Title

NtML1, an HD-Zip IV transcription factor, regulates long-stalked glandular trichome development in *Nicotiana tabacum*.

Short title

NtML1 regulates long-stalked trichome formation in tobacco.

Authors

Louis Gueuning[1#], Alice Berhin[1#], Gabriel Walckiers[1], Maxime Jacques Laurent[1], Aldana Ramirez[1], Belkacem El Amraoui[1] and Charles Hachez[1*]

[#] *These two authors contributed equally to this work*

[1] Louvain Institute of Biomolecular Science and Technology, UCLouvain, Louvain-la-Neuve, Belgium

*Corresponding author, charles.hachez@uclouvain.be

Emails:

Louis Gueuning: louis.gueuning@uclouvain.be
Alice Berhin: alice.berhin@uclouvain.be
Gabriel Walckiers: gabriel.walckiers@uclouvain.be
Maxime Jacques Laurent : laurent.maxime@outlook.com
Aldana Ramirez: ramirezdana@yahoo.com
Belkacem El Amraoui : belkacem.elamraoui@uclouvain.be
Charles Hachez : charles.hachez@uclouvain.be

ORCID:
Louis Gueuning: 0009-0001-7095-434X
Alice Berhin: 0000-0002-5214-6533
Gabriel Walckiers: 0009-0002-5693-9052
Maxime Jacques Laurent 0009-0001-1364-8223
Aldana Ramirez: 0000-0002-0332-0531
Belkacem El Amraoui : 0009-0004-25158100
Charles Hachez: 0000-0002-3688-7614

Submission date: July 16, 2025

0 table, 7 Fig., 4607 words, 7 Supplemental Fig. and 5 Supplemental Tables.

Highlights

Glandular trichomes are hair-like structures that contribute to plant defense. This study identifies NtML1, an HD-Zip IV transcription factor, as a key regulator of long-stalked glandular trichome development in tobacco.

Abstract

Glandular trichomes are specialized epidermal structures that contribute to plant defense by secreting protective specialized metabolites. In this study, we investigated the functional role of homeobox-leucine zipper protein MERISTEM L1-like (NtML1), a class IV HD-Zip transcription factor, in regulating glandular trichome development in *Nicotiana tabacum*. Using an integrated approach combining transcriptomics, reverse genetics, functional genomics, and protein interaction assays, we found that NtML1 is essential for the formation of long-stalked glandular trichomes. Loss-of-function mutation in *NtML1* leads to reduced long-stalked glandular trichome density and altered metabolite profiles, underscoring its regulatory role in both trichome initiation and secondary metabolism. Furthermore, NtML1 interacts with the R2R3-MYB transcription factor REGULATORY AXILLARY MERSITEM 2 (NtRAX2), jointly modulating trichome developmental genes. RNA-seq analysis uncovered the extensive transcriptional network influenced by *NtML1*, offering novel insights into the molecular mechanisms underlying glandular trichome formation. These findings may inform future biotechnological approaches aimed at improving pest resistance and enhancing metabolite production in *Solanaceae* crops.

Keywords: Epidermal Differentiation, Glandular trichomes, *Nicotiana tabacum*, NtML1, NtRAX2, Specialized metabolites.

Abbreviations:

Long-stalked glandular trichomes – LGT

Non-glandular trichome – NGT

Short-stalked glandular trichome – SGT

Yeast Two-Hybrid - Y2H

Bimolecular fluorescence complementation - BiFC

DNA Affinity Purification sequencing - DAP-seq

Electrophoretic mobility shift assay – EMSA

Differentially expressed genes – DEGs

Jasmonic acid - JA

Introduction

Plants have evolved an array of specialized epidermal structures, including trichomes, to enhance their adaptability to biotic and abiotic stresses (Agrawal, 2007; Wagner *et al.*, 2004). Trichomes are hair-like outgrowths that develop on aerial organs and are categorized into unicellular or multicellular forms, as well as glandular or non-glandular types (Schuurink and Tissier, 2020). Non-glandular trichomes primarily provide physical defenses, forming a dense barrier that deters herbivores, mitigates excessive transpiration and enhances tolerance to environmental stresses. In addition to similar protective roles, glandular trichomes also act as biochemical factories, synthesizing, storing or secreting a wide variety of specialized metabolites, including terpenoids, flavonoids and phenylpropanoids (Huchelmann *et al.*, 2017; Tian *et al.*, 2017). These metabolites play a crucial role in plant defense against herbivores and pathogens, facilitate plant-pollinator interactions, and have extensive applications in agriculture, medicine and industry (Glas *et al.*, 2012; Huchelmann *et al.*, 2017; Tian *et al.*, 2017; Tissier, 2012).

Glandular trichomes exhibit remarkable morphological and functional diversity, typically categorized into peltate and capitate trichomes based on their structure and mode of secretion. Peltate trichomes possess a broad glandular head atop a short stalk, accumulating secretory products in a subcuticular space before release, whereas capitate trichomes have a smaller glandular head on a longer stalk and actively extrude metabolites to the external environment (Glas *et al.*, 2012). Environmental cues, including light intensity, temperature, and water availability, significantly influence glandular trichome density and metabolic composition, modulating the plant defensive capabilities (Yan *et al.*, 2016).

Despite their ecological significance, the molecular mechanisms regulating glandular trichome differentiation and specialization remain only partially understood. Compared to the well-characterized non-glandular trichome regulatory network in *Arabidopsis thaliana*, glandular trichome development in *Solanaceae* species such as tomato (*Solanum lycopersicum*) and tobacco (*Nicotiana tabacum*) follows distinct genetic pathways that involve a complex interplay of transcription factors, hormonal signals and cell cycle regulators (Chalvin *et al.*, 2020). Advances in transcriptomic and genetic studies have identified some key regulatory components, shedding light on the intricate mechanisms governing glandular trichome initiation, patterning and functional specialization.

In *N. tabacum* leaves, three distinct trichome types are visible: non-glandular trichomes (NGTs), long- and short-stalked glandular trichomes (LGTs and SGTs respectively). A few genes have been identified as involved in LGT development in *N. tabacum*: *NtHD9* controls glandular head formation, while *NtHD12* influences both initiation and gland formation (Zhang *et al.*, 2023), and *NtCYCB2* is a negative regulator of their formation (Wang *et al.*, 2022).

Much of the current understanding of the transcriptional network governing glandular trichome development originates from studies in *Solanum lycopersicum* where eight distinct types of trichomes have been identified. Trichomes of types I to V are classified as digitate, with types I and IV being glandular and capable of secreting metabolites, whereas types II, III, and V are non-glandular. Types VI and VII are peltate glandular trichomes, distinguished by their multicellular secretory heads and short stalks (Wu *et al.*, 2023). A key regulator in this species is the *Woolly* (*SlWOOLLY*) gene, which encodes an HD-Zip IV transcription factor essential for the formation of Type I and Type VI glandular trichomes (Yang *et al.*, 2011a). SlWOOLLY orchestrates spatially polarized cell division and cell fate determination through its interactions with multiple genetic regulators (Wu *et al.*, 2023). High *SlWOOLLY* expression promotes apical cell division via activation of SlWOX3B, a WUSCHEL-related homeobox transcription factor, and SlMX1, a R2R3-MYB transcription factor essential for trichome initiation (Wu *et al.*, 2023). Conversely, in basal cells, SlWOOLLY activates SlBRANCHED2a (SlBRC2a), a transcription factor that mediates the transition from cell division to expansion through cytokinin signaling. SlBRC2a interacts with SlWOOLLY to inhibit further cell division, triggering endoreduplication and stalk elongation in multicellular trichomes (Wu *et al.*, 2024). SlWOOLLY also forms regulatory complexes with zinc finger proteins to modulate trichome development. ZINC FINGER PROTEIN 8 (SlZFP8), a key regulator of type I trichome formation, functions as a heterodimer with SlWOOLLY (Chang *et al.*, 2018). A homolog, SlZFP8-like (SlZFP8L), similarly forms complexes with SlWOOLLY; its overexpression enhances trichome density and elongation across multiple trichome types (Zheng *et al.*, 2022). Together, SlWOOLLY and SlZFP8L also regulate *SlZFP6* expression, further refining the network controlling trichome initiation and elongation (Zheng *et al.*, 2022).

Beyond its role in trichome morphogenesis, SlWOOLLY is a central component of the regulatory network controlling metabolite biosynthesis within glandular cells. It interacts with SlMYC1, a bHLH transcription factor, to activate terpene biosynthesis genes in a process

tightly regulated by jasmonic acid (JA) signaling (Xu *et al.*, 2018). SlJAZ2, a JA signaling repressor, binds to SlWOOLLY and SlMYC1 to suppress their activity. Upon JA perception, SlJAZ2 is degraded, relieving this repression and allowing SlWOOLLY and SlMYC1 to activate terpene production (Hua *et al.*, 2021; Xu *et al.*, 2018). Hormonal crosstalk extends beyond JA signaling; SlbHLH95 negatively regulates trichome initiation, particularly of type I trichomes on stems, by modulating gibberellin biosynthesis (Chen *et al.*, 2020).

The integration of SlWOOLLY into multiple transcriptional and hormonal pathways underscores the complexity of glandular trichome development and highlights the intricate balance between genetic regulation and environmental signaling.

Similarly, in *Nicotiana benthamiana*, the HD-Zip transcription factor NbWOOLLY, homologous to SlWOOLLY, is a positive regulator of trichome development through *NbCYCB2* regulation and feedback loop (Wu *et al.*, 2020).

Our work focuses on homeobox-leucine zipper protein MERISTEM L1-like (NtML1), a member of the HD-Zip IV family and a close homolog of SlWOOLLY from *S. lycopersicum* (84% protein identity) and of AtPDF2 from *Arabidopsis thaliana* (76% protein identity), this latter being specifically expressed in the outermost cell layer (L1) of shoot meristems (Abe *et al.*, 2003). In *Arabidopsis*, the double mutant of *AtPDF2* and its paralog *AtML1* displays severe defects in shoot epidermal cell differentiation, indicating that both genes are essential for maintaining L1 cell identity and proper differentiation (Abe *et al.*, 2003).

In this study, we investigated the role of NtML1 in LGT development in *N. tabacum*, with a particular focus on its interaction with transcription factor REGULATORY AXILLARY MERSITEM 2-like (NtRAX2), an R2R3 MYB transcription factor. Our findings shed light on the molecular mechanisms governing glandular trichome development and underscore the central role of NtML1 and its potential protein partners in this process. Investigating these interactions will not only enhance our understanding of trichome biology in tobacco but may also offer insights transferable to other *Solanaceae* species, enabling targeted manipulation of trichome-related traits. Given the agricultural and industrial relevance of glandular trichome-derived metabolites, advancing our knowledge of the genetic and molecular pathways controlling glandular trichome development holds strong promise for crop improvement, pest resistance, and metabolic engineering.

Materials and Methods

Generation of constructs

For the transcriptional reporter lines, promoter sequences of *NtML1* (*LOC107766155*) and *NtRAX2-like* (LOC107813213) (defined as the 3 kb region upstream of the initiating ATG) were amplified by PCR from genomic DNA (Supplemental Table S1) and inserted into the pDONR P4P1r via a Gateway using BP Clonase™ II (Invitrogen). *pNtMl1::nlsGFP-GUS* and *pNtRAX2::nlsGFP-GUS* were generated by recombining *pENTRY L4-pNtML1-R1* or *pENTRY L4-pNtRAX2-R1* respectively with *pENTRY L1-nlsGFPGUS-L2* (Berhin *et al.*, 2019) into the *pH7m34GW* (Karimi *et al.*, 2005).

For the *NtML1* and *NtRAX2* CRISPR-Cas9, one guide RNA (Supplemental Table S2) was chosen for each. It was cloned by restriction digest using bbs1 and ligated into *pSF539* (U6-26gRNA) to form *P2P3r-AtU6-tRNA-gRNA-scaffold*. This plasmid, *pENTRY-L4-pUBQ10-L1* and *pENTRY-CAS9-tagRFP-T35S* (Wang *et al.*, 2020) were transferred to the *pH7m34GW* plasmid via an LR Gateway reaction. For the *NtRAX2* overexpression lines, the *NtRAX2* coding sequence was amplified by PCR from cDNA generated based on leaf primordia mRNA (Supplemental Table S1) and recombined into a *pDONR221* vector to create the *pENTRY L1-NtRAX2-L2. p35S::NtRAX2* was produced by LR Gateway reaction recombining the corresponding entry clone into the *pH7WG2D* plasmid (Karimi *et al.*, 2002).

