scientific reports

Check for updates

OPEN

Gene editing, metabolomics, network pharmacology strategies to explore terpenoid content and anti-TMV activity in *NtSPS1* knockout *Nicotiana tabacum*

Jia-Meng Dai[1], Jian-Duo Zhang[1], Xin Liu[1], Ling-Fang Zhang[2], Jin Wang[1], Yong Xu[1], Guang-Yu Yang[1,2], Jing Li[1]✉, Ming-Li Chen[3]✉ & Qiu-Fen Hu[1,2]✉

The content of terpenoids in tobacco can alter its resistance to TMV. *NtSPS1*, a pivotal structural gene in tobacco, is capable to regulate the terpenoid content. In this study, we investigated the effect of *NtSPS1* knockout in HD on the content of terpenoids and the anti-TMV activity of this mutant using gene editing, widely targeted metabolomics, network pharmacology, and molecular docking. 48 terpenoids (six up-regulated and five down-regulated) in *NtSPS1* knockout tobacco compared with WT leaves. Notably, solanesol was remarkable downregulation which was lowered by fourfold and compounds 1 ($log_2FC = 18.2$), 8 ($log_2FC = 16.7$) were significant upregulation between the mutants and wild-type line leaves. The 46 terpenoid's target network encompassed 150 nodes, 509 edges and their underlying mechanisms in the therapeutic management of TMV are discussed. Furthermore, the network pharmacology and molecular docking revealed that compounds 16, 18, 23, 27, and 36 exhibited significant affinity in their respective interactions. Ultimately, five compounds were assayed for their anti-TMV effects, noteworthily, compounds 36 showed potential anti-TMV activity. Above all, we adopted a multifaceted approach to gain a comprehensive understanding of the terpenoid content and anti-TMV properties in *NtSPS1* knockout HD. It enlightens the therapeutic potential of *NtSPS1* knockout tobacco and it is helpful to find undescribed anti-TMV activity inhibitors, as well as searching for new anti-TMV candidates from the mutants.

Keywords *Nicotiana tabacum*, CRISPR/Cas9, *NtSPS1*, Terpenoid, Network pharmacology, anti-TMV activities

Abbreviations

NtSPS1	A solanesyl diphosphate synthases gene
HD	Nicotiana tabacum L. variety Hong Hua Da Jin Yuan
TMV	Tobacco mosaic virus
WT	Wild-type

Nicotiana tabacum (*N. tabacum*) is well known for its medicinal value purposes[1–3]. Previous phytochemical findings on *N. tabacum* showed terpenoids play a crucial role in multiple plant functions, for instance response to against biotic and abiotic stresses[4–6]. Tobacco mosaic virus (TMV) has a wide host range and naturally causes extensive damage in tobacco (*N. tabacum*)[7–10], tomato (*S. lycopersicum*)[11], and pepper (*C. annuum*)[12,13]. Plant metabolites, particularly terpenoids, have demonstrated the ability to counter TMV infection[14–17]. Such as solanesol, a non-cyclic trisesquiterpenoid alcohol, is mainly found in solanaceous plants: tobacco[18–20], potato[21,22], and tomato[23], which might participate in the immune response to TMV infection[24,25]. *NtSPS1* functions as a

[1]Yunnan Key Laboratory of Tobacco Chemistry, China Tobacco Yunnan Industrial Co., Ltd, Kunming 650231, P.R. China. [2]Key Laboratory of Chemistry in Ethnic Medicinal Resources, State Ethnic Affairs Commission & Ministry of Education, Yunnan Minzu University, Kunming 650031, P. R. China. [3]Tobacco Research Institute, Chinese Academy of Agricultural Sciences, Qingdao 266101, China. ✉email: lijing_1107@163.com; chenmingli@caas.cn; huqiufena@aliyun.com

key enzyme in the biosynthetic pathway of terpenoids, especially solanesol[26–28]. After *NtSPS1* overexpression in *N. tabacum* the total content of solanesol increase significantly[29,30]. The specific accumulation patterns and regulatory processes of terpenoids in *NtSPS1*-deficient Hong Hua Da Jin Yuan tobacco (*Nicotiana tabacum* L., HD) are still unclear. Additionally, there have been no prior studies on the terpenoid levels and anti-TMV effectiveness in this variety following *NtSPS1* gene knockout. The present study aimed to systematically compare the terpenoid content and anti-TMV activities of *NtSPS1* knockout HD and WT lines by multiple methods. *NtSPS1* knockout HD resulted in 48 terpenoids: six up-regulated, five down-regulated, especially remarkable downregulation of solanesol which was lowered by fourfold and significantly upregulation of compound **1** and **8** in mutants compared with the wild-type lines.

The 48 terpenoids were divided into five classes, involving eight monoterpenes (two up-regulated metabolites), 23 sesquiterpenes (four up-regulated, one down-regulated metabolites), ten diterpenoids, six triterpenoids (three down-regulated metabolites) and one trisesquiterpenoid (down-regulated metabolites) between the mutants and wild-type lines. Network pharmacology provides a comprehensive framework for examining the complex relationships between terpenoids and TMV, thus pioneering a new approach to decipher the therapeutic principles[31]. Guided by network pharmacology and molecular docking, we selected compounds **16**, **18**, **23**, **27**, and **36** from our laboratory for testing their anti-TMV activities. The results showed that compound **36** exhibited a protective response against HD, with inhibiting rates of 36.7% (IC_{50} = 32.1 μM).

Materials and methods

Plant material

The methods employed for construction of *NtSPS1* knockout *Nicotiana tabacum* L. variety Hong Hua Da Jin Yuan (HD) and growth conditions were the same as those described by Zhang et al. (Supplementary Method S1).

Chemical profiles

LC-MS/MS methods are employed to determine the total terpenoid content (Supplementary Method S2). Extraction of solanesol from mature leaves of *NtSPS1* knockout and WT HD was performed according to the method described by Zhao et al. (Supplementary Method S2.1.1, S2.2)[32]. The LC-MS/MS-based widely targeted metabolomics analysis was used to analyze the terpenoid content in *NtSPS1* knockout HD, as reported by Li et al. (Supplementary Method S2.1.2, S2.3)[33]. The T_1 plants (T_1 mutant HD), which were free of transgenes and carried a homozygous mutation resulting in premature stop codons, were allowed to self-fertilize for further inheritance in the T_2 generation (T_2 mutant HD). To determine the effect of *NtSPS1* knockout on HD, T_2 *ntsps1* mutants (T_2 mutant HD) and wild-type (WT) HD were cultivated in a glasshouse. The terpenoid content in the leaves of each HD at maturity stage were determined by analyzing their MS fragmentations and/or comparing them with reference samples (Supplementary Tables S1, S2). The PCA scatter plots were generated for all the samples to identify potential biomarkers after data normalization using Principal Component Analysis (PCA). This analysis provides an overview of the chemical relationship between leaves of *NtSPS1* knockout and WT HD.

