iTRAQ-based proteomic analysis reveals the effect of ribosomal proteins on essential-oil accumulation in *H. cordata* Thunb.

Dandan Guo[1,2,#], Beixuan He[3,#], Fei Feng[1], Diya Lv[1,2], Ting Han[1], Xiaofei Chen[1,2*]

[1]School of Pharmacy, Naval Medical University (Second Military Medical University), 325 Guohe Road, Shanghai, 200433, China

[2]Shanghai Key Laboratory for Pharmaceutical Metabolite Research, Shanghai, 200433, China

[3]State Key Laboratory of Oncogenes and Related Genes, Shanghai Cancer Institute, Shanghai, 200032, China

[#] These authors contribute equally.

*Corresponding Authors:

Xiaofei Chen,

325 Guohe Road, Shanghai, 200433, China.

E-mail address: xfchen2010@163.com.

Abstract

Houttuynia cordata Thunb., referred to as Yuxingcao in Chinese, holds a central position in Asian traditional medicine and cuisine. Remarkable chemical diversity, particularly in essential oil, exists between the aerial parts and underground stems of *H. cordata*. Despite this, the mechanisms governing essential oil biosynthesis in *H. cordata* remain enigmatic. In this study, we present a quantitative overview of the proteomes across four tissues (flower, stem, leaf, and underground stem) of *H. cordata*, achieved through the application of the isobaric tag for relative and absolute quantitation (iTRAQ). Our research findings suggest that certain crucial ribosomal proteins and their interactions could significantly impact the production of essential oils in *H. cordata*. These results offer novel insights for investigating and understanding the roles of ribosomal proteins and their associations in essential oil

Comment [11]: Full name in the title

Comment [12]: Delete

biosynthesis across various organisms of *H. cordata*.

INTRODUCTION

Herbal medicine boasts a rich history spanning various cultures, with numerous contemporary pharmaceuticals originally derived from these medicinal plants (Drasar & Khripach 2019). Over 70 to 95% of citizens in developing countries use medicinal plants for their health care, which shown in reports of The World Health Organization(Farnsworth et al. 1985). With the increasing interest in the application of medicinal plants, there arises a pressing demand for the analysis of frequently utilized botanical remedies. A multitude of herbs exhibiting a "Medicine-food homology", which ensures their safety for clinical use, have been comprehensively investigated on a global scale(Hou & Jiang 2013).

H. cordata, a perennial herbaceous plant categorized under the *Saururaceae* family, thrives in the damp and shaded locales of southwestern China. In traditional Chinese medicine, various parts of *H. cordata* are employed to alleviate swelling and pain, reduce inflammation, suppress coughs, and enhance diuretic effects(Wu et al. 2021). Furthermore, the rhizomes and tender leaves of the plant are gathered as delectable spices and nutritious vegetables. Recent research has unveiled its diverse array of attributes, including anti-allergic, anti-inflammatory, antiviral, antioxidant, anti-leukemic, and anti-cancer properties (Chen et al. 2013; Chiang et al. 2003; Kim et al. 2001; Li et al. 2005; Lu et al. 2006b; Ng et al. 2007; Zhuang et al. 2015). These diverse pharmacological activities are frequently linked to the chemical composition of H. cordata, encompassing compounds such as alkaloids, essential oils, and flavonoids (Bauer et al. 1996). Moreover, the bioactivities of *H. cordata*'s oil have been documented since 1921, encompassing compounds such as decanoyl acetaldehyde, myrcene, ethyl decanoate, ethyl dodecanoate, and more (Xu 2012). Research into the extraction methods, bioactivities, and distribution of phytochemicals in *H. cordata*'s oil has been underway for nearly a century (Xu 2012). A notable distinction was evident between the aerial stems and the underground parts,

Comment [13]: Space

Comment [14]: Reference is so old the number may differ at now a day

Comment [15]: Space

Comment [16]: Space

Comment [17]: The references should be ordered from the oldest to the newest

Comment [18]:

Comment [19]: The same reference for 2 respective phrases. Delete one

with elevated levels of 2-undecanone, myrcene, ethyl decanoate, ethyl dodecanoate, 2-tridecanone, and decanal observed in the aerial parts compared to the underground components. Additionally, 11 constituents were exclusively isolated from the leaves, whereas seven of the identified components in the underground-stems were absent from the leaves (Verma et al. 2017).

