

# Transcriptome Analysis Reveals Key Genes Involved in the Regulation of Nicotine Biosynthesis at Early Time Points after Topping in Tobacco (*Nicotiana tabacum* L.)

Guiling Sun (✉ [sung@vip.henu.edu.cn](mailto:sung@vip.henu.edu.cn))

Henan University <https://orcid.org/0000-0001-7580-8301>

Yan Qin

Henan University

Shenglong Bai

Henan University

Wenzheng Li

Yunnan Academy of Tobacco Science

Ting Sun

Henan University

David W. Galbraith

The University of Arizona

Zefeng Yang

Yangzhou University Agricultural College

Yun Zhou

Henan University

Bingwu Wang

Yunnan Academy of Agricultural Sciences

---

## Research article

**Keywords:** Tobacco; topping; transcriptome analysis; differentially expressed genes; nicotine biosynthesis and regulation

**Posted Date:** October 16th, 2019

**DOI:** <https://doi.org/10.21203/rs.2.16109/v1>

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

[Read Full License](#)

**Version of Record:** A version of this preprint was published on January 20th, 2020. See the published version at <https://doi.org/10.1186/s12870-020-2241-9>.

# Abstract

**Background** *Nicotiana tabacum* is an important economic crop. Topping, a common agricultural practice employed with flue-cured tobacco, is designed to increase leaf nicotine contents. Many genes are found to be differentially expressed in response to topping, particularly genes involved in nicotine biosynthesis, but comprehensive analyses of early transcriptional responses induced by topping are not yet available. To develop a detailed understanding of the mechanisms regulating nicotine biosynthesis after topping, we have sequenced the transcriptomes of *N. tabacum* roots at seven time points following topping. **Results** Differential expression analysis revealed that 4,830 genes respond to topping across all time points. Amongst these, nine gene families involved in nicotine biosynthesis and two gene families involved in nicotine transport showed significant changes during the immediate 24 hour period following topping. No obvious preference to the parental species was detected in the differentially expressed genes (DEGs). Significant changes in transcript levels of nine genes involved in nicotine biosynthesis and phytohormone signal transduction were validated by qRT-PCR assays. 549 genes encoding transcription factors (TFs), found to exhibit significant changes in gene expression after topping, formed 15 clusters based on similarities of their transcript level time-course profiles. 336 DEGs involved in phytohormone signal transduction, including genes functionally related to the phytohormones jasmonic acid, abscisic acid, auxin, ethylene, and gibberellin, were identified at the earliest time point after topping. **Conclusions** Our research provides the first detailed analysis of the early transcriptional responses to topping in *N. tabacum*, and identifies excellent candidates for further detailed studies concerning the regulation of nicotine biosynthesis in tobacco roots.

## Background

For tobacco (*Nicotiana tabacum* L.) plants, topping (defined as the removal of the flowering head and young leaves) is an essential cultivation practice. Topping switches the plant from a seed reproductive to a leaf vegetative phase, and this significantly increases leaf nicotine contents [1]. A number of studies, employing a variety of experimental techniques, have investigated tobacco responses to topping [2-4]. An up-regulation of nicotine biosynthesis, found to occur exclusively in roots, and particularly in growing root tips, is one of the typical responses of tobacco plants to topping [5].

Nicotine plays pivotal roles both in establishing the commercial quality of tobacco, and in defending plants against herbivores. Biosynthesis of nicotine, a secondary metabolite associated with the tobacco stress response, is reproducibly promoted by topping in tobacco roots [6]. Nicotine comprises two main nitrogen-containing rings, the pyrrolidine ring and the pyridine ring [7]. Biosynthesis of the pyrrolidine ring involves arginine decarboxylase (ADC) [8], ornithine decarboxylase (ODC) [9], S-adenosylmethionine decarboxylase (SAMDC), S-adenosyl-L-methionine synthetase (SAMS), putrescine N-methyltransferase (PMT) [10, 11], and N-methylputrescine oxidase (MPO) [12, 13]. Biosynthesis of the pyridine ring starts with the nicotinic acid dinucleotide (NAD) biosynthetic pathway. Enzymes participating in the early metabolic conversion steps of this pathway include aspartate oxidase (AO), quinolinate synthase (QS), and quinolinic acid phosphoribosyl transferase (QPT) [14-16]. The A622 gene (encoding an isoflavone

reductase-like protein) is responsible for nicotine ring coupling, and the BBL genes (encoding berberine bridge enzyme-like proteins) are involved in the subsequent oxidation step that leads to nicotine [17, 18].

A recent report [19] employed suppression subtractive hybridization (SSH) techniques to further examine the transcriptional responses of tobacco roots during the first 24 hours after topping. Of the 129 high quality expressed sequence tags identified as representing DEGs, most were involved in stress/defense, in secondary metabolism, and in signaling/transcription [19]. The regulation of nicotine biosynthesis has long been considered a complex physiological response, and many TFs are directly or indirectly involved in its regulation [20, 21]. Further insights into the transcriptional regulation of the nicotine biosynthetic pathway have come from the analysis of two subtractive cDNA libraries of jasmonate-treated *Nicotiana benthamiana* roots, and through examination of the effects of virus-induced gene silencing (VIGS) technologies. Of the sixty-nine TFs, six (from three TF families) affect nicotine metabolism, with *NbbHLH1* and *NbbHLH2* (basic helix-loop-helix) genes positively regulating the jasmonate activation of nicotine biosynthesis, as evidenced by overexpression [22].

Although specific genes regulating nicotine synthesis after tobacco topping have been identified, a detailed description of the transcriptional regulatory network that responds to topping is not available. The situation is further complicated by the allotetraploid status of *N. tabacum*, formed through the hybridization of *N. sylvestris* (S-subgenome) and *N. tomentosiformis* (T-subgenome), and how these two subgenomes respond to topping is unclear. In this study, we have sequenced tobacco root transcriptomes at seven different time points (0, 0.5, 1, 3, 5, 8 and 24 h) after topping. These time points were chosen to identify candidate genes associated with the regulation of nicotine biosynthesis at the earliest stages, as well as to allow discovery of upstream regulators of nicotine synthesis through clustering of the time-course profiles of TF gene expression, and to compare the responses of the two subgenomes to topping. This comprehensive approach to characterization of the transcriptional responses of tobacco, especially focusing on the early regulation of nicotine biosynthesis, should serve to advance genetic improvement in this crop.

## Results

### Transcriptome sequencing and quality assessment

Total RNA of tobacco roots, isolated separately from 18 individual plants, was employed for RNA sequencing (RNA-Seq) library construction. The 18 RNA-Seq libraries were sequenced using the Illumina platform. After filtering out low quality sequences (quality scores <25), 105 Gb of cleaned data was obtained, representing approximately 6 Gb per sample. The cleaned sequence GC content varied from 42.1% to 42.7% (Table S1). The mapping rates for the cleaned sample reads aligned against the reference genome sequence ranged from 91.6% to 97.8% (Table S1). The sequencing quality and gene expression levels were generally consistent across the sequenced samples (Figure S1).

# Identification and verification of differentially expressed genes (DEGs)

The expression levels of the genes from the tobacco transcriptomes were calculated and normalized to FPKM values (Fragments Per Kilobase of transcript per Million fragments mapped). The values of Pearson's Correlation Coefficient across biology replicates exceeded 0.82. In terms of the correlation between samples from different time points, some samples displayed higher values with those from other time points. For example, BWR3-2A showed a correlation coefficient of 0.96 as compared to BWR24-1A, and 0.95 with BWR5-2A (Figure S2). Further experiments will be required to elucidate this unexpected observation. Through comparison of the samples at each time point to the t=0 sample, and using a fold-change (FC) > 2, and a false discovery rate (FDR) < 0.05 as selection criteria, 4,830 DEGs were identified after topping. An almost identical number (2,082 and 2,075 genes) came from the *N. tomentosiformis* and *N. sylvestris* genomes, respectively (Table S2). Notably, the number of DEGs at 0.5 h (2,562) was much greater than those at any other time point, indicating that more genes respond to topping at earlier times. The number of DEGs dropped to its lowest level (815) at 1 h after topping (Figure 1a). However, a second burst of differential gene expression was observed at t=8 h (1,756), followed by a decrease at t=24 h (Figure 1a). The results imply that the *N. tabacum* root produces two discrete peaks of transcriptional activity, at 0.5 h and 8 h after topping. This result is consistent with the gene numbers identified as specifically induced at each of the six time points after topping, the largest number being 1,186 at t=0.5 h and the second highest number being 585 at t=8 h after topping (Figure 1b).

To validate the transcriptional results obtained by RNA-Seq, we selected nine genes related to nicotine biosynthesis and phytohormone signal transduction, and examined their transcriptional responses by qRT-PCR. The expression trends of these genes analyzed by qRT-PCR were consistent with the RNA-Seq analysis performed at the corresponding time points (Figure 2). The changes of selected DEGs obtained by RNA-Seq analysis had good correlations with those obtained by qRT-PCR ( $R^2 = 0.674$ ). These results confirm that the alterations in gene expression detected by RNA-Seq accurately reflect transcript differences at the different time points after topping.

## Functional classification and enrichment analysis of DEGs

4,830 DEGs showing significant variation at the different time points after topping were selected for further analysis. Based on their relative expression levels, the DEGs were divided into different categories using hierarchical clustering, being distinguishable in terms of the temporal patterns of the transcriptional responses of the roots at the various time points after topping (Figure S3). The predicted functions of the DEGs were then obtained from their GO (Gene Ontology) annotations, and using KEGG (Kyoto Encyclopedia of Genes and Genomes) pathway analysis. According to GO annotation, the DEGs were distributed across 42 functional terms, as follows: 19 terms for biological process, 12 terms for molecular functions, and 11 terms for cellular component (Figure S4).

GO enrichment analyses were performed to classify the putative functions of DEGs in the comparisons of libraries prepared from the different time points. The DEGs in GO enriched categories of biological process were mainly involved in response to oxidative stress (GO:0006979), the phenylpropanoid metabolic process (GO:0009698), the lignin metabolic process (GO:0009808), and response to abiotic stimulus (GO:0009628). The DEGs of GO enriched categories of cellular component were mainly involved in the apoplast (GO:0048046), the extracellular region (GO:0005576), the external encapsulating structure (GO:0030312), and the cell wall (GO:0005618). The DEGs of GO enriched categories of molecular function were mainly associated with peroxidase activity (GO:0004601), antioxidant activity (GO:0016209), and a series of transporter activities (GO: GO:0006857, GO:0008272, GO:0008509, and GO:0008271) (Figure 3).

