

Integrity transcriptome and proteome analyses provide new insights into the mechanisms regulating peel cracking in *Akebia trifoliata* fruit

Juan Niu

Chinese Academy of Agricultural Sciences Institute of Bast Fiber Crops

Yaliang Shi

Chinese Academy of Agricultural Sciences Institute of Bast Fiber Crops

Kunyong Huang

Chinese Academy of Agricultural Sciences Institute of Bast Fiber Crops

Yicheng Zhong

Chinese Academy of Agricultural Sciences Institute of Bast Fiber Crops

Jing Chen

Chinese Academy of Agricultural Sciences Institute of Bast Fiber Crops

Zhimin Sun

Chinese Academy of Agricultural Sciences Institute of Bast Fiber Crops

Mingbao Luan (✉ luanmingbao@caas.cn)

Chinese Academy of Agricultural Sciences Institute of Bast Fiber Crops

Jianhua Chen

Chinese Academy of Agricultural Sciences Institute of Bast Fiber Crops

Research

Keywords: Akebia trifoliata, fruit peel structure, fruit cracking, transcriptome, proteome, pentose and glucuronate interconversions, galactose metabolism

Posted Date: December 18th, 2019

DOI: <https://doi.org/10.21203/rs.2.19126/v1>

License: © ⓘ This work is licensed under a Creative Commons Attribution 4.0 International License.
[Read Full License](#)

Version of Record: A version of this preprint was published at Biotechnology for Biofuels on August 20th, 2020. See the published version at <https://doi.org/10.1186/s13068-020-01789-7>.

Abstract

Background

Akebia trifoliata (Thunb.) Koidz, can be used as a new potential candidate biofuel and bioenergy crop due to its high productivity, adaptability and tolerance to cultivation conditions. However, the pericarp of *A. trifoliata* cracks open longitudinally and disperses its seeds along the ventral suture during fruit ripening, which is a serious problem that limits the usefulness of its biofuel feedstocks and causes significant losses of yield and commercial value. However, there have been no known previous investigations of the fruit cracking and its molecular mechanisms in *A. trifoliata*.

Results

The dynamic structural changes of the fruit peels were observed. In the non-cracking stage of growth, the exocarp was dense, had an orderly arrangement, and the cuticle was complete and distributed continuously. However, the cells became thinner, had reduced integrity, lost cell wall structures, and there was cell wall break down, in the fruit cracking stage. Moreover, analysis of the complementary RNA sequencing based transcriptomes and tandem mass tag based proteomes at different development stages of the fruit ripening, were performed to detect the genes and proteins related to the fruit cracking in *A. trifoliata*. A total of 20 differentially expressed genes and 17 differentially abundant proteins were identified from the transcriptomics and proteomics data that contribute to the fruit cracking, by participating in the biosynthesis of the phenylpropanoid pathway, galactose metabolism, pentose, and glucuronate interconversions, starch and sucrose metabolism, amino sugar and nucleotide sugar metabolism. Several candidate genes and proteins related to pentose and glucuronate interconversions (pectate lyases and pectinesterase) and galactose metabolism (β -galactosidases 1 and β -galactosidases 2) may play key roles in *A. trifoliata* fruit cracking.

Conclusions

Complementary transcriptome and proteome level analyses indicated that a complex molecular network was controlling the fruit cracking process. This study provides new insights into the molecular basis of fruit cracking in *A. trifoliata* fruits. The candidate genes/proteins identified in this study may be useful for the genetic improvement of *A. trifoliata* and other crops.

Background

Akebia trifoliata (Thunb.) Koidz is a perennial woody liana, belonging to the family Lardizabalaceae and the subgenus *Akebia* Decne [1]. The fruit of *A. trifoliata* is a promising biofuel resource, due to its high total sugar content (50.32%), high yields, wide adaptability, tolerance to both drought and heat, and ease of management [2]. Moreover, it produces many seeds (100–300 seeds), and the seed contain 38.83% oil [3]. The seed oil from *A. trifoliata* is not only one of the most important biodiesels for industry, as it has a similar structure to petrol-diesel, but it also has an abundance of nutritional properties, and is commonly

used as an edible oil for human consumption [4, 5]. Moreover, it has high adaptability for natural conditions, and has wide geographic and vertical distributions, ranging from the subtropical to temperate regions, and from flat ground to elevations of 250 to 2500 m [6]. *A. trifoliata* is also considered to be edible, have medicinal value, be ornamental, and have high economic and developmental values [7]. The advantages of *A. trifoliata* have made it a great prospect for industrial applications, human nutrition, and other potential applications.

However, the pericarp of *A. trifoliata* cracks open longitudinally when matured in Chinese lunar August, which was also called 'Bayuezha' or 'Bayuegua' by the local people. Fruit cracking is a serious problem that increases the incidence of pathogen infection, affects the utilization rates of biofuel feedstocks, and causes significant losses of yield and commercial value. The cracking of fruits has been observed in many species, such as tomato [8], litchi [9], pomegranate [10], soybean [11], *Brassica napus* [12], *Arabidopsis thaliana* [13], and many other species. Various studies have been carried out to elucidate the mechanisms of fruit cracking. Previous studies have shown that it is a complex phenomenon caused by several factors, such as peel morphology, physiobiochemical characteristics, environmental and genetic factors [14]. For example, peel tensile strength, cracking turgor thresholds, and the hormone and soluble solid contents were higher in the cracked fruits than in the uncracked fruits [15]. Environmental factors, such as rainfall, mineral nutrition, temperature, light intensity, and humidity also influence the fruit cracking, but these factors are difficult to control. The loss caused by fruit cracking can be reduced by cultivation techniques, but such treatments are temporary and expensive. Consequently, genetic analysis of fruit cracking and improved understanding of its molecular mechanisms is an effective strategy for breeding varieties that will not crack.

Previous studies have indicated that cell wall modifying proteins, including polygalacturonases (PGs), pectin esterase (PE), β -galactosidases (β -GAL), expansins (EXPs), and xyloglucan endotransglycosylase (XET) proteins, were reported to be associated with fruit cracking [16, 17]. Research suggested that pod shatter in Brassicaceae is influenced by mechanical processes in the valve margin structure and the molecular pathways associated with the valve development [18]. Seed coat cracking in soybean was caused by the separation of the epidermal and hypodermal cells, which exposed the underlying parenchyma tissues [19]. Sorefan et al. [20] found that INDEHISCENT (IND) coordinated the regulation of PINOID (PID), WAG2, and PIN-FORMED3 (PIN3), established a dynamic pattern of auxin accumulation at the valve margin, which appears to play a dominant role in the dehiscence regulation. Dong et al. [21] found that pod shattering resistance in soybean was mediated by the NAC gene.

Although fruit cracking has been researched for many years, there has been little progress in our understanding of its molecular mechanisms. Especially in *A. trifoliata*, in which no previous studies have yet reported on its fruit cracking. Even molecular research of *A. trifoliata* are rare, only two studies about the use of transcriptome sequencing have recently been reported in *A. trifoliata* [1, 22]. Lack of knowledge about the molecular basis of *A. trifoliata* has made it difficult to recommend preventive measures for fruit cracking. Next generation sequencing (NGS) technologies, such as transcriptome and proteome technologies for measuring gene expression and protein abundance, have become powerful tools for the

discovery of novel genes and their functions, molecular markers, and physiological stress responses in plants [23, 24]. Therefore, in this study, integrative analysis of the transcriptomes and proteomes were performed to illuminate the mechanism of *A. trifoliata* fruit cracking at the molecular level by RNA-seq and tandem mass tag (TMT) technologies. Our comprehensive parallel analyses will provide several new and interesting insights for understanding the molecular mechanisms in fruit cracking, enhancing fruit utilization, and biomass production.

Results

Changes in pericarp structure

The pericarp of *A. trifoliata* is known to crack longitudinally as it matures. The crack in the fruit and subsequent dispersal of the seeds from along the ventral suture, which is similar to the cracking seen in other legume species, so the dynamic structures of the fruit peels in different development stages were observed (Fig. 1). In the non-cracking stage (PS), the arrangement of pericarp cell and cuticle were dense, small intercellular space and distributed continuously. The fruit cracking however, was accompanied by the cells becoming thinner and bigger, a reduction in the number of cell layers, and furthermore, the arrangement of the cells was loose and poor integrity, and they began to degrade in the initial cracking stage (PM). The cells were arranged irregularly and continue to reduce cell layers, and there was atrophy degradation in the total cracking stage (PL).

Transcriptomic analysis overview

To obtain an overview of the *A. trifoliata* transcriptome during fruit development and ripening, three cDNA libraries (ie., PS, PM and PL) were constructed. A total of 47.05, 46.92, and 54.00 million raw sequence reads were produced from the PS, PM and PL libraries, respectively. After removing reads with indeterminate base ratios of > 10%, low-quality reads and adaptor sequences, 46.45, 46.35, 53.46 million clean reads with the percentage of Q30 bases and GC contents of 91.58 – 93.75% and 45.94%–48.48%, were obtained respectively (Table S1). The resultant *A. trifoliata* transcriptome contains 241 376 transcripts, ranging from 201 to 2000 bp, and a total of 186 054 unigenes were identified (Table 1), and the details of size distribution of the transcripts and unigenes are shown in Fig. S1.

Table 1

Summary of the transcriptome and proteome data in *Akebia trifoliata* fruits.

RNA-seq data		MS data based on transcriptome	
Total number of transcripts	241,376	Total spectra	812625
Mean length of transcripts (bp)	515	Identified spectra	68151
Total number of unigenes	186,054	Identified peptides	12456
Mean length of unigenes (bp)	447	unique peptides	10572
N50 length of transcripts (bp)	713	Identified proteins	2839
N50 length of unigenes (bp)	518		

To determine the putative functions of the assembled transcripts, all unigenes were annotated using Basic Local Alignment Search Tool (BLAST) searches against the five databases, including National center for biotechnology information non-redundant protein sequences database (NR) (100 329, 100%), SwissProt (56346, 56.16%), Protein families database (Pfam) (34428, 34.32%), Gene Ontology database (GO) (44558, 44.41%), and Kyoto Encyclopedia of Genes and Genomes pathway database (KEGG) (30298, 30.20%); indicating that the NR database provided the largest number of annotations. 100 329 unigenes were shown to correspond with sequences from at least one of the public databases, and 7283 unigenes annotated in all databases, resulting in the annotation of 100,329 unigenes (41.57% of all the assembled transcripts) in *A. trifoliata* pericarp (Table S2).

Among of these unigenes, 11 205 were identified as differently expressed genes (DEGs) using absolute log2 fold change > 1 with $p < 0.05$ during the fruit ripening. There were 779 up-regulated and 1924 down-regulated unigenes in the PM compared with the PS group, which are presented in a volcano plot in Fig. S2a; 4623 up-regulated and 1975 down-regulated in PL, compared with the PM group, and are presented in a volcano plot in Fig. S2b. There were 1904 DEGs that were co-expressed at PM_PS and PL_PM (Table 2).

Table 2

Summary of proteins and transcripts detected from TMT and RNA sequence data.

	Protein		Transcriptome	
	PM_PS	PL_PM	PM_PS	PL_PM
Unique proteins/genes detected	2839	2839	100329	100329
Significantly DAPs /DEGs	190	50	4607	6598
Up-regulated	84	28	1945	4623
Down-regulated	106	22	2662	1975
Shared proteins/genes	17		1904	
Shared proteins/genes (up-regulated)	9	8	1123	808
Shared proteins/genes (down-regulated)	8	9	781	1096

Functional classification of the identified DEGs

To further understand the function of the identified DEGs, bioinformatics analysis was performed on the basis of gene functional classification and hierarchical cluster analysis. GO analysis indicated that most of the DEGs in the biological processes were involved in the cellular amide metabolic processes and amide metabolic processes; structural molecular activity and oxidoreductase activity were the highest portion of the DEGs in the category of molecular functions, both in PM_PS and PL_PM; cytoplasmic parts and intracellular ribonucleoprotein complexes in PM_PS cells and PL_PM cell parts, were the highest portion of the DEGs in category of cell components respectively (Fig. 2a-2b).

