

Systematic Qualitative Proteome-wide Analysis of Lysine Malonylation Profiling in *Platycodon grandiflorus*

Qingshan Yang

Anhui University of Chinese Medicine

Shaowei Xu

Anhui University of Chinese Medicine

Weimin Jiang

Hengyang Normal University

Fei Meng

Anhui University of Chinese Medicine

Shuting Wang

Anhui University of Chinese Medicine

Zongping Sun

Chinese Herbal Medicine, Fuyang Normal University

Liangping Zha

Anhui University of Chinese Medicine

Shuangying Gui

Anhui University of Chinese Medicine

Daiyin Peng

Anhui University of Chinese Medicine

Juan Liu

Anhui University of Chinese Medicine

Shihai Xing (✉ xshshihai@163.com)

Anhui University of Chinese Medicine

Research Article

Keywords: *Platycodon gradiflorus*, Lysine malonylation, Saponin biosynthesis, Polysaccharide biosynthesis, Heavily-malonylated

Posted Date: December 8th, 2022

DOI: <https://doi.org/10.21203/rs.3.rs-2345793/v1>

Abstract

Background

The modification of lysine malonylation found in recent years can affect biological metabolism and play an important role in life activities. *Platycodon grandiflorus*, as an economic crop and medicinal plant with a wide range of applications and has broad prospects for development. However, there is no reports on malonylation in the related research of *P. grandiflorus*. This study qualitatively introduces the malonylation of lysine in *P. grandiflorus*.

Results

888 lysine malonylation modified proteins in *P. grandiflorus* were identified, with a total of 1755 modification sites. According to the functional annotation, malonylation proteins were closely related to catalysis, binding and other reactions. Subcellular localization showed that related proteins were enriched in chloroplasts, cytoplasm and nuclei, indicating that this modification could regulate a variety of biological metabolic processes. Motif analysis showed that the enrichment of Alanine (A), Cysteine (C), Glycine (G) and Valine (V) amino acids surrounding malonylated lysine residues. Metabolic pathway and protein-protein interaction network analyses suggested modification is mainly involved in plant photosynthesis. Moreover, malonylated proteins also involved in stress and defense responses.

Conclusion

This study shows that lysine malonylation can affect a variety of biological processes and metabolic pathways, and these contents are reported for the first time in *P. grandiflorus*, which can provide important information for further research on *P. grandiflorus* and lysine malonylation.

Introduction

Protein modification plays an important role in biological organisms. It changes the composition and structure of proteins through methylation, glycosylation, hydroxylation and glyoxylation etc. (Lin and Carroll 2018), so as to regulate the corresponding biosynthetic process. So far, more than 600 kinds of protein post translational modifications (PTMs) have been found. Unlike animals and bacteria, the protein PTMs process in plants is more to resist external invasion or adapt to environmental changes (Taguchi et al. 2010). Malonylation is active in the biosynthesis of flavonoids and isoflavones, and there are clear traces of malonylation in soybeans (Ahmad et al. 2021; Ahmad et al. 2017), corn (Xu et al. 2021), rice (Mujahid et al. 2018) and tobacco (Manjasetty et al. 2012). If some small molecules are to undergo enzymatic acylation reaction, they need some external assistance, that is, high-energy acyl donors (Yu et al. 2008), such as acetyl coenzyme A thioester and 1-O-acyl glucoside. Moreover, 30% –

45% malonylation occurs in chloroplasts (Berghold et al. 2004; Xu et al. 2021), serving for plant photosynthesis.

Lysine malonylation is a reversible reaction adding malonyl to lysine residues. It was found earlier in microorganisms and mammals, such as *Toxoplasma gondii*, *Escherichia coli* (Nie et al. 2020) and human (Nahomi et al. 2020) and its evolution is conservative. Malonylation can affect many biochemical processes, including a variety of cell functions and disease processes. The existing data show that the protein modified by malonylation enriches in the pathways of energy metabolism and fatty acid synthesis (Zhang et al. 2019; Qian et al. 2016), in cells with damaged mitochondrial function and the process of fatty acid oxidation (Wang et al. 2017), which shows that malonylation is related to energy metabolism. Only few studies about malonylation in plants were studied, including tobacco (Manjasetty et al. 2012), corn (Xu et al. 2021), rice (Mujahid et al. 2018), and wheat (Liu et al. 2018), but there is no relevant report on an important medicinal plant *P. grandiflorus*. The research in plants suggests that malonylation is also closely related to energy metabolism in plants, similar to that in microorganisms and mammals. Although malonylation can affect a variety of biochemical processes, due to the dynamic characteristics and low abundance of malonylation (Chen et al. 2018; Liu et al. 2020), malonylation sites on proteins in plants cannot be enough analyzed. The matching cell annotation, that is, the recognition of modification sites, is the most important step to clarify the relevant mechanism (Wang et al. 2017), which limits the functional study of malonylation. Although malonylation sites prediction tools, such as Long Short-Term Memory (LSTM) (Al-Barakati et al. 2020), Malonylation Prediction (Mal-Prec) (Liu et al. 2020) and LSTM-based Ensemble Malonylation Predictor (LEMP) (Chen et al. 2018) were developed by scientists in recent years, conventional mass-based proteomic approach is a more reliable way in malonylation sites identification.

Beside malonylation, there is also another modification of lysine (Nahomi et al. 2020). Malonylation and succinylation have similar modification structures (Taherzadeh et al. 2018). So malonylation may be reversible, dynamic and conservative and able to influence the physiology and pathology of animals, the photosynthesis and Tricarboxylic acid cycle (TCA cycle) of plants, cell cycle and growth in the process of life like succinylation (Dong et al. 2021; Sreedhar et al. 2020; Wang et al. 2019). Malonylation and succinylation are both short chain acylation modifications and can change the charge state of amino acids (Colak et al. 2015). This acidic modification may be quite different from those widely studied modification such as acetylation (Qian et al. 2016; Colak et al. 2015).

P. grandiflorus is the only plant in the genus *Platycodon* in the family Campanulaceae. Nowadays, the main production areas of *P. grandiflorus* are located in Shandong, Anhui and Inner Mongolia of China, and the seed resources are relatively rich (WANG Zhi-fen et al. 2006). So far, more than 100 kinds of secondary metabolites have been isolated from *P. grandiflorus*, including triterpenoids, saponins, flavonoids, polysaccharides, etc. (Jia et al. 2022), and more than 70 kinds of major active ingredients have been identified (Huang et al. 2021), among which triterpenoid saponins are the effective compounds, including platycodin D, platycodin A, platycodonic acid, etc. (Su et al. 2021). These organics are used as the chemical basis for clinical treatment of cancer (Deng et al. 2020b) and bronchitis (Deng

et al. 2020a) and other diseases. In addition, *P. grandiflorus* can also be materials for gardening and food (Jia et al. 2022). Therefore, its potential economic value is very high. Nevertheless, it is still limited in the study about the metabolic pathway and protein modification of *P. grandiflorus*.

Lysine malonylation of *P. grandiflorus* has not been reported in this high-value economic plant. In this study, we assess the lysine malonylation in *P. grandiflorus* by Western blot and mass-spectrum quality control analysis. Conservative motifs surrounding modification sites, the Gene Ontology (GO) enrichment, and the Kyoto Encyclopedia of Genes and Genomes (KEGG) of modified proteins were analyzed. Our study is the first qualitative investigation of lysine malonylation modification in *P. grandiflorus*. Those achievements are probable to serve as the foundation of follow-up research and accelerate the development of potential application of malonylation in plant physiology and biology in the future.

Materials And Methods

Plant Materials

Plant used as experiment materials is identified as *Platycodon grandiflorus* (Jacq.) A. DC. by Qingshan Yang who is an associate professor in Anhui University of Chinese Medicine. Before detected, those sample include root, stem and leaf of *P. grandifloras* must be washed by distilled water and alcohol and stored in liquid nitrogen.

Protein Extraction

Protein extraction

Proteins were extracted according to the procedures of our previous study (Palomba et al. 2021; Shi et al. 2021) by slight modification. Grind three duplicates of appropriate amount of samples (-80°C) into powder in mortars which were precooled by liquid nitrogen. Each sample group was ultrasonically split after adding buffer (10 mM DTT, 1% protease inhibitor cocktail, 3 µM TSA, 50 mM NAM) which was 4 times volume of sample. Before centrifugation (4°C, 10 min 5500 g), the same volume Tris-saturated phenol was added. While the supernatant was precipitated by ammonium acetate/methanol, the remaining precipitate was washed by ammonium acetate/methanol and determined with BCA kit to identify its protein concentration after redissolved in 8 M urea.

