



Research Article

Spatial protein expression of *Panax ginseng* by in-depth proteomic analysis for ginsenoside biosynthesis and transportation

Xiaoying Li^{1,☆}, Xianhui Cheng^{1,☆}, Baosheng Liao², Jiang Xu², Xu Han³, Jinbo Zhang⁴, Zhiwei Lin⁴, Lianghai Hu^{1,*}

¹ Key Laboratory for Molecular Enzymology and Engineering, The Ministry of Education, National Engineering Laboratory of AIDS Vaccine, School of Life Sciences, Jilin University, Changchun, China

² Institute of Chinese Materia Medica, China Academy of Chinese Medical Sciences, Beijing, China

³ Jilin Zixin Pharmaceutical Industrial Co., Ltd, China

⁴ Nanjing Novogene Bio Technology Co., Ltd, China

ARTICLE INFO

Article history:

Received 22 May 2019

Accepted 21 January 2020

Available online 6 April 2020

Keywords:

Cytochromes P450

Mass spectrometry

Panax ginseng

Proteomics

UDP-glycosyltransferase

ABSTRACT

Background: *Panax ginseng*, as one of the most widely used herbal medicines worldwide, has been studied comprehensively in terms of the chemical components and pharmacology. The proteins from ginseng are also of great importance for both nutrition value and the mechanism of secondary metabolites. However, the proteomic studies are less reported in the absence of the genome information. With the completion of ginseng genome sequencing, the proteome profiling has become available for the functional study of ginseng protein components.

Methods: We optimized the protein extraction process systematically by using SDS-PAGE and one-dimensional liquid chromatography mass spectrometry. The extracted proteins were then analyzed by two-dimensional chromatography separation and cutting-edge mass spectrometry technique.

Results: A total of 2,732 and 3,608 proteins were identified from ginseng root and cauline leaf, respectively, which was the largest data set reported so far. Only around 50% protein overlapped between the cauline leaf and root tissue parts because of the function assignment for plant growing. Further gene ontology and KEGG pathway revealed the distinguish difference between ginseng root and leaf, which accounts for the photosynthesis and metabolic process. With in-deep analysis of functional proteins related to ginsenoside synthesis, we interestingly found the cytochrome P450 and UDP-glycosyltransferase expression extensively in cauline leaf but not in the root, indicating that the post glucoside synthesis of ginsenosides might be carried out when growing and then transported to the root at withering.

Conclusion: The systematically proteome analysis of *Panax ginseng* will provide us comprehensive understanding of ginsenoside synthesis and guidance for artificial cultivation.

© 2020 The Korean Society of Ginseng. Publishing services by Elsevier B.V. This is an open access article under the CC BY-NC-ND license (<http://creativecommons.org/licenses/by-nc-nd/4.0/>).

1. Introduction

Ginseng is one of the oldest and most important medicinal plants in the Orient. It is a perennial herb of the family *Araliaceae*, which has the reputation of "king of hundred herbs." Ginseng has remarkable and powerful medicinal effects, which can relieve fatigue, regulate blood pressure and endocrine, as well as protect the heart. It can also be used to restore vitality, regulate the body and prolong life for the weak patients [1–5]. Ginseng usually takes 4–6

years to mature, and the requirements for the soil, temperature, precipitation, and illumination of the growing environment are high, which further reflected their preciousness and rarity. In recent years, unreasonable large-scale excavation and planting had caused serious damage to ginseng resources and environmental resources. Therefore, scientific cultivation of ginseng and research on its active ingredients brook no delay. The root is the official essence and the most important part of ginseng. After extensive

* Corresponding author. School of Life Sciences, Jilin University, Changchun, 130023, China

E-mail address: lianghaiu@jlu.edu.cn (L. Hu).

☆ These authors contributed equally to this work.

analysis, ginsenosides were considered to be the main pharmacological ingredients of ginseng to reflect its biological activity [6,7].

In recent years, more and more research groups have put the focus of ginseng research on the level of proteomics, and the related research methods and the establishment of protein databases have been gradually improved. The proteomes of different varieties of ginseng have been analyzed comparatively, the identification of stress-responsive proteins and the researches on proteins related to ginsenoside biosynthesis have been also carried out [8–10]. Although there are numbers of reports on the protein component of ginseng, the systematical study of ginseng proteome has been limited because of the uncompleted genome resources [11]. Recently, the Chen and Yang teams have released the genome sequences individually, which greatly push the understanding of the biosynthetic pathway of ginsenosides [12,13]. The completion of whole genome sequencing provided the possibility for proteomic analysis. Protein is the direct executor of life activities and the ultimate embodiment of complexity. In-depth studies on physiological and metabolic pathways of ginseng are inseparable from protein-level analysis. Proteomics can be used to scientifically isolate and identify the proteins in different parts of ginseng tissues, so as to study the role of abundant proteins of ginseng in metabolic activities according to their different functions, explore the markers and related metabolic pathways involved in ginsenoside biosynthesis at the protein level, to provide new thought for ginseng cultivation and ginsenoside biosynthesis.

In this study, proteomic methods were used to identify and analyze the proteins in ginseng cauline leaf and root as shown in the workflow (Fig. 1). Firstly, the extraction methods of protein for different tissue part were optimized by different extraction buffers to increase the extraction efficiency and to eliminate the non-protein contaminants [14]. Secondly, protein samples were digested and fractionated by high-pH reverse chromatography to reduce the complexity [15–18]. Thirdly, high-resolution mass spectrometry (MS) and latest database were used for protein identification. Finally, the biological functions and metabolic pathways of proteins in cauline leaf and root were classified and compared. We have preliminarily selected UDP-glycosyltransferase (UGT), cytochromes P450 (CYP), and related proteins involved in ginsenoside

biosynthesis to better understand the related biological functions of cauline leaf and root in the process of ginsenoside biosynthesis.