Moreover, a KEGG pathway analysis was carried out to further evaluate the DEGs. In the comparison between the PM_PS and PL_PM, many DEGs were enriched in metabolic pathways, ribosomes, and biosynthesis of the secondary metabolites. Notably, most of the DEGs involved in the cell wall-related DEGs, were downregulated in the PM, compared with the PS group, but however they were upregulated in the PL compared with the PM group (Fig. 2e-2f).

A hierarchical cluster analysis was performed to further understand the expression changes in the cell wall-related DEGs, (Fig. 3). In all, 285 cell wall related DEGs were clustered closely both in PM_PS and PL_PM group. Most of these were involved in pentose and glucuronate interconversions, phenylpropanoid pathway, galactose metabolism, starch and sucrose metabolism, amino sugar and nucleotide sugar metabolism and transcription factors (Fig. 3a-3f).

Confirmation of fruit cell-wall related genes in *A. trifoliata* by Reverse transcription real-time quantitative PCR (qPCR)

To validate the results of the RNA-Seq data, an expression and correlation analysis between the qPCR and the fragments per kilobase per million reads mapped (FPKM) values obtained from the RNA-seq were performed. The 20 selected genes had shown differential expression patterns in the PM_PS and PL_PM groups, and the results of the qPCR are shown in Fig. 4. Specifically, in the PM_PS group, most of these DEGs were down-regulated, including the phenylpropanoid pathway related genes- 4-coumarate-COA-ligase (4CL), peroxidase (PRX), and PRX2; galactose metabolism related genes- β -galactosidases (β -GAL1 and β -GAL2); amino sugar and nucleotide sugar metabolism related gene-beta-D-xylosidase (BXL), starch and sucrose metabolism related genes- cellulase (CEL), cellulose synthase-like protein (CSLG) and Glucan endo-1,3-beta-D-glucosidase (ENDOB), and the transcription factor and cell wall metabolism genes-NAC, NAC like and EXP1. While the phenylpropanoid pathway related genes- cinnamyl-alcohol dehydrogenase (CAD) and shikimate O-hydroxycinnamoyltransferase (HCT), transcription factor and cell wall metabolism genes-BHLH and dirigent protein (DIR2), and pentose and glucuronate interconversion related genes (PL, PG and PE) were significantly up-regulated. In the PL_PM group, most of these genes were significantly up-regulated in the PL_PM group, except for 4CL, CAD, β -GAL, and EXP1. Moreover, the expression of 13 candidate genes, including DIR2 ($r = 0.7977$, $p < 0.05$), NAC like ($r = 0.9464$, $p < 0.01$), EXP1 ($r = 0.8582$, $p < 0.01$), CAD ($r = 0.8015$, $p < 0.01$), β -GAL1 ($r = 0.8440$, $p < 0.05$), β -GAL2 ($r = 0.6675$, $p < 0.05$), 4CL ($r = 0.7466$, $p < 0.05$), ENDOB ($r = 0.7052$, $p < 0.05$), PE ($r = 0.8042$, $p < 0.05$), BHLH ($r = 0.7485$, $p < 0.05$), PG3 ($r = 0.6819$, $p < 0.05$), CEL ($r = 0.8732$, $p < 0.05$), and PRX2 ($r = 0.8325$, $p < 0.01$), showed strong correlations with the RNA-Seq data, and 7 genes showed poor correlation with the corresponding protein expression (Fig. 4; Table 3).

Table 3
Correlation analysis between cell wall related gene and protein expression levels

Gene/Protein name	Gene ID	Pearson correlation efficient	P-value	Numbers
Gene				
DIR2	TRINITY_DN135342_c3_g3	0.7977	0.0100	18
HCT	TRINITY_DN138969_c0_g6	0.2848	0.4577	18
CSLG	TRINITY_DN136551_c1_g11	-0.06589	0.8663	18
NAClike	TRINITY_DN136342_c2_g5	0.9464	0.0001	18
EXP1	TRINITY_DN141308_c1_g4	0.8582	0.0031	18
CAD	TRINITY_DN141875_c0_g4	0.8015	0.0094	18
β -GAL1	TRINITY_DN138388_c1_g1	0.8440	0.0169	18
β -GAL2	TRINITY_DN142386_c5_g1	0.6675	0.0495	18
PG	TRINITY_DN196976_c0_g1	0.7863	0.0636	18
4CL	TRINITY_DN141686_c0_g5	0.7466	0.0208	18
ENDOB	TRINITY_DN138197_c1_g5	0.7052	0.0338	18
NAC	TRINITY_DN131789_c1_g4	0.1755	0.6515	18
PE	TRINITY_DN143028_c0_g1	0.8042	0.0292	18
PRX	TRINITY_DN142336_c1_g1	0.4798	0.1912	18
BXL	TRINITY_DN141432_c1_g2	0.1761	0.6505	18
PL	TRINITY_DN143250_c1_g6	0.6021	0.0862	18
BHLH	TRINITY_DN138043_c3_g7	0.7485	0.0327	18
PG3	TRINITY_DN142042_c0_g2	0.6819	0.0430	18
CEL	TRINITY_DN76417_c0_g1	0.8732	0.0103	18
PRX2	TRINITY_DN141264_c1_g1	0.8325	0.0054	18
Protein				
DIR1	TRINITY_DN135837_c0_g4	0.8316	0.0402	18
PG2	TRINITY_DN142943_c1_g1	0.8336	0.0101	18
EXP1	TRINITY_DN141308_c1_g4	-0.7482	0.0204	18

Gene/Protein name	Gene ID	Pearson correlation efficient	P-value	Numbers
PRX3	TRINITY_DN139379_c0_g3	-0.1040	0.7900	18
F26G	TRINITY_DN142424_c1_g1	0.8907	0.0013	18
BGLU33	TRINITY_DN137437_c3_g1	0.7248	0.0272	18
PE	TRINITY_DN143028_c0_g1	0.7596	0.0176	18
PRX	TRINITY_DN142336_c1_g1	0.7489	0.0202	18
BXL	TRINITY_DN141432_c1_g2	0.6766	0.0453	18
α -HY	TRINITY_DN141662_c1_g4	0.8270	0.0060	18
PL	TRINITY_DN143250_c1_g6	0.4460	0.2289	18
PRX2	TRINITY_DN141264_c1_g1	0.5422	0.1315	18
PRX4	TRINITY_DN137008_c1_g6	0.7481	0.0204	18
PRX5	TRINITY_DN139660_c0_g1	0.6748	0.0462	18
ENBG	TRINITY_DN141880_c0_g1	0.8917	0.0029	18
β -GAL1	TRINITY_DN138388_c1_g1	-0.8215	0.0234	18
β -GAL2	TRINITY_DN142386_c5_g1	0.7797	0.0132	18

Quantitative proteome analysis

To understand the molecular mechanisms of peel cracking in *A. trifoliata* fruits, a quantitative proteomics analysis was also performed using the TMT platform and LC-MS/MS analysis during fruit development, to complement the transcriptome analysis. Accordingly, a total of 812 625 spectra, 68 151 identified spectra, 12 456 peptides, and 10 572 unique peptides, were found by proteomic analysis and 2839 proteins were identified (Table 1). In terms of protein mass distribution, proteins with molecular weight greater than 9 kDa have a wide range and good coverage, with the maximum distribution area of 10–40 kDa (Fig. S3a). The peptide quantitative analysis of the proteins showed that the protein quantity decreased with the increase of the matching peptide (Fig. S3b).

Among of these proteins, 240 were identified as differentially abundant proteins (DAPs) using a fold-change > 1.2 and < 0.83 with $p < 0.05$ as the up-regulated and down-regulated threshold, respectively. In the comparison between PM and PS, 84 proteins were more abundant and 106 proteins were less abundant in the PM than PS and are shown in a volcano plot in Fig. S2c; 20 DAPs were more abundant and 13 DAPs were less abundant in the PL than PM, and are shown in a volcano plot in Fig. S2d, and 17 were co-expressed at PM_PS and PL_PM.

Functional classification of the identified DAPs

Bioinformatics analysis of DAPs was carried out based on protein functional classifications and hierarchical cluster analysis. GO analysis showed that most of the DAPs in the biological processes (BP) were involved in cellular responses to the chemical stimulus and cellular oxidant detoxification processes in the PM_PS, and the metabolic and macromolecule metabolic processes in the PL_PM, respectively. The oxidoreductase and antioxidant activity in PM_PS and the cytoplasmic part and protein-containing complex in PL_PM, were the highest portions of DAPs in the category of molecular functions (MF). Extracellular regions in the PM_PS, and cytoplasmic parts of the PL_PM, were the highest portions of DAPs in the category of cell components (CC) (Fig. 2c-2d).

Moreover, a KEGG pathway analysis was carried out to further evaluate the DAPs. In the comparison between the PM_PS and PL_PM, many DAPs were enriched in two-component systems and ribosome pathways. Notably, most of the DAPs involved in cell wall-related DAPs were upregulated in both the PM_PS and PL_PM groups. While those DAPs involved in phenylpropanoid pathways and peroxisomes were downregulated in the PM_PS and PL_PM groups, respectively (Fig. 2g-2h).

Hierarchical cluster analysis was performed to further explore the expression changes in the cell wall-related DAPs. A total of 40 cell wall-related DAPs were clustered closely, in both the PM_PS and PL_PM groups. Most of these were involved in pentose and glucuronate interconversions, the phenylpropanoid pathway, galactose metabolism, starch and sucrose metabolism, amino sugar and nucleotide sugar metabolism, and cell wall metabolism related proteins (Fig. 3g-l).

Confirmation of fruit cell-wall related proteins in *A. trifoliata* by qPCR

To validate the results of the TMT data, an expression and correlation analysis between the qPCR and the FPKM values was obtained when the TMT was performed. There were 17 genes selected that had shown differential expression patterns in the PM_PS and PL_PM groups, and the results of the qPCR of these genes are shown in Fig. 6. Specifically, in the PM_PS group, the phenylpropanoid pathway related genes (PRX, PRX3, PRX4 and PRX5), galactose metabolism related genes (β -GAL1 and β -GAL2), amino sugar and nucleotide sugar metabolism related gene (BXL), and pentose and glucuronate interconversion related genes (PG and PG2), were significantly down-regulated. While starch and sucrose metabolism related genes- glucan endo-1,3-beta-glucosidase (ENBG), Furostanol glycoside 26-O-beta-glucosidase (F26G), pentose and glucuronate interconversion related genes (PL and PE), and cell wall metabolism related genes (EXP1, DIR1) were significantly up-regulated. In the PL_PM group, most of the genes were up-regulated, except for DIR1, EXP1, BGLU33, PG, PG2, β -GAL1, and alpha/beta hydrolase (α -HY). Moreover, the expression of the 14 candidate genes, including DIR1 ($r = 0.8316$, $p < 0.05$), PG2 ($r = 0.8336$, $p < 0.05$), EXP1 ($r = -0.7482$, $p < 0.05$), F26G ($r = 0.8907$, $p < 0.01$), BGLU33 ($r = 0.7248$, $p < 0.05$), PE ($r = 0.7596$, $p < 0.05$), PRX ($r = 0.7489$, $p < 0.05$), BXL ($r = 0.6766$, $p < 0.05$), α -HY ($r = 0.8270$, $p < 0.01$), PRX5 ($r =$

0.7481, $p < 0.05$), PRX6 ($r = 0.6748$, $p < 0.05$), β -GAL2 ($r = 0.7797$, $p < 0.05$), ENBG ($r = 0.8917$, $p < 0.01$) and β -GAL1 ($r = -0.8215$, $p < 0.05$), showed strong correlations with the RNA-Seq data, 3 genes showed poor correlation with their corresponding protein expression (Fig. 5; Table 3).

