BMC Genomics

RESEARCH Open Access

Understanding the molecular regulation of flavonoid 3′-hydroxylase in anthocyanin synthesis: insights from purple qingke

Lupeng Chen[1,2,3,4], Youhua Yao[1,2,3,4], Yongmei Cui[1,2,3,4], Xin Li[1,2,3,4], Likun An[1,2,3,4], Yixiong Bai[1,2,3,4], Xiaohua Yao[1,2,3,4*] and Kunlun Wu[1,2,3,4*]

Abstract

Background The Flavonoid 3′-hydroxylase gene(*F3′H*) is an important structural gene in the anthocyanin synthesis pathway of plants, which has been proven to be involved in the color formation of organs such as leaves, flowers, and fruits in many plants. However, the mechanism and function in barley are still unclear.

Results In order to explore the molecular mechanism of the grain color formation of purple qingke, we used the cultivated qingke variety Nierumzha (purple grain) and the selected qingke variety Kunlun 10 (white grain) to conduct transcriptomic sequencing at the early milk, late milk and soft dough stage. Weighted Gene Co-expression Network Analysis (WGCNA) was used to construct weighted gene co-expression network related to grain color formation, and three key modules (brown, yellow, and turquoise modules) related to purple grain of qingke were selected. *F3′H* (HORVU1Hr1G094880) was selected from the hub gene of the module for the yeast library, yeast two-hybrid (Y2H), subcellular localization and other studies. It was found that in purple qingke, HvnF3′H mainly distributed in the cytoplasm and cell membrane and interacted with several stress proteins such as methyltransferase protein and zinc finger protein.

Conclusions The results of this study provide reference for the regulation mechanism of anthocyanin-related genes in purple grain qingke.

Keywords Qingke, Anthocyanin synthesis, Weighted gene co-expression network, Yeast two-hybrid, Subcellular localization, *F3′H*

*Correspondence:
Xiaohua Yao
yaoxiaohua009@126.com
Kunlun Wu
wklqaaf@163.com
[1]Academy of Agricultural and Forestry Sciences, Qinghai University, Xining, Qinghai 810016, China
[2]Laboratory for Research and Utilization of Qinghai Tibet Plateau Germplasm Resources, Xining, Qinghai 810016, China
[3]Qinghai Key Laboratory of Hulless Barley Genetics and Breeding, Xining, Qinghai 810016, China
[4]Qinghai Subcenter of National Hulless Barley Improvement, Xining, Qinghai 810016, China

Background

Qingke (*Hordeum vulgare* L. var. *nudum* Hook. F.) being a characteristic crop of the Qinghai-Tibet Plateau, plays a crucial role in providing food and forage for the Tibetan people [1]. In botanical classification, qingke belongs to a variety of barley, whose genome is highly similar to that of barley, and the inner lemma of qingke can be separated from the caryopsis at maturity, which is also known as naked barley [2]. As the main grain crop in the Qinghai-Tibet Plateau, with the gradual improvement of people's living standards, the dietary demand has changed from focusing on food and clothing to emphasizing health, and its edible value has attracted much attention [3]. Qingke is rich in protein and contains most of the essential amino acids for the human body. Meanwhile, qingke has more dietary fiber and less fat than other crops, and the content of β-glucan in qingke is the highest among barley crops [4]. In addition, qingke is rich in phenols and anthocyanins, which play an important role in anti-aging and cancer prevention [5].

The biosynthesis pathway of anthocyanins involves multiple enzymes and intermediate products [6]. Phenylalanine aminotransferase (PAL) is the first enzyme in the anthocyanin biosynthesis pathway, converting phenylalanine to transaminate coumaric acid, the subsequent steps include converting to cinnamic acid, coumarin acid, and ultimately synthesizing anthocyanins [7]. In addition, the anthocyanin synthesis pathway is affected by various regulatory factors [8]. Among them, transcription factors are one of the most important regulatory factors [9]. The complex of MYB, bHLH and WD40 transcription factors is called MYB-bHLH-WD40 (MBW) complex, which plays a key role in anthocyanin biosynthesis pathway [10, 11]. MYB transcription factors regulate the expression of specific anthocyanin synthesis genes by interacting with bHLH and WD40 transcription factors [12]. In addition, other transcription factors and hormone signals are also involved in the regulation of anthocyanin biosynthesis, and plant hormones such as ethylene can affect anthocyanin biosynthesis through signal transduction pathways [13]. These hormones can regulate anthocyanin synthesis pathway by altering the expression of related genes [14]. Not only that, environmental factors also have a certain impact on anthocyanin biosynthesis [15]. Environmental factors such as light, temperature and adversity can affect anthocyanin accumulation [16–18]. For example, UV exposure can stimulate plants to synthesize more anthocyanins as a defense mechanism against UV radiation [19].

The anthocyanin regulatory pathway is a complex network involving the interaction of many factors, including gene expression regulation, substrate supply, environmental factors and signal transduction pathways [20, 21]. These regulatory mechanisms together affect anthocyanin synthesis and accumulation in plants. At the same time, the anthocyanin regulatory pathway may be different in different plant species and environmental conditions. At present, the mechanism of regulating anthocyanin synthesis of qingke grain is still unclear.

Recent high-throughput sequencing analyses, especially in combination with metabolome and transcriptome analysis, have characterized anthocyanin synthesis pathways in a variety of non-model plants such as asparagus, turnip, and grape [22–24]. However, previous studies of barley seed pigments have mainly focused on comparisons of color differences among seeds; to our knowledge, no studies of the molecular mechanisms of seed anthocyanin formation are available and understanding this mechanism is important for the improvement of barley seedcoat color. Here, genes associated with anthocyanin biosynthesis and its regulations were identified using WGCNA based on transcriptomics. of these, the hub genes, accumulation-related structural genes and transcription factors were further examined. The results improve our understanding of the molecular mechanisms underlying anthocyanin biosynthesis in qingke and provide valuable information for the genetic improvement of purple qingke and other purple crops.

Materials and methods

Plant material and data acquisition

Nierumuzha (purple grain) and Kunlun 10 (white grain) were obtained from Academy of Agriculture and Forestry Sciences, Qinghai University. According to the growth scale, the seeds of two varieties were collected at early milk, late milk and soft dough stage for gene expression analysis [25]. Transcriptome sequencing was performed using three biological replicates of each sample [26].

Construction and analysis of weighted gene co-expression network

The WGCNA package of R language was used to construct gene co-expression network analysis (https://horvath.genetics.ucla.edu/html/CoexpressionNetwork/Rpackages/WGCNA/index.html). Then the weighting coefficient β was calculated by the pickSoftThreshold function to make the WGCNA conform to the scale-free network distribution. The pickSoftThreshold function in WGCNA package is used to calculate the weighting coefficient β to make WGCNA conform to the scale-free network distribution. The β value with correlation coefficient squared close to 0.8 and certain gene connectivity was selected as the soft threshold of WGCNA. The dynamic tree cut method and blockwiseModules are used to identify coexpression modules and construct the network.

Principal component analysis was carried out for all genes in each module. PC1 value was used as the module

feature vector ME to calculate the correlation coefficient r and corresponding p value between the module feature vector and different traits in each module. Where, $r>0$ represented positive correlation, and $r<0$ represents negative correlation. Then we selected $|r|>0.70$, $p<0.001$ as correlation module.

SoftConnectivity function was used to calculate gene connectivity, and hub genes of each module (top 5 genes of connectivity) were screened. The weight values between different genes in each module are calculated according to the topological overlap matrix and visualized using Cytoscape v3.9.1.

RNA-seq analysis and functional annotation

We used the MiniBEST Universal RNA Extraction Kit (TaKaRa, Tokyo, Japan) to extract RNA from the seedcoat of Nierumuzha and Kunlun10 at three stages, and set three biological replicates for each sample. RNA seq and functional annotation by Novogene (Beijing, China). We functionally annotated DEGs using KEGG (Kyoto Encyclopedia of Genes and Genomes; www.kegg.jp/kegg/kegg1.html) and GO (Gene Ontology; http://wego.genomics.org.cn/cgi-bin/wego/index.pl) enrichment analyses and associated analyses using the OmicStudio tools (https://www.omicstudio.cn/tool/62).

Reverse transcription-polymerase chain reaction (RT-PCR) and quantitative-real-time polymerase chain reaction (qRT-PCR) analysis

The reverse transcription of the total RNA of the sample was completely carried out according to the steps of the cDNA synthesis kit (PrimeScript™ II 1st Strand cDNA Synthesis Kit) and stored at -80℃. qRT-PCR were performed using the SYBR Premix Ex Taq II Kit (Takara, Tokyo, Japan) on a Roche LightCycler480 II (Roche Molecular Biochemicals, Mannheim, Germany). Design primers for 15 hub genes using Primer 5.0 (Table S1). The qRT-PCR operation program refers to the method of Yao et al. [26]. Using TC139056 as an internal reference gene for qRT-PCR analysis [27]. Gene expression levels were calculate using the $2^{-\Delta\Delta Ct}$ method.

