bioRxiv preprint doi: https://doi.org/10.1101/620021; this version posted May 22, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Unveiling two new trichome-specific promoters of interest for metabolic engineering in *Nicotiana tabacum*

Mathieu Pottier[$], *Raphaëlle Laterre, Astrid Van Wessem, Aldana M. Ramirez, Xavier Herman, Marc Boutry and Charles Hachez*[*]

Louvain Institute of Biomolecular Science and Technology, University of Louvain, 1348 Louvain-la-Neuve, Belgium

[$] Current address: InBioS-PhytoSYSTEMS, Laboratory of Plant Physiology, University of Liège, B-4000, Liège, Belgium

* Corresponding author:
Charles Hachez,
charles.hachez@uclouvain.be,
+ 32 10 47 37 96

ORCID:
Mathieu Pottier : 0000-0003-1551-4699
Marc Boutry : 0000-0002-2315-6900
Charles Hachez : 0000-0002-3688-7614

Acknowledgments

The authors are grateful to Joseph Nader for his technical contribution. This work was supported by the Belgian Fund for Scientific Research (Grant ID: MIS – F.4522.17), the Interuniversity Poles of Attraction Program (Belgian State, Scientific, Technical and Cultural Services), and an EU Marie Skłodowska-Curie fellowship (Project ID: 658932) to MP.

Author contributions

MP, RL, CH and MB designed the experiments and analyzed the data. MP, RL, AVW, AR, XH and MB performed experiments. CH, MP and MB wrote the manuscript.

Conflict of Interest: The authors declare that they have no conflict of interest.

bioRxiv preprint doi: https://doi.org/10.1101/620021; this version posted May 22, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Key message

pRbcS-T1 and *pMALD1*, two new trichome-specific promoters of *Nicotiana tabacum*, were identified and their strength and specificity were compared to those of previously described promoters in this species.

Abstract

Nicotiana tabacum has emerged as a suitable host for metabolic engineering of terpenoids and derivatives in tall glandular trichomes, which actively synthesize and secrete specialized metabolites. However, implementation of an entire biosynthetic pathway in glandular trichomes requires the identification of trichome-specific promoters to appropriately drive the expression of the transgenes needed to set up the desired pathway. In this context, RT-qPCR analysis was carried out on wild-type *N. tabacum* plants to compare the expression pattern and gene expression level of *NtRbS-T1* and *NtMALD1*, two newly identified genes expressed in glandular trichomes, with those of *NtCYP71D16, NtCBTS2α, NtCPS2*, and *NtLTP1*, which were reported in the literature to be specifically expressed in glandular trichomes. The latter were previously investigated separately, preventing any accurate comparison of their expression level. We show that *NtRbcS-T1* and *NtMALD1* are specifically expressed in glandular trichomes like *NtCYP71D16, NtCBTS2α,* and *NtCPS2*, while *NtLTP1* was also expressed in other leaf tissues as well as in the stem. Transcriptional fusions of each of the six promoters to the *GUS-VENUS* reporter gene were introduced in *N. tabacum* by *Agrobacterium*-mediated transformation. Most transgenic lines displayed GUS activity in tall glandular trichomes. In some transgenic lines, except for *pNtLTP1:GUS-VENUS*, this expression was specific. In other transgenic lines, GUS expression was extended to other tissues, probably resulting from a position effect during transgene integration. We discuss alternatives to overcome this lack of tissue specificity in some transgenic lines, should some of these promoters be used in the context of metabolic engineering in *N. tabacum*.

Keywords: Rubisco small subunit, Major Allergen Mal D 1.0501, Cembratrien-ol Synthase, Copal-8-ol diphosphate Synthase, Lipid Transfer Protein, Cytochrome P450 oxygenase

bioRxiv preprint doi: https://doi.org/10.1101/620021; this version posted May 22, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Introduction

Trichomes, the epidermal outgrowths covering most of aerial plant tissues are found in a very large number of plant species. Several types of trichomes (unicellular or multicellular, glandular or non-glandular) can be observed in a single plant species. Among those, glandular trichomes are characterized by the presence of cells forming a glandular structure that has the ability to secrete or store large quantities of specialized (also called secondary) metabolites (e.g., phenylpropanoids, flavonoids, acyl sugars, methylketones, and terpenoids),which possess antimicrobial and anti-fungal properties or which act as a defense barrier against herbivorous insects (Schilmiller et al. 2008).

The specialized metabolites secreted by glandular trichomes, which might represent up to 17 % of the leaf dry weight in *Nicotiana tabacum* (tobacco), have been largely exploited over centuries (Wagner et al. 2004). One of their most ancient uses originates from the aromatic properties and fragrance of some of those secretions referred to as essential oils. Besides, many of these specialized metabolites constitute an interesting source of pharmaceuticals and food additives. However, one of the common issues with some specialized metabolites is their low production yield in their natural host plant (Van Agtmael et al. 1999; Yoon et al. 2013). Some of them are highly praised molecules only found in a single plant species or even a single plant cultivar and often at low concentration (e.g., taxol found in *Taxus sp.*, artemisinin in *Artemisia annua* or cannabinoids in *Cannabis sativa*). Therefore, natural resources are often insufficient to reach the global need, while the complex stereochemistry of these compounds often prevents their full chemical synthesis in a cost-effective way.

In order to increase the overall yield, metabolic engineering strategies are undertaken to synthesize such metabolites through homologous or heterologous expression systems (Kirby and Keasling 2009; Marienhagen and Bott 2013). Advances in plant biotechnology and increasing knowledge in specialized metabolism also make it possible to exploit plants as production platforms for specific metabolites. One of the main advantages of such a strategy is that plants are photoautotrophic organisms, therefore requiring simple and cheap growth conditions, which accounts for a cost-effective biomass production (Kempinski et al. 2015). Besides, another benefit of using plant hosts is their ability to deal with membrane proteins such as P450 enzymes and posttranslational modifications such as glycosylation, two key features that are frequently limiting in prokaryotic hosts (van Herpen et al. 2010).

Among plant specialized metabolites, terpenoids and derivatives are the most abundant in terms of sheer number and chemical diversity (for review, see Croteau et al. 2000; Bouvier et

bioRxiv preprint doi: https://doi.org/10.1101/620021; this version posted May 22, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

al. 2005; Gershenzon and Dudareva 2007). Different plant species have been tested to perform metabolic engineering. From these trials, *N. tabacum* has emerged as one of the most suitable plant hosts (Moses and Pollier 2013; Lange et al. 2013; Wang et al. 2016). *N. tabacum* synthesizes an important pool of terpenoid precursors (IPP/DMAPP) and, besides the essential metabolites derived from the isoprenoid biosynthesis pathways, it also produces a very high amount of a limited range of specialized metabolites (Huchelmann et al. 2017). This combined to its high biomass, its fast growth rate, its easy genetic transformation and regeneration make it an interesting host to implement the biosynthesis pathways of terpenoid compounds and derivatives thereof.

However, engineering terpenoid biosynthetic pathways using ubiquist promoters frequently leads to severe phenotypes including dwarfism, chlorosis, and decreased seed production due to the cytotoxicity of these compounds or detrimental impact on the biosynthesis of essential metabolites (Saxena et al. 2014; Gwak et al. 2017; reviewed in Huchelmann et al. 2017). To avoid these adverse effects on the plant physiology, a fine control of the spatiotemporal expression of the transgenes, physically restricting the biosynthesis of potentially cytotoxic metabolites to specialized organs, is desirable (Huchelmann et al. 2017). Since glandular trichomes are non-essential organs, which are involved in the biosynthesis, storage and/or secretion of specialized metabolites and which naturally evolved to efficiently deal with high concentrations of these metabolites, they make ideal targets to develop such a metabolic engineering approach. For this purpose, identification of trichome-specific promoters and therefore of trichome-specific genes is required.

The promoters of four genes from *Nicotiana* species have previously been shown to be trichome-specific: CYtochrome P450 oxygenase 71D16 (*NtCYP71D16*), Copal-8-ol diPhosphate Synthase 2 (*NtCPS2*), Lipid Transfer Protein 1 *(NtLTP1*) of *N. tabacum*, as well as CemBraTrien-ol Synthase 2α (*NsCBTS2α*) of *Nicotiana sylvestris*. All these genes are exclusively related to specialized metabolism (Wang et al. 2002; Ennajdaoui et al. 2010; Choi et al. 2012; Sallaud et al. 2012).

A proteomic comparison was recently performed in *N. tabacum* between proteins extracted from tall glandular trichomes, which produce large amounts of terpenes, and those extracted from other plant organs (Laterre et al. 2017). This led to the identification of 47 proteins that were more abundant in tall glandular trichomes, the most enriched ones being a putative PR-10 type pathogenesis-related protein, namely Major Allergen Mal D 1.0501 (MALD1) and a

bioRxiv preprint doi: https://doi.org/10.1101/620021; this version posted May 22, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

small subunit of ribulose-1,5-bisphosphate carboxylase/oxygenase (RbcS-T1) (Laterre et al. 2017). The transcriptional promoter of a *Nicotiana benthamiana RbcS-T* homolog coupled to a reporter gene was shown to be trichome-specific in leaf tissues (Laterre et al. 2017). In *N. tabacum*, the trichome-specific localization of *NtMALD1* and *NtRbcS-T* transcripts was supported by semi-quantitative RT-PCR (Harada et al. 2010; Laterre et al. 2017).

