

Analysis of Regulation Networks in the Formation of Green Pigment in Garlic

Ningyang Li (✉ ningyangli@126.com)

Shandong Agricultural University <https://orcid.org/0000-0002-6103-3303>

Xiaoming Lu

shandong agricultural university

Xiudong Sun

shandong agricultural university

Xuguang Qiao

shandong agricultrual university

Research article

Keywords: garlic, green pigment, porphobilinogen, citrate cycle, salicylic acid

Posted Date: March 17th, 2020

DOI: <https://doi.org/10.21203/rs.3.rs-17429/v1>

License: © ⓘ This work is licensed under a Creative Commons Attribution 4.0 International License.

[Read Full License](#)

Abstract

Background: Garlic (*Allium sativum* L.) has high medical and edible value, but is prone to producing green pigment during storage, which seriously affects product quality. Herein, we found multiple metabolic pathways involved in green pigment formation by analyzing the transcriptome and proteome changes of garlic stored at 4 °C for 0 day (DS0), 10 days (DS1), 15 days (DS2) and 40 days (DS3).

Results: In this work, we found that porphobilinogen is a precursor of garlic green pigment and the citrate cycle (TCA cycle) can promote the formation of garlic green pigment by increasing the content of porphobilinogen. The salicylic acid (SA) inhibits garlic greening, whereas the lignin synthesis pathway positively correlates with greening. In addition to the above metabolic pathways, other pathways may be associated with the formation of garlic's green pigment, such as glycolysis/gluconeogenesis, pyruvate metabolism and fatty acid metabolism. It has been reported that pyruvate and amino acids play a role in garlic greening, which were also demonstrated in this study.

Conclusion: our results have important implications for understanding the mechanism of green pigment formation in garlic.

Background

Garlic is a Liliaceae allium plant that is used worldwide in the gastronomic traditions of many countries, its production accounts for millions of tons each year. All parts of garlic are edible and the bulb is the most important part for the market. The beneficial properties of garlic for human health are well known, such as reducing the risk of cardiovascular problems and cancer (1–4). The best ideal conditions for storing garlic for over six months is that the storage temperature is between – 5 °C to 0 °C and at a relative humidity between 60–80% (5, 6). Dormancy plays an important role in garlic storage, which can consist of three stages: pre-hibernation, deep dormancy and forced dormancy. If storage conditions change, dormancy will end, which can lead to rooting, germination and greening phenomena (7). The greening phenomenon of garlic can seriously reduce its medicinal and edible value; therefore, the mechanism of garlic greening has been widely studied. Since the 1940s, scholars have published much research on the causes and formation mechanisms of green pigmentation in garlic. The green pigment of garlic is not caused by chlorophyll or related porphyrin ring pigments, and is not related to phenolase and peroxidase (8, 9). There are numerous causes for the formation of green pigment in garlic, such as PH, temperature, amino acids and enzymes. Garlic forms different pigments at different pH conditions, which are bluer at $\text{pH} \geq 5$ than at $\text{pH} \leq 4$ (10, 11). This phenomenon may be explained by changes in the composition of different pigment mixtures or pH affecting the formation of pigment. Storing garlic bulbs for a month at $> 23^\circ\text{C}$ prior to processing was shown to prevent the production of green pigment in garlic puree. The amino acid S-(1-propenyl) cysteine sulfoxide was essential for the development of green color (12, 13).

To a certain extent, ascorbic acid can inhibit the greening of garlic, but the effect is not complete, ascorbic acid can effectively inhibit browning of garlic products however (14). Some enzymes in garlic accelerate the greening process, such as γ -glutamyl transpeptidase (γ -GTP) and alliinase, and the latter participates in the garlic greening mechanism (15). In this mechanism, cysteine reacts with garlic thiosulfinates to form S-alk(en)yl mercaptocysteines (SAMCs), which then react with color-developing thiosulfinates to form pigments, but only when residual thiosulfinates are available after complete cysteine exhaustion (16). According to current research, the greening process of garlic is mainly divided into the following six steps: First, pyruvate and S-1-propenyl 1-propenethiosulfinate is produced from S-(1-propenyl)-L-cysteine sulfoxide (1-PeCSO) by the action of alliinase. Second, pyrrolyl amino acid is generated as an intermediate by the reaction of S-1-propenyl 1-propenethiosulfinate with amino acids. Third, S-allyl-L-propenethiosulfinate is generated from S-allyl-L-cysteine sulfoxide (2-PeCSO) by the action of alliinase. Fourth, the pyrrolyl intermediate reacts with pyruvate to form a yellow pigment. Fifth, blue pigment is generated from the reaction of pyrrolyl amino acid with s-allyl-L-propenethiosulfinate. Finally, blue pigment is degraded to yellow pigment which combines with remaining blue pigment, forming the characteristic garlic green pigment (17, 18). These research efforts have concentrated on physical factors affecting garlic greening to prolong its storage period to improve the quality of garlic products.

Knowledge of regulatory mechanisms underlying garlic greening is useful for maintaining garlic quality. Although the greening process of garlic has been well explained, we assume that there are other factors that affect the greening. In order to further improve the mechanism of garlic greening, we investigate the garlic greening in more depth. According to previous research in our laboratory, it is proved that low temperature induced stress-responsive enzyme-encoding genes are possibly responsible for garlic discoloration by comparative transcriptome analysis (19). Herein, we reanalyzed previous RNA-seq data (SRP148664) and obtained protein quantization data from all garlic samples. A method combining bioinformatics analysis and experimental validation was used to find multiple metabolic pathways associated with the formation of garlic green pigment.

Results

Physiological characteristics of garlic green pigment

The absorption spectrum of a methanol extract containing garlic green pigment has previously shown two absorbance maxima at 440 and 590 nm (13). To support this observation, we measured the absorbance of crude methanol extracts of pigment at different concentrations. Although green pigment concentration varied, the absorbance maxima of green pigment remained at 440 nm and 590 nm (Fig. 1a). Our conclusions are consistent with previous reports, so we proceeded to use the absorbance maxima 440 and 590 nm as indicators to measure the degree of garlic greening. During the dormant period, the activity of metabolism is weak and metabolic rate is low, therefore the content of metabolites and green pigment is low. However, a low temperature can break dormancy and promote the formation of green pigment (12). In our previous study, we measured the changes in green pigment content of these

garlic samples. We found that as the storage period increases, green pigment content gradually increases (Fig. 1b).

Dynamic Transcriptome Of Garlic Samples

To investigate changes in the transcriptome of different samples, high-quality reference transcriptomes were obtained using a single-molecule long-read sequencer from Pacific Biosciences to sequence transcriptomes from a subsection of garlic samples (DS0 and DS3). For Pacific Biosciences sequencing, 0.95 million circular consensus reads were obtained from six libraries of different fragment sizes, and 60878 non-redundant transcripts, accounting for 1.24 billion bases in total, were generated using CD-HIT software. Transcript length ranged from 101 – 15,077 bp, and the average length was 2,037 bp (Supplementary Fig. 1a). The full-length transcript was predicted using Trans Decoder software to obtain a coding domain (CDS) of 36,776 unique protein-encoding transcripts. These proteins were functionally annotated with NR, Swiss Prot, Pfam, NOG, GO and KEGG databases, and a total of 30,082 full-length transcripts with functional annotations were obtained (Supplementary Table 1).

Illumina RNA sequencing revealed a total of 51,041 co-expression transcripts across DS0, DS1, DS2 and DS3 samples (Supplementary Table 2). Previously, 20 genes were selected to validate the expression profiling by Illumina sequencing. To obtain highly similar genes, BLAST searches were carried out to search all gene sequences generated in this work with the 20 gene sequences as queries. A total of 14 highly similar genes were generated and the correlations of the expression patterns of these genes were calculated (Supplementary Fig. 2). There were 11/14 genes with correlations above 0.8, indicate that the accurate analysis results can be used for further functional analysis. To understand the dynamics of the garlic transcriptome in different samples, the Short Time-series Expression Miner (STEM) (20) program was used to analyze the total co-expressing genes. There were 10 significant temporal gene expression profiles generated by STEM ($p < 0.05$) (Supplementary Fig. 3). Profiles 21, 22, 24 and 25 were similar in expression profile and displayed the same color profile, and genes of these four profiles presented a tendency for upregulation. The down-regulated profiles 1, 2 and 4 were similar in expression profile and displayed the same color profile. Genes in up-regulated (21, 22, 24 and 25) and down-regulated (1, 2 and 4) profiles were subjected to KEGG pathway analysis and the significant pathways ($p < 0.05$) enriched were shown (Supplementary Fig. 1b). The genes in up-regulated profiles involved 18 significant KEGG pathways, including carbon metabolism, glycolysis/gluconeogenesis, amino acid biosynthesis, pyruvate metabolism and the citrate cycle (TCA cycle). Genes in down-regulated profiles involved 11 significant KEGG pathways, including phenylalanine metabolism, fatty acid biosynthesis and phenylpropanoid biosynthesis (Supplementary Table 4). These results indicate that the 18 significant KEGG pathways associated with up-regulated profiles may promote garlic greening, while 11 pathways associated with down-regulated profiles may play a negative role in garlic greening.