Isobaric tags for relative or absolute quantitation (iTRAQ), recognized as one of the premier mass spectrometry techniques for facilitating high-throughput proteomic analysis, offers the benefits of heightened sensitivity and robustness(Evans et al. 2012). iTRAQ has found extensive application across various domains including animals, plants, and microorganisms(Baslam et al. 2020; Lü et al. 2021). In recent times, iTRAQ analysis has been employed in some medical plants to investigate proteomic shifts in *Panax notoginseng* seeds (Ge et al. 2021), assess protein abundance in the underground stems of *Rehmannia glutinosa*(Chen et al. 2021), explore proteomic changes during the development haustorium of *Taxillus chinensis* (Pan et al. 2022), and conduct proteome profiling of *Anemone flaccida* (Zhan et al. 2016). In this current study, we undertook tandem transcriptome and proteome profiling of distinct medicinal parts in *H. cordata*. Consequently, a total of 90,067 transcripts were assembled from Flower (F), Stem (S), Leaf (L), and Underground-stem (Us) tissues, and were subsequently used to construct a Coding Sequence (CDS) library for proteomic analysis. Overall, 3,930 proteins were successfully identified, allowing for a comparative examination of protein expression profiles among the diverse parts of *H. cordata* through iTRAQ-based proteomics analysis. This investigation represents the inaugural exploration into the variation of proteins across the distinct parts in *H. cordata*, thereby deepening our comprehension of the molecular mechanisms underlying their development and the synthesis of active ingredients.

Comment [110]: Space

Comment [111]: Space

Comment [112]: Space

Comment [113]: This part should be replaced by the objective of your research

MATERIALS AND METHODS

Plant material and sample source

H. cordata wild type plants were grown in medical plant garden of pharmacy

college, Navy Medical University, under the natural climate. Different tissues samples for flower, stem, leave and underground-stem were collected from flourishing plants in summer. They were put in a 1.5ml tube and immediately in liquid nitrogen and stored at -80°C until further use.

CDS library prediction

The RNA of flower, stem, leave and underground-stem tissues were extracted by Trizol kit as described. Total amounts and integrity of RNA were assessed by the RNA Nano 6000 Assay Kit of the Bioanalyzer 2100 system (Agilent Technologies, CA, USA). cDNA library construction started with mRNA purification and then doble- strand cDNA synthesis and was quantified to ensure the quality of library. The separate libraries are pooling equivalently according to the concentration of samples, then being sequenced by the use of the Illumina NovaSeq 6000. The transcriptome assembly was performed using Trinity software (version 2.6.6) (Grabherr et al. 2011). Several databases were used to annotate gene functional annotation: Nr (NCBI non-redundant protein sequences); Nt (NCBI non-redundant nucleotide sequences); Pfam (Protein family); KOG/COG (Clusters of Orthologous Groups of proteins); Swiss-Prot (A manually annotated and reviewed protein sequence database); KO (KEGG Ortholog database); GO (Gene Ontology).

CDS represents the sequence encoding protein. CDS prediction firstly is mapped into NR and Swissprot. protein library. If mapped successfully, the Open Reading Frame (ORF) of the transcript is isolated and the coding region sequence is translated into amino acid sequence (5'- > 3'). For the sequences without successfully, the ORF was predicted by TransDecoder (3.0.1) software.

Protein extraction

Protein extraction was carried out according to the protocol of Isaacson et al. (Isaacson et al. 2006) with some modifications. Frozen tissues were homogenized with a plant-tissue grinder. Protein extraction buffer was added to the sample and mixed gently in a vortex for 5 ~ 10 min. Adding two volume of phenol saturated with Tris-HCl (pH 8.0) to each sample and then mixing for 10 min, followed by centrifugation at 14000g for 10 min at 4°C. The upper phenolic phase was picked and

mixed with 1.2 volume of protein extraction buffer, and then centrifuged. The upper phenolic phase was again collected and mixed with pre-cooled 0.1M ammonium acetate in methanol, followed by the incubation for 6h at -20°C. The precipitate was obtained by centrifugation at 14000g for 10min and washed 1-2 times by pre-cooled methanol. Then the precipitate was washed 2-3 times by acetone containing 0.07% β-Mercaptoethanol and dried at room temperature. Protein concentration was determined by the BCA method (Smith et al. 1985).

