A Comparative Transcriptome Analysis of Purple and Yellow Fleshed Potato Tubers Reveals Long Non-coding RNAs and their Targets Functioned in Anthocyanin Biosynthesis

Ruimin Tang
Shanxi Agricultural University
Haitao Dong
Shanxi Agricultural University
Wanyi Wu
Shanxi Agricultural University
Cailiang Zhao
Shanxi Agricultural University
Xiaoyun Jia (jiaxiaoyun@sxau.edu.cn)
Shanxi Agricultural University
Qing Yang
Nanjing Agricultural University
Jie Zhang
Shanxi Agricultural University
Liheng He
Shanxi Agricultural University
Hong'e Xie
Shanxi Academy of Agricultural Sciences
Zongxin Wu
Shanxi Academy of Agricultural Sciences

Research Article

Keywords: potato, anthocyanin biosynthesis, long non-coding RNA (lncRNA), transcriptome sequencing, co-expression network

Posted Date: May 26th, 2021

DOI: https://doi.org/10.21203/rs.3.rs-515121/v1

Abstract

Background: Purple fleshed potato tubers accumulate large amounts of anthocyanin content, servicing as functional foods and high-value feedstock. Long non-coding RNAs (lncRNAs) have been reported to play an important role in anthocyanin synthesis by regulating gene expression in various action modes. However, the mechanism underlying anthocyanin accumulation mediated by lncRNAs in underground organs remains unclear.

Results: To excavate the differentially expressed lncRNAs (DE lncRNAs) between purple and yellow fleshed potato tubers, the transcriptome sequencing was performed and a total of 1421 DE lncRNAs were identified. Gene Ontology and KEGG (Kyoto Encyclopedia of Genes and Genomes) enrichment analyses showed that the target genes of these DE lncRNAs were involved in diverse biological processes and pathways for anthocyanin biosynthesis, reflecting the functional diversity of the corresponding lncRNAs. A lncRNA-mRNA interaction network was created based on their correlation to investigate the regulatory relationship among them. Notably, lncRNAs like XLOC_060098 and XLOC_017372 might contribute to anthocyanin synthesis by targeting the key enzyme genes and transcription factor genes in the pathway.

Conclusions: The construction of expression profiling of DE lncRNAs and lncRNA-mRNA relationship network is helpful for further unraveling the molecular mechanisms of lncRNAs in anthocyanin synthesis in potato tubers, and provides theory basis for the cultivation of functional potato varieties and the improvement of nutritional quality of other underground crops.

Background

Potato (*Solanum tuberosum* L.), the fourth largest food crop in the world following rice, wheat and corn, is an important resource of raw materials for food and industrial processing[1, 2]. As a special potato cultivar, purple fleshed potato is a desirable resource for functional food development because of the abundant anthocyanin accumulated in its tubers[3]. Anthocyanin is one of the main secondary metabolites in the biosynthesis of plant flavonoid, which makes flowers, fruits and other organs show various colors under different pH conditions in plant vacuole[4, 5]. Due to its outstanding free radical scavenging capacity, anthocyanin was demonstrated to have healthcare effects such as antioxidant, anti-aging, anti-tumor and immune activity regulation[6-8]. Different from the anthocyanin in aboveground organs of plants like flowers, leaves and fruits, the anthocyanin in purple potato tubers is synthesized and accumulated without direct action of light[9], which makes potato an ideal research material for investigating the mechanism of anthocyanin accumulation in underground organs.

Anthocyanin biosynthesis occurs in the endoplasmic reticulum of the cytoplasm, starting from phenylalanine and being transported to vacuole or cell wall for storage after multiple enzymatic reactions and modifications of methylation, glycosylation, and acylation[10]. This process is mediated by numerous structural genes, regulatory genes and other regulatory factors[11]. The structural genes encoding a series of enzymes, such as phenylalanine ammonia lyase (PAL), chalcone synthase (CHS), dihydroflavonol reductase (DFR), anthocyanidin synthase (ANS), UDP-glucose flavonoid 3-o-glycosyltransferase (UFGT), have been studied extensively in a variety of plants including potato[12]. In recent years, some new structural genes have been found to be involved in anthocyanin synthesis like *supressor of phytochrome A-105* (*SPA*) in *Arabidopsis*[13] and *glutathione transferase 4* (*LCGST4*) in litchi[14]. These structural genes are usually transcriptionally modulated by the regulatory genes, which are primarily transcription factors. R2R3-MYB, the most widely studied transcription factor, usually interacts with bHLH and WDR proteins to form the WBM protein complex, and then binds to the cis-acting elements in the enzyme gene promoter region to regulate anthocyanin synthesis[15]. Besides, other transcription factors, such as phytochrome-interacting factor 3 (PIF3), ethylene insensitive (EIN3) and squamosa promoter binding proteinlike (SPL) have also been reported to participate in the regulation of anthocyanin synthesis in plants[16, 17]. In addition to mRNAs, long non-coding RNAs (lncRNAs) have also been discovered to mediate anthocyanin synthesis.

LncRNAs with a length of more than 200 nt contain the cap structure and poly-A tail like mRNAs, but have no protein coding capacity[18, 19]. LncRNAs regulate the expression of their downstream genes in different action modes at the transcriptional or post-transcriptional levels[20]. The cis-acting target genes, also known as co-located genes, are considered to be influenced by the activity of their adjacent lncRNAs; while the trans-acting target genes are defined as the co-expressed genes of lncRNA, which could be regulated by distant lncRNAs [21]. Additionally, lncRNAs were identified to act as targets of miRNAs and regulate gene expression by competitively interact with miRNAs[22]. Some lncRNAs can function as precursors of miRNAs and TasiRNA (trans-acting small interfering) and mediate RdDM (RNA-directed DNA methylation) scilencing[23]. Also, some antisense lncRNAs may regulate gene silencing, transcription and stabilization by binding to mRNAs on the sense strand via complementary pairing of bases[5].

In recent years, several lncRNAs have been found to take part in plant growth, flower and fruit development, stress response, and other biological processes[24, 25]. For example, COOLAIR (Cold induced antisense intragenic RNA) and COLDAIR (Cold assisted intronic noncoding RNA) were reported to inhibit the flowering genes to achieve rapid flowering transition in *Arabidopsis*[26, 27]. In rice, a lncRNA named LDMAR (Long-day-specific male-fertility associated RNA) was revealed as a participant in DNA methylation process, which resulted in male sterility[28]. TAS3 (trans-acting siRNA3), which was considered as a nitrogen (N)-responsive lncRNA in *Arabidopsis*, has been confirmed to promote lateral root development[29]. In addition to these biological processes, lncRNAs have also been reported to be involved in anthocyanin synthesis pathways. At present, the anthocyanin synthesis associated lncRNAs were identified in a few plant species like sea buckthorn, apple and strawberry [21, 30, 31]. For instance, during fruit ripening stage in sea buckthorn, LNC1 and LNC2 mediated the synthesis of anthocyanin by regulating the expression of *SPL9* and *MYB114*[31]. In apple, MLNC3.2 and MLNC4.6 promoted the expression of *SPL2-like* and *SPL33* during light-induced anthocyanin synthesis[30]. To date, the in-depth researches on the lncRNAs involved in anthocyanin synthesis mostly focused on the aboveground organs. The functions and regulatory mechanism of lncRNAs during anthocyanin biosynthesis in underground organs are still obscure.

In this study, transcriptome sequencing between yellow fleshed (‘Jin-16’) and purple fleshed potato tubers (‘Xisen-8’) were carried out to excavate the potential lncRNAs and their target genes that associated with anthocyanin biosynthesis. Furthermore, the functional enrichment analyses were conducted on the target genes of differentially expressed lncRNAs (DE lncRNAs) to reveal the potential biological processes and pathways in which lncRNAs involved. A lncRNA-mRNA co-expression network was constructed based on their correlation to unravel the regulatory mechanism of lncRNAs during anthocyanin accumulation. These findings replenish the molecular mechanism of non-coding RNAs in the regulation of anthocyanin synthesis, which will be of great guiding significance for the cultivation of new potato varieties with high added value.

Results

Phenotype comparison and anthocyanin content determination in potato tubers of cultivar ‘Jin-16’ and ‘Xisen-8’

Tubers appearance of potato cultivar ‘Jin-16’ and ‘Xisen-8’ showed that both of the two cultivars were of long oval shape with shallow bud eyes and smooth skin. The tuber skin and flesh of cultivar ‘Xisen-8’ were observed dark purple (Fig. 1b) compared with those of ‘Jin-16’, which were yellow in color (Fig. 1a).

The anthocyanin contents of tuber flesh in ‘Jin-16’ and ‘Xisen-8’ were detected, respectively. As shown in Fig. 1c, the relative anthocyanin content of tuber flesh in ‘Xisen-8’ was significantly higher than that in ‘Jin-16’ (~ 103.5 fold), indicating that there may exist a series of key regulatory factors in ‘Xisen-8’ which promoted anthocyanin accumulation.

Sequencing and assembly of the lncRNA-Seq datasets

To identify the lncRNAs involved in the process of anthocyanin biosynthesis, the tubers of 'Jin-16' and 'Xisen-8' were used to construct lncRNA libraries and sequenced on Illumina novoseq6000 platform. After eliminating the low quality reads, Illumina adapters and reads with unidentifiable base information, the clean reads were obtained from each sample and accounted for more than 95% of the raw reads (Fig. 2a). Among these clean reads, the exonic reads made up more than 60%, followed by intergenic reads (less than 35%) and intronic reads (less than 5%) (Fig. 2b).

The filtered reads were subsequently aligned against a potato reference genome (*S. tuberosum* assembly 4.03) by Hisat2 software[32]. The sequence alignment results of reads mapped to the reference genome were presented in Additional file 1. Finally, 6371 non-coding transcripts corresponding 4376 lncRNAs were identified as the candidate lncRNA datasets based on the analyses of CPC2, PFAM and CNCI softwares (Additional file 2). According to the location relationship between lncRNAs and known mRNAs, lncRNAs could be classified into 4 categories: long intergenic lncRNAs (lincRNAs), antisense lncRNAs, sense overlapping lncRNAs and sense intronic lncRNAs[33]. In this study, lincRNAs contained the largest number with 3461 lncRNAs (79.1%), followed by antisense and sense overlapping lncRNAs with 705 (16.1%) and 210 (4.8%) lncRNAs, respectively. And no sense intronic lncRNAs were identified (Fig. 2c).

Construction of expression profile and Chromosomal localization of differentially expressed lncRNAs in potato tubers

After difference significance analysis, a total of 1421 DE lncRNAs were acquired between 'Jin-16' and 'Xisen-8' tuber samples. The expression levels of these 1421 DE lncRNAs were shown in Additional file 3. In this study, hierarchical clustering analysis was conducted using the FPKM value of DE lncRNAs to cluster the genes with similar expression patterns, which may play analogical roles or participate in common metabolic pathways (Fig. 3a).

Of those DE lncRNAs, 735 lncRNAs were identifed in both 'Jin-16' and 'Xisen-8' tuber samples; while 325 and 361 DE lncRNAs were found to be specifically expressed in 'Jin-16' (like XLOC_029213 and XLOC_053592) and 'Xisen-8' (like XLOC_026093 and XLOC_085889), respectively (Fig. 3b). Compared with 'Jin-16' tubers, there were 735 up-regulated and 686 down-regulated DE lncRNAs in 'Xisen-8' tubers (Fig. 3c). The identification of these DE lncRNAs is conducive to further investigation on anthocyanin synthesis mechanism in potato tubers.

The distribution of these 1421 DE lncRNAs on potato chromosomes and their expression levels were intuitively displayed in Fig. 3d. These DE lncRNAs were unevenly distributed in all potato chromosomes, which were most distributed on CHR1 (191) and least distributed on CHR2 (84). Other chromosomes, such as CHR0, 4, 6, 7, 8, 9 and 11 contained more than 100 DE lncRNAs; while CHR3, 5, 10, 12 contained less than 100 DE lncRNAs. The exact position of the 1421 DE lncRNAs on chromosomes were presented in Additional file 3. The chromosomal localization of lncRNAs contributes to obtaining complete lncRNA sequences and provides basis for better understanding of their functions.

Functional category of DE lncRNA targets by GO enrichment analysis

The co-located and co-expressed targets of all identified lncRNAs in potato tubers were presented in Additional file 4 and Additional file 5. The GO enrichment analysis of genes co-located and co-expressed with DE lncRNAs were performed to predict the potential functions of DE lncRNAs between 'Jin-16' and 'Xisen-8'. In this study, the top 40 significantly enriched GO terms were shown in Fig. 4 and Additional file 6. Of the GO annotations associated with mRNAs co-located with DE lncRNAs, 22 GO terms were assigned to biological process (BP), 8 and 10 GO terms belonged to molecular function (MF) and cellular component (CC), respectively. These genes were mainly enriched in "cellular metabolic process" (GO:0044237), "phosphorus metabolic process" (GO:0006793), "hydrolase activity, acting on ester bonds" (GO:0016788), "transferase activity, transferring acyl groups" (GO:0016746), "membrane-enclosed lumen" (GO:0031974) and "organelle lumen" (GO:0043233).