For yeast two-hybrid assays, coding sequences of *NtML1* and *NtRAX2* were PCR-amplified using primers with BsaI site overhangs (Supplemental Table S1) and cloned into *CK011* and *CK012*, containing respectively the GAL4 activation and binding domain, under the *ADH1* promoter (Gantner *et al.*, 2018), using Golden Gate cloning technology (Kirchmaier *et al.*, 2013). *CK011-nlsGFP* and *CK012-nlsGFP* constructs were used as a control (Berhin *et al.*, 2025).

For the BIFC, coding sequences of *NtML1* and *NtRAX2* were amplified (Supplemental Table S1) and cloned via BP Gateway reaction into a *pUC57-R1R4* and *pUC57-R3R2* respectively. Then they were transferred together into the *pBiFCt-2in1-NN* expression vector via LR reaction (Grefen and Blatt, 2012). In this plasmid, each CDS is under the control of the *35S* promoter and is fused to a truncated YFP (either the C-terminus or the N-terminus) at its N-terminal part.

For protein expression in the DAP-seq assay, the coding sequences of *NtML1* and *NtRAX2* were amplified by PCR from cDNA (Supplemental Table S1) and recombined into the *pDONR221* vector to generate the entry clones *pENTRY L1-NtML1-L2* and *pENTRY L1-NtRAX2-L2*. These entry clones were then individually introduced into the *pIX-HALO* expression vector via LR recombination. *pIX-HALO-nlsGFP* construct was used as control (Berhin *et al.*, 2025).

For the dual luciferase assay, promoters identified in DAP-seq were amplified from *N. tabacum* genomic DNA (Supplemental Table S1) and inserted into the *pDONR221* entry vector before being recombined into the *pGWL7* plasmid via Gateway cloning. *pENTRY L1-NtML1-L2* and *pENTRY L1-NtRAX2-L2* were also both independently transferred to the *p2GW7* plasmid by LR Gateway cloning (Karimi *et al.*, 2002).

For EMSA, the CDS of *NtML1* or *NtRAX2-like* were separately introduced in the *pQE-60* expression vector (Qiagen) via digestion/ligation (Supplemental Table S1). This vector contains the inducible promoter driving heterologous protein expression in *E. coli*. This plasmid was introduced in *E. coli* BL21(DE3)

Gene accession

All genes mentioned in the manuscript are referred to by their official gene identifiers (LOC) upon first mention in the text. These LOC identifiers may correspond to different gene names depending on the annotation database. For consistency, we refer to genes and proteins using their UniProt names throughout the manuscript. The raw data from RNA-seq and DAP-seq analyses retain the original annotations from Sierro et al. (2014) as the LOC identifiers remain consistent across datasets.

Plant Material and Growth Conditions

N. tabacum cv Petit Havana SR1 (Maliga 1973) was used as the wild-type (WT) plant material. For in vitro cultures, seeds were sterilized by exposure to chlorine gas for 30 minutes, produced by mixing 37% HCl with bleach (1:5 v:v). Sterilized seeds were then placed on solid Murashige and Skoog (MS) medium [4.4 g/l MS basal medium (MP Biochemicals), 30 g/l sucrose, 8 g/l agar, pH 5.6 (KOH)] and cultured at 25°C with a 16-hour photoperiod (35 µmol photons/m^2s). After one week, seedlings were transferred to Jiffy pots for a week, followed by transplantation into potting soil (DCM) and grown in a phytotron at 25°C with a 16-hour photoperiod (200 µmol photons/m^2s). All constructs in destination vectors were introduced into *Agrobacterium tumefaciens* strain GV3101pmp90.

Transient Plant Transformation

Tobacco plants were transiently transformed by classical leaf infiltration using *Agrobacterium tumefaciens* strain. Infiltrated areas were harvested after 72 hours for analysis.

Stable Plant Transformation

After centrifugation (8000 RPM, 2 min), the pellet of the overnight culture of *Agrobacterium tumefaciens* was washed twice in 1 ml resuspension solution [10 mM $MgSO_4$, 1 mM

acetosyringone], resuspended in 300 µl solution, and incubated in the dark for 2 hours at room temperature. Four-week-old *N. tabacum* leaf explants (1 cm^2) were pricked four times with a syringe needle dipped in *A. tumefaciens* suspension under sterile conditions and incubated for 5 days at 25°C with a 16-hour photoperiod. After rinsing with liquid MS medium, explants were placed on regeneration medium [solid MS medium, 500 mg/l cefotaxime, 400 mg/l carbenicillin, 0.2 mg/l Indole-3-Acetic Acid (IAA), 2.2 mg/l 6-Benzylaminopurine (BAP), 25 mg/l hygromycin] for 5–6 weeks until calli formed. Calli were transferred to shoot regeneration medium (without IAA), followed by rooting medium (without BAP). Regenerated seedlings (T0) were transplanted into Jiffy pots, then transferred to potting soil in a phytotron to produce T1 seeds.

Tissue Harvesting

For gene expression profiling, tissues for RNA extraction were collected from various plant organs, including primary and secondary roots, leaf primordia, expanding leaves (approximately 10 cm), mature leaves (approximately 20 cm), leaf trichomes, mature leaves without trichomes, stems, petals, and sepals. Leaf trichomes were scraped off using the protocol of Wang et al. (2001). Tissues were harvested from plants at 6 weeks old (approximately 20 cm tall), with four biological replicates per tissue type. Samples were stored at -80°C prior to RNA extraction.

RNA Extraction and cDNA Synthesis

RNA was extracted following the protocol of the Spectrum™ Plant Total RNA Kit (Sigma-Aldrich, protocol A), and genomic DNA was removed using the On-Column DNase I Digestion set (Sigma-Aldrich). Reverse transcription was performed using Moloney Murine Leukemia Virus Reverse Transcriptase (Promega) using Oligo dT(18) primers according to the supplier's protocol.

RT-qPCR

Primers used for qPCR are listed in Supplemental Table S1. The qPCR was performed using the qPCR master mix plus for SYBR assay ROX (Eurogentec, RT-SN2X-06-+).

Relative expression levels were calculated using the $2^{-\Delta\Delta CT}$ method (Schmittgen and Livak, 2008) with the geometric mean of three reference genes: *Elongation factor α (NtEF1a*,), *Actin (NtTac9*), and *Ubiquitin-conjugating enzyme E2 (NtUbc2*).

GUS Staining Assay

For GUS staining, organs (roots, leaves, primordia, stems, petals, sepals) from transgenic lines were placed in a cell culture plate with acetone (90%) for 20 minutes at -20°C. The organs were then transferred to a GUS staining solution [50 mM Na_2HPO_4/NaH_2PO_4 pH 7.2, 2 mM Ferricyanide, 2 mM Ferrocyanide, 1% Triton X-100, and 1 mM X-Gluc] and incubated at 37°C overnight in the dark. After incubation, the organs were washed in ethanol solutions (70%, 80%, 90%, and 100%) for 20 minutes each with gentle shaking. The organs were observed under a stereomicroscope (Zeiss Axio Imager) with visible light.

Confocal Imaging

To observe GFP expression, organs were imaged using a confocal fluorescence microscope (LSM710, Zeiss) equipped with a spectral detector. Excitation was performed at 488 nm, and emission was detected between 493 and 540 nm, corresponding to the GFP emission spectrum.

For the BIFC, after transient transformation of the BIFC plasmid, leaf epidermis (abaxial side) was observed by confocal fluorescence microscopy (LSM710 confocal microscope equipped with a spectral detector, Zeiss) with laser set at 514nm (YFP) and 561nm (mCherry) for the excitation wavelength. The detection wavelength was set to range from 520nm to 544nm for YFP and from 583nm to 635nm for mCherry.

Yeast Two Hybrid Assay

The yeast two-hybrid assay followed Paiano *et al.* (2019), using the *pJ696* yeast strain (Lafuente *et al.*, 2000). Briefly, a single yeast colony was inoculated in 50 mL of YPDA medium and grown overnight at 30°C with shaking (250 rpm). The culture was diluted to an OD_{600} of 0.2–0.3 in fresh YPDA and grown to mid-log phase ($OD_{600} \approx 0.5$). Cells were harvested (5 min at 1000 g), washed, and resuspended in 1.5 ml of TE/LiAc buffer (10 mM Tris-HCl, 1 mM EDTA, 100 mM lithium acetate, pH 7.5). For each transformation, 1 µg of bait plasmid, 1 µg of prey plasmid, and 10 µL of denatured carrier DNA were mixed with 100 µL of competent yeast cells. After addition of 750 µL of PEG/LiAc solution, the mixture was incubated at 30°C for 30 min, followed by a 20 min heat shock at 42°C in the presence of 100 µL DMSO. Cells were harvested (5 min at 1000 g) and resuspended in TE solution before being plated on SD/–WL selective medium to select for co-transformants. After 3 days of incubation at 30°C, three independent colonies were picked and resuspended in TE buffer. Each suspension was spotted onto: SD/-Leu/-Trp and SD/-Leu/-Trp/-Ade/-His/+AT-3 (5 mM) plates to assess activation of the HIS3 and ADE2 reporter genes, indicative of weak and strong interactions, respectively. To confirm

specificity, each prey plasmid was also co-transformed with a bait vector containing nlsGFP as a negative control.

Transgenic Lines Validation

KO lines were validated through genomic characterization using PCR and sequencing of the targeted gene region, with the TIDE software (Brinkman *et al.*, 2014).

Trichome density

Trichome quantification was performed using a Zeiss Observer Z1 epifluorescence microscope equipped with GFP, DsRed, and DAPI filter sets. From each of four independent plants, three circular sections were excised from the central region of leaves measuring 12–15 cm in length. The samples were fixed in ethanol and sequentially stained: overnight with Calcofluor White/Fluorescent Brightener 28 (VWR, ICNA0215806701), followed by 1 hour with 0.01% Fluorol Yellow 088 (Santa Cruz Biotechnology, 81-37-8), and then 1 hour with 0.5% Rhodamine B (Merck, 7559).

Metabolite quantification

One 10 cm long leaf was harvested per biological replicate (6-week-old) and immersed in 20ml of dichloromethane (for 5 sec each time, 4 times per leaf). 1-heptadecanol and sucrose acetate was added to each sample as internal controls. Samples were derivatied with 25µL of N,N-dimethylformamide dimethyl acetal (DMF) and 25µL of N,O-Bis(trimetylsilyl)trifluoroacetamide (BSTFA)

Gas chromatography coupled with a flame ionization detector (FID) (Trace 1300 Gas Chromatograph, Thermo Scientific) was used to analyses metabolites extracted from N. tabacum leaves. The injector temperature was 300°C. Injection volume was 1µl in split mode. Compounds were separated with Restek Rxi 5 Sil MS (30m length, 0.25mm diameter, 0.25µm film thickness). Helium was used as carrier gas and set at a constant flow rate of 1ml/min. Oven initial temperature was 60°C and ramped to 210°C at 20°C/min, then ramped to 230°C at 1°C/min and finally ramped to 300°C at 20°C/min and kept this final temperature until the end of the run (51min). For FID, temperature was 300°C, sampling rate 10Hz, dihydrogen flow was 35ml/min and air flow was 350ml/min.