Protein digestion

Protein digestion was carried out according to the method of FASP (Wiśniewski et al. 2009). Sample dried protein powered was resolved in 8M urea with Tris-Hcl (pH 8.0). 1M reducing reagent (DTT) was added to cell protein lysates and incubated at 55°C for 1 h, followed by 5 μL 1M cysteine-blocking reagent (IAA) for 30 min at 25°C in dark. Then the samples were transferred into 10 KDa ultrafiltration tube and centrifuged 14000 g for 15 min at 4°C. Then 100μL 8M urea was added twice to the ultrafiltration membrane to protein denaturation completely. 0.5 M TEAB was added and centrifuged three times at 14000 g for 15 min to clean denaturation reagent. Sequencing grade trypsin (Enzyme : Substrate = 1:50) was incubated with samples at 37°C overnight. Next day, the peptide samples were collected and added 1% formic acid to stop the reaction.

iTRAQ labeling and high pH reverse phase separation by HPLC

The peptides were in dried in a vacuum centrifugal concentrator. Then the dried peptides were suspended in TEAB buffer and labeled by iTRAQ 8-plex kits according to the manufacturer's guidebook (AB SCIEX Inc., USA). Labelling reaction was performed by adding one reagent vial to digested peptides and proceeded for 2 h at room temperature, followed by adding 100 μL water to stop the reaction. The labelling scheme was as follows: Tags 113 and 114, F1 and F2; Tags 115 and 116, L1 and L2; Tags 117 and 118, S1 and S2; Tags 119 and 121, R1 and R2. after that, all the samples were pooled for mass spectrometry analysis. High pH reverse phase separation by HPLC was performed as previous decribed (You et al. 2017) with some modifications. Combining Agilent 1100 HPLC with Agilent Zorbax Extend C18

Comment [114]: Delete

Comment [115]: Where it described ?

column (2.1×150 mm, 5 µm) were used for peptide separation by 1.0 ml/min. The mobile phases A was 20 mM ammonium formate (pH = 10) and the mobile phases A was 80% acetonitrile with 20mM ammonium formate (pH=10). The entire gradient course was 65min (5min, 5% B; 30min, 15% B; 45min, 38% B; 46min, 90%B; 54.5 min, 90% B; 55 min, 5% B; 65 min, 5% B).

LC–MS/MS analysis

Microflow LC-MS/MS was performed by coupling Eksigent Micro LC-1D plus system to Triple TOF 5600 System (ABSCIEX). Samples for the proteome analysis were resuspended in 2% acetonitrile containing 0.1% formic acid. Peptide loading and washing were done on a trap column (ChromeXP C18_CL-3µm, 120A, 350µm ×0.5µm, ABSCIEX) at a flow rate of 10µL/min in 2% acetonitrile (0.1% formic acid) for 7 min. Peptide separation was performed on an analytical column (0.075×150 mm, 3µm, 120A, ABSCIEX) at a flow rate of 5µL/min using a 120 min gradient from 5% to 80% solvent B (solvent A: 2% acetonitrile with 0.1% formic acid in LC-MS grade water; solvent B: 98% acetonitrile with 0.1% formic acid in acetonitrile) for proteome analysis. A spray voltage of 2.3 kV and ESI ion source temperature of 150°C were used to ionize peptides. A switch from MS to MS2 scanning was automatically initiated based on the data collected from the instrument. For the full proteome samples, full scan MS spectra (m/z 350 -1,500) were acquired. In Time of Fight, high-resolution MS2 spectra for the 40 top precursor ions were acquired. This analysis focused on fragmenting precursors with charge states between 2 and 5.

Analysis of proteomic data

Peptide and protein identification and quantification data files were searched with Protein Pilot software (v.5.0) using default parameters(Zhang et al. 2014). MS/MS spectra were searched against CDS library (Hc. blast.pep.fasta). Trypsin was the only specific protease and iTRAQ-8-plex (peptide labelled) was chosen as sample labeling type (FDR ≤ 1%). The annotations of the identified proteins, combined with the result of BLAST alignments. The screening criteria for Differential Expression Proteins (DEPs) was |Fold Change (FC)| ≥ 2 and a p-value ≤0.05 in aerial parts (Leaf, Stem, Flower) compared to underground -stem. The Weighted Gene CoExpression

Comment [116]: According to reference or according to what ?

