Parallel analysis of global garlic gene expression and alliin content following leaf wounding

Xuqin Yang
Jiangsu Normal University
Yiren SU
Jiangsu Normal University https://orcid.org/0000-0002-6908-2580
jiaying wu
Jiangsu Normal University
wen wan
Jiangsu Normal University
Huijian chen
XuZhou Nuote chemical Co Ltd
Xiaoying cao
Jiangsu Normal University
Junjuan Wang
Jiangsu Normal University
Zhong zhang
XuZhou Nute Chemicals Co Ltd
Youzhi wang
Xuzhou Nuote Chemicals Co Ltd
Deliang Ma
Xuzhou Chemicals Co Ltd
Gary Gloake
The University of Edinburgh
JiHong Jiang (j18361391330@163.com)
Jiangsu Normal University https://orcid.org/0000-0002-0830-1487

Research article

Keywords: Garlic, Allium sativum, Transcriptomics, Alliin, Gene Prediction

Posted Date: September 1st, 2020

DOI: https://doi.org/10.21203/rs.3.rs-50719/v1

Version of Record: A version of this preprint was published at BMC Plant Biology on April 10th, 2021. See the published version at https://doi.org/10.1186/s12870-021-02948-0.

Abstract

Background

Allium sativum (garlic) is both an important food and medicinal plant of economic significance. This plant is rich in sulfides, especially alliin, which is a precursor for the synthesis of allicin. At present, there are few reports on the determination of alliin content in different parts of garlic under abiotic stress.

Results

Our data determining the accumulation of alliin in different organs showed that the content of alliin in garlic root was the lowest level recorded, while the content of alliin within a garlic bud was the highest level determined. Further, alliin levels decreased in mature leaves following wounding. Further, transcriptomic data generated over time following wounding of mature garlic leaves showed genes integral to the biosynthetic pathways of cysteine (CYS) and serine (SER) formation were significantly up-regulated.

Conclusions

This differential expression could underpin the accumulation of alliin and its precursors in garlic. Thus, our results provide a platform to help elucidate the biosynthetic pathway alliin biosynthesis.

Background

Garlic *(Allium sativum L.*), a diploid (2n = 2x = 16) plant species, is one of the most important members of the Allium genus whether for medicinal value or economic value. Garlic has been widely cultivated internationally for more than 5,000 years [1–3]. According to a recent study, garlic has properties that convey strong antioxidant propential[4], stabilize blood pressure[5], reduce cancer risk[6], provide cardiovascular protection[7], and reduced hyperlipidemia[8] among other therapeutic effects. Garlic roots, bulbs, leaves and sprouts are all important agronomically. Bulbs consist of several abnormal axillary buds and exhibit the characteristic clove shape of garlic, these bulbs are an important economic trait[2]. A key feature of garlic is that it produces secondary metabolites rich in sulfur, such as S-methyl-L-cysteine sulfoxide (MCSO, methiin) and S-propyl-L-cysteine sulfoxide (PCSO, propiin), S-trans-1-propenyl-L-cysteine sulfoxide (PECSO, isoalliin) and the most significant, S-allyl-L-cysteine sulfoxide (ACSO, alliin)[12]. These substances are also taste precursors, which regulate the taste and odor of garlic[13]. Allicin, derived from alliin, is another sulfur compound that is a key bioactive molecule in garlic[14, 15]. Various *in vivo* and *in vitro* studies have demonstrated the anti-apoptosis and anti-oxidation potential of allicin[16]. However, allicin is extremely unstable due to rapid decomposition to diallyl disulfides and sulfur dioxide[17]. Allicin is produced from alliin by the action of alliinase[18]. In garlic bulb cells these two molecules are separated:

allinase is located in the vacuole, while alliin is located in the cytoplasm. When cells are broken or damaged, allinase is released, converting alliin to allicin[19, 20].

Alliin is a stable and odorless molecule[21], of significant nutritional and medicinal value [22]. For example, alliin is known to exhibit antioxidant[23] and anti-inflammatory[24] activity, promote cardiovascul ar function[25] and convey beneficial effects on some intestinal diseases[26]. Alliin is first synthesized in garlic leaves and subsequently transferred to garlic bulbs[27, 28]. Despite the importance of alliin to the commercial potential of garlic, the details of alliin biosynthesis remain to be fully elucidated.

Garlic has a large genome[29] and cultivars are sterile. Consequently, there has been no real classical breeding and genetic studies of garlic[30]. Thus, despite its agronomic importance, garlic remains largely undomesticated, which has hampered the commercial potential of garlic. The diploid garlic genome (2n = 2x = 16) is estimated to be 15.9G, 32 times the size of the rice genome and remains to be fully sequenced. Unlike genome analysis, transcriptome analysis has the advantages of high speed, low cost and no limitation of genome complexity[31]. In recent years, transcriptome has been used for the correlation analysis of traits, extending the study of genetic association to many species, especially complex polyploid species[32–34]. Recently, RNA-Seq has been used for transcript profiling in various non-model plant species, including: Salvia miltiorrhiza[35], Pinellia ternata[36], Cicer arietinum L[37].

Recently, RNA-Seq has been used for transcript profiling of garlic. The garlic bulbs correlation transcriptome research[2], reported by related studies, showed that twenty-two transcripts were identified as candidate transcripts with complex interactions among them. In addition, transcriptome revealed that the corresponding genetic changes of allinase during the whole cycle[38]. The transcription of two enzymes was the highest during sprouting. The growth of garlic flowers and pollen was analyzed by transcriptome and proteome[29], proposing the potential molecular markers for male fertility and sterility in garlic. However, according to detailed studies of wound treatment and correlation transcriptome, the content of alliin in leaf has not been reported. Alliin is commonly found in garlic, we need to further optimize the method of determining the alliin content.

However, transcriptomics analysis of garlic post wound treatment of leaf tissue correlated with the production of alliin has not been previously reported. Alliin is commonly found in garlic, we need to further optimize the method of determining the alliin content. According to the study and the transcriptome corresponding pathways, we speculated the pathway of alliin synthesis. To establish the molecular basis for the synthesis mechanism of alliin, we found the differential genes expression and enrichment of the corresponding pathways, and inferred the unknown pathway corresponding to proteins and genes through the data.

Results

By Sykam S-433d measuring the content of alliin

Sykam S-433d was used to determine the alliin content of different tissues and wounded leaves of garlic. Our data showed that the content of alliin in garlic root was the lowest and the content of alliin in garlic bulbs was the highest. In addition, the content of alliin decreased over time post wounding (Fig. 1c).

