

First comprehensive proteomics analysis of lysine crotonylation in leaves of peanut (*Arachis hypogaea* L.)

Manlin Xu

Shandong Peanut Research Institute

Jianda Luo

Chinese Academy of Agricultural Sciences Institute of Tobacco Research

Ying Li

Chinese Academy of Agricultural Sciences Institute of Tobacco Research

Lili Shen

Chinese Academy of Agricultural Sciences Institute of Tobacco Research

Xia Zhang

Shandong Peanut Research Institute

Jing Yu

Shandong Peanut Research Institute

Zhiqing Guo

Shandong Peanut Research Institute

Juxiang Wu

Shandong Peanut Research Institute

Yucheng Chi (✉ 87626681@163.com)

Shandong Peanut Research Institute

Jinguang Yang (✉ yangjinguang@caas.cn)

Chinese Academy of Agricultural Sciences Institute of Tobacco Research

Research article

Keywords: Post-translational modification, Lysine crotonylation, peanut (*Arachis hypogaea* L.)

Posted Date: March 20th, 2020

DOI: <https://doi.org/10.21203/rs.3.rs-18151/v1>

License:   This work is licensed under a Creative Commons Attribution 4.0 International License.

[Read Full License](#)

Version of Record: A version of this preprint was published at PROTEOMICS on February 23rd, 2021. See the published version at <https://doi.org/10.1002/pmic.202000156>.

Abstract

Background: Lysine crotonylation is an important post-translational modification (PTM) process. Most research in this area has been carried out on mammals and yeast, but there has been no research published on crotonylated proteins on peanut.

Results: In the current study, large-scale lysine crotonylation analysis was performed by a combination of affinity enrichment and high-resolution LC-MS/MS analysis. Altogether, 6051 lysine crotonylation sites were identified in 2508 protein groups. Bioinformatics analysis showed that lysine-crotonylated proteins were involved in many biological processes, such as carbon fixation in photosynthetic organisms, photosynthesis, biosynthesis of amino acids, ribosomes, etc. In particular, subcellular localization analysis showed that 43% of the crotonylated proteins were located in the chloroplast and that there were 29 crotonylated proteins associated with photosynthesis. In addition, 26 crotonylated proteins were identified in photosynthesis network and 145 proteins were mapped to ribosome network, indicating the diverse functions for lysine crotonylation in peanut.

Conclusion: These data show that crotonylated proteins play a major part in peanut biological functions including carbon fixation in photosynthetic organisms, photosynthesis, biosynthesis of amino acids, ribosomes, etc. A lot of proteins related to photosynthesis and ribosome suggest that lysine crotonylation may play important regulatory roles in their structure and function. This dataset is the first comprehensive proteomics analysis of lysine crotonylation in peanut and will serve as an important resource with which to study the biosynthesis and function of lysine crotonylation in peanut and related plants. Based on these results, further studies to expand on the lysine crotonylation analysis were suggested. Data are available via ProteomeXchange with the identifier PXD017675.

Background

Post-translational modifications (PTMs) are important methods for the regulation of gene expression, and have a very important effect on protein function, including cell growth and differentiation, metabolism, etc. Most important are the reversible PTMs, and lysine (K) is one of the most frequently modified amino acid residues [1, 2]. Such modifications include lysine acetylation, succinylation, crotonylation, malonylation, 2-hydroxyisobutyrylation, and β -hydroxybutyrylation [3–8]. PTM is one of the important targets for epigenetic studies, in which histone lysine modification changes dynamically under the action of regulatory enzymes. On one hand, such modification is essential for the control of gene expression by complex interactions of transcription factors binding to regulatory DNA elements. On the other hand, it plays an important role in transcriptional regulation by modifying the chromatin remodelling, and by activating or repressing regulatory DNA elements [1, 9]. Lysine-modified histone proteins contain important biological information, and their identification is key to understanding their function and understanding the mechanism of epigenetic effects.

Lysine crotonylation was first reported as a PTM process in 2011, and is a robust indicator of active promoters, which could be an important signal in the control of male germ cell differentiation. The original paper reported 67 new histone modification markers, with further structural and genomic localization analysis confirming histone lysine crotonylation (Kcr) to be evolutionarily highly conserved and to have different biological functions from histone lysine acetylation (Kac) [3]. Crotonyl-CoA catalyzes crotonylation by the transcriptional coactivator p300, which is also influenced by the concentration of crotonyl-coA in cells. Therefore, both genetic and environmental factors can be used to regulate histone crotonylation [10, 11]. Some writers, erasers, and readers of Kcr have been identified [12]. Recently, a number of human proteins were identified as being crotonylated [13–17]. Global profiling of crotonylation has been reported in tobacco [18], zebrafish [19], rice [20], yeast [21], and papaya [22], which showed that crotonylation plays an important role in the regulation of various biological processes across kingdoms.

Peanut (*Arachis hypogaea* L.) is one of the most important oil crops in the world, and the harvested seeds have high concentrations of lipids and proteins. The peanut genome was sequenced in 2016, revealing peanut to be a major source of candidate genes for fructification, oil biosynthesis, and allergens, providing considerable information on plant biology [23]. No research has been published on crotonylated proteins on peanut. In the current study, using a combination of affinity purification and high-resolution liquid chromatography–tandem mass spectrometry (LC–MS/MS) analysis, 6051 lysine crotonylation sites were identified in 2508 protein groups. Bioinformatics analysis showed that lysine crotonylated proteins were involved in many biological processes, including carbon fixation in photosynthetic organisms, photosynthesis, biosynthesis of amino acids, ribosomes, etc. This information is the first comprehensive proteomics analysis of lysine crotonylation in peanut and will serve as an important resource with which to study the biosynthesis and function of Kcr in peanut and related plants.

Methods

Plant material The peanut variety “Huayu 20” plants were arranged in a randomized complete block design, and grown at 25 °C with a 12 h photoperiod. The leaves excised from four week old peanut plants for protein extraction.

Protein Extraction Sample was sonicated three times on ice using a high intensity ultrasonic processor (Scientz) in lysis buffer (8 M urea, 2 mM EDTA, 3 μM TSA, 50 mM NAM, 10 mM DTT and 1% Protease Inhibitor Cocktail). The remaining debris was removed by centrifugation at 20,000 g at 4 °C for 10 min. Finally, the protein was precipitated with cold 15% TCA for 2 h at -20 °C. After centrifugation at 4 °C for 10 min, the supernatant was discarded. The remaining precipitate was washed with cold acetone for three times. The protein was redissolved in buffer (8 M urea, 100 mM NH₄CO₃, pH 8.0) and the protein concentration was determined with 2-D Quant kit according to the manufacturer’s instructions.

Trypsin Digestion For digestion, the protein solution was reduced with 10 mM DTT for 1 h at 37 °C and alkylated with 20 mM IAA for 45 min at room temperature in darkness. For trypsin digestion, the protein

sample was diluted by adding 100 mM NH_4CO_3 to urea concentration less than 2M. Finally, trypsin was added at 1:50 trypsin-to-protein mass ratio for the first digestion overnight and 1:100 trypsin-to-protein mass ratio for a second 4 h-digestion.

HPLC Fractionation The sample was then fractionated into fractions by high pH reverse-phase HPLC using Agilent 300Extend C18 column (5 μm particles, 4.6 mm ID, 250 mm length). Briefly, peptides were first separated with a gradient of 2–60% acetonitrile in 10 mM ammonium bicarbonate pH 10 over 80 min into 80 fractions, Then, the peptides were combined into 6 fractions and dried by vacuum centrifuging.