DNA Affinity Purification Sequencing (DAP-seq)

The DNA affinity purification sequencing (DAP-seq) experiment was conducted following the protocol outlined by Bartlett *et al.* (2017). Three independent DNA libraries were generated from genomic DNA (gDNA) extracted from leaf primordia of *N. tabacum* plants NtML1-HALO-

tagged was expressed in triplicate using the TnT® Coupled Wheat Germ Extract System (Promega, L41030). Bound genomic DNA fragments were prepared and ligated with adapters and indices using the NEBNext® Multiplex Oligos for Illumina® (NEB, E6440S). The resulting DNA-seq libraries were sequenced by Macrogen on the Illumina NovaSeq 6000 platform, generating a minimum of 50 million 150 bp paired-end reads per sample. Sequencing reads were processed and mapped; Peak calling was performed using the Transcription Factor ChIP-seq tool within the CLC Genomics Workbench, comparing each sample to the negative control (HALO-nlsGFP). Significant peaks were identified based on a peak score threshold of >1.5 and a p-value <0.05

Dual Luciferase Assay

Protoplasts were prepared from leaves of 5-week-old *N. tabacum* plants. The epidermis of the abaxial side were incubated overnight at 25°C in the dark in a digestion solution containing 5 mM $CaCl_2$, 0.5 M sucrose, 1 g/L BSA, 1.25 g/L macerozyme Onozuka R-10, 2 g/L cellulase Onozuka R-10, pH 5.5. After digestion, the leaf pieces were removed, and the solution was diluted with an ML06 solution (15 mM $CaCl_2$, 0.6 M sucrose, 1.5 g/L MES, pH 6). The protoplast solution was filtered through a sterile 100 µm cell strainer and centrifuged at 700 RPM for 7 minutes. The upper phase, containing viable protoplasts, was collected, diluted with MEL solution, and centrifuged again. The protoplast pellet was resuspended in MEL solution for counting using a Brand Burker chamber. The protoplast concentration was adjusted to 100,000 protoplasts per transformation and transferred to a 2 mL Eppendorf tube. After centrifugation at 100 g for 2 minutes, the supernatant was removed, and the protoplasts were resuspended in 100 µL of transfection solution (12.5 µM $CaCl_2$, 109.3 g/L mannitol, 1 g/L MES, pH 5.7). Three pmol of each plasmid was added, mixed, and incubated for 5 minutes. Next, 110 µL of PEG solution (54.5 g/L mannitol, 0.1 M $Ca(NO_3)_2$, 400 g/L PEG 4000) was added, and the mixture was incubated for 2 minutes. Transformation was terminated by adding 440 µL of W5 buffer (0.15 M NaCl, 0.16 M $CaCl_2$, 5 mM KCl, 0.43 g/L MES, pH 5.7), followed by centrifugation at 100 g for 1 minute. The transformed protoplasts were resuspended in 1 mL of culture solution (0.04 M KCl, 109.3 g/L mannitol, 0.085 g/L MES, 90 mg/L cefotaxime, pH 5.7) and incubated overnight at room temperature in the dark. Following incubation, culture medium was removed by centrifugation (3 min at 100 g). The protoplasts were lysed using the passive lysis buffer from the Dual-Luciferase Reporter Assay kit (Promega). Luciferase activities were assessed by measuring luminescence intensity, following the manufacturer's

protocol, using a GloMax 20/20 luminometer (Promega). For promoter activity assessment, the *p2GW7* empty vector was used as a negative control. Each transformation was performed in triplicate. The ratio of firefly luciferase activity to Renilla luciferase activity was calculated to normalize the promoter activity and reduce experimental variability.

RNA-seq

RNA-seq was performed to identify differentially expressed genes (DEGs) between conditions. Total RNA was extracted from leaf primordia, and libraries were prepared using the TruSeq Stranded Total RNA Library Prep Kit (Illumina). Paired-end reads (150 bp) were generated on the Illumina HiSeq 2500 platform, with 50 million reads per sample. Data analysis was performed using the nf-core RNA-seq pipeline (Ewels *et al.*, 2020). Reads were trimmed using Trim Galore (Babraham Bioinformatics) and aligned with STAR (Dobin *et al.*, 2013) to the *N. tabacum* genome assembly (Ntab-TN90) (Sierro *et al.*, 2014). Gene expression levels were estimated as Transcripts Per Million (TPM) using Salmon (Patro *et al.*, 2017), and DEGs were identified using DESeq2 (Love *et al.*, 2014). Statistical significance was determined using an adjusted p-value (q-value) below an FDR threshold of 0.05. Gene Ontology (GO) enrichment analysis was conducted using DAVID software (Sherman *et al.*, 2022).

Electrophoretic mobility shift assay

The transformed cells were cultured at 30°C to OD600 0.3–0.4 and induced with isopropyl β-D-1-thiogalactopyranoside (IPTG) [0.5mM] for 5 hours. Cells were then harvested by centrifugation and lysed in a lysis solution [50mM Tris-HCl pH8, 0.5M NaCl, 30mM imidazole, 0.5mM DTT, 1 mM PMSF, 10mg/ml lysozyme, and protease inhibitors] by sonication (Ultrasonic Processor Vibra Cell 75022). Centrifugation was performed to remove the cellular debris (pellet). Proteins were purified using a Ni-NTA column (HisPur Ni-NTA Resin, Thermo Scientific). After three washes with the washing solution [10 mM Tris-HCl pH7.5, 0.1M NaCl, 30/40/50mM imidazole, 10% glycerol], proteins were eluted in the elution solution [10 mM Tris-HCl pH7.5, 0.1M NaCl, 300mM imidazole, 10% glycerol]. Complementary oligonucleotides encoding the MEME motif determined previously were synthesized and annealed, with one oligo tagged with Cyanine3 (Supplemental Table S3). The double-stranded DNA [5nM] was then incubated with the purified protein at different concentration [0 to 30ng/µl] in a binding solution [25mM Tris-HCl pH 7.5, 9% glycerol, 75mM NaCl, 5 mM MgCl2 5mM DTT,1mg/mL BSA] for 30min at room temperature in the dark. The solution was put in an 8% polyacrylamide

gel for electrophoresis (2h at 50mA in a cold room 4°C). Gel was visualized using a Typhoon scanner (Cytiva).

Results

NtML1 is specifically expressed in epidermal tissues of aerial organs

To determine the spatial and temporal expression pattern of the homeobox-leucine zipper protein MERISTEM L1-like (*NtML1*; LOC107766155), we conducted RT-qPCR analyses across a comprehensive range of *N. tabacum* tissues, encompassing both vegetative and reproductive organs. *NtML1* expression peak in leaves (Fig. 1A) is consistent with the developmental pattern of its *A. thaliana* orthologs, *AtML1* and *AtPDF2*, which regulate L1 cell identity and differentiation (Abe *et al.*, 2003; Takada *et al.*, 2013). Similar expression level was also observed in sepals and petals, organs which are also covered in trichomes (Supplemental Fig. S1). Trichomes from mature leaves showed a lower *NtML1* expression level than in leaf primordia.

We complemented this transcriptional profiling with histochemical and fluorescence-based reporter assays, using a transcriptional fusion of the native *NtML1* promoter with a nuclear-localized GFP-GUS reporter (*pNtML1::nlsGFP-GUS*). In seedlings and young plants, strong GUS staining was observed exclusively in emerging leaf primordia, consistent with *NtML1* expected role in L1 layer identity specification (Supplemental Fig. S2). As leaves expanded, GUS signal persisted in the epidermis of both adaxial and abaxial surfaces, but gradually diminished in mature leaves, paralleling the RT-qPCR findings. No GFP or GUS signal was detected in roots of the reporter lines suggesting a shoot-specific and epidermally localized expression profile. Confocal imaging of GFP fluorescence confirmed, with cell-specific resolution, that NtML1 expression is restricted to epidermal derivatives, including pavement cells, stomatal guard cells, and various trichome types (Fig. 1B). GFP signal was clearly detected in developing long glandular trichomes (LGTs), short glandular trichomes (SGTs), and non-glandular trichomes (NGTs) in early leaf primordia (Fig. 1B), indicating that *NtML1* is not only expressed in the epidermis broadly, but is also active in specialized epidermal structures during their development. Expression persisted in the glandular cells of mature LGTs (Fig. 1B), suggesting a potential functional role in gland maintenance or metabolite production beyond initial cell fate specification.

CRISPR-Cas9 knockout of *NtML1* results in loss of multicellular trichomes

To investigate the functional role of *NtML1*, we used CRISPR-Cas9 genome editing to generate loss-of-function mutants. Since *Nicotiana tabacum* is an allotetraploid species, it contains two homeologs of the *NtML1* gene—*NtML1_1* (LOC107766155) and *NtML1_2* (LOC107785796). A single guide RNA (gRNA) targeting a conserved region within the first exon of both homeologs was designed to induce targeted mutations. Stable transgenic *NtML1* knockout (KO) lines were generated and confirmed by Sanger sequencing and TIDE analysis (Brinkman et al., 2014). Genotyping revealed several kinds of biallelic edits in both homeologs, all resulting in frameshift mutations that introduced premature stop codons within the HD-Zip domain, thereby disrupting gene function (Fig. 2A-B).

Morphologically, *ntml1* mutant plants appeared normal in overall stature and leaf morphology under standard growth conditions, indicating that *NtML1* is not essential for general shoot development. However, detailed examination of leaf epidermal surfaces revealed a striking loss of multicellular trichomes in *ntml1* knockout lines (Fig. 3A). The long glandular trichomes (LGTs), which are typically abundant on the surface of leaves, were nearly absent in mutant plants (Fig. 3B). Similarly, non-glandular trichomes (NGTs), were drastically reduced (Fig. 3B). By contrast, the density of short glandular trichomes (SGTs) was unaffected (Fig. 3B), suggesting that *NtML1* is specifically required for the initiation and differentiation of LGTs but not of SGTs.

Pavement cell density and stomatal indices remained unchanged, indicating that the reduction in trichome number was not due to altered epidermal cell proliferation or leaf expansion (Supplemental Fig. S3). These data strongly support a trichome-specific role for NtML1 in multicellular trichome initiation or development.

Secondary metabolite production is compromised in *ntml1* mutant leaves

Given the known role of glandular trichomes in the biosynthesis and secretion of specialized metabolites, we next investigated whether loss of trichomes in *ntml1* mutants impacted chemical composition of the leaf surface. Surface metabolite extractions from leaves of WT and *ntml1* mutants were subjected to quantitative GC-FID profiling. We observed significant reductions in total diterpenoid content, including α-cembrenediol, a hallmark tobacco trichome diterpene (Fig. 3C). Acyl sugar levels, another class of trichome-derived specialized metabolites, were also significantly reduced. These chemical deficits correlated with the near absence of LGTs in the same lines.

Transcriptomic profiling reveals widespread misregulation of epidermal and floral genes in *ntml1* mutants

To explore the broader molecular consequences of *NtML1* loss, we performed RNA-seq analysis on mRNA isolated from leaf primordia of wild-type and *ntml1* plants. RNA-seq analysis yielded high-confidence alignments to the *N. tabacum* reference genome (Sierro *et al.*, 2014). Differential expression analysis identified 201 genes significantly upregulated and 78 downregulated in *ntml1* mutants (adjusted $p < 0.05$, $|\log2FC| > 1$) (Fig. 4A, Supplemental Fig. S4 and Supplemental Table S4).

Notably, *NtML1* was among the most significantly downregulated genes, suggesting that it may regulate its own expression. Another possible explanation for the reduction in *NtML1* mRNA levels in the mutant lines involves the Nonsense-Mediated mRNA Decay (NMD) pathway. NMD is a eukaryotic quality-control mechanism that regulates the stability of both aberrant and normal transcripts by targeting faulty mRNAs for degradation during translation (Lykke-Andersen and Jensen, 2015). This process leads to decreased levels of aberrant mRNAs within the cell. Therefore, the observed reduction in *NtML1* transcript abundance in the *ntml1* lines may result from a combination of disrupted positive feedback and NMD-triggered degradation of the mutant transcript.

Other canonical trichome development regulators, including the homeobox leucine zipper protein GLABRA 2 (NtGL2, LOC107819990 and LOC107815348), homeobox-leucine zipper HDG11 (NtHDG11, LOC107770385 and LOC107767107) and G2/mitotic-specific cyclin-2-like (NtCYCB2*,* LOC107804360 and LOC107767057*)* did not show differential expression (Supplemental Table S4), suggesting that *NtML1* operates either downstream or independently from these pathways. Gene ontology enrichment analyses revealed overrepresentation of terms related to the regulation of gene transcription and sterol metabolic process, consistent with its role in promoting glandular trichome development and an involvement in secondary metabolism (e.g. biosynthesis of terpenoid compounds from sterol precursors) (Fig. 4B).