Subcellular localization

We constructed p35S::HvnF3'H: GFP recombinant vector by reconstructing *HvnF3'H* CDS sequence onto a carrier containing green fluorescent protein (Table S2). The recombinant vector was transformed into Agrobacterium tumefaciens EHA105, and the agrobacterium carrying the recombinant vector was transferred into Nicotiana benthamiana by injection, cultured in the dark for 2–3 d, and sliced by laser scanning confocal microscope (Nikon A1R+, Japan).

Construction of qingke cDNA library

The total RNA of Nierumuzha seed skin was extracted by Trizol Kit (Invitrogen™, Shanghai). The mRNA was isolated and purified using the Oligotex mRNA Midi Kit (Qiagen, Shanghai). The cDNA library (Uncut type) was constructed using the CloneMiner II cDNA library construction kit (ThermoFisher, Shanghai) by reverse transcription of RNA.

We recombined cDNA onto pDONR222 vector and transformed the recombinant vector into *Escherichia coli* (*E. coli*) DH10B for primary library construction. We reconstructed the plasmid extracted from the primary library onto the pGADT7-DEST vector, transformed the recombinant vector into Escherichia coli DH10B, then cultured it at 37℃ to construct the secondary library. Further, the secondary library plasmid was transformed into yeast Y187 to construct a yeast library, and 100 μL diluter 1/10, 1/100, 1/1000, 1/10,000 was applied to 100 mm SD/-Leu plate to calculate the conversion efficiency. Library capacity identification was performed using 10 μL of converted *E. coli* stock solution, which was diluted 100 times with SOC medium, 50 μL of the dilute solution was applied to LB plate (containing 50 μg/mL kanamycin) and cultured in a constant temperature incubator at 37℃ overnight. The clones on the plate were counted on the second day.

Yeast two-hybrid (Y2H) verification

We used pGBKT7-F3'H as bait protein for yeast library screening, sequenced the positive clones, and recombined them into pGADT7 vector according to the sequence. Two germplasms containing AD and BK vectors were co-transferred into the yeast receptive state, and 600 μL PEG/LiAC solution was added with high-speed vortex. Shake at 30°C at 200 rpm for 30 min, then add 70 μL DMSO, slowly invert and mix; Heat shock at 42°C for 15 min, place on ice for 2 min (quickly insert into ice bath), centrifuge at 14,000 rpm for 5s, remove supernatant, and add 100 μL 1×TE suspensible bacteria. Apply to SD/-Leu-Trp for screening, grow at 30°C for 2–3 d. The monoclones were scraped and gradient diluted with SD/-Leu-Trp liquid Dilute the bacterial solution onto SM/-Leu-Trp for self-activation validation, and apply it onto SD/-Trp His-Ade deficient culture medium. Grow at 30℃ for 2–3 d for protein interaction validation.

Bimolecular fluorescence complementation (BiFC) assays

Cloning the CDS sequence of *HvnF3'H* and the gene screened by Y2H using primers with *Pac I* and *Spe I* double enzyme cutting sites. Recombine HvnF3'H onto the pBiFC-YC vector, and recombine the genes screened by Y2H onto the pBiFC-YN vector (Table S2). The extracted plasmid was transformed into *Agrobacterium tumefaciens EHA105*. *Agrobacterium* with recombinant vector

was injected into *Nicotiana benthamiana* and cultured in the dark for 3 d. The sections were observed using laser confocal microscopy (Nikon A1R+, Japan).

Results

Data processing and WGCNA

To better obtain genes associated with anthocyanin synthesis, we generated RNA-seq data for the seedcoats of Kunlun10 and Nierumuzha at the three different developmental stages (18 samples). In total, we generated 112.62 Gb clean data, an average of 6.26 Gb per sample. The average Q20 ratio of each sample was 97.41–98.08%, and the Q30 was 93.17–94.36%. The average GC content was 53.88%. Across the six samples (early milk, late milk and soft dough of Kunlun 10; early milk, late milk and soft dough of Nierumuzha), 37.34 (91.01%), 37.39(88.89%), 36.53(87.32%), 37.35(90.32%), 36.35(89.21%), and 36.76 (85.02%)million reads, respectively, were mapped to the reference barley (H. vulgare) genome in total, while 33.34(81.25%), 33.17(78.87%), 32.44(77.53%), 33.09 (80.01%), 31.44 (77.14%), and 31.92 (73.83%) million reads, respectively, were uniquely mapped.

After removal of all genes that were unexpressed in 18 samples (FPKM=0), the remaining 25,375 genes were used for WGCNA. The optimal soft threshold of 10 is selected to construct an approximate scale-free topological overlap matrix (Fig. 1A and B). A scale-free network was constructed to obtain 25 colored modules, corresponding to clusters of co-expressed genes (Fig. 1C). The cluster modules ranged in size from 49 genes (orange) to 5801 genes (turquoise).

Identification of color-related modules in seedcoat

Due to the large number of genes in each module, representative genes (module characteristic gene, ME) in each module were selected for clustering. We conducted cluster analysis on 25 co-expression modules, and the results showed that the 25 co-expression modules were

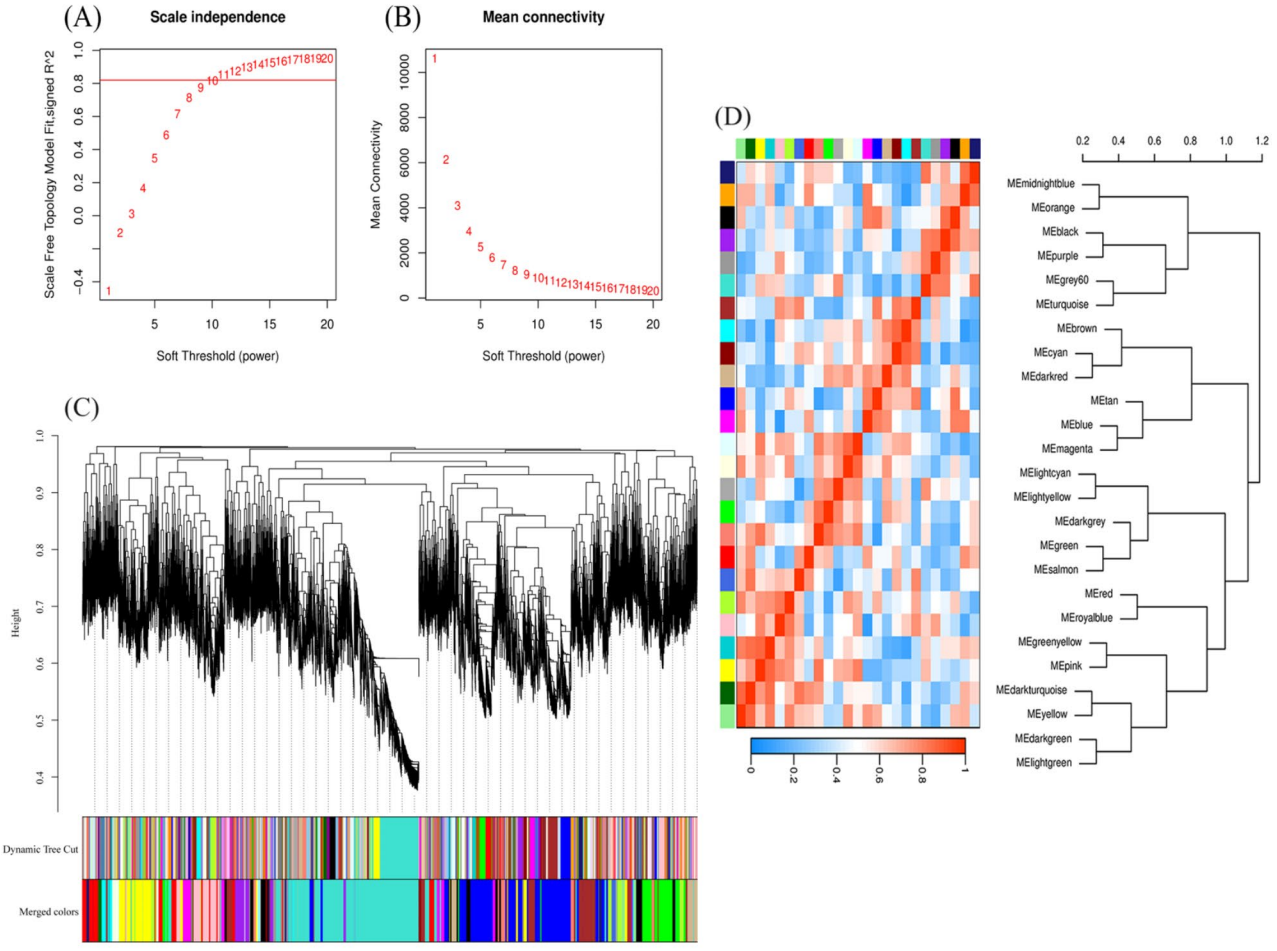

Fig. 1 Constructions of gene co-expression network. **A**: scale-free topology fit index as a function of the soft-thresholding power, the red line represents that R^2 is equal to 0.8. **B**: mean connectivity as a function of the soft-thresholding power. **C**: Clustering dendrograms of genes and module detecting. **D**: Gene clustering heat map of module

divided into 3 categories, each of which contained two main branches (Fig. 1D). At the same time, the heat map of correlation between modules shows that there are different degrees of correlation between the 25 modules.