Affinity Enrichment To enrich Kcro peptides, tryptic peptides dissolved in NETN buffer (100 mM NaCl, 1 mM EDTA, 50 mM Tris-HCl, 0.5% NP-40, pH 8.0) were incubated with pre-washed antibody beads (PTM Biolabs) at 4 °C overnight with gentle shaking. The beads were washed four times with NETN buffer and twice with ddH₂O. The bound peptides were eluted from the beads with 0.1% TFA. The eluted fractions were combined and vacuum-dried. The resulting peptides were cleaned with C18 ZipTips (Millipore) according to the manufacturer's instructions, followed by LC-MS/MS analysis.

LC-MS/MS Analysis The peptides were dissolved in 0.1% FA, directly loaded onto a reversed-phase pre-column (Acclaim PepMap 100, Thermo Scientific). Peptide separation was performed using a reversed-phase analytical column (Acclaim PepMap RSLC, Thermo Scientific). The gradient was comprised of an increase from 6–22% solvent B (0.1% FA in 98% ACN) for 24 min, 22–40% for 8 min and climbing to 80% in 5 min then holding at 80% for the last 3 min, all at a constant flow rate of 300 nl/min on an EASY-nLC 1000 UPLC system, the resulting peptides were analyzed by Q Exactive™ Plus hybrid quadrupole-Orbitrap mass spectrometer (ThermoFisher Scientific).

The peptides were subjected to NSI source followed by tandem mass spectrometry (MS/MS) in Q Exactive™ plus (Thermo) coupled online to the UPLC. Intact peptides were detected in the Orbitrap at a resolution of 70,000. Peptides were selected for MS/MS using NCE setting as 30; ion fragments were detected in the Orbitrap at a resolution of 17,500. A data-dependent procedure that alternated between one MS scan followed by 20 MS/MS scans was applied for the top 20 precursor ions above a threshold ion count of 5E3 in the MS survey scan with 15.0 s dynamic exclusion. The electrospray voltage applied was 2.0 kV. Automatic gain control (AGC) was used to prevent overfilling of the orbitrap; 5E4 ions were accumulated for generation of MS/MS spectra. For MS scans, the m/z scan range was 350 to 1800. Fixed first mass was set as 100 m/z.

Database Search The resulting MS/MS data was processed using MaxQuant with integrated Andromeda search engine (v.1.5.2.8). Tandem mass spectra were searched against *Arachis duranensis* database (from NCBI, 42660 entries) concatenated with reverse decoy database. Trypsin/P was specified as cleavage enzyme allowing up to 4 missing cleavages, 5 modifications per peptide and 5 charges. Mass error was set to 10 ppm for precursor ions and 0.02 Da for fragment ions. Carbamidomethylation on Cys was specified as fixed modification and oxidation on Met, crotonylation on Lys and acetylation on protein

N-terminal were specified as variable modifications. False discovery rate (FDR) thresholds for protein, peptide and modification site were specified at 1%. Minimum peptide length was set at 7. All the other parameters in MaxQuant were set to default values. The site localization probability was set as > 0.75.

Bioinformatics Methods Gene Ontology (GO) annotation proteome was derived from the UniProt-GOA database ([www. http://www.ebi.ac.uk/GOA/](http://www.ebi.ac.uk/GOA/)). Proteins were classified by Gene Ontology annotation based on three categories: biological process, cellular component and molecular function. Identified proteins domain functional description were annotated by InterProScan (a sequence analysis application) based on protein sequence alignment method, and the InterPro domain database was used. InterPro (<http://www.ebi.ac.uk/interpro/>) is a database that integrates diverse information about protein families, domains and functional sites. Kyoto Encyclopedia of Genes and Genomes (KEGG) database was used to annotate protein pathway. Firstly, using KEGG online service tools KAAS to annotated protein's KEGG database description. Then mapping the annotation result on the KEGG pathway database using KEGG online service tools KEGG mapper. There, we used wolfpsort a subcellular localization predication soft to predict subcellular localization. Wolfpsort an updated version of PSORT/PSORT II for the prediction of eukaryotic sequences. Soft motif-x was used to analysis the model of sequences constituted with amino acids in specific positions of modify-21-mers (10 amino acids upstream and downstream of the site) in all protein sequences. And all the database protein sequences were used as background database parameter, other parameters with default. Cytoscape (version 3.0) software was used to display the network [24].

Data Availability The mass spectrometry proteomics data have been deposited to the ProteomeXchange Consortium via the PRIDE [25] partner repository with the dataset identifier PXD017675.

Results

Identification and analysis of lysine crotonylation sites and crotonylated proteins in peanut leaves.

The experimental procedure is presented diagrammatically in Fig. 1a. In order to identify the crotonylated proteins and crotonylation sites, we used sensitive immunoaffinity purification and high-resolution LC-MS/MS. Altogether, 6051 crotonylation sites were identified in 2508 proteins. The statistics of the data are presented in Additional file 1: Table S1. Data are available via ProteomeXchange (<https://www.ebi.ac.uk/pride>), with the identifier PXD017675. Among the 2508 crotonylated proteins, the majority of the proteins had fewer than three crotonylation sites, with only a few proteins having seven or more crotonylation sites. The length of most of the peptides was between 8 and 28 residues, which agrees with the size distribution of tryptic peptides (Fig. 1c), indicating that sample preparation reached the standard required. The mass errors of the proteins was almost zero, meaning that the mass spectrometry data met the required standard (Fig. 1b).

Analysis of crotonylation site motifs

To investigate any patterns of amino acids adjacent to the Kcr sites, we used the Motif-X program. A total of six clearly conserved motifs (with a motif score >25) were identified (Fig. 2a), namely EKcrG, AKcrE, EKcrA, DKcrI, EKcrV and KcrE.....K; a substantial bias in amino acid distribution was observed from the -10 to +10 positions around the Kcr sites in peanut (Fig. 2b). As shown in Fig. 2a, the first five motifs contained a residue with acidic groups (E, G, A, D, I or V) at the -1 or +1 position. The sixth motif contained a residue with acidic groups (E or K) at the +1 or +5 position.

Structural analysis of all the crotonylated proteins

To understand the relationship between Kcr and protein structure, the structures of all the crotonylated proteins were analyzed. A total of 46.5% of the crotonylated sites were located in α -helices, 10.6% of the sites were in β -strands and the rest of the sites were in disordered coils. The distribution pattern between the crotonylated lysines and all lysines was very similar (Fig. 2c). In addition, analysis of surface accessibility of crotonylation sites of peanut proteins showed that Kcr may not affect the surface properties of the modified proteins, because the enrichment of crotonylated sites on the protein surface is very similar to that of all lysine residues (Fig. 2d).

Functional annotation and cellular localization of crotonylated proteins in peanut

To obtain further information on the distribution and function of lysine crotonylated proteins identified in peanut, we summed the number of the proteins in each level 2 Gene Ontology (GO) term, namely biological process, cellular component and molecular function (Fig. 3, Additional file 2: Table S2). The GO analysis showed that the crotonylated proteins were involved in diverse biological processes, cellular components and molecular functions. The identified crotonylated proteins with respect to biological process being related to metabolic process, cellular process, single-organism process, localization, biological regulation, response to stimulus, cellular component organization and biogenesis and others, accounting for 1106 (36%), 852 (28%), 671 (22%), 151 (5%), 110 (4%), 88 (3%), 69 (2%) and 28 (1%), respectively (Fig 3a, Additional file 3: Table S3). In the cellular component classification, 443 (38%), 268 (23%), 254 (21%), 197 (17%) and 13 (1%) were associated with cell, macromolecular complex, organelle, membrane and others, respectively (Fig. 3b, Additional file 3: Table S3). In the molecular function classification, proteins related to binding, catalytic activity, structural molecule activity, transporter activity, and others, accounted for 1088 (46%), 1015 (42%), 121 (5%), 78 (3%) and 88 (4%), respectively (Fig. 3c, Additional file 3: Table S3). These data indicated that crotonylated proteins take part in a range of biological processes. According to subcellular location annotations of the crotonylated proteins, we summed the number of proteins in each subcellular location. The crotonylated proteins identified in peanut leaf were located in the chloroplast (1068, 43%), cytosol (618, 25%) and nucleus (428, 17%), with a further 116 (5%) proteins being located in the plasma membrane, with 112 (4%) proteins located in the mitochondria, 53 (2%) proteins located to the extracellular space and 30 (1%) proteins located in the cytoskeleton (Fig. 3d, Additional file 4: Table S4).

