scientific reports

Check for updates

OPEN

Identification of transcription factors associated with leaf senescence in tobacco

Zhenbiao Zhang[1,2,3,5], Muhammad Shan[1,3,5], Liyu Yang[4,5], Songxiao Cao[1], Jie Wang[1], Wei Li[1]✉ & Yongfeng Guo[1]✉

Leaf senescence represents the final stage of leaf development, involving transcription factors (TFs)-mediated genetic reprogramming events. The timing of crop leaf senescence has a major influence on the yield and quality of crop in agricultural production. As important regulator of plant growth, the significance of TFs in the regulation of leaf senescence have been highlighted in various plant species by recent advances in genetics. However, studies on underlying molecular mechanisms are still not adequately explained. In this study, for analyzing the regulation of TFs on senescence of tobacco leaves, we combined gene differential expression analysis with weighted gene co-expression network analysis (WGCNA) to analyze the time-series gene expression profiles in senescing tobacco leaf. Among 3517 TF genes expressed in tobacco leaves, we identified 21, 35, and 183 TFs that were associated with early, middle, and late stages of tobacco leaf senescence, respectively. The Gene Ontology (GO) and Kyoto Encyclopedia of Genes and Genomes (KEGG) annotation results reveal that these senescence response TFs are correlated with several biological pathways such as plant hormone signal transduction, ubiquitin mediated proteolysis and MAPK signaling pathway, indicating the roles of TFs in regulating leaf senescence. Our results provide implications for future studies of the potential regulatory mechanisms of TFs involved in senescence of tobacco leaves.

Leaf senescence is the final stage of leaf development. During this process, active changes of cell structure, metabolism and gene expression occur in leaf cell, resulting in the chlorophyll decomposition, protein degradation, cell structure deformation or even cell death[1,2]. Superficially, leaf senescence is commonly perceived as a degenerative physiological phenomenon, however, it is, in essence, a pivotal component contributing to the evolutionary fitness of plants. In the progress of leaf senescence, macromolecules such as proteins, lipids, and nucleic acids are broken down into small molecular nutrients, which are relocated to the younger, growing organs (new buds, young leaves, and seeds), contributing to mobilization and reutilization of resources[3]. In agricultural production, abnormal leaf senescence could lead to the disorder of leaf functional phase and nutrient reallocation, which influence the production and quality of offspring. Thus, well-timed leaf senescence is not only beneficial to the fitness of whole plant, but also has important influence on the crop yield and quality[4,5]. In recent years, regulators of leaf senescence are widely investigated in various plants. Internal factors including leaf aging, variation of endogenous hormone content and external factors including nutrient limitation, drought stress and pathogens attack have been identified to have a major impact on leaf senescence[6–8]. Given the regulatory complexity of leaf senescence, current knowledge has proposed that leaf senescence is associated with various signaling pathways. Furthermore, the molecular mechanism of leaf senescence needs more extensive exploration.

Transcription factosr are delineated as nuclear proteins capable of recognizing specific sequences within the promoters of target genes, facilitating binding interactions that subsequently modulate the activation or suppression of gene expression. In past decades, with the help of high throughout sequencing and microarray, numerous senescence-associated genes (SAGs) were identified in a wide range of plants species including Arabidopsis, Populus and Wheat[9–11]. A considerable proportion of SAGs annotated as having transcription factor activity have been reported to play important roles in leaf senescence regulation[7,12]. AtNAP, a member of NAC family, has been proved highly expressed in the senescing Arabidopsis leaves, and loss-of-function mutants in AtNAP

[1]Tobacco Research Institute, Chinese Academy of Agricultural Sciences, Keyuanjing Forth Road, Qingdao 266101, Shandong, China. [2]Department of Plant Physiology, Wageningen University & Research, Wageningen, The Netherlands. [3]Graduate School of Chinese Academy of Agricultural Sciences, Beijing 100081, China. [4]Shandong Peanut Research Institute/Key Laboratory of Peanut Biology, Genetic & Breeding, Ministry of Agriculture and Rural Affairs, Qingdao, Shandong, China. [5]These authors contributed equally: Zhenbiao Zhang, Muhammad Shan and Liyu Yang. ✉email: liwei06@caas.cn; guoyongfeng@caas.cn

exhibited an obviously delayed senescence phenotype during natural developmental process compared to wild type plants, suggesting AtNAP to be a positive modulator of leaf senescence[13]. By analyzing the expression pattern and interaction effects of WRKY genes, Besseau, et al.[14] proposed that WRKY54 and WRKY70 cooperate as negative regulators of leaf senescence in Arabidopsis. Despite that previous studies have demonstrated the importance of TFs as potential regulators for leaf senescence, such simplex studies on transcription factor could not dig the regulated function of transcription factor families comprehensively and efficiently due to the versatility of transcription factor families and regulated functions.

Tobacco (*Nicotiana tabacum*) is an important commercial crop as well as a prominent model plant that is constantly being researched in the field of molecular biology. Based on the high-throughput sequencing technologies, Li, et al.[15] have provided exhaustive information about gene expression profiles in leaf senescence of tobacco. However, these studies are limited to analyzing the functions and participation of metabolic pathways of differentially expressed genes in different comparisons among developmental stages of leaf. There has been no systematic analysis on dynamics of tobacco transcription factor regulatory networks that potentially contribute to leaf senescence. In the present study, we identified transcription factor genes expressed in tobacco leaves and revealed the transcriptional dynamics of these transcription factor genes during leaf senescence process. Furthermore, we conducted co-expression network-based analysis to identify a series of transcription factors that regulate tobacco leaf senescence at different developmental stages. In addition, we predicted the putative targets of these senescence-associated TFs and constructed the corresponding regulatory network to explore their biological function and downstream pathways. These results deepen our understanding of the important regulatory function of transcription factors in tobacco leaf senescence.

Materials & methods

Sequencing data resource and processing

The sequencing data of tobacco leaves were downloaded from Sequence Read Archive (SRA) database, with accession number PRJNA376075. This dataset contains a total of 20 samples (2 technical replicates per sample). All the leaf samples from common tobacco (*Nicotiana tabacum* L.) variety Zhongyan100 were collected from 15 days after topping (DAT). The upper leaf (UL) samples (leave at the 16th leaf position) and middle leaf (ML) samples (leave at the 9th to 10th positions) were collected at 8 different time points (15, 25, 35, 45, 55, 65, 75 and 85 DAT) and the lower leaf (LL) samples (leave at the 4th position) were collected at 4 different time points (15, 25, 35 and 45 DAT). These SRA files were transformed into FASTQ files using the SRA Toolkit from NCBI and then processed as the methods previously described by Li et al.[15]. After data processing, a total of 272, 920 unigenes were identified for further analysis. The expression level of each unigene was normalized using the FPKM (Fragments per Kilobase Million) algorithm in each sample. Unigenes with a fold change in expression level between two samples of more than 2 and a false discovery rate (FDR) of < 0.001 were considered as differentially expressed genes (DEGs).

Identification of modules related to the tobacco leaf senescence

WGCNA (weighted gene co-expression network analysis) is a method to identify clusters (modules) of genes that are highly correlated by analyzing the correlation coefficients of the expression profiles of genes[16]. Given the large number of differentially expressed genes (DEGs) generated in the present study, only DEGs with expression values in the top 25 percent were retained to ensure the significance of the mining results. The Pearson's correlation matrix was constructed by calculating the absolute Pearson correlations between pairs of remaining DEGs. After determining the appropriate power value (power value = 10) based on the scale-free topology criterion, the WGCNA network was constructed. Densely interconnected and co-expressed genes were then clustered into 21 modules, each assigned a different color. After calculating the Pearson correlations between the module eigengene and traits, the module-trait relationships were estimated. The module-trait relationship could help us discover the modules closely related to leaves in different senescence stages.