Transcriptomic Alterations In Garlic At Different Storage Periods

To assess alterations in gene expression during different storage periods of garlic, we globally analyzed the genes expressed in DS0, DS1, DS2 and DS3 samples. The changes in gene expression level were identified by comparing DS3, DS2, DS1 with DS0 (termed DS3/DS0, DS2/DS0, DS1/DS0, respectively). Genes with a p-value < 0.05 and absolute value of log₂ fold change ($|\log_2FC| \geq 1$) were selected as differentially expressed genes (DEGs). By further analyzing DEGs, we found 9,131 DEGs (5,697 up and 3,434 down) between DS1 and DS0 garlic, 8,262 DEGs (5,349 up and 2,913 down) between DS2 and DS0 garlic and 5,923 DEGs (3,509 up and 2,414 down) between DS3 and DS0 garlic (Supplementary Fig. 4a, b, and c, Supplementary Table 4). The distribution of co-expressed DEGs were calculated and are presented in a Venn-diagram (Supplementary Fig. 4d). Time course sequencing data analysis (TCseq), an R package for analysis of different time course types, was performed to analyze the hierarchical cluster of co-expressed DEGs. Hierarchical cluster analysis indicated that all DEGs can be grouped into six clusters (Supplementary Fig. 5). Genes in Cluster 1 showed increased expression level during storage from DS0 to DS3. Genes in Cluster 6 showed decreased expression level during storage from DS0 to DS3. Genes in Clusters 4 and 5 showed decreased expression level from DS0 to DS2 and from DS0 to DS1, respectively. In addition, genes in Cluster 2 showed decreased expression level during storage from DS0 to DS1 and increased expression level during storage from DS1 to DS3. By contrast, genes in Cluster 3 showed increased expression from DS0 to DS1 and slightly decreased expression during subsequent days.

GO enrichment analysis was performed to investigate biological functions of crucial DEGs. It was revealed that Cluster 1 genes were enriched in GO terms (biological process subcategory) associated with the tricarboxylic acid cycle (TCA), glycolysis/gluconeogenesis, respiration, and pyruvate metabolism. Genes in Cluster 3 with a similar expression pattern to Cluster 1 were associated with transport processes of elements such as iron, hydrogen and protons (Fig. 2). Cluster 4 genes were enriched in GO terms related to fatty acid metabolism, phenylpropanoid metabolism, chitin metabolism, response to abiotic stimulus, the gibberellic acid signaling pathway, and organic metabolism. Cluster 5 genes were enriched in GO terms related to biological regulation, response to abiotic stimulus, and regulation of gene expression. Cluster 6 genes were enriched in GO terms related to biological regulation, the red-light signaling pathway, and response to abiotic stimuli. It is of interest to note that Cluster 2, 5 and 6 genes were enriched in GO terms related to biological regulation and Cluster 4, 5 and 6 genes were enriched in GO terms related to response to abiotic stimuli.

Proteome Analysis Of All Samples

The formation of garlic green pigment relies on a complex set of traits related to morphological, physiological, developmental, and cellular processes that are linked to multiple pathways. In addition, to understand transcriptome changes in garlic, it is also necessary to understand the proteome changes. Thus, to better understand the molecular mechanisms of garlic greening, proteome data from all samples

was identified and quantified using Label-free analysis. Accordingly, 30,234 unique peptides and 15,860 proteins were identified with a false discovery rate (FDR) $\leq 1\%$ (Table 1). Perseus software was used to identify differentially abundant proteins (DAPs) (FDR < 0.05 , absolute value of \log_2 fold change ($|\log_2FC| \geq 1$) between DS3, DS2, DS1 and DS0. Among these proteins, there were 621 (546 up and 75 down), 657 (562 up and 95 down) and 1254 (1,027 up and 227 down) DAPs in DS1/DS0, DS2/DS0 and DS3/DS0, respectively (Table S5). DAPs from all comparisons were subjected to KEGG analysis and significant pathways are listed (Supplementary Fig. 6, Supplementary Table 6). Up-regulated DAPs related to the TCA cycle were observed in DS1/DS0 and DS2/DS0. The up-regulated DAPs in DS1/DS0 were enriched in KEGG pathways related to pyruvate metabolism and glycolysis/gluconeogenesis. The down-regulated DAPs from all comparisons were enriched in KEGG pathways related to primary metabolic pathways, such as ribosome formation, DNA replication and purine metabolism (Supplementary Fig. 6).

Table 1
Overview of the proteome.

Samples	Number of proteins	Peptides	Unique peptides
DS0_1	15860	24132	13560
DS0_2	15860	24244	13756
DS0_3	15860	24310	13822
DS1_1	15860	25971	14877
DS1_2	15860	26571	15173
DS1_3	15860	26418	15180
DS2_1	15860	26417	15264
DS2_2	15860	26164	15139
DS2_3	15860	26259	15157
DS3_1	15860	26706	15489
DS3_2	15860	26771	15427
DS3_3	15860	26594	15342
Total	15860	45984	30234

Tca Cycle Promotes The Development Of Garlic Green Pigment

Since GO enrichment analysis of co-expressed DEGs in Cluster1 showed that the TCA cycle may play a role in the formation of garlic green pigment, we paid special attention to genes involved in the TCA cycle.

Several TCA cycle intermediates between fresh garlic and green garlic were detected. We found that the fumaric acid, citric acid and ketoglutaric acid contents in fresh garlic were significantly lower than that in green garlic (Table 2). We identified 51 unigenes in co-expressed up-regulated DEGs encoding critical enzymes of TCA cycle (Supplementary Table 7).

Table 2
The content of several key intermediate objects of TCA cycle in garlic cloves.

Garlic samples	fumaric acid ($\times 10^{-3}$ mg/g)	citric acid (mg/g)	ketoglutaric acid ($\mu\text{g/g}$)
Fresh garlic	34	0.0857	66.3
Greened garlic	63.5	0.12	248

From the identified 51 unigenes, 5 genes encoding pyruvate dehydrogenase E2 component (DLAT), 1 gene encoding pyruvate dehydrogenase E1 component alpha subunit (PDHA), 4 genes encoding pyruvate dehydrogenase E1 component beta subunit (PDHB) and 1 gene encoding citrate synthase (CS) were identified (Supplementary Table 7). The expression levels of these 11 genes at DS0 was significantly lower than that at DS1, DS2 and DS3 (Fig. 3a). In addition, the activity of citrate synthase and the citric acid content increased from DS0 to DS3 (Fig. 3b). There were 8 genes encoding isocitrate dehydrogenase (NAD⁺), and the expression level of these genes at DS1, DS2 and DS3 was higher than that at DS0 (Fig. 3c). The activity of isocitrate dehydrogenase was enhanced and the content of α -ketoglutarate gradually increased from DS0 to DS3 (Fig. 3d). In addition, there were 2 genes encoding the succinyl-CoA synthetase alpha subunit, 2 genes encoding the succinyl-CoA synthetase beta subunit, 1 gene encoding the 2-oxoglutarate dehydrogenase E2 component and 2 genes encoding the succinate dehydrogenase (ubiquinone) iron-sulfur subunit. Compared with DS0, the expression level of these 7 genes and fumarate content at DS1, DS2 and DS3 was increased (Fig. 3e, f). Although the activity of succinate dehydrogenase decreased from DS1 to DS3 gradually, at DS1, DS2 and DS3 it was higher than that at DS0 (Fig. 3f).

