Food Chemistry: Molecular Sciences 8 (2024) 100187

Contents lists available at ScienceDirect

Food Chemistry: Molecular Sciences

journal homepage: www.sciencedirect.com/journal/food-chemistry-molecular-sciences

Profiles of phenolics and their synthetic pathways in *Asparagus officinalis* L.

Yuping Li [a], Sylvia E. Brown [a], Yunbin Li [a], Qin Cheng [a], He Wu [a], Shugu Wei [b], Xingyu Li [c,d], Chun Lin [a,c], Zhengjie Liu [a,c,*], Zichao Mao [a,c,d,*]

[a] College of Agronomy and Biotechnology, Yunnan Agricultural University (YNAU), Kunming, Yunnan 650201, China
[b] Industrial Crop Research Institute, Sichuan Academy of Agricultural Sciences, Chengdu, Sichuan 610023, China
[c] Institute of Improvement and Utilization of Characteristic Resource Plants, YNAU, Kunming, China
[d] The Laboratory for Crop Production and Intelligent Agriculture of Yunnan Province, Kunming, China

ARTICLE INFO

Keywords:
Asparagus officinalis L.
Phenolics biosynthesis pathway
Transcriptomics
Metabolomics

ABSTRACT

The synthetic pathways of some phenolics compounds in asparagus have been reported, however, the diversified phenolics compounds including their modification and transcription regulation remains unknown. Thus, multi-omics strategies were applied to detect the phenolics profiles, contents, and screen the key genes for phenolics biosynthesis and regulation in asparagus. A total of 437 compounds, among which 204 phenolics including 105 flavonoids and 82 phenolic acids were detected with fluctuated concentrations in roots (Rs), spears (Ss) and flowering twigs (Fs) of the both green and purple cultivars. Based on the detected phenolics profiles and contents correlated to the gene expressions of screened synthetic enzymes and regulatory TFs, a full phenolics synthetic pathway of asparagus was proposed for the first time, essential for future breeding of asparagus and scaled healthy phenolics production using synthetic biological strategies.

1. Introduction

Phenolics are ubiquitously present across green plants ranging from simple phenolic acids, polyphenols to their complex integrated polymers such as lignins and tannins (Cocuron et al., 2019). Phenolics are essential for plant development including: flower development (Ledesma-Escobar et al., 2018), senescence (Xi et al., 2016), extracellular linking (Brady & Fry, 1997), pathogens and herbivores resistance, UV protection and photodamage recovery (Tuominen & salminen, 2017). Phenolics also play important roles by acting as both underground (mycorrhizal symbiosis, actinorhizal symbiosis, plant growth-promoting rhizobacteria (PGPR)) and above ground part signaling molecules (Desmedt et al., 2021). Although the functions of the majority of phenolics are structural components of cell walls, a vast array of phenolics are the toxins and antifeedants for plant defense, coloring agents and aroma chemicals of flowers, fruits, seeds and woods. Since the origin of green plants can be traced from the ocean, phenolics as structural components in plant cell play crucial importance for green plants support, water and inorganic nutrient transportation via vascular system from soils. Other functions such as signal molecules and enzymatic activities regulations are essential for the continued survival and flourishment of vascular plants especially angiosperms.

Asparagus is a well-known healthy vegetable cultivated worldwide for thousands of years, enrichened with nutrient components and the accumulation of bioactive medicinal compounds, such as phenolics and steroids (Cheng et al., 2023). Flavonoids, the major group of active phenolics compounds in asparagus, possesses the potential of regulating multiple processes such as: apoptosis, angiogenesis, tumor differentiation, and cell proliferation due to its modulation of cancer related kinase activities. Flavonoids have additional effects of improving cardiovascular disease treatments, antidiabetic as well as having antiradical efficacy (Liu et al., 2017). The antiradical efficacy of phenolics involves direct scavenging of radicals by upregulating and activating enzymes such as: catalase, peroxidase and superoxide dismutase (Renuka & Arumughan, 2007). Other activities of phenolics includes: enzyme activities, regulation of human diseases related pathways, such as protein kinase C, phosphoinositide-3-kinases, protein tyrosine kinase, cell-cycle arrest at G1/G2 phase, apoptosis induction and rehabilitation of cellular homeostats (Bhosale et al., 2022, Pang et al., 2022, Rana et al., 2022, Shao et al., 2022). Other studies have reported that even simple phenolics such as coumarins were observed to have lipopolysaccharide-induced nitric oxide inhibition, antiproliferative and anti-leishmanial properties (Fernandez-Pena et al., 2022, Santibanez et al., 2022).

In studying phenolics profiles, their biosynthesis and regulations in

* Corresponding authors at: College of Agronomy and Biotechnology, Yunnan Agricultural University (YNAU), Kunming, Yunnan 650201, China.
E-mail addresses: lzj1022@163.com (Z. Liu), zmao@ynau.edu.cn (Z. Mao).

https://doi.org/10.1016/j.fochms.2023.100187
Received 8 October 2023; Received in revised form 30 November 2023; Accepted 9 December 2023
Available online 14 December 2023

garden asparagus, Liu et al (2022) reported, a total of 1045 metabolites annotated in asparagus in which flavonoids and phenolic acids accounted for 37.51 % of the total in three green and one white asparagus cultivars at different harvesting stages. Ma et al (2023) reported, lignins and their simple phenolic compounds such as cinnamic acid, caffeic acid, ferulic acid and sinapyl alcohols, had reduced contents due to downregulation of expressions in their biosynthetic genes with shading treatments. By using transcriptomic strategies, the biosynthetic pathway of rutin (one of major asparagus flavonoids) was proposed (Yi et al., 2019). With similar strategies, anthocyanins biosynthetic pathway was also proposed (Dong et al., 2019). As showed in their published results, the key synthetic genes and regulatory TFs related to biosynthesis of phenolics lacks the use of the published genome of asparagus, and/or full length transcripts data, therefore, it is still necessary to analyze the contents and profiles of phenolics and to screen out their key synthetic and regulated genes at transcriptional isoforms levels with the available genome of *A. officinalis* L. (Harkess et al., 2017).

In this study, both metabolomic and transcriptomic strategies were applied to green "Guelph Millennium" and purple "Purple Passion" cultivars respectively to analyze their profiles, distribution of phenolics and other excluded phenolics nutrient elements which aims to i) define a suitable harvesting time of asparagus vegetables and/or processing for phenolics purposes used in new healthy food development; ii) screen key genes for phenolics synthesis and regulatory TFs ; iii) propose a full biosynthetic pathway with its regulation networks in garden asparagus at their transcriptional isoform levels.

2. Materials and methods

2.1. Plant materials

The green (G) and purple (P) asparagus cultivars were planted in the field of Yunnan Agricultural University. The three years roots (Rs), spears (Ss) and flowering twigs (Fs), labeled as GRs, GSs, GFs, PRs, PSs and PFs respectively, were sampled with 3 biological replications and store at −80℃, or used directly for detection of genes expressions and their isoform transcripts using methods of Illumina RNAseq and PacBio ISOseq, and metabolites analysis using UPLC-MS-MS respectively.

2.2. Illumina RNAseq and data processing

RNAs were extracted for RNAseq by Illumina Nova Seq 6000 platform, quality control of obtained raw data (Supplementary files1 Table S1) were analyzed with FastQC (de Sena Brandine & Smith, 2019) then processed by Fastp (https://github.com/OpenGene/fastp) to get clean data. The clean reads were mapped to the reference genome of *A. officinalis* downloaded from Phytozome (https://phytozome-next.jgi.doe.gov/) using Hisat2 (Kim et al., 2015), and genes expression quantification were performed using scripts of Featurecounts R (Liao et al., 2014) to obtain the matrix of both counts and TPM (normalizing expression unit of transcripts per kilobase of exon per million mapped fragments) of genes. Gene expression were analyzed based on TPM matrix, while the differential expression genes (DEGs) were analyzed by DEseq2 (Love et al., 2014) with the resulted counts matrix of genes. The predicted genes functions from the asparagus were annotated using eggNOG (https://eggnog-mapper.embl.de/). The expression of TPM matrix of 18 asparagus samples were filtered by median absolute deviation (MAD) and used to perform co-expression analysis by WGCNA for co-expressed gene modules (MEs) (Langfelder & Horvath, 2008). Principal components analysis (PCA) were performed using the expression matrix (TPM) of all genes, abundance of phenolics metabolites and additional non phenolics nutrients (ANu) by PCAtools (Groth et al., 2013), respectively. GO and KEGG enrichment of differential expression genes (DEGs), MEs or intersection of interested DEGs sets and MEs correlation to phenolics synthesis were performed using ClusterProfiler (Wu et al., 2021). The DEGs sets were visualized with the R package UpSetR (https://github.com/hms-dbmi/UpSetR).

2.3. PacBio IsoSeq and data processing

To optimize asparagus genome annotation and detection of the possible alternatives exon splicing and genes function of the phenolics synthetic pathway (PSP), full length transcripts were detected by ISOseq with the RNAs extracted from the Rs, Ss and Fs samples of purple passion male, female and adromonoecious respectively. PacBio raw data (Supplementary file1 Table S2) obtained were analyzed by IsoSeq3 pipelines (https://github.com/ErisonChen/IsoSeq), and the polished subreads were generated by CCS (https://github.com/PacificBiosciences/ccs) with a minimum quality of 0.9 and minimum number of full subreads (FSRs) ($n \geq 3$). Full length transcripts (FL) were obtained when the sequences had poly (A), both 5′ and 3′ cloning primers with Lima scripts (https://github.com/fluffos/lima), then the FLs with non-chimeric transcripts (FLNC) were obtained by Isoseq3 refine scripts. The FLNC reads were then mapped to the *A. officinalis* reference genome using both minimap2 (Li, 2018) and desalt (Liu et al., 2019) respectively. The obtained BAM files were sorted and further used to collapse redundant isoforms using tama (https://github.com/GenomeRIK/tama). The annotated gene transfer format (GTF) files from both annotated genome and GTF results from IsoSeq data, *A. officinalis* genome and short Illumina sequences of Rs, Ss and Fs were used to annotate the genes at transcriptional isoform levels by SQANTI3 (Tardaguila et al., 2018). Transcript isoforms including alternative splicing (AS) and the frequencies of the AS events including intron retaining (IR), exon skipping (ES), alternative of both 5′ and 3′ splicing, were defined as partially overlapping with exons/introns of an annotated gene, whereas fusion transcripts span two annotated gene loci with SQANTI machine learning classifier methods. ORFs from the corrected transcripts were predicted by GeneMarkS-T (Tang et al., 2015), while the transcripts lacking ORFs or harboring ORFs less than 50 amino acids were aligned to all miRNA databases of miRbase6 (release 22.1) and Rfam (14.4) database to identify possible pre-miRNAs and other non-coding RNAs. The matched and remaining transcripts which were more than 200 bp were assessed for coding potential using CPC (CPC score ≤ 1) (Kong et al., 2007). The resulting transcripts were considered as putative lncRNA transcripts. The obtained optimized coding genes were further annotated with EggNOG Database (https://eggnog5.embl.de/) and the TFs and the promoter binding sites of target gene(s) were annotated by PlanTFDB database (https://planttfdb.gao-lab.org/).