Western Blot Analysis

Extracted Proteins (10 µg) of *P. grandiflorus* was separated using 15.0% SDS-PAGE gel and then electro-transferred to a polyvinylidene fluoride (PVDF) membrane (Millipore) for Western blot analysis. Malonylated proteins were tested by primary anti-malonyllysine antibody (PTM-902, Lot: 23056103M119) in a 1: 2000 (v/v) dilution. The secondary antibody goat anti-mouse IgG [Thermo, Pierce, (H + L), Peroxidase Conjugated, 31430] was diluted in 1: 10,000 (v/v).

Protein Trypsin Digestion

Every equal volume sample was enzymatically broken, then their volume was adjusted to the same value by lysis solution. After putting TCA (20%) in sample and mixing, they were precipitated (4°C, 2 h). The precipitation was washed by acetone which had been precooled and was left to dry in the air. We made the precipitation been broken by ultrasonic with 200 mM TEAB and hydrolyzed by trypsin (protease/protein: 1/50). It had been rested 30 min (56°C) after its volume was adjusted to 2 mM by DTT. We adjust its volume to 11 mM by IAA and stored it in darkness at room temperature (15 min).

Modification Enrichment

The supernatant which was got when putting peptide in IP buffer (100 mM NaCl, 1 mM EDTA, 50 mM Tris-HCl, 0.5% NP-40, pH 8.0) was moved into malonyl resin (Article Number: PTM-904). We made it incubate on rotary shaker for one night (4°C). It had been flushed with IP buffer (4 times), deionized water (2 times) and 0.1% TFA (3 times) successively. Then it was drained by vacuum freezing. After desalting according to instruction named C18 ZipTips, we drained it again. Finally, it was sent to analysis by LC-MS/MS.

Lc-ms/ms Analysis

The tryptic peptides were dissolved in solvent A (0.1% formic acid, 2% acetonitrile/in water) and separated solvent B (0.1% formic acid, 100% acetonitrile/in water) with a gradient (0–42 min, 6%~24% B; 42–56 min, 24%~32% B; 56–58 min, 32%~80% B; 58–60 min, 80% B) at a constant flowrate of 450 nL/min on NanoElute. After ionization separation (1.6 kV), mass spectrometer collected its mass spectra for 1 times and secondary mass spectrogram for 10 times (the dynamic elimination time of tandem mass spectrometry scanning is 30 seconds to avoid repeated scanning of parent ions). Tryptic peptides parent ions and secondary fragments were detected and analyzed by TOF. The scanning range of the secondary branch is 100–1700, and the data acquisition mode is parallel cumulative serial fragmentation mode.

Analysis data shows that most the first-order mass error of spectra were under 4ppm, which means that is completely up to the precision standard of Orbitrap MS. Consequently, subsequent qualitative and quantitative analysis would not be influenced by quality error. In addition, Because the score of the matching peptide in the spectrogram (representing the reliability of peptide identification) is negatively correlated with the distribution of quality deviation, the higher the score, the smaller the quality deviation.

Database Search

We used Maxquant (v1.6.15.0) (Palomba et al. 2021) to retrieve the original data file and searched the tandem mass spectrometry data according to the *P. grandifloras* protein database which was uploaded

by us. The database has a reverse bait database and a contamination database, which can eliminate the influence of contaminated proteins.

The enzyme digestion mode is set to Trypsin/P. The number of missing bits is set to 4; The minimum length of the peptide was set as 7 amino acid residues; The maximum number of modified peptides was set as 5; The mass error tolerance of primary parent ions of first search and main search are set to 20 ppm and 20 ppm respectively, and the mass error tolerance of secondary fragment ions is set to 20 ppm. Cysteine alkylation Carbamidomethyl was set as fixed modification, and variable modification was methionine oxidation, protein N-terminal acetylation, and lysine malonylation. FDR for protein identification and PSM identification was set to 1%.

Protein Functional Enrichment And Protein Annotation Method

We made analysis of protein annotation and functional enrichment by comprehensive approach, for example, gene ontology, protein domain, KEGG pathway, COG/KOG functional classification and subcellular localization.

GO analysis is a bioinformatics analysis method which can hold every kind information of gene and gene product together (Sun et al. 2022). It integrates information based on eggno database to provide statistical information. What it includes are cell components, molecular functions and biological processes. It not only can produce much information of gene and protein and classify them based on analysis content, but also become a database of the information which we have got.

Protein domain is a kind of structural element which have molecular function so that it can work independently and is conservatism in the order of protein. We made annotation of protein domain which had been found based on Pfam database and PfamScan.

KEGG can integrate known protein interaction network information, such as pathways and related complexes, genes and gene products, biological complexes and related reactions, etc. The annotated contents of our study are protein pathways, such as metabolism, genetic information processing, cellular processes, and environmental information processing. The highest score of blast comparison of modified protein was selected to annotate the results.

In addition, we also used the KOG database to construct the phylogenetic tree and annotated the functions of homologous gene clusters. Moreover, because the proteins in eukaryotic tissue cells can be localized to various elements in cells, Wolf Psort software can be used to annotate the submitted proteins for subcellular localization.

Protein Interaction Network Analysis

After comparing the identified protein sequence or modified protein database number with the STRING protein interaction network database, we got the protein interaction relationship according to the confidence score > 0.7 (high confidence), and then used the R package "network D3" tool to visually display it.

Results And Discussion

Identification of lysine malonylation modified protein in *P. grandiflorus*

In order to ensure that the protein modification data used meet the standards, proteome scale analysis was carried out on the existing data (Fig. 1). The existence of lysine malonylation in *P. grandiflorus* was confirmed by Western blot with antimalonyllysine antibody (Fig. 1A). The overall quality error is less than 5 ppm (Fig. 1B), which indicates that the MS data involved in this study is reliable. And most peptide segments only have been modified once (Fig. 1C). After checking the distribution of all peptides, it was found that most of the peptides were distributed in 7–20 amino acids (Fig. 1D), which was in line with the general law based on enzymatic hydrolysis and mass spectrometry fragmentation, indicating that the length distribution of peptides identified by this mass spectrometry met the quality control requirements. Through statistics, 1755 lysine malonylation modification sites in 888 proteins were identified in this study (Supplementary Table S1). A representative MS/MS diagram is revealed in Fig. 1E. Although lysine malonylation modification has been reported in some other plant species (Nie et al. 2020; Nahomi et al. 2020), similar to pathogens (Shi et al. 2021) there is no relevant record in *P. grandiflorus*. Our results the first report of lysine malonylation in *P. grandiflorus*.

Functional annotation and subcellular localization of lysine malonylation modified proteins in *P. grandiflorus*

In order to further understand the functional role of lysine malonylation modification in *P. grandiflorus*, the identified modified proteins were classified based on the Gene Ontology (GO) annotation according to biological processes, cellular components and molecular functions (Supplementary Table S2). It can be seen from Fig. 2A that the proteins related to cellular processes (24%), metabolic processes (20%) and stimulation reactions (17%) have been modified by malonylation, which is consistent with the view that the modification of lysine malonylation affects the biological metabolic process in recent years (Liu et al. 2018). The data of cell composition (Fig. 2B) shows that most of the modified proteins are located in cells (45%) and cells (44%), and only a few are located in protein containing complexes (11%). Based on the analysis of molecular function, it is obvious that malonylation modification is closely related to the activity of Catalysis (45%) and binding (37%) (Fig. 2C), which is also similar to the results obtained in wheat (Liu et al. 2018). The above data show that malonylation modification can affect a variety of biological metabolic processes of *P. grandiflorus*, and plays an indispensable role in various life activities of plants.

After further subcellular localization (Fig. 2D, Supplementary Table S2) of the modification sites, it was found that not only most of the modification sites existed in chloroplasts (38%) and cytoplasm (34%), but

there are few of in mitochondria (6%), indicating that the metabolic process of photosynthesis and ATP transformation were both affected by malonylation modification. In addition, this modification sites also exists in the nucleus (11%) and extracellular (3%). Considering the above results, it can be speculated that malonylation is regulated at each stage of the process of energy generation, transformation and application. Its specific influence range can be the follow-up key research direction of this modification in *P. grandiflorus*.