To further investigate the functions of differentially expressed transcripts in response to topping, we performed enrichment analyses by mapping the sequences to the KEGG database categories. The DEGs with KEGG annotation were assigned to 28 classes, mainly related to signal transduction (221), carbohydrate metabolism (212), biosynthesis of other secondary metabolites (166), and metabolism of terpenoids and polyketides (69) (Figure S5). KEGG enrichment analyses also indicated the DEGs were significantly enriched in the main pathways of phenylpropanoid biosynthesis (ko00940), of starch and sucrose metabolism (ko00500), and in the plant MAPK signaling pathway (ko04016) (Figure 4).

## DEGs involved in nicotine synthesis and transport

We further investigated whether the genes activated by topping were involved in nicotine biosynthesis and transport. As expected, nine gene families involved in nicotine biosynthesis (*AO*, *QS*, *ODC*, *ADC*, *SAMS*, *PMT*, *A622*, *MPO*, and *BBL*) (Figure 5), and two gene families involved in nicotine transport (*MATE*, *NUP*), as identified by showing at least 93% identities with the primary sequences of previously reported enzymes, were found in the DEG dataset (Table S3). All genes showed transcriptional up-regulation, with most being up-regulated at 8h and 24h after topping; our qPCR assay also verified the expression changes of four genes (*PMT1a*, *PMT1b*, *MPO*, *ODC*) at the corresponding time points (Figure 2). Both gene families encoding *MATE* and *NUP* in nicotine transport were found to be up-regulated (Table S3). Similar to the situation across all DEGs, most of those involved in nicotine synthesis and transport were found in both subgenomes. One DEG encoding *AO* was derived from the T-subgenome and all the DEGs encoding *MPO* and *ODC* were from the S-subgenome.

## Transcription factors (TFs) of DEGs, and gene clustering by expression patterns

To investigate the upstream regulatory mechanisms of nicotine biosynthesis after topping, we next focused on the types of TFs represented in the DEGs from the tobacco root transcriptome. In our study, a total of 549 DEGs encoding TFs were identified (Table S4), being divided into 49 TF families. Amongst

these, the number of TFs was highest at t=0.5 h (355), accounting for 65% of all TFs, with 240 being up-regulated and 115 down-regulated. This was followed by the t=8 h timepoint (253), accounting for 46% of all TFs, with 94 being up-regulated and 159 down-regulated. This suggests that many TFs genes participate in instant early gene activation. Notably, and representing most of these TFs, 18 families were found to contain more than 10 gene members: AP2-EREBP (75), MYB (69), bHLH (44), NAC (30), bZIP (30), Orphans (26), HB (24), WRKY (23), HSF (19), C2C2-Dof (17), GRAS (16), LOB (15), MADS (12), GNAT (11), AUX / IAA (11), G2-like (11), C3H (10) and C2H2 (10) (Figure 6).

To further examine the contributions of specific TFs to the regulatory network of nicotine biosynthesis, we performed clustering using the 549 TFs and the up-regulated structural genes associated with the nicotine biosynthesis pathway. Fifteen clusters showing similar expression profiles were obtained (Figure 7 and Table S5). It can be observed that several clusters are similar but with minor differences. For instance, the TFs of Clusters 2 and 12 were up-regulated at t=0.5 h, and the TFs in Cluster 9 and 10 were up-regulated at 0.5-1 h. They then returned to the expression levels found before topping (Figure 7). Notably, most of the up-regulated DEGs in nicotine biosynthesis were in Cluster 11 (20 DEGs), which displayed the greatest up-regulation at t=8h and at t=24 h after topping (Figure 7 and Table S5). Seventeen genes from the bHLH family and the AP2-EREBP family were found in Cluster 11, including ERF189 (Nitab4.5\_0003090g0030 and Nitab4.5\_0015055g0010), and ERF91 (Nitab4.5\_0004620g0030) (Table S5).

## DEGs connected with phytohormone signal transduction

As phytohormones are known to quickly respond to tobacco topping and to also influence nicotine biosynthesis in tobacco roots, we examined the role of phytohormone signal transduction in the transcriptional responses induced by topping. We identified 336 DEGs, including those related to the biosynthesis, metabolism and action of auxin (IAA), abscisic acid (ABA), ethylene, gibberellin (GA), and jasmonic acid (JA) (Table S6). The 53 DEGs involved in IAA signal transduction included the ARF (auxin response factor) family (4), the AUX/IAA (auxin responsive protein) family (11), the AUX1 (amino acid transporter protein) family (22), the GH3 (GH3 auxin-responsive promoter) family (7), and the SAUR (auxin responsive SAUR protein) family (9). Most of the DEGs associated with the IAA signaling pathway showed significant up-regulated expression changes, 21 of 36 genes being up-regulated at t=0.5 h, and 10 of 18 genes being up-regulated at t=8 h. For the ABA signal transduction pathway, six gene families were identified, including the PYL/PYR (abscisic acid receptor) family (5), the SAPK (Serine threonine protein kinase) family (2), the PP2C (protein phosphatase 2C) family (28), the CIPK (CBL-interacting protein kinase) family (19), the CDPK (Calcium-dependent protein kinase) family (9), and the Calmodulin (Calmodulin-like protein) family (8). 43 expression changes were detected at t=0.5 h, and 19 at t=8 h. 83 DEGs were implicated in ethylene signaling, including the AP2-EREBP (ethylene responsive transcription factor) family (75), and the ETR (ethylene receptor) family (8), with most DEGs identified at t=0.5 h (55). The GA and the JA signaling pathways (four and three gene families, respectively) also showed significant transcriptional changes after topping.

## Quantification of phytohormones and nicotine

Phytohormones play a vital role in regulating plant defense and development. To gain insights into the mechanisms whereby phytohormones affect the responses of tobacco to topping, we measured the levels of IAA, JA, JA-Ile, and ABA in the root samples at the various timepoints after topping. Both JA and auxin signaling pathways were induced by topping at t=3 h. JA levels at t=3 h were significantly increased by almost 34% ( $P=0.035$ , paired  $t$  test), and reduced by 23.5% and 18.9% at t=8h and t=24h (Figure 8). The dynamics of JA-Ile levels elicited by topping closely followed those of JA, the levels of JA-Ile significantly increasing to about 3-fold at t=24 h ( $P=0.014$ , paired  $t$  test) as compared with untreated plants. The levels of IAA significantly increased at t=3 h ( $P=0.024$ , paired  $t$  test), whilst declining to initial levels at t=24h (Figure 8). The levels of ABA gradually increased to 2.3-fold at t=8 h ( $P=0.0003$ , paired  $t$  test), and to 1.6-fold at t=24 h ( $P=0.009$ , paired  $t$  test), as compared to the untreated plants. We also measured nicotine levels after topping. Our analyses indicated the nicotine levels significantly increased to 1.5-fold at t=24 h ( $P=0.01$ , paired  $t$  test) after topping (Figure 8).

## Discussion

### Transcriptome sequencing and DEGs responses to topping

Nicotine is a characteristic secondary product of tobacco. In most *Nicotiana* species, it is synthesized in the roots, being then transported to the leaves where it accumulates [23]. Nicotine synthesis and accumulation is controlled not only by various environmental cues but also by managerial practices including topping [2, 6]. The factors controlling the topping-induced increase in alkaloid biosynthesis are not well understood, but involve a complex physiological response in the plant as a result of altered phytohormone induced signaling [24]. In order to better understand the mechanism of the tobacco response to topping, we have sequenced and analyzed the transcriptomes of *N. tabacum* roots at early time points after topping. We identified a total of 4,830 topping-responsive DEGs, with representatives being distributed across a number of different molecular functional categories, including secondary metabolism, plant hormone signaling transduction, stress defense, and other metabolism.

### DEGs involved in nicotine biosynthesis and subgenome transcriptional preference

We detected 1.5-fold changes in nicotine levels at t=24 h after topping with each biological replicates containing 4 individual plants, which is consistent with previous report [25]. It was worth to mention that no significant changes of nicotine content were detected at this time point with each biological replicates containing a single plant (data not shown), indicating that the individual plants response differently in the short time of decapitation. We then inspected more closely the transcriptional changes of genes known to be involved in nicotine biosynthesis and transport. [Nine gene families in the nicotine biosynthesis pathway were identified within the DEGs. Their functions included pyridine ring synthesis \(AO and QS\),](#)

pyrrolidine ring synthesis (*ODC*, *ADC*, *PMT*, *SAMS*, *MPO*), and the coupling of the two nicotine rings (*A622* and *BBL*). Ornithine decarboxylase (*ODC*) catalyzes the first and rate-limiting step of polyamine biosynthesis that converts ornithine into putrescine. Down-regulation of *ODC* transcript levels using RNAi led to lower leaf levels of nicotine in *N. tabacum* [26, 27]. Correspondingly, in our study, one *ODC* gene was significantly up-regulated at t=8 h and t=12 h. A second example involves *PMT*, which converts putrescine into N-methylputrescine [28]. In that previous study, five *PMT* genes were investigated (*NtPMT1a*, *NtPMT1b*, *NtPMT2*, *NtPMT3* and *NtPMT4*). Transcripts derived from *NtPMT2* and *NtPMT1b* showed the greatest increase in abundance (about 3-fold) during the first 24 h after topping [29]. In line with these findings, the expression levels of all the five *PMT* genes in our study were significantly up-regulated at different time points after topping (Table S2). A further example is provided by *SAMS*, which indirectly contributes to nicotine biosynthesis by supplying the S-adenosylmethionine cofactor for the *PMT* reaction [30]. In our study, five *SAMS* genes were significantly up-regulated at early time points after topping (Table S3).

The enzyme *QPT* plays a critical role in the synthesis of the pyridine moiety of nicotine in *Nicotiana*, in addition to its ubiquitous role in NAD(P)(H) synthesis [31]. The tobacco genome contains two duplicated *QPT* genes (designated *QPT1* and *QPT2*). *QPT1* is expressed at a constitutive basal level in all plant tissues, with somewhat higher levels of expression within the apical meristem. In comparison, *QPT2* is expressed exclusively in the tobacco root and is regulated coordinately with other structural genes for nicotine biosynthesis [31, 32]. Although *QPT1* and *QPT2* were not present in our list of DEGs, the gene of *QPT2* (Nitab4.5\_0000742g0010) showed significant up-regulation at 3 h and 24 h in our qPCR assay (data not shown).

Finally, we consider *A622*, which is expressed in the root, and may be involved in the final condensation reaction of nicotine biosynthesis [33]. The capacity of *N. glauca* to produce anabasine was markedly reduced when an RNAi approach was used to down-regulate gene expression, thereby decreasing *A622* protein levels. This resulted in plants having almost undetectable levels of pyridine alkaloids in their leaves, even after suffering damage to apical tissues [18]. In our work, the expression of *A622* was significantly up-regulated at t=8 h and t=24 h after topping (Table S2), implying that *A622* positively regulates the biosynthesis of nicotine in the final coupling ring step. Consistent with the results for *A622*, *BBL* genes were also found to be induced after topping, as previously reported [17].