2.4. Metabolic compounds detection

The fresh Rs, Ss and Fs samples of asparagus were freeze-dried using a lyophilizer (Scientz-100F, Ningbo, China) and crushed to powder. ~ 100 mg powder was extracted overnight using 1 mL of 70 % methanol at 4℃ and centrifuged at 10,000 x g for 10 min. The resulted supernatants were applied to an UPLC-ESI-MS/MS system (www.shimadzu.com.cn/ and www.appliedbiosystems.com.cn/) equipped with Waters column of ACQUITY UPLC HSS T3 C18 (1.8 μm, 2.1 mm*100 mm). The analytic mobile system, consists of solvent A (water with 0.04 % acetic acid) and solvent B (acetonitrile with 0.04 % acetic acid). The separation of metabolites was performed with a linear gradient program of elution starting from 95 % A, 5 % B to 5 % A, 95 % B within 10 min, then a composition of 5 % A, 95 % B kept for 1 min. Subsequently, a composition of 95 % A, 5 % B were adjusted within 0.5 min and kept for 5 min with an injected volume of 2 μl extracted samples. The integrated and corrected chromatographic peaks of each metabolite were corrected with MultiaQuant version 3.0.3 (Sciex, Darmstadt, Germany). Based on the MetWare metabolism self-built DataBase (MWDB) (https://cloud.metware.cn/), the qualitative analyses of substances were carried out according to the secondary mass spectrometry information removing the isotopic and repeated noise signals such as K^+, Na^+, and NH_4^+. The multiple reaction monitoring (MRM) mode of triple quadruple mass

spectrometry quantified the metabolites (Fraga et al., 2010) using Analyst 1.6.3 (AB Sciex) with default parameters. Quality control samples were prepared by mixing sample extracts and analyzing the repeatability using the same methods. Significantly different metabolites were further determined by R packages MetaboAnalystR (https://github.com/xia-lab/MetaboAnalystR/) and ropls (https://www.bioconductor.org/packages/release/bioc/html/ropls.html), judged by critical standard VIP ≥ 1 and absolute Log2FC (fold change) ≥ 1. Different metabolites (DMs) sets were also visualized by the R package UpSetR (https://github.com/hms-dbmi/UpSetR).

2.5. Determination of total flavonoids contents (TFCs)

The TFCs were determined by sodium nitrate-aluminum and nitrate-sodium hydroxide colorimetry method (Song et al., 2022) using 4 biological duplicates of Rs, Ss and Fs of both green (G) and purple (P) asparagus cultivars. Rutin was selected as the standard chemical to obtain a linear regression equation, $y = 1.0576x + 0.0002$, with $R^2 = 0.9998$, using a wavelength of 510 nm for quantification.

2.6. Genes expression analysis using qRT-PCR assay

The RNAs were extracted using the RNA Easy Fast Plant Tissue Kit, reverse transcriptions to cDNA using the Fasting RT Kit, and qRT-PCR were analyzed using SuperReal PreMix Plus kit with SYBR Green respectively (Tiangen Biotech, Beijing, China). The PCR primers were designed based on the fine annotated genome of asparagus following chemical synthesis from Shengong Biotech (Shanghai, China) as listed in supplementary file1 Table S3.

3. Results and analysis

3.1. The TFCs, phenolics and ANus profiles in garden asparagus

The TFCs in Rs, Ss and Fs of green and purple (Fig. 1A ~ 1F) asparagus had no significant differences between the Rs in both cultivars of asparagus, however, the TFCs of top parts (Ts, consisting of Ss and Fs) of purple asparagus were significantly higher than that of the green asparagus counterparts. Within cultivars, TFCs of Fs > Ss > Rs (Fig. 1G). A total of 437 metabolites were detected in 6 samples, in which 204 were phenyl compounds including 105 flavonoids, 82 phenolic acids, 13 lignans/coumarins, 2 quinones and 2 tannins with different metabolites (Supplementary file1 Fig. S1 and Supplementary file 2). The most abundant phenolics types of different metabolites (DMs) were flavonoids which are richened in the Ts of asparagus (Supplementary file2 and Supplementary file1 Fig. S1 and Fig. S3A) which is consistent with the TFCs determined with the spectrophotometric assays (Fig. 1).

Further PCA analyses showed that, the GFs and GSs of phenolics overlapped while that of PFs, PSs, PRs, GRs were separated from each other, and they separated from the overlapped GFs and GSs as well in the asparagus cultivars (Supplementary file1 Fig. S2A). The PCA results of ANus showed that aside GRs and PRs, ANus of other organs including GSs, GFs, PSs and PFs were separated from each other and also separated from overlapped ANus of GRs and PRs. Based on these results, the different metabolites including phenolics (DPs) and ANus (DANus) of Rs vs Ss, Rs vs Fs (GRs vs GSs, GRs vs GFs, PRs vs PSs and PRs vs PFs) within cultivars, and Fs and Ss (GSs vs PSs and GFs vs PFs) among cultivars were chosen for further DMs analyses. The results showed that the intersect of DMs between GR vs GS and GR vs GF were united with intersect of DMs between PRs vs PSs and PRs vs PFs, resulting in R_T_DMs set with 350 DMs, in which 180 are DANus, while 170 are DPs including 64 phenolic acids, 40 flavanols, 20 flavones, 9 flavone carbonosides, 7 anthocyanins and 30 other phenolic compounds. To detect the DPs and DANus between Ts of green (GTs) and purple (PTs) asparagus cultivars, the union of DMs of GS vs PS and GF vs PF (GT_PT_DMs) were analyzed, the results show 268 DMs in which 118 are DPs and 150 DANus, including; amino acids & derivatives (37), lipids (25), nucleotides & derivatives (22), organic acids (20), sugars and alcohols (18), vitamins (6), alkaloids (4) and terpenoids (8) (Supplementary file1 Fig. S2D ~ S2F, Fig. 2, Fig. 3 and supplementary file1 Fig. S3). Interestingly a total of 26 DMs (i.e. 22 phenolics and 4 ANus) were found to have intersected DMs sets amongst GRs vs GSs, GRs vs GFs, PRs vs PSs and PRs vs PFs, while 13 DMs (1 phenolics and 12 ANus) were found to have intersected metabolites sets between GSs vs PSs and GFs vs PFs (Supplementary file1 Fig. S2). Further comparison of the DPs between GSs and PSs, which are the commercially eatable vegetable parts of asparagus, showed to have 24 compounds of phenolics, i.e. C009 (naringenin chalcone), C010 (fisetin), C013 (hesperetin), C022 (kaempferol), C023 (kaempferol-7-O-glucosdide), C024 (trifolin), C033 (quercetin), C035 (isorhamnetin), C055 (quercetin-3-O-methyl ether), C062 (7-O-methxyl quercetin), C064 (juglanin), C076 (5-hydroxy-7-methoxyflavone), C077 (luteolin-7-O-glucoside), C081 (hesperetin-7-O-glucoside), C092 (Luteolin-6-C-2-glucuronylglucoside), C098 (Tricin), C103 (Prunetin), C128 (Caffeic acid), C147 (Salvianolic acid F), C148 (Benzyl β-D-Glucopyranoside), C171 (3,4′-Dihydroxy-3′-methoxybenzene pentanoic acid), C176 (tuberonic Acid), C272 (syringaresnol-4-O-β-D-apiofuranoside) and C407 (6-hydroxyrumicin-8 anthocyanins-O-D-glucopyranoside) were higher in GSs than in PSs. However 5 anthocyanins, i.e. C001 (cyanidin 3-O-glucoside), C004 (pelargonidin-3-O-rutinoside), C005 (cyanidin 3-O-rutinoside), C006 (peonidin 3,5-O-diglucoside chloride) and C007 (cyanidin-3-O-rutinoside-5-O-glucoside) and 1 Flavanol (C019 (epicatechin)) and 1 chalone (C008 ((E)-chalcone) were lower in GSs than in PSs of asparagus which is consistent with its purple color among purple cultivars and higher TFCs in their Ts (PSs and PFs) (Fig. 1, Fig. 2 Supplementary file1 Fig. S1).Interestingly 3 phenolic acid, C155 (β-D-furanofructosyl-α-D-(6-mustardacyl) glucoside), C118 (calceorioside B) & C142 (4-hydroxybenzoic acid) and 2 flavonoids including; C078 (diosmetin-6-C-glucoside) & C034 (quercetin-7-O-(6′-O-malonyl)-β-D-glucoside) were higher in all organs of purple asparagus cultivars as compared to its green counterparts. Within green asparagus cultivars, aside a few phenolics contents such as: C005 (cyanidin 3-O-rutinoside), C148 (Benzyl β-D-Glucopyranoside) and C103 (prunetin), other phenolics compounds are either non-significantly different or higher in GFs than in GSs (Fig. 2). When comparing the DPs in PSs and PFs, 17 phenolics including: C009 (naringenin chalcone), C010 (fisetin), C013 (hesperetin), C022 (kaempferol), C023 (kaempferol-7-O-glucosdie), C024 (trifolin), C033 (quercetin), C055 (quercetin-3-O-methylether), C062 (rhamnetin (7-O-methxyl quercetin), C064 (juglanin), C077 (luteolin-7-O-glucoside), C081 (hesperetin-7-O-glucoside), C098 (tricin), C123 (anthranilic acid), C171 (3,4′-dihydroxy-3′-methoxybenzenepentanoic acid), C176 (tuberonic acid), and C407 (6-hydroxyrumicin-8-O-D-glucopyranoside) were higher in PFs than in PSs. While 9 phenolics including: C021 (gallocatechin-Gallocatechin), C036 (kaempferol-3-O-neohesperidoside), C053 (6-hydroxykaempferol-3,6-O-diglucoside), C059 (isorhamnetin-O-hexoside-O-hexoside-O-pentoside), C070 (isovitexin), C072 (isoorientin), C073 (orientin), C074 (vitexin-2-O-D-glucopyranoside) and C090 (luteolin-6,8-di-C-glucoside) were lower in PFs than in PSs (Supplementary files1 Fig. S1).