To limit cytotoxicity issues and/or adverse effects on plant metabolism, elucidating whether gene promoters reported to be active in glandular trichomes are specific to glandular trichomes or display a broader activity in plant tissues is critical if those promoters are to be used in metabolic engineering approaches. The six promoters mentioned above had been investigated separately, preventing one to compare their transcript levels. In addition, for some of them, their cell-type specificity monitored by the GUS reporter gene had not been described in other organs than leaf tissues. Yet, these are key parameters to determine which are the most suitable ones for metabolic engineering in *N. tabacum* tall glandular trichomes.

The present study thus aimed at comparing the expression patterns and expression levels of *NtCYP71D16, NtCBTS2α, NtCPS2, NtLTP1, NtRbS-T1*, and *NtMALD1* promoters in *N. tabacum*. Their transcript levels in trichomes and different organs were compared. Transcriptional fusions of each of the six promoters to *GUS-VENUS* were expressed in transgenic *N. tabacum* plants. Most of the lines obtained with the six reporter constructs displayed GUS activity in the tall glandular trichomes. This expression was trichome-specific in a few lines, except for *pNtLTP1:GUS-VENUS*, corroborating transcripts data. The other lines exhibited GUS activity in other tissues, pointing out to random transgene insertion and positional effects as the most probable causes of such discrepancies. We therefore discuss alternatives to overcome this lack of tissue specificity in some lines, should some of these promoters be used in genetic constructs to drive the expression of specific transgenes.

bioRxiv preprint doi: https://doi.org/10.1101/620021; this version posted May 22, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Results

In a 2D gel analysis of glandular trichome proteins from *N. tabacum*, several spots were identified as trichome-specific proteins, among which RbcS-T1 and MALD1 (Laterre et al. 2017). The RNA levels of *NtRbS-T1* and *NtMALD1* as well as of *NtLTP1*, *NtCYP71D16*, *NtCBTS2α*, *and NtCPS2*, previously reported as genes specifically expressed in tall glandular trichomes, were compared in trichomes and different other *N. tabacum* organs. To do so, leaves were frozen in liquid nitrogen and carefully scratched with a brush to collect the trichomes. RT-qPCR assays were then performed on RNA extracted from trichomes, roots, trichome-free leaves, and trichome-free stems of six-week-old plants as well as from flowers of 10-week-old plants. Unlike for leaves and stems, trichomes could not be retrieved from flower sepals and petals. For each gene, the relative expression level in trichomes was arbitrarily set to one. The stability of the ubiquitin (*NtUBQ*) and the ATP-synthase β-subunit (*NtATP2*) control genes used to normalize the data is shown in Supplemental Fig. S1. All six studied genes showed a significantly ($p < 0.05$) higher relative expression level in isolated trichomes compared to the levels observed in roots, leaves, stems or flowers (Fig. 1). *NtCYP71D16, NtCBTS2α, NtCPS2, NtRbS-T1*, and *NtMALD1* exhibited very low expression in roots, trichome-free leaves and trichome-free stems, while higher transcript levels were found for *NtLTP1* in leaves and stems. The expression was high in flowers for the six genes but, as noted above, trichomes were not removed from these organs.

As most of these genes are involved in the biosynthesis (*NtCYP71D16, NsCBTS2α,* and *NtCPS2*) or transport (*NtLTP1*) of specialized metabolites secreted by mature glands, we wondered whether the leaf developmental stage could impact their expression in trichomes. Thus, glandular trichomes were isolated from leaves at different developmental stages, arbitrarily defined by the leaf length: < 2.5 cm (stage I), between 2.5 cm and 6.5 cm (stage II), between 6.5 cm and 15 cm (stage III), and > 15 cm (stage IV). The transcript levels of the six genes increased with leaf development (Fig. 2). The opposite trend was observed for elongation factor α (*EF1α)*, which peaked at stage I, confirming that the observed increasing level of all six genes is not an artifact of the normalization method (Supplemental Fig. S2). While the transcript level of *NtLTP1* appeared stable during leaf development, expression of the other five genes steadily increased until stage III where it reached a plateau (Fig. 2). Among them, *NtRbcS-T1* was the gene for which the transcript level increased the most with leaf development (4-fold increase). Expression of *NtCBTS2α* and *NtCYP71D16* involved in the biosynthesis of cembrenes, the major subgroup of diterpenes produced by *N. tabacum*

bioRxiv preprint doi: https://doi.org/10.1101/620021; this version posted May 22, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

glandular trichomes, also exhibited a large increase (3.8- and 3.6-fold, respectively) (Fig. 2). A more moderate increase was found for *NtMALD1* (2.6-fold) and *NtCPS2* (2.4-fold) transcripts, the latter being involved in the biosynthesis of another subgroup of diterpenes, namely labdanes.

The absolute expression levels of all six genes of interest was then determined in isolated trichomes for developmental stage III using an absolute standard curve method (Fig. 3, see Material and methods for details). Several control genes, some of which were used to normalize the relative expression data shown in Figures 1 and 2, were also added to the study for comparison purposes. Among control genes, the absolute expression levels (Fig. 3) were in agreement with previously published data in other Solanaceae species (Lu et al. 2012; Lacerda et al. 2015). Genes involved in cembrene production, *NtCBTS2α* (78.0 copies/pg), and *NtCYP71D16* (67.9 copies/pg), were the most expressed genes at stage III (Fig. 3), while a lower expression was found at this stage for *NtMALD1* (40.8 copies/pg), *NtLTP1* (28.2 copies/pg), *NtCPS2* (labdane *diterpenes*, 11.1 copies/pg) and *NtRbS-T1* (5.1 copies/pg).

To further confirm the trichome-specific expression pattern observed by RT-qPCR, we generated transcriptional reporter lines using a *GUS-VENUS* coding sequence. In the 2D gel analysis which led to the identification of trichome-specific proteins, two spots had been identified as trichome-specific RbcS (Laterre et al. 2017). At that time, only the *N. benthamiana* genome sequence was available and a RbcS transcription promoter (*pNbRbcS-T*) corresponding to the minor RbcS spot had been retrieved from this species and characterized (Laterre et al. 2017). Once the sequence of a *N. tabacum* genome became available, we identified *pNtRbcS-T1* (accession: MG493459.1) as the promoter of the gene corresponding to the major RbcS spot (NtRbcS-T1; accession: DV157962). The *pNtMALD1* promoter (accession: MG493458.1), corresponding to the NtMALD1 spot (accession: FS387666) was identified as well. The *GUS-VENUS* coding sequence was fused to *N. tabacum* genomic fragments of 1993 bp and 1974 bp upstream of the translation initiation codon of *NtRbcS-T1* and *NtMALD1*, respectively (Fig. 4). For the other genes, the previously published promoter regions, i.e. 985 bp (*NsCBTS2α*), 849 bp (*NtLTP1*), 1852 bp *(NtCYP71D16*), and 1448 bp (*NtCPS2*) (Wang et al. 2002; Ennajdaoui et al. 2010; Choi et al. 2012; Sallaud et al. 2012) were isolated and similarly fused to the *GUS-VENUS* coding sequence. These constructs were introduced in *N. tabacum* through *Agrobacterium tumefaciens*-mediated transformation. For each construct, 24 to 45 independent T_0 transgenic lines were generated and their GUS activity was monitored in leaves, stems, and roots of six-week-old transgenic plants after 16 h of incubation at 37°C to challenge the cell-type

bioRxiv preprint doi: https://doi.org/10.1101/620021; this version posted May 22, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

specificity (Table 1). GUS expression patterns were then confirmed on T_1 lines. Several observations can be made. i/ For each promoter, a large majority (83-100%) of the lines displayed GUS activity in tall glandular trichomes (Table 1). This indicates that appropriate cis-sequences required for expression in tall glandular trichomes are present in the sequences fused to the reporter gene. ii/ Some lines displayed strict trichome specificity even after extended GUS assay (see Table I as well as Figure 5 for *pNtRbcS-T1* and *pMALD1* as well as for *pNtCYP71D16* as a control). As an exception, none of the *NtLTP1* lines displayed trichome-specific GUS expression (Table I). This agrees with the RT-qPCR data (Fig. 1). iii/ In many lines, however, GUS activity was also found in other organs such as root, stem, leaf and/or flower (Table 1). In this case, the expression profile was variable according to the line (examples are displayed in Figure 6), probably as a consequence of the position effect (see discussion).

bioRxiv preprint doi: https://doi.org/10.1101/620021; this version posted May 22, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Discussion

NtRbS-T1 and *NtMALD1* extend the list of trichome-specific genes in *Nicotiana tabacum*

In this project, the tissue-specific expression of six *N. tabacum* genes, namely *NtLTP1*, *NtCYP71D16*, *NtCBTS2α*, *NtCPS2*, *NtRbS-T1*, and *NtMALD1*, was analyzed. Their trichome-specific expression at the transcript level had not yet been quantified and compared by RT-qPCR. We performed this comparison by RT-qPCR, using tall glandular trichome RNA as well as RNA of different plant organs. Except for *NtLTP1*, all these genes were found to be specifically expressed in tall glandular trichomes in *N. tabacum* (Fig. 1). Apart from *NtLTP1*, whose expression was almost constant during leaf development, that of the other five genes was lower at an early stage of leaf development and reached a maximum at stage III (Fig. 2), presumably when the specialized metabolism is fully operating. This is also true for *RbcS-T1* and this observation is in agreement with the hypothesis that in glandular trichomes, Rubisco recycles the CO_2 released by the specialized metabolism (Pottier et al. 2018).

