Plant Science 328 (2023) 111573

Contents lists available at ScienceDirect

Plant Science

journal homepage: www.elsevier.com/locate/plantsci

Engineering *Nicotiana tabacum* trichomes for triterpenic acid production

Nicola Gossart [a], Alice Berhin [a], Kjell Sergeant [b], Iftekhar Alam [a,c], Christelle André [b,d], Jean-François Hausman [b], Marc Boutry [a], Charles Hachez [a,*]

[a] Louvain Institute of Biomolecular Science and Technology, UCLouvain, Louvain-la-Neuve, Belgium
[b] Environmental Research and Innovation, Luxembourg Institute of Science and Technology, Esch-sur-Alzette, Luxembourg
[c] Plant Biotechnology Division, National Institute of Biotechnology, Ganakbari, Ashulia, Savar, Dhaka 1349, Bangladesh
[d] The New Zealand Institute for Plant and Food Research Ltd (PFR), Private Bag 92169, Auckland, New Zealand

ARTICLE INFO

Keywords:
Glandular trichomes
Nicotiana tabacum
Metabolic engineering
Terpenes

ABSTRACT

In this work, we aimed at implementing the biosynthesis of triterpenic acids in *Nicotiana tabacum* glandular trichomes. Although endogenous genes coding for enzymes involved in such biosynthetic pathway are found in the *Nicotiana tabacum* genome, implementing such pathway specifically in glandular trichomes required to boost endogenous enzymatic activities. Five transgenes coding for a farnesyl-diphosphate synthase, a squalene synthase, a squalene epoxidase, a beta-amyrin synthase and a beta-amyrin 28-monooxygenase were introduced in *N. tabacum*, their expression being driven by *pMALD1*, a trichome-specific transcriptional promoter. This study aimed at testing whether sinking isoprenoid precursors localized in plastids, by exploiting potential cross-talks allowing the exchange of terpenoid pools from the chloroplast to the cytosol, could be a way to improve overall yield. By analyzing metabolites extracted from entire leaves, a low amount of ursolic acid was detected in plants expressing the five transgenes. Our study shows that the terpene biosynthetic pathway could be, in part, redirected in *N.tabacum* glandular trichomes with no deleterious phenotype at the whole plant level (chlorosis, dwarfism,…). In light of our results, possible ways to improve the final yield are discussed.

1. Introduction

Terpenes and their derivatives are the most abundant class of secondary metabolites in plants (Croteau et al., 2000). These lipophilic molecules represent one of the most structurally diverse classes of natural products. Their synthesis is mediated by prenyltransferases that generally start with one unit of dimethylallyl diphosphate (DMAPP) and add one or more isopentenyl diphosphate (IPP) units.

While about 80.000 terpenes have been identified so far and clustered based on their number of C5 isoprene units, many more still remain to be characterized (Zhou and Pichersky, 2020). Monoterpenes are formed by the condensation of two isoprene units (C10), sesquiterpenes of three (C15), diterpenes of four (C20) while triterpenes are formed by the condensation of six isoprene units (C30).

In plants, a balanced terpene profile is important for both generalized and specialized metabolism. Terpenes perform a variety of functions and their biosynthesis is under tight metabolic control to optimize the plant fitness, for example by forming a chemical barrier against pathogens (bacteria, fungi) and herbivore attacks, and by helping attract pollinators (Ninkuu et al., 2021).

As some terpenes of economic interest have a complex stereochemistry and are produced in limited amount in their natural host, research has focused on trying to improve their production by metabolic engineering using a variety of chassis organisms (Lange and Ahkami, 2013).

In the early 2000s, significant advances were made in *E. coli* by exploiting the methylerythritol phosphate (MEP) pathway (Schmidt--Dannert et al., 2000). This pathway produces the IPP/DMAPP intermediates from pyruvate and glyceraldehyde-3-phosphate. Further improvements in yield (from μg/l to mg/l) were obtained when the mevalonate pathway (MVA pathway), which produces IPP/DMAPP from acetyl-CoA, was fully implemented in *E. coli* using yeast genes (Martin et al., 2003). Overexpression of 1-deoxy-D-xylulose-5-phosphate (DXP) synthase gene (DXS) and DXP reductoisomerase gene (DXR), the first two genes of the MEP pathway, from *Bacillus subtilis* in *E. coli* then led to a drastic increase in yield, with more than 300 mg/l of isoprene (Zhao et al., 2011).

Despite these advances, the use of bacteria as chassis organisms has quickly become limiting because many enzymes involved in the

* Corresponding author.
E-mail address: charles.hachez@uclouvain.be (C. Hachez).

https://doi.org/10.1016/j.plantsci.2022.111573
Received 10 August 2022; Received in revised form 13 December 2022; Accepted 15 December 2022
Available online 21 December 2022

biosynthesis of terpenes or responsible for further modification of these compounds, such as cytochromes P450, require an internal membrane system which is not present in bacteria (Wu et al., 2012). Even if it is possible to express some plant cytochromes P450 in *E. coli*, to obtain a protein with a proper folding and membrane topology remained challenging (Chang et al., 2007).