2. Materials and methods

2.1. Protein extraction

The 5-yr-old ginseng samples were from ginseng culture base (Tonghua, Jilin Province, China) of Jilin Zixin Pharmaceutical Industrial Co., Ltd. The root and cauline leaf from 30 strains of ginseng were pooled together for the following experiments. Three biological replicates were used for each extracted protein condition. We designed and compared different extraction methods for ginseng cauline leaf and root samples as described in Table S1. The extract was treated with a sonicator at room temperature for 60 min, stirred at 4°C for 15 h, and centrifuged to remove the precipitation. For methanol–chloroform precipitation, the crude protein extract was sequentially mixed with four volumes of methanol, four volumes of chloroform, and three volumes of ddH₂O. The suspension was centrifuged at 16,000g for 3 min. The precipitate was washed with four volumes of methanol and then reconstituted with distilled water to obtain a total protein extract. For acetone precipitation, the crude ginseng protein extract was added to the precooled acetone at a volume ratio of 1:5, left at –20°C overnight, and centrifuged at 5,000r/min for 30 min. After discarding the supernatant and evaporating the remaining acetone, it was reconstituted with distilled water to obtain a total protein extract. Each protein sample was loaded at the equivalent amounts and separated on sodium dodecyl sulfate polyacrylamide gel electrophoresis (SDS-PAGE) gel for comparison.

2.2. Protein digestion and fractionation

In solution, digestion was performed as following: 1 mg of above ginseng protein samples were added, 200 µL denaturant (100mM NH₄HCO₃, 8M urea) and 10 µL DTT (1 mg/mL) shaking 2 h at 37°C to disrupt the disulfide bonds. Then, 1.4 mg iodoacetamide (IAA) were added in samples for 40 min at room temperature in the absence of light for alkylation, followed by addition of seven volumes of

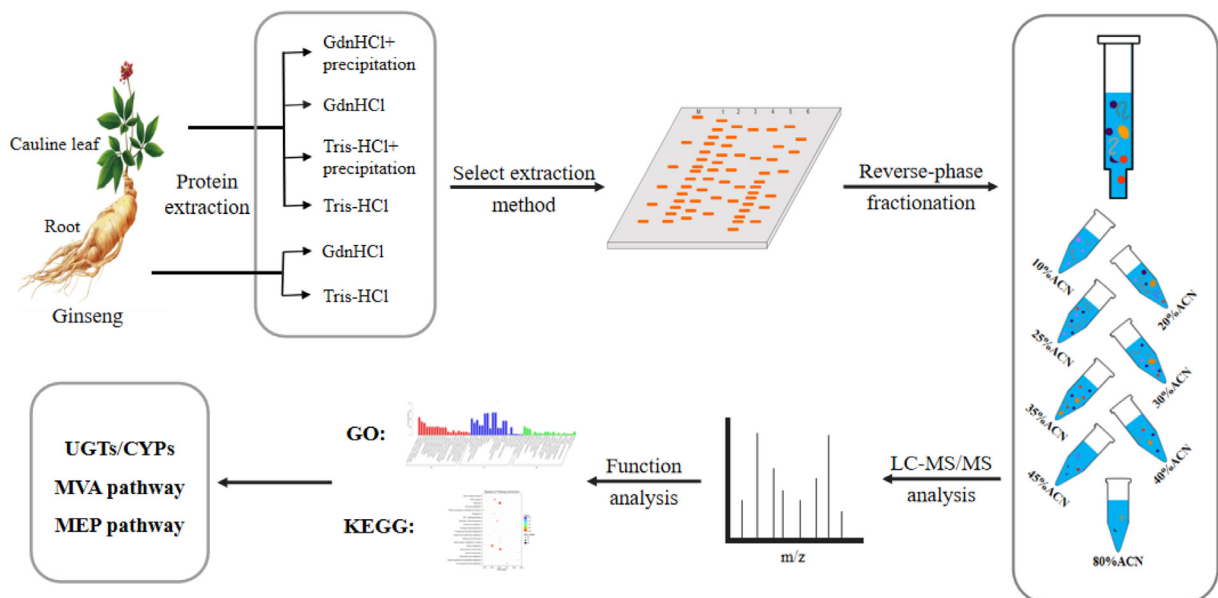


Fig. 1. Workflow illustration for ginseng protein extraction and analysis. UGTs, UDP-glycosyltransferase; CYPs, cytochromes P450; MVA, mevalonate; MEP, methyl-D-ERYTHRITOL-4-PHOSPHORYLATION; LC-MS/MS, LIQUID CHROMATOGRAPHY–TANDEM MASS SPECTROMETRY.

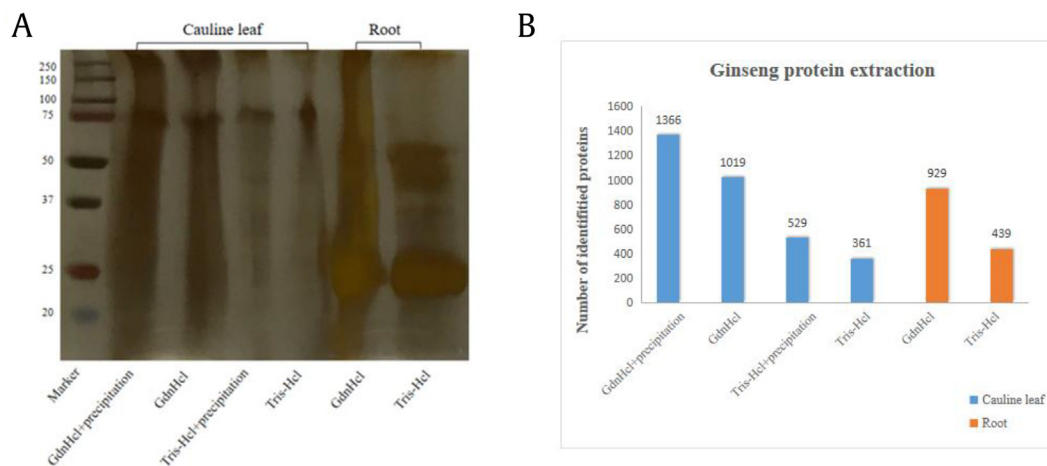


Fig. 2. Proteomic analysis results of ginseng cauline leaf and root protein samples prepared using different extraction methods. (A) SDS-PAGE separation results. (B) Number of proteins identified by MS.

100mM NH_4HCO_3 to dilute the denaturant. Then 20 μg of trypsin was added to at 1:50 (w/w) enzyme-to-protein ratio to digest the protein samples at 37°C overnight. The digested peptides were desalted using Sep-Pak C18 column and dried by speed vacuum for analysis.