In order to further understand the relationship between modules and samples, we used the data of six samples in three grain formation periods of two varieties, white grain and purple grain, to conduct correlation analysis with 25 modules obtained, and constructed the heat map of module phenotypic correlation. When the correlation is red, it means that the characteristic gene of the module is positively correlated with the phenotype; on the contrary, when the correlation is blue, this means that there is a negative correlation, with the correlation getting deeper as the color gets darker. Among the 25 gene modules, black, red and green modules were positively correlated with early milk maturation (r=0.95, P=4e-03), late milk maturation (r=0.90, P=0.02) and soft dough stage (r=0.95, P=4e-03) of Kunlun10 (white), respectively. brown, yellow and turquoise modules were positively correlated with early milk ripening (r=0.96, P=2e-3), late milk ripening (r=0.92, P=9e-3) and soft dough ripening (r=0.95, P=4e-3) of Nierumuzha (purple), respectively (Fig. 2).

Analysis of gene expression patterns of key modules

After obtaining the key modules, we analyzed the expression levels of all genes and ME genes in these six modules in all samples respectively, and found that the gene expression levels in each module were highly correlated, and the ME expression level was also highly correlated with the overall expression level of the module, indicating that the ME of the target module could fully represent the whole gene in its module. Red represents high expression, while green represents low expression. The results showed that genes of black, red and green modules were highly expressed in early milk, late milk and soft dough stage of Kunlun10 (white grain), respectively. Genes of brown, yellow and turquoise modules were highly expressed in early milk, late milk and soft dough stage of Nierumuzha (purple grain), respectively (Fig. 3).

Functional enrichment analysis of key module genes

With the progress of the period, the grain color of purple variety Nierumuzha gradually deepened. In order to further explore the function of gene modules related to anthocyanin synthesis of qingke grains, GO enrichment was analyzed using brown, yellow and turquoise modules related to purple grains. They were divided into three categories: biological process, cell component and molecular function. As you can see, in terms of biological processes, brown module was mainly enriched in intracellular transport (GO: 0046907), establishment of localization in cell (GO: 0051649) and protein transport (GO: 0015031) and other regulation and colligation; yellow module was mainly enriched in DNA integration (GO:

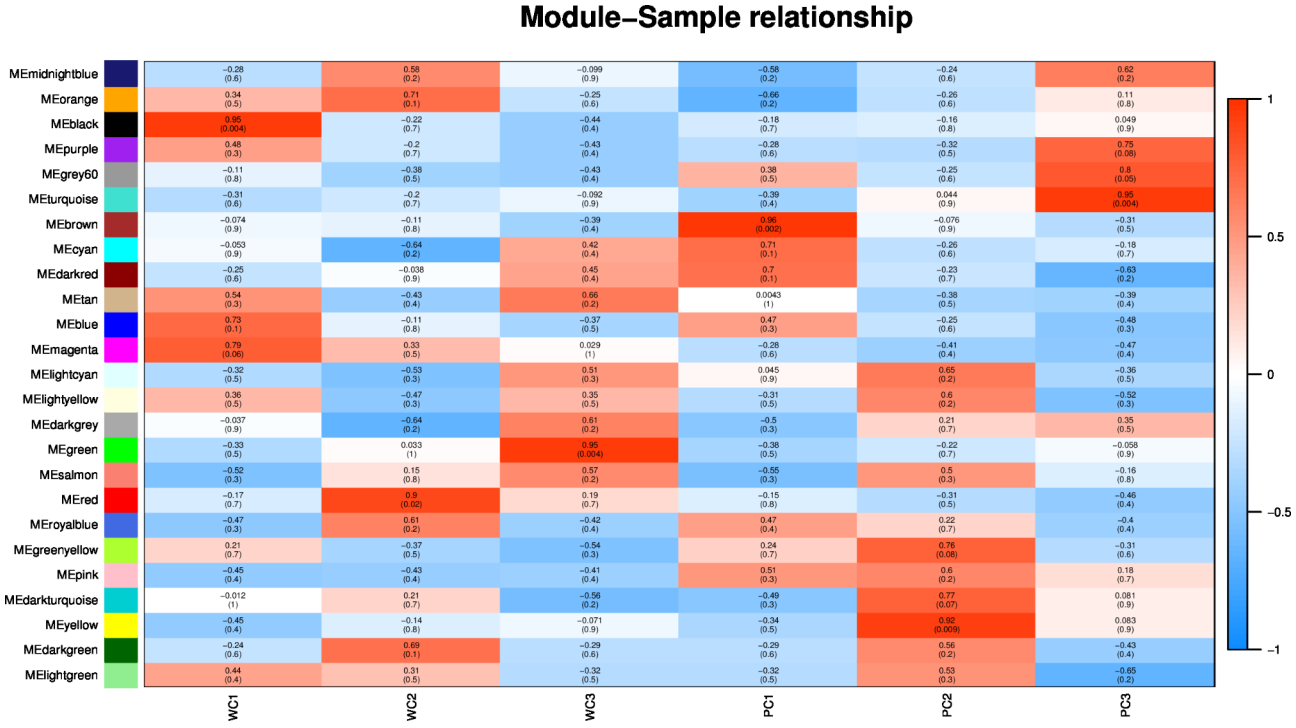

Fig. 2 Correlation between modules and traits. Each row represents a module and each column represents a sample. The numbers in the rectangular box represent the correlation coefficient between the module and the sample and the corresponding *p* value