Because of these limitations, yeast could appear as a promising alternative in the metabolic engineering of terpenes. However, yeast does not possess an active MEP pathway and only produces, via the MVA pathway, a limited amount of IPP/DMAPP, which is then mostly used to sustain the biosynthesis of sterols, in particular ergosterol, the most important sterol in yeast. By using mutated yeast capable of assimilating exogenous ergosterol, high yields of sesquiterpenes (more than 50 mg/l) were achieved (Ro et al., 2006,Takahashi et al., 2007). Further modification of the regulation of both the MVA pathway and sterol biosynthesis allowed to obtain up to 25 g/l of artemisinic acid, a sesquiterpene precursor of artemisinin, an anti-malaria drug (Paddon et al., 2013).

In plants, two different biosynthetic pathways are found that generate two different pools of IPP/DMAPP precursors. These two pools are produced in different subcellular compartments and are used to drive the biosynthesis of different classes of terpenes. Indeed, while the MVA pathway produces a cytosolic pool of IPP/DMAPP, the MEP pathway forms a plastidial pool of these precursors. Mono- and diterpenes are mostly produced using the MEP-derived plastidial pool of IPP/DMAPP, while sesqui- and triterpenes are mostly produced using MVA-derived precursors. This is however a simplified view. Indeed, although both pathways are physically separated and function independently, cross-talks were reported in *Nicotiana tabacum* (Hemmerlin et al., 2003), *Daucus carota* (Hampel et al., 2005) and *Lavandula latifolia* (Mendoza--Poudereux et al., 2015). This means that precursors used in the synthesis of a specific terpene could actually be derived from both pathways.

For example, even though sesquiterpenes are synthesized in the cytosol, they could potentially be synthesized using a plastidial pool of IPP/DMAPP precursor (Bartram et al., 2006, Dudareva et al., 2005). Similarly, the IPP pool used to sustain the biosynthesis of mono- and diterpenes, that generally occurs in plastids, can also be of cytosolic origin (MVA-derived), and not just plastidial, as originally thought (Hampel et al., 2007, Itoh et al., 2003, Wungsintaweekul and De-Eknamkul, 2005).

The cross-talks between the MVA and MEP pathways are generally described as an exchange of prenyl diphosphates between the cytosol and the plastids (reviewed by Hemmerlin et al. (2012)). It involves exchange of IPP/DMAPP (Hemmerlin et al., 2003; Bick and Lange, 2003) but also geranylgeranyl diphosphate (Gerber et al., 2009), geranyl diphosphate (Adam et al., 1999) and some farnesol metabolites (Huchelmann et al., 2016). The precursor exchange mechanism has not been discovered yet and the actual extent of these cross-talks is actually not totally clear. Other terpenoid compounds could indeed be concerned by such cross-talks. For example, leakage of squalene across the chloroplast envelope was recently reported (Zhao et al., 2018). Could such observed "leakage" be mediated by a transporter in the chloroplast envelope? While active transport of terpenoids across plastidial membranes has been postulated, this principle is not sufficiently understood to generalize it. Identification and characterization of proteins involved in the transport of precursors would help shedding light on these exchange processes (Huchelmann et al., 2017). However, simple diffusion of the hydrophobic products across the plastidial membranes into the ER could be fully sufficient in many scenarios.

Our work focuses on the biosynthesis of triterpenic acids. Many triterpenic acids such as ursolic acid and oleanolic acid present interesting pharmaceutical properties such as antitumoral activity (Liu, 2005) and beneficial effects on the cardiovascular system (Somova et al., 2003). In plants, several enzymes, located in the cytosol, are required to form such triterpenes from IPP/DMAPP. The first one, farnesyl diphosphate synthase (FDS) is involved in the synthesis of farnesyl diphosphate (FPP, C15) from one molecule of DMAPP and two molecules of IPP which are grafted onto it. This IPP/DMAPP to FPP reaction takes place in two steps, the first one forming a GPP intermediate on which the second IPP is grafted to form the FPP (Davis and Croteau, 2000). In *Arabidopsis thaliana*, FPP does not accumulate in the cytosol and appears to serve as a regulator for the MVA pathway (Closa et al., 2010). The second enzyme involved in the synthesis of triterpenes is squalene synthase (SQS). This enzyme is anchored in the ER membrane with the catalytic site facing the cytosol (Linscott et al., 2016). SQS catalyzes the condensation of two FPP molecules into squalene (C30) (Nakashima et al., 1995). Squalene will then serve as a substrate for squalene epoxidase (SQE) to form 2,3-oxidosqualene, the last common intermediate for all plant sterols and triterpenes (Thimmappa et al., 2014).

Oxidosqualene can be cyclized by an oxidosqualene cyclase (OSC) into sterols or triterpenes (Thimmappa et al., 2014). Some OSCs are multifunctional, in the way that they can generate several different scaffolds. For example, baruol synthase 1 (BARS1) from *Arabidopsis thaliana* forms 23 different products, with baruol as the major (90%) one (Lodeiro et al., 2007). More than one hundred triterpene scaffolds have already been identified in plants (Xu et al., 2004). These scaffolds can then be modified by enzymes (for example sugar transferases, CYP450, etc.) to synthesize new triterpenoids. All this explains the great diversity of existing triterpenoids.

