

Transcriptome Analysis of Black and White Sesame Seed Reveals Candidate Genes Associated with Black Seed Development in Sesame (*Sesamum Indicum*)

Senouwa Segla Koffi Dossou

Oil Crops Research Institute Chinese Academy of Agricultural Sciences <https://orcid.org/0000-0002-7556-9555>

Linhai Wang

Oil Crops Research Institute Chinese Academy of Agricultural Sciences

Xin Wei

Shanghai Normal University

Yanxin Zhang

Oil Crops Research Institute Chinese Academy of Agricultural Sciences

Donghua Li

Oil Crops Research Institute of the Chinese Academy of Agricultural Sciences

Jingyin Yu

Oil Crops Research Institute Chinese Academy of Agricultural Sciences

Xiurong Zhang (✉ zhangxr@oilcrops.cn)

Oil Crops Research Institute Chinese Academy of Agricultural Sciences

Research article

Keywords: *Sesamum indicum*, Seed coat color, Transcriptome, Candidate genes, Black pigment

Posted Date: July 29th, 2020

DOI: <https://doi.org/10.21203/rs.3.rs-46165/v1>

License: © ⓘ This work is licensed under a Creative Commons Attribution 4.0 International License. [Read Full License](#)

Abstract

Background: Seed coat color is a key agronomic characteristic in sesame (*Sesamum indicum*) since it is strongly linked to seed oil, proteins, and lignans content and it influences consumer preferences. Even though some QTL and candidate genes have been detected for sesame seed coat color, the mechanism and regulation of black pigmentation are not entirely understood. This study provides an overview of developing seeds transcriptome of two varieties of sesame “Zhongfengzhi No.1” (white seed) and “Zhongzhi No.33” (black seed) and shed light on genes involving in black seed formation.

Results: Both black and white sesame showed similar trend expressed genes with the numbers increased at the early stages of seed development. The differentially expressed genes (DEGs) number increased with seed development in the two sesame varieties. We examined the DEGs and uncovered that the early stage, which is from 8 to 17 days post-anthesis (DPA) plays an important role in black pigment biosynthesis and accumulation. The gene expression patterns were consistent with the seed color change. Besides, we studied the shared DEGs between the black and white sesame. We figured out 17 candidate genes associated with pigments biosynthesis in black sesame seed including 2 chalcone synthase genes SIN_1018961 and SIN_1018959 which may function in the phenylpropanoid pathway. 5 of these candidate genes, SIN_1006242 and SIN_1016759/PPO, SIN_1026689 and SIN_1006025, SIN_1025056 are located on chromosomes 4, 8 and 11 respectively, in conformity with previous QTL mapping. These genes were believed to play a major role in black seed development in sesame.

Conclusion: This work illuminated the different expression profiles in black and white sesames and unfolded pivotal stages and a catalog of candidate genes associated with black seed formation in sesame. These findings provide a vast transcriptome dataset and list of genes that will be targeted for functional studies related to the molecular mechanism involved in biosynthesis and regulation of seed coat color in sesame and for molecular breeding of high-quality sesame varieties.

Background

Sesame (*Sesamum indicum*) owned to the Pedaliaceae family, is a vital oil crop widely grown in tropical, subtropical, and southern temperate regions due to its high-quality nutritional and therapeutic seeds. Sesame is an ideal plant model for studying many traits related to the oil content in crops because of its diploid nature and tiny genome size of 357 Mb [1]. Moreover, it exhibits contrasting genetic dissimilarity in morpho-economic traits and quality attributes (seed shape, color, length) [2]. Maximum variation was revealed in seed coat color and many other traits [3]. Sesame seed color varies from black, brown, yellow to white. Black sesame seeds are used as food and medicine in Asia, whereas white sesame seeds are primarily used to produce oil [4]. Previous studies revealed that white sesame seeds contained higher sesamin and sesamolin [5][6], oil, protein, and linoleic acid [7], whereas black sesame seeds generally possess high level of ash and carbohydrate [8], total phenolic and total antioxidant [9]. Due to the various health-promoting abilities of natural antioxidants, medical and pharmaceutical industries researchers have been engaged to develop food products with high amount of antioxidants principally for nutritional goals [10][11][12]. Black sesame products attract good acceptance in East Asia and is used in the treatment of diverse diseases [13] [14]. Recently, Wang et al. [13] established a link between black sesame nutritional components and its proprieties recorded in traditional Chinese medicine. Indeed, they detected that black sesame seeds contain a higher level of indole-3-carboxylic acid, hesperidin, 2-methoxycinnamic acid, vitamin B2, coniferyl aldehyde, phloretin and hyoscyamine than white sesame seeds. These compounds especially hesperidin, 2-methoxycinnamic acid, and vitamin B2 were reported to exhibit various biological activities including anti-cancer, anti-oxidative, anti-bacterial, anti-inflammatory, and anti-infective proprieties [15][16][17].

Several years ago, sesame research entered the “Omics” era [18]. Since then, a list of QTL and candidate genes have been identified related to oil, sesamin and sesamolin content, fatty acid biosynthesis and yield, plant height and seed coat color [1][19][4][20]. Seed coat color is a complex polygenetic trait that has been reported to be the result of several pigments

including flavonoids (flavonols and proanthocyanidin), lignin and melanin [21]. Zhang et al.(2013) have detected 4 QTL (QTL1-1, QTL11-1, QTL11-2 and QTL13-1) for seed coat color [14], whereas Wang et al. [2016] have detected 9 QTL (qSCa-8.2, qSCb-4.1, qSCb-8.1, qSCb-11.1, qSCI-4.1, qSCI-8.1, qSCI-11.1, qSCa-4.1 and qSCa-8.1) located on chr4, chr8 and chr11 [4]. However, no relationship has been established between the two groups of QTLs due to the differences in markers and methods used. Hence, genes involved in biosynthesis and regulation of the sesame seed coat color and how they operate during seed development still needs to be clarified, as only SiPPO (SIN_1016759), a gene encoding polyphenol oxidase, had been identified as the regulatory gene that controls seed coat color in sesame [20].

RNA-Seq is an advanced approach in transcriptome profiling that uses deep-sequencing technologies. It has become an essential tool for transcriptome-wide analysis of differential gene expression and differential splicing of mRNAs [22]. RNA-seq was successfully used to identify candidate genes for seed coat color in many plants, including *Arabidopsis thaliana* [23], *Brassica rapa* [24], *Brassica napus* [25] and soybean [26]. In addition, it has been used to detect candidate genes that shaped oil content variation in sesame [27]. In this study, we investigated the transcriptome during the seed development stages at the 5, 8, 11, 14, 17, 20, 23, 26 and 30 days post-anthesis (DPA) of two sesame varieties different in seed coat color “Zhongfengzhi No.1” (white seed) and “Zhongzhi No.33” (black seed). Differentially expressed genes (DEGs) and some candidate genes for sesame seed coat color were identified. This study will facilitate further investigations on seed coat color modulation in sesame and will help to improve sesame varieties quality in the future.

Results

Overview the changes of seed and expressed genes in black and white sesame

The seed samples showed the black sesame start to accumulate black pigment or related compositions from 11 DPA (Days Post Anthesis) (Fig. 1a). The Typical change was observed at 11 DPA as the seeds became brown and then black at 14 DPA. For the expressed genes, the black and white showed similar trends with the numbers increased at early stages then decreased. The black sesame reached the maximum expressed gene number with 20,253 genes over 0.1 in FPKM (Fragments Per Kilobase of transcripts per Million fragments mapped) at 11 DPA. The inflection point in white sesame appeared later than the black. The number of expressed genes increased again from 23 DPA to 30 DPA in black sesame, which may indicate some biosynthesis pathways were activated newly (Fig. 1b).