The 251 co-expressed up-regulated DAPs were generated by performing proteomic analysis. In these DAPs, 10 DAPs were phosphoenolpyruvate carboxykinases (ATP) associated with the TCA cycle (Supplementary Table 7). The content of these proteins at DS0, DS1, DS2 and DS3 were significantly higher than that at DS0 and the highest content was observed at DS3 (Supplementary Fig. 7). These results suggest that the TCA cycle promotes the formation of green pigment in garlic.

Genes Regulating Respiration Play A Role In Garlic Greening

GO enrichment analysis showed that genes in Clusters 1 were enriched in GO terms related to respiration, and the TCA cycle is closely related to respiration, we suspected that respiration may affect garlic greening. To test this hypothesis, garlic was treated with different concentrations of sodium azide (0.1 mol/L and 0.5 mol/L) as a respiratory chain inhibitor, and distilled water (DW) as control, then the

intensity of green pigment was measured. The intensity of greening was significantly lower in samples treated with different concentrations of sodium azide than those treated with DW, suggesting that respiration promotes the formation of garlic green pigment (Supplementary Fig. 8).

Porphobilinogen As A Precursor Of Garlic Green Pigment

Adding synthetic pyrrolyl amino acids to garlic can cause different degrees of change in fresh garlic greening, it is considered that pyrrolyl compounds are an intermediate formed by garlic green pigment (21, 22). The pyrrolyl compound porphobilinogen (PBG), as a precursor of various biological natural tetrapyrrole pigments, is likely to participate biosynthetic pathways as a precursor of garlic green pigment (Fig. 4a). By measuring the content of PBG, we found that as storage time increased, PBG content gradually increased, and there was a significant positive correlation with the intensity of garlic greening ($P < 0.01$) (Fig. 4b). These results indicate that PBG is closely related to the biosynthetic pathway of green pigment. PBG is a pyrrole compound generated by aminolevulinate dehydratase (ALAD) that catalyzes asymmetric dehydration condensation of aminolevulinic acid (ALA). ALAD activity increased from DS0 to DS3 (Fig. 4b). ALA is formed by aminolevulinate synthetase, which is catalyzed by succinyl-CoA (23). Therefore, we suspect that the TCA cycle promotes the formation of garlic green pigment by increasing the content of PBG (Fig. 4a). To test this hypothesis, garlic was treated with malonic acid, a competitive inhibitor of succinate dehydrogenase (SDH), and fumarate content and SDH activity were measured. Fumarate content and SDH activity decreased as the concentration of malonic acid increased, suggesting malonic acid can prevent the transition from succinyl-CoA to fumarate (Fig. 4c). Subsequently, we found that from 0–0.4% malonate concentration, the green intensity and PBG content increased, then from 0.4% – 1% malonate, a gradual decrease in green intensity and PBG content were observed (Fig. 4d). When the concentration of malonic acid is less than 0.4%, malonic acid lowers the activity of SDH, leading to succinic acid accumulation, thereby reducing the conversion rate of succinyl-CoA to succinic acid. Malonic acid is conducive to the conversion of succinic acid CoA to pyrrolyl compounds, therefore garlic green intensity increased. However, when the concentration of malonic acid is higher than 0.4%, it can act as a cellular respiratory inhibitor, blocking the TCA cycle and hindering the conversion of succinic acid CoA to pyrrolyl compound, resulting in a decrease in garlic greening intensity and PBG content. The above results indicate that TCA promotes the formation of garlic green pigment by increasing the content of PBG.

Sa Inhibits The Formation Of Garlic Green Pigment

GO enrichment analysis showed that genes in Cluster 4 were enriched in GO terms related to phenylpropanoid metabolic pathways. Furthermore, we identified 17 co-expressed DEGs related to phenylalanine metabolism and phenylpropanoid biosynthesis by performing KEGG enrichment analysis of all co-expressed DEGs. In these 17 genes, 6 were involved in both phenylalanine metabolism and phenylpropanoid biosynthesis, therefore, a total of 11 genes were involved these two pathways

(Supplementary Table 7). There were 5 key genes in salicylic acid synthesis, and the expression level of these 5 genes in DS1, DS2 and DS3 was significantly lower than that at DS0 (Supplementary Fig. 9a). Based on the above result, we suspected that SA may negatively regulate the process of garlic greening. In order to test this hypothesis, garlic was treated with different concentrations of SA and the green color intensity was measured. We found that as the concentration of SA increased, the green intensity of garlic gradually decreases, indicating that SA inhibited garlic greening (Supplementary Fig. 9b).

Gene Co-expression Network Construction And Module Identification

Expression level of all co-expressed genes in 12 samples were used to construct the gene co-expression module by WGCNA package. A total of 9 distinct gene co-expression modules (70–6560 genes) were generated (Supplementary Fig. 10a). In these modules, antiquewhite1 and brown4 were positively correlated with changes in garlic green intensity ($r > 0.8$, $p < 0.05$) and blue4 module was negatively correlated with it ($r < -0.8$, $p < 0.05$) (Supplementary Fig. 10b). Therefore, only antiquewhite1, brown4 and blue4 modules were selected for further analysis. Hub genes, as the most important part of co-expression network, were identified from these three modules by performing Cytoscape software. We obtained 27, 37 and 40 hub genes from antiquewhite1, brown4, and blue4 modules, respectively, based on node degree centrality. These hub genes involved multiple KEGG pathways (Supplementary Table 8), such as two hub genes c225444/f1p7/1575 and c278010/f1p43/1816 in the antiquewhite1 module encoding citrate synthase, hub genes c26311/f1p49/2860 and c60/f1p394/3763 in brown4 module encoding citrate synthase and aldehyde dehydrogenase (NAD⁺), respectively. In the blue4 module, two hub genes (c272/f1p17/4342, c3369/f1p14/2256) encoding phenylalanine ammonia-lyase (PAL), which was the key gene for SA synthesis. The above results further indicate that the TCA promotes the formation of green pigment, while SA has a suppressive function on process of garlic greening.

Discussion

Advantages of PacBio sequencing in transcriptome analysis on species without a reference genome

The greening phenomenon of garlic seriously reduces the medicinal and edibility value of garlic. Therefore, studying the greening mechanism at the transcriptome level can prolong the storage time of garlic and maintain its original value. However, the lack of genome and transcriptome information severely impedes the study of garlic's greening mechanism. Sequencing of the full-length garlic transcriptome could help us to understand and use this important species. When performing transcriptome analysis on species without a reference genome, many complex and unclear problems are often encountered, especially from the assembly of sequencing reads. For example, transcripts of different gene isoforms show high sequence similarity to one another, and repeat sequences are present

in some transcripts. Therefore, a high-confidence transcriptome atlas can play a key role in solving these problems. In this study, we successfully obtained the first high-quality collection of transcripts in garlic using single-molecule long-read PacBio sequencing.

The Relationship Between Tca Cycle And Garlic Greening

The TCA cycle, as a universal feature of the metabolism, is closely related to the growth and development of most organisms. The TCA cycle plays a central role in mitochondrial respiratory metabolism, which oxidizes organic carbon substrate to generate the reducing equivalents NADH and FADH₂ the fuel ATP synthesis by oxidative phosphorylation. In addition to respiratory metabolism, the TCA cycle provides a carbon skeleton for biosynthetic processes, and is involved in metabolizing organic acids generated from other pathways, such as the glyoxylate cycle during lipid mobilization (24, 25). Although the cycle's operation was demonstrated in plants decades ago (26), information of the relationship between TCA cycle and formation of garlic green pigment is unknown. In this study, we found that key enzyme activities in the TCA cycle and the expression levels of TCA-related genes increased with the increase of garlic green intensity (Fig. 3), indicating that the TCA cycle promotes formation of garlic green pigment.