Transcriptomic analysis of garlic

The sequencing of twelve garlic leaf cDNA libraries generated raw reads was qualified, and the adapter removed. Approximately, 0.03% of raw reads were removed post filtering of adapter sequences, including low quality reads and short reads. This resulted in 34855947 clean reads at 0 hours post-wounding, 30830543 at 3 hours post-wounding, 29001419 at 6 hours post-wounding and 34278922.33 in 12 hours post-wounding (Additional file 1). For each sample there were three biological replicates: Pearson's Correlation Coefficient of three replicates × four garlic biological samples. The results showed that the sequencing data was very reproducible (Fig. 2a). At the same time, it can be seen from the control group and each treatment group that the comparison of individual genes was large (Fig. 2b). The data indicated that there is a significant difference in reliability among genes. All reads of the 12 samples were assembled using Trinity[39]. The result showed that 194,627 transcripts (N50:1,667) had an average length of 1157.22 bp. The total number of unigenes (N50:1,394) was 94,144, with a mean length of 933.08 bp. The assembly and size distribution of unigenes and transcripts are displayed (Fig. 2c). For transcript data analysis, a read length of > 2000 bp was found in 15.38% of transcripts, 28.78% of transcripts were less than 1000 bp. A minimum read length of 200–300 bp about occurred in 11.56% transcripts. The reads length of unigenes of more than 2000 bp, accounted for 10.36% of the total. The 16,882 unigenes were annotated in the Swiss-Prot database. Nr database displayed 26,662 unigenes and the Pfam database found 19,206 unigenes, with significant similarities with known proteins. In the KEGG and KOG databases, 9677 and 16,197 unigenes were annotated (Fig. 2d). The distribution of unigene species annotation is shown (Fig. 2e).

Unigene classification

Unigenes were divided into KEGG metabolism and signaling pathways from five aspects. These three pathways, including 'Protein processing in endoplasmic reticulum' (75 unigenes), 'Plant hormone signal transduction' (92 unigenes), and 'Photosynthesis'(56 unigenes), are the most abundant KEGG pathways (Fig. 3a), detail data other see(Additional file 2). 1714 unigenes were up regulated and 1135 unigenes down regulated in the T1/3-T4/6 comparison post-wounding. T1/3-T10/12 identified 688 highly-expressed unigenes of up regulated and 1375 highly-expressed unigenes of down regulated (Fig. 3b) (Additional file 3). Venn diagrams were used to represent the number of differentially identified genes for each treatment compared with the control (Fig. 3c). By this standard of ($P < 0.05$) in the pathways. Compared with T1/3-T4/6, we sought four significant enrichment pathways. In the T1/3-T7/9 group, four pathways were also significantly enriched. While in the T1/3-T10/12 group, only three pathways were significantly enriched (Fig. 3d). Unigenes are contrasted to the COG database, to predict possible functions and orthology classify. A total of 12,375 sequences in the comparison of sample T1/3-T4/6 were assigned to twenty-five COG categories. There is usually only one category of functional prediction:

general functional prediction (2191;17.71%) represented the largest group. The next three are: Replication, recombination and repair (1350;10.91%). Transcription (1099;8.88%)Signal transduction mechanisms(909;7.35%). Additional samples of the COG classification are detailed in (Additional file 4).

Expression of DEGs related to alliin biosynthesis

Currently, the details of alliin biosynthesis remain unclear[40]. But, some of the corresponding molecules have been inferred including the precursors, glutathione, glycine, serine[41], cysteine[42] and sulfur[43] which feed into a series of hydrocarbylation, alkylation and oxidation reactions (Fig. 3e). 'Sulfur compound biosynthetic process' (GO:0044272), 'sulfur amino acid metabolic process' (GO:0000096), 'cysteine biosynthetic process' (GO:0019344), 'L-serine biosynthetic process' (GO:0006564), 'glutathione peroxidase activity'(GO:0004602), are five GO terms related to alliin biosynthesis, which provided a basis to assess different genes related to alliin biosynthesis. Meanwhile, the expression of alliin biosynthesis pathway was also treated. In the following pathways. For example the sulfur compound biosynthetic process, the genes (c131842.graph_c0c93579.graph_c0c95474.graph_c0) encoding the sulfur compound are expressed at high levels in the sample. Secondly, for the cysteine biosynthetic process steps, a series of genes are most of highly expressed in sample:(c102143.graph_c0c107612.graph_c1c109017.graph_c1c111121.graph_c0c111539.graph_c0c114448.graph_c0c116925.graph_c0) (Additional file 5).

Transcriptomic analysis reveals the differences cysteine(CYS) pathway-related genes

Through the analysis of transcriptomes, we observed the situation of differentially expressed genes significant enrichment and the corresponding pathways about eight cysteine-related GO terms were found: 'D-cysteinecatabolicprocess'(GO:0019447),'cysteine biosynthetic process'(GO:0019344), 'cysteine biosynthetic process from serine'(GO:0006535), 'cysteine desulfurase activity' (GO:0031071),'D-cysteine desulfhydrase activity' (GO:0019148), 'glutamate-cysteine ligase activity' (GO:0004357), 'peptidyl-cysteine modification' GO:0018198'cysteine metabolic process'(GO:0006534), were confirmed. Among them, the significant values of GO:0019344 in the T1/3-T4/6 and T1/3-T7/9 and T1/3-T10/12 comparisons were 0.00168 and 0.0013 and 0.00012, respectively. For GO:0006535, the significant values in the T1/3-T4/6 and T1/3-T7/9 and T1/3-T10/12 comparisons were 0.46558 and 0.45829 and 0.23682, respectively. Corresponding values of other pathways is in (Additional file 6). In all sample comparisons, there was a significant difference in (GO:0019344) (Fig. 4a) that shown more significant the enrichment. Although the significance value of (GO:0006535) was not obvious, transcriptome data analysis showed that the value is a downward trend, indicating an upward trend of difference significance enrichment in the(GO:0006535).The alliin synthesis pathway includes the serine to cysteine pathway, indicating changes of related differential genes. With the wound exposure time of garlic leaf, the enrichment of different genes in cysteine synthesis was more and more significant. The change of cysteine was accordant and verified that cysteine was involved in alliin synthesis. At the same time, we analyzed all the

differential genes related to cysteine and drew a heat map, from which we could clearly see the changes of genes in different samples (Fig. 4b) (Additional file 7).