Stage differences of the expressed genes along with seed development

As the black sesame starts to accumulate pigment or related compounds at 11 DPA, the 5 and 8 DPA samples were used as controls to study how the expressed genes change. At the early stage from 5 to 8 DPA, the seeds were similar in color in the black sesame (Fig. 1a). We compared the other samples to the two points respectively and found some differences in the black sesame with 8 DPA as control. When the 5 DPA was used as the control, a sharp increase was observed from 8 DPA to 11 DPA, but it becomes slightly when 8 DPA was used as control (Figure S1 and S2 in Addition file 1). It may suggest some genes involved in the pigment biosynthesis had been initiated at 8 DPA and 5 DPA can be used as a control to compare the differentially expressed genes for black sesame. When we checked the white sesame seed, though the numbers of differentially expressed genes tended to increase with seed development, it showed similar scenes with 5 DPA and 8 DPA as controls. We also found the down-regulated genes increased more than the up-regulated in both black and white sesame seeds.

We then investigated how the expressed genes change by comparing the seeds to the former ones. It showed a great difference between the black and white sesame. In the black sesame, there were more genes up or down-regulated at the

stages of 8, 11 and 30 DPA, and it was increased from 8 to 11 DPA (Fig. 2a). In the white sesame, both the up and down-regulated genes decreased in number from 8 to 23 DPA and then increased slightly (Fig. 2b). The result also suggested the essential roles of 8 and 11 DPA in the black pigment biosynthesis and accumulation, which is consistent with the seed color change.

The differentially expressed genes between black and white sesame

We compared the differentially expressed genes between black and white sesames at different stages. The DEG numbers changed with seed development. From 5 to 8 DPA, more up-regulated genes highlighted the early stages, then it decreased to a low level with only 225 up-regulated DEG at 20 DPA. Interestingly many DEGs were observed at the later stage of 30 DPA when the seed reached the maturity stage (Fig. 3a). It may suggest more biosynthesis pathways were strengthened to accumulate different storage energy or components in black and white sesame. We also studied the shared DEGs by comparing the adjacent two points to reduce the effect of factitious sampling stages. In total, it also showed the black sesame seed had more genes up-regulated at early stages before 17 DPA, and more genes down-regulated at the later stages from 20 to 30 DPA (Fig. 3b).

The candidate gene responsible for the coat color of black sesame

The above analysis indicated the seed coat color difference not only exist between black and white sesame but also during seed development in the black sesame. We took the 5 DPA as the initiated control to identify the candidate genes and group 11 to 23 DPA for black pigment biosynthesis and accumulation. We firstly figured out the shared up and down-regulated DEGs between 11, 14, 17, 20 and 23 DPA against 5 DPA in the black sesame, and there were 1,254 up and 1,617 down-regulated DEGs respectively (Fig. 4 and Figure S3 in Additional file 1). Then the shared DEGs between black and white sesames at 11, 14, 17, 20 and 23 DPA were also figured out, with 53 up-regulated and 57 down-regulated, respectively (Fig. 4 and Additional file 1: Figure S3). Finally, the two sets up and down-regulated DEGs were checked and the shared genes were found out, with only 31 up and 11 down-regulated left (Fig. 4).

We looked into the 11 common down-regulated DEGs in the black and white sesame seed. It showed all the genes tended to decrease during seed development, and all the genes in the black sesame expressed less than in the white (Figure S4 in Additional file 1). In consideration of the characteristics of black seed development, these genes were unlikely to be involved in the biosynthesis of black pigment or related compounds, as all the genes expressed in sesames and none was out of the ordinary in white or black sesame.

For the 31 up-regulated genes in the black sesame, we firstly filtered out the ones (8 genes) with high FPKM over 100 in white sesame and those less than 5 in both black and white sesame, because these genes were unlikely to associate with the biosynthesis of black pigment. The 23 left genes were then grouped into 4 subgroups using the hierarchical clustering method (Fig. 5). Subgroup 1 consisted of only 2 genes SIN_1003674 and SIN_1009127, but the expression level of the later one increased in the white sesame and should be filtered out. Subgroup 2 consisted of 3 genes which all expressed in the two sesames with different FPKM and should not include in the candidate genes for black sesame. Subgroups 3 and 4 consisted of 7 and 9 genes, respectively. The 7 genes SIN_1024143, SIN_1016759, SIN_1018917, SIN_1018959, SIN_1001138, SIN_1026689, SIN_1006892 in Subgroup 3 and the 9 genes SIN_1002392, SIN_1006025, SIN_1025056, SIN_1013986, SIN_1006242, SIN_1018543, SIN_1020696, SIN_1018961, SIN_1022200 in subgroup 4 were selected as the candidate genes for black sesame according to their expression patterns and level (Figures S5 and S6 in Additional file 1).

Wang et al. [4] mapped 4 loci on linkage groups 4, 8 and 11, which were predicted to associate with the a*, b* and l* value of seed coat color. Here we figured out 17 candidate genes (Table 1), and their expression differences were validated with qRT-PCR (Figures S7 in Additional file 1). Among them, 5 SIN_1006242, SIN_1016759/PPO, SIN_1026689, SIN_1006025

and SIN_1025056 are located on chromosomes 4, 8 and 11. Especially, SIN_1016759/PPO had been reported to be responsible for black sesame based on GWAS [20]. Also, the 17 genes also included 2 chalcone synthase genes SIN_1018961 and SIN_1018959 that may function in the phenylpropanoid pathway. Thus, some of these genes were believed to involve in black pigment or related compounds biosynthesis in black sesame.

Table 1
Information of the 17 candidate genes

Chromosome	Start	End	Regulated	Gene_ID	NCBI annotation	NCBI annotation-blast
Chr1	18909556	18911695	Up	SIN_1013986	XP_011091088.1	Ferruginol synthase-like [Sesamum indicum]
Chr2	6742573	6744985	Up	SIN_1002392	XP_011069671.2	Probable carotenoid cleavage dioxygenase 4, chloroplastic[Sesamum indicum]
Chr4	625034	626985	Up	SIN_1006242	XP_011075178.1	Cytochrome P450 93A3-like [Sesamum indicum]
Chr4	12428252	12430048	Up	SIN_1016759	XP_011076536.1	Polyphenol oxidase I, chloroplastic-like [Sesamum indicum]
Chr6	5785721	5787326	Up	SIN_1020696	XP_011081935.1	Cytochrome P450 71D95-like [Sesamum indicum]
Chr6	14054614	14056241	Up	SIN_1018543	XP_011081404.2	Transcription factor WER-like [Sesamum indicum]
Chr7	11180578	11181996	Up	SIN_1001138	XP_011101592.1	Croctin glucosyltransferase, chloroplastic-like [Sesamum indicum]
Chr8	15145987	15147743	Up	SIN_1026689	XP_011086525.1	WAT1-related protein At3g28050-like [Sesamum indicum]
Chr8	19512267	19517197	Up	SIN_1006025	XP_011087218.1	Isochorismate synthase, chloroplastic [Sesamum indicum]
Chr9	5392467	5398195	Up	SIN_1024143	XP_011089887.1	MLO-like protein 6 [Sesamum indicum]
Chr9	19701218	19702926	Up	SIN_1003674	XP_011099188.1	3-ketodihydrosphingosine reductase [Sesamum indicum]
Chr10	14971538	14973984	Up	SIN_1018961	XP_011091402.1	Chalcone synthase [Sesamum indicum]
Chr10	15004626	15007378	Up	SIN_1018959	XP_011091399.1	Chalcone synthase-like [Sesamum indicum]
Chr10	15394225	15396302	Up	SIN_1018917	XP_011091358.1	Cytochrome P450 71A1-like [Sesamum indicum]
Chr11	6138991	6148011	Up	SIN_1025056	XP_011093954.1	Ceta-glucosidase 12-like isoform X3 [Sesamum indicum]