For both GSs and PSs, which are harvested spears as vegetables, ANus including steroidal saponins are also important in both green and purple asparagus. Our previous reports (Cheng et al., 2023) are consistent with our current findings which shows that steroids are richened in the roots and not in the eatable parts of asparagus (both GSs and PSs). In the comparison of ANus among GSs vs PSs, excluding some lipids, modified amino acids and peptides and nucleosides such as: C191 (5-hydroxy-L-tryptophan), C310 (MAG (18:4) isomer1), C288 (9-KODE), C354 (uracil), C222 (N-(3-indolylacetyl)-L-alanine) and C225 (N-acetyl-DL-tryptophan) were higher in PSs than PFs, most detected ANus are without significant differences between GSs and PSs indicating major different metabolites between GSs and PSs are phenolics but not ANus. Further analysis showed that in green asparagus, compounds of ANus including: C299 (lysoPC 17:0), C312 (palmitoleic acid), C377

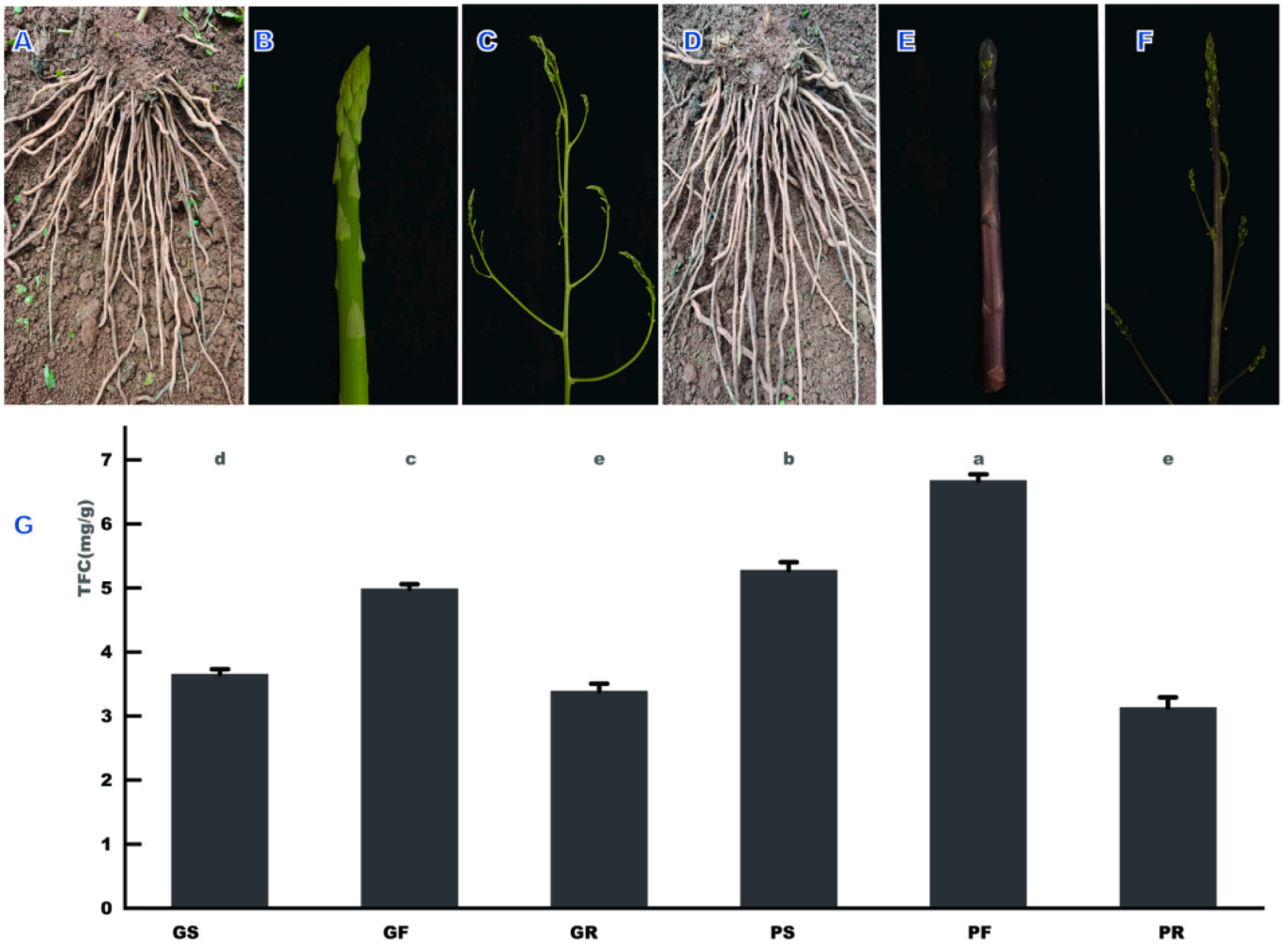

Fig. 1. Flavonoids metabolites in both asparagus cultivars; **A** ~ **F**, samples of roots (Rs), spears (Ss) and flowering twigs (Fs) of green (G) and purple (P) asparagus, respectively; **G**, total flavonoids content (TFC) in Rs, Ss and Fs of green and purple asparagus. Bars which do not share the same letters are significantly different with p-values < 0.05. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

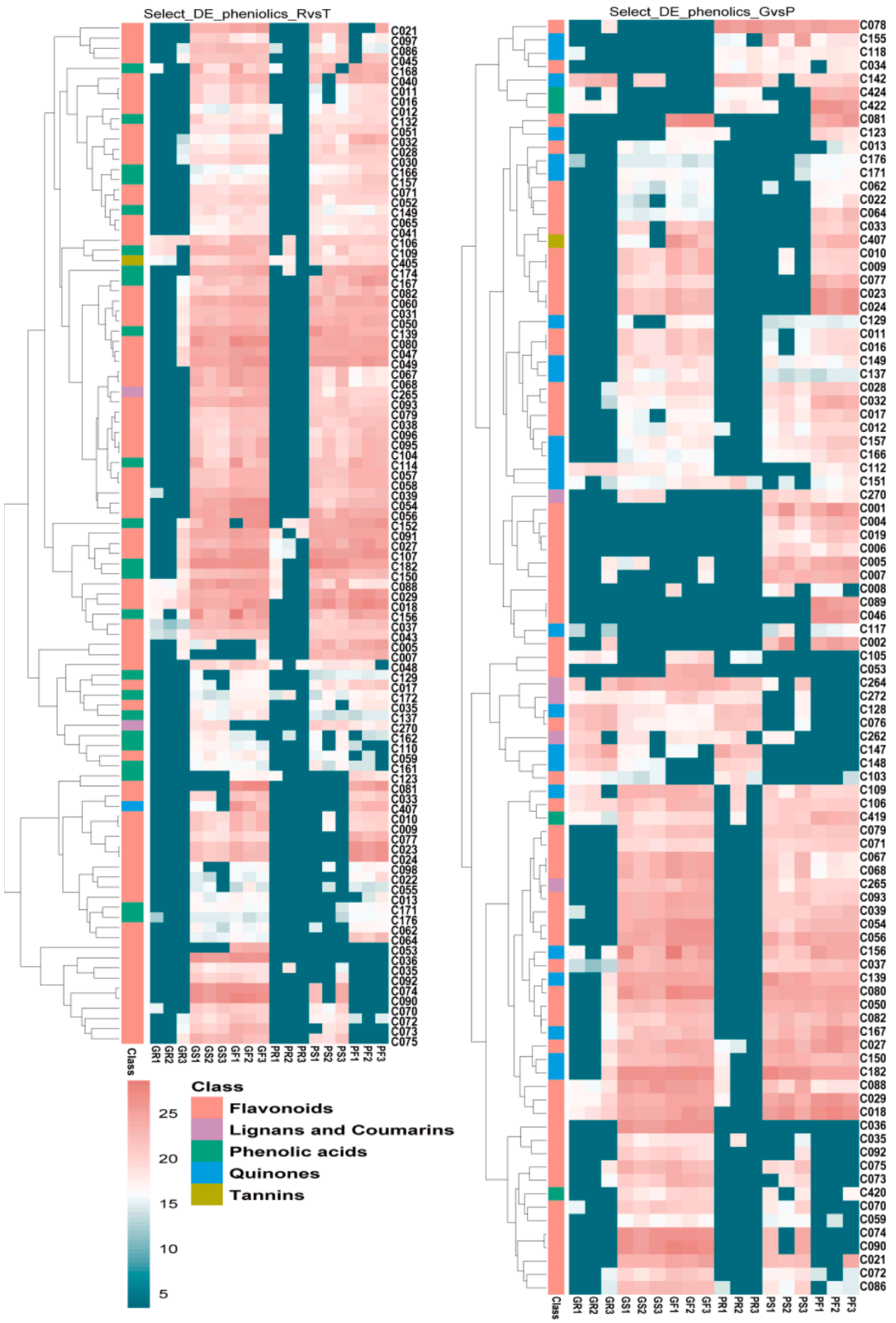

Fig. 2. Heatmaps of abundance selective differential phenolics accumulation analysis of **A**: Phenolics metabolites between Rs vs Ts; **B**: phenolics metabolites between green and purple cultivars; note, the names and structures of corresponding assigned chemical numbers are shown in Supplementary file1 Fig. S1, full set of differential phenolics showed in supplementary file1 Fig S3 A. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

Fig. 3. Heatmaps of abundance selective differential accumulation of additional non-phenolics nutrients (ANus) analysis of **A**: differential ANus between Rs vs Ts; **B**: differential ANus between green and purple cultivars; note, the names and structures of corresponding assigned chemical numbers are shown in Supplementary file1 Fig. S1, full set of differential ANus showed in Supplementary file1 Fig. S3 B. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

(maltotetraose), C366 (D-galacturonic acid), C389 (trehalose 6-phosphate), C429 (β-pseudouridine) were lower in GFs than GSs, while only two ANus compounds C192 (5-oxoproline) and C254 (4-pyridoxic acid) were lower in GSs than in GFs (supplementary file1 Fig. S5). Comparing PSs and PFs samples, ANus including: C320 (aromadendrin-7-glucoside), C420 (N1-dihydrocaffeoyl-N10-coumaroyl spermidine), S_C09 (sweroside), C255 (4-pyridoxic acid O-hexoside), C301 (lysoPC 18:2(isomer)), C305 (lysoPE 18:1 (2n isomer)), C235 ((S)-(-)-2-hydroxyisocaproic acid) were higher in PSs than in PFs, while ANus such as C384 (mannitol), C424 (p-coumaroylputrescine), C422 (N-feruloylputrescine), S_C05 (diosgeninrha-glc-glc), S_C06 (Pseudoprotodioscin), S_C13 (podecdysoneC), C238 (2-methylglutaric acid), C237 (2-aminoethanesulfinic acid), C298 (lysoPC 16:2(2n isomer), C344 (cordycepin), C241 (3-hydroxybutyrate), C346 (deoxyguanosine), C191 (5-hydroxy-L-tryptophan), C310 (MAG(18:4)isomer1), C288 (9-KODE), C354 (uracil), C428 (α-linolenic Acid), C218 (L-pyroglutamic acid) and C290 (choline alfoscerate) were higher in PFs than in PSs. Interestingly 2 steroidal saponins (i.e. S_C05 and S_C06) were higher in PFs than commercial parts of both GSs (green) and PSs (purple) asparagus (supplementary files1 Fig. S2B), indicating PFs may have more phenolics as well as steroidal saponins which is normally accumulated in Rs of asparagus.

Interestingly, there are 16 phenolics including; C020 (Gallocatechin), C063 (3,7-Di-O-methylquercetin), C069 (Apigenin-8-C-glucoside), C083 (Rhoifolin), C102 (Demethyltexasin), C103 (Prunetin), C127 (Caffeoyl-p-coumaroyltartaric acid),C141 (Protocatechuic aldehyde), C142 (4-Hydroxybenzoic acid),C147 (Salvianolic acid F), C148 (Benzyl β-D-Glucopyranoside), C164 (3-Hydroxy-4-isopropylbenzylalcohol 3-

glucoside), C261 (1-Hydroxyterpinin monO-glucoside), C268 (Matairesinol), C316 (Ethyl caffeate) and C003 (Peonidin 3-O-(6′′-malonylglucoside) found with higher accumulation and some of them (e.g. C141, C127, C102 and C316) specifically accumulated in Rs of asparagus (Supplementary file1 Fig. S1 and Fig. S3).