Then, the functional classification of the modified proteins is performed (Fig. 2E, Supplementary Table S2). Based on the analysis of homologous clusters (COGs) in the COG database, it can be found that the number of modified proteins related to post-translational modification, protein turnover and chaperone ranks first, with 132, followed by translation, ribosomal structure and biogenesis (92). This proves once again that malonylation modification will affect the gene expression process. At the same time, a large number of modified proteins are predicted to be related to carbohydrate transport, metabolism (73) and energy production and conversion (67).

Functional Enrichment Analysis Of Lysine Malonylation Modified Proteins

In order to further explore the properties of the modified proteins, GO and KEGG enrichment analysis were carried out (Supplementary Table S3). GO enrichment plots (Figure 3A) indicated that there were many modified proteins related to biological processes as well as molecular functions, including cellular processes, stimulus responses, as well as transcription translation, etc. In which catalysis, as well as binding, accounts for the vast majority of the modified proteins (Figure 3A). Combined with its high level of expression inside the cell as well as on the chloroplast (Figure 3B), it is difficult to deny that malonylation is able to catalyze the execution of photosynthesis. To verify the real reliability of the resulting data, it was similarly done enrichment analysis of KEGG pathways. The results showed that malonylated proteins were greatly enriched in carbon fixation, tricarboxylic acid (TCA) cycle as well as glucose metabolism pathways, which illustrated that the modified proteins were closely related to them, sufficient to confirm the above points. While domain enrichment plots (Figure 3C, Supplementary Table S3) suggest that proteins with structures similar to transferases, dehydrogenases as well as pyruvate kinase would be more likely to be modified, resulting in the corresponding malonylated proteins.

Lysine Malonylation Site Pattern Analysis

To understand whether there were specific amino acid motifs in the vicinity of the modified lysine (Supplementary Table S4), it was analyzed for frequency, from - 10 to + 10. As can be seen from the heat map (Fig. 4A), alanine (a), cysteine (c), glycine (g), as well as valine (V) have relatively high frequencies. Residue a distribution is more diffuse, occurring at -1, + 1 as well as + 10, while C, G as well as V are all relatively concentrated between - 4 and + 4. These suggest that the surrounding sequence of lysine residues may influence whether malonylation occurs or not. Six of these motifs were more frequent and

are represented after their columns in a graphic (Fig. 4B), including C...Kma, C...Kma, Kma...C, C...Kma and KmaA, suggesting that malonylation is more likely to occur when these motifs resembled near lysine.

In order to further understand the role of malonylated protein in *P. grandiflorus*, the secondary structure of malonylated protein was analyzed. As can be seen from Fig. 4C, malonylation occurs preferentially in coil (67.8%), followed by α -Helix (26.0%), only about 6.17% of the sites appeared in β -In the spiral. Moreover, the probability of their appearance on the protective surface is small, and the relative surface accessibility can reach 98.2%, so malonylation modification may not change or affect the surface properties of proteins. Thus, the modification of lysine malonylation may be directional rather than random.

Protein interaction network of lysine malonylation modified proteins in *P. grandiflorus*

All the identified associated proteins were characterized using database network (Figure 5), and after the protein interaction network was constructed, it was able to make a clear observation of the relationship between the target protein and other proteins. Based on this interaction network, three high score enriched pathways were picked and the corresponding action network (Figure 5, Supplementary Table S5) including glucose metabolism pathways (Figure 5A), TCA cycle (Figure 5B) and carbon fixation (Figure 5C), according to the Cytoscape algorithm and enrichment analysis of KEGG pathways. These data suggest that multiple pathways may be regulated by lysine malonylation in *P. grandiflorus*. But the current study does not cover this aspect and lacks corresponding biological information, which should be studied more closely and perhaps uncover new key enzymes in known biological metabolic pathways or unprecedented special mechanisms.

Lysine Malonylation Involved In Pathway Analysis Of Carbon Fixation, Tca Cycle And Glucose Metabolism

The Calvin cycle and the tricarboxylic acid cycle belong to chloroplasts and mitochondria respectively. Both of them play an important role in *P. grandiflorus*. Studies have shown that lysine malonylation is closely related to the Calvin cycle (Liu et al. 2018) and the tricarboxylic acid cycle (Qian et al. 2016). According to KEGG pathway enrichment analysis (Figure 3B), carbon fixation, tricarboxylic acid cycle and sugar metabolism pathways are among the pathways with the highest malonylation enrichment. Subcellular localization results (Figure 2C) also showed that malonylated proteins were distributed in chloroplasts and mitochondria. These results indicate that this modification will affect the energy conversion process of plants. With further research, it can be seen (Figure 6) that most enzymes in carbon fixation, tricarboxylic acid cycle and sugar metabolism pathway have more or less malonylation modifications, and these three are interconnected by phosphoenolpyruvate, oxaloacetate and acetyl CoA as bridges. Although the presence of modified proteins in mitochondria in subcellular localization is not in the forefront, it has a high level of enrichment in the TCA cycle, which may indicate that malonylation selectively acts on this cycle in mitochondria, without affecting other parts or having a small impact. The above results indicate that lysine malonylation may regulate the energy conversion process in *P. grandiflorus*, including energy collection, production, fixation and application.

Malonylated Proteins Involved In Stress And Defense Responses

When plants are exposed to unfavorable conditions for their survival, the stress response is called environmental stress response. At present, it is clear that environmental stress can be divided into biological stress and abiotic stress, among which abiotic stress includes heavy metals, ultraviolet light, strong light, drought, salinity, extreme temperature, ozone, etc. (Saijo and Loo 2020). Stress is likely to cause plants to produce reactive oxygen species (ROS) (Waszczak et al. 2018), causing damage to plant organism, resulting in slow growth and yield decline. Moreover, in the natural environment, plants generally face a variety of different environmental stresses (Suzuki et al. 2014). Correspondingly, when plants respond to external stress, they will regulate gene expression and produce a variety of special proteins or secondary metabolites to ensure their own survival (Ramakrishna and Ravishankar 2011).

At present, there is no special report on *P. grandiflorus* in response to environmental stress. In this study, we found there were many malonylated proteins response to environmental stress in this plant. It can be seen from the GO analysis that *P. grandiflorus* has obvious biological reactions to cadmium ions and cytokinins, and has a variety of molecular functions to deal with metal ions and antioxidant (Fig. 7). A large number of heat shock proteins (1.58%) and antioxidant related proteins (2.58%) can be found malonylated in *P. grandiflorus*, including HSP22, HSP70, HSP90, APX1, APX3, and APX6. HSP22, HSP70, and HSP90 are common plant heat shock proteins (Lund et al. 1998; Berka et al. 2022; Li et al. 2020), while APX1, APX3 and APX6 are related proteins against reactive oxygen species in plants (Chen et al. 2021; Maruta et al. 2012; Wang et al. 1999). HSP22 has been proved to be a member of the small heat shock protein superfamily, and it is highly expressed in maize (Lund et al. 1998), pea (Avelange-Macherel et al. 2019), tomato (Banzet et al. 1998) and other plants. In order to ensure the smooth progress of photosynthesis, the active oxygen in chloroplasts needs to be rapidly removed, and the key component involved is ascorbic acid. However, Monodehydroascorbate reductase (Maynard et al. 2020), the key enzyme for ascorbic acid regeneration in *P. grandiflorus*, was not significantly enriched, indicating that malonylation may not have obvious scavenging effect on active oxygen in chloroplasts. Interestingly, 2.23% of the proteins are elongation factor proteins, which belong to the molecular functions involved in metal ion binding. Perhaps elongation factor proteins have some correlation with the heavy metal ion stress response of plants. Not only that, APX1 was found to be related to lead stress response (Jiang et al. 2017), maybe heavy metal ion stress response intersected with other stress responses. The results show that although many proteins of *P. grandiflorum* have close relationship with environmental stress, the specific mechanism and main target types of environmental stress response have not been clearly explained. In addition, *P. grandiflorum* is used as food and medicine, so the direction of biological stress promoting certain secondary metabolites of *P. grandiflorum* can be taken as the subsequent research content.