Identification and expression analysis of TF genes

PlantTFDB is a public database used to identify and catalog all plant genes involved in transcriptional control. In this study, TF genes were predicted using TF prediction server (http://planttfdb.gao-lab.org/prediction.php) provided by this database. This tool used HMMER 3.0[17], downloaded from Pfam (version 27.0)[18], to identify TFs and assigned them into different TF families based on the sequence characterization of DNA binding domain. The details of TF prediction pipeline were described by Jin, et al.[19].

Functional annotation and biological pathway analysis of unigenes

To understand the functions of unigenes, we performed GO annotation analysis using the Blast2GO (v2.5.0) software. Then unigenes were then classified into different categories using WEGO 2.0 (Web Gene Ontology Annotation Plotting) website[20]. KEGG (Kyoto Encyclopedia of Genes and Genomes) database provides the information about the biological pathways in which genes are involved. The pathway of unigene was annotated by assigning the sequences of unigenes to the KEGG database via BLASTx (E-value < 0.00001).

To determine the most relevant function and biological pathway of TFs and TF targets associated with tobacco leaf senescence, we performed GO and KEGG enrichment pathway for the TF genes or TF target genes by using TBtools software[21]. The calculated p-value underwent Bonferroni correction. GO terms or KEGG pathways with a corrected p-value < 0.05 were considered significantly enriched.

Construction of TF and TF target network

PlantRegMap (http://plantregmap.gao-lab.org/) is a database integrating the transcription factor regulation information of 132 plant species. By using the regulation prediction tool of this database, the regulatory interactions of tobacco TFs were revealed. In the present study, unigenes were rematched based on the *Nicotiana tabacum* L. symbol data from PlantRegMap. Then the early, middle, and late senescence-response TF regulatory network were constructed according to the TF regulation network of *Nicotiana tabacum* L. downloaded from PlantRegMap database by using Cytoscape software (vision 3.6.1)[22]. The early, middle, and late senescence response hub TF genes (degree > 300) were identified by using Centiscape (Version 2.2) (Shannon et al., 2003).

qRT-PCR validation of TF genes

Total RNA was extracted from leaves of tobacco (CV-ZY100) by using Trizol (Invitrogen, Burling-ton, CA) and then was reverse transcriptased to cDNA by using PrimeScript™ RT reagent Kit with gDNA Eraser (Perfect Real Time) (TaKaRa, Tokyo, Japan). Primers used for qRT-PCR were designed by Primer5 software. The qRT-PCR was carried out using the ABI Step One Plus Real-Time PCR System (Applied Biosystems, USA) with the TB Green™ Premix Ex Taq™ II (Tli RNase H Plus) (TaKaRa, Tokyo, Japan). The PCR amplification reactions were 95 ℃ for 10 min followed by 40 cycles of 95 ℃ for 15 s and 60 ℃ for 35 s. Three biological replicates were implemented for each sample. Results were calculated using using $2^{-\Delta\Delta CT}$ method. Actin7, cited from Li, et al.[15], was used as the internal reference control to standardize the results. All primers used in this study were listed in Table S1.

Results

Sequencing data processing and TF genes identification

The global gene expression profiles of 20 tobacco leaf sample collected from different developmental stages were downloaded from the SRA database and were processed as the methods previously described by Li, et al.[15]. Following data processing, 272,920 unigenes were identified for subsequent analysis. In this study, TF prediction server (http://planttfdb.gao-lab.org/prediction.php) was used to predict the ability of identified unigenes to encode transcription factor based on their sequence characteristics. As a result, a total of 3617 transcription factor genes were identified to be expressed in tobacco leaf and classified into 56 TF families using the classification criterion in PlantTFDB (Fig. 1, Table S2). The number of TFs belonging to bHLH, B3, MYB_related, C3H, bZIP, ERF, WRKY, C2H2, FAR1, and NAC families was very large, and they were widely expressed in tobacco leaves at different senescence stages.

Among, some members of TF families expressed at specific stages during tobacco leaf senescence. As shown in Fig. 2, 239 senescence-associated TF genes (21 early senescence-associated TF genes, 35 middle senescence-associated TF genes, and 183 late senescence-associated TF genes) belong to 35 TF families (Fig. 2 and Table S2). Members of the C3H, ERF, HD-ZIP, and WRKY families were found to be associated with tobacco leaf senescence

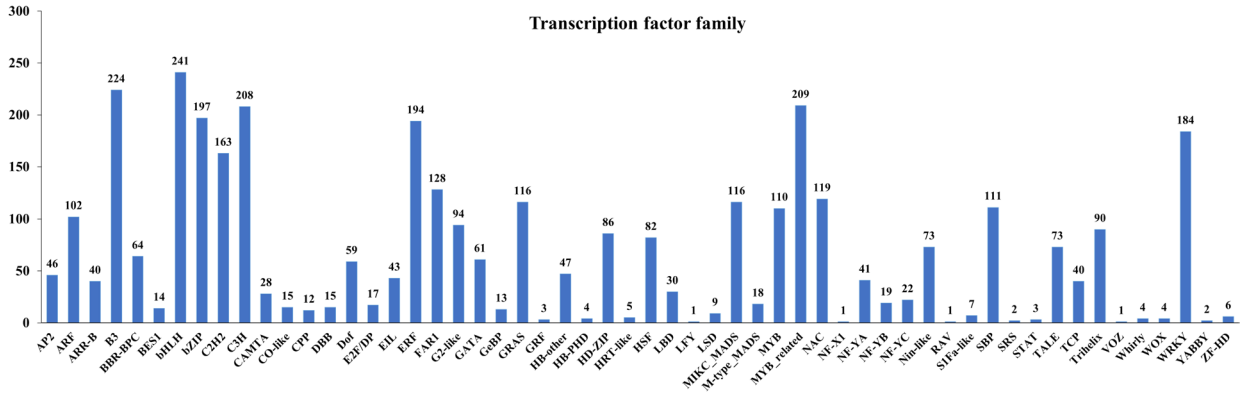

Fig. 1. The distribution of tobacco transcription factors in different families.

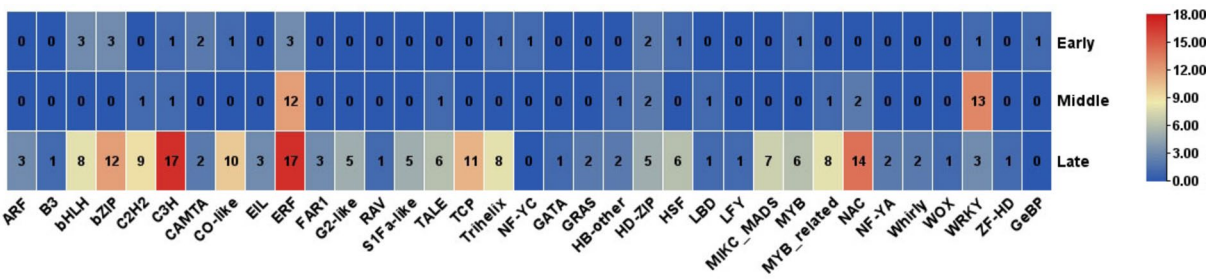

Fig. 2. The distribution of early, middle, and late senescence-associated transcription factors in different families.

at all three senescing stages, suggesting that TF genes in these families play an extensive regulatory role in the process of tobacco leaf senescence.