3.2. Different expression genes analysis

A total of ~1.3 G reads with average length of ~150 bp ranging from 43 ~ 110 M reads for each sample, were obtained by quality control from raw Illumina RNAseq data from 6 samples of Rs, Ss and Fs in both green and purple asparagus cultivars with 3 replications. A total of 192 Gb clean bases with high average quality (Q) value of 35.9 were obtained (supplementary file1 Table S1). The mixture of RNAs of Rs, Ss and Fs of male, female and andromonoecious asparagus "purple passion" were further sequenced with PACBIO RSII platform to obtain 43.98, 16.98 and 28.58 Gb Pacbio raw data respectively (supplementary file1 Table S2). The raw reads of insertion (ROI) were processed with IsoSeq3 pipelines to get 0.68, 0.33 & 0.59 M circular consensus sequence (CCS), which resulted in 0.6, 1.3 and 0.23 M CCS and isoforms of male, andromonoecious and female purple asparagus with an average length of 2.4, 2.0 & 1.7 kb respectively (Supplementary file 1 Table S2). The Pacbio Iso-Seq data were used for optimizing annotation of the detected synthetic and regulatory phenolics genes in asparagus, and detection of their possible exon alternative splicing (ASs) and genes fusions as shown in supplementary file 4 sheet 1 ~ 4. The PCA results of genes expression (based on TPM matrix) from Illumia RNAseq data showed that Rs, Ss and Fs clustered together and separated from each other within and between both asparagus cultivars (Supplementary file1 Fig. S2 C). The PCA results of gene expressions are similar with the PCA results of abundance detected in phenolics and ANu compounds (Supplementary file1 Fig. S2A and S2B), indicating that, the DMs are resulted from different expression genes (DEGs) of organs in both asparagus cultivars. Judged by detected chemical profiles and abundance, the most significantly different compounds are phenolics with higher accumulation concentration in the Ts of asparagus, especially in PFs of purple asparagus (Supplementary file1 Fig. S3 and Fig. 2) suggesting that, there is a significant difference in nutrient value among the purple and green asparagus which are mainly due to more accumulation of phenolics especially anthocyanins derived from aglycones of peonidin, cyanidin and pelargonidin which are consistent with the deep purple color observed in both PSs and PFs of purple cultivars (Fig. 1E &1F). Based on these phenotypes, the key genes for phenolics biosynthesis pathway and its regulators (TFs) are focused on different expression genes (DEGs) between Rs vs Ts (including GRs vs GSs, GRs vs GFs, PRs vs PSs and PRs vs PFs) in both green and purple asparagus (Supplementary file1 Fig. S2), while the key genes of anthocyanin synthesis and regulation are focused on DEGs between GTs and PTs. The DEGs analyses results showed that, 18127, 18091, 18899, 19963, 19007and 20,048 DEGs (selected by critical standards of absolute log (Fold changes) > 1 and P values < 0.05) were detected between GRs vs GSs, GRs vs GFs, PRs vs PSs, PRs vs PFs, GSs vs PSs and GFs vs PFs respectively. By using these detected DEGs sets, the intersected DEGs between GRs vs GSs & GRs vs GFs, PRs vs PSs & PRs vs PFs were obtained to result in DEGs sets of both GRs vs GTs and PRs vs PTs with 6740 and 6647 DEGs respectively (Supplementary file1 Fig. S2G). The DEGs set of GTs vs PTs (named GT_PT_DEGs) with 20,384 DEGs, and DEGs set of Rs vs Ts (named R_T_DEGEs) with 19,671 DEGs were obtained by uniting the DEGs sets of GSs vs PSs & GFs vs PFs and GRs vs GTs & PRs vs PTs respectively (Supplementary file1 Fig. S2 G II). Interestingly the intersect between GSs vs PSs & GFs vs PFs and GRs vs GTs & PRs vs PTs resulted in 731 and 314 DEGs sets respectively (Supplementary file1 Fig. S2 I).

3.2.1. Screening of key genes for biosynthesis and regulation of phenolics metabolites

To screen the key genes for the biosynthesis and regulation of phenolics, the gene TPM matrix of the 18 detected RNAseq data by co-expression network analysis strategies were performed. The results showed that all expressed genes were aggregated in 19 co-expressed gene modules (MEs) with 19,689 genes marked by different colors (Supplementary file1 Fig S4), and MEs correlated to the abundance of phenolics metabolites were further analyzed. The results showed that five modules labeled by colors blue (with 3722 genes) brown (2932), red (1224), magenta (603), and salmon (211) were selected with a critical standard of mean absolute correlation value ≥ 0.5 and p value ≤ 0.05 (Supplementary file1 Fig. S4).

The genes in these five modules were united to get phenolics related gene modules (PMEs) set with a total of 8692 genes, then PMEs intersected with DEGs of different groups (Rs vs Ts and GTs vs PTs) were selected for further KEGG and GO genes enrichment analysis. The GO enrichment results of intersects between PMEs and Rs vs Ts DEGs sets (5067 genes) showed that these genes were enriched in pigment binding, antioxidant activity, oxidoreductase activity etc, while the KEGG enrichment analysis showed that these genes were mainly enriched with flavonoid biosynthesis, phenylpropanoid biosynthesis and drug metabolism-cytochrome P450 pathways which are essential for the biosynthesis of phenolics metabolites (Supplementary file1 Fig. S5). Likewise, the GO enrichment analysis results of intersects of PMEs between DEGs sets of GTs vs PTs (with 3528 genes) showed that these genes were mainly enriched in pathways related to cell wall synthesis, phenylpropanoid biosynthesis, glycosyltransferase and glycosyl-hydroxylase related activities, while the KEGG enrichment result analysis showed that these genes were enriched with phenylpropanoid biosynthesis, amino acids and nucleotide sugar metabolism, stilbene diarylheptanoid and gingerol biosynthesis which are also related to the biosynthesis of phenolics metabolites (Supplementary file1 Fig. S6). Based on the enrichment analyses of both GOs and KEGGs, the key candidate genes for the biosynthesis and regulation of phenolics metabolites were collected by combing both KEGG and GO rich pathways related phenolics biosynthesis and regulation from the above richened pathways, which resulted in the final phenolics synthesis and regulatory gene sets (with 114 genes) which included coding genes for CYP450 type dioxygenase/ hydroxylase (P450s), 2-oxoglutarate-dependent dioxygenases (OGDs), glycosyltransferase (GTs), glycosyl-hydroxylases (GHs), amino transferases (ATs), acetyltransferases (ACTs), methyltransferases (MTs/OMTs) families as well as genes of TFs such as MYBs, bHLHs, WD40s, WRKYs and AP2/ERFs (Supplementary file 4, Fig. 4). At the same time, the resulted protein sequences of enzymes and TFs for phenolics synthesis and regulation in asparagus have been further characterized by homologs searched with NR (https://www.ncbi.nlm.nih.gov/), UniProt database (https://www. uniprot.org/) and/or protein sequences of functionally confirmed phenolics synthetic enzymes and regulatory TFs in other plants by BLAST analysis. The total 114 selected key candidate genes consisting of 97 synthetic genes and 17 TFs genes, were further optimizely annotated at isoform transcript levels with ISOseq data as described in materials and methods. The 114 genes expressions in different organs of asparagus were demonstrated by heatmaps as shown in Fig. 4A and their information of transcriptional isoforms and expressions in different sampled asparagus organs are listed in supplementary file 4 sheet 1 ~ 4. For a convenience, we definite pathways of phenolics synthesis for simple phenols (including phenolic acids) and their resulted polymers i.e. lignins as phenylpropanoid pathway (PPS), in which before p-Coumaroyl-CoA is referred to as the upstream pathway (UPPS), while after p-Coumaroyl-CoA, is referred to as downstream pathway (DPPS). Similarly, we describe all flavonoids including their resulted polymers, i.e. tannins synthesis pathway as the phenylpropanoid-acetate pathway (PPAS), and the pathway before and after chalcone were referred to as upstream (UPPAS) and downstream (DPPAS) respectively. From ISOSeq data a total 23 genes related to phenolics synthesis and regulation in which 3 TFs (*bHLH_3/4* and *WD40_1*) were detected with exon alternative splicing (ASs), while 5 genes (*V1_05.3379|AROH_2, V1_07.1354|CHI_2, V1_10.1948|PAL_4,*

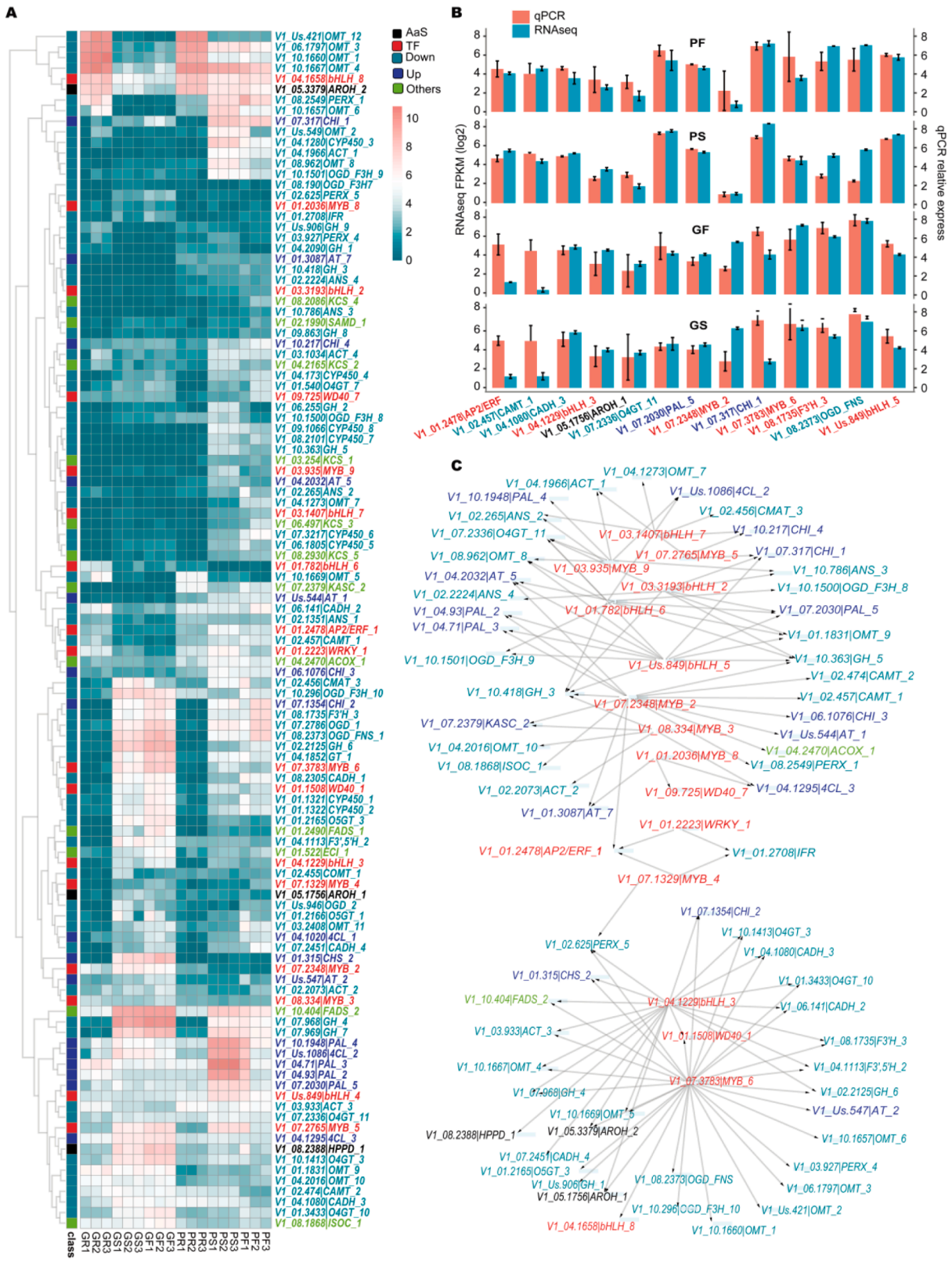

Fig. 4. Identification of key genes in phenolics metabolites biosynthesis pathway; **A**: heat map analysis of selected candidate genes related to the biosynthesis and regulation of phenolics metabolites in the Rs, Ss and Fs of both green and purple asparagus cultivars respectively; **B**: qPCR analysis of the selected key genes and TFs of phenolics biosynthesis and regulation in asparagus; **C:** the regulatory network based on gene expression of TFs and phenolics genes in asparagus, where black, red, green, blue and lighter green fonts represent amino acid synthesis genes (AaS), TFs, downstream genes, upstream genes and other genes from enrichment analysis respectively. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

V1_10.418|GH_3 and *V1_Us.1086|4CL_2*) were detected with gene fusion (Supplementary file1 sheet 1 ~ 3). The results of transcript isoforms detection indicate that, both ASs and fusion of phenolics synthetic and regulatory genes also contribute to regulating the phenolics accumulation in asparagus.