Functional enrichment analysis of crotonylated proteins

In order to identify the types of proteins that were involved in lysine crotonylation, we performed GO enrichment and Kyoto Encyclopedia of Genes and Genomes (KEGG) pathway analysis. GO enrichment analysis included biological process, cellular component and molecular function classifications (Fig. 4a, Additional file 5: Table S5). The enrichment analysis of the biological process grouping indicated that crotonylated proteins were enriched with respect to biosynthetic processes and catabolic processes. In the cellular component category, the crotonylated proteins were significantly enriched with respect to the cytoplasmic component. There were also many crotonylated proteins associated with thylakoids, intracellular non-membrane-bounded organelles and the proteasome core complex. With respect to the enrichment of the molecular function category, most crotonylated proteins were enriched with respect to ligase activity, cofactor binding, structural constituents of ribosomes, oxidoreductase activity, and isomerase activity (Fig. 4a, Additional file 5: Table S5).

The KEGG pathway indicated that Kcr was associated with many proteins involved in carbon fixation in photosynthetic organisms, aminoacyl-tRNA biosynthesis, glutathione metabolism, the pentose phosphate pathway, biosynthesis of amino acids, ribosomes, carbon metabolism, proteasomes, photosynthesis, glyoxylate and dicarboxylate metabolism, oxidative phosphorylation, pyruvate metabolism, glycolysis/gluconeogenesis, the citrate cycle (TCA cycle), photosynthesis-antenna proteins, and alanine, aspartate, and glutamate metabolism (Fig. 4b, Additional file 6: Table S6). These results imply that crotonylated proteins play an important regulatory role in a wide range of important processes. In line with these results, the ClpP/crotonase-like domain, Glutathione S-transferase, C-terminal-like and Glutathione S-transferase N-terminal domains were clearly enriched in crotonylated proteins (Fig. 4c, Additional file 7: Table S7).

Crotonylated proteins involved in photosynthesis

Photosynthesis is a vital biological process in all plants, including peanut. A total of 29 crotonylated proteins were involved in photosynthesis. LC-MS/MS analysis showed that many subunits of these proteins were crotonylated (Fig. 5a): nine subunits of photosystem II (PsbC/B/O/P/Q/R/S/27/28), six subunits of photosystem I (PsaD/E/F/G/H/N), one subunit of the cytochrome b6/f complex (PetC), two subunits of photosynthetic electron transport (PetE/H) and three subunits of F-type ATPase (gamma, delta, b). The data from the functional enrichment analysis showed that crotonylated proteins may play an important role in photosynthesis. LC-MS/MS analysis also showed that a large number of subunits were crotonylated in ribosomes (Fig. 5b).

Interactive networks among crotonylated proteins in peanut

To further investigate the regulatory role of crotonylation in photosynthesis and ribosomes, we used the algorithm from the Cytoscape software to establish protein interaction networks (Fig. 6a, Additional file 8: Table S8). In the peanut photosynthesis network, 26 crotonylated proteins were identified as nodes in the protein interaction database. In the peanut ribosome network, 145 proteins were mapped to the protein interaction database (Fig. 6b, Additional file 9: Table S9). These maps provide a view of how crotonylated proteins regulate photosynthetic and ribosome-associated biological processes.

Discussion

Several recent reports have shown that crotonylation plays an important role in the regulation of various biological processes in humans, animals [19], plants [18, 22] and microorganisms [21]. Crotonylation has been reported in several plants, such as tobacco [18], rice [20], papaya [22] and tea [26], but there has been no report of crotonylation in leguminous plants, such as peanut. In the present study, we used a combination of immunoaffinity purification and high-resolution LC-MS/MS analysis to identify 6051 lysine crotonylation sites in 2508 protein groups. Most of these proteins had fewer than three crotonylation sites and involved diverse biological processes, cellular components, and molecular functions. These data show that crotonylated proteins play a very important role in peanut biological functions.

The Motif-X program showed that the conserved motif adjacent to the Kcr site is clearly conserved; the EK and DK sites have also been associated with Kcr sites in tobacco, rice, and animals [18–20]. With respect to the subcellular localization of crotonylated proteins, the distributions reported for rice, papaya, tea and tobacco were identical, with the order of location frequencies, from high to low, being in the chloroplast, cytosol, nuclear and mitochondrion. The pattern in peanut was very similar, with the first three locations being the same as for the earlier-reported plants, but the fourth location on the list in peanut was the plasma membrane, rather than the mitochondrion.

Photosynthesis is an essential biological process for plants, converting light energy into chemical energy and storing it in the bonds of sugars, the process taking place in the structures of chloroplasts. Similar to other plants, the highest frequency of crotonylated proteins in subcellular fractions of peanut leaves (43%) was located in the chloroplast. These data implied that these crotonylated proteins play an important part in the photosynthesis process. The subsequent GO enrichment analysis showed that crotonylated proteins were associated with biosynthetic processes in the thylakoid parts of the cell. The KEGG pathway-based enrichment analysis showed that many crotonylated proteins take part in photosynthetic carbon fixation. In the peanut photosynthesis network, 29 crotonylated proteins were identified as nodes in the protein interaction database. These proteins were involved in photosystem II, photosystem I, the cytochrome b6/f complex, photosynthetic electron transport, and F-type ATPase. The KEGG pathway enrichment analysis and the network of Kcr also showed that many ribosomal proteins were crotonylated. A similar phenomenon was reported in tobacco [18] and rice [20]. These findings suggest that lysine crotonylation may play important regulatory roles in peanut photosynthesis and ribosome structure and function.

There is considerable research interest in Kac, which has been reported in many plants, such as *Arabidopsis* [27, 28], rice [29], wheat [27], and soybean [30]. A number of acetylated proteins have been localized in the chloroplast and have been shown to take part in photosynthesis. Kcr and Kac can occur on the same protein, but they may not have exactly the same substrate and modification sites [14]. In rice, there were 21 acetylated proteins and 16 overlapping proteins, which were both crotonylated and acetylated [20, 29]. The intracellular Kcr could promote transcription; compared with acetylation,

crotonylation has a greater ability to promote gene expression, and the balance between crotonylation and acetylation (another PTM) has functional effects on gene expression [10]. Histone crotonylation reprogrammed the function of the nucleosome, distinguishing it from histone acetylation, by interacting with a specific set of chromatin modifiers [23, 31–33]. These lysine modifications can be changed by environmental (for instance, starvation or submergence) or metabolic cues, controlling the relative proportions in histone-acylated plants, which, in turn, control the expression of responsive genes [34].

In animals, the relative proportions of mixed histone acylations have a functional association with different physiological stages [10, 32, 35]. Kcr occurs at a large number of histone sites [3], where it may affect the structure of chromatin [36]. Histone crotonylation is abundant in many human tissues, is linked to regulation of gene expression and is highly dynamic, being regulated during the cell cycle [37]. It has been reported that the crucial role of Kac and methylation in DNA-templated processes is likely to have significant consequences for chromatin structure and function [8]. In turn, chromatin structure affects gene expression to a considerable extent [27]. More research needs to be carried out to determine whether Kcr, Kac and other PTMs play important roles in peanut biological progresses.