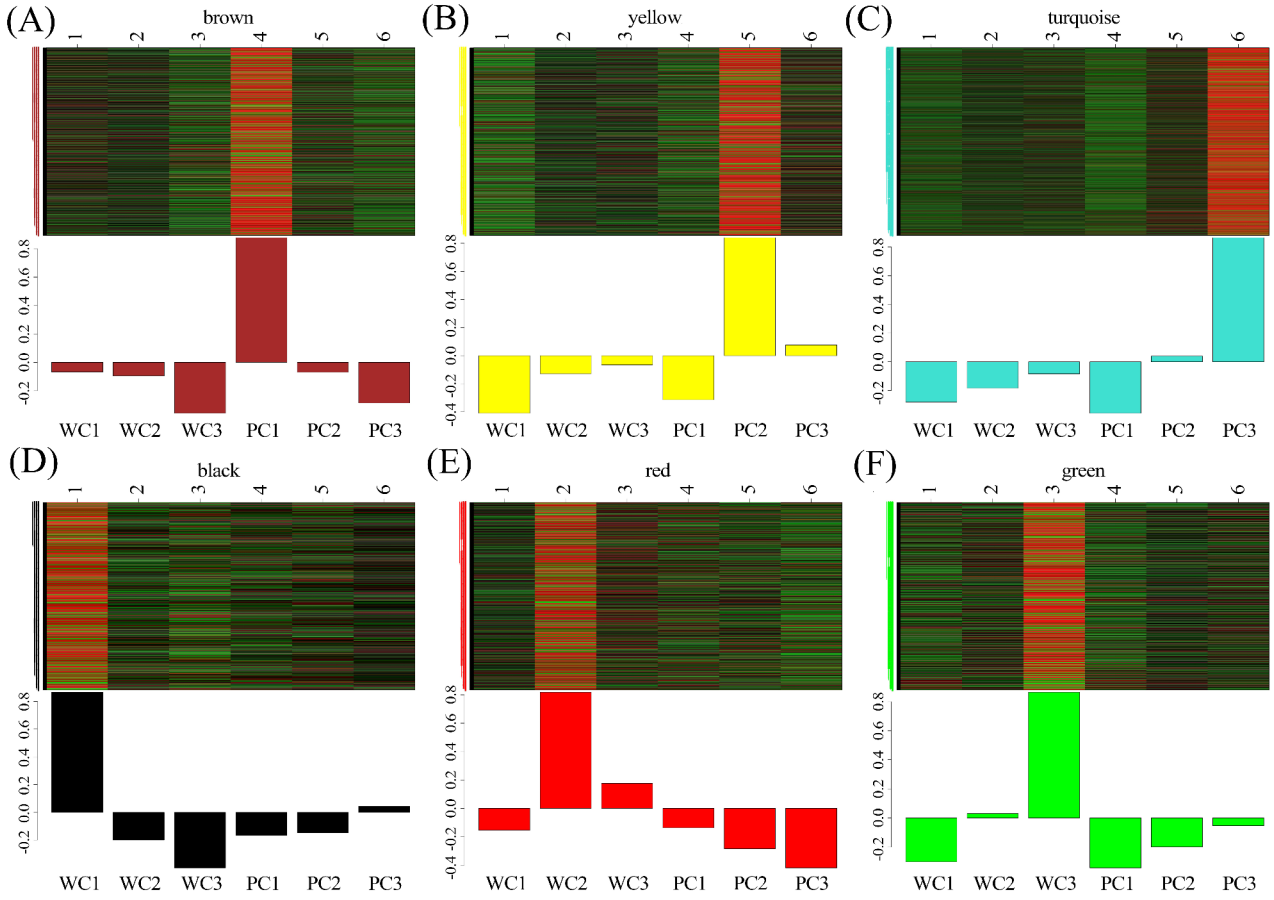

Fig. 3 Expression levels of all genes and corresponding ME in different modules of each sample. The above figure (heat map) shows the expression level of all genes in modules. The row represents all genes in the module, and the column represents each sample. **A**: brown module; **B**: yellow module; **C**: turquoise module; **D**: black module; **E**: red module; **F**: green module. The below figures show the ME expression level of the module in the samples

0015074). Turquoise module was mainly enriched to protein ubiquitination (GO: 0016567), protein modification by small protein conjugation (GO: 0032446) and nitrogen compound transport (GO: 0071705). In terms of molecular functions, brown module was mainly enriched in microtubule binding (GO: 0008017) and tubulin binding (GO: 0015631). Yellow module was mainly enriched in histone binding (GO: 0042393) and transcription factor activity, protein binding (GO: 0000988) and RNA binding (GO: 0003723); turquoise module was mainly enriched in ubiquitin-protein transferase activity (GO: 0004842) and ubiquitin-like protein transferase activity (GO: 0019787). In terms of cell components, brown module was mainly enriched in three cell components, endomembrane system (GO: 0012505), membrane coat (GO: 0030117) and coated membrane (GO: 0048475). Yellow module mainly was enriched in transferase complex (GO: 1990234), nucleoplasm (GO: 0005654) and nucleoplasm part (GO: 0044451). Turquoise module was mainly enriched in peroxisome (GO: 0005777) and microbody (GO: 0042579). We analyzed the KEGG pathway in order to explore the pathway of key module genes. Brown module was mainly enriched in the Glycine, serine and threonine metabolism (ath00260) pathway. Yellow module was mainly concentrated in phagosome (ath04145), oxidative phosphorylation (ath00190) and lysine degradation (ath00310). Turquoise module was mainly enriched in biosynthesis of amino acids (ath01230) and plant-pathogen interaction (ath04626) pathways (Fig. 4).

Construction of gene interaction network

In order to obtain the hub genes in the three purple granulose-related modules, Cytoscape software was used to visualize the gene regulatory network. Five genes with the highest connectivity in each module were selected as the hub genes (red dots) (Fig. 5). The hub genes in brown module were *GST* (HORVU4Hr1G081110), *MFP1* (HORVU4Hr1G069210), *GR-RBPs* (HORVU5Hr1G040380), *UBXD5* (HORVU6Hr1G016460) and *d-Cdes* (HORVU6Hr1G079140). The hub genes ID in yellow module were *IAA*(HORVU1Hr1G028170), *PMEI* (HORVU1Hr1G093370),

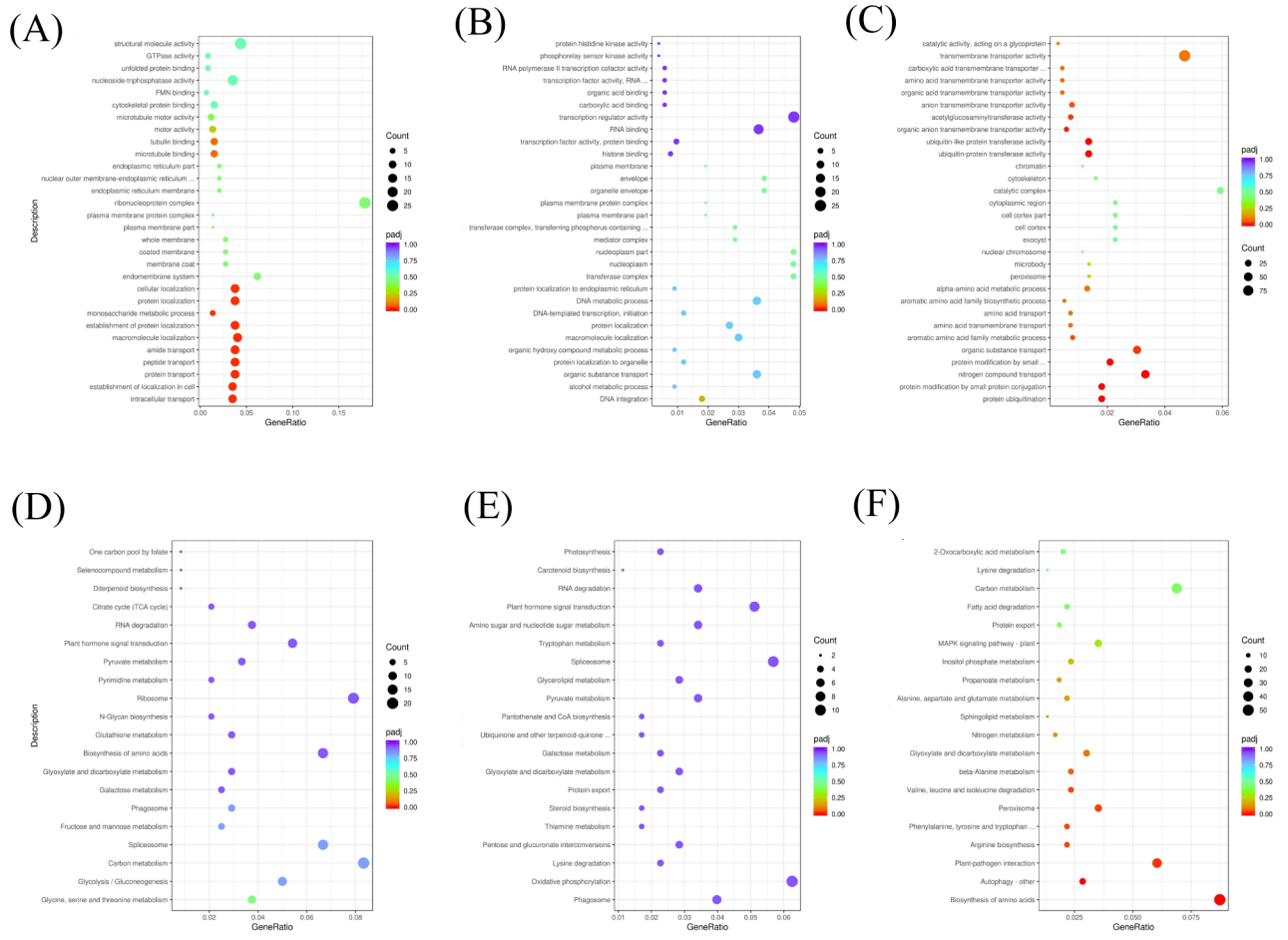

Fig. 4 GO functional and KEGG pathway enrichment analysis in the three modules. **A**, **B** and **C**: GO enrichment analysis of three modules; **D**, **E** and **F**: KEGG enrichment analysis of three modules. The vertical axis represents the enriched GO term or KEGG pathway, and the horizontal axis represents the ratio of enriched genes in the module. The point size represents the gene number enriched in the pathway, and the color represents the *p*-value corrected for multiple testing