Of interest, zinc finger transcription factor and MADS-box genes were particularly affected, suggesting that *NtML1* may act as a transcriptional hub integrating developmental cues (Fig. 4C, Supplemental Table S4).

NtML1 Physically Interacts with R2R3-MYB Transcription Factors

Having established a clear role of NtML1 in leaf primordia epidermal cell differentiation, we hypothesized that NtML1 could form functional complexes with R2R3-MYB transcription factors to initiate trichome development. This hypothesis draws support from similar interactions observed in other species, such as GbML1-GbMYB25 in cotton (Zhang *et al.*, 2010), AaHD8-AaMIXTA in *Artemisia annua* (Yan *et al.*, 2018), and SlWOOLLY-SlMYB31 in tomato (Xiong *et al.*, 2020), where HD-Zip IV and R2R3-MYB factors co-regulate trichome development.

We identified the R2R3-MYB transcription factor REGULATORY AXILLARY MERISTEM 2-like (NtRAX2, LOC107813213), as a potential NtML1 partner. Using RT-qPCR, we investigated *NtRAX2* expression pattern (Fig. 1A). In trichomes, *NtRAX2* exhibits significantly higher transcript levels compared to leaf primordia, developing leaves, and mature leaves, where expression levels remain relatively similar. (Fig. 1A). To complement those data, reporter lines expressing the *pRAX2::nlsGFP-GUS* construct were generated. Promoter reporter assays with GUS showed that *NtRAX2* promoter is specifically active in developing and mature trichomes (Supplemental Fig. S2). GFP-driven confocal imaging confirmed that *NtRAX2* expression is localized to the stalk cells of LGTs, where it persists from early development through maturity (Fig. 1B). In NGTs and SGTs, *NtRAX2* is active in all trichome cells. Interestingly, promoter activity was also observed in stomatal guard cells, but not in pavement cells. This expression pattern, which partially overlaps with that of *NtML1*, is compatible with a potential role for NtRAX2 as a partner or co-regulator of NtML1 in specific epidermal cell types.

To explore the role of *NtRAX2*, we employed CRISPR-Cas9 to disrupt both homeologs, *NtRAX2_1* (LOC107813213) and *NtRAX2_2* (LOC107809127), using a single common guide RNA. The resulting loss-of-function mutants were validated through Sanger sequencing and TIDE analysis (Supplemental Fig. S5A) (Brinkman *et al.*, 2014). In addition, we generated *NtRAX2* overexpression lines (*p35S::NtRAX2*), confirmed by RT-qPCR (Supplemental Fig. S6A), to complement the knockout approach and assess potential gain-of-function effects. The analysis of KO and overexpression lines focused on quantifying three trichome types: LGT, NGT, and SGT. Based on the trichome index (number of trichomes per pavement cell), neither the knockout nor the overexpression lines showed significant differences compared to the wild type (WT) (Supplemental Fig. S5B–C and Supplemental Fig. S6B–C). However, the LGT/NGT ratio was significantly reduced in the overexpression lines, suggesting that *NtRAX2* may influence trichome specialization into glandular or non-glandular rather than overall

trichome density (Supplemental Fig. S6C). The absence of a phenotype in the KO lines could indicate functional redundancy within the *NtRAX* gene family, implying that a higher order mutant may be required to reveal a clear phenotype.

NtML1 and NtRAX2 physically interact

To test protein-protein interaction, we performed Yeast Two-Hybrid (Y2H) assays that demonstrated NtML1 capability to interact with NtRAX2 (Fig. 5A). The NtML1–NtRAX2 interaction supported yeast growth on selective medium, indicating strong activation of the HIS3 reporter gene. Those results were confirmed in *N. tabacum* leaves using bimolecular fluorescence complementation (BiFC) assays. Reconstitution of YFP fluorescence, indicating physical interaction between proteins, was observed only for the NtML1–NtRAX2 pair (Fig. 5B). This interaction was localized specifically to the nucleus, consistent with the predicted roles of both proteins as transcriptional regulators.

Together, this suggests that NtML1 acts in concert with NtRAX2 through direct interaction in the nucleus. These results uncover a novel HD-Zip IV–MYB interaction in tobacco that may underlie key transcriptional programs driving epidermal specialization and metabolic differentiation.

Promoter Activation and Functional Synergy between NtML1 and NtRAX2

To uncover the direct targets of NtML1, DNA Affinity Purification sequencing (DAP-seq) analysis was performed using genomic DNA from leaf primordia. Sequencing reads were processed and aligned to the *N. tabacum* reference genome (Sierro *et al.*, 2014). This approach provided a genome-wide view of binding site preferences with 79.3% of the detected intragenic interactions found in genomic regions located upstream of the START codon, primarily within promoters and 5’ untranslated regions (UTRs), as well as within intronic sequences (Fig. 6A).

NtML1 showed widespread binding, targeting over 10,000 genes (peak shape score>1.5 and a p-value<0.01) (Supplemental Table S5), including known trichome regulators such as *NtGL2 (LOC107819990)*, the ZINC FINGER PROTEIN 8 (*NtZFP8, LOC107828820)*, the TRANSPARENT TESTA GLABRA-1-like (*NtTTG1, LOC107803204)*, *NtRAX2* and *NtML1* itself (Supplemental Fig. S7).

MEME and TomTom analyses (Bailey *et al.*, 2015; Gupta *et al.*, 2007) confirmed the presence of a conserved binding motif overlapping the L1 box (TAAATG(C/T)A), known to mediate L1-

specific gene expression (Fig. 6B). This highlights a transcriptional cascade potentially driven by autoregulatory feedback and L1 box motif recognition.

To validate these findings, an electrophoretic mobility shift assay (EMSA) was performed using a 50 bp Cy3-labeled probe centered on the DAP-seq peak located in *NtZFP8* promoter region and containing a L1 motif. Increasing NtML1 concentrations produced a shift in the DNA band (Fig. 6C), confirming protein–DNA complex formation. To assess binding specificity, a second probe containing the L1 motif only yielded a similar shift (Fig. 6D). Finally, mutating two bases within the L1 motif abolished binding (Fig. 6E), demonstrating sequence specificity. Together, these results confirm that NtML1 directly and specifically binds the L1 box, offering molecular insight into its role in trichome gene activation.

Dual luciferase assays were conducted to validate DAP-seq predictions, focusing on promoters of genes linked to epidermal development. *NtZFP8* and *NtML1* promoter activation by NtML1 was confirmed, supporting its role as a transcriptional activator and its ability to self-regulate via a positive feedback loop (Fig. 6F). Notably, co-expression of NtRAX2 with NtML1 further enhanced activation of the *NtZFP8* promoter, indicating a synergistic interaction likely mediated by their physical association. This cooperative regulation supports a model where the NtML1/NtRAX2 complex promotes LGT initiation through direct activation of *NtZFP8* in trichome precursor cells (Fig. 6F).

To assess the functional relevance of NtML1 targets, we compared genes bound by NtML1 in DAP-seq with DEGs from RNA-seq analysis of *ntml1* mutants. 30 genes overlapped between datasets (Fig. 7A), suggesting that many DEGs are indirectly regulated and that NtML1 is able to bind to some genes *in vitro* but not *in vivo.* Among these 30 shared genes, six encode transcription factors, some of which are implicated in trichome and/or floral development: developmental protein SEPALLATA 1-like (LOC107826782), B3 domain-containing transcription factor NGA1-like (LOC107829012), protein SRG1 (LOC107766747), WD repeat-containing protein 82 (LOC107818263) and NtML1 itself. Several genes encoding RNA-binding proteins (Protein MEI2-like 7, (LOC107759864); RNA-binding protein 42-like, (LOC107776982)) and enzymes involved in secondary metabolism (Beta-amyrin 28-oxidase-like (LOC107804010); cis-abienol synthase (LOC107830721)) were also direct transcriptional targets (Fig. 7B). These findings highlight a small but functionally diverse subset of direct targets, pointing to an integrated regulation of epidermal fate and metabolic pathways by NtML1.

The limited overlap between DAP-seq and RNA-seq results may stem from factors such as cell-type-specific expression of transcription factors, altered protein conformation during in vitro expression, and the absence of chromatin context in DAP-seq assays (Mahendrawada *et al.*, 2025). In addition, DAP-seq data should be interpreted with caution and complemented by functional assays and expression analyses to reliably identify biologically relevant targets, such as those validated through dual luciferase assays.

These results highlight NtML1 as a central regulator of epidermal fate and trichome initiation, with NtRAX2 potentially acting on a more restricted, possibly co-regulated gene subset.

Discussion

This study identifies NtML1, a member of the HD-Zip IV transcription factor family in *N. tabacum*, as a fundamental regulator of epidermal development, particularly in multicellular trichome formation. NtML1 expands the known roles of HD-Zip IV genes beyond general epidermal specification, highlighting its involvement in specialized epidermal structures development. Such insights establish NtML1 as a crucial developmental switch interfacing both structural and metabolic aspects of epidermal cell fate.

Role of *NtML1* in Multicellular Trichome Initiation

Prior research in *Arabidopsis thaliana* has elucidated that HD-Zip IV family members, such as *AtML1* and *AtPDF2*, function redundantly to specify the L1 layer identity and encourage epidermal differentiation throughout embryogenesis and organogenesis (Abe *et al.*, 2003 ; Takada *et al.*, 2013). Comparable roles for orthologs of *ML1-like* genes have been recognized in other plant species; for instance, *GhHD1* in cotton (Walford *et al.*, 2012), and *AaHD8* in *Artemisia annua* (Yan *et al.*, 2018).

Our mutant analysis and expression studies demonstrate that *NtML1* is essential for the formation of both LGTs and NGTs in tobacco (Fig. 1A-B, Fig. 2, Fig. 3A-B, Supplemental Fig. S1, S2, S3). Interestingly, SGTs remained unaffected in *ntml1* mutants, suggesting the presence of distinct regulatory networks governing specific trichome types (Tian *et al.*, 2020). Given that *SlWOOLLY*, the closest homolog of *NtML1* in tomato, has already been shown to induce the formation of type IV glandular trichomes when overexpressed (Wu *et al.*, 2023; Yang *et al.*, 2011b), it is reasonable to hypothesize that overexpression of *NtML1* in tobacco may similarly trigger the development of supernumerary long glandular trichomes.

Notably, the absence of differential expression of well-known trichome regulators such as *NtGL2, NtHDG11* and *NtCYCB2* in our transcriptomic data (Fig. 4, Supplemental Table S4)

suggests that NtML1 may act downstream or in a parallel pathway, regulating a distinct branch of the trichome developmental network. This points to a broader and more intricate regulatory framework in multicellular trichome-bearing species like tobacco, where epidermal identity and trichome fate are shaped by multiple, converging transcriptional circuits (Wu *et al.*, 2023; Yang *et al.*, 2011a).

HD-Zip IV-MYB Module in Epidermal Specialization

Despite their functional significance, the genetic mechanisms controlling trichome subtype diversity remain relatively underexplored. Emerging evidence indicates that specific combinations of HD-Zip IV, MYB, and WOX transcription factors direct the differentiation of trichome types, such as type I and type VI in tomato (Wu *et al.*, 2023; Yang *et al.*, 2011a). Building on those interactions, our findings support a similar modular organization in tobacco, where NtML1 likely governs a subset of trichome types, potentially through interactions with cell-specific MYB partners like NtRAX2. In the L1 layer, *NtRAX2* is expressed in trichome precursor cells and developing trichomes, its expression persisting in the stalk cells of long glandular trichomes throughout development (Fig. 1A-B, Supplemental Fig. S1–S2). The physical interaction between NtML1 and the R2R3-MYB transcription factor NtRAX2 confirmed by Y2H and BIFC (Fig. 5A-B) supports the well-established model in which HD-Zip IV and MYB transcription factors cooperatively regulate epidermal cell fate (Chen and Wang, 2019). This mechanism is further supported by findings in other species, including GbML1–GbMYB25 in cotton (Zhang *et al.*, 2010) and AaHD8–AaMIXTA1 in *Artemisia annua* (Yan *et al.*, 2018); all of which demonstrate the importance of HD-Zip IV–MYB complexes in trichome initiation and elongation.