Comment [117]: ??!!

Comment [118]: Space

Network Analysisn (WGCNA) of identified proteins was performed with the R package WGCNA (https://cran.rstudio.com/web/packages/WGCNA), and further bioinformatic analysis of DEP was performed using the OECloud tools (https://cloud.oebiotech.cn/), based on the Gene Ontology (GO) categories and KEGG Ortholog database.

Bioinformatic analysis of different expression proteins

The KEGG pathway enrichment profile of DEPs was analyzed by OECloud tools with FDRs less than 0.05, and DEPs data was displayed using volcano plots and Wayne diagram with the help of the image GP. The prediction of functional protein association networks was analyzed by the STRING database (http://string-db.org/).

RESULTS

Plant morphology and chemical composition

The general morphology of the *H. cordata* plant during its flowering period is illustrated (Fig. 1A). The plant is segmented into four parts: flowers, stems, leaves, and underground stems (Us). Flowers, stems, and leaves are classified as aerial parts (AP). Both AP and Us find application as medicinal components, recognized for their heat-clearing and detoxifying properties in Traditional Chinese Medicine. Among the constituents of *H. cordata* essential oil, the most predominant compounds were 2-undecanone (23.96-36.07%), β-myrcene (12.57-14.29%), bornyl acetate (6.03-8.61%), and β-pinene (trace - 23.29%) (chemical structures shown in Fig. 1B). The previous findings indicated that the essential oil content is higher in the aerial parts of *H. cordata* compared to its underground stems(Lu et al. 2006a). These findings suggest that the substantial variation of essential oil content could potentially be linked to the differential expression of proteins across distinct plant parts in *H. cordata*.

Comment [119]: Space

Assembled transcriptome analysis

To perform proteomics analysis, RNA-seq is carried out to build non-model plant proteins identification search database. Based on transcriptome result, 90,067 transcripts were assembled in tissues pool (flower, stem, leave and underground-stem)

of *H. cordata*, of which 33,493 were unigenes that represents the longest transcript of each gene. More than 50% of unigenes were distributed between 300-1000bp (Fig. S1A). 21,661 unigenes (64.67%) were annotated by Nr (NCBI non-redundant protein sequences) database, while annotated 13,290 unigenes were screened out by Nt (NCBI nucleotide sequences) database. 15,756 were annotated by GO (Gene Ontology) and involved in cellular and metabolic process (Fig. S1B and Fig.1C). 8,928 unigenes were mapped to KO and KEGG (Kyoto Encyclopedia of Genes and Genomes) pathway which mainly participated in translation and signal transduction (Fig. 1C and Fig. 1D). In total of 4,177 unigenes existed in multiple databases. 19,271 peptides sequence of CDS were predicted through Nr and SwissProt. Assembled transcriptome database serves as an extensive public repository, serving as a valuable search resource for proteomic studies of *H. cordata*.

Comment [120]: Which one. You have 1, 2, 3, and 4 which one is S1

Figure 1 Phenotype and chemical components and RNA-seq analysis of *H. cordata*

(A) Tissue samples analyzed according to morphology group, F, Flower, S, Stem, L, Leaf, and Us, Underground-stem. AP: Aerial parts, Us: Underground-stem. (B) Structures of the major chemical components of *H. cordata*. (C) GO annotation classification includes biological process (BP), cellular component (CC), molecular Function (MF). (D) KEGG pathway classification was divided into five branches.