These expression data may help choose appropriate trichome-specific promoters to drive the expression of a transgene for metabolic engineering purposes. Indeed, although *NtCYP71D16* and *NtCBTS2α* reach higher expression level in trichomes at stage III of leaf development, *NtCPS2* and *NtMALD1* promoters should lead to a more homogenous expression of transgenes among leaves at different developmental stages. *NtLTP1* was an exception since its transcripts were identified in leaf and stem tissues devoid of trichomes. Although the *NtLTP1* promoter was claimed to confer trichome-specific expression, examination of the GUS reporter activity reported by Choi et al. (2012) revealed some activity in other cell types than trichomes and the expression in the stem was not displayed. Our observations are in line with previously published semi-quantitative RT-PCR which showed that *NtLTP1* is expressed in different organs (Harada et al. 2010).

A variety of expression patterns was detected for the different transcriptional reporters

Analysis of *GUS-VENUS* reporter lines revealed that, in most of them, the six promoters drove gene expression in the head cells of tall glandular trichomes of *N. tabacum.* However, expression was rarely exclusively observed in trichomes (Table 1). Moreover, for a given reporter construct, variability in expression patterns was observed between independent lines. Concerning *NtCYP71D16*, Wang et al (2002) noted that after overnight incubation (like in the present work) with the GUS substrate, faint GUS staining was found in guard cells and veins

bioRxiv preprint doi: https://doi.org/10.1101/620021; this version posted May 22, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

in some areas of some leaves, and at the cut edges of some leaf blades. No data was reported for the GUS reporter driven by the *NtCBTS-2α* and *NtCPS2* promoters in other organs than leaves (Ennajdaoui et al. 2010; Sallaud et al. 2012). Based on RT-qPCR data, *NtCYP71D16*, *NtCBTS2α*, *NtCPS2*, *NtRbS-T1*, and *NtMALD1* can, however, be considered as specifically expressed in trichomes. A possible reason for the partial discrepancy between RT-qPCR and gene reporter data is that the promoter regions that were used in the constructs are incomplete. This is unlikely to be the case for six promoters, especially for those that are close to 2 kb long (i.e. *pNtRbcS-T1*, *pNtMALD1*, and *pNtCYP71D16*), unless cis regulatory sequences are located downstream of the translation start. In addition, this hypothesis fails to explain the wide diversity in expression patterns observed between lines expressing the same construct. A more likely explanation is a position effect due to the random insertion of the T-DNA in the plant cell genome. Indeed, the genomic environment surrounding the integrated cassette (structure of chromatin, presence of enhancers/silencers near the insertion site) is known to alter the expression level and profile of transgenes (Kohli et al. 2010; Hernandez-Garcia and Finer 2014). Between independent lines, and thus different insertion sites, those position effects might differ according to the proximal endogenous regulatory elements.

Integration of insulators in the vector, at both ends of the transgene (as well as between different expression cassettes within the construct in case of a multigene construct) could be a way to prevent undesirable effects (Hasegawa and Nakatsuji 2002; Biłas et al. 2016). Indeed, insulators are sequences that stabilize gene expression by guaranteeing gene autonomy (Hasegawa and Nakatsuji 2002; Biłas et al. 2016). Another option could be to insert the transgenes in a specific locus, preferentially one promoting high expression of genes (Abdel-ghany et al. 2015). Genome editing technologies such as CRISPR-Cas9, which allows insertions in a desired locus through homologous end joining, could be a way to cope with such an issue (Cao et al. 2016; Khatodia et al. 2016). Finally, a reporter gene coding for a fluorescent protein such as GFP could be included in the genetic construct to facilitate the screening of transformed plants in order to select transgenic lines that only express the reporter in glandular trichomes. If the same promoter is used to drive the expression of both the fluorescent reporter and the transgene(s) of interest, the rationale is that the fluorescence pattern and intensity of the reporter mirror those of the transgene(s) of interest.

It would have been interesting to confirm, in the reporter lines, the quantitative variations of activity between the different promoters. However, because of the position effect, quantitative variations do not necessarily reflect the actual strengths of the promoters. The

bioRxiv preprint doi: https://doi.org/10.1101/620021; this version posted May 22, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

positive consequence of this variation is that transgenic plants can be screened to find those that express the gene of interest at the appropriate level.

Cloning strategy for implementation of an entire metabolic pathway in *Nicotiana tabacum* glandular trichomes

The use of *N. tabacum* glandular trichomes for metabolic engineering purposes is a very attractive strategy. However, heterologous expression of an entire metabolic pathway into such specific plant organ may be a tedious process that poses several technical challenges. This includes promoting cell type-specific expression at appropriate levels of multiple genes (coding for enzymes or transporters) that segregate together in the progeny.

For this purpose, the use of several independent T-DNA to generate plants expressing an entire metabolic pathway is questionable and an approach allowing delivery of multiple genetic constructs in a unique vector appears to be a far better option. In this way, a single DNA molecule is transferred into the plant cells, at the same locus, which limits the intergenic variation in expression due to positional effects in the genome. Such an approach also ensures that all genes cloned in this molecule are inherited together (Dafny-Yelin and Tzfira 2007).

However, generating a single T-DNA vector bearing several genes each placed under the same trichome-specific promoter can lead to a dilution effect since transcription might be limited by the amount of available transcription factors recruited by cis-elements present in this promoter (Biłas et al. 2016). To prevent such dilution effect, genes coding for each enzyme could be placed downstream different trichome-specific promoters. The use of different promoters will also prevent homologous recombination between similar promoters in the construct and loss of the sequence in-between.

In addition to that, the level of expression is obviously an essential trait. Beside avoiding dilution effects of trans elements, using different promoters may be needed to adapt the level of expression to the efficiency of the enzyme encoded by the transgene. This explains why identifying different trichome-specific gene promoters is so important. By identifying and comparing different trichome-specific promoters such as *pNtCYP71D16*, *pNtCBTS2α*, *pNtCPS2*, *pNtRbS-T1*, and *pNtMALD1*, our work provides the community with a useful toolkit to express a gene in glandular trichomes at different levels.

In conclusion, a key and unique feature of glandular trichomes is their ability to synthesize and secrete large amounts of a limited panel of specialized metabolites. Taking advantage of the pool of natural precursors to produce specific metabolites in glandular trichomes by metabolic engineering would therefore be of high biotechnological interest. This requires the

bioRxiv preprint doi: https://doi.org/10.1101/620021; this version posted May 22, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

availability of transcriptional promoters specifically active in these structures that could be used to efficiently drive the expression of the transgenes coding for the enzymes needed to implement the pathway in a cell-type specific way. In this respect, the identification of the *NtMALD1* and *NtRbS-T1* promoters and their comparison with previously identified trichome-specific promoters are promising tools for expressing entire biosynthesis pathways in glandular trichomes of *N. tabacum.* However, use of trichome-specific promoters should ideally be associated with the addition of insulators at the end of each expression cassette as well as with site-directed genome edition technologies to ensure cell-type specific expression at an appropriate level. In the long term, detailed knowledge of the gene network existing in plant glandular trichomes will generate new leads to tap the largely unexploited potential of these specialized plant organs to increase plant resistance to pests, or to lead to improved production of important specialized metabolites via finely tuned metabolic engineering approaches.

bioRxiv preprint doi: https://doi.org/10.1101/620021; this version posted May 22, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Material and methods

Plant material and plant growth conditions

Nicotiana tabacum cv Petit Havana SR1 (Maliga *et al.* 1973) plants were used in this work. For the *in vitro* cultures, seeds were sterilized by immersion in 1 ml 70% (v/v) ethanol for 1 min and then in 1 ml 50% (v/v) commercial bleach for 2 min. Seeds were then washed three times with 1 ml of sterile MilliQ water and kept at 4°C, in the dark, during 48 h for stratification. Sterilized seeds were sown on solid Murashige and Skoog (MS) medium [4.33 $g.l^{-1}$ MS salts (MP Biochemicals, Solon, OH, USA; www.mpbio.com), 3% (w/v) sucrose, 1% (w/v) agar, pH 5.8 (KOH)] and placed in the growth chamber at 25°C under a 16 h photoperiod (50 µmol photon m^{-2} sec^{-1}). For the soil cultures, seeds were stratified before being sown in potting soil (DCM, Grobbendonk, Belgium; dcm-info.com). Isolated plantlets coming from potting soil or *in vitro* conditions were transferred to Jiffy pots (Gronud, Norway; www.jiffypot.com) before being transferred to bigger pots containing potting soil (DCM). Plants on soil were grown under controlled conditions, in a phytotron set at 25°C and with a 16 h photoperiod (300 µmol photon m^{-2} sec^{-1}).