PTMs impact a diverse array of cellular processes, and Kac is one of the main PTMs studied. It has been reported that Kac plays a central role in the host–pathogen interaction [38]. Research into plant defense signaling has established that modulation of histone acetylation is very important for immune responses [39, 40]. Glutathione (GSH) is an important reducing substance commonly present in plant, which plays an important role in the antioxidant defense against the oxidation of membrane lipids by free radicals. The GSH-conjugating enzymes, glutathione S-transferases (GSTs), are present in different subcellular compartments, and up-regulation appear to be an evolutionary response by cells for the protection against chemical toxicity and oxidative stress [41–43]. Plant GST genes are expressed differently when plants encounter biotic and abiotic stresses. The stress-induced plant growth regulator, nitric oxide (NO), differentially affects the expression level of GSTs [44, 45]. Our research showed that there were 19 lysine-crotonylated proteins involved in glutathione metabolism (Additional file 10: Figure S1). Glutathione S-transferase, C-terminal-like and Glutathione S-transferase N-terminal were obviously enriched in crotonylated proteins. These results suggested that crotonylated proteins in GSH may play an important role in plant response to biotic and abiotic stresses. There has been some research on Kcr involvement in response to stress in different plant species [26, 34]. We need to narrow down the list of candidate proteins of interest and then to perform functional studies on the target candidates.

Conclusion

In peanut leaves, we identified 6051 lysine-crotonylation sites in 2508 protein groups. Bioinformatics analysis showed that lysine crotonylated proteins were involved in many biological processes, including carbon fixation in photosynthetic organisms, photosynthesis, biosynthesis of amino acids, ribosomes, etc. We demonstrated that Kcr may play an important role in peanut growth and tolerance to adverse environments. The status of peanut Kcr may represent a crucial role for this type of PTM, opening up new research avenues in this area. Future work should expand on the results from bioinformatic analyses,

selecting those markedly enriched proteins/clusters in particular pathways, processes, molecular functions, localizations and protein complexes, etc. Our results provide the foundation for future studies of the effects of crotonylation on performance of the important cash crop, the peanut plant.

Declarations

Acknowledgments

This research was funded by the National Natural Science Foundation of China (31801711); the Natural Science Foundation of Shandong Province (ZR2018LC015); Agricultural Science and Technology Innovation Project of Shandong Academy of Agricultural Sciences (CXGC2018E21, CXGC2016B11); Shandong Agriculture Research System (SDAIT-04-07); Sichuan Provincial Tobacco Company Science and Technology Project (SCYC201804); Yunnan Provincial Tobacco Company Science and Technology Project (2018530000241014); Major Tobacco Green Prevention and Control Project 110201901041(Is-04).

Funding

Dr Manlin Xu receives funding from the National Natural Science Foundation of China (31801711).

Ethics approval and consent to participate

Not applicable. The research conducted in this study required neither approval from an ethics committee, nor involved any human or animal subjects.

Consent for publication

Not applicable.

Competing interests

The authors declare that they have no competing interests.

Authors' contributions

Manlin Xu, Jianda Luo, Yucheng Chi and Jinguang Yang carried out, design the experiment, collected and organized data and wrote the manuscript. Ying Li, Lili Shen, Xia Zhang, Jing Yu, Zhiqing Guo and Juxiang Wu participated in designing, carried out the experiment and reviewed the manuscript.

References

1. Boussouar F, Rousseaux S, Khochbin S: **A new insight into male genome reprogramming by histone variants and histone code.** *Cell Cycle* 2008, **7**(22):3499-3502.

2. Chen Y, Chen W, Cobb MH, Zhao Y: **PTMap—a sequence alignment software for unrestricted, accurate, and full-spectrum identification of post-translational modification sites.** *Proc Natl Acad Sci U S A* 2009, **106**(3):761-766.
3. Tan M, Luo H, Lee S, Jin F, Yang JS, Montellier E, Buchou T, Cheng Z, Rousseaux S, Rajagopal N *et al*: **Identification of 67 histone marks and histone lysine crotonylation as a new type of histone modification.** *Cell* 2011, **146**(6):1016-1028.
4. Park J, Chen Y, Tishkoff DX, Peng C, Tan M, Dai L, Xie Z, Zhang Y, Zwaans BM, Skinner ME *et al*: **SIRT5-mediated lysine desuccinylation impacts diverse metabolic pathways.** *Mol Cell* 2013, **50**(6):919-930.
5. Meyer JG, D'Souza AK, Sorensen DJ, Rardin MJ, Wolfe AJ, Gibson BW, Schilling B: **Quantification of Lysine Acetylation and Succinylation Stoichiometry in Proteins Using Mass Spectrometric Data-Independent Acquisitions (SWATH).** *J Am Soc Mass Spectrom* 2016, **27**(11):1758-1771.
6. Chen L, Miao Z, Xu X: **beta-hydroxybutyrate alleviates depressive behaviors in mice possibly by increasing the histone3-lysine9-beta-hydroxybutyrylation.** *Biochem Biophys Res Commun* 2017, **490**(2):117-122.
7. Huang J, Luo Z, Ying W, Cao Q, Huang H, Dong J, Wu Q, Zhao Y, Qian X, Dai J: **2-Hydroxyisobutyrylation on histone H4K8 is regulated by glucose homeostasis in *Saccharomyces cerevisiae*.** *Proc Natl Acad Sci U S A* 2017, **114**(33):8782-8787.
8. Xie Z, Dai J, Dai L, Tan M, Cheng Z, Wu Y, Boeke JD, Zhao Y: **Lysine succinylation and lysine malonylation in histones.** *Mol Cell Proteomics* 2012, **11**(5):100-107.
9. Binder H, Steiner L, Przybilla J, Rohlf T, Prohaska S, Galle J: **Transcriptional regulation by histone modifications: towards a theory of chromatin re-organization during stem cell differentiation.** *Phys Biol* 2013, **10**(2):026006.
10. Sabari BR, Tang Z, Huang H, Yong-Gonzalez V, Molina H, Kong HE, Dai L, Shimada M, Cross JR, Zhao Y *et al*: **Intracellular crotonyl-CoA stimulates transcription through p300-catalyzed histone crotonylation.** *Mol Cell* 2015, **58**(2):203-215.
11. Sabari BR, Tang Z, Huang H, Yong-Gonzalez V, Molina H, Kong HE, Dai L, Shimada M, Cross JR, Zhao Y *et al*: **Intracellular Crotonyl-CoA Stimulates Transcription through p300-Catalyzed Histone Crotonylation.** *Mol Cell* 2018, **69**(3):533.
12. Wei W, Liu X, Chen J, Gao S, Lu L, Zhang H, Ding G, Wang Z, Chen Z, Shi T *et al*: **Class I histone deacetylases are major histone decrotonylases: evidence for critical and broad function of histone crotonylation in transcription.** *Cell Res* 2017, **27**(7):898-915.
13. Xu W, Wan J, Zhan J, Li X, He H, Shi Z, Zhang H: **Global profiling of crotonylation on non-histone proteins.** *Cell Res* 2017, **27**(7):946-949.
14. Wei W, Mao A, Tang B, Zeng Q, Gao S, Liu X, Lu L, Li W, Du JX, Li J *et al*: **Large-Scale Identification of Protein Crotonylation Reveals Its Role in Multiple Cellular Functions.** *J Proteome Res* 2017, **16**(4):1743-1752.