Comparative analysis between protein abundance and gene expression levels

To evaluate the relationships between the transcriptomic and proteomic changes during fruit pericarp cracking, the quantitative data for DEGs and DAPs was used for correlation analysis. According to the association analysis, 1904 shared DEGs and 17 shared DAPs were identified in the comparison between the PM_PS and PL_PM. Among the shared DEGs, 1123 were up-regulated and 781 were down-regulated in the PM_PS, whereas 808 were up-regulated and 1096 were down-regulated in PL_PM. Of the common DAPs, 9 were more abundant and 8 were less abundant in PM_PS than in PL_PM, while 8 were more abundant and 9 were less abundant in PL_PM than in the PM_PS. (Table 3). Furthermore, 14 and 4 DAPs and their corresponding DEGs were identified in the PM_PS and PL_PM groups, respectively. Of these, 12 DAPs (4 with increased abundance and 8 with decreased abundance) and 4 DAPs (2 with increased abundance and 2 with decreased abundance) were regulated in the same direction as their corresponding DEGs in the PM_PS and PL_PM groups, respectively (Fig. 6a-6b). There were more DEGs than DAPs in both the PM_PS and PL_PM groups, with significant differences of the trends in transcript levels and protein abundance.

Furthermore, the fold-changes of the DAPs were significantly negatively correlated with their corresponding DEGs by Pearson's correlation tests ($r = 0.03$ and 0.11 , $p < 0.01$, in PM_PS and PL_PM, respectively), indicating a poor correlation between transcript levels and protein abundance (Fig. 6c-6d).

Identification of DAPs and DEGs associated with candidate pathways

To further clarify the biological functions of the co-regulated DEGs-DAPs genes, an enrichment analysis was conducted based on the GO analysis. The largest groups within the biological processes category were those linked with metabolic and cellular processes; catalytic activity and binding mainly included membranes, cell and cell parts were predominant in the category of molecular functions, both in PM_PS and PL_PM (Fig. S4a-b).

Moreover, a pathway enrichment analysis was conducted in the PM_PS and PL_PM groups, based on the KEGG database. In the PM_PS, 14 DAPs were significantly enriched in 7 pathways, both in DEGs and DAPs, which including fructose and mannose metabolism pathway; phenylpropanoid biosynthesis; glutathione metabolism; ubiquinone and other terpenoid-quinone biosynthesis; pentose and glucuronate interconversions; amino sugar and nucleotide sugar metabolism, and galactose metabolism. In the

PL_PM group, 4 DAPs were significantly enriched in the 3 pathways, both in the DEGs and DAPs, which included a calcium signaling pathway, pentose and glucuronate interconversions, and galactose metabolism pathway (Fig S4c-4d). The comparative analysis showed that 2 pathways, including pentose and glucuronate interconversions (2 DEGs and DAPs), and galactose metabolism pathways (2 DEGs and DAPs) were shared in the transcriptome and proteome data, for both the PM_PS and PL_PM groups. While the 2 DEGs and DAPs involved in the phenylpropanoid biosynthesis pathways were only shared by DAPs and DEGs in the PM_PS group. Therefore, involved in the shared DAPs and DEGs of the two pathways were further investigated as candidate genes related to the *A. trifoliata* fruit peel cracking.

Analysis of proteins expressed in *A. trifoliata* fruits identifies genes that might play relevant roles in fruit peel cracking

Of the two candidate pathways, a total of 13 and 3 DAPs and 28 and 46 DEGs were detected in the PM_PS and PL_PM groups, respectively, where indicated that more DEGs than DAPs were involved in the peel cracking of *A. trifoliata* fruit (Fig. 3a-3H). Moreover, most of these DEGs and DAPs were down-regulated in the PM_PS group, and up-regulated in the PL_PM group, both in the transcriptome and proteome data. Notably, PL was up-regulated in the PM_PS group, and the PE and β -GAL2 genes were up-regulated in the PL_PM both in the transcriptome and proteome data and showed strong positive correlations. While β -GAL1 had a negative correlation between the transcriptome and proteome data, as it was down-regulated in the transcriptome and up-regulated in the proteome. Furthermore, the results indicated that most of the genes encoding DAPs were not included in the DEGs, which was accorded with the observed differences between the proteome and transcriptome data.

Additionally, the signal transmissions underlying the *A. trifoliata* fruit peel cracking were studied by analyzing the protein–protein interactions using the STRING database. In the interaction network of these DEGs and DAPs (Fig. 7), there were two pathways identified, PE interacted with PL and β -fructofuranosidase was interacted with raffinose synthase. These results indicated that PE and PL should be further investigated as candidate proteins in the development of *A. trifoliata* fruit peel cracking.

Discussion

The structural changes of the pericarp

Fruit cracking is a key factor that influences the utilization rate of biofuel feedstock, marketability of fruits, and causes significant loss of yield and commercial value. So elucidating the molecular mechanisms that regulate fruit cracking played key roles in the utilization of *A. trifoliata* for biofuels. But however there have been no studies and no identified genes in *A. trifoliata* that are related to fruit cracking. In this study, the structures of the different development stages of the fruits were observed, and there were significant differences in the different stages. In the non-cracking stage, the arrangement of

pericarp cell and cuticle were dense, small intercellular space and distributed continuously. While the cells became thinner and the number of cell layers reduced in the initial cracking stages, and these characteristics increased at the total cracking fruit stage; the arrangement of the cells was loose and the cells began to degrade, compared with those not cracking (Fig. 1). These results were consistent with previous results that showed that the cell wall structures of the fruit pericarp had poor integrity, loose cell wall structures, deformed cell layers, presented larger spaces and break down cell wall during fruit cracking [25]. Studies have demonstrated that the arrangement of the subcutaneous layers of the cells were relatively regular, and that cell layers had a closer arrangement in the cracking-resistant tomato genotype [26]. The biomechanical behaviors of the plant cell walls changed due to the pectin degradation [27, 28]. Moreover, the modifications reduced the strength of the fruit pericarp, because the cell polysaccharides degraded the cell wall hydrolases, and the formation of phenolic cross-linking cell wall structural components catalyzed by cell wall peroxidase, resulted in changes to the pericarp development and fruit cracking [25]. The jujube fruit cracking might be related to the changes of cell wall structure, and the rearrangement of the cell wall at the later stages of fruit ripening [29]. Therefore, the changes of the structures and the ultrastructures of the pericarp cell wall between the unripe and ripe fruits may play a key role in the tendency of *A. trifoliata* fruits to crack.

The reference transcriptome and proteome generated by RNA-seq and TMT

Fruit cracking is a complex phenomenon that is caused by a series of environmental, physiological, biochemical, and genetic changes during fruit ripening. In this study, the differences in the transcriptome and proteome were firstly investigated based on RNA-seq and TMT during different development stages. Moreover, the transcriptome database was used for protein identification in this study, so the quality of the sequencing and assembly of the transcriptome data was crucial for the subsequent analyses. A total of 46.45, 46.35, 53.46 million clean reads for PS, PM and PL, respectively, and 186,054 unigenes (> 200 bp) were assembled and identified in a reference transcriptome of *A. trifoliata* pericarp, which is much more than those of the other Ranunculales data such as *A. trifoliata* (11,749 by Yang et al. [22]; 65,757 by Niu et al. [1]), *Dysosma aurantiocaulis* (53,929) [30], and *Dysosma versipellis* (44,855) [31]. A total of 100% and 56.16% of the unigenes were identified in the present study, and matched to the NR and SwissProt databases, respectively. The level of identification was higher than that presented in other similar studies that have used the same technologies, such as 32.14% and 21.91% [22], and 29.95% [1]. The percentages of the Q30 bases (91.58 – 93.75%) and GC (45.94–48.48%) that were obtained were similar to the content levels reported in other studies using RNA-seq approach, such as (89.06 – 93.33%) and (43.20–43.93%) [22], and 45.10%, 96.31% [1]. As there were no previous proteome studies reported on the Lardizabalaceae family, a total of 812 625 spectra, 68 151 identified spectra, 12 456 identified peptides, and 10 572 unique peptides, were found by proteomic analysis and 2839 proteins were identified in a reference proteome of the *A. trifoliata* pericarp. These results indicated that the *A. trifoliata*

pericarp transcriptome and proteome presented here were comprehensive, accurate, and useful tools for future genetic research of *A. trifoliata* fruit cracking and of other Lardizabalaceae and fruit species.

Identification of potential regulators and metabolism pathways involved in fruit cracking

In this study, more DEGs were detected in the PL_PS group, while more DAPs were detected in the PM_PS group. The enrichment analysis of GO and KEGG pathway revealed that most of the DEGs and DAPs were involved in the metabolic process and ribosome metabolites. Moreover, more down-regulated and up-regulated DEGs and DAPs were detected at the fruit cracking stage than the non-cracking stage. Previous studies indicated that numerous changes, such as increased respiration, fruit softening, metabolic activities in compounds, structural polysaccharides, and a softening of textures generated, accompanied the fruit ripening processes [32], suggesting that these changes may have some correlation with *A. trifoliata* fruit cracking.

The main components of plant cell walls, include lignin, cellulose, hemicellulose, and pectin, and these would be modified or degraded during cell differentiation and dehiscence in plants, and are linked to fruit cracking [32, 33]. Previous studies have suggested that several cell-wall related genes are susceptible or resistant to fruit cracking, including CAD, 4CL, HCT; cell-wall modification genes- β -GAL, EXP, PE and PG; [34–36]; cellulase gene CELLULASE6 (CEL6) was also shown to be essential for silique dehiscence in *Arabidopsis* [37]. Additionally, cell-wall related transcription factor NAC also affected cell differentiation, seed abscission and fruit dehiscence [38]. In this study, cell wall-related DEGs and DAPs, including the galactose metabolism, phenylpropanoid pathway, pentose and glucuronate interconversions, starch and sucrose metabolism, amino sugar and nucleotide sugar metabolism showed consistent downregulation and upregulation in the cracking stage, compared with the non-cracking stage. Therefore, DEGs and DAPs, which are related to the cell wall synthesis and degradation was selected as candidates, and their gene expression levels were further investigated. Among them, 27 (13 in DEGs and 14 in DAPs) showed strongly correlated expression at the gene and protein expression levels, indicating that the cell-wall metabolic pathways may play a crucial role in the regulation of *A. trifoliata* fruit cracking, which is in accordance with the results of the structure changes generated in *A. trifoliata* pericarp.

DEGs and DAPs involved in candidate pathways determine *A. trifoliata* fruit cracking

In order to identify the candidate DEGs and DAP involved in the fruit cracking, an integrated quantitative proteomics and transcriptomics analysis was performed in this study. There were more DEGs (11205) than DAPs (240) in both the PM_PS and PL_PM groups, and 1904 shared DEGs and 17 shared DAPs were identified during the comparisons between the PM_PS and PL_PM groups. There was discordance between the transcript levels and protein abundance, which was similar to the reports for pomegranate

[39] and rice [40]. The inconsistency between mRNA and proteins could be attributed to the post-transcriptional and post-translational regulation, reversible phosphorylation, splicing events in cells, and translation efficiency [39, 41]. This inconsistency showed that the efficiency of gene translation and posttranslation processes is an important regulatory factor during *A. trifoliata* fruit development. Therefore, the consistency between the transcriptional level and the trend of protein abundance between the transcript levels and protein abundance were performed using Pearson's correlation tests. In summary, a poor correlation ($r = 0.03$ and 0.11) between the protein abundance and the expression of the corresponding genes in the PM_PS and PL_PM groups. A similar result, showing a relatively poor correlation ($r = 0.27-0.40$) was identified in a previous study [42]. A possible explanation is that the regulation from mRNA to protein is a complex process, and the changes of the protein abundance were generated after its corresponding transcript was stabilizing [43]. Therefore, both the transcriptomic and proteomic data play key roles in deciphering the molecular processes involved in fruit cracking.