For the genes co-expressed with DE lncRNAs, the top 40 significantly enriched GO terms were grouped into BP (12 GO terms) and MF (28 GO terms). And no GO term was classified in CC. In the category of BP, “biological_process” (GO:0008150), “metabolic process” (GO:0008152), and “organic substance metabolic process” (GO:0071704) including 7712 (3.26%), 5905 (2.50%), 4606 (1.95%) and 36 (2.54%) genes respectively, were the predominant GO terms. In MF category, “molecular_function” (GO:0003674), “catalytic activity” (GO:0003824) and “transferase activity” (GO:0016740) were the most representative terms and contained 9381 (3.97%), 5444 (2.30%) and 2115 (0.89%) genes, respectively. The results of GO analysis implied that genes enriched in these GO terms were probably involved in anthocyanin biosynthesis. Therefore, DE lncRNAs were the potential participants in these biological processes, mediating the regulation of anthocyanin metabolism in potato tubers through cis- and trans-regulation of the expression of their target genes.

KEGG pathway enrichment of DE lncRNA targets

To delve into the metabolic pathways in which the DE lncRNAs were involved, the genes co-located and co-expressed with these DE lncRNAs were subjected to the KEGG database. Totally, the co-located genes and co-expressed genes were identified to participate in 121 and 119 metabolic pathways, respectively (Additional file 7). According to enrichment significance, the top 20 enriched biological pathways with p value < 0.05 were presented in Fig. 5. The genes co-located and co-expressed with DE lncRNAs were both mainly involved in “DNA replication”, “Homologous recombination”, “Tyrosine metabolism”, “Fatty acid degradation”, “Carbon metabolism” and “Nucleotide excision repair”.

Except for these common pathways, the co-located targets of DE lncRNAs were primarily enriched in the pathway of “Sesquiterpenoid and triterpenoid biosynthesis”, “Zeatin biosynthesis”, “Fatty acid biosynthesis”, “Biosynthesis of secondary metabolites”, “Starch and sucrose metabolism”, etc. While, the co-expressed targets of DE lncRNAs were related to “Photosynthesis-antenna proteins”, “Plant hormone signal transduction”, “Carotenoid biosynthesis”, “RNA transport” and so on. The results of KEGG enrichment analysis further illustrated that these lncRNAs and their target genes might be involved in these specific biological pathways to mediate anthocyanin biosynthesis in potato tubers.

According to the expression patterns of the target genes involved in “Flavonoid biosynthesis” (sot00941) and “Anthocyanin biosynthesis” (sot00942), a schematic for anthocyanin synthesis pathway was proposed in potato tubers (Fig. 6). In the pelargonidin synthesis pathway, the expression levels of genes encoding CHS, F3H and DFR were significantly up-regulated, while the expression levels of CHI and ANS had no significantly difference between ‘Jin-16’ and ‘Xisen-8’. The down-regulated expression of *FLS* might reduce kaempferol synthesis and enhance pelargonidin production. Although the biosynthesis of caffeoyl-CoA was inhibited due to the down-regulated genes encoding HQT and C3’H, the significantly up-regulated genes such as CHS, F3’5’H, F3H and DFR might promote the preferred synthesis of delphinidin and cyanidin from caffeoyl-CoA. The DE lncRNAs, which were possibly involved in anthocyanin synthesis by regulating these enzyme genes, were listed in Table 1. Each enzyme gene was regulated by multiple lncRNAs and the same lncRNA could mediate the expression of many different genes.

Table 1

The expression of enzyme genes and their corresponding DE lncRNAs involved in anthocyanin synthesis in ‘Jin-16’ and ‘Xisen-8’ tubers.

Gene name	Annotation	Gene ID	Gene expression		Corresponding LncRNA ID	LncRNA expression	
			Xisen-8	Jin-16		Xisen-8	Jin-16
CHS	Chalcone synthase 1B	PGSC0003DMG400029620	0.808	0.000	XLOC_046144	0.943	0.163
					XLOC_038153	3.103	1.442
					XLOC_047085	1.337	0.258
F3H	Flavanone 3 beta-hydroxylase	PGSC0003DMG400003563	1.037	0.047	XLOC_038153	3.103	1.442
					XLOC_070549	0.000	2.344
DFR	Dihydroflavonol 4-reductase	PGSC0003DMG400003605	4.644	0.012	XLOC_100183	1.443	0.057
					XLOC_107693	10.329	64.517
					XLOC_038153	3.103	1.442
					XLOC_047085	1.337	0.258
					XLOC_075817	32.616	18.880
F3’5’H	Flavonoid 3’,5’-hydroxylase	PGSC0003DMG400000425	1.785	0.387	XLOC_038153	3.1027	1.4422
					XLOC_087286	2.998	0.409
					XLOC_095453	4.643	0.181
					XLOC_015689	4.722	8.089
					XLOC_063942	23.342	10.796
FLS	flavonol synthase	PGSC0003DMG400014093	0.051	0.759	XLOC_030593	0.000	9.902
					XLOC_085889	0.000	10.540
					XLOC_010366	0.000	5.913
					XLOC_025110	0.002	4.598
					XLOC_066673	0.000	4.863
					XLOC_004180	48.485	7.441
					XLOC_075621	46.023	5.710
					XLOC_058142	7.446	0.016
					XLOC_090403	0.065	4.005
					XLOC_040817	0.036	4.267
					XLOC_004081	2.628	9.961
					XLOC_012867	0.148	6.039
					XLOC_094680	0.315	4.110
					XLOC_107693	10.329	64.517

					XLOC_067296	2.898	9.005
					XLOC_043939	1.460	5.937
					XLOC_089954	2.283	12.322
					XLOC_049163	4.807	7.406
					XLOC_019471	3.488	6.027
C3'H	P-coumaroyl quinate/shikimate 3'-hydroxylase	PGSC0003DMG400007178	1.546	3.476	XLOC_047468	1.966	6.560
					XLOC_060098	16.316	2.663
					XLOC_004180	48.4847	7.441
					XLOC_005340	42.397	8.543
					XLOC_015288	62.074	8.770
					XLOC_107693	10.329	64.517
					XLOC_096998	8.304	1.676
					XLOC_105729	8.198	1.047
HQT	shikimate O-hydroxycinnamoyl-transferase	PGSC0003DMG400011189	8.448	46.876	XLOC_029213	4.977	0.000
					XLOC_056281	2.895	6.228
					XLOC_107693	10.329	64.517
					XLOC_093630	0.959	14.693
					XLOC_006486	4.765	1.466

Analysis of lncRNA-mRNA co-expression network and identification of anthocyanin synthesis related lncRNAs and genes

The interaction network of lncRNAs and their co-expressed mRNAs was constructed and visualized to investigate the potential regulatory relationship between them and explore the relevant lncRNAs and genes involved in anthocyanin synthesis (Fig. 7). The lncRNAs and mRNAs in the network were represented by squares and circles, respectively. The results illustrated that 310 DE lncRNAs and 352 target mRNAs were contained in the co-expression network, establishing 771 anthocyanin-responsive lncRNA-mRNA pairs (Additional file 8). The intricate interaction between lncRNAs and their co-expressed genes was manifested in the following aspects: on the one hand, the same lncRNA could regulate the expression of different mRNAs, such as XLOC_060098, XLOC_004180, XLOC_030593 and so on. For example, XLOC_060098 was predicted to be correlated with NAC (PGSC0003DMG400011891), ERF7 (Ethylene-responsive transcription factor 7, PGSC0003DMG401013892), UGT (UDP-glucosyltransferase, PGSC0003DMG400011492) and so on, which were previously reported as important mediators of anthocyanin metabolism. On the other hand, the expression of most mRNAs could be influenced by multiple lncRNAs. For instance, CCoAOMT5 (Caffeoyl-CoA O-methyltransferase 5, PGSC0003DMG400006448) was found to be correlated with a series of lncRNAs, like XLOC_032852, XLOC_032892, XLOC_039335, XLOC_047085, XLOC_048038, XLOC_096885 and XLOC_100659. Similarly, Peroxidase (PGSC0003DMG402025083) was also regulated by numerous lncRNAs, such as XLOC_086959, XLOC_005760, XLOC_027548, XLOC_016209 and XLOC_017372. In addition to the interaction between lncRNAs and their trans target mRNAs, lncRNAs could also interact with each other directly or indirectly. As shown in Fig. 6, XLOC_060098 was associated with XLOC_053197, XLOC_096423, XLOC_027380, XLOC_084558 and XLOC_011664. These results provide some basis for

further screening candidate lncRNAs and genes to study their functions and regulatory mechanisms in anthocyanin synthesis.

Among these 352 target genes regulated by DE lncRNAs, 309 and 43 genes were up-regulated and down-regulated in 'Xisen-8' tubers respectively. The up-regulated genes included the mRNAs encoding CCoAOMT5, AnAT (anthocyanin acyltransferase), NAC2, UGT (UDP-glucose:glucosyltransferase), F3'H and so on, which were recognized as key positive elements in anthocyanin synthesis. While the genes coding for C3H (P-coumarate 3-hydroxylase), ARF (Auxin response factor), SPL (Squamosa promoter binding proteinlike) and HQT (Hydroxycinnamoyl-CoA: quinate hydroxycinamoyl transferase) were down-regulated. These genes have been reported to negatively regulate anthocyanin synthesis. Most genes were observed to interact with XLOC_060098, indicating that XLOC_060098 might be an important regulator participating in various biological processes during anthocyanin synthesis.

The qRT-PCR validation of anthocyanin synthesis-related DE lncRNAs and their targets in potato tubers

In order to validate the expression patterns of potential anthocyanin-associated DE lncRNAs and mRNAs, 6 DE lncRNAs and their corresponding targets, including lncRNAs and mRNAs, were selected from the co-expressed network for qRT-PCR analysis utilizing the same RNA samples as used for RNAseq. The specific primer sequences of the lncRNAs and genes were shown in Additional file 9. The results showed that the expression trends of the DE lncRNAs and their co-expressed targets examined by qRT-PCR were basically consistent with the RNAseq data (Fig. 8a). Although there existed quantitative differences in expression levels between the RNA-seq data and qRT-PCR assay, a high linear correlation (R^2=0.8226) was obtained under two different detection methods, suggesting that the RNA sequencing data were credible (Fig. 8b).

As shown in Fig. 8a, the expression levels of XLOC_060098, XLOC_017372, XLOC_038153 and XLOC_046144 showed the same patterns with that of their corresponding targets, indicating that these lncRNAs positively regulated the expression of their targets, respectively. In addition, lncRNAs could also negatively regulate their target genes. For instance, the expression of XLOC_100954 and its target genes ARF8 and SPL witnessed a downward trend in 'Xisen-8' compared to 'Jin-16', which was in contrast to the expression pattern of Ubiquitin ligase. Similarly, HQT, one of the co-expressed target genes of XLOC_029213, presented a different expression pattern from XLOC_029213 and its other target genes MYC4 (Myelocytomatosis4) and AT (Acyltransferase), whose expression amounts were up-regulated in 'Xisen-8'. The forward and inverse correlations between lncRNAs and their corresponding target genes revealed that the lncRNAs might participate in anthocyanin biosynthesis by regulating the expression of anthocyanin-related genes positively or negatively. Besides, lncRNAs, such as XLOC_017372 and XLOC_046144, could also regulate other lncRNAs via co-expression to mediate anthocyanin biosynthesis.

Discussion

The construction of expression profile of 1421 DE lncRNAs provides new insights into the regulation of anthocyanin biosynthesis in potato tubers

Anthocyanin is a kind of water-soluble pigment with strong antioxidant capacity in plants, contributing to a variety of vibrant colors in different organs and promoting the ornamental and nutritional value of plants[30]. Purple fleshed potatoes are favored by consumers because they not only contain the general nutrients in white or yellow fleshed potatoes, but also enrich high level of anthocyanin content[5]. With the development of processed food and health products from purple fleshed potatoes, the increase of anthocyanin content in tubers has become the aim of potato breeders and growers.

Investigating the regulatory mechanism of anthocyanin biosynthesis is of great guiding significance and practical application value for improving potato tuber quality and facilitating the breeding of potato germplasm resources.