We sought to better understand the functional role of NtRAX2 through expression profiling and reverse genetics. The expression pattern of NtRAX2 is compatible with its involvement in the development of LGTs and NGTs, similar to NtML1 (Fig. 1A, Supplemental Fig. S1, S2). The elevated *NtRAX2* transcript levels observed in the overexpression lines resulted in a shift in the ratio of LGTs to NGTs, with a noticeable reduction in LGTs upon overexpression. This observation suggests that *NtRAX2* may not be directly involved in initiating trichome formation, but rather in guiding their specialization. The increased prevalence of NGTs could reflect a default developmental pathway that emerges in the absence of signals promoting glandular identity.

Thus, NtRAX2 may act as a modulator of trichome fate, influencing the balance between glandular and non-glandular trichome differentiation. Interestingly, this dose-dependent effect is reminiscent of the regulatory behavior of SlWOOLLY, which promotes digitate trichomes at high expression levels and peltate glandular trichomes at lower levels (Wu *et al.*, 2024). Given the homology between SlWOOLLY and NtML1, and the reported interaction between NtRAX2 and NtML1, a similar regulatory mechanism may operate in tobacco. We hypothesize that NtRAX2 contributes to trichome differentiation in a dose-dependent manner, potentially influencing the developmental trajectory by forming a transcription factor complex with NtML1. At high expression levels, NtRAX2 may titrate NtML1 activity, thereby favoring the formation of nonglandular trichomes.

To test this hypothesis, we characterized NtRAX2 knock-out mutants. Surprisingly, no distinct trichome phenotype was observed, suggesting possible functional redundancy among NtRAX2 family members (Supplemental Fig. S5). This lack of phenotype contrasts with findings in *Cucumis sativ*us, where CsRAX2 was downregulated in the *mict* mutant, which displays severe trichome defects (Zhao *et al.*, 2015), supporting a role for RAX2 homologs in trichome development. These observations imply that RAX2-related transcription factors may influence trichome initiation or specialization in a species-specific and possibly context-dependent manner. Further functional validation, including combinatorial knock-outs or overexpression of other NtRAX family members, will be essential to fully elucidate the role of NtRAX2 in trichome development and its potential interaction with NtML1.

NtML1-NtRAX2, a complex regulating *NtZFP8* expression

NtML1 exhibits extensive genome-wide binding, targeting over 10,000 genes including key trichome regulators such as *NtGL2*, *NtZFP8*, *NtTTG1, NtRAX2* and itself (Fig. 6A, 6C-D-E, Supplemental Fig. S7, Supplemental Table S5). Motif analysis revealed a conserved L1 box at many binding sites (Fig. 6B), supporting L1-specific transcriptional regulation and autoregulatory feedback. EMSA assays confirmed that NtML1 directly and specifically binds the L1 motif, notably within the *NtZFP8* promoter (Fig 6C–E). Dual luciferase assays validated NtML1 role as a transcriptional activator, confirming its activation of both *NtZFP8* and its own promoter (Fig. 6F). Notably, co-expression with NtRAX2 significantly enhanced *NtZFP8* promoter activation, indicating a synergistic interaction likely mediated by physical association in trichome precursor cells. Together, these findings establish NtML1 as a central

regulator of epidermal fate and trichome initiation, with NtRAX2 acting as a co-regulator that fine-tunes activation of targets like *NtZFP8* (Fig. 6F).

Connection Between Structural Development and Metabolic Output

We found a pronounced reduction in trichome-derived metabolites, including labdanoids, cembranoids and acyl sugars, in *ntml1* knockout lines (Fig. 3C). These metabolites play key roles in fungal and herbivore resistance, UV protection, and inter-plant signaling (Chien *et al.*, 2022; Kennedy, 2003; Liu *et al.*, 2024; Vendemiatti *et al.*, 2024). The observed reduction is likely due to the absence of trichomes and thus secretory cells or the inability of residual glandular structures to activate key biosynthetic pathways. In addition to this, DAP-seq and RNA-seq analyses revealed that NtML1 directly and indirectly targets several genes encoding enzymes involved in secondary metabolism, such as beta-amyrin oxidase and cis-abienol synthase (Fig. 7). These findings point to an integrated regulatory role of NtML1 in both epidermal fate determination and specialized metabolite biosynthesis.

Implications for Metabolic Engineering and Crop Improvement

From a practical standpoint, identifying NtML1 as a crucial coordinator of trichome development and specialized metabolite production heralds new opportunities for biotechnology applications in tobacco and related *Solanaceae* crops. By manipulating the expression of *NtML1* or its interaction partners, strategies may emerge to enhance trichome density or increase the yield of trichome-derived metabolites. This has tremendous relevance for bioengineering efforts aimed at producing pharmaceuticals, deterrents to pests or value-added aromatic compounds within trichomes (Gossart *et al.*, 2023).

In conclusion, our research shows that NtML1 serves as a key developmental regulator, orchestrating multicellular trichome development in *N. tabacum*. Through its epidermal specificity, the ability to regulate different trichome type ratios, its physical interactions with MYB partners, and its impact on both anatomical and metabolic outcomes, NtML1 emerges as a critical control point. These revelations enhance our comprehension of trichome biology and open promising pathways for metabolic engineering and trait optimization in economically significant crop species.

<u>Supplementary data</u>

Supplemental Fig. S1. Global expression patterns of *NtML1* and *NtRAX2.*

Supplemental Fig. S2. GUS histochemical analysis of transcriptional reporter lines.

Supplemental Fig. S3. Epidermal cell-type regulation by *NtML1.*

Supplemental Fig. S4. RNA-seq analysis of WT and *ntml1* lines.

Supplemental Fig. S5. Characterization of the *NtRAX2* CRISPR mutant lines

Supplemental Fig. S6. Characterization of the *NtRAX2* overexpression lines

Supplemental Fig. S7. NtML1 binding near trichome development-related genes.

Supplemental Table 1 : Primer list used for any PCR.

Supplemental Table 2 : gRNA used for CRISPR-Cas9 induced KO.

Supplemental Table 3 : Oligonucleotide list used for EMSA.

Supplemental Table 4: RNA-seq

Supplemental Table 5 : DAP-seq

Acknowledgments

We are grateful to Dr. Robertas Ursache from CRAG (Barcelona, Spain) and Dr. Satoshi Fujita from LRSV (Toulouse, France) for providing the pSF539 plasmid. We thank Manon Peeters for reading the manuscript.

Author contributions

LG and AB designed, performed the experiments and analyzed the results. GW, MJL and AR performed some experiments/data analysis. BEA offered essential support in analyzing the RNA-seq and DAP-seq data. CH conceived and supervised the whole project. LG, AB and CH drafted the manuscript.

Conflict of interest

The authors declare that there are no competing interests.

Funding

This work was supported by the Belgian National Fund for Scientific Research (PDR/PGY grant #T.0116.20 to LG, AB, AR and CH, Charge de recherche #FC46171 to AB, FRIA #FC 58225 to GW). LG benefitted from funding from the UCLouvain Special Research Funds (UCLouvain FSR 2020), AB from the Swiss National Science Foundation (FNS; Grant#P2LAP3_191271) and M.J.L. from the Fonds pour la Formation à la Recherche dans l'Industrie et dans l'Agriculture (FRIA, grant#1.E.086.21F).

Data availability

RNA-seq and DAP-seq data sets are MIAME compliant and were submitted to the Gene Expression Omnibus (GEO) database under the GEO Series accession number GSE297065 and GSE297067.

References:

Abe M, Katsumata H, Komeda Y, Takahashi T. 2003. Regulation of shoot epidermal cell differentiation by a pair of homeodomain proteins in Arabidopsis. Development **130**, 635-643.

Agrawal AA. 2007. Macroevolution of plant defense strategies. Trends in Ecology & Evolution **22**, 103-109.

Bailey TL, Johnson J, Grant CE, Noble WS. 2015. The MEME Suite. Nucleic Acids Research **43**, W39-W49.

Bartlett A, O'Malley RC, Huang S-sC, Galli M, Nery JR, Gallavotti A, Ecker JR. 2017. Mapping genome-wide transcription-factor binding sites using DAP-seq. Nature Protocols **12**, 1659-1672.

Berhin A, de Bellis D, Franke RB, Buono RA, Nowack MK, Nawrath C. 2019. The Root Cap Cuticle: A Cell Wall Structure for Seedling Establishment and Lateral Root Formation. Cell **176**, 1367-1378.e1368.

Berhin A, Ramirez A, Peeters M, Walckiers G, Vannieuwenhuyze M, Legay S, El Amraoui B, Hachez C. 2025. Insight into the genetic network governing long-stalked glandular trichome development in *Nicotiana tabacum*. Planta **262**, 123.

Brinkman EK, Chen T, Amendola M, van Steensel B. 2014. Easy quantitative assessment of genome editing by sequence trace decomposition. Nucleic Acids Research **42**, e168-e168.

Chalvin C, Drevensek S, Dron M, Bendahmane A, Boualem A. 2020. Genetic Control of Glandular Trichome Development. Trends in Plant Science **25**, 477-487.

Chang J, Yu T, Yang Q, Li C, Xiong C, Gao S, Xie Q, Zheng F, Li H, Tian Z, Yang C, Ye Z. 2018. Hair, encoding a single C2H2 zinc-finger protein, regulates multicellular trichome formation in tomato. Plant Journal **96**, 90-102.

Chen S, Wang S. 2019. GLABRA2, a Common Regulator for Epidermal Cell Fate Determination and Anthocyanin Biosynthesis in Arabidopsis. International Journal of Molecular Sciences **20**, 4997.

Chen Y, Su D, Li J, Ying S, Deng H, He X, Zhu Y, Li Y, Chen Y, Pirrello J, Bouzayen M, Liu Y, Liu M. 2020. Overexpression of bHLH95, a basic helix–loop–helix transcription factor family member, impacts trichome formation via regulating gibberellin biosynthesis in tomato. Journal of Experimental Botany **71**, 3450-3462.

Chien Y-H, Yu Y-H, Ye S-R, Chen Y-W. 2022. Antibacterial and Antioxidant Activity of the Fruit of Macaranga tanarius, the Plant Origin of Taiwanese Green Propolis. Antioxidants **11**, 1242.

Dobin A, Davis CA, Schlesinger F, Drenkow J, Zaleski C, Jha S, Batut P, Chaisson M, Gingeras TR. 2013. STAR: ultrafast universal RNA-seq aligner. Bioinformatics **29**, 15-21.

Ewels PA, Peltzer A, Fillinger S, Patel H, Alneberg J, Wilm A, Garcia MU, Di Tommaso P, Nahnsen S. 2020. The nf-core framework for community-curated bioinformatics pipelines. Nature Biotechnology **38**, 276-278.

Gantner J, Ordon J, Ilse T, Kretschmer C, Gruetzner R, Lofke C, Dagdas Y, Burstenbinder K, Marillonnet S, Stuttmann J. 2018. Peripheral infrastructure vectors and an extended set of plant parts for the Modular Cloning system. PLOS ONE **13**, e0197185.

Glas J, Schimmel B, Alba J, Escobar-Bravo R, Schuurink R, Kant M. 2012. Plant Glandular Trichomes as Targets for Breeding or Engineering of Resistance to Herbivores. International Journal Molecular Sciences **13**, 17077-17103.

Gossart N, Berhin A, Sergeant K, Alam I, André C, Hausman J-F, Boutry M, Hachez C. 2023. Engineering Nicotiana tabacum trichomes for triterpenic acid production. Plant Science **328**, 111573.

Grefen C, Blatt MR. 2012. A 2in1 cloning system enables ratiometric bimolecular fluorescence complementation (rBiFC). Biotechniques **53**, 311-314.

Gupta S, Stamatoyannopoulos JA, Bailey TL, Noble WS. 2007. Quantifying similarity between motifs. Genome Biology **8**, R24.

