution of succinylation sites in developing rice seed succinylome, the number of modification sites in each protein was calculated (Supplementary Fig. S4). Among the identified 347 succinylated proteins, approximately 52% (181/347) modified proteins contained a single putative succinylation site, and the average degree of succinylation was 2.46 (854/347). Remarkably, endoplasmic reticulum chaperone protein disulfide isomerase-like 1-1 (PDI) (Q53LQ0) possessed the most substantial succinylated lysine sites of 15 in a single protein. The function of the modification on this protein remains to be tested.

To investigate the coverage of succinylated lysine residues in a modified protein, the frequency of succinylated lysines within each modified protein was calculated by dividing the number of succinylated lysine residues with a total number of lysine residues and the results were recorded in Supplementary Figure S4. Most identified proteins had fewer than 20% of the lysine residues experiencing succinylation. However, 22 proteins with succinylated lysine residues exceeded 40% of the total lysines, and 2 proteins obtained succinylated lysine rates surpassing 70% of the total lysine residues. One of the two exceptional proteins is non-specific lipid-transfer protein 1(Q7XBA6). Observations in *Mycobacterium tuberculosis* and mice suggested a
connection between lysine succinylation and fatty acid metabolism (7, 17, 38). Sirt5-deficient mice exhibited accumulated lysine succinylation level and defective fatty acid metabolism (38). It will be interesting to investigate the connection between non-specific lipid-transfer protein succinylation and fatty acid metabolism in plants.

**Characterization of lysine succinylated peptides**

To understand the regulation and amino acid residue preference at the sites surrounding the succinylated lysine, we carried out succinylation site motif analysis by examining the sequences from -10 to +10 of the 854 succinylation sites. Six distinguished motifs were identified: KsuXXXXR, KsuXI, KsuXXR, LKsu, KsuXXXXXXXXR, and KXXXXKsu (Ksu indicates the succinylated lysine, and X indicates a random amino acid residue) (Fig. 3a). The frequency of amino acid residues flanking succinylated lysine was analyzed to investigate the enrichment or depletion of various amino acids (Fig. 3b). Arginine at positions from +4 to +9, especially at +4, +5 and +7, and isoleucine at +2 position, leucine at -1 position, were substantially preferred and these patterns agreed with the identified conserved motifs reported in this study. Interestingly, lysines with an arginine positioned at -1 or another lysine located from -2 to +2 have the lowest frequency to be succinylated according to flanking sequence analysis (Fig. 3b). Remarkably, motif LKsu had also been identified in the succinylomes of *E.coli* BW25113, *T.gondii*, mouse liver mitochondria, and *S.cerevisiae* (4, 7, 15). Motif KsuXI was also reported in *B.subtilis* succinylome (14). In addition, some enriched amino acids identified in this study also shared high similarities with the preferred sequence around the succinylation sites in mouse liver, *V.parahaemolyticus*, *M.tuberculosis*, *S.lycopersicum*, *T.gondii* and *B.distachyon* although the precise residue position
might be slightly different (4, 15, 19-21, 23), suggesting some degree of conservation in the succinylation sites in both prokaryotes and eukaryotes.

A noteworthy motif preference was discovered in proteins associated with different biological processes, molecular functions, cellular components, pathways, and domains (Fig. 3c-e). For example, the proteins expressing binding activities prefer to harbor motifs of KXXXXKsu and KsuXI (Fig. 3c). The storage-related proteins in developing rice seeds habitually adopt cupin 1 domain, RmlC-like cupin domain, and RmlC-like jelly roll fold domain. They prefer to harbor motifs containing arginine (KsuXXXR, KsuXXXXR, KsuXXXXXXXR) (Fig. 4e). The recognition of conserved motifs and motif preferences in different protein clusters verifies that protein lysine succinylation is a highly regulated modification process instead of non-specific modifications, and its function is influenced by the environment of neighboring amino acid residues.

**Enrichment analyses of lysine succinylated proteins**

Enrichment analyses of GO annotations, KEGG pathways, and protein domains were performed to expand our knowledge on the biological regulations and functions of succinylated proteins in developing rice seeds (Supplementary Fig. S5 and Supplementary Table S2).

GO enrichment analysis indicated that proteins related to molecular functions of oxidoreductase activities, nutrient reservoir activities, binding activities, translation elongation factor activity, antioxidant activity, etc. were significantly enriched in rice succinylome with corrected p-value < 0.05 and fold enrichment ≥ 2 (Supplementary Fig. S5). Moreover, GO
enrichment analysis suggested that the succinylated proteins in developing rice seeds are significantly enriched in various biological processes of biotic/abiotic stresses and responses to inorganic/ionic substances (Supplementary Fig. S5). Therefore, succinylation may have a fundamental engagement in stress regulatory procedures and defense responses.

KEGG pathway enrichment analysis substantiated the idea that most succinylated proteins were essential participants in carbon, energy, and amino acid metabolic pathways (corrected p-value < 0.05 and fold enrichment ≥ 2) (Supplementary Fig. S5). Outstandingly, succinylated proteins recognized in other reported eukaryotic and prokaryotic succinylomes were acknowledged as significant contributors in analogous pathways (7, 13-15, 17, 19-21). Perception of this data portion provided indications of a certain degree of conservation embraced by succinylation among mammals, microbes, and plants. Enzymes involved in fatty acid oxidation pathways were strikingly succinylated in mouse liver cells (4, 7). However, the succinylation status of the homologous proteins in developing rice seeds is incompatible with mouse liver cells, suggesting that a particular extent of functional differentiation in succinylation may transpire between plants and mammals.