As a class of non-coding RNAs, lncRNAs which were initially regarded as “transcriptional noise”[23], have been reported to play crucial roles in various biological processes such as plant development[34], fruit ripening[35, 36], stress responses[37, 38] and anthocyanin biosynthesis[31]. Although thousands of lncRNAs related to anthocyanin synthesis were identified in fruits of several plant species[21, 30, 31], the lncRNAs regulating anthocyanin biosynthesis in potato tubers have not yet been reported.

In this study, two potato cultivars ‘Jin-16’ (the yellow fleshed potato) and ‘Xisen-8’ (the purple fleshed potato with high anthocyanin content) were used as the experimental materials to excavate the lncRNAs involved in anthocyanin biosynthesis (Fig. 1). Finally, 4376 lncRNAs were identified in potato tubers including 3461 lincRNAs, 705 anti-sense lncRNAs and 210 sense overlapping lncRNAs (Fig. 2c). LincRNAs accounted for the largest proportion of the total lncRNAs in potato tubers, which were similar to those identified in other plants[24, 39], suggesting that the overwhelming majority of lncRNAs are located in intergenic regions. Totally, 1421 DE lncRNAs, including 735 up-regulated and 686 down-regulated lncRNAs, were identified between ‘Xisen-8’ and ‘Jin-16’ libraries by analysis of transcriptome sequencing (Fig. 3a and 3c). LncRNAs in sea buckthorn and strawberry fruits have been reported to display expression specificity in tissue types and developmental stages[21, 31]. In potato tubers, 361 DE lncRNAs were only expressed in ‘Xisen-8’; while 325 DE lncRNAs were observed to be specifically expressed in ‘Jin-16’, suggesting that the expression of lncRNAs also has evident varietal specificity (Fig. 3b). The result of chromosome localization showed that 1312 of the 1421 DE lncRNAs were distributed in CHR1 ~ CHR12, and 109 DE lncRNAs which might be located on the unanchored scaffolds were placed on CHR0 (Fig. 3d). The global identification of DE lncRNAs between the yellow and purple fleshed potatoes provided a novel perspective on the regulation mechanism of anthocyanin biosynthesis in underground organs.

The DE lncRNAs and their target genes participate in diverse biological processes to mediate anthocyanin biosynthesis in potato tubers

LncRNAs perform their functions by binding to DNA, RNA or proteins via multiple different regulatory mechanisms[23]. LncRNAs can act as cis-acting or trans-acting factors [34,35], miRNA target mimicry[30], miRNA precursor[36]; and can be involved in histone modification[37] as well as DNA methylation[38]. Although the functions and regulation mechanisms of most lncRNAs have not been studied clearly, previous studies indicated that lncRNAs are likely to partake in diversified biological processes by regulating the expression of their target genes either in cis- or in trans-acting[30, 40, 41]. Therefore, the function of lncRNAs depends on the function of their corresponding target genes. In this research, the cis (co-located) and trans (co-expressed) target genes of DE lncRNAs between ‘Jin-16’ and ‘Xisen-8’ were identified and subjected to GO and KEGG pathway enrichment analyses to predict the functions of these DE lncRNAs.

For the cis-acting target genes, most of them were associated with “cellular metabolic process”, “phosphorus metabolic process” and “phosphate-containing compound metabolic process” which were belonged to BP category, suggesting that these genes might be involved in phosphorus metabolism to promote anthocyanin synthesis (Fig. 4a). Numerous researches have demonstrated that the anabolism of anthocyanin in plants is usually affected by multifarious environmental factors such as temperature, light, water, nitrogen and phosphorus[30, 42, 43]. Phosphorus is one of the most important factors influencing anthocyanin biosynthesis and Pi deficiency results in purple color in leaves and stems due to the anthocyanin accumulation in plants[44]. Therefore, these related DE lncRNAs and their corresponding target genes probably play a systematic role in regulating phosphorus balance and anthocyanin synthesis in plants. Anthocyanin is synthesized from phenylalanine through a series of enzymatic reactions in cytoplasm and then transported to vacuoles for storage[11]. However, the stability of anthocyanin is low at neutral pH environment and must be enhanced by

glycosidylation and acylation[45]. The involvement of the genes that were significantly enriched in the terms of “hydrolase activity” and “transferase activity” may enhance the anthocyanin stabilization and contribute to the biosynthesis of different anthocyanins. In addition, some genes were clustered into the terms of “membrane-enclosed lumen” and “organelle lumen” in CC category, indicating that these genes might participate in repairing the cellular components (especially lumen regions) under the regulation of their corresponding lncRNAs to provide a suitable microenvironment for the synthesis and accumulation of anthocyanin. For the trans-acting target genes of DE lncRNAs, in addition to the genes involved in phosphorus metabolism, catalytic reactions and other basic metabolic processes as described above, numerous genes were also linked to “binding” like “small molecule binding”, “nucleotide binding”, “nucleoside phosphate binding” and so on (Fig. 4b). This result was in accordance with that in cineraria[46], and suggested that the trans-acting genes of DE lncRNAs were likely to mediate the synthesis of biomacromolecules, the generation of copolymers and the regulation of gene expression during anthocyanin synthesis in potato tubers. Notably, these trans-acting target genes were not classified into the terms of CC category among the top 40 significantly enriched GO terms, revealing that they may not be primarily involved in the synthesis of subcellular structures[47].

The KEGG enrichment analysis showed that the target genes of DE lncRNAs were significantly enriched in “DNA replication”, “Homologous recombination” and “Nucleotide excision repair” (Fig. 5), suggesting that these genes and their corresponding DE lncRNAs might participate in cell proliferation and regulate the formation and development of potato tubers. Similar results have been found in the development of tomato flowers and fruits[48]. Anthocyanin is a kind of most important secondary metabolites, and its synthesis is a quite complicated process which is based on various primary metabolic pathways, such as glucose metabolism, fatty acid metabolism, amino acid metabolism and so on[49]. It has been reported that starch degradation could supply abundant substrates for anthocyanin synthesis in tuberous roots of purple fleshed sweet potato[50]. Therefore, the genes enriched in “Biosynthesis of secondary metabolites”, “Starch and sucrose metabolism”, “Fatty acid biosynthesis and degradation”, “Glycolysis/Gluconeogenesis” and different amino acid metabolism like “Glycine, Serine and Threonine metabolism” are possibly to play necessary roles in anthocyanin synthesis. In addition, the genes involved in “plant hormone signal transduction” may also take part in the regulation of anthocyanin metabolism (Fig. 5b). DELLA protein, a key factor in gibberellin signal transduction pathway, has been revealed to regulate anthocyanin synthesis[51]. These results help to predict the functions of DE lncRNAs and provide a basis for the further study of the regulation mechanism of lncRNAs in anthocyanin biosynthesis.

The DE lncRNAs contribute to anthocyanin biosynthesis by targeting the key enzyme genes and transcription factor genes in potato tubers

Numerous genes encoding enzymes and transcription factors have been recognized to be involved in the anthocyanin metabolic process[52-54]. Enzyme genes such as CHS, CHI, F3H, DFR, ANS, ANT and UFGT play indispensable roles in synthesis, modification and accumulation of anthocyanins in different plants[11, 55, 56]. During anthocyanin biosynthesis, p-Coumaroyl-CoA was synthesized from phenylalanine via a series of enzymes and competed by CHS and HQT[57]. In ‘Xisen-8’ tubers, the expression levels of CHS and F3H were up-regulated; while HQT was down-regulated, enhancing the production of dihydrokaempferol which is the substrate for synthesizng pelargonidin and flavonol[58] (Fig. 6). Some researchers have proposed that DFR could catalyze DHK (dihydrokaempferol), DHQ (dihydroquercetin) and DHM (Dihydromynicetin) to synthesize different anthocyanins[59]. Due to the high expression of DFR and low expression of FLS, pelargonidin production was promoted and flavonol biosynthesis was inhibited. Caffeoyl-CoA is the substrate for generating anthocyanins and chlorogenic acid under the action of CHS and HQT, respectively[60, 61]. Although the synthesis of caffeinyl-CoA might be inhibited by down-regulated expression of HQT and C3’H in ‘Xisen-8’, the high expression levels of CHS, F3H, DFR and F3’5’H promoted the synthesis of delphinidin and cyanidin instead of chlorogenic acid (Fig. 6). Different ratio and content of pelargonidin, delphinidin and cyanidin result in different color of potato tubers[62]. During this process, the expression of these anthocyanin-related enzyme genes would be regulated by multiple

lncRNAs in different plant species[21, 30]. TCONS_01039552, a lncRNA in sea buckthorn fruit, was observed to regulate the expression of F3H[31]. In strawberry fruit, lncRNAs TRINITY_DN48515_c0_g3_i1 and TRINITY_DN1328_c0_g1_i1 were predicted to positively and negatively correlated with CHI and CHS, respectively[21]. In potato tubers, multiple lncRNAs such as XLOC_046144, XLOC_038153, XLOC_047085, XLOC_060098, XLOC_070549 were also predicted to influence the anthocyanin content by regulating the expression of enzyme genes displayed in Fig. 6 (Table 1). Besides, CCoAOMT and AnAT, which have been reported to be involved in methylation and acylation during anthocyanin metabolism respectively[63, 64], were also predicted to be regulated by lncRNAs like XLOC060098 and XLOC038153 in potato tubers (Fig. 8).

Additionally, a number of transcription factors that mediate anthocyanin metabolism are often regulated by lncRNAs as well[30]. In this study, the transcription factor genes such as MYB, SPL, ARF, NAC and ERF were identified to be targeted by different lncRNAs through the analysis of co-expression network (Fig. 7 and Fig. 8). MYB transcription factor usually interacts with bHLH and WD40 to form the ternary WBM complex which subsequently regulates the expression of related enzyme genes like F3'5'H, DFR and UFGT, thereby participating in anthocyanin synthesis[15, 65-67]. In sea buckthorn fruit, LNC2 was identified as the endogenous target mimic (eTM) of miR828a, thereby inducing MYB114 expression[31]. It has been found in rice that lncRNA TWISTED LEAF could affect OsMYB60 expression and leaf phenotype[5]. Similarly, in potato tubers, the changes of MYB expression under the guidance of the endogenous lncRNAs, such as XLOC_017372, XLOC_029213, and XLOC_038153, might influence the anthocyanin accumulation and tuber colors (Additional file 8).

SPL, which was considered as the target of miR156, has been reported to negatively regulate anthocyanin biosynthesis under the action of lncRNAs in different plant species [11, 16, 68]. In banana, the lncRNA could modulate the expression of SPL2-like gene[68]. LNC1 was predicted to be a decoy for miR156a and reduce SPL9 expression, thus increasing anthocyanin content in sea buckhorn fruit[31]. In apple, MLNC3.2 and MLNC4.6 were recognized as eTMs for miRNA156a and facilitated the expression of the SPL2-like and SPL33 by preventing cleavage of these two SPL genes[30]. They also proposed that the regulatory mechanism by which lncRNAs act as eTMs for miR156 may be conserved and universal in plants, and these lncRNAs may influence various aspects of plant development [30]. In 'Xisen-8', the down-regulated XLOC_100954 reduced the expression of SPL and promoted anthocyanin production (Fig. 7, Fig. 8). Therefore, XLOC_100954 might function as the eTM for miR156 in potato and prevent breakage of SPL by miR156. However, the specific regulation mechanism of lncRNAs on SPL needs further validation in potato tubers.

ARF was regarded as a key regulator of auxin responsive signaling pathway and plays a crucial role in anthocyanin biosynthesis in plants[69]. MdARF13 was proved to suppress MdDFR expression to decrease anthocyanin accumulation[70]. Recently, 6 PeARFs and 5 lncRNAs were predicted to be targeted by pei-miR160a in populus[42] . But the regulatory relationship between lncRNAs and ARFs has not been clarified. In potato tubers, ARF8 was identified as the co-expressed target of XLOC_100954 (Fig. 7). The low expression of XLOC_100954 affected ARF8 expression and thus increased the anthocyanin content in purple fleshed potato tubers (Fig. 8). Other anthocyanin-related transcription factor genes like NAC2 and ERF4 were also predicted to be regulated by XLOC_060098 and XLOC_017372, respectively. Although several lncRNAs were predicted to regulate the expression of NAC and EFR in other plants[71], the specific action mechanism remains unclear.

The regulatory relationship between specific lncRNAs and mRNAs is quite complicated due to the involvement of many other factors like miRNAs and other genes. In this study, DE lncRNAs and their potential target genes were identified and predicted by comparing transcriptome data from purple and yellow fleshed potato tubers. The target genes co-expressed with XLOC_060098 and XLOC_017372 were the most, indicating that XLOC_060098 and XLOC_017372 may be involved in the most biological processes and play crucial roles during anthocyanin metabolism. These results provide a basis for further investigation of the function of potato lncRNAs in anthocyanin synthesis. The analysis of the regulation mechanism of anthocyanin biosynthesis in potato tubers enriches the existing function mechanism of anthocyanin biosynthesis in different plant organs especially the underground organs.