For two-dimensional separation, the digested samples were loaded into the equilibrated Sep-Pak C18 column, and eight fractions were subsequently eluted with 200 μL 0.05% triethanolamine (TEA) buffer solution containing 10%, 20%, 25%, 30%, 35%, 40%, 45%, and 80% acetonitrile (ACN), respectively. Each fractionated sample was dried completely using speed vacuum and stored at -80°C until analysis.

2.3. LC-MS/MS analysis

The dried peptides were resolubilized in 0.1% (v/v) formic acid with 2% (v/v) acetonitrile and injected into nano-liquid chromatography system (UltiMate 3000 RSLCnano System; Thermo Scientific, Waltham, MA, US). Samples were separated on a 45-cm in-house packed column containing C18 resin (2.2 μm) by a constant flow rate of 0.35 $\mu\text{L}/\text{min}$ with column oven temperature of 60°C. The mobile phase buffer consisted of 0.1% formic acid in ultrapure water (buffer A) with an eluting buffer of 0.1% formic acid in 80% (v/v) acetonitrile (buffer B). The separating gradient ramped linearly from 1% buffer B to 34% buffer B for 63 min, then 34% to 47% for 10 min, followed by 47% to 90% for 5 min, and finally 90% buffer B for 12 min. The UltiMate 3000 RSLCnano was coupled online with a hybrid higher solution Orbitrap Fusion Lumos Tribrid Mass Spectrometer (Thermo Fisher Scientific). Full scan mass spectra were acquired in the Orbitrap cell in the mass range 300–1,500 m/z (positive polarity, with the resolution of 240,000, automatic gain control (AGC) = 4×10^5), followed by MS/MS of the 10 most intense ions.

2.4. Data processing

The MS data were searched against ginseng database [13] (<https://www.ncbi.nlm.nih.gov/bioproject/PRJNA385956/>) using the Proteome Discoverer 2.2 software (Thermo Fisher Scientific). Enzyme was set with trypsin with three missed cleavages. Precursor mass tolerance was set at 10 ppm, and fragment mass tolerance was set at 0.6 Da. Search criteria included a oxidation (+15.9949 Da) of methionine residues in dynamic modifications, carbamidomethylation of cysteine (+57.0214 Da) in static modifications, and

acetylation (+42.011 Da) at N-terminus of proteins in dynamic modifications (protein terminus). The false discovery rate was set at 0.01.

3. Results and discussion

3.1. Comparative analysis of the ginseng root and cauline leaf proteome

Several extraction approaches were compared to get the systematical analysis of different tissue proteome components. Optimizing the preparation of ginseng samples is critical to reducing impurities, increasing detection sensitivity, and expanding ginseng proteome coverage. For example, cell wall rupture effects and cell membrane lysate selection directly affect the efficiency of plant protein extraction, and the method of removing interfering substances directly affects the extent of protein purification. GdnHCl lysis buffer is reported to be an effective reagent for decomposing cell walls and extracting proteins from cells, and provides optimal yield for protein identification, because GdnHCl can eliminate interference from nucleic acids, phospholipids, light and pigments, and secondary metabolites, allowing good separation of proteins from interfering substances, thereby increasing the number of identified proteins [14]. To explore the efficiency of ginseng root and cauline leaf lysis strategies, we tested two commonly used lysis buffers GdnHCl and tris-HCl as shown in Table S1. Compared with root, cauline leaf protein detection is greatly disturbed by factors such as chlorophyll and secondary metabolites. Therefore, methanol-chloroform and acetone precipitation steps were further performed to purify the protein components. SDS-PAGE analysis on the protein extracts showed that the extraction efficiency of GdnHCl is significantly higher than that of tris-HCl (Fig. 2A). And for the cauline leaf, the combination of GdnHCl with methanol-chloroform precipitation got the most comprehensive extraction efficiency.

Each extract was then analyzed by nano-liquid chromatography tandem MS for protein identification. The results were consistent with the SDS-PAGE results as shown in Fig. 2B. For cauline leaf sample, 1,366 and 1,019 proteins were identified with GdnHCl extraction followed by methanol-chloroform precipitation, whereas only 529 and 361 proteins were identified from tris-HCl protocol with or without precipitation, respectively. For the extraction of ginseng root protein, 929 proteins were identified from GdnHCl protocol, which is over twice of that of tris-HCl lysis

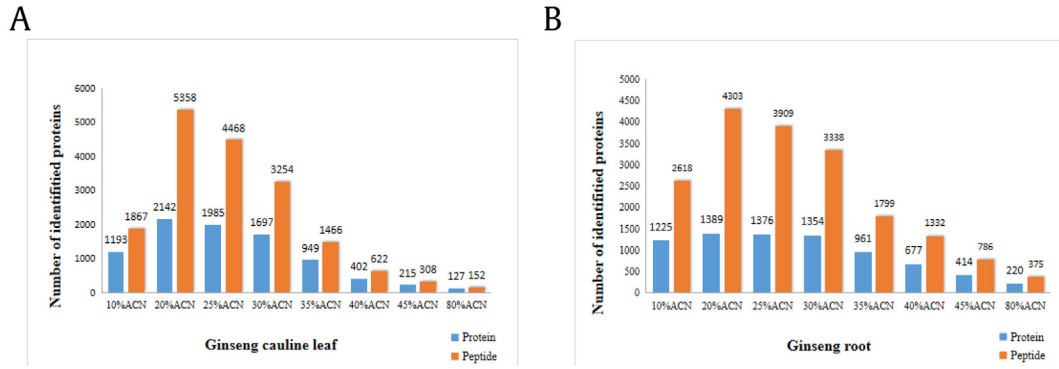


Fig. 3. The numbers of protein and peptide identified from each fraction by two-dimensional LC-MS/MS proteomic analysis of (A) cauline leaf and (B) root of ginseng samples.

method (439 proteins). Although the cauline leaf has much more small molecules such as chlorophyll, the total protein identification is significantly higher than root, which is the main ingredient for herbal medicines. The SDS-PAGE results also showed very clearly that the cauline leaf proteins existed in the whole range of molecular weight, whereas the root proteins dominated around the 25 kDa. Therefore, we summarized the proteomic results of different region of ginseng to find out the highest abundant protein expressions.

