Integrity transcriptome and proteome analyses provide new insights into the mechanisms regulating pericarp 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 pericarp structure, fruit cracking, transcriptome, proteome, pentose and glucuronate interconversions, galactose metabolism

Posted Date: June 22nd, 2020

DOI: https://doi.org/10.21203/rs.2.19126/v2

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, a perennial wild woody liana, can be used as biofuel to generate bioenergy, as well as a traditional Chinese medicine plant,and new potential edible fruit crop, due to its high yields in fields, wide adaptability, high economic, medicinal and nutritive values, and tolerance tocultivation conditions. However, the pericarp of A. trifoliata cracks longitudinallyalong the ventral suture during fruit ripening, which is a serious problem that limits its usefulness and causes significant losses in yield and commercial value. Furthermore, there have been no known investigations on fruit cracking and its molecular mechanisms in A. trifoliata .

Results: In this study, the dynamic structural changes in fruit pericarps were observed, revealing that the cell wall of fruit pericarp became thinner, and had reduced integrity, and that the cell walls began to degrade in the cracking fruits compared to those observed in non-cracking fruits. Moreover, analyses of the complementary RNA- sequencing-based transcriptomes and tandem mass tag-based proteomes at different development stages during fruit ripening were performed, and the expression of various genes and proteins was found to be changed after cracking, The mRNA levels of 20 differentially expressed genes and 17 differentially abundant proteins (DAPs) involved in cell wall metabolism were further analyzed; 20 DAPs were also validated through parallel reaction monitoring analysis. Among these, pectate lyases and pectinesterase involved in pentose and glucuronate interconversions, β-galactosidases 2 involved in galactose metabolism, were significantly up-regulated in cracking fruits compared to levels in non-cracking fruits, suggesting that they might play crucial roles in A. trifoliata fruit cracking.

Conclusions: This study provides new insights into the molecular basis of fruit cracking in A. trifoliata fruits and important clues for further studies on the genetic improvement of A. trifoliata and the breeding of non-cracking varieties.

Background

Akebia trifoliata (Thunb.) Koidz is a perennial wild woody vine, belonging to the family Lardizabalaceae and the subgenus *Akebia Decne* [1]. The fruit of *A. trifoliata* has a high content of sugar (50.32 %), which is the focus of industrial bioethanol production because it can boost the amount of ethanol produced at the end of fermentation [2-3]. It also has many seeds (300–800), which account for approximately 85 % of the whole fruit weight in wild plants, as well as high seed oil content, and its physicochemical properties meet the biodiesel standard [4-5]. Similar in structure to petrol-diesel, biodiesel prepared with C15–C18 alkane fractions with a high cetane number of 91 can be obtained from *A. trifoliata* seed oil, which was similar to that of biodiesel prepared using *Sapium sebiferum* oil (40.2) and *Vernicia fordii* (53) [6-8]. Moreover, as a traditional Chinese medicine, after ChCl-formic acid pretreatments, a 98.0 % and 100 % level of enzymatic saccharification and ethanol yield have been achieved, respectively, from the solid residues of *Akebia,* suggesting that it can be used as feedstock for cellulosic ethanol production [9]. The advantage of *A. trifoliata*, such as high total sugar content of fruit, abundance of seeds, high content of seed oil, huge yields in fields, wide adaptability, tolerance to both drought and heat, and ease of management, provide new resources for biodiesels and feedstocks for cellulosic ethanol production [5, 9]. *A. trifoliata* can also be exploited as a new high-value fruit crop due to its high nutritional, ornamental, economic, medicinal, and potential development value [10-11].

However, the pericarp of *A. trifoliata* cracks longitudinally when matured in Chinese lunar August, which was also called ‘Bayuezha’. Fruit cracking is a serious problem that increases the incidence of pests and disease infection, leading to fruit decay and deterioration and affecting the utilization rate of fruit and seeds, reducing the storage and shelf life of fruits, and causing significant losses in yield and commercial value [12-13]. Studies indicated that the dry mass of insect-infested fruits (33.1 mg/fruit), as well as the oil content (3.9 %), was less than that in sound fruits (67.4 mg/fruit and 39.9 %, respectively); further, the observed fruit (2.9 kg/tree) and oil yield (0.6 kg/tree) was less than the excepted yield (4.7 kg/tree, and 1.9 kg/tree, respectively) [14]. Fruit cracking is a complex phenomenon caused by numerous factors, including pericarp morphology, physio-biochemical characteristics, and environmental and genetic factors [15]. For example, pericarp tensile strength, cracking turgor thresholds, the hormone and soluble solid contents were found to be higher in the cracked fruits

than in the uncracked fruits [16]. Environmental factors, such as rainfall, mineral nutrition, temperature, light intensity, and humidity also influence the fruit cracking, but these factors are difficult to control; cultivation techniques could reduce the loss caused by fruit cracking, but such treatments are temporary and expensive. Consequently, genetic analysis of fruit cracking and an improved understanding of its molecular mechanisms are effective strategies to breed non-cracking varieties.

Fruit cracking has been studied in some species to elucidate the mechanisms underlying this process, such as tomato [17], litchi [18], soybean [19], and *Arabidopsis thaliana* [20]. Studies have indicated that cell wall-modifying proteins, including polygalacturonases (PGs), pectinesterase (PE), β-galactosidases (β-GAL), expansins (EXPs), and xyloglucan endotransglycosylase proteins, were reported to be associated with fruit cracking [21-22]. Dong et al. [23] found that pod shattering resistance in soybean was mediated by the *NAC* gene. Sorefan et al. [24] found that INDEHISCENT coordinates the regulation of PINOID, WAG2, and PIN-FORMED3 and establishes a dynamic pattern of auxin accumulation at the valve margin, which appears to play a dominant role in the dehiscence regulation of *Arabidopsis*.

Although fruit cracking has been researched for many years, there has been little progress in our understanding of its molecular mechanisms. In particular, there have been no reports on fruit cracking in *A. trifoliata*. Even molecular research on *A. trifoliata* is rare, and only two studies on the use of transcriptome sequencing have recently been reported for this species [1, 25]. A lack of knowledge about the molecular characteristics of *A. trifoliata* has made it difficult to recommend preventive measures for fruit cracking. Next generation sequencing methods, 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 [26, 27]. Therefore, in this study, an integrative analysis of the transcriptomes and proteomes was performed to illuminate the mechanism of *A. trifoliata* fruit cracking at the molecular level based on RNA-seq and tandem mass tag (TMT) technologies. Our comprehensive parallel analyses will provide several new and interesting insights into the molecular mechanisms of *A. trifoliata* fruit cracking and genetic improvement.

Results

Changes in pericarp structure

The pericarp of *A. trifoliata* is known to crack longitudinally at maturity, and the seeds disperse from along the ventral suture with fruit cracking. First, the dynamic structures of the fruit pericarps in different development stages were observed in this study (Fig. 1). In the non-cracking stage (PS), the arrangements of pericarp cells and cuticles were dense, with small intercellular spaces, and were distributed continuously (Fig. 1a, 1d, 1g). However, the cell wall became thinner, the cell volume became larger, the number of cell layers decreased, and the arrangement of cells was loose with poor integrity; further, the spacing between the cells became bigger and the cell of exocarp and mesocarp began to degrade in the initial cracking stage (PM) (Fig. 1b, 1e, 1h). The cells were arranged irregularly and were not compact, and the cell layers continued to reduce, with larger cell spaces, and cells continued to degrade in the total cracking stage (PL) (Fig. 1c, 1f, 1i).

Transcriptomic analysis overview

To obtain an overview of the *A. trifoliata* transcriptome during fruit development and ripening, nine cDNA libraries (i.e. PS, PM and PL, each with three repeats) 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 > 10 %, as well as 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 %, respectively, were obtained (Table S1). The resultant *A. trifoliata* transcriptome contained 241 376 transcripts, ranging from 201 to 2000 bp, and 186 054 unigenes (> 200 bp; Table 1 and Fig. S1); the details of the size distribution of the transcripts and unigenes are shown in Fig. S1.

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; 53.9 % and 41.57 % of all identified unigenes and transcripts, respectively), SwissProt (56 346; 30.3 % and 23.3 %), Protein families database (Pfam) (34 428; 18.5 % and 14.3 %), Gene Ontology database (GO) (44 558; 23.9 % and 18.5 %), and Kyoto Encyclopedia of Genes and Genomes pathway database (KEGG) (30 298; 16.3 % and 12.6 %). This indicated that the NR database provided the largest number of annotations, suggesting that 100 329 unigenes correspond with sequences from at least one of the public databases, and 7283 unigenes were annotated to all databases. Moreover, a total of 17 601, 19 281, and 13 525 unique unigenes and 45 866, 48 104, and 43 303 absent unigenes were identified in PS, PM and PL, respectively. Most of these were uncharacterized, unknown, and hypothetical proteins in the annotation results (Table S2-S3).

Among these unigenes, 11 205 were identified as differently expressed genes (DEGs) using an absolute $\log_2$ fold-change >1 with $p < 0.05$ during fruit ripening. There were 779 up-regulated and 1924 downregulated genes in the PM compared to levels in the PS group, which are presented in a volcano plot in Fig. S2a, and 4623 upregulated and 1975 downregulated genes in PL compared to expression in the PM group, and these are presented in a volcano plot in Fig. S2b. There were 1904 DEGs that were co-expressed in PM_PS and PL_PM combinations (Table 2).

Functional classification of the identified DEGs

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

Moreover, a KEGG pathway analysis was carried out to further evaluate the DEGs. Many were enriched in metabolic pathways, ribosomes in PM_PS, and biosynthesis of secondary metabolites and metabolic pathways in PL_PM (Fig. 2e-2f). Additionally, 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, including those associated with pentose and glucuronate interconversions, the phenylpropanoid pathway, galactose metabolism, starch and sucrose metabolism, amino sugar and nucleotide sugar metabolism and transcription factors were clustered closely both in PM_PS and PL_PM group. Notably, most cell wall-related DEGs were downregulated in the PM_PS group but were upregulated in PL_PM group (Fig. 3a-3f).

Quantitative proteome analysis

To understand the molecular mechanisms of pericarp cracking in *A. trifoliata* fruits, a quantitative proteomics analysis was performed using the TMT platform and LC-MS/MS analysis during fruit ripening, to complement the transcriptome analysis. Accordingly, totals 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 and Table S3). In terms of protein mass distribution, proteins with molecular weights greater than 9 kDa had a wide range and good coverage, with a maximum distribution area of 10–40 kDa (Fig. S3a). Peptide quantitative analysis of the proteins showed that protein quantity decreased with an increase in 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 thresholds for upregulated and downregulated, respectively. Further, 84 proteins were more abundant and 106 proteins were less abundant in the PM_PS group and are shown in a volcano plot in Fig. S2c; 20 DAPs were more abundant and 13 DAPs were less abundant in the PL_PM group, and these are shown in a volcano plot in Fig. S2d, whereas 17 were co-expressed in 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 DAPs in the BP category were involved in cellular responses to the chemical stimulus and cellular oxidant detoxification processes in the PM_PS group, as well as the metabolic and macromolecule metabolic processes in the PL_PM. The highest portions of DAPs in MF category were the oxidoreductase activity and antioxidant activity in PM_PS, as well as structural constituent of ribosome, and structural molecule activity in PL_PM group. Extracellular region in the PM_PS group and cytoplasmic parts in the PL_PM group comprised the highest portions of DAPs in CC category (Fig. 2c-2d).