Overview of proteomics analysis

According to the proteomics analysis, a total of 3,261 proteins were identified across the four parts (Flower, Stem, Leaf, and Underground-stem) of *H. cordata*. Notably, the expression clusters exhibited significant distribution shifts among the various plant parts as shown in Fig.2A. In terms of GO categories, 1,470 proteins participated in biological processes, 1,194 proteins were linked to cellular components, and 2,247 proteins were associated with molecular functions (Fig. S2A). Moreover, 1,684 proteins were assigned to KEGG pathways, with 31 notable signaling pathways exhibiting significance (p value ≤ 0.05). The leading three pathways encompassed dihydrolipoamide dehydrogenase, photosystem I P700 chlorophyll apoprotein A1, and

the 26S proteasome complex subunit DSS1(Fig. S2B). To assess the correlation coefficients among the expressed proteins across various parts of H. cordata, WGCNA analysis was conducted. This analysis categorized the 3,930 proteins into five distinct modules, revealing correlations that ranged from high to low among the MEturquoise proteins. Notably, the proteins expressed in the underground-stem exhibited a lower correlation coefficient with those in the aerial parts (Fig. 2B and Fig. S3).

Comment [121]: Space

Comment [122]: Italic

Comment [123]: ???

Figure 2 Cluster maps of identified proteins in each tissue of *H. cordata*.
(A) Heatmaps of proteins in different tissues. (B) Visualization of WGCNA results.

DEPs analysis of aerial parts and underground-stem

A significance cutoff was established with criteria including a fold change ≥ 1.5 and a p-value ≤ 0.05 (t-test), along with a false discovery rate of ≤ 1%. Compared with the proteins in underground-stem, 306 proteins were up-regulated in leaf, 192 were up-regulated in flower, and 266 were up-regulated in stem, there were 61 proteins increased both in leaf and stem (Fig. 3A). In comparison to the underground-stem, ribosomal proteins L23A and S28 displayed notable upregulation in flower tissue. Similarly, proteins L1, L30e, L23A, L4, and L29 exhibited upregulation in both the stem and leaf. Conversely, P3-like and S8 were found to be downregulated in the stem and leaf (Fig. 3B-3D). Furthermore, the Differential Abundance Proteins (DAPs) were associated with KEGG pathways, with 251 proteins being mapped to these pathways. Notably, this analysis revealed 12 significant signaling pathways (p-value ≤ 0.05), where the Differentially Expressed Proteins (DEPs) were most prominently enriched within the large subunit ribosomal protein group (Fig. 3E). The findings suggest that large subunit ribosomal proteins might play a role in influencing essential oil accumulation in aerial parts (AP).

Figure 3 Differentially expressed proteins analysis of different tissues in *H. cordata*
(A) Venn diagrams showing the numbers of differentially expressed proteins. (B, C, D) Volcano plot of differentially expressed proteins. (E) KEGG pathway enrichment analysis of differentially

expressed proteins.

The expression profile of interesting proteins

The ribosome, an ancient, intricate, and complex organelle, maintains a remarkably conserved structure and composition across prokaryotes and eukaryotes(Ramakrishnan & White 1998). In a eukaryotic cell, the ribosome comprises four ribosomal RNAs and 79-81 ribosomal proteins. It is widely acknowledged that ribosomal proteins contribute to the maintenance of RNA conformation (Barakat et al. 2001). Nevertheless, an increasing number of studies have uncovered that ribosomal proteins are engaged not solely in rRNA processing, folding, assembly, and ribosomal subunit transportation, but also in the reinforcement of subunit architectures and the interaction between ribosomes and diverse translation factors. Furthermore, they contribute to the folding and positioning of nascent peptides and potentially possess additional biological functions beyond the ribosomal context(Wilson & Doudna Cate 2012).

Comment [124]: Space

Comment [125]: Remove to the discussion part

In this study, a total of 34 ribosomal proteins were identified, comprising 10 small subunit ribosomal proteins and 24 large subunit ribosomal proteins (Table 1). The majority of these exhibited higher expression levels in the leaf and stem, a pattern consistent with the spatial distribution of active components (Fig. 4A) (Lu et al. 2006a). The internet analysis revealed the majority of ribosomal proteins in *H. cordata* exhibited robust interactions with those in *Arabidopsis thaliana* according to STRING model data. (Fig. 4B, Table 1).

Figure 4 Differentially expressed analysis of ribosomal proteins

(A) Ribosomal proteins expression profiling in different tissues of *H. cordata*. (B) Interaction network of subunit ribosomal protein compared to *Arabidopsis thaliana* (confidence was set ≥ 0.9).