N. tabacum is often used as a model species to study terpene metabolism in plants. In *N. tabacum*, sesquiterpenes are produced from IPP/DMAPP synthesized mainly via the MVA pathway (Back and Chappell, 1996, Ralston et al., 2001, Starks et al., 1997) while diterpenes are mainly formed from a MEP-derived pool of precursors (Keene and Wagner, 1985, Kandra and Wagner, 1988, Guo and Wagner, 1995, Wang and Wagner, 2003). Even though both MEP and MVA pathways are actively producing IPP/DMAPP, most terpenes precursors come from the plastidial MEP pathway. Naturally, *N. tabacum* produces up to 100 μg/g of fresh weight (FW) of the sesquiterpene capsidiol (Dokládal et al., 2012) and up to 5 mg/g $_{FW}$ of diterpenes (Sallaud et al., 2012).

Given that it naturally produces terpenes, *N. tabacum* appears as an appropriate chassis organism to produce more or new terpenes using metabolic engineering approaches. To be economically interesting, metabolic engineering of terpenes in *N. tabacum* should reach yields close to those naturally occurring in plants currently used to extract these compounds (Huchelmann et al., 2017). However, by using constitutive and ubiquitous promoters, yields reach, at best, a few μg/g $_{FW}$ (Lücker et al., 2004, Wei et al., 2004, Farhi et al., 2011).

Considering, as mentioned previously, cross-talks between the cytosolic MVA and the plastidial MEP pathways, new metabolic engineering strategies emerged that aimed at implementing a biosynthetic pathway to sink the most abundant pool of IPP/DMAPP directly in the subcellular compartment of interest. *N. tabacum* produces a large amount of diterpenes in the plastids, where an abundant IPP/DMAPP pool is found. Therefore, provided with a transit peptide, a farnesyl diphosphate synthase and an appropriate terpene synthase were targeted to plastids and redirected the metabolic flux in these organelles. Such strategy increased synthesis of the sesquiterpenes patchoulol and amorpha-4,11-diene more than 1000-fold as well as that of the monoterpene limonene between 10 and 30-fold (Wu et al., 2006). However, along with these yield improvements, severe adverse phenotypes were observed, such as dwarfism and chlorosis (Wu et al., 2006). Such detrimental phenotypes were also observed in the production of artemisinic acid (Saxena et al., 2014) and ginsenoside saponin (Gwak et al., 2017). As terpene precursors are extremely important for the growth and development of plants (especially for chlorophyll and phytohormones synthesis; Lichtenthaler, 1999), the problem probably arose from the use of constitutive and ubiquitous transcription promoters and the resulting depletion of essential terpene precursors (IPP/DMAPP) in the whole plant. To limit side effects, a better control of transgene expression is needed. One strategy to improve this aspect, and to make metabolic engineering more efficient, is to use a cell-type specific promoter, to direct the biosynthesis to specific cells, group of cells, or organs, such

as trichomes.

Trichomes are unicellular or multicellular outgrowths derived from the protoderm. They are found on the surface of aerial plant tissues. There are several trichome classification systems, but usually two major classes are distinguished: glandular and non-glandular trichomes. In a single plant, different types of trichomes, both glandular and/or non-glandular can be found. Glandular trichomes, observed on almost 30% of all vascular plant species, are characterized by the presence of one or more glandular cells able to synthesize, store and/or secrete various secondary (also called specialized) metabolites which are involved in the adaption of the plants to their environment (Huchelmann et al., 2017). Secondary metabolites synthesized by glandular trichomes include, among others, acyl sugars (Schilmiller et al., 2010; Kroumova and Wagner, 2003) and mainly terpenes (Gershenzon and Dudareva, 2007; Gershenzon et al., 1992).

A key and unique feature of glandular trichomes is their ability to synthesize and secrete large amounts of a limited number of metabolites relative to their size. In *N. tabacum*, up to 17% of the leaf dry weight comes from molecules secreted by glandular trichomes (Wagner et al., 2004). Metabolic engineering in trichomes is therefore particularly interesting. Many studies have already been conducted in tomato (Kortbeek et al., 2016, Wang et al., 2022) and in tobacco, producing up to 100 μg/g $_{FW}$ of diterpenes (Tissier et al., 2012, Zhang et al., 2018) and up to 1760 μg/g $_{FW}$ of triterpenes (Wu et al., 2012). To obtain such yields, certain authors used a hybrid promoter, combining a trichome-specific promoter (the cembratrien-ol synthase or the cembratriene-ol hydroxylase promoter) and a transcription enhancer from the CaMV 35 S promoter fused to the 5' terminus. By choosing to use a hybrid promoter, the authors were able to obtain higher yields than those obtained with trichome-specific promoters (Tissier et al., 2012). However, deleterious phenotypes (chlorosis and dwarfism) have been observed for the plants with the highest yields, explained by a lack of tissue specificity of the transcriptional promoters used in the constructs.

In the present effort, our objectives are threefold. The first is to redirect the metabolic flux of terpenes to produce triterpenic acids in *N. tabacum* while limiting detrimental effect at the whole plant level. Our second objective is to evaluate the potential of the *pMALD1* promoter, a newly identified trichome-specific promoter, for metabolic engineering purposes (Pottier et al., 2020). Our last objective is to evaluate the impact of various subcellular localizations of the enzymes (plastidial or cytosolic) to produce triterpenic acids from the existing pools of precursors.