Statistical analysis

The orthogonal partial least squares discriminant analysis (OPLS-DA) and principal component analysis (PCA) were carried out on all the samples to identify the putative biomarkers after data normalization. Unsupervised PCA (principal component analysis) was performed by statistics function prcomp within R (www.r-project.org). The data was unit variance scaled before unsupervised PCA. VIP values were extracted from OPLS-DA result, which also contain score plots and permutation plots, was generated using R package MetaboAnalystR. The data was log transform and mean centering before OPLS-DA. In order to avoid overfitting, a permutation test (200 permutations) was performed. The HCA (hierarchical cluster analysis) results of samples and metabolites were presented as heatmaps with dendrograms, while pearson correlation coefficients (PCC) between samples were caculated by the cor function in R and presented as only heatmaps. Both HCA and PCC were carried out by R package ComplexHeatmap. For HCA, normalized signal intensities of metabolites (unit variance scaling) are visualized as a color spectrum. For two-group analysis, differential metabolites were determined by VIP ($VIP > 1$), *P*-value ($P\text{-value} < 0.05$, Student's *t* test) and absolute Log_2FC ($|Log_2FC| \geq 0.58$).

Network pharmacological analysis

Screening natural product and targets

The database of Swiss Target Prediction (http://www.swisstargetprediction.ch/) (accessed on 1 October 2024) was utilized to search for the compounds and their respective targets of *N. tabacum*. Subsequently, the molecules and their corresponding target proteins of *N. tabacum* were determined. By accessing the Uniprot database (https://www.uniprot.org) (accessed on 1 October 2024), the gene names linked to these protein targets were efficiently obtained.

Construction of compound-target network

Cytoscape is a powerful implement for visualizing network pharmacology and analyzing associated results[34,35]. By employing Cytoscape 3.10.0 (https://github.com/cytoscape/cytoscape/releases/3.10.0/), we successfully integrated the terpenoids of *N. tabacum* with their corresponding targets, thereby robustly constructing a compound-target network for this medicinal plant. The integrated target network is constructed with nodes representing either compounds or targets, and the interactions between them are illustrated by lines.

Identification of predicted targets of tobacco mosaic virus

The disease targets of tobacco mosaic virus were retrieved in the Genecards databases (https://www.genecards.org/) (accessed on 1 October 2024)[36], and OMIM databases (https://www.omim.org/)(accessed on 1 October

2024)[37]. The procured targets from the two databases were merged, removing any duplicate entries in the process for further investigation.

Construction of PPI network

We imported the target data of those molecules as well as for TMV, into the Venn Diagram (http://bioinformatics.psb.ugent.be/webtools/Venn/) (accessed on 1 October 2024). This program was utilized to pinpoint the relationship between the compound targets and the TMV targets. The overlapping parts, which are involved in both the disease pathology of TMV and the molecular pharmacodynamics, can be tentatively identified as the effective therapeutic targets where the medicinal impact of drugs is genuinely manifested. We constructed an integrated network of molecules and disease targets to identify the interactions between the compounds and their associated therapeutic targets. Subsequently, he integrated network was uploaded to the STRING database (https://string-db.org/) (accessed on 1 October 2024) to obtain information on the protein-protein interaction network[38]. The application of this method is relatively rare in TMV, hence no core targets were identified. So, we utilized the previously identified targets for the subsequent work[39–41].

Molecular docking

We accessed ChemDraw to obtain the molecule structures, the 3D structural files were converted by Chem3D version 18.0. The RSK structure, identified by the PDB code 2OM3, was chosen and prepared for further analysis. Hydrogen was removed, and polar hydrogen atoms were introduced using AutoDock version 4.2.6 (https://autodock.scripps.edu/download-autodock4/) tools. Active sites were constructed, and each one was subsequently docked with the molecule. Active pockets were built, and each active site was docked with the molecule[42].

Anti-tobacco mosaic virus activity assays

Using half-leaf method detecte the anti-TMV activities and using ningnanmycin as a positive control (Supplementary Method S2.4). The antiviral inhibition rates of molecules **16**, **18**, **23**, **27**, and **36** at 20.0 μM are shown in Table 1.

Results

NtSPS1 mutants obtained by using CRISPR/Cas9 gene editing

The HD plants with *NtSPS1* gene knockout were successfully generated using CRISPR/Cas9 technology. An optimized gene editing vector for tobacco, developed by our group, was used to drive the expression of the *NtSPS1* sgRNA and the 2 × 35 S promoter was used to drive the expression of *Streptococcus pyogenes Cas9 (SpCas9)*[43]. *NPTII*, neomycin phosphotransferase gene; LB, left border; RB, right border. (Fig. 1A–E). The sequencing data revealed that two distinct lines exhibited two different types of mutation sites, which deletion of 2 bases (Fig. 1F).

No off-target effects were observed for the three genes in the *ntlht1* mutant strains, including *Ntab0818090* and the *NtLHT1* gene (Fig. 2A–C). Based on the results, *NtSPS1* mutants were successfully obtained, and a dark-red marker was introduced to facilitate the generation of transgene-free plants in the T_1 generation and obtained enough homozygous T_2 population. The phenotypes of the *NtSPS1* mutants in the T_2 generation were same as those of WT HD (Fig. S2).

Chemical profiles

Comparison of the metabolic profiles between HDK and HDS group based on widely targeted metabolomics

The overall metabolic profiles of T_2 mutant plants (HDK) and WT plants (HDS) at the mature stage show a clear separation tendency, indicating that the terpenoids in the tobacco leaves changed after the knockout of the *NtSPS1* gene (Fig. 3A). And the largest principal component PC1 contributes 43.67%, and PC2 and PC3 make the contributions of 12.10% and 10.70%, respectively. The cumulative contribution rate of TOP3 principal components is greater than 60%, which indicated the overall explanation ability of the variance of original variables was acceptable.

Functional pathway analysis related to NtSPS1 knockout

The pathway enrichment analysis of 11 metabolic pathways including biosynthesis of flavonoids, alkaloids and terpenoids metabolism, etc. based on metabolic were performed (Fig. 3B). *NtSPS1* knockout leads to a significant decrease in the content of downstream terpenoids such as solanesol. Due to limitations like incomplete database coverage and insufficient terpenoid concentrations leading to a low detection rate, the terpenoids enrichment factor is lower than that of flavonoid and alkaloid. Nevertheless, the enriched result the terpenoid pathway

No.	% inhibition at 20 μM	IC_{50} (μM)
16	20.5 ± 2.6	53.9
18	25.2 ± 3.0	65.6
23	25.8 ± 2.8	65.8
27	26.8 ± 3.0	48.9
36	36.7 ± 3.2	32.1
Ningnanmycin	33.2 ± 2.2	36.1

Table 1. Anti-TMV activities of terpenoids[a]. [a]All results are expressed as mean ± SD; $n = 3$.