Chromosome	Start	End	Regulated	Gene_ID	NCBI annotation	NCBI annotation-blast
Chr12	7689485	7691828	Up	SIN_1022200	XP_011083203.2	LOW QUALITY PROTEIN: dihydroflavonol 4-reductase [Sesamum indicum]
Chr12	12203164	12204704	Up	SIN_1006892	XP_011095911.1	Uncharacterized protein LOC105175232 [Sesamum indicum]

Discussion

The seed coat is the external protective layer of seed and develops from the integument initially surrounding the ovule and is maternal in origin [28]. It protects the embryo and endosperm from external factors such as mechanical injuries, desiccation and infections [29]. Moreover, it helps developing seed to regulate its metabolism in response to changes in its external environment by transmitting environmental signals to the interior of the seed [30]. In sesame, seed coat color is strongly associated with seed quality [5][7][31]. Therefore, genetic resources on pigmentation mainly black seed coats will help to improve the sesame seed quality. In this study, RNA-seq was used to scrutinize transcriptome differences between “Zhongfengzhi No.1” (white seed) and “Zhongzhi No.33” (black seed) at different stages of seed development. DEGs differently regulated during seed coat development were screened, and candidate genes associate with black pigmentation were detected.

The key stage for black pigment biosynthesis and accumulation in black sesame

In this study, we analyzed seeds samples and observed that seeds started browning and changed to black from 11 DPA. A great difference was observed between the expression profiles of black and white sesame. The black sesame reached the maximum expressed gene number with 20,253 genes over 0.1 in FPKM at 11 DPA and more genes were up or down-regulated at the early stages before 17 DPA. These results are consistent with the findings of Wei et al. [19][20], who reported that the gene SIN_1016759/SiPPO was highly expressed in the seeds from 11 to 20 DPA. PPO encodes polyphenol oxidase and generates black pigments via the browning reaction in plant [32].

As in previous studies, our results confirmed the pivotal role of later stages in biosynthesis of nutrients (oil, protein, and lignans) in sesame [33][34][27]. Indeed, we observed that more genes were active and up or down-regulated at the later stages from 23 to 30 DPA in the two varieties especially in black sesame. Our findings provide the support that in sesame developing seed, early stages (before 23 DPA) play an important role in pigments biosynthesis and substrates preparation for nutrients biosynthesis in the later stages. Taken altogether, we thus suggested that black pigment is biosynthesized and accumulates in seed at early stages of seed development that is from 8 to 17 DPA mainly.

Candidate genes controlling black pigment biosynthesis in sesame

Flavonoids, including anthocyanins and proanthocyanidins, lignin and melanin are secondary metabolites that influence seed color in plants [21]. They are derived from the phenylpropanoid pathway and are controlled by a complex regulatory network with multiple transcription factors [35]. This pathway involves many genes such as chalcone synthase (CHS), chalcone isomerase (CHI), flavonol 3-hydroxylase (F3H), flavonol 3'-hydroxylase (F3'H), dihydroflavonol-4-reductase (DFR), anthocyanidin synthase (ANS) anthocyanidin reductase (ANR) and laccase. Some of these genes have been cloned from Arabidopsis and many plants [36]. Previous studies in sesame detected that two major genes with additive-dominant-epistatic effects plus polygenes with additive-dominant-epistatic effects control the seed coat color, and several major have been identified QTL [14][4]. Besides, the gene SIN_1016759/SiPPO had been reported as the candidate gene for black sesame [19][20]. Here, we studied the shared DEGs between black and white sesame and identified 17 candidate genes

associated with black pigmentation in sesame. Among the 17 candidate genes, 5 SIN_1016759/PPO, SIN_1006242, SIN_1026689, SIN_1006025 and SIN_1025056 are located on chromosomes 4, 8 and 11. This finding is congruent with the work of Wang et al. [4]. SIN_1016759/PPO encodes polyphenol oxidase; SIN_1006242 is a cytochrome P450 gene; SIN_1026689 is a WAT1-related protein; SIN_1006025 encodes isochorismate synthase, and SIN_1025056 encodes a beta-glucosidase isoform. In plants, isochorismate synthase converts chorismate the last product in the shikimate pathway into isochorismate, a precursor of phyloquinone (vitamin K1) and salicylic acid [37]. In addition, the 17 candidate genes also included 2 chalcone synthase genes SIN_1018961 and SIN_1018959. Three chalcone synthase-like genes have been identified in *A. thaliana* [23]. Chalcone synthase is the first committed enzyme in the biosynthesis of all flavonoids which function in the phenylpropanoid pathway [23]. It catalyzes the reaction leading to naringenin chalcone formation from p-coumaroyl-CoA and three molecules of malonyl-CoA [38]. All the 17 candidate genes will be targeted in the future for functional genomic study. We preferentially suggested that the genes SIN_1016759/SiPPO, SIN_1018961, and SIN_1018959 may play a major role in pigments biosynthesis especially black pigment in sesame.

The compound responsible for black seed color in sesame

In plants, browning reactions on seed coat pigments are often induced by the oxidation of phenolic compounds by polyphenol oxidases (PPO) such as laccases and tyrosinases and result in melanin formation mostly [39][32]. Tyrosinase is the enzyme that catalyzes the first two steps in the melanin synthesis pathway and controls the rate and yield of melanin production [40]. Xiaoli et al. [41] had reported that free tyrosine and polyphenols were needed for melanin biosynthesis and compared to anthocyanin, the melanin content in brown seed was higher than in yellow seed rapes. The metabolomic analysis demonstrated that phenylpropanoid biosynthesis, tyrosine metabolism, and riboflavin metabolism were the main pathways differentially activated between black and white sesame and were responsible for the color difference [13]. Furthermore, the black pigments present in the skin of banana and sunflower have been assumed to be melanin [42][43]. Wan et al. [29] detected that brown and dark peanut seed (mutants) had lower level of lignin, anthocyanin, proanthocyanidin content, and a higher level of melanin content compared to wild type with light color. Also, they reported that the expression of polyphenol oxidases was significantly activated in mutants developing seed. The above results indicated that PPO might be one of the dominant genes responsible for black pigment biosynthesis in sesame. Therefore, we suggested that the black pigment in sesame seed would be melanin. Future studies will help to confirm whether melanin is the black pigment or not.

