A peer-reviewed version of this preprint was published in PeerJ on 27 June 2018.

View the peer-reviewed version (peerj.com/articles/4976), which is the preferred citable publication unless you specifically need to cite this preprint.

Hu P, Li G, Zhao X, Zhao F, Li L, Zhou H. 2018. Transcriptome profiling by RNA-Seq reveals differentially expressed genes related to fruit development and ripening characteristics in strawberries (*Fragaria* × *ananassa*) PeerJ 6:e4976 https://doi.org/10.7717/peerj.4976

PeerJ Preprints NOT PEER-REVIEWED

Transcriptome profiling by RNA-Seq reveals differentially expressed genes related to fruit development and ripening characteristics in strawberry (*Fragaria* × *ananassa*)

Panpan Hu [1], **Gang Li** [1], **Xia Zhao** [1], **Fengli Zhao** [1], **Liangjie Li** [1], **Houcheng Zhou** Corresp. [1]

[1] Zhengzhou Fruit Research Institute, Chinese Academy of Agricultural Sciences, Zhengzhou, Henan, China

Corresponding Author: Houcheng Zhou
Email address: zhouhoucheng@caas.cn

Strawberry (*Fragaria* × *ananassa*) is an ideal plant for fruit development and ripening research due to the rapid substantial changes in fruit color, aroma, taste and softening. To gain deeper insights into the genes that play a central regulatory role in strawberry fruit development and ripening characteristics, transcriptome profiling was performed for the large green fruit, white fruit, turning fruit, and red fruit stages of strawberry. A total of 6,608 differentially expressed genes (DEGs) with 2,643 up-regulated and 3,965 down-regulated genes were identified in the fruit development and ripening process. The DEGs related to fruit flavonoid biosynthesis, starch and sucrose biosynthesis, the citrate cycle, and cell-wall modification enzymes played important roles in the fruit development and ripening process. Particularly, some candidate genes related to the ubiquitin mediated proteolysis pathway and MADS-box were confirmed to be involved in fruit development and ripening according to their possible regulatory functions. Five *ubiquitin-conjugating enzymes* and ten *MADS-box transcription factors* were differentially expressed between the four fruit ripening stages. The expression levels of DEGs relating to color, aroma, taste, and softening of fruit were confirmed by quantitative real-time polymerase chain reaction. Our study provides important insights into the complicated regulatory mechanism underlying the fruit ripening characteristics in *Fragaria* × *ananassa*.

PeerJ Preprints NOT PEER-REVIEWED

Transcriptome profiling by RNA-Seq reveals differentially expressed genes related to fruit development and ripening characteristics in strawberry (*Fragaria* × *ananassa*)

Panpan Hu[1], Gang Li[1], Xia Zhao[1], Fengli Zhao[1], Liangjie Li[1], Houcheng Zhou[1*]

[1] Zhengzhou Fruit Research Institute, Chinese Academy of Agricultural Sciences, Zhengzhou, Henan, China

*Correspondence: Houcheng Zhou

Email: zhouhoucheng@caas.cn

Abstract

Strawberry (*Fragaria* × *ananassa*) is an ideal plant for fruit development and ripening research due to the rapid substantial changes in fruit color, aroma, taste and softening. To gain deeper insights into the genes that play a central regulatory role in strawberry fruit development and ripening characteristics, transcriptome profiling was performed for the large green fruit, white fruit, turning fruit, and red fruit stages of strawberry. A total of 6,608 differentially expressed genes (DEGs) with 2,643 up-regulated and 3,965 down-regulated genes were identified in the fruit development and ripening process. The DEGs related to fruit flavonoid biosynthesis, starch and sucrose biosynthesis, the citrate cycle, and cell-wall modification enzymes played important roles in the fruit development and ripening process. Particularly, some candidate genes related to the ubiquitin mediated proteolysis pathway and MADS-box were confirmed to be involved in fruit development and ripening according to their possible regulatory functions. Five *ubiquitin-conjugating enzymes* and ten *MADS-box transcription factors* were differentially expressed between the four fruit ripening stages. The expression levels of DEGs relating to color, aroma, taste, and softening of fruit were confirmed by quantitative real-time polymerase chain reaction. Our study provides important insights into the complicated regulatory mechanism underlying the fruit ripening characteristics in *Fragaria* × *ananassa*.

PeerJ Preprints NOT PEER-REVIEWED

Introduction

The octoploid strawberry (*Fragaria* × *ananassa*) is the dominant cultivatedspecie of its high yield and nutritional value, including vitamin C, sugar and organic acid, and anthocyanin contents (*Tanaka et al.*, *2008*; *Giampieri et al.*, *2012*; *Chen et al.*, *2016b*). The strawberry fruit development and ripening process involves intricate metabolic event and is divided into four distinct phases: the green fruit, white fruit, turning fruit and red fruit stages (*Fait et al.*, *2008*). In the green fruit stage, fruits undergo cell division and cell expansion. In the white fruit stage, fruit growth is nearly complete, and fruits begin to enter the maturation process. Subsequently, fruit development enters the turning fruit stage, as indicated by slight coloration. During the red fruit stage, the characteristics of ripening such as color, aroma, taste and softening, increase rapidly along with a massive accumulation of pigments, amino acids, and organic acids, among other compounds. In addition, strawberry is an ideal model plant for studying the fruit development and ripening process in non-climacteric fruit (*Giovannoni*, *2004*; *Zhang et al.*, *2011*).

Following the sequencing of the genome of diploid woodland strawberry (*Fragaria vesca*) (*Shulaev et al.*, *2011*; *Edger et al.*, *2018*), the sequence of the octoploid cultivated strawberry (*Fragaria* × *ananassa*) was also completed (*Hirakawa et al.*, *2014*). However, the sequence information of genes published on *Fragaria* × *ananassa* is insufficient and cannot be wholly used as an available reference genome for studies of the octoploid strawberry in at the molecular level.

The use of transcriptome sequencing technology for gene detection and markers in different strawberry tissues and in response to various environmental stresses has increased (*Li et al.*, *2013*; *Hollender et al.*, *2014*; *Chen et al.*, *2016a*; *Wang et al.*, *2017a*). To date, RNA-Seq has also been widely used to study gene expression in the strawberry fruit development and ripening process. The genome-scale transcriptomic analysis of hormone signaling in early strawberry fruit developmental stages from floral anthesis to enlarged fruit suggests that the biosynthesis genes for indole-3-acetic acid (IAA) and gibberellin (GA) are most highly and specifically expressed in

PeerJ Preprints NOT PEER-REVIEWED

endosperm and seed coat and play a most prominent role for fruit set (*Kang et al.*, *2013*). *FaTCP11*, *FaPCL1-like* and *FaSCL8* modulate the metabolism of strawberry flavonoids by regulating the expression of flavonoid pathway genes based on a transcriptome correlation network analysis of ripe strawberry fruits (*Pillet et al.*, *2015*). Another application of transcriptome analysis in the strawberry anthocyanin biosynthesis pathway reveals that exogenous hematin promotes fruit coloring through multiple related metabolic pathways including anthocyanin biosynthesis and hormone signaling transduction, among others (*Li et al.*, *2016*). In red-fruited and natural white-fruited strawberry varieties, transcriptome analysis showed that the genes related to the polyphenol biosynthesis pathway may interact with anthocyanin biosynthesis, flavor formation and fruit softening to regulate the fruit ripening process (*Hartl et al.*, *2017*). For postharvest strawberry fruit, transcriptome profiling showed that exogenous IAA delays the fruit ripening process, whereas abscisic acid (ABA) promotes the postharvest ripening by regulating the expressions of genes related to receptor-like kinases, ubiquitin ligases, and IAA and ABA hormone signaling pathways (*Chen et al.*, *2016b*). Transcriptomic analysis of strawberry endogenous IAA suggests that the candidate genes of *FaTAA1*, *FaTAR2*, *FaAux/IAA11* and *FaARF6a* are involved in active IAA biosynthesis in the strawberry ripe receptacle (*Estrada-Johnson et al.*, *2017*). RNA-Seq is also used to study the polymorphisms of the octoploid strawberry. According to transcriptional analyses of the *FaERF* family in ripening strawberry fruits, *FaERF3*, *FaERF6* and *FaERF71a* as candidates were identified to play a primary role in the ripening receptacle (*Sanchez-Sevilla et al.*, *2017*). A recent study suggests that the down regulation of the key gene *PDHE1α* of the pyruvate dehydrogenase for glycolysis derived oxidative phosphorylation inhibits respiration and ATP biosynthesis but promotes the accumulation of sugar, ABA, ethylene (ETH) and polyamines, and ultimately accelerates the strawberry fruit ripening (*Wang et al.*, *2017b*).

The functions of the ubiquitin mediated proteolysis pathway in the regulation of fruit ripening have been studied in banana, tomato, papaya and barbarum, and these studies confirm the regulatory role of the ubiquitin proteasome in the fruit ripening process (*Liu et al.*, *2013*;

PeerJ Preprints NOT PEER-REVIEWED

Wang et al., *2014*; *Bi et al.*, *2015*; *Zeng et al.*, *2015*). Research on the regulatory mechanism of fruit ripening indicates that MADS-box transcription factors have a pivotal effect on fruit ripening by regulating carotenoid synthesis, the ETH signaling pathway, cell wall metabolism, flavonoid and lignin biosynthesis, and cuticle development in apple, banana, tomato and peach (*Youssef et al.*, *2012*; *Ireland et al.*, *2013*; *Liu et al.*, *2013*; *Feng et al.*, *2016*; *Garceau et al.*, *2017*; *Yin et al.*, *2017*; *Hu et al.*, *2017*).

In this study, based on the characteristics changes in fruit development and ripening, a global expression analysis by RNA-Seq at four stages of strawberry fruit ripening was performed to discover additional candidate genes in ubiquitin mediated proteolysis for MADS-box transcription factors and for other aspects. In this paper, the different expression patterns of DEGs related to coloring, aroma, taste, softening, and other aspects among different fruit ripening stages in strawberry are outlined. The purpose of this study was to understand the molecular mechanisms controlling the characteristics of strawberry fruit ripening according to transcriptome profiling analysis and to provide a theoretical foundation for the cultivation of strawberry varieties.

Materials and methods

Plant materials

The fruits used in this study were obtained from the strawberry cultivar 'Toyonoka' cultivated in the greenhouse (8-28°C, relative humidity 55-70%, and without supplemental lighting) in Zhengzhou, Henan, China. Fruits of large green fruit (l-GF), white fruit (WF), turning fruit (TF), and red fruit (RF) stages were selected as the sequencing materials (Fig. 1). Ten uniformly sized fruits were sampled at each stage. For quantitative real-time polymerase chain reaction **(**qRT-PCR), the green fruit stage was subdivided into small green fruit (s-GF), middle green fruit (m-GF), and l-GF stages. In total, fruits of six different ripening stages (s-GF, m-GF, l-GF, WF, TF and RF) were prepared for qRT-PCR. Three uniform fruits were sampled at each of the six stages for RNA isolation and cDNA synthesis (three replicates). The experimental materials

PeerJ Preprints NOT PEER-REVIEWED

were placed immediately in liquid N_2 and stored at -80°C for RNA extraction.

Total RNA extraction, library preparation, and transcriptome sequencing

Total RNA was extracted using a Spin Column Plant total RNA Purification Kit (Order No. B518661; Sangon Biotech, China) according to the manufacturer's instructions. DNase digestion with Dnase I (Promega) was performed to remove contaminating DNA. Briefly, mRNA was purified from total RNA using poly-T oligo-attached magnetic beads (Novogene, Beijing, China) and then broken into short fragments. With these fragments as templates, cDNA were synthesized. To select cDNA fragments of 150 to 200 bp in length, the library fragments were purified with an AMPure XP system (Beckman Coulter, Beverly, USA). Then, those fragments were selected for PCR amplification as sequencing templates. The PCR products were purified and library quality was assessed on an Agilent Bioanalyzer 2100 system (Agilent Technologies, CA, USA). The clustering of the index-coded samples was performed on a cBot Cluster Generation System using a TruSeq PE Cluster Kit v3-cBot-HS (Illumina) according to the manufacturer's instructions. After cluster generation, the library preparations were sequenced on an Illumina HiSeq 4000 platform (Novogene, Beijing, China) and paired-end reads were generated. Each RNA sample was ligated with a separate adapter and sequenced together in a single run. The whole transcriptome file was composed of separate RNA sample data for analyzing gene expression. Transcriptome assembly was accomplished based on the transcripts and unigenes using Trinity (*Grabherr et al.*, *2011*) with min_kmer_cov set to 2 by default and all other default parameters set.

Data analysis

The raw image data files from the Illumina HiSeq 4000 were transformed into the original sequenced reads (raw reads) by CASAVA 1.8 base calling analysis and processed through in-house Perl scripts. A certain error rate is observed in the sequencing process. If e represents the sequencing error rate and Qphred represents the base quality, then Qphred = $-10\log_{10}$ (e). When

PeerJ Preprints NOT PEER-REVIEWED

Qphred = Q20, the correct recognition rate of a base reaches 99%. Clean data (clean reads) were obtained by eliminating the low-quality reads (reads containing an adapter, reads containing ploy-N, and reads with Qphred ≤ 20) from raw reads. All the downstream analyses were based on clean data with high quality.

Gene functions were annotated based on the following seven databases (Table S1): NCBI non-redundant protein sequences (Nr), NCBI non-redundant nucleotide sequences (Nt), Protein family (Pfam), EuKaryotic Orthologous Groups (KOG), a manually annotated and reviewed section of the UniProt Knowledgebase database (Swiss-Prot), KEGG Ortholog (KO), and Gene Ontology (GO). The URLs, annotation methods and parameters of the seven databases are shown in Table S1, and the information of all software versions and parameters is shown in Table S2.