Conclusions

In this study, fresh tubers of purple fleshed potato (‘Xisen-8’) and yellow fleshed potato (‘Jin-16’) were used as experimental materials to excavate lncRNAs related to potato anthocyanin biosynthesis through transcriptome sequencing. The cis- and trans-acting target genes of the DE lncRNAs were predicted via bioinformatics analyses. Totally, 4376 lncRNAs were identified, among which 1421 were significantly differentially expressed with 735 up-regulated and 686 down-regulated lncRNAs in ‘Xisen-8’. Chromosomal mapping analysis showed that these DE lncRNAs were unevenly distributed on potato chromosomes: the number of lncRNAs on chromosome 1 was the highest, and the number of lncRNAs on chromosome 2 was the lowest. The target genes of DE lncRNAs could be clustered into lots of GO types and enriched into multiple metabolic pathways, thus refecting the functional diversity of their corresponding lncRNAs. The co-expression network and the expression patterns of DE lncRNAs and their target genes related to anthocyanin biosynthesis in potato tubers were analyzed to reveal their potential regulatory relationship. Results showed that most anthocyanin-associated enzyme genes and transcription factor genes were predicted to be regulated by lncRNAs such as XLOC_060098 and XLOC_017372, indicating that these lncRNAs might play crucial roles in anthocyanin metabolism. The results of this study lay the foundation for further elucidation of regulatory mechanism of lncRNAs in anthocyanin biosynthesis and provide new insights for quality improvement strategies of potato and other crops.

Materials And Methods

Plant materials and growth conditions

Potato cultivars ‘Jin-16’ and ‘Xisen-8’ were used as materials in this study. The tissue culture seedlings of cultivar ‘Jin-16’ were preserved in the College of Agriculture, Shanxi Agricultural University. The tissue culture plantlets of the other cultivar ‘Xisen-8’ were kindly gifted by Leling Xisen Potato Industry Co. Ltd. (Leling, Shandong, China). The 4-week-old tissue culture plantlets were transferred into the clay pots with soil and grown in the greenhouse under 14 h light/10 h dark regime at 22 ± 1 ℃. Three months later, the fresh tubers were harvested for anthocyanin content determination and transcriptome sequencing.

Determination of anthocyanin contents in potato tubers

Three potato tubers with similar sizes (5-6 cm in length) were selected from ‘Jin-16’ and ‘Xisen-8’, respectively. Anthocyanin was extracted according to the method used by Wang et al.[70]. The potato flesh from each tuber was ground into powder and then exposed to HCl-methanol solution (1:99 by volume) at 4 ℃ for 6-8 h under darkness until the tissues were completely decolorized. After centrifuging at 12000 rpm for 10 min, the absorbance values of supernatants were determined at 530 nm using a UV-2450 spectrophotometer (Shimadzu, Kyoto, Japan). Each sample had three replicates to ensure the results reliable.

Total RNA extraction, library preparation and lncRNA sequencing

Total RNA was isolated from 0.5 g potato flesh samples using the Quick RNA Isolation Kit (Huayueyang, Beijing, China). Subsequently, electrophoresis was performed with 1% agarose gel to monitor the presence of RNA degradation and DNA contamination. The purity and concentration of RNA samples were measured using Nanodrop 1000 spectrophotometer (Thermo Scientific). After integrity testing by Agilent 2100 BioAnalyzer (Agilent Technologies), the total RNA samples were used for the construction of lncRNA libraries and validation of deep sequencing results.

Strand-specific library was constructed using the deoxy-UTP (dUTP) strand-marking method[72]. Three biological replicates were set for each potato cultivar. Therefore, 6 strand-specific libraries were constructed, including Jin-16_1, Jin-16_2, Jin-

16_3, Xisen-8_1, Xisen-8_2 and Xisen-8_3. After quantitative and qualitative determination of all libraries, lncRNA deep sequencing were carried out on an Illumina novaseq 6000 platform by Novogene Bioinformatics Technology Co. Ltd. (Beijing, China). The obtained raw reads were processed by getting rid of the low quality reads, the reads with sequencing adapters and poly-N sequences. The clean reads were acquired and aligned to a potato reference genome (S. tuberosum assembly 4.03) using Hisat2 software[32]. The mapped reads were spliced and assembled into transcripts using Stringtie[73] and Cuffmerge software.

Identification of differentially expressed lncRNAs

The identification of lncRNAs was based on their structural characteristics and non-coding functional characteristics. The transcripts with less than two exons and transcripts less than 200 nt in length were removed. CPC2, PFAM and CNCI tools were employed to estimate the coding potential of the spliced transcripts. Based on the comprehensive analysis results of the three softwares, the transcripts without coding potential were regarded as candidate lncRNAs. The differential expression analysis of identified lncRNAs was performed using DESeq [18] and edgeR [19]. The lncRNAs with q value less than 0.05 were designated as differentially expressed lncRNAs.

Expression analysis and chromosomal localization of DE lncRNA

General expression patterns of DE lncRNAs in ‘Jin-16’ and ‘Xisen-8’ were displayed in a Hcluster heatmap using FPKM (fragments per kilobase of transcript sequence per millions base pairs sequenced) values of DE lncRNAs. The drawing of venn diagram also used FPKM values to show the DE lncRNA numbers in ‘Jin-16’ and ‘Xisen-8’ tubers. The numbers of up-regulated and down-regulated DE lncRNAs were expressed in volcano plot utilizing the $\log_2$ (fold change) values. These three drawing tools are available on the Novomagic online platform for data analysis (https://magic.novogene.com). Distribution of DE lncRNAs on potato chromosomes was implemented by SvgWrite package in Python based on the chromosome location information and the $\log_{10}$ (FPKM) values of DE lncRNAs.

Prediction and functional enrichment of lncRNA target genes

LncRNAs regulate target genes through a variety of mechanisms, and the most common ways of acting on downstream target genes are co-location and co-expression regulation. The coding genes which located in 100 kb upstream and downstream of lncRNAs were searched and classified into co-located target genes. The co-expressed target genes were predicted by correlation analysis between lncRNAs and mRNAs with the correlation coefficient above 0.95.

To predict the main functions of lncRNAs, the functional enrichment analyses were performed on the co-located and co-expressed target genes, respectively. The sequences of the target genes were blasted to the Gene Ontology (GO) database to obtain their GO annotation information. Then GO enrichment analysis was implemented by GOseq software[74]. The GO terms with p-adjust < 0.05 were considered significantly enriched.

The gene sequences were also mapped to the KEGG database to detect the main pathways in which they participate. KOBAS (2.0) software[75] was used to perform the pathway enrichment analysis of target genes of DE lncRNAs. The pathways with p value < 0.05 were regarded as the significant enriched pathways.

Construction of co-expression network between DE lncRNAs and their co-expressed target genes

According to the predicted co-expression relationship between DE lncRNAs and their co-expressed mRNAs related to anthocyanin biosynthesis, lncRNA-mRNA pairs with correlation coefficient greater than 0.98 were screened out. The lncRNA-mRNA interaction network was graphically presented using Cytoscape software[76].

Validation of the RNA-seq results by quantitative real-time PCR

RNA samples were extracted from the tuber flesh of 'Jin-16' and 'Xisen-8' as described above. Quantitative real-time PCR (qRT-PCR) was performed with the TB GreenTM Premix Ex TaqTM (Tli RNase H Plus) (Takara, Dalian, China) on CFX96 PCR System (Bio-Rad, USA). The specificity of these primers which designed by Primer-Blast in NCBI website (https://www.ncbi.nlm.nih.gov/tools/primer-blast/) was tested using dissociation curve analysis. The 10 μl reaction volume samples, containing 5 μL TB Green, 1 μL diluted template, 0.4 μL 10 μM solution of each primer and 3.2 μL ddH_2O, were used for PCR using the following cycling program: 95 ℃ for 3 min, followed by 40 cycles of 95 ℃ for 10 s, 60 ℃ for 30 s, and 72 ℃ for 20 s. The relative expression of selected lncRNAs and genes was calibrated against the reference gene *EF1α* using the method of $2^{-\Delta\Delta Ct}$[1]. For each sample, three biological repeats and two experimental replicates were performed to make sure the results reliable.

Abbreviations

AT: Acyltransferase; ANS: Anthocyanidin synthase; ARF: Auxin response factor; AnAT: Anthocyanin acyltransferase; CHI: Chalcone isomerase; CHS: Chalcone synthase; COOLAIR: Cold induced antisense intragenic RNA; COLDAIR: Cold assisted intronic noncoding RNA; CCoAOMT5: Caffeoyl-CoA O-methyltransferase 5; C3H: P-coumarate 3-hydroxylase; DE: Differentially expressed; DFR: Dihydroflavonol 4-reductase; DHK: Dihydrokaempferol; DHQ: Dihydroquercetin; DHM: Dihydromynicetin; EIN3: Ethylene insensitive; ERF: Ethylene-responsive transcription factor; ERF7: Ethylene-responsive transcription factor 7; F3H: Flavanone 3-hydroxylase; F3'H: Flavanone 3-hydroxylase; F3'5'H: Flavonoid 3',5'-hydroxylase; GO: Gene ontology; GST4: glutathione transferase 4; HQT: Hydroxycinnamoyl-CoA: quinate hydroxycinamoyl transferase; KEGG: Kyoto encyclopedia of genes and genomes; lncRNA: Long non-coding RNA; LDMAR: Long-day-specific male-fertility associated RNA; MYC4: Myelocytomatosis4; PIF3: Phytochrome-interacting factor 3; PAL: Phenylalanine ammonia lyase; RdDM: RNA-directed DNA methylation; SPA: Supressor of phytochrome A-105; SPL: Squamosa promoter binding proteinlike; TasiRNA: Trans-acting small interfering; TAS3: Trans-acting siRNA3; UFGT: UDP-glucose flavonoid 3-o-glycosyltransferase; UGT: UDP-glucose:glucosyltransferase.

Declarations

Ethics approval and consent to participate

Not applicable.

Consent for publication

Not applicable.

Availability of data and materials

The datasets generated and/or analyzed for this work were deposited in the NCBI Sequence Read Archive under the Bioproject accession PRJNA729884, available from https://dataview.ncbi.nlm.nih.gov/object/PRJNA729884?reviewer=ntlkjmravag9c9ousg57ps9k86.

Competing interests

The authors declare that they have no competing interests.

Funding

This work was supported by the National Natural Science Foundation for Young Scientists of China (31900450), Science and Technology Innovation Project of Higher Education of Shanxi Province (2019L0388), and Science and Technology Innovation Fund of Shanxi Agricultural University (2018YJ28), Science and Technology Innovation Fund of Shanxi Agricultural University (2018yz001), Applied Basic Research Program of Shanxi Academy of Agricultural Sciences (YGC2019FZ4) and Shanxi Provincial Leading Talents in Emerging Industries Project.

Authors' contributions

XYJ, RMT and QY conceived and designed the original research plans; RMT, WYW collected the samples; WYW and JZ determined the content of anthocyanin; RMT, HTD, LHH, HEX and ZXW analyzed the data and performed the bioinformatics analysis; CLZ, HTD and LHH performed qRT-PCR analysis; RMT and HTD constructed and analyzed the co-expression network; XYJ and RMT wrote the manuscript; QY edited the English language in this manuscript. All authors have read and approved the manuscript.

Acknowledgments

The authors are grateful to Yonghong Duan (College of Agriculture in Shanxi Agricultural University) and Leling Xisen Potato Industry Co. Ltd. (Leling, Shandong, China) for providing the experimental materials 'Jin-16' and 'Xisen-8', respectively.

References

1. Tang R, Gupta SK, Niu S, Li XQ, Yang Q, Chen G, Zhu W, Haroon M. Transcriptome analysis of heat stress response genes in potato leaves. Mol Biol Rep 2020, 47(6):4311-4321.
2. Bajji M, Hamdi MM, Gastiny F, Rojas-Beltran JA, Jardin PD. Catalase inhibition accelerates dormancy release and sprouting in potato (Solanum tuberosum L.) tubers. Biotechnol Agron Soc Environ 2007, 11(2):121-131.
3. Liu Y, Lin-Wang K, Espley RV, Wang L, Li Y, Liu Z, Zhou P, Zeng L, Zhang X, Zhang J, Allan AC. StMYB44 negatively regulates anthocyanin biosynthesis at high temperatures in tuber flesh of potato. J Exp Bot 2019, 70(15):3809-3824.
4. Li L, He Y, Ge H, Liu Y, Chen H. Functional characterization of SmMYB86, a negative regulator of anthocyanin biosynthesis in eggplant (Solanum melongena L.). Plant Sci 2021, 302(1):110696.
5. Liu F, Yang Y, Gao J, Ma C, Bi Y. A comparative transcriptome analysis of a wild purple potato and its red mutant provides insight into the mechanism of anthocyanin transformation. PLoS One 2018, 13(1):e0191406.
6. He W, Zeng M, Chen J, Jiao Y, Niu F, Tao G, Zhang S, Qin F, He Z. Identification and quantitation of anthocyanins in purple-fleshed sweet potatoes cultivated in China by UPLC-PDA and UPLC-QTOF-MS/MS. J Agric Food Chem 2016, 64(1):171-177.
7. Sehitoglu MH, Farooqi AA, Qureshi MZ, Butt G, Aras A. Anthocyanins: targeting of signaling networks in cancer cells. Asian Pac J Cancer Prev 2014, 15(5):2379-2381.
8. Sarkar B, Kumar D, Sasmal D, Mukhopadhyay K. Antioxidant and DNA damage protective properties of anthocyanin-rich extracts from Hibiscus and Ocimum: a comparative study. Nat Prod Res 2014, 28(17):1393-1398.