Protein domain enrichment analysis for succinylated proteins showed that domains of 11-S seed storage protein conserved site, RmlC-like jelly roll fold, RmlC-like cupin, plant 11-S storage protein, and cupin 1 were significantly enriched (corrected p-value < 0.05 and fold enrichment ≥ 2) (Supplementary Fig. S5), and these domains were consistently positioned in the nutrient reservoir proteins of seeds.

*Lysine succinylated proteins in central carbon metabolism*
Extensive evidence revealed that lysine succinylation targets metabolic enzymes for carbon source utilization and metabolic flux coordination in both prokaryotes and eukaryotes (4, 7, 14). Central carbon metabolism (TCA cycle and glycolysis/gluconeogenesis pathway) accomplishes oxidative degradation of fatty acids, amino acids, and monosaccharides for energy production. Grain filling is mainly a process of carbohydrate accumulation. Thus, carbon metabolism is crucial for the development of cereal seeds. Gluconeogenesis pathway generates glucose as the primary substrate for starch biosynthesis. Proteomic and transcriptional studies revealed that protein expression in glycolysis and TCA cycle was less abundant in inferior spikelets than in superior spikelets during grain filling (39, 40).

In developing rice seeds, the majority of enzymes engaged in TCA cycle were succinylated, and most were additionally modified by acetylation (28), malonylation (29), 2-hydroxyisobutyrylation (30) and crotonylation (Meng et al., unpublished result) (Fig. 4). In yeast, the upregulation of the enzyme Kgd1 (α-ketoglutarate dehydrogenase, which catalyzes the formation of succinyl-CoA) or the blocking of the enzyme Lsc1 (succinyl-CoA ligase, which catabolizes the succinyl-CoA in TCA cycle), triggered elevated succinylation levels. This investigation insinuates that succinyl-CoA produced in the TCA cycle alters succinylation levels throughout the cells (4). Meanwhile, Gibson et al. revealed that the E2 subunit of KGDHC could act as succinyltransferase, inducing succinylation in an α-ketoglutarate-dependent manner (6). Additionally, KGDHC stimulates fumarate hydratase activities and its succinylation levels, but it inhibits isocitrate dehydrogenase activity in the TCA cycle (6). The gathered investigation data indicates that the TCA cycle is the critical donor for global succinylation in cells through either a
succinyl-CoA-guided non-enzymatic mechanism or a KGDHC-guided enzymatic mechanism. In our examinations, all three KGDHC subunits in the TCA cycle were succinylated in plants. Given that conserved motifs and motif preferences were observed in different protein clusters of succinylated proteins and the patterns of artificially succinylated proteins were different from isolated cellular proteins in Western blot, our results suggest that protein lysine succinylation in developing rice is a highly regulated modification process instead of non-specific modifications.

The multitude of glycolysis/gluconeogenesis pathway enzymes was discovered to harbor modification of succinylation (Fig. 4). And these enzymes extensively overlapped with enzymes possessing acetylation, malonylation, 2-hydroxyisobutyrylation, and crotonylation (Fig. 4). Pyruvate dehydrogenase complex (PDC) catalyzes the conversion of pyruvate into acetyl-CoA. Thus, PDC links the glycolysis metabolic pathway to the TCA cycle. We discovered 5 succinylation sites on the PDC E1 α subunit (PDHA, pyruvate dehydrogenase α subunit), 16 sites on the E2 subunit (dihydrolipoyl transacetylase, DLAT), and 11 sites on the E3 subunit (dihydrolipoyl dehydrogenase, DLD). Moreover, the E1 and the E3 subunits are additionally modified by acetylation (Fig. 4). In cancer cells, phosphorylation and acetylation play an imperative role in the joint regulation of PDHA activities. The varying acetylation status of PDHA directly changes PDC activities, which stimulates glycolysis in cancer cells (41, 42). Furthermore, decreasing the succinylation levels by KGDHC inhibition leads to a substantial reduction of PDC activities and acetylation levels in neuronal cell lines (6). Down-regulation of succinylation levels may hinder PDC E2 (operating as acetyltransferase) activities, consequently preceding towards acetylation diminution (6). Broadening investigation is intriguing if succinylation and acetylation
work simultaneously in PDC and glycolysis pathway regulation in correspondence to phosphorylation and acetylation.

**Succinylation of starch biosynthesis enzymes and regulators**

In rice grain, the majority of accumulated carbohydrates is in starch form, significantly contributing to its nutritional values, milling properties, appearances, and cooking qualities (27). The constituents and contents of storage starch are predominantly determined during seed development. We found that the enzymes in the starch synthesis and regulatory pathways, including ADP-glucose pyrophosphorylases (OsAGPS2 and OsAGPL2), starch branching enzymes (OsBEI and OsBEIIb), starch debranching enzymes (OsPUL), starch phosphorylase (OsPHOL), sucrose synthase (SUS2 and SUS3), UDP-glucose pyrophosphorylase (UGP) and phosphoglycerate mutase (PGM) were succinylated in developing rice seeds (Table 1 and Fig. 4). Moreover, many of these starch biosynthesis proteins were found subjected to acetylation, crotonylation, malonylation, and 2-hydroxyisobutyrylation (Fig. 4). OsAGPs catalyzes the initial step of starch biosynthesis, which has correlations with rice grain filling rates and starch accumulations (43). OsBEs, OsPUL, and OsPHOL catalyze amyllopectin formation. Starch quality regulator FLO4 was similarly succinylated (Table 1). FLO4 mutation induced the formation of smaller grain dimensions and floury-white endosperms (44). Since starch biosynthetic enzymes are succinylated, it is commendable to intensify upcoming evaluations that delve into the concept of potential regulatory responsibilities of succinylation during starch synthesis in seed development.
Storage proteins are targeted by succinylation

Storage proteins and starches are both core components of rice endosperm. Seed storage proteins are not only critical as a nutritional reservoir for rice grain, but they also participate as vital machinery for seed germination, initial seedling development, endosperm structure, and culinary value (27, 45). For our investigations, 13 proteins with storage specialization were succinylated, including 9 rice glutelin proteins, 3 globulin proteins, and 1 patatin-like storage protein (Table 2). Glutelins are storage proteins that are encoded explicitly by 15 genes in the rice genome, and they are considered the prevailing storage proteins in rice (27, 46). GluA-1, GluA-2, GluA-3, GluB-1a, GluB-5, GluB-6, GluB-7, GluC-1, and GluD-1 are 9 of the 15 constituting glutelin proteins that are succinylated. Further analysis showed that succinylations were detected in over 30% of the lysine residues of storage proteins GluA-1, GluA-3, GluB-1a, GluB-5, GluC-1, GluD-1, and globulin (Q75GX9), respectively.