Expression of DEGs related to Serine pathway

Serine is thought to play an important role in the alliin synthesis pathway. From the perspective of transcriptome analysis, the annotation significance of different genes found the corresponding pathway about seven serine-related GO terms. ‘protein serine/threonine/tyrosine kinase activity’(GO:0004712), ‘protein tyrosine/serine/threonine phosphatase activity’ (GO:0008138), ‘L-serine biosynthetic process’ (GO:0006564), ‘L-serine metabolic process’ (GO:0006563), ‘serine-type endopeptidase inhibitor activity’ (GO:0004867), ‘serine family amino acid biosynthetic process’ (GO:0009070), and‘D-serine metabolic process’ (GO:0070178), were identified. At length, the significant values of GO:0008138 in the T1/3-T4/6 and T1/3-T7/9 and T1/3-T10/12 comparisons were 0.01758 and 0.01673 and 0.01689. However, GO:0004712 only in the T1/3-T4/6 and T1/3-T7/9 respectively were 0.01939 and 0.01911. Corresponding values of other pathways (Additional file 8). Significant differences for GO:0008138 were observed in the all sample comparisons (Fig. 4c). Heat maps of SER metabolism-related genes and signaling pathway related genes were analyzed in different samples (Fig. 4d) other detail figure see (Additional file 9).

Sulfur-related differences genes in transcriptomic analysis

Sulfur compounds are important organic compounds in garlic. Alliin itself is a sulfur compound. There are many sulfur compounds in the synthesis pathway of alliin. The annotation significance of different genes found the corresponding pathway about six sulfuret-related GO terms,including ‘sulfate transmembrane transporter activity’ (GO:0015116), ‘sulfate transport’ (GO:0008272), ‘sulfur compound transport’ (GO:0072348), ‘ligase activity, forming carbon-sulfur bonds’ (GO:0016877), and‘sulfur compound transmembrane transporter activity’ (GO:1901682), ‘sulfur compound metabolic process’(GO:0006790). The significant values of GO:0015116 in the T1/3-T4/6 and T1/3-T7/9 and T1/3-T10/12 comparisons were 0.13427 and 0.13008 and 0.03515. However, the significant value of GO:0016877 only in the T1/3-T4/6 and T1/3-T7/9 were 0.02281 and 0.02219, respectively. Corresponding values of other pathways (Additional file 10). Significant differences for GO:0015116 were observed in the all sample comparisons (Fig. 4e). Heat maps of Sulfuret-related genes and signaling pathway related genes were analyzed in different samples(Fig. 4f) other detail figure see (Additional file 11).

Identifications of TFs families in alliin

Many TFs play important roles in the biosynthesis of alliin. In this study, 452 presumed TF coding genes in 47 major TF families were analyzed. Among them, the BHLH family (40 genes), FAR1 family (36 genes) and NAC family (36 genes) contain a large number of TFs. In order to advance the process of alliin biosynthesis, we screened the key regulatory factors and predicted 8 transcription factors (Additional file 12). Looking for regulatory genes is prepared for the subsequent work of alliin synthesis pathway.

Discussion

The key sulphoxide, allicin, is produced from the precursor alliin through the action of alliinase. Our measured results showed that the amino acid composition of different tissue parts and wound leaves contains a lot of amino acid including: cysteine, serine, alliin. Furthermore, the consequence showed that the content of alliin was significantly different with the time of wound varied. According to our experimental data, the content of allin is the lowest in the root and the highest in the inner bundle and leaf. According to previous study[27], the conclusion is consistent. At the same time, our study shows that the content of alliin gradually decreases with the prolongation of wound time. It can be speculated that when leaves are damaged, cells break up and alliin reacts with alliinase to synthesize allicin. Under normal conditions alliin precursor is being degraded producing allicin. Analysing these genes may help us understand the allicin synthesis pathway.

In the comparison of T1/3-T4/6□T1/3-T7/9 and T1/3-T10/12, we obtained a large number of DEGs, of which 12 major pathways were enriched. We listed the up-regulated and down-regulated genes post-wounding in each treatment group and control group (Fig. 3b, Table 1). In our study, the protein processing process was significantly changed in T1/3-T4/6. There was no significant change in the comparison of T1/3-T7/9 and T1/3-T10/12. With the change of the differential expression of alliinase, we can concluded that alliin could happen to change. The differential expression of these important pathways may provide some important related genes for the molecular mechanism of allicin and its precursor.

CYS is an effective substance for synthesizing alliin. The content of CYS increases with the increasing of the corresponding yield of alliin. Data show that a large number of DEGs are enriched in ‘cysteine biosynthetic process’. According to the comparison of T1/3-T4/6□T1/3-T7/9 and T1/3-T10/12. It proves that the enrichment of CYS is increased over time post wounding. Meanwhile, ‘peptidyl-cysteine modification’ is another term which is quite obvious, when the content of CYS changes in the process of synthesis of alliin, CYS needs to make corresponding structural changes, which will change the peptide group and transform it into the next substance. Through the above two and other GO terms related to CYS, we make the heat map for the corresponding differential genes, providing a basis for the search of CYS related genes in the subsequent alliin synthesis pathway.

Garlic contains the substances of Sulfur [48] and SER[41]. Sulfur plays an important role in the synthesis of CSOs. Moreover sulfur is an important component of the flavor composition of allium plants. The basis for the formation of flavor are plants at the early stage by absorbing and assimilating sulfur, and alliin itself is sulfide [49]. According to the analysis, the differential genes were mainly enriched in the ‘sulfur compound transmembrane transporter activity’ and ‘the sulfate transmembrane transporter activity’ that the significance was relatively high in all samples. It was inferred that when the cytomembrane was broken after treatment, the sulfur compound (alliin) transmembrane transporter content will change in the middle transmembrane. Because the process of alliin and the corresponding sulfide’s releasing process need the corresponding receptor binding to transport. Similarly, the‘sulfur compound transport’is

significant in T1/3-T10/12. We guess when sulfur compound binding transmembrane transporter then the corresponding sulfur compound will transport, leading to the improve of transport activity to participate the allicin synthetic in garlic leaves. About SER were enriched in ‘protein serine/threonine/tyrosine kinase activity’ and ‘protein tyrosine/serine/threonine phosphatase activity’. In particular, these two enzymes were opposite of function. We infer that phosatase is hydrolyzed to yield numerous free hydroxyl and ions. The corresponding substances will happen hydroxylation or phosphorylation through kinase in the synthesis of alliin. It will promote the synthesis of alliin and precursor substances. In conclusion, firstly, the related precursor substances structure of alliin synthesis are modified. Secondly, reaction achieves new substances. Finally, the matter was bound with receptors then transported to cells .

Conclusions

At present, the synthesis pathway of alliin is still unclear. Researchers through the study of radioactive tracers have achieved the precursor substances in the corresponding synthesis pathway, mainly cysteine, serine and sulfide.

The differential expression of alliin synthesis pathway related genes, including CYS synthesis and metabolism, SER synthesis and enzyme activity, Sulfur formation and transport. Our results are helpful to understand the related regulation genes change of alliin synthesis precursor substances. At the same time, the related genes could provide a theoretical basis and a favorable molecular foundation for further study of the alliin synthesis pathway .

Methods

Plant materials and RNA extraction

Garlic samples were collected form Pizhou, China (PW). PW garlic was cultivated in the test farm (Xuzhou city, Jiangsu province). The garlic was grown over 100 m^2.