2. Materials and methods

2.1. Genetic constructs for stable plant transformation

The auxiliary plasmid - binary plasmid system of Goderis et al. (2002) was used as basis for all constructions. The promoter region of *NtMALD1* (1974 bp; GenBank accession: MG493458.1) (Pottier et al., 2020) was inserted into pAUX3132, pAUX3133, pAUX3169 by restriction with *Not*I and *Kpn*I. The NOS terminator from *Agrobacterium tumefaciens* (Navarre et al., 2011) was inserted into these pAUX by a *Bam*HI/*Age*I restriction.

The *Arabidopsis thaliana* farnesyl-diphosphate synthase (*AtFPS2*) gene (FDS, AGI code: AT4G17190), as well as the squalene synthase (*ScERG9*) (SQS, Gene ID: 853086) and squalene epoxidase (*ScERG1*) (SQE, Gene ID: 853086) genes from *Saccharomyces cerevisiae* were ordered as synthetic DNA blocks (Integrated DNA Technologies, Supplemental Data 1). For *ScERG9*, a full length or truncated version (a soluble version of the protein without the 3 part coding for the last 24 amino acid residues forming a membrane domain (Lograsso et al., 1993)) were ordered. These synthetic DNA sequences were tagged at the 3' end with a 6 histidines, FLAG or Myc tag, respectively. When chloroplastic localization of the recombinant protein was required, the plastid targeting (TP) sequence (first 177 nucleotides) of *N. tabacum* ribulose bisphosphate carboxylase oxygenase small subunit (LOC107766567) was inserted directly upstream of the initiating ATG of the corresponding coding sequence. Synthetic DNA blocks containing the transit peptide (if present), *ScERG1*, *ScERG9* (truncated or full-length version) or *AtFPS2* coding sequence and a tag were placed in *pAUX3132*, *3133* or *3169* (respectively) containing the *pMALD1* promoter and *tNOS* terminator by restriction with *Kpn*I and *Bam*HI. By restriction with the corresponding homing endonucleases (Goderis et al., 2002), the expression cassettes were then inserted into a *pPZP-RCS2-nptII* plant expression vector already containing the *pMALD1-Gus/Venus-tNOS* expression cassette (Pottier et al., 2020). Three versions were prepared to have constructs in *pPZP* plasmid sending one (AtFPS2), two (AtFPS2 and ScERG9) or all three enzymes (AtFPS2, ScERG9 and ScERG1) to the plastids (Supplemental Fig. 1A-C).

The coding sequences of *Malus domestica beta-amyrin synthase MdOSC1* (Gene ID: LOC103443360) and of *M. domestica beta-amyrin 28-monooxygenase MdCYP90-like* (Gene ID: LOC103453334) were amplified from plasmids (Andre et al., 2016). A hemagglutinin (HA) tag was added in 3' of the *MdOSC1* coding sequence while a Streptavidin (Strep) tag was added in 3' of the *MdCYP90-like* coding sequence. The tagged coding sequences were inserted into *pAux vectors* containing the *pMALD1* promoter and *tNOS* terminator by restriction with *Asc*I and *Pac*I. By restriction with the corresponding homing endonucleases, the expression cassettes were then inserted into a *pPZP-RCS2-hpt* plant expression vector containing a *pMALD1-mCherry-tNOS* expression cassette (Supplemental Fig. 1D).

A. thaliana hydroxy methylglutaryl CoA reductase 1 (*AtHMG1*, also called HMGR, AGI code: AT1G76490) and *A. thaliana* deoxyxylulose-5-phosphate synthase (*AtCLA1*, also called DXS, AGI code: AT4G15560) coding sequences were amplified from *A. thaliana* leaf cDNA and inserted into a pAux3133 containing the pMALD1 promoter and tNOS terminator by restriction with *Kpn*I and *Bam*HI. Using the corresponding homing endonuclease, each expression cassette was then inserted into a *pPZP-RCS2-hpt* plant expression vector.

2.2. Genetic constructs for transient plant transformation

The *AtFPS2*, *ScERG9*, *ScERG1* coding sequences (with or without the TP sequence, Supplemental Data 1) as well as *MdOSC1* and *MdCYP90-like* coding sequences were PCR amplified and inserted in frame with GFP under control of a 35 S promoter in pMDC83 vector (Curtis et al., 2003). A *HDEL-CFP* construct was used as endoplasmic reticulum marker (Napier et al., 1992).

2.3. Plant material and growth conditions

N. tabacum cv Petit Havana SR1 (Maliga et al., 1973) was used for stable plant transformation. For in vitro cultures, seeds were surface sterilized by exposing them to chlorine gas for 45 min. The sterilized seeds were then placed on solid Murashige and Skoog (MS) medium [4.4 g/l MS basal medium (MP biochemicals: www.mpbio.com), 30 g/l sucrose, 8 g/l agar, pH 5.6 (KOH)] and cultured in a growth chamber (25 °C, 16 h photoperiod, 50 μmol photon/m^2s).