Furthermore, a pathway enrichment analysis indicated that 14 genes from the detected DAPs and DEGs were significantly enriched in 7 pathways in the PM_PS, and 4 DAPs were significantly enriched in 3 pathways in the PL_PM group for both the DEGs and DAPs, based on the KEGG database results. Among which, two cell wall related pathways, including the pentose and glucuronate interconversions (2), and galactose metabolism pathway (2), were the common pathways shared by the DAPs and DEGs in both the PM_PS and PL_PM groups, while the phenylpropanoid biosynthesis (2) pathway was only shared by the DAPs and DEGs in the PM_PS group. These detected proteins belong to corresponding pathways, especially the glucuronate interconversions and galactose metabolism pathways that could be core regulators in *A. trifoliata* fruit cracking.

Candidate DEGs and DAPs might play key roles in fruit peel cracking

Cell wall hydrolases, such as BG, PL and PE, could reduce the strength of the fruit peel by cooperatively degrading the cell wall polysaccharide networks. While the cell wall PRX could catalyze the lignification of the phenolic cross-linking through the cell wall structural components. These modifications contributed to changes in the pericarps developmental status and mechanical properties, fruit cracking, and fruit-splitting processes [25, 44]. For example, CanBGal3 displayed significant hydrolytic activity in cell wall pectin-degradation, and might play role in cell wall loosening [45]. Studies found that suppressing the PL gene could not only greatly increase the firmness of the full ripe fruits and reduce the postharvest softening in strawberry [46, 47]; but also increased the cell separation, and the cellulose and hemicellulose contents, but reduced the water-soluble and total pectin in SIPL-RNAi fruits, compared with the WT, suggesting that the SIPL gene participated in the pericarp cell wall rearrangement during fruit softening [48]. Additionally, suppressing PE gene expression also significantly altered fruit cracking and viscosity, but had little effect on the fruit firmness. PE and β -Gals were differentially expressed in the cracked fruits, compared with the non-cracked fruits of the litchi pericarp, which were identified as candidate genes for fruit cracking in litchi [49]. Schuch et al. [50] indicated that suppressing PE gene

expression significantly reduce cracking in tomato. Considered a key enzymes in lignin-degradation, PRX can induce cell wall loosening [51, 52]. Consistent with these findings, in this study, 6 proteins were functionally annotated among of thousands of identified DEGs and DAPs, both in transcriptome and proteome data. Among which, PE and PL gene were up-regulated both at gene and protein expression levels in both the PM_PS and PL_PM groups, during fruit ripening, and PE interacted with the PL protein by protein–protein interactions using the STRING database (Fig. 7). Moreover, D-Gal, PRX, and PRX2 were down-regulated in the PM_PS group and up-regulated in PL_PM group, both at gene and protein expression levels during fruit ripening. The significantly increased expression of these cell-wall related genes, especially the PE, PL and D-Gal genes, suggests their dynamic roles in *A. trifoliata* fruit cracking. However, the BGal gene showed an opposite trend of expression at gene and protein expression levels in both the PM_PS and PL_PM groups, which may be because the changes in the protein abundance were generated after its corresponding transcript had been stabilized, so the role of the BGal gene in fruit cracking needs to be further studied. The regulatory functions of the other identified DEGs and DAPs in the pathways related to pentose and glucuronate interconversions, phenylpropanoid biosynthesis and galactose metabolism pathways have not been characterized, combined analysis of the other data may further illuminate their functions in relation to fruit cracking.

Conclusions

In conclusion, this study provides the first comprehensive transcriptome and proteome data that indicate that there is a complex transcriptional and translational network in the regulating of fruit cracking in *A. trifoliata*. These results suggested that these candidate genes/proteins may play important roles in fruit ripening and cracking of *A. trifoliata* fruits. The results described here provide important insights into *A. trifoliata* fruit ripening and indicate that cell wall related genes/proteins play key roles in the process of fruit cracking and pave the way for further investigations into their molecular mechanisms and applications of *A. trifoliata* as a bioenergy crop.

Materials And Methods

Plant materials

The *A. trifoliata* sample Nong No.8 was used as the research material in this study. Nine year-old trees of Nong No.8 were growing in the nursery at Hunan Academy of agricultural sciences, Changsha, P. R. China. According to our observations, in Hunan Province, the blooming stage for the germplasm of Nong No.8 was in early April, when 50% of the *A. trifoliata* flowers were in bloom, and the fruit usually ripen in early October. Fruit of different developmental stages, including PS, PM and PL were randomly taken from the same Nong No.8 tree as a biological replicate when the fruits were ripe; in total, three biological replicates were collected for each stage. The pericarps of the fruits were cut longitudinally into several parts after harvesting and frozen in liquid nitrogen and stored at -80 °C, prior to transcriptome, proteome, and qPCR analyses.

Anatomical structure of pericarp

Anatomy of the pericarp samples taken from the Nong No.8 fruits at different development stages including PS, PM, and PL, were prepared for paraffin sections and scanning electron microscopy (SEM) was carried out according to Chen et al. [53]. Pericarp samples were fixed directly in the field using FAA [70% ethyl alcohol + 38% methyl aldehyde + 25% acetic acid (16:1:1)]. Then the tissues were subsequently dehydrated through an ethanol series with increasing ethanol concentrations and embedded in the paraffin. Subsequently, paraffin sections were stained with Safranin O-staining and observed by Axio Imager (Zeiss, Oberkochen, Germany), upright microscopy and images were displayed using Image-pro Plus 6.0 software.

Pericarp samples were fixed in 2.5% glutaraldehyde (pH 7.4) for 4 h under syringe suction, and washed subsequently in phosphate buffer (PBS, 0.1M, pH 7.2). Following fixation, samples were dehydrated with 30, 50, and 70% ethanol for 20 min in turn, then 100% ethanol for 30 min, and then they were dried in liquid carbon dioxide. Dried samples were placed on stubs and sputter-coater with gold film (FEI, America) and observed under a SEM (FEI Verios 460, America).

RNA isolation, library construction and sequencing

Total RNA used for the RNA-seq assays was isolated from three independent replicates of pericarp in the PS, PM, and PL stages, as described by Tao et al [54]. The RNA samples were detected by absorbance ratio of A260/A280 with a Nanodrop ND-1000 system (Thermo Scientific). Pair-end Libraries were prepared using NEBNext® Ultra™ RNA Library Prep Kit for Illumina® (NEB, USA) following the manufacturer's instructions. The mRNA was purified from 3 µg of the total RNA using oligo (dT) magnetic beads followed by fragmentation carried out using divalent cations at elevated temperatures in NEBNext First Strand Synthesis Reaction Buffer. Subsequently, first strand cDNAs were synthesized by random hexamer primer and Reverse Transcriptase (RNase H-) using mRNA fragments, as templates followed by the Second strand cDNA synthesis using DNA Polymerase I, RNaseH, buffer and dNTPs. The synthesized double-stranded cDNA fragments were then purified with AMPure XP system (Beckman Coulter, Beverly, USA). The purified double-stranded cDNA was polyadenylated and adapter-ligated for preparation of the paired-end library. Adaptor-ligated cDNA and adaptor primers were used for PCR amplification. PCR products were purified (AMPure XP system) and library quality was assessed on the Agilent Bioanalyzer 2100 system. Finally, sequencing was performed by Illumina HiSeq2500 instrument by Shanghai Applied Protein technology (Shanghai, China).

Quality control and transcriptome assembly

The raw paired-end reads of fastq format produced from the sequencing were first processed with the in house Perl scripts. Those reads containing adapters, the excess "N" nucleotides with more than 10% of the bases, and reads of low-quality (reads with quality values ≤ 10) were removed by filterfq software.

The Q20, Q30, GC-content, and sequence duplication levels of the obtained clean reads were calculated. The assembly of clean reads to unigene collections was performed using the Trinity software package (<https://github.com/trinityrnaseq/trinityrnaseq/releases>) [55]. The Trinity software consisted of three independent software modules, including Inchworm, Chrysalis, and Butterfly, and the transcripts less than 200 bp in length were discarded. Sequences containing the longest cluster transcripts without redundancy extracted from transcripts can be considered unigenes.

Bioinformatics analyses

The de novo assembled unigenes were annotated in five databases, which include NR, Pfam, the Swiss-Prot, GO, and KEGG pathway database, based on BLAST search with an E-value threshold of $1e^{-5}$. Moreover, in order to further analyze the annotation results, GO and KEGG with an E-value of $1E^{-5}$ were used for functional gene annotation. GO terms could be classified into three categories, including BP, MF and CC. In addition to the GO terms, the pathway maps were determined by KEGG database.

The normalized transcript abundance of the genes was estimated using the FPKM based on the length of the gene and reads count mapped to this gene. DESeq2 R package (1.16.1) software was used to identify the differential expression of the genes (DEGs), and the false discovery rate (FDR) was controlled using Benjamini and Hochberg's approach to adjusted P-value. Genes with an adjusted P-value < 0.05 and absolute fold change of 2 were deemed to be differentially expressed between the two samples. In addition, GO and KEGG pathways enrichment analysis of DEGs was implemented by the clusterProfiler R package. Transcription factor (TF) analysis of DEGs was performed against the PlantTFDB database (<http://planttfdb.cbi.pku.edu.cn/>). The heat map was visualized using heatmap 2.0 in the gplot R package.

Protein extraction

Protein extraction from *A. trifoliata* pericarp was performed from each sample as described previously [56]. The samples were frozen in liquid nitrogen and ground into powder. Add 5 times volume of TCA/acetone (1:9), vortex and mix, place at -20°C for 4 h, $6000 \times g$, centrifuge at 4°C for 40 min. Discard the supernatant and wash the precipitate 3 times with pre-cooling acetone. After the precipitation air drying, the precipitate was redissolved in a buffer (4% SDS, 100 mM Tris-HCl and 1 mM DTT) (pH 7.6). After sonicated and boiled for 15 min, the lysate was centrifuged 40 min, filter the supernatant, quantify the filtrate using the BCA Protein Assay Kit (Bio-Rad, USA).

Trypsin Digestion and TMT Labeling

For digestion, the samples were added to the buffer (4% SDS, 100 mM DTT, 150 mM Tris-HCl pH 8.0) and UA buffer (8 M Urea, 150 mM Tris-HCl pH 8.0) by repeated ultrafiltration (Microcon units, 10 kD). Then

iodoacetamide (100 mM IAA in UA buffer) was putted into the samples to block reduced cysteine residues, incubateing for 30 min in darkness at room temperature. After the filters washed by UA buffer and triethylamine borane (TEAB) buffer in turn, the suspensions were digested with trypsin (Promega, Madison, WI) in TEAB buffer overnight. After trypsin digestion, the samples (100 µg of protein) were categorized to label with 129-tag (PS), 130-tag (PM) and 131-tag (PL) (Thermo Fisher Scientific, Waltham, MA, USA) respectively. Finally, TMT-labeled peptide aliquots were pooled for subsequent fractionation using the Pierce high pH reversed-phase fractionation kit (Thermo scientific).