*N. tabacum* is hypothesized to be a consequence of hybridization of two parental genomes (*N. sylvestris* and *N. tomentosiformis*). *N. tomentosiformis* displayed much lower nicotine levels than *N. sylvestris* in both roots and leaves [34]. All DEGs, especially those involved in nicotine biosynthesis and transport, showed no obvious preference to any of the two subgenomes. More experimental work will be required to elucidate the molecular basis of heterosis and the dramatic domestication selection following the hybridization.

# Differential expressed transcription factors related to regulation of nicotine synthesis

The pattern of expression changes induced by topping for the structural genes of nicotine biosynthesis was initially derived from bioinformatics analyses of sequence data, but was confirmed by the qRT-PCR experiments. It was therefore reasonable to speculate that the sequence data could also be used to identify the TF(s) that actively regulate the production of nicotine at the early time points after topping. As reported previously, many TFs play important roles in regulating nicotine biosynthesis, including members of the AP2/EREBP, bHLH, ARF, and WRKY families [35, 36]. The AP2/EREBP family is the largest TF family in the tobacco genome [37], and ERF-type TFs of the group IX subfamily, including *ERF1*, *ERF189*, and *ERF32*, have been recently identified as direct regulators of the structural genes of nicotine biosynthesis [22, 31, 38]. The second largest class of TFs shown to induce alkaloid biosynthesis in *Nicotiana* is the MYC2-like bHLH family. *MYC2*, belonging to the bHLH-family of TFs, is a key component in conserved jasmonate signaling [39]. It positively regulates nicotine production either directly, through G-box-mediated binding and activation of nicotine structural genes, or indirectly, through the activation of the *ERF* genes [40, 41]. Wang et al. [42] found that overexpression of *NtMYC2a* led to great enhancement, under field testing, of nicotine levels in the transgenic lines. Although the mechanism by which *ARF1* regulates nicotine synthesis remains unclear, VIGS (virus-induced gene silencing) of *bHLH3* and *ARF1* results in a significant increase in nicotine content as compared to control plants [22]. In addition, *WRKY-R1*, the Group II member of the WRKY family, was specifically and highly expressed in tobacco roots. This suggests it regulates the expression of genes related to nicotine synthesis, like *PMT* [35].

Screening the DEGs in our study led to the identification of 549 DEGs annotated as TFs, including members of the *AP2/EREBP* (75), *bHLH* (44), *WRKY* (23), and *ARF* (4) families. (Figure 6 and Table S4). To elucidate patterns of co-regulation of TFs, we clustered all DEG TFs along with the structural genes involved in nicotine biosynthesis. We found that Cluster 11 contained 20 of 28 structure and transporter genes in the DEGs, as well as a total of 17 genes from the bHLH and AP2-EREBP families. We selected one of seven genes in the bHLH family (*Nitab4.5\_0000093g0110*) for study using RNAi, finding the resultant plant showed an altered nicotine level (unpublished data), and implying an important regulatory role in nicotine biosynthesis. Other newly identified *AP2/EREBP*, *bHLH*, *WRKY*, *ARF*, *MYB*, and *NAC* TF genes might also be involved in nicotine biosynthesis, since all of these TF families have been described as functioning in the regulation of plant secondary metabolism [43-45]. These are therefore good targets for further experiments.

## Involvement of DEGs responsive to topping in phytohormone signal transduction

Plant hormones play pivotal roles in regulating numerous aspects of plant growth and development, including secondary metabolism. Five major classes of plant hormones are known to be involved in the

regulation of nicotine biosynthesis, including JA, IAA, GA, ethylene, and ABA [30, 46-49]. It is known that JA treatment effectively induces nicotine biosynthesis in tobacco by regulating genes encoding enzymes of nicotine biosynthesis, including *QPT*, *ODC*, and *PMT* [32, 36, 50, 51]. Yang et al. [52] reported an analysis of the effects of JA on tobacco BY-2 cells, finding that transcript levels were increased for 12 ethylene response factors and 4 basic helix-loop-helix factors associated with alkaloid formation. JAs are known to operate in complex networks with crosstalk to other phytohormone signaling pathways in the regulation of tobacco nicotine biosynthesis. Examples include the JA ZIM-domain 1 (JAZ1) protein (a key repressor of JA signaling) which interacts *in vivo* with DELLA proteins (repressors of the GA pathway) [46], and *NtPYL4* (a functional ABA receptor) whose transcription is regulated by JAs [47].

In the present study, we propose that topping, as a form of mechanical wounding, activates JA signaling as one of the earliest events. JA-Ile, as the only JA derivative known to be involved in JA signaling, displayed continuously increasing levels over the first 24 h after topping, indicating a vital role in the regulation of nicotine biosynthesis. The expression levels of the two JAZ-encoding genes were significantly up-regulated at t=0.5 h and t=8 h, respectively, which is consistent with the observed increases in JA and JA-Ile levels. Our experimental data also indicated similar patterns of changes in concentrations of JA and IAA following topping. This is consistent with the report that JA and auxin have very similar signal transmission mechanisms. However, how JA and IAA collaboratively respond to topping stimulation and regulate nicotine biosynthesis remains unclear [35], and will require further investigation.

As a consequence of phytohormone changes, genes in the downstream of regulation network also showed corresponding variation. Notably, all the genes encoding AUX/IAA were found to be up-regulated at t=0.5 h, which implies that IAA levels increased at an early time point. Similarly, all ABA receptor PYL genes exhibited lower expression levels at t=0.5 h and t=8 h, suggesting that altered *PYL* expression affects the JA response to nicotine biosynthesis in tobacco. Finally, in our study, we found that a large number of regulatory genes associated with auxin signaling, ABA signaling, ethylene signaling, and gibberellin signaling showed significantly altered expression levels at t=0.5 h after topping, indicating the complex networks with crosstalk are promptly constructed across these phytohormonal signaling pathways as an early response to topping (Table S6).

## Methods

### RNA extraction, library preparation, and transcriptome sequencing

The seeds of *Nicotiana tabacum* Yunyan 87 were kindly provided by the Tobacco Seed Bank of the Yunnan Academy of Tobacco Agricultural Sciences. The plants were identified by Dr. Zhongbang Song. The voucher specimen is accessible at the Herbarium of the Kunming Institute of Botany, Chinese Academy of Sciences (accessions No. SGL-001-1). All the materials were grown for 8 weeks in commercial potting soil in a Percival PGC-10 incubator set for a 16 hour day/8 hour night cycle at 28°C.

Individual plants with the most similar morphology were selected for topping by removing the apices above the youngest unfolded leaf. Roots were washed with water, and tissues excised using a surgical blade, 18 samples being taken 0, 0.5, 1, 3, 5, 8 and 24 h after topping, with 2-3 replicates at each time point (Table S1). For RNA extraction, root tissues were frozen in liquid nitrogen and ground into a powder using a mortar and pestle. RNA was extracted using TRIzol (Invitrogen Life Technologies) according to the manufacturer's instructions. Extracted RNA was assessed for quality and quantity using an Agilent 2100 Bioanalyzer (Agilent Technologies). Dried RNA pellets were diluted in RNase-free distilled water and stored at -80°C until used. The integrity of total RNA was determined by running samples on 1% denaturing agarose gels. RNA concentrations and quality were initially assessed using a spectrophotometer (NanoDrop, Technologies Inc.) measuring absorbance ratios of A260/230 and A260/280 nm. A Bioanalyzer (Agilent) was used to perform a final assessment of RNA quality prior to sequencing.

Magnetic beads with Oligo (dT) are used to isolate mRNA from total RNA. Purified mRNA was then fragmented in fragmentation buffer. Using the short fragments as templates, random hexamer-primers were used to synthesize the first-strand cDNA. The second-strand cDNA was synthesized using buffer, dNTPs, RNase H and DNA polymerase I. Short double-stranded cDNA fragments (200 bp) were purified with a QIAquick PCR extraction kit and eluted with EB buffer for end repair and the addition of an "A" base. Next, the short fragments were ligated to Illumina sequencing adaptors. DNA fragments of a selected size were gel-purified and amplified by PCR. The amplified library was sequenced on an Illumina HiSeq™ 2000 sequencing machine in October 2013 at BGI (Shenzhen, China). Reads were generated in 90 bp paired-end format. The sequencing data was deposited in the National Center for Biotechnology Information (NCBI) Sequence Read Archive (SRA) database with accession number SRP154415.

## Mapping of sequencing reads and quantification of gene expression

Before further analysis, the raw reads were filtered to remove adaptor sequences, low-quality reads, and reads containing poly-N, using CutAdapt (<http://code.google.com/p/cutadapt/>) [53] and Btrim [54]. Next, we processed and mapped the clean reads onto the *Nicotiana tabacum* genome [55](Edwards, 2017 #656;Edwards, 2017 #656) using Hisat2 version 2.1.0 [56] with default parameters. The reference genome can be found on the Sol Genomics Network website ([ftp://ftp.solgenomics.net/genomes/Nicotiana\\_tabacum/edwards\\_et\\_al\\_2017/assembly/Nitab-v4.5\\_genome\\_Chr\\_Edwards2017.fasta.gz](ftp://ftp.solgenomics.net/genomes/Nicotiana_tabacum/edwards_et_al_2017/assembly/Nitab-v4.5_genome_Chr_Edwards2017.fasta.gz)). Gene expression levels were estimated as FPKM (fragments per kilobase of exon per million fragments mapped reads) values [57] using Cufflinks software version 1.2.1 [58].

## Identification and enrichment Analysis of DEGs

Differential expression of genes between the t=0 h sample and other samples at different time points were calculated with the cuffdiff program in Cufflinks [58]. Gene abundance differences among these samples were established on basis of the fold change (FC) of the FPKM values. The genes with an absolute value of  $\log_2$  fold changes (FC)  $\geq 1$  and an adjusted p-value (q-value)  $\leq 0.05$  were defined as differentially expressed genes (DEGs). The FPKM values of the DEGs were normalized and clustered by heatmap with parameters "scale=row, cluster\_rows=T, cutree\_rows=8", where cutree function in R was used to estimate category number and eight categories were obtained when h=4.8.

Gene Ontology (GO) and Kyoto Encyclopedia of Genes and Genomes (KEGG) analyses were performed to identify the enrichment of DEGs in GO terms and metabolic pathways, respectively. For functional categorization and pathway visualization of DEGs, the metabolic pathways of DEGs were predicted using the KEGG Automated Annotation Server (KAAS) with default parameters. The DEGs in the KEGG pathways and GO analysis were enriched using ClusterProfile R and WEGO [59] [60]. A corrected *p*-value (*q*-value)  $\leq 0.05$  was chosen as the threshold for significantly enriched KEGG pathways and GO terms.