Differentialy expression analysis

Clean reads of each library were compared with transcriptome reference sequences. Gene expression levels were evaluated by RNA-Seq by Expectation Maximization (RSEM) with the bowtie2 parameters (*Li & Dewey*, *2011*) for each sample (Table S2). The read_count for each gene was obtained from the mapping results of clean reads back onto the assembled transcriptome. The read_count of each gene was normalized data of the fragments per kilobaseof exon per million fragments mapped (FPKM) which is the most commonly used method of estimating gene expression levels (*Trapnell et al.*, *2010*). Those genes whose FPKM > 0.3 were considered to be expressed (Fig. S1, Table S3) (*Mortazavi et al*., *2008*; *Trapnell et al.*, *2010; Ho et al.*, *2012*). For those samples with biological replicates, differential expression of unigenes was analyzed and calculated based on the read_count value using the DESeq R package (*Anders & Huber*, *2010*). Based on the negative binomial distribution model, DESeq provided statistical routines for determining differential expression in digital gene expression data. The *p*-values in statistics were adjusted using Benjamini and Hochberg's approach for controlling the false discovery rate (*Benjamini & Hochberg*, *1995*). The thresholds for judging significant difference

PeerJ Preprints NOT PEER-REVIEWED

of gene expression level between any two groups were padj < 0.05 and $|\log_2$ (fold change)$| \geq 1$. The *p*-adjusted (padj) was the corrected *p*-value, and a small padj value of DEG indicated high significance of the differential expression.

qRT-PCR analysis

Total RNA extraction and reverse transcription PCR were performed as previously described for RNA extraction and library preparation of RNA-Seq. All qRT-PCR samples were run on a Light Cycler 480 system (Roche, Switzerland). Each reaction was performed with a total volume of 20 μL that contained 5 μL of first-strand cDNA as a template, with a pre-incubation program of 5 min at 95 °C, followed by 45 cycles of 10 s at 95 °C and 30 s at 60 °C, according to the Light Cycler 480 SYBR Green I Master protocol (Cat. No.04707516001). Gene-specific primers were designed with Primer Premier 5 (Table S4). The *FaACTIN* gene was used as an internal reference for gene expression. Gene expression levels were calculated using the $2^{-\Delta\Delta Ct}$ method (*Livak & Schmittgen*, *2001*). The mean threshold cycle values for each gene were obtained from three independent PCR reactions.

Results

RNA-Seq

A total of 45.48 G of data with two biological replicates of each library were generated in this study (Table 1). A total of 172,799 transcripts which spliced by Trinity (*Grabherr et al.*, *2011*), were assembled based on the raw reads with an average length of 951 bp. Then, 91,790 valid unigenes were obtained, with an average length of 714 bp. Figure S2 shows the length distributions of the transcripts and unigenes.

Functional annotation of unigenes

Of the total 91,790 unigenes, 57,200 unigenes were annotated to the seven databases (Table S5). Among all the databases, 40.53% of unigenes were aligned to the Nr protein database with an *e*-

PeerJ Preprints NOT PEER-REVIEWED

value threshold of e^{-5}. The similarity of gene sequence and the function information of genes between strawberry and other species were obtained through the Nr annotation database. The results of species classification, *e*-value distribution, and sequence similarity distribution are shown in Fig. S3A-C, respectively.

GO annotation results primarily describe gene functions. A total of 26,523 unigenes in the GO database were classified into 57 functional categories, among which 22,087 unigenes were assigned to biochemical processes, 10,259 genes were assigned to cellular components, and 16,418 unigenes were assigned to molecular functions (Table S6).

To evaluate the effectiveness of the annotation process and possible functions of unigenes, 13,442 unique sequences were noted on the KOG database, based on their ortholog relationship. KOG was segmented into 26 orthologous groups (Table S7). Among the 26 KOG groups, 2,263 and 1,808 unigenes were enriched to the 'general function prediction only' and 'post-translational modification, protein turnover, chaperones' clusters, respectively. Based on the same ortholog gene function in the KOG classification, we could effectively analyze the functions of DEGs in fruit ripening.

The KEGG database is available to systematically analyze the metabolic pathways and functions of gene products and compounds in cells by integrating the genome, molecular chemical and biochemical systems data. Annotated to the KEGG database, 10,932 unigenes were assigned to 274 KEGG pathways using BLASTx with an *e*-value threshold of e^{-10} (Table S8). KEGG results provided a good transcription platform for investigating the related metabolic pathways in the strawberry development and ripening process.

Analysis of differentially expressed genes in the fruit development and ripening process

In different comparative combinations, volcano plot (Fig. 2A-F, Table S9) can visually demonstrate the relationship among padj, $\log_2$ (fold change) and the number of up/down-regulated DEGs. A total of 6,608 DEGs with 2,643 up-regulated and 3,965 down-regulated, were differentially expressed in the six combinations (WF/l-GF, TF/l-GF, RF/l-GF, TF/WF, RF/WF

PeerJ Preprints NOT PEER-REVIEWED

and RF/TF). The number of up/down-regulated DEGs in each combination is displayed in Fig. 2A-F, which shows that the most DEGs were detected when comparing l-GF with RF and TF and WF with RF. The WF and TF libraries possessed similar gene expression patterns, and therefore the fewest DEGs were detected in the TF/WF combination. Of these DEGs, in each combination, the genes were predominantly down-regulated. For the different combinations, figure 3A-C shows the numbers of specific and common DEGs. In the comparison of l-GF with WF, TF, and RF, 785, 2,157 and 5,271 DEGs were identified, respectively (Fig. 3A). In the comparison of WF with TF and RF, 40 and 2,748 DEGs were identified (Fig. 3B). Compared TF with RF, 781 DEGs were identified (Fig. 3C). Subsequent analyses focused on these DEGs related to fruit development and ripening characteristics.

Enrichment pathway analysis of DEGs

The functional enrichment analyses of DEGs are based on the GO and KEGG databases. GO enrichment analysis of the DEGs was performed by the GOseq R packages (*Young et al.*, *2010*). KEGG (*Kanehisa et al.*, *2008*) enrichment analysis was used to test the statistical enrichment of DEGs with KOBAS software (*Mao et al.*, *2005*). GO and KEGG pathway enrichment analyses (padj < 0.05) were used to categorize the biological functions of DEGs. The expression patterns of the DEGs and their enrichment results in different combinations showed that the down-regulated expression of DEGs and metabolic pathways was predominant in the strawberry fruit development and ripening process.

Genes related to color, aroma, taste, and softening in the fruit development and ripening process

Research into non-climacteric fruit color is concentrated on flavonoid biosynthesis, and the types of anthocyanins in strawberry are pelargonidin, delphinidin, and cyanidin (Fig. 4). In this study, the expression of most genes in anthocyanin biosynthesis such as *chalcone synthase* (*CHS*) (c51804_g1, c78983_g2, and c98687_g1), *chalcone isomerase* (*CHI*) (c78027_g1), *naringenin*

PeerJ Preprints NOT PEER-REVIEWED

3-dioxygenase (*N3D*) (c71005_g1), *dihydroflavonol-4-reductase (DFR*) (c63190_g1, c64617_g1, and c69531_g1), and *leucoanthocyanidin dioxygenase* (*LAD*) (c70308_g1) were up-regulated with strawberry ripening (Fig. 4). However the down-regulated expression of *flavonoid 3'-monooxygenase* (*F3'M*) (c72378_g2) decreased the synthesis of cyanidin and accelerated the accumulation of pelargonidin in anthocyanin biosynthesis (Fig. 4). Cluster analysis was used to analyze 36 unigenes involved in flavonoid biosynthesis (Fig. S4A, Table S10). Among the 36 genes, the relative expression analysis, which revealed the expression patterns of genes over time, showed that the expression of 1 gene was up-regulated and that of 5 genes was down-regulated with fruit ripening (Fig. S4B, Table S11). Figure S4C showed the differential expression patterns of 10 DEGs in flavonoid biosynthesis (Table S11), and most DEGs played important roles in anthocyanin biosynthesis (Fig. 4).

The MYB-bHLH-WD40 transcription complex also regulates anthocyanin biosynthesis. Of the genes encoding MYB transcription factors in this data set, one up-regulated unigene was *R2R3 MYB transcription factor* (*FaMYB10*) (c76851_g2), which can positively control the biosynthesis of anthocyanin (*Lin-Wang et al.*, *2014*; *Medina-Puche et al.*, *2014*). Among the bHLH transcription factors, two down-regulated unigenes (c75633_g2 and c78773_g1) were annotated as bHLH33 and bHLH3, respectively, which can interact with MYB10 to play important roles in proanthocyanidin and anthocyanin biosynthesis (*Schaart et al.*, *2013*). Figure S5A/C shows the up-regulated and down-regulated expression of MYB and bHLH transcription factor genes (Table S11), and between of different combinations, 14 and 27 DEGs were found (Fig. S5B/D, Table S11). The two DEGs MYB113-like (c76114_g1) and R2R3 MYB transcription factor (c76851_g2) were up-regulated with strawberry ripening. The down-regulated DEGs of MYB transcription factors included MYB39-like, MYB12-like and MYB1R1-like (c67743_g1, c68086_g1 and c75011_g1). Three DEGs of transcription factors bHLH104-like, bHLH135-like and bHLH122-like (c71077_g1, c76460_g2, and c78358_g1, respectively) were up-regulated, although more DEGs were down-regulated including bHLH33 (c75633_g2). The expression pattern of these genes was closely related to fruit coloring. Ten

PeerJ Preprints NOT PEER-REVIEWED

WD40 repeat-containing protein genes were described in the RNA-Seq data, but none had significantly differential expression at the four fruit ripening stages.

DEGs of *CHS* (c78983_g2) and *DFR* (c63190_g1) involved in anthocyanin synthesis were verified by qRT-PCR (Fig. 5A/B). The results showed that the expression levels of genes were consistent with the results of transcriptome analysis (Fig. S6A/B): their expression levels increased and promoted the biosynthesis and accumulation of anthocyanin with fruit ripening.

Strawberry fruit is rich in characteristic aromatics in the later stages of fruit ripening. The primary aromatic compounds are derived from ester metabolism. The precursors of esters such as amino acids, sugars and lipids are converted to acids, alcohols, and aldehydes in ester biosynthesis (Fig. 6A). The decreased expression level of *alcohol dehydrogenase* (*FaADH*) (c60055_g1, c70375_g1, c70503_g2, c74014_g1, c78458_g1, c80660_g1, and c81069_g4) with strawberry fruit ripening (Fig. 6A) is consistent with previous research results from tests on peach *PpADH1*, *PpADH2* and *PpADH3* (*Zhang et al.*, *2010*). The expression level of *alcohol acyltransferase* (*FaAAT*) (c70507_g1) was significantly difference in the fruit ripening process, and the expression values in WF, TF and RF increased to 49.5, 174.5 and 380.8 times, respectively, than those in l-GF (Fig. S6C, Table S11). The qRT-PCR result for *FaAAT* showed a significant increase with fruit ripening (Fig. 5C); therefore, the *FaAAT* gene was considered to play a vital role in the metabolism of esters. To study the functions of additional genes on aromatics, the expression patterns of genes in the degradation of aromatic compound pathways was analyzed based on the transcriptome data (Fig. 6B/C, Table S11), and all those genes were down-regulated with fruit ripening.

Sugar and acidity are the primary components of fruit soluble solids governing fruit quality, which depend on starch and sucrose metabolism (Fig. 7) and citrate cycle metabolic pathways (Fig. 8), respectively. In the qRT-PCR, the up-regulated *SPS 1-like* (*FaSPS*) (c79838_g1) had the highest level in RF (Fig. 5D). The down-regulated *FaCES* (c75759_g1) decreased from l-GF to WF but increased from WF to RF (Fig. 5E). The up-regulated expression of *FaACC* (c77811_g1) in the qRT-PCR was consistent with the transcriptome expression pattern (Fig. 5F). The

298 expression patterns of *FaSPS*, *FaCES* and *FaACC* in transcriptome data were shown in figure
299 S6D-F. Confirming that the expression levels of most genes decreased with fruit ripening, figure
300 S7A-E shows the expression pattern of additional genes related to starch and sucrose metabolism.
301 The expression patterns of genes participating in the citrate cycle pathway are identified in Fig.
302 S7F/G, which shows that more genes were up-regulated with fruit ripening. More detailed
303 information on these genes is listed in Table S11. The expression of three *CS* (c74887_g1,
304 c78658_g1 and c78658_g3) and one *ACS* gene (c74238_g1) was up-regulated (Fig. S7F), which
305 indicated that the synthesis of citric acid increased with fruit ripening. The up-regulated
306 succinate dehydrogenase gene (c77175_g1) and down-regulated malate dehydrogenase gene
307 (c70484_g1) illustrated that the accumulation of malic acid increased with fruit ripening.
308 The research on strawberry fruit texture focuses on the cell wall modifying enzymes. In this
309 paper, two DEGs of *endoglucanase CX-like* (*Fa^{a}EG*) (c8256_g1) and *endoglucanase 24-like*
310 (*Fa^{b}EG*) (c66070_g2) were selected to verify their expression patterns in strawberry ripening
311 process. The results showed that the expression level of *Fa^{a}EG* and *Fa^{b}EG* was higher in the TF
312 and WT (Fig. 5G/H), which was not inconsistent with the expression pattern in transcriptome
313 data (Fig. S6G/H). So these two genes cannot be used to study the softening of strawberry fruit.
314
315 **Genes involved in ubiquitin mediated proteolysis associated with the fruit development and**
316 **ripening process**
317 Ubiquitin-activating enzyme (E1), ubiquitin-conjugating enzyme (E2), and ubiquitin-protein
318 ligase (E3) are the three major enzymes in ubiquitin mediated proteolysis. The specificity of
319 target proteins is determined by E2 and E3 in ubiquitin mediated proteolysis (*Schwechheimer &*
320 *Calderon Villalobos*, *2004*; *Stone & Callis*, *2007*; *Wang et al.*, *2014*). Only nine E1 proteins were
321 identified in this transcriptome data, and one E1 DEGs (c69468_g2) was only up-regulated in
322 RF/l-GF. Some E2 and E3 proteins were analyzed based on their expression pattern in the
323 transcriptome data (Fig. 9A-H, Tables S11). Ten and 6 E2 genes were up- and down-regulated,
324 respectively, with strawberry ripening. Among the DEGs annotated as E2, the expression of two