Comparative analysis of lysine malonylation in *P. grandiflorus* and other plants

As a type of PTMs, malonylation, similar to succinylation, acetylation and other modifications, is not a PTM unique to *P. grandiflorus*. Although there is a wealth of experimental data for malonylation

modification studies in microbes as well as in mammals, studies in plants in this context are still in their infancy. The advent of technologies such as LEMP (Chen et al. 2018), MalPrec (Liu et al. 2020) and LETM (Al-Barakati et al. 2020) has facilitated malonylation modification studies in plants as diverse as *Oryza sativa* (Mujahid et al. 2018), *Triticum aestivum* (Liu et al. 2018) and *Zea mays* (Xu et al. 2021). However, there is no *P. grandiflorus* literature on this aspect reported in the available studies.

As can be seen in the table (Table 1) done according to the related data, 1755 malonylation modification sites (the most) among 888 modified proteins were found in *P. grandiflorus*. Maize, which is not much different from *P. grandiflorus*, 1722 modification sites were found among 810 modified proteins, and the least is wheat, with only 342 modification sites. Whereas there was a strong enrichment of malonylation in photosynthesis in either plant, suggesting that the site and function of malonylation in different plants may be conserved. Regardless, the data obtained for malonylation modifications in general are rather sparse relative to common post protein modifications such as acetylation and phosphorylation, suggesting that this class of modifications still holds great research value and potential.

Table 1
Comparison of the number and recognition sites of malonylation proteins in *P. grandiflorus* with previous plant studies

Species	Tissues	proteins	sites	References
<i>Oryza sativa</i>	Seeds	247	421	(Mujahid et al. 2018)
<i>Triticum aestivum</i>	Whole plants	233	342	(Liu et al. 2018)
<i>Zea mays</i>	Leaves	810	1722	(Xu et al. 2021)
<i>Platycodon grandiflorus</i>	Whole plants	888	1755	

Conclusion

In the present study, processing analysis of whole tissues of *P. grandiflorus* led to the identification of a total of 888 lysine malonylation proteins, which harbor 1755 special modification sites, indicating that lysine malonylation modification is widely present in *P. grandiflorus*. Annotation, subcellular localization, enrichment analysis, and pattern site analysis of the identified modified proteins proved that the modification of lysine malonylation in *P. grandiflorus* involves multiple biological processes as well as metabolic pathways and plays an important role in life activities. The subsequent PPI network analysis and comparison analysis of malonylation in other plants suggested that lysine malonylation had a great research value and was worthy of deeper exploration. Our study expands the understanding of lysine malonylation in medicinal plants, which provides an important resource for subsequent research on lysine malonylation as well as metabolisms in *P. grandiflorus*.

Declarations

AUTHOR CONTRIBUTIONS

Conceived, designed, and implemented the study: Q.-S. Y., S.-H. X.; Statistics analysis: Q.-S. Y., W.-M. J., F. M., S.-T. W., L.-P. Z.; Reagents/materials/analysis tools: Q.-S. Y., W.-M. J., S.-H. X., Z.-P. S., S.-Y. G.; and D.-Y. P.; Drafted the manuscript: Q.-S. Y., S.-W. X., J. L., and S.-H. X.; All authors edited the manuscript and approved the final version.

CONFLICTS OF INTEREST

All authors have no conflicts of interest.

FUNDING

This work was supported by National Natural Science Foundation of China (No. U21A20406), the NSF of Anhui Province (Grant No. 1908085MH268), Key Natural Science Research Projects in Anhui Universities (No. xxxxxxxx, No. KJ2019A0453 and KJ2018A0275), Open Project of Provincial and Ministerial Scientific Research Platform, Fuyang Normal University (No. FSKFKT010D) and Anhui University Collaborative Innovation Project (Grant No. GXXT-2019-043, Grant No. GXXT-2019-049).

ACKNOWLEDGEMENTS

We would like to thank Dr. Zhang Zhang from Jingjie PTM BioLab (Hangzhou) Co. Ltd in China for MS analysis and some bioinformatics of proteome. We also thank Dr. Shuai Liu from the University of Hawaii Cancer Center and Dr. Xiaoxi Meng from the Department of Horticultural Science at University of Minnesota for helping to revise the manuscript.

DATA AVAILABILITY

The supplementary materials including five supplementary tables and five supplementary figures for this article. The mass spectrometry proteomics dataset in this manuscript has been deposited into the ProteomeXchange Consortium via the PRIDE partner repository with the dataset identifier PXD037395 (<http://proteomecentral.proteomexchange.org/cgi/GetDataset?ID=PX037395>).

References

1. Ahmad MZ, Li P, Wang J, Rehman NU, Zhao J (2017) Isoflavone Malonyltransferases GmIMaT1 and GmIMaT3 Differently Modify Isoflavone Glucosides in Soybean (*Glycine max*) under Various Stresses. *Front Plant Sci* 8:735. doi:10.3389/fpls.2017.00735
2. Ahmad MZ, Zhang Y, Zeng X, Li P, Wang X, Benedito VA, Zhao J (2021) Isoflavone malonyl-CoA acyltransferase GmMaT2 is involved in nodulation of soybean by modifying synthesis and secretion of isoflavones. *J Exp Bot* 72 (4):1349–1369. doi:10.1093/jxb/eraa511
3. Al-Barakati H, Thapa N, Hiroto S, Roy K, Newman RH, Kc D (2020) RF-MaloSite and DL-Malosite: Methods based on random forest and deep learning to identify malonylation sites. *Comput Struct Biotechnol J* 18:852–860. doi:10.1016/j.csbj.2020.02.012

4. Avelange-Macherel MH, Rolland A, Hinault MP, Tolleter D, Macherel D (2019) The Mitochondrial Small Heat Shock Protein HSP22 from Pea is a Thermosoluble Chaperone Prone to Co-Precipitate with Unfolding Client Proteins. *Int J Mol Sci* 21 (1). doi:10.3390/ijms21010097
5. Banzet N, Richaud C, Deveaux Y, Kazmaier M, Gagnon J, Triantaphylidès C (1998) Accumulation of small heat shock proteins, including mitochondrial HSP22, induced by oxidative stress and adaptive response in tomato cells. *Plant J* 13 (4):519–527. doi:10.1046/j.1365-313x.1998.00056.x
6. Berghold J, Eichmüller C, Hörtensteiner S, Kräutler B (2004) Chlorophyll breakdown in tobacco on the structure of two nonfluorescent chlorophyll catabolites. *Chem Biodivers* 1 (4):657–668. doi:10.1002/cbdv.200490057
7. Berka M, Kopecka R, Berkova V, Brzobohaty B, Cerny M (2022) Regulation of heat shock proteins 70 and their role in plant immunity. *J Exp Bot* 73 (7):1894–1909. doi:10.1093/jxb/erab549
8. Chen C, Galon Y, Rahmati Ishka M, Malihi S, Shimanovsky V, Twito S, Rath A, Vatamaniuk OK, Miller G (2021) ASCORBATE PEROXIDASE6 delays the onset of age-dependent leaf senescence. *Plant Physiol* 185 (2):441–456. doi:10.1093/plphys/kiaa031
9. Chen Z, He N, Huang Y, Qin WT, Liu X, Li L (2018) Integration of A Deep Learning Classifier with A Random Forest Approach for Predicting Malonylation Sites. *Genomics Proteomics Bioinformatics* 16 (6):451–459. doi:10.1016/j.gpb.2018.08.004
10. Colak G, Pougovkina O, Dai L, Tan M, Te Brinke H, Huang H, Cheng Z, Park J, Wan X, Liu X, Yue WW, Wanders RJ, Locasale JW, Lombard DB, de Boer VC, Zhao Y (2015) Proteomic and Biochemical Studies of Lysine Malonylation Suggest Its Malonic Aciduria-associated Regulatory Role in Mitochondrial Function and Fatty Acid Oxidation. *Mol Cell Proteomics* 14 (11):3056–3071. doi:10.1074/mcp.M115.048850
11. Deng Y, Ren H, Ye X, Xia L, Liu M, Liu Y, Yang M, Yang S, Ye X, Zhang J (2020a) Integrated Phytochemical Analysis Based on UPLC-Q-TOF-MS/MS, Network Pharmacology, and Experiment Verification to Explore the Potential Mechanism of Platycodon grandiflorum for Chronic Bronchitis. *Front Pharmacol* 11:564131. doi:10.3389/fphar.2020.564131
12. Deng Y, Ye X, Chen Y, Ren H, Xia L, Liu Y, Liu M, Liu H, Zhang H, Wang K, Zhang J, Zhang Z (2020b) Chemical Characteristics of Platycodon grandiflorum and its Mechanism in Lung Cancer Treatment. *Front Pharmacol* 11:609825. doi:10.3389/fphar.2020.609825
13. Dong Y, Li P, Li P, Chen C (2021) First comprehensive analysis of lysine succinylation in paper mulberry (*Broussonetia papyrifera*). *BMC Genomics* 22 (1):255. doi:10.1186/s12864-021-07567-5
14. Huang W, Zhou H, Yuan M, Lan L, Hou A, Ji S (2021) Comprehensive characterization of the chemical constituents in Platycodon grandiflorum by an integrated liquid chromatography-mass spectrometry strategy. *J Chromatogr A* 1654:462477. doi:10.1016/j.chroma.2021.462477
15. Jia Y, Chen S, Chen W, Zhang P, Su Z, Zhang L, Xu M, Guo L (2022) A Chromosome-Level Reference Genome of Chinese Balloon Flower (*Platycodon grandiflorus*). *Front Genet* 13:869784. doi:10.3389/fgene.2022.869784