Moreover, a KEGG pathway analysis was carried out to further evaluate the DAPs. Many were enriched in two-component system and ribosome pathways 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, including those associated with 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 were clustered closely in both PM_PS and PL_PM groups. Notably, most cell wall-related DAPs were upregulated in both the PM_PS and PL_PM groups, whereas those DAPs involved in phenylpropanoid pathways and galactose metabolism were downregulated in the PM_PS group and up-regulated in the PL_PM group (Fig. 3g-l).

Comparative analysis between protein abundance and gene expression levels

There were more DEGs (4607 and 6598) than DAPs (190 and 50; PM_PS and PL_PM, respectively) and many more shared DEGs (1904) than shared DAPs (17) in cracking fruit compared to those in non-cracking fruit. Most of these DEGs and DAPs were downregulated in the PM_PS group but upregulated in the PL_PM group, suggesting that greater changes in fruit cracking occurred during fruit ripening. Among the shared DEGs, 1123 were upregulated and 781 were downregulated in the PM_PS group, whereas 808 were upregulated and 1096 were downregulated in PL_PM group. Of the shared DAPs, nine were increased in abundance and eight were decreased in abundance the PM_PS group, whereas eight exhibited increased abundance and nine showed decreased abundance in the PL_PM group (Table 2). To evaluate the relationships between the transcriptomic and proteomic changes during fruit pericarp cracking, the quantitative data for DEGs and DAPs were used for correlation analysis. According to this, 14 and four DAPs and their corresponding DEGs were identified in the PM_PS and PL_PM groups, respectively. Of these, 12 DAPs (four with increased abundance and eight with decreased abundance) and four DAPs (two with increased abundance and two with decreased abundance) were regulated in the same direction as their corresponding DEGs in the PM_PS and PL_PM groups, respectively (Fig. 4a-4b); meanwhile, two were regulated in the opposite direction as their corresponding DEGs in the PM_PS group. There were more DEGs than DAPs in both the PM_PS and PL_PM groups, with significant differences in the trends of transcript levels and protein abundance.

Furthermore, the fold-changes of the DAPs indicated weak positive correlations with their corresponding DEGs based on Pearson's correlation tests ($r = 0.03$ and 0.11, $p < 0.01$, in PM_PS and PL_PM, respectively; Fig. 4c-4d). The fold-changes in the DAPs were positively correlated with the DEG showing the same trend ($r = 0.9161$ and 0.8, $p < 0.01$, in PM_PS and PL_PM, respectively; Fig. 4e-4f).

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 and KEGG pathways analyses. The largest groups within the BP category were those linked to metabolic process and cellular process; catalytic activity and binding and cell and cell part were predominant in MF and CC categories, respectively, both in PM_PS and PL_PM groups (Fig. S4a-4b). In the PM_PS group, 14 DAPs were significantly enriched in seven pathways, both with respect to DEGs and DAPs, which included 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, four DAPs were significantly enriched in the three pathways, both in terms of 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 two pathways, including pentose and glucuronate interconversions and galactose metabolism pathways, were shared between the transcriptome and proteome data, for both PM_PS and PL_PM groups. However, the phenylpropanoid biosynthesis pathway was only shared by DAPs and DEGs in the PM_PS group. Therefore, those shared metabolic pathways might play potential roles in *A. trifoliata* fruit cracking.

Moreover, the protein–protein interaction (PPI) network was analyzed to predict the biological functions of *A. trifoliata* fruit cracking using the STRING database. A total of 85 DAPs, including 29 upregulated (28 and one in PM_PS and PL_PM, respectively) and 56 downregulated DAPs (46 and 10 in PM_PS and PL_PM, respectively) were assigned to the interaction network (Fig. 5). Only four cell-wall related proteins, including endoglucanase 8 (TRINITY_DN136333_c1_g2), β-glucosidase 33 (BGLU33; TRINITY_DN137437_c3_g1), peroxiredoxin-2 (TRINITY_ DN137008_c1_g6), and β-galactosidase1 (β-GAL1; TRINITY_DN138388_c1_g1) were determined to interact with other proteins.

Validation of data reliability through reverse transcription real-time quantitative PCR (qPCR)

To validate the results of the RNA-Seq and TMT data, expression and correlation analyses between the qPCR and the fragments per kilobase per million reads mapped (FPKM) values obtained from the transcriptome and proteome data were performed. The 20 selected DEGs, involved in phenylpropanoid biosynthesis, pentose and glucuronate interconversions, amino sugar and nucleotide sugar metabolism, galactose metabolism, and starch and sucrose metabolism, as well as other cell wall metabolism-related DEGs, had shown differential expression patterns in the PM_PS and PL_PM groups, and the results of the qPCR are shown in Fig. 6. Specifically, in the PM_PS group, phenylpropanoid pathway-related genes 4-coumarate-COA-ligase (*4CL*), peroxidase (*PRX*), and *PRX2* were downregulated, but cinnamyl-alcohol dehydrogenase (*CAD*) and shikimate O-hydroxycinnamoyltransferase (*HCT*) were upregulated. The 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*) were downregulated. Meanwhile, the cell wall metabolism genes *NAC*, *NAC*-like, and *EXP1* were downregulated, but the *BHLH* transcription factor and dirigent protein (*DIR2*) were upregulated. Further, pentose and glucuronate interconversion-related genes *PL*, *PG*, and *PE* were upregulated. Most of these genes were significantly upregulated 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 seven genes showed poor correlations with the corresponding protein expression (Fig. 6; Table 3).

Then, the corresponding genes of 17 selected DAPs were analyzed by qPCR to further validate the proteomic data, and the qPCR results of these genes are shown in Fig. 7. Specifically, *PRX*, *PRX3*, *PRX4*, *PRX5*, *β-GAL1*, *β-GAL2*, *BXL*, *PG*, and *PG2* were significantly downregulated, whereas glucan endo-1,3-beta-glucosidase (*ENBG*) and furostanol glycoside 26-O-beta-glucosidase (*F26G*), involved in starch and sucrose metabolism, and *PL*, *PE*, *EXP1*, and *DIR1* were significantly upregulated in the PM_PS group. In the PL_PM group, most of the genes were upregulated, except for *DIR1*, *EXP1*, *PG*, *PG2*, *β-GAL1*, β-glucosidase33 (*BGLU33*), and alpha/beta hydrolase (*α-HY*). Moreover, 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-GAL1* (r = −0.8215, $p < 0.05$), *β-GAL2* (r = 0.7797, $p < 0.05$), and *ENBG* (r = 0.8917, $p < 0.01$) showed strong correlations with the TMT data. However, levels of three genes showed poor correlations with their corresponding protein expression (Fig. 7; Table 3). In general, the qPCR results confirmed the gene expression patterns obtained by transcriptomic and proteomic data, suggesting that our results are reliable.

Validation of data reliability by parallel reaction monitoring (PRM)

The protein expression levels obtained by TMT were confirmed by quantifying their expression by PRM analysis. 20 proteins that exhibited significantly different levels by TMT analysis were selected for PRM analysis, of which 18 were successfully quantified (Table 4). Here, 14 of the 18 (77.8%) proteins showed the same trend as that observed when the protein levels were quantified by TMT, including PE, PL, PG2, F26G, β-GAL2, Auxin efflux carrier, α-HY, PRX2, PG4, PRX5, PRX3, endoglucanase 19, endoglucanase 8, and DIR1. Meanwhile the mean expression levels of PRX, beta-fructofuranosidase, BXL, and BGLU33 proteins were inconsistent with the protein levels quantified by TMT. In general, the trends in the expression changes measured by PRM and TMT were basically consistent.

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

To illuminate which genes might play key roles in cell wall metabolism pathways during fruit cracking, the expression profiles of 10 genes were analyzed. Among these, *PL* (TRINITY_DN143250_c1_g6) was significantly upregulated in the PM_PS group, *PE* (TRINITY_DN143028_c0_g1), *β-GAL2* (TRINITY_DN142386_c5_g1), *F26G* (TRINITY_ DN142424_c1_g1), *PG* (TRINITY_DN196976_c0_g1), *PG3* (TRINITY_DN142042_c0_g2), and *BXL* (TRINITY_DN141432_c1_g2) were upregulated in the PL_PM group, and *BGLU33* (TRINITY_DN137437_c3_g1) was downregulated in the IS_NS group. *PG2* (TRINITY_ DN142943_c1_g1) and *PG4* (TRINITY_DN141074_c0_g1) were downregulated in the PL_PM group based on both the transcriptome and proteome (Fig. 8). Notably, *PL*, *PE*, and *β-GAL2* were upregulated in the PM_PS, and PL_PM groups, respectively, based on transcriptome, proteome, qPCR, and PRM data.

Discussion

Structural changes in the pericarp cell wall might affect *A. trifoliata* fruit cracking

Fruit cracking is a complex phenomenon that is caused by a series of environmental, physiological, biochemical, and genetic changes during fruit ripening. Fruit cracking happens when the stress exerted on the pericarp from the enlarged aril is greater than the strength of the fruit skin, and the mechanical strength of the pericarp depends largely on its cell wall [28]. Studies showed that 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]. Arrangement of the subcutaneous layers of the cells was found to be relatively regular, and cell layers had a closer arrangement in the cracking-resistant tomato genotype [30]. In this study, the pericarp cell structure and ultrastructure in *A. trifoliata* fruits of different development stages were observed. Compared to those in the non-cracking fruit, the cell wall structures of the fruit pericarp had poor integrity, loose cell wall structures, deformed and reduced cell layers, and larger spaces, and the pericarp cells began to degrade during fruit cracking (Fig. 1e and 1h); these changes were more pronounced during the total cracking stage (Fig. 1f and 1i), which was consistent with previous results in grapes and orange [31-32]. These results indicated that the structural changes in the cell wall of *A. trifoliata* pericarp might play a key role in the occurrence of fruit cracking.

General features of the transcriptomes and proteomes of different *A. trifoliata* pericarps

Fruit cracking is a key factor that affects marketability of fruits, reduces their market acceptability, and causes significant losses in field yield and commercial value [33]. Elucidating the molecular mechanisms that regulate fruit cracking could aid in the utilization of *A. trifoliata* for biofuels. However, there have been no specific studies on *A. trifoliata* fruit cracking and the underlying mechanisms are still largely unknown. In this study, the differences in the transcriptome and proteome were investigated based on RNA-seq and TMT data during different development stages. As the transcriptome database was used for protein identification in this study, the quality of the sequencing and assembly of the transcriptome data was crucial for subsequent analyses. A total of 186 054 unigenes (> 200 bp) were assembled in the transcriptome of *A. trifoliata* pericarp, which is much more than that previously obtained for this species and those of other Ranunculales such as *A. trifoliata* (11 749 by Yang et al. [25]; 65 757 by Niu et al. [1]), *Dysosma aurantiocaulis* (53 929) [34], and *Dysosma versipellis* (44 855) [35]. The obtained percentages of the Q30 bases (91.58–93.75 %) and GC content (45.94–48.48 %) were similar to the levels

reported in other studies for *A. trifoliata* transcriptome studies using an RNA-seq approach, such as 89.06–93.33 % and 43.20–43.93 % [25] and 96.31 % and 45.10 %, respectively [1]. In total, 100 329 (53.9 %) and 56 346 (30.3 %) of all unigenes identified in the present study were matched to the NR and SwissProt databases, respectively. These results are similar to values reported for *Dysosma aurantiocaulis* (29 497 and 18 029; 54.70% and 33.43%) [34] and were higher than those presented in other similar *A. trifoliata* transcriptome studies, including 34 245 and 23 352 (32.14 % and 21.91 %) [25] and 19 096 (29.95 %) in SwissProt [1]. The results provide extensive sequencing and unigene resources for *A. trifoliata*.