PeerJ Preprints NOT PEER-REVIEWED

DEGs (c65857_g1 and c69752_g1) was down-regulated in RF/l-GF and that of one DEG (c76267_g5) was down-regulated in RF/WF. The expression of two E2 DEGs (c69865_g1 and c80589_g1) were all up-regulated in RF/l-GF and RF/WF (Fig. 9B). The expression of 16 E3 genes was up-regulated (Fig. 9C, Table S11), and that of 23 genes decreased with fruit ripening. The differential expression analysis results for E3 showed that the expression of 3 DEGs (c67240_g1, c80832_g1 and c68571_g1) decreased and that of one DEG (c37206_g1) increased in TF/l-GF. In the RF/l-GF combination, the expression of 6 DEGs (c67240_g1, c77964_g1, c68571_g1, c77964_g1, c70427_g1 and c79627_g3) decreased and that of 3 DEGs (c73766_g1, c80901_g1 and c81107_g2) increased with fruit ripening. The expression of two E3 DEGs (c63405_g1 and c68571_g1) decreased in the TF/WF combination. The expression of an E3 DEG (c73766_g1) increased in both RF/WF and RF/TF and that of two DEGs (c70427_g1 and c77964_g1) decreased in RF/WF and RF/TF, respectively (Table S11). Based on the above results, the expression quantity of E2 DEGs in the later stage (RF) was significantly different from that of the early stages (l-GF and WF), and no E2 DEG was identified in any other combination. The down-regulated and up-regulated DEGs of E2 and E3 were possibly closely related to the fruit ripening process.

The expression patterns of MADS-box transcription factors were studied (Fig. 9I/J, Tables S11). The transcriptional level of most MADS-box transcription factors was down-regulated with fruit ripening. The expression pattern analysis showed that 3 and 14 MADS-box transcription factors increased and decreased with fruit ripening, respectively. Among DEGs of MADS-box transcription factors in each combination, the expression of 5 DEGs (c70741_g4, c72369_g2, c71360_g1, c69175_g1 and c77683_g2) was down-regulated and that of 3 DEGs was up-regulated in the RF/l-GF comparison (Table S11). In the RF/WF comparison, the expression of 3 DEGs (c77683_g2, c71360_g3 and c69157_g2) was down-regulated and that of 2 DEGs (c70335_g1 and c62694_g1) was up-regulated with fruit ripening. In the TF/l-GF comparison, the expression of one DEG (c69157_g2) was down-regulated and that of 2 DEGs (c70335_g1 and c66335_g1) was up-regulated (Table S11). The expression of 2 MADS-box

PeerJ Preprints NOT PEER-REVIEWED

DEGs (c62694 and c66335) was up-regulated in the RF/TF and WF/l-GF comparisons (Table S11). In terms of the above results, more DEGs were found in RF/l-GF and RF/WF than in other comparisons, thus the MADS-box transcription factor DEGs were related to fruit ripening to some extent. According to the known functions of MADS-box transcription factors in fruit ripening, further study of MADS-box transcription factors might lead to a new discovery pertinent to the regulation of fruit ripening.

The expression patterns of some *E2*, *E3* and *MADS-box* genes were analyzed by qRT-PCR. The results showed that the expression levels of *E2* and *^{a}E3* were the highest at the TF stage of strawberry and that of *^{b}E3* was up-regulated with fruit ripening (Fig. 5I-N). The expression patterns of *Fa^{a}EG*, *Fa^{a}E2*, *Fa^{b}E2*, and *Fa^{b}E3* were similar, suggesting that they might have the same function in the fruit ripening process. The expression of *FaMADS-box* decreased significantly with fruit ripening, as shown in Fig. 5O. Combining gene expression patterns in the transcriptome data (Fig. S6I-O), More work is required to discover and verify the regulatory mechanisms and functions of E2, E3 and MADS-box transcription factors in the fruit development and ripening process.

Discussion

In previous studies, RNA-Seq technology has been used to study fruit development and ripening (*Kang et al.*, *2013*; *Pillet et al.*, *2015*; *Li et al.*, *2016*; *Estrada-Johnson et al.*, *2017*; *Hartl et al.*, *2017*). In this study, 91,790 unigenes were obtained, and 56,606 unigenes were annotated to known proteins in the Swiss-Prot or Nr database. In addition, 6,608 DEGs were identified to analyze the changes in fruit characteristics with strawberry development and ripening. When our transcriptome data were compared with the transcriptome assembly results of octoploid strawberry in a previous study (*Sanchez-Sevilla et al.*, *2017*), fewer clean reads were retained in this data set (Table 1) due to the sequencing technology at that time and the experimental design, but more genes with FPKM $\geq$ 0.3 were identified than in the previous study (*Sanchez-Sevilla et al.*, *2017*) (Table S3). The unigenes with FPKM > 0.3 were considered to be expressed

PeerJ Preprints NOT PEER-REVIEWED

(*Mortazavi et al.*, *2008*; *Trapnell et al.*, *2010; Ho et al.*, *2012*), which were the satisfactory reference sequences for enrichment and differential expression analyses of DEGs.

The most intuitive indicator of strawberry ripening is the coloring. The synthesis mechanism of anthocyanins derived from the plant secondary metabolite pathway of flavonoid biosynthesis, has been extensively studied in strawberry (*Manning*, *1998*; *Castellarin & Di Gaspero*, *2007*; *Niu et al.*, *2010*). The high expression of genes such as *CHS*, *CHI*, *F3H*, and *DFR* increases the accumulation of anthocyanin content with fruit ripening (*Almeida et al.*, *2007*; *Salvatierra et al.*, *2010*; *Jiang et al.*, *2012*; *Zhang et al.*, *2015*; *Hartl et al.*, *2017*). Except for the down-regulated *F3'M* (c72378_g2), which accelerated the accumulation of pelargonidin, the other up-regulated DEGs in anthocyanin biosynthesis promoted fruit coloring and ripening (Fig. 4). The MYB-bHLH-WD40 transcription factors complex regulates the biosynthesis of anthocyanins (*Schwinn et al.*, *2006*; *Allan et al.*, *2008*; *Hichri et al.*, *2011*; *Schaart et al.*, *2013*). *FaMYB10* plays a positive regulatory role in the flavonoid/phenylpropanoid pathway (*Lin-Wang et al.*, *2014*; *Medina-Puche et al.*, *2014*). *FaMYB1* is described as a transcriptional repressor and represses the biosynthesis of anthocyanins in strawberry (*Aharoni et al.*, *2001*). Among the transcription factors of bHLH and WD40, FabHLH33, FabHLH3 and FaTTG1 transcription factors interact with the MYB transcription factors to play important roles in proanthocyanidin and anthocyanin biosynthesis (*Schaart et al.*, *2013*). In this study, the up-regulated expression of R2R3 MYB transcription factor (MYB10) (c76851_g2) was positively correlated with its function in anthocyanin biosynthesis. Consistent with a negative regulatory function in anthocyanin biosynthesis (*Aharoni et al.*, *2001*), FaMYB1R1 (c75011_g1) was down-regulated. The expression of bHLH33 (c75633_g2) and bHLH3-like (c78773_g1) was down-regulated with fruit ripening. No difference was detected in expression of WD40 during the four fruit ripening stages. The function of those genes related to anthocyanin biosynthesis requires future verification. Strawberry fruits release a special fragrance in the ripening process. AAT participates in the synthesis of strawberry fruit aroma because of the maximum gene expression and increasing activity throughout the ripening process (*Perez et al.*, *1996*; *Cumplido-Laso et al.*,

PeerJ Preprints NOT PEER-REVIEWED

406 *2012*). The expression pattern of AAT was significantly up-regulated with fruit ripening, based
407 on qRT-PCR, which was consistent with the expression pattern in this transcriptome data set and
408 that of the previous study.

409 Little is known of the functional mechanism of ubiquitin mediated proteolysis in strawberry
410 fruit ripening. In a previous study, *MuUBA*, the ubiquitin-activating enzyme E1 gene, and
411 *MuMADS1* showed high expression in the 4 ovule stage, and the expression levels were
412 stimulated by exogenous ETH and suppressed by 1-methylcyclopropene in banana. These results
413 indicated that the interaction of MuMADS1 and MuUBA might play an important role in post-
414 harvest banana fruit ripening (*Liu et al.*, *2013*). In the tomato *rin* mutant, SlUBC32 encodes an
415 E2 ubiquitin-conjugating enzyme and five E2s as direct targets of RIN were identified, which
416 uncovered a novel regulatory function of proteins in ubiquitin mediated proteolysis in fruit
417 ripening (*Wang et al.*, *2014*). Based on the above findings, 34 putative *CpUBC* genes are
418 identified in the papaya genome. The expression patterns of these genes showed the expression
419 level of 13 *CpUBC* genes increased at one ripening stage and that of 2 *CpUBC* genes decreased
420 at two ripening stages, which indicated the possible regulatory function of E2s in papaya fruit
421 ripening (*Jue et al.*, *2017*). Additionally, ubiquitin mediated proteolysis participates in fruit
422 ripening found based on microRNA analysis (*Bi et al.*, *2015*; *Zeng et al.*, *2015*). In this study, the
423 analysis of E2 DEGs in different comparative combinations of fruit ripening stages showed that
424 the expression levels of E2 28-like and E2 4-like decreased from l-GF to RF. The expression of
425 E2 5-like and E2 23-like increased from l-GF to RF (Table S11). The expression of E3 DEGs of
426 S-phase kinase-associated protein 1 and ubiquitin-protein ligase TRIP12 increased and that of the
427 other E3 DEGs decreased with fruit ripening (Table S11). The expression patterns of E2 and E3
428 DEGs in qRT-PCR were not consistent with those in the transcriptome data. Based on the
429 differentially expressed patterns of these genes in the transcriptome data, their functional
430 mechanisms in regulating fruit ripening require in-depth research.

431 The texture of strawberry fruit changes significantly changes with fruit ripening. The
432 regulatory factors that regulate the synthesis of enzymes related to fruit softening, play important

PeerJ Preprints NOT PEER-REVIEWED

roles in fruit ripening (*Youssef et al.*, *2012*). MADS-box transcription factors are key elements of the genetic networks that control flower and fruit development, and currently, a pivotal regulatory effect of these transcription factors on fruit ripening is widely repoeted. Recently, *MdMADS1* was found to inhibit fruit coloration and regulate apple fruit ripening (*Ireland et al.*, *2013*; *Feng et al.*, *2016*). *TAGL1*, a MADS-box transcription factor gene, controls several aspects of tomato fruit ripening by regulating carotenoid synthesis, ETH signaling pathway, cell cycle regulation, flavonoid and lignin biosynthesis, and cuticle development (*Garceau et al.*, *2017*). The suppressed expression of *SlMBP8*, a MADS-box gene, promotes carotenoid and ETH biosynthesis and induces the expressions of cell wall metabolism genes, which ultimately accelerate tomato fruit ripening (*Yin et al.*, *2017*). The MADS-box genes of *MaMADS24* and *MaMADS49* regulate the fruit development and ripening process by interacting with MaMADS proteins themselves and ETH signal transduction, biosynthesis-related proteins, starch biosynthesis proteins and metabolism-related proteins (*Liu et al.*, *2013*; *Hu et al.*, *2017*). The *PrupeSEP1* gene, a subfamily of MADS-box transcription factors, regulates fruit ripening and softening by exhibiting similar expression patterns of cell wall modification-related genes and N-glycan processing genes in melting flesh peach (*Li et al.*, *2017a*). RIN and MC, two truncated transcription factors, fuse with one another because of a genomic DNA deletion in the *rin* mutant. The over-expression and down-regulated expression patterns of RIN-MC in tomato transgenic wild-type and in the *rin* mutant, respectively, indicate a negative role of RIN-MC in fruit ripening (*Li et al.*, *2017b*). Transcriptome profiles analysis revealed that the silence of fruit-related gene *SEP1/2-like* (*FaMADS9*) leads to the inhibition of normal development and ripening in strawberry achenes (*Seymour et al.*, *2011*; *Qin et al.*, *2012*; *Wang et al.*, *2014*). In our study, the differential expression of MADS-box proteins SVP-like, ZMM17-like, CMB1-like, and MADS-box 17-like, among others which has not been reported in other studies, was identified in the strawberry fruit development and ripening process (Table S11).

RIN is a key regulator of the ripening gene expression network and has hundreds of target genes that can regulate changes in fruit color, flavor, texture and taste with tomato fruit ripening.

PeerJ Preprints NOT PEER-REVIEWED

Research on the *rin* mutants found that *RIN* encodes a MADS-box transcription factor that binds to the promoter of SlUBC32, an E2 ubiquitin-conjugating enzyme (*Wang et al.*, *2014*). Based on the relevance and possible regulatory role of E2 and MADS-box DEGs in the strawberry fruit ripening process, further work must be performed to verify the function and relationship between ubiquitin mediated proteolysis and MADS-box transcription factors in the fruit ripening process.