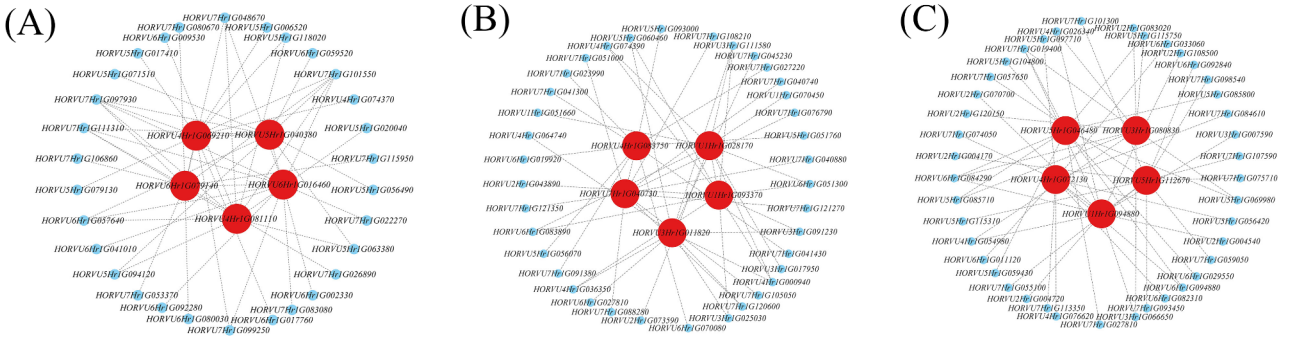

Fig. 5 Gene co-expression networks within specific modules related to anthocyanin synthesis. **A**, **B**, and **C** represent brown, yellow, and turquoise modules respectively. The red circles in each network map represent the core genes of each module. Genes are connected by line segments, indicating that genes are related to each other

IMO(HORVU3Hr1G011820), *ROC8* (HORVU4Hr1G083750), and *PRB1-2* (HORVU7Hr1G040730). The hub genes ID in turquoise module were *CHI* (HORVU5Hr1G046480, HORVU5Hr1G112670), *F3'H* (HORVU1Hr1G094880), *C4H* (HORVU3Hr1G080830) and *4CL* (HORVU4Hr1G072130) (Table 1).

The expression of hub genes by qRT-PCR

15 hub genes related to anthocyanin in qingke with purple grains, and 15 of them were selected for real-time fluorescence quantitative PCR verification. We used the same batch of RNA in transcriptome sequencing for reverse transcription, and designed primers to detect the expression of these genes in the three stages of the Nierumuzha and Kunlun10 by qRT-PCR (Fig. 6). The test results showed that the expression levels and changing trends of the tested genes in different periods were basically consistent with the transcriptome results, indicating that the selected core genes had important research value in affecting the formation of purple grains. In addition, the qRT-PCR results also proved the reliability of the transcriptome sequencing results.

Table 1 Functional annotations of hub genes in different modules

Module	Hub gene	Gene function
Brown	HORVU4Hr1G081110	Encodes Glutathione S-transferase (*GST*)
	HORVU4Hr1G069210	MFP1 attachment factor 1(*MFP1*)
	HORVU5Hr1G040380	Glycine-rich RNA-binding protein (*GR-RBPs*)
	HORVU6Hr1G016460	UBX domain-containing protein 6 (*UBXD6*)
	HORVU6Hr1G079140	Putative D-cysteine desulfhydrase (*d-Cdes*)
Yellow	HORVU1Hr1G028170	Auxin-responsive protein IAA1 (*IAA*)
	HORVU1Hr1G093370	Putative pectin esterase/pectin esterase inhibitor 38 (*PMEI*)
	HORVU3Hr1G011820	Indole-2-monooxygenase (*IMO*)
	HORVU4Hr1G083750	Homeodomain transcription factor ROC8 (*ROC*)
	HORVU7Hr1G040730	Pathogenesis-related protein PRB1-2 (*PRB*)
Turquoise	HORVU5Hr1G046480	Probable chalcone-flavonone isomerase 3 (*CHI*)
	HORVU1Hr1G094880	Flavonoid 3'-monooxygenase (*F3'H*)
	HORVU3Hr1G080830	Cinnamic acid 4-hydroxylase (*C4H*)
	HORVU4Hr1G072130	4-coumarate–CoA ligase-like 1 (*4CL*)
	HORVU5Hr1G112670	Chalcone-flavonone isomerase (*CHI*)

Spatiotemporal expression analysis of *HvnF3'H*

Quantitative real time Polymerase Chain Reaction (qRT-PCR) was used to detect the relative expression of this gene in the early milk, late milk, and soft dough stage of Nierumzha and Kunlun 10 seed coat color formation. The results showed that the expression of *HvnF3'H* of Kunlun 10 was lower in early milk, late milk and soft dough of grain color formation, but there were significant differences. The expression level of *HvnF3'H* in Nierumzha showed a significant increase in the late milk and soft dough stage ($P<0.01$). The expression of *HvnF3'H* in Nierumzha was significantly higher than that in Kunlun 10 in late milk and soft dough stage ($P<0.01$). It is speculated that with anthocyanin synthesis in seed coat, the expression of *HvnF3'H* is gradually increased, and this gene may positively regulate anthocyanin accumulation. In addition, qRT-PCR was performed on the roots, stems, leaves and awns of the two varieties of Nierumzha and Kunlun 10. The results showed that *HvnF3'H* was not expressed in the roots, stems, leaves and awns of the two varieties, but was significantly increased in the seed coats of Nierumzha (Fig. 7A and B).

Subcellular localization of HvnF3'H

In order to investigate the subcellular localization of the HvnF3'H, we successfully transformed the recombinant vector into Agrobacterium tumefaciens EHA105, and then injected the Agrobacterium tumefaciens containing the recombinant vector into tobacco leaves for microscopic observation. The green fluorescence signal of the HvnF3'H mainly exists in the cytoplasm and cell membrane of *Nicotiana benthamiana*, and there is also a small distribution in the nucleus. The green fluorescence signal of the control spread throughout the entire cell, indicating that HvnF3'H is a protein located in the cell membrane and cytoplasm (Fig. 7C).

Y2H screening of cDNA library

To explore the specific regulatory mechanism of HvnF3'H, we first determined the insertion fragment size and recombination rate of the cDNA library. The findings revealed that the capacity of the primary library is 1.28×107 cfu and that of the secondary library is 1.12×107 cfu. The recombination rate of the two libraries reached 100%, and the average inserted fragment length exceeded 1000 bp, indicating that the constructed libraries met the quality requirements (Fig.S1). We amplified and constructed the *HvnF3'H* into a pGBKT7 vector as a bait protein (pGBKT7-F3'H), and screened it using a yeast nuclear library. In the protein toxicity test, the yeast transformed by the recombinant bait vector grew normally on SD/-Leu/-Trp medium and showed similar colony density as the blank control (pGBKT7 Lam). In the self-activation test, we set a negative control