## Clustering of differentially expressed TFs and the DEGs involved in nicotine biosynthesis by Expression Profiles

Clustering of differentially expressed TFs and the DEGs involved in nicotine biosynthesis was performed using R package mfuzz with the cluster number assigned to 15 using the Dmin function and others as defaults on the profiles of FPKM at the different time points after topping.

## Subgenome Assignment of All the DEGs

All DEGs were used as queries to search against all the annotated proteins in the genomes of *N. tomentosiformis* and *N. sylvestris*. Subgenomes were assigned based on the first hit species if the first hit had at least 90% identity and the second hit showed a lower identity value than the first hit. For the DEGs involved in nicotine biosynthesis and transport, a manual check was performed based on previous reports.

## Quantitative Real-Time PCR analysis (qRT-PCR)

To confirm the gene expression levels of the RNA-Seq assay, 1  $\mu$ g total RNA mixed equally from 3 plants for each RNA-Seq sample was used for the qRT-PCR assay. The RNA, following elimination of residual genomic DNA using DNase I, was reverse transcribed into cDNA using the PrimeScript RT reagent Kit with gDNA Eraser (Perfect Real Time) (Takara, Japan). qRT-PCR was performed in a 20  $\mu$ L reaction volume using the LC480 system (Roche, SUI) and with three biological replicates. Three technical replicates were used for the samples at different time points with the tobacco *Actin* gene as an internal standard for normalization, and SYBR Green as the fluorochrome. A three-step PCR process was performed with a pre-

denaturation at 95° C for 30 s, followed by 45 cycles of denaturation at 95° C for 30 s, annealing at the optimal temperature of each primer pair for 20 s, and elongation at 72° C for 20 s, and finally for melting point curve analysis (95° C for 15 s, 60° C for 1 min and 95° C for 15 s) to test amplicon specificity. Relative quantification of gene expression level was carried out using the  $2^{-\Delta\Delta C_t}$  method. The specific primer sequences of the selected genes for qRT-PCR validation were designed based on the divergent regions among the orthologous genes and listed in Table S7.

## Measurement of Nicotine Content and Hormone Levels

For measurement of nicotine content, roots were dried at 105° C for 30 minutes and at 60° C for 3d. Each sample (100 mg) equally mixed from 4 individual plants was mixed with 1 mL of extraction solution (40% methanol containing 0.5% acetic acid (v/v)), and nicotine was quantified by HPLC as previously described [61]. Three biological replicates each with four technical replicates were performed. For quantification of phytohormones, approximately 200 mg of tobacco roots were ground in liquid nitrogen and 1 mL of ethyl acetate spiked with the internal standards was added to each sample.  $^{13}\text{C}_2$ -JA,  $^{13}\text{C}_2$ -JA-Ile,  $\text{D}_4$ -SA,  $\text{D}_6$ -ABA and  $\text{D}_5$ -IAA were used as the internal standards for JA, JA-Ile, SA, ABA and IAA, respectively. Phytohormone extraction and quantification were described as previously by HPLC-MS/MS (LCMS-8040, Shimadzu) system [62]. Three biological replicates were used for each sample.

## Abbreviations

DEGs, differentially-expressed genes; ADC, arginine decarboxylase; ODC, ornithine decarboxylase; SAMDC, S-adenosylmethionine decarboxylase; SAMS, S-adenosyl-L-methionine synthetase; PMT, putrescine N-methyltransferase; MPO, N-methylputrescine oxidase, NAD, nicotinic acid dinucleotide, AO, aspartate oxidase; QS, quinolinate synthase; QPT, quinolinic acid phosphoribosyl transferase; A622, isoflavone reductase-like protein; BBL, berberine bridge enzyme-like protein; SSH, suppression subtractive hybridization; VIGS, virus induced gene silencing; FPKM, Fragments Per Kilobase of transcript per Million fragments mapped; GO, Gene Ontology; KEGG, Kyoto Encyclopedia of Genes and Genomes; TF, transcription factors; IAA, auxin; ABA, abscisic acid; JA, jasmonic acid; ARF, auxin response factor; PP2C, protein phosphatase 2C; CIPK, CBL-interacting protein kinase; CDPK, Calcium-dependent protein kinase; AP2-EREBP, ethylene responsive transcription factor; ETR, ethylene receptor; JAZ1, JA ZIM-domain 1 containing protein

## Declarations

# Ethics approval and consent to participate

Not Applicable.

## Consent for publication

Not Applicable.

## Availability of data and material

The sequencing data were then deposited in the National Center for Biotechnology Information (NCBI) Sequence Read Archive (SRA) database with the accession number SRP154415.

## Competing interests

The authors declare that they have no competing interests.

## Funding

This work is supported by a grant from the National Natural Science Foundation of China (31771414) and the Program for Science and Technology Innovation Talents in Universities of Henan Province (18HASTIT041) and a grant from Yunnan Tobacco Company (2018530000241001). The funders made no significant role in study design, data collection and analysis, decision to publish, or writing of the manuscript.

## Authors' contributions

GS and BW conceived and designed the research. YQ, SB, YZ, TS, ZY, and WL generated the data and performed the analysis. WL and BW contributed reagents and materials. QY, GS, DG, and BW wrote the manuscript. All authors read and approved the final manuscript.

## Acknowledgements

We thank Dr. Zhongbang Song for his constructive suggestions and helpful discussion.

## References

1. Baldwin IT: **An ecologically motivated analysis of plant-herbivore interactions in native tobacco.** *Plant Physiol* 2001, **127**(4):1449-1458.
2. Baldwin IT, Karb MJ, Ohnmeiss TE: **Allocation of N-15 from Nitrate to Nicotine - Production and Turnover of a Damage-Induced Mobile Defense.** *Ecology* 1994, **75**(6):1703-1713.

3. Baldwin IT, Schmelz EA, Ohnmeiss TE: **Wound-Induced Changes in Root and Shoot Jasmonic Acid Pools Correlate with Induced Nicotine Synthesis in *Nicotiana-Sylvestris* Spegazzini and Comes.** *J Chem Ecol* 1994, **20**(8):2139-2157.
4. Kutchan TM: **Alkaloid Biosynthesis - the Basis for Metabolic Engineering of Medicinal-Plants.** *Plant Cell* 1995, **7**(7):1059-1070.
5. Baldwin IT: **Inducible nicotine production in native *Nicotiana* as an example of adaptive phenotypic plasticity.** *J Chem Ecol* 1999, **25**(1):3-30.
6. Wang SS, Shi QM, Li WQ, Niu JF, Li CJ, Zhang FS: **Nicotine concentration in leaves of flue-cured tobacco plants as affected by removal of the shoot apex and lateral buds.** *J Integr Plant Biol* 2008, **50**(8):958-964.
7. Hakkinen ST, Tilleman S, Swiatek A, De Sutter V, Rischer H, Vanhoutte I, Van Onckelen H, Hilson P, Inze D, Oksman-Caldentey K-M *et al.*: **Functional characterisation of genes involved in pyridine alkaloid biosynthesis in tobacco.** *Phytochemistry* 2007, **68**(22-24):2773-2785.
8. Tiburcio AF, Galston AW: **Arginine decarboxylase as the source of putrescine for tobacco alkaloids.** *Phytochemistry* 1986, **25**(1):107-110.
9. Imanishi S, Hashizume K, Nakakita M, Kojima H, Matsubayashi Y, Hashimoto T, Sakagami Y, Yamada Y, Nakamura K: **Differential induction by methyl jasmonate of genes encoding ornithine decarboxylase and other enzymes involved in nicotine biosynthesis in tobacco cell cultures.** *Plant molecular biology* 1998, **38**(6):1101-1111.
10. Saunders JW, Bush LP: **Nicotine Biosynthetic Enzyme Activities in *Nicotiana tabacum* L. Genotypes with Different Alkaloid Levels.** *Plant Physiol* 1979, **64**(2):236-240.
11. Hibi N, Higashiguchi S, Hashimoto T, Yamada Y: **Gene expression in tobacco low-nicotine mutants.** *The Plant cell* 1994, **6**(5):723-735.
12. Katoh A, Shoji T, Hashimoto T: **Molecular cloning of N-methylputrescine oxidase from tobacco.** *Plant Cell Physiol* 2007, **48**(3):550-554.
13. Heim WG, Sykes KA, Hildreth SB, Sun J, Lu RH, Jelesko JG: **Cloning and characterization of a *Nicotiana tabacum* methylputrescine oxidase transcript.** *Phytochemistry* 2007, **68**(4):454-463.
14. Wagner R, Feth F, Wagner KG: **Regulation in tobacco callus of enzyme activities of the nicotine pathway : II. The pyridine-nucleotide cycle.** *Planta* 1986, **168**(3):408-413.
15. Sinclair SJ, Murphy KJ, Birch CD, Hamill JD: **Molecular characterization of quinolinate phosphoribosyltransferase (QPRTase) in *Nicotiana*.** *Plant Molecular Biology* 2000, **44**(5):603-617.
16. Katoh A, Uenohara K, Akita M, Hashimoto T: **Early steps in the biosynthesis of NAD in *Arabidopsis* start with aspartate and occur in the plastid.** *Plant Physiol* 2006, **141**(3):851-857.
17. Kajikawa M, Shoji T, Kato A, Hashimoto T: **Vacuole-Localized Berberine Bridge Enzyme-Like Proteins Are Required for a Late Step of Nicotine Biosynthesis in Tobacco.** *Plant Physiol* 2011, **155**(4):2010-2022.