Conclusions

A transcriptome analysis identified many DEGs associated with fruit ripening characteristics. These DEGs were involved in multiple metabolic pathways of flavonoid biosynthesis, ester biosynthesis, starch and sucrose biosynthesis, the citrate cycle, MADS-box transcription factors, and the ubiquitin mediated proteolysis pathway, among others, in the fruit ripening process. The functional analysis and expression patterns of DEGs related to fruit development and ripening characteristics lay the foundation for the development of molecular markers in the cultivation of new strawberry varieties. The results of this study will also contribute to strawberry molecular breeding.

ADDITIONAL INFORMATION AND DECLARATIONS

Availability of data and materials

The Illumina reads have been deposited in the Sequence Read Archive (SRA) database at NCBI (http://www.ncbi.nlm.nih.gov/sra) and are available under study accession number SRP 111905.

Acknowledgements

The authors are grateful for the comments of several anonymous reviewers on the manuscript.

REFERENCES

Aharoni A, De Vos CH, Wein M, Sun Z, Greco R, Kroon A, Mol JN, O'Connell AP. 2001. The strawberry *FaMYB1* transcription factor suppresses anthocyanin and flavonol accumulation in transgenic tobacco. *The Plant Journal* **28**:319-332 DOI：10.1104/pp/.112.208678.

Allan AC, Hellens RP, Laing WA. 2008. MYB transcription factors that colour our fruit. *Trends in Plant Science* **13**:99-102 DOI: 10.1016/j.tplants.2007.11.012.

Almeida JR, D'Amico E, Preuss A, Carbone F, de Vos CH, Deiml B, Mourgues F, Perrotta G, Fischer TC, Bovy AG, Martens S, Rosati C. 2007. Characterization of major enzymes and genes involved in flavonoid and proanthocyanidin biosynthesis during fruit development in strawberry (*Fragaria* × *ananassa*). *Archives of Biochemistry and Biophysics* **465**:61-71 DOI: 10.1016/j.abb.2007.04.040.

Anders S, Huber W. 2010. Differential expression analysis for sequence count data. *Genome Biology* **11**:R106 DOI: 10.1186/gb-2010-11-10-r106.

Benjamini Y, Hochberg Y. 1995. Controlling the false discovery rate - a practical and powerful approach to multiple testing. *Journal of the Royal Statistical Society. Series B (Methodological)* **57**:289-300 DOI: jstor.org/stable/2346101.

Bi F, Meng X, Ma C, Yi G. 2015. Identification of miRNAs involved in fruit ripening in Cavendish bananas by deep sequencing. *BMC Genomics* **16**:776 DOI: 10.1186/s12864-015-1995-1.

Castellarin SD, Di Gaspero G. 2007. Transcriptional control of anthocyanin biosynthetic genes in extreme phenotypes for berry pigmentation of naturally occurring grapevines. *BMC Plant Biology* **7**:46 DOI: 10.1186/1471-2229-7-46.

Chen J, Zhang H, Feng M, Zuo D, Hu Y, Jiang T. 2016a. Transcriptome analysis of woodland strawberry (*Fragaria vesca*) response to the infection by *Strawberry vein banding virus* (SVBV). *Virology Journal* **13**:128 DOI: 10.1186/s12985-016-0584-5.

Chen JX, Mao LC, Lu WJ, Ying TJ, Luo ZS. 2016b. Transcriptome profiling of postharvest strawberry fruit in response to exogenous auxin and abscisic acid. *Planta* **243**:183-197

DOI: 10.1007/s00425-015-2402-5.

Cumplido-Laso G, Medina-Puche L, Moyano E, Hoffmann T, Sinz Q, Ring L, Studart-Wittkowski C, Caballero JL, Schwab W, Munoz-Blanco J, Blanco-Portales R. 2012. The fruit ripening-related gene *FaAAT2* encodes an acyl transferase involved in strawberry aroma biogenesis. *Journal of Experimental Botany* **63**:4275-4290 DOI: 10.1093/jxb/ers120.

Edger PP, VanBuren R, Colle M, Poorten TJ, Wai CM, Niederhuth CE, Alger EI, Ou S, Acharya CB, Wang J, Callow P, McKain MR, Shi J, Collier C, Xiong Z, Mower JP, Slovin JP, Hytonen T, Jiang N, Childs KL, Knapp SJ. 2018. Single-molecule sequencing and optical mapping yields an improved genome of woodland strawberry (*Fragaria vesca*) with chromosome-scale contiguity. *Gigascience* **7**:1-7 DOI: 10.1093/gigascience/gix124.

Estrada-Johnson E, Csukasi F, Pizarro CM, Vallarino JG, Kiryakova Y, Vioque A, Brumos J, Medina-Escobar N, Botella MA, Alonso JM, Fernie AR, Sanchez-Sevilla JF, Osorio S, Valpuesta V. 2017. Transcriptomic analysis in strawberry fruits reveals active auxin biosynthesis and signaling in the ripe receptacle. *Frontiers in Plant Science* **8**:889 DOI: 10.3389/fpls.2017.00889.

Fait A, Hanhineva K, Beleggia R, Dai N, Rogachev I, Nikiforova VJ, Fernie AR, Aharoni A. 2008. Reconfiguration of the achene and receptacle metabolic networks during strawberry fruit development. *Plant Physiology* **148**:730-750 DOI: 10.1104/pp.108.120691.

Feng X, An Y, Zheng J, Sun M, Wang L. 2016. Proteomics and SSH analyses of ALA-promoted fruit coloration and evidence for the involvement of a MADS-Box gene, *MdMADS1*. *Frontiers in Plant Science* **7**:1615 DOI: 10.3389/fpls.2016.01615.

Garceau DC, Batson MK, Pan IL. 2017. Variations on a theme in fruit development: the *PLE* lineage of MADS-box genes in tomato (*TAGL1*) and other species. *Planta* **246**:313-321 DOI: 10.1007/s00425-017-2725-5.

PeerJ Preprints NOT PEER-REVIEWED

Giampieri F, Tulipani S, Alvarez-Suarez JM, Quiles JL, Mezzetti B, Battino M. 2012. The strawberry: composition, nutritional quality, and impact on human health. *Nutrition* **28**:9-19 DOI: 10.1016/j.nut.2011.08.009.

Giovannoni JJ. 2004. Genetic regulation of fruit development and ripening. *The Plant Cell* **16 Suppl**:S170-180 DOI: 10.1105/tpc.019158.

Grabherr MG, Haas BJ, Yassour M, Levin JZ, Thompson DA, Amit I, Adiconis X, Fan L, Raychowdhury R, Zeng Q, Chen Z, Mauceli E, Hacohen N, Gnirke A, Rhind N, di Palma F, Birren BW, Nusbaum C, Lindblad-Toh K, Friedman N, Regev A. 2011. Full-length transcriptome assembly from RNA-Seq data without a reference genome. *Nature Biotechnology* **29**:644-652 DOI: 10.1038/nbt.1883.

Hartl K, Denton A, Franz-Oberdorf K, Hoffmann T, Spornraft M, Usadel B, Schwab W. 2017. Early metabolic and transcriptional variations in fruit of natural white-fruited *Fragaria vesca* genotypes. *Scientific Reports* **7**:45113 DOI: 10.1038/srep45113.

Hichri I, Barrieu F, Bogs J, Kappel C, Delrot S, Lauvergeat V. 2011. Recent advances in the transcriptional regulation of the flavonoid biosynthetic pathway. *Journal of Experimental Botany* **62**:2465-2483 DOI: 10.1093/jxb/erq442.

Hirakawa H, Shirasawa K, Kosugi S, Tashiro K, Nakayama S, Yamada M, Kohara M, Watanabe A, Kishida Y, Fujishiro T, Tsuruoka H, Minami C, Sasamoto S, Kato M, Nanri K, Komaki A, Yanagi T, Guoxin Q, Maeda F, Ishikawa M, Kuhara S, Sato S, Tabata S, Isobe SN. 2014. Dissection of the octoploid strawberry genome by deep sequencing of the genomes of *Fragaria species*. *DNA research* **21**:169-181 DOI: 10.1093/dnares/dst049.

Hollender CA, Kang C, Darwish O, Geretz A, Matthews BF, Slovin J, Alkharouf N, Liu Z. 2014. Floral transcriptomes in woodland strawberry uncover developing receptacle and anther gene networks. *Plant Physiology* **165**:1062-1075 DOI: 10.1104/pp.114.237529.

Ho DW, Yang ZF, Yi K, Lam CT, Ng MN, Yu WC, Lau J, Wan T, Wang X, Yan Z, Liu H, Zhang Y, Fan ST. 2012. Gene expression profiling of liver cancer stem cells by RNA-

sequencing. *Plos One* **7**:e37159 DOI: 10.1371/journal.pone.0037159.

Hu W, Yan Y, Shi HT, Liu JH, Miao HX, Tie WW, Ding ZH, Ding XP, Wu CL, Liu Y, Wang JS, Xu BY, Jin ZQ. 2017. The core regulatory network of the abscisic acid pathway in banana: genome-wide identification and expression analyses during development, ripening, and abiotic stress. *Bmc Plant Biology* **17**:145 DOI: 10.1186/S12870-017-1093-4.

Ireland HS, Yao JL, Tomes S, Sutherland PW, Nieuwenhuizen N, Gunaseelan K, Winz RA, David KM, Schaffer RJ. 2013. Apple *SEPALLATA1/2-like* genes control fruit flesh development and ripening. *The Plant Journal***73**:1044-1056 DOI: 10.1111/tpj.12094.

Jiang F, Wang JY, Jia HF, Jia WS, Wang HQ, Xiao M. 2012. RNAi-mediated silencing of the flavanone 3-hydroxylase gene and its effect on flavonoid biosynthesis in strawberry fruit. *Journal of Plant Growth Regulation* **32**:182-190 DOI: 10.1007/s00344-012-9289-1.

Jue D, Sang X, Shu B, Liu L, Wang Y, Jia Z, Zou Y, Shi S. 2017. Characterization and expression analysis of genes encoding ubiquitin conjugating domain-containing enzymes in *Carica papaya*. *PLOS ONE* **12:**e0171357 DOI: 10.1371/journal.pone.0171357.

Kanehisa M, Araki M, Goto S, Hattori M, Hirakawa M, Itoh M, Katayama T, Kawashima S, Okuda S, Tokimatsu T, Yamanishi Y. 2008. KEGG for linking genomes to life and the environment. *Nucleic Acids Research* **36**:D480-484 DOI: 10.1093/nar/gkm882.

Kang CY, Darwish O, Geretz A, Shahan R, Alkharouf N, Liu ZC. 2013. Genome-scale transcriptomic insights into early-stage fruit development in woodland strawberry *Fragaria vesca*. *The Plant Cell* **25**:1960-1978 DOI: 10.1105/tpc.113.111732.

Li B, Dewey CN. 2011. RSEM: accurate transcript quantification from RNA-Seq data with or without a reference genome. *Bmc Bioinformatics* **12**:323 DOI: 10.1186/1471-2105-12-323.

Li HE, Yao WJ, Fu YR, Li SK, Guo QQ. 2015. De novo assembly and discovery of genes that are involved in drought tolerance in Tibetan *Sophora moorcroftiana*. *PLOS ONE* **10**:e111054 DOI: 10.1371/journal.pone.0111054.

595 **Li J, Zhang QY, Gao ZH, Wang F, Duan K, Ye ZW, Gao QH. 2013.** Genome-wide
596 identification and comparative expression analysis of NBS-LRR-encoding genes upon
597 *Colletotrichum gloeosporioides* infection in two ecotypes of *Fragaria vesca*. *Gene*
598 **527**:215-227 DOI: 10.1016/j.gene.2013.06.008.
599 **Li JJ, Fang L, Qian M, Han MY, Liu HK, Zhang D, Ma JJ, Zhao CP. 2017a.** Characteristics
600 and regulatory pathway of the *PrupeSEP1 SEPALLATA* gene during ripening and
601 softening in peach fruits. *Plant Science* **257**:63-73 DOI: 10.1016/j.plantsci.2017.01.004.
602 **Li S, Xu H, Ju Z, Cao D, Zhu H, Fu D, Grierson D, Qin G, Luo Y, Zhu B. 2017b.** The RIN-
603 MC fusion of MADS-box transcription factors has transcriptional activity and modulates
604 expression of many ripening genes. *Plant Physiology* **176**:891-909 DOI:
605 10.1104/pp.17.01449.
606 **Li Y, Li HY, Wang FD, Li JJ, Zhang YH, Wang LJ, Gao JW. 2016.** Comparative
607 transcriptome analysis reveals effects of exogenous hematin on anthocyanin biosynthesis
608 during strawberry fruit ripening. *International Journal of Genomics* **2016**:6762731 DOI:
609 10.1155/2016/6762731.
610 **Lin-Wang K, McGhie TK, Wang M, Liu Y, Warren B, Storey R, Espley RV, Allan AC.**
611 **2014.** Engineering the anthocyanin regulatory complex of strawberry (*Fragaria vesca*).
612 *Frontiers in Plant Science* **5**:651 DOI: 10.3389/fpls.2014.00651.
613 **Liu JH, Zhang J, Jia CH, Zhang JB, Wang JS, Yang ZX, Xu BY, Jin ZQ. 2013.** The
614 interaction of banana MADS-box protein MuMADS1 and ubiquitin-activating enzyme E-
615 MuUBA in post-harvest banana fruit. *Plant Cell Reports* **32**:129-137 DOI:
616 10.1007/s00299-012-1347-4.
617 **Livak KJ, Schmittgen TD. 2001.** Analysis of relative gene expression data using real-time
618 quantitative PCR and the $2^{-\Delta\Delta Ct}$ method. *Methods* **25**:402-408 DOI:
619 10.1006/meth.2001.1262.**Manning K. 1998.** Isolation of a set of ripening-related genes
620 from strawberry: their identification and possible relationship to fruit quality traits.
621 *Planta* **205**:622-631 DOI: 10.1007/s004250050365.