Strategies to improve black sesame quality

Sesame is especially widely grown for its high-quality nutritional seeds [18]. However, compared with white sesame, black sesame content less oil, protein, linoleic acid, sesamin, and sesamol [5][6][7][31]. Hence the necessity to improve black sesame quality. The study carried out by Wei et al. [19] revealed that in sesame, the genes SiPPO and SiNST1 (SIN_1005755) associated respectively with black pigmentation and lignification in the seed coat are strongly associated oil, protein, sesamin, and sesamol content variation in seeds. Furthermore, in sesame aromatic amino acids L-phenylalanine (Phe) and L-tyrosine (Tyr) are needed for protein biosynthesis and served as precursors for numerous compounds including flavonoids, melanin, lignin, lignans, quinones, and condensed tannins [44]. These amino acids are produced from chorismate, the final product of the shikimate pathway, which involves many genes [44]. Here, we figured out 17 candidate genes associated with black pigment synthesis in sesame including the genes function for PPO, chalcone, isochorismate, and so on. Thus, functional analysis coupling with the genetic transformation of these genes simultaneously with other pivotal genes, may be sufficient to improve black sesame quality.

Conclusion

Overall, our study revealed the transcriptome difference between black and white seed coats during seed development in sesame using RNA-seq analysis. Our results provide valuable information on the complex transcriptome dynamics

involved in the control of seed coat color in sesame. The early stages play a crucial role in the biosynthesis and accumulation of the black pigment in *S. indicum* black seed. The phenylpropanoid and flavonoid biosynthetic pathways genes previously identified in other plants are also involved in the formation of seed coat color in sesame. Notably, 17 candidate genes controlling black pigmentation in sesame were identified and will be targeted in future studies for validation. As seed coat color in sesame is strongly associated with seed biochemistry and disease resistance, functional studies (cloning, genome editing, and transformation in sesame or Arabidopsis) of these candidate genes will help to understand molecular mechanisms involved in these correlations and for breeding high-quality sesame varieties.

Materials And Methods

Sesame varieties

Two sesame varieties “Zhongfengzhi No.1” (white seed) and “Zhongzhi No.33” (black seed) were used in this study. The seeds were given by the national sesame medium-term genebank (Wuhan, China).

Planting and sampling

The two varieties were shown under identical growth and experimental conditions in Wuhan in 2019. Flowers were labeled every three days post-anthesis (DPA). At 5, 8, 11, 14, 17, 20, 23, 26, and 30 DPA, capsules for each variety were sampled from 10 plants (Figure 1), and seeds were separated from the capsules on ice. Different plant seeds were therefore mixed equally and represented samples at 5, 8, 11, 14, 17, 20, 23, 26, and 30 DPA. All samples were prepared for two repeats and were subjected to RNA-seq analysis.

RNA extraction and library preparation

The 18 seed samples RNA were extracted and sequenced as per Wang et al. [27]. Briefly, each sample total RNA was extracted with the TRIzol reagent (Invitrogen Corp.). After that, using the Oligotex mRNA Midi Kit (Qiagen, Germany) we purified the mRNA from the total RNA. The quantity and quality of the mRNA were investigated with Invitrogen Qubit2.0 and Agilent 2100. All the mRNAs were then transcribed into double-stranded cDNAs with the SMART cDNA Library Construction kit (Clontech, USA) following the user guide. Finally, appropriate fragments (200 ± 25 bp) were chosen for PCR amplification, and adapters were ligated to the targeted fragments.

Data generation and quality assessment

The libraries of 18 cDNA generated from the sesame seeds were sequenced for paired-end reads using the Illumina HiSeq 2000 platform. The FastQC software (<http://www.bioinformatics.babraham.ac.uk/projects/fastqc/>) was used to check the reads base qualities. We then removed all paired-end reads with more than 5% ambiguous residues (Ns) and those which contained more than 10% bases with a Phred quality score of less than 20. The remaining reads were considered as “clean reads” [45]. Finally, 16.3–32.7 million clean reads with a 90 bp length were acquired for each sample. 77.2–85.2% of the clean reads were mapped uniquely to the sesame reference genome after verification of the parameter that no more than one mismatch was accepted in the alignment using TopHat2 (Kim D, Pertea G, Trapnell C, Pimentel H, Kelley R, Salzberg SL. TopHat2: accurate alignment of transcriptomes in the presence of insertions, deletions and gene fusions. *Genome Biology*. 2013, 14: R36). Only 70.4–79.2% of the reads were uniquely mapped to the predicted gene model regions.

Statistical analysis of gene expression

Samples gene expression levels were evaluated based on the read numbers which were uniquely mapped to the sesame genome sequence [1]. Each gene expression level were then normalized to the number of FPKM as per Trapnell et al. [46] using the Cufflinks 2.0 software. The differentially expressed genes (DEGs) were found out for sesame seed samples following the method described by Chen et al. [47] and Wang et al. [48]. The threshold P-value in multiple tests was

determined using the Poisson distribution [49] and false discovery rate (FDR). The significance of the DEGs was determined using an $FDR \leq 0.01$ and absolute value of \log_2 Ratio ≥ 1 [50].

Real-time quantitative PCR (qRT-PCR)

The expression profiles of the 17 candidate genes were validated with qRT-PCR referring to Wang et al. [48] using LightCycler® 480II Real-Time PCR Detection System (Roche Diagnostics, Rotkreuz, Switzerland). Five stages (5, 11, 17, 23, and 30 DPA) seed samples were run in triplicate on the same plate with a negative control that lacked cDNA. The gene actin7 of sesame was used as the positive control. The relative expression levels of the target genes were calculated using the $2^{-\Delta\Delta CT}$ method [51].

List Of Abbreviations Used

DPA: Days Post Anthesis, FPKM: Fragments Per Kilobase of transcript per Million fragments mapped, PPO: polyphenol oxidase

Declarations

Competing interests

The authors declare that they have no competing interests.

Author contributions

X.R.Z. contributed to the design of the research. L.H.W and D.S.S.K. wrote the manuscript. X.W., J.Y.Y. and L.H.W. performed the sequencing experiments and data analysis. Y.X.Z. R.Z. and D.H.L. prepared materials and performed the experiments. All authors read and approved the final manuscript.

Acknowledgements

Studies conducted by the Oil Crops Research Institute of the Chinese Academy of Agricultural Sciences were supported by the Natural Science Foundation of Hubei Province, China (2019CFB574), the Agricultural Science and Technology Innovation Project of the Chinese Academy of Agricultural Sciences (CAAS-ASTIP-2013-OCRI) China Agriculture Research System (CARS-14) the Wuhan cutting-edge application technology fund (2018020401011303) and the Science and Technology Innovation Project of Hubei province (201620000001048).

References

1. L. Wang *et al.*, "Genome sequencing of the high oil crop sesame provides insight into oil biosynthesis," *Genome Biol.*, vol. 15, no. 2, 2014, doi: 10.1186/gb-2014-15-2-r39.
2. R. Anjay Tripathi, P. A. Ravindra, P. Seema, S. Roshni, and A. R. . Ranganatha, "Study on Genetic Divergence in Sesam (*Sesamum indicum* L .) Germplasm Based on Morphological and Quality Traits," *Int. Quarterly J. Life Sci.*, vol. 8, no. 4, pp. 1387–1391, 2013.
3. A. A. Dar, S. Mudigunda, P. K. Mittal, and N. Arumugam, "Comparative assessment of genetic diversity in *Sesamum indicum* L. using RAPD and SSR markers," *3 Biotech*, vol. 7, no. 1, pp. 1–12, 2017, doi: 10.1007/s13205-016-0578-4.
4. L. Wang *et al.*, "Updated sesame genome assembly and fine mapping of plant height and seed coat color QTLs using a new high-density genetic map," *BMC Genomics*, vol. 17, no. 1, pp. 1–13, 2016, doi: 10.1186/s12864-015-2316-4.
5. L. Wang *et al.*, "HPLC analysis of seed sesamin and sesamolin variation in a sesame germplasm collection in China," *JAACS, J. Am. Oil Chem. Soc.*, vol. 89, no. 6, pp. 1011–1020, 2012, doi: 10.1007/s11746-011-2005-7.