15. Wu Q, Li W, Wang C, Fan P, Cao L, Wu Z, Wang F: **Ultradeep Lysine Crotonylome Reveals the Crotonylation Enhancement on Both Histones and Nonhistone Proteins by SAHA Treatment.** *J Proteome Res* 2017, **16**(10):3664-3671.
16. Huang H, Wang DL, Zhao Y: **Quantitative Crotonylome Analysis Expands the Roles of p300 in the Regulation of Lysine Crotonylation Pathway.** *Proteomics* 2018, **18**(15):e1700230.
17. Chen W, Tang D, Xu Y, Zou Y, Sui W, Dai Y, Diao H: **Comprehensive analysis of lysine crotonylation in proteome of maintenance hemodialysis patients.** *Medicine (Baltimore)* 2018, **97**(37):e12035.
18. Sun H, Liu X, Li F, Li W, Zhang J, Xiao Z, Shen L, Li Y, Wang F, Yang J: **First comprehensive proteome analysis of lysine crotonylation in seedling leaves of *Nicotiana tabacum*.** *Sci Rep* 2017, **7**(1):3013.
19. Kwon OK, Kim SJ, Lee S: **First profiling of lysine crotonylation of myofilament proteins and ribosomal proteins in zebrafish embryos.** *Sci Rep* 2018, **8**(1):3652.
20. Liu S, Xue C, Fang Y, Chen G, Peng X, Zhou Y, Chen C, Liu G, Gu M, Wang K *et al*: **Global Involvement of Lysine Crotonylation in Protein Modification and Transcription Regulation in Rice.** *Mol Cell Proteomics* 2018, **17**(10):1922-1936.
21. Yang Q, Li Y, Apaliya MT, Zheng X, Serwah BNA, Zhang X, Zhang H: **The Response of *Rhodotorula mucilaginosa* to Patulin Based on Lysine Crotonylation.** *Front Microbiol* 2018, **9**:2025.
22. Liu K, Yuan C, Li H, Chen K, Lu L, Shen C, Zheng X: **A qualitative proteome-wide lysine crotonylation profiling of papaya (*Carica papaya* L.).** *Sci Rep* 2018, **8**(1):8230.
23. Xiong X, Panchenko T, Yang S, Zhao S, Yan P, Zhang W, Xie W, Li Y, Zhao Y, Allis CD *et al*: **Selective recognition of histone crotonylation by double PHD fingers of MOZ and DPF2.** *Nat Chem Biol* 2016, **12**(12):1111-1118.
24. Cline MS, Smoot M, Cerami E, Kuchinsky A, Landys N, Workman C, Christmas R, Avila-Campilo I, Creech M, Gross B *et al*: **Integration of biological networks and gene expression data using Cytoscape.** *Nat Protoc* 2007, **2**(10):2366-2382.
25. Perez-Riverol Y, Csordas A, Bai J, Bernal-Llinares M, Hewapathirana S, Kundu DJ, Inuganti A, Griss J, Mayer G, Eisenacher M *et al*: **The PRIDE database and related tools and resources in 2019: improving support for quantification data.** *Nucleic Acids Res* 2019, **47**(D1):D442-D450.
26. Sun J, Qiu C, Qian W, Wang Y, Sun L, Li Y, Ding Z: **Ammonium triggered the response mechanism of lysine crotonylome in tea plants.** *BMC Genomics* 2019, **20**(1):340.
27. Li L, Shi L, Yang S, Yan R, Zhang D, Yang J, He L, Li W, Yi X, Sun L *et al*: **SIRT7 is a histone desuccinylase that functionally links to chromatin compaction and genome stability.** *Nat Commun* 2016, **7**:12235.
28. Wu X, Oh MH, Schwarz EM, Larue CT, Sivaguru M, Imai BS, Yau PM, Ort DR, Huber SC: **Lysine acetylation is a widespread protein modification for diverse proteins in *Arabidopsis*.** *Plant Physiol* 2011, **155**(4):1769-1778.
29. Xiong Y, Peng X, Cheng Z, Liu W, Wang GL: **A comprehensive catalog of the lysine-acetylation targets in rice (*Oryza sativa*) based on proteomic analyses.** *J Proteomics* 2016, **138**:20-29.

30. Smith-Hammond CL, Swatek KN, Johnston ML, Thelen JJ, Miernyk JA: **Initial description of the developing soybean seed protein Lys-N(epsilon)-acetylome.** *J Proteomics* 2014, **96**:56-66.
31. Zhao D, Guan H, Zhao S, Mi W, Wen H, Li Y, Zhao Y, Allis CD, Shi X, Li H: **YEATS2 is a selective histone crotonylation reader.** *Cell Res* 2016, **26**(5):629-632.
32. Li Y, Sabari BR, Panchenko T, Wen H, Zhao D, Guan H, Wan L, Huang H, Tang Z, Zhao Y *et al*: **Molecular Coupling of Histone Crotonylation and Active Transcription by AF9 YEATS Domain.** *Mol Cell* 2016, **62**(2):181-193.
33. Andrews FH, Shinsky SA, Shanle EK, Bridgers JB, Gest A, Tsun IK, Krajewski K, Shi X, Strahl BD, Kutateladze TG: **The Taf14 YEATS domain is a reader of histone crotonylation.** *Nat Chem Biol* 2016, **12**(6):396-398.
34. Lu Y, Xu Q, Liu Y, Yu Y, Cheng ZY, Zhao Y, Zhou DX: **Dynamics and functional interplay of histone lysine butyrylation, crotonylation, and acetylation in rice under starvation and submergence.** *Genome Biol* 2018, **19**(1):144.
35. Xie Z, Zhang D, Chung D, Tang Z, Huang H, Dai L, Qi S, Li J, Colak G, Chen Y *et al*: **Metabolic Regulation of Gene Expression by Histone Lysine beta-Hydroxybutyrylation.** *Mol Cell* 2016, **62**(2):194-206.
36. Mersfelder EL, Parthun MR: **The tale beyond the tail: histone core domain modifications and the regulation of chromatin structure.** *Nucleic Acids Res* 2006, **34**(9):2653-2662.
37. Fellows R, Denizot J, Stellato C, Cuomo A, Jain P, Stoyanova E, Balazsi S, Hajnady Z, Liebert A, Kazakevych J *et al*: **Microbiota derived short chain fatty acids promote histone crotonylation in the colon through histone deacetylases.** *Nat Commun* 2018, **9**(1):105.
38. Song G, Walley JW: **Dynamic Protein Acetylation in Plant-Pathogen Interactions.** *Front Plant Sci* 2016, **7**:421.
39. Defraia CT, Wang Y, Yao J, Mou Z: **Elongator subunit 3 positively regulates plant immunity through its histone acetyltransferase and radical S-adenosylmethionine domains.** *BMC Plant Biol* 2013, **13**:102.
40. Wang Y, An C, Zhang X, Yao J, Zhang Y, Sun Y, Yu F, Amador DM, Mou Z: **The Arabidopsis elongator complex subunit2 epigenetically regulates plant immune responses.** *Plant Cell* 2013, **25**(2):762-776.
41. Raza H: **Dual localization of glutathione S-transferase in the cytosol and mitochondria: implications in oxidative stress, toxicity and disease.** *FEBS J* 2011, **278**(22):4243-4251.
42. Hayes JD, Flanagan JU, Jowsey IR: **Glutathione transferases.** *Annu Rev Pharmacol Toxicol* 2005, **45**:51-88.
43. Allocati N, Federici L, Masulli M, Di Ilio C: **Glutathione transferases in bacteria.** *FEBS J* 2009, **276**(1):58-75.
44. Moons A: **Regulatory and functional interactions of plant growth regulators and plant glutathione S-transferases (GSTs).** *Vitam Horm* 2005, **72**:155-202.

45. Francis F, Vanhaelen N, Haubruge E: **Glutathione S-transferases in the adaptation to plant secondary metabolites in the Myzus persicae aphid.** *Arch Insect Biochem Physiol* 2005, **58**(3):166-174.

Supplemental Information Note

Additional file 1: Table S1 Basic analysis of MS identified information.