3.2.2. Asparagus roots phenolics biosynthesis related genes

As described in 3.1, the major phenolics were mainly accumulated in Ts of asparagus, which may be due to most phenolic synthetic genes being positively regulated by light signaling, however, 16 phenolics accumulated higher in roots of asparagus. The gene expression results showed that, 4 TFs genes (*V1_01.2223|WRKY_1*, *V1_01.2478|AP2/ERF_1*, *V1_01.782|bHLH_6*, *V1_04.1658|bHLH_8*), 1 aromatic amino acid synthesis (AaS) gene (*V1_05.3379|AROH_2*), an AT (*V1_Us.544|AT_1*), a *CHI* (*V1_06.1076|CHI_3)* and 3 DPPS genes (*V1_04.2470| ACOX_1*, *V1_06.141|CADH_2*, *V1_02.457|CAMT_1*), 6 methyltransferases (OMTs) (*V1_10.1660|OMT_1*, *V1_06.1797|OMT_3*, *V1_10.1667|OMT_4*, *V1_10.1669|OMT_5*, *V1_10.1657|OMT_6*, *V1_Us.421|OMT_12*), taking part in both DPPS and DPPAS modification of methylation, and a peroxidase gene (*V1_08.2549|PERX_1*) playing roles in lignin and tannin biosynthesis through their monomers linking by oxidation. These listed genes have higher expressions in Rs of asparagus indicating that, they are related to phenolics accumulation in Rs of asparagus. (Supplementary Fig. S3A and Fig. 4).

3.2.3. The genes related to anthocyanins biosynthesis in asparagus

The results showed that, the DMs of both GSs vs PSs and GFs vs PFs of 5 anthocyanins (C001, C00, C4, C005, 006, C007) had higher accumulation in PTs. The genes expression analyses showed that, 35 genes including 4 regulatory TFs (*V1_09.725|WD40_7*, *V1_03.935|MYB_9*, *V1_03.1407| bHLH_7 and V1_03.3193|bHLH_2*), an AT (*V1_04.2032| AT_5)* 2 CHIs (*V1_07.317|CHI_1*, *V1_10.217|CHI_4*), 6 CYP450s (*V1_04.1280|CYP450_3*, *V1_04.173|CYP450_4*, *V1_06.1805|CYP450_5*, *V1_07.3217|CYP450_6*, *V1_08.2101|CYP450_7*, *V1_09.1066|CYP450_8*), 2 OGD type F3Hs (*V1_10.1500|OGD_F3H_8*, V1_10.1501|OGD_F3H_9), 2 acetyltransferases (*V1_03.1034|ACT_4*, *V1_04.1966|ACT_1*), a glycosyltransferase (*V1_01.540|O4GT_7*), 2 glycosyl-hydroxylases (*V1_06.255| GH_2*, *V1_10.363| GH_5*), 4 MTs (*V1_Us.549|OMT_2*, *V1_10.1657|OMT_6*, *V1_04.1273|OMT_7*, *V1_08.962|OMT_8*) and a leucoanthocyanidin dioxygenase (LODX) or anthocyanidin synthase (ANS) (*V1_02.265| ANS_2),* have higher expressions in PTs asparagus. These DEGs results suggests that, the above mentioned synthetic and regulatory genes are related to anthocyanin biosynthesis in purple asparagus (Supplementary Fig. S3A, Fig. 2 and Fig. 4).

3.2.4. The genes related to specific accumulated phenolics in top part of green asparagus

There are 23 phenolics which have higher accumulation in top parts of green asparagus than its purple counterpart, gene expression results showed that there are 34 genes including; 7 TFs (*V1_07.2348|MYB_2*, *V1_08.334|MYB_3*, *V1_07.1329|MYB_4*, *V1_01.2036|MYB_8, V1_04.1229|bHLH_3*, *V1_01.782|bHLH_6*, *V1_01.1508|WD40_1*), an AaS (*V1_05.1756|AROH_1)*, 3 UPPS/UPPAS genes (*V1_Us.547|AT_2*, *V1_01.315|CHS_2*, *V1_04.1020|4CL_1*), 1 *CHI* (*V1_07.1354|CHI_2*), 2 tandem repeat CYP450s (*V1_01.1321|CYP450_1*, *V1_01.1322| CYP450_2*), 4 OGDs (*V1_07.2786|OGD_1*, *V1_Us.946|OGD_2*, *V1_10.296| OGD_F3H_10*, *V1_08.2373| OGD_FNS_1*), 2 CAD(H)s (*V1_08.2305| CADH_1*, *V1_07.2451|CADH_4*), a CMAT (*V1_02.456|CMAT_3*), a COMT (*V1_02.455|COMT_1*), an ACT (*V1_02.2073|ACT_2*), 4 GTs (*V1_04.1852| GT_1*, *V1_01.2166|O5GT_1*, *V1_01.2165|O5GT_3*), 2 GHs (*V1_07.968| GH_4*, *V1_02.2125|GH_6*), a CYP450 F3′,5′H (*V1_04.1113|F3′,5′H_2*) and a CYP450 F3′H (*V1_08.1735|F3′H_3)* have higher expressions in GSs or GFs asparagus than its purple counterparts. The results indicate that, these genes are responding to specific higher accumulated phenolics biosynthesis in GTs of asparagus (Supplementary Fig. S3A, Fig. 2 and Fig. 4).

3.3. qRT-PCR quantification to confirm the gene expressions detected by RNAseq

In order to confirm the correctness of key phenolics synthetic and regulatory gene expressions in asparagus with RNASeq data, 13 candidate genes including 5 TF genes were selected for further qRT-PCR amplification using primers shown in supplementary file 1 Table S3 with RNAs extracted from organs of GSs and GFs and PSs and PFs of asparagus. The results of the expression patterns of RNAseq are nearly consistent with the expression patterns of qRT-PCR analysis (Fig. 4B).

3.4. Phenolics synthetic regulatory genes network in asparagus

For proposing the possible regulatory networks, among 17 selected TFs and between these TFs and 97 putative genes coding the phenolics synthetic enzymes, a total of 114 promoters (2000 bp sequence which is 5′ flanking starting with ATG) were extracted and submitted to PlanTFDB (http://planttfdb.gao-lab.org/) to find the cis element for binding AP2/ERFs, WRKYs, MYBs, bHLHs and WD40s TFs. Judged by i) promoters having TF binding sites (TFBSs), and ii) among TFs, and between TFs and phenolics synthetic gene expressions (TPM) were positively or negatively correlated with absolute correlation value > 0.7 and p values < 0.05 amongst TFs, and between TFs and synthetic genes based on Pearson method with R package psych (https://personality-project.org/r/psych), the regulatory gene pairs (including TFs to TFs or TFs to synthetic genes (SGs)) were obtained, and the resulted correlation regulatory network was further visualized by Cystoscopes (https://github.com/cytoscape/ cytoscape/) (Fig. 4C).

Based on RNAseq data*, MYB_5/6, bHLH_3*, *WD40_1* were found to have higher expressions in Ss and Fs but not Rs of asparagus, while *WRKY_1* and *AP2/ERF1* have lower expressions in GSs and GFs but have higher expressions in Rs of green asparagus and all organs (PRs and PSs and PFs) of purple asparagus. *MYB_9*, *bHLH_2/7* have higher expressions in PSs and PFs, but *MYB_2/3/6/8*, bHLH_3*/6, WD40_1* have higher expressions in GSs and GFs than their purple counterparts (PSs and PFs) (Fig. 4A). As demonstrated with cystoscopes, the regulatory gene pairs show that MYB_6 and bHLH_3 have many putative target genes, and they may co-regulate the TF gene of *WD40_1*, 2 AaS genes of *AROH1/2* encoding chorismate mutase (Eberhard et al., 1993) and a 4-hydroxyphenylpyruvic acid oxidase (HPPD) for 2,5-dihydroxyphenylacetate synthesis (Yang et al., 2004) and an aminotransferase (*AT_4*) for nitrogen cycling in phenolics biosynthesis (Singh et al., 1998). The results suggest that, these TFs (MYB_6/bHLH_3/WD4-1) mainly regulate the genes of both DPPS and DPPAS such as, *CHI_2,* CYP450 of *F3′H_3*, *F3′5′H_2*, the DPPS genes of *CAD (H)_2/3*, the DPPAS / DPPS modification genes of *GH_4/6, O4GT_3/10, OMT_4* /5, *ACT_3* and a *PERX_5* encoding peroxidase for lignin, suberization (tannin) biosynthesis by the oxidation and ligation of monomers of simple phenolics or flavonoids respectively. Based on the results of these regulatory gene pairs, it's reasonable to suggest that MYB_6 and bHLH_3 may have integrated WD40_1 to form MBW (i.e. MYBs, bHLH and WD40s) complexes (Lloyd et al., 2017), take general regulatory roles for phenolics synthesis in asparagus. The regulatory gene pairs also showed that MYB_2 and bHLH_5 also take central roles in the regulatory networks of phenolics synthesis, importantly they co-regulate the genes of *AT-5*, *PAL2/3* (Phenylalanine ammonia-lyase), *GH_3/5*, suggesting they form different MBWs complexes with unknown WD40 (e.g. *WD40-1 or WD40-7*) to regulate the process of nitrogen cycle engaged in pathways of both UPPS and UPPAS, which involves glutamine synthetases (GSs), glutamate synthetases (GOGATs) and ATs for removing the accumulation of ammonia by nitrogen cycle during phenolic synthesis (highlighted by red circle in Fig. 5).