Phenylpropanoid Metabolism Affects The Green Pigment In Garlic

Phenylpropanoid metabolism is an important pathway involved in plant secondary metabolic pathways, including phenylalanine metabolism and synthesis of secondary metabolites such as anthocyanins and lignin (27). In transcriptome analysis, there were 5 co-expressed down-regulated DEGs encoding phenylalanine ammonia-lyase (PAL) (Supplementary Fig. 9). The PAL mediated phenylalanine metabolic pathway is the main pathway for SA synthesis (28). Treatment of garlic with different concentrations of SA can inhibit greening, indicating that SA can inhibit the formation of green pigment (Supplementary Fig. 9b). SA plays a variety of roles across the plant's entire growth cycle, such as thermogenesis, flower induction, uptake of ions, ethylene biosynthesis, stomatal movement, and it can also reverse the effect of ABA on leaf abscission. In addition, it also enhances photosynthetic rate, photosynthetic pigments, and can modify the activity of important enzymes (29). In this study, we suspect that SA may inhibit the formation of garlic green pigment by lowering the respiratory rate.

Based on proteomic analysis, the DAPs of all samples involved in phenylpropanoid biosynthesis were up-regulated. Through further analysis, we found that the DAPs involved in phenylpropanoid biosynthesis were mainly cinnamyl-alcohol dehydrogenases and beta-glucosidases, which are the main enzymes of lignin synthesis (30). Therefore, the lignin content gradually increases during the storage of garlic at 4 °C. These results suggest that different pathways of phenylpropanoid metabolism have different effects on the formation of garlic green pigment.

The role of other metabolic pathways in the greening process of garlic

In addition to the above metabolic pathways, other pathways that may be associated with garlic greening were also significantly enriched. The genes in up-regulated profiles (21, 22, 24 and 25) and up-regulated DAPs of DS1/DS0 and Cluster1 DEGs were associated with glycolysis/gluconeogenesis and pyruvate metabolism (Fig. 2, Supplementary Table 3, 6). In addition, there were 3 and 2 hub genes in the antiquewhite1 and brown4 modules involved glycolysis/gluconeogenesis, respectively, and one hub gene in the brown4 module involved pyruvate metabolism (Supplementary Table 3, 6).

Glycolysis/gluconeogenesis provides energy and material for most metabolic pathways in the organism; thus, it is an indispensable part of the garlic greening process. K Adler and S Imai demonstrated that pyruvate plays an important role in the greening mechanism of garlic, which was further confirmed in this study. The genes in up- and down-regulated profiles and DEGs in Cluster4 were associated with fatty acid metabolism, and there was one hub gene in each of the brown4 and blue4 modules involved in fatty acid metabolism (Fig. 2, Supplementary Table 3, 6, 8). The structure, function, synthesis and metabolic pathways of fatty acids have previously been thoroughly studied, and excellent results have been achieved. However, the role of fatty acids in the mechanism of garlic greening has not been studied. This work demonstrates that multiple fatty acid pathways were involved in the process of garlic greening, indicating that fatty acids may play a significant role in the mechanism of garlic greening.

There were multiple genes in up-regulated profiles and up-regulated DEGs in all comparison groups involved in biosynthesis of amino acids. Five hub genes in the antiquewhite1 module and four hub genes in the brown4 module were related to biosynthesis of amino acids, suggesting that amino acids promote the formation of garlic green pigment. Garlic green pigment is comprised of yellow pigment and blue pigment. Many amino acids are involved in the blue color formation, such as glycine (Gly), asparagine (Asn), and glutamine (Gln), Alanine (Ala) and valine (Val) (17), therefore the amino acid metabolism pathway is involved in garlic greening.

Based on the above conclusions, a regulatory network for the formation of garlic green pigment was established. The TCA cycle increases the greening intensity of garlic by increasing the content of PBG. In the phenylpropanoid metabolic pathway, lignin may promote the formation of garlic green pigment, while SA, may inhibit it by inhibiting respiration. In addition, pyruvate plays a very important role in the process of garlic greening, which perhaps promotes the formation of garlic green pigment by participating in various metabolic pathways (Fig. 5).

Conclusions

In this study, we obtained high-quality reference transcriptomes of garlic by using a single-molecule long-read sequencer from Pacific Biosciences. The regulation mechanism of garlic green pigment formation has been further improved. Porphobilinogen (PBG) is the precursor of garlic green pigment. TCA cycle positively regulates the formation of garlic green pigment by increasing the content of PBG. Different pathways of phenylpropanoid metabolism have different effects on the formation of garlic green pigment, such as salicylic acid (SA) inhibits formation of garlic green pigment, lignin may positively regulate garlic greening.

Methods

Plant treatments and RNA isolation

The garlic stored in dark conditions for 0 (DS0), 10 (DS1), 15 (DS2) and 40 (DS3) days at 4 °C. When samples were collected, they were immediately frozen in liquid nitrogen, and stored at -80 °C until use. Total RNA from each sample was extracted using a GeneJET Plant RNA Purification Mini Kit (#K0801) according to the manufacturer's protocol.

Pacbio Library Construction And Single-molecule Sequencing

Garlic samples at DS0 and DS3 were selected for single-molecule sequencing. SMRT libraries were prepared according to the Isoform Sequencing protocol (Iso-Seq™). First, equal amounts of RNA (1 µg per sample) from 6 single plants were pooled together and 3 µg was used for reverse transcription (RT) using a Clontech SMARTer cDNA synthesis kit. Second, the optimal amplification cycle number was utilized to generate large-scale double-strand cDNA. Third, after dsDNA was obtained, a BluePippin System was used to select the DNA size for subsequent re-amplification. Finally, three libraries (1–2, 2–3, and 3–6 kb) generated by the products were subjected to an Iso-Seq SMRTBell library preparation (<https://pacbio.secure.force.com/SamplePrep>). Three libraries were sequenced on the PacBio RSII platform using P5-C3 chemistry.

Data analysis of PacBio sequencing reads

The sequence data was analyzed using SMRT analysis software (www.pacb.com/products-and-services/analytical-software/devnet/). Circular consensus sequence (CCS) reads were generated by using the following parameters: min_predicted_accuracy, 0.8; max_length, 15000; min_length, 50; min_passes, 0; max_drop_fraction, 0.8. According to the parameter minSeqLength 100, CCS reads were classified into full-length and non-full-length sequences. Non-redundant sequences were obtained using CD-HIT software (31), which were identified as transcripts. The coding domains (CDs) of the full-length transcript were predicted by TransDecoder software, and a total of 36,776 transcripts with protein-encoding functions were obtained. The CDs were subjected to functional annotation by searching against six public databases, including eukaryotic ortholog groups (KOG), National Center for Biotechnology Information (NCBI) non-redundant (NR) protein sequences, Kyoto Encyclopedia of Genes and Genomes ortholog (KEGG), Swiss-Prot protein, Gene Ontology (GO), and protein family (PFAM) databases.

Illumina RNA-sequencing library construction

The total RNA from all garlic samples was subjected to Illumina RNA sequencing individually. The total RNA from all garlic samples was subjected to Illumina RNA sequencing individually. Firstly, the purified RNA was randomly broken into short fragments using fragmentation buffer to construct cDNA libraries. Second, the short-fragmented RNA was used as a template to synthesize single-stranded cDNA with six-base random primers (Random hexamers), which generates double-stranded cDNA in the presence of buffer, dNTPs, RNaseH and DNA Polymerase I. Finally, after adapter was added to purified double-stranded cDNA, PCR amplification was performed to obtain the final sequencing library.

Expression level analysis of genes

The expression of each gene in all garlic samples was quantified in Bowtie2 (32) using the transcriptome generated by PacBio sequencing as reference transcriptome. To analyze the expression level of each gene in all samples, RSEM (33) was performed to estimate the expected number of fragments per kilobase of transcript sequence per million base pairs sequenced (FPKM).

Protein extraction and quantification

Proteins from all garlic samples were extracted using cell lysis buffer (P0013), which contains 20 mM Tris (pH 7.5), 150 mM NaCl, 1% Triton X-100, and various inhibitors such as sodium pyrophosphate, EDTA and β -glycerophosphate. The protein in all garlic samples was quantified using label-free quantitation, in which the relative content of a corresponding protein was obtained using MaxQuant software to analyze quantitative information of peptides generated by LC-MS detection. Differentially abundant proteins (DAPs) were generated using Perseus software (34).

Extraction of green pigment from garlic

(1) Garlic peeled, crushed (2) added 2% citric acid (3) 80 °C water bath heated for 30 min (4) methanol leached (5) filtration (anhydrous Na₂SO₄ dehydration) (6) petroleum ether extraction degreased (7) 35 °C rotary evaporation concentration (8) added 95% Ethanol precipitation protein (9) centrifugation (10) suction filtration (11) rotary evaporation to remove some solvent (12) crude extraction concentrate.