A growing number of functions of PTMs on protein stability, localization, and activity have been reported (47-49), but the role of storage protein succinylation is not clear thus far. Storage proteins are nutritional reservoirs for rice seeds, but they lack active positions in cellular metabolism and cellular development. Storage proteins possessing PTMs will be vastly investigated to determine if their obligations are confined to protein trafficking or protein accumulation within organized protein bodies.

Comparison of succinylome in different organisms

Global proteome analysis for lysine succinylation in germinating rice embryo identified 665
succinylated sites on 261 proteins (22). Compared with the 854 succinylation sites reported in this study in developing seeds, we identified 236 common sites on 141 succinylated proteins in the two rice succinylome studies (Supplementary Table S3). The heavy succinylation for glutelins in developing rice seeds was found absent in germinating embryo succinylomes. Moreover, protein biosynthetic enzymes, protein disulfate isomerase (Q53LQ0), aspartate aminotransferase (P37833), and alanine aminotransferase (Q338N8) contain grander modification sites in developing rice seeds in comparison with germinating embryo. Many sites on starch biosynthetic enzymes (OsAGP, OsBE, OsDBE, and SUS) are succinylated in developing rice seeds, but they are not succinylated in germinating rice embryos (22).

A BLAST analysis was performed to recognize conserved succinylated proteins possessed by various organisms. The proportion of conserved succinylated proteins was substantial in comparison to the succinylomes of *B.distachyon* L. (23) and *S.lycopersicum* (21) (Table 3 and Supplementary Table S3). We discovered that proteins for carbon metabolism, amino acid biosynthesis, and oxidative phosphorylation pathways are conserved in plant succinylomes. Orthologs of 131, 75, 39, and 68 succinylated proteins of developing rice seed were detected in the succinylomes of Animalia, Fungi, Protista, and Bacteria, respectively. Moreover, 18 developing rice seed proteins contain succinylated homologs in all examined kingdoms. Among these conserved proteins, 11 proteins are classified as ATP synthases and DnaK family proteins. Two are classified as aldehyde dehydrogenases with involvement in the cellular response to oxidative stress. One is classified as adenylate kinase enzyme with specializations in energy homeostasis and cell growth. Three are TCA cycle enzymes (OGDH, DLST, and IDH) which
catalyze succinyl-CoA formation; DLST additionally operates as succinyltransferase (6). Ortholog detection in the succinylomes of various kingdoms provides noteworthy evidence of succinylation conservation across diversified organisms.

Proteins co-modified by succinylation, acetylation, malonylation, crotonylation, and 2-hydroxyisobutyrylation

Emerging evidence proposes that PTMs may incorporate crosstalk and substitutions between alterations, which contributes to fluctuating functional consequences of proteins (47). Certain lysine acylation regulatory enzymes could have influences on various lysine acylation categories in a comparable manner (1, 8, 12, 50, 51). A plausible notion is that the active acetyltransferases randomly utilize acyl-CoAs to modify lysine substrates, and diverse lysine modifications may experience competition for protein lysine sites during specific biological regulations. Proper investigations were performed to determine if rice proteins, as well as individual lysine sites, could have the capability of sheltering distinctive PTMs. We achieved comparative examinations of the developing rice seed succinylomes by utilizing acetylome (28), malonylome (29), crotonylome (Meng et al., unpublished result), and 2-hydroxyisobutyrylome (30). A remarkable PTM overlapping was detected on proteins in rice (Fig. 5a). We discovered 169 existing proteins in 15 dpa developing rice seeds were modified by both lysine succinylation and acetylation. This obtained data was accountable for approximately 48.7% (169/347) of succinylome reported here. Additionally, 133, 250, and 300 succinylated rice proteins were modified by malonylation, crotonylation, and 2-hydroxyisobutyrylation, respectively. The quantity of proteins embraces two
and three particular PTM types are 1902 and 613, respectively (Fig. 5a). Remarkably, 67 rice proteins encompass the complete array of all the analyzed PTMs. These proteins are extensively involved in central carbon metabolism, amino acid biosynthesis, starch biosynthesis, and plant defense response mechanisms (Supplementary Table S4). Moreover, 85.1% of succinylated lysine sites identified in this study could also be modified by at least one of the other PTMs: acetylation, crotonylation, malonylation, and 2-hydroxyisobutyrylation (Fig. 5b). Interestingly, 14 lysine sites were found can harbor all the five types of PTMs in developing rice seed proteome (Fig. 5b). The PTM modification sites on representative protein phosphoglycerate kinase (Q6H6C7) are highlighted in Figure 5c.