Moreover, a total of 812 625 spectra, 10 572 unique peptides, and 2839 proteins were identified based on transcriptome data from *A. trifoliata* pericarp. The RNA-seq and protein sequencing methods identified and annotated many genes and proteins, providing the basis for a more precise and detailed description of molecular processes, as well as the elucidation and better understanding of complex physiological processes and their genetic regulation [36]. It can be seen that the *A. trifoliata* pericarp sequencing and assembly quality presented here were high enough and useful tools for future genetic research on fruit cracking in *A. trifoliata* and other Lardizabalaceae species. Additionally, there are many ‘uncharacterized’, ‘predicted’, or ‘putative’ transcripts and proteins in the annotation results that might be limited by the lack of genomic information; these were not associated with a definitive annotation and might also exist in other species [37]. Thus, the roles of these unknown or uncharacterized genes and proteins in *A. trifoliata* fruit cracking constitute an important issue for future research.

Translational and posttranslational regulation of *A. trifoliata* fruit cracking

Some inevitable passive processes will occur in the cracked pericarp following fruit cracking, such as oxidative stress and microbial invasion [37]. Therefore, the differences in the expression of genes and proteins among PS, PM, and PL alone cannot accurately reflect the cause of fruit cracking. In this study, a comparative analysis of protein abundance and gene expression levels was performed, and more DEGs (11205) than DAPs (240), as well as more shared DEGs (1904) than shared DAPs (17), were identified in cracking fruits compared to those in non-cracking fruits. A possible explanation for the much lower number of DAPs is the limitation of the proteomic method of choice (MS/MS) in identifying the presence and abundance of the proteins, as compared to that with transcriptomics (RNA-seq) [38]. Moreover, a different trend was observed between protein and transcription levels. The integration of transcriptomic and proteomic data revealed that only 14 and four DAPs corresponding to DEGs and a poor correlation ($r = 0.03$ and 0.11) between the protein abundance and the corresponding gene expression were identified in the PM_PS and PL_PM groups, respectively. These results indicated that there was discordance between the transcript levels and protein abundance, which was similar to that observed in previous reports, suggesting that post-transcriptional and post-translational regulation, reversible phosphorylation, splicing events in cells, and translation efficiency might also play key roles in the regulation of fruit ripening [39-40]. Moreover, the PPI analysis indicated that the interactions of proteins were weak in this study, and only four proteins associated with cell wall metabolism were determined to interact with other proteins. This was associated with the results indicating that the interactions are often weak for many cellular processes, which are regulated by post-translational modifications that are recognized by specific domains in protein binding partners [41]. This inconsistency showed that the regulation from mRNA to proteins is a complex process, and the changes in protein abundance were generated after its corresponding transcript was stabilized [42]. Thus, gene translation and post-translation processes might be important regulatory factors during *A. trifoliata* fruit ripening and cracking, which was consistent with the result that many unigenes were assigned to the category of posttranslational modification [35].

Potential regulators and metabolism pathways during fruit cracking

GO functional enrichment and KEGG pathways provide prediction information of inner-cell metabolic pathways, as well as the genetic and biologic behaviors of genes. The GO enrichment analysis indicated that most DEGs and DAPs were both involved in oxidoreductase activity and structural molecule activity in MF category in the PM_PS and PL_PM groups, respectively (Fig. 2a-2d). These results were associated with previous studies and could classify the target genes into different categories based on their functions and amounts, such as peroxidase and cell wall polysaccharide [18, 43]. Cell polysaccharides are degraded by the cell wall hydrolases, and the formation of phenolic cross-linking cell wall structural

components is catalyzed by cell wall peroxidase. These modifications reduce the strength of the fruit pericarp, resulting in changes in pericarp development and fruit cracking [44]. In this study, the hierarchical cluster analysis of cell wall-related DEGs and DAPs indicated that most were upregulated in the PL_PM groups both in the transcriptome and proteome (Fig. 3). The pathway enrichment analysis indicated that two cell wall-related pathways, including the pentose and glucuronate interconversions and galactose metabolism pathway, were common pathways shared by the DAPs and DEGs in both the IS_NS and TS_IS groups. Further, the phenylpropanoid biosynthesis pathway was shared by the DAPs and DEGs in the IS_NS group, suggesting that cell well metabolism might have an important role in *A. trifoliata* fruit ripening and cracking. Moreover, these results revealed that the proteomic data and the transcriptome data were complementary, and that the proteome could confirm the transcriptome data; in addition, genes perform the same function at the transcriptome and proteome levels [42]. The functional classification and integrative analysis of the transcriptome and proteome would provide a good basis to better understand the molecular physiology of fruit ripening and cracking.

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

Fruit cracking is a complex physiological phenomenon that is controlled by many genes working

together, rather than a single gene directly controlling the process [16, 45]. Researchers have found that several cell wall modification genes, including *β-GAL*, *β-GLU*, *PE*, and *PG* are differentially expressed in cracked fruits compared to levels in non-cracked litchi fruits [46]. The suppression of *PpBGAL* expression can reduce *PpPG* transcription and activity, suggesting that the downregulation of *PpBGAL* expression delays peach fruit softening [47]. Antisense inhibition of PE and PG activity in tomato was also found to reduce fruit cracking [48]. Further, silencing the *SlPL* gene can enhance fruit firmness and reduce the content of pectin, suggesting that this gene participates in the pericarp cell wall rearrangement during fruit softening [49]. Fruit softening is mainly caused by hemicellulose and pectin degradation proteins, including XYL, BGAL, PE, and PG [50-51]. As a major cellulose degradation enzyme, BGLU is expressed in ripening fruits, and the downregulation of this protein in strawberry can delay fruit maturation [52]. PRXs are well-known cell wall enzymes and involved in the rearrangements of cell wall polysaccharides during plant development [53].

In this study, 20 DEG and 17 DAPs involved in cell wall metabolism were validated using qPCR, and 27 of 37 (13 in DEGs and 14 in DAPs) showed strong correlations with protein expression levels. Moreover, 20 DAPs involved in cell wall metabolism were also validated using PRM and 14 showed the same trend as TMT protein levels, which was associated with the results that cell polysaccharide metabolism plays key roles in fruit ripening and cracking [43, 45]. Notably, three proteins, including PE, PL, and β-GAL2, were functionally annotated among thousands of identified DEGs and DAPs, and the expression of these three enzymes was significantly upregulated in cracking fruit comparing to that in non-cracking fruit. PE, involved in the pentose and glucuronate interconversions pathway, and β-GAL2, involved in galactose metabolism, were upregulated in the PL_PM group; meanwhile, *PL*, involved in the pentose and glucuronate interconversions pathway was upregulated in the PM_PS group with the same expression trends observed in transcriptome, proteome, qPCR, and PRM data (Fig. 8), suggesting that PE, PL, and β-GAL2 might play important roles in the regulation of *A. trifoliata* fruit cracking. Moreover, F26G, PG, PG3, XYL, PRX3, and PRX5 also showed higher protein and gene expression in cracking fruit than in non-cracking fruit. The significantly increased expression of these cell wall metabolism proteins is in accordance with the results indicating that a network of cell wall cellulose, hemicellulose, and pectin might be correlated with fruit cracking, suggesting their dynamic roles in fruit ripening and cracking [52, 54].

Conclusions

This study revealed the structural changes in the cell wall during different fruit development stages of *A. trifoliata* pericarp, suggesting that the structural changes between unripe and ripe fruit might be an important factor in fruit tendency to crack. Therefore, bagging treatment could be performed to strengthen the structure of the pericarp and reduce fruit cracking [55]. Moreover, this study presented comprehensive transcriptome and proteome data to screen the potential effectors of fruit cracking in *A. trifoliate*, and various genes and proteins were found to be differentially expressed after cracking, providing a

global view of the molecular mechanism of *A. trifoliate* fruit cracking. The co-expressed genes and proteins were found to be implicated in various signaling pathways in fruit development, and those genes involved in cell wall metabolism might play important roles in *A. trifoliate* fruit ripening and cracking. Therefore, cell wall structure and modifications might be correlated with the strength of the pericarp and fruit cracking. Moreover, degradation of the cell wall also depends on ethylene during fruit softening. For example, the activities of β-GAL and PG could be suppressed in 1-methylcyclopropene (1-MCP)-treated fruit in response to ethylene treatment [56]. Therefore, 1-MCP treatment could be performed to delay fruit cracking during fruit ripening. In general, these omics data might provide a new perspective to further study the molecular mechanism of fruit cracking and the breeding of non-cracking varieties of Lardizabalaceae and other species. The CRISPR/Cas9 system could be used to determine the molecular functions of these co-differentially expressed proteins/genes in further studies to reveal the mechanisms of fruit cracking in Lardizabalaceae.

Materials And Methods

Plant materials

The wild germplasm Nong No.8, which had been transplanted for 9 years in the nursery at Hunan Academy of agricultural sciences, Changsha, P. R. China, was used as the research material in this study. According to our observations, in Changsha, 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 with a flowering period of approximately 30–45 d. The fruit development lasts for 5 months, from the ovary inflation through the fruit setting, and longitudinal and transverse elongation, flesh softening, peel cracking, complete the development process [57]. The Nong No.8 variety usually ripens in early October. Since the fruit of *A. trifoliata* do not develop uniformly, fruits were harvested from different stages of the same Nong No.8 tree at the same time and were then sorted according to their developmental stage, considered to be a sample. Different fruits, including the non-cracking fruits (PS), the initial cracking fruits (PM), and the total cracking fruits (PL) were randomly taken every 10 days at the ripening stage (September 18, September 28 and October 8, 2018) and three fruits were mixed into one biological replicate for further analysis, and in total, three biological replicates were collected for each stage. Sampled pericarps were rapidly collected and immediately frozen in liquid nitrogen and stored at −80 ºC until use for transcriptome, proteome, qPCR, and PRM 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 previous studies [58-59]. 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 paraffin. Subsequently, paraffin sections were stained with Safranin O and fast green staining and observed with an 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 7.0). After post-fixing with 1 % osmium for 1–2 h and washing three times in PBS, samples were dehydrated with 30, 50, and 70 % ethanol for 20 min in turn and then ethanol and iso-amyl acetate (V:V=1:1) for 30 min; they were finally dried in liquid carbon dioxide. Dried samples were coated with gold-palladium in a Hitachi Model E-1010 ion sputter for 4–5 min and observed via SEM with a Hitachi Model SU-8010 (Hitachi, SU8010, Japan).