HPLC Fractionation and LC-MS/MS Analysis

For the fractionation of labeled peptides, samples were loaded onto a reverse phase trap column (Thermo Scientific Acclaim PepMap100, 100 µm * 2 cm, nanoViper C18) connected to the C18-reversed phase analytical column (Thermo Scientific Easy Column, 10 cm long, 75 µm inner diameter, 3 µm resin) in buffer A (0.1% Formic acid) and separated with a linear gradient of buffer B (84% acetonitrile and 0.1% formic acid) at a flow rate of 300 nl/min controlled by IntelliFlow technology. The resultant peptides were further processed using a Q Exactive mass spectrometer (Thermo Scientific) that was coupled to Easy nLC (Thermo Fisher Scientific). Mass spectrometry analysis was performed in positive ion mode, and MS data was acquired using a data-dependent top10 method dynamically choosing the most abundant precursor ions from the survey scan (300–1800 m/z) by the higher-energy collision dissociation (HCD) fragmentation method. Automatic gain control (AGC) was set as 3E6, and maximum inject time to 10 ms; 40.0-s dynamic exclusion duration; 70,000 resolutions with survey scans at m/z 200 and resolution for HCD spectra at 17500 at m/z 200; 2 m/z of isolation width. 30 eV of Normalized collision energy and the underfill ratio was defined as 0.1%.

Sequence database search and data analysis

The raw data were processed by MASCOT engine (Matrix Science, London, UK; version 2.2) and Proteome Discoverer 1.4 software was used to process MS/MS spectra. The search was performed using the following settings based on the *A. trifoliata* database: trypsin for the enzyme and 2 as the maximum missed cleavage allowed; fixed modifications of Carbamidomethyl (C), TMT-6plex (N-term) and TMT-6plex (K), variable modification of oxidation (M); mass tolerance for fragment ions of 0.1 Da while 20 ppm for peptide ions and both at peptide and protein levels of FDR less than 0.01, and only unique peptides of the protein were employed for the proteins identification and quantification. Proteins with a P value less than 0.05 and fold change ≥ 1.2 or ≤ 0.83 within a comparison were recognized as differentially abundant proteins (DAPs).

Functional categorization was performed using GO and KEGG pathways database with a P value ≤ 0.05 . The protein functional network was performed by STRING 9.0 software (<http://string-db.org>). Clustering analysis of the DEPs was performed using Cluster 3.0 (<http://bioservices.capitalbio.com/xzzq/rj/3885.shtml>) and the Java Treeview software

(<http://jtreeview.sourceforge.net>). Correlations were analyzed based on the DEGs and DEPs, and Person correlation tests were conducted for each comparison group, including PM vs PS, PL vs PM.

Reverse transcription real-time quantitative PCR

The method of Total RNA extraction and synthesis of cDNA were described previously. Bio-Rad CFX96 Touch detection system (Bio-Rad, Richmond, CA, USA) with SYBR Green PCR master mix (Aidlab Biotechnologies, Co., Ltd) were used for the reactions of each sample. In this study, EF-a gene was used as the internal control gene which was detected by de novo transcriptome sequencing of *A. trifoliata* [57]. Primers for qPCR experiments were designed using Primer 5.0 software (Supplementary Table 1) and those gene sequences were blasted against the NCBI database. The amplification reactions contained 12.5 μ L SYBR Green PCR master mix, 1 μ L cDNA, and 0.5 μ L of each primer in a final reaction volume of 25 μ L. The thermal cycling program began with 3 min at 95 °C, followed by 40 cycles of 95 °C for 10 s, 55 °C for 30 s, Melt Curve 65 to 95 °C, increment 0.5 °C for 5 s. After PCR amplification, the quantitative variation was analyzed using Delta Ct method, and the analysis of statistically significant differences from gene expression was performed by the independent samples t-test analysis at $P < 0.05$ using GraphPad Prism 8 software. Correlation analysis between cell wall-related gene and protein expression was performed by Pearson's correlation coefficient analysis [58].

Abbreviations

Akebia trifoliata: *A. trifoliata*; TMT: tandem mass tag; PL: pectate lyases; PE: pectinesterase; β -GAL: β -galactosidases; PGs: polygalacturonases; EXPs: expansins; XET: xyloglucan endotransglycosylase; IND: INDEHISCENT; PID: PINOID; PIN3: PIN-FORMED3; NGS: Next generation sequencing; PS: the non-cracking stage; PM: the initial cracking stage; PL: the total cracking stage; BLAST: Basic Local Alignment Search Tool; NR: non-redundant protein sequences database; Pfam: Protein families database; GO: Gene Ontology database; KEGG: Kyoto Encyclopedia of Genes and Genomes pathway database; DEGs: differently expressed genes; qPCR: Reverse transcription real-time quantitative PCR; FPKM: fragments per kilobase per million reads mapped; 4CL: 4-coumarate-COA-ligase; PRX: peroxidase; CAD: cinnamyl alcohol dehydrogenase; 4CL: 4-coumarate-COA-ligase; HCT: shikimate O-hydroxycinnamoyl transferase; CEL6: CELLULASE6; MAN7: MANNANASE7; DAPs: differentially abundant proteins; BP: biological processes; MF: molecular functions; cell components: CC.

Declarations

Authors' contributions

JN designed and performed the study, analyzed the data, and drafted the manuscript. YS, YZ, YS and JC assisted in analysis and interpretation of data, ZS, ML, and JC provided direction for the experimental design, studies, and revised the manuscript. All authors read and approved the final manuscript.

Author details

Institute of Bast Fiber Crops, Chinese Academy of Agricultural Sciences/Key Laboratory of Stem-Fiber Biomass and Engineering Microbiology, Ministry of Agriculture, Changsha 410205, P. R. China.

Acknowledgements

Not applicable.

Competing interests

The authors declare that they have no competing interests.

Availability of data and materials

The data sets used and/or analyzed during the current study are available from the corresponding author on reasonable request.

Consent for publication

Not applicable.

Ethics approval and consent to participate

Not applicable.

Funding

This research was financially supported by the Central Public-interest Scientific Institution Basal Research Fund (1610242019005) and Agricultural Science, Technology Innovation Program (ASTIP) of CAAS (Grant No.2017IBFC).

Publisher's Note Springer

Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

References

1. Niu J, Wang Y, Shi Y, Wang X, Sun Z, Huang K, et al. Development of SSR markers via de novo transcriptome assembly in *Akebia trifoliata* (Thunb.) Koidz. *Genome*, 2019;999:1-15.
2. Zou S, Yao X, Zhong C, Zhao T, Huang H. Effectiveness of recurrent selection in *Akebia trifoliata* (Lardizabalaceae) breeding. *Sci Hortic*, 2019;246:79-85.
3. Du Y, Jiang Y, Zhu X, Xiong H, Shi S, Hu J, Sun W. Physicochemical and functional properties of the protein isolate and major fractions prepared from *Akebia trifoliata* var. *australis* seed. *Food*

- chemistry, 2012;133(3): 923-929.
4. Kitaoka F, Kakiuchi N, Long C, Itoga M, Yoshimatsu H, Mitsue A, Mikage M. Difference of ITS sequences of *Akebia* plants growing in various parts of Japan. *J Nat Med*, 2009;63(3): 368-374.
 5. Ma C, Dong X, Cheng Y. Components of *Akebia trifoliata* seed and the physicochemical indexes of the seed oil. *China Oils and Fats*, 2009;34(9):77-79.
 6. Li L, Yao X, Zhong C, Chen X, Huang H. *Akebia*: a potential new fruit crop in China. *Hortic Sci*, 2010; 45: 4–10.
 7. Zhou Y, Jie B, Wu G, Cao Y, Wu W. Research progress of *Akebia trifoliata* (Thunb.) Koidz. *Agr Biotechnol*, 2017;6(6): 1–8.
 8. Jiang F, Lopez A, Jeon S, de Freitas ST, Yu Q, Wu Z, Mitcham E. Disassembly of the fruit cell wall by the ripening-associated polygalacturonase and expansin influences tomato cracking. *Hortic Res*, 2019;6(1):1-15.
 9. Wang J, Gao X, Ma Z, Chen J, Liu Y. Analysis of the molecular basis of fruit cracking susceptibility in *Litchi chinensis* cv. Baitangying by transcriptome and quantitative proteome profiling. *J Plant Physiol*, 2019;234:106-116.
 10. Chater JM, Garner LC. Foliar nutrient applications to Wonderful pomegranate (*Punica granatum* L.). I. Effects on fruit mineral nutrient concentrations and internal quality. *Sci Hortic*, 2019;244:421-427.
 11. Kang X, Cai J, Chen Y, Yan Y, Yang S, He R, Zhu Y. Pod-shattering characteristics differences between two groups of soybeans are associated with specific changes in gene expression. *Funct Integr Genomic*, 2019;1-10.
 12. Zhai Y, Cai S, Hu L, Yang Y, Amoo O, Fan C, Zhou Y. CRISPR/Cas9-mediated genome editing reveals differences in the contribution of *INDEHISCENT* homologues to pod shatter resistance in *Brassica napus* L. *Theor Appl Genet*, 2019;132(7):2111-2123.
 13. Dawson J, Sözen E, Vizir I, Van Waeyenberge S, Wilson ZA, Mulligan BJ. Characterization and genetic mapping of a mutation (ms35) which prevents anther dehiscence in *Arabidopsis thaliana* by affecting secondary wall thickening in the endothecium. *New Phytologist*, 1999;144(2):213-222.
 14. Mustafa M, Syukur M, Sutjahjo SH. Inheritance of Radial Fruit Cracking Resistance in Tomatoes (*Solanum lycopersicum* L.). In *IOP Conference Series: Earth and Environmental Science*, 2019; 270(1):012032.
 15. Khadivi-Khub A. Physiological and genetic factors influencing fruit cracking. *Acta Physiol Plant*; 2015; 37(1):1718.
 16. Li J, Huang X, Huang H. Comparison of the activities of enzymes related to cell-Wa metabolism in pericarp between litchi cultivars susceptible and resistant to fruit cracking. *J Plant physiol Mol Biol* , 2003; 29(2):141-146.
 17. Khurnpoon L, Siriphanich J, Labavitch JM. Cell wall metabolism during durian fruit dehiscence. *Postharvest Biol Tec*, 2008;48(3):391-401.

18. Zhang Y, Shen YY, Wu XM, Wang JB. The basis of pod dehiscence: anatomical traits of the dehiscence zone and expression of eight pod shatter-related genes in four species of Brassicaceae. *Biol Plantarum*, 2016;60(2):343-354.
19. Ha BK, Kim HK, Kang ST. Mapping QTLs with epistatic effects and QTL-by-environment interactions for seed coat cracking in soybeans. *Euphytica*, 2012;186(3):933-942.
20. Sorefan K, Girin T, Liljegren SJ, Ljung K, Robles P, Galván-Ampudia CS, Østergaard L. A regulated auxin minimum is required for seed dispersal in *Arabidopsis*. *Nat*, 2009;459(7246):583.
21. Dong Y, Yang X, Liu J. Wang BH, Liu BL, Wang YZ. Pod shattering resistance associated with domestication is mediated by a NAC gene in soybean. *Nat Commun*, 2014;5:3352.
22. Yang H, Liu H, Shi X, Ge F, Zhao Z, Luo H. Transcriptome analysis of *Akebia trifoliata* in ripening process. *J Mountain Agric Biol*. 2016;35(2):46-51.
23. Dai F, Wang Z, Li Z, Luo G, Wang Y, Tang C. Transcriptomic and proteomic analyses of mulberry (*Morus atropurpurea*) fruit response to *Ciboria carunculoides*. *J proteomics*, 2019;193:142-153.
24. Voelckel C, Gruenheit N, Lockhart P. Evolutionary transcriptomics and proteomics: insight into plant adaptation. *Trends Plant Sci*, 2017;22(6):462-471.
25. Li J, Chen JZ. Citrus fruit-cracking: Causes and occurrence. *HPJ*, 2017;3(6):255-260.
26. Yang Z, Wu Z, Zhang C, Hu E, Zhou R, Jiang F. The composition of pericarp, cell aging, and changes in water absorption in two tomato genotypes: mechanism, factors, and potential role in fruit cracking. *Acta Physiol Plant*, 2016;38(9):215.
27. Emaga TH, Robert C, Ronkart SN, Wathélet B, Paquot M. Dietary fibre components and pectin chemical features of peels during ripening in banana and plantain varieties. *Bioresource Technol*, 2008;99(10):4346-4354.
28. Waldron KW, Parker ML, Smith AC. Plant cell walls and food quality. *Compr Rev Food Sci F*, 2003;2(4):128-146.
29. Wang BM, Ding GX, Wang XY, Fu CB, Qin GJ, Yang JQ, Wen PF. Changes of histological structure and water potential of huping jujube fruit cracking. *Scientia Agr Sinica*, 2013;46(21):4558-4568.
30. Mao Y, Zhang Y, Xu C, Qiu Y. Comparative transcriptome resources of two *Dysosma species* (Berberidaceae) and molecular evolution of the *CYP 719A* gene in Podophylloideae. *Mol Ecol Resour*, 2016;16(1):228-241.
31. Guo R, Mao YR, Cai JR, Wang JY, Wu J, Qiu YX. Characterization and cross-species transferability of EST–SSR markers developed from the transcriptome of *Dysosma versipellis* (Berberidaceae) and their application to population genetic studies. *Mol Breeding*, 2014;34(4):1733-1746.
32. Keegstra K. Plant cell walls. *Plant Physiol*, 2010;154(2):483-486.
33. Balbontín C, Ayala H, M Bastías R, Tapia G, Ellena M, Torres C, Silva H. Cracking in sweet cherries: A comprehensive review from a physiological, molecular, and genomic perspective. *Chil J Agr Res*, 2013;73(1):66-72.