Hua B, Chang J, Xu Z, Han X, Xu M, Yang M, Yang C, Ye Z, Wu S. 2021. HOMEODOMAIN PROTEIN8 mediates jasmonate-triggered trichome elongation in tomato. New Phytologist **230**, 1063-1077.

Huchelmann A, Boutry M, Hachez C. 2017. Plant Glandular Trichomes: Natural Cell Factories of High Biotechnological Interest. Plant Physiology **175**, 6-22.

Karimi M, De Meyer B, Hilson P. 2005. Modular cloning in plant cells. Trends in Plant Science **10**, 103-105.

Karimi M, Inzé D, Depicker A. 2002. GATEWAY™ vectors for Agrobacterium-mediated plant transformation. Trends in Plant Science **7**, 193-195.

Kennedy GG. 2003. Tomato, pests, parasitoids, and predators: tritrophic interactions involving the genus Lycopersicon. Annual Review Entomology **48**, 51-72.

Kirchmaier S, Lust K, Wittbrodt J. 2013. Golden GATEway Cloning – A Combinatorial Approach to Generate Fusion and Recombination Constructs. PLOS ONE **8**, e76117.

Lafuente MJ, Gancedo C, Jauniaux J-C, Gancedo JM. 2000. Mth1 receives the signal given by the glucose sensors Snf3 and Rgt2 in Saccharomyces cerevisiae. Molecular Microbiology **35**, 161-172.

Liu J, Wang J, Du Y, Yan N, Han X, Zhang J, Dou Y, Liu Y. 2024. Application and Evaluation of the Antifungal Activities of Glandular Trichome Secretions from Air/Sun-Cured Tobacco Germplasms against Botrytis cinerea. Plants **13**, 1997.

Love MI, Huber W, Anders S. 2014. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. Genome Biology **15**, 550.

Lykke-Andersen S, Jensen TH. 2015. Nonsense-mediated mRNA decay: an intricate machinery that shapes transcriptomes. Nature Reviews Molecular Cell Biology **16**, 665-677.

Mahendrawada L, Warfield L, Donczew R, Hahn S. 2025. Low overlap of transcription factor DNA binding and regulatory targets. Nature **642**, 796-804.

Paiano A, Margiotta A, De Luca M, Bucci C. 2019. Yeast Two-Hybrid Assay to Identify Interacting Proteins. Current Protocols in Protein Science **95**, e70.

Patro R, Duggal G, Love MI, Irizarry RA, Kingsford C. 2017. Salmon provides fast and bias-aware quantification of transcript expression. Nature Methods **14**, 417-419.

Schmittgen TD, Livak KJ. 2008. Analyzing real-time PCR data by the comparative CT method. Nature Protocols **3**, 1101-1108.

Schuurink R, Tissier A. 2020. Glandular trichomes: micro-organs with model status? New Phytologist **225**, 2251-2266.

Sherman BT, Hao M, Qiu J, Jiao X, Baseler MW, Lane HC, Imamichi T, Chang W. 2022. DAVID: a web server for functional enrichment analysis and functional annotation of gene lists (2021 update). Nucleic Acids Research **50**, W216-W221.

Sierro N, Battey JND, Ouadi S, Bakaher N, Bovet L, Willig A, Goepfert S, Peitsch MC, Ivanov NV. 2014. The tobacco genome sequence and its comparison with those of tomato and potato. Nature Communications **5**, 3833.

Takada S, Takada N, Yoshida A. 2013. ATML1 promotes epidermal cell differentiation in Arabidopsis shoots. Development **140**, 1919-1923.

Tian N, Liu F, Wang P, Zhang X, Li X, Wu G. 2017. The molecular basis of glandular trichome development and secondary metabolism in plants. Plant Gene **12**, 1-12.

Tian YS, Dai TW, Fan CX, Zhou J, Ren XL, Xu Y. 2020. An HD-Zip IV transcription factor protein NbGL3 regulates glandular trichome initiation in tobacco. Biologia plantarum **64**, 378-384.

Tissier A. 2012. Trichome Specific Expression: Promoters and Their Applications. Transgenic Plants, Advances and Limitations, InTech, 353-378

Vendemiatti E, Hernández-De Lira IO, Snijders R, Torne T, Therezan R, Prants G, Lopez-Ortiz C, Reddy U, Bleeker P, Schenck CA, Peres LEP, Benedito VA. 2024. *Woolly* mutation with *Get02* locus overcomes the polygenic nature of trichome-based pest resistance in tomato. Plant Physiology **195**, 911-923.

Wagner GJ, Wang E, Shepherd RW. 2004. New approaches for studying and exploiting an old protuberance, the plant trichome. Annals Bototany **93**, 3-11.

Walford SA, Wu Y, Llewellyn DJ, Dennis ES. 2012. Epidermal cell differentiation in cotton mediated by the homeodomain leucine zipper gene, *GhHD-1*. The Plant Journal **71**, 464-478.

Wang X, Ye L, Lyu M, Ursache R, Löytynoja A, Mähönen AP. 2020. An inducible genome editing system for plants. Nature Plants **6**, 766-772.

Wang Z, Yan X, Zhang H, Meng Y, Pan Y, Cui H. 2022. NtCycB2 negatively regulates tobacco glandular trichome formation, exudate accumulation, and aphid resistance. Plant Molecular Biology **108**, 65-76.

Wu M-L, Cui Y-C, Ge L, Cui L-P, Xu Z-C, Zhang H-Y, Wang Z-J, Zhou D, Wu S, Chen L, Cui H. 2020. NbCycB2 represses Nbwo activity via a negative feedback loop in tobacco trichome development. Journal of Experimental Botany **71**, 1815-1827.

Wu M, Bian X, Hu S, Huang B, Shen J, Du Y, Wang Y, Xu M, Xu H, Yang M, Wu S. 2024. A gradient of the HD-Zip regulator Woolly regulates multicellular trichome morphogenesis in tomato. The Plant Cell **36**, 2375-2392.

Wu M, Chang J, Han X, Shen J, Yang L, Hu S, Huang B-B, Xu H, Xu M, Wu S, Li P, Hua B, Yang M, Yang Z, Wu S. 2023. A HD-ZIP transcription factor specifies fates of multicellular trichomes via dosage-dependent mechanisms in tomato. Developmental Cell **58**, 278-288.e275.

Xiong C, Xie Q, Yang Q, Sun P, Gao S, Li H, Zhang J, Wang T, Ye Z, Yang C. 2020. WOOLLY, interacting with MYB transcription factor MYB31, regulates cuticular wax biosynthesis by modulating CER6 expression in tomato. The Plant Journal **103**, 323-337.

Xu J, van Herwijnen ZO, Drager DB, Sui C, Haring MA, Schuurink RC. 2018. SlMYC1 Regulates Type VI Glandular Trichome Formation and Terpene Biosynthesis in Tomato Glandular Cells. Plant Cell **30**, 2988-3005.

Yan N, Du Y, Liu X, Zhang H, Liu Y, Zhang P, Gong D, Zhang Z. 2016. Chemical structures, biosynthesis, bioactivities, biocatalysis and semisynthesis of tobacco cembranoids: An overview. Industrial Crops and Products **83**, 66-80.

Yan T, Li L, Xie L, Chen M, Shen Q, Pan Q, Fu X, Shi P, Tang Y, Huang H, Huang Y, Huang Y, Tang K. 2018. A novel HD-ZIP IV/MIXTA complex promotes glandular trichome initiation and cuticle development in *Artemisia annua*. New Phytologist **218**, 567-578.

Yang C, Li H, Zhang J, Luo Z, Gong P, Zhang C, Li J, Wang T, Zhang Y, Lu Y, Ye Z. 2011a. A regulatory gene induces trichome formation and embryo lethality in tomato. Proceedings of the National Academy of Sciences **108**, 11836-11841.

Yang C, Li H, Zhang J, Wang T, Ye Z. 2011b. Fine-mapping of the woolly gene controlling multicellular trichome formation and embryonic development in tomato. Theoretical and Applied Genetics **123**, 625-633.

Zhang F, Zuo K, Zhang J, Liu X, Zhang L, Sun X, Tang K. 2010. An L1 box binding protein, GbML1, interacts with GbMYB25 to control cotton fibre development. Journal of Experimental Botany **61**, 3599-3613.

Zhang H, Xu H, Xu M, Yan X, Wang Z, Yu J, Lei B, Cui H. 2023. Transcription factors *NtHD9* and *NtHD12* control long glandular trichome formation via jasmonate signaling. Plant Physiology **191**, 2385-2399.

Zhao J-L, Pan J-S, Guan Y, Nie J-T, Yang J-J, Qu M-L, He H-L, Cai R. 2015. Transcriptome analysis in Cucumis sativus identifies genes involved in multicellular trichome development. Genomics **105**, 296-303.

Zheng F, Cui L, Li C, Xie Q, Ai G, Wang J, Yu H, Wang T, Zhang J, Ye Z, Yang C. 2022. Hair (H) interacts with SlZFP8-like to regulate the initiation and elongation of trichomes by modulating SlZFP6 expression in tomato. Journal of Experimental Botany, **73**, 228-244.

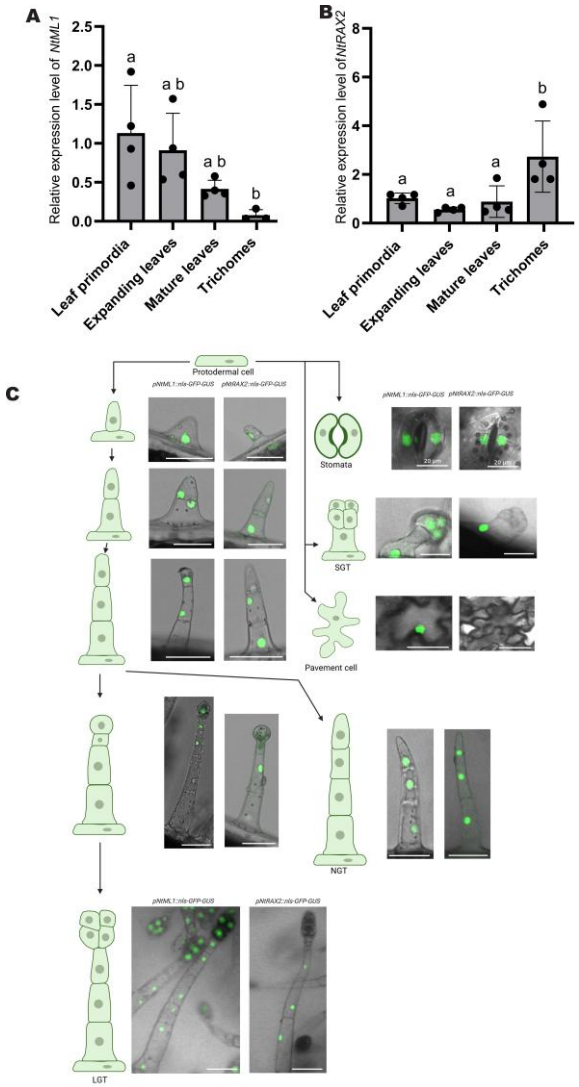

Figure 1 : Expression of *NtML1* and *NtRAX2* in leaf epidermal cells, including trichomes.

(A-B) RT-qPCR analysis of *NtML1* (A) and *NtRAX2* (B) in leaf primordia (<2 cm), expanding leaves (~10 cm), mature leaves (~20 cm), and isolated trichomes (from mature leaves) of *Nicotiana tabacum*. Expression levels are shown relative to those in leaf primordia (set to 1) and normalized to the geometric mean of three reference genes: *NtEF2*, *NtACTIN*, and *NtUBC2*. Relative expression values were calculated using the ΔCt method. Data represent means ± standard deviation from 3–4 biological replicates. Significant differences were determined by one-way ANOVA followed by Tukey's HSD test; different letters denote statistically significant differences between tissues ($p < 0.05$). (C) GFP fluorescence in transgenic tobacco plants expressing *pNtML1::nlsGFP-GUS* and *pNtRAX2::nlsGFP-GUS*, highlighting signal in specific epidermal cell types of leaf primordia (stomata, SGT, pavement cells, trichomes at various developmental stages, LGT, and NGT). Scale bars: 50 µm unless otherwise indicated; green indicates GFP signal. Created in BioRender. Berhin, A. (2025) https://BioRender.com/8u9cw3v.