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 [60]. The RNA samples were detected based on the A260/A280 absorbance ratio with a Nanodrop ND-1000 system (Thermo Scientific). Paired-end libraries were prepared using NEBNext® UltraTM 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 with random hexamer primers 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 an 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 with an Illumina HiSeq2500 instrument by Shanghai Applied Protein technology (Shanghai, China).

Quality control and transcriptome assembly

The raw paired-end reads in fastq format produced from the sequencing were first processed with the in-house Perl scripts. Those reads containing adapters, excess “N” nucleotides with more than 10 % of the bases and reads of low-quality (reads with quality values ≤ 10) were removed by filter_fq 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) [61]. The Trinity software consists 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 databases, based on a BLAST search with an E-value threshold of $1E^{-5}$. Moreover, to further analyze the annotation results, GO and KEGG results with E-values 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 based on the KEGG database.

The normalized transcript abundances of the genes were estimated using the FPKM based on the length of the gene and read counts 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 the Benjamini and Hochberg’s approach to adjust the *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 pathway enrichment analysis of DEGs was implemented with the clusterProfiler R package. Transcription factor 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 [62]. The samples were frozen in liquid nitrogen and ground into powder. A five times volume of TCA/acetone (1:9) was added, and the sample was vortexed and mixed and placed at −20 ℃ for 4 h; it was then centrifuged at 6000 × *g* at 4 ℃ for 40 min. The supernatant was discarded, and the precipitate was washed three times with pre-cooled acetone. After the precipitation and air drying, the precipitate was redissolved in buffer (4 % SDS, 100 mM tris–HCl and 1 mM DTT; pH 7.6). After sonication and boiling for 15 min, the lysate was centrifuged for 40 min, and the supernatant was filtered and quantified 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 added into the samples to block reduced cysteine residues, incubating for 30 min in darkness at room temperature. After the filters were washed with 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 them with 129-tag (PS), 130-tag (PM), and 131-tag (PL) (Thermo Fisher Scientific, Waltham, MA, USA). 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 were acquired using a data-dependent top10 method dynamically choosing the most abundant precursor ions from the survey scan (300–1800 m/z) through the higher-energy collision dissociation (HCD) fragmentation method. Automatic gain control (AGC) was set as 3E6 and the maximum inject time was 10 ms, with the following parameters: 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 the 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, and 20 ppm for peptide ions, as well as both peptide and protein levels of FDR less than 0.01, and only unique peptides of the protein were employed for the protein 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 DAPs.

Functional categorization was performed using GO and KEGG pathway databases with P values ≤ 0.05. The protein functional network was performed with 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 IS vs NS and TS vs IS.

Reverse transcription real-time quantitative PCR

The method of total RNA extraction and synthesis of cDNA were described previously. The 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, the *EF-1α* gene was used as the internal control gene, which was detected by the *de novo* transcriptome sequencing of *A. trifoliata* [63]. 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 μL SYBR Green PCR master mix, 1 μL cDNA, and 0.5 μL of each primer in a final reaction volume of 25 μL. The thermal cycling program began with 3 min at 95 °C, followed by 40 cycles of 95 °C for 10 s and 55 °C for 30 s, with a melt curve from 65 to 95 °C based on increments of 0.5 °C for 5 s. After PCR amplification, the quantitative variation was analyzed using the 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 [64].

PRM analysis

To verify the protein expression levels obtained by TMT analysis, the expression levels of selected proteins were further quantified by PRM analysis. The AQUA stable isotope peptide was spiked in each sample as an internal standard reference. The tryptic peptides were directly loaded on C18 STAGE-tips for desalting prior to reversed-phase chromatography on an Easy nLC-1200 system (Thermo Scientific). Acetonitrile was used in 45 min from 5% to 35% for a 1-h liquid chromatography gradient, and PRM analysis was performed on a Q Exactive ^{TM} Plus mass spectrometer (Thermo Scientific). The mass spectrometer was operated in positive ion mode and the full MS1 scan was acquired with a resolution of 60 000 (at 200 m/z). AGC target values and maximum ion injection times were set at 3e6 and 200 ms, respectively; full MS scans were followed by 20 PRMs (MS2 scans) at 30 000 resolution (at m/z 200) with AGC of 3e6 and the maximum injection time set as 120 ms. The targeted peptides were isolated with a 2 Th window. Ion activation/dissociation was performed at a normalized collision energy of 27 in the HCD collision cell. Skyline version 3.7.0 was used to analyze the MS data where signal intensities for individual peptide sequences for each of the significantly altered proteins were quantified relative to each sample and normalized to a standard reference [65].

Abbreviations

Akebia trifoliata: *A. trifoliata*; TMT: tandem mass tag; PL: pectate lyases; PE: pectinesterase; β-GAL: β-galactosidases; PGs: polygalacturonases; EXPs: expansins; NS: the non-cracking stage; IS: the initial cracking stage; TS: 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; PRX: peroxidase; CAD: cinnamyl alcohol dehydrogenase; 4CL: 4-coumarate-COA-ligase; HCT: shikimate O-hydroxycinnamoyl transferase; PL: pectate lyase; PE: pectinesterase; PG: polygalacturonase; F26G: Furostanol glycoside 26-O- beta-glucosidase; β-GAL: Beta-galactosidase; DAPs: differentially abundant proteins; BP: biological processes; MF: molecular functions; CC: cell component.

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 raw sequence data have been deposited in the National Center for Bio-technology Information (NCBI) database (https://www.ncbi.nlm.nih.gov/sra/SRP246017). The mass spectrometry proteomics data have been deposited to the PRIDE partner repository with the dataset identifier PXD017282.

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 (1610242020010) and Agricultural Science, Technology Innovation Program (ASTIP) of CAAS (Grant No.2017IBFC).

Publisher's Note Springer

References

1. Niu J, Wang YJ, Shi YL, Wang XF, Sun ZM, Huang KY, Gong C, Luan MB, Chen JH. Development of SSR markers via de novo transcriptome assembly in *Akebia trifoliata* (Thunb.) Koidz. Genome. 2019;999:1-15.
2. Wang DZ, Li F, Yan J, Zhong HM. Study and application of nutritional components of wild plant Var *australis* (Diels) Rehd. Amino Acids Biot Resourc. 2004;26:16-17.
3. Ishmayana S, Learmonth RP, Kennedy UJ. Fermentation performance of the yeast Saccharomyces cerevisiae in media with high sugar concentration. The 2nd International Seminar on Chemistry. 2011; p. 379-85.
4. Zhong CH, Bu FW, Wang ZY, Peng DF. Fruit development and biological characteristics of seedling progenies of *Akebia trifoliata*. Hunan Agricultural Sciences. 2006;1:27-29.
5. Luo LP, Lan YH, Hu BF, Zhang X, Guo XL. Physicochemical properties and biodiesel preparation of *Akebia trifoliate* seed oil. Natural product research & development. 2013;25(8):1095-1100.
6. Liu C, Jiang SH, Xu LL, Li TJ, Han XJ, Wang JX, Liao L, Rong L. Biodiesel Prepared by Hydrocracking *Akebia trifoliata* Oil. CCOA. 2015;30(1):76-80.
7. Wang R, Hanna MA, Zhou WW, Bhadury PS, Chen Q, Song BA. Production and selected fuel properties of biodiesel from promising non-edible oils: *Euphorbia lathyris* L. *Sapium sebiferum* L. and *Jatropha curcas* L. bioresource Technol. 2011;102(2):1194-9.
8. Chung KH. Transesterification of Camellia japonica and *Vernicia fordii* seed oils on alkali catalysts for biodiesel production. J Ind Eng Chem. 2010;16(4):506-9.
9. Yu Q, Zhang A, Wang W, Chen, L, Ruxue B, Zhuang XS, Wang Q, Wang ZM, Yuan ZH . Deep eutectic solvents from hemicellulose-derived acids for the cellulosic ethanol refining of Akebia' herbal residues. Bioresource Technol. 2017;247:705-10.
10. Li L, Yao X, Zhong C, Chen X, Huang H. Akebia: a potential new fruit crop in China. Hortic Sci. 2010;45:4-10.
11. 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.

12. 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;127:436-45.
13. Huang X, Wang HC, Li J, Yin J, Yuan W, Lu J, Huang HB. An overview of calcium's role in lychee fruit cracking. Acta Hortic. 2003;665:1026-35.
14. Liu QX, Wen LZ, Zhou HJ, Wu Q, Marvin K H, Xiao ZS. Effects of seed wasps and seedless fruits on fruit and oil yields of *Pistacia chinensis* as a biofuel tree. Acta Entomologica Sinica. 2011;54(10): 1133-9.
15. 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.
16. Khadivi-Khub A. Physiological and genetic factors influencing fruit cracking. Acta Physiol Plant. 2015;37(1):1718.
17. 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.
18. 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.
19. 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;6:1-10.
20. 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 Phytol. 1999;144(2):213-22.
21. Li J, Huang X, Huang H. Comparison of the activities of enzymes related to cell-Wall metabolism in pericarp between litchi cultivars susceptible and resistant to fruit cracking. J Plant physiol Mol Biol. 2003;29(2):141-46.
22. Khurnpoon L, Siriphanich J, Labavitch JM. Cell wall metabolism during durian fruit dehiscence. Postharvest Biol Tec. 2008;48(3):391-401.
23. 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.
24. 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.
25. Yang H, Liu H, Shi X, Ge F, Zhao Z, Luo H. Transcriptome analysis of *Akebia trifoliate* in ripening process. J Mountain Agrie Biol. 2016;35(2):46-51.
26. 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;193:142-53.
27. Voelckel C, Gruenheit N, Lockhart P. Evolutionary transcriptomics and proteomics: insight into plant adaptation. Trends Plant Sci. 2017;22(6):462-71.
28. Huang XM, Yuang WQ, Wang HC, Li JG, Huang HB, Luo S, Yin JH. Linking cracking resistance and fruit desiccation rate to pericarp structure in litchi (*Litchi Chinensis* Sonn.). J Hortic Sci Biotech. 2004;79:897-905.
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):4558-68.
30. 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.
31. Li J, Chen JZ, Wan JF, Huang YJ, Yao Q. The correlation of cracking fruits types and peel structure between 'Washington' and 'Bonanza' navel oranges. Chin J Trop Crop. 2011;32:921-25.
32. Zhang C, Guan L, Fan X, et al. Anatomical characteristics associated with different degrees of berry cracking in grapes. Sci Hortic. 2020;261:108992.