DISCUSSION

As known well, leaves, stems, and underground-stems of H. cordata hold considerable importance in traditional Chinese medicine due to their applications in heat-clearing and detoxification therapies. Moreover, distinct variations in chemical composition were observed among different plant parts (Verma et al. 2017). The essential oil, which constitutes the primary active constituents, exhibits substantial variation between the aerial and underground parts (Verma et al. 2017). However, the distinctive protein profiles that form the groundwork for subsequent investigations into molecular mechanisms remain elusive. In order to shed light on the variations in protein expression in *H. cordata* tissues, a comparative proteomic analysis utilizing iTRAQ was conducted. A total of 306 and 266 Differentially Abundant Proteins (DAPs) were successfully identified in the leaf and stem, respectively, exhibiting upregulated expression compared to the underground-stem. These DAPs were categorized based on their functions and were found to be enriched in pathways such as photosynthesis, glycolysis, and ribosomal processes. These findings lay the foundation for identifying and functionally analyzing potential proteins linked to the interspecific variation of essential oil in *H. cordata*.

Until now, only a limited number of studies have unveiled the significant role of ribosomal proteins in essential oil accumulation, potentially governing the biosynthesis of secondary metabolites and even impacting plant development and abiotic stress tolerance (Wilson & Doudna Cate 2012). According to reports, ribosomal proteins uS2c, uS4c, bL20c, and bL33c have been shown to impact abnormal leaf morphologies (Rogalski et al. 2008). Additionally, ribosomal protein bS1c is known to affect nuclear and plastid genes encoding proteins involved in chlorophyll biosynthesis, chloroplast development, and photosynthesis (Gong et al. 2013; Zhou et al. 2021). Furthermore, Rogalski *et al*. discovered that tobacco's PRP bL33c plays a role in enhancing plant tolerance to low-temperature stress (Rogalski et al. 2008). In addition, PRP bS1c was observed to impede HsfA2-dependent heat stress responses in the chloroplasts of Arabidopsis(Yu et al. 2012).

While substantial strides have been achieved in recent years, a comprehensive understanding of the individual functions of each PRP and their involvement in

Comment [126]: Italic

Comment [127]: The same reference for 2 respective phrases. Delete one

Comment [128]: ???

Comment [129]: Space

specific biological processes still eludes us. In this current study, 18 upregulated Differentially Abundant Proteins (DAPs) were recognized as ribosomal proteins, encompassing large subunit ribosomal proteins L13, L19, and L3. In the forthcoming research, there may be endeavors to manipulate ribosomal protein expression using biotechnological tools, aiming to stimulate essential oil biosynthesis in *H. cordata*. These findings lay the groundwork for guiding the utilization of seed resources and enhancing the chemical diversity of components in both the aerial parts and underground-stem of *H. cordata*.

CONCLUSION

In this study, we reported iTRAQ-based quantitative proteomics of *H. cordata*, identifying a total of 3,261 proteins and mapping the differential expression pattern across multiple tissues. By correlating analysis with essential oil content, our results showed that ribosomal proteins could significantly influence essential oil production in *H. cordata*. The research provides a novel insight into the study of protein function involved in essential oil in non-model plants.

ACKNOWLEDGEMENTS

We acknowledge the assistance of staff at the Pharmaceutical Analysis and Testing center.

References

Barakat A, Szick-Miranda K, Chang IF, Guyot R, Blanc G, Cooke R, Delseny M, and Bailey-Serres J. 2001. The organization of cytoplasmic ribosomal protein genes in the Arabidopsis genome. *Plant physiology* 127:398-415.

Baslam M, Kaneko K, and Mitsui T. 2020. iTRAQ-Based Proteomic Analysis of Rice Grains. *Methods in molecular biology (Clifton, NJ)* 2139:405–414.

Bauer R, Pröbstle A, Lotter H, Wagner-Redecker W, and Matthiesen U. 1996. Cyclooxygenase Inhibitory Constituents from Houttuynia Cordata. *Phytomedicine*

2:305–308.

Chen HY, Lin YH, Thien PF, Chang SC, Chen YC, Lo SS, Yang SH, and Chen JL. 2013. Identifying Core Herbal Treatments for Children with Asthma: Implication from a Chinese Herbal Medicine Database in Taiwan. . *Evid-Based Complement Altern Med*:125943.