Tissue isolation, RNA extraction and cDNA synthesis

Trichomes were removed from tissues of 6-week-old plants following the cold-brushing method (Wang *et al.* 2001). For the analysis of gene expression in trichomes during leaf development, trichomes were isolated from leaves at different developmental stages defined here by leaf length: < 2.5 cm (stage I), between 2.5 cm and 6.5 cm (stage II), between 6.5 cm and 15 cm (stage III), and > 15 cm (stage IV). For the analysis of gene expression in different tissues, we harvested roots, trichomes-free stems, trichomes-free leaves, and leaf trichomes (pool of leaves from stage I to stage III) from 6-week-old plants, and flowers from 10-week-old plants. For each biological replicate (except for isolated trichomes), 100 mg of material was pre-ground in liquid nitrogen using a mortar and pestle. Pre-ground tissues and isolated trichomes were ground in 2 mL Precellys tubes containing 200 µL of ceramic beads Zirmil (0.5 mm, Saint Gobain Zipro, Le Pontet, France) and 500 µL of lysis/2-Mercaptoethanol solution of the Spectrum™ Plant Total RNA Kit (Sigma-Aldrich, St. Louis, Missouri, USA; http://www.sigmaaldrich.com). Samples were subjected to four consecutive 30 s grinding periods at 6,000 rpm using a Precellys 24 (Bertin Technologies, Montigny-le-Bretonneux, France). The homogenates were centrifuged at 1,000 *g* for 3 min (Eppendorf 5430, Hamburg,

bioRxiv preprint doi: https://doi.org/10.1101/620021; this version posted May 22, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Germany). The subsequent steps of the RNA extraction were performed on the supernatants according to the manufacturer's specifications, except that the 56 °C incubation step was omitted. RNA was eluted in 50 µl elution buffer and quantified using a spectrophotometer (Nanodrop® ND-1000, Isogen Life Science, The Netherlands; www.isogen-lifescience.com). Genomic DNA contamination was eliminated by using the On-Column DNase I Digestion Set (Sigma-Aldrich, St. Louis, Missouri, USA; www.sigmaaldrich.com). The RNA was finally flash frozen in liquid nitrogen and stored at -80°C. DNA-free RNA (500 µg) was used for reverse transcription using the Moloney Murine Leukemia Virus Reverse transcriptase (Promega, Madison, Wisconsin, USA; be.promega.com) and oligo$(dT)_{18}$. Reverse transcription mixture was added according to the manufacturer's specifications. After adding the transcriptase, samples were incubated for 5 min at 25°C, followed by 1 h at 42°C and 5 min at 85°C, placed on ice for 5 min, aliquoted, and stored at -20°C.

Gene expression

Gene-specific RT-qPCR primers listed in Table S1 were designed at the 3'end of the coding sequence, (size, about 100 bp; melting temperature, 60°C) using OligoPerfect™ Designer (www.thermofisher.com). cDNA (5 µl, 17 fold diluted) was used as a template in 20 µl RT-qPCR reaction, which also contained 10 µl of qPCR master mix plus for SYBR Green I (Eurogentec, Seraing, Belgium, https://secure.eurogentec.com/eu-home.html) and 5 µl of primer mix (1.3 µM each). Amplification was performed on an ABI 7500 Real-Time PCR system (Waltham, Massachusetts, USA; http://www.thermofisher.com). Primer specificity was confirmed by analysis of the melting curves. For each tissue, primer amplification efficiency (≥ 95%) was determined using five standards from serial dilutions of a cDNA pool of the biological replicates used for gene expression analysis. Relative transcript levels were calculated following the $2^{-\Delta\Delta Ct}$ method (Livak and Schmittgen 2001) with the geometric mean of mitochondrial ATP-synthase β-subunit (*ATP2*) and ubiquitin (*UBQ*) used as references for comparison between different tissues, and of *ATP2*, *UBQ*, and actin (*ACTIN*), for comparison between different leaf developmental stages. For absolute quantification, PCR products amplified by gene-specific RT-qPCR primers listed in Supplemental Table S1 were cloned in pGEM-T Easy vector (Promega, Madison, Wisconsin, USA) prior to their sequencing. Constructs were linearized by PstI restriction, purified using Nucleospin Extract II kit (Macherey-Nagel, Düren, Germany) and rigorously quantified using a spectrophotometer (Nanodrop® ND-1000, Isogen Life Science, The Netherlands; www.isogen-lifescience.com). For each quantified purified linear plasmid, the copy number was determined according to the

bioRxiv preprint doi: https://doi.org/10.1101/620021; this version posted May 22, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

following equation: copy number = (vector amount [g]) × 6.023 × 10^{23} [molecules/mole]) / (660 [g/mole/base] × size of the vector+insert [bases]. Absolute transcript levels were determined through the absolute standard curve method. Thus, for each studied gene, standards (2.10^6, 2.10^5, 2.10^4, 2.10^3 copies) obtained by serial dilution of the purified linear plasmids were included in duplicate in q-PCR plates used to study gene expression during trichome development.

Generation of plants expressing promoter-*GUSVENUS* Fusions

The transcription promoter regions of *NtRbcS-T1* (1993 pb; GenBank accession: MG493459.1) and *NtMALD1* (1974 pb; GenBank accession: MG493458.1) were identified blasting the EST corresponding to *NtRbcS-T1* (GenBank accession: DV157962) and *NtMALD1* (GenBank accession: FS387666) coding sequences to the genome of *N. tabacum* TN90 in the Solgenomics database (http://solgenomics.net). The promoter regions of *NsCBTS2α* (985 bp; GenBank accession: HM241151.1), *NtLTP1* (849 bp; GenBank accession: AB625593.1), *NtCYP71D16* (1852 pb; GenBank accession: AF166332.1), and *NtCPS2* (1448 bp; GenBank accession: HE588139.1) were defined as previously (Wang et al. 2002; Ennajdaoui et al. 2010; Choi et al. 2012; Sallaud et al. 2012). Promoter regions were amplified by PCR using as a template genomic DNA prepared from *N. tabacum* or *N. sylvestris* leaves and the primers listed in Supplemental Table S2. The amplified fragments were inserted in the pGEM®-T Easy Vector (Promega, Madison, Wisconsin, USA; www.promega.com) and sequenced. Cloned fragments were cleaved using HindIII (or NotI for *pNtMALD1* and *pNtRbcS-T1*) and KpnI, prior to their insertion in a pAUX3131 construct (Navarre *et al.* 2011), upstream of the *GUSVENUS* coding sequence. The fusion construct was excised using I-SceI and inserted into the pPZP-RCS2-nptII plant expression vector (Goderis *et al.* 2002), also cut with I-SceI. The construct was introduced into *Agrobacterium tumefaciens* LBA4404 virGN54D (van der Fits *et al.* 2000) for subsequent *N. tabacum* leaf disc transformation (Horsch *et al.* 1986). The regenerated plants were finally transferred to soil to be analyzed by GUS staining.

GUS histochemical analysis

Histochemical staining of plant tissues for GUS activity was conducted as described previously (Bienert *et al.* 2012). To determine the tissue expression frequencies among independent lines transformed with the same construct (Table 1), GUS staining was performed for 16h on 6-week-old T0 plants growing in soil. Results were then confirmed in

bioRxiv preprint doi: https://doi.org/10.1101/620021; this version posted May 22, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

the corresponding T1 plants of the same age except for flowers tissues, which were analyzed on 10-week-old plants (Fig. 5 and Fig. 6). Stained tissues were washed with 70% ethanol for chlorophyll extraction, transferred to 50% glycerol, observed under a light binocular (Carl Zeiss MicroImaging) and photographed (Moticam 2300).

Statistical analysis

All tests were performed using the R software. For q-PCR, data were analyzed using *kruskal.test* (Kruskal–Wallis) function for multiple comparisons. For multiple comparisons, *nparcomp* package was used to perform Tukey post-hoc test when significant differences were detected ($P < 0.05$). Different letters indicate significant differences between samples.

Supplementary data

Supplemental Fig. 1 Expression of *NtATP2* and *NtUBQ* control genes used to normalize the data in Figure 1.

Supplemental Fig. 2 Expression of *NtEF1α* in trichomes according to leaf developmental stage.

Supplemental Table S1 List of primers used for RT-qPCR.