Using vernier caliper to measure garlic leaf length, width, thickness when garlic matures, and ensuring the selected material phenotype keep similar. Materials, after 0 h, 3 h, 6 h, 12 h wound dispose, were all obtained. The materials selection of garlic root, clove, gailic inner bud, garlic sprout were collected,then we assembled corresponding plant samples(Fig. 1a), immediately frozen in liquid nitrogen, stored at 80 °C, until use. Total RNAs were extracted using the Plant Total RNA Isolation Kit according to its manual. By adding DNase I to the mixture to remove DNA contamination. Purified RNA was detected by 1% agarose gel electrophoresis. RNA was quantitatively detected by Implen nanometer photometer using Nanodrop with an RNA integrity number > 7.0.

Library preparation and transcriptomic analysis

The samples of approprite quality total RNA is prepared. The main process of library construction is as follows: magnetic beads containing Oligo (dT) are enriched with eukaryotic mRNA; The mRNA is randomly fragmented by Fragmentation Buffer; Acting as templates, the mRNA combined random hexamers primers to synthesize the first cDNA chain, then added the buffer, dNTPs, RNase H and DNA polymerase I to compound the second cDNA chain, and used AMPure XP beads to purify cDNA; Using purified double-stranded cDNA to repair and adding A-tails to connected to the sequencing beads. Then using AMPure XP beads to select the fragment size; Finally, the cDNA library was constructured according to PCR enrichione. When the library construction was completed, Qubit 2.0 and Agilent 2100 were used to detect the library concentration and Insert Size. To ensure the quality of the library, we use Q-PCR method to accurately quantify the effective concentration. After qualifying inspect library, we use HiSeq2500 for High-throughput sequencing. The reading length of sequencing was PE125. Clean Data that filter Raw Data is high quality data. It needs to be sequentially assembled. Trinity software[39] will assemble Clean Data. In genetic identification and expression analysis, reads of different species were assembled together [40] (Fig. 1b).

Functional annotation and analysis

BLAST of NCBI was used to compare unigene sequences with databases, such as Non-redundant (NR) protein, Swiss-prot, Gene Ontology (GO), Clusters of Orthologous Groups of proteins (COG), EuKaryotic Orthologous Groups(KOG) and Kyoto Encyclopedia of Genes and Genomes(KEGG). Using KOBAS 2.0 to achieve the results of Unigene in KEGG of KEGG Orthology. We use HMMER and PFAM to compare the sequence after achieving annotation information of Unigene.

Differentially expressed unigene (DEGs) analysis

Bowtie[14] compared the sequencing Reads of each sample with the Unigene library and estimated the expression level by RSEM[15] based on the comparison results. The abundance of corresponding Unigene is indicated by FPKM value. The DEGs were screened with a criterions: FDR 0.01 and FC (Fold Change) ≥ 2. The abundance values of transcripts were normalized. Using the MultiExperiment Viewer (version 4.9.0) to draw heat maps based on the transformation values. The figure showed that different columns represent different samples and different rows represent different genes .

Homology analysis and CDS prediction

TransDecoder software is used to compare the length of open reading frame, logarithmic likelihood function value and amino acid sequence with protein structure domain sequence in Pfam database. Predicted full-length sequences of the key genes, involved in the alliin synthetic pathway, were used for alignments.

Analysis of alliin contents

We prepare the materials of PW garlic roots, bulbs, inner bulb, garlic sprout, and wound leaves. Grinding the sample of liquid nitrogen, adding 4% of sulfosalicylic acid and ddH_2O, 25℃ 30 min. and using centrifugal (12000rap, 20 min) to achieve supernatant. Then detect alliin content. To ensure the accuracy of the data,we measured the content of alliin at least three replicates ± standard error.

Statistical analysis

We adopted this effective method of Benjamini-Hochberg to correct p-value that hypothesis tests getting in the process of differential expression analysis. Finally, we uesed the adjusted p-values that FDR (False Discovery Rate). FDR is a key indicator for screening differentially expressed genes. It is an important methods to reduce the false positives of a large number of genes expression. Statistical analysis software is SPSS version 22.0. The method of comparing the differences uses ONE WAY ANOVA analysis of variance.

Abbreviations

Sykam S-433d
Amino acid measuring instrument;MCSO:S-methyl-L-cysteine sulfoxide; PCSO S-propyl-L-cysteine sulfoxide; ACSO:S-allyl-L-cysteine sulfoxide; PECSO:S-trans-1-propenyl-L-cysteine sulfoxide; T1/3-T4/6:T01/02/03-T04/05/06; T1/3-T7/9:T01/02/03-T07/08/09; T1/3-T10/12:T01/02/03-T10/11/12; NR:Non-redundant protein; GO:Gene Ontology; COG:Clusters of Orthologous Groups of proteins; KOG:EuKaryotic Orthologous Groups; KEGG:Kyoto Encyclopedia of Genes and Genomes; CYS:Cysteine; SER:Serine; PW:Pizhou white garlic

Declarations

Ethics approval and consent to participate

Garlic were collected form pizhou, China(PW). PW garlic was cultivated in the test farm (site in sanbao town, xuzhou city, jiangsu province). This project uses plant materials and does not utilize transgenic technology.

Consent to publish

Not applicable.

Availability of data and materials

The datasets generated and analysed during the current study are available

Competing interests

The authors statement that they have no competing interests.

Funding

Our work was funded by the National Natural Science Foundation of China (31672148, 31770613), the reserve talents of preeminent youth fund of Jiangsu province by Jiangsu Normal University (HB2016016), the Student Innovation Training Program of Jiangsu (SJKY19-2046) and Science and technology special project of north jiangsu (XZ-SZ201839). The funding body played no role in our study.

Authors' Contributions

JJ and G. J. Loake made substantial contributions to design, was participated in drafting the manuscript, and given final approval of the version to be published. YQ, YS, JW, and WW made substantial contributions to acquisition and analysis of data, was participated in revising the manuscript, and given final approval of the version to be published. HC, XC, JW, ZZ, YW, and DM made substantial contributions to analysis and annotations of data, was participated in revising the manuscript, and given final approval of the version to be published. All authors have read and approved the manuscript, and ensure that this is the case.

Acknowledgements

We thank Cai Run (Shanghai Jiao Tong University) and Lihuang Zhu (Institute of Genetics and Developmental Biology, Chinese Academy of Sciences CAS) for their technical assistance, manuscript review and suggestions.