The top 20 high-abundance proteins in MS identification results were listed in (Tables S2 and S3). To our surprise that, five of the top 20 proteins from root are ribonuclease-like storage proteins, which exactly corresponded to the 25 kDa band for SDS-PAGE analysis. It was reported that the ginseng root 28 kDa protein was composed of four abundant proteins with molecular weights of 28, 26, 21, and 20 kDa, also known as ginseng major proteins. Ginseng major proteins are also ribonuclease-like storage proteins whose main biological function is to store nutrients for the main root [19]. The root of ginseng stores energy to meet the needs of the fast-growing season (1–3 yrs), and the slow growth period (>5 yrs) of secondary metabolite synthesis and stress resistance because of harsh environments consumes a lot of energy. The rest of the high-abundance proteins in ginseng root are mainly involved in energy metabolism and stress response. In the related process of energy metabolism, mainly included glycolysis (enolase, fructose-diphosphate aldolase 6, glyceraldehyde-3-phosphate dehydrogenase), carbohydrate

chemical reaction (β -amylase, α -1,4-glucan phosphorylase L-1 isozyme, 1,4- α -glucan branching enzyme), amino acid metabolism (adenosine homocysteine, 5-methyltetrahydropropionylglutamate-homocysteine methyltransferase), proton transport (ATP synthase subunit β), and so on. The heat shock homologous 70 kDa protein is a cytoprotective protein and these types of protein mainly act on the stress response of ginseng [20].

For the high-abundance proteins in cauline leaf, identification results showed that four of the top 20 proteins are ATP synthase subunits, which generated an electrochemical gradient through the transport of protons across a membrane that powers ATP synthesis. Another four proteins are subunits of glyceraldehyde-3-phosphate dehydrogenase, which mainly participates in the chemical reactions and pathways involving glucose with the concomitant removal or addition of electrons to or from substance [21]. Three of them are ribulose biphosphate carboxylase family, which is the product controlled by both nuclear and chloroplast genes with main function for carbon dioxide fixation [22]. Others are catalase and heat shock cognate 70 kDa protein for stress response [23] and ribonuclease-like storage protein to store nutrients. Most high-abundance proteins play an important role in energy metabolism-related processes. To systematically analyze the protein functions, we further performed two-dimensional chromatography separation for comprehensive identification of the protein components.

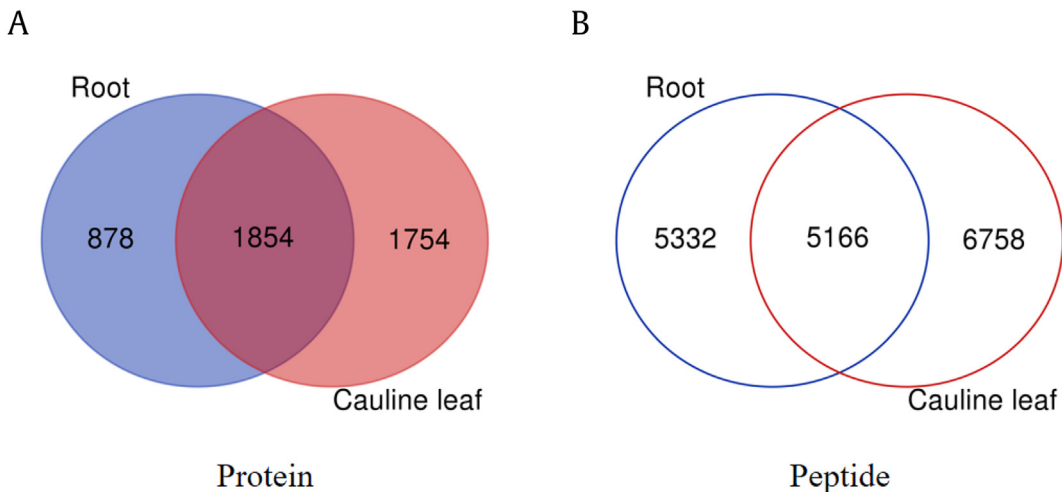


Fig. 4. The Venn diagram of (A) protein and (B) peptide identified from ginseng cauline leaf and root.