33. Liao NQ, Hu ZY, Li YY, Hao GF, Chen SN, Xue Q, Ma YY, Zhang KJ, Mahmoud A, AliA, Malangisha GK, Lyu XL, Yan JH. Ethylene-responsive factor 4 is associated with the desirable rind hardness trait conferring cracking resistance in fresh fruits of watermelon. PBJ. 2020;18(4):1-12.
34. 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):228-41.
35. 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):1733-46.
36. Blencowe BJ, Ahmad S, Lee LJ. Current-generation high-throughput sequencing: deepening insights into mammalian transcriptomes. Genes Dev. 2009;23:1379-86.
37. Wang J, Gao XM, Ma ZL, Jing C, Liu YA. Analysis of the molecular basis of fruit cracking susceptibility in *litchi chinensis* cv. baitangying by transcriptome and quantitative proteome profiling. J Plant Physiol. 2019;234:106-16.
38. Ning K, Nesvizhskii AI. The utility of mass spectrometry-based proteomic data for validation of novel alternative splice forms reconstructed from RNA-Seq data: A preliminary assessment. BMC Bioinformatics. 2010;S14:1-6.
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, TaoY, 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. Tripsianes K, Chu NK, Friberg A, Sattler M, Becker CF. Studying weak and dynamic interactions of posttranslationally modified proteins using expressed protein ligation. Acs Chemical Biol. 2014;9(2):347-52.
42. Muers M. Gene expression: Transcriptome to proteome and back to genome. Nat Rev Genet. 2011;12(8):518.
43. Xue LZ, Sun MT, Wu Z, Yu L, Yu QH, Tang YP, Jiang FL. LncRNA regulates tomato fruit cracking by coordinating gene expression via a hormone-redox-cell wall network. BMC Plant Biology. 2020;20:162.
44. Li J, Chen JZ. Citrus fruit-cracking: Causes and occurrence. HPJ. 2017;3(6):255-60.
45. Chen J, Duan Y, Hu Y, Li W, Xie J. Transcriptome analysis of atemoya pericarp elucidates the role of polysaccharide metabolism in fruit ripening and cracking after harvest. BMC Plant Biology. 2019;19(1):2-19.
46. 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):17667-85.
47. Liu H, Qian M, Song CH, Li JJ, Zhao CP, Li GF, Wang AZ, Han MY. Down-regulation of *Pp GAL10* and *Pp GAL16* delays fruit softening in peach by reducing polygalacturonase and pectin methylesterase activity. Front Plant Sci. 2018;9:1015.
48. Brummell DA, Harpster MH. Cell wall metabolism in fruit softening and quality and its manipulation in transgenicplants. Plant Mol. Biol. 2001;47:311-18.
49. Yang L, Huang W, Xiong F, Xian Z, Su D, Ren M, et al. Silencing of SlPL, 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):1544-55.
50. Bustamante CA, Civello PM, Martínez GA, Cloning of the promoter region of β-xylosidase (*FaXyl1*) gene and effect of plant growth regulators on the expression of *FaXyl1* in strawberry fruit. Plant Sci. 2009;177:49-56.
51.Song XX, Hu Z, Bi Y, Jiang AL. Research progress in softening mechanism and regulation of fresh-cut fruits and vegetables. Sci Technol Food Industry.2013;34(4):397-9.
52. Sara F, Simona M, Chiara M. Fruit ripening: the role of hormones, cell wall modifications, and their relationship with pathogens. J Exp Bot. 2019;70(11):2993-3006.

53. Francoz E, Ranocha P, Nguyen-Kim H, Jamet E, Burlat V, Dunand C. Roles of cell wall peroxidases in plant development. Phytochemistry. 2015:112:15-21.
54. Huang XM, Wang HC, Lu XJ, Yuan W Q, Lu JM, Li JG. Huang HB. Cell wall modifications in the pericarp of litchi (*Litchi chinensis* sonn.) cultivars that differ in their resistance to cracking. J Hortic Sci Biotech. 2006;81(2):231-7.
55. Son IC, Kim DI . Effects of bagging periods on pericarp characteristics and berry cracking in 'Kyoho' Grape (Vitis sp.). Korean J Hortic Sci Technol. 2010;28(3):381-6.
56. Jeong J, Huber DJ. Suppression of avocado (*Persea americana* Mill.) fruit softening and changes in cell wall matrix polysaccharides and enzyme activities: differential responses to 1-MCP and delayed ethylene application. J Am Soc Hortic Sci. 2004;129:752-9.
57. Zhang XF. Study on the physiological and ecological characteristics and qualityof wild *Akebia trifoliate* (Thunb.) Koidz. Guizhou University. 2017.
58. Zhang EZ, Liao F, Yang RG, He Qg, Huang MH, Xin M, Huang ZY. Comparison of cell structure of the fruit peel in seventeen varieties of *Mangifera indica* L during different postharvest stages. Chinese J Tropical Crops. 2016;37(11):2238-43.
59. 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.
60. 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.
61. 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.
62. 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.
63. Yang H. Analysis of gene expression on ethylene biosynthesis of *Akebia trifoliate* (Thunb.) Koidz in ripping process. Guizhou University. 2016.
64. 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;14:581-9.
65. MacLean B, Tomazela DM, Shulman N, Chambers M, Finney GL, Frewen B, Kern R, Tabb DL, Liebler DC, MacCoss MJ. Skyline: an open source document editor for creating and analyzing targeted proteomics experiments. Bioinformatics. 2010;26(7):966-8.

Additional Files

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

Additional file 2: Figue S2. Volcano plot depicting the proteomics data of *A. trifoliate*. 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 btween proteins having absolute log2 fold change of 1. Red dots represent proteins with p<0.05 and absoulte log2 fold changes above 1. Black dots indicate no difference in protein expression.

Additional file 3: Figue S3. 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 4: Figue S4. GO and KEGG pathway functional enrichment analysis of co-regulated genes and proteins in *A. trifoliate*. 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. Table S1 Sequencing statistics for *A. trifoliata*

Additional file 6: Table S2 Statistical data of unigenes annotation

Additional file 7: Table S3.The summary of the total number of transcripts and proteins identified from different stages and replicates.

Additional file 8: Table S4. Sequences of specific primers used for qPCR experiment.

Tables

Table 1 Summary of the transcriotome and proteome data in *Akebia trifoliata fruits.*

RNA-seq data		MS data based on transcriptome	
Total number of transcripts	241376	Total spectra	812625
Mean length of transcripts (bp)	515	Identified spectra	68151
Total number of unigenes	186054	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		

Table 2 Summary of transcripts and proteins detected from RNA and TMT sequence data.

	Transcriptome		Protein	
	PM_PS	PL_PM	PM_PS	PL_PM
Unique proteins/genes detected	100329	100329	2839	2839
Significantly DEGs/DAPs	4607	6598	190	50
Up- regulated	1945	4623	84	28
Down- regulated	2662	1975	106	22
Shared genes/proteins	1904	1904	17	17
Shared genes/proteins (up- regulated)	1123	808	9	8
Shared genes/proteins (down- regulated)	781	1096	8	9
Co-regulated DEGs-DAPs	14	4	14	4
Co-regulated DEGs-DAPs with the same trends	12	4	12	4

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

Accession	Protein description	Gene/Protein name	Pearson correlation efficient	*P*-value	Numbers
		Gene			
TRINITY_DN135342_c3_g3	Dirigent protein 22-like	DIR2	0.7977	0.0100	18
TRINITY_DN138969_c0_g6	Shikimate O-hydroxycinnamoyltransferase	HCT	0.2848	0.4577	18
TRINITY_DN136551_c1_g11	Cellulose synthase-like protein G3	CSLG	-0.06589	0.8663	18
TRINITY_DN136342_c2_g5	NAC domain-containing protein 100-like	NAClike	0.9464	0.0001	18
TRINITY_DN141308_c1_g4	expansin-A1	EXP1	0.8582	0.0031	18
TRINITY_DN141875_c0_g4	Cinnamyl alcohol dehydrogenase	CAD	0.8015	0.0094	18
TRINITY_DN138388_c1_g1	Beta-galactosidase 3-like	β-GAL1	0.8440	0.0169	18
TRINITY_DN142386_c5_g1	Beta-galactosidase	β-GAL2	0.6675	0.0495	18
TRINITY_DN196976_c0_g1	Polygalacturonase	PG	0.7863	0.0636	18
TRINITY_DN141686_c0_g5	4-coumarate--CoA ligase-like 5	4CL	0.7466	0.0208	18
TRINITY_DN138197_c1_g5	Glucan endo-1,3-beta-glucosidase	ENDOB	0.7052	0.0338	18
TRINITY_DN131789_c1_g4	NAC domain-containing protein	NAC	0.1755	0.6515	18
TRINITY_DN143028_c0_g1	Pectinesterase	PE	0.8042	0.0292	18
TRINITY_DN142336_c1_g1	Peroxidase	PRX	0.4798	0.1912	18
TRINITY_DN141432_c1_g2	Beta-D-xylosidase 2	BXL	0.1761	0.6505	18
TRINITY_DN143250_c1_g6	Pectate lyase	PL	0.6021	0.0862	18
TRINITY_DN138043_c3_g7	Transcription factor bHLH66	BHLH	0.7485	0.0327	18
TRINITY_DN142042_c0_g2	Polygalacturonase	PG3	0.6819	0.0430	18
TRINITY_DN76417_c0_g1	Glucan 1,3-beta- glucosidase	CEL	0.8732	0.0103	18
TRINITY_DN141264_c1_g1	Peroxidase N1	PRX2	0.8325	0.0054	18
		Protein			
TRINITY_DN135837_c0_g4	Dirigent-like protein	DIR1	0.8316	0.0402	18
TRINITY_DN142943_c1_g1	Polygalacturonase	PG2	0.8336	0.0101	18
TRINITY_DN141308_c1_g4	Expansin-A1	EXP1	-0.7482	0.0204	18
TRINITY_DN139379_c0_g3	Peroxidase	PRX3	-0.1040	0.7900	18
TRINITY_DN142424_c1_g1	Furostanol glycoside 26-O-beta-glucosidase	F26G	0.8907	0.0013	18
TRINITY_DN137437_c3_g1	Beta-glucosidase 33	BGLU33	0.7248	0.0272	18
TRINITY_DN143028_c0_g1	Pectinesterase	PE	0.7596	0.0176	18
TRINITY_DN142336_c1_g1	Peroxidase	PRX	0.7489	0.0202	18
TRINITY_DN141432_c1_g2	Beta-D-xylosidase 2	BXL	0.6766	0.0453	18
TRINITY_DN141662_c1_g4	Alpha/beta hydrolase	α-HY	0.8270	0.0060	18
TRINITY_DN143250_c1_g6	Pectate lyase	PL	0.4460	0.2289	18
TRINITY_DN141264_c1_g1	Peroxidase N1	PRX2	0.5422	0.1315	18
TRINITY_DN137008_c1_g6	Peroxidase	PRX4	0.7481	0.0204	18
TRINITY_DN139660_c0_g1	Peroxidase 53	PRX5	0.6748	0.0462	18
TRINITY_DN141880_c0_g1	glucan endo-1,3-beta-glucosidase	ENBG	0.8917	0.0029	18
TRINITY_DN138388_c1_g1	Beta-galactosidase 3-like	β-GAL1	-0.8215	0.0234	18
TRINITY_DN142386_c5_g1	Beta-galactosidase	β-GAL2	0.7797	0.0132	18