34. Li WC, Wu JY, Zhang HN, Shi SY, Liu LQ, Shu B, Wei YZ. De novo assembly and characterization of pericarp transcriptome and identification of candidate genes mediating fruit cracking in Litchi chinensis Sonn. *Int J Mol Sci*, 2014;15(10):17667-17685.
35. Alkio M, Jonas U, Sprink T, van Nocker S, Knoche M. Identification of putative candidate genes involved in cuticle formation in *Prunus avium* (sweet cherry) fruit. *Ann Bot*, 2012;110(1):101-112.
36. Joshi M, Baghel RS, Fogelman E, Stern RA, Ginzberg I. Identification of candidate genes mediating apple fruit-cracking resistance following the application of gibberellic acids 4+ 7 and the cytokinin 6-benzyladenine. *Plant Physiol Bioch*, 2018;127:436-445.
37. He H, Bai M, Tong P, Hu Y, Yang M, Wu H. CELLULASE6 and MANNANASE7 affect cell differentiation and silique dehiscence. *Plant Physiol*, 2018;176(3):2186-2201.
38. Balanzà V, Roig-Villanova I, Di Marzo M, Masiero S, Colombo L. Seed abscission and fruit dehiscence required for seed dispersal rely on similar genetic networks. *Development*, 2016;143(18):3372-3381.
39. Luo X, Cao D, Li H, Zhao D, Xue H, Niu J, Cao S. Complementary iTRAQ-based proteomic and RNA sequencing-based transcriptomic analyses reveal a complex network regulating pomegranate (*Punica granatum* L.) fruit peel colour. *Sci Rep*, 2018;8(1):12362.
40. Chen X, Tao Y, Ali A, Zhuang Z, Guo D, Guo Q, Wang J. Transcriptome and Proteome Profiling of Different Colored Rice Reveals Physiological Dynamics Involved in the Flavonoid Pathway. *Int J Mol Sci*, 2019;20(10):2463.
41. Niu J, Cao D, Li H, Xue H, Chen L, Liu B, Cao S. Quantitative proteomics of pomegranate varieties with contrasting seed hardness during seed development stages. *Tree Genet Genomes*, 2018;14(1):1-12.
42. Muers M. Gene expression: Transcriptome to proteome and back to genome. *Nat Rev Genet*, 2011;12(8): 518.
43. Chen J, Liu SS, Kohler A, Yan B, Luo HM, Chen XM, Guo SX. iTRAQ and RNA-Seq analyses provide new insights into regulation mechanism of symbiotic germination of *Dendrobium officinale* seeds (Orchidaceae). *J proteome Res*, 2017;16(6):2174-2187.
44. Dautt-Castro M, Ochoa-Leyva A, Contreras-Vergara CA, Pacheco-Sanchez MA, Casas-Flores S, Sanchez-Flores A, Islas-Osuna MA. Mango (*Mangifera indica* L.) cv. Kent fruit mesocarp de novo transcriptome assembly identifies gene families important for ripening. *Front Plant Sci*, 2015;6(62):1-11.
45. Martín I, Hernández-Nistal J, Albornos L, Labrador E, Dopico B. β III-Gal is involved in galactan reduction during phloem element differentiation in chickpea stems. *Plant Cell Physiol*, 2013;54(6):960-970.
46. Jiménez-Bermudez S, Redondo-Nevado J, Muñoz-Blanco J, Caballero JL, López-Aranda JM, Valpuesta V, Mercado JA. Manipulation of strawberry fruit softening by antisense expression of a pectate lyase gene. *Plant Physiol*, 2002;128(2):751-759.
47. Uluisik S, Chapman NH, Smith R, Poole M, Adams G, Gillis RB, Samsulrizal N. Genetic improvement of tomato by targeted control of fruit softening. *Nat Biotechnol*, 2016;34(9):950.

48. Yang L, Huang W, Xiong F, Xian Z, Su D, Ren M, et al. Silencing of slpl (solyc03g111690), which encodes a pectate lyase in tomato, confers enhanced fruit firmness, prolonged shelf-life, and reduced susceptibility to gray mold. *Plant Biotechnol J*, 2017;15(9):1544-1555.
49. Li WC, Wu JY, Zhang HN, Shi SY, Liu LQ, Shu B, Wei YZ. De novo assembly and characterization of pericarp transcriptome and identification of candidate genes mediating fruit cracking in Litchi chinensis Sonn. *Int J Mol Sci*, 2014;15(10):17667-17685.
50. Schuch W, Kanczler J, Robertson D, Hobson G, Tucker G, Grierson D, Bird C. Fruit quality characteristics of transgenic tomato fruit with altered polygalacturonase activity. *HortScience*, 199;126(12):1517-1520.
51. Falade AO, Nwodo UU, Iweriebor BC, Green E, Mabinya LV, Okoh AI. Lignin peroxidase functionalities and prospective applications. *MicrobiologyOpen*, 2017;6(1):e00394.
52. Adav SS, Ravindran A, Sze SK. Quantitative proteomic analysis of lignocellulolytic enzymes by *Phanerochaete chrysosporium* on different lignocellulosic biomass. *J proteomics*, 2012;75(5):1493-1504.
53. Chen L, Zhang J, Li H, Niu J, Xue H, Liu B, Cao S. Transcriptomic analysis reveals candidate genes for female sterility in pomegranate flowers. *Front Plant Sci*, 2017;8:1430.
54. Tao X, Fang Y, Xiao Y, Jin YL, Ma XR, Zhao Y, Wang HY. Comparative transcriptome analysis to investigate the high starch accumulation of duckweed (*Landoltia punctata*) under nutrient starvation. *Biotechnol Biofuels*, 2013;6(1):72.
55. Grabherr MG, Haas BJ, Yassour M, Levin JZ, Thompson DA, Amit I, Chen Z. Full-length transcriptome assembly from RNA-Seq data without a reference genome. *Nat Biotechnol*, 2011;29(7):644.
56. Liu Y, Cao D, Ma L, Jin X, Yang P, Ye F, Wei C. TMT-based quantitative proteomics analysis reveals the response of tea plant (*Camellia sinensis*) to fluoride. *J Proteomics*, 2018;176: 71-81.
57. Yang H. Analysis of gene expression on ethylene biosynthesis of *Akebia trifoliata* (Thunb.) Koidz in rippling process. Guizhou University, 2016.
58. Sato A, Zhang T, Yonekura L, Tamura H. Antiallergic activities of eleven onions (*Allium cepa*) were attributed to quercetin 4'-glucoside using QuEChERS method and Pearson's correlation coefficient. *J Funct Foods*, 2015;14:581-589.

Additional Files

Additional file 1: Figure S1. The size distribution of the assembled transcript and unigene sequences of *A. trifoliata*

Additional file 2: Figure S2. Molecular weight and peptide count distribution were identified from TMT proteomics by searching against the database. a Distribution of the proteins that were identified among different molecular weights. b Distribution of peptide count of the proteins were identified from TMT data.

Additional file 3: Figure S3. Volcano plot depicting the proteomics data of *A. trifoliata*. a Volcano plot depicting the proteomics data in PM_PS. b Volcano plot depicting the proteomics data in PL_PM.

Absolute log₂ fold change and protein expression ratio were plotted on the y-axis and x-axis, respectively. Horizontal dotted line presents p values of 0.05 cut-off position while the vertical dotted lines discriminate between proteins having absolute log₂ fold change of 1. Red dots represent proteins with $p < 0.05$ and absolute log₂ fold changes above 1. Black dots indicate no difference in protein expression.

Additional file 4: Figure S4. GO and KEGG pathway functional enrichment analysis of co-regulated genes and proteins in *A. trifoliata*. a-b GO enrichment analysis of co-regulated genes and proteins in PM_PS and PL_PM, respectively. c-d KEGG pathways enrichment analysis of co-regulated genes and proteins in PM_PS and PL_PM, respectively.

Additional file 5: Table S1. Sequences of specific primers used for qPCR experiments.

Additional file 6: Table S2 Sequencing statistics for *A. trifoliata*.

Additional file 7: The statistical data of unigenes annotation.