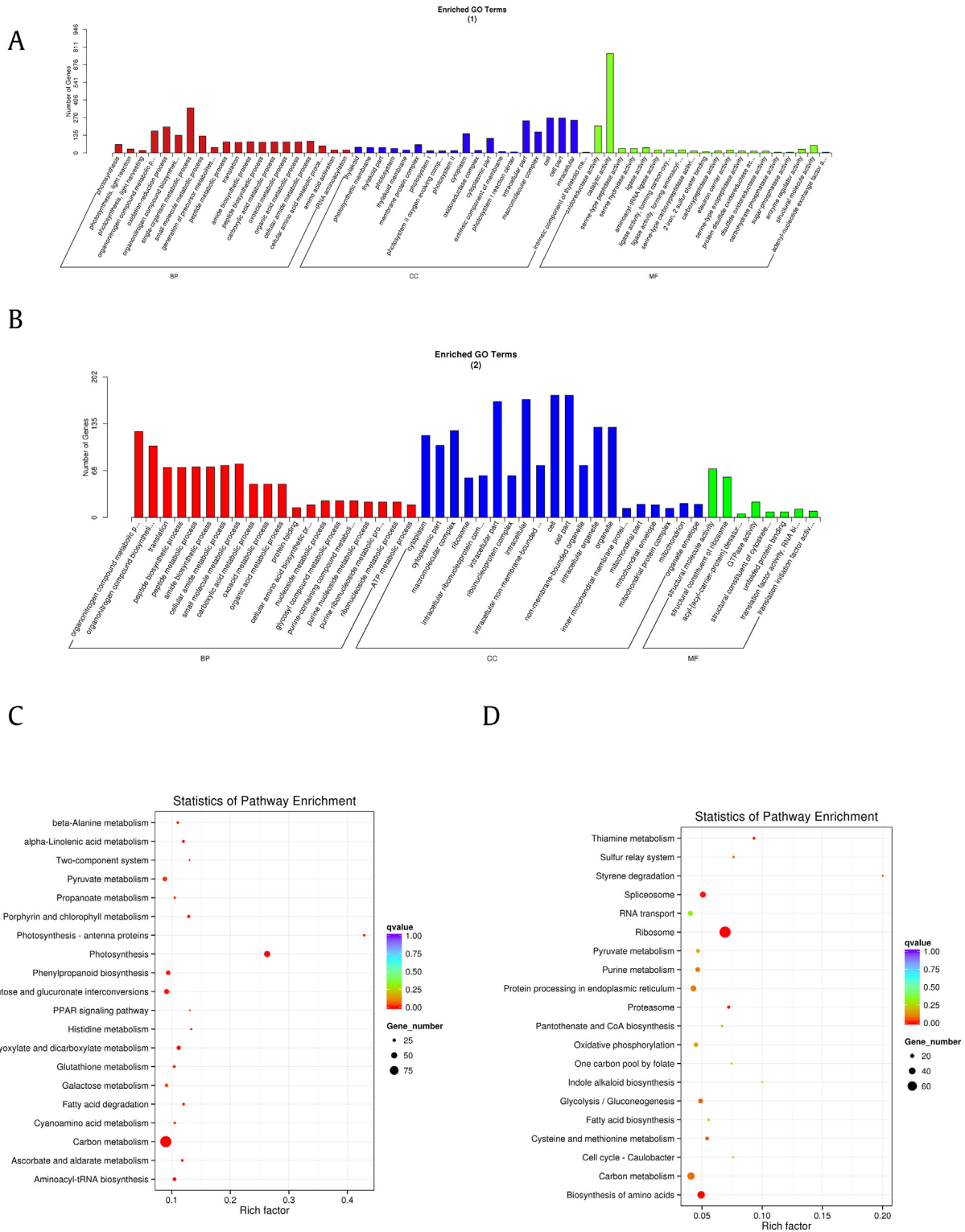


Fig. 5. GO annotation of differentially expressed proteins in (A) cauline leaf and (B) root. KEGG enrichment of differentially expressed proteins in (C) cauline leaf and (D) root. GO, gene ontology.

3.2. In-depth analysis of the ginseng proteome

Currently, there are several approaches to improve the protein identification, such as prefraction on protein stage using SDS-PAGE or peptide level separation by different type of chromatography

[15]. Recently, high-pH reversed-phase based fractionation method was explored as the first dimensional separation because of the high orthogonality with normally used low pH liquid chromatography, which is more comparable with the ionization process of MS [16–18]. The protein samples of root and cauline leaf were

fractionated by high-pH C18 cartridge into eight fractions, and then each fraction was submitted for MS analysis. The number of protein and peptide identified in each fraction was shown in Fig. 3. It was found that the elution trend of protein in the samples of cauline leaf and root were basically consistent: most proteins were identified with the gradient from 20% to 30% ACN. After combination of all the proteins identified from eight fractions, totally 12,552 peptides corresponding to 3,608 proteins and 11,290 peptides corresponding to 2,732 proteins were identified in the cauline leaf and root protein samples, respectively. However, only around half peptides (5,166) overlapped between ginseng root and cauline leaf as shown in the Venn diagram of Fig. 4, which indicated significant difference between the tissue parts of ginseng. To explore the function difference between root and cauline leaf, gene ontology (GO) annotation and KEGG metabolic pathway analyses were carried out on the identified proteins.

The KEGG enrichment and GO analysis of biological process (BP), cellular components, and molecular function were shown in Fig. 5. In the function analysis results of overlapped proteins between root and cauline leaf, we found that the GO results were mainly covered 20 BPs, which were basically involved in the biosynthesis and metabolism of organic compounds such as nucleosides, peptides, proteins, amides, and oxo acids, in which the highest protein abundance was the single-organism metabolic process, followed by organonitrogen compound and small molecule metabolic process. The molecular functions of common proteins mainly included structural molecule activity, cofactor and RNA binding, and as an enzyme to catalyze various enzymatic reactions such as oxidation–reduction reaction. These proteins are mainly present as cells, intracellular and cytoplasm parts, and include many macromolecular complexes (Fig. S1A). KEGG results showed that the pathway with the most common protein participation was ribosome and carbon metabolism, followed by biosynthesis and metabolism of amino acids, which is consistent with GO analysis results (Fig. S1B). These consensus proteins maintain the basic biological functions and metabolic activities of ginseng, provide the necessary substances for each tissue and organ to allow the functionalized reactions to proceed smoothly.

Based on the GO function annotation and KEGG metabolic pathway analysis of specific proteins in cauline leaf and root (Fig. 5), we found that these specific proteins in different parts were involved in some of the same basic BPs to maintain the normal state of cells and tissues, such as biosynthesis and metabolism of organonitrogen compounds, metabolic processes of small molecules, organic acids, peptides and cellular amino acids and the biosynthetic process of peptides and amides. In addition, we found that the proteins unique to cauline leaf mainly specifically possess the BPs related to photosynthesis, such as the photoreaction phase and the light capture process. Quite a few cellular components in the annotation also belong to photosynthesis, like thylakoid and light reaction center. For example, the ribulose-1,5-bisphosphate carboxylase/oxygenase (Rubisco) was present in the chloroplast matrix, which is not only an important carboxylase in CO₂ fixation, but also an oxygenase that catalyzes the reaction with O₂ in photorespiration [22]. In terms of biological function, these proteins were mainly used as an enzyme to catalyze various redox reactions to maintain normal metabolism of the cauline leaf part and normal photosynthesis (Fig. 5A). The specific proteins in the root mainly involved in various nucleoside and ATP metabolic processes, and acted on the modification processes such as protein folding, mainly existed as the structural molecule in cells, and the cell components were also dominated by cell, intracellular, organelle, and cytoplasm part (Fig. 5B).

For the KEGG enrichment analysis of ginseng cauline leaf (Fig. 5C), we found that the most involved pathway was carbon

metabolism, next were pentose and glucuronate interconversions and pyruvate metabolism, followed by the biosynthesis of aminoacyl–tRNA and phenylpropanoid, metabolism of galactose, glutathione, glyoxylate, and dicarboxylate, etc. Moreover, the pathways to photosynthesis and metabolism of chlorophyll got the highest enrichment as expected. Therefore, we supposed that the most important metabolic pathway of specific proteins in cauline leaf was the metabolism of various organic substances and the transformation and fixation of energy around photosynthesis. In the KEGG pathway analysis of root specific proteins (Fig. 5D), the most abundant of protein participation is ribosome, which mainly store nutrients and corresponded to the high abundant ribonuclease-like storage protein in root. The second one is carbon metabolism, followed by spliceosome, biosynthesis of amino acids and protein processing in endoplasmic reticulum, which are indispensable metabolic processes in the biosynthesis of peptides and proteins. In addition, root proteins also act on pathways such as proteasome, oxidative phosphorylation, glycolysis/gluconeogenesis, RNA transport, thiamine metabolism, and biosynthesis of some active substances, etc. These KEGG enrichment results reflected some of the major metabolic processes in the different parts of ginseng. Overall, the specific proteins in cauline leaf mainly acted on photosynthesis and related energy conversion, whereas the root had more kinds of proteins, had functions on the biosynthesis and modification of biomacromolecules, and provided energy for these processes in addition. These results were consistent with the functional analysis of high-abundance proteins in one-dimensional analysis.