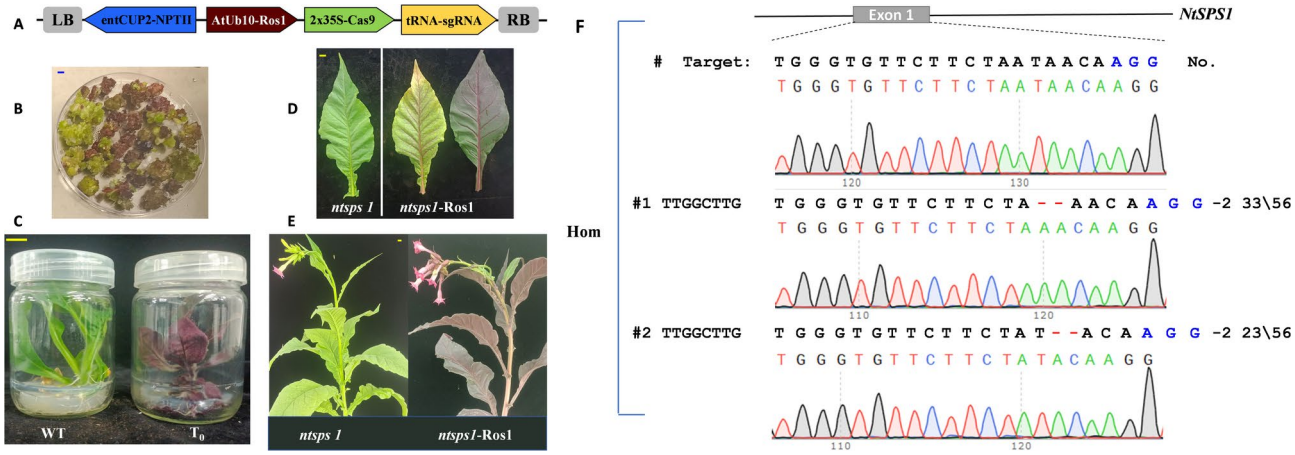

Fig. 1. Authentic *NtSPS1* of tobacco mutants and wild-type HD. (**A**) Schematic diagram of pOREU3TR vectors, (**B**–**E**) The phenotype of T_0 mutant plants (T_0 mutant HD) and wild type HD in different periods. WT, wild-type HD. *ntsps1*, *NtSPS1* editing HD. *ntsps1*-Ros1, a transgenic plant carrying a red marker. (**F**) Mutations for the 56 green tobacco in the T_1 generation (T_1 mutant HD). The protospacer adjacent motif (PAM) site "AGG" needed for Cas9 fragmentation is marked in blue. *Hom* homozygous. #, related plant. *No.* number.

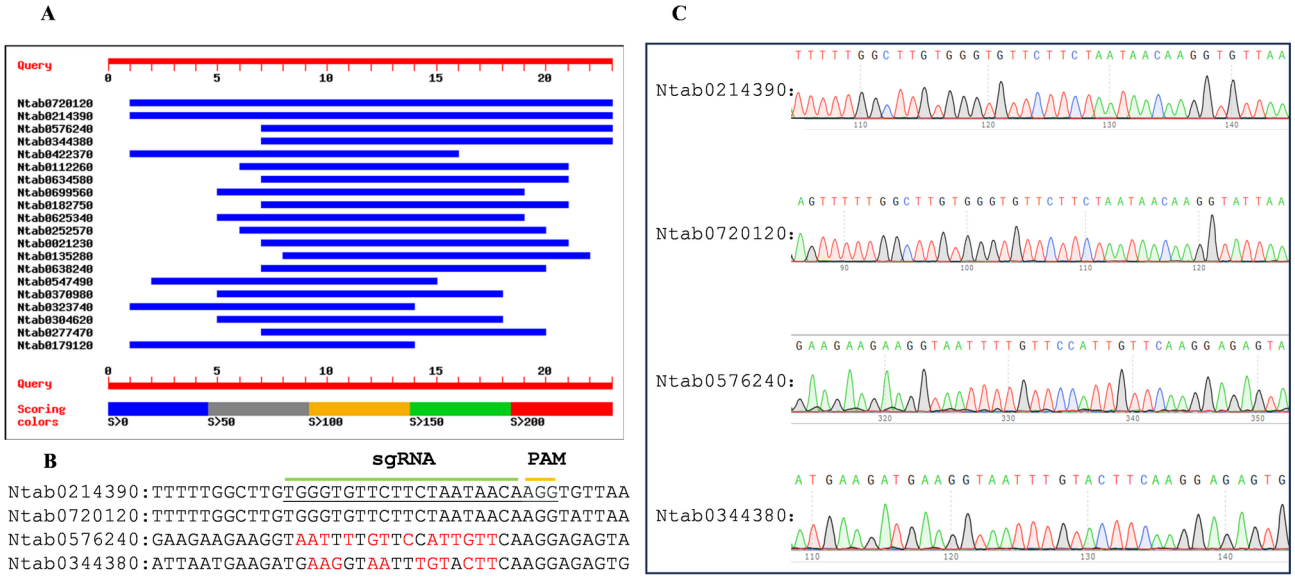

Fig. 2. Exploration of the off-target sites. (**A**) Blastn result for the *NtSPS1* sgRNA in *China Tobacco Genome* Database. (**B**) DNA sequences cognition of the 3 off-target genes among the sgRNA site. (**C**) There were no off-target items detected for the 4 genes in *ntsps1* mutants. *Ntab0214390*, *NtSPS1* gene.

ranked TOP3 as its *p*-value was 0.053 (almost significant) and the FDR is lower than 0.2, demonstrating the regulation of terpenoid pathway after the *NtSPS1* knockout. Moreover, since there are currently few reports on the study of terpenoids and related pathways, further study was carried out.

Identification and confirmation of terpenoids based on LC-MS/MS analysis

HDK and HDS were measured by LC-MS/MS based on widely targeted metabolomics analysis in positive ion mode. The results demonstrated that the mutants showed obvious regulation of terpenoid content. The assignment of the peaks for 47 terpenoids were shown in Fig. 4 and Table S1. These molecules including 8 monoterpenes (**1**, **5**, **6**, **7**, **11**, **13**, **22**, **29**), 23 sesquiterpenes (**2**, **3**, **4**, **8**, **9**, **10**, **12**, **14**, **15**, **16**, **17**, **18**, **20**, **23**, **25**, **26**, **27**, **30**, **31**, **32**, **33**, **34**, **36**), 10 diterpenoids (**19**, **21**, **24**, **28**, **35**, **38**, **41**, **42**, **43**, **44**), 6 triterpenoids (**37**, **39**, **40**, **45**, **46**, **47**) were characterized based on MS^n fragmentation and reference materials (Fig. 4A and Tables S1, S2, Fig. S2).