Additional file 2: Table S2 Annotation combine of identified proteins.

Additional file 3: Table S3 Detailed information of GO terms of level 2 distribution of all proteins.

Additional file 4: Table S4 Detailed information of subcellular localization of the identified protein.

Additional file 5: Table S5 GO enrichment analysis of crotonylated proteins.

Additional file 6: Table S6 Detailed information of KEGG pathway.

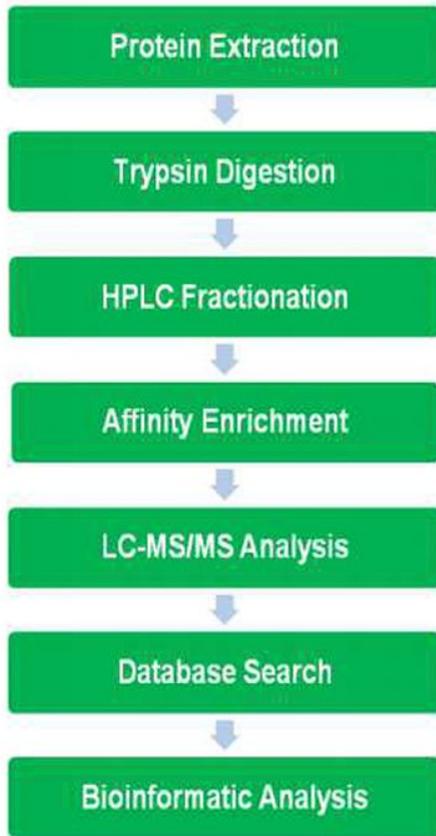
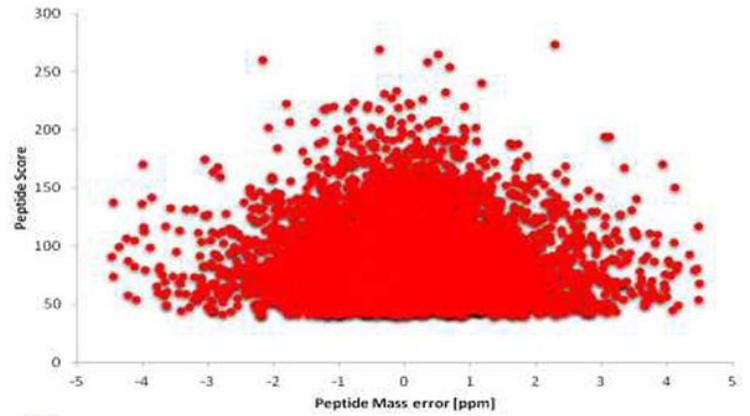
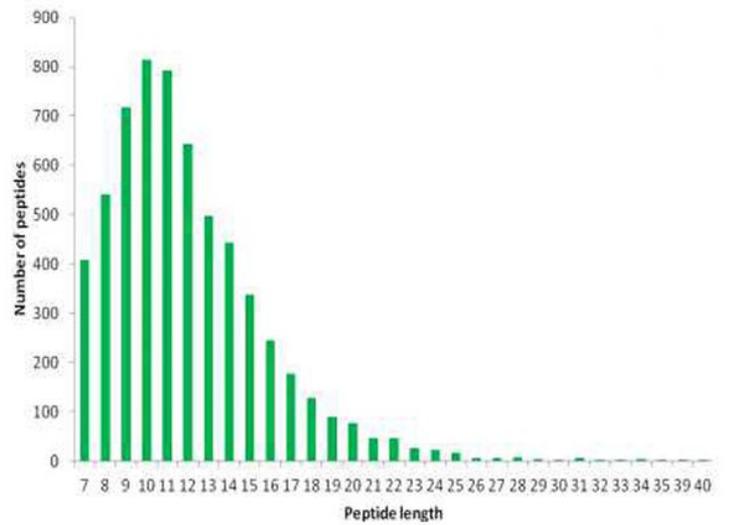
Additional file 7: Table S7 Detailed information of protein domain.

Additional file 8: Table S8 Detailed information of photosynthesis related protein.

Additional file 9: Table S9 Detailed information of ribosome related protein.

Additional file 10: Figure S1 Glutathione metabolism pathway enriched from KEGG pathway analysis. The proteins in white are background, the proteins in red are identified in this work.

Figures

A**B****C****Figure 1**

Technical route and the basic information of LC-MS/MS data. a The general technical route for lysine crotonylation identification was showed. All crotonylated peptides were enriched by an antibody and analyzed by LC-MS/MS. b Mass error distribution of all identified peptides. c Peptide length distribution.

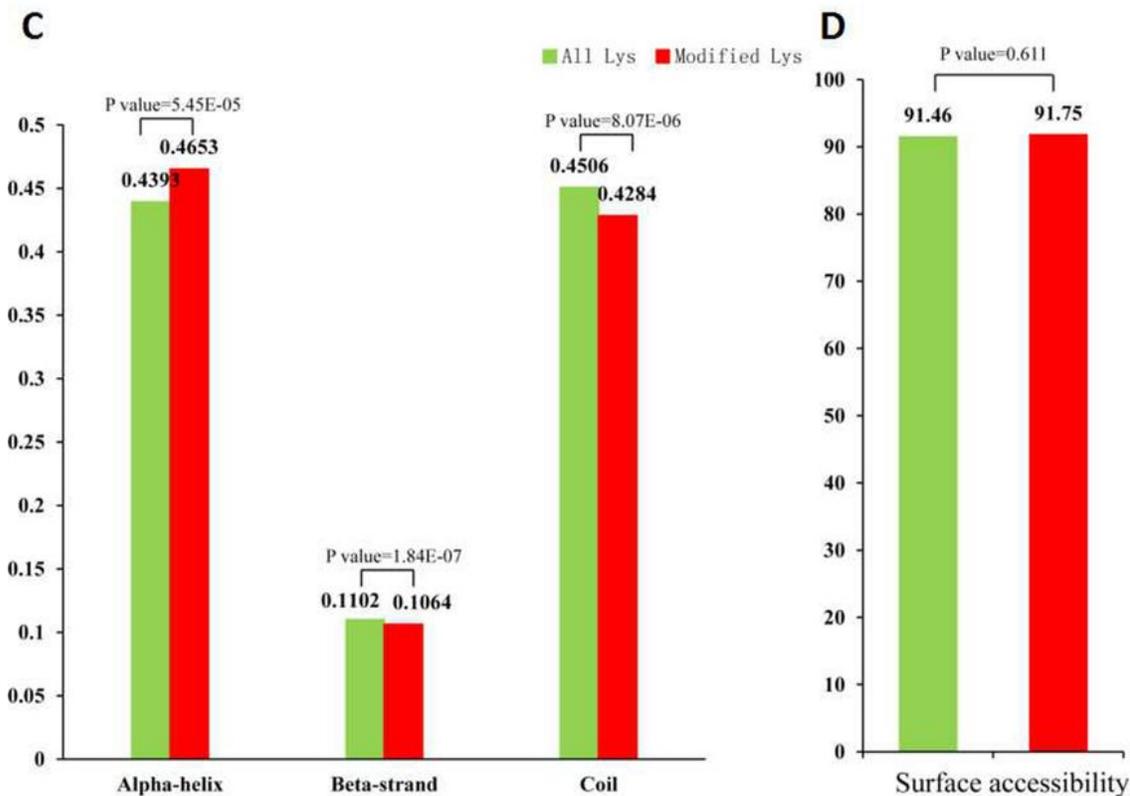
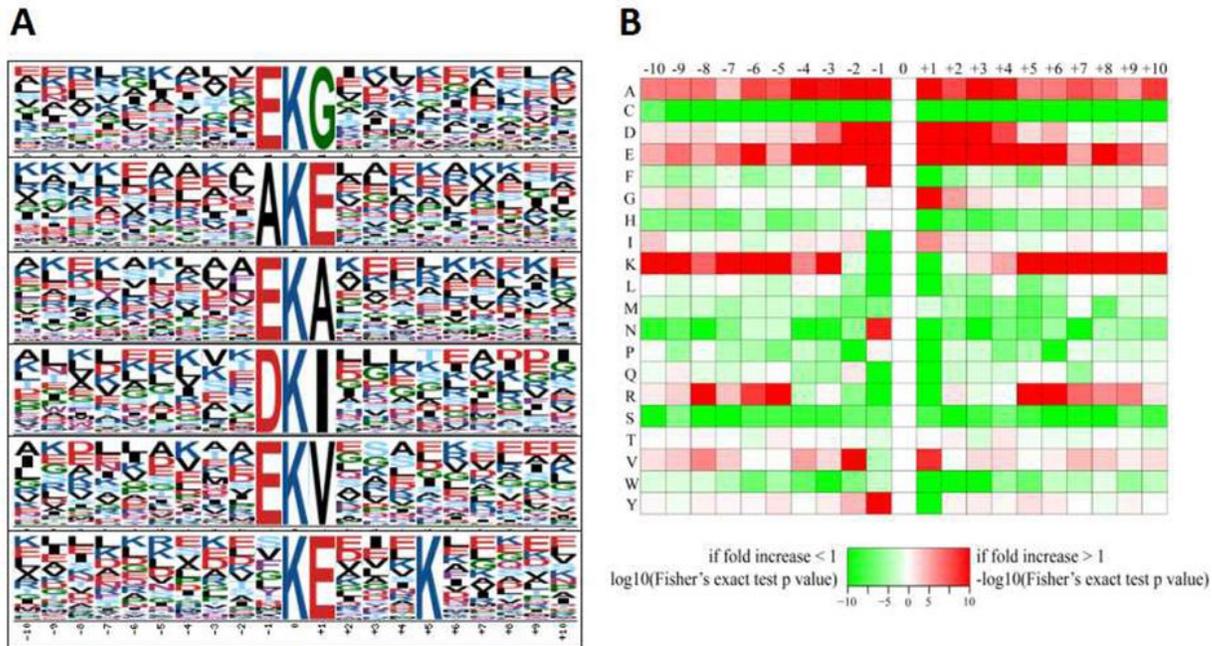