3.3. Analysis of related proteins and pathways of ginsenoside biosynthesis

Ginsenosides are considered as the main bioactive components for *Panax ginseng* and the quality is highly related with the types and content of different ginsenosides. Therefore, we further screened proteins related to ginsenoside biosynthesis, including the well-known mevalonate (MVA) and 2C-methyl-D-erythritol-4-phosphorylation (MEP) pathway proteins, cytochromes P450 and UGT enzymes. Cytochrome P450 is a complex family of proteins that have powerful functions. As terminal oxygenases, they can insert an oxygen atom into hydrophobic molecules to change their hydrophilicity and activity. CYPs participate in a variety of secondary metabolic reactions in plants and play an important biological role [24]. During the biosynthesis of ginsenosides, they mainly carried out the catalysis of the hydroxylation process [25]. Glycosyltransferase family can catalyze the glycosylation of receptor molecules, thereby altering their biological activity, solubility, and stability *in vivo*, thus play an important role in cell biological regulation and metabolite synthesis and storage [26]. UGTs mainly use UDP-glucose and UDP-glucuronic acid as donors to glycosylate receptor molecules and participate in important glycosylation steps in the biosynthesis of ginsenosides [27,28]. These complex biomolecular modification reactions enable the generation of diverse and complex ginsenoside molecules via MVA and MEP pathways. As listed in Tables S4 and S5, a total of 32 proteins related to ginsenoside synthesis were identified in ginseng root, including six enzymes in the MVA pathway, six enzymes in the MEP pathway and one FPS, one PPDS, one PPTS, eight CYPs, and nine UGTs in the downstream pathway. A total of 78 proteins related to ginsenoside synthesis were identified in the ginseng cauline leaf, including five enzymes in the MVA pathway, six enzymes in the MEP pathway and one FPS, one PPDS, one PPTS, 33 CYPs, and 31 UGTs in the downstream pathway. AACT (PG11912), MVD (PG37213), and IDI (PG19600) were co-identified in root and cauline leaf in the MVA pathway. HMGS (PG38955), MVK (PG20228), and MVD (PG02263)

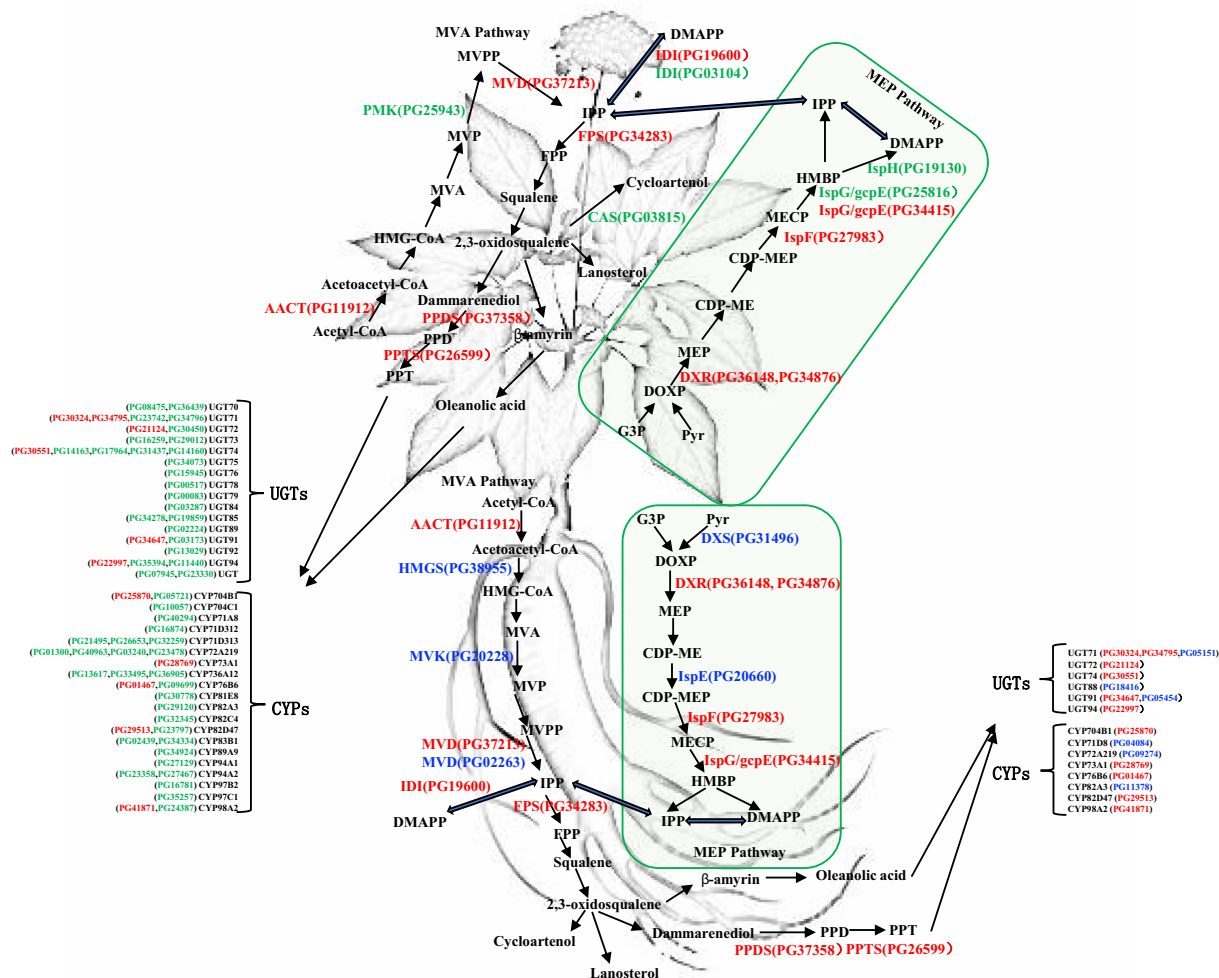


Fig. 6. Spatial distribution of proteins identified related to ginsenoside biosynthesis pathway. The proteins are marked in red for co-identified, green for caudex leaf, and blue for ginseng root only.

were only identified from root. PMK (PG25943) and IDI (PG03104) were only identified in caudex leaf. DXR (PG36148, PG34876), IspF (PG27983), and IspG/gcpE (PG34415) were co-identified in root and caudex leaf in the MEP pathway. DXS (PG31496) and IspE (PG20660) were only identified in the root. IspG/gcpE (PG25816) and IspH (PG19130) were only identified in caudex leaf. FPS (PG34283), PPDS (PG37358), and PPTS (PG26599) in the downstream pathway were co-identified from root and caudex leaf.