For plant propagation and phenotyping purposes, seedlings from in vitro cultures were placed in Jiffy pots (Gronud: www.jiffypot.com) for a week, before being transferred to pots containing potting soil (DCM: dcm-info.com) and placed in a phytotron (25 °C, 16 h photoperiod, 200 μmol photon/m^2s).

2.4. Generation of transgenic plant lines

Electrocompetent *A. tumefaciens* strain GV3101 was transformed by electroporation (400 Ω, 1.25 kV) using 50 ng of plasmid DNA. The constructs were introduced into *A. tumefaciens* GV3101 for subsequent *N. tabacum* leaf disk transformation (Horsch et al. 1989). For each

construct, 15–20 independent transgenic lines were generated and finally transferred to soil (25 °C, 16 h photoperiod, 50 μmol photon/m^2s). All T1 transgenic lines presumably possessed multiple T-DNA insertion events in their genomes as no 1:3 lethality ratio was observed when plated on herbicide-containing MS medium. Transgenic lines (T0 generation) were further selected based on the expression pattern of the fluorescent reporters (Venus and mCherry). Only lines showing trichome-specific fluorescent reporter signals were selected and further analyzed, T0 lines displaying a broader expression pattern of the reporters due to position effects of the T-DNA in the genome were discarded.

2.5. Transient transformation via agroinfiltration of leaves

Transformed GV3101 cells were used for agroinfiltration of *N. benthamiana* or *N. tabacum* leaves following the protocol described in (Hachez et al., 2014). The agroinfiltrated zones were harvested after 4 days for observations under the confocal microscope and after 7 days for UPLC analysis.

2.6. Confocal microscopy

All the observations were carried out according to standard imaging procedures on a LSM710 confocal microscope (Zeiss: www.zeiss.com) equipped with a spectral detector.

2.7. Plant tissue isolation

Trichomes from T1 lines over-expressing three (*AtFPS2/ScERG9/ScERG1*) or five (*AtFPS2/ScERG9/ScERG1/MdOSC1/MdCYP90-like*) transgenes were isolated according to the protocol of Wang et al. (2001). Wild type lines were used as control. For each T1 line, trichomes from three 10-week-old plants were collected and pooled as one biological replicate.

Leaf disks (d ≈ 1 cm, one disc ≈ 50 mg) from 10-week-old plants (T0 or T1 lines) over-expressing three (*AtFPS2/ScERG9/ScERG1*) or five (*AtFPS2/ScERG9/ScERG1/MdOSC1/MdCYP90-like*) transgenes and lines over-expressing HMGR were used for chromatographic analyses. Two leaf discs were collected per mature leaf (larger than 15 cm) on three leaves per plant. Three 10-week-old plants were analyzed per line. After flash freezing in liquid nitrogen, all tissues (trichomes and leaves of *N. tabacum*) were ground using glass beads (d = 5 mm, Sigma-Aldrich) and an MM 400 stirrer (Retsch; 30 Hz, 2 * 30 s).

2.8. RNA extraction, and cDNA synthesis

RNA extraction was carried out on tissues from WT (SR1 ecotype) or T1 transgenic lines, using the Spectrum Plant Total RNA Kit (Sigma-Aldrich) with on-Column DNase I Digestion (Sigma-Aldrich). Moloney Murine Leukemia Virus Reverse transcriptase (Promega: be.progema.com) and oligo$(dT)_{18}$ were used for the reverse transcription of 500 μg of DNA-free RNA according to the manufacturer's protocol.

2.9. Gene expression analysis

All RT-qPCR assays were carried out on RNA extracted from isolated trichomes (except for Supplemental Fig. 4 where RNA extracts from whole leaves were used instead).

The primers used for RT-qPCR (supplemental Table 1) were designed using Primer3 (https://bioinfo.ut.ee/primer3/) according to the protocol described by (Thornton and Basu, 2011). qPCRs were performed within a volume of 20 μL, comprising 1 μL of cDNA (corresponding to 40 ng of RNA), 2.2 μL of each primer (3 μM) and 10 μL of the Power SYBR green PCR master mix (Eurogentec: https://www.eurogentec.com). An ABI 7500 Real-Time PCR system (Waltham: www.thermofischer.com) was used for amplification. The relative expression levels were determined based on the $2^{-\Delta\Delta CT}$ method (Livak and Schmittgen, 2001) using the geometric mean of 3 reference genes for normalization (*ELONGATION FACTOR α* (*NtEF1a, AF120093*), *ACTIN* (*NtTac9, X69885*) and *UBIQUITIN-CONJUGATING ENZYME E2* (*NtUbc2, AB026056*) (Vandesompele et al., 2002). The primers and reference genes used were chosen according to the recommendations of Schmidt and Delaney (2010).

2.10. Triterpene extraction and GC/UPLC-MS analysis

Leaf tissue (100 mg) from T0 lines expressing AtFPS2, ScERG9 and ScERG1 were first freeze-dried for 16 h before extraction. Fresh tissue from leaves was used as starting material for all other constructs.

For the extraction, approximately 100 mg of tissue was mixed with 1 mL of methanol (sample-solvent ratio of 1:10) by vortexing for 30 s, sonicated in continuous mode for 10 min and shaken at room temperature for 4 h. After a centrifugation step (10,000 *g* for 15 min), the supernatant was recovered and stored at 4 °C.