622 **Mao XZ, Cai T, Olyarchuk JG, Wei LP. 2005.** Automated genome annotation and pathway
623 identification using the KEGG Orthology (KO) as a controlled vocabulary.
624 *Bioinformatics* **21**:3787-3793 DOI: 10.1093/bioinformatics/bti430.
625 **Medina-Puche L, Cumplido-Laso G, Amil-Ruiz F, Hoffmann T, Ring L, Rodriguez-Franco**
626 **A, Caballero JL, Schwab W, Munoz-Blanco J, Blanco-Portales R. 2014.** *MYB10*
627 plays a major role in the regulation of flavonoid/phenylpropanoid metabolism during
628 ripening of *Fragaria ananassa* fruits. *Journal of Experimental Botany* **65**:401-417 DOI:
629 10.1093/jxb/ert377.
630 **Mortazavi A, Williams BA, McCue K, Schaeffer L, Wold B. 2008.** Mapping and quantifying
631 mammalian transcriptomes by RNA-Seq. *Nature Methods* **5**:621-628 DOI:
632 10.1038/nmeth.1226.
633 **Niu SS, Xu CJ, Zhang WS, Zhang B, Li X, Lin-Wang K, Ferguson IB, Allan AC, Chen KS.**
634 **2010.** Coordinated regulation of anthocyanin biosynthesis in Chinese bayberry (*Myrica*
635 *rubra*) fruit by a R2R3 MYB transcription factor. *Planta* **231**:887-899 DOI:
636 10.1007/s00425-009-1095-z.
637 **Pei MS, Niu JX, Li CJ, Cao FJ, Quan SW. 2016.** Identification and expression analysis of
638 genes related to calyx persistence in Korla fragrant pear. *Bmc Genomics* **17**:132 DOI:
639 10.1186/S12864-016-2470-3.
640 **Perez AG, Sanz C, Olias R, Rios JJ, Olias JM. 1996.** Evolution of strawberry alcohol
641 acyltransferase activity during fruit development and storage. *Journal of Agricultural and*
642 *Food Chemistry* **44**:3286-3290 DOI: 10.1021/Jf960040f.
643 **Pillet J, Yu HW, Chambers AH, Whitaker VM, Folta KM. 2015.** Identification of candidate
644 flavonoid pathway genes using transcriptome correlation network analysis in ripe
645 strawberry (*Fragaria* × *ananassa*) fruits. *Journal of Experimental Botany* **66**:4455-4467
646 DOI: 10.1093/jxb/erv205.
647 **Qin G, Wang Y, Cao B, Wang W, Tian S. 2012.** Unraveling the regulatory network of the
648 MADS box transcription factor RIN in fruit ripening. *The Plant Journal* **70**:243-255 DOI:

PeerJ Preprints NOT PEER-REVIEWED

10.1111/j.1365-313X.2011.04861.x.

Salvatierra A, Pimentel P, Moya-Leon MA, Caligari PD, Herrera R. 2010. Comparison of transcriptional profiles of flavonoid genes and anthocyanin contents during fruit development of two botanical forms of *Fragaria chiloensis* ssp. *chiloensis*. *Phytochemistry* **71**:1839-1847 DOI: 10.1016/j.phytochem.2010.08.005.

Sanchez-Sevilla JF, Vallarino JG, Osorio S, Bombarely A, Pose D, Merchante C, Botella MA, Amaya I, Valpuesta V. 2017. Gene expression atlas of fruit ripening and transcriptome assembly from RNA-seq data in octoploid strawberry (*Fragaria* × *ananassa*). *Scientific Reports* **7**:13737 DOI: 10.1038/s41598-017-14239-6.

Schaart JG, Dubos C, De La Fuente IR, van Houwelingen AMML, de Vos RCH, Jonker HH, Xu WJ, Routaboul JM, Lepiniec L, Bovy AG. 2013. Identification and characterization of MYB-bHLH-WD40 regulatory complexes controlling proanthocyanidin biosynthesis in strawberry (*Fragaria* × *ananassa*) fruits. *The New Phytologist* **197**:454-467 DOI: 10.1111/nph.12017.

Schwechheimer C, Calderon Villalobos LI. 2004. Cullin-containing E3 ubiquitin ligases in plant development. *Current Opinion in Plant Biology* **7**:677-686 DOI: 10.1016/j.pbi.2004.09.009.

Schwinn K, Venail J, Shang Y, Mackay S, Alm V, Butelli E, Oyama R, Bailey P, Davies K, Martin C. 2006. A small family of *MYB*-regulatory genes controls floral pigmentation intensity and patterning in the genus *Antirrhinum*. *The Plant Cell* **18**:831-851 DOI: 10.1105/tpc.105.039255.

Seymour GB, Ryder CD, Cevik V, Hammond JP, Popovich A, King GJ, Vrebalov J, Giovannoni JJ, Manning K. 2011. A *SEPALLATA* gene is involved in the development and ripening of strawberry (*Fragaria* × *ananassa* Duch.) fruit, a non-climacteric tissue. *Journal of Experimental Botany* **62**:1179-1188 DOI: 10.1093/jxb/erq360.

Shima Y, Kitagawa M, Fujisawa M, Nakano T, Kato H, Kimbara J, Kasumi T, Ito Y. 2013. Tomato FRUITFULL homologues act in fruit ripening via forming MADS-box

PeerJ Preprints NOT PEER-REVIEWED

transcription factor complexes with RIN. *Plant Molecular Biology* **82**:427-438 DOI: 10.1007/s11103-013-0071-y.

Shulaev V, Sargent DJ, Crowhurst RN, Mockler TC, Folkerts O, Delcher AL, Jaiswal P, Mockaitis K, Liston A, Mane SP, Burns P, Davis TM, Slovin JP, Bassil N, Hellens RP, Evans C, Harkins T, Kodira C, Desany B, Crasta OR, Jensen RV, Allan AC, Michael TP, Setubal JC, Celton JM, Rees DJ, Williams KP, Holt SH, Ruiz Rojas JJ, Chatterjee M, Liu B, Silva H, Meisel L, Adato A, Filichkin SA, Troggio M, Viola R, Ashman TL, Wang H, Dharmawardhana P, Elser J, Raja R, Priest HD, Bryant DW, Jr., Fox SE, Givan SA, Wilhelm LJ, Naithani S, Christoffels A, Salama DY, Carter J, Lopez Girona E, Zdepski A, Wang W, Kerstetter RA, Schwab W, Korban SS, Davik J, Monfort A, Denoyes-Rothan B, Arus P, Mittler R, Flinn B, Aharoni A, Bennetzen JL, Salzberg SL, Dickerman AW, Velasco R, Borodovsky M, Veilleux RE, Folta KM. 2011. The genome of woodland strawberry (*Fragaria vesca*). *Nature Genetics* **43**:109-116 DOI: 10.1038/ng.740.

Smaczniak C, Immink RG, Angenent GC, Kaufmann K. 2012. Developmental and evolutionary diversity of plant MADS-domain factors: insights from recent studies. *Development* **139**:3081-3098 DOI: 10.1242/dev.074674.

Stone SL, Callis J. 2007. Ubiquitin ligases mediate growth and development by promoting protein death. *Current Opinion in Plant Biology* **10**:624-632 DOI: 10.1016/j.pbi.2007.07.010.

Tanaka Y, Sasaki N, Ohmiya A. 2008. Biosynthesis of plant pigments: anthocyanins, betalains and carotenoids. *The Plant journal* **54**:733-749 DOI: 10.1111/j.1365-313X.2008.03447.x.

Trapnell C, Williams BA, Pertea G, Mortazavi A, Kwan G, van Baren MJ, Salzberg SL, Wold BJ, Pachter L. 2010. Transcript assembly and quantification by RNA-Seq reveals unannotated transcripts and isoform switching during cell differentiation. *Nature Biotechnology* **28**:511-515 DOI: 10.1038/nbt.1621.

Wang F, Zhang F, Chen M, Liu Z, Zhang Z, Fu J, Ma Y. 2017a. Comparative

transcriptomics reveals differential gene expression related to colletotrichum gloeosporioides resistance in the octoploid strawberry. *Frontiers in Plant Science* **8**:779 DOI: 10.3389/fpls.2017.00779.

Wang QH, Zhao C, Zhang M, Li YZ, Shen YY, Guo JX. 2017b. Transcriptome analysis around the onset of strawberry fruit ripening uncovers an important role of oxidative phosphorylation in ripening. *Scientific Report* **7**:41477 DOI: 10.1038/srep41477.

Wang Y, Wang W, Cai J, Zhang Y, Qin G, Tian S. 2014. Tomato nuclear proteome reveals the involvement of specific E2 ubiquitin-conjugating enzymes in fruit ripening. *Genome Biology* **15**:54 DOI: 10.1186/s13059-014-0548-2.

Wei H, Chen X, Zong X, Shu H, Gao D, Liu Q. 2015. Comparative transcriptome analysis of genes involved in anthocyanin biosynthesis in the red and yellow fruits of sweet cherry (*Prunus avium* L.). *PLOS ONE* **10**:e0121164 DOI: 10.1371/journal.pone.0121164.

Wu H, Zhang Y, Zhang W, Pei X, Zhang C, Jia S, Li W. 2015. Transcriptomic analysis of the primary roots of *Alhagi sparsifolia* in response to water stress. *PLOS ONE* **10**:e0120791 DOI: 10.1371/journal.pone.0120791.

Yin WC, Hu ZL, Cui BL, Guo XH, Hu JT, Zhu ZG, Chen GP. 2017. Suppression of the MADS-box gene *SlMBP8* accelerates fruit ripening of tomato (*Solanum lycopersicum*). *Plant Physiology and Biochemistry* **118**:235-244 DOI: 10.1016/j.plaphy.2017.06.019.

Young MD, Wakefield MJ, Smyth GK, Oshlack A. 2010. Gene ontology analysis for RNA-seq: accounting for selection bias. *Genome Biology* **11**:R14 DOI: 10.1186/gb-2010-11-2-r14.

Youssef SM, Amaya I, López-Aranda JM, Sesmero R, Valpuesta V, Casadoro G, Blanco-Portales R, Pliego-Alfaro F, Quesada MA, Mercado JA. 2012. Effect of simultaneous down-regulation of pectate lyase and endo-β-1,4-glucanase genes on strawberry fruit softening. *Molecular Breeding* **31**:313-322 DOI: 10.1007/s11032-012-9791-y.

Zeng S, Liu Y, Pan L, Hayward A, Wang Y. 2015. Identification and characterization of miRNAs in ripening fruit of *Lycium barbarum* L. using high-throughput sequencing.

PeerJ Preprints NOT PEER-REVIEWED

Frontiers in Plant Science **6**:778 DOI: 10.3389/fpls.2015.00778.

Zhang B, Shen JY, Wei WW, Xi WP, Xu CJ, Ferguson I, Chen K. 2010. Expression of genes associated with aroma formation derived from the fatty acid pathway during peach fruit ripening. *Journal of Agricultural and Food Chemistry* **58**:6157-6165 DOI: 10.1021/jf100172e.

Zhang J, Wang X, Yu O, Tang J, Gu X, Wan X, Fang C. 2011. Metabolic profiling of strawberry (*Fragaria* × *ananassa* Duch.) during fruit development and maturation. *Journal of Experimental Botany* **62**:1103-1118 DOI: 10.1093/jxb/erq343.

Zhang YC, Li WJ, Dou YJ, Zhang JX, Jiang GH, Miao LX, Han GF, Liu YX, Li H, Zhang ZH. 2015. Transcript quantification by RNA-Seq reveals differentially expressed genes in the red and yellow fruits of *Fragaria vesca*. *PLOS ONE* **10**:e0144356 DOI: 10.1371/journal.pone.0144356.

Table and figure legends

Table 1 Throughput and quality of RNA-Seq data.

[a,b] Q20 and Q30 indicate the percentage of bases whose Qphred > 20, 30. Error rate, Q20, Q30 and GC content distribution are used to reflect the quality of sequencing data.

Figure 1 Tissues of strawberry 'Toyonoka' used in deep sequencing.

(A) l-GF; (B) WF; (C) TF; (D) RF. Yellow bar = 10mm.

Figure 2 The volcano plots of DEGs in six combinations.

(A) WF/l-GF; (B) TF/l-GF; (C) RF/l-GF; (D) TF/WF; (E) RF/WF; (F) RF/TF. The *x-axis* represents the gene expression times. The *y-axis* represents the statistically significant degree of gene expression change. The smaller the corrected pvalue, the larger the $-\log_{10}$ (padj), and the more significant the difference. The scattered dots represent each gene, the blue dots indicate

genes with no significant differences, the red dots indicate up-regulated genes with significant differences, and the green dots indicate down-regulated genes with significant differences.

Figure 3 Venn diagrams for the different DEGs between each combination.

(A) Number of common and specific DEGs in WF/l-GF, TF/l-GF and RF/l-GF. (B) Number of common and specific DEGs in TF/WF and RF/WF. (C) Number of common and specific DEGs in six combinations (WF/l-GF, TF/l-GF, RF/l-GF, TF/WF, RF/WF and RF/TF). The sum of the numbers in each large circle represents the total number of DEGs in the comparison, and the overlapping parts of the circle represent the number of common DEGs among the combinations.