The aglycone skeleton of ginsenoside is classified into a dammarane type and an oleanane type. The triterpene skeleton is converted into various ginsenosides by hydroxylation of CYP and glycosylation of uridine diphosphate-dependent glycosyltransferase (UGT). A total of 33 CYPs were identified in the ginseng caudex leaf, whereas only eight CYPs were identified from the root. A total of five CYPs were co-identified, including CYP704B1 (PG25870), CYP73A1 (PG28769), CYP76B6 (PG01467), CYP82D47 (PG29513), and CYP98A2 (PG41871). There are three CYPs unique to ginseng root, including CYP71D8 (PG04084), CYP72A219 (PG09274), and CYP82A3 (PG11378). There are 28 unique CYPs in the caudex leaf, including CYP704B1 (PG05721), CYP704C1 (PG10057), CYP71A8 (PG40294), CYP71D312 (PG16874), CYP71D313 (PG21495, PG26653, PG32259), CYP72A219 (PG01300, PG40963, PG03240, PG23478), CYP736A12 (PG13617, PG33495, PG36905), CYP76B6 (PG09699), CYP81E8 (PG30778), CYP82A3 (PG29120), CYP82C4

(PG32345), CYP82D47 (PG23797), CYP83B1 (PG02439, PG34334), CYP89A9 (PG34924), CYP94A1 (PG27129), CYP94A2 (PG23358, PG27467), CYP97B2 (PG16781), CYP97C1 (PG35257), and CYP98A2 (PG24387). A total of 31 UGTs were identified in the ginseng caudex leaf, whereas only nine UGTs were identified from the root. A total of six UGTs were co-identified, including UGT71 (PG30324, PG34795, PG05151), UGT72 (PG21124), UGT74 (PG30551), UGT91 (PG34647), and UGT94 (PG22997). There are three UGTs unique to ginseng root, including UGT71 (PG05151), UGT88 (PG18416), and UGT91 (PG05454). There are 25 unique UGT proteins in the caudex leaf, including UGT70 (PG08475, PG36439), UGT71 (PG23742, PG34796), UGT72 (PG30450), UGT73 (PG16259, PG29012), UGT74 (PG14163, PG17964, PG31437, PG14160), UGT75 (PG34073), UGT76 (PG15945), UGT78 (PG00517), UGT79 (PG00083), UGT84 (PG03287), UGT85 (PG34278, PG19859), UGT89 (PG02224), UGT91 (PG03173), UGT92 (PG13029), UGT94 (PG35394, PG11440), UGT (PG07945), and UGT (PG23330). Fig. 6 showed the spatial distribution of proteins related to ginsenoside biosynthesis pathway according to the location of the identified proteins [27,29–32]. We hypothesized that the synthesis of ginsenosides are carried out in the root and caudex leaf simultaneously, whereas the post glucoside adduction of ginsenosides might be more conducted when growing in the caudex leaf and then transported to the root at withering every year. Our hypothesis is also consistent to the

findings of ginsenoside distribution profile on the tissue part of ginseng. We believe that the realization of the protein expression localization will greatly help us on the understanding of the biosynthesis pathway of the ginsenosides.

4. Conclusion

In this report, we explored different methods to extract proteins from ginseng cauline leaf and root. State-of-art nano-LC-MS/MS was used to analyze the extracted proteins. The biological functions of the proteins expressed in different parts of ginseng were compared, and the enzymes involved in ginsenoside biosynthesis and modification were fished out for ginsenoside biosynthesis pathway analysis. Spatial distribution of CYPs, UGTs, and the proteins for MVA and MEP pathways was described to better understand the related processes of ginsenoside biosynthesis. Interestingly, a significant higher number of CYPs and UGTs were identified from ginseng cauline leaf than ginseng root, which indicated that this important glycosylation step in ginsenoside biosynthesis might be more conducted in cauline leaf. We believe that the in-depth proteomic analysis will help us on the knowledge of ginsenoside synthesis and improvement of ginseng cultivation.

Conflicts of interest

All authors have no conflicts of interest to declare.

Acknowledgments

Financial supports from National Science Foundation (21675061) and Jilin Provincial Science & Technology Department (20170101151JC) are gratefully acknowledged.

Appendix A. Supplementary data

Supplementary data to this article can be found online at <https://doi.org/10.1016/j.jgr.2020.01.009>.