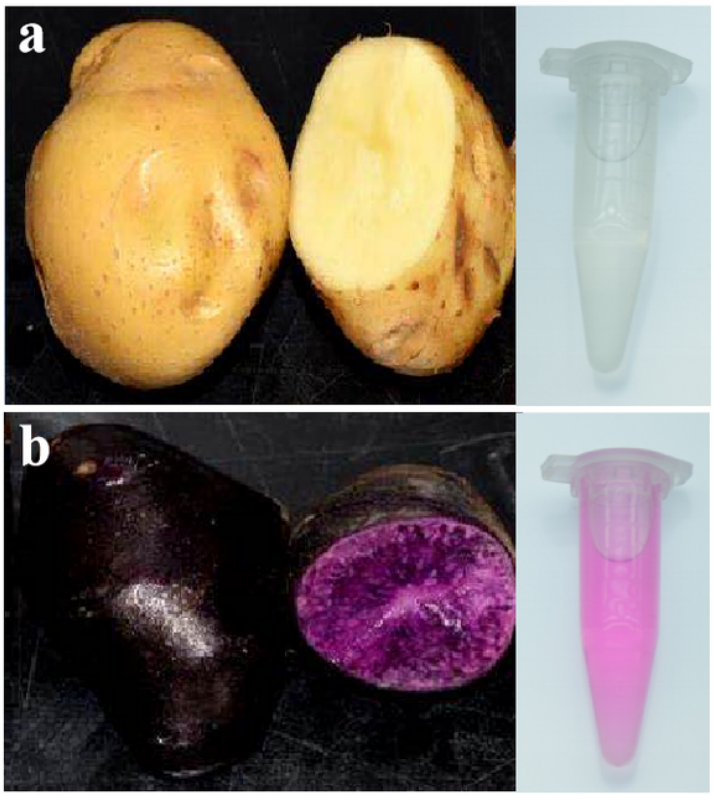
9. Xu P, Zawora C, Li Y, Wu J, Liu L, Liu Z, Cai R, Lian H. Transcriptome sequencing reveals role of light in promoting anthocyanin accumulation of strawberry fruit. J Plant Growth Regul 2018, 86(1):121-132.
10. Castañeda-Ovando A, Pacheco-Hernández MdL, Páez-Hernández ME, Rodríguez JA, Galán-Vidal CA. Chemical studies of anthocyanins: A review. Food Chem 2009, 113(4):859-871.
11. He L, Tang R, Shi X, Wang W, Cao Q, Liu X, Wang T, Sun Y, Zhang H, Li R, Jia X. Uncovering anthocyanin biosynthesis related microRNAs and their target genes by small RNA and degradome sequencing in tuberous roots of sweetpotato. BMC Plant Biol 2019, 19(1):232.
12. Castellarin SD, Pfeiffer A, Sivilotti P, Degan M, Peterlunger E, G DIG. Transcriptional regulation of anthocyanin biosynthesis in ripening fruits of grapevine under seasonal water deficit. Plant Cell Environ 2007, 30(11):1381-1399.
13. Maier A, Hoecker U. COP1/SPA ubiquitin ligase complexes repress anthocyanin accumulation under low light and high light conditions. Plant Signal Behav 2015, 10(1):e970440.
14. Hu B, Zhao J, Lai B, Qin Y, Wang H, Hu G. LcGST4 is an anthocyanin-related glutathione S-transferase gene in Litchi chinensis Sonn. Plant Cell Rep 2016, 35(4):831-843.
15. Xu W, Dubos C, Lepiniec L. Transcriptional control of flavonoid biosynthesis by MYB-bHLH-WDR complexes. Trends Plant Sci 2015, 20(3):176-185.
16. Gou JY, Felippes FF, Liu CJ, Weigel D, Wang JW. Negative regulation of anthocyanin biosynthesis in *Arabidopsis* by a miR156-targeted SPL transcription factor. Plant Cell 2011, 23(4):1512-1522.
17. Shin J, Park E, Choi G. PIF3 regulates anthocyanin biosynthesis in an HY5-dependent manner with both factors directly binding anthocyanin biosynthetic gene promoters in *Arabidopsis*. Plant J 2007, 49(6):981-994.
18. Hung T, Chang HY. Long noncoding RNA in genome regulation: prospects and mechanisms. RNA Biol 2010, 7(5):582-585.
19. Marcel TR, Dinger ME, Mattick JS. Long non-coding RNAs: insights into functions. Nat Rev Genet 2009, 10(3):155-159.
20. Li Z, Rana TM. Molecular mechanisms of RNA-triggered gene silencing machineries. Acc Chem Res 2012, 45(7):1122-1131.
21. Lin Y, Jiang L, Chen Q, Li Y, Zhang Y, Luo Y, Zhang Y, Sun B, Wang X, Tang H. Comparative transcriptome profiling analysis of red- and white-fleshed strawberry (Fragariai 1/2 ananassa) provides new insight into the regulation of the anthocyanin pathway. Plant Cell Physiol 2018, 59(9):1844-1859.
22. Wang J, Meng X, Dobrovolskaya OB, Orlov YL, Chen M. Non-coding RNAs and their roles in stress response in plants. Genomics Proteomics Bioinformatics 2017, 15(5):301-312.
23. Urquiaga MCO, Thiebaut F, Hemerly AS, Ferreira PCG. From trash to luxury: the potential role of plant lncRNA in DNA methylation during abiotic stress. Front Plant Sci 2020, 11:603246.
24. Xiao L, Shang XH, Cao S, Xie XY, Zeng WD, Lu LY, Chen SB, Yan HB. Comparative physiology and transcriptome analysis allows for identification of lncRNAs imparting tolerance to drought stress in autotetraploid cassava. BMC Genomics 2019, 20(1):514.
25. Hou X, Du Y, Liu X, Zhang H, Liu Y, Yan N, Zhang Z. Genome-wide analysis of long non-coding RNAs in potato and their potential role in tuber sprouting process. Int J Mol Sci 2017, 19(1):101.
26. Csorba T, Questa JI, Sun Q, Dean C. Antisense COOLAIR mediates the coordinated switching of chromatin states at FLC during vernalization. Proc Natl Acad Sci U S A 2014, 111(45):16160-16165.
27. Yamaguchi A, Abe M. Regulation of reproductive development by non-coding RNA in *Arabidopsis*: to flower or not to flower. J Plant Res 2012, 125(6):693-704.
28. Ding J, Shen J, Mao H, Xie W, Li X, Zhang Q. RNA-directed DNA methylation is involved in regulating photoperiod-sensitive male sterility in rice. Mol Plant 2012, 5(6):1210-1216.

29. Fukuda M, Nishida S, Kakei Y, Shimada Y, Fujiwara T. Genome-wide analysis of long intergenic noncoding RNAs responding to low-nutrient conditions in *Arabidopsis thaliana*: possible involvement of trans-ccting siRNA3 in response to Low Nitrogen. Plant Cell Physiol 2019, 60(9):1961-1973.
30. Yang T, Ma H, Zhang J, Wu T, Song T, Tian J, Yao Y. Systematic identification of long noncoding RNAs expressed during light-induced anthocyanin accumulation in apple fruit. Plant J 2019, 100(3):572-590.
31. Zhang G, Chen D, Zhang T, Duan A, Zhang J, He C. Transcriptomic and functional analyses unveil the role of long non-coding RNAs in anthocyanin biosynthesis during sea buckthorn fruit ripening. DNA Res 2018, 25(5):465-476.
32. Kim D, Langmead B, Salzberg SL. HISAT: a fast spliced aligner with low memory requirements. Nat Methods 2015, 12(4):357-360.
33. Wu L, Liu S, Qi H, Cai H, Xu M. Research progress on plant long non-coding RNA. Plants 2020, 9(4):408.
34. Zhang YC, Chen YQ. Long noncoding RNAs: new regulators in plant development. Biochem Biophys Res Commun 2013, 436(2):111-114.
35. Tang W, Zheng Y, Dong J, Yu J, Yue J, Liu F, Guo X, Huang S, Wisniewski M, Sun J, Niu X, Ding J, Liu J, Fei Z, Liu Y. Comprehensive transcriptome profiling reveals long noncoding RNA expression and alternative splicing egulation during fruit development and ripening in Kiwifruit (*Actinidia chinensis*). Front Plant Sci 2016, 29(3):335.
36. Li R, Fu D, Zhu B, Luo Y, Zhu H. CRISPR/Cas9-mediated mutagenesis of lncRNA1459 alters tomato fruit ripening. Plant J 2018, 94(3):513-524.
37. Sun X, Zheng H, Li J, Liu L, Zhang X, Sui N. Comparative transcriptome analysis reveals new lncRNAs responding to salt stress in sweet sorghum. Front Bioeng Biotechnol 2020, 15(4):331.
38. Kindgren P, Ard R, Ivanov M, Marquardt S. Transcriptional read-through of the long non-coding RNA SVALKA governs plant cold acclimation. Nat Commun 2019, 10(1):5141.
39. Zhao X, Li J, Lian B, Gu H, Li Y, Qi Y. Global identification of *Arabidopsis* lncRNAs reveals the regulation of MAF4 by a natural antisense RNA. Nat Commun 2018, 9(1):5056.
40. Zhang G, Duan A, Zhang J, He C. Genome-wide analysis of long non-coding RNAs at the mature stage of sea buckthorn (*Hippophae rhamnoides* Linn) fruit. Gene 2017, 596:130-136.
41. Fatica A, Bozzoni I. Long non-coding RNAs: new players in cell differentiation and development. Nat Rev Genet 2014, 15(1):7-21.
42. Liu S, Wu L, Qi H, Xu M. LncRNA/circRNA–miRNA–mRNA networks regulate the development of root and shoot meristems of Populus. Ind Crop Prod 2019, 133(7):333-347.
43. Zoratti L, Karppinen K, Luengo Escobar A, Haggman H, Jaakola L. Light-controlled flavonoid biosynthesis in fruits. Front Plant Sci 2014, 5:534.
44. He Y, Zhang X, Li L, Sun Z, Li J, Chen X, Hong G. SPX4 interacts with both PHR1 and PAP1 to regulate critical steps in phosphorus-status-dependent anthocyanin biosynthesis. New Phytol 2021, 230(1):205-217.
45. Tanaka Y, Brugliera F, Kalc G, Senior M, Dyson B, Nakamura N, Katsumoto Y, Chandler S. Flower color modification by engineering of the flavonoid biosynthetic pathway: practical perspectives. Biosci Biotechnol Biochem 2010, 74(9):1760-1769.
46. Lu C, Li Y, Cui Y, Ren J, Qi F, Qu J, Huang H, Dai S. Isolation and functional analysis of genes involved in polyacylated anthocyanin biosynthesis in Blue Senecio cruentus. Front Plant Sci 2021, 12:640746.
47. Roncaglia P, Martone ME, Hill DP, Berardini TZ. The gene ontology (GO) cellular component ontology: integration with SAO (Subcellular Anatomy Ontology) and other recent developments. J Biomed Semant 2013, 4(1):20.
48. Zhu B, Yang Y, Li R, Fu D, Wen L, Luo Y, Zhu H. RNA sequencing and functional analysis implicate the regulatory role of long non-coding RNAs in tomato fruit ripening. J Exp Bot 2015, 66(15):4483-4495.