Frequency of amino acid residues flanking lysines with at least three types of PTMs was analyzed to investigate the enrichment or depletion of various amino acids (Fig. 5d). Lysines that could be modified by multiple PTMs has a strong sequence preference for the charged amino acids of lysine (K) and glutamic acid (E) flanking the modification sites (Fig. 5d). Moreover, amino acids of D, I, V, H, F, Y, and R were found overly presented near to the modification sites as well (Fig. 5d). The discovery of multi-acylations at the same lysine sites and the patterned amino acid arrangement around the lysine residues (Fig. 5b and d) suggested that the same or similar enzymes are involved in the establishment and removal of these different acylation moieties. A plausible notion is that active acetyltransferases utilize acyl-CoAs, based on their cellular concentrations and affinity, to modify lysine substrates; and the diverse lysine modifications, whose modification moieties come from the different metabolic pathway, may be subjected to competition for the lysine sites for specific biological regulations. The identification of substantial overlaps of
modifications at the same sites indicates that these modifications may interact and regulate biological processes via crosstalk. The possible competition, crosstalk, and functional interplay among different acylation moieties on the same sites of a protein require further investigation to validate.

To completely understand the protein regulatory mechanisms with multi-modifications, KEGG pathway enrichment analysis, and protein-protein interaction networks were created for rice proteins with at least three analyzed modifications (Fig. 5e-f). Proteins embracing multiple modifications are associated with numerous cell biological processes. The superior six protein-protein interaction clusters are located in the highlighted portion of Figure 5f. Proteins engaging in interactions with ribosomes, protein processing in the endoplasmic reticulum, proteasome, glycolysis/gluconeogenesis, the TCA cycle, and oxidative phosphorylation provides evidence that cellular processes with strong connections to protein metabolism, carbon metabolism, and energy metabolism are subjected to various post-translational regulations.

CONCLUSION

Throughout this investigative study, 854 lysine succinylation sites on 347 proteins in developing rice seeds have been discovered. Western blot analysis provided significant evidence that lysine succinylation is a protein-specific reaction in developing rice seeds, and the global succinylation occupancy estimated from developing rice seed proteins was from 2% to 10%. Amino acid position predilections adjacent to the succinylated sites are obtained via motif and flaking sequence analyses. Six distinguished motifs were identified. LKsu and KsuXI are shown to be conserved for
both eukaryotes and prokaryotes. Additionally, proteins with different functions displayed a clear preference for distinct motifs in succinylation. Functional enrichment of GO annotations, KEGG pathways, and protein domain analyses revealed various cellular functions and diverse subcellular localizations of the succinylated proteins. A substantial amount of succinylated proteins in developing rice seeds were associated with carbon metabolism, starch biosynthesis, and seed storage proteins. Our analyses suggested a high level of conserved succinylation throughout distinctive organisms. Moreover, rice proteins expressing multiple post-translational modifications of succinylation, acetylation, malonylation, crotonylation, and 2-hydroxyisobutyrylation were identified. Sixty-seven rice proteins were shown to be modified by all the five modifications evaluated in this particular analysis and many of them share the same modification sites. Overall, our investigations successfully extended the knowledge and awareness of the plant succinylome.

ACKNOWLEDGEMENTS

This project is partly supported by the Agriculture and Food Research Initiative competitive grant (2011-67013-30125) of the USDA National Institute of Food, Agriculture, Mississippi Rice Promotion Board, and National Natural Science Foundation of China (Grant No. 31771761). We thank Mr. John Hunt and Ms. Lindsay Rexrode for proof-reading of the manuscript.

COMPETING INTERESTS

The authors declare that they have no competing interests.

DATA AVAILABILITY
The raw data and annotated MS spectra are available in the proteomics repository PRIDE with ProteomeXchange dataset accession number of PXD005582 and PXD013664 (Username: reviewer68213@ebi.ac.uk, Password: YpmXohf1), respectively.

REFERENCES

1. Hirschey, M. D., and Zhao, Y. (2015) Metabolic regulation by lysine malonylation, succinylation, and glutarylation. Mol Cell Proteomics 14, 2308-2315
2. Zhang, Z., Tan, M., Xie, Z., Dai, L., Chen, Y., and Zhao, Y. (2011) Identification of lysine succinylation as a new post-translational modification. Nat Chem Biol 7, 58-63
3. Papanicolaou, K. N., O’Rourke, B., and Foster, D. B. (2014) Metabolism leaves its mark on the powerhouse: recent progress in post-translational modifications of lysine in mitochondria. Front Physiol 5, 301
4. Weinert, B. T., Schölz, C., Wagner, S. A., lesmantavicius, V., Su, D., Daniel, J. A., and Choudhary, C. (2013) Lysine succinylation is a frequently occurring modification in prokaryotes and eukaryotes and extensively overlaps with acetylation. Cell Rep 4, 842-851
5. Wagner, G. R., and Payne, R. M. (2013) Widespread and enzyme-independent Nϵ-acetylation and Nϵsuccinylation of proteins in the chemical conditions of the mitochondrial matrix. J Biol Chem 288, 29036-29045
6. Gibson, G. E., Xu, H., Chen, H. L., Chen, W., Denton, T. T., and Zhang, S. (2015) Alpha-ketoglutarate dehydrogenase complex-dependent succinylation of proteins in neurons and neuronal cell lines. J Neurochem 134, 86-96
7. Rardin, M. J., He, W., Nishida, Y., Newman, J. C., Carrico, C., Danielson, S. R., Guo, A., Gut, P., Sahu, A. K., and Li, B. (2013) SIRT5 regulates the mitochondrial lysine succinyloyme and metabolic networks. Cell Metab 18, 920-933
8. Park, J., Chen, Y., Tishkoff, D. X., Peng, C., Tan, M., Dai, L., Xie, Z., Zhang, Y., Zwaans, B. M., and Skinner, M. E. (2013) SIRT5-mediated lysine desuccinylation impacts diverse metabolic pathways. Mol Cell 50, 919-930
9. Peng, C., Lu, Z., Xie, Z., Cheng, Z., Chen, Y., Tan, M., Luo, H., Zhang, Y., He, W., and Yang, K. (2011) The first identification of lysine malonylation substrates and its regulatory enzyme. Mol Cell Proteomics 10, M111.012658
10. Li, L., Shi, L., Yang, S., Yan, R., Zhang, D., Yang, J., He, L., Li, W., Yi, X., and Sun, L. (2016) SIRT7 is a histone desuccinylase that functionally links to chromatin compaction and genome stability. Nat Commun 7,12235
11. Du, J., Zhou, Y., Su, X., Yu, J. J., Khan, S., Jiang, H., Kim, J., Woo, J., Kim, J. H., and Choi, B. H. (2011) Sirt5 is a NAD-dependent protein lysine demalonylase and desuccinylase. Science 334, 806-809
12. 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 Metab 19, 605-617
13. Colak, G., Xie, Z., Zhu, A. Y., Dai, L., Lu, Z., Zhang, Y., Wan, X., Chen, Y., Cha, Y. H., and Lin, H. (2013) Identification of lysine succinylation substrates and the succinylation regulatory enzyme CobB in Escherichia coli. Mol Cell Proteomics 12, 3509-3520
14. Kosono, S., Tamura, M., Suzuki, S., Kawamura, Y., Yoshida, A., Nishiyama, M., and Yoshida, M. (2015) Changes in the Acetylome and Succinylome of Bacillus subtilis in Response to Carbon Source. PloS One 10, e0131169
15. Li, X., Hu, X., Wan, Y., Xie, G., Li, X., Chen, D., Cheng, Z., Yi, X., Liang, S., and Tan, F. (2014) Systematic Identification...
of the Lysine succinylation in the Protozoan Parasite Toxoplasma gondii. *J Proteome Res* 13, 6087-6095