Transcription factors are involved in the growth and development of plants by regulating target genes. The expression levels of TF genes have an important impact on the degree of regulation. Therefore, the expression trends of transcription factor genes during leaf senescence are crucial to the study of the regulatory role of transcription factors during leaf senescence. As shown in Fig. 3, most identified early, middle, and late senescence response TF genes, respectively, exhibit relatively higher expression levels at their corresponding stages than at the other two senescence stages, indicating the potential regulatory roles of these senescence response TF genes in the early, middle, and late stages of tobacco leaf senescence. These TF genes may be associated with the variations of leaf morphology and physiological metabolism in the tobacco leaf senescence process.

Weighted gene co-expression networks construction, identification, and expression analysis of senescence-related TF genes

To investigate the important regulatory genes of tobacco leaf senescence, we screened for genes that were differentially expressed during the senescence of tobacco leaves. As UL, ML, and LL samples were all collected at 15 DAT (the starting time point), the gene expressions at 15 DAT were defined as the reference. All subsequent time points were compared to the starting point. As a result, a total of 64,964 differentially expressed genes (DEGs) were identified through pair comparison between different stages of UL, ML, and LL, under the threshold of fold change ≥ 2.0 and FDR ≤ 0.001 (Table S3). Of all the identified senescence-associated DEGs, 1,359 were TF genes, accounting for 30% of the identified TF genes expressed in tobacco leaves, indicating tight regulation of the transcriptional activity (Table S4). Although the systemic analysis provides sufficient information about genes differentially expressed during tobacco senescence, the large number of identified DEGs resulted in the lack of specificity in distinguishing key senescence-related genes. To further excavate the co-expressed modules and key genes associated with leaf senescence, we constructed a co-expression network based on the identified DEGs. As a result, DEGs were clustered into 21 modules (module1-module 21), with modules ranging in size from 61 to 8876 unigenes (Table S5). Each module represented genes with highly correlated expression profiles. The correlation coefficients between modules and developmental stages were shown in Fig. 4. According to the senescence phenotype of the leaf and the expression feature of senescence marker genes (CP1, SAG12 homolog in tobacco) in all samples, the upper and middle leaves collected at 15 DAT, 25 DAT, and 35 DAT were considered to be in the early senescence stage. The upper and middle leaves collected at 45 DAT, 55 DAT, 65 DAT, and lower leaves collected at 15 DAT and 25 DAT were considered to be in the middle senescence stage. The upper and middle leaves collected at 75 DAT, 85 DAT, and lower leaves collected at 35 DAT and 45 DAT were considered to be in the late senescence stage. We found that most samples at the particular senescence stage were associated with the same module. The most relevant modules of samples at different senescence stages were different from each other. Therefore, the co-expression analyses also highlight notable differences among different leaf senescence stages. By module-trait correlation analysis, relative to other modules, 6 (modules 2, 3, 5, 6, 8, 9), 8 (modules 1, 4, 11, 12, 13, 14, 15, 16), and 6 (modules 7, 10, 17, 18, 19, 20) modules were found to be the most relevant modules with leaf samples at early, middle, and late senescence stage, respectively. 21, 35, and 183 TF genes were clustered into modules associated with early, middle, and late senescence stages, respectively. Hence, these TF genes were regarded as early, middle, and late senescence-associated TF genes.

Gene ontology and pathway analysis of TF genes associated with different senescence stages

To characterize the functional changes of tobacco leaf senescence-associated TFs, we conducted GO annotation of these TFs (Fig. 5). Based on the functional annotation of senescence-associated TF genes, we assigned 21 early senescence-associated TF genes, 35 middle senescence-associated TF genes, and 183 late senescence-associated TF genes with 39, 46, and 196 GO terms, respectively. These genes were grouped into three main GO categories: molecular function, cellular component, and biological process (Fig. 5A, Table S2). Among them, the most abundant GO terms assigned by these senescence-associated TF genes in the cellular component, molecular function, and biological process category were nucleus (GO:0005634), sequence-specific DNA binding transcription factor activity (GO:0003700), and regulation of transcription, DNA-dependent (GO:0006355), indicating that TFs in tobacco leaf at different senescence stages were active. In the biological process category, many senescence-associated TFs were correlated with leaf development, suggesting that these TFs are key regulators of

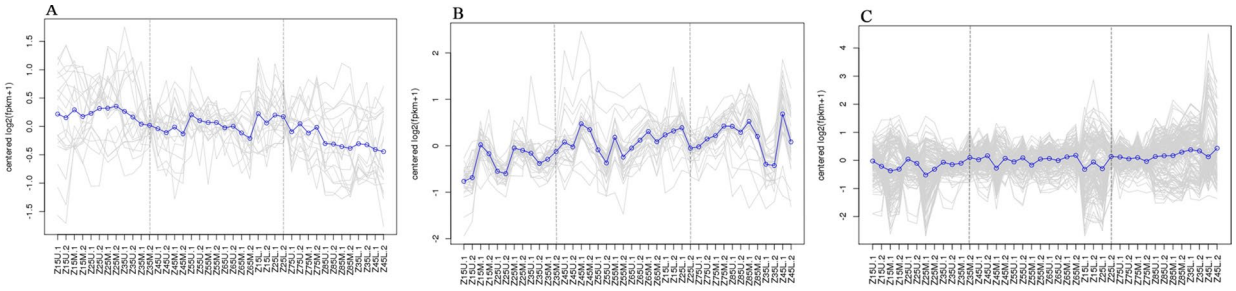

Fig. 3. Trend plots of the normalized expression levels of early, middle, and late senescence-associated TF genes in different leaf samples. The expression value in the broken line graphs were log2-transformed FPKM values.

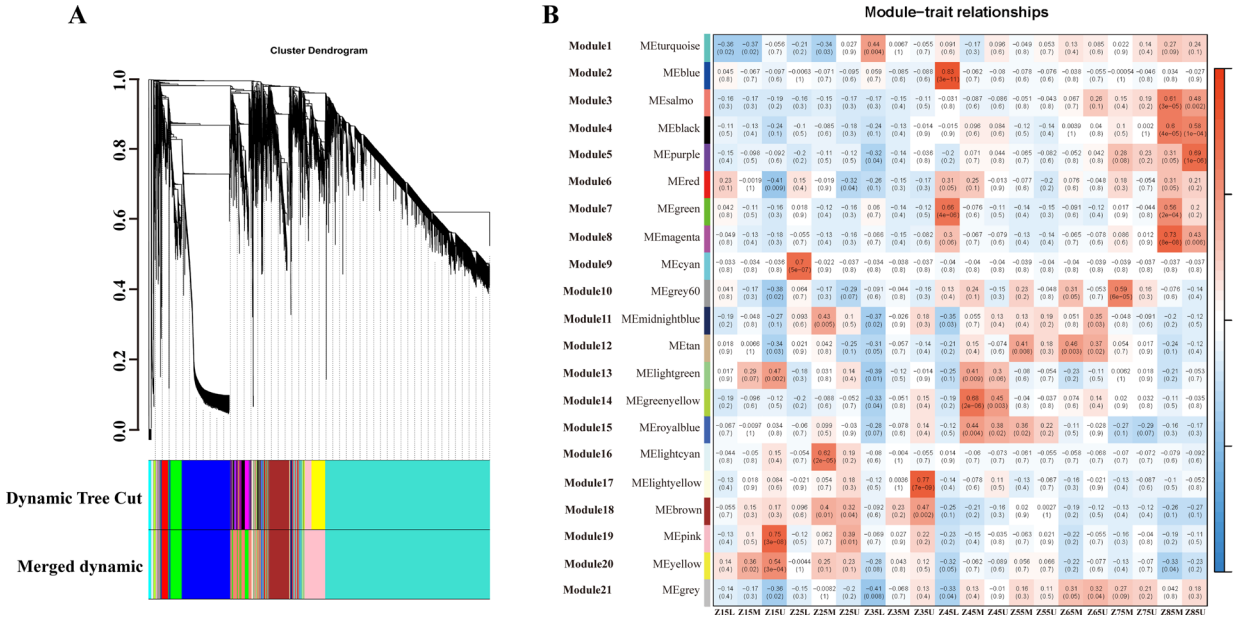

Fig. 4. Network analysis of tobacco leaf senescence. (**A**) Gene clustering tree (dendrogram) obtained by hierarchical clustering of adjacency-based dissimilarity. The colored row below the dendrogram indicates different modules identified by WGCNA. (**B**) Correlations between module eigengenes and leaf sample. Biological traits (rows) were regrouped in 3 categories (early senescence trait, middle senescence trait and late senescence trait). In the heatmap, positive correlation and negative correlation were represented by red color and blue color.