The principal component analysis (PCA) loading plot based on all terpenoids was shown in Fig. 4B. *NtSPS1* knockout leads to a significant decrease in the content of downstream terpenoids such as solanesol. We have examined solanesol content using the method reported by Zhang et al. (see Supplementary Method S2.1.1), and

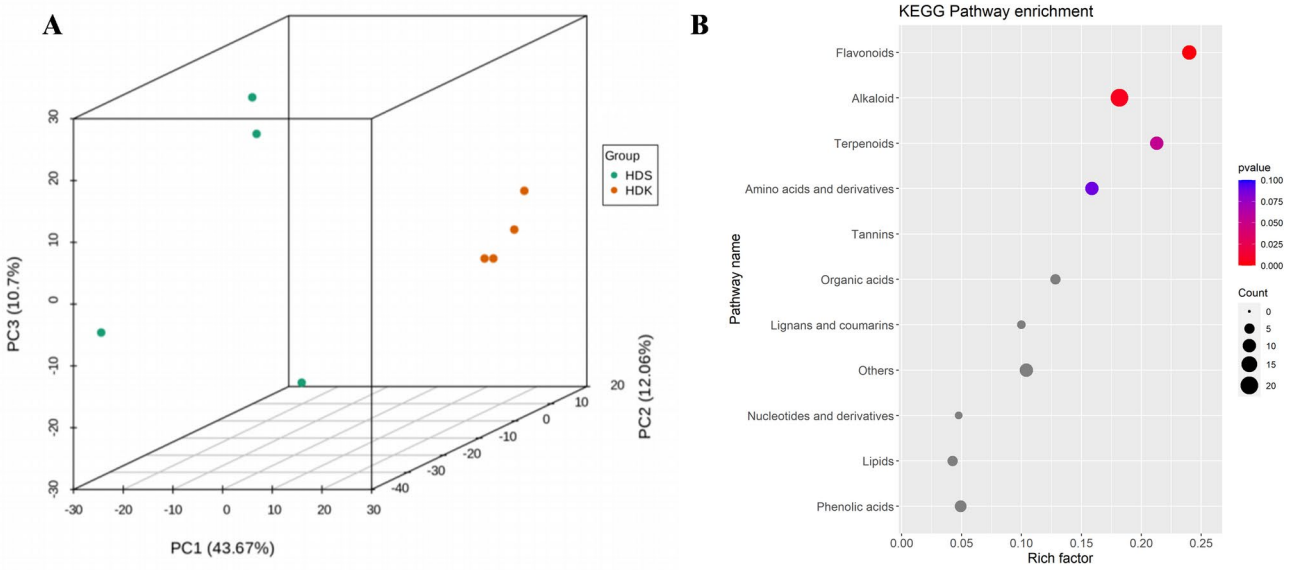

Fig. 3. The largest principal component, pathway enrichment, heat map analysis of HDS and HDK based on untargeted metabolomics. (**A**) The largest principal component plot of HDS (in green) and HDK (in red). (**B**) Pathway enrichment bubble plot related to HDS. The bubbles' size and darkness pf the color represented the count and level of the enrichment for terpenoids. The mutation types (HDK) and wild types (HDS) at the mature stage.

we found that solanesol content has decreased by fourfold. There was a clear distinction between the HDK and HDS groups, with the replicate samples of each cluster closely grouped together. This suggests that the variations in terpenoids were consistent across the samples within each group and strongly associated with the *NtSPS1* knockout of the HDK group.

Metabolic changes in HDS and HDK

The changes of 48 terpenoids before and after *NtSPS1* knockout in HD are presented. Among them, 10 terpenoids were differential between HDS and HDK (|Fold change| ≥ 2 and VIP ≥ 1), 6 terpenoids were upregulated and 5 were downregulated, 37 did not show significant changes (Fig. 4C, and D, Supplementary Method S2.2).

Network pharmacological analysis

Compound-target network and analysis

We have created a detailed network of component-target interactions for these molecules, considering their unique structural features and specific receptors as identified by Swiss target prediction. As shown in Fig. 5A, a total of 150 nodes and 509 edges were assigned, with green ellipses indicating potential targets for anti-TMV. We obtained 36 compounds and their respective targets for *N. tabacum* from the Swiss target prediction database.

PPI network construction and analysis

After screening the active targets of TMV through the disease database and eliminating duplicate values, we identified 390 disease targets. Subsequently, we employed this data to create a Venn chart (Fig. 5B), comparing the active target compounds associated with the disease to those found in the compound formulations. The overlap between these sets revealed 115 common targets between the molecules and TMV. To clarify the complex interactions between pharmaceutical constituents and TMV targets, we established an extensive network linking compounds to disease targets. To depict these shared targets, we constructed a PPI network featuring 7 nodes, yet without any connecting edges. So, we utilized the previously identified targets for the subsequent work.

Molecular docking

Utilizing network pharmacology, we have previously pinpointed five essential active ingredients: compounds **16**, **18**, **23**, **27** and **36**. Subsequently, we performed molecular docking analysis of these compounds with RSK protein (Fig. 6). This analysis sought to confirm the critical function of the compounds in treating TMV. Following that, we conducted a detailed examination of the docking outcomes for proteins and ligands individually. Subsequently, we scrutinized the docking results of proteins and ligands separately, molecular docking revealed that **16**, **18**, **23**, **27**, and **36** were well located within the bonding site of anti-TMV by contacting ARG92, ILE93, SER15, VAL11, SER14, PHE10, GLY85, ASP116, and ALA86 residues. Our results show that the compounds have a strong affinity for binding to the RSK protein. Additionally, the interactions between each ligand-macromolecule complex are facilitated by hydrogen bonds, ensuring the stability of these interactions.