16. Jiang L, Wang W, Chen Z, Gao Q, Xu Q, Cao H (2017) A role for APX1 gene in lead tolerance in *Arabidopsis thaliana*. *Plant Sci* 256:94–102. doi:10.1016/j.plantsci.2016.11.015
17. Li W, Chen Y, Ye M, Wang D, Chen Q (2020) Evolutionary history of the heat shock protein 90 (Hsp90) family of 43 plants and characterization of Hsp90s in *Solanum tuberosum*. *Mol Biol Rep* 47 (9):6679–6691. doi:10.1007/s11033-020-05722-x
18. Lin H, Carroll KS (2018) Introduction: Posttranslational Protein Modification. *Chem Rev* 118 (3):887–888. doi:10.1021/acs.chemrev.7b00756
19. Liu J, Wang G, Lin Q, Liang W, Gao Z, Mu P, Li G, Song L (2018) Systematic analysis of the lysine malonylome in common wheat. *BMC Genomics* 19 (1):209. doi:10.1186/s12864-018-4535-y
20. Liu X, Wang L, Li J, Hu J, Zhang X (2020) Mal-Prec: computational prediction of protein Malonylation sites via machine learning based feature integration: Malonylation site prediction. *BMC Genomics* 21 (1):812. doi:10.1186/s12864-020-07166-w
21. Lund AA, Blum PH, Bhatramakki D, Elthon TE (1998) Heat-stress response of maize mitochondria. *Plant Physiol* 116 (3):1097–1110. doi:10.1104/pp.116.3.1097
22. Manjasetty BA, Yu XH, Panjekar S, Taguchi G, Chance MR, Liu CJ (2012) Structural basis for modification of flavonol and naphthol glucosylconjugates by *Nicotiana tabacum* malonyltransferase (NtMaT1). *Planta* 236 (3):781–793. doi:10.1007/s00425-012-1660-8
23. Maruta T, Inoue T, Noshi M, Tamoi M, Yabuta Y, Yoshimura K, Ishikawa T, Shigeoka S (2012) Cytosolic ascorbate peroxidase 1 protects organelles against oxidative stress by wounding- and jasmonate-induced H₂O₂ in *Arabidopsis* plants. *Biochim Biophys Acta* 1820 (12):1901–1907. doi:10.1016/j.bbagen.2012.08.003
24. Maynard D, Kumar V, Sprii J, Dietz KJ (2020) 12-Oxophytodienoic Acid Reductase 3 (OPR3) Functions as NADPH-Dependent α,β -Ketoalkene Reductase in Detoxification and Monodehydroascorbate Reductase in Redox Homeostasis. *Plant Cell Physiol* 61 (3):584–595. doi:10.1093/pcp/pcz226
25. Mujahid H, Meng X, Xing S, Peng X, Wang C, 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. doi:10.1016/j.jprot.2017.08.021
26. Nahomi RB, Nandi SK, Rakete S, Michel C, Fritz KS, Nagaraj RH (2020) Lysine malonylation and propionylation are prevalent in human lens proteins. *Exp Eye Res* 190:107864. doi:10.1016/j.exer.2019.107864
27. Nie LB, Liang QL, Du R, Elsheikha HM, Han NJ, Li FC, Zhu XQ (2020) Global Proteomic Analysis of Lysine Malonylation in *Toxoplasma gondii*. *Front Microbiol* 11:776. doi:10.3389/fmicb.2020.00776
28. Palomba A, Abbondio M, Fiorito G, Uzzau S, Pagnozzi D, Tanca A (2021) Comparative Evaluation of MaxQuant and Proteome Discoverer MS1-Based Protein Quantification Tools. *J Proteome Res* 20 (7):3497–3507. doi:10.1021/acs.jproteome.1c00143

29. Qian L, Nie L, Chen M, Liu P, Zhu J, Zhai L, Tao SC, Cheng Z, Zhao Y, Tan M (2016) Global Profiling of Protein Lysine Malonylation in *Escherichia coli* Reveals Its Role in Energy Metabolism. *J Proteome Res* 15 (6):2060–2071. doi:10.1021/acs.jproteome.6b00264
30. Ramakrishna A, Ravishankar GA (2011) Influence of abiotic stress signals on secondary metabolites in plants. *Plant Signal Behav* 6 (11):1720–1731. doi:10.4161/psb.6.11.17613
31. Saijo Y, Loo EP (2020) Plant immunity in signal integration between biotic and abiotic stress responses. *New Phytol* 225 (1):87–104. doi:10.1111/nph.15989
32. Shi Y, Zhu J, Xu Y, Tang X, Yang Z, Huang A (2021) Malonyl-proteome profiles of *Staphylococcus aureus* reveal lysine malonylation modification in enzymes involved in energy metabolism. *Proteome Sci* 19 (1):1. doi:10.1186/s12953-020-00169-1
33. Sreedhar A, Wiese EK, Hitosugi T (2020) Enzymatic and metabolic regulation of lysine succinylation. *Genes Dis* 7 (2):166–171. doi:10.1016/j.gendis.2019.09.011
34. Su X, Liu Y, Han L, Wang Z, Cao M, Wu L, Jiang W, Meng F, Guo X, Yu N, Gui S, Xing S, Peng D (2021) A candidate gene identified in converting platycoside E to platycodin D from *Platycodon grandiflorus* by transcriptome and main metabolites analysis. *Sci Rep* 11 (1):9810. doi:10.1038/s41598-021-89294-1
35. Sun T, Wang T, Qiang Y, Zhao G, Yang J, Zhong H, Peng X, Yang J, Li Y (2022) CBL-Interacting Protein Kinase OsCIPK18 Regulates the Response of Ammonium Toxicity in Rice Roots. *Front Plant Sci* 13:863283. doi:10.3389/fpls.2022.863283
36. Suzuki N, Rivero RM, Shulaev V, Blumwald E, Mittler R (2014) Abiotic and biotic stress combinations. *New Phytol* 203 (1):32–43. doi:10.1111/nph.12797
37. Taguchi G, Ubukata T, Nozue H, Kobayashi Y, Takahi M, Yamamoto H, Hayashida N (2010) Malonylation is a key reaction in the metabolism of xenobiotic phenolic glucosides in *Arabidopsis* and tobacco. *Plant J* 63 (6):1031–1041. doi:10.1111/j.1365-313X.2010.04298.x
38. Taherzadeh G, Yang Y, Xu H, Xue Y, Liew AW, Zhou Y (2018) Predicting lysine-malonylation sites of proteins using sequence and predicted structural features. *J Comput Chem* 39 (22):1757–1763. doi:10.1002/jcc.25353
39. Wang J, Zhang H, Allen RD (1999) Overexpression of an *Arabidopsis* peroxisomal ascorbate peroxidase gene in tobacco increases protection against oxidative stress. *Plant Cell Physiol* 40 (7):725–732. doi:10.1093/oxfordjournals.pcp.a029599
40. Wang LN, Shi SP, Xu HD, Wen PP, Qiu JD (2017) Computational prediction of species-specific malonylation sites via enhanced characteristic strategy. *Bioinformatics* 33 (10):1457–1463. doi:10.1093/bioinformatics/xxxxx
41. Wang X, Chen X, Li J, Zhou X, Liu Y, Zhong L, Tang Y, Zheng H, Liu J, Zhan R, Chen L (2019) Global analysis of lysine succinylation in patchouli plant leaves. *Hortic Res* 6:133. doi:10.1038/s41438-019-0216-5
42. WANG Zhi-fen, SU Xue-he, YAN Shu-lin, ZHU Lian-xian, Chun-qing S (2006) Comparison of the Fruit Characteristics of *Platycodon grandiflorum* from Different Habitats in China. *Res Pract Chin Med*