leaf development (Fig. 5B). According to the GO enrichment analysis results, six GO terms, including transcription regulator activity (GO:0140110), organic cyclic compound binding (GO:0097159), heterocyclic compound binding (GO:1901363), binding (GO:0005488), biosynthetic process (GO:0009058), and nitrogen compound metabolic process (GO:0051171) (corrected p-value < 0.05) (Fig. 5C, Table S6), were significantly enriched by early, middle, and late senescence-associated TF genes.

To investigate which metabolic pathways the identified TFs were associated with, we obtained the pathway annotations of TF genes using the KEGG database[23–25]. As a result, we found that early, middle, and late senescence-associated TF genes were associated with 10, 2, and 13 metabolic pathways, respectively (Fig. 6, Table S7). According to the KEGG enrichment analysis results, no pathways were enriched by the early senescence-associated TF genes. Middle senescence-associated TF genes were enriched in the plant pathogen interaction pathway. Pathways of circadian rhythm plant, plant hormone signal transduction, MAPK signaling pathway plant, and plant pathogen interaction were enriched by late senescence-associated TF genes (Fig. 6, Table S7). These data suggest that complex regulatory mechanisms underlie tobacco leaf senescence via TF genes.

Transcription factor regulatory network in senescent tobacco leaves

Transcription factors are indispensable regulatory factors in the life activities of higher plants. Some transcription factor genes may regulate the expression of many key genes associated with important metabolic pathways. Therefore, identifying the target genes of senescence-associated transcription factors is important for unraveling regulatory networks in senescing tobacco leaves. PlantRegMap is a database that integrates a series of regulatory data, including regulatory information of transcription factors and other gene regulatory elements obtained from relative references and experimental data. Based on the information of TF-TF-target interactions in tobacco extracted from the PlantRegMap database, 797, 756, and 4552 unigenes were found to be regulated by early, middle, and late senescence-associated TF genes, respectively (Table S8). To explore the dynamics of TF regulatory networks across tobacco leaf senescence, we constructed early, middle, and late senescence-associated TF gene regulatory networks using Cytoscape software[22]. (The early, middle, and late senescence-associated TF gene regulatory networks were shown in Figure S1). Hub genes, which are strongly related to numerous genes, have been shown to play important regulatory roles in gene expression networks. Based on the degree of the node, 2, 1, and 6 hub genes in the early, middle, and late senescence-associated TF regulatory networks, respectively, were identified. Genes with high values of degree (> 300), including Unigene34594 (HSF family member), Unigene29760 (ERF family member), CL11135.Contig1 (WRKY family member), CL11180.Contig2 (TCP family member), CL20581.Contig2 (C3H family member), Unigene47914 (HSF family member), CL7305.Contig4 (HSF family member), Unigene47579 (CAMTA family member), and CL17451.Contig2 (ERF family member), were observed as hub genes in the network.

To further excavate the TF regulating effects on the tobacco leaf senescence, we focused on analyzing the functions and biological pathways associated with targets of senescence-associated TFs at different developmental stages. As a result, 797, 756, and 4552 targets of early, middle, and late senescence-associated TFs were assigned to

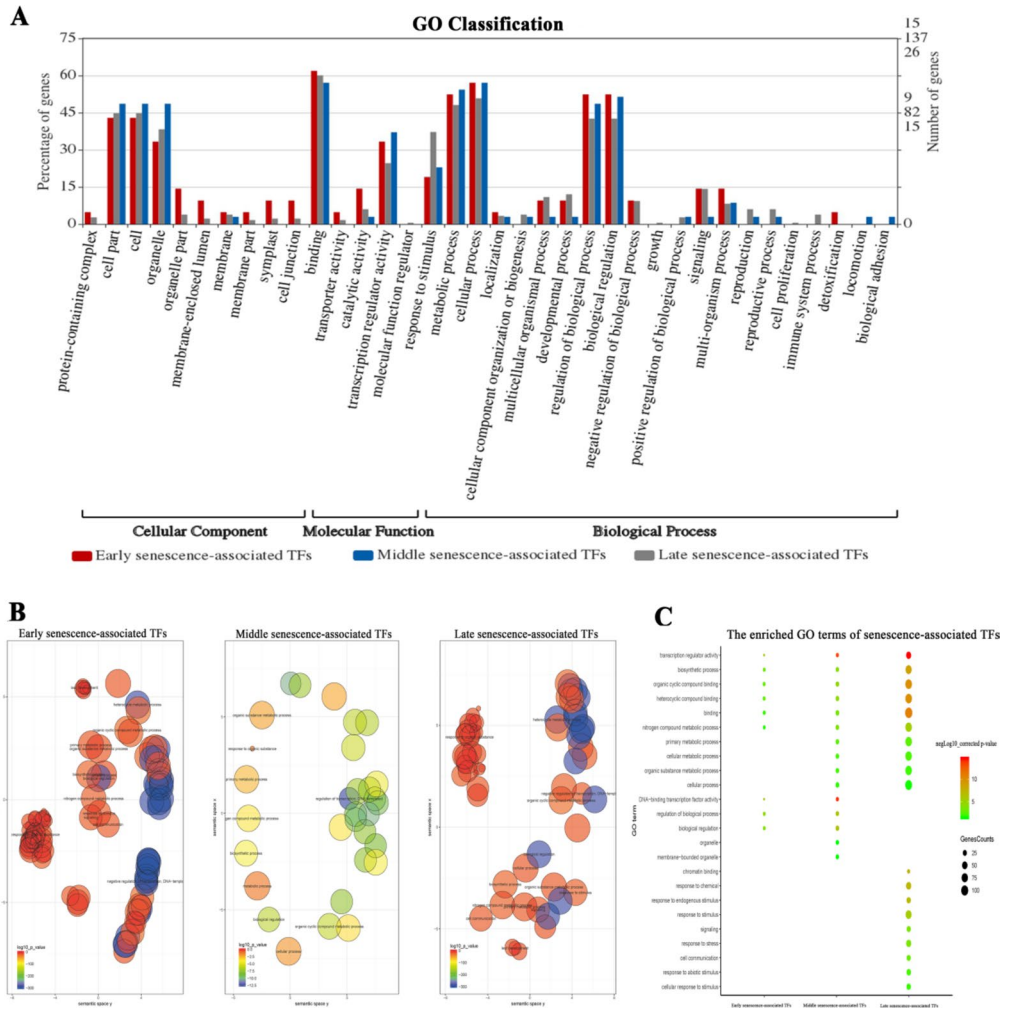

Fig. 5. GO analysis of senescence-associated TFs. (**A**) GO classifications of early, middle, and late senescence-associated TF genes. (**B**) The enriched GO terms (corrected P-value ≤ 0.05) annotated by senescence-associated TF genes. Each dot indicates a GO term. The dot sizes indicate the number of genes. Colors indicate the significance of the GO term. (**C**) The results of enrichment analysis of GO terms annotated by senescence-associated TFs in biological process category are shown as REVIGO scatterplots in which similar GO terms are grouped in arbitrary two-dimensional space based on semantic similarity. Each dot indicates a GO term. The dot sizes indicate the number of genes. Colors indicate the significance of the GO term.