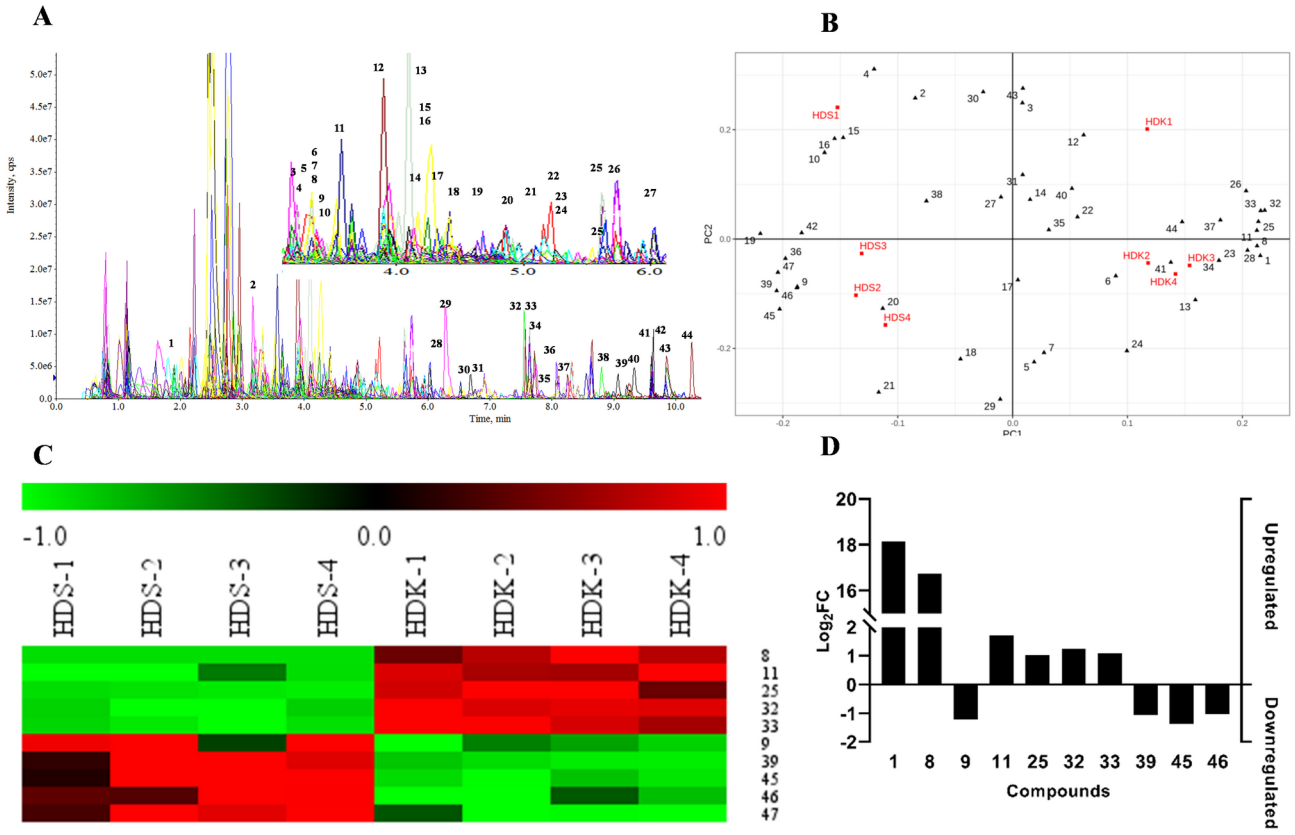

Fig. 4. Untargeted metabolomics analysis. (**A**) LC-MS/MS-base multiple reaction monitoring mode (MRM) of one *NtSPS1* knockout HD (HDK) in positive ion mode. (**B**) PCA loading plot of the 47 terpenoids. HDS, the wild types. Numbers are the IDs of 47 terpenoids. (**C**) Heatmap analysis was conducted on 10 terpenoids which change significantly between HDS and HDK, the content of terpenoids were normalized and visualized where the red and green color represented up- and down-regulation. (**D**) The variation in terpenoid content (Log_2FC) between HDS and HDK. For each test, three specialized replicates were managed. At least 6 tobacco were evaluated in each sample. Data interpreted are mean ± SE of three specialty tests.

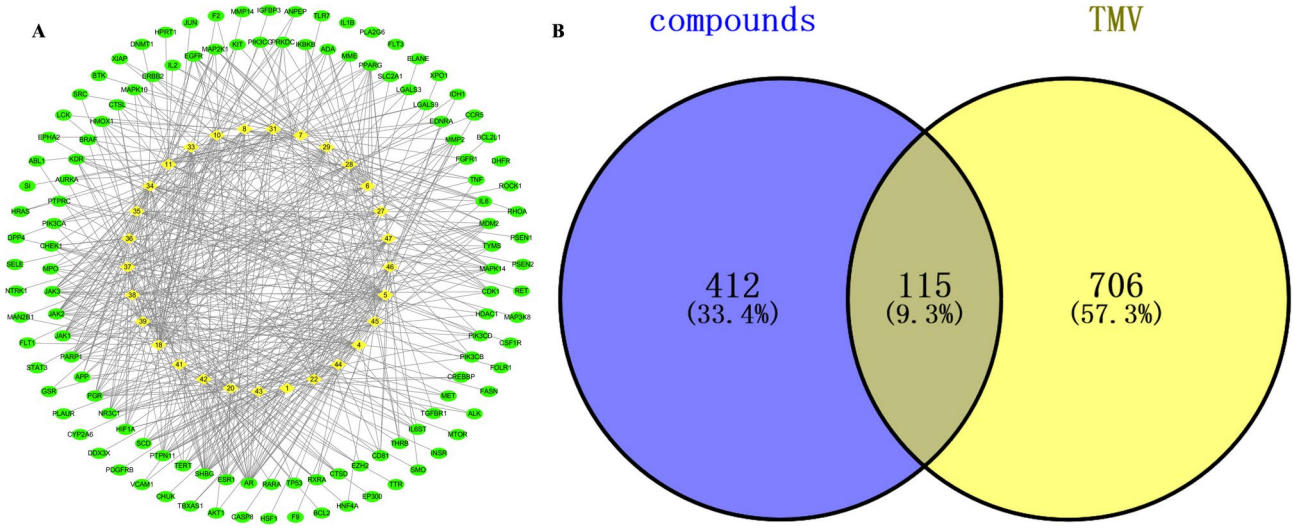

Fig. 5. Network pharmacological analysis of terpenoids in TMV. (**A**) Compound-target network (36 yellow diamonds represent the terpenoids of *N. tabacum*, the circles around the diamonds represent drug targets). (**B**) Venn diagram of the target of compounds and the target of TMV.

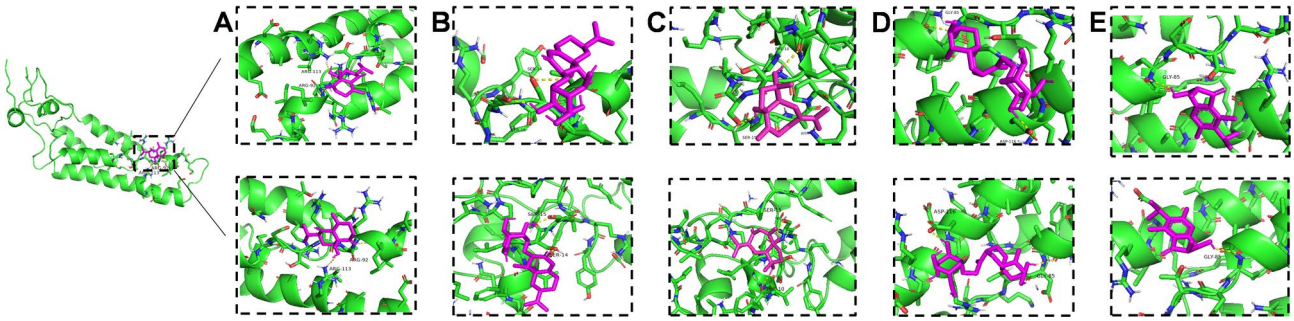

Fig. 6. Interactions between TMV-helicase (sticks are used to represent key residues while solid lines indicate H-bonds) and **16** (**A**), **18** (**B**), **23** (**C**), **27** (**D**), and **36** (**E**).

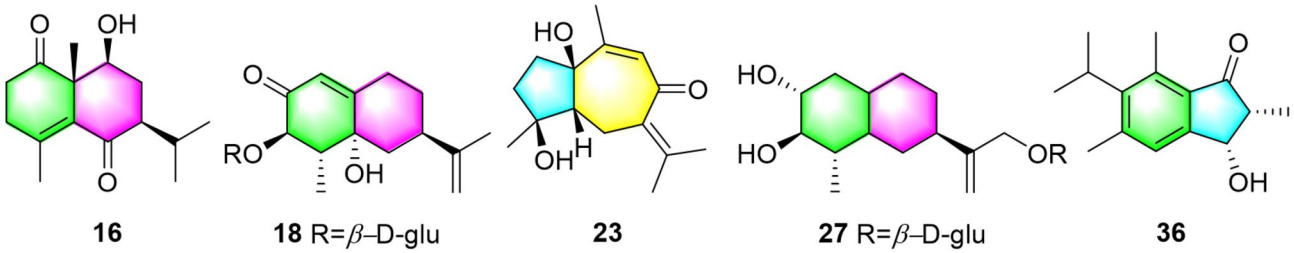

Fig. 7. Structures of compounds for anti-TMV evaluation.