(05):7–9

43. Waszczak C, Carmody M, Kangasjarvi J (2018) Reactive Oxygen Species in Plant Signaling. *Annu Rev Plant Biol* 69:209–236. doi:10.1146/annurev-arplant-042817-040322
44. Xu M, Tian X, Ku T, Wang G, Zhang E (2021) Global identification and systematic analysis of lysine malonylation in Maize (*Zea mays* L.). *Front Plant Sci* 12. doi:10.3389/fpls.2021.728338
45. Yu X-H, Chen M-H, Liu C-J (2008) Nucleocytoplasmic-localized acyltransferases catalyze the malonylation of 7-O-glycosidic (iso)flavones in *Medicago truncatula*. *The Plant Journal* 55 (3):382–396. doi:10.1111/j.1365-313X.2008.03509.x
46. Zhang Y, Xie R, Wang J, Leier A, Marquez-Lago TT, Akutsu T, Webb GI, Chou KC, Song J (2019) Computational analysis and prediction of lysine malonylation sites by exploiting informative features in an integrative machine-learning framework. *Brief Bioinform* 20 (6):2185–2199. doi:10.1093/bib/bby079

Figures

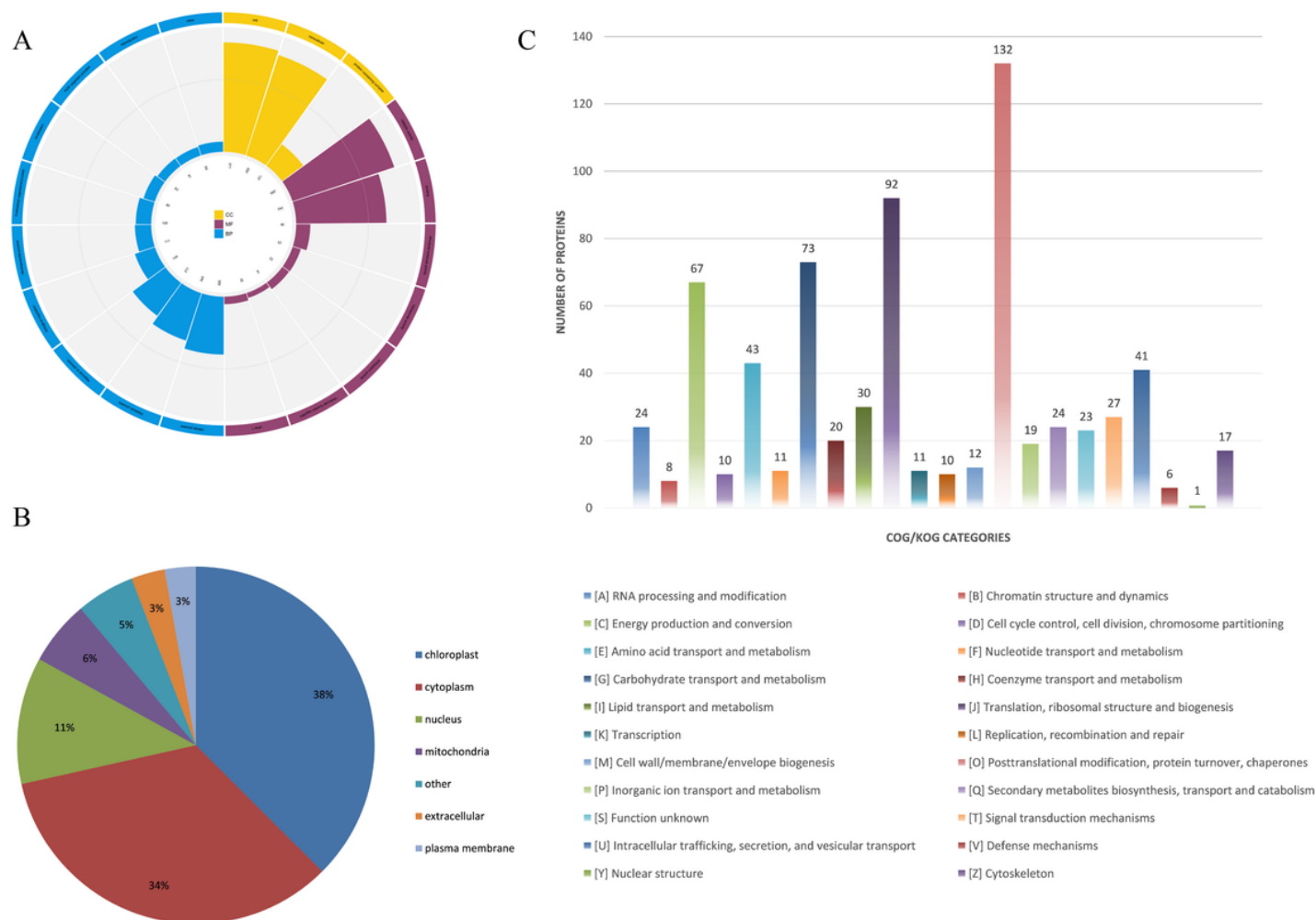


Figure 2

Classification of proteins in *P. grandiflorus* corresponding to identified sites. **(A)** Classification of malonylated proteins according to biological. **(B)** Classification of malonylated proteins according to cellular component. **(C)** Classification of malonylated proteins according to molecular. **(D)** Subcellular localization chart of proteins malonyting to modification sites. **(E)** COG (COK) functional classification of malonylated proteins