18. DeBoer KD, Lye JC, Aitken CD, Su AKK, Hamill JD: **The A622 gene in *Nicotiana glauca* (tree tobacco): evidence for a functional role in pyridine alkaloid synthesis.** *Plant Molecular Biology* 2009, **69**(3):299-312.
19. Qi YC, Guo HX, Li K, Liu WQ: **Comprehensive analysis of differential genes and miRNA profiles for discovery of topping-responsive genes in flue-cured tobacco roots.** *FEBS J* 2012, **279**(6):1054-1070.
20. Xu S, Brockmoller T, Navarro-Quezada A, Kuhl H, Gase K, Ling Z, Zhou W, Kreitzer C, Stanke M, Tang H *et al.*: **Wild tobacco genomes reveal the evolution of nicotine biosynthesis.** *Proc Natl Acad Sci U S A* 2017, **114**(23):6133-6138.
21. Kajikawa M, Siero N, Kawaguchi H, Bakaher N, Ivanov NV, Hashimoto T, Shoji T: **Genomic Insights into the Evolution of the Nicotine Biosynthesis Pathway in Tobacco.** *Plant Physiol* 2017, **174**(2):999-1011.
22. Todd AT, Liu EW, Polvi SL, Pammett RT, Page JE: **A functional genomics screen identifies diverse transcription factors that regulate alkaloid biosynthesis in *Nicotiana benthamiana*.** *Plant J* 2010, **62**(4):589-600.
23. Chintapakorn Y, Hamill JD: **Antisense-mediated down-regulation of putrescine N-methyltransferase activity in transgenic *Nicotiana tabacum* L. can lead to elevated levels of anatabine at the expense of nicotine.** *Plant Molecular Biology* 2003, **53**(1):87-105.
24. Hibi N, Higashiguchi S, Hashimoto T, Yamada Y: **Gene-Expression in Tobacco Low-Nicotine Mutants.** *Plant Cell* 1994, **6**(5):723-735.
25. Fu Y, Guo H, Cheng Z, Wang R, Li G, Huo G, Liu W: **NtNAC-R1, a novel NAC transcription factor gene in tobacco roots, responds to mechanical damage of shoot meristem.** *Plant Physiol Biochem* 2013, **69**:74-81.
26. Dalton HL, Blomstedt CK, Neale AD, Gleadow R, DeBoer KD, Hamill JD: **Effects of down-regulating ornithine decarboxylase upon putrescine-associated metabolism and growth in *Nicotiana tabacum* L.** *J Exp Bot* 2016, **67**(11):3367-3381.
27. DeBoer KD, Dalton HL, Edward FJ, Hamill JD: **RNAi-mediated down-regulation of ornithine decarboxylase (ODC) leads to reduced nicotine and increased anatabine levels in transgenic *Nicotiana tabacum* L.** *Phytochemistry* 2011, **72**(4-5):344-355.
28. Xu BF, Timko MP: **Methyl jasmonate induced expression of the tobacco putrescine N-methyltransferase genes requires both G-box and GCC-motif elements.** *Plant Molecular Biology* 2004, **55**(5):743-761.
29. Riechers DE, Timko MP: **Structure and expression of the gene family encoding putrescine N-methyltransferase in *Nicotiana tabacum*: new clues to the evolutionary origin of cultivated tobacco.** *Plant Molecular Biology* 1999, **41**(3):387-401.
30. Shoji T, Nakajima K, Hashimoto T: **Ethylene Suppresses Jasmonate-Induced Gene Expression in Nicotine Biosynthesis.** *Plant Cell Physiol* 2000, **41**(9):1072-1076.
31. Ryan SM, Cane KA, DeBoer KD, Sinclair SJ, Brimblecombe R, Hamill JD: **Structure and expression of the quinolinate phosphoribosyltransferase (QPT) gene family in *Nicotiana*.** *Plant Science* 2012,

188:102-110.

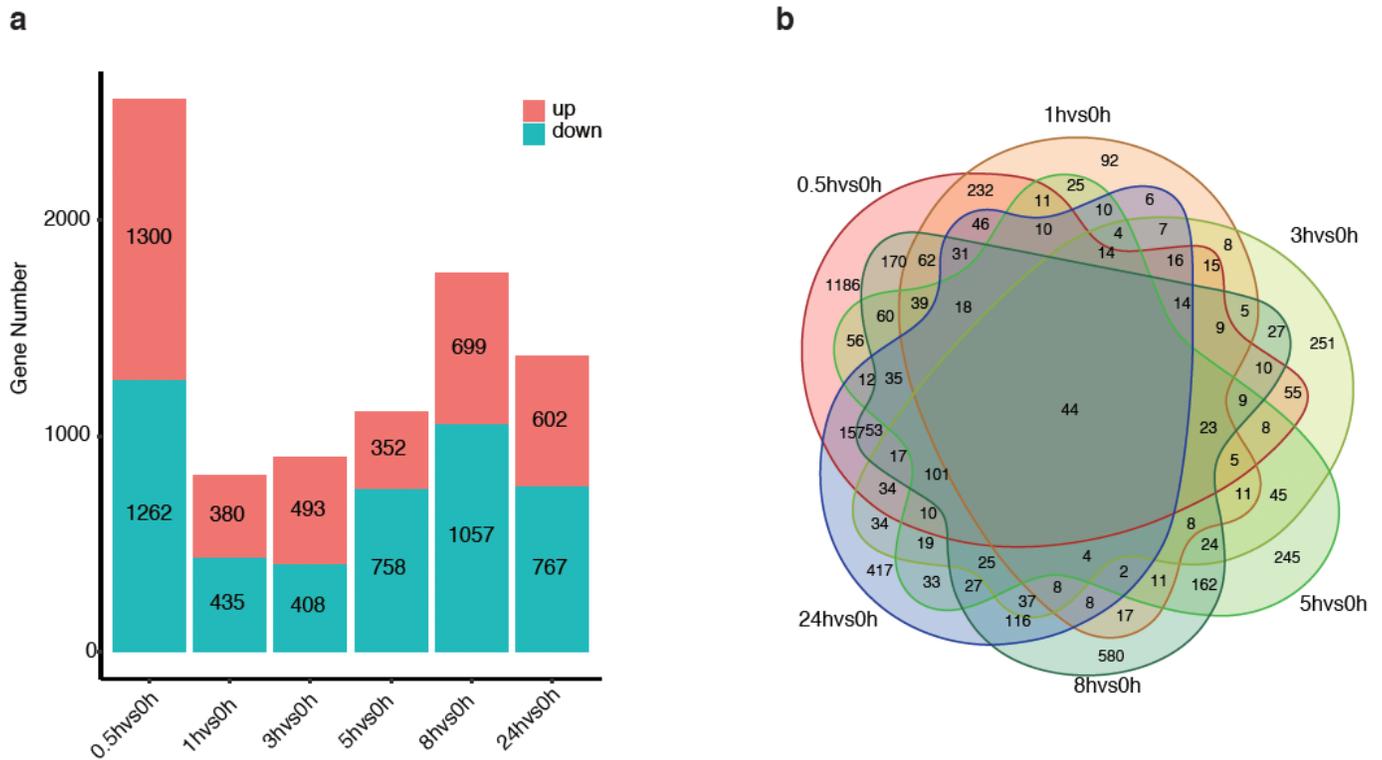
32. Shoji T, Hashimoto T: **Recruitment of a duplicated primary metabolism gene into the nicotine biosynthesis regulon in tobacco.** *Plant J* 2011, **67**(6):949-959.
33. Kajikawa M, Hirai N, Hashimoto T: **A PIP-family protein is required for biosynthesis of tobacco alkaloids.** *Plant Mol Biol* 2009, **69**(3):287-298.
34. Siervo N, Battey JND, Ouadi S, Bovet L, Goepfert S, Bakaher N, Peitsch MC, Ivanov NV: **Reference genomes and transcriptomes of *Nicotiana sylvestris* and *Nicotiana tomentosiformis*.** *Genome Biol* 2013, **14**(6).
35. Jin WH, Zhou Q, Wei YF, Yang JM, Hao FS, Cheng ZP, Guo HX, Liu WQ: **NtWRKY-R1, a Novel Transcription Factor, Integrates IAA and JA Signal Pathway under Topping Damage Stress in *Nicotiana tabacum*.** *Front Plant Sci* 2018, **8**.
36. Wang XW, Bennetzen JL: **Current status and prospects for the study of *Nicotiana* genomics, genetics, and nicotine biosynthesis genes.** *Mol Genet Genomics* 2015, **290**(1):11-21.
37. Rushton PJ, Bokowiec MT, Han S, Zhang H, Brannock JF, Chen X, Laudeman TW, Timko MP: **Tobacco transcription factors: novel insights into transcriptional regulation in the Solanaceae.** *Plant Physiol* 2008, **147**(1):280-295.
38. Sears MT, Zhang HB, Rushton PJ, Wu M, Han SC, Spano AJ, Timko MP: **NtERF32: a non-NIC2 locus AP2/ERF transcription factor required in jasmonate-inducible nicotine biosynthesis in tobacco.** *Plant Molecular Biology* 2014, **84**(1-2):49-66.
39. Goossens J, Fernandez-Calvo P, Schweizer F, Goossens A: **Jasmonates: signal transduction components and their roles in environmental stress responses.** *Plant Molecular Biology* 2016, **91**(6):673-689.
40. Zhang HB, Bokowiec MT, Rushton PJ, Han SC, Timko MP: **Tobacco Transcription Factors NtMYC2a and NtMYC2b Form Nuclear Complexes with the NtJAZ1 Repressor and Regulate Multiple Jasmonate-Inducible Steps in Nicotine Biosynthesis.** *Mol Plant* 2012, **5**(1):73-84.
41. Shoji T, Hashimoto T: **Tobacco MYC2 Regulates Jasmonate-Inducible Nicotine Biosynthesis Genes Directly and By Way of the NIC2-Locus ERF Genes.** *Plant Cell Physiol* 2011, **52**(6):1117-1130.
42. Wang B, Lewis RS, Shi J, Song Z, Gao Y, Li W, Chen H, Qu R: **Genetic Factors for Enhancement of Nicotine Levels in Cultivated Tobacco.** *Scientific Reports* 2015, **5**:17360.
43. De Geyter N, Gholami A, Goormachtig S, Goossens A: **Transcriptional machineries in jasmonate-elicited plant secondary metabolism.** *Trends Plant Sci* 2012, **17**(6):349-359.
44. Yang CQ, Fang X, Wu XM, Mao YB, Wang LJ, Chen XY: **Transcriptional Regulation of Plant Secondary Metabolism.** *J Integr Plant Biol* 2012, **54**(10):703-712.
45. Zhou ML, Memelink J: **Jasmonate-responsive transcription factors regulating plant secondary metabolism.** *Biotechnol Adv* 2016, **34**(4):441-449.
46. Hou XL, Lee LYC, Xia KF, Yen YY, Yu H: **DELLAs Modulate Jasmonate Signaling via Competitive Binding to JAZs.** *Developmental Cell* 2010, **19**(6):884-894.