Anti-tobacco mosaic virus activity assays

Research indicates that solanesol could be involved in the resistance anti-TMV, but upon TMV infection, mutants (solanesol content has decreased by fourfold) and wild-type plants exhibit comparable characteristics (Fig. 8A and B). Other substances might be responsible for anti-TMV effects; to verify this hypothesis, guided by network pharmacology and molecular docking, the anti-TMV properties of compounds **16**, **18**, **23**, **27**, and **36** were evaluated (Fig. 7). The HD was pre-treated with 20.0 μM solutions of the compounds or a DMSO solution for 6 h prior to inoculation with TMV. The consequences clearly show that **36** offered protection against TMV, with an IC_{50} value of 32.1 μM (Table 1). For the practical application of our lab findings to tobacco cultivation, we chose mutant tobacco at the 7–8 leaf stage to evaluate the therapeutic efficacy of **36** against TMV with ningnanmycin as positive control. The negative control group showed a severe infection of TMV, the lesions were clearly visible, the growth rate was slow, and the leaves exhibited noticeable deformities (Fig. 8B), this may substantially affect the tobacco leaf yield and quality. However, after therapy with **36** (compared with ningnanmycin, Fig. 8C), the lesions showed a decrease compared to that of ningnanmycin (Fig. 8D). This result demonstrates that **36** exerts a significant therapeutic effect on HDK infected with TMV, and it can substantially alleviate the detrimental effects of TMV on the host plants.

Discussion

We identified 48 terpenoid metabolites and categorized them into five distinct classes: monoterpenoids, sesquiterpenes, diterpenoids, triterpenoids, and polyisoprenoid alcohol. Among these, compounds **1**, **8**, **9**, **11**, **25**, **32**, **33**, **39**, **45**, **46**, and solanesol exhibited differential accumulation between the mature leaves of the *NtSPS1* knockout plants and wild-type plants (Tables S1, S2, and S3). Knockdown of *NtSPS1* may inhibit the synthesis of compounds **9**, **39**, **45**, **46**, and solanesol, potentially leading to the accumulation of compounds **1**, **8**, **11**, **25**, **32**, and **33**, which could consequently result in the observed effects. The PCA scatter plots summarize the chemical connection between mutants and WT plants at the mature stage. Meanwhile, it can be observed from Fig. 4B that HDK samples were clustered together, while HDS samples were grouped together, which indicated that our analysis was stable and repeatable.

Additionally, we investigated the potential healing properties of the 36 metabolites for anti-TMV infections. These discoveries highlight the key role of terpenoids from *N. tabacum* in managing TMV. By extracting TMV-related targets from the Genecards database, we pinpointed 115 potential targets that overlapped with the compounds. The analysis of docking poses indicated that compounds **16**, **18**, **23**, **27**, and **36** exhibited strong interactions with the catalytic pocket of TMV (Fig. 6). Compound **16**, the oxygen of the carbonyl group formed hydrogen bond with ARG92 and ILE93 residues (Fig. 6A). Compound **18**, the hydroxyl groups on the ligand backbone and sugar formed hydrogen bonds with SER15 and VAL11 residues (Fig. 6B). Compound **23**, the hydroxyl groups on the ligand backbone formed hydrogen bonds with SER14 and PHE10 residues (Fig. 6C). Compound **27**, the hydroxyl groups on the ligand backbone and sugar formed hydrogen bonds with GLY85,

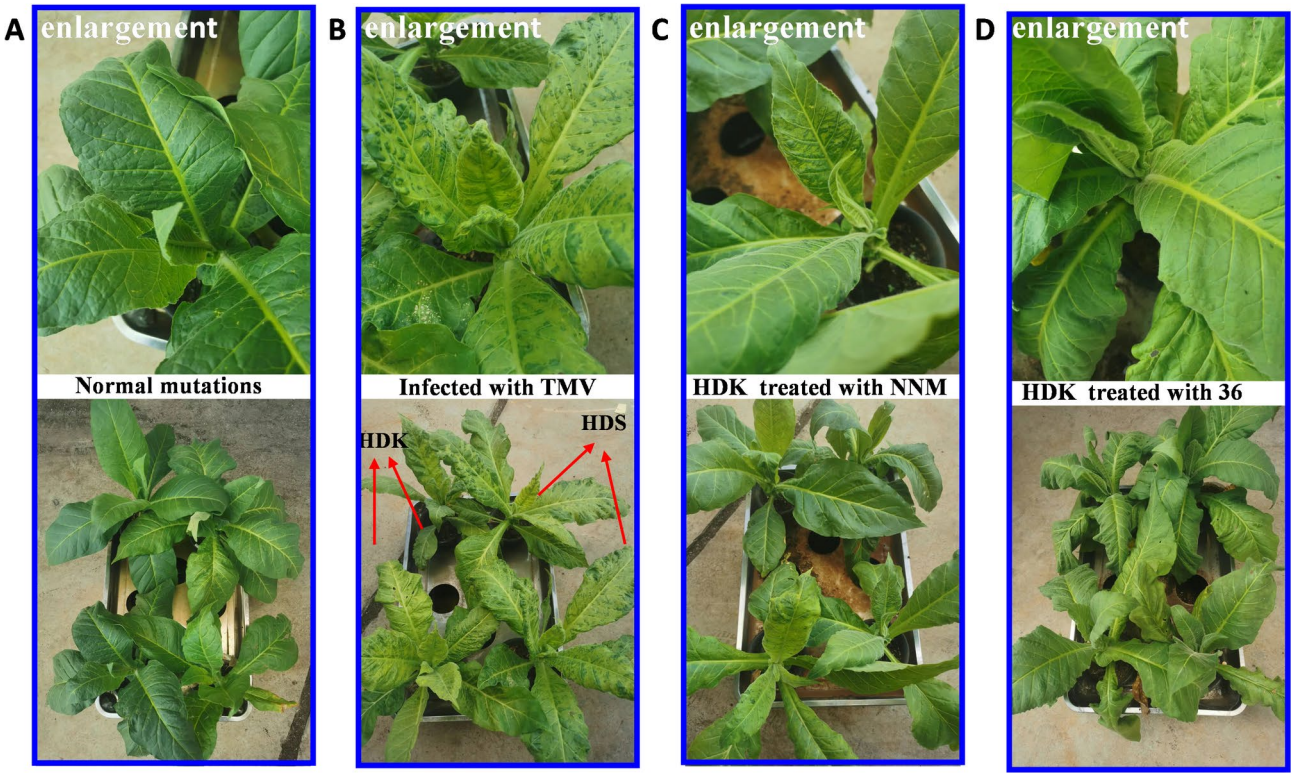

Fig. 8. The curative effect of compound **36** on TMV after Photos were taken 20 days after treatment (down). (**A**) Normal mutations, (**B**) HDK and HDS at the 7–8 leaf stage infected with TMV to be treated, (**C**) HDK infection of TMV treated with 100.0 μg/mL ningnanmycin (NNM), (**D**) HDK infection of TMV treated with compound **36**.