38, 41, and 51 GO terms, respectively (Table S8). A high percentage of targets of senescence-associated TFs from GO terms of cell part (GO:004446), binding (GO:0005488), and response to stimulus (GO:0050896) (Fig. 7A). GO terms of response to abiotic stimulus (GO:0009628), response to stress (GO:0006950), and response to stimulus (GO:0050896) were significantly enriched (corrected p-value < 0.05, see Fig. 7B, Table S6) by the targets of early and late senescence-associated TFs, suggestive of protective mechanisms against potential stresses during tobacco leaf senescence. No significant enrichment of GO term was observed among the targets of middle senescence-associated TFs.

Previous studies have revealed many biological pathways associated with leaf senescence, including stress resistance, plant hormone signal transduction, and protein metabolism pathways. In the present study, the targets of early, middle, and late senescence-associated TFs were assigned to 94, 85, and 130 metabolic pathways according to KEGG annotation (Table S7). Based on the KEGG enrichment analysis results, targets of early senescence-associated TFs were enriched in protein processing in the endoplasmic reticulum, zeatin biosynthesis, and spliceosome pathways. Middle senescence-associated TFs enriched pathways of flavone and flavonol biosynthesis, zeatin biosynthesis, and plant pathogen interaction. Late senescence-associated TFs were enriched in three pathways, including protein processing in the endoplasmic reticulum, ubiquitin-mediated proteolysis, and flavonoid biosynthesis (Fig. 8, Table S7). These results revealed the targets of senescence-associated TFs responsible for essential biological functions during leaf senescence. These senescence-associated TFs may control tobacco leaf senescence by regulating the expression of genes involved in these metabolic pathways.

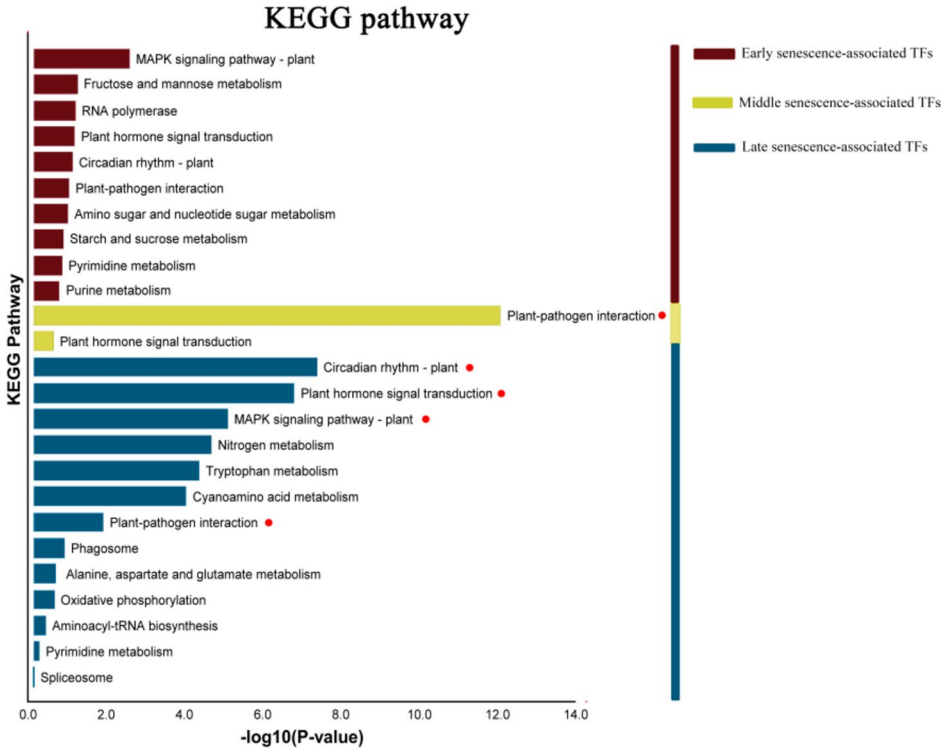

Fig. 6. Barcharts represent the KEGG pathway annotated by early, middle, and late senescence response TF genes. X axis is the normalized corrected p-value. Y axis is the KEGG pathway. The KEGG pathways enriched by early, middle, and late senescence response TF genes were marked by the red dot.

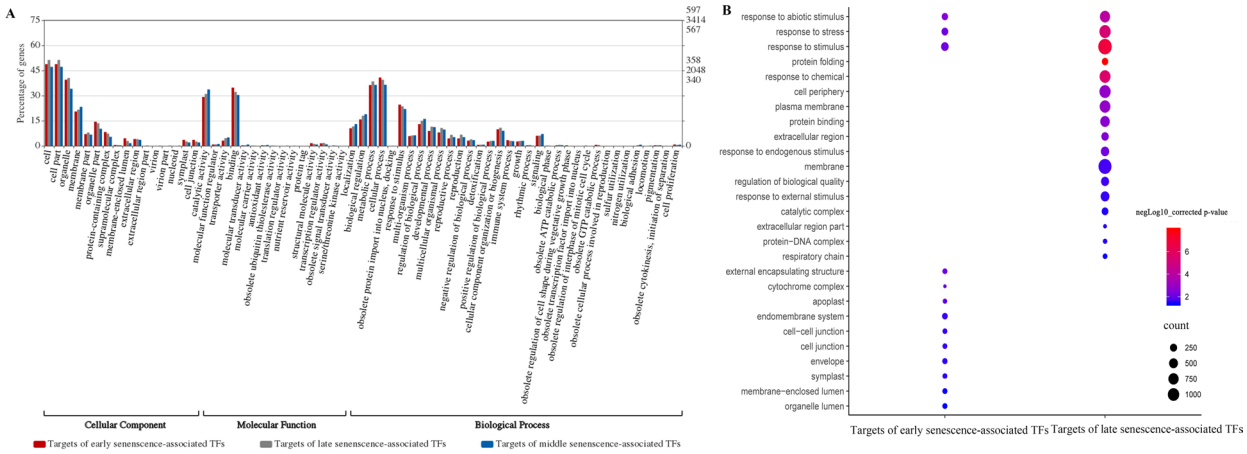

Fig. 7. GO analysis of senescence-associated TFs. (**A**) GO classifications of targets of early, middle, and late senescence-associated TFs. (**B**) The enriched GO terms (corrected P-value ≤ 0.05) annotated by targets of senescence-associated TF genes. Each dot indicates a GO term. The dot sizes indicate the number of genes. Colors indicate the significance of the GO term.

qRT-PCR analysis of hub TF genes selected from network

In the current study, hub genes selected from senescence response TF-TF-target networks were subjected to qRT-PCR analysis using samples collected from mature leaves (ML) at different time points (15, 25, 35, 45, 55, 65, 75, and 85 (DAT)). As a result, the ex-pression patterns of these hub TF genes determined by qRT-PCR were basically consistent with the transcriptome sequencing data, confirming the reliability of the transcriptome data used in this study. The hub TF genes detected from different TF-TF-target networks exhibited different time-course expression characteristics during leaf senescence. The results will help infer their potential roles in regulating the senescence of tobacco leaves. The qRT-PCR results are shown in Fig. 9, and the primer sequences are available in Table S1.