16. Xie, Z., Dai, J., Dai, L., Tan, M., Cheng, Z., Wu, Y., Boeke, J. D., and Zhao, Y. (2012) Lysine succinylation and lysine malonylation in histones. *Mol Cell Proteomics* 11, 100-107

17. Yang, M., Wang, Y., Chen, Y., Cheng, Z., Gu, J., Deng, J., Bi, L., Chen, C., Mo, R., and Wang, X. (2015) Succinylome analysis reveals the involvement of lysine succinylation in metabolism in pathogenic Mycobacterium tuberculosis. *Mol Cell Proteomics* 14, 796-811

18. Arnaudo, A. M., and Garcia, B. A. (2013) Proteomic characterization of novel histone post-translational modifications. *Epigenetics Chromatin* 6, 24

19. Pan, J., Chen, R., Li, C., Li, W., and Ye, Z. (2015) Global Analysis of Protein Lysine Succinylation Profiles and Their Overlap with Lysine Acetylation in the Marine Bacterium Vibrio parahemolyticus. *J Proteome Res* 14, 4309-4318

20. Xie, L., Liu, W., Li, Q., Chen, S., Xu, M., Huang, Q., Zeng, J., Zhou, M., and Xie, J. (2014) First succinyl-proteome profiling of extensively drug-resistant Mycobacterium tuberculosis revealed involvement of succinylation in cellular physiology. *J Proteome Res* 14, 107-119

21. Jin, W., and Wu, F. (2016) Proteome-Wide Identification of Lysine Succinylation in the Proteins of Tomato (Solanum lycopersicum). *PloS One* 11, e0147586

22. 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. *J Proteome Res* 15, 879-890

23. 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. *Sci Rep* 6, 31576

24. Zhang, Y., Wang, G., Song, L., Mu, P., Wang, S., Liang, W., and Lin, Q. (2017) Global analysis of protein lysine succinylation profiles in common wheat. *BMC genomics* 18, 309

25. Shen, C., Xue, J., Sun, T., Guo, H., Zhang, L., Meng, Y., and Wang, H. (2016) Succinyl-proteome profiling of a high taxol containing hybrid Taxus species (Taxus× media) revealed involvement of succinylation in multiple metabolic pathways. *Sci Rep* 6, 21764

26. Bhullar, N. K., and Gruissem, W. (2013) Nutritional enhancement of rice for human health: the contribution of biotechnology. *Biotechnol Adv* 31, 50-57

27. Chen, Y., Wang, M., and Ouwerkerk, P. B. (2012) Molecular and environmental factors determining grain quality in rice. *Food Energy Secur* 1, 111-132

28. Meng, X., Lv, Y., Mujahid, H., Edelmann, M. J., Zhao, H., Peng, X., and Peng, Z. (2018) Proteome-wide lysine acetylation identification in developing rice (Oryza sativa) seeds and protein co-modification by acetylation, succinylation, ubiquitination, and phosphorylation. *BBA-Proteins Proteomics* 1866, 451-463

29. Mujahid, H., Meng, X., Xing, S., Peng, X., Wang, C., and Peng, Z. (2018) Malonylome analysis in developing rice (Oryza sativa) seeds suggesting that protein lysine malonylation is well-conserved and overlaps with acetylation and succinylation substantially. *J Proteomics* 170, 88-98

30. Meng, X., Xing, S., Perez, L. M., Peng, X., Zhao, Q., Redoña, E. D., Wang, C., and Peng, Z. (2017) Proteome-wide Analysis of Lysine 2-hydroxyisobutyrylation in Developing Rice (Oryza sativa) Seeds. *Sci Rep* 7, 17486

31. Lee, T.-J., Shultz, R. W., Hanley-Bowdoin, L., and Thompson, W. F. (2004) Establishment of rapidly proliferating rice cell suspension culture and its characterization by fluorescence-activated cell sorting analysis. *Plant Mol Biol Rep* 22, 259-267