ASP116, and ALA86 residues (Fig. 6D). Compound **36**, the hydroxyl group on the ligand backbone formed hydrogen bond with ALA86 residue (Fig. 6E).

This research inevitably has some inherent limitations. First, the chemical and pharmacological basis for *NtSPS1* knockout on tobacco is still to be fully understood. In addition, the leaves of *N. tabacum* are also not fully used due to the absence of chemical and pharmacological supports. In this investigation, *NtSPS1* silencing significantly downregulated solanesol content (decreased to 25% compared with wild type control) in tobacco leaves. As well as we used LC-MS/MS-based widely targeted metabolomics to understand terpenoid content of mutant and WT plants at the mature stage. But the precision and currency of database information need to be enhanced, given that the predictive results of network pharmacology heavily depend on comprehensive bioinformatics data. Moreover, although network pharmacology mainly concentrates on investigating the interactions between natural products and biological systems, it does not completely account for the intricate and holistic nature of *N. tabacum*. In addition, molecular docking is primarily used to simulate drug-target interactions, but *N. tabacum* often involve multiple targets and therapeutic pathways. Moreover, our study has only identified the key targets of terpenoids from *N. tabacum* in TMV treatment, hinting at the biological mechanisms and important pathways without further exploration.

Conclusion

48 terpenoids, including six up-regulated and five down-regulated terpenoids compared with the wild type leaves control. We carried out an extensive investigation into the use of terpenoids for anti-TMV infections. Employing a combination of network pharmacology and molecular docking techniques, this exhaustive study sought to thoroughly comprehend the medicinal properties of *N. tabacum*. In the future, guided by network pharmacology and molecular docking, compounds **16**, **18**, **23**, and **27** were found to have moderate anti-TMV activities. Furthermore, compound **36** exhibited potential anti-TMV activity, with inhibition rates in the range of 36.7% (IC_{50} = 32.1 μM, Table 1). This detailed analysis of gene knockout mutants and LC-MS/MS analysis will offer new insights into the chemical differences between the mutants and the wild type (WT), as well as their extensive application.

In this study, we investigated the effect of *NtSPS1* knockout in HD on the content of terpenoids and the anti-TMV activity of this mutant. This provides valuable insight into the regulation of terpenoids, specifically solanesol, biosynthesis in tobacco. In recent years, research on terpenoid metabolism has been accelerated by advancements in synthetic biology. Advancements have led to engineering platforms for high-value terpenoid production, but engineering terpenoids in HD still face challenges due to complex metabolic and regulatory networks. Due to the limited number of standard substances, we did not obtain more comprehensive information

about the correlation between *NTSPS1* and other terpenoid content. In further research, we will isolate and identify more compounds from *NtSPS1* knockout HD to search for relevant connections of the knockout and WT plants.

Data availability

The datasets generated and/or analyzed during the current study are available in the PDB data of 2OM3 (https://www.rcsb.org/structure/2OM3), Full wwPDB EM Validation Report of 2OM3 (https://files.rcsb.org/pub/pdb/validation_reports/om/2om3/2om3_full_validation.pdf). The datasets used and/or analyzed during the present study are available from the corresponding author on reasonable request.

Received: 27 October 2024; Accepted: 14 April 2025
Published online: 25 April 2025