6. S. Ghotbzadeh Kermani, G. Saeidi, M. R. Sabzalian, and A. Gianinetti, "Drought stress influenced sesamin and sesamol content and polyphenolic components in sesame (*Sesamum indicum* L.) populations with contrasting seed coat colors," *Food Chem.*, vol. 289, no. October 2018, pp. 360–368, 2019, doi: 10.1016/j.foodchem.2019.03.004.
7. K. Dossa *et al.*, "Near-infrared reflectance spectroscopy reveals wide variation in major components of sesame seeds from Africa and Asia," *Crop J.*, vol. 6, no. 2, pp. 202–206, 2018, doi: 10.1016/j.cj.2017.10.003.
8. P. J. Kanu, "Biochemical analysis of black and white sesame seeds from China," *American Journal of Biochemistry and Molecular Biology*, vol. 1, no. 2, pp. 145–157, 2011, doi: 10.3923/ajbmb.2011.145.157.
9. L. Zhou, X. Lin, A. M. Abbasi, and B. Zheng, "Phytochemical Contents and Antioxidant and Antiproliferative Activities of Selected Black and White Sesame Seeds," *Biomed Res. Int.*, vol. 2016, 2016, doi: 10.1155/2016/8495630.
10. S. Bryngelsson, L. H. Dimberg, and A. Kamal-Eldin, "Effects of commercial processing on levels of antioxidants in oats (*Avena sativa* L.)," *J. Agric. Food Chem.*, vol. 50, no. 7, pp. 1890–1896, 2002, doi: 10.1021/jf011222z.
11. L. W. Chang, W. J. Yen, S. C. Huang, and P. Der Duh, "Antioxidant activity of sesame coat," *Food Chem.*, vol. 78, no. 3, pp. 347–354, 2002, doi: 10.1016/S0308-8146(02)00119-X.
12. F. Shahidi, C. M. Liyana-Pathirana, and D. S. Wall, "Antioxidant activity of white and black sesame seeds and their hull fractions," *Food Chem.*, vol. 99, no. 3, pp. 478–483, 2006, doi: 10.1016/j.foodchem.2005.08.009.
13. D. Wang *et al.*, "Identification of nutritional components in black sesame determined by widely targeted metabolomics and traditional chinese medicines," *Molecules*, vol. 23, no. 5, pp. 1–12, 2018, doi: 10.3390/molecules23051180.
14. H. Zhang, H. Miao, L. Wei, C. Li, R. Zhao, and C. Wang, "Genetic Analysis and QTL Mapping of Seed Coat Color in Sesame (*Sesamum indicum* L.)," *PLoS One*, vol. 8, no. 5, pp. 1–10, 2013, doi: 10.1371/journal.pone.0063898.
15. V. Aggarwal *et al.*, "Minireview Molecular mechanisms of action of hesperidin in cancer: Recent trends and advancements," *Exp. Biol. Med.*, vol. 0, pp. 1–12, 2020, doi: 10.1177/1535370220903671.
16. W. Xia *et al.*, "Photo-Activated Psoralen Binds the ErbB2 Catalytic Kinase Domain , Blocking ErbB2 Signaling and Triggering Tumor Cell Apoptosis," vol. 9, no. 2, pp. 1–10, 2014, doi: 10.1371/journal.pone.0088983.
17. L. Yu, Y. Tan, and L. Zhu, "Dietary vitamin B2 intake and breast cancer risk: a systematic review and meta-analysis," *Arch. Gynecol. Obstet.*, 2016, doi: 10.1007/s00404-016-4278-4.
18. K. Dossa *et al.*, "The emerging oilseed crop *Sesamum indicum* enters the 'Omics' era," *Front. Plant Sci.*, vol. 8, no. June, pp. 1–16, 2017, doi: 10.3389/fpls.2017.01154.
19. X. Wei *et al.*, "Genetic discovery for oil production and quality in sesame," *Nat. Commun.*, vol. 6, no. May, pp. 1–10, 2015, doi: 10.1038/ncomms9609.
20. X. Wei *et al.*, "Identification of sesame genomic variations from genome comparison of landrace and variety," *Front. Plant Sci.*, vol. 7, no. AUG2016, pp. 1–12, 2016, doi: 10.3389/fpls.2016.01169.
21. C. Y. Yu, "Molecular mechanism of manipulating seed coat coloration in oilseed Brassica species," *J. Appl. Genet.*, vol. 54, no. 2, pp. 135–145, 2013, doi: 10.1007/s13353-012-0132-y.
22. R. Stark, M. Grzelak, and J. Hadfield, "RNA sequencing: the teenage years," *Nat. Rev. Genet.*, vol. 20, no. 11, pp. 631–656, 2019, doi: 10.1038/s41576-019-0150-2.
23. K. Saito *et al.*, "The flavonoid biosynthetic pathway in Arabidopsis: Structural and genetic diversity," *Plant Physiol. Biochem.*, vol. 72, no. February, pp. 21–34, 2013, doi: 10.1016/j.plaphy.2013.02.001.
24. H. Zhao, U. Basu, B. Kebede, C. Qu, J. Li, and H. Rahman, "Fine mapping of the major QTL for seed coat color in Brassica rapa var. Yellow Sarson by use of NIL populations and transcriptome sequencing for identification of the candidate genes," *PLoS One*, vol. 14, no. 2, pp. 1–25, 2019, doi: 10.1371/journal.pone.0209982.
25. M. Hong *et al.*, "Transcriptomic analysis of seed coats in yellow-seeded Brassica napus reveals novel genes that influence proanthocyanidin biosynthesis," *Front. Plant Sci.*, vol. 8, no. October, pp. 1–14, 2017, doi: 10.3389/fpls.2017.01674.