32. Hurkman, W. J., and Tanaka, C. K. (1986) Solubilization of plant membrane proteins for analysis by two-
dimensional gel electrophoresis. Plant Physiol 81, 802-806
33. Xing, S., Meng, X., Zhou, L., Mujahid, H., Zhao, C., Zhang, Y., Wang, C., and Peng, Z. (2016) Proteome Profile of Starch Granules Purified from Rice (Oryza sativa) Endosperm. PloS One 11, e0168467
34. Udeshi, N. D., Svinkina, T., Mertins, P., Kuhn, E., Mani, D., Qiao, J. W., and Carr, S. A. (2013) Refined preparation and use of anti-diglycine remnant (K-ε-GG) antibody enables routine quantification of 10,000 s of ubiquitination sites in single proteomics experiments. Mol Cell Proteomics 12, 825-831
35. Cox, J., Matic, I., Hilger, M., Nagaraj, N., Selbach, M., Olsen, J. V., and Mann, M. (2009) A practical guide to the MaxQuant computational platform for SILAC-based quantitative proteomics. Nat Protoc 4, 698-705
36. Tyanova, S., Temu, T., and Cox, J. J. N. p. (2016) The MaxQuant computational platform for mass spectrometry-based shotgun proteomics. Nat Protoc 11, 2301-2319
37. Meyer, J. G., D’Souza, A. K., Sorensen, D. J., Rardin, M. J., Wolfe, A. J., Gibson, B. W., and Schilling, B. (2016) Quantification of lysine acetylation and succinylation stoichiometry in proteins using mass spectrometric data-independent acquisitions (SWATH). J Am Soc Mass Spectrom 27, 1758-1771
38. Sadhukhan, S., Liu, X., Ryu, D., Nelson, O. D., Stupinski, J. A., Li, Z., Chen, W., Zhang, S., Weiss, R. S., and Locasale, J. W. J. p. o. t. N. A. o. S. (2016) Metabolomics-assisted proteomics identifies succinylation and SIRT5 as important regulators of cardiac function. Proc Natl Acad Sci USA 113, 4320-4325
39. Zhang, Z., Zhao, H., Tang, J., Li, Z., Li, Z., Chen, D., and Lin, W. (2014) A proteomic study on molecular mechanism of poor grain-filling of rice (Oryza sativa L.) inferior spikelets. PloS One 9, e89140
40. Sun, H., Peng, T., Zhao, Y., Du, Y., Zhang, J., Li, J., Xin, Z., and Zhao, Q. (2015) Dynamic Analysis of Gene Expression in Rice Superior and Inferior Grains by RNA-Seq. PloS One 10, e0137168
41. Saunier, E., Benelli, C., and Bortoli, S. (2016) The pyruvate dehydrogenase complex in cancer: an old metabolic gatekeeper regulated by new pathways and pharmacological agents. Int J Cancer 138, 809-817
42. Fan, J., Shan, C., Kang, H.-B., Elf, S., Xie, J., Tucker, M., Gu, T.-L., Aguiar, M., Lonning, S., and Chen, H. (2014) Tyr phosphorylation of PDP1 toggles recruitment between ACAT1 and SIRT3 to regulate the pyruvate dehydrogenase complex. Mol Cell 53, 534-548
43. Bao, J., Lu, Y., Yang, F., Zhang, G., Shao, Y., Corke, H., and Sun, M. (2012) Nucleotide polymorphisms in OsAGP genes and their possible association with grain weight of rice. J Cereal Sci 55, 312-317
44. Kang, H. G., Park, S., Matsuoka, M., and An, G. (2005) White – core endosperm floury endosperm – 4 in rice is generated by knockout mutations in the C4 – type pyruvate orthophosphate dikinase gene (OsPPDKB). Plant J 42, 901-911
45. Shewry, P. R., and Halford, N. G. (2002) Cereal seed storage proteins: structures, properties and role in grain utilization. J Exp Bot 53, 947-958
46. Kawakatsu, T., Hirose, S., Yasuda, H., and Takaawa, F. (2010) Reducing rice seed storage protein accumulation leads to changes in nutrient quality and storage organelle formation. Plant Physiol 154, 1842-1854
47. Spange, S., Wagner, T., Heinzel, T., and Krämer, O. H. (2009) Acetylation of non-histone proteins modulates cellular signalling at multiple levels. Int J Biochem Cell Biol 41, 185-198
48. Sadoul, K., Boyault, C., Pabion, M., and Khochbin, S. (2008) Regulation of protein turnover by acetyltransferases and deacetylases. Biochimie 90, 306-312
49. Sears, R., Nuckolls, F., Haura, E., Taya, Y., Tamai, K., and Nevins, J. R. (2000) Multiple Ras-dependent phosphorylation pathways regulate Myc protein stability. Genes Dev 14, 2501-2514
50. Sabari, B. R., Tang, Z., Huang, H., Yong-Gonzalez, V., Molina, H., Kong, H. E., Dai, L., Shimada, M., Cross, J. R., and
Zhao, Y. (2015) Intracellular crotonyl-CoA stimulates transcription through p300-catalyzed histone crotonylation. *Mol cell* 58, 203-215

51. Chen, Y., Sprung, R., Tang, Y., Ball, H., Sangras, B., Kim, S. C., Falck, J. R., Peng, J., Gu, W., and Zhao, Y. (2007) Lysine propionylation and butyrylation are novel post-translational modifications in histones. *Mol Cell Proteomics* 6, 812-819