47. Lackman P, Gonzalez-Guzman M, Tilleman S, Carqueijeiro I, Perez AC, Moses T, Seo M, Kanno Y, Hakkinen ST, Van Montagu MCE *et al.*: **Jasmonate signaling involves the abscisic acid receptor PYL4 to regulate metabolic reprogramming in Arabidopsis and tobacco.** *P Natl Acad Sci USA* 2011, **108**(14):5891-5896.
48. De Boer K, Tilleman S, Pauwels L, Vanden Bossche R, De Sutter V, Vanderhaeghen R, Hilson P, Hamill JD, Goossens A: **APETALA2/ETHYLENE RESPONSE FACTOR and basic helix-loop-helix tobacco transcription factors cooperatively mediate jasmonate-elicited nicotine biosynthesis.** *Plant J* 2011, **66**(6):1053-1065.
49. Palazon J, Pinol MT, Altabella T, Cusido R, Serrano M: **Auxin-Induced Regulation of Amino-Acid and Putrescine in the Free State and Nicotine Content in Cultured Tobacco Callus.** *Journal of Plant Physiology* 1987, **128**(1-2):153-159.
50. Xu BF, Sheehan MJ, Timko MP: **Differential induction of ornithine decarboxylase (ODC) gene family members in transgenic tobacco (*Nicotiana tabacum* L. cv. Bright Yellow 2) cell suspensions by methyl-jasmonate treatment.** *Plant Growth Regul* 2004, **44**(2):101-116.
51. Shoji T, Yamada Y, Hashimoto T: **Jasmonate Induction of Putrescine N-Methyltransferase Genes in the Root of *Nicotiana glauca*.** *Plant Cell Physiol* 2000, **41**(7):831-839.
52. Yang Y, Yan P, Yi C, Li W, Chai Y, Fei L, Gao P, Zhao H, Wang Y, Timko MP *et al.*: **Transcriptome-wide analysis of jasmonate-treated BY-2 cells reveals new transcriptional regulators associated with alkaloid formation in tobacco.** *J Plant Physiol* 2017, **215**:1-10.
53. Martin M: **Cutadapt removes adapter sequences from high-throughput sequencing reads.** *Embnet Journal* 2011, **17**(1).
54. Kong Y: **Btrim: A fast, lightweight adapter and quality trimming program for next-generation sequencing technologies.** *Genomics* 2011, **98**(2):152-153.
55. Edwards KD, Fernandez-Pozo N, Drake-Stowe K, Humphry M, Evans AD, Bombarely A, Allen F, Hurst R, White B, Kernodle SP *et al.*: **A reference genome for *Nicotiana tabacum* enables map-based cloning of homeologous loci implicated in nitrogen utilization efficiency.** *BMC Genomics* 2017, **18**(1):448.
56. Kim D, Langmead B, Salzberg SL: **HISAT: a fast spliced aligner with low memory requirements.** *Nat Methods* 2015, **12**(4):357-360
57. Mortazavi A, Williams BA, McCue K, Schaeffer L, Wold B: **Mapping and quantifying mammalian transcriptomes by RNA-Seq.** *Nat Methods* 2008, **5**(7):621-628.
58. Roberts A, Pimentel H, Trapnell C, Pachter LJB: **Identification of novel transcripts in annotated genomes using RNA-Seq.** 2011, **27**(17):2325-2329.
59. Ye J, Fang L, Zheng H *et al.*: **WEGO: a web tool for plotting GO annotations.** *Nucleic acids research* 2006, **34**(suppl\_2):W293-W297.
60. Yu G, Wang LG, Han Y, He QY: **clusterProfiler: an R package for comparing biological themes among gene clusters.** *Omic* 2012, **16**(5):284-287.
61. Onkokesung N, Gaquerel E, Kotkar H, Kaur H, Baldwin IT, Galis I: **MYB8 controls inducible phenolamide levels by activating three novel hydroxycinnamoyl-coenzyme A:polyamine transferases**

in *Nicotiana attenuata*. *Plant Physiol* 2012, **158**(1):389-407.

62. Wu J, Hettenhausen C, Meldau S, Baldwin IT: **Herbivory rapidly activates MAPK signaling in attacked and unattacked leaf regions but not between leaves of *Nicotiana attenuata***. *The Plant cell* 2007, **19**(3):1096-1122.

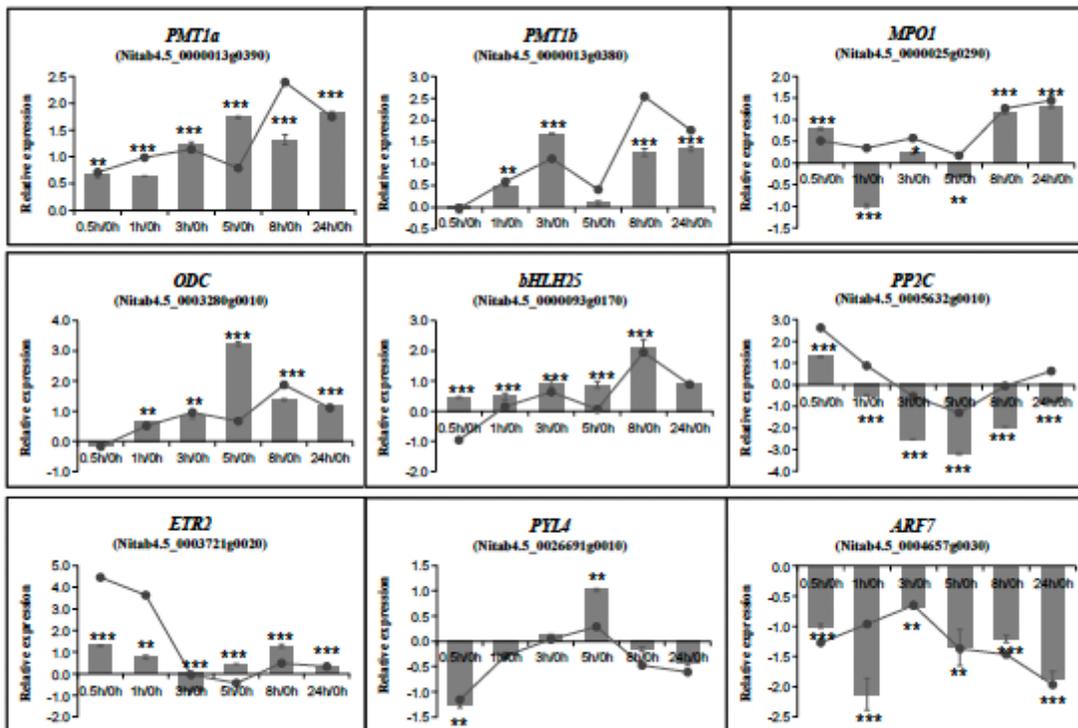
## Figures



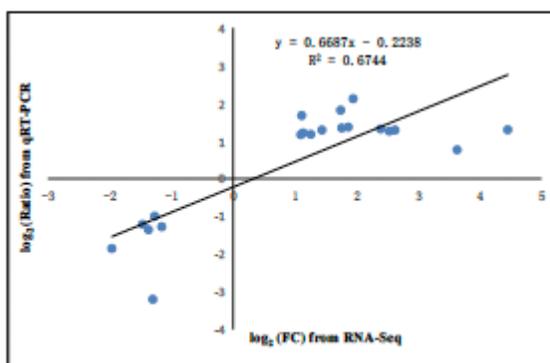
**Figure 1**

DEGs statistics at different stages after tobacco topping. a In the stacked bar-graphs, the up-regulated DEGs are located in the red regions, and the down-regulated DEGs in the blue regions. b Venn diagram of DEGs at different time points after topping in tobacco.