References

1. Yan, B. W. et al. Extraction, purification and anti-TMV effects of α (β)-2,7,11-cembratriene-4,6-diol from tobacco leaves. *Ind. Crop Prod.* **174**, 114197 (2021).
2. Nagai, A., Yamamoto, T. & Wariishi, H. Identification of fructo- and malto-oligosaccharides in cured tobacco leaves (*Nicotiana tabacum*). *J. Agric. Food Chem.* **60**, 6606–6612 (2012).
3. Yang, G. Y. et al. Isoindolin-1-ones from the stems of *Nicotiana tabacum* and their antiviral activities. *Arch. Pharm. Res.* **45**, 572–583 (2022).
4. Häkkinen, S. T. et al. Differential patterns of dehydroabietic acid biotransformation by *Nicotiana tabacum* and *Catharanthus roseus* cells. *J. Biotechnol.* **157**, 287–294 (2012).
5. Jassbi, A. R., Zare, S., Asadollahi, M. & Schuman, M. C. Ecological roles and biological activities of specialized metabolites from the genus *Nicotiana*. *Chem. Rev.* **117**, 12227–12280 (2017).
6. Dewey, R. E. & Xie, J. H. Molecular genetics of alkaloid biosynthesis in *Nicotiana tabacum*. *Phytochemistry* **94**, 10–27 (2013).
7. Dai, J. M. et al. The anti-TMV potency of the tobacco-derived fungus *Aspergillus versicolor* and its active alkaloids, as anti-TMV activity inhibitors. *Phytochemistry* **205**, 113485 (2023).
8. Yang, G. Y. et al. Cyclopiazonic acid type Indole alkaloids from *Nicotiana tabacum*-derived fungus *Aspergillus versicolor* and their anti-tobacco mosaic virus activities. *Phytochemistry* **198**, 113137 (2022).
9. Yang, F. X. et al. Hu, Q. F. Isochromenes from the *Nicotiana tabacum*-derived endophytic fungus *Aspergillus versicolor* and their anti-tobacco mosaic virus activities. *Nat. Prod. Res.* **37**, 1608–1616 (2023).
10. Ohnishi, J. et al. The coat protein of tomato mosaic virus L11Y is associated with virus-induced chlorosis on infected tobacco plants. *J. Gen. Plant. Pathol.* **75**, 297–306 (2009).
11. Abdelkhalek, A., Al-Askar, A. A., Arishi, A. A. & Behiry, S. I. Trichoderma Hamatum strain Th23 promotes tomato growth and induces systemic resistance against tobacco mosaic virus. *J. Fungi*. **8**, 228 (2022).
12. Scholthof, K. B. G. et al. Top 10 plant viruses in molecular plant pathology. *Mol. Plant. Pathol.* **12**, 938–954 (2011).
13. Shin, R. et al. Isolation of pepper mRNAs differentially expressed during the hypersensitive response to tobacco mosaic virus and characterization of a proteinase inhibitor gene. *Plant. Sci.* **161**, 727–737 (2001).
14. Zhao, L. H. et al. Anti-TMV activity and functional mechanisms of two sesquiterpenoids isolated from *Tithonia diversifolia*. *Pestic Biochem. Phys.* **140**, 24–29 (2017).
15. Xu, J. et al. Characterization of diterpenoids from *Caesalpinia decapetala* and their anti-TMV activities. *Fitoterapia* **113**, 144–150 (2016).
16. Shang, S. Z. et al. Antiviral sesquiterpenes from leaves of *Nicotiana tabacum*. *Fitoterapia* **108**, 1–4 (2016).
17. Ma, Y. et al. Diterpene alkaloids and diterpenes from *Spiraea japonica* and their anti-tobacco mosaic virus activity. *Fitoterapia* **109**, 8–13 (2016).
18. Yan, N. et al. Solanesol Biosynth. *Plants Mol.* **22**, 510 (2017).
19. Yan, N. et al. Bioactivities and medicinal value of Solanesol and its accumulation, extraction technology, and determination methods. *Biomolecules* **2019**(9), 334 (2019).
20. Yan, N. et al. Solanesol: a review of its resources, derivatives, bioactivities, medicinal applications, and biosynthesis. *Phytochem Rev.* **14**, 403–417 (2015).
21. Lan, T. et al. A simple and standardized method for the determination of total Solanesol in potato leaves and its extracts based on HPLC-MS. *J. AOAC Int.* **104**, 479–484 (2021).
22. Chen, T. et al. Optimization of microwave-assisted extraction of Solanesol from potato leaves and stems. *Med. Chem. Res.* **19**, 732–742 (2010).
23. Taylor, M. A., Fraser, P. D. & Solanesol Added value from *Solanaceous* waste. *Phytochemistry* **72**, 1323–1327 (2011).
24. Akinyemi, I. A., Wang, F., Zhou, B. G., Qi, S. S. & Wu Q. F. Ecogenomic survey of plant viruses infecting tobacco by next generation sequencing. *Virol. J.* **13**, 181 (2016).
25. Bailey, J. A., Carter, G. A., Burden, R. S. & Wain, R. L. Control of rust diseases by diterpenes from *Nicotiana glutinosa*. *Nature* **255**, 328–329 (1975).
26. Campbell, R., Freitag, S., Bryan, G. J., Stewart, D. & Taylor, M. A. Environmental and genetic factors associated with Solanesol accumulation in potato leaves. *Front. Plant. Sci.* **7**, 1263 (2016).
27. Xiang, D. H. et al. Analysis on Solanesol content and genetic diversity of Chinese flue-cured tobacco (*Nicotiana tabacum* L). *Crop Sci.* **57**, 847–855 (2017).
28. Yan, N. et al. RNA sequencing provides insights into the regulation of Solanesol biosynthesis in *Nicotiana tabacum* induced by moderately high temperature. *Biomolecules* **8**, 165 (2018).
29. Yan, N. et al. Effects of *NtSPS1* overexpression on Solanesol content, plant growth, photosynthesis, and metabolome of *Nicotiana tabacum*. *Plants* **9**, 518 (2020).
30. Yan, N. et al. Organ- and growing stage-specific expression of Solanesol biosynthesis genes in *Nicotiana tabacum* reveals their association with Solanesol content. *Molecules* **21**, 1536 (2016).
31. Cong, S., Feng, Y. & Tang, H. P. Network Pharmacology and molecular Docking to explore the potential mechanism of urolithin A in combined allergic rhinitis and asthma syndrome. *Naunyn-Schmiedeberg's Arch. Pharmacol.* **396**, 2165–2177 (2023).
32. Zhao, C. J., Li, C. Y. & Zu, Y. G. Rapid and quantitative determination of Solanesol in *Nicotiana tabacum* by liquid chromatography–tandem mass spectrometry. *J. Pharm. Biomed. Anal.* **44**, 35–40 (2007).
33. Li, S. P. et al. Metabolic and transcriptomic analyses reveal different metabolite biosynthesis profiles between leaf buds and mature leaves in *Ziziphus jujuba* mill. *Food Chem.* **347**, 129005 (2021).
34. Paul Shannon; Andrew Markiel; Owen Ozier, Nitin, S., Baliga; Jonathan, T. & Wang Daniel Ramage; Nada Amin; Benno Schwikowski; Ideker, T. Cytoscape: A software environment for integrated models of biomolecular interaction networks. *Genome Res.* **13**, 2498–2504 (2003).

35. Otasek, D., Morris, J. H., Bouças, J., Pico, A. R. & Demchak, B. Cytoscape automation: empowering workflow-based network analysis. *Genome Biol.* **20**, 185 (2019).
36. Stelzer, G. et al. The genecards suite: from gene data mining to disease genome sequence analyses. *Curr. Protoc. Bioinf.* **54**, 1301–13033 (2016).
37. Amberger, J. S. & Hamosh, A. Searching online mendelian inheritance in man (OMIM): A knowledgebase of human genes and genetic phenotypes. *Curr. Protoc. Bioinf.* **58**, 1.2.1–1.2.12 (2017).
38. Szklarczyk, D. et al. The STRING database in 2017: quality-controlled protein–protein association networks, made broadly accessible. *Nucleic Acids Res.* **45**, D362–D368 (2017).
39. Yang, F. X. et al. Extraction and characterization of anti-virus anthraquinones from *Nicotiana tabacum*-derived *Aspergillus oryzae* YNCA1220. *Pestic Biochem. Physiol.* **196**, 105613 (2023).
40. Ma, Y. Y. et al. X. Indole alkaloids isolated from the *Nicotiana tabacum*-derived *Aspergillus* fumigatus 0338 as potential inhibitors for tobacco powdery mildew and their mode of actions. *Pestic Biochem. Physiol.* **200**, 105814 (2024).
41. Liu, H. Y. et al. Characterization of anti-TMV Indole alkaloid and isocoumarin derivatives from *Aspergillus* versicolor YNCA0363. *Chem. Biol. Technol. Agric.* **10**, 138 (2023).
42. Trott, O., Olson, A. J., AutoDock & Vina Improving the speed and accuracy of Docking with a new scoring function, efficient optimization, and multithreading. *J. Comput. Chem.* **31**, 455–461 (2010).
43. Zhang, J. D. et al. Highly efficient transgene-free genome editing in tobacco using an optimized CRISPR/Cas9 system, pOREU3TR. *Plant. Sci.* **326**, 111523 (2023).

Author contributions

JL, MLC, and QFH conceived and designed the project; JMD performed the experiment of T0 mutant plants; XL, LFZ, JDZ, and JW performed the experiment of T1 and T2 mutant plants; YX and JL performed the experiment of LC-MS/MS analysis; GYY completed the biological evaluation; JMD, XL, JDZ, JW and GYY analyzed the experimental data; JMD wrote the initial draft; All the authors discussed the results and contributed to the preparation of the final paper.

Funding

The authors are grateful to the project of Yunnan Daguan Laboratory (No. YNDG202301YZ04 and No. YNDG202401YZ01), the key projects of China Tobacco Yunnan Industrial (No. 2022JY02 and No. 2023YJ05) and the Project from Yunnan Key Laboratory of Tobacco Chemistry (No. 2021539200340239) for the financial support. The authors declare that they have no competing interests with the funding company.

Declarations

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/s41598-025-98745-y.

Correspondence and requests for materials should be addressed to J.L., M.-L.C. or Q.-F.H.