Figures

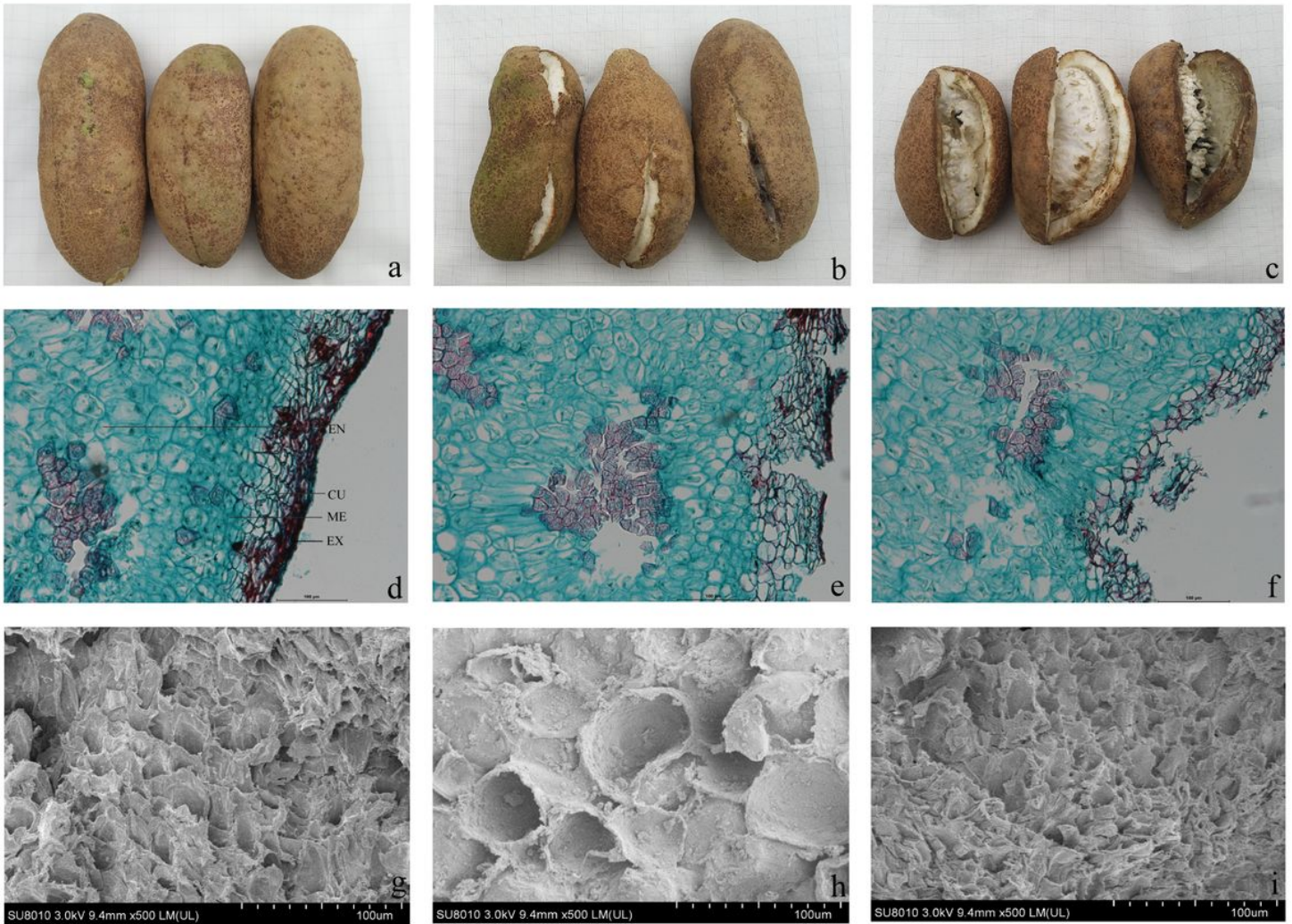


Figure 1

Changes in pericarp structure during fruit development. a-c The morphologies of PS, PM and PL, respectively. d-f These semi-thin slices of pericarp in PS, PM and PL, respectively. g-i The scanning electron microscope pictures of pericarp in PS, PM and PL, respectively. EX. Exocarp; ME. Mesocarp; EN. Endocarp; CU. Cuticle.

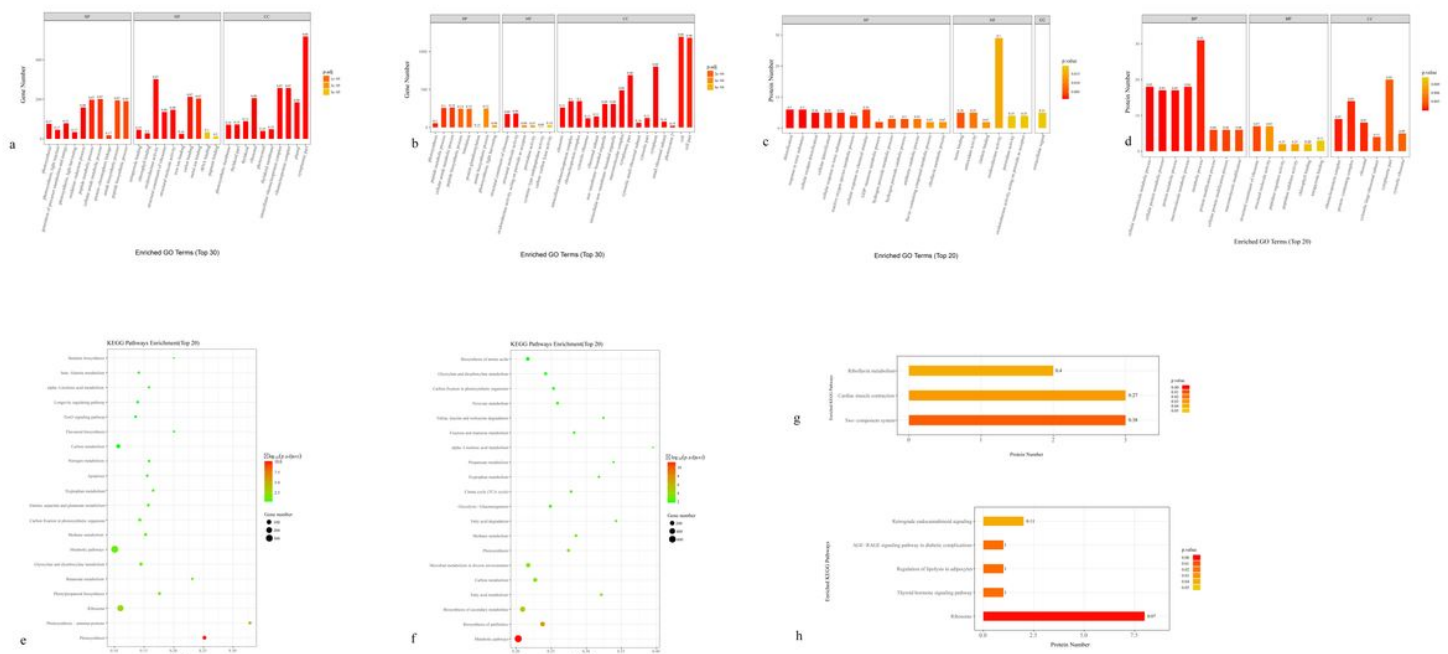


Figure 2

GO classifications and KEGG pathways associated with DEGs and DAPs in PM_PS and PL_PM. a-d GO classifications of DEGs and DAPs in PM_PS and PL_PM, respectively. e-h KEGG pathways of DEGs and DAPs in PM_PS and PL_PM, respectively.

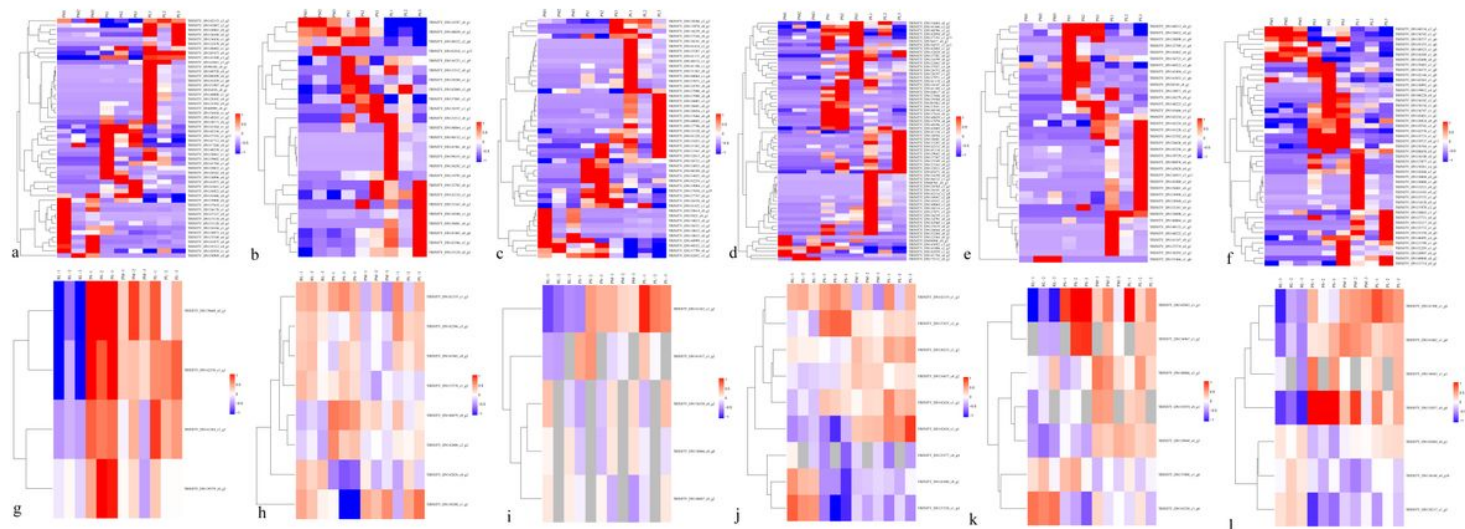


Figure 3

Heatmap analysis of DEGs and DAPs based on transcriptomic and proteomic, which are associated with cell wall metabolic processes. Red indicates significantly upregulated proteins, and green indicates significantly downregulated proteins. Black indicates proteins with no significant changes. a and g Heat map of phenylpropanoid biosynthesis-associated gene and protein expression. b and h Heat map of

galactose metabolism-associated gene and protein expression. c and i Heat map of amino sugar and nucleotide sugar metabolism-associated gene and protein expression. d and j Heat map of starch and sucrose metabolism-associated protein expression. e and k Pentose and glucuronate interconversions metabolism-associated gene and protein expression. f Cell wall related transcription factor. l cell wall metabolism-associated protein expression.

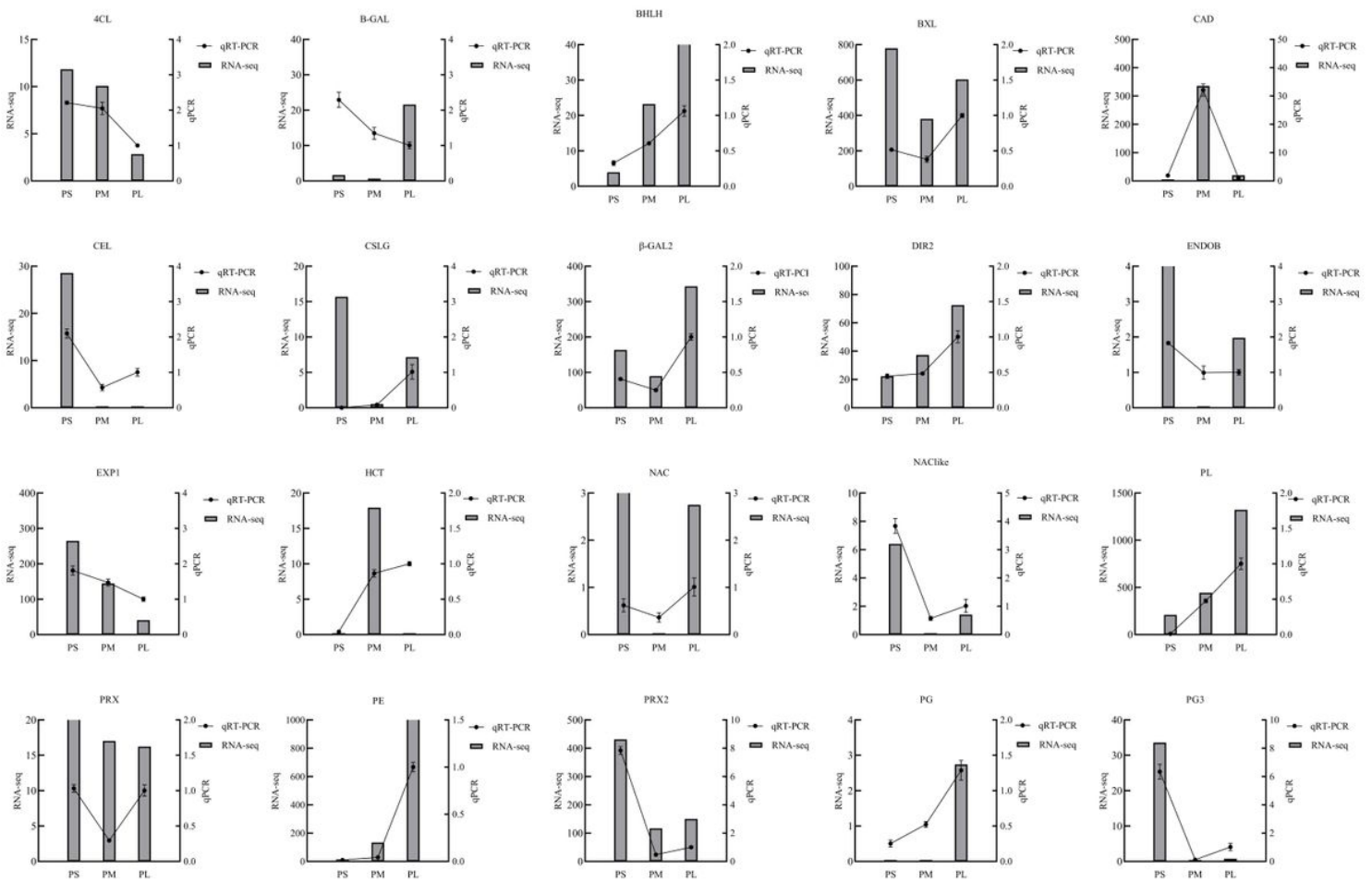


Figure 4

Validation and expression analysis of selected genes and proteins using qPCR and RNA-seq. The expression levels of the genes revealed by RNA-seq (Left y-axis) and qPCR (right y-axis). Histograms were gene expression detected by RNA-seq. Line graphs were relative expression validated by qPCR.