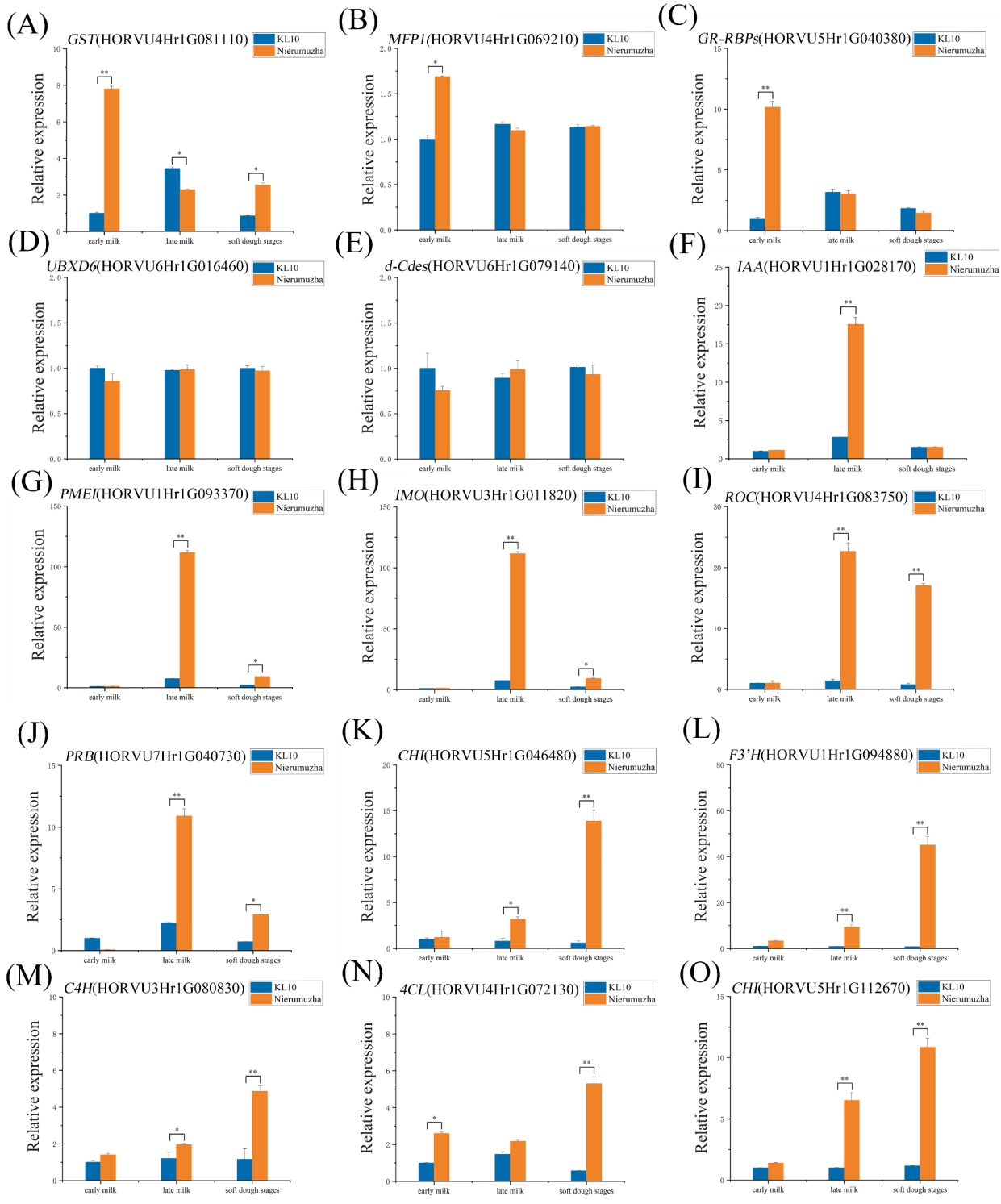

Fig. 6 Expression patterns of the 15 hub genes in Kunlun10 and Nierumuzha. The TC139056 were used as internal. Blue represents the relative expression level of Kunlun 10, while orange represents the relative expression level of Nierumuzha. (**A**): HORVU4Hr1G081110; (**B**): HORVU4Hr1G069210; (**C**): HORVU5Hr1G040380; (**D**): HORVU6Hr1G016460; (**E**): HORVU6Hr1G079140; (**F**): HORVU1Hr1G028170; (**G**): HORVU1Hr1G093370; (**H**): HORVU3Hr1G011820; (**I**): HORVU4Hr1G083750; (**J**): HORVU7Hr1G040730; (**K**): HORVU5Hr1G046480; (**L**): HORVU1Hr1G094880; (**M**): HORVU3Hr1G080830; (**N**): HORVU4Hr1G072130; (**O**): HORVU5Hr1G112670. Analysis of differences in relative expression levels between two varieties at different stages using one-way ANOVA. *: $p < 0.05$; **: $p < 0.01$

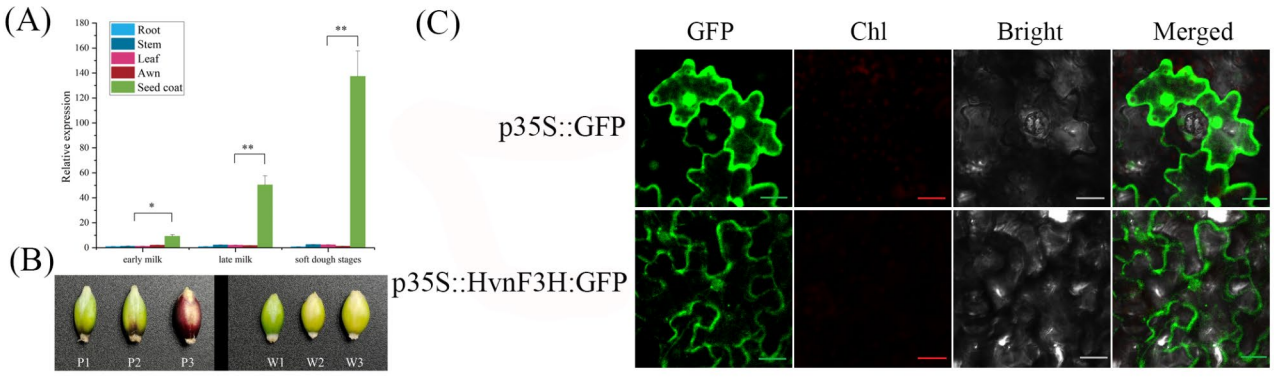

Fig. 7 Expression and subcellular localization of *HvnF3'H*. **A**: Expression of *HvnF3'H* in various tissues. The *TC139056* were used as internal gene; W1, W2, W3 represent the early milk, late milk and soft dough stage of Kunlun10; P1, P2, P3 represent the early milk, late milk and soft dough stage of Nierumuzha; *: $p < 0.05$; **: $p < 0.01$; **B**: Observation of grain phenotypes of two varieties; **C**: *HvnF3'H* was fused with GFP and expressed in *Nicotiana benthamiana* leaves. The fluorescence signal was observed by laser scanning confocal microscope. Scale bars = 50 μm

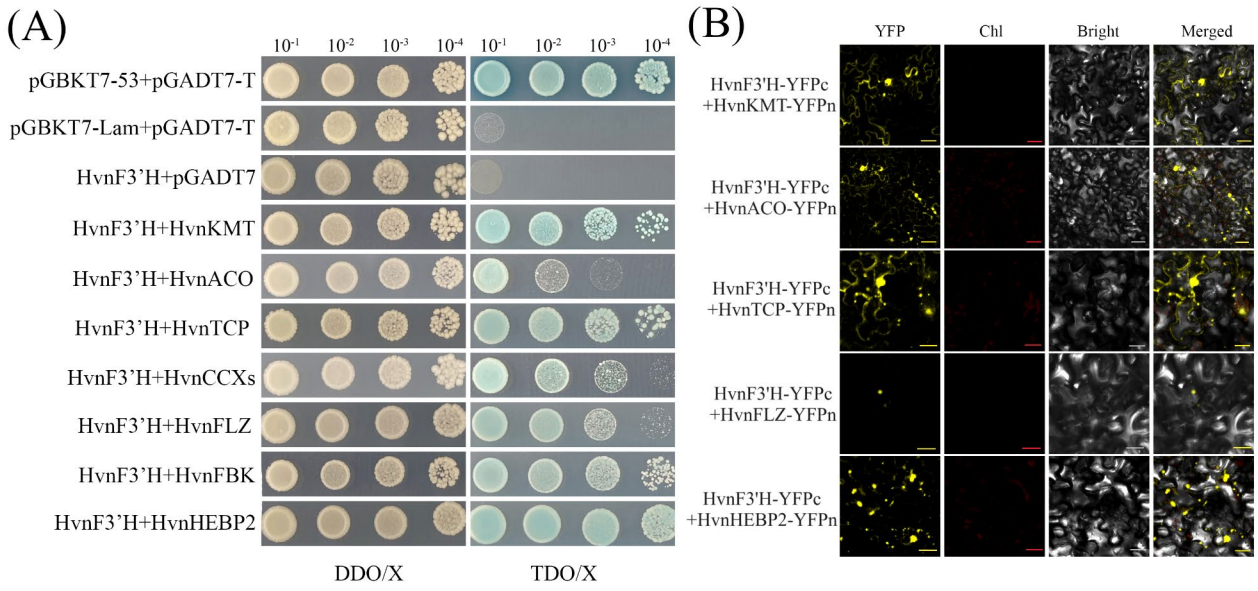

Fig. 8 Y2H screening and BiFC analysis. **A**: The positive clones screened by yeast library were verified one-to-one by yeast two-hybridization, and the dilute concentrations of 10^{-1}, 10^{-2}, 10^{-3} and 10^{-4} bacterial solution were placed on the medium with DDO/X and TDO/X. DDO/X: SD/-Leu/-Trp/X-α-gal; TDO/X: SD/-Leu/-Trp/-His/X-α-gal. **B**: The in vivo interaction between HvnF3'H and screened protein in tobacco leaves was detected by BiFC method, and the fluorescence signal was observed under fluorescence microscope. Scale bars = 50 μm

(pGBKT7 Lam) and a positive control (pGBKT7-53), and through comparison, we found that the recombinant vector yeast did not grow on SD/-Leu/-Trp/-His/-Ade/X-α-gal/AbA medium. The concentration of AbA is 125 ng/mL, and the concentration of X-α-gal is 40 mg/L. The above results indicate that the HvnF3'H has no significant toxicity to yeast and there is no self-activation phenomenon, which can be used for subsequent yeast library screening and one-on-one validation. The pGBKT7-F3'H was co screened with the yeast cDNA library of qingke (DOE20223637), and a total of 34 blue clones were found on the SD/-Leu/-Trp/-His/X-α-gal screening plate. The blue clones on the screening plate were selected for re-screening, and a total of 27 clones grew and turned blue.

We selected genes that were "in-frame" in the library vector for yeast two hybrid one-on-one validation, and obtained a total of seven proteins that interacted with HvnF3'H (Fig. 8A). On this basis, we performed functional annotation on seven genes encoding candidate proteins. Among them, HORVU4Hr1G003060 encodes Lysine N-methyltransferase (KMT), HORVU7Hr1G100810 encodes 1-aminocyclopane-1-carboxylate oxidase-1-like protein (ACO), HORVU6Hr1G093960 belongs to the TCP family transcription factor (TCP), HORVU5Hr1G006550 encodes Cation/calcium exchange 1

(CCXs), HORVU0Hr1G021850 encodes FCS-Like Zinc finger 6-like (FLZ), HORVU7Hr1G026770 encodes F-box protein At5g67140-like isoform X (FBK), and HORVU3Hr1G028810 encodes Heme binding protein 2 (HEBP2) (Table 2).