26. C. P. Leisner, C. R. Yendrek, and E. A. Ainsworth, "Physiological and transcriptomic responses in the seed coat of field-grown soybean (*Glycine max* L. Merr.) to abiotic stress," *BMC Plant Biol.*, vol. 17, no. 1, pp. 1–11, 2017, doi: 10.1186/s12870-017-1188-y.
27. L. Wang *et al.*, "Gene expression profiles that shape high and low oil content sesames," *BMC Genet.*, vol. 20, no. 1, pp. 1–11, 2019, doi: 10.1186/s12863-019-0747-7.
28. A. Kour, A. M. Boone, and L. O. Vodkin, "RNA-Seq profiling of a defective seed coat mutation in *Glycine max* reveals differential expression of proline-rich and other cell wall protein transcripts," *PLoS One*, vol. 9, no. 5, 2014, doi: 10.1371/journal.pone.0096342.
29. L. Wan *et al.*, "Transcriptome analysis of a new peanut seed coat mutant for the physiological regulatory mechanism involved in seed coat cracking and pigmentation," *Front. Plant Sci.*, vol. 7, no. OCTOBER2016, 2016, doi: 10.3389/fpls.2016.01491.
30. V. Radchuk and L. Borisjuk, "Physical, metabolic and developmental functions of the seed coat," *Front. Plant Sci.*, vol. 5, no. OCT, pp. 1–18, 2014, doi: 10.3389/fpls.2014.00510.
31. P. K. Kancharla and N. Arumugam, "Variation of Oil, Sesamin, and Sesamolin Content in the Germplasm of the Ancient Oilseed Crop *Sesamum indicum* L.," *JAOCs, J. Am. Oil Chem. Soc.*, vol. 97, no. 5, pp. 475–483, 2020, doi: 10.1002/aocs.12346.
32. A. M. Mayer, "Polyphenol oxidases in plants and fungi: Going places? A review," *Phytochemistry*, vol. 67, no. 21, pp. 2318–2331, 2006, doi: 10.1016/j.phytochem.2006.08.006.
33. E. Ono *et al.*, "Formation of two methylenedioxy bridges by a *Sesamum* CYP81Q protein yielding a furofuran lignan, (+)-sesamin," *Proc. Natl. Acad. Sci. U. S. A.*, vol. 103, no. 26, pp. 10116–10121, 2006, doi: 10.1073/pnas.0603865103.
34. T. Ke *et al.*, "Analysis of expression sequence tags from a full-length-enriched cDNA library of developing sesame seeds (*Sesamum indicum*)," *BMC Plant Biol.*, vol. 11, 2011, doi: 10.1186/1471-2229-11-180.
35. W. Xu, C. Dubos, and L. Lepiniec, "Transcriptional control of flavonoid biosynthesis by MYB-bHLH-WDR complexes," *Trends Plant Sci.*, vol. 20, no. 3, pp. 176–185, 2015, doi: 10.1016/j.tplants.2014.12.001.
36. E. L. Santos, B. H. L. N. S. Maia, A. P. Ferriani, and S. D. Teixeira, "Flavonoids: Classification, Biosynthesis and Chemical Ecology," *Flavonoids - From Biosynth. to Hum. Heal.*, pp. 3–16, 2017, doi: 10.5772/67861.
37. S. Yokoo *et al.*, "Comparative analysis of plant isochorismate synthases reveals structural mechanisms underlying their distinct biochemical properties," *Biosci. Rep.*, vol. 38, pp. 1–13, 2018.
38. M. B. Austin, J. P. Noel, and J. P. Noel, "The chalcone synthase superfamily of type III polyketide synthases," *R. Soc. Chem.*, vol. 20, pp. 79–110, 2003, doi: 10.1039/b100917f.
39. X. Cai *et al.*, "Mutant identification and characterization of the laccase gene family in *Arabidopsis*," *J. Exp. Bot.*, vol. 57, no. 11, pp. 2563–2569, 2006, doi: 10.1093/jxb/erl022.
40. R. M. Slominski, M. A. Zmijewski, and A. T. Slominski, "Commentary from the Editorial Board The role of melanin pigment in melanoma," *www.wileyonlinelibrary.com/journal/EXD*, pp. 258–259, 2015, doi: 10.1111/exd.12618.
41. Y. Xiaoli, L. Xuegang, and L. Jiana, "Mechanism of melanin synthesis in seed coat of *Brassica napus* L.," *Zuo wu xue bao*, vol. 28, no. 5, p. 638–643, 2002.
42. L. J. Min and J. F. Li, "Study on extraction technology and stability of Melanin from Banana Peel," *Adv. Mater. Res.*, vol. 815, pp. 287–292, 2013, doi: 10.4028/www.scientific.net/AMR.815.287.
43. E. Solem, F. Tuczec, and H. Decker, "Tyrosinase versus Catechol Oxidase: One Asparagine Makes the Difference," *Angew. Chemie - Int. Ed.*, vol. 55, no. 8, pp. 2884–2888, 2016, doi: 10.1002/anie.201508534.
44. H. Maeda and N. Dudareva, "The Shikimate Pathway and Aromatic Amino Acid Biosynthesis in Plants," *Annu. Rev. Plant Biol.*, vol. 63, pp. 73–105, 2012, doi: 10.1146/annurev-arplant-042811-105439.

45. F. Yang, W. Li, and H. J. L. Jørgensen, "Transcriptional reprogramming of wheat and the hemibiotrophic pathogen *Septoria tritici* during two phases of the compatible interaction," *PLoS One*, vol. 8, no. 11, pp. 1–15, 2013, doi: 10.1371/journal.pone.0081606.
46. C. Trapnell *et al.*, "Transcript assembly and quantification by RNA-Seq reveals unannotated transcripts and isoform switching during cell differentiation," *Nat. Biotechnol.*, vol. 28, no. 5, pp. 511–515, 2010, doi: 10.1038/nbt.1621.
47. S. Chen, P. Yang, F. Jiang, Y. Wei, Z. Ma, and L. Kang, "De Novo analysis of transcriptome dynamics in the migratory locust during the development of phase traits," *PLoS One*, vol. 5, no. 12, 2010, doi: 10.1371/journal.pone.0015633.
48. L. Wang, Y. Zhang, X. Qi, D. Li, W. Wei, and X. Zhang, "Global gene expression responses to waterlogging in roots of sesame (*Sesamum indicum* L.)," *Acta Physiol. Plant.*, vol. 34, no. 6, pp. 2241–2249, 2012, doi: 10.1007/s11738-012-1024-9.
49. S. Audic and J. M. Claverie, "The significance of digital gene expression profiles," *Genome Res.*, vol. 7, no. 10, pp. 986–995, 1997, doi: 10.1101/gr.7.10.986.
50. A. Mortazavi, B. A. Williams, K. McCue, L. Schaeffer, and B. Wold, "Mapping and quantifying mammalian transcriptomes by RNA-Seq," *Nat. Methods*, vol. 5, no. 7, pp. 621–628, 2008, doi: 10.1038/nmeth.1226.
51. K. J. Livak and T. D. Schmittgen, "Analysis of relative gene expression data using real-time quantitative PCR and the 2- $\Delta\Delta$ CT method," *Methods*, vol. 25, no. 4, pp. 402–408, 2001, doi: 10.1006/meth.2001.1262.