References

1. Satyal P, Craft JD, Dosoky NS, et al. The chemical compositions of the volatile oils of garlic (*Allium sativum L.*) and wild garlic (*Allium sativum L.*). Foods. 2017;6(8):63–1.
2. Xiaojun C, Xia L, Siyuan Z, et al. Transcriptome-referenced association study of clove shape traits in garlic. DNA Res. 2018;6(12):588.
3. Petropoulos S, Fernandes, Ângela, Barros L, et al. Antimicrobial and antioxidant properties of various Greek garlic genotypes. Food Chem. 2018;245:7.
4. Lu X, Li N, Qiao X, et al. Composition analysis and antioxidant properties of black garlic extract. Journal of Food Drug Analysis. 2016;25(2):340.
5. Mcmahon FG, Vargas R. Can garlic lower blood pressure? A pilot study. Pharmacotherapy the Journal of Human Pharmacology Drug Therapy. 2012;13(4):406.

6. Lv Y, So KF, Wong NK, et al. Anti-cancer activities of S-allylmercaptocysteine from aged garlic. Chinese journal of natural medicines. 2019;17(1):43.
7. Varshney R, Budoff MJ. Garlic and Heart Disease. J Nutr. 2016;13(1):417.
8. Gupta V. Hyolipidemic Effect Of Garlic: On Experimentally Induced Hyperlipidemia. Journal of Advance Researches in Biological Sciences. 2012.
9. Sun X, Zhou S, Meng F, et al. De novo assembly and characterization of the garlic (*Allium sativum L.*) bud transcriptome by Illumina sequencing. Plant Cell Rep. 2012;31(10):1824.
10. Marta Corzo-Martínez. Corzo N, Villamiel M. Biological properties of onions and garlic. Trends Food Sci Technol. 2007;18(12):0–625.
11. Kharde KR, Soujanya S. Effect of garlic and neem leaf powder supplementation on growth performance and carcass traits in broilers. Veterinary World. 2014;7(10):800.
12. Jones MG, Hughes J, Tregova A, et al. Biosynthesis of the flavour precursors of onion and garlic. J Exp Bot. 2004;55(404):1906.
13. Kubec R, VelíŠek J, Musah RA. The amino acid precursors and odor formation in society garlic (Tulbaghia violaceaHarv.). Phytochemistry. 2002;60(1):23.
14. Lou Z, Wei QQ, Wang DW, et al. Effect of allicin on proliferation and apoptosis of KG-1 cells and its molecular mechanism. China Journal of Chinese Materia Medica. 2018;43(12):2612.
15. Andualem B. Combined antibacterial activity of stingless bee (Apis mellipodae) honey and garlic (*Allium sativum L.*) extracts against standard and clinical pathogenic bacteria. Asian Pacific Journal of Tropical Biomedicine. 2013;3(9):726.
16. Wu X, Li X, Song Y, et al. Allicin protects auditory hair cells and spiral ganglion neurons from cisplatin - Induced apoptosis. Neuropharmacology. 2017;116:431.
17. Leontiev R, Hohaus N, Jacob C, et al. A Comparison of the Antibacterial and Antifungal Activities of Thiosulfinate Analogues of Allicin. Sci Rep. 2018;8(1):6763.
18. Lawson L, Hunsaker S. Allicin Bioavailability and Bioequivalence from Garlic Supplements and Garlic Foods. Nutrients. 2018;10(7):812.
19. Wang J, Cao Y, Sun B, et al. Effect of ultrasound on the activity of alliinase from fresh garlic. Ultrason Sonochem. 2011;18(2):536.
20. Borlinghaus J, Albrecht F, Gruhlke M, et al. Allicin: Chemistry Biological Properties Molecules. 2014;19(8):125999.
21. Ide N, Ryu K, Ogasawara K, et al. Antioxidants in processed garlic: I. Fructosyl arginine identified in aged garlic extract. International Congress. 2002;1245(02):447.
22. Zhai B, Zhang C, Sheng Y, et al. Hypoglycemic and hypolipidemic effect of S-allyl-cysteine sulfoxide (alliin) in DIO mice. Sci Rep. 2018;8(1):3527.
23. Helen A, Krishnakumar K, Vijayammal PL, et al. A comparative study of antioxidants S-Allyl cysteine sulfoxide and vitamin E on the damages induced by nicotine in rats. Pharmacology. 2003;67(3):115.

24. Saray Quintero-Fabián, Daniel Ortuño-Sahagún, Manuel Vázquez-Carrera, et al. Alliin, a garlic (*Allium sativum L.*) compound, prevents LPS-induced inflammation in 3T3-L1 adipocytes. Mediators of Inflammation. 2013;26(12):381815.
25. Gunther WHH. Garlic and other alliums-the lore and the science. Sulfur Reports. 2010;34(1–2):208.
26. Singh MR, Singh D, Srivastava S. Formulation and evaluation of NLCS encapsulated with S-Allyl-L-Cysteine sulfoxide for treatment of inflammatory bowel disease. Planta Medica. 2015;81(5).
27. Jones MG, Collin HA, Tregova A, et al. The biochemical and physiological genesis of alliin in garlic. Medicinal Aromatic Plant Science Biotechnology. 2007;1(1):22.
28. Nasim SA, Dhir B, Samar F, et al. Sulphur treatment alters the therapeutic potency of alliin obtained from garlic leaf extract. Food Chemical Toxicology. 2009;47(4):890.
29. Einat SM, Tomer BM, Neta R, et al. Garlic (*Allium sativum L.*) fertility: transcriptome and proteome analyses provide insight into flower and pollen development. Front Plant Sci. 2015;6(4):271.
30. Kamenetsky R, Faigenboim A, Shemesh ME, et al. Integrated transcriptome catalogue and organ-specific profiling of gene expression in fertile garlic (*Allium sativum L.*). Bmc Genomics. 2015;16(1):12.
31. Jayasena AS, Secco D, et al. Next generation sequencing andde novotranscriptomics to study gene evolution. Plant Methods. 2014;10(1):34.
32. Mutz KO, Heilkenbrinker A, Lönne M, et al. Transcriptome analysis using next-generation sequencing. Curr Opin Biotechnol. 2013;24(1):25.
33. Harper AL, Trick M, Higgins J, et al. Associative transcriptomics of traits in the polyploid crop species Brassica napus. Nat Biotechnol. 2012;30(8):79.
34. Lu G, Harper AL, Trick M, et al. Associative transcriptomics study dissects the genetic architecture of seed glucosinolate content in Brassica napus. DNA Res. 2014;21(6):620.
35. Zhou W, Huang Q, Wu X, et al. Comprehensive transcriptome profiling of Salvia miltiorrhiza for discovery of genes associated with the biosynthesis of tanshinones and phenolic acids. Sci Rep. 2017;7(1):10554.
36. Guang-Hui Z, Ni-Hao J, Wan-Ling S, et al. De novo sequencing and transcriptome analysis of pinellia ternata identify the candidate genes involved in the biosynthesis of benzoic acid and ephedrine. Front Plant Sci. 2016;7:1209.
37. Kumar M, Chauhan AS, Kumar M, et al. Transcriptome sequencing of Chickpea (Cicer arietinum L.) genotypes for identification of Drought-Responsive genes under drought stress condition. Plant Molecular Biology Reporter. 2019;37(3):187.
38. Mitrová K, Svoboda P, Milella L, et al. Alliinase and cysteine synthase transcription in developing garlic (*Allium sativum L.*) over time. Food Chem. 2018;251:108.
39. Grabherr MG, Haas BJ, Yassour M, et al. Full length transcriptome assembly from RNA Seq data without a reference genome. Nature Biotechnology Italic. 2011;(29):648.