Stable T0 and T1 lines over-expressing HMGR or AtFPS2/*ScERG9/ScERG1* or/and *MdOSCI/MdCYP90-like* were analyzed by gas chromatography-mass spectrometry (GC-MS), while ultra-performance liquid chromatography - mass spectrometry (UPLC-MS) was used for plants transiently overexpressing AtFPS2-GFP, ScERG9-GFP and ScERG1-GFP.

For GC-MS analysis, 200 μL of extracts were evaporated to dryness under nitrogen gas and resuspended with 100 μL of N,O-Bis (trimethylsilyl)trifluoroacetamide with trimethylchlorosilane (99/1). To allow derivatization, samples were kept at 60 °C during 30 min. Gas chromatography coupled with a triple quadrupole mass spectrophotometer (Trace GC-TSQ Quantum, Thermo Fisher Scientific) was used in this study. The injector temperature was programmed to start at 50 °C for 5 s ramped to 250 °C at 14.5 °C/min and held for 20 min. Injection volume was 1 μL in splitless mode. Compounds were separated with RXi-5Sil MS column (20 m, 0.18 mm ID, 0.18 μm df, Restek). Helium was used as gas carrier and set at a constant flow rate of 1 mL/min. Oven initial temperature was 175 °C (for 2 min) and increased to 300 °C at 15 °C/min and held for 6 min. For MS/MS, electron ionization (EI) mode was selected, and the electron voltage was 70 eV. Collision inductive dissociation (CID) was assisted with argon gas. Two ion transitions for Selected Reaction Monitoring (SRM) of each compound were determined as follows: Ursolic acid: TIC SRM ms2:203.140 [105.109–105.111; 133.169–133.171], squalene: TIC SRM ms2:136.11 [93.159–93.161;121.539–121.541].

For UHPLC-HRMS analysis, extracts were evaporated to dryness using a centrifugal vacuum evaporator (Centrivap, Labconco). Dried extracts were made up in 600 μL of methanol:water (95:5, v/v), sonicated 10 min and filtered through a syringe filter (0.2 μm, PTFE Millex-LG (Merck KGaA). Ultra-high pressure liquid chromatography (Acquity UPLC I-class ultra-high pressure liquid chromatography (UHPLC)) coupled with a triple quadrupole time-of-flight mass spectrometer (Triple TOF 6600, Sciex) was used in this study in positive mode. The column and the sample temperatures were 40 °C and 4 °C, respectively. Injection volume was 5 μL. Compounds were separated with a reverse-phase Acquity UPLC BEH C18 column (2.1 × 100 mm, 1.7 μm particle size, Waters). Solvents used were water + 0.1% formic acid (solvent A) and acetonitrile + 0.1% formic acid (solvent B). Solvents were used following linear gradient (Supplemental Table 2). For MS1, Duospray ion source was used with the following parameters: 30 psi for curtain gas, 55 psi for gas1, 50 psi for gas2, 4500 V for ion spray voltage and 650 °C for source temperature. Mass range (m/z) was between 100 and 2000 and parameters were 15 V for collision energy and 60 V for declustering potential. For MS2 accumulation, mass range (m/z) was between 50 and 2000 and parameters were 15 V for collision energy, 15 V for collision energy spread and 60 V for declustering potential.

For GC and UPLC analysis, data processing was conducted using Xcalibur software (Thermo Fisher Scientific). Quantification was based

on external calibration curve obtained with standards purchased from Merck and Extrasynthese.

3. Results

Nicotiana tabacum has an active secondary metabolism and large amounts of diterpenes are produced in the glandular heads of long-stalked trichomes from the pool of IPP/DMAPP produced by the MEP pathway (Huchelmann et al., 2017). Unlike diterpenes, triterpenic acids are synthesized in the endoplasmic reticulum from the pool of IPP/DMAPP produced by the MVA pathway. To implement the biosynthesis pathway of triterpenic acids from IPP in *N. tabacum*, five enzymes are needed (AtFPS2, ScERG9, ScERG1, MdOSC1 and MdCYP90-like). The first three enzymes, which synthesize oxidosqualene from IPP/DMAPP, are common for all triterpenes while the two last enzymes specifically turn oxidosqualene into specific triterpenic acids (ursolic and oleanolic acids).

Actually, endogenous *FDS*, *SQS* and *SQE* genes are found in the *N. tabacum* genome. Our idea was to complement endogenous activities to overcome possible rate limiting steps in the pathway and to test whether sinking isoprenoid precursors localized in plastids, by exploiting potential cross-talks allowing the transport of terpenoids from the chloroplast to the cytosol, could be a way to improve overall yield.

3.1. Transient overexpression of AtFPS2, ScERG9 and ScERG1 in plastids

In order to use the plastidial isoprenoid flux derived from the MEP pathway (which is the most active in *N. tabacum*), we targeted all three enzymes (AtFPS2, ScERG9 and ScERG1) to the plastids by fusing them to a transit peptide.