References

- Colzani M, Altomare A, Caliendo M, Aldini G, Righetti PG, Fasoli E. The secrets of Oriental panacea: *Panax ginseng*. *J Proteomics* 2016;130:150–9.
- Kim SW, Gupta R, Min CW, Lee SH, Cheon YE, Meng QF, Jang JW, Hong CE, Lee JY, Jo IH, et al. Label-free quantitative proteomic analysis of *Panax ginseng* leaves upon exposure to heat stress. *J Ginseng Res* 2019;43:143–53.
- Helms S. Cancer prevention and therapeutics: *Panax ginseng*. *Altern Med Rev* 2004;9:259–74.
- Kiefer D, Pantuso T. *Panax ginseng*. *Am Fam Physician* 2003;68:1539–42.
- Kim JH, Yi YS, Kim MY, Cho JY. Role of ginsenosides, the main active components of *Panax ginseng*, in inflammatory responses and diseases. *J Ginseng Res* 2017;41:435–43.
- Kim SK, Park JH. Trends in ginseng research in 2010. *J Ginseng Res* 2011;35:389–98.
- Shin BK, Kwon SW, Park JH. Chemical diversity of ginseng saponins from *Panax ginseng*. *J Ginseng Res* 2015;39:287–98.
- Kim SW, Lee SH, Min CW, Jo IH, Bang KH, Hyun D-Y, Agrawal GK, Rakwal R, Zargar SM, Gupta R, et al. Ginseng (*Panax sp.*) proteomics: an update. *Applied Biol Chem* 2017;60:311–20.
- Ma R, Sun L, Chen X, Jiang R, Sun H, Zhao D. Proteomic changes in different growth periods of ginseng roots. *Plant Physiol Biochem* 2013;67:20–32.
- Nam MH, Kim SI, Liu JR, Yang DC, Lim YP, Kwon KH, Yoo JS, Park YM. Proteomic analysis of Korean ginseng (*Panax ginseng* C.A. Meyer). *J Chromatogr B Analyt Technol Biomed Life Sci* 2005;815:147–55.
- Jayakodi M, Choi BS, Lee SC, Kim NH, Park JY, Jang W, Lakshmanan M, Mohan SVG, Lee DY, Yang TJ. Ginseng Genome Database: an open-access platform for genomics of *Panax ginseng*. *BMC Plant Biol* 2018;18:62.
- Kim NH, Jayakodi M, Lee SC, Choi BS, Jang W, Lee J, Kim HH, Waminal NE, Lakshmanan M, van Nguyen B, et al. Genome and evolution of the shade-requiring medicinal herb *Panax ginseng*. *Plant Biotechnol J* 2018;16:1904–17.
- Xu J, Chu Y, Liao B, Xiao S, Yin Q, Bai R, Su H, Dong L, Li X, Qian J, et al. *Panax ginseng* genome examination for ginsenoside biosynthesis. *Gigascience* 2017;6:1–15.
- Hsu CC, Zhu Y, Arrington JV, Paez JS, Wang P, Zhu P, Chen IH, Zhu JK, Tao WA. Universal plant phosphoproteomics workflow and its application to tomato signaling in response to cold stress. *Mol Cell Proteomics* 2018;17:2068–80.
- Yang F, Shen Y, Camp 2nd DG, Smith RD. High-pH reversed-phase chromatography with fraction concatenation for 2D proteomic analysis. *Expert Rev Proteomics* 2012;9:129–34.
- Wang Y, Yang F, Gritsenko MA, Wang Y, Claus T, Liu T, Shen Y, Monroe ME, Lopez-Ferrer D, Reno T, et al. Reversed-phase chromatography with multiple fraction concatenation strategy for proteome profiling of human MCF10A cells. *Proteomics* 2011;11:2019–26.
- Song C, Ye M, Han G, Jiang X, Wang F, Yu Z, Chen R, Zou H. Reversed-phase liquid chromatography approach with high orthogonality for multi-dimensional separation of phosphopeptides. *Anal Chem* 2010;82:53–6.
- Han D, Moon S, Kim Y, Kim J, Jin J, Kim Y. In-depth proteomic analysis of mouse microglia using a combination of FASP and StageTip-based, high pH, reversed-phase fractionation. *Proteomics* 2013;13:2984–8.
- Kim SI, Kweon SM, Kim EA, Kim JY, Kim S, Yoo JS, Park YM. Characterization of rRNA-like major storage protein from the ginseng root by proteomic approach. *J Plant Physiol* 2004;161:837–945.
- Duck NB, Folk WR. Hsp70 heat shock protein cognate is expressed and stored in developing tomato pollen. *Plant Mol Biol* 1994;26:1031–9.
- Gani Z, Boradia VM, Raghu Ram J, Suryavanshi PM, Patil P, Kumar S, Singh R, Raje M, Raje CI. Purification and characterization of glyceraldehyde-3-phosphate-dehydrogenase (GAPDH) from pea seeds. *Protein Expr Purif* 2016;127:22–7.
- Andersson I, Backlund A. Structure and function of rubisco. *Plant Physiol Biochem* 2008;46:275–91.
- Yuan HM, Liu WC, Lu YT. CATALASE2 coordinates SA-mediated repression of both auxin accumulation and JA biosynthesis in plant defenses. *Cell Host Microbe* 2017;21:143–55.
- Siminszky B, Corbin FT, Ward ER, Fleischmann TJ, Dewey RE. Expression of a soybean cytochrome P450 monooxygenase cDNA in yeast and tobacco enhances the metabolism of phenylurea herbicides. *Proc Natl Acad Sci U S A* 1999;96:1750–5.
- Wang Y, Li X, Lin Y, Wang Y, Wang K, Sun C, Lu T, Zhang M. Structural variation, functional differentiation, and activity correlation of the cytochrome P450 gene superfamily revealed in ginseng. *Plant Genome* 2018:11.
- Hu Y, Walker S. Remarkable structural similarities between diverse glycosyltransferases. *Chem Biol* 2002;9:1287–96.
- Kang KB, Jayakodi M, Lee YS, Nguyen VB, Park HS, Koo HJ, Choi IY, Kim DH, Chung YJ, Ryu B, et al. Identification of candidate UDP-glycosyltransferases involved in protopanaxadiol-type ginsenoside biosynthesis in *Panax ginseng*. *Sci Rep* 2018;8:11744.
- Luo SL, Dang LZ, Zhang KQ, Liang LM, Li GH. Cloning and heterologous expression of UDP-glycosyltransferase genes from *Bacillus subtilis* and its application in the glycosylation of ginsenoside Rh1. *Lett Appl Microbiol* 2015;60:72–8.
- Lu J, Li J, Wang S, Yao L, Liang W, Wang J, Gao W. Advances in ginsenoside biosynthesis and metabolic regulation. *Biotechnol Appl Biochem* 2018;65:514–22.
- Zhao S, Wang L, Liu L, Liang Y, Sun Y, Wu J. Both the mevalonate and the non-mevalonate pathways are involved in ginsenoside biosynthesis. *Plant Cell Rep* 2014;33:393–400.
- Kim YJ, Zhang D, Yang DC. Biosynthesis and biotechnological production of ginsenosides. *Biotechnol Adv* 2015;33:717–35.
- Kim YJ, Jeon JN, Jang MG, Oh JY, Kwon WS, Jung SK, Yang DC. Ginsenoside profiles and related gene expression during foliation in *Panax ginseng* Meyer. *J Ginseng Res* 2014;38:66–72.