49. Akira Iwase KM, Masaru Ohme-Takagi. Manipulation of plant metabolic pathways by transcription factors. Plant Biotechnology 2009, 26(1):29–38.
50. Wang S, Pan D, Lv X, Song X, Qiu Z, Huang C, Huang R, Chen W. Proteomic approach reveals that starch degradation contributes to anthocyanin accumulation in tuberous root of purple sweet potato. J Proteomics 2016, 143:298-305.
51. Xie Y, Tan H, Ma Z, Huang J. DELLA proteins promote anthocyanin biosynthesis via sequestering MYBL2 and JAZ suppressors of the MYB/bHLH/WD40 Complex in *Arabidopsis thaliana*. Mol Plant 2016, 9(5):711-721.
52. Tian J, Chen MC, Zhang J, Li KT, Song TT, Zhang X, Yao YC. Characteristics of dihydroflavonol 4-reductase gene promoters from different leaf colored Malus crabapple cultivars. Hortic Res 2017, 4:17070.
53. Ban Y, Kondo S, Ubi BE, Honda C, Bessho H, Moriguchi T. UDP-sugar biosynthetic pathway: contribution to cyanidin 3-galactoside biosynthesis in apple skin. Planta 2009, 230(5):871-881.
54. Espley RV, Hellens RP, Putterill J, Stevenson DE, Kutty-Amma S, Allan AC. Red colouration in apple fruit is due to the activity of the MYB transcription factor, MdMYB10. Plant J 2007, 49(3):414-427.
55. Ju Z, Yuan Y, Liu C, Wang Y, Tian X. Dihydroflavonol reductase activity and anthocyanin accumulation in 'Delicious', 'Golden Delicious' and 'Indo' apples. Sci Hortic 1997, 70(1):31-43.
56. Lister CE, Lancaster JE, Walker JR. Developmental changes in enzymes biosynthesis in the skins of red and of flavonoid green apple cultivars. J Food Agric Environ 1996, 71(1):313-320.
57. Saito K, Yonekura-Sakakibara K, Nakabayashi R, Higashi Y, Yamazaki M, Tohge T, Fernie AR. The flavonoid biosynthetic pathway in *Arabidopsis*: structural and genetic diversity. Plant Physiol Biochem 2013, 72:21-34.
58. Davies KM, Schwinn KE, Deroles SC, Manson DG, Lewis DH, Bloor SJ, Bradley JM. Enhancing anthocyanin production by altering competition for substrate between flavonol synthase and dihydroflavonol 4-reductase. Euphytica 2003, 7(131):259-268.
59. Whang SS, Um WS, Song I-J, Lim PO, Choi K, Park K-W, Kang K-W, Choi MS, Koo JC. Molecular Analysis of Anthocyanin Biosynthetic Genes and Control of Flower Coloration by Flavonoid 3',5'-Hydroxylase (F3'5'H) in Dendrobium moniliforme. Journal of Plant Biology 2011, 54(3):209-218.
60. Chen C, Zhou G, Chen J, Liu X, Lu X, Chen H, Tian Y. Integrated metabolome and transcriptome analysis unveils novel pathway involved in the formation of yellow peel in cucumber. Int J Mol Sci 2021, 22(3):1494.
61. Sonnante G, D'Amore R, Blanco E, Pierri CL, De Palma M, Luo J, Tucci M, Martin C. Novel hydroxycinnamoyl-coenzyme A quinate transferase genes from artichoke are involved in the synthesis of chlorogenic acid. Plant Physiol 2010, 153(3):1224-1238.
62. Lee W, Yeo Y, Oh S, Cho KS, Park YE, Park SK, Lee SM, Cho HS, Park SY. Compositional analyses of diverse phytochemicals and polar metabolites from different-colored potato (Solanum tubersum L.) tubers. Food Sci Biotechnol 2017, 26(5):1379-1389.
63. Rinaldo AR, Cavallini E, Jia Y, Moss SM, McDavid DA, Hooper LC, Robinson SP, Tornielli GB, Zenoni S, Ford CM, Boss PK, Walker AR. A grapevine anthocyanin acyltransferase, transcriptionally regulated by VvMYBA, can produce most acylated anthocyanins present in grape skins. Plant Physiol 2015, 169(3):1897-1916.
64. Hugueney P, Provenzano S, Verries C, Ferrandino A, Meudec E, Batelli G, Merdinoglu D, Cheynier V, Schubert A, Ageorges A. A novel cation-dependent O-methyltransferase involved in anthocyanin methylation in grapevine. Plant Physiol 2009, 150(4):2057-2070.
65. Feller A, Machemer K, Braun EL, Grotewold E. Evolutionary and comparative analysis of MYB and bHLH plant transcription factors. Plant J 2011, 66(1):94-116.
66. Xie XB, Li S, Zhang RF, Zhao J, Chen YC, Zhao Q, Yao YX, You CX, Zhang XS, Hao YJ. The bHLH transcription factor MdbHLH3 promotes anthocyanin accumulation and fruit colouration in response to low temperature in apples. Plant Cell Environ 2012, 35(11):1884-1897.

67. Patra B, Pattanaik S, Yuan L. Ubiquitin protein ligase 3 mediates the proteasomal degradation of GLABROUS 3 and ENHANCER OF GLABROUS 3, regulators of trichome development and flavonoid biosynthesis in *Arabidopsis*. Plant J 2013, 74(3):435-447.
68. Muthusamy M, Uma S, Backiyarani S, Saraswathi MS. Genome-wide screening for novel, drought stress-responsive long non-coding RNAs in drought-stressed leaf transcriptome of drought-tolerant and -susceptible banana (Musa spp) cultivars using Illumina high-throughput sequencing. Plant Biotechnology Reports 2015, 9(5):279-286.
69. Tiwari SB, Wang XJ, Hagen G, Guilfoyle TJ. AUX/IAA proteins are active repressors, and their stability and activity are modulated by auxin. Plant Cell 2001, 13(12):2809-2822.
70. Wang Y, Gao L, Li J, Zhu B, Zhu H, Luo Y, Wang Q, Zuo J. Analysis of long-non-coding RNAs associated with ethylene in tomato. Gene 2018, 674:151-160.
71. Xue L, Sun M, Wu Z, Yu L, Yu Q, Tang Y, Jiang F. LncRNA regulates tomato fruit cracking by coordinating gene expression via a hormone-redox-cell wall network. BMC Plant Biol 2020, 20(1):162.
72. Levin JZ, Yassour M, Adiconis X, Nusbaum C, Thompson DA, Friedman N, Gnirke A, Regev A. Comprehensive comparative analysis of strand-specific RNA sequencing methods. Nat Methods 2010, 7(9):709-715.
73. Pertea M, Kim D, Pertea GM, Leek JT, Salzberg SL. Transcript-level expression analysis of RNA-seq experiments with HISAT, StringTie and Ballgown. Nat Protoc 2016, 11(9):1650-1667.
74. Young MD, Wakefield MJ, Smyth GK, Oshlack A. Gene ontology analysis for RNA-seq: accounting forselection bias. Genome Biology 2010, 11(2):R14.
75. Mao X, Cai T, Olyarchuk JG, Wei L. Automated genome annotation and pathway identification using the KEGG Orthology (KO) as a controlled vocabulary. Bioinformatics 2005, 21(19):3787-3793.
76. Shannon P, Markiel A, Ozier O, Baliga NS, Wang JT, Ramage D, Amin N, Schwikowski B, Ideker T. Cytoscape: a software environment for integrated models of biomolecular interaction networks. Genome Res 2003, 13(11):2498-2504.

Figures



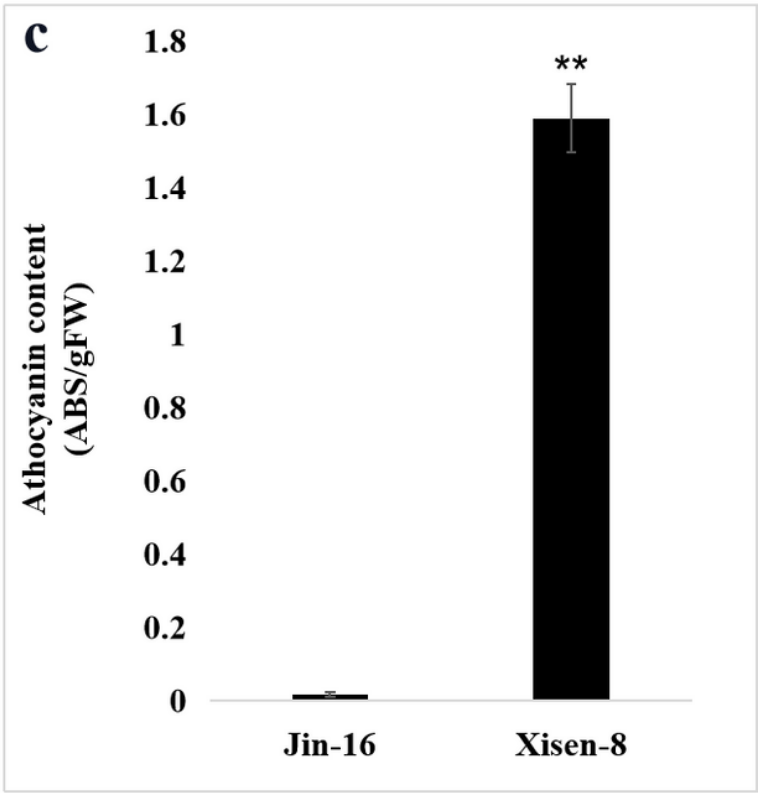


Figure 1

The phenotypes and anthocyanin contents of ‘Jin-16’ and ‘Xisen-8’ tubers. (a, b) The appearance, cross profile and anthocyanin extraction of ‘Jin-16’ and ‘Xisen-8’ tubers; (c) Relative anthocyanin content calculated as absorbance at 530 nm/fresh weight (g).

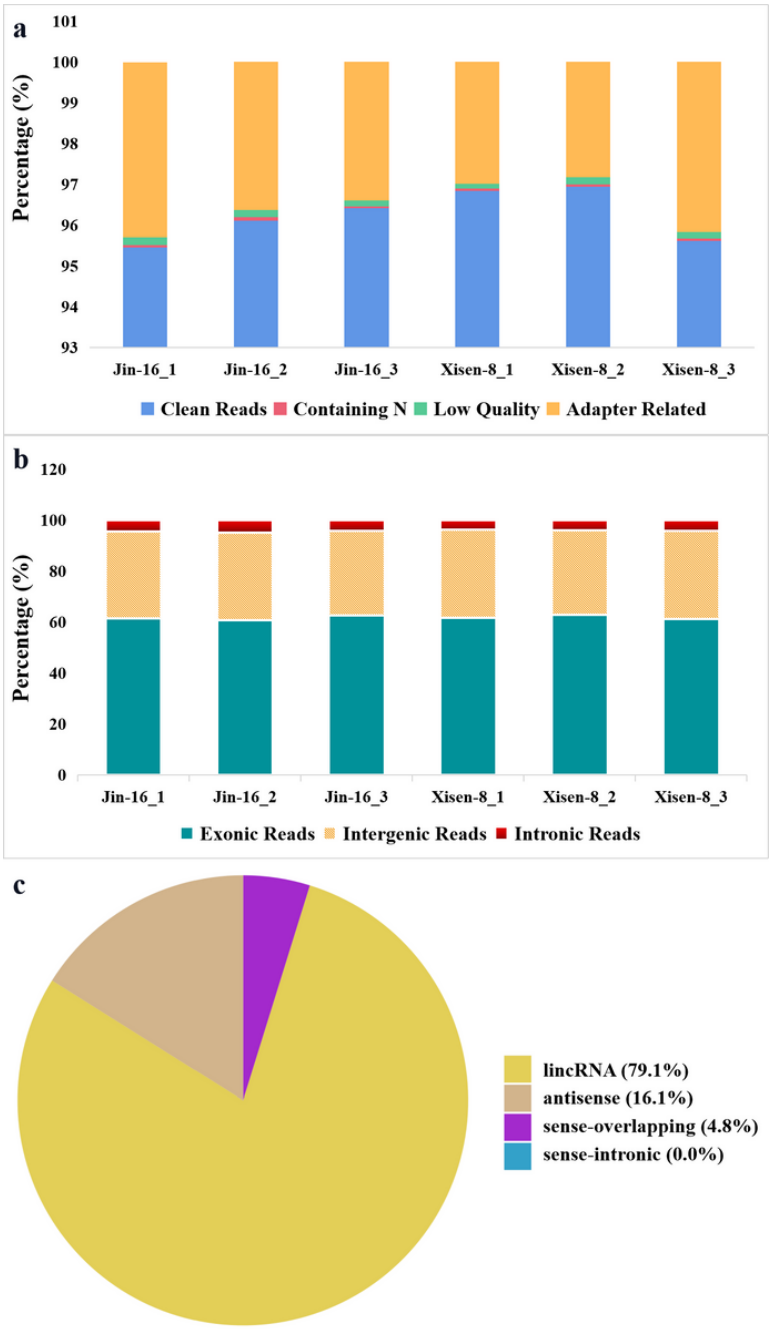


Figure 2

Features of reads and lncRNAs obtained from transcriptome sequencing. (a) Classification of raw reads; (b) Percentage of reads mapped to genome regions; (c) Classification of 4376 identified lncRNA. LincRNA: long intergenic lncRNAs; antisense: the lncRNAs located in the coding regions at the antisense direction; sense overlapping: the lncRNAs located in the coding regions at the sense direction and overlapped with one or more exons of a gene; sense intronic: the lncRNAs located in the intronic region of a gene at the sense direction.