Table 4 Comparison of PRM and TMT quantification results

Accession	Protein description	Gene name	PM_PS ratio		PL_PM ratio	
			PRM	TMT	PRM	TMT
TRINITY_DN143250_c1_g6	Pectate lyase	PL	43.325	1.392	2.029	#N/A
TRINITY_DN143028_c0_g1	Pectinesterase	PE	1.470	0	1.695	1.293
TRINITY_DN142943_c1_g1	Polygalacturonase	PG2	0.799	0.566	0.993	#N/A
TRINITY_DN142424_c1_g1	Furostanol glycoside 26-O-beta-glucosidase	F26G	4.390	1.853	1.185	1.301
TRINITY_DN142386_c5_g1	Beta-galactosidase	β-GAL2	1.620	0	1.071	1.394
TRINITY_DN142336_c1_g1	Peroxidase	PRX	1.303	0.594	0.707	1.304
TRINITY_DN142120_c0_g1	Auxin efflux carrier	AEC	2.593	1.245	0.535	0
TRINITY_DN142119_c1_g3	Beta-fructofuranosidase	β-FRU	1.497	0.655	0.500	0
TRINITY_DN141662_c1_g4	Alpha/beta hydrolase	α-HY	3.192	1.243	0.306	0
TRINITY_DN141432_c1_g2	Beta-D-xylosidase	XYL	1.717	0	0.606	1.398
TRINITY_DN141264_c1_g1	Peroxidase N1	PRX2	0.828	0.619	1.651	0
TRINITY_DN141074_c0_g1	Polygalacturonase	PG4	1.620	1.464	0.680	0
TRINITY_DN139660_c0_g1	Peroxidase 53	PRX5	0.703	0.644	1.239	0
TRINITY_DN139379_c0_g3	Cationic peroxidase 1	PRX3	0.421	0.546	1.208	0
TRINITY_DN137437_c3_g1	Beta-glucosidase 33	BGLU33	1.291	0.636	0.317	0
TRINITY_DN137338_c1_g4	Endoglucanase 19	ENDOG19	15.470	1.502	0.232	0
TRINITY_DN136333_c1_g2	Endoglucanase 8	ENDOG8	3.207	1.220	0.696	0
TRINITY_DN135837_c0_g4	Dirigent-like protein	DIR1	0.733	0.597	0.927	0

Figures

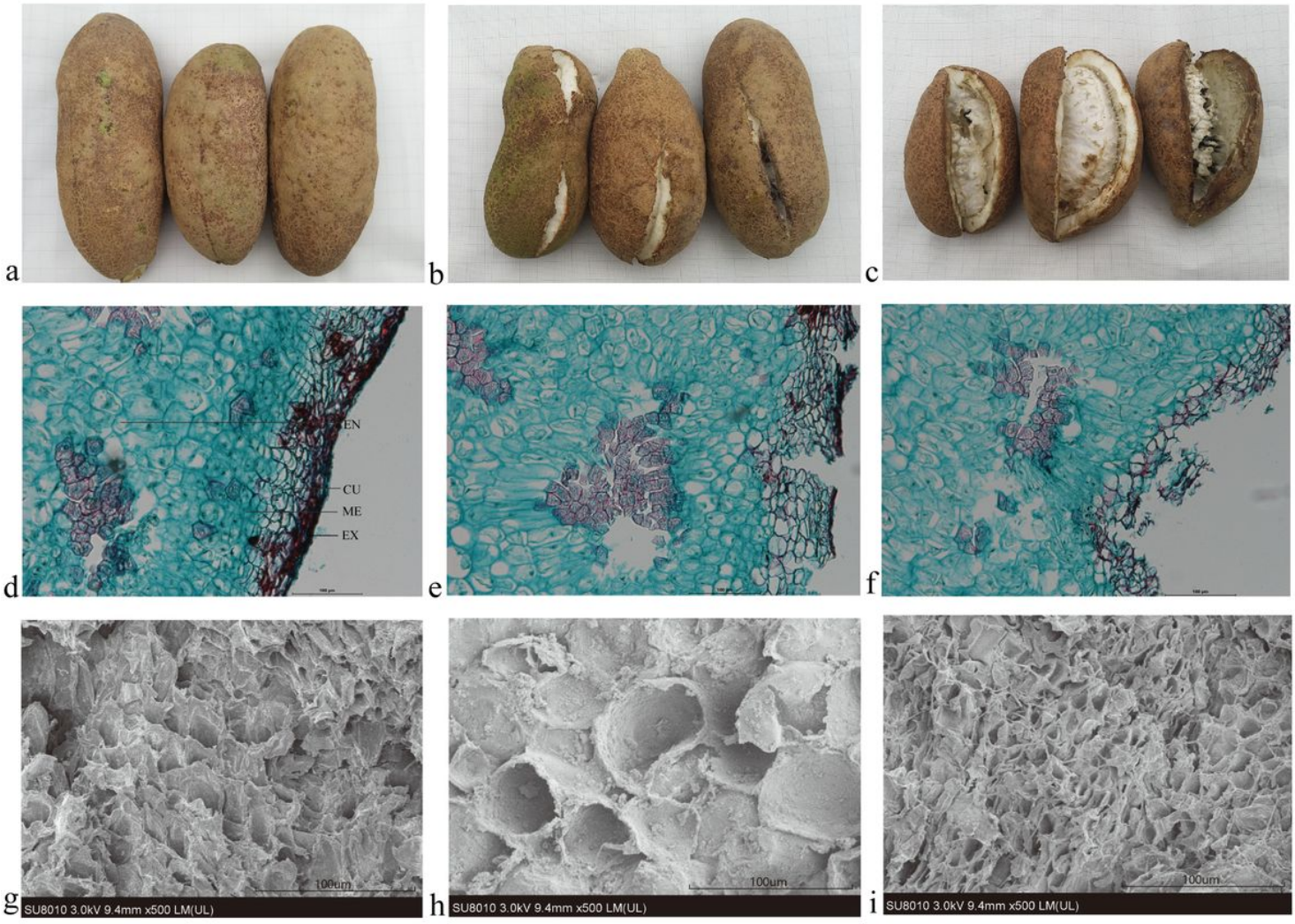

Figure 1

The morphology and structure Changes of ‘Nong No.8’ pericarp at different development stages. a-c The morphology changes of A. trifoliate pericarp in non-cracking stage (PS) (Fig. 1a), initial cracking stage (PM) (Fig. 1b) and total cracking stage (PL) (Fig. 1c), respectively. d-f Pericarp structure of ‘Nong No.8’ at different development stages using the semi-thin slices method and stained with Safranin O-staining in PS (Fig. 1d), PM (Fig. 1e) and PL (Fig. 1f). g-i Pericarp structure of ‘Nong No.8’ at different development stages using the scanning electron microscope method in PS (Fig. 1g), PM (Fig. 1h) and PL (Fig. 1i). EX. Exocarp; ME. Mesocarp; EN. Endocarp; CU. Cuticle.

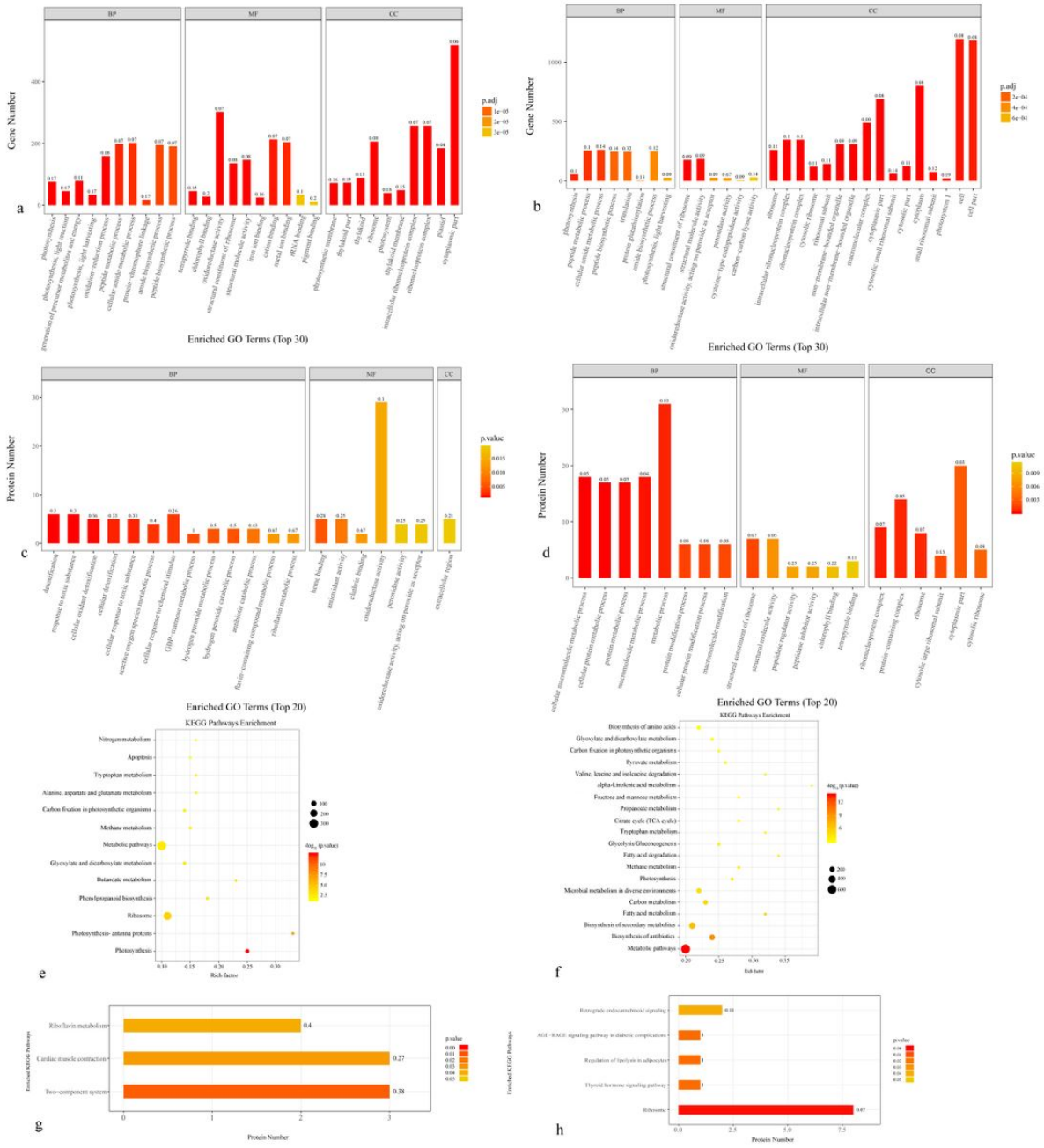

Figure 2

The enrichment analysis of the DEGs and DAPs using GO and KEGG pathways. a-b GO classifications of DEGs in PM_PS (Fig. 2a) and PL_PM (Fig. 2b), respectively. The horizontal axis shows the top 30 enriched GO terms in the biological process, molecular function, and cellular component categories. The vertical axis shows the number of genes enriched in each term. c-d GO classifications of DAPs in PM_PS (Fig. 2c) and PL_PM (Fig. 2d), respectively. The horizontal axis shows the top 20

enriched GO terms in the biological process, molecular function, and cellular component categories. The vertical axis shows the number of genes enriched in each term. e-f KEGG pathways enrichment of DEGs in PM_PS and PL_PM, respectively. g-h KEGG pathways enrichment of DAPs in PM_PS and PL_PM, respectively. The horizontal axis shows the enrichment factors of the KEGG pathways terms (e-f), and the number of genes enriched in each term (g-h). The vertical axis shows the top enriched KEGG pathways terms. The size and color of the circle indicates the number of genes and P value of the enriched term, respectively.