Figure 4 The expression pattern of genes involved in anthocyanin biosynthesis.

Enzyme names, unigene ids and expression patterns are indicated on each step. The *y-axis* represents average read_count value of each library. No gene is found in the gray line step. TCM: trans-cinnamate 4-monooxygenase. SHT: shikimate O-hydroxycinnamoyl transferase. CHS: chalcone synthase. CHI: chalcone isomerase. F3'M: flavonoid 3'-monooxygenase. N3D: naringenin 3-dioxygenase. DRF: bifunctional dihydroflavonol 4-reductase. LAD: leucoanthocyanidin dioxygenase. ANS: anthocyanidin reductase.

Figure 5 Expression profile of candidate genes during different fruit development and ripening stages in qRT-PCR.

FaActin were used as an internal control. Result shows expression value of candidate genes relative to s-GF stage. The experiments were repeated three times and provided consistent results. The mean values and error bars were obtained from three biological and three technical replicates. The left fruits represents the materials of six different ripening stages in qRT-PCR.

Figure 6 The expression pattern of genes involved in ester biosynthesis.

(A) Ester biosynthesis pathway. Enzyme names, unigene ids and expression patterns are

PeerJ Preprints NOT PEER-REVIEWED

indicated on each step. The *y-axis* represents average read_count value of each library. ADH: alcohol dehydrogenase. AAT: alcohol acyltransferase. (B) The relative expression of down-regulated genes in the degradation of aromatic compound pathway. Black Fonts indicate the up-regulated gene ID. (C) The expression pattern of DEGs in the degradation of aromatic compound pathway.

Figure 7 The expression pattern of genes involved in starch and sucrose biosynthesis.

Enzyme names, unigene ids and expression patterns are indicated on each step. The *y-axis* represents average read_count value of each library. β-FRU: beta-fructofuranosidase. α-GLU: alpha-glucosidase. SUT: sucrose translocase. SUS: sucrose synthase. SPS: sucrose-phosphate synthase. TPS: trehalose 6-phosphate synthase. TPP: trehalose 6-phosphate phosphatase. α-TRE: alpha-trehalase. UGP: UTP--glucose-1-phosphate uridylyltransferase. ASD: ADP-sugar diphosphatase. STP: starch phosphorylase. GPA: glucose-1-phosphate adenylyltransferase. STS: starch synthase. GOP: glycogen operon protein. α-AMY: alpha-amylase. β-AMY: beta-amylase. α-GLU: alpha-glucosidase.

Figure 8 The expression pattern of genes involved in citrate cycle.

Enzyme names, unigene ids and expression patterns are indicated on each step. The *y-axis* represents average read_count value of each library. PYC: acetyl-CoA C-acetyltransferase. MDH: malate dehydrogenase. FH: fumarate hydratase. SDH/α-SCS: succinate dehydrogenase/succinyl-CoA synthetase alpha subunit. DST/β-SCS: dihydrolipoamide succinyltransferase/succinyl-CoA synthetase beta subunit. DLST: dihydrolipoamide succinyltransferase. OGDH: 2-oxoglutarate dehydrogenase E1 component. IDH: isocitrate dehydrogenase. ACH: aconitate hydratase. CS: citrate synthase. ACL: ATP citrate (pro-S)-lyase.

Figure 9 The expression pattern of genes involved in the ubiquitin mediated proteolysis pathway and MADS-box transcript factors.

PeerJ Preprints NOT PEER-REVIEWED

(A) The relative expression of up- and down-regulated genes of ubiquitin conjugating enzyme. Black Fonts indicate the up-regulated gene ID. (B) The expression pattern of DEGs of ubiquitin conjugating enzyme. (C) The relative expression of up-regulated and down-regulated genes of ubiquitin protein ligase. Black Fonts indicate the up-regulated gene ID. (D-F) The expression pattern of DEGs of ubiquitin protein ligase. (G) The relative expression of up- and down-regulated genes of MADS-box transcript factors. Black Fonts indicate the up-regulated gene ID. (H) The expression pattern of DEGs of MADS-box transcript factors.

Supplementary material

Table S1 URLs, annotation methods and parameters of seven databases.

The comprehensive information of gene function comes from seven databases.

Table S2 The information of software version and parameter.

Each data indicate the information of software that produces all the transcriptome data.

Table S3 The distribution of FPKM values of each library.

FPKM: fragments per kilobaseof exon per million fragments mapped. FPKM is the most commonly used method of estimating gene expression level, which eliminates the expression level of technical deviation.

Table S4 Primers used in this study.

Each data indicates the detail information of candidate genes in quantitative real-time polymerase chain reaction.

Table S5 The annotation results of KOG classification.

Each data indicates the annotation result in seven databases.

PeerJ Preprints NOT PEER-REVIEWED

Table S6 The GO classification of unigenes.

Each data indicates the classification of unigenes in GO database.

Table S7 The KOG classification of unigenes.

Each data indicates the classification of unigenes in KOG database.

Table S8 The KEGG classification of unigenes.

Each data indicates the classification of unigenes in KEGG database.

Table S9 Differential analysis results of genes in different combinations.

Each data is used to determine the differentially expressed genes (DEGs). The DEGs with padj < 0.05 and $\log_2$ (fold change) ≥ 1 are up-regulated, and those with padj < 0.05 and $\log_2$ (fold change) ≤ -1 are down-regulated. The other genes that do not meet the conditions of padj < 0.05 and $|\log_2 \text{(fold change)}| \geq 1$ are not DEGs.

Table S10 Detailed information of genes in the flavonoid biosynthesis pathway.

Each data indicates the average read_count of genes in each library.

Table S11 Detailed information of genes in results.

Each data indicates the corrected read_count value, differential analysis results and annotation information of genes in each library.

Figure S1 FPKM interval of all samples.

FPKM: fragments per kilobaseof exon per million fragments mapped. The percentage of each sample's corresponding FPKM interval can be used to measure the difference in expression between samples.

PeerJ Preprints NOT PEER-REVIEWED

Figure S2 Length distribution of transcripts and unigenes.

The *x-axis* represents the length interval of transcript/unigene, and the *y-axis* represents the number of times for each length of the transcript/unigene.

Figure S3 Characteristics of homology search of Illumina sequences against the Nr database.

(A) Percentage of the total homologous sequences of 5 top species against the Nr database; (B) E-value distribution of the top BLASTx hits against the Nr database; (C) Similarity distribution of the top BLASTx hits for each sequence.

Figure S4 Expression pattern of genes in the flavonoid biosynthetic pathway.

(A) Cluster analysis of genes in flavonoid biosynthetic pathway. Expression level was showed by different colors, the redder the higher expression and the bluer the lower. The values of red to blue is log10 (read_count). (B) The relative expression of up- and down-regulated genes in flavonoid biosynthetic pathway. (C) The expression pattern of DEGs in flavonoid biosynthetic pathway.

Figure S5 Expression pattern of MYB and bHLH transcription factors.

(A/C) The relative expression of up- and down-regulated MYB and bHLH transcription factors. Black Fonts indicate the up-regulated gene ID. (B/D) The expression pattern of DEGs of MYB and bHLH transcription factors.

Figure S6 The expression level of candidate genes in transcriptome data.

Each data indicates the expression pattern of candidate genes with strawberry ripening in transcriptome data.

Figure S7 Expression pattern of genes in starch and sucrose biosynthesis and citrate cycle.

PeerJ Preprints NOT PEER-REVIEWED

(A/B) The relative expression of up- and down-regulated genes in starch and sucrose biosynthesis. (C-E) The expression pattern of DEGs in starch and sucrose biosynthesis. (F) The relative expression of up- and down-regulated genes in citrate cycle. (G) The expression pattern of DEGs in citrate cycle.

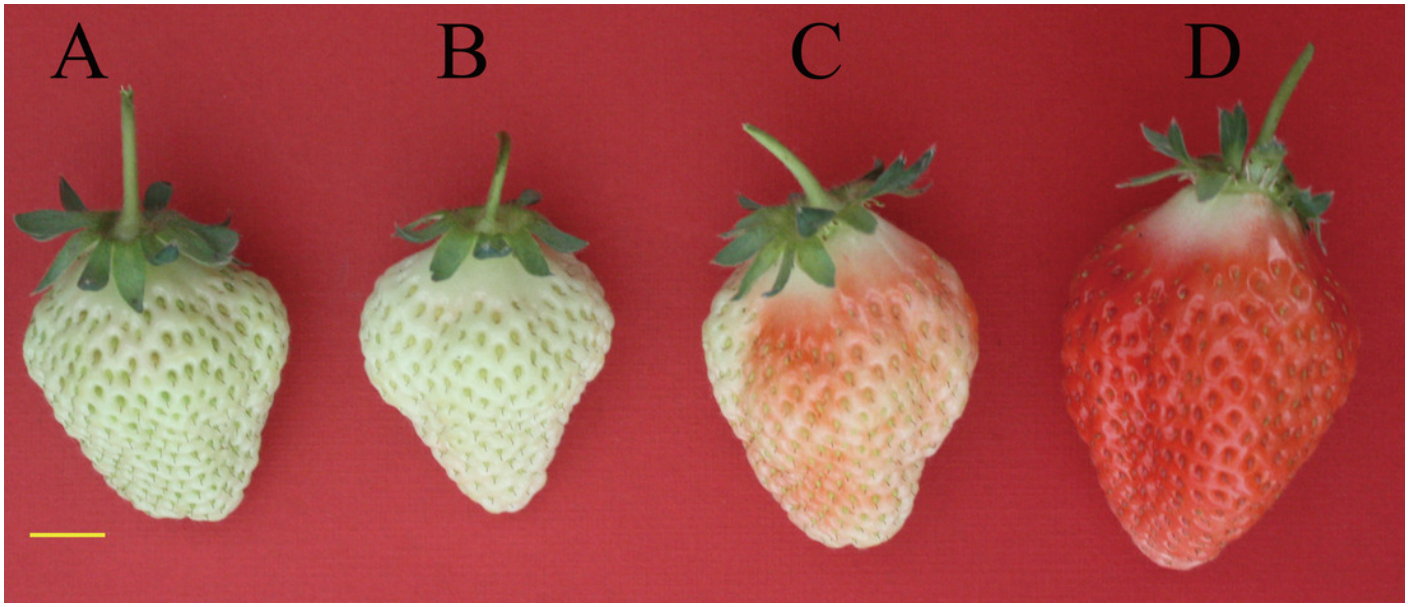
PeerJ Preprints | https://doi.org/10.7287/peerj.preprints.26870v1 | CC BY 4.0 Open Access | rec: 19 Apr 2018, publ: 19 Apr 2018

PeerJ Preprints NOT PEER-REVIEWED

Figure 1

Tissues of strawberry ‘Toyonoka’ used in deep sequencing.

(A) l-GF; (B) WF; (C) TF; (D) RF. Yellow bar = 10mm.



PeerJ Preprints NOT PEER-REVIEWED

Figure 2

The volcano plots of DEGs in six combinations.

(A) WF/l-GF; (B) TF/l-GF; (C) RF/l-GF; (D) TF/WF; (E) RF/WF; (F) RF/TF. The *x-axis* represents the gene expression times. The *y-axis* represents the statistically significant degree of gene expression change. The smaller the corrected pvalue, the larger the -$\log_{10}$ (padj), and the more significant the difference. The scattered dots represent each gene, the blue dots indicate genes with no significant differences, the red dots indicate up-regulated genes with significant differences, and the green dots indicate down-regulated genes with significant differences.