Table 2 Functional annotation of screening library genes

gene ID	gene name
HORVU4Hr1G003060	Lysine N-methyltransferase (*KMT*)
HORVU7Hr1G100810	1-aminocyclopropane-1-carboxylate oxidase-1-like protein (*ACO*)
HORVU6Hr1G093960	TCP family transcription factor (*TCP*)
HORVU5Hr1G006550	Cation/calcium exchanger 1 (*CCXs*)
HORVU0Hr1G021850	FCS-Like Zinc finger 6-like (*FLZ*)
HORVU7Hr1G026770	F-box protein At5g67140-like isoform X (*FBK*)
HORVU3Hr1G028810	Heme-binding protein 2 (*HEBP2*)

In order to avoid false positives in Y2H, we detected seven proteins interacting with HvnF3'H through biomolecular fluorescence complementarity. We fused the HvnF3'H with the C-terminal yellow fluorescent protein YFPc, and the candidate protein fused with the N-terminal yellow fluorescent protein YFPn, successfully constructing a recombinant vector for *Agrobacterium tumefaciens* transformation. We co transformed *Agrobacterium tumefaciens* containing recombinant vectors into tobacco leaves for microscopic observation, and the results showed that among the seven candidate proteins, except for HORVU5Hr1G006550 and HORVU7Hr1G026770, the remaining five candidate proteins all interacted with the HvnF3'H (Fig. 8B).

Discussion

Anthocyanins were an important class of plant phenolic pigments, which play an important role in the color formation of many plant organs and are the key factors for the red, blue or purple appearance of plant organs [28, 29]. In addition to participating in color formation, anthocyanins were also able to improve plant stress resistance [30]. Previous studies had shown that anthocyanins play an important role in high temperature resistance, cold resistance, drought resistance and ultraviolet radiation resistance [31]. Purple grain qingke is a kind of important germplasm resources, the study of the formation of the color and the synthesis pathway of anthocyanin of purple grain qingke will play the basic role in the improvement of qingke varieties.

In the study of anthocyanins in qingke, we conducted in-depth analysis of the selected hub genes. Functional annotation showed that the hub genes of Nierumuzha in soft dough phase encoded key enzymes in anthocyanin synthesis pathway, among which C4H encoded by HORVU3Hr1G080830 belonged to plant P450 monooxygenase, acting on the first stage of anthocyanin synthesis and catalyzing the second step of phenylalanine pathway [26]. CHI encoded by HORVU5Hr1G046480 is one of the rate-limiting enzymes of flavonoid metabolism pathway and acts on the second stage of anthocyanin synthesis [32]. Previous studies had shown that *CHI* is involved in the formation of plant flower color. By inhibiting *CHI* expression in tobacco using RNA interference technology, the color of petals and pollen became lighter [33]. The GST encoded by HORVU4Hr1G081110 is transported to plant vacuoles through the combination of glutathione (GSH) and Anthocyanidin to promote the accumulation of Anthocyanidin [34]. HORVU6Hr1G016460 was related to ubiquitination. Promoting ubiquitin degradation also had a negative regulation effect on Anthocyanidin accumulation [35]. HORVU1Hr1G028170 and HORVU3Hr1G011820 were related to photosensitive pigments. The transcription factor *HY5* encoding photosensitive pigments could positively regulate anthocyanidin accumulation by combining with anthocyanidin structural gene [36].

In order to further study these hub genes, we found that the expression of *HvnF3'H* was in line with expectations through fluorescence quantification, so we focused on this gene. HORVU1Hr1G094880 encoded flavonoid 3'-hydroxylase, which catalyzes the single oxygenation reaction involving NADPH and O_2, and is another rate limiting enzyme in the flavonoid metabolic pathway [37]. Its metabolites are important intermediates in the process of Anthocyanidin synthesis [38]. Brugliera et al. detected the expression of *F3'H* gene in petunias petals, ovary, calyx, stem and other organs [39]. Castellarin et al. found that *F3'H* was highly expressed in the roots, stems, leaves, flowers and seeds of grape [40]. Compared with the expression of these plants, the expression of *F3'H* in purple grain qingke was relatively simple, and only showed high expression in the seed skin of purple grain qingke. Through subcellular localization, we found that the green fluorescence signal was mainly distributed in the cell membrane and cytoplasm, indicating that HvnF3'H mainly plays a role in the cell membrane and cytoplasm.

Functional annotation was performed on the coding genes of 5 candidate proteins, among which HORVU4Hr1G003060 encoded Lysine N-methyltransferase (KMT). Research has shown that colorless dihydroflavonols are formed under the catalysis of the *F3'H*, which are then catalyzed by dihydroflavonol reductase (DFR) and anthocyanin synthase (ANS) to form colored anthocyanin glycosides. Finally, different anthocyanin glycosides are catalyzed by methyltransferase (MT), glycosyltransferase (GT), and acyltransferase (AT) to form different anthocyanin glycosides [41]. HORVU7Hr1G100810 encodes 1-aminocyclopropane-1-carboxylate oxidase (ACO). Studies showed that the *ACO* had a promoting effect on the synthesis of ethylene compounds; the accumulation

of anthocyanins significantly increases after treated by ethylene [42]. An et al. found that ethylene responsive factors ERFs are involved in regulating plant Anthocyanidin synthesis [43]. HORVU6Hr1G093960 belonged to the *TCP* transcription factor family. Xu et al. found *TCP* might be involved in anthocyanin accumulation, and the *TCP* played an inhibitory role in anthocyanin synthesis in *camellia* [44]. We found *HvnF3'H* was an important structural gene in the early stage of purple qingke in anthocyanin synthesis [26]. Based on our study, *HvnF3'H* is not only a structural gene in the anthocyanin synthesis pathway in purple qingke, but the specific mechanism of action still needs further exploration.

Conclusions

Through the construction of a gene co-expression network, this study identified hub genes related to anthocyanin, most of which directly or indirectly affect the synthesis or accumulation of anthocyanin. HvnF3'H is located in the cell membrane and cytoplasm. Through Y2H and BiFC experiments, it was found that HvnF3'H interacts with various proteins related to anthocyanin synthesis and metabolism. The experimental results indicated that HvnF3'H regulated the synthesis of anthocyanin in purple-grained qingke through interactions with related proteins, but the specific regulatory mechanism requires further exploration. These results provide theoretical guidance for the in-depth analysis of the molecular mechanism and biological function of anthocyanin-related genes in purple qingke.

Abbreviations

DEGs	Differentially Expressed Genes
GO	Gene Ontology
KEGG	The Kyoto Encyclopedia of Genes and Genomes
WGCNA	Weighted Correlation Network Analysis
ME	Module Characteristic Gene
UV	Ultraviolet
F3'H	Dihydrofavonol-3'-Hydrogenase
LB	Lysogeny Broth
Y2H	Yeast Two-Hybrid
DMSO	Dimethyl Sulfoxide
SD	Synthetic Dropout
CDS	Coding Sequence
BiFC	Bimolecular Fluorescence Complementation

Supplementary Information

The online version contains supplementary material available at https://doi.org/10.1186/s12864-024-10738-9.

Supplementary Material 1

Supplementary Material 2

Supplementary Material 3

Supplementary Material 4

Supplementary Material 5

Supplementary Material 6

Supplementary Material 7

Supplementary Material 8

Acknowledgements

We thank LetPub (www.letpub.com) for linguistic assistance and pre-submission expert review.

Author contributions

L.C. contributed in experimental design, methodology, software, investigation, data analysis, and writing original draft. Y.Y. contributed in experimental design, software, data analysis and editing. Y.C. contributed in investigation, review and editing. X.L. contributed in software, investigation, and writing. L.A. contributed in software, data analysis, and writing. Y.B. contributed in software, supervision, data analysis, and writing. X.Y. and K.W. contributed in supervision, funding acquisition, writing, review and editing. All authors read and approved the final manuscript.

Funding

The Construction Project for Innovation Platform of Qinghai Province (2023_1_5), this research was the National Key Research and Development Project (2022YFD2301300), and the Agriculture Research System of China (CARS-05).

Data availability

Qingke reference genome sequence(https://www.ncbi.nlm.nih.gov/datasets/genome/GCA_916098225.1/); Large-scale molecular datasets generated by genome sequencing and other high-throughput experimental technologies (http://www.genome.jp/kegg/).