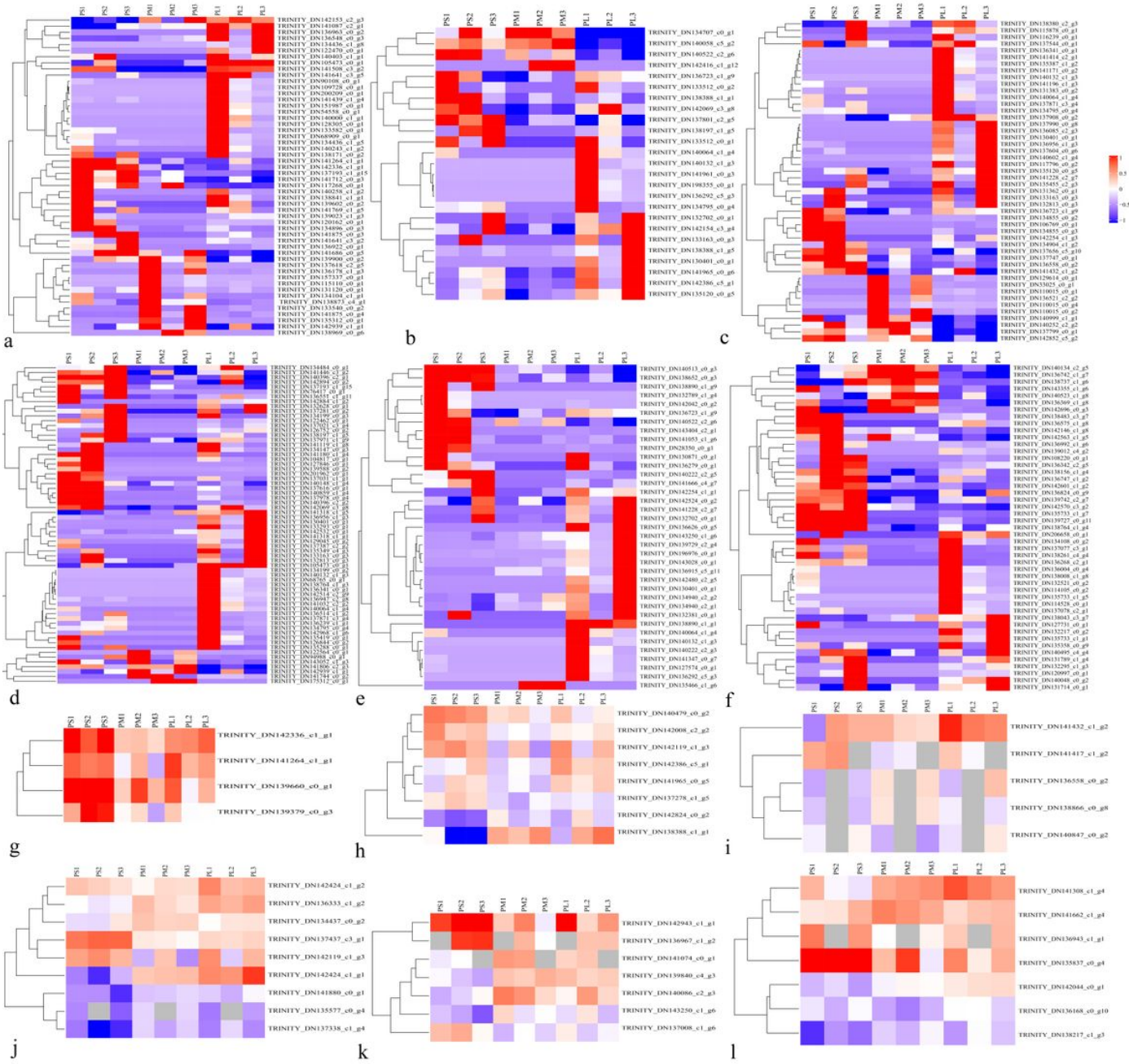

Figure 3

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 blue indicates significantly downregulated proteins. White 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.

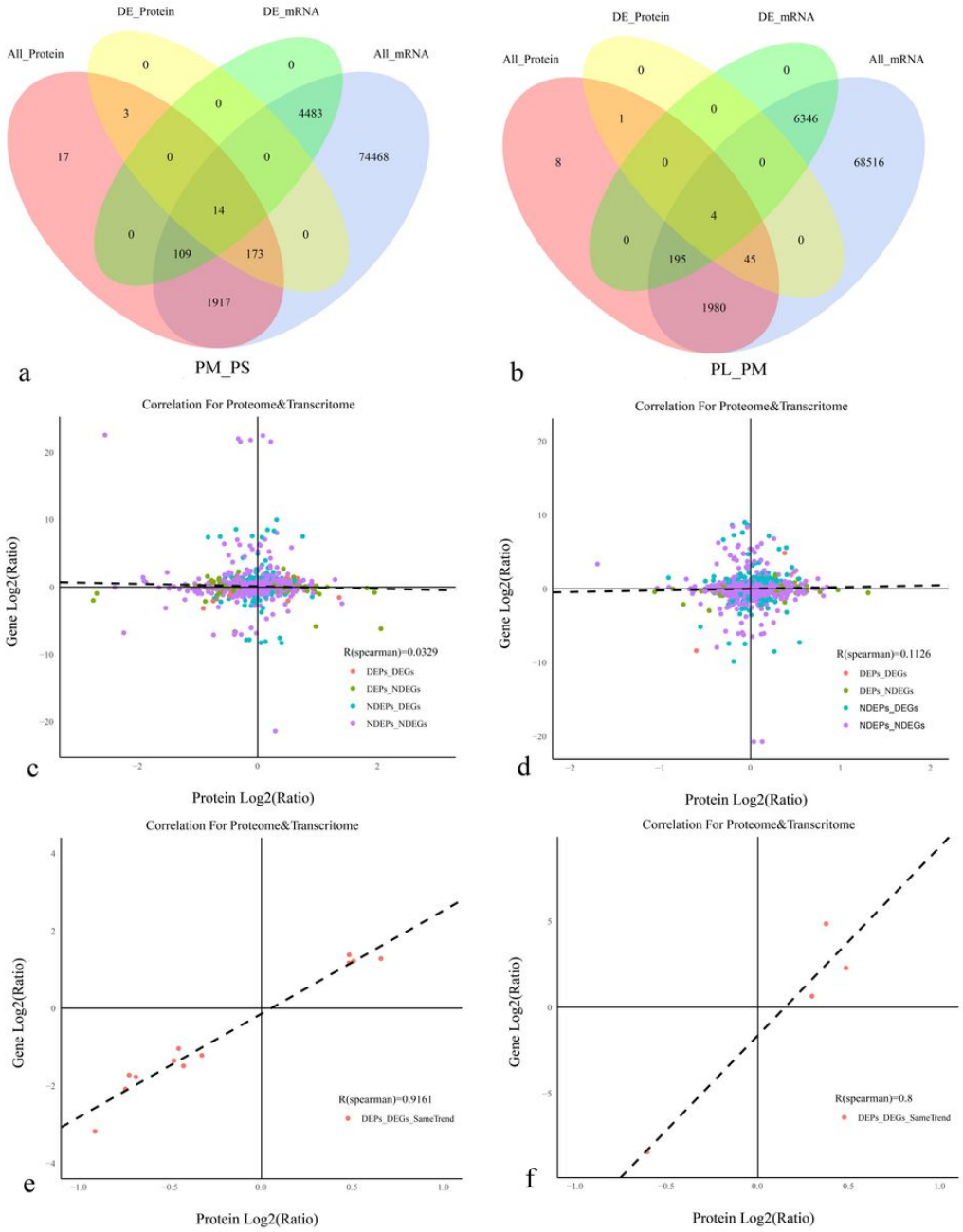

Figure 4

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 betweengenes 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.

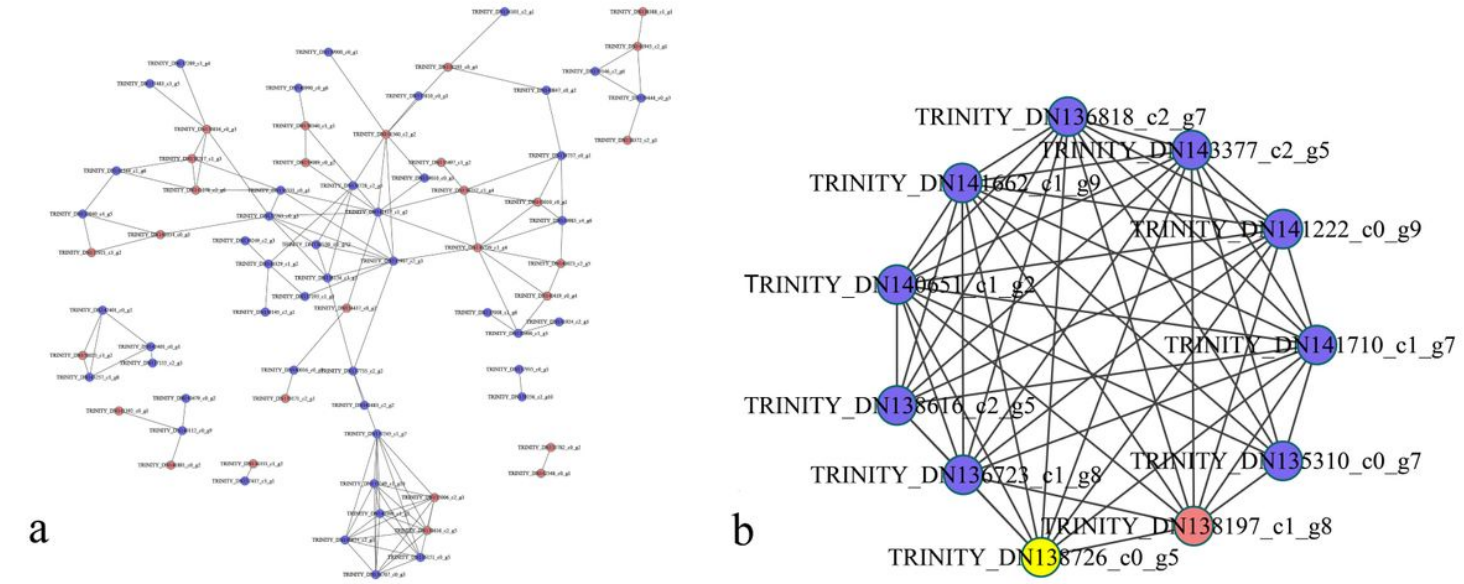

Figure 5

Analysis of the functional network by STRING 9.0 of DAPs. a Analysis of the functional network by STRING 9.0 of DAPs in PM_PS. b Analysis of the functional network by STRING 9.0 of DAPs in PL_PM. Red indicates significantly upregulated proteins, and blue indicates significantly downregulated proteins.