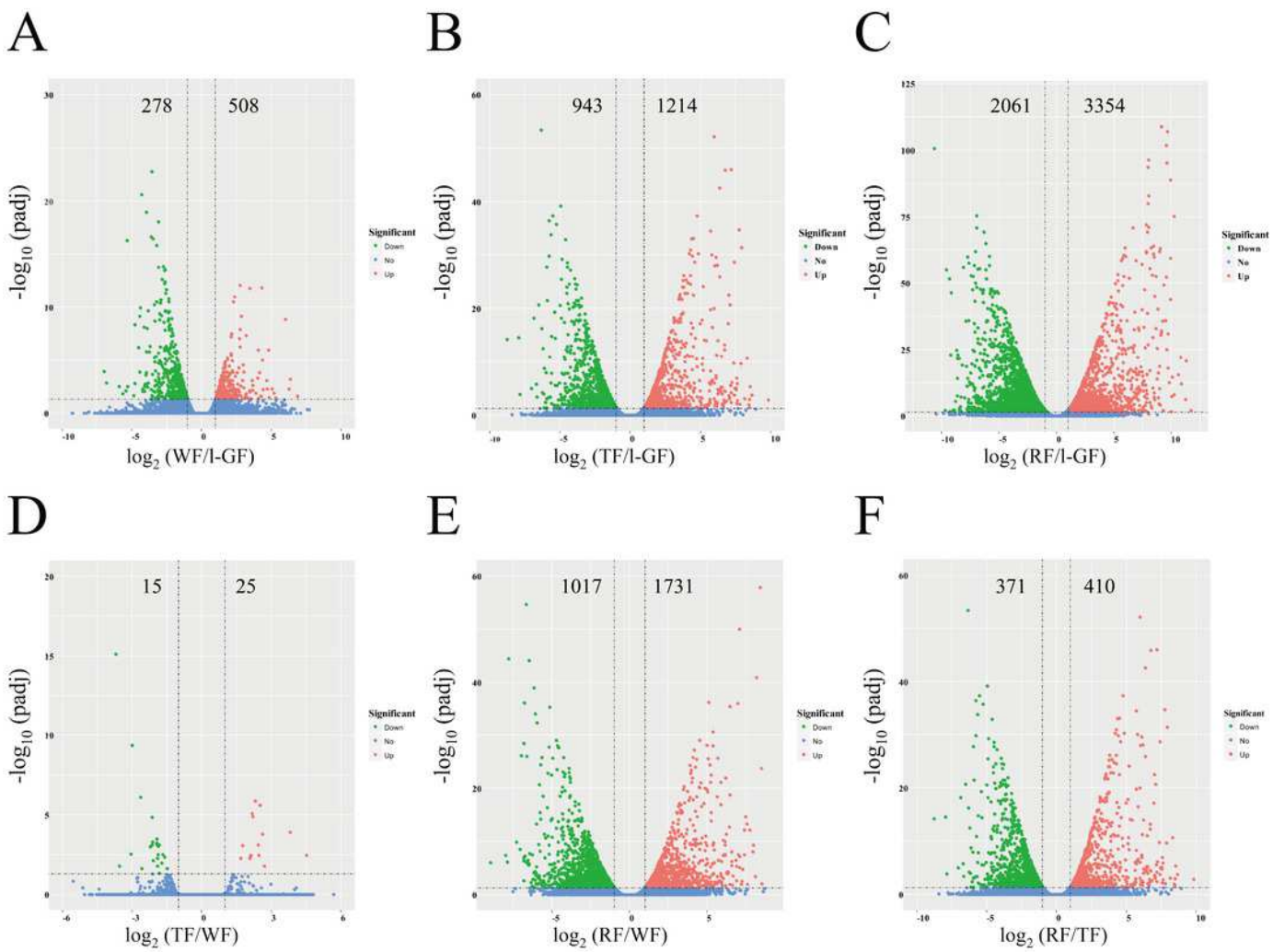
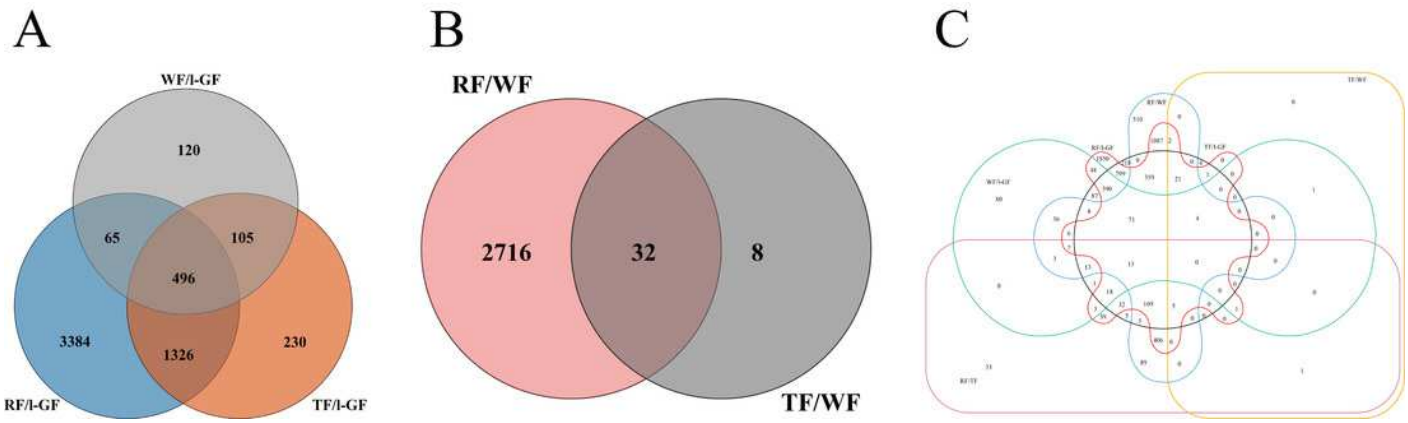


Figure 3

Venn diagrams for the different DEGs between each combination.

(A) Number of common and specific DEGs in WF/l-GF, TF/l-GF and RF/l-GF. (B) Number of common and specific DEGs in TF/WF and RF/WF. (C) Number of common and specific DEGs in six combinations (WF/l-GF, TF/l-GF, RF/l-GF, TF/WF, RF/WF and RF/TF). The sum of the numbers in each large circle represents the total number of DEGs in the comparison, and the overlapping parts of the circle represent the n umber of common DEGs among the combinations.



PeerJ Preprints NOT PEER-REVIEWED

Figure 4

The expression pattern of genes involved in anthocyanin biosynthesis.

Enzyme names, unigene ids and expression patterns are indicated on each step. The *y-axis* represents average read_count value of each library. No gene is found in the gray line step. TCM: trans-cinnamate 4-monooxygenase. SHT: shikimate O-hydroxycinnamoyl transferase. CHS: chalcone synthase. CHI: chalcone isomerase. F3'M: flavonoid 3'-monooxygenase. N3D: naringenin 3-dioxygenase. DRF: bifunctional dihydroflavonol 4-reductase. LAD: leucoanthocyanidin dioxygenase. ANS: anthocyanidin reductase.

PeerJ Preprints | https://doi.org/10.7287/peerj.preprints.26870v1 | CC BY 4.0 Open Access | rec: 19 Apr 2018, publ: 19 Apr 2018

PeerJ Preprints NOT PEER-REVIEWED

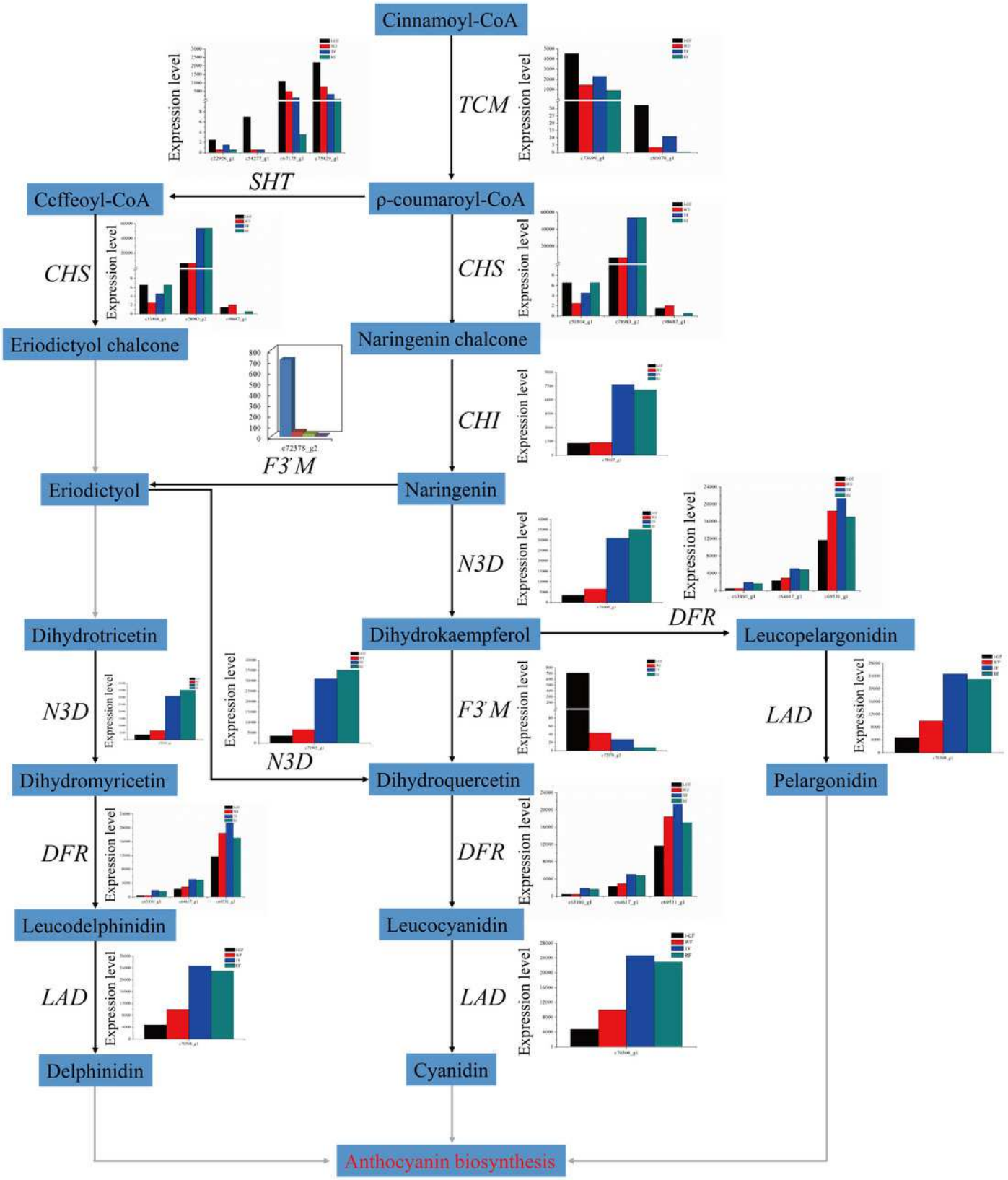


Figure 5

Expression profile of candidate genes during different fruit development and ripening stages in qRT-PCR.

FaActin were used as an internal control. Result shows expression value of candidate genes relative to s-GF stage. The experiments were repeated three times and provided consistent results. The mean values and error bars were obtained from three biological and three technical replicates. The left fruits represents the materials of six different ripening stages in qRT-PCR.

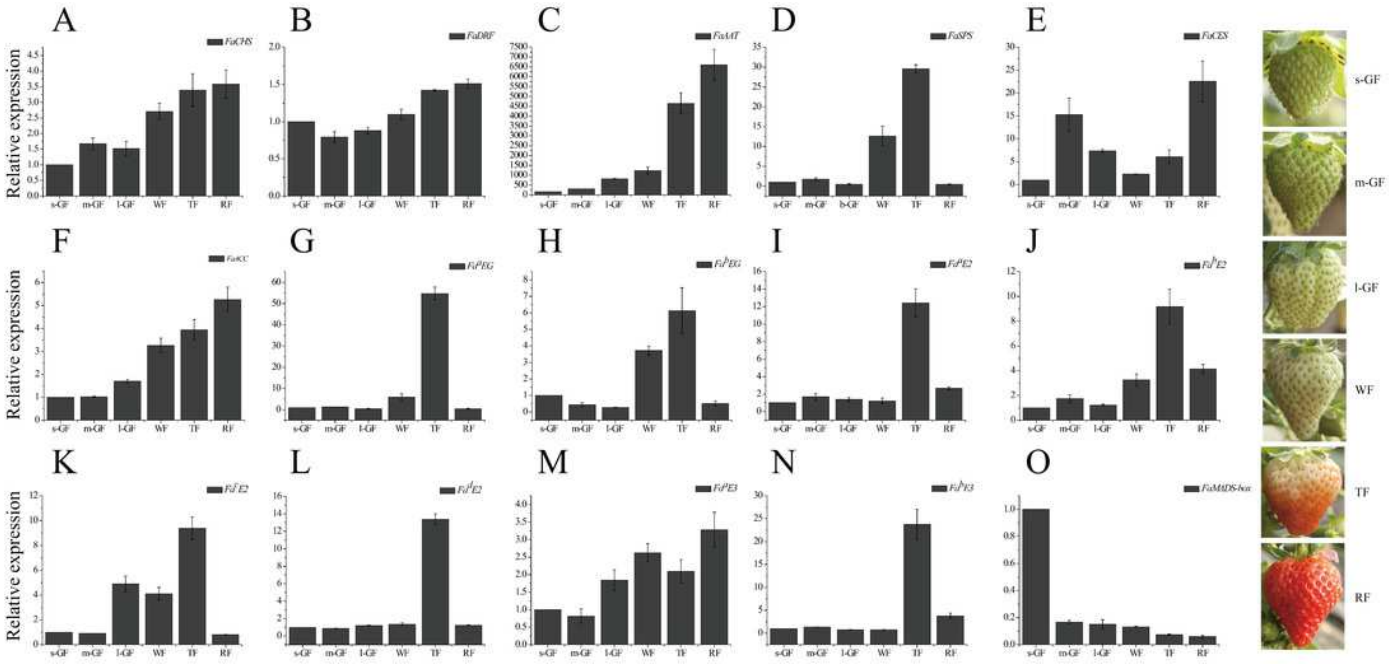


Figure 6

The expression pattern of genes involved in ester biosynthesis.

(A) Ester biosynthesis pathway. Enzyme names, unigene ids and expression patterns are indicated on each step. The *y-axis* represents average read_count value of each library. ADH: alcohol dehydrogenase. AAT: alcohol acyltransferase. (B) The relative expression of down-regulated genes in the degradation of aromatic compound pathway. Black Fonts indicate the up-regulated gene ID. (C) The expression pattern of DEGs in the degradation of aromatic compound pathway.