40. Hughes J, Tregova A, Tomsett AB, et al. Synthesis of the flavour precursor, alliin, in garlic tissue cultures. Phytochemistry. 2005;66(2):190.
41. Langmead B, Trapnell C, Pop M, et al. Ultrafast and memory-efficient alignment of short DNA sequences to the human genome. Genome Biology. 2009;10(3).
42. Dewey CN, Bo L. RSEM: accurate transcript quantification from RNA-Seq data with or without a reference genome. BMC Bioinformatics. 2011;12(1):323.
43. Valencia A. STAR: ultrafast universal RNA-seq aligner. Bioinformatics. 2014;22(1):1.
44. Mckenna A, Hanna M, Banks E, et al. The Genome Analysis Toolkit: A MapReduce framework for analyzing next-generation DNA sequencing data. Genome Res. 2010;20(9):1300.
45. Lawson L. Garlic. The Science and Therapeutic Application of (*Allium sativum L.*) and Related Species. HerbalGram. 1996;38:65.
46. Warrilow AGS. Modulation of cyanoalanine synthase and O-acetylserine (thiol) lyases A and B activity by beta-substituted alanyl and anion inhibitors. J Exp Bot. 2002;53(368):441.
47. Leustek T, Martin MN, Bick JA, et al. Pathways and regulation of sulfur metabolism revealed through molecular and genetic studies. Annual Review of Plant Physiology Plant Molecular Biology. 2000;51(4):141.
48. Sun XQ, Huan LI, Li LI, et al. Component concentration of organic sulfides in garlic leaf extracts at different developmental stages. Acta Agriculturae Zhejiangensis. 2014;26(4):1065.
49. Abbey L, Aked J, Joyce DC. Discrimination amongst Alliums using an electronic nose. Ann Appl Biol. 2015;139(3):340.

Figures

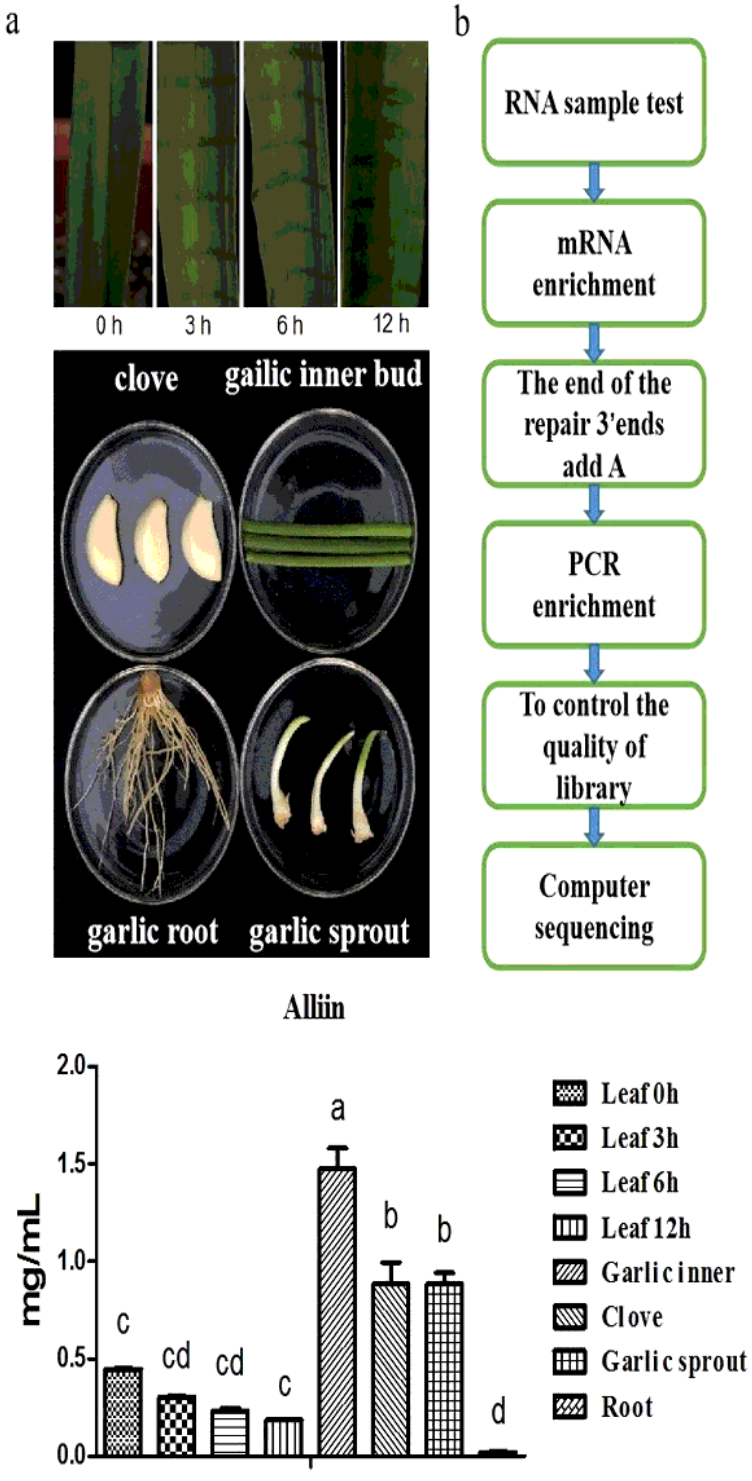

Figure 1

Variation of the contents of alliin among sample. a Mature garlic tissue: root, bulb, garlic leaf, garlic inner bud garlic sprout. The fresh tissue was collected. b The transcriptome sequencing basic process. c By S-4330D measure the contents of alliin in garlic. Different alphabet indicate significant differences in the drawing($P < 0.05$).