Based on the results of genes expression analyses, we also found higher expressions of *V1_01.2223|WRKY_1, V1_01.2478|AP2/ERF-1, MYB_2/3,* and *bHLH_6/8* in both GRs and PRs, and *MYB_2/3,* and *bHLH_6* have higher expressions in top part of green asparagus (GSs and

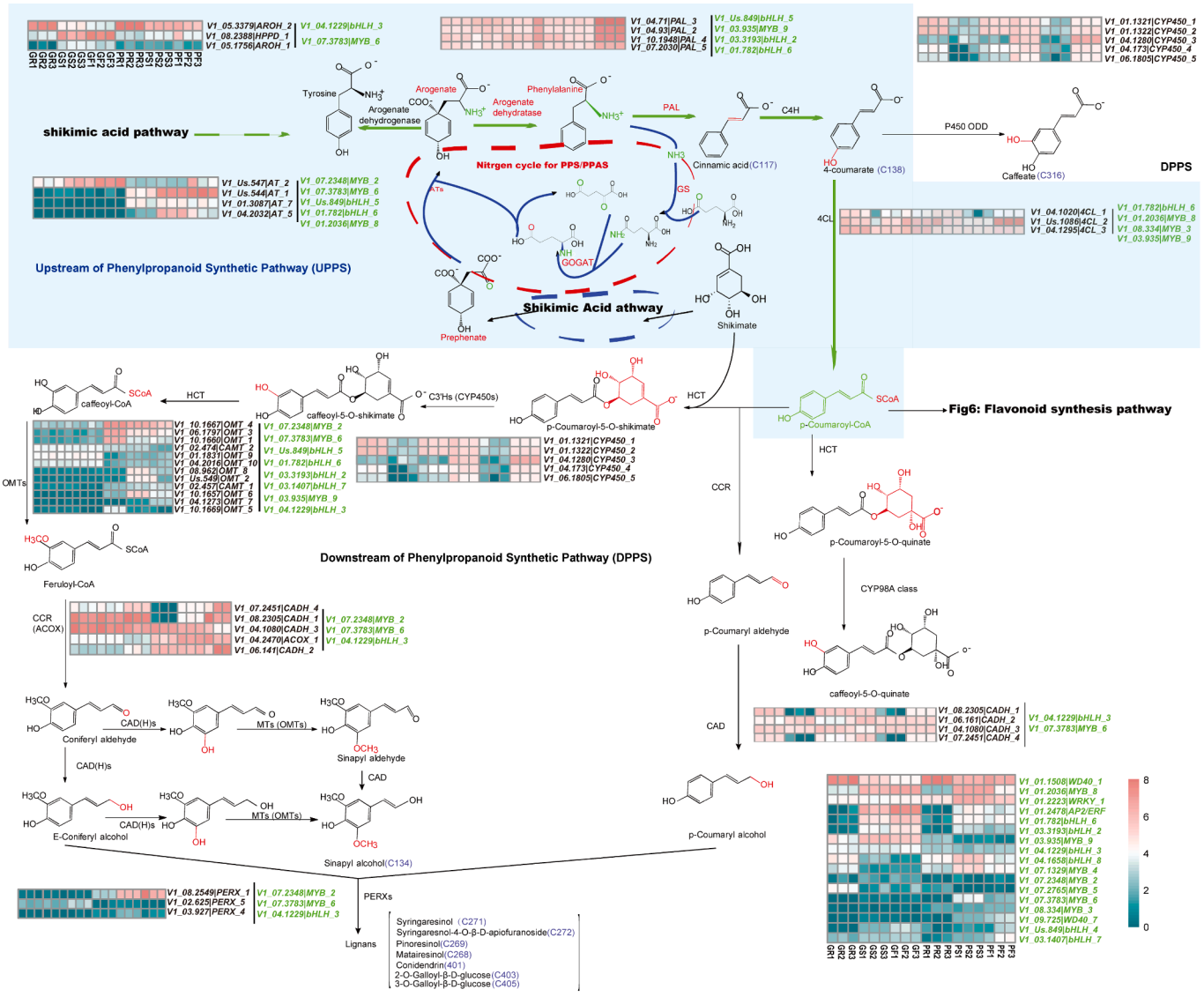

Fig. 5. The main hypothetical biosynthetic pathway of phenylpropanoid pathway (PPS) for the synthesis of simple phenols and their resulting polymers such as lignins in asparagus, where changes in metabolites in each reaction are highlighted in red in the structural formula and average expression of genes of Rs, Ss and Fs of green and purple asparagus respectively are shown in the heat maps. One step reaction is represented with solid lines while the dash lines represent multiple steps reaction, light green module: the upstream of PPS (UPPS), white module: the downstream of PPS(DPPS), the red and blue circle in light green module representing the nitrogen cycle for PPS /phenylpropanoid-acetate (PPAS) and shikimic acid pathway respectively. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

GFs), while *WRKY_1, ERF-1 and bHLH8* have higher expressions in PSs and PFs of asparagus, suggesting that, they may integrate together to regulate the specific phenolic accumulation in the Rs and different phenolic contents in Ts of both asparagus cultivars. Interestingly gene regulatory pair analyses showed that, MYB_2/4 and WRKY_1 directly regulate *AP2/ERF_1* suggesting that asparagus may increase the stress tolerance by upregulation of specific phenolic accumulation by mediation of ethylene signaling which may be induced by abiotic/biotic stress signals. These results also showed that both MYB_4 and WRKY1 have regulatory gene pairs with IFR (isoflavone reductase) for synthesis of isoflavonoids which are signaling molecules for responding to microorganism. The regulatory gene paired results showed that MYB_4/6 and bHLH_3 regulates *PERX_5,* suggesting these TFs are major regulators for enzyme expressions for lignins and/or tannins synthesis in asparagus. With higher expression *of MYB-2/4/8* and *bHLH_6* detected in both GSs and GFs, while higher expressions of *MYB_9 and bHLH_2/7* only detected in top parts of purple (PSs and PFs) asparagus cultivars, suggesting they may concertedly regulate different accumulations of phenolics between green and purple asparagus. MYB_2 and bHLH_5, which have many regulatory gene pairs as described above, therefore the accumulation of specific phenolics especially flavonoids in green asparagus may be due to interaction of MYB_2 with bHLH_5 (Fig. 4C). The MYB_9, bHLH_2/7 also have many regulatory pairs genes of both upstream pathways of PPS and PPAS, including *PAL_4/5*, *4CL_2*, and DPPAS genes such as *CHI_1/4*, *GH_3/5*, *OMT_7/8/9*, *F3H_8*, *4OGT-11*, *ACT_1*, *CMAT_3 and ANS_2/3*, indicating they are key TFs to regulate the phenolics synthetic genes such as *F3H_8 and ANS_2* to channel more metabolic flux into anthocyanin synthesis, and the resulted anthocyanin compounds may be further modified by glycosylation by GTs (O4GT-11) or Glycosyl-hydrolase (GH_3/5) and methylation of OMTs (OMT_7/8/9) for diversified anthocyanin structure accumulation in purple asparagus.

3.5. The proposed full synthetic and regulatory PPS and PPAS pathway in asparagus

3.5.1. The synthesis and regulation of PPS

The UPPS, starting from phenylalanine (Phe) is synthesized by shikimic acid pathway. Phe is first converted to cinnamic acid by action of PAL2/3/4/5 with nitrogen cycle for PPS/PPAS which are catalyzed by the enzymes: GSs, GOGATs, AT1/2/5/7 and arogenate dehydratases (highlighted red cycle of Fig. 6). The resulted cinnamic acid is further converted into p-coumarin by actions of p-coumarin synthase (C4Hs), which is then converted to p-coumarinol-CoA by action of 4-hydroxycinnamyl-CoA ligase (4CL_1/2/4) (showed in light-green module of Fig. 5). The downstream pathway of phenylpropanoid (DPPS, showed in white module of Fig. 5) starting from the p-coumarinol-CoA which can form diversified monophenols (e.g. caffetol, coniferol, 5-hydroxyconiferol, or glusinosol), and polymers of lignin, through 4-coumaryaldehyde, to cinnamyl alcohol by concerted action of ACOXs/CCRs (ACOX_1), CAD (H)s) (CADH_1/2/3/4) and CYP98As/C3Hs (CYP450_1/2/3/4/5). These resulted monolignols can be converted to each other and ligated with shikimate as well, for instance; p-coumarol CoA can link together with shikimate to form p-coumaryl-5-O-shikimic acid via shikimate hydroxycinnamoyl-transferases (HCTs). While coumaryl-5-O-shikimic acid can also be split into 4-coumarol and shikimic acid again by the reversible action of HCTs. The 4-coumaryl-5-O-shikimic acid can be

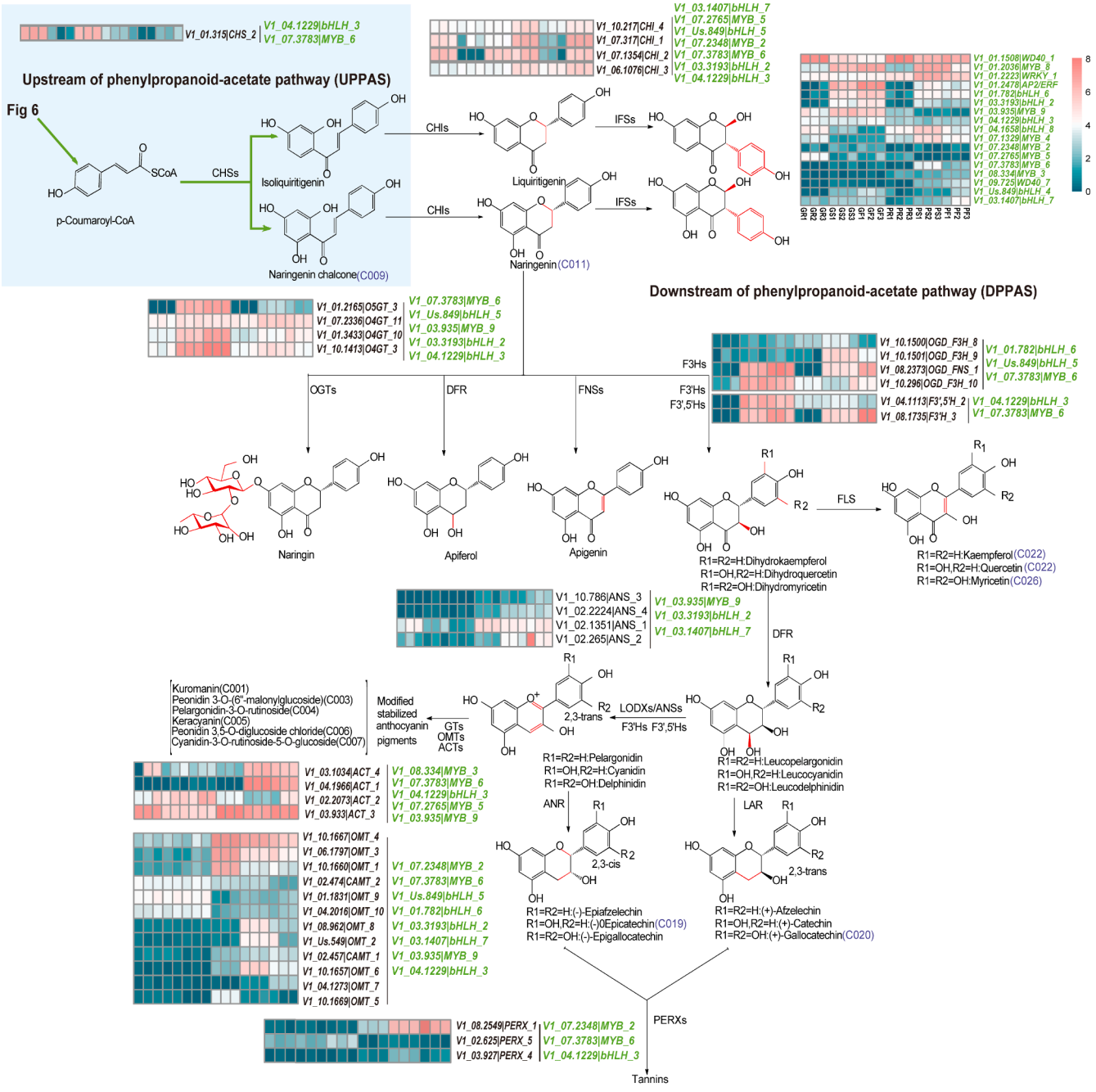

Fig. 6. The main hypothetical biosynthetic pathway of phenylpropanoid-acetate (PPAS) for the synthesis of flavonoids and their resulting polymers such as tannins in asparagus, where changes in metabolites in each reaction are highlighted in red in the structural formula and average expression of genes of Rs, Ss and Fs of green and purple asparagus respectively are shown in the heat maps, light green module: the upstream of PPAS (UPPAS), white module: the downstream of PPAS(DPPAS). (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