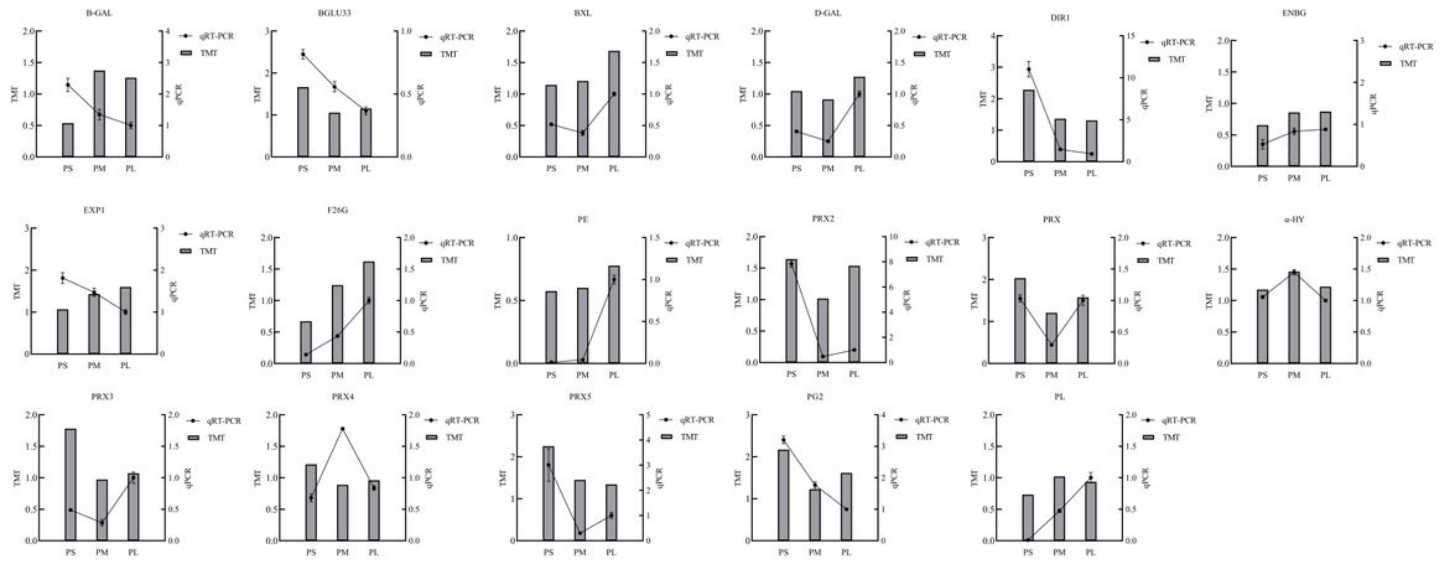


Figure 5

Validation and expression analysis of selected proteins and proteins using qPCR and TMT. The expression levels of the genes revealed by TMT (Left y-axis) and qPCR (right y-axis). Histograms were protein abundance detected by TMT. Line graphs were relative expression validated by qRT-PCR.

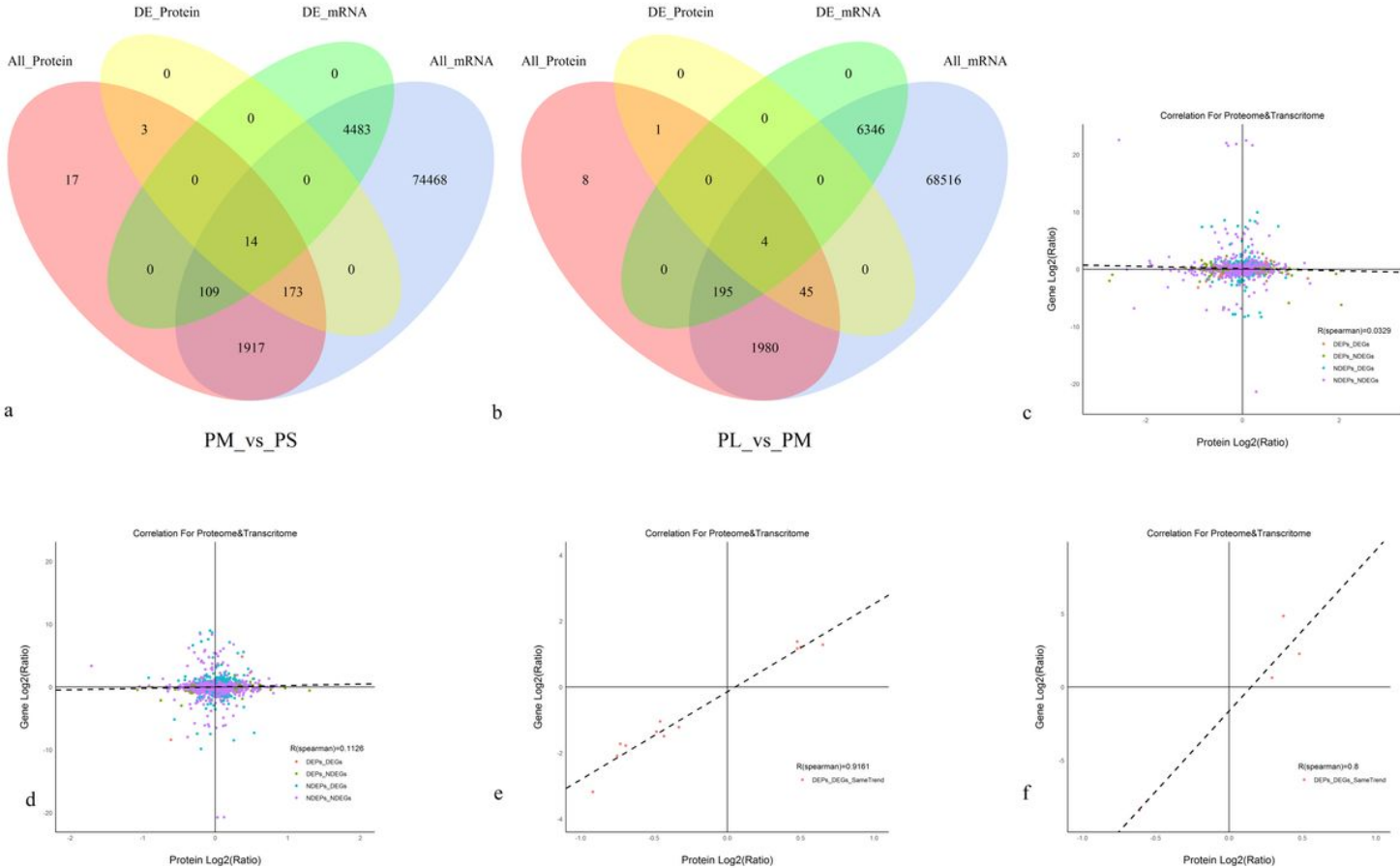


Figure 6

Correlations between mRNA and protein expression. a Venn diagram of genes quantified in the transcriptome (blue) and proteome (pink), DEGs (green) and DAPs (yellow) in PM_PS. b Venn diagram of genes quantified in the transcriptome (blue) and proteome (pink), DEGs (green) and DAPs (yellow) in PL_PM. c Scatterplot of the relationship between genes identified in both the transcriptome and proteome in PM_PS. d Scatterplot and correlation coefficients between DEGs and DEPs in PL_PM. e Scatterplot and correlation coefficients between DEGs and DEPs (the same trend) in PM_PS. F Scatterplot and correlation coefficients between DEGs and DEPs (the same trend) in PL_PM.

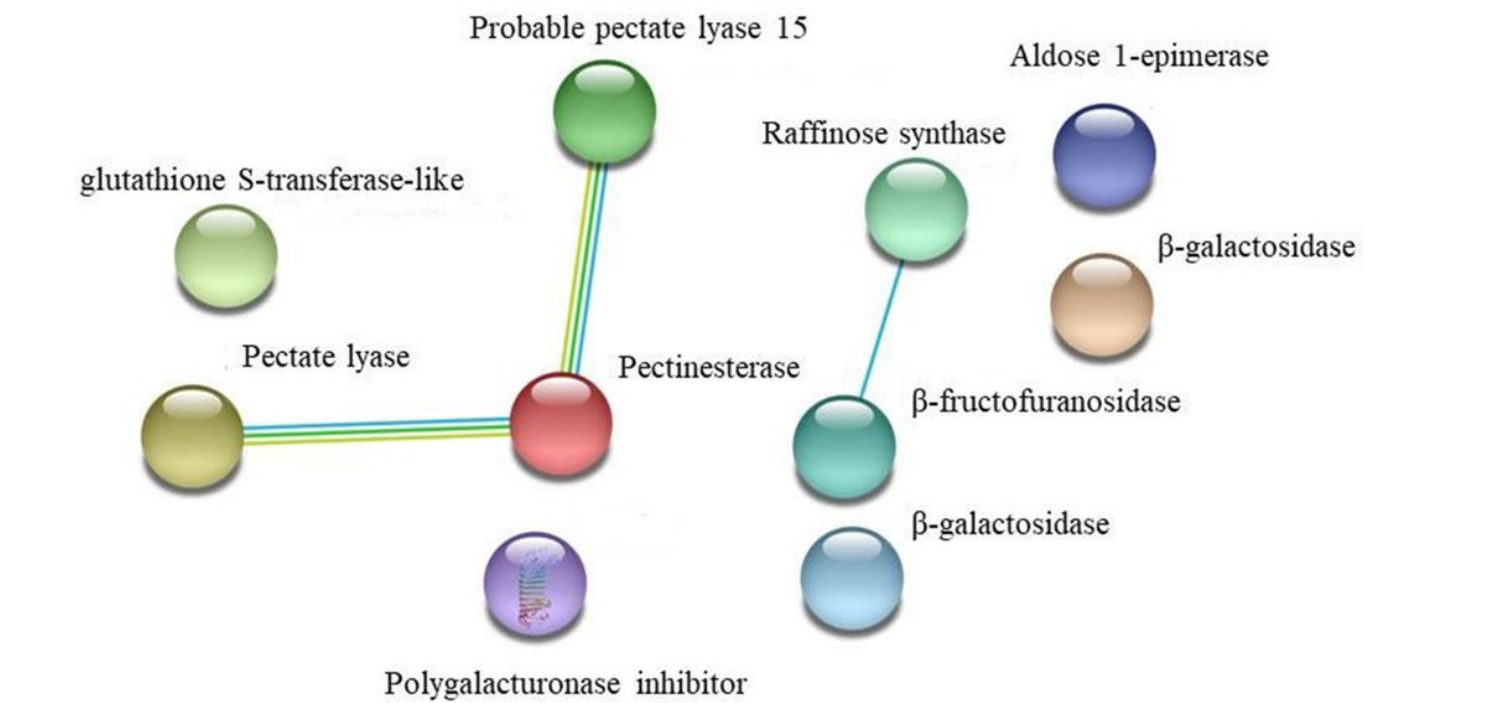


Figure 7

Analysis of the functional network by STRING 9.0 of the cell-wall related proteins, which are associated with pentose and glucuronate interconversions and galactose metabolism.

Supplementary Files

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

- [supply.rar](#)