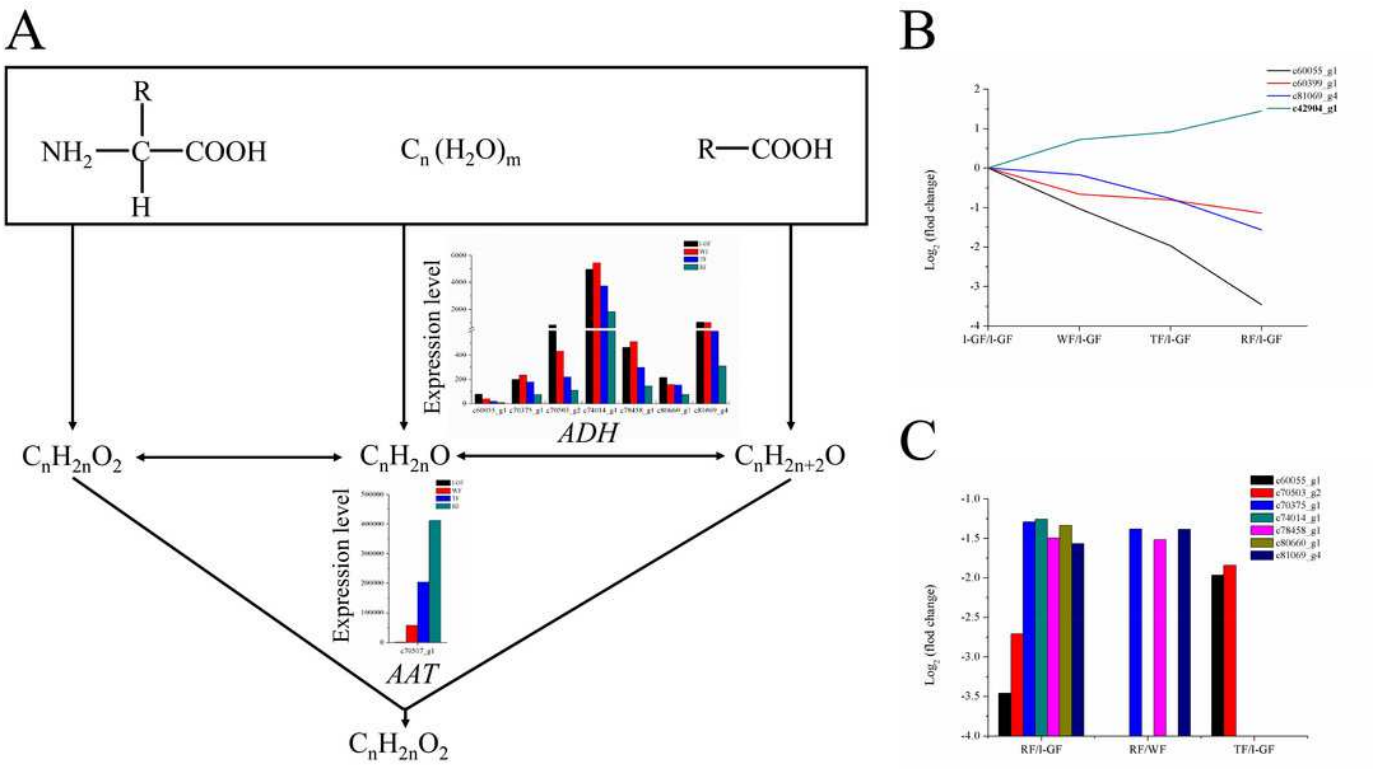


Figure 7

The expression pattern of genes involved in starch and sucrose biosynthesis.

Enzyme names, unigene ids and expression patterns are indicated on each step. The *y-axis* represents average read_count value of each library. β-FRU: beta-fructofuranosidase. α-GLU: alpha-glucosidase. SUT: sucrose translocase. SUS: sucrose synthase. SPS: sucrose-phosphate synthase. TPS: trehalose 6-phosphate synthase. TPP: trehalose 6-phosphate phosphatase. α-TRE: alpha-trehalase. UGP: UTP--glucose-1-phosphate uridylyltransferase. ASD: ADP-sugar diphosphatase. STP: starch phosphorylase. GPA: glucose-1-phosphate adenylyltransferase. STS: starch synthase. GOP: glycogen operon protein. α-AMY: alpha-amylase. β-AMY: beta-amylase. α-GLU: alpha-glucosidase.

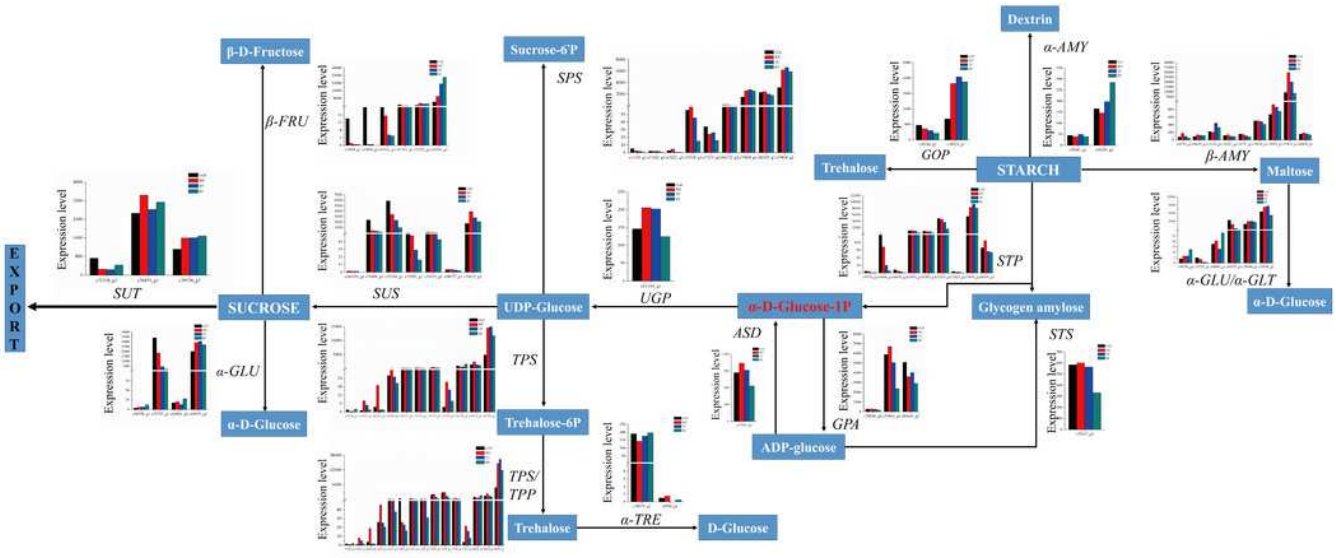
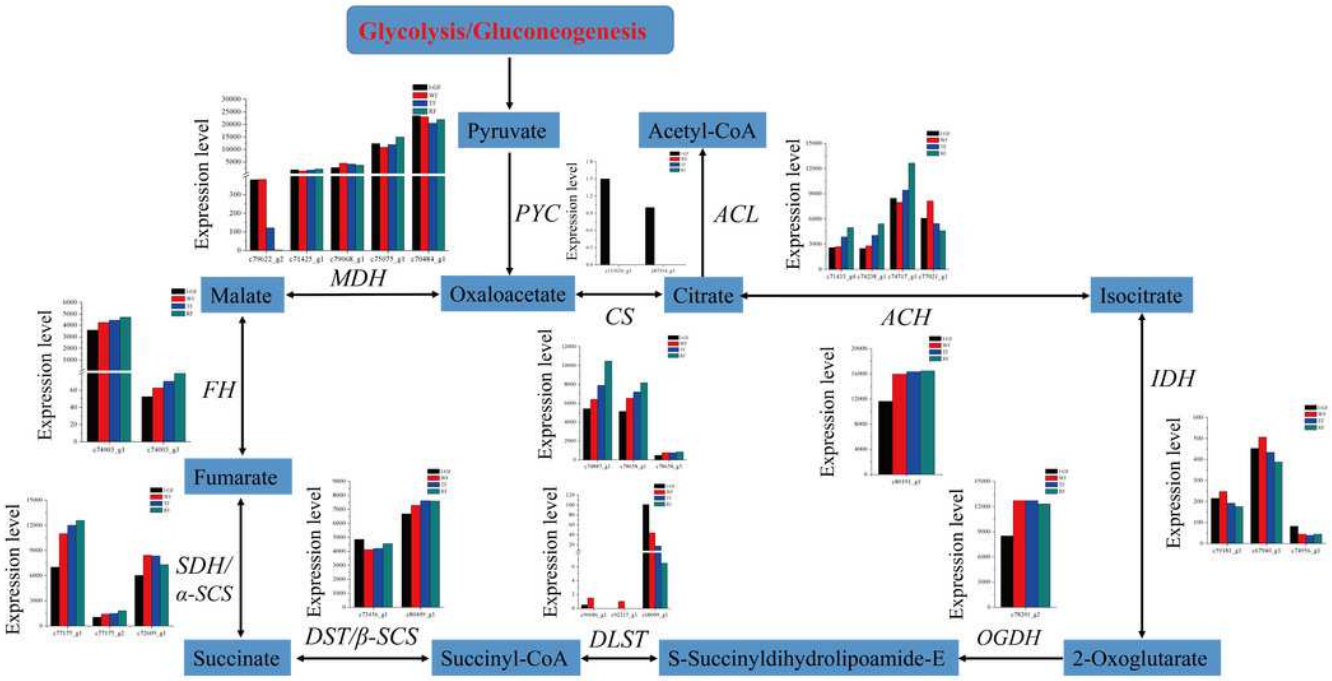


Figure 8

The expression pattern of genes involved in citrate cycle.

Enzyme names, unigene ids and expression patterns are indicated on each step. The *y-axis* represents average read_count value of each library. PYC: acetyl-CoA C-acetyltransferase. MDH: malate dehydrogenase. FH: fumarate hydratase. SDH/α-SCS: succinate dehydrogenase/succinyl-CoA synthetase alpha subunit. DST/β-SCS: dihydrolipoamide succinyltransferase/succinyl-CoA synthetase beta subunit. DLST: dihydrolipoamide succinyltransferase. OGDH: 2-oxoglutarate dehydrogenase E1 component. IDH: isocitrate dehydrogenase. ACH: aconitate hydratase. CS: citrate synthase. ACL: ATP citrate (pro-S)-lyase.

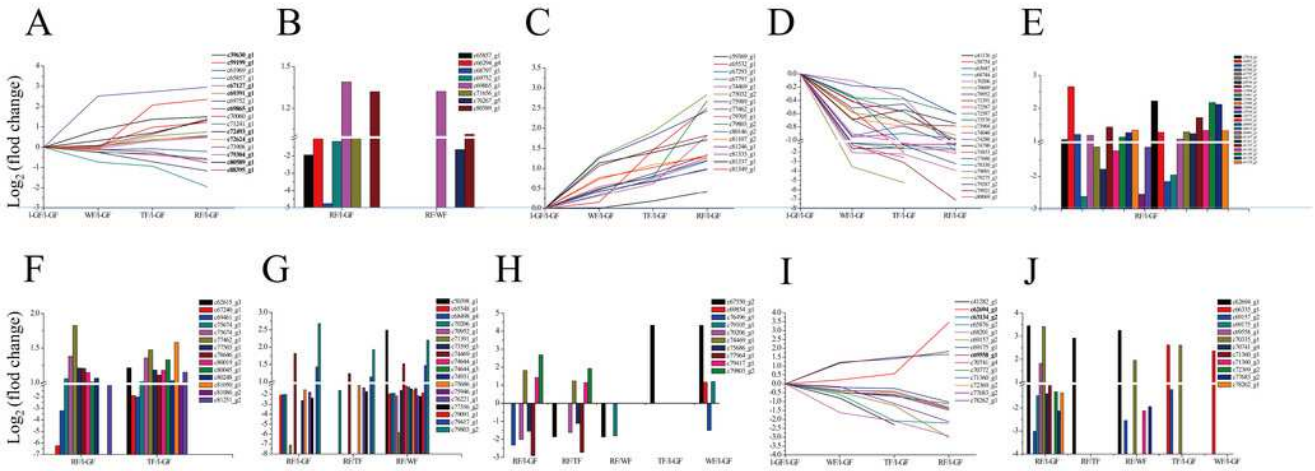


PeerJ Preprints NOT PEER-REVIEWED

Figure 9

The expression pattern of genes involved in the ubiquitin mediated proteolysis pathway and MADS-box transcript factors.

(A) The relative expression of up- and down-regulated genes of ubiquitin conjugating enzyme. Black Fonts indicate the up-regulated gene ID. (B) The expression pattern of DEGs of ubiquitin conjugating enzyme. (C) The relative expression of up-regulated and down-regulated genes of ubiquitin protein ligase. Black Fonts indicate the up-regulated gene ID. (D-F) The expression pattern of DEGs of ubiquitin protein ligase. (G) The relative expression of up- and down-regulated genes of MADS-box transcript factors. Black Fonts indicate the up-regulated gene ID. (H) The expression pattern of DEGs of MADS-box transcript factors.



PeerJ Preprints NOT PEER-REVIEWED

Table 1(on next page)

Throughput and quality of RNA-Seq data.

[a,b] Q20 and Q30 indicate the percentage of bases whose Qphred > 20, 30. Error rate, Q20, Q30 and GC content distribution are used to reflect the quality of sequencing data.

PeerJ Preprints NOT PEER-REVIEWED

1

Sample	Raw reads	Clean reads	Clean bases	Error rate (%)	*Q*20[a] (%)	*Q*30[b] (%)	GC (%)
l-GF1	58541836	57209502	8.58G	0.02	96.48	91.19	46.74
l-GF2	60581866	59222064	8.88G	0.02	96.27	90.72	46.93
WF1	66696962	65070548	9.76G	0.02	96.56	91.35	46.8
WF2	61783100	60066380	9.01G	0.02	96.55	91.29	46.84
TF1	63081374	61671990	9.25G	0.02	96.55	91.29	46.66
TF2	61345068	59832880	8.97G	0.02	96.48	91.16	46.42
RF1	61847198	60455548	9.07G	0.02	96.68	91.53	46.04
RF2	59579024	58261550	8.74G	0.02	96.67	91.53	45.88

[a,b] Q20 and Q30 indicate the percentage of bases whose Qphred > 20, 30 in the overall bases

2

3

4

5