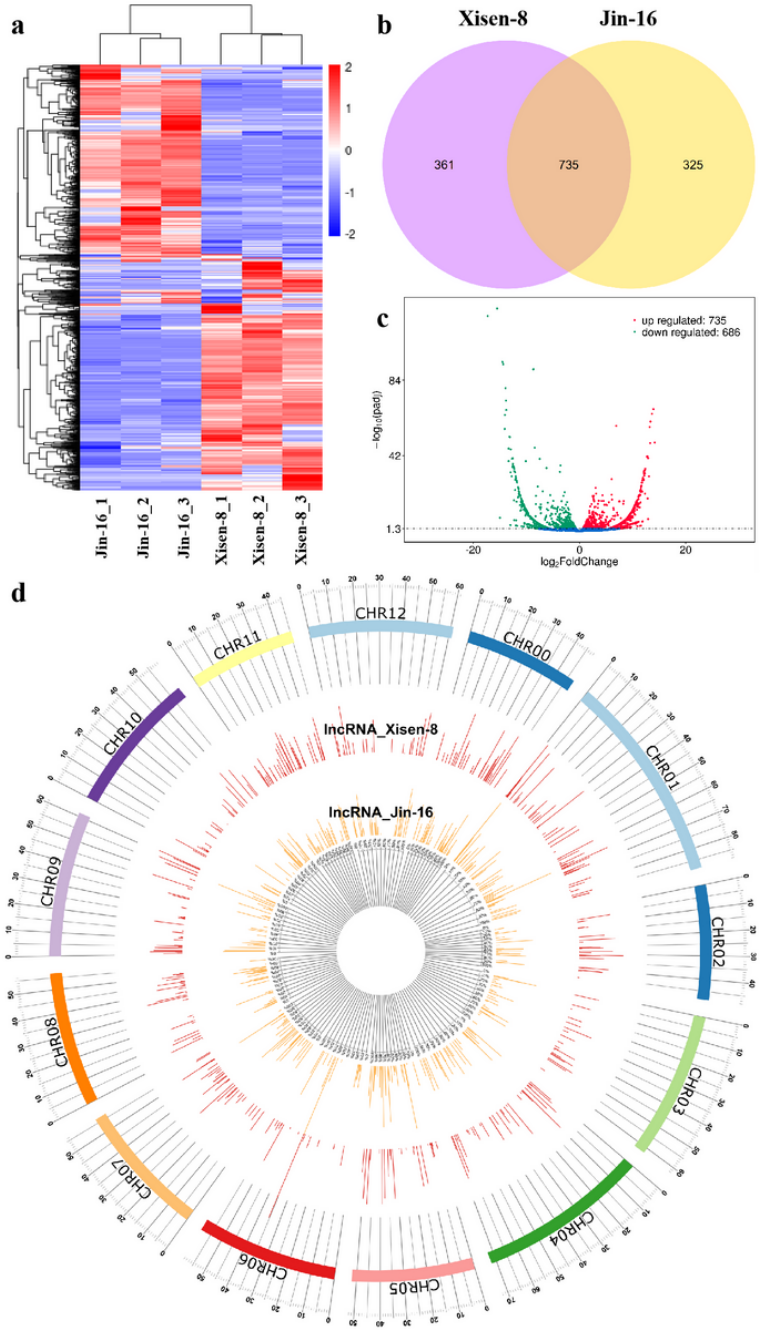


Figure 3

The expression and chromosomal localization of 1421 DE lncRNAs between ‘Jin-16’ and ‘Xisen-8’ potato tubers. (a) Hierarchical clustering analysis of DE lncRNAs in three samples of ‘Jin-16’ and ‘Xisen-8’; (b) Venn diagram showing the counts of specific and common DE lncRNAs in ‘Jin-16’ and ‘Xisen-8’; (c) Volcano plot showing the numbers of up-regulated and down-regulated DE lncRNAs in ‘Xisen-8’. The X-axis represented the log2 (fold change) values of DE lncRNAs in different samples; the Y-axis represented the -log10 (padj) value in expression between the samples; the red dots and green dots represented the up-regulated and down-regulated DE lncRNAs, respectively. (d) The distribution of DE lncRNAs on different chromosomes. The outside ring represented different potato chromosomes. The orange and red columns inside presented the expression levels of corresponding DE lncRNAs in ‘Jin-16’ and ‘Xisen-8’ tubers, respectively.

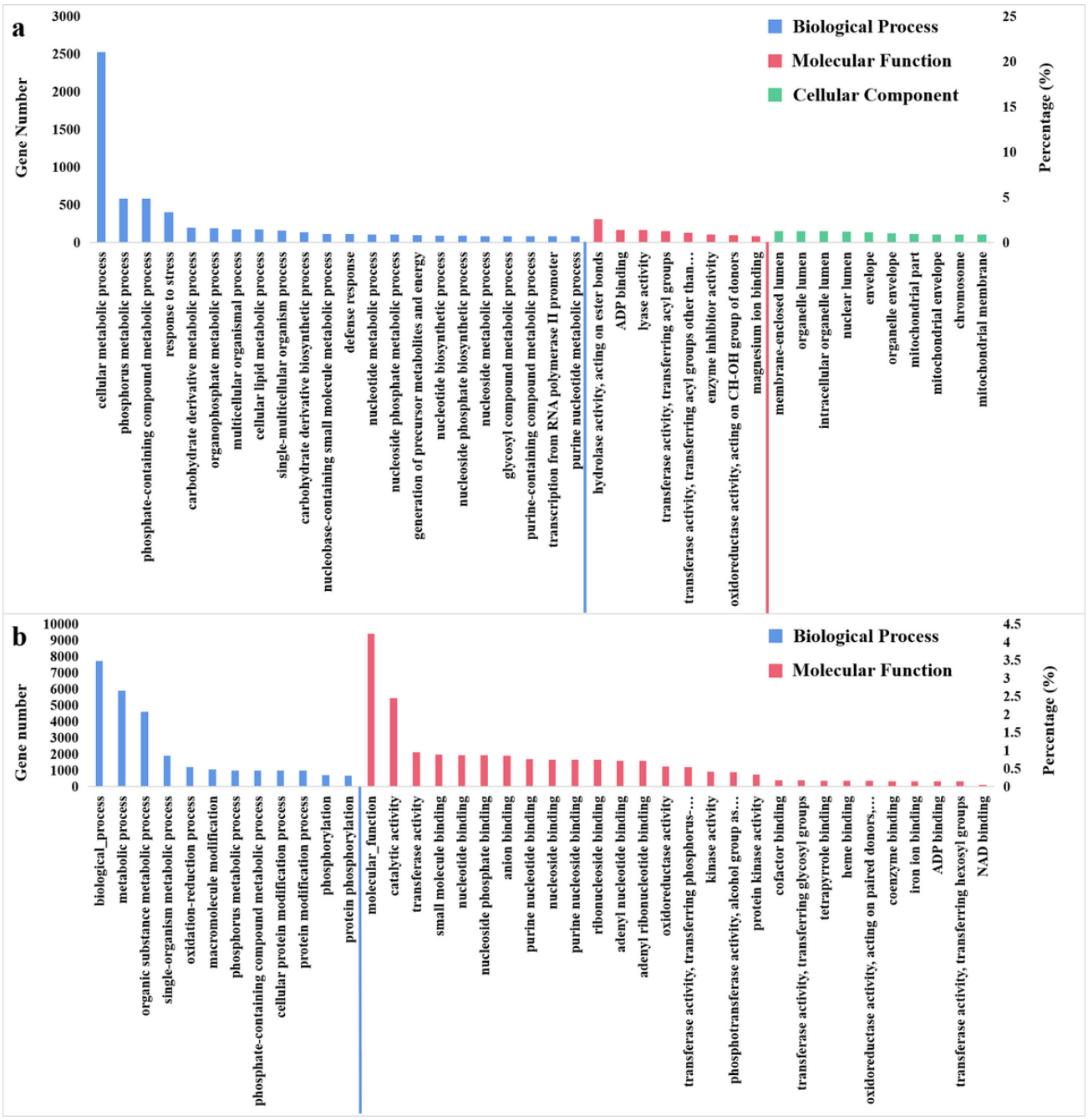


Figure 4

Gene Ontology classification of the co-located target genes (a) and co-expressed target genes (b) of DE lncRNAs. GO functions were presented in X-axis; the number and the percentage of annotated target genes in each GO term were presented in left and right Y-axis, respectively.

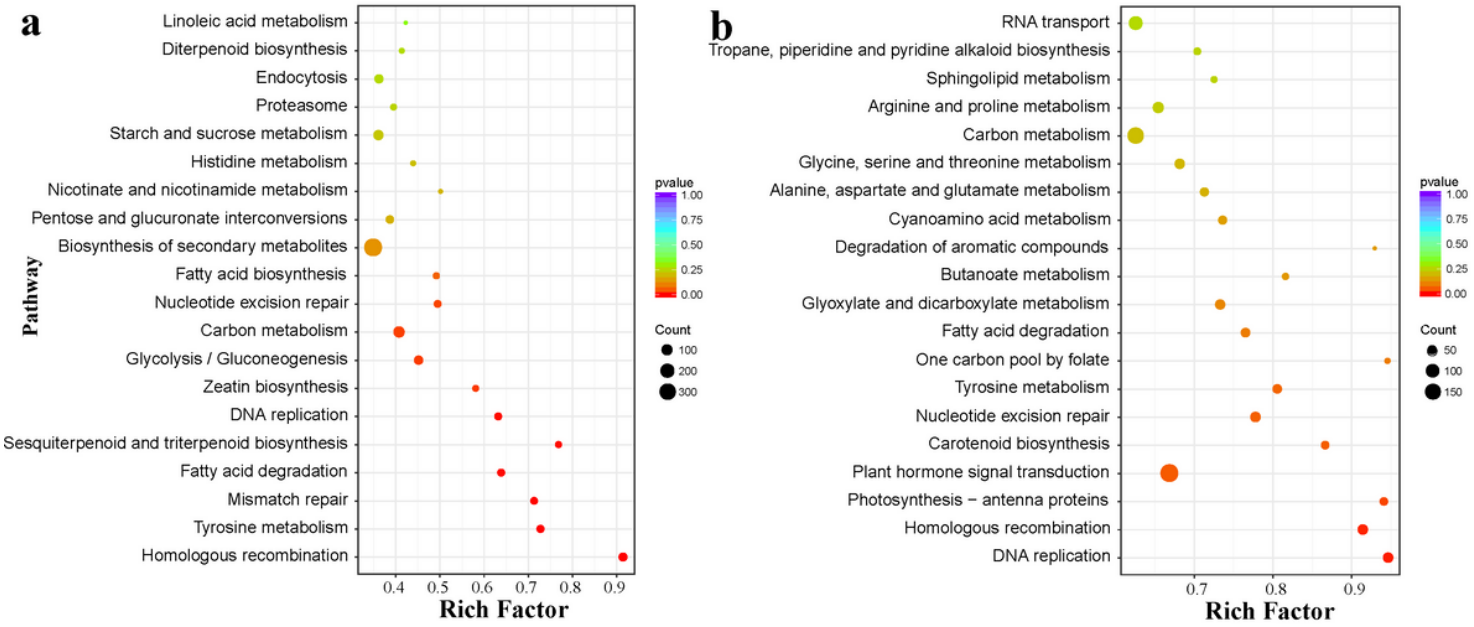


Figure 5

Top 20 biological pathways enrichment of the target genes regulated by the DE lncRNAs. The dot size was proportional to the target gene numbers. The different dot color represented the p value of each pathway. Rich Factor referred to the ratio between the number of target genes of DE lncRNAs and the total number of annotated genes in each pathway.