converted to caffeoyl-5-O-shikimate by C4Hs (CYP450_1/2/3/4/5), and similarly 4-coumarate, 4-coumarylaldehyde and 4-coumaryl alcohol converted into caffeoyl acid, caffeoyl aldehyde and caffeoyl alcohol by C3Hs (CYP450_1/2/3/4/5) respectively. These phenols can be further modified by OMTs (OMT_1/2/3/6/7/8/9/10 or CAMT1/2) to form sinapyl alcohol/ aldehyde/ acid, or condensed by oxidation of peroxidase (e.g. PERX1/5) to form different lignins. To further investigate the possible regulators of PPS pathway, it was observed that MYB_6 and bHLH_3 regulated nearly all steps of PPS, suggesting they combined with either WD40_1 or WD40_7 to form MBWs complex(s) to regulate the PPS synthetic genes for monophenols and different lignins accumulation in asparagus (Fig. 5).

3.5.2. *The synthesis and regulation of PPAS*

The UPPAS pathway starts from Phe to naringenin chalcone or isoliquiritigenin as showed in light-green modules of Fig. 5 and Fig. 6, while DPPAS starts from naringenin chalcone (showed in white module of Fig. 6). The typical flavonoids have three interconnected rings in which the A ring is derived from hydroxycinnamoyl-CoA, the B ring is formed by the condensation reactions from malonyl-CoA (an activated form of acetyl-coA) by CHS_2, while the C ring is formed by closure of the chalcone which is catalyzed by chalcone isomerases (CHI_1/2/3/4). After the core flavanone (naringenin (C011)) is formed, several modifications take place that lead to diversified structures of isoflavonoids catalyzed by isoflavone synthase (IFS), which catalyzes an oxygenation and an unusual C-2 to C-3 aryl migration to yield the 2-hydroxyisoflavanones. and further action of 2-hydroxyisoflavanone dehydratase (IFR), catalyzing the dehydration of the 2-hydroxyisoflavanones to form isoflavanols. The naringenin and its derived flavonoids can be further modified with methylation, acetylation and glycosylation/deglycosylation by OMTs, ACTs and OGTs/GHs to produce compounds with diversified structures (Fig. 6). In the second (major) conversion, C ring of flavanone can be aromatized by action of the OGD type flavone synthase (FNS_1), or further oxygenated at position 3 (on the C ring) by OGDs F3H_8/9/10 (Fig. 6) to form dihydroflavonols (i.e. dihydrokaempferol). The A ring of dihydrokaempferol can be further oxygenated by CYP450 type F3′5′H or F3′H to form compounds (dihydroquercetin and dihydromyricetin) with higher orders of hydroxylation. The C ring of dihydroflavonols can be aromatized by dihydroflavonol dehydrogenase (FLS) to kaempferol, quercetin and myricetin respectively. The keto group of C ring of dihydroflavonol can also be catalyzed by dihydroflavonol reductase (DFR) to form the leucoanthocyanidin compounds (leucopelargonidin, leucocyanidin and leucodelphinidin), which start another important branch in the flavonoid biosynthetic network. Leucoanthocyanidins as precursors for anthocyanidins (cyanidin, pelargonidin and delphinidin) are catalyzed by anthocyanidin synthases (ANS_1/2/4), and condensed to tannins (proanthocyanidins) probably by peroxidase action of PERX_1/5. Finally, the stable and water-soluble anthocyanins (C001 ~ C007) were formed by glycosylation by GTs. As described above, with higher expressions of *MYB_9 and bHLH_2/7* in top parts of purple (PSs and PFs) asparagus suggesting they are major players in regulating anthocyanins accumulation in top parts of purple asparagus.

4. Discussion and conclusion

Phenolics has been reported as the most abundant group of organic compounds of phytochemicals and play several roles including growth and reproduction, providing protection against pathogens as well as used as antifeedants of plants. Phenolics also plays roles in plant structural support, protection against biotic and abiotic stresses, as well as food resource. Humans need phenolics to provide protection against non-transmissible chronic diseases such as cancer, diabetes, obesity and cardiovascular diseases by means of their antioxidant activities as well as regulating many cellular processes such as protein phosphorylation, gene expression modification, enzyme regulation etc (De la Rosa et al., 2019). In this study, integrated strategies to divergent organs of both green and purple asparagus cultivars were used to detect their phenolics profiles, contents and key genes for both biosynthesis and regulation which resulted in 204 phenolics containing 47 % of 437 total detected metabolite (Figs. 2 & 3, Supplementary file 2, Supplementary file1 Fig. S2 and S3) with fluctuated concentrations in Rs, Ss and Fs of both purple and green asparagus cultivars. The phenolics are mostly accumulated in the Ts of both cultivars with most abundant type being flavonoids, which are consistent with the TFCs determined with the spectrophotometric methods suggesting that phenols biosynthesis is regulated by light. (Figs. 1 & 2). Between cultivars (green and purple) 5 out of 7 anthocyanins were derived from 3 types of anthocyanidins (pelargonidin, cyanidin and delphinidin) which are highly accumulated in PTs especially in the PFs of purple asparagus, while both the Ss and Fs have non-significant differences of both phenolics and ANus within each cultivar. These results suggest a slightly late harvesting time for Ss (unless the spears have over lignified) of both green and purple cultivars, might be a good practice for increased yields due to further biomass accumulation either transported from storage roots or by direct photosynthetic accumulation during periods when Ss develop into Fs stages, especially the asparagus products only used for processing functional foods development, e.g. dried asparagus powders and additives for functional food development (e.g. fermentation of dairy products).

Asparagus roots are traditionally used as medicinal products, mainly due to their accumulation of steroidal saponins and functional polysaccharides (e.g. fructus). The asparagus root system consists of fleshy storage roots attached to an underground rhizome, as well as feeder roots attached to storage roots for absorption of mineral nutrients and water. The storage roots continue to grow throughout the plant's life, while that of the feeder roots have relatively shorter life time. Roots of asparagus exert considerable phenolics and steroidal saponins with antimicrobial activities, allow perennial survival and shapes their rhizospheral microbiota over their lifespan. The metabolomic data showed that 46 phenolics were present in all organs, 8 phenolics (C152,C088, C018,C029, C037,C043, C066,C156) which are higher accumulated in Ts of parts of purple asparagus, have higher accumulation in GR than PRs, while 16 phenolics (e.g.; C020, C261, C141, C127, C102, C316, C069, C083, C147, C103, C148, C164, C101) were highly accumulated in both green and purple asparagus roots (Supplementary file 3, supplementary file1 Fig. S3). These root accumulated phenolics were either transported from top parts of asparagus or directly synthesized in the Rs with regulation of *bHLH_6/8* only or their combining with unknown TFs, such as *MYBs*, WD40s, *WRKY_1* and *AP2/ERF_1* (Fig. 4).

Anthocyanins are the major different phenolics with higher accumulation in Ts of purple asparagus. As showed in Fig. 6, dihydroflavonols can be converted to leucoanthocyanidin class (e.g. leucopelargonidin, leucocyanidin and leucodelphinidin) by dihydroflavonol reductase (DFR), which further form the anthocyanidins by actions of anthocyanidin synthases (ANSs) In this study, DFR(s) were not identified, however 3 *ANSs* (*ANS_2/3/4*) were detected to have higher expressions in the PTs with co-expressions of TF genes of *MYB_9* and *bHLH_2/7* (Fig. 4). Further gene regulatory pairs show that both MYB_9 and bHLH_2 have regulatory pairs with *ANS_2*, while both MYB_9 and bHLH_7 have regulatory pairs with *ANS_3* suggesting that anthocyanidins synthesis by ANS2/3 may be key steps in purple asparagus and these *ANSs* are more likely regulated by MYB_9 and bHLH_2/7, which may have interacted with unknown WD40s (e.g. WD40-7) to form regulatory MBW complex(s).

This work confirms the high important phenolics compounds biosynthesis and regulation for healthy asparagus production as well as other important agricultural compounds as previously discovered (Abd-Ella et al., 2022, Min et al., 2019, Hijaz et al., 2020). In summary, as an adaptation to the varying environments, asparagus can amend the production and preservation of diverse sets of phenolics in different organs. The non-significance of phenolics and ANus in Ss and Fs of both purple and green asparagus provide chemical evidence for a slightly late

harvesting time for commercial asparagus production. A total 23 phenolic synthetic and regulatory genes were detected with ASs or fusions suggesting the isoform levels of transcripts of these genes taking part in regulating the differential phenolics accumulation in organs of asparagus. Phenolics are mainly accumulated in top part of asparagus, in which anthocyanins which are key different phenolics between green and purple asparagus, and are highly accumulated in purple asparagus may be mainly regulated by TFs of both MYB_9 and bHLH_2/7.

CRediT authorship contribution statement

Yuping Li: . **Sylvia E. Brown:** Formal analysis, Writing – original draft, Writing – review & editing. **Yunbin Li:** . **Qin Cheng:** Data curation, Formal analysis. **He Wu:** Data curation, Formal analysis. **Shugu Wei:** Formal analysis. **Xingyu Li:** Funding acquisition, Project administration, Supervision. **Chun Lin:** Data curation, Funding acquisition, Supervision. **Zhengjie Liu:** Funding acquisition, Investigation, Supervision. **Zichao Mao:** .

Declaration of competing interest

The authors declare the following financial interests/personal relationships which may be considered as potential competing interests: Zichao Mao reports financial support was provided by Yunnan Provincial Science and Technology Department and Sichuan Provincial Science and Technology Department.

Data availability

Data will be made available on request.

Acknowledgements

The authors are grateful to Biomarker Technologies (Beijing, China) for assisting with high-throughput sequencing and Metware Biotechnology Inc (Wuhan, China) for widely-targeted metabolomics analysis. The datasets generated during the study are available in the National Genomics Data Center repository with No GSA CRA009175, https://ngdc.cncb.ac.cn/

Funding

This work was supported by Yunnan Provincial Science and Technology Department (202101BD070001-027, 202305A0350012) and National Natural Science Foundation of China (No. 32360089).

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.fochms.2023.100187.