While AtFPS2 is a soluble enzyme that can be easily addressed to the stroma of the chloroplasts and remain functional (Bhatia et al., 2015, Wu et al., 2006), ScERG9 and ScERG1 are membrane proteins that naturally localize to the endoplasmic reticulum. While ScERG9 possesses one transmembrane domain located at its C-terminus, ScERG1 possesses two transmembrane domains at its N- and C- termini. Interestingly, it is possible to truncate the last 24 amino acid residues of the C-terminal end of ScERG9, forming a membrane anchoring domain, to obtain a soluble SQS without loss of activity (Lograsso et al., 1993), while no functional characterization of such truncated version has been reported for ScERG1. As a first line of study, AtFPS2, truncated ScERG9 and ScERG1 were transiently expressed in leaves with a plastid targeting (TP) sequence to determine if it is possible to accumulate a significant amount of oxidosqualene in this compartment. The three constructs, each containing the coding sequence of those three enzymes, fused to a TP sequence and to GFP and placed under control of a 35 S promoter, were co-infiltrated in *N. benthamiana* leaves using *Agrobacterium*-mediated transformation. Unlike the commercial standard of 2,3-oxidosqualene, the chromatograms of the samples (Fig. 1) have several peaks whose exact mass is 427.3940 Da (ESI (+)). These different peaks presumably correspond to different conformations of 2,3-oxidosqualene. However, only the peak at 21.13 min was considered as it is the only one that could be validated by MS-MS with the commercial standard of 2,3-oxidosqualene (Supplemental Fig. 1). In infiltrated leaves, we obtained a concentration of up to 12.3 μg/g fresh weight (FW) of extracted oxidosqualene compared to 0.15 μg/g $_{FW}$ in non-infiltrated plants (WT), corresponding to an 82-fold increase in oxidosqualene accumulation.

In the literature, AtFPS2 and ScERG9 had already been targeted to the plastids successfully, showing an accumulation of squalene as a result of their overexpression (Wu et al., 2012). With this transient expression assay, we demonstrated that ScERG1 also remains active once sent to the plastids. These preliminary results showed that transiently overexpressing these three enzymes with a TP sequence increases the accumulation of oxidosqualene.

3.2. Moving the route into stably engineered transgenic plants

To determine the impact of the subcellular localization of these enzymes, we expressed, in a stable way, these enzymes under the control of the trichome-specific promoter, *pMALD*1 (Pottier et al., 2020). Three different configurations of the constructs, with different subcellular localizations of the recombinant proteins, were tested. The first configuration consisted of sending all three enzymes to the plastids (AtFPS2, truncated ScERG9 and ScERG1, Fig. 2 A).

As ScERG1 is an endoplasmic reticulum (ER)-localized membrane protein that could negatively impact the integrity of the plastid membrane compartments (Babiychuk et al., 2008), we kept, as a second line of study, the ER localization of ScERG1 and tested the influence of different subcellular localizations of ScERG9: in the plastids (C-terminal truncated form with a transit peptide, Fig. 2B) or in the ER (native protein, Fig. 2C). The architecture of the different constructs is described in Supplemental Fig. 2. The subcellular localizations of the recombinant proteins were confirmed by transient expression in *N. benthamiana* or in BY-2 cells (Fig. 3). Addition of a transit peptide to AtFPS2 resulted in a plastidic localization of the recombinant protein (Fig. 3A). When no

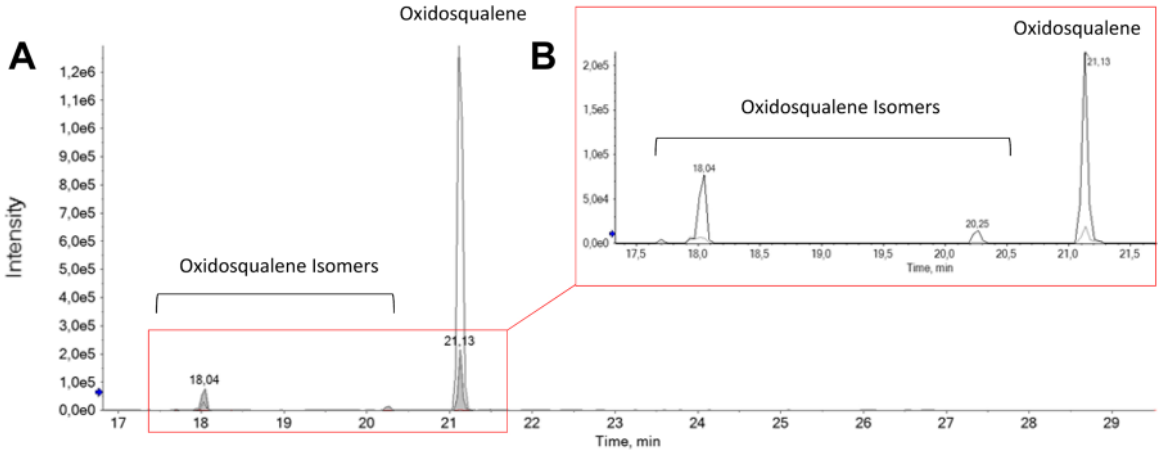

Fig. 1. Analysis of oxidosqualene production upon transient expression of *AtFPS2*, *ScERG9*, *ScERG1*. (A) Superposition of the chromatograms of a methanol extract of *N. benthamiana* leaves from non-transformed control (red line), a sample transiently expressing *AtFPS2*, ScERG9 and ScERG1 with a plastid targeting sequence (grey surface) and of an authentic standard of 2,3-oxidosqualene (10 μg/mL, black line). (B) Zoom on the chromatogram of the non-transformed control and the transiently transformed sample shown in A. Non-infiltrated control (grey line), transformed sample (black line). Final concentrations of extracted oxidosqualene are 0.15 μg/g $_{FW}$ (non-infiltrated control) and 12.4 μg/g $_{FW}$ (transformed sample). The chromatograms are for selected ions at *m/z*: 427.394 + /- 0.005.