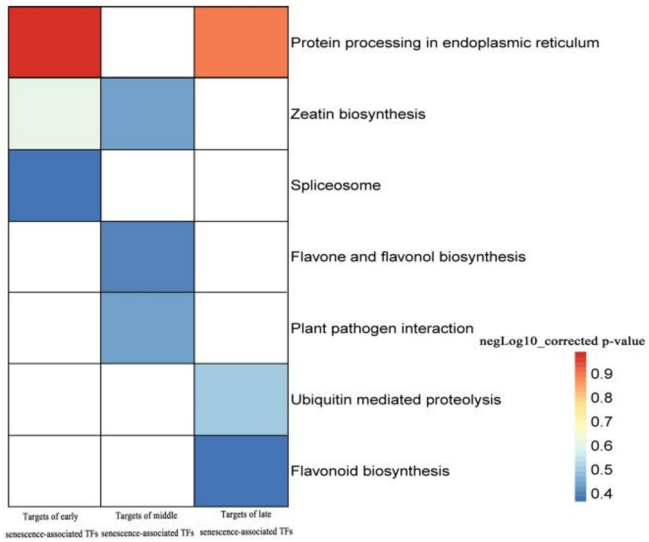

Fig. 8. Heatmap of KEGG pathway which enriched by the targets of early, middle, and late senescence-associated TF genes. The color from red to blue represent the − Log10 corrected p-values.

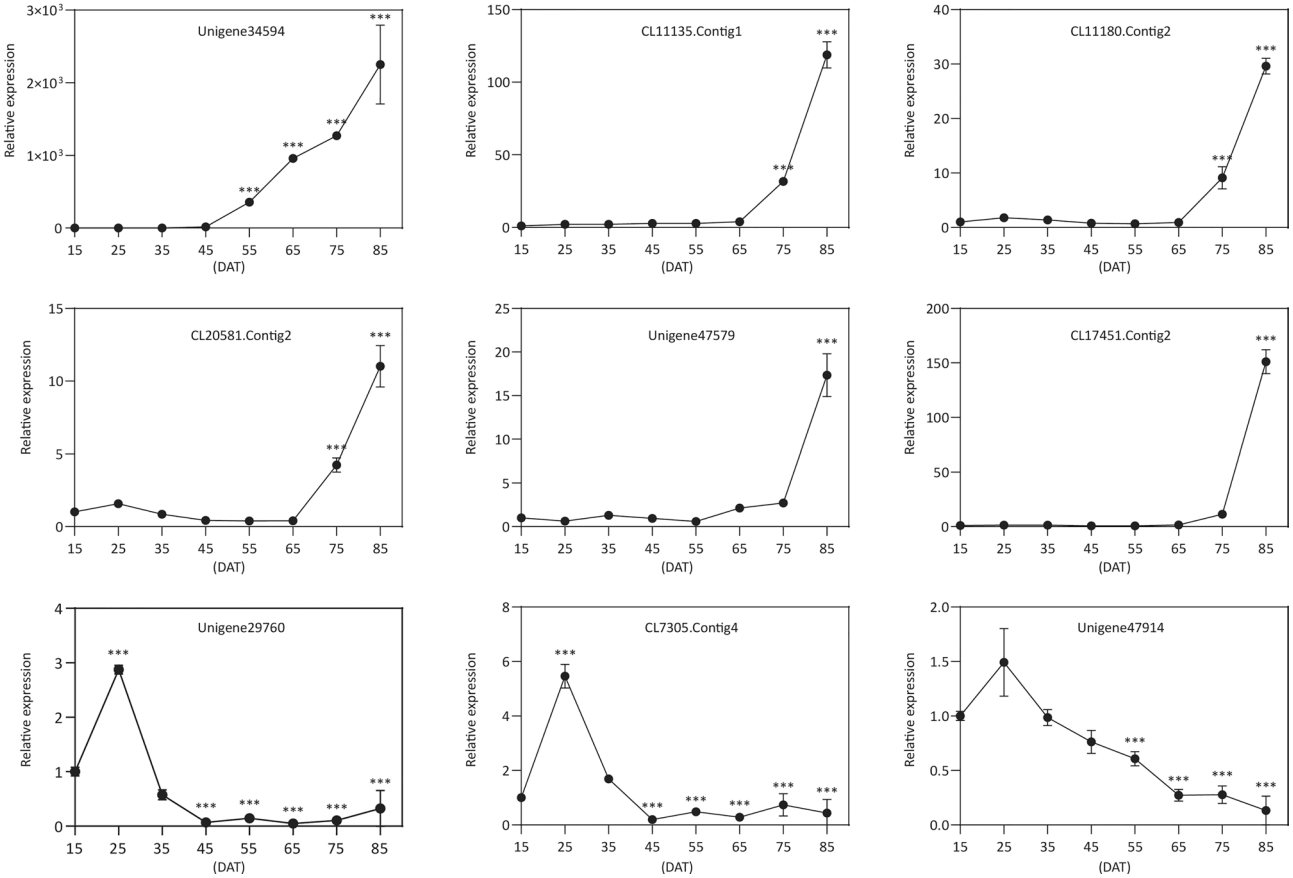

Fig. 9. Expression analysis of hub genes associated with senescence response TF-TF-targets networks at indicated time points. The error bars above line present ± SD with three biological replicates and asterisks denote a statistically significant difference compared with expression level at 15 DAT according to one-way ANOVA analysis (***$P < 0.01$).

Discussion

Based on transcriptome data of tobacco leaves at three different positions in eight different developmental phases, this study identified transcription factors in the leaves of tobacco and analyzed variations in function and expression. By constructing co-expression network modules, we systematically investigated specific modules during senescence and focused on transcription factor genes that play important regulatory roles. We attempted to explain the internal regulation mechanism of transcription factors during tobacco leaf senescence by constructing a regulatory network involving transcription factors relevant to senescence and their target genes. According to our analyses, we can better understand the regulatory functions played by transcription factor genes during to bacco leaf senescence.

As a gradual process, leaf senescence can be generally divided into three stages early, middle, and late senescence stage. In recent years, numerous studies have reported that transcription factors are crucial transcription regulatory factors for the onset of leaf senescence[26–29]. In rice, SUBMERGENCE 1A (SUB1A) is a member of ERF family. Conditional and ectopic overexpression of SUB1A significantly delayed the onset of leaf senescence. Thus, SUB1A is a negative regulator of the onset of leaf senescence[30]. In Arabidopsis, WRKY53 was highly expressed in rosette leaf at early stage of leaf senescence and attenuated thereafter. RNAi-mediated knockdown of WRKY53 induces the delay of leaf senescence, while overexpression of WRKY53 causes accelerated leaf senescence. These results strongly suggest that the function of WRKY53 is as a positive regulator for the initiation of leaf senescence[31,32]. In the present study, based on the results of WGCNA analysis, the expression of three members (CL18684.Contig1_All, CL7330.Contig1_All, Unigene29760_All, and Unigene47139_All) of the ERF family and 1 member (CL18684.Contig1_All) of the WRKY family exhibited high correlation with the early senescence of tobacco leaves. These TF genes of ERF and WRKY families may have similar putative functions as regulators of the onset of leaf senescence. Previous studies have demonstrated that the content of glucose in Arabidopsis could influence the photosynthetic activity through sugar sensing and signaling pathways and thus trigger the senescence of leaves[33,34]. Metabolic analyses prove that substantial changes in sugar metabolism occur during tobacco leaf senescence. The content of sugar in upper tobacco leaves exhibits a completely opposite trend from early to middle senescence, decreasing first from 15 to 35 DAT and then starting to increase significantly at 45 DAT[15]. Therefore, we speculate that carbohydrate metabolism is an important pathway affecting the onset of tobacco leaf senescence. In the present study, we found that 12 targets of 2 unigenes (Unigene34594_All and CL18684.Contig1_All), which encode proteins that belong to HSF and ERF families, were assigned to the pathway of starch and sucrose metabolism. These two TFs may function in regulating the early senescence of tobacco leaves by regulating sugar metabolism.