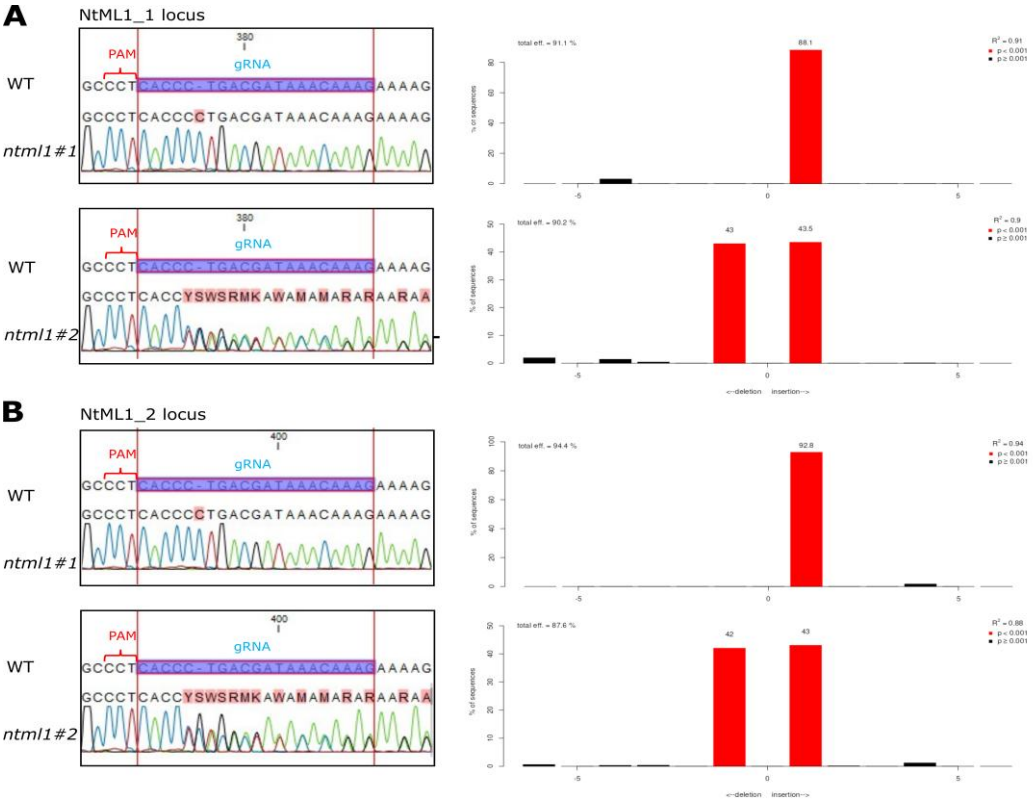

Figure 2: Genotypic characterization of *ntml1* mutants.

Sequencing chromatograms were analyzed and aligned with the wild-type (WT) sequence to assess editing efficiency. (A) Sequencing around the gRNA target site for *NtML1_1* in *ntml1 #1* revealed a single-nucleotide insertion, while *ntml1 #2* displayed a heterogeneous population of edited alleles with either a one-nucleotide insertion or deletion. These mutations caused frameshifts leading to premature stop codons, likely abolishing NtML1_1 protein function and confirming successful gene editing. (B) Similar editing patterns were observed at the *NtML1_2* locus in the same knockout lines. Left panels show sequencing chromatograms; right panels show TIDE deconvolution analysis.

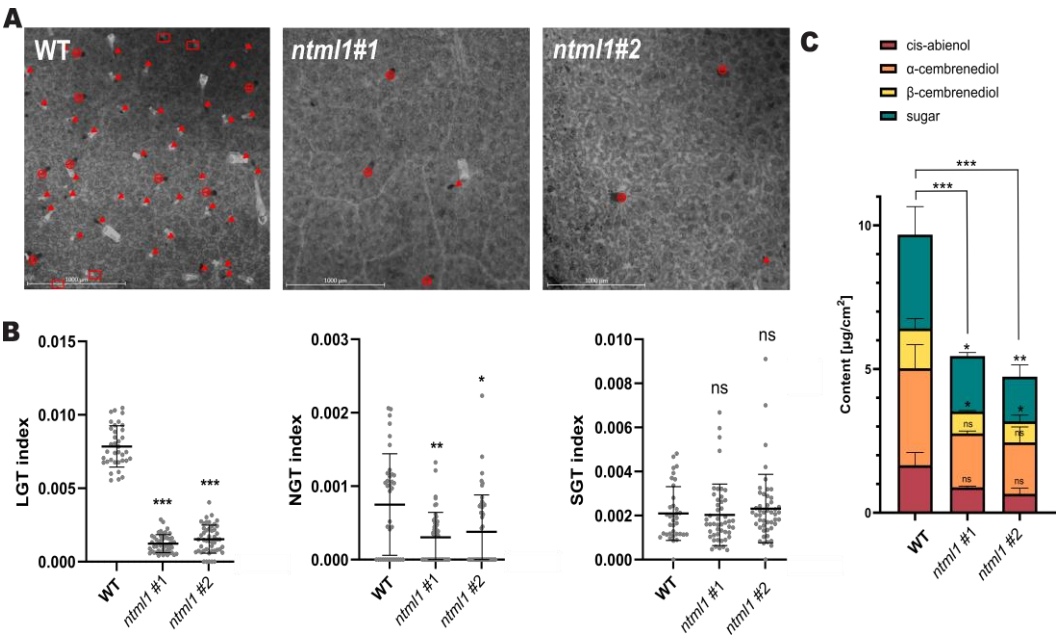

Figure 3 : *ntml1* mutants exhibit a reduction in LGTs and NGTs and altered secondary metabolite content.

(A) Leaf epidermis of wild-type (WT) and *ntml1* mutant lines. Images were taken from the abaxial side of mature leaves of 6-week-old plants stained with Calcofluor White. Red triangles indicate LGT, red circles indicate SGT, and red squares indicate NGT. (B) Quantification of trichome index (LGT, NGT, and SGT) in WT and *ntml1* lines. The index was calculated as the number of trichomes per pavement cell. Data represent means ± standard deviation. Statistical differences were assessed using one-way ANOVA followed by Dunnett's T3 test; asterisks indicate significance relative to WT (*:$p < 0.05$, **:$p < 0.01$, ***:$p < 0.001$, ns: not significant). $n = 48$. (C) Quantification of secondary metabolite levels in leaves of 6-week-old WT and *ntml1* mutants, including *cis*-abienol, α-cembrenediol, β-cembrenediol and total sugars. Data are presented as means ± standard deviation, $n = 4$. Statistical significance determined by ANOVA with Dunnett's T3 test (*:$p < 0.05$, **:$p < 0.01$, ***:$p < 0.001$; ns: not significant).

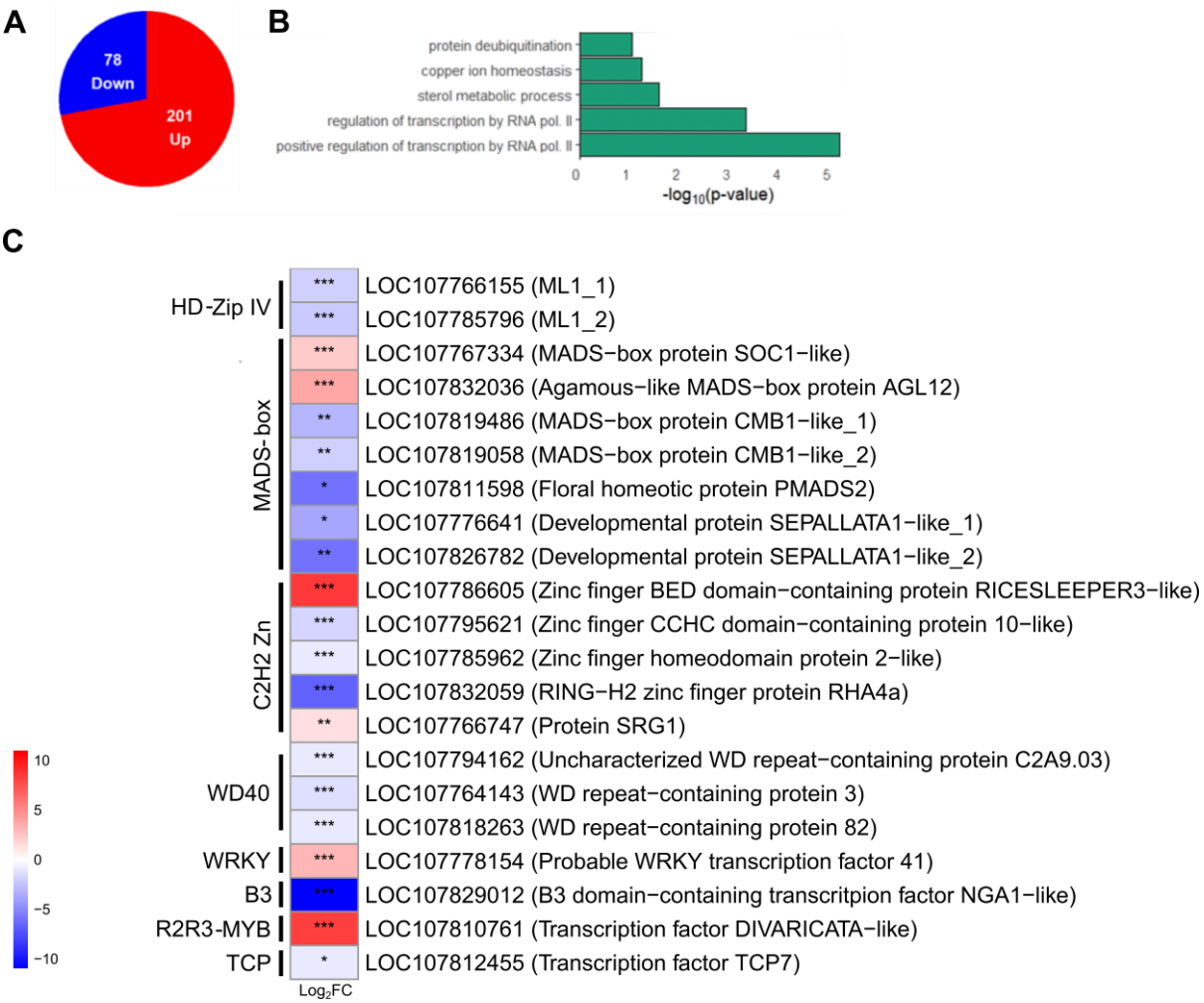

Figure 4 : Differentially expressed genes (DEGs) between WT and *ntml1* lines.

(A) Proportion of genes upregulated and downregulated in *ntml1* mutants compared to WT. (B) Gene Ontology (GO) enrichment analysis of DEGs associated with biological processes. (C) Heatmap of differentially expressed transcription-related genes identified by RNA-seq. Genes are labeled with their LOC identifiers and corresponding protein names (based on UniProt annotations). They are grouped according to protein family. Expression levels are represented on a $\log_2$ scale. Statistical significances are presented based on the adjusted pvalue *:$p < 0.05$, **:$p < 0.01$, ***:$p < 0.001$.