Garlic green intensity definition and detection method

A Total of 20.00 g garlic bulb was weighed, then mashed after germination. 2% citric acid was added then the sample was stirred evenly and heated at 80 °C in a water bath for 30 min, then cooled to room temperature. Samples were then made up to 50 mL with 95% ethanol, and the leaching was performed for 24 h at 4 °C. The supernatant was measured for absorbance A at 440 nm and 590 nm after filtration.

Definition of green intensity: Green intensity = $A_{590(440)} \times 10$.

Treatment of garlic with sodium azide

A total of 1 g garlic bulb was weighed and soaked in sodium azide solution at a concentration of 0.1 mol/L or 0.5 mol/L for 24 h at room temperature. Garlic was also soaked in deionized water as a control. After draining, samples were stored at 4 °C, and the green intensity of garlic was measured at DS0, DS1, DS2 and DS3.

Effect of salicylic acid on the greening of garlic

Equal sizes of garlic were picked, adding different concentrations of salicylic acid (SA) according to the quality of garlic after peeling and slicing. Garlic extracts without salicylic acid were used as a control. The green intensity of the garlic was measured after treatment with different concentrations of salicylic acid.

The pyrrolyl compound porphobilinogen (PBG) content detection

Total 1.00 g germinated garlic bulbs as added to 5 mL of phosphate buffer pH 6.8, and the mixture was centrifuged at 12000 r/min for 10 min. The supernatant was added to Ehrlich's reagent for color development, and the absorbance was measured at 553 nm.

Method for determining organic acid content

Organic acid extraction: 1.00 g of garlic bulb was added to 2 mL of deionized water then ground to a slurry of germinated garlic bulbs. 3 mL deionized water was used to wash the sample into a centrifuge tube before centrifugation at 9000r/min for 30 min. The supernatant was filtered through a microporous membrane (0.45 Mm) before the determination of organic acids in the TCA cycle.

Detection method: UV detector, Alltima C18 Column (250 mm × 4.6 mm, 5 μm) column, flow rate: 0.4 mL/min, injection volume: 20 μL, detection wavelength: 214 nm, column temperature: 30 °C, mobile phase: 5% methanol 0.01 mol/L H₃PO₄-KH₂PO₄ buffer (pH = 2.35).

Construction of a weighted gene co-expression network

After the data was standardized or processed accordingly, weighted gene co-expression network analysis (WGCNA) was performed (35). First, in order to ensure the results of weighted gene co-expression network construction were reliable, all samples were clustered and outlier samples were filtered. Second, the soft threshold power (soft power) in accordance with standard scale-free networks was selected, by which the adjacencies between genes of all samples were calculated by a power function. Then, the adjacency was transformed into a topological overlap matrix (TOM), and the corresponding dissimilarity (1-TOM) was calculated and the scale-free topology network was constructed from the dissimilarity values. Gene modules in the resulting networks were accomplished with the dynamic tree cut method (<http://www.coxdocs.org/doku.php?id=perseus:start>).

GO term and KEGG pathway enrichment analyses

Gene Ontology (GO, www.geneontology.org) is an international standard classification system for gene function. It facilitates study of the distribution of genes in Gene Ontology groups and elucidation of the functional representation of genes in different modules. The analysis method of GO enrichment was top

GO and the GO term enrichment p-values were calculated using the Fisher's exact test in the TopGO R package (<http://www.bioconductor.org/packages/release/bioc/html/topGO.html>).

Different genes in vivo coordinate with each other to perform their biological function, and KEGG pathway analysis helps to further understand the biological functions of genes. KEGG is the main public database for pathways. A pathway's significant enrichment analysis uses KEGG Pathway as a unit to apply hypergeometric tests to find a pathway that is significantly enriched in significant differentially expressed genes compared to the entire genomic background (36).

$$p = 1 - \sum_{i=0}^{m-1} \binom{m}{i} \binom{N-m}{n-i} / \binom{N}{n}$$

N is the total number of genes and n is the number of differentially expressed genes in N. M is the number of genes annotated as a particular Pathway and m is the number of differentially expressed genes annotated as a particular Pathway.

Declarations

Acknowledgments

We are grateful to all co-authors who participated in the studies mentioned in the text that were published by our groups. All authors declared no conflict of interest.

Funding

This work was supported by the National Key R&D Program of China (2016YFD040024); Natural Science Foundation of China (31972000;31772323) and Natural Science Foundation of Shandong Province (ZR2016CM33).

Availability of data and materials

The dataset supporting the conclusions of this article is available in the NCBI Sequence Read Archive database with accession number of SRR8865936 and SRR8865937 for PacBio data, and SRR7196287, SRR7196288, SRR7196289, SRR7196290, SRR7196291, SRR7196292, SRR7196293, SRR7196294, SRR7196295, SRR7196296, SRR7196297, and SRR7196298 for Illumina data.

Authors' contributions

NYL, XML and XDS contributed to the study design, data collection, and data analysis, wrote the first draft, and revised the manuscript. XGQ was the supervisor of the project and contributed to the study

design, data analysis, and manuscript revision. All authors reviewed and accepted the content of the final manuscript.

Consent for publication

Not applicable.

Competing interests

The authors declare that they have no competing interests.

Author details

¹Key Laboratory of Food Processing Technology and Quality Control in Shandong Province, College of Food Science and Engineering, Shandong Agricultural University, No. 61, Daizong Road, Tai'an, Shandong 271018, PR China. ²State Key Laboratory of Crop Biology, College of Horticulture Science and Engineering, Shandong Agricultural University, No. 61, Daizong Road, Tai'an, Shandong 271018, PR China

References

1. Tattelman EP. Health effects of garlic. *American Family Physician*. 2005;72(1):103-6.
2. Milner JA. Garlic: Its Anticarcinogenic and Antitumorigenic Properties. *Nutrition Reviews*. 2009;54(11).
3. Banerjee SK, Maulik SK. EFFECT OF GARLIC ON CARDIOVASCULAR DISORDERS: A REVIEW. *Nutrition Journal*. 2002;1(1):4-.
4. Rahman K, Lowe GM. Garlic and Cardiovascular Disease: A Critical Review. *Journal of Nutrition*. 2006;136(3).
5. Miccolis V, Monteleone M, Candido V. Garlic Bulbing as influenced by pre-planting low temperature storage. *Advances in horticultural science*. 2000;14(4):159-68.
6. Cantwell M, Kang J, Hong G. Heat treatments control sprouting and rooting of garlic cloves. *Postharvest Biology and Technology*. 2003;30(1):57-65.
7. Ichikawa M, Ide N, Ono K. Changes in organosulfur compounds in garlic cloves during storage. *Journal of Agricultural and Food Chemistry*. 2006;54(13):4849-54.
8. Lee JY, Lee HJ, Kyung KH. Green pigmentation characteristics of amino compounds with thiosulfinates obtained from garlic (*Allium sativum* L.). *Food Science and Biotechnology*. 2013;22(2):373-81.
9. Lee EJ, Cho J, Lee SK. Effect of Storage Temperature and S-Alk(en)yl-L-Cysteine Sulfoxide on Green Pigment Formation in Crushed Garlic (*Allium sativum* L.) Cloves. *Horticulture Environment and Biotechnology*. 2006;47(4):165-73.