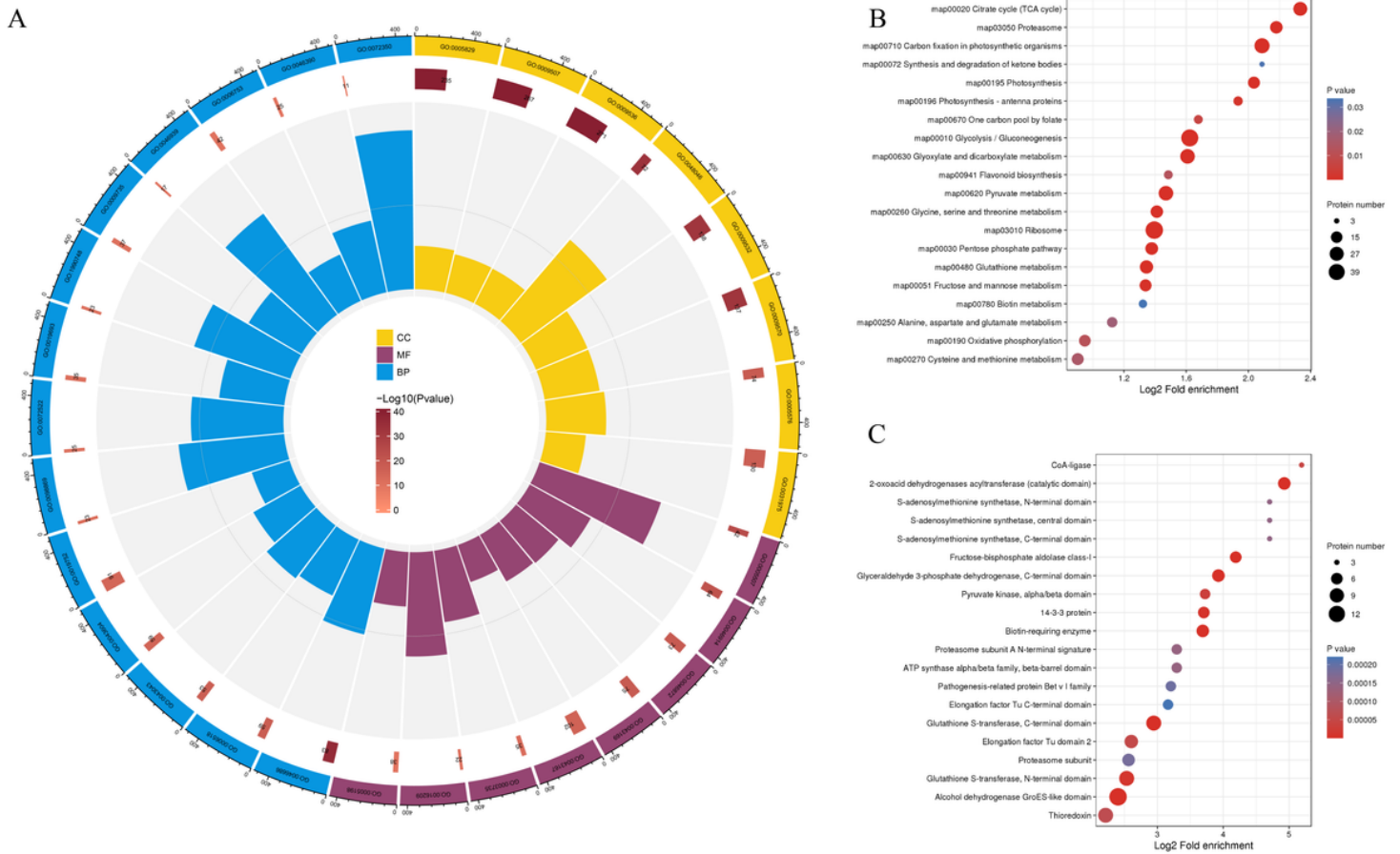


Figure 3

Go classification, KEGG pathway enrichment, domain enrichment of lysine malonylation in *P. grandiflorus*. **(A)** The GO enrichment plot of proteins corresponding. **(B)** The KEGG pathway enrichment of proteins malonylation. **(C)** Protein domain enrichment bubble plot.

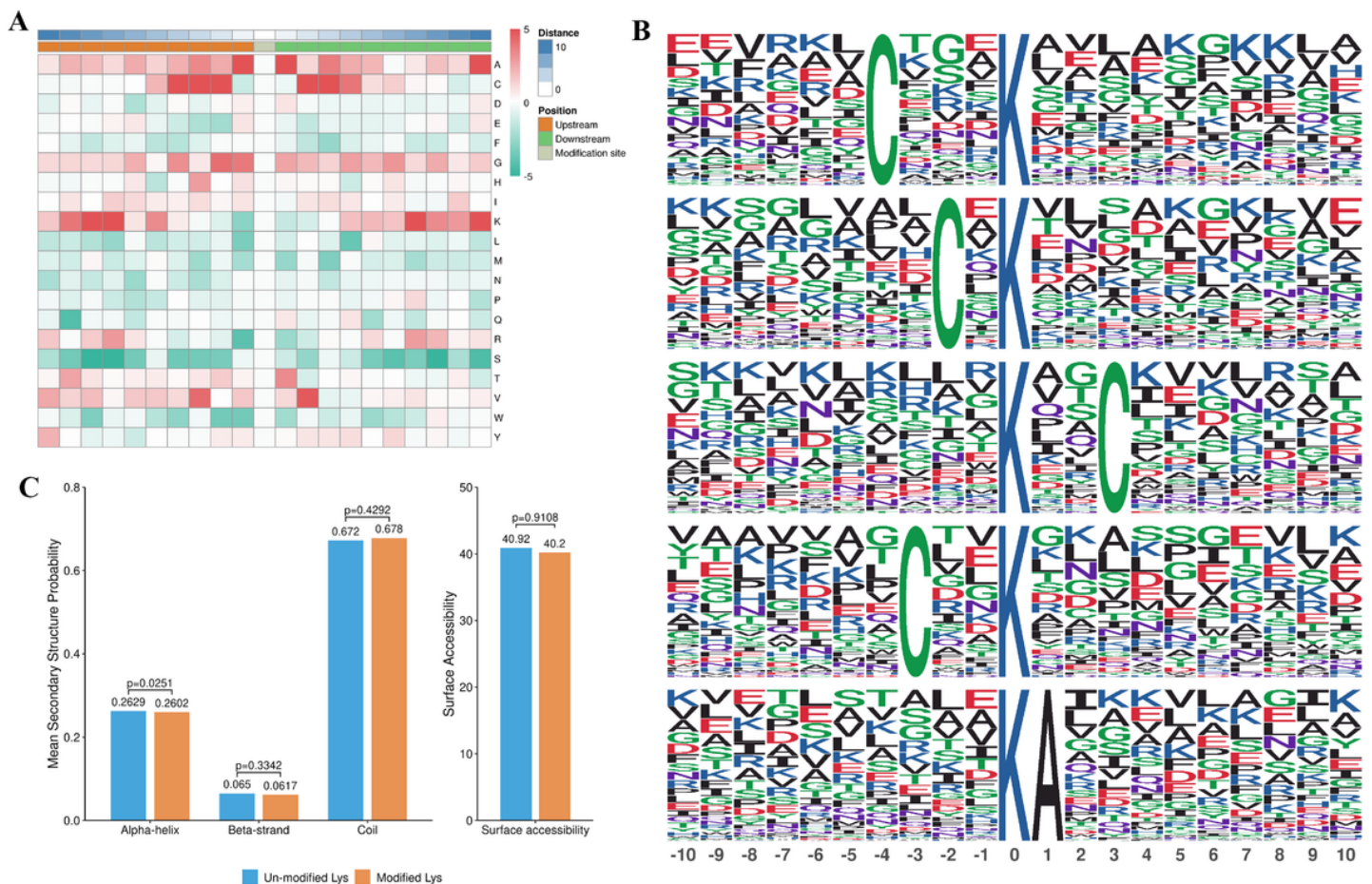


Figure 4

Properties of the lysine malonylation sites in *P. grandiflorus*. **(A)** Heat map of the amino acid compositions around the Kcr sites showing the frequency of different types. **(B)** Sequence probability logos of the five enriched malonylation site motifs around the lysine malonylation sites. **(C)** Probabilities of lysine malonylation in different protein secondary structures (alpha helix, beta-strand

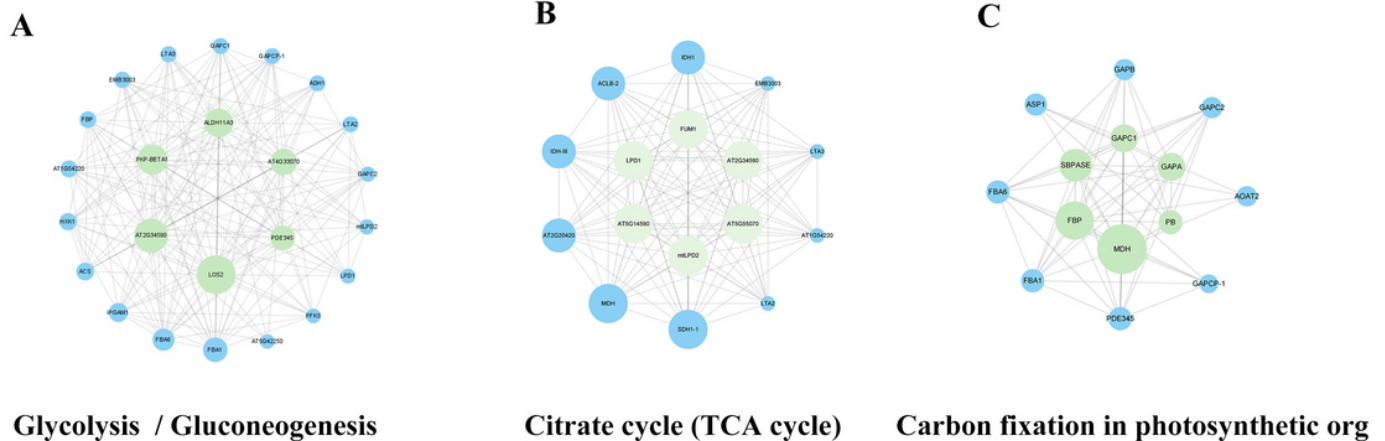


Figure 5

The three KEGG pathways with the highest score in the protein–protein interaction (PPI) networks of Kma proteins in *P. grandiflorus*. **(A)** Glycolysis/Gluconeogenesis. **(B)** Citrate cycle (TCA cycle). **(C)** Carbon fixation in photosynthesis.