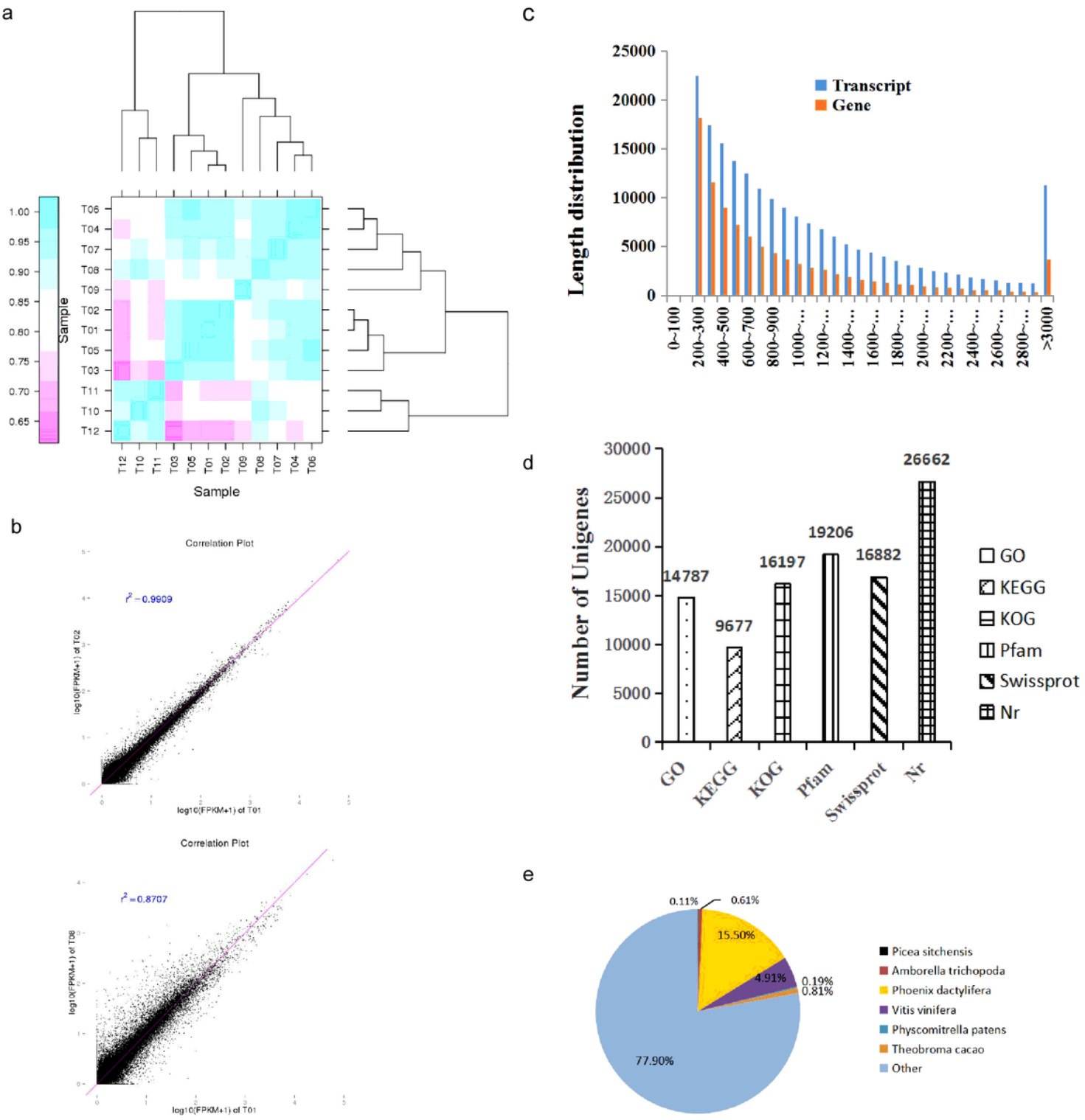

Figure 2

Transcriptome data analysis of garlic. a The biological repeat correlation of garlic sequencing data was determined by Pearson correlation coefficient. b The scatter diagram show that the degree of samples correlation gene deviates from the diagonal. c The size distributions of transcripts and unigenes of gailic. d The annotation of unigenes base on various databases. e The notes of unigenes species distribution .