References

Abd-Ella, A., Metwally, S., El-Ossaily, Y., Elrazek, F., Aref, S., Naffea, Y., et al. (2022). A review on recent advances for the synthesis of bioactive pyrazolinone and pyrazolidinedione derivatives. *Current Chemistry Letters, 11*(2), 157–172.

Brady, J. D., & Fry, S. C. (1997). Formation of Di-Isodityrosine and Loss of Isodityrosine in the Cell Walls of Tomato Cell-Suspension Cultures Treated with Fungal Elicitors or H2O2. *Plant Physiology, 115*(1), 87–92.

Cheng, Q., Zeng, L., Wen, H., Brown, S. E., Wu, H., Li, X., et al. (2023). Steroidal saponin profiles and their key genes for synthesis and regulation in Asparagus officinalis L. by joint analysis of metabolomics and transcriptomics. *BMC Plant Bio, 23*(1), 207.

Cocuron, J. C., Casas, M. I., Yang, F., Grotewold, E., & Alonso, A. P. (2019). Beyond the wall: High-throughput quantification of plant soluble and cell-wall bound phenolics by liquid chromatography tandem mass spectrometry. *Journal of Chromatography. A, 1589*, 93–104.

De la Rosa LA, Moreno-Escamilla JO, Rodrigo-García J, Alvarez-Parrilla E: Chapter 12 - Phenolic Compounds. In: Postharvest Physiology and Biochemistry of Fruits and Vegetables. Edited by Yahia EM: Woodhead Publishing; 2019: 253-271.

de Sena Brandine, G., & Smith, A. D. (2019). Falco: high-speed FastQC emulation for quality control of sequencing data. F1000Research, 8, 1874. Bhosale, P. B., Abusaliya, A., Kim, H. H., Ha, S. E., Park, M. Y., Jeong, S. H., Vetrivel, P., Heo, J. D., Kim, J. A., Won, C. K., Kim, H. W., and Kim, G. S., 2022, Apigetrin Promotes TNFalpha-Induced Apoptosis, Necroptosis, G2/M Phase Cell Cycle Arrest, and ROS Generation through Inhibition of NF-kappaB Pathway in Hep3B Liver Cancer Cells: Cells, v. 11, no. 17.

Desmedt, W., Jonckheere, W., Nguyen, V. H., Ameye, M., De Zutter, N., De Kock, K., et al. (2021). The phenylpropanoid pathway inhibitor piperonylic acid induces broad-spectrum pest and disease resistance in plants. *Plant, Cell & Environment, 44*(9), 3122–3139.

Dong, T., Han, R., Yu, J., Zhu, M., Zhang, Y., Gong, Y., et al. (2019). Anthocyanins accumulation and molecular analysis of correlated genes by metabolome and transcriptome in green and purple asparaguses (Asparagus officinalis, L.). *Food Chemistry, 271*, 18–28.

Eberhard, J., Raesecke, H. R., Schmid, J., & Amrhein, N. (1993). Cloning and expression in yeast of a higher plant chorismate mutase. Molecular cloning, sequencing of the cDNA and characterization of the Arabidopsis thaliana enzyme expressed in yeast. *FEBS Letters, 334*, 233–236.

Fernandez-Pena, L., Matos, M. J., and Lopez, E., 2022, Recent Advances in Biologically Active Coumarins from Marine Sources: Synthesis and Evaluation: Mar Drugs, v. 21, no. 1.

Fraga, C. G., Clowers, B. H., Moore, R. J., & Zink, E. M. (2010). Signature-discovery approach for sample matching of a nerve-agent precursor using liquid chromatography-mass spectrometry. *XCMS, and chemometrics: Anal Chem, 82*(10), 4165–4173.

Groth, D., Hartmann, S., Klie, S., & Selbig, J. (2013). Principal components analysis. Computational. *Toxicology, II*, 527–547.

Harkess, A., Zhou, J., Xu, C., Bowers, J. E., Van der Hulst, R., Ayyampalayam, S., et al. (2017). The asparagus genome sheds light on the origin and evolution of a young Y chromosome. *Nature Communications, 8*(1), 1279.

Hijaz, F., Al-Rimawi, F., Manthey, J. A., & Killiny, N. (2020). Phenolics, flavonoids and antioxidant capacities in citrus species with different degree of tolerance to huanglongbing. *Plant Signaling & Behavior, 15*(5), 1752447.

Kim, D., Langmead, B., & Salzberg, S. L. (2015). HISAT: A fast spliced aligner with low memory requirements. *Nature Methods, 12*(4), 357–360.

Kong, L., Zhang, Y., Ye, Z. Q., Liu, X. Q., Zhao, S. Q., Wei, L., and Gao, G., 2007, CPC: assess the protein-coding potential of transcripts using sequence features and support vector machine: Nucleic Acids Res, v. 35, no. Web Server issue, p. W345-349.

Langfelder, P., & Horvath, S. (2008). WGCNA: An R package for weighted correlation network analysis. *BMC Bioinformatics, 9*, 559.

Ledesma-Escobar, C. A., Priego-Capote, F., Robles Olvera, V. J., & Luque de Castro, M. D. (2018). Targeted Analysis of the Concentration Changes of Phenolic Compounds in Persian Lime (Citrus latifolia) during Fruit Growth. *Journal of Agricultural and Food Chemistry, 66*(8), 1813–1820.

Li, H. (2018). Minimap2: Pairwise alignment for nucleotide sequences. *Bioinformatics, 34*(18), 3094–3100.

Liao, Y., Smyth, G. K., & Shi, W. (2014). featureCounts: An efficient general purpose program for assigning sequence reads to genomic features. *Bioinformatics, 30*(7), 923–930.

Liu, B., Liu, Y., Li, J., Guo, H., Zang, T., & Wang, Y. (2019). deSALT: Fast and accurate long transcriptomic read alignment with de Bruijn graph-based index. *Genome Biology, 20*(1), 274.

Liu, H. J., Jiang, X. X., Guo, Y. Z., Sun, F. H., Kou, X. H., Bao, Y., et al. (2017). The flavonoid TL-2-8 induces cell death and immature mitophagy in breast cancer cells via abrogating the function of the AHA1/Hsp90 complex. *Acta Pharmacologica Sinica, 38*(10), 1381–1393.

Liu, P., Gao, R., Gao, L., Bi, J., Jiang, Y., Zhang, X., et al. (2022). Distinct Quality Changes of Asparagus during Growth by Widely Targeted Metabolomics Analysis. *Journal of Agricultural and Food Chemistry, 70*(50), 15999–16009.

Lloyd, A., Brockman, A., Aguirre, L., Campbell, A., Bean, A., Cantero, A., et al. (2017). Advances in the MYB-bHLH-WD Repeat (MBW) Pigment Regulatory Model: Addition of a WRKY Factor and Co-option of an Anthocyanin MYB for Betalain Regulation. *Plant & Cell Physiology, 58*, 1431–1441.

Love, M. I., Huber, W., & Anders, S. (2014). Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biology, 15*(12), 550.

Ma, J., Li, X., He, M., Li, Y., Lu, W., Li, M., Sun, B., and Zheng, Y., 2023, A Joint Transcriptomic and Metabolomic Analysis Reveals the Regulation of Shading on Lignin Biosynthesis in Asparagus: Int J Mol Sci, v. 24, no. 2.

Min, T., Bao, Y., Zhou, B., Yi, Y., Wang, L., Hou, W., et al. (2019). Transcription profiles reveal the regulatory synthesis of phenols during the development of lotus rhizome (nelumbo nucifera gaertn). *International Journal of Molecular Sciences, 20*(11).

Pang, H., Wu, T., Peng, Z., Tan, Q., Peng, X., Zhan, Z., et al. (2022). Baicalin induces apoptosis and autophagy in human osteosarcoma cells by increasing ROS to inhibit PI3K/Akt/mTOR, ERK1/2 and beta-catenin signaling pathways: J Bone. *Oncol, 33*, Article 100415.

Rana, A. K., Sharma, S., Saini, S. K., & Singh, D. (2022). Rutin protects hemorrhagic stroke development via supressing oxidative stress and inflammatory events in a zebrafish model. *European Journal of Pharmacology, 925*, Article 174973.

Renuka Devi., R., & Arumughan, C., 2007, Antiradical efficacy of phytochemical extracts from defatted rice bran. Food and chemical toxicology : an international journal published for the British Industrial Biological Research Association, 45(10).

Santibanez, A., Herrera-Ruiz, M., Gonzalez-Cortazar, M., Nicasio-Torres, P., Sharma, A., and Jimenez-Ferrer, E., 2022, Pharmacokinetics and Tissue Distribution of Coumarins from Tagetes lucida in an LPS-Induced Neuroinflammation Model: Plants (Basel), v. 11, no. 21.

Shao, X., Chen, X., Wang, Z., Zhu, C., Du, Y., Tang, D., et al. (2022). Diprenylated flavonoids from licorice induce death of SW480 colorectal cancer cells by promoting autophagy: Activities of lupalbigenin and 6,8-diprenylgenistein. *Journal of Ethnopharmacology, 296*, Article 115488.

Singh, S., Lewis, N. G., & Towers, G. H. (1998). Nitrogen recycling during phenylpropanoid metabolism in sweet potato tubers. *Journal of Plant Physiology, 153*, 316–323.

Song, F., Wang, L., Mu, J., & Ma, H. (2022). Protective Effects and Molecular Mechanism of Total Flavonoids from Lycium Barbarum Leaves on Photoaged Human Dermal Fibroblasts. *Evidence-Based Complementary and Alternative Medicine, 2022*, 4156330.

Tang, S., Lomsadze, A., & Borodovsky, M. (2015). Identification of protein coding regions in RNA transcripts. *Nucleic Acids Research, 43*(12), e78.

Tardaguila, M., de la Fuente, L., Marti, C., Pereira, C., Pardo-Palacios, F. J., Del Risco, H., et al. (2018). SQANTI: Extensive characterization of long-read transcript sequences for quality control in full-length transcriptome identification and quantification. *Genome Research, 28*(3), 396–411.

Tuominen, A., & Salminen, J. P. (2017). Hydrolyzable Tannins, Flavonol Glycosides, and Phenolic Acids Show Seasonal and Ontogenic Variation in Geranium sylvaticum. *Journal of Agricultural and Food Chemistry, 65*(31), 6387–6403.

Wu, T., Hu, E., Xu, S., Chen, M., Guo, P., Dai, Z., et al. (2021). clusterProfiler 4.0: A universal enrichment tool for interpreting omics data. *Innovation (Camb), 2*(3), Article 100141.

Xi, Y., Cheng, D., Zeng, X., Cao, J., & Jiang, W. (2016). Evidences for Chlorogenic Acid–A Major Endogenous Polyphenol Involved in Regulation of Ripening and Senescence of Apple Fruit. *PLoS One1, 11*(1), e0146940.

Yang, C., Pflugrath, J. W., Camper, D. L., Foster, M. L., Pernich, D. J., & Walsh, T. A. (2004). Structural basis for herbicidal inhibitor selectivity revealed by comparison of crystal structures of plant and mammalian 4-hydroxyphenylpyruvate dioxygenases. *Biochemistry, 43*, 10414–10423.

Yi, T. G., Yeoung, Y. R., Choi, I. Y., & Park, N. I. (2019). Transcriptome analysis of Asparagus officinalis reveals genes involved in the biosynthesis of rutin and protodioscin. *PLoS One1, 14*(7), e0219973.