Once senescence starts, the obvious visible morphological alteration in the processional phase of leaf senescence is the yellowing of leaves. This change is mainly due to the degradation of chlorophylls, proteins, lipids, and other macromolecules in the leaf cell[35]. Previously, a proteomic study of senescing leaves indicated that leaf proteins, especially the chloroplast proteins, exhibit a decline in abundance during senescence, which could be degraded by pathways of ubiquitin-mediated proteolysis, SAVs (senescence-associated vacuoles) mediated proteolysis, and autophagy[36]. In our tobacco leaf transcriptome database, among all the identified 35 TF genes associated with the procession of leaf senescence, the targets of 4 TFs (CL11135.Contig1_All, CL12202.Contig4_All, CL10608.Contig2_All, and Unigene58069_All), which encode the proteins belonging to HD-zip and WRKY families, were found to be associated with various protein degradation pathways, including ubiquitin-mediated proteolysis, autophagy, phagocytosis, and protein export. This indicates that these TF genes are probably important for chloroplast protein degradation in the process of leaf senescence in tobacco. In the present study, it is appealing that a considerable portion of middle senescence-associated TFs targets assigned to the plant hormone signal transduction pathway based on the KEGG analysis. In fact, much evidence supports the role of phytohormones interacting with TFs as important endogenous regulators of leaf senescence. WRKY57, whose target mRNAs encode repressors of JA and auxin signaling pathway, JAZ4/8 and IAA29, functions in jasmonic acid-induced leaf senescence in Arabidopsis[37]. CL11135.Contig1_All (a WRKY family member), as the hub gene in middle senescence-associated TF genes regulatory network, was predicted to target genes correlated with ethylene, cytokinin, and ABA-mediated signaling pathway, and its targets take a considerable portion of targets assigned to the plant hormone signal transduction pathway. This indicates that this WRKY gene may function as an important regulator during the procession of leaf senescence by regulating the expression of plant hormone-related genes. Further studies will be needed to investigate the regulatory mechanism underlying leaf senescence of tobacco in this process.

At the terminal phase of leaf senescence, the entire leaf turns almost entirely yellow. The morphological characteristics of leaf cells change significantly, including volume reduction, chromatin condensation, nucleus shrinking, DNA fragmentation, and plastoglobuli accumulation[38]. The overwhelming majority of leaf cells undergo programmed cell death (PCD). In the current study, many late senescence-associated TF genes of the G2-like and bZIP families were found to be closely correlated with PCD-associated metabolic pathways, including porphyrin and chlorophyll metabolism, phagosome, oxidative phosphorylation, and natural killer cell-mediated cytotoxicity pathways. This suggests that these TFs may be involved in the leaf senescence of tobacco by participating in the regulation of pathways associated with PCD. By analyzing the expression profiles of genes expressed in the senescing leaf of Arabidopsis, Breeze, et al.[39] found that more than 30 NAC TFs were induced during senescence process, indicating their participation in regulating the leaf senescence. Kim, et al.[40] reported that ORE1, which is a NAC transcription factor, positively regulates aging-induced cell death in Arabidopsis leaves. Findings of Matallana-Ramirez, et al.[41] proved that BFN1 (BIFUNCTIONAL NUCLE-ASE1), which encodes a nuclease protein, is the target of ORE1. At the terminal stage of leaf senescence, BFN1 localizes together with fragmented nuclei in membrane-coated vesicles. These findings imply that ORE1 participates in DNA fragmentation during senescence-associated PCD by regulating the expression of BFN1. Therefore, NAC TF, ORE1 is an important

regulator for the terminal stage of leaf senescence. In this study, among all the identified TFs associated with leaf senescence, the number of identified senescence-associated NAC TFs increased gradually as leaf senescence proceeded. The vast majority of NAC TFs (88%) were associated with the terminal stage of leaf senescence. These findings suggest that there may be a relationship between NAC TFs functioning to regulate the terminal senescence process of tobacco leaf.

Studies on genetics have demonstrated that TFs play significant roles in plant development, including leaf senescence. However, the specific regulation has largely been unknown. In this study, by analyzing tobacco leaf gene expression profiles using mRNA co-expression network methods, TFs associated with different processes of tobacco leaf senescence were identified. Through putative TF target prediction, we also discovered etensive interactions between senescence-associated TFs and their targets and realized that hundreds of tobacco genes are subject to their regulatory effects. This study contributes to the understanding of potential regulatory mechanisms of TFs involved in the senescence process in tobacco leaf. Future studies are needed to functionally validate the involvement of TFs during tobacco leaf senescence.

Conclusions

Based on co-expression network-based analysis, a time-series gene expression profile in senescing tobacco leaf was explored, revealing that a total of 21, 35, and 183 TFs were associated with early, middle, and late stages of tobacco leaf senescence, respectively. Also, we predicted the putative targets of these senescence-associated transcription factors and constructed the corresponding regulatory network, which benefits its further functional characterization and identification of downstream pathways.

Statement

All the methods were carried out in accordance to relevant local/national/international guidelines and regulations. The permissions using tobacco materials in scientific research has been obtained in presented study.

Data availability

All data generated and/or analysed during the current study are included in this published article or in the accompanying Supplementary Information file.

Received: 20 May 2023; Accepted: 2 September 2024
Published online: 16 September 2024