Figures

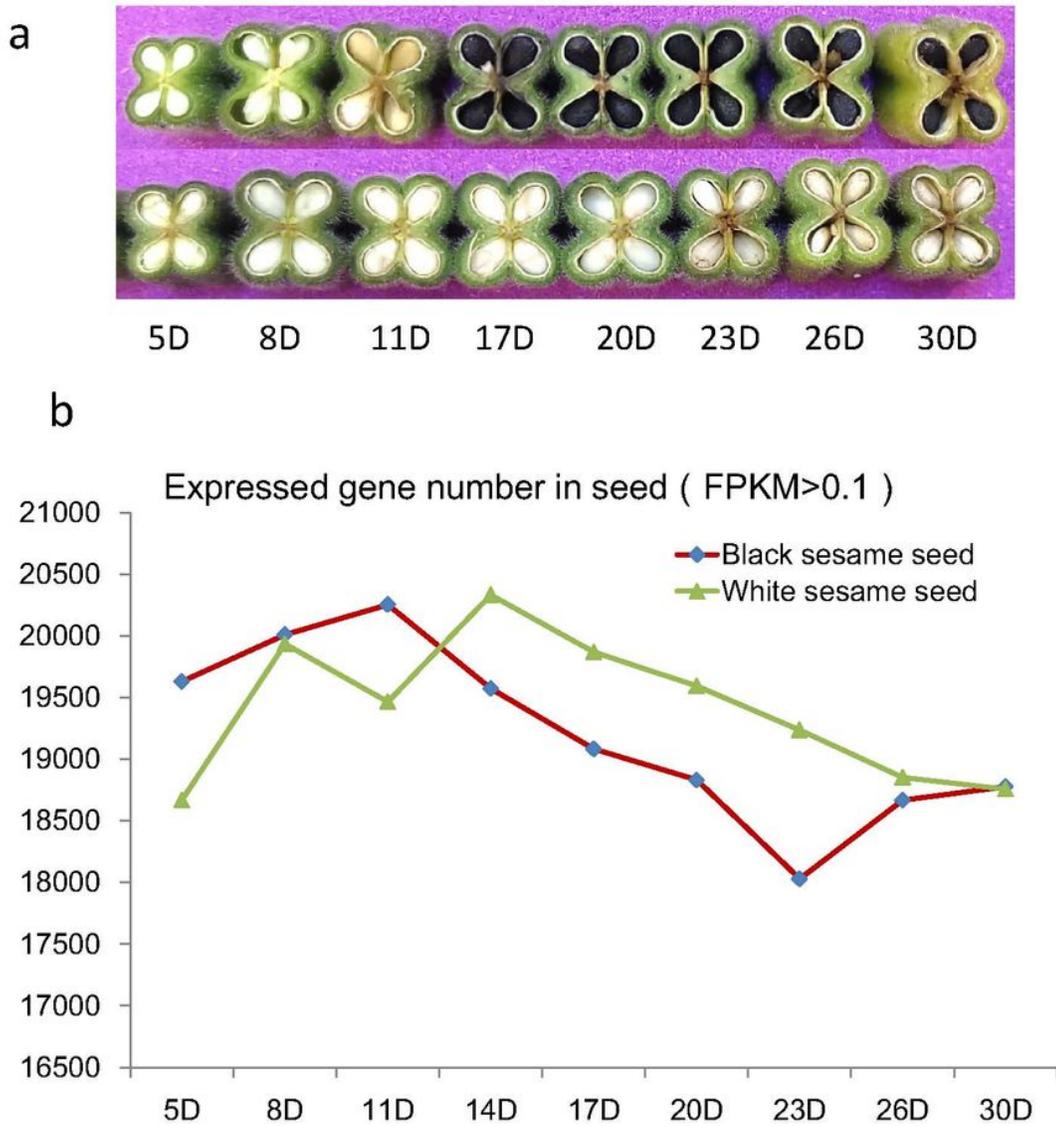


Figure 1

The development characteristics of the seed. a Seeds sampled at different DPA. b Numbers of genes expressed at different DPA in black and white sesame.

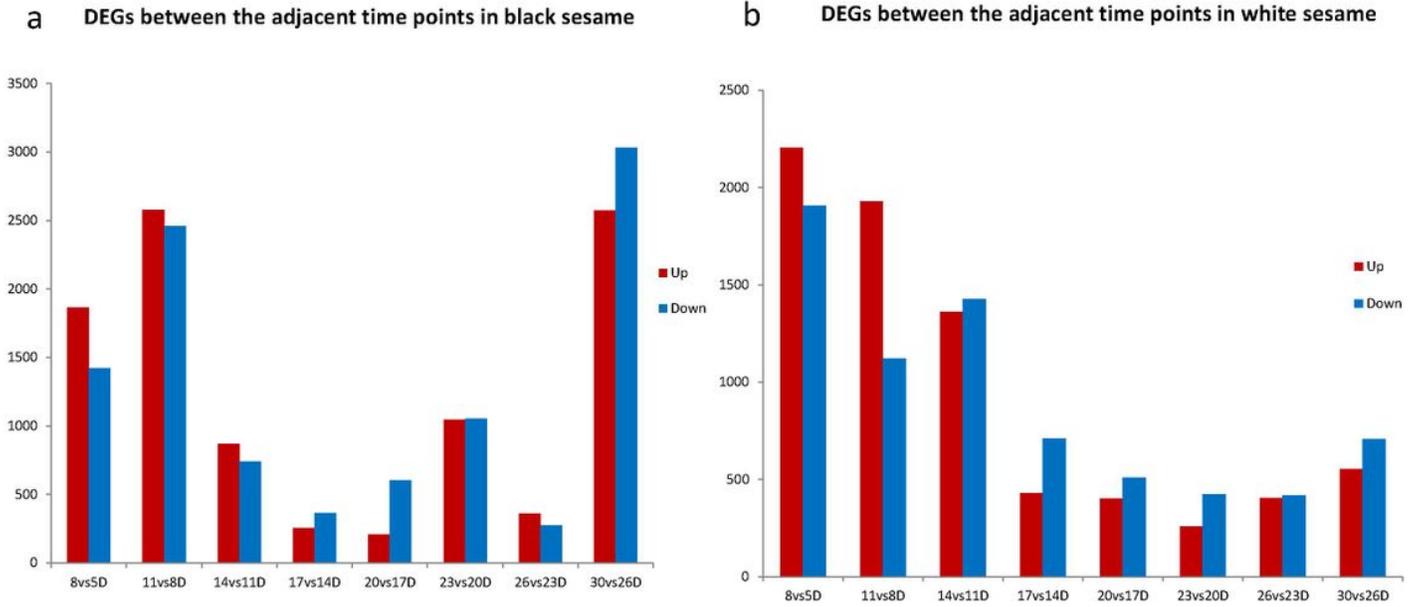


Figure 2

Dynamic change of the differentially expressed genes in black (a) and white (b) sesame seeds.

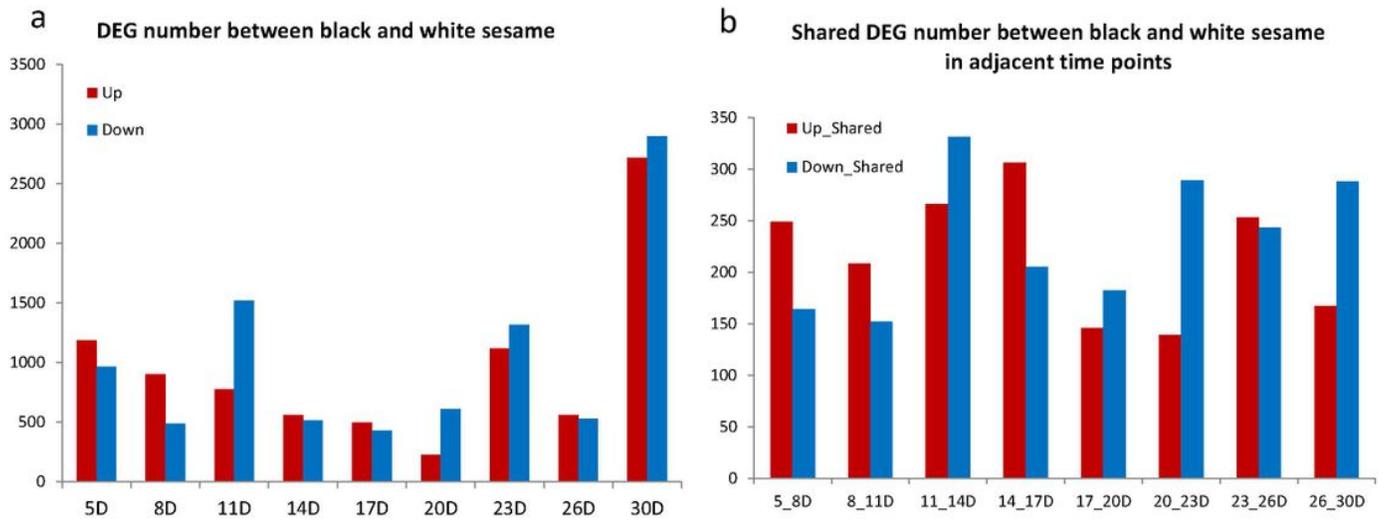


Figure 3

The differentially expressed genes between black and white sesame at different stages. a DEG number between black and white sesame. b The shared differentially expressed genes between black and white sesame at two contiguous stages.