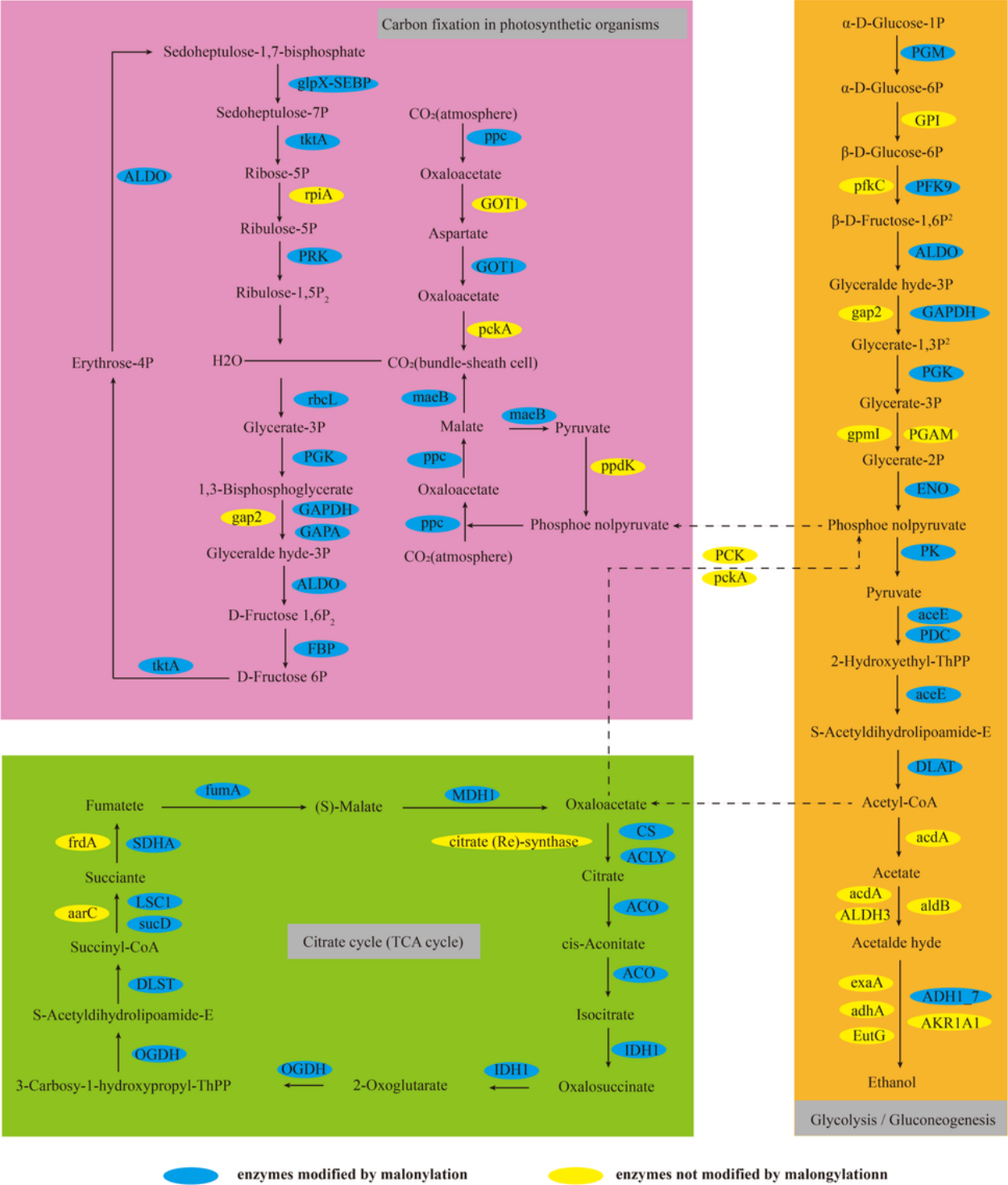


Figure 6

Malonylated enzymes were involved in Citrate cycle (TCA cycle), Glycolysis gluconeogenesis, and carbon fixation in photosynthetic organisms of *P. grandiflorus*. Yellow oval referred to enzymes not modified by malonylation. Blue oval referred to enzymes modified by malonylation.

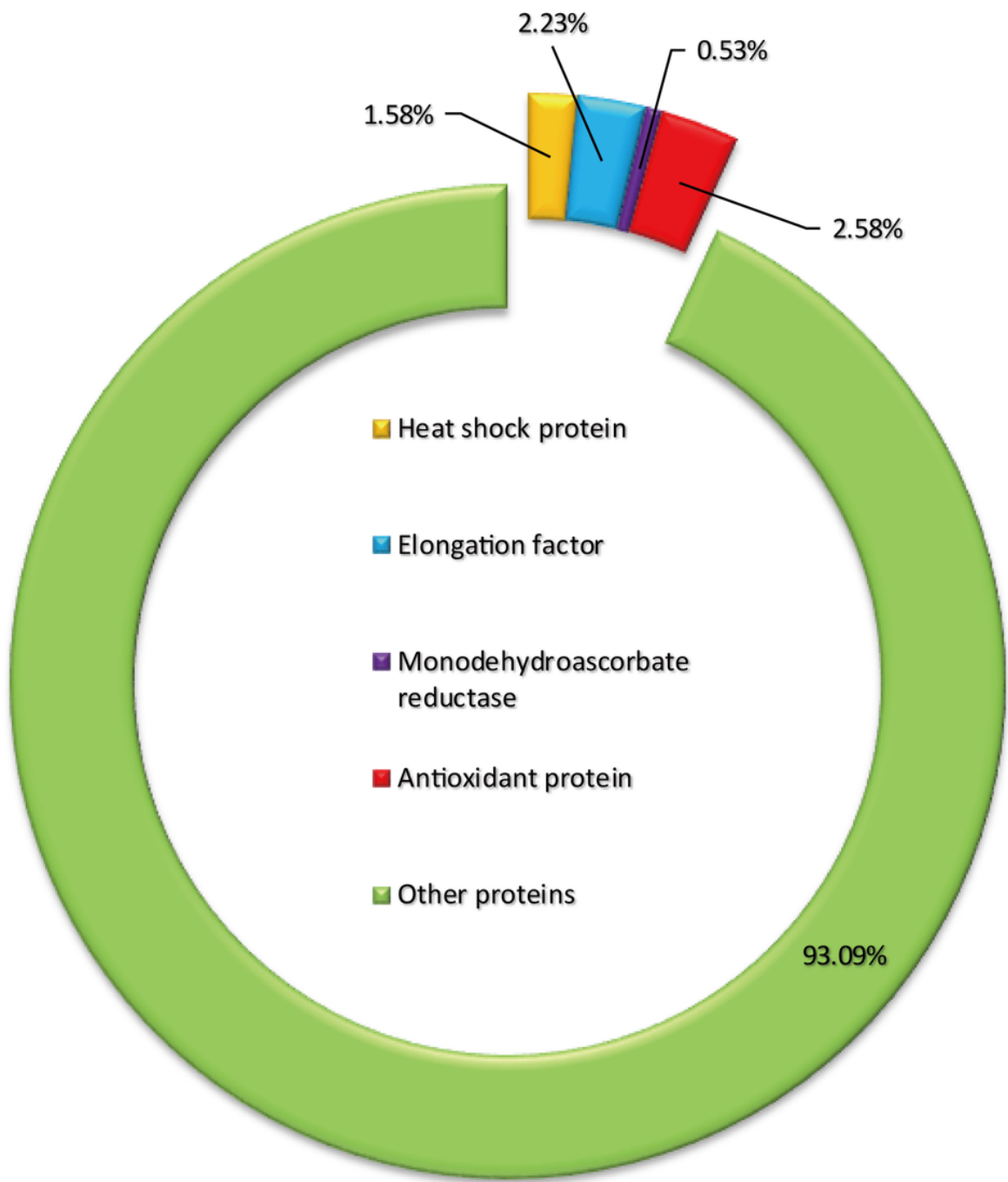


Figure 7

The percentage of types of proteins malonylated in *P. grandifloras*

Supplementary Files

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

- [TableS1MSidentifiedinformation.xlsx](#)
- [TableS2Functionalclassification.xlsx](#)
- [TableS3Functionalenrichment.xlsx](#)
- [TableS4Motifanalysis.xlsx](#)
- [TableS5theProteinProteinInteractionNetworks.xlsx](#)