Chen P, Wei X, Qi Q, Jia W, Zhao M, Wang H, Zhou Y, and Duan H. 2021. Study of Terpenoid Synthesis and Prenyltransferase in Roots of Rehmannia glutinosa Based on iTRAQ Quantitative Proteomics. *Frontiers in plant science* 12:693758.

Chiang LC, Chang JS, Chen CC, Ng LT, and Lin CC. 2003. Anti-Herpes Simplex Virus Activity of Bidens Pilosa and Houttuynia Cordata. *Am J Chin Med* 31:355–362.

Drasar PB, and Khripach VA. 2019. Growing Importance of Natural Products Research. *Molecules* 25:6.

Evans C, Noirel J, Ow SY, Salim M, Pereira-Medrano AG, Couto N, Pandhal J, Smith D, Pham TK, Karunakaran E, Zou X, Biggs CA, and Wright PC. 2012. An insight into iTRAQ: where do we stand now? *Analytical and bioanalytical chemistry* 404:1011–1027.

Farnsworth NR, Akerele O, Bingel AS, Soejarto DD, and Guo ZB. 1985. Medicinal Plants in Therapy. *World Health Organ* 63:965–981.

Ge N, Yang K, Yang L, Meng ZG, Li LG, and Chen JW. 2021. iTRAQ and RNA-seq analyses provide an insight into mechanisms of recalcitrance in a medicinal plant Panax notoginseng seeds during the after-ripening process. *Functional plant biology : FPB* 49:68-88.

Gong X, Jiang Q, Xu J, Zhang J, Teng S, Lin D, and Dong Y. 2013. Disruption of the rice

plastid ribosomal protein s20 leads to chloroplast developmental defects and seedling lethality. *G3 (Bethesda, Md)* 3:1769–1777.

Grabherr MG, Haas BJ, Yassour M, Levin JZ, Thompson DA, Amit I, Adiconis X, Fan L, Raychowdhury R, Zeng Q, Chen Z, Mauceli E, Hacohen N, Gnirke A, Rhind N, di Palma F, Birren BW, Nusbaum C, Lindblad-Toh K, Friedman N, and Regev A. 2011. Full-length transcriptome assembly from RNA-Seq data without a reference genome. *Nature biotechnology* 29:644–652.

Hou Y, and Jiang JG. 2013. Origin and concept of medicine food homology and its application in modern functional foods. *Food & function* 4:1727–1741.

Isaacson T, Damasceno CM, Saravanan RS, He Y, Catalá C, Saladié M, and Rose JK. 2006. Sample extraction techniques for enhanced proteomic analysis of plant tissues. *Nature protocols* 1:769–774.

Kim SK, Ryu SY, No J, Choi SU, and Kim YS. 2001. Cytotoxic Alkaloids from Houttuynia Cordata. *Arch Pharm Res* 24:518–521.

Li GZ, Chai OH, Lee MS, Han EH, Kim HT, and Song CH. 2005. Inhibitory Effects of Houttuynia Cordata Water Extracts on Anaphylactic Reaction and Mast Cell Activation. *Biol Pharm Bull* 28:1864–1868.

Lü D, Xu P, Hou C, Li R, Hu C, and Guo X. 2021. iTRAQ-based quantitative proteomic analysis of silkworm infected with Beauveria bassiana. *Molecular immunology* 135:204–216.

Lu H, Wu X, Liang Y, and Zhang J. 2006a. Variation in chemical composition and antibacterial activities of essential oils from two species of Houttuynia THUNB. *Chem Pharm Bull*

(Tokyo) 54:936-940. 10.1248/cpb.54.936

Lu HM, Liang YZ, Yi LZ, and Wu XJ. 2006b. Anti-Inflammatory Effect of Houttuynia Cordata Injection. *J Ethnopharmacol* 104:245–249.

Ng LT, Yen FL, Liao CW, and Lin CC. 2007. Protective Effect of Houttuynia Cordata Extract on Bleomycin-Induced Pulmonary Fibrosis in Rats. *Am J Chin Med* 35:465–475.