References

1. Buet, A., Costa, M. L., Martínez, D. E. & Guiamet, J. J. Chloroplast protein degradation in senescing leaves: Proteases and lytic compartments. *Front. Plant Sci.* **10**, 451973 (2019).
2. Tamary, E. *et al.* Chlorophyll catabolism precedes changes in chloroplast structure and proteome during leaf senescence. *Plant Direct* **3**, e00127 (2019).
3. Lim, P. O., Kim, H. J. & Gil Nam, H. Leaf senescence. *Annu. Rev. Plant Biol.* **58**, 115–136 (2007).
4. Gregersen, P. L., Culetic, A., Boschian, L. & Krupinska, K. Plant senescence and crop productivity. *Plant Mol. Biol.* **82**, 603–622 (2013).
5. Guo, Y. & Gan, S.-S. Translational researches on leaf senescence for enhancing plant productivity and quality. *J. Exp. Bot.* **65**, 3901–3913 (2014).
6. Jibran, R., Hunter, A. D. & Dijkwel, P. P. Hormonal regulation of leaf senescence through integration of developmental and stress signals. *Plant Mol. Biol.* **82**, 547–561 (2013).
7. Li, W. & Guo, Y. Transcriptome, transcription factors and transcriptional regulation of leaf senescence. *JBCG* **1**, 1 (2014).
8. Woo, H. R., Kim, H. J., Lim, P. O. & Nam, H. G. Leaf senescence: Systems and dynamics aspects. *Annu. Rev. Plant Biol.* **70**, 347–376 (2019).
9. Andersson, A. *et al.* A transcriptional timetable of autumn senescence. *Genome Biol.* **5**, 1–13 (2004).
10. Buchanan-Wollaston, V. *et al.* Comparative transcriptome analysis reveals significant differences in gene expression and signalling pathways between developmental and dark/starvation-induced senescence in Arabidopsis. *Plant J.* **42**, 567–585 (2005).
11. Gregersen, P. L. & Holm, P. B. Transcriptome analysis of senescence in the flag leaf of wheat (*Triticum aestivum* L.). *Plant Biotechnol. J.* **5**, 192–206 (2007).
12. Riester, L., Köster-Hofmann, S., Doll, J., Berendzen, K. W. & Zentgraf, U. Impact of alternatively polyadenylated isoforms of ETHYLENE RESPONSE FACTOR4 with activator and repressor function on senescence in *Arabidopsis thaliana* L.. *Genes* **10**, 91 (2019).
13. Guo, Y. & Gan, S. AtNAP, a NAC family transcription factor, has an important role in leaf senescence. *Plant J.* **46**, 601–612 (2006).
14. Besseau, S., Li, J. & Palva, E. T. WRKY54 and WRKY70 co-operate as negative regulators of leaf senescence in *Arabidopsis thaliana*. *J. Exp. Bot.* **63**, 2667–2679 (2012).
15. Li, W. *et al.* Intergrative metabolomic and transcriptomic analyses unveil nutrient remobilization events in leaf senescence of tobacco. *Sci. Rep.* **7**, 12126 (2017).
16. Langfelder, P. & Horvath, S. WGCNA: An R package for weighted correlation network analysis. *BMC Bioinform.* **9**, 1–13 (2008).
17. Eddy, S. R. Accelerated profile HMM searches. *PLoS Comput. Biol.* **7**, e1002195 (2011).
18. Punta, M. *et al.* The Pfam protein families database. *Nucleic Acids Res.* **40**, D290–D301 (2012).
19. Jin, J., Zhang, H., Kong, L., Gao, G. & Luo, J. PlantTFDB 3.0: A portal for the functional and evolutionary study of plant transcription factors. *Nucleic Acids Res.* **42**, D1182–D1187 (2014).
20. Ye, J. *et al.* WEGO 2.0: A web tool for analyzing and plotting GO annotations, 2018 update. *Nucleic Acids Res.* **46**, W71–W75 (2018).
21. Chen, C. *et al.* TBtools: An integrative toolkit developed for interactive analyses of big biological data. *Mol. Plant* **13**, 1194–1202 (2020).
22. Shannon, P. *et al.* Cytoscape: A software environment for integrated models. *Genome Res.* **13**, 426 (1971).
23. Kanehisa, M. Toward understanding the origin and evolution of cellular organisms. *Protein Sci.* **28**, 1947–1951. https://doi.org/10.1002/pro.3715 (2019).
24. Kanehisa, M. & Goto, S. KEGG: Kyoto encyclopedia of genes and genomes. *Nucleic Acids Res.* **28**, 27–30. https://doi.org/10.1093/nar/28.1.27 (2000).

25. Kanehisa, M., Furumichi, M., Sato, Y., Kawashima, M. & Ishiguro-Watanabe, M. KEGG for taxonomy-based analysis of pathways and genomes. *Nucleic Acids Res.* **51**, D587-d592. https://doi.org/10.1093/nar/gkac963 (2023).
26. Balazadeh, S., Riaño-Pachón, D. & Mueller-Roeber, B. Transcription factors regulating leaf senescence in *Arabidopsis thaliana*. *Plant Biol.* **10**, 63–75 (2008).
27. Jan, S., Abbas, N., Ashraf, M. & Ahmad, P. Roles of potential plant hormones and transcription factors in controlling leaf senescence and drought tolerance. *Protoplasma* **256**, 313–329 (2019).
28. Kim, H. J., Nam, H. G. & Lim, P. O. Regulatory network of NAC transcription factors in leaf senescence. *Curr. Opinion Plant Biol.* **33**, 48–56 (2016).
29. Koyama, T. *et al.* A regulatory cascade involving class II ETHYLENE RESPONSE FACTOR transcriptional repressors operates in the progression of leaf senescence. *Plant Physiol.* **162**, 991–1005 (2013).
30. Fukao, T., Yeung, E. & Bailey-Serres, J. The submergence tolerance gene SUB1A delays leaf senescence under prolonged darkness through hormonal regulation in rice. *Plant Physiol.* **160**, 1795–1807 (2012).
31. Miao, Y., Laun, T., Zimmermann, P. & Zentgraf, U. Targets of the WRKY53 transcription factor and its role during leaf senescence in Arabidopsis. *Plant Mol. Biol.* **55**, 853–867 (2004).
32. Zentgraf, U., Laun, T. & Miao, Y. The complex regulation of WRKY53 during leaf senescence of *Arabidopsis thaliana*. *Eur. J. Cell Biol.* **89**, 133–137 (2010).
33. Smeekens, S. & Hellmann, H. A. Sugar sensing and signaling in plants. *Front. Plant Sci.* **5**, 113. https://doi.org/10.3389/fpls.2014.00113 (2014).
34. Stessman, D., Miller, A., Spalding, M. & Rodermel, S. Regulation of photosynthesis during Arabidopsis leaf development in continuous light. *Photosynth. Res.* **72**, 27–37 (2002).
35. Huang, P., Li, Z. & Guo, H. New advances in the regulation of leaf senescence by classical and peptide hormones. *Front. Plant Sci.* **13**, 923136 (2022).
36. Wilson, K. A., McManus, M. T., Gordon, M. E. & Jordan, T. W. The proteomics of senescence in leaves of white clover, *Trifolium repens* (L.). *Proteomics* **2**, 1114–1122 (2002).
37. Jiang, Y., Liang, G., Yang, S. & Yu, D. Arabidopsis WRKY57 functions as a node of convergence for jasmonic acid–and auxin-mediated signaling in jasmonic acid–induced leaf senescence. *Plant Cell* **26**, 230–245 (2014).
38. Simeonova, E., Sikora, A., Charzyńska, M. & Mostowska, A. Aspects of programmed cell death during leaf senescence of mono- and dicotyledonous plants. *Protoplasma* **214**, 93–101 (2000).
39. Breeze, E. *et al.* High-resolution temporal profiling of transcripts during Arabidopsis leaf senescence reveals a distinct chronology of processes and regulation. *Plant Cell* **23**, 873–894 (2011).
40. Kim, J. H. *et al.* Trifurcate feed-forward regulation of age-dependent cell death involving miR164 in Arabidopsis. *Science* **323**, 1053–1057 (2009).
41. Matallana-Ramirez, L. P. *et al.* NAC transcription factor ORE1 and senescence-induced BIFUNCTIONAL NUCLEASE1 (BFN1) constitute a regulatory cascade in Arabidopsis. *Mol. Plant* **6**, 1438–1452 (2013).

Acknowledgements

We thank the current Guo lab members at the Tobacco Research Institute, Chinese Academy of Agricultural Sciences (CAAS) provide valuable comments on the revision of the manuscript.

Author contributions

Z. Z. and L. Y. conducted the research and drafted the manuscript. M. S. contributed to proofs and revisions in the final manuscript. W. L., S. C. and J. W assisted to the data analysis. Y. G. conceived the research, designed the experiments, and drafted the manuscript. All authors read and approved the final manuscript.

Funding

Natural Science Foundation of Shandong Province (ZR2021QC096), Fundamental Research Funds for Chinese Academy of Agricultural Sciences (2013ZL024), Agricultural Science and Technology Innovation Program (ASTIP-TRIC02).

Competing interests

The authors declare no competing interests.

Additional information

Supplementary Information The online version contains supplementary material available at https://doi.org/10.1038/s41598-024-71941-y.

Correspondence and requests for materials should be addressed to W.L. or Y.G.