**TABLES**

**Table 1. Identified succinylated proteins involved in starch biosynthesis pathway.**

| Gene Name | Encoded Protein | Enzyme Entry | Protein Entry* | Succinylation site(s) |
|-----------|-----------------|--------------|----------------|-----------------------|
| *OsAGPS2* | ADP-glucose pyrophosphorylase small subunit 2 | [EC:2.7.7.27] | P15280 | 217, 261, 263, 403, 447, 476 |
| *OsAGPL2* | ADP-glucose pyrophosphorylase large subunit 2 | [EC:2.7.7.27] | Q5VNT5 | 37, 250, 312, 449, 459, 504 |
| *OsBEI*   | Starch branching enzyme I | [EC:2.4.1.18] | Q0D9D0 | 89, 103, 164, 324, 372, 500, 524, 697 |
| *OsBEIIb* | Starch branching enzyme IIb | [EC:2.4.1.18] | Q6H6P8 | 134, 191, 587 |
| *OsPUL*   | Starch debranching enzyme: Pullulanase | [EC:3.2.1.41] | Q7X834 | 732 |
| *OsPHOL*  | Starch phosphorylase L | [EC:2.4.1.1] | Q9AU8V | 259, 255, 657, 734 |
| *SUS2*    | Sucrose synthase 2 | [EC:2.4.1.13] | P30298 | 38, 319, 380, 725 |
| *SUS3*    | Sucrose synthase 3 | [EC:2.4.1.13] | Q43009 | 68, 172, 215, 218, 327, 358, 588, 755, 797 |
| *UGP*     | UDP-glucose pyrophosphorylase | [EC:2.7.7.9] | Q9X08 | 8, 80, 180, 257, 285, 346, 402, 412 |
| *PGM*     | Phosphoglucomutase | [EC:5.4.2.2] | Q9AUQ4 | 568 |
| *FLO4*    | Flouy endosperm 4 / phosphate dikinase 1 | [EC:2.7.9.1] | Q6AVA8 | 238, 259, 440, 462, 555, 609, 803, 810, 815, 889 |

a: Protein entry is protein Uniprot ID.
Table 2. Identified succinylated seed storage proteins.

| Storage Protein | Protein Entry* | RAP Locus | Succinylation site(s) |
|-----------------|----------------|-----------|-----------------------|
| GluA-1          | Q0JJ36         | Os01g0762500 | 154, 210, 395, 418, 444 |
| GluA-2          | P07730         | Os10g0400200 | 418                   |
| GluA-3          | Q10JA8         | Os03g0427300 | 146, 153, 210, 394, 417, 443 |
| GluB-1a         | Q0E2D2         | Os02g0249800 | 149, 197, 263, 312, 391, 414, 427, 440 |
| GluB-5          | Q6ERU3         | Os02g0268100 | 245, 341, 392, 415, 416, 428, 441 |
| GluB-6          | Q6T725         | Os02g0248800 | 261                   |
| GluB-7          | Q0E2G5         | Os02g0242600 | 147, 384, 408, 433    |
| GluC-1          | Q6K7K6         | Os02g0453600 | 95, 152, 214, 290, 331, 456 |
| GluD-1          | Q6K508         | Os02g0249000 | 410, 467, 261, 190    |
| gb1             | Q6ZK46         | Os08g0127900 | 373                   |
| Globulin        | Q75GX9         | Os03g0663800 | 257, 272, 265, 367, 249, 524, 488 |
| Globulin        | Q852L2         | Os03g0793700 | 434, 274, 240, 216    |
| Patatin         | Q8S0E1         | Os01g0898500 | 48                    |

*a: Protein entry is protein Uniprot ID.

Table 3. Comparison of the developing rice seed succinyllome with published succinyllomes of different species.

| Kingdom   | Organism                      | NO. of Ksu sites | NO. of Ksu proteins | No. of conserved Ksu protein in rice | Percent | Reference |
|-----------|-------------------------------|------------------|---------------------|--------------------------------------|---------|-----------|
| Plantae   | Oryza sativa (developing seeds) | 854              | 347                 | —                                    | 40.63%  | (22)      |
|           | Oryza sativa (embryo)         | 665              | 261                 | 141                                  | 59.37%  | (23)      |
|           | Brachypodium distachyon L.    | 605              | 262                 | 206                                  | 32.28%  | (21)      |
|           | Solanum lycopersicum          | 347              | 202                 | 112                                  |         |           |
| Animalia  | Homo sapiens (Hela cell)      | 2004             | 738                 | 117                                  | 33.72%  | (4)       |
|           | Mus musculus (MEFs)           | 2565             | 779                 | 66                                   | 19.02%  | (8)       |
|           | Mus musculus (liver)          | 2140             | 750                 | 113                                  | 32.56%  | (4)       |
| Fungi     | Saccharomyces cerevisiae      | 1345             | 474                 | 75                                   | 21.61%  | (4)       |
| Protista  | Toxoplasma gondii             | 425              | 147                 | 39                                   | 11.24%  | (15)      |
| Bacteria  | Mycobacterium tuberculosis (H37Rv) | 1537            | 627                 | 45                                   | 12.97%  | (17)      |
|           | Bacillus subtilis (strain 168) | 327              | 204                 | 32                                   | 9.22%   | (14)      |
|           | Escherichia coli (BW25113)    | 2572             | 990                 | 51                                   | 14.70%  | (4)       |
|           | Vibrio parahaemolyticus       | 1931             | 642                 | 54                                   | 15.56%  | (19)      |
**FIGURES AND FIGURE LEGENDS**

![Figure 1](https://www.mcponline.org)

**Fig. 1** Succinylation profile in different rice organs/tissues revealed by Western blots. Molecular weight is labeled on the left. The samples are labeled on the top. M: size marker; 1: suspension cell protein; 2: root protein; 3: leave protein; 4: flower protein; 5: pollen protein; 6: protein from 7 dpa seeds; 7: protein from 15 dpa seeds; 8: protein from 21 dpa seeds; 9: protein from mature dry seeds. **a.** Image of SDS-PAGE stained with Coomassie Brilliant Blue G-250. **b.** Western blot image of protein succinylation. The primary antibodies used were rabbit-derived pan anti-succinyl lysine antibody (PTM-401, PTM Biolabs, Chicago, IL, USA). The same amount of proteins (25 μg per lane) were loaded in panels a and b. The original images of SDS-PAGE and western blotting are shown in Supplementary Fig. S1.
Fig. 2 Estimation of global protein succinylation occupancy in developing rice seeds.