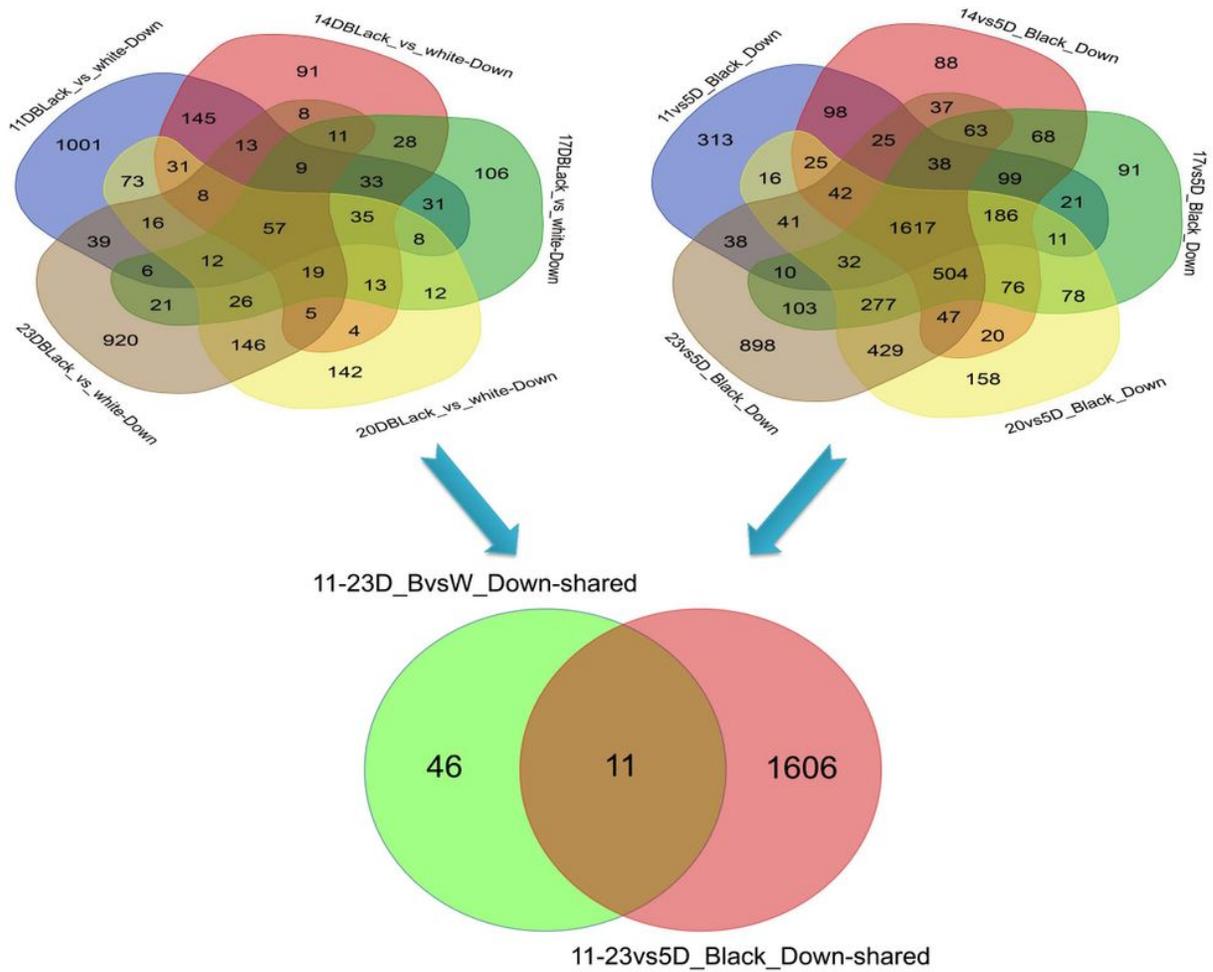


Figure 4

The common up-regulated DEGs. The gene sets were generated from the shared DEGs between black and white sesame from 11 to 23 DPA and the shared DEGs with 5 DPA as a control in black sesame from 11 to 23 DPA.

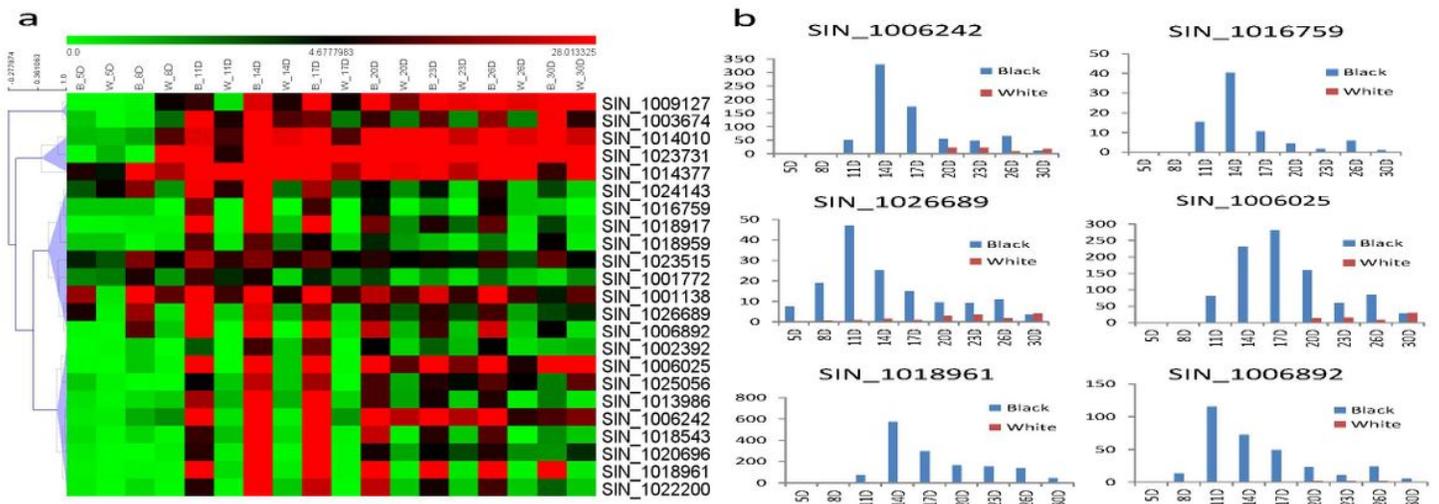


Figure 5

Hierarchical clustering and expression profiles of the 23 up-regulated DEGs in black and white sesame. (a) Hierarchical clustering of the 23 DEGs; (b) Expression profiles of the 23 DEGs in black and white sesame.

Supplementary Files

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

- [Additionalfile01.docx](#)