Supplemental Table S2 List of primers used to amplify the promoter sequences.

bioRxiv preprint doi: https://doi.org/10.1101/620021; this version posted May 22, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

References

Abdel-ghany SE, Golovkin M, Reddy ASN (2015) Engineering of Plants for the Production of Commercially Important Products: Approaches and Accomplishments. In: Bahadur B (ed) Plant Biology and Biotechnology. Springer India, New Delhi, pp 551–577

Bienert MD, Delannoy M, Navarre C, Boutry M (2012) NtSCP1 from Tobacco Is an Extracellular Serine Carboxypeptidase III That Has an Impact on Cell Elongation. Plant Physiol 158:1220–1229 . doi: 10.1104/pp.111.192088

Biłas R, Szafran K, Hnatuszko-Konka K, Kononowicz AK (2016) Cis-regulatory elements used to control gene expression in plants. Plant Cell. Tissue Organ Cult. 127:269–287

Bouvier F, Rahier A, Camara B (2005) Biogenesis, molecular regulation and function of plant isoprenoids. Prog Lipid Res 44:357–429 . doi: 10.1016/j.plipres.2005.09.003

Cao HX, Wang W, Le HTT, Vu GTH (2016) The power of CRISPR-Cas9-induced genome editing to speed up plant breeding. Int J Genomics. doi: 10.1155/2016/5078796

Choi YE, Lim S, Kim H-J, Han JY, Lee M-H, Yang Y, Kim J-A, Kim Y-S (2012) Tobacco NtLTP1, a glandular-specific lipid transfer protein, is required for lipid secretion from glandular trichomes. Plant J 70:480–91 . doi: 10.1111/j.1365-313X.2011.04886.x

Croteau R, Kutchan TM, Lewis NG (2000) Natural Products (Secondary Metabolites). In: Buchanan B, Gruissem W, Jones R (eds) Biochemistry & Molecular Biology of Plants. American Society of Plant Physiologists, pp 1250–1318

Dafny-Yelin M, Tzfira T (2007) Delivery of Multiple Transgenes to Plant Cells. Plant Physiol 145:1118–1128 . doi: 10.1104/pp.107.106104

Ennajdaoui H, Vachon G, Giacalone C, Besse I, Sallaud C, Herzog M, Tissier A (2010) Trichome specific expression of the tobacco (Nicotiana sylvestris) cembratrien-ol synthase genes is controlled by both activating and repressing cis-regions. Plant Mol Biol 73:673–85 . doi: 10.1007/s11103-010-9648-x

Gershenzon J, Dudareva N (2007) The function of terpene natural products in the natural world. Nat Chem Biol 3:408–414 . doi: 10.1038/nchembio.2007.5

Goderis IJWM, De Bolle MFC, François IEJA, Wouters PFJ, Broekaert WF, Cammue BPA (2002) A set of modular plant transformation vectors allowing flexible insertion of up to six expression units. Plant Mol Biol 50:17–27 . doi: 10.1023/A:1016052416053

Gwak YS, Han JY, Adhikari PB, Ahn CH, Choi YE (2017) Heterologous production of a ginsenoside saponin (compound K) and its precursors in transgenic tobacco impairs the vegetative and reproductive growth. Planta 245:1105–1119 . doi: 10.1007/s00425-017-

bioRxiv preprint doi: https://doi.org/10.1101/620021; this version posted May 22, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

2668-x

Harada E, Kim J-AA, Meyer AJ, Hell R, Clemens S, Choi Y-EE (2010) Expression profiling of tobacco leaf trichomes identifies genes for biotic and abiotic stresses. Plant Cell Physiol 51:1627–1637 . doi: 10.1093/pcp/pcq118

Hasegawa K, Nakatsuji N (2002) Insulators prevent transcriptional interference between two promoters in a double gene construct for transgenesis. FEBS Lett 520:47–52 . doi: 10.1016/S0014-5793(02)02761-8

Hernandez-Garcia CM, Finer JJ (2014) Identification and validation of promoters and cis-acting regulatory elements. Plant Sci 217–218:109–119 . doi: 10.1016/j.plantsci.2013.12.007

Horsch RB, Klee HJ, Stachel S, Winans SC, Nester EW, Rogers SG, Fraley RT (1986) Analysis of Agrobacterium tumefaciens virulence mutants in leaf discs. Proc Natl Acad Sci U S A 83:2571–2575 . doi: 10.1073/pnas.83.8.2571

Huchelmann A, Boutry M, Hachez C (2017) Plant glandular trichomes: natural cell factories of high biotechnological interest. Plant Physiol 175:6–22 . doi: 10.1104/pp.17.00727

Kempinski C, Jiang Z, Bell S, Chappell J (2015) Metabolic engineering of higher plants and algae for isoprenoid production. In: Schrader J, Bohlmann J (eds) Biotechnology of Isoprenoids. Springer International Publishing, Cham, pp 161–199

Khatodia S, Bhatotia K, Passricha N, Khurana SMP, Tuteja N (2016) The CRISPR/Cas Genome-Editing Tool: Application in Improvement of Crops. Front Plant Sci 7:1–13 . doi: 10.3389/fpls.2016.00506

Kirby J, Keasling JD (2009) Biosynthesis of Plant Isoprenoids: Perspectives for Microbial Engineering. Annu Rev Plant Biol 60:335–55 . doi: 10.1146/annurev.arplant.043008.091955

Kohli A, Miro B, Twyman RM (2010) Chapter 7 Transgene Integration, Expression and Stability in Plants: Strategies for Improvements. In: C. Kole et al. (ed) Transgenic Crop Plants. Springer-Verlag, Berlin, Heidelberg, pp 201–237

Lacerda ALM, Fonseca LN, Blawid R, Boiteux LS, Ribeiro SG, Brasileiro ACM (2015) Reference gene selection for qPCR analysis in tomato-bipartite begomovirus interaction and validation in additional tomato-virus pathosystems. PLoS One 10:1–17 . doi: 10.1371/journal.pone.0136820

Lange BM, Ahkami A, Markus Lange B, Ahkami A (2013) Metabolic engineering of plant monoterpenes, sesquiterpenes and diterpenes-current status and future opportunities. Plant Biotechnol J 11:169–96 . doi: 10.1111/pbi.12022

bioRxiv preprint doi: https://doi.org/10.1101/620021; this version posted May 22, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

536 Laterre R, Pottier M, Remacle C, Boutry M (2017) Photosynthetic Trichomes Contain a
537 Specific Rubisco with a Modified pH-Dependent Activity. Plant Physiol 173:2110–2120
538 . doi: 10.1104/pp.17.00062
539 Livak KJ, Schmittgen TD (2001) Analysis of relative gene expression data using real-time
540 quantitative PCR and the 2-ΔΔCT method. Methods 25:402–408 . doi:
541 10.1006/meth.2001.1262
542 Lu Y, Xie L, Chen J (2012) A novel procedure for absolute real-time quantification of gene
543 expression patterns. Plant Methods 8:1–11 . doi: 10.1186/1746-4811-8-9
544 Maliga P, Sz-Breznovits A, Márton L (1973) Streptomycin-resistant plants from callus culture
545 of haploid tobacco. Nat New Biol 244:29–30
546 Marienhagen J, Bott M (2013) Metabolic engineering of microorganisms for the synthesis of
547 plant natural products. J Biotechnol 163:166–178 . doi: 10.1016/j.jbiotec.2012.06.001
548 Moses T, Pollier J (2013) Bioengineering of plant (tri) terpenoids: from metabolic
549 engineering of plants to synthetic biology in vivo and in vitro. New Phytol 200:27–43 .
550 doi: 10.1111/nph.12325
551 Navarre C, Sallets A, Gauthy E, Maîtrejean M, Magy B, Nader J, Pety de Thozée C, Crouzet
552 J, Batoko H, Boutry M (2011) Isolation of heat shock-induced Nicotiana tabacum
553 transcription promoters and their potential as a tool for plant research and biotechnology.
554 Transgenic Res 20:799–810 . doi: 10.1007/s11248-010-9459-5
555 Pandey A, Misra P, Bhambhani S, Bhatia C, Trivedi PK (2014) Expression of arabidopsis
556 myb transcription factor, atmyb111, in tobacco requires light to modulate flavonol
557 content. Sci Rep 4:5018 . doi: 10.1038/srep05018
558 Pottier M, Gilis D, Boutry M (2018) The Hidden Face of Rubisco. Trends Plant Sci 23:382–
559 392 . doi: 10.1016/j.tplants.2018.02.006
560 Sallaud C, Giacalone C, Töpfer R, Goepfert S, Bakaher N, Rösti S, Tissier A (2012)
561 Characterization of two genes for the biosynthesis of the labdane diterpene Z-abienol in
562 tobacco (Nicotiana tabacum) glandular trichomes. Plant J 72:1–17 . doi: 10.1111/j.1365-
563 313X.2012.05068.x
564 Saxena B, Subramaniyan M, Malhotra K, Bhavesh NS, Potlakayala SD, Kumar S (2014)
565 Metabolic engineering of chloroplasts for artemisinic acid biosynthesis and impact on
566 plant growth. J Biosci 39:33–41 . doi: 10.1007/s12038-013-9402-z
567 Schilmiller AL, Last RL, Pichersky E (2008) Harnessing plant trichome biochemistry for the
568 production of useful compounds. Plant J 54:702–11 . doi: 10.1111/j.1365-
569 313X.2008.03432.x

bioRxiv preprint doi: https://doi.org/10.1101/620021; this version posted May 22, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