**a**

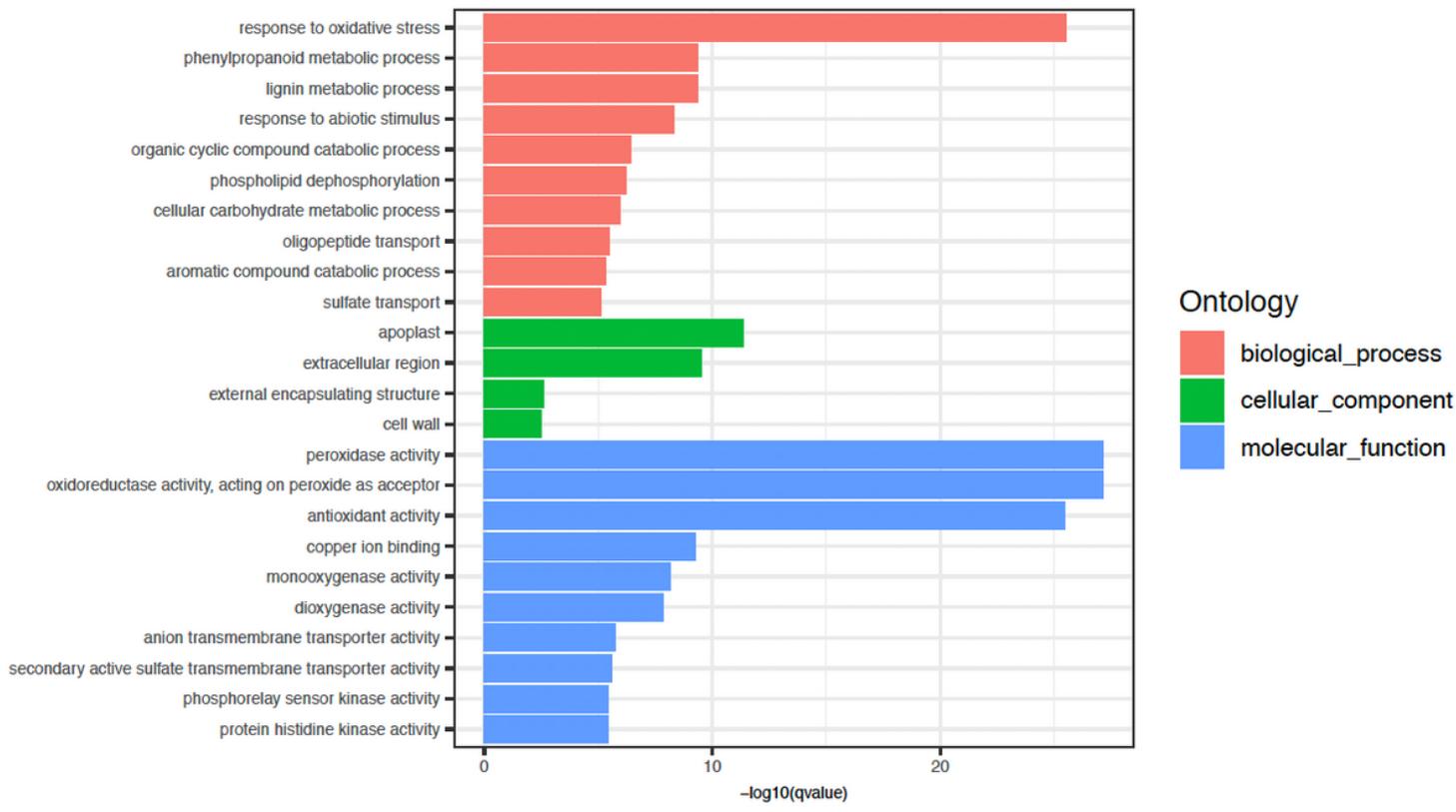


**b**



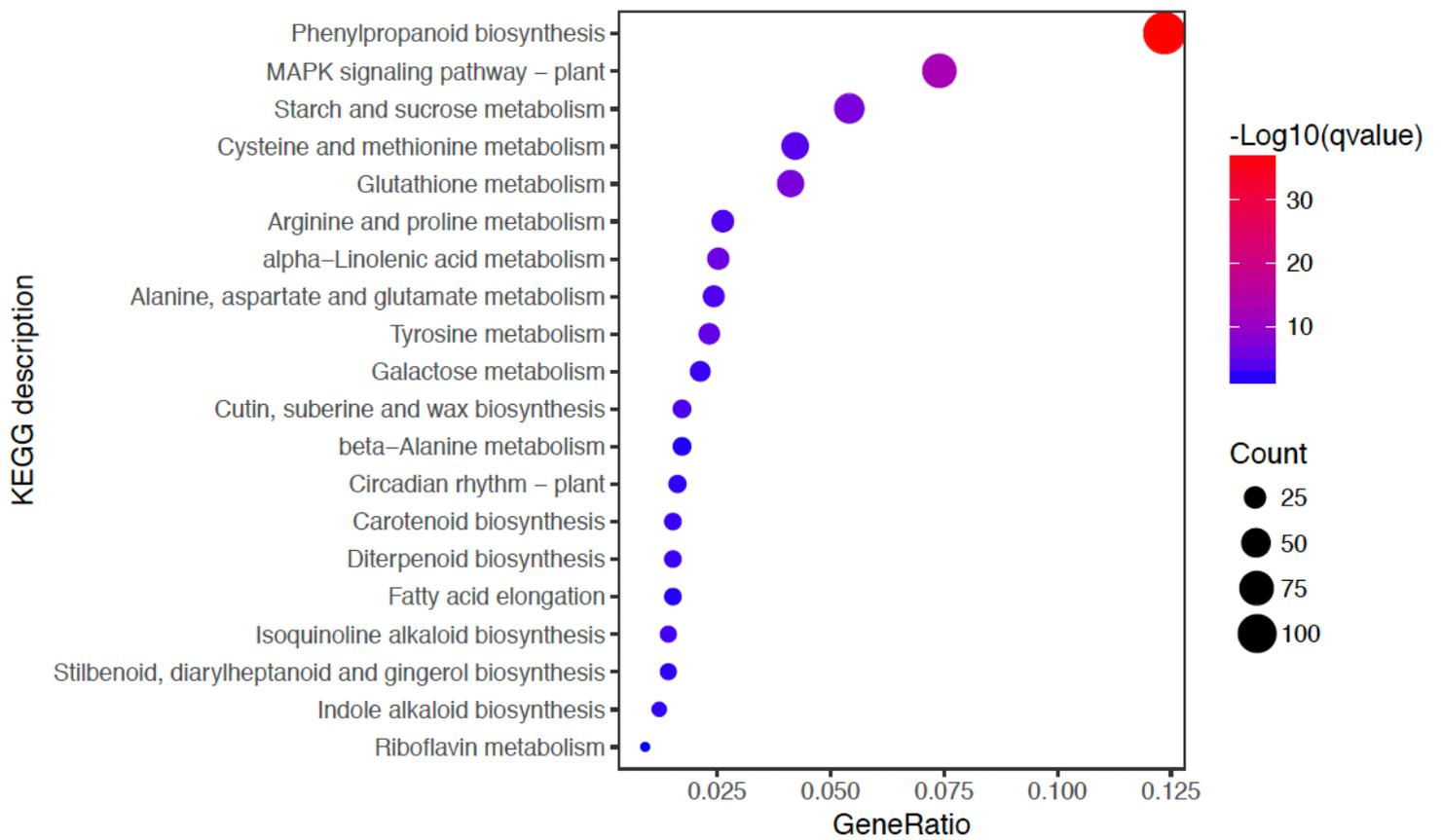
**Figure 2**

Validation of RNA-Seq data by qRT-PCR. a Expression levels of 9 randomly-selected DEGs of the nicotine anabolic pathway as measured by qRT-PCR (the columns) and the corresponding expression trends measured by RNA-Seq (the lines). The error bars represent SDs (n = 3). Asterisks represent significantly different transcript levels between the topping treatment and control plants at the indicated times. (t-test; \*, P < 0.05; \*\*, P < 0.01; \*\*\*, P < 0.001). b Correlation analysis of fold-change data between qRT-PCR and RNA-Seq. Scatterplots are generated from the  $\log_2$  expression ratios of qRT-PCR analyses (x-axis) and from RNA-Seq analyses (y-axis). Each scatter point depicts a time point at which significant differences in gene expression levels were found. The equation of the linear regression relationship and the associated correlation coefficient (R<sup>2</sup>) are provided.



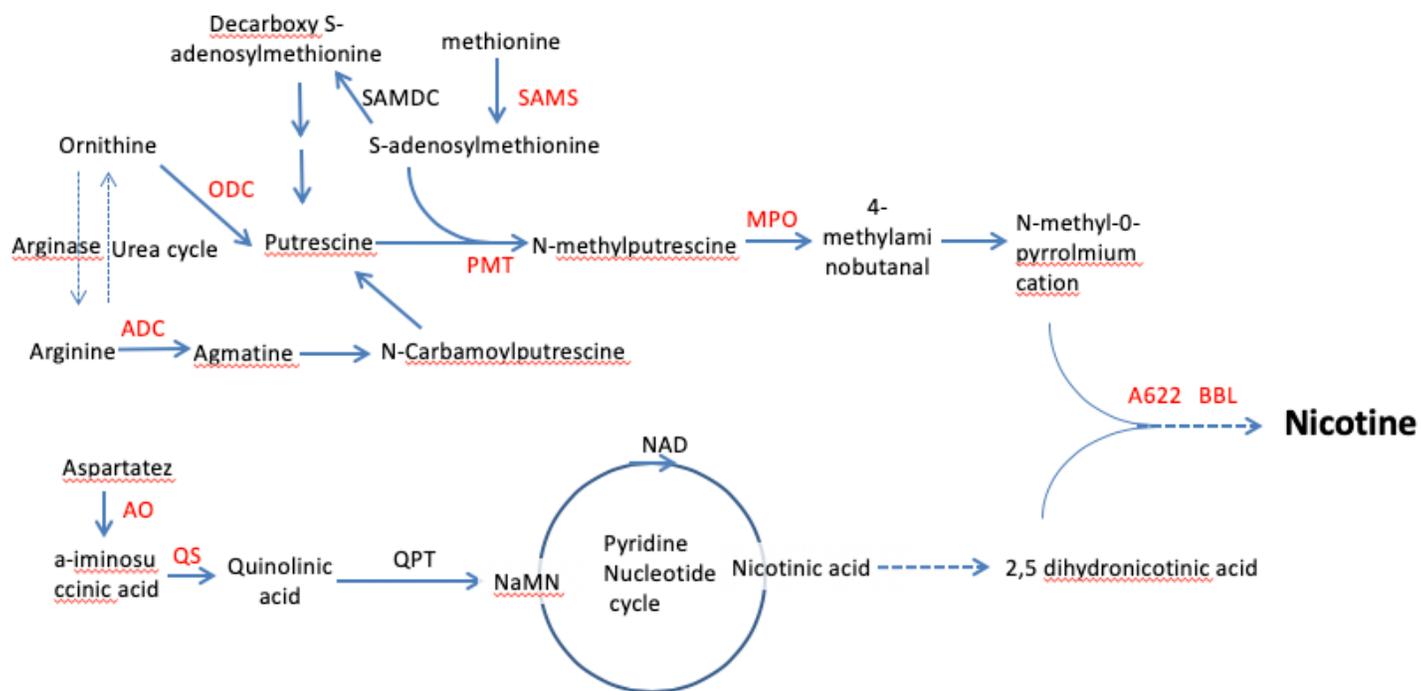
**Figure 3**

Gene Ontology (GO) term enrichment analysis. Significantly enriched GO terms were selected based on a FDR < 0.05. GO terms of the categories of Biological Processes, Cellular Components, and Molecular Functions are depicted in red, green, and blue, respectively.



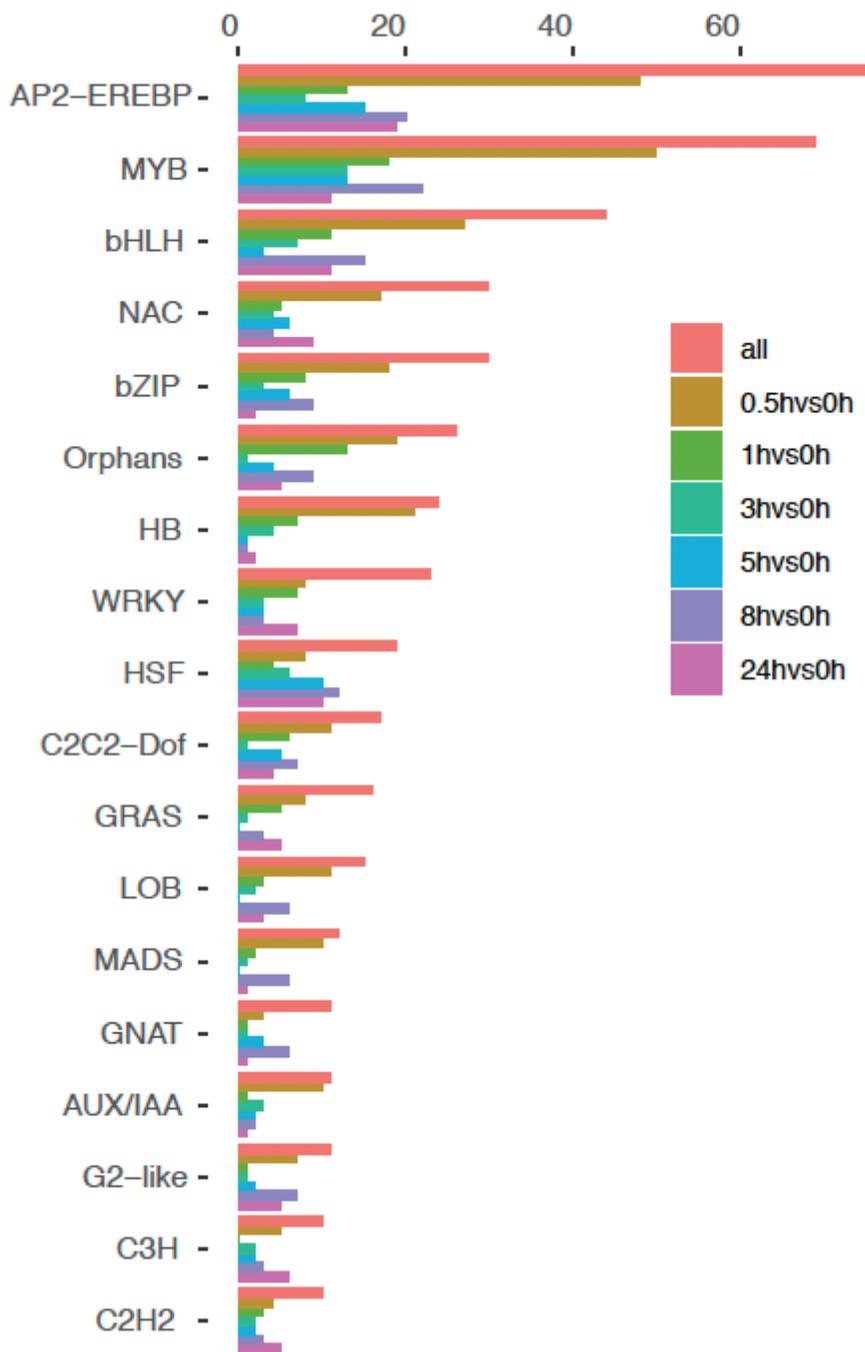
**Figure 4**

KEGG enrichment analysis. Each circle in the figure represents a KEGG metabolic pathway, and the number of genes enriched in a pathway corresponds to the size of the circle. The degree of significance of the enrichment of DEGs in a pathway is represented by  $-\log_{10}(\text{qvalue})$ . The abscissa indicates the ratio of the number of DEGs annotated to a particular pathway to the number of the DEGs annotated to all pathways.