Figure 2

Motif analysis of lysine crotonylation peptides. a Motif analysis showed crotonylated peptide motif and conservation of Kcr sites. The intensity map shows enrichment of amino acids in particular positions of crotonylated lysine (10 amino acids upstream and downstream of the site). b Motif analysis of all the identified sites. Probability sequence motifs of crotonylation sites consisting of 10 residues surrounding the targeted lysine residue using Motif-X. c Probabilities of lysine crotonylation in different protein

secondary structures (alpha helix, beta-strand and coil). d Predicted surface accessibility of crotonylation sites.

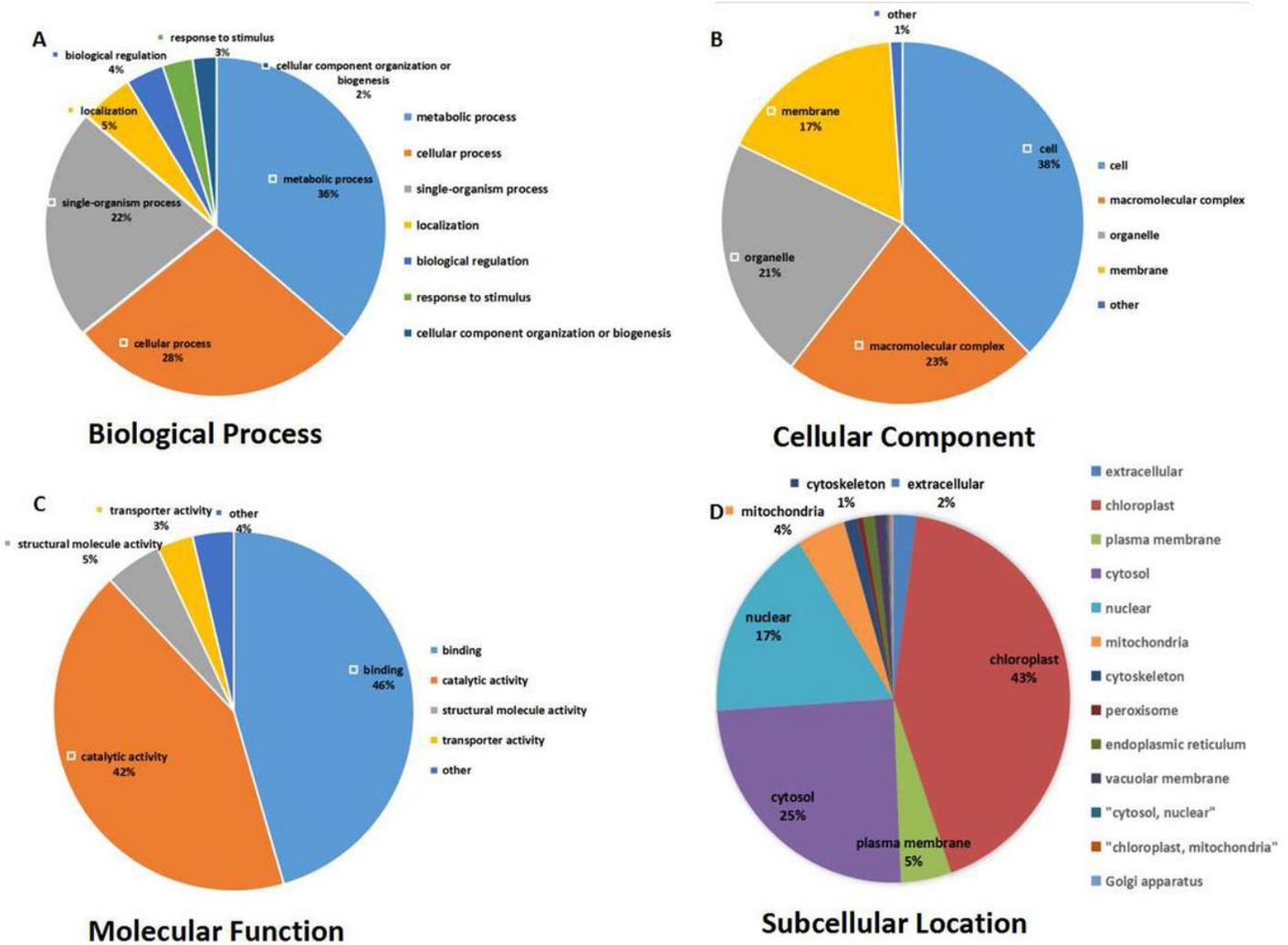


Figure 3

Functional classification of crotonylated proteins. Functional classification of the identified crotonylated proteins based on a biological processes. b cellular component. c molecular function. d subcellular localization.