10. Lee Hj, Shin Y, Kyung K, Chung C-H. Green pigmentation and pH change of homogenized garlic. *Food Science and Biotechnology*. 2013;23:121-4.
11. Joslyn M, Sano T. The formation and decomposition of green pigment in crushed garlic tissue. *Journal of Food Science*. 2006;21:170-83.
12. Lukes TM. Factors Governing the Greening of Garlic Puree. *Journal of Food Science*. 1986;51(6):1577-.
13. Cho J, Lee EJ, Yoo KS, Lee S, Patil BS. Identification of Candidate Amino Acids Involved in the Formation of Blue Pigments in Crushed Garlic Cloves (*Allium sativum* L.). *Journal of Food Science*. 2009;74(1).
14. Kim WJ, Cho JS, Kim KH. Stabilization of ground garlic color by cysteine, ascorbic acid, trisodium phosphate and sodium metabisulfite. *Journal of Food Quality*. 2007;22:681-91.
15. Thuwapanichayanan R, Prachayawarakorn S, Soponronnarit S. Heat and moisture transport behaviour and quality of chopped garlic undergoing different drying methods. *Journal of Food Engineering*. 2014;136:34-41.
16. Shin YK, Kyung KH. Cysteine reacts to form blue-green pigments with thiosulfinates obtained from garlic (*Allium sativum* L.). *Food Chemistry*. 2014;142:217-9.
17. Imai S, Akita K, Tomotake M, Sawada H. Identification of two novel pigment precursors and a reddish-purple pigment involved in the blue-green discoloration of onion and garlic. *Journal of Agricultural and Food Chemistry*. 2006;54(3):843-7.
18. Imai S, Akita K, Tomotake M, Sawada H. Model Studies on Precursor System Generating Blue Pigment in Onion and Garlic. *Journal of Agricultural and Food Chemistry*. 2006;54(3):848-52.
19. Li N, Qiu Z, Lu X, Shi B, Sun X, Tang X, et al. Comparative Transcriptome Analysis of Temperature-Induced Green Discoloration in Garlic. *Comparative and Functional Genomics*. 2018;2018:1-8.
20. Ernst J, Barjoseph Z. STEM: a tool for the analysis of short time series gene expression data. *BMC Bioinformatics*. 2006;7(1):191-.
21. Hu D, Zhang J, Zhang T, Dong J, Leng X, Zhao G. Characterization of yellow pigments formed on reaction of 2-(1H-pyrrolyl)carboxylic acids with pyruvic acid in garlic greening model systems. *Food Research International*. 2010;43(3):915-21.
22. Wang D, Nanding H, Han N, Chen F, Zhao G. 2-(1H-Pyrrolyl)carboxylic Acids as Pigment Precursors in Garlic Greening. *Journal of Agricultural and Food Chemistry*. 2008;56(4):1495-500.
23. Battersby AR. Tetrapyrroles: the pigments of life. *Natural Product Reports*. 2000;17(6):507-26.
24. O'Leary BM, Plaxton WC. Mechanisms and Functions of Post-translational Enzyme Modifications in the Organization and Control of Plant Respiratory Metabolism. In: Tcherkez G, Ghashghaie J, editors. *Plant Respiration: Metabolic Fluxes and Carbon Balance*. Cham: Springer International Publishing; 2017. p. 261-84.
25. Tcherkez G, Gauthier PPG, Buckley T, Busch FA, Barbour MM, Bruhn D, et al. Leaf day respiration: Low CO₂ flux but high significance for metabolism and carbon balance. *New Phytologist*.

- 2017;216(4):986-1001.
26. Reich PB, Tjoelker MG, Machado J, Oleksyn J. Universal scaling of respiratory metabolism, size and nitrogen in plants. *Nature*. 2006;439(7075):457-61.
 27. Gramazio P, Prohens J, Plazas M, Andujar I, Herraiz FJ, Castillo E, et al. Location of chlorogenic acid biosynthesis pathway and polyphenol oxidase genes in a new interspecific anchored linkage map of eggplant. *BMC Plant Biology*. 2014;14(1):350-.
 28. Lee EJ, Cho J, Kim J, Lee SK. Green pigment in crushed garlic (*Allium sativum* L.) cloves: Purification and partial characterization. *Food Chemistry*. 2007;101(4):1677-86.
 29. Yusuf M, Hayat S, Alyemeni M, Fariduddin Q, Ahmad A. Salicylic Acid: Physiological Roles in Plants. 2013. p. 15-30.
 30. Van Acker R, Vanholme R, Storme V, Mortimer JC, Dupree P, Boerjan W. Lignin biosynthesis perturbations affect secondary cell wall composition and saccharification yield in *Arabidopsis thaliana*. *Biotechnology for Biofuels*. 2013;6(1):46-.
 31. Li W, Godzik A. Cd-hit: a fast program for clustering and comparing large sets of protein or nucleotide sequences. *Bioinformatics*. 2006;22(13):1658-9.
 32. Langmead B, Salzberg SL. Fast gapped-read alignment with Bowtie 2. *Nature Methods*. 2012;9(4):357-9.
 33. Li B, Dewey CN. RSEM: accurate transcript quantification from RNA-Seq data with or without a reference genome. *BMC Bioinformatics*. 2011;12(1):323-.
 34. Tyanova S, Cox J. Perseus: A Bioinformatics Platform for Integrative Analysis of Proteomics Data in Cancer Research. *Methods of Molecular Biology*. 2018;1711:133-48.
 35. Langfelder P, Horvath S. WGCNA: an R package for weighted correlation network analysis. *BMC Bioinformatics*. 2008;9(1):559-.
 36. Thorvaldsdottir H, Robinson J, Mesirov JP. Integrative Genomics Viewer (IGV): high-performance genomics data visualization and exploration. *Briefings in Bioinformatics*. 2013;14(2):178-92.

Supplementary Files Legend

Supplementary Table 1. Functional enrichment analysis of all genes obtained from single-molecule long-read sequencing.

Supplementary Table 2. Total co-expressed transcripts of DS0, DS1, DS2 and DS3 samples.

Supplementary Table 3. KEGG pathway enrichment analysis of genes in up- and down- regulated profiles ($p < 0.05$).

Supplementary Table 4. DEGs in DS1/DS0, DS2/DS0 and DS3/DS0.

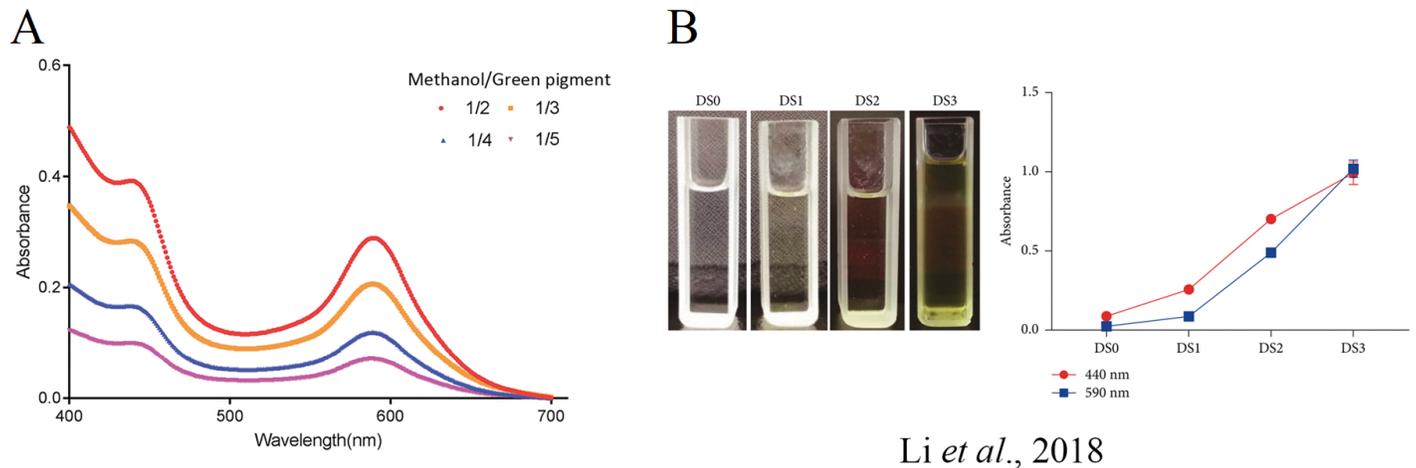
Supplementary Table 5. DAPs in DS1/DS0, DS2/DS0 and DS3/DS0.

Supplementary Table 6. KEGG pathway enrichment analysis of genes in DAPs.

Supplementary Table 7. Phenylpropanoid biosynthesis and TCA cycle involved in co-expression DAPs and DEGs.

Supplementary Table 8. Hub genes in antiquewhite1, brown4 and blue4 involved in KEGG pathways.

Figures



Li *et al.*, 2018

Figure 1

Detection of physiological characteristics of garlic green pigment. (A) Measuring absorption spectrum of garlic green pigment. (B) Appearance of pigment extracts prepared from DS0 to DS3 (left). Colorimetric measurement of pigment content at 440 nm and 590 nm (right).