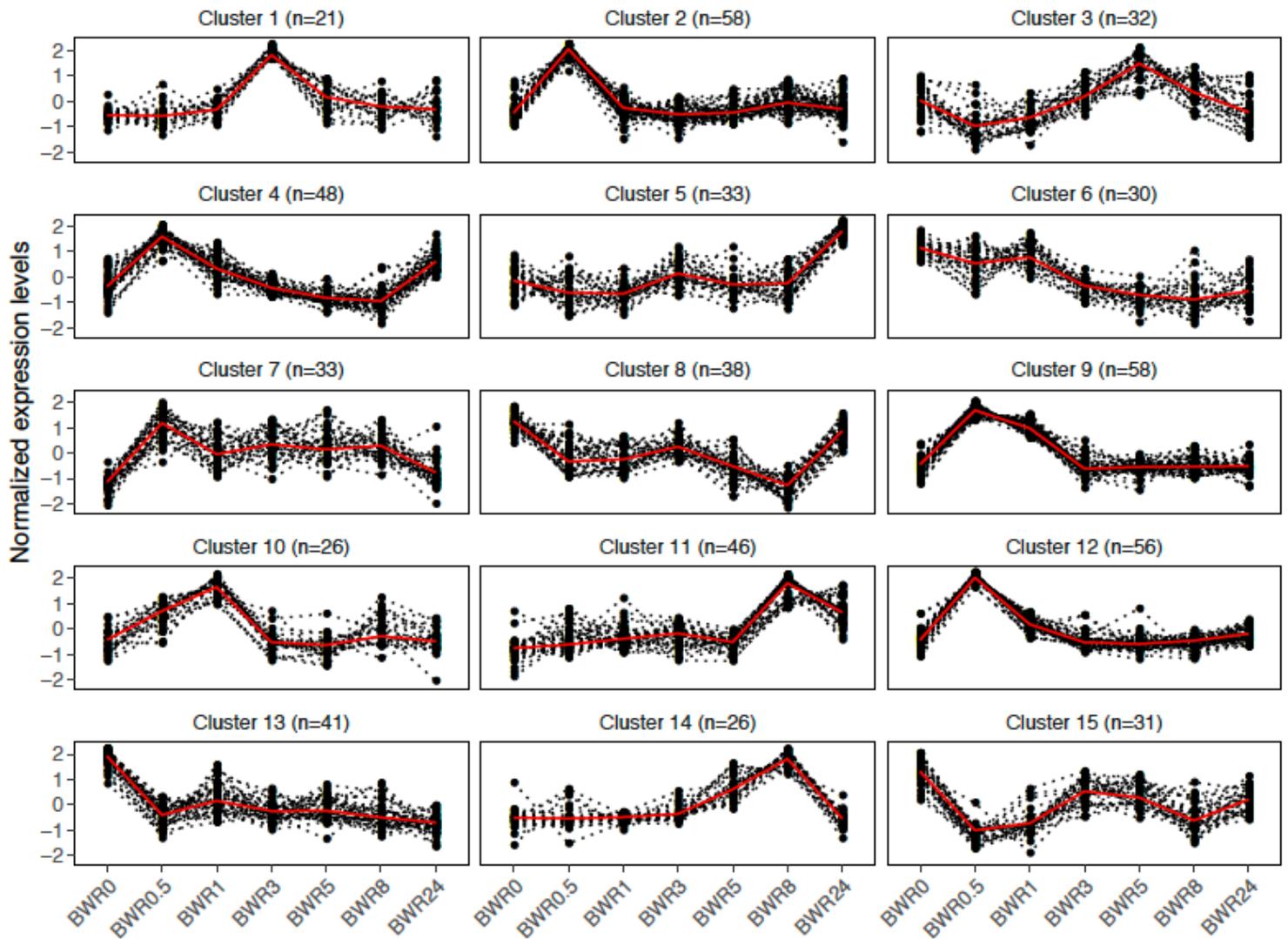
**Figure 5**

The DEGs involved in nicotine biosynthesis. Solid and dashed lines indicate defined and undefined reactions, respectively, with the DEGs shown in red. Abbreviations: ADC, arginine decarboxylase; ODC, ornithine decarboxylase; SAMS, S-adenosyl-L-methionine synthetase; SAMDC, S-adenosylmethionine decarboxylase; PMT, putrescine N-methyltransferase; MPO, N-methylputrescine oxidase; AO, aspartate oxidase; QS, quinolinate, synthase; QPT, quinolinic acid phosphoribosyl transferase; A622, isoflavone reductase-like protein; BBLs, berberine bridge enzyme-like proteins. The genes with significantly up-regulated transcription levels are shown in red.



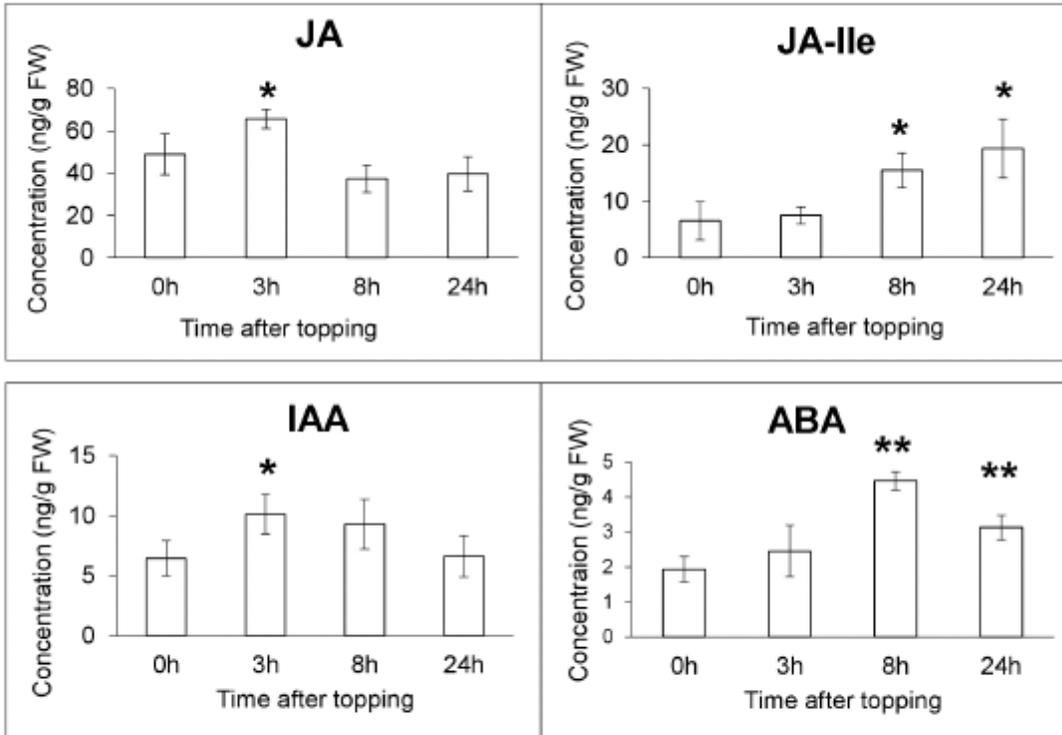
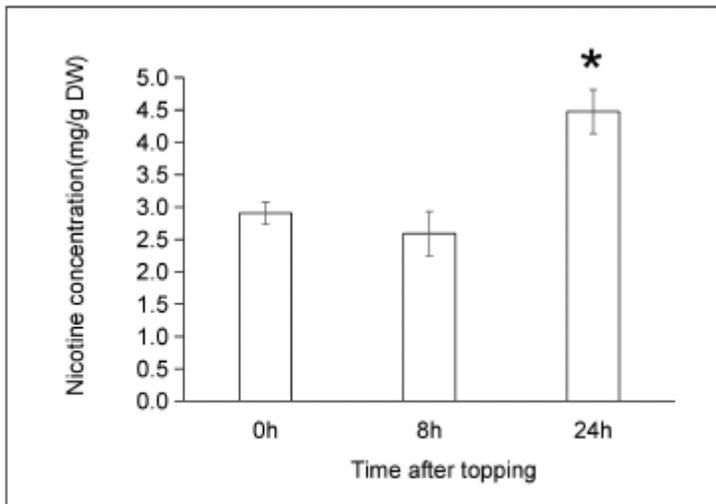
**Figure 6**

TFs classification of DEGs in tobacco. DEGs at different time points after topping are represented by different colors, the abscissa representing the number of transcription factors at each time point.



**Figure 7**

The clustering of gene expression pattern of DEGs on the TFs and the genes involved in nicotine biosynthesis and transport across different time points after topping in tobacco. The x-axis represents treatment conditions and the y-axis represents centralized and normalized expression values. The red lines indicate the mean expression trends of the TFs (dotted lines) belonging to each cluster. The gene number is marked following the cluster ID.

**a****b****Figure 8**

Mean (+SE) concentrations of phytohormones and nicotine from three replicates of roots harvested after topping treatment within the indicated times. a Mean (+SE) concentrations of JA, JA-Ile, IAA, and ABA from *N. tabacum* roots were measured using HPLC-MS/MS within 24 hours after topping. b The level of nicotine from roots within 24 hours after topping treatment, untreated plants served as controls. Asterisks represent significantly different hormone and nicotine levels between control and treatment plants after the indicated times. (t-test; \*,  $P < 0.05$ ; \*\*,  $P < 0.01$ ).

## Supplementary Files

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

- [TableS3.xlsx](#)
- [TableS1qualityandmappingrate.xlsx](#)
- [TableS7qPCRprimers.docx](#)
- [FigureS2.pdf](#)
- [FigureS1.pdf](#)
- [FigureS4.pdf](#)
- [TableS5.xlsx](#)
- [FigureS5.pdf](#)
- [TableS6.xlsx](#)
- [TableS2.xlsx](#)
- [FigureS3.pdf](#)
- [TableS4.xlsx](#)