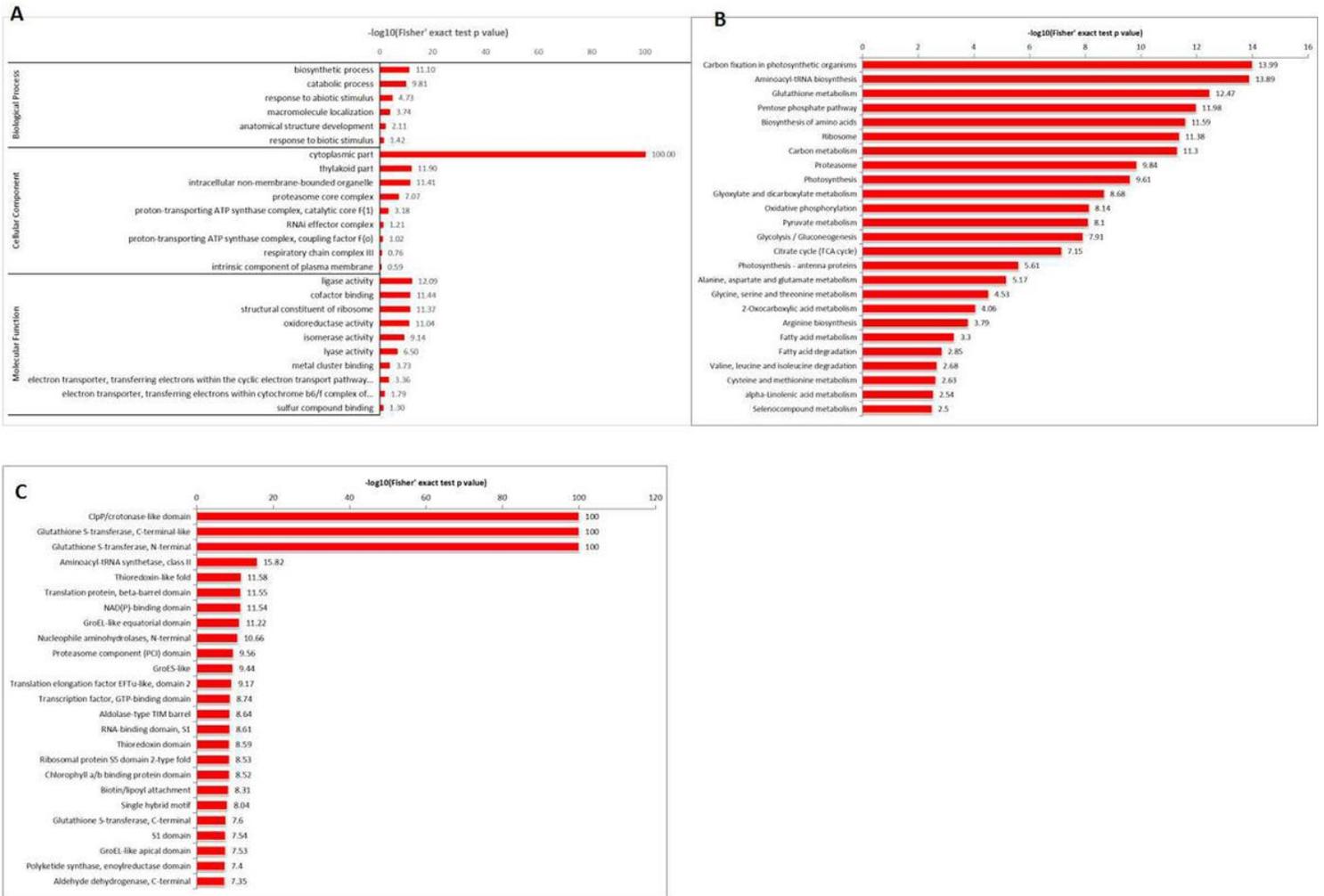
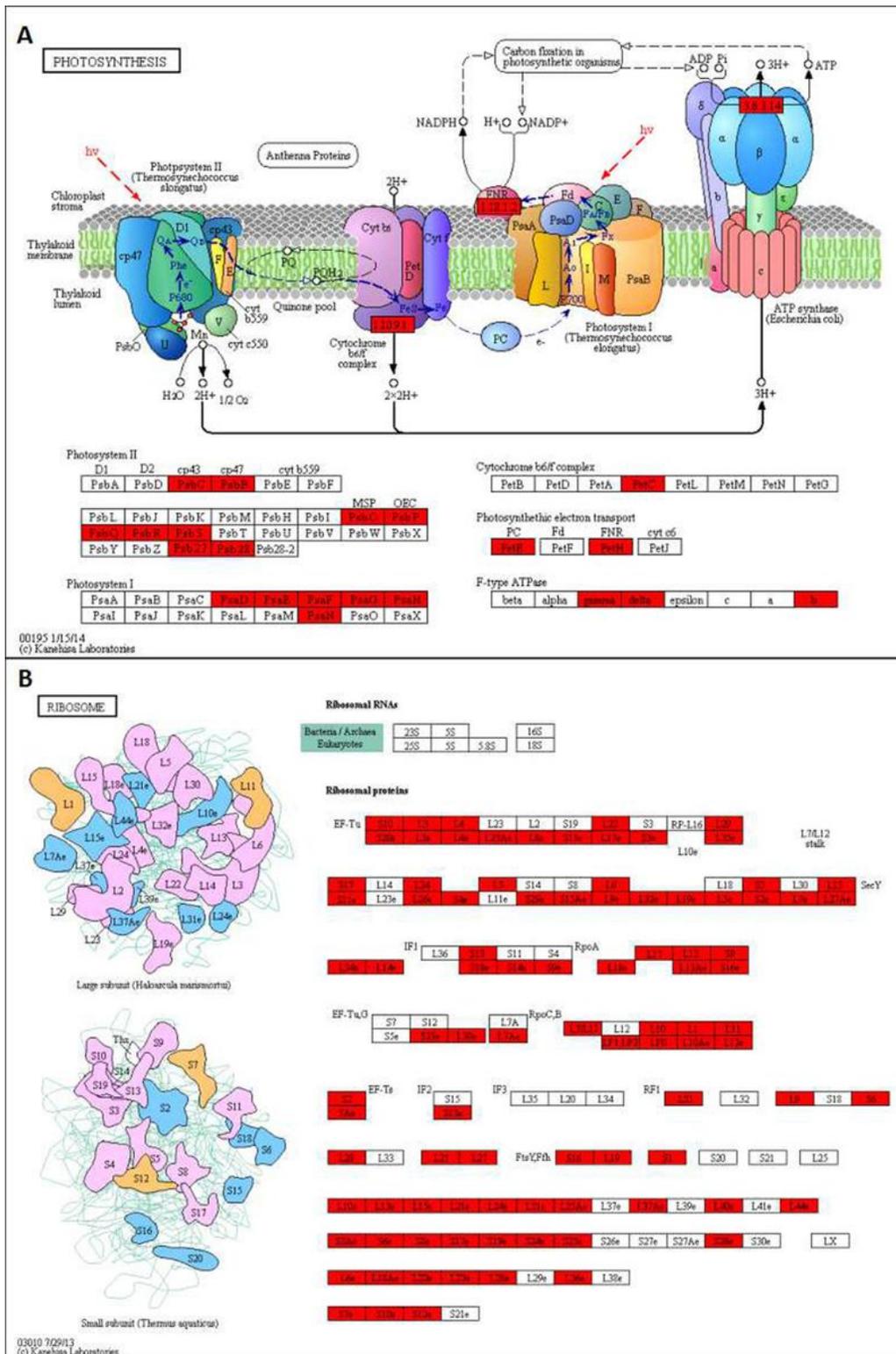


Figure 4

Enrichment analysis of crotonylated proteins. a GO-based enrichment analysis of identified proteins. b KEGG pathway-based enrichment analysis. c Protein domain enrichment analysis of all identified proteins



Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- [Additionalfile2TableS2.xlsx](#)
- [Additionalfile9TableS9.xlsx](#)
- [Additionalfile10FigureS1.jpg](#)
- [Additionalfile4TableS4.xlsx](#)
- [Additionalfile5TableS5.xlsx](#)
- [Additionalfile6TableS6.xlsx](#)
- [Additionalfile7TableS7.xlsx](#)
- [Additionalfile8TableS8.xlsx](#)
- [Additionalfile3TableS3.xlsx](#)
- [Additionalfile1TableS1.xlsx](#)