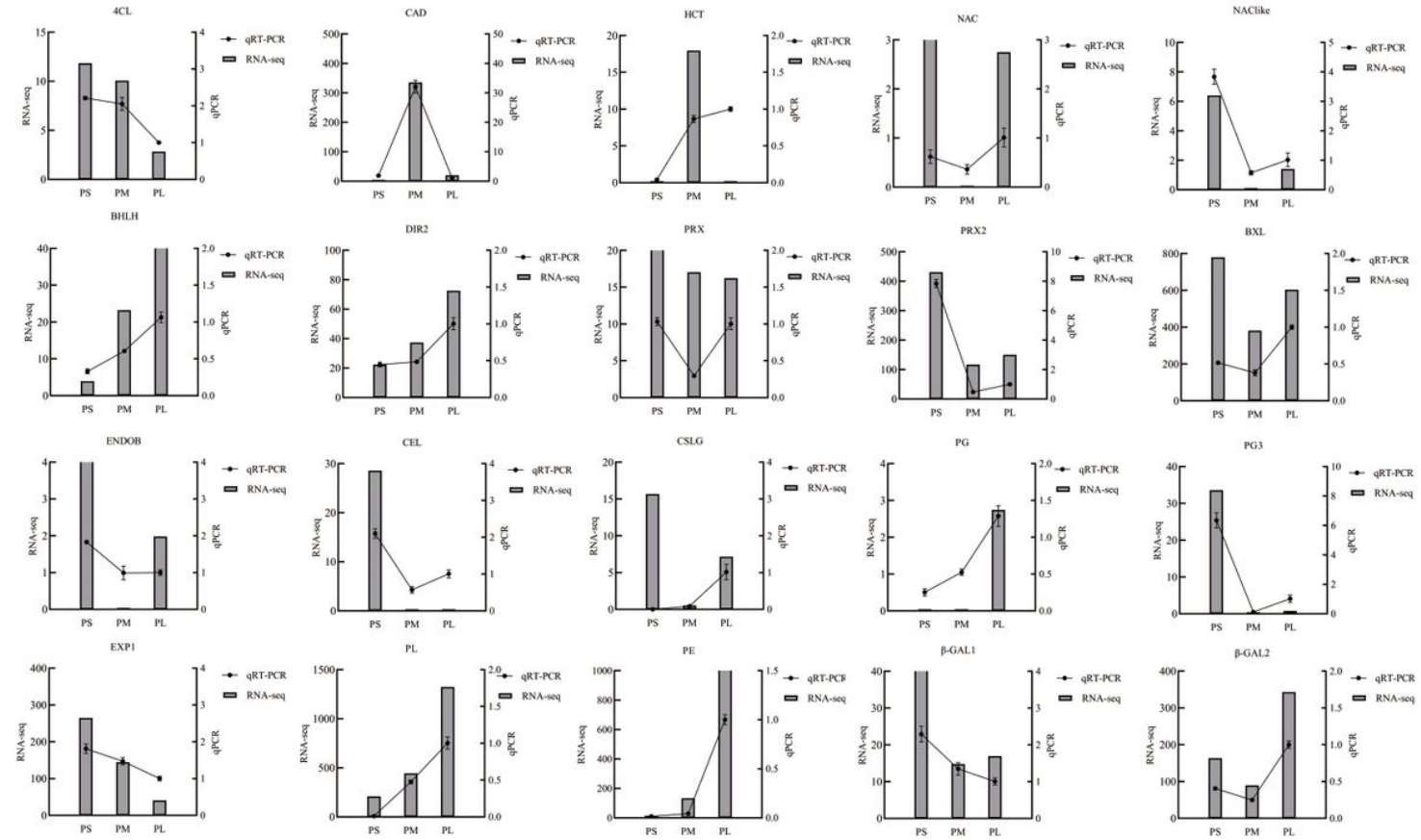

Figure 6

Validation and expression analysis of selected genes using qPCR. The expression levels of the genes revelaed 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.

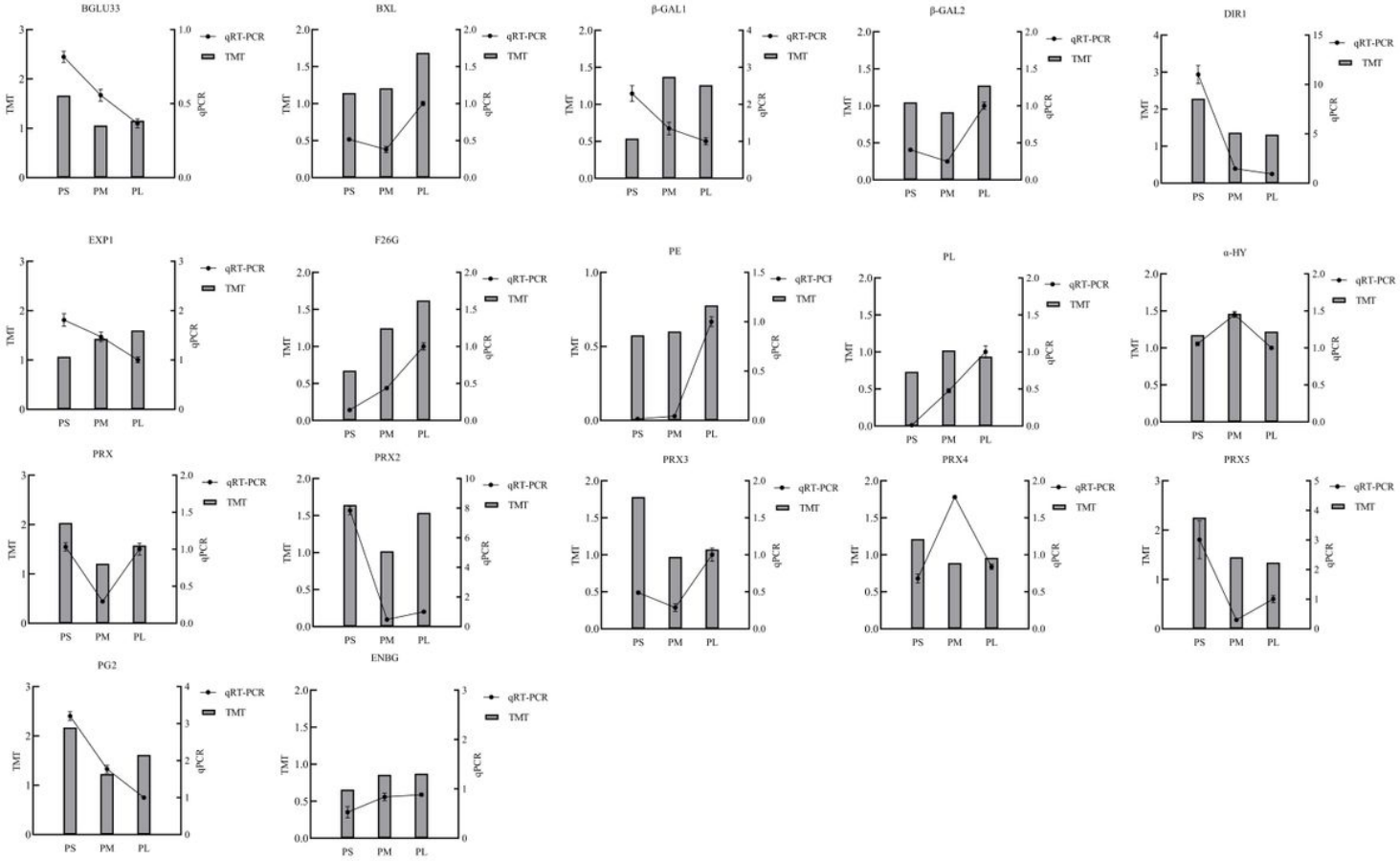

Figure 7

Validation and expression analysis of selected proteins using qPCR. The expression levels of the genes revelaed 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

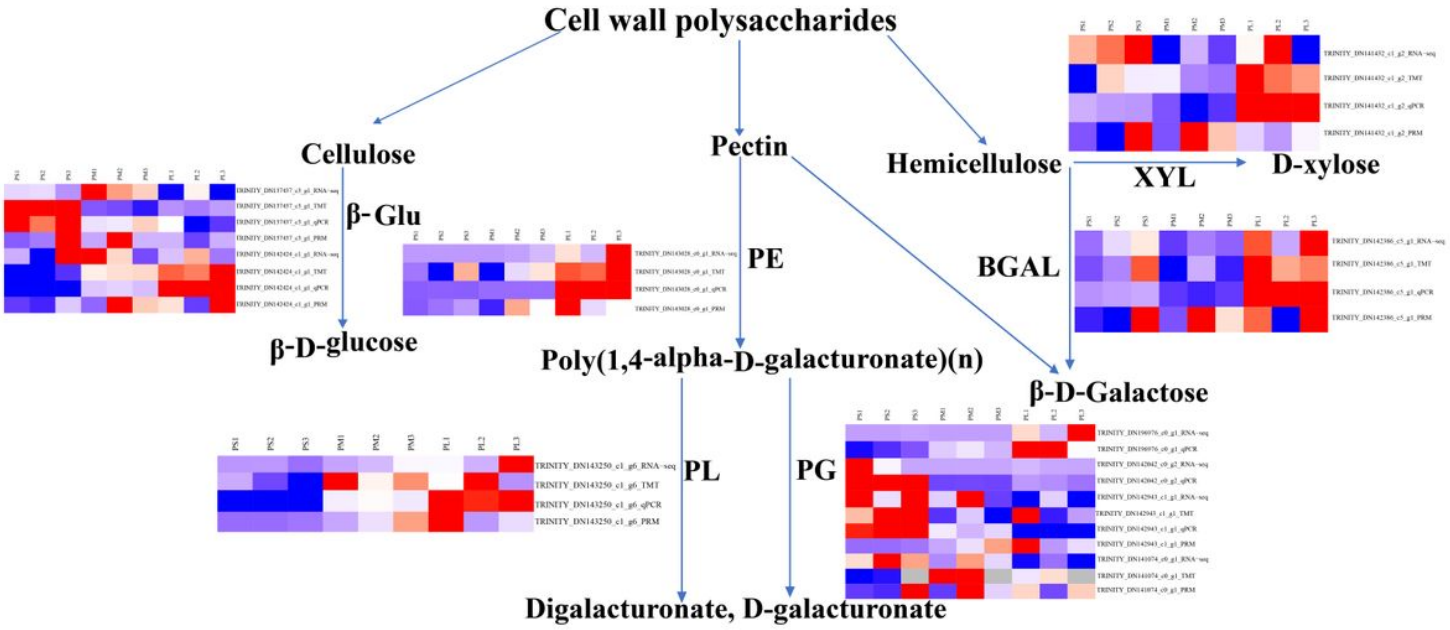

Figure 8

Summary of some of the biological pathways involved in A. trifoliate fruit ripening and cracking. TRINITY_DN143250_c1_g6 (PL); TRINITY_DN196976_c0_g1 (PG), TRINITY_DN142042_c0_ g2 (PG3); TRINITY_DN142943_c1_g1 (PG2); TRINITY_DN143028_c0_g1 (PE); TRINITY_ DN141074_c0_g1 (PE2); TRINITY_DN137437_c3_g1 (BGLU33); TRINITY_DN141880_c0_g1 (ENBG); TRINITY_DN142424_c1_g1 (F26G); TRINITY_DN141432_c1_g2 (BXL). Red indicates significantly upregulated proteins, and blue indicates significantly downregulated proteins. White indicates proteins with no significant changes.

Supplementary Files

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

- supply.rar