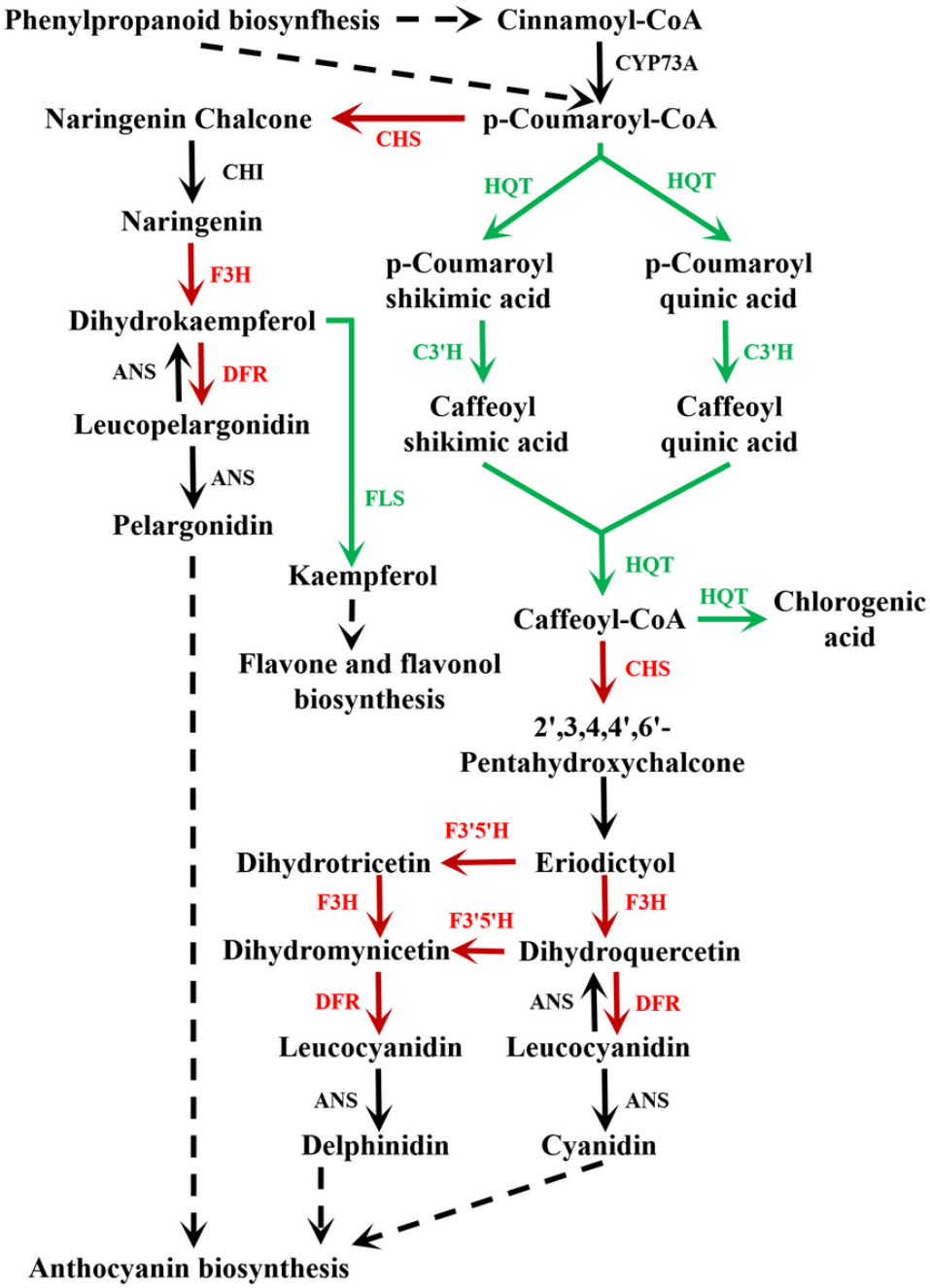


Figure 6

Expression patterns of structural genes in the synthesis pathway of anthocyanin in ‘Xisen-8’ tubers. The enzyme genes that were up-regulated in ‘Xisen-8’ were colored red, and down-regulated enzyme genes were colored green; genes in black letters were insignificantly expressed. The red and green arrows indicated that the reactions were promoted and inhibited, respectively. CYP73A, Cinnamate 4-hydroxylase; CHS, chalcone synthase; CHI, chalcone isomerase; F3H, flavanone 3-hydroxylase; ANS, anthocyanidin synthase; DFR, dihydro-flavonol 4-reductase; FLS, flavonol synthase; C3’H, P-coumaroyl quinate/shikimate 3’-hydroxylase; HQT, shikimate O-hydroxycinnamoyltransferase; F3’5’H, Flavonoid 3’,5’-hydroxylase.

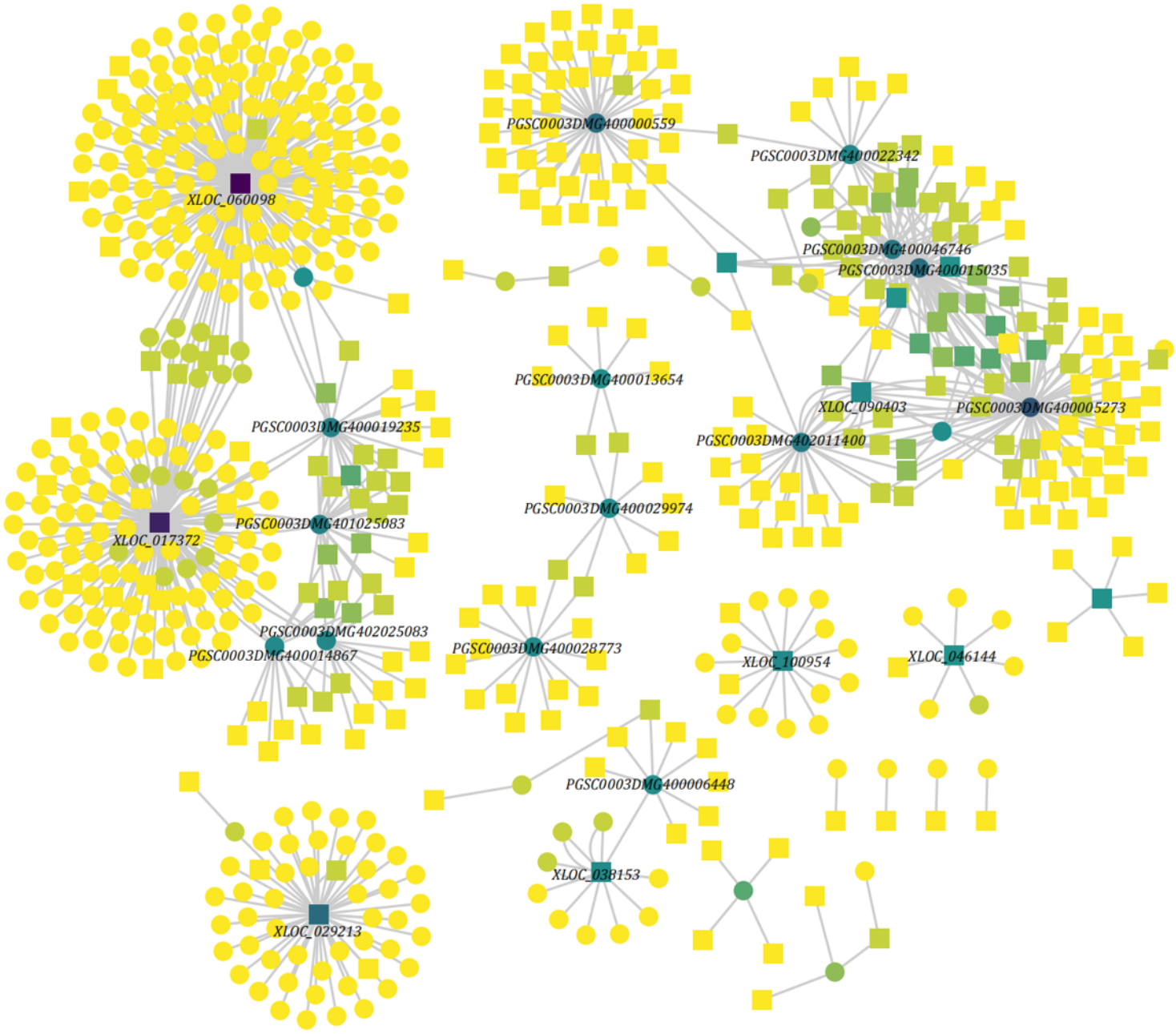


Figure 7

Co-expression network between DE lncRNAs and their co-expressed targets related to anthocyanin synthesis. Nodes represented DE lncRNAs and mRNAs in potato, edges indicated pairwise correlation. LncRNAs and mRNAs were represented as squares and circles, respectively. The different colors of the nodes indicated the numbers of lncRNA or mRNA involved in the interaction: the darker the color, the more interactions were involved. The colors were in order from dark to light: purple, dark green, light green and yellow.

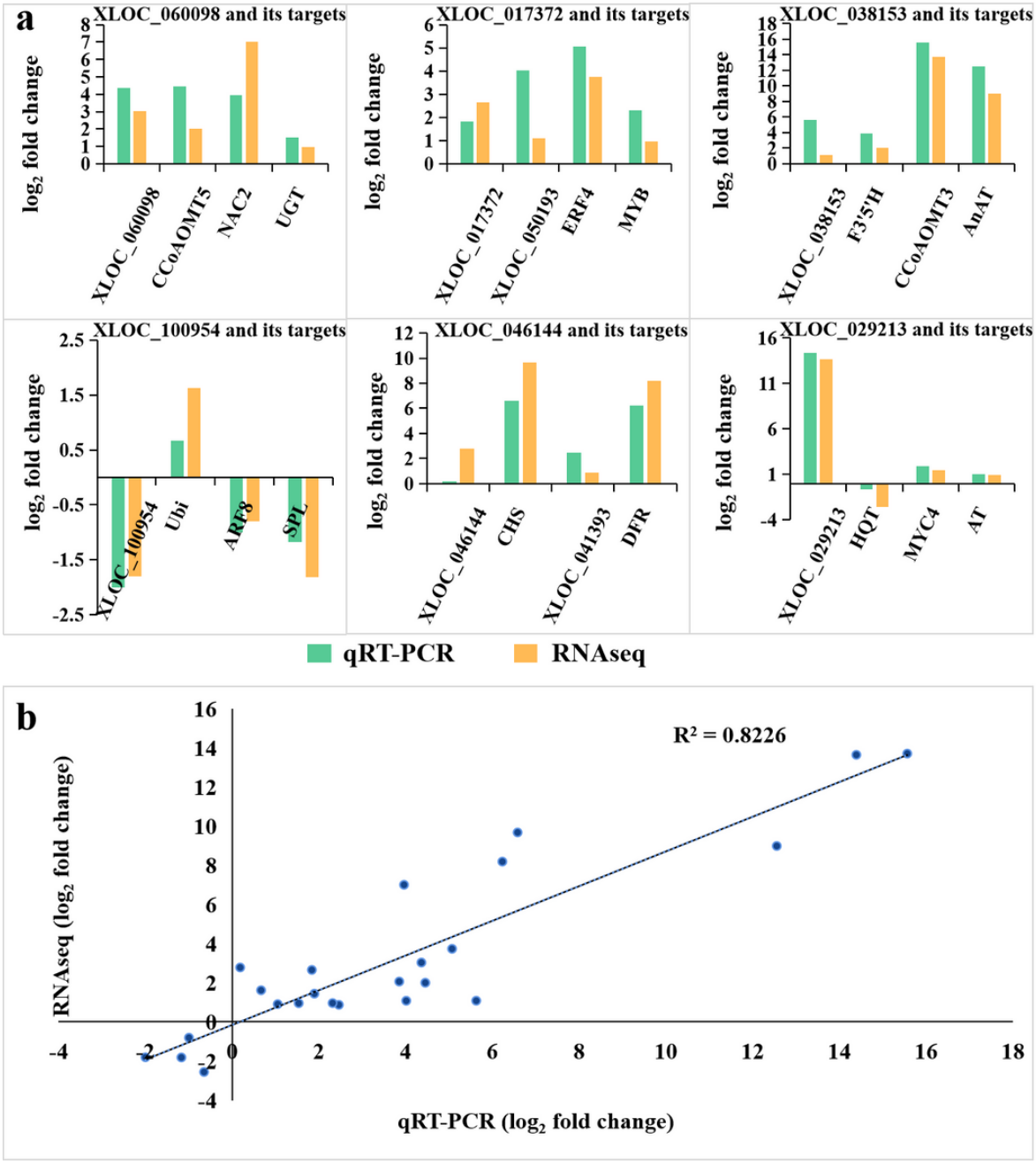


Figure 8

Analysis of the expression levels of DE lncRNAs and their corresponding targets by comparing the results obtained from RNA-seq and qRT-PCR. (a) Expression fold changes of 8 DE lncRNAs and 16 mRNAs by RNA-seq and qRT-PCR. (b) Correlation analysis between RNA-seq and qPCR results. The value showed the expression changes of lncRNA or mRNA under two different methods, that is log2 (lncRNA or mRNA expression in ‘Xisen-8’ tubers / its expression in ‘Jin-16’ tubers).

Supplementary Files

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

- Additionalfiles.xlsx