Pan L, Wan L, Song L, He L, Jiang N, Long H, Huo J, Ji X, Hu F, Fu J, and Wei S. 2022. Comparative Proteomic Analysis Provides New Insights into the Development of Haustorium in Taxillus chinensis (DC.) Danser. *BioMed research international* 2022:9567647.

Ramakrishnan V, and White SW. 1998. Ribosomal protein structures: insights into the architecture, machinery and evolution of the ribosome. *Trends in biochemical sciences* 23:208-212.

Rogalski M, Schöttler MA, Thiele W, Schulze WX, and Bock R. 2008. Rpl33, a nonessential plastid-encoded ribosomal protein in tobacco, is required under cold stress conditions. *The Plant cell* 20:2221–2237.

Smith PK, Krohn RI, Hermanson GT, Mallia AK, Gartner FH, Provenzano MD, Fujimoto EK, Goeke NM, Olson BJ, and Klenk DC. 1985. Measurement of protein using bicinchoninic acid. *Analytical biochemistry* 150:76-85.

Verma RS, Joshi N, Padalia RC, Singh VR, Goswami P, and Kumar A. 2017. Chemical Composition and Allelopathic, Antibacterial, Antifungal, and Antiacetylcholinesterase Activity of Fish-Mint (Houttuynia cordataThunb.) from India. *Chem Biodivers* 14:e1700189.

419 Wilson DN, and Doudna Cate JH. 2012. The structure and function of the eukaryotic ribosome.
420 *Cold Spring Harbor perspectives in biology* 4:a011536.
421 Wiśniewski JR, Zougman A, Nagaraj N, and Mann M. 2009. Universal sample preparation
422 method for proteome analysis. *Nature methods* 6:359–362.
423 Wu Z, Deng X, Hu Q, Xiao X, Jiang J, Ma X, and Wu M. 2021. Houttuynia cordata Thunb: An
424 Ethnopharmacological Review. *Frontiers in pharmacology* 12:714694.
425 Xu YW. 2012. Monoterpene secondary metablites in Houttuynia cordata Thunb. Doctor.
426 SICAU.
427 You C, Chen L, He H, Wu L, Wang S, Ding Y, and Ma C. 2017. iTRAQ-based proteome profile
428 analysis of superior and inferior Spikelets at early grain filling stage in japonica Rice.
429 *BMC plant biology* 17:100.
430 Yu HD, Yang XF, Chen ST, Wang YT, Li JK, Shen Q, Liu XL, and Guo FQ. 2012.
431 Downregulation of chloroplast RPS1 negatively modulates nuclear heat-responsive
432 expression of HsfA2 and its target genes in Arabidopsis. *PLoS genetics* 8:e1002669.
433 Zhan C, Li X, Zhao Z, Yang T, Wang X, Luo B, Zhang Q, Hu Y, and Hu X. 2016.
434 Comprehensive Analysis of the Triterpenoid Saponins Biosynthetic Pathway in
435 Anemone flaccida by Transcriptome and Proteome Profiling. *Frontiers in plant science*
436 7:1094.
437 Zhang P, Li C, Zhang P, Jin C, Pan D, and Bao Y. 2014. iTRAQ-Based Proteomics Reveals
438 Novel Members Involved in Pathogen Challenge in Sea Cucumber Apostichopus
439 japonicus. *PLoS ONE* 9:e100492.
440 Zhou K, Zhang C, Xia J, Yun P, Wang Y, Ma T, and Li Z. 2021. Albino seedling lethality 4;

Chloroplast 30S Ribosomal Protein S1 is Required for Chloroplast Ribosome Biogenesis and Early Chloroplast Development in Rice. *Rice (New York, NY)* 14:47.

Zhuang T, Li F, Huang LR, Liang JY, and Qu W. 2015. Secondary Metabolites from the Plants of the Family Saururaceae and Their Biological Properties. *Chem Biodivers* 12:194–220.

Fig.S1. the characterization of Transcriptome from *H. cordata.* (A) Length distribution of unigenes. (B) Venn map of annotated unigenes in Nt, Nr, KOG, Go, Pfam database. (C) number of unigenes in different annotated database.

Fig. S2. Overview of proteomics analysis (A). GO annotation classification includes biological process (BP), cellular component (CC), molecular Function (MF). (B). Top 30 KEGG pathway classification.

Fig. S3. Visualization of WGCNA result.