a. Molecular weight is labeled on the left. The samples are labeled on the top. M: size marker; 1. Protein (15μg) dissolved in SDS buffer; 2: Protein (15μg) dissolved in PBS buffer; 3: Protein (15μg) dissolved in PBS buffer and treated with 0.5mM succinyl-CoA; 4: Protein (15μg) dissolved in PBS...
buffer and treated with 1mM succinyl-CoA; 5-12: 60%, 45%, 30%, 15%, 10%, 5%, 2%, 1% of protein loaded comparing to lane 4 (protein source was the same as lane 4). Upper panel: Image of SDS-PAGE stained with Coomassie Brilliant Blue G-250. Lower panel: Western blot image of protein succinylation. The primary antibodies used were rabbit-derived pan anti-succinyl lysine antibody (PTM-401, PTM Biolabs, Chicago, IL, USA). b. Succinylation intensity for the two major bands in lane 1, 2, 9, 10, 11, and 12 of Western blot image in panel A. Succinylation intensity was quantified by Image Studio Lite software. The bar plot was created from three technical replicates for succinylation intensity measurement shown in Supplementary Figure S2.
Fig. 3 Motif and logo-based clustering analyses of the succinylation sites. 

a. Conserved motifs of succyl-21-mers flanking succinylation sites (“K” at position 0). The size of each letter correlates to the frequency of that amino acid residue occurring in that position. 

b. Heat map of the amino acid compositions around the succinylation sites. The $-\log_{10}$ (Fisher’s exact test p-value) for every amino acid in each position (from -10 to +10) is shown. Motif logo-based clustering analyses: GO annotation enrichment (c), KEGG pathway enrichment analysis (d), and domain enrichment analysis (e).
Fig. 4 Succinylated enzymes in the TCA cycle, glycolysis/gluconeogenesis, and starch biosynthesis pathways. MDH: malate dehydrogenase; CS: citrate synthase; ACO: aconitate hydratase; IDH: isocitrate dehydrogenase; OGDH: 2-oxoglutarate dehydrogenase E1 component; DLD: dihydrolipoamide dehydrogenase; LSC: succinyl-CoA synthetase; DLST: 2-oxoglutarate dehydrogenase E2 component (dihydrolipoamide succinyltransferase); SDHA: succinate dehydrogenase; FH: fumarate hydratase; PGM: phosphoglucomutase; PGI: phosphoglucoisomerase; ALDO: fructose-bisphosphate aldolase; GAPDH: glyceraldehyde-3-phosphate dehydrogenase; PRPS: pyridoxal phosphate synthetase; PRPP: phosphoribosyl pyrophosphate.
dehydrogenase; TPI: triosephosphate isomerase; PGK: phosphoglycerate kinase; GPML: phosphoglycerate mutase; ENO: enolase; PK: pyruvate kinase; PDHA: pyruvate dehydrogenase; DLAT: pyruvate dehydrogenase E2 component (dihydrolipoamide acetyltransferase); UGP: UDP-glucose pyrophosphorylase; OsSUS: sucrose synthase; OsAGP: ADP-glucose pyrophosphorylase; OsGBSS: granule-bound starch synthase; OsBE: starch branching enzyme; OsSS: starch synthase; OsISA: starch debranching enzyme: isoamylase; OsPUL: starch debranching enzyme: pullulanase; OsPHO: starch phosphorylase; G6PD: glucose-6-phosphate dehydrogenase; PGLS: 6-phosphogluconolactonase; PGD: 6-phosphogluconate dehydrogenase, decarboxylating 2; RPE: Ribulose-phosphate 3-epimerase; RPIA: ribose-5-phosphate isomerase; TKT: transketolase 1. Solid circles in green, red, yellow, blue and pink colors represent enzymes modified by succinylation, crotonylation, acetylation, malonylation, and 2-hydroxyisobutyrylation, respectively.
Fig. 5 Overlapping rice proteins between lysine succinylation, acetylation, malonylation, crotonylation, and 2-hydroxyisobutyrylation. a. Venn diagram showing the number of proteins overlapping among succinylation, acetylation, malonylation, crotonylation, and 2-hydroxyisobutyrylation. b. Venn diagram showing the number of lysine sites overlapping among...
succinylation, acetylation, malonylation, crotonylation, and 2-hydroxyisobutyrylation.  

c. Succinylated, acetylated, malonylated, crotonylated, and 2-hydroxyisobutyrylated sites on representative protein phosphoglycerate kinase (Q6H6C7). su: lysine succinylation; ac: lysine acetylation; cr: lysine crotonylation; mal: lysine malonylation; hib: lysine 2-hydroxyisobutyrylation.  

d. Heat map represents the frequency of amino acid compositions around (-10 to +10 position) the lysines (K in position 0) of proteins with at least three types of PTMs identified.  

e. KEGG pathway enrichment analysis (p-value < 0.05) for proteins could be modified by at least types of PTMs. -log10(fisher’s exact p-value) shown as x-axis.  

f. Protein-protein interaction network of rice proteins with at least three PTMs among succinylation, acetylation, malonylation, crotonylation, and 2-hydroxyisobutyrylation. Protein-protein interaction network was built against the STRING database (version 10.5). Identified interactions with confidence score ≥ 0.7 (high confidence) were fetched and visualized by Cytoscape software.