570 Van Agtmael MA, Eggelte TA, Van Boxtel CJ (1999) Artemisinin drugs in the treatment of
571 malaria: From medicinal herb to registered medication. Trends Pharmacol Sci 20:199–
572 205 . doi: 10.1016/S0165-6147(99)01302-4
573 van der Fits L, Deakin E, Hoge J, Memelink J (2000) The ternary transformation system:
574 constitutive virG on a compatible plasmid dramatically increases Agrobacterium-
575 mediated plant transformation. Plant Mol Biol 43:495–502
576 van Herpen TWJM, Cankar K, Nogueira M, Bosch D, Bouwmeester HJ, Beekwilder J (2010)
577 Nicotiana benthamiana as a production platform for artemisinin precursors. PLoS One
578 5:e14222 . doi: 10.1371/journal.pone.0014222
579 Wagner GJ, Wang E, Shepherd RW (2004) New approaches for studying and exploiting an
580 old protuberance, the plant trichome. Ann Bot 93:3–11 . doi: 10.1093/aob/mch011
581 Wang B, Kashkooli AB, Sallets A, Ting H-MMH-M, de Ruijter NCAA, Olofsson L,
582 Brodelius P, Pottier M, Boutry M, Bouwmeester H, van der Krol AR (2016) Transient
583 production of artemisinin in Nicotiana benthamiana is boosted by a specific lipid transfer
584 protein from A. annua. Metab Eng 38:159–169 . doi: 10.1016/j.ymben.2016.07.004
585 Wang E, Gan S, Wagner GJ (2002) Isolation and characterization of the CYP71D16
586 trichome-specific promoter from Nicotiana tabacum L. J Exp Bot 53:1891–1897 . doi:
587 10.1093/jxb/erf054
588 Wang E, Wang R, DeParasis J, Loughrin JH, Gan S, Wagner GJ (2001) Suppression of a
589 P450 hydroxylase gene in plant trichome glands enhances natural-product-based aphid
590 resistance. Nat Biotechnol 19:371–374 . doi: 10.1038/86770
591 Yoon JM, Zhao L, Shanks J V (2013) Metabolic Engineering with Plants for a Sustainable
592 Biobased Economy. Annu Rev Chem Biomol Eng 4:211–37 . doi: 10.1146/annurev-
593 chembioeng-061312-103320
594
595

bioRxiv preprint doi: https://doi.org/10.1101/620021; this version posted May 22, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Table 1 Frequency of GUS activity in different organs of transgenic *N. tabacum* lines showing expression in tall glandular trichomes[a].

PROMOTERS	SGT[b]	LEAVES	ROOTS	STEMS	ONLY IN TGT
pNtRbcS-T1	11/20	18/20	10/20	16/20	2/20
pNtMALD1	10/26	20/26	21/26	22/26	1/26
pNsCBTS2α	7/44	29/44	27/44	41/44	2/44
pNtCYP71D16	12/32	21/32	15/32	25/32	3/32
pNtCPS2	21/26	17/26	16/26	24/26	1/26
pNtLTP1	18/23	23/23	22/23	23/23	0/23

[a] The GUS expression frequency was determined as indicated in the Material and methods among independent lines of transgenics plants expressing the *GUS-VENUS* reporter under the control of the indicated promoters.

[b] GUS activity was observed in tall glandular trichomes (TGT), short glandular trichomes (SGT), leaves (besides trichomes), roots, and stems (besides trichomes) of 6-week-old T0 plants.

Figure legends

Fig. 1 Transcript levels in different organs of *N. tabacum*. Normalized transcript levels were determined as indicated in the Material and methods on the indicated organs. Results are shown as mean ± SD of three to five repeats. Different letters indicate significant differences according to a Kruskal-Wallis test ($p < 0.05$) followed by a Tukey post hoc test.

Fig. 2 Transcript levels in trichomes isolated from *N. tabacum* leaves at different developmental stages. Normalized transcript levels were determined as indicated in the Material and methods. St: leaf developmental stage. Stage 1: leaf length < 2.5 cm; stage II: leaf length between 2.5 cm and 6.5 cm; stage III: leaf length between 6.5 cm and 15 cm; stage IV: leaf length > 15 cm. Results are shown as mean ± SD of three repeats. Different letters indicate significant differences according to a Kruskal-Wallis test ($p < 0.05$) followed by a Tukey post hoc test.

Fig. 3 Absolute transcript levels at stage III of leaf development in *N. tabacum*. Absolute transcript levels were determined as indicated in the Material and methods. Results are shown as mean ± SD of three repeats. Different letters indicate significant differences according to a Kruskal-Wallis test ($p < 0.05$) followed by a Tukey post hoc test.

bioRxiv preprint doi: https://doi.org/10.1101/620021; this version posted May 22, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Fig. 4 Molecular constructs used to generate transgenic *N. tabacum* expressing the *GUSVENUS* reporter gene under the control of trichome-specific promoters. The transcription promoter regions of *NtRbcS-T1* (MG493459.1), *NtMALD1* (MG493458.1), *NsCBTS2α* (HM241151.1), *NtLTP1* (AB625593.1), *NtCYP71D16* (AF166332.1), and *NtCPS2* (HE588139.1) were amplified and cloned as described in the Material and methods.

Fig. 5 Specific GUS activity in trichomes of *N. tabacum*. GUS staining was performed in 6-week-old T1 lines.

Fig. 6 Diversity of GUS activity of different transgenic *N. tabacum* lines in leaf (A), stem (B), and root (C) tissues. GUS staining was performed in 6-week-old T1 lines.

bioRxiv preprint doi: https://doi.org/10.1101/620021; this version posted May 22, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

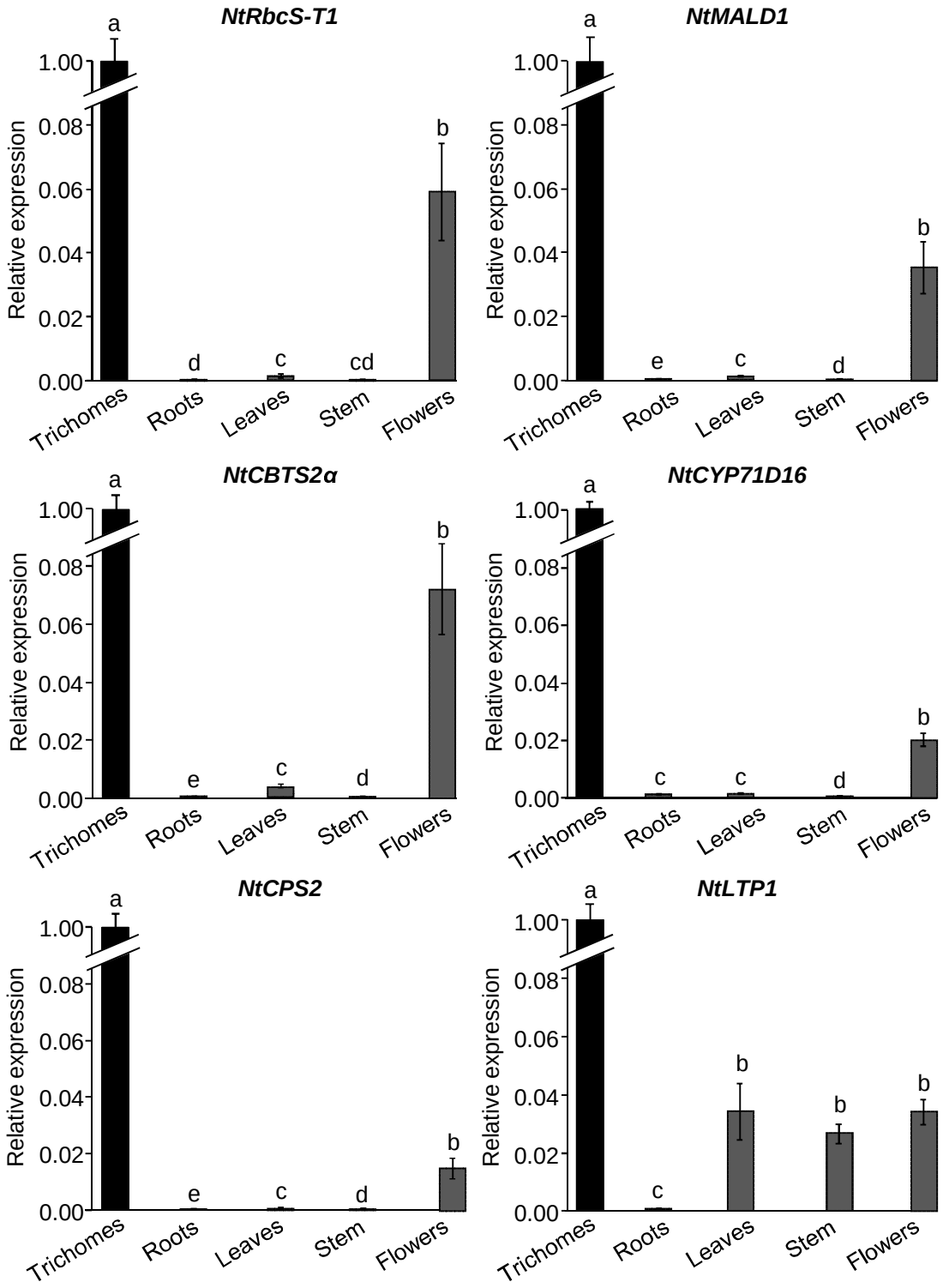

Fig. 1 Transcript levels in different organs of *N. tabacum*. Normalized transcript levels were determined as indicated in the Material and methods on the indicated organs. Expression was measured as described in the Material and methods Results are shown as mean ± SD of three to five repeats. Different letters indicate significant differences according to a Kruskal-Wallis test ($p < 0.05$) followed by a Tukey post hoc test.