Declarations

Ethics approval and consent to participate

The seeds of Nierumuzha and Kunlun10 used in this study were preserved by Qinghai Academy of Agriculture and Forestry Sciences. The collection and use of these seeds, including experimental research and field studies, were performed completely in accordance with relevant institutional, national, and international guidelines and legislation.

Consent for publication

Not applicable.

Competing interests

The authors declare no competing interests.

Received: 11 January 2024 / Accepted: 23 August 2024
Published online: 02 September 2024

References

1. Sosulski F, Krygier K, Hogge L. Free, esterified, and insoluble-bound phenolic acids. 3. Composition of phenolic acids in cercal and potato flours. J Agric Food Chem. 1982;30:337–40.
2. Deng J, et al. Effect of highland barley-glucan on starch digestibility in vitro. Shipin Kexue/Food Sci. 2018;39:106–11.
3. Obadi M, Sun J, Xu B. Highland barley: chemical composition, bioactive compounds, health effects, and applications. Food Res Int. 2021;140:110065.
4. Wang CP, et al. Starch granule-associated proteins of hull-less barley (Hordeum vulgare L.) from the Qinghai-Tibet Plateau in China. J Sci Food Agric. 2011;91:616–24.
5. Gao WL, Gong LX, Zhang Y. The development potential of tibetan hull-less barley as China plateau characteristic grain resource. J Cereals Oils. 2015;28:1–4.
6. Chaves-Silva S, et al. Understanding the genetic regulation of anthocyanin biosynthesis in plants-tools for breeding purple varieties of fruits and vegetables. Phytochemistry. 2018;153:11–27.
7. Jende-Strid B. Genetic control of flavonoid biosynthesis in barley. Hereditas. 1993;119:187–204.

8. Wu YQ, et al. Research progress in understanding the biosynthesis and regulation of plant anthocyanins. Sci Hortic-Amsterdam. 2023;321:112374.
9. Chen R, et al. Transcription factor SmSPL7 promotes anthocyanin accumulation and negatively regulates phenolic acid biosynthesis in Salvia miltiorrhiza. Plant Sci. 2021;310:110993.
10. Liu BL, Zhu YC, Zhang TZ. The R3-MYB gene GhCPC negatively regulates cotton fiber elongation. PLoS ONE. 2015;10:e0116272.
11. Hichri I, et al. The basic helix-loop-helix transcription factor MYC1 is involved in the regulation of the flavonoid biosynthesis pathway in grapevine. Mol Plant. 2010;3:509–23.
12. Xu WJ, Dubos C, Lepiniec L. Transcriptional control of flavonoid biosynthesis by MYB-bHLH-WDR complexes. Trends Plant Sci. 2015;20:176–85.
13. Yu JX, et al. A long noncoding RNA functions in high-light-induced anthocyanin accumulation in apple by activating ethylene synthesis. Plant Physiol. 2022;189:66–83.
14. Nath P, et al. Ripening of fleshy fruit: molecular insight and the role of ethylene. Biotechnol Adv. 2010;28:94–107.
15. Honda C, Moriya S. Anthocyanin biosynthesis in apple fruit. Horticult J. 2018;87:305–14.
16. Carbone, et al. Developmental, genetic and environmental factors affect the expression of flavonoid genes, enzymes and metabolites in strawberry fruits. Plant Cell Environ. 2009;32:1117–31.
17. Jenkins GI, et al. Plant responses to UV and blue light: biochemical and genetic approaches. Plant Sci. 1995;112:117–38.
18. Ni JB, et al. Ethylene mediates the branching of the jasmonate-induced flavonoid biosynthesis pathway by suppressing anthocyanin biosynthesis in red Chinese pear fruits. Plant Biotechnol J. 2019;18:1223–40.
19. Takos AM, et al. Light-induced expression of a MYB gene regulates anthocyanin biosynthesis in red apples. Plant Physiol. 2006;142:1216–32.
20. Tian J, et al. Characteristics of dihydroflavonol 4-reductase gene promoters from different leaf colored Malus crabapple cultivars. Hortic Res. 2017;4:17070.
21. Ban Y, et al. Isolation and functional analysis of a MYB transcription factor gene that is a key regulator for the development of red coloration in apple skin. Plant Cell Physiol. 2007;48:958–70.
22. Dong TT, et al. Anthocyanins accumulation and molecular analysis of correlated genes by metabolome and transcriptome in green and purple asparaguses (Asparagus officinalis, L). Food Chem. 2018;271:18–28.
23. Yang B, et al. Transcriptomics integrated with metabolomics reveals the effect of regulated deficit irrigation on anthocyanin biosynthesis in Cabernet Sauvignon grape berries. Food Chem. 2020;314:11–3.
24. Zhuang HM, et al. Differential regulation of anthocyanins in green and purple turnips revealed by combined de novo transcriptome and metabolome analysis. Int J Mol Sci. 2019;20:4387.
25. Chen L, et al. Genome-wide identification of WD40 transcription factors and their regulation of the MYB-bHLH-WD40 (MBW) complex related to anthocyanin synthesis in Qingke (Hordeum vulgare L. var. Nudum hook. F). BMC Genomics. 2023;24:166.
26. Yao XH, et al. Accumulation and regulation of anthocyanins in white and purple tibetan Hulless Barley (Hordeum vulgare L. var. Nudum hook. f.) revealed by combined de novo transcriptomics and metabolomics. BMC Plant Bio. 2022;22:391.
27. Faccioli P, et al. From single genes to co-expression networks: extracting knowledge from barley functional genomics. Plant Mol Bio. 2005;58:739–50.
28. Goto T, Kondo T. Structure and molecular stacking of anthocyanins flower color variation. Angew Chem Int Ed. 1991;30:17–33.
29. Bogs J, et al. Identification of the flavonoid hydroxylases from grapevine and their regulation during fruit development. Plant Physiol. 2006;140(1):279–91.
30. Kim J, et al. High accumulation of anthocyanins via the ectopic expression of AtDFR confers significant salt stress tolerance in Brassica napus L. Plant Cell Rep. 2017;36:1215–24.
31. Kobayashi K, et al. MYB3Rs, plant homologs of myb oncoproteins, control cell cycle-regulated transcription and form DREAM-like complexes. Transcription. 2015;6:106–11.
32. Chao N, et al. Functional characterization of two chalcone isomerase (CHI) revealing their responsibility for anthocyanins accumulation in mulberry. Plant Physiol Bioch. 2021;161:65–73.
33. Nishihara M, Nakatsuka T, Yamamura S. Flavonoid components and flower color change in transgenic tobacco plants by suppression of chalcone isomerase gene. FEBS Lett. 2005;579:6074–8.
34. Marrs KA, Alfenito MR, Lloyd AM, Walbot V. A glutathione S-transferase involved in vacuolar transfer encoded by the maize gene Bronze-2. Nature. 1995;375:397–400.
35. An JP, et al. An apple MYB transcription factor regulates cold tolerance and anthocyanin accumulation and undergoes MIEL1mediated degradation. Plant Biotechnol J. 2019;18:337–53.
36. Song XH, et al. Molecular and metabolic insights into anthocyanin biosynthesis during leaf coloration in autumn. Environ Exp Bot. 2021;190:104584.
37. Sun RZ, Pan QH, Duan CQ, Wang J. Light response and potential interacting proteins of a grape flavonoid 3'-hydroxylase gene promoter. Plant Physiol Biochem. 2015;97:70–81.
38. Baba SA, Ashraf N. Functional characterization of flavonoid 3'-hydroxylase, CsF3'H, from Crocus sativus L: insights into substrate specificity and role in abiotic stress. Arch Biochem Biophys. 2019;667:70–8.
39. Brugliera F, Barri-Rewell G, Holton TA, Mason JG. Isolation and characterization of a flavonoid 3'-hydroxylase cDNA clone corresponding to the Ht1 locus of Petunia hybrida. Plant J. 1999;19(4):441–51.
40. Castellarin SD, et al. Colour variation in red grapevines (Vitis vinifera L.): genomic organisation, expression of flavonoid 3'-hydroxylase, flavonoid 3', 5'-hydroxylase genes and related metabolite profiling of red cyanidin-/blue delphinidin-based anthocyanins in berry skin. BMC Genomics. 2006;7:1–17.
41. Zhao D, Tao J. Recent advances on the development and regulation of flower color in ornamental plants. Front Plant Sci. 2015;261.
42. Barba-Espin G, et al. Ethephon-induced changes in antioxidants and phenolic compounds in anthocyanin-producing black carrot hairy root cultures. J Exp Bot. 2020;71:7030–45.
43. An JP, et al. The ERF transcription factor MdERF38 promotes drought stress-induced anthocyanin biosynthesis in apple. Plant J. 2020;101:573–89.
44. Fu M, et al. Unraveling the regulatory mechanism of color diversity in Camellia japonica petals by integrative transcriptome and metabolome analysis. Front Plant Sci. 2021;12:685136.

Publisher's note