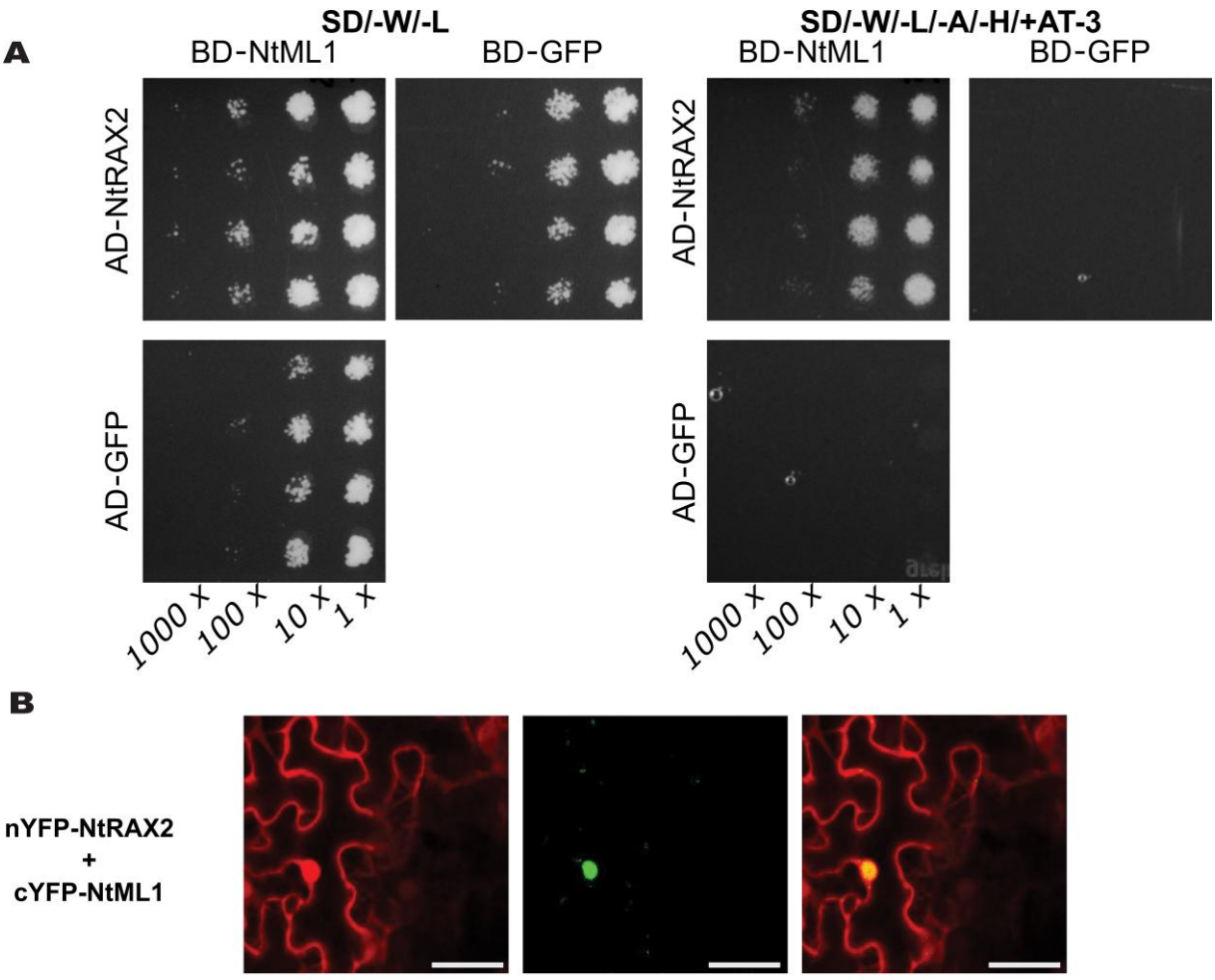

Figure 5: Interaction between NtML1 and NtRAX2 transcription factors. (A) Yeast two-hybrid (Y2H) assay showing interaction between NtML1 and NtRAX2. SD/-Leu/-Trp is the non-selective medium (positive control). SD/-Leu/-Trp/-Ade/-His + 5 mM of 3-amino-1,2,4-triazole (3-AT) is the selective medium. Co-transformation of BD-nlsGFP + AD-NtRAX2 and BD-NtML1 + AD-GFP was used as negative control in the selective medium. BD-NtML1 combined with AD-NtRAX2 conferred yeast cell growth on selective media. This experiment was done for four colonies/replicates. The colonies used for the spotting were diluted 1, 10, 100 and 1000 times. BD, Binding Domain; AD, Activation Domain; SD, synthetically defined medium. (B) Bimolecular complementation (BiFC) assay in tobacco leaves showing interaction between NtML1 and NtRAX2. NtML1 was fused to the C-terminal part of yellow fluorescent protein (cYFP) to form cYFP-NtML1. NtRAX2 was fused to the N-terminal part of yellow fluorescent protein (nYFP) to form nYFP-NtRAX2. mCherry signal is in red (left), YFP signal is in green (middle) and the merged signal is on the right.

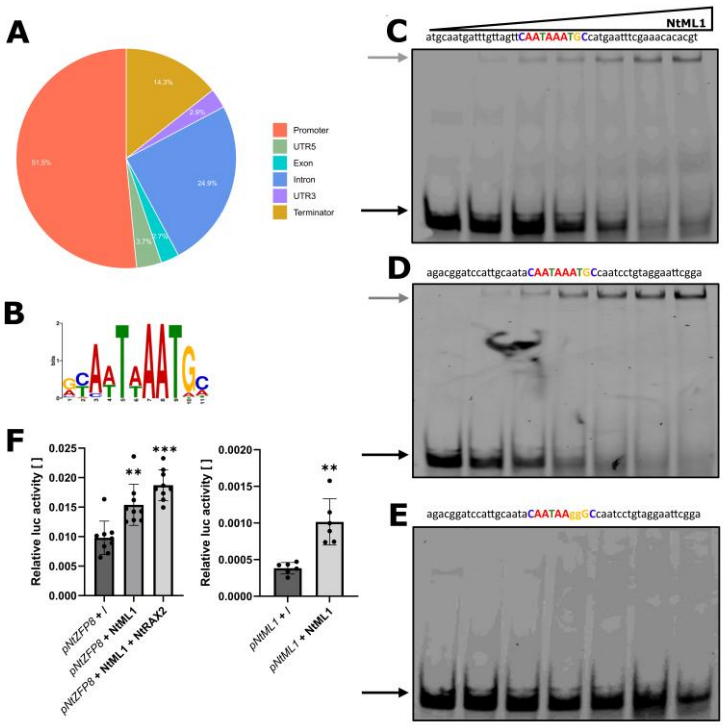

Figure 6 : NtML1 DNA-binding activity and target validation.

(A) Distribution of *NtML1*-binding peaks identified by DAP-seq relative to gene features. "Promoter" refers to the region from −3 kb to the Transcription Start Site; "Terminator" refers to the region from the 3' UTR to 1 kb downstream of the 3’ UTR. (B) DNA motif bound by *NtML1* identified through MEME analysis of DAP-seq data analysis (q-value = 1.2 e-^{381}).(C–E) Electrophoretic mobility shift assays (EMSAs) confirming *NtML1* binding to the CAATAAATGC motif; (C) 50 bp DNA probe centered on the DAP-seq peak in the *NtZFP8* promoter, including 10 bp motif sequence; (D) a 10 bp motif sequence (CAATAAATGC);(E) a mutated version of the motif. Black arrows indicate free DNA; grey arrows indicate DNA–protein complexes. Increasing concentrations (0 to 30 ng/ul) of *NtML1* were incubated with a constant amount of DNA from left to right. (F) Dual-luciferase assay showing activation of *pNtZFP8* by *NtML1* and *NtRAX2* (left) and activation of *pNtML1* by *NtML1* (right). Empty vector ("/") was used as a negative control. Luciferase activity was normalized to renilla luciferase. Data represents mean ± standard deviation; *n* = 6-9. Statistical significance was determined using ANOVA with Dunnett’s T3 test or Welch’s t-test (**:$p < 0.01$, ***:$p < 0.001$; ns: not significant).

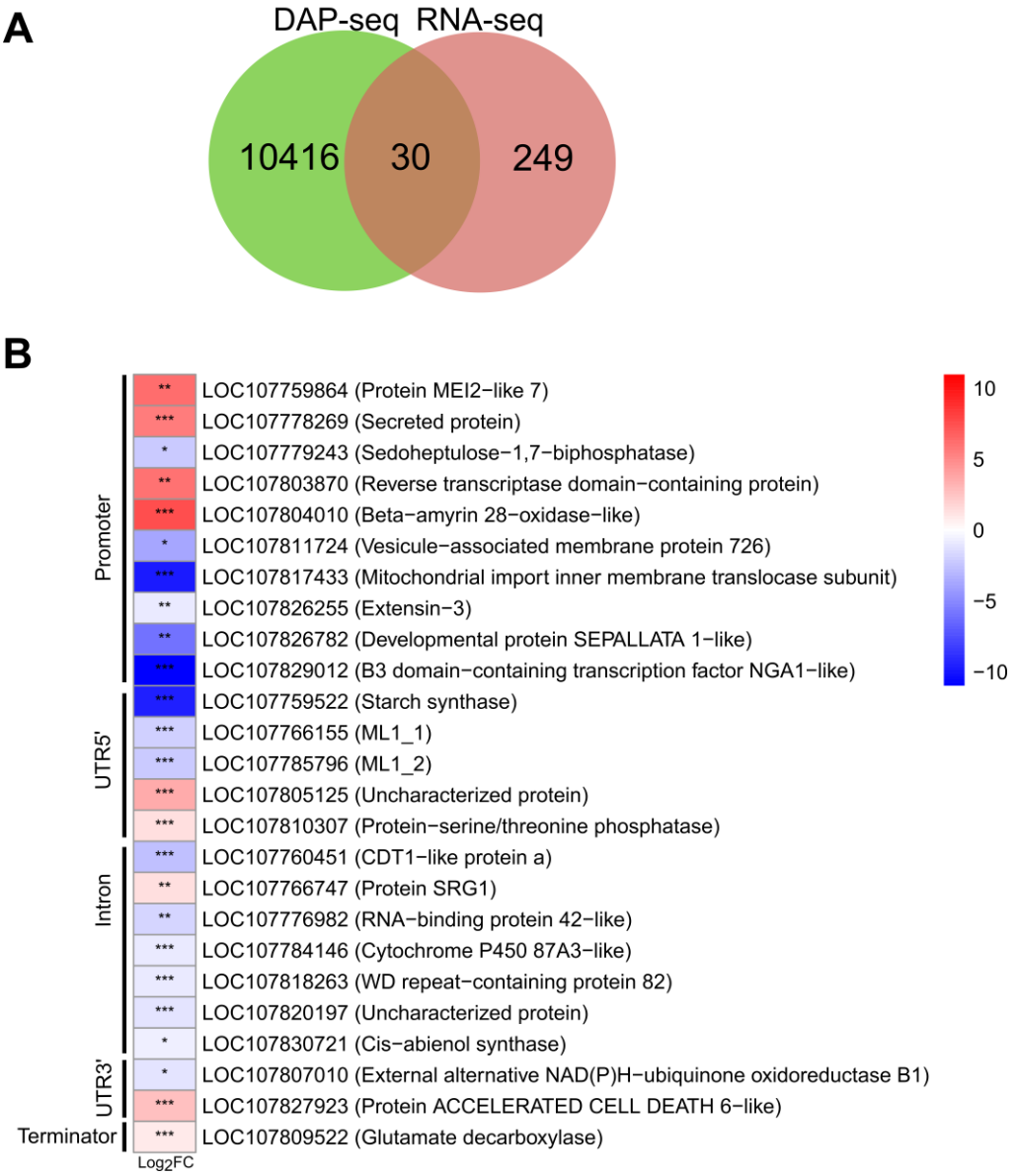

Figure 7: *NtML1* target genes identified by integrating DAP-seq and RNA-seq analyses.

(A) Venn diagram showing the overlap between genes directly bound by *NtML1* (identified by DAP-seq) and genes differentially expressed between WT and *ntml1* mutant lines (identified by RNA-seq). (B) Heatmap of expression profiles (RNA-seq) for genes common to RNA-seq and DAP-seq datasets. Genes are labeled with their LOC identifiers and corresponding protein names (based on UniProt annotations). Genes are grouped based on the location of NtML1 binding site (e.g., promoter, intron, UTR5', UTR3', terminator). Non-coding RNAs are not presented in the heat map. Expression levels are represented on a $\log_2$ scale. Statistical significances are presented based on the adjusted p-value *:$p < 0.05$, **:$p < 0.01$, ***:$p < 0.001$.