bioRxiv preprint doi: https://doi.org/10.1101/620021; this version posted May 22, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

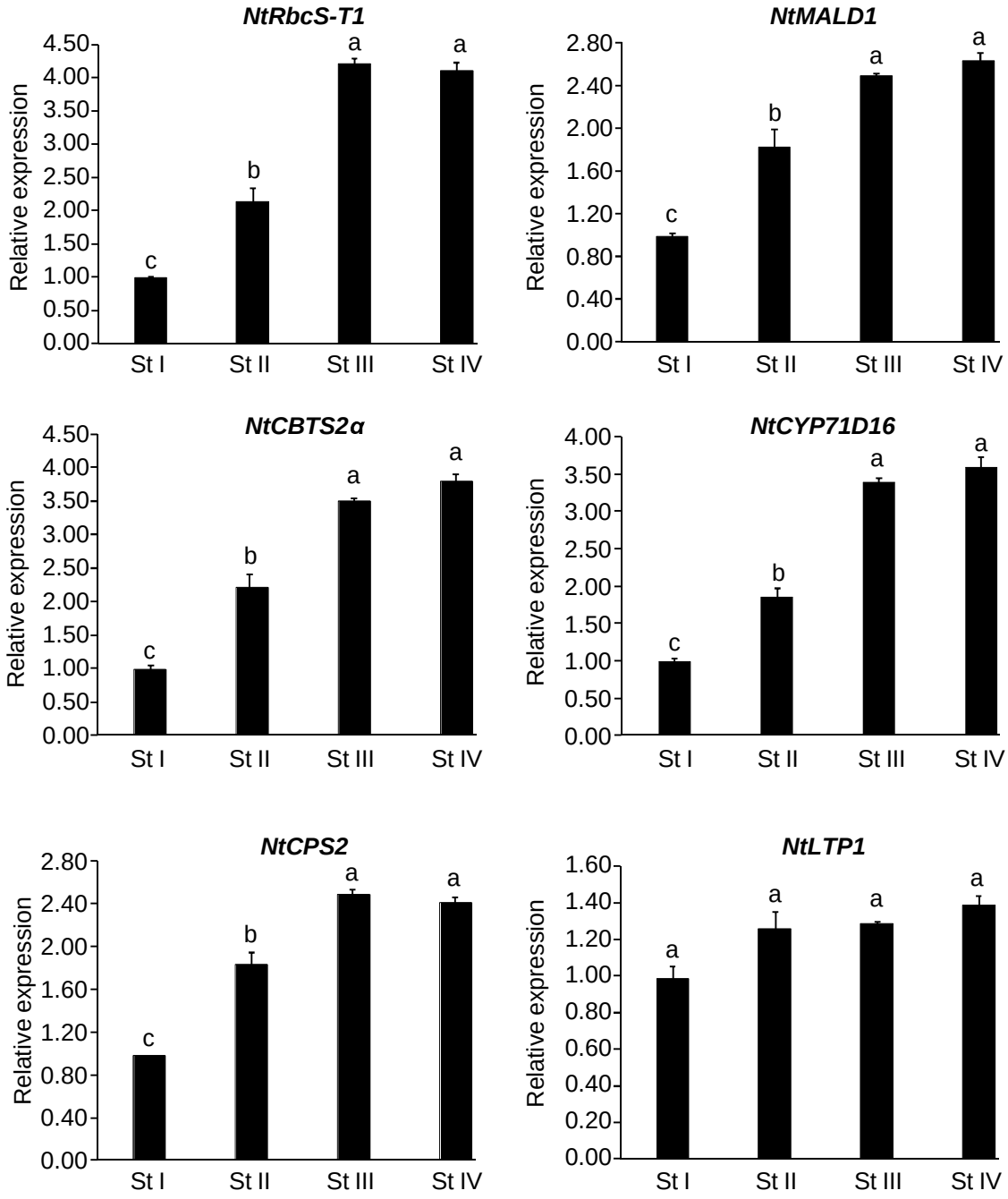

Fig. 2 Transcript levels in trichomes isolated from *N. tabacum* leaves at different developmental stages. Normalized transcript levels were determined as indicated in the Material and methods. St: leaf developmental stage. Stage 1: leaf length < 2.5 cm; stage II: leaf length between 2.5 cm and 6.5 cm; stage III: leaf length between 6.5 cm and 15 cm; stage IV: leaf length > 15 cm. Results are shown as mean ± SD of three repeats. Different letters indicate significant differences according to a Kruskal-Wallis test ($p < 0.05$) followed by a Tukey post hoc test.

bioRxiv preprint doi: https://doi.org/10.1101/620021; this version posted May 22, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

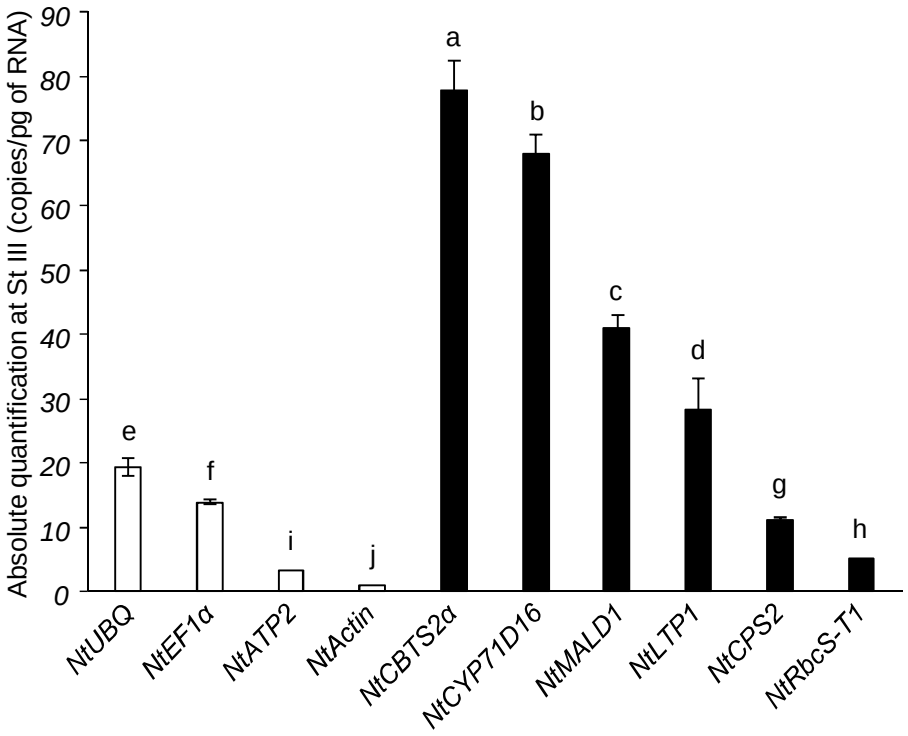

Fig. 3 Absolute transcript levels at stage III of leaf development in *N. tabacum*. Absolute transcript levels were determined as indicated in the Material and methods. Results are shown as mean ± SD of three repeats. Different letters indicate significant differences according to a Kruskal-Wallis test ($p < 0.05$) followed by a Tukey post hoc test.

bioRxiv preprint doi: https://doi.org/10.1101/620021; this version posted May 22, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

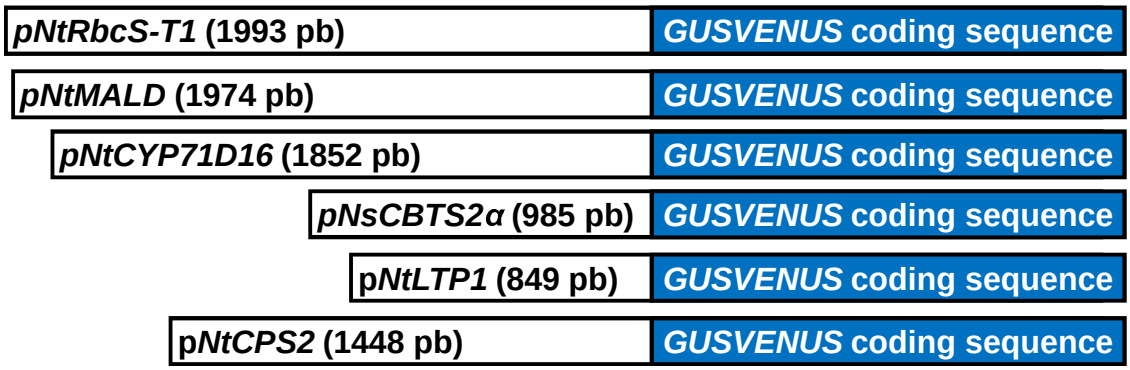

Fig. 4 Molecular constructs used to generate transgenic *N. tabacum* expressing the *GUSVENUS* reporter gene under the control of trichome-specific promoters. The transcription promoter regions of *NtRbcS-T1* (MG493459.1), *NtMALD1* (MG493458.1), *NsCBTS2α* (HM241151.1), *NtLTP1* (AB625593.1), *NtCYP71D16* (AF166332.1), and *NtCPS2* (HE588139.1) were amplified and cloned as described in the Material and methods.