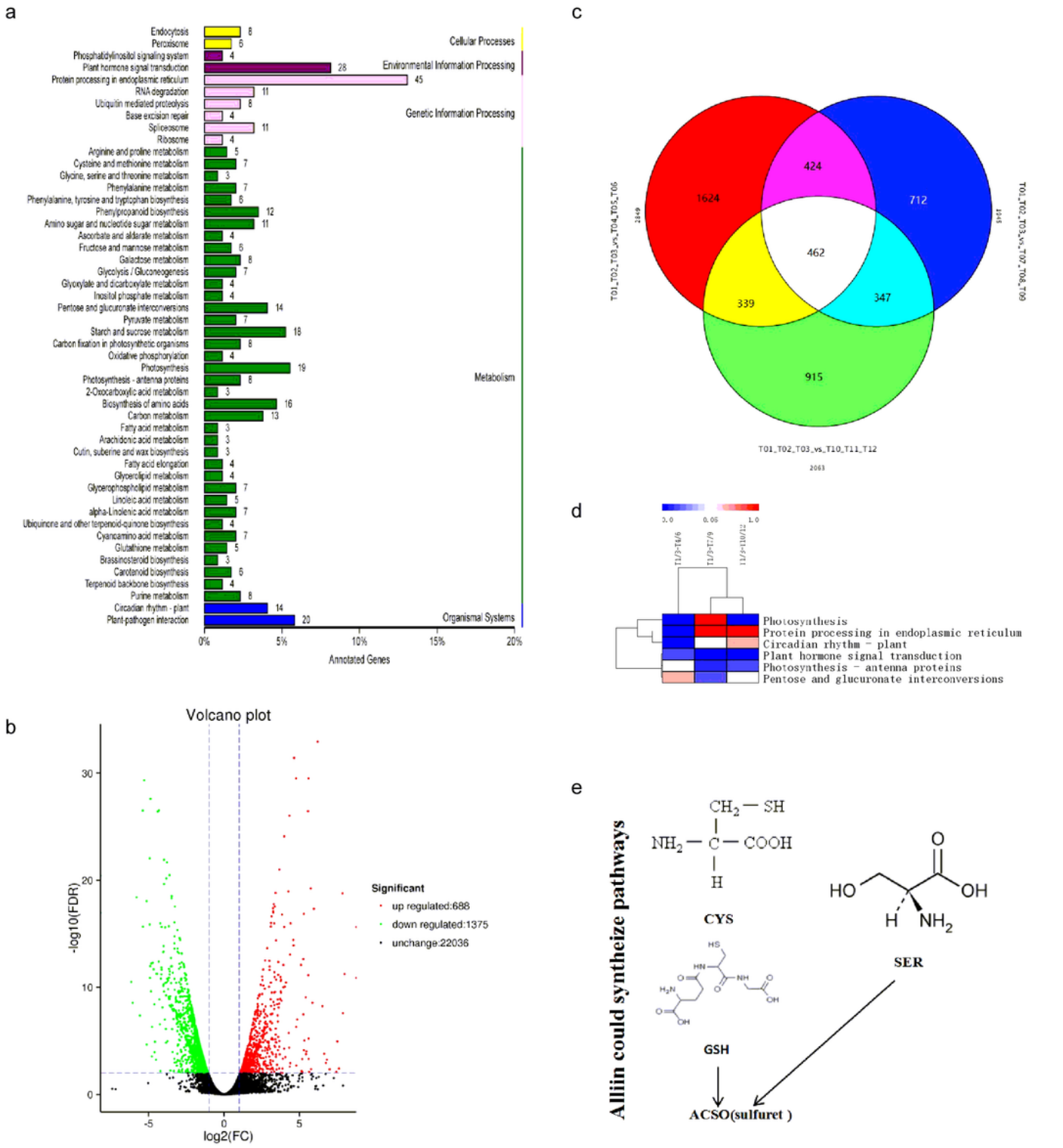

Figure 3

Recognition of the DEGs among all sample. a The annotation results of differentially expressed KEGG genes were classified, according to the KEGG pathway types for organismal systems, metabolism, cellular process, environmental information processing, andgenetic information processing. b The Volcano Plot showed the differences expressed level and number of single genes in each of sample comparisons. c The Venn diagram shows that the number of DEG in the T1/3-T4/6 and T1/3-T7/9 and

T1/3-T10/12 three comparisons. d KEGG enrichment analysis of three samples DEGs comparison. By heat maps show that all the significant P-values of KEGG in the three groups of comparisons were represented. e Alliin could synthesize pathways.

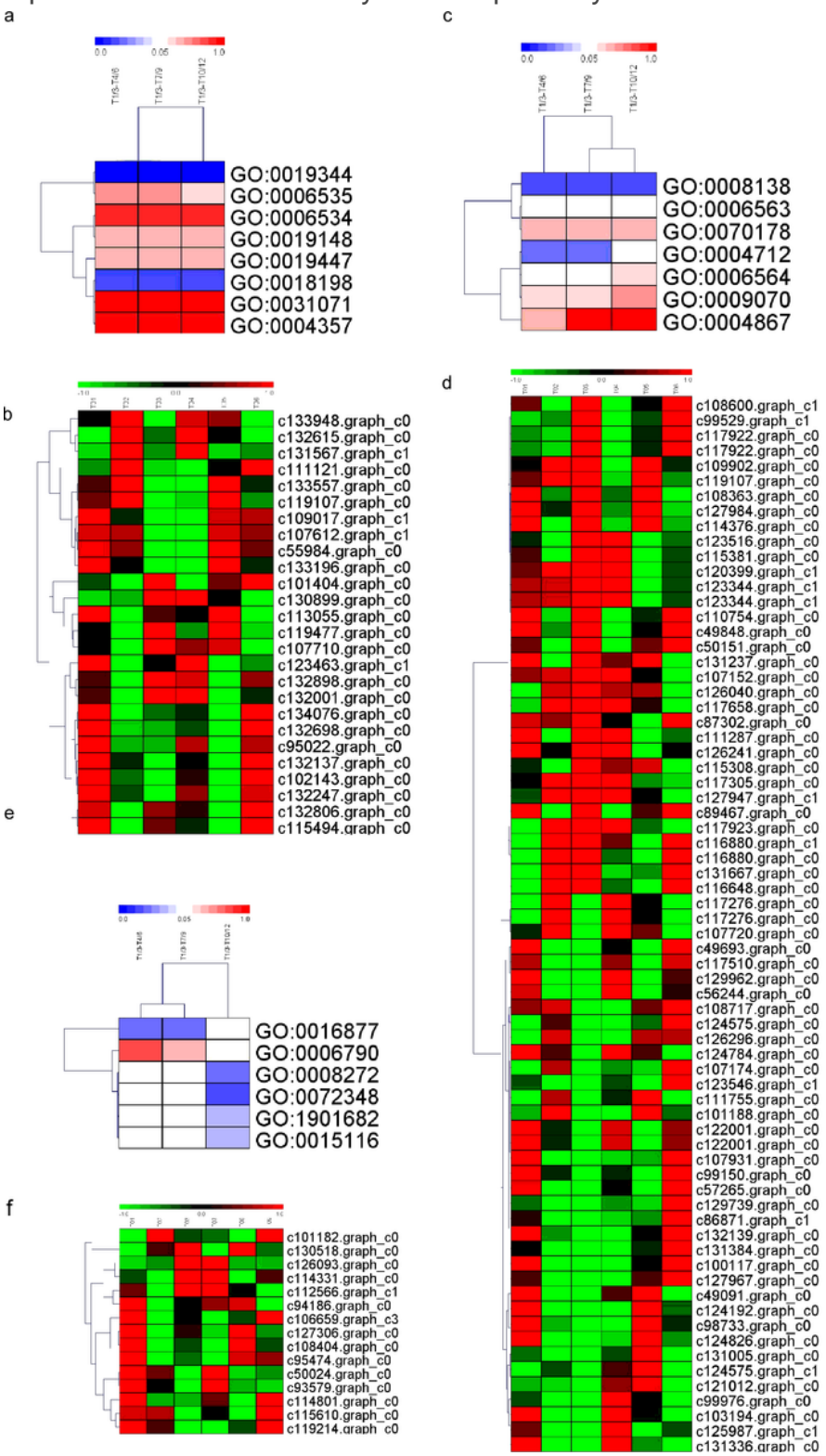

Figure 4

Transcriptomic analysis reveals the differences CYS/SER/Sulfuret pathway-related genes. a Transcriptomic analysis reveals the CYS pathway-related differences genes. b By a heatmap shown that

the analysis of eight cysteine-related GO terms in T1/3-T4/6 and T1/3-T7/9 and T1/3-T10/12 comparisons. c Significance analysis of seven serine-related GO terms in T1/3-T4/6 and T1/3-T7/9 and T1/3-T10/12 comparisons were shown by a heat map. d The heat map showed the expression of different genes related to SER signaling pathway. e The GO terms of sulfur-related found six in T1/3-T4/6 and T1/3-T7/9 and T1/3-T10/12 comparisons by a heat map. f The heat map shown differentially expressed genes related of the sulfur pathway.

Supplementary Files

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

- Additionalfile9.doc
- Additionalfile8.xlsx
- Additionalfile7.doc
- Additionalfile6.xlsx
- Additionalfile5.xlsx
- Additionalfile4.doc
- Additionalfile3.doc
- Additionalfile2.doc
- Additionalfile1.xlsx
- Additionalfile12.xlsx
- Additionalfile11.doc
- Additionalfile10.xlsx