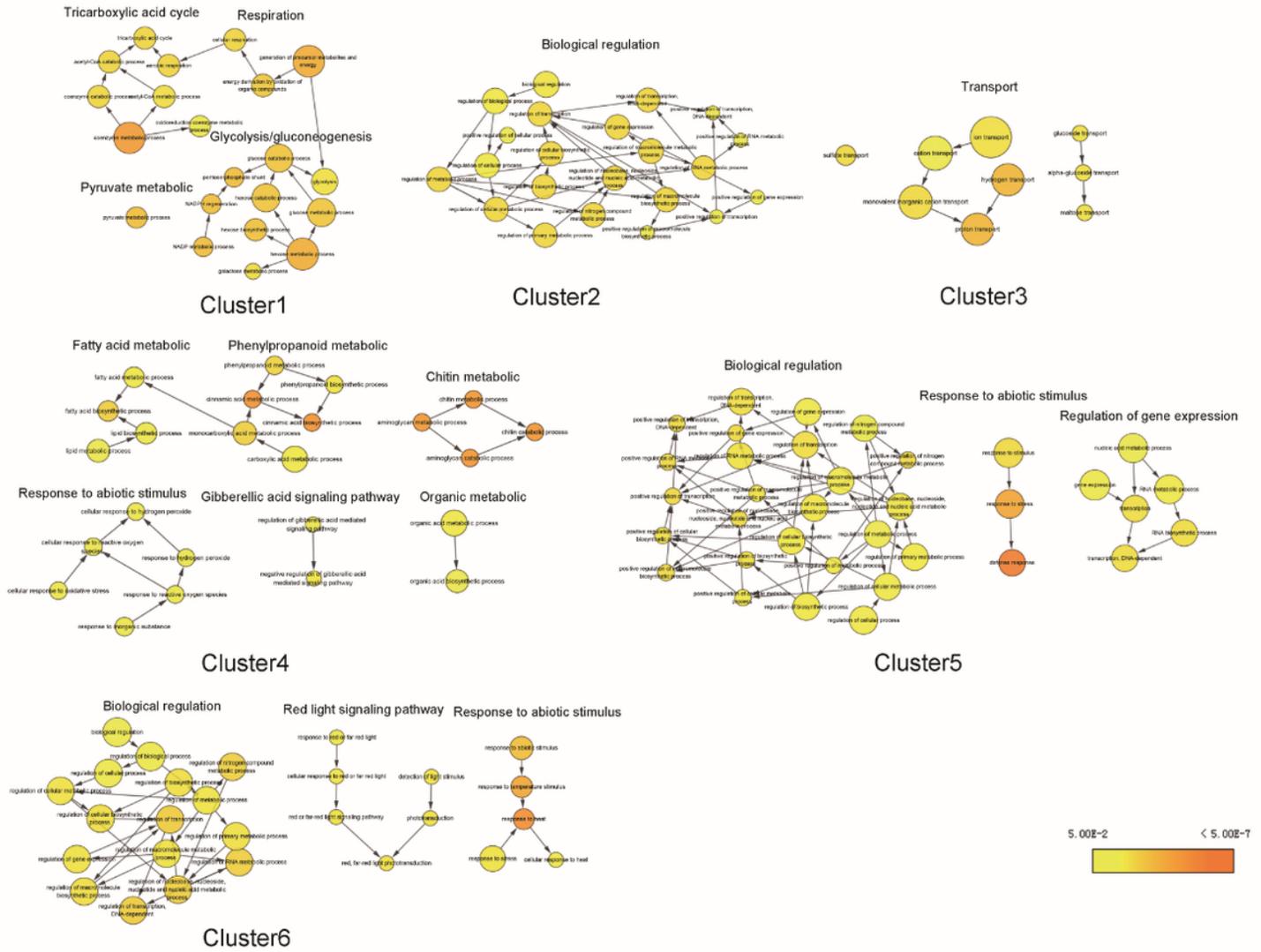


Figure 2

Enriched GO (BP) terms for the six gene clusters.

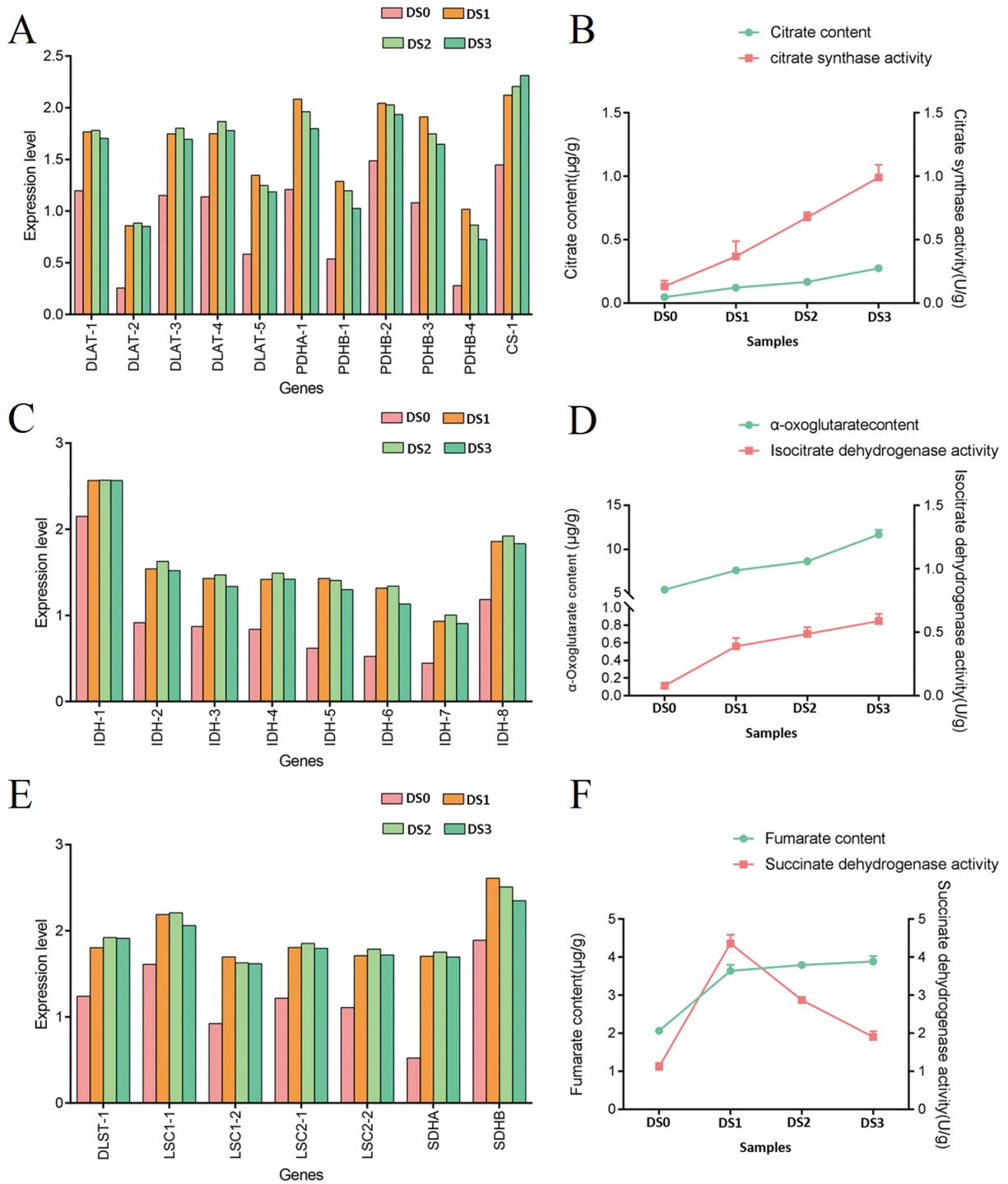


Figure 3

TCA promotes the process of garlic green. (A), (C), (E) The expression level of 51 unigenes encoding critical enzymes of TCA cycle. (B), (D), (F) The content of several key intermediate objects of Krebs cycle and the activity in critical enzymes of TCA cycle.