bioRxiv preprint doi: https://doi.org/10.1101/620021; this version posted May 22, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

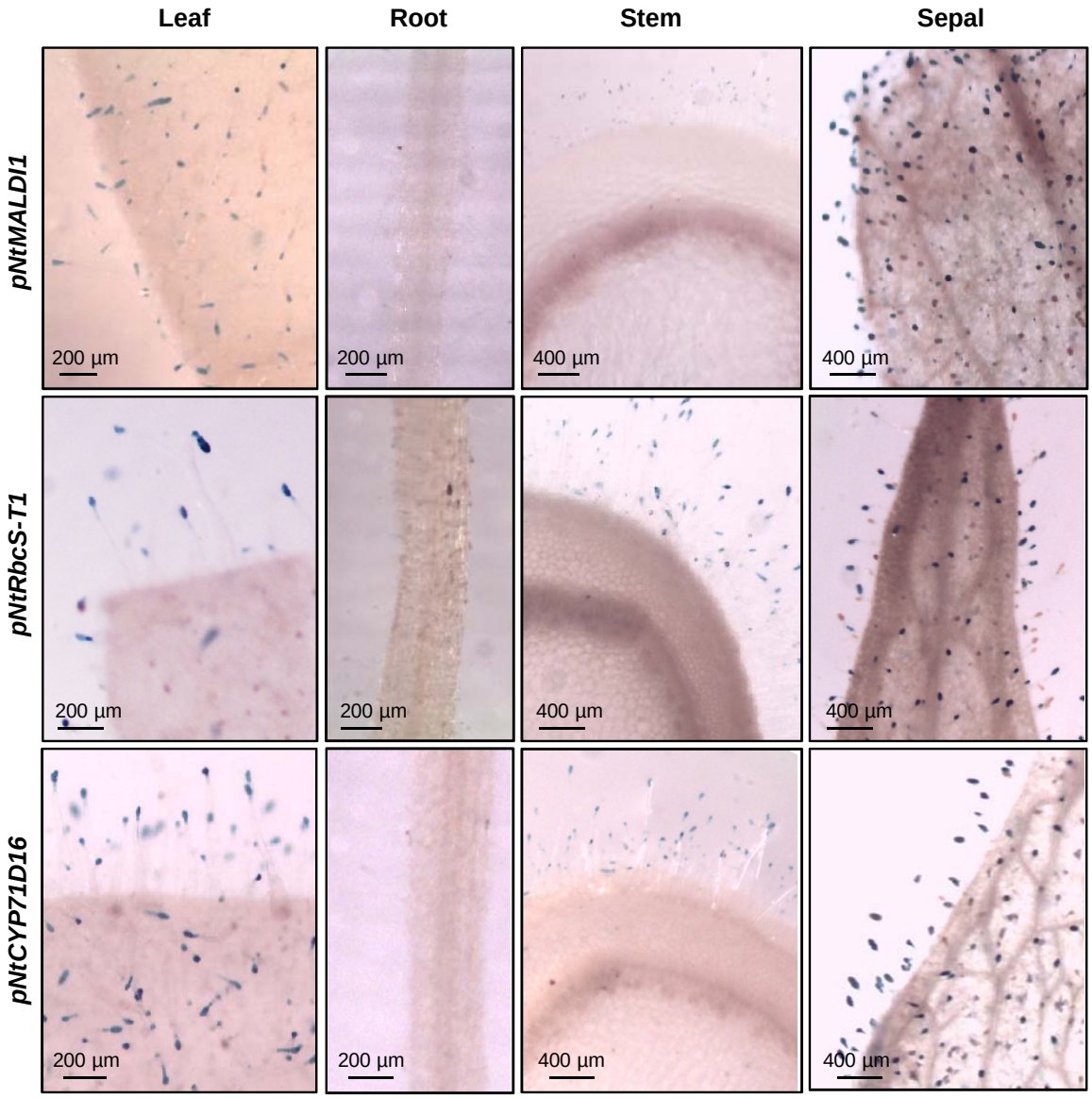

Fig. 5 Specific GUS activity in trichomes of *N. tabacum.* GUS staining was performed in 6-week-old T1 lines.

bioRxiv preprint doi: https://doi.org/10.1101/620021; this version posted May 22, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

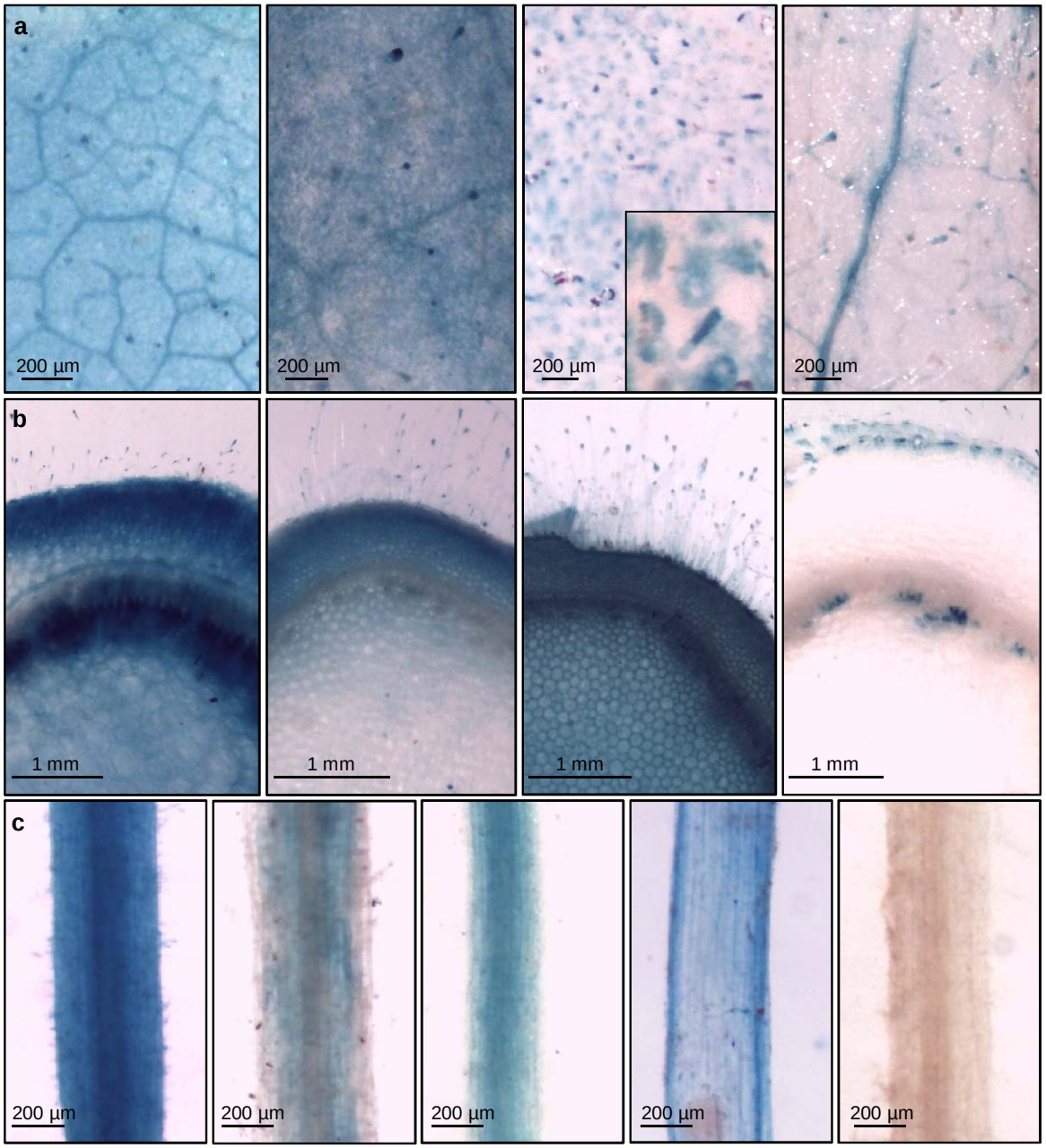

Fig. 6 Diversity of GUS activity of different transgenic *N. tabacum* lines in leaf (a), stem (b), and root (c) tissues. GUS staining was performed in 6-week-old T1 lines.