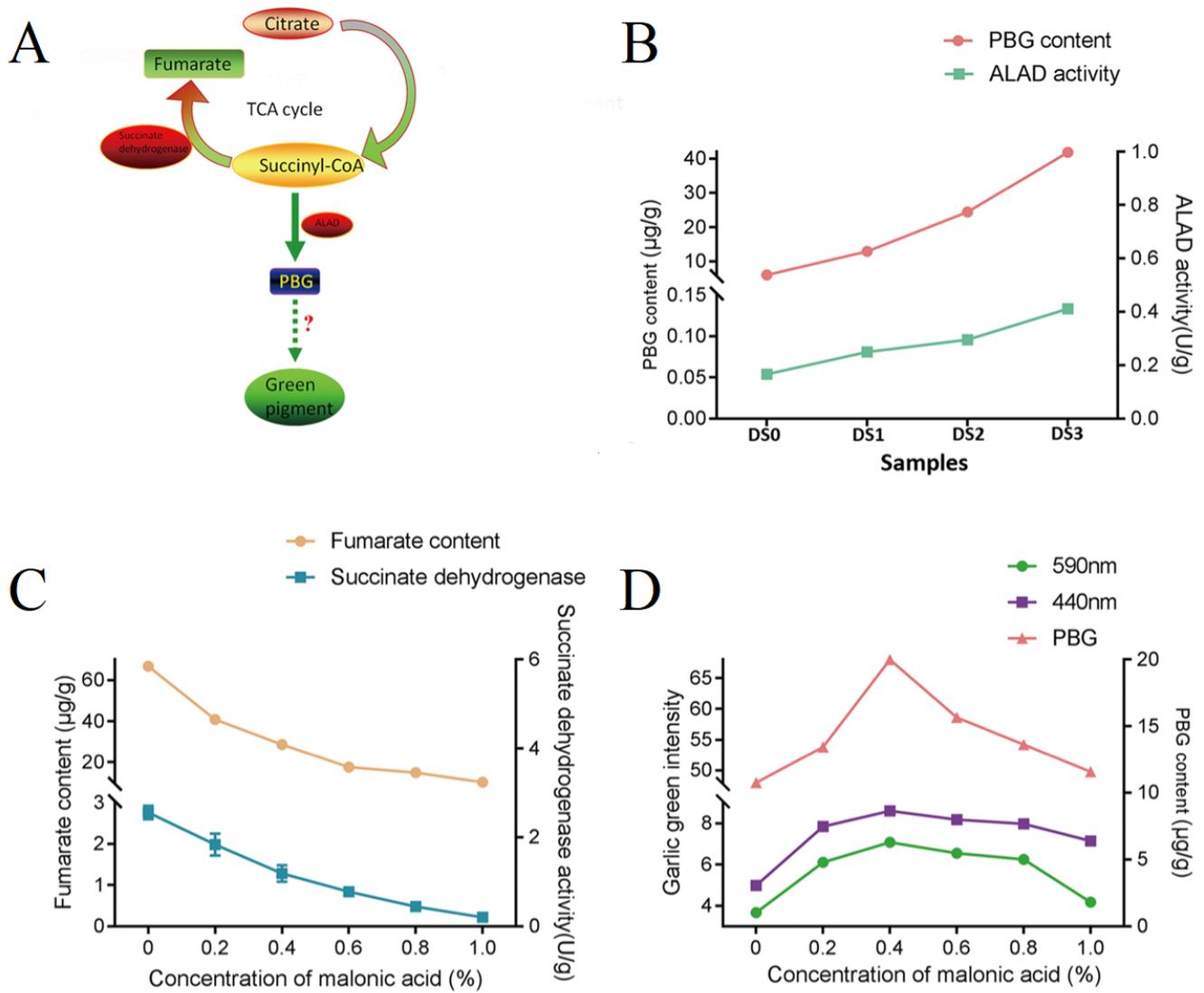


Figure 4

Porphobilinogen is a precursor of garlic green pigment. (A) PBG may promote the formation of garlic green pigment. (B) Changes of porphobilinogen (PBG) content and ALAD activity in all samples. (C) Changes of fumarate content and succinate dehydrogenase activity in garlic processed by malonic acid. (C) Changes of garlic green intensity and PBG content in garlic processed by malonic acid.

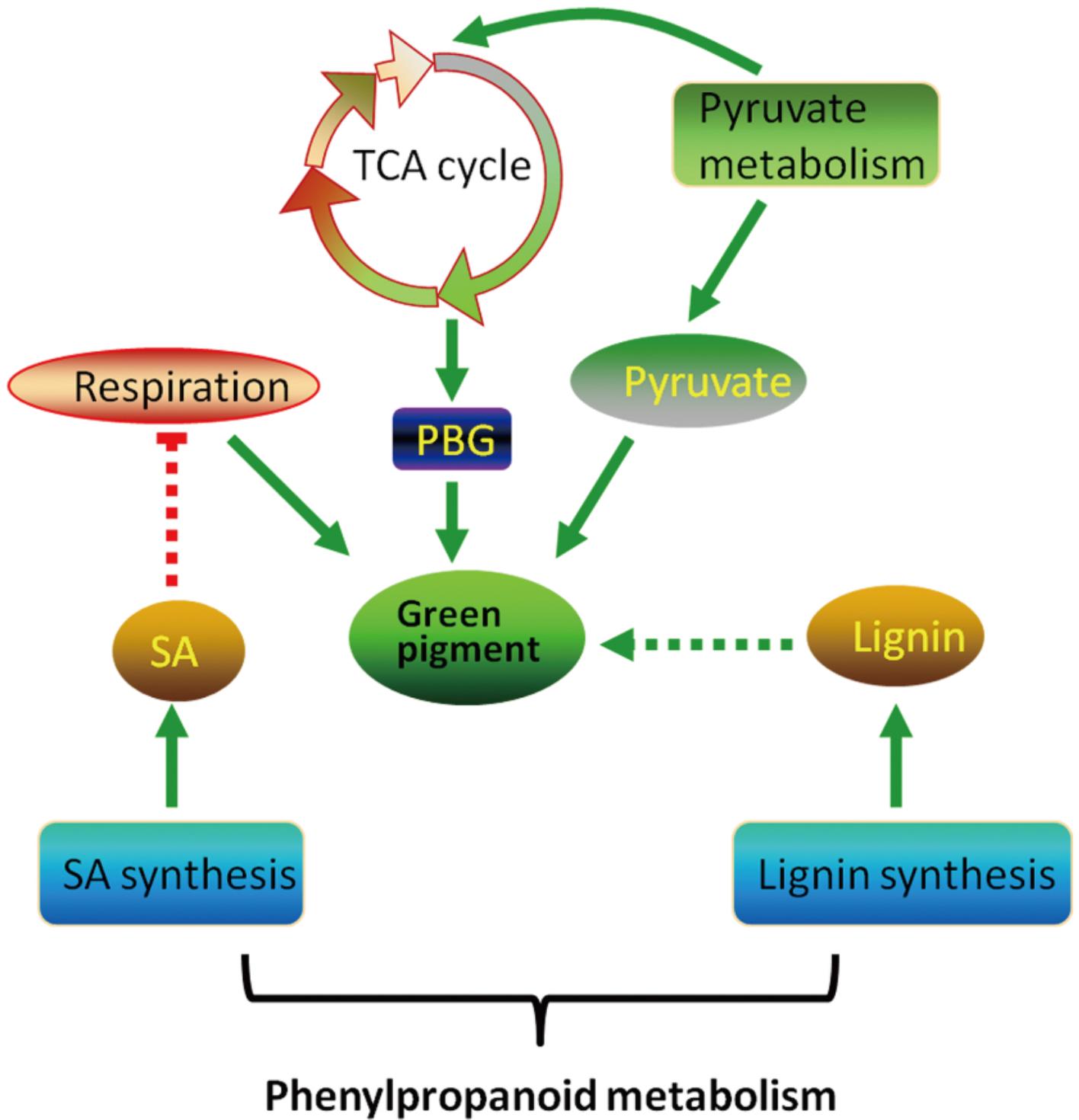


Figure 5

The regulation network for the formation of garlic green pigment.

Supplementary Files

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

- TableS2.xls
- TableS6.xls
- FigureS4.tif
- FigureS2.tif
- TableS7.xls
- FigureS6.tif
- FigureS3.tif
- FigureS1.tif
- FigureS5.tif
- SupplementaryMaterial.docx
- TableS3.xls
- FigureS7.tif
- TableS5.xls
- FigureS10.tif
- FigureS9.tif
- FigureS8.tif
- TableS4.xls
- TableS8.xls
- TableS1.xls