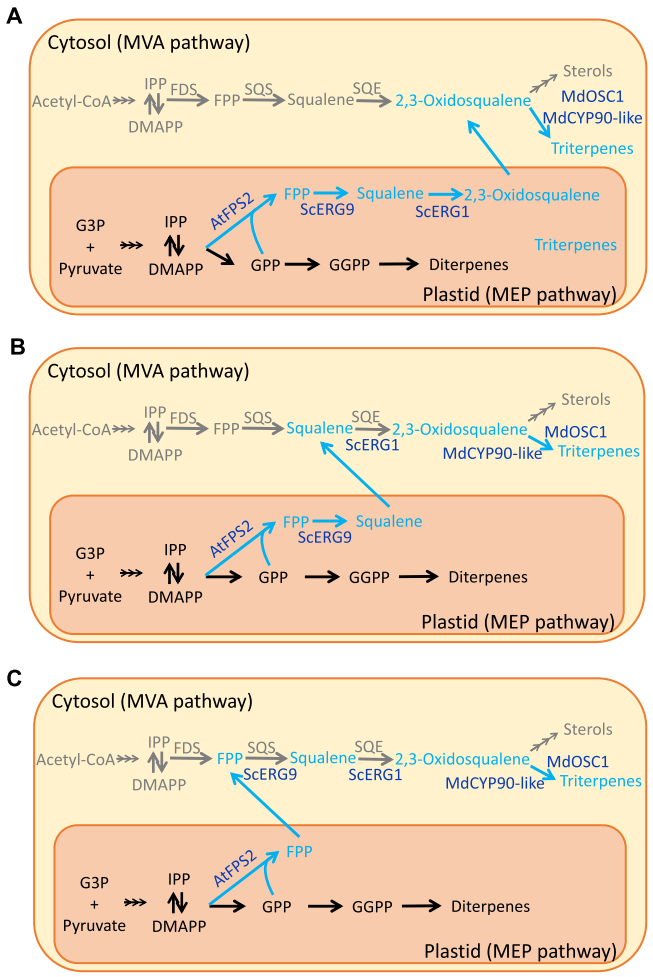

Fig. 2. Terpene metabolism in engineered *N. tabacum* trichomes. The grey pathway represents the normal biosynthetic route for triterpenes and sterols. Overexpressed enzymes are denoted in dark blue. New products deriving from metabolic engineering are denoted in light blue. (A) Oxidosqualene production in plastids. All the enzymes are targeted to directly produce 2,3-oxidosqualene in the plastids. (B-C) Oxidosqualene production using a plastidial AtFPS2 to sink the plastidial pool of IPP/DMAPP precursors. In this configuration, both AtFPS2 and ScERG9 (B) or only AtFPS2 (C) are targeted to the plastids to enhance the cross-talks and sink the plastids from its precursors. The rest of the pathway takes place in the cytosol.

transit peptide is present, recombinant ScERG9 and ScERG1 proteins colocalized with the HDEL protein (ER marker) as expected (Fig. 3B and D, respectively). Addition of a transit peptide triggers a relocalization of these latter to the chloroplasts (Fig. 3C and E, respectively).

To determine if these transgenes were expressed in trichomes, their relative gene expression (RGE) was determined by RT-qPCR on RNA extracted from trichomes (Fig. 4A). In all the transgenic lines we analyzed, the *AtFPS2* expression level (0.5–1.2) was significantly higher than the one of *SQS* and *SQE* (RGE: 0.05–0.11, Fig. 4A).

In an attempt to compare the expression level of the *ScERG9* and *ScERG1* transgenes to the endogenous *SQS* and *SQE* genes, we identified the endogenous forms. Two endogenous *SQS* paralogs are found in the *Nicotiana tabacum* genome (LOC107790933 and LOC107791556 are identical at 90%) as well as six endogenous *SQE* genes present as 3 pairs of highly similar paralogs (LOC107828762 and LOC107771747: 93.5% identity, LOC107821041 and LOC107761131: 93.4% identity, LOC107831617 and LOC107786596: 82.5% identity). Primers in the conserved regions of each pair of endogenous *SQS* and *SQE* genes were designed and RT-qPCR assays were carried out on RNA extracted from isolated glandular trichomes (from WT plantas well as transgenic lines with different subcellular localizations of the recombinant AtFPS2, ScERG9 and ScERG1 proteins; Supplemental Fig. 3).

In the different transgenic lines we tested, the expression level of endogenous *SQS* genes was close to the one of ScERG9 (the *ScERG9 / endogenous SQS* ratio ranged between 0.97 and 2.68 depending on the line; Supplemental Fig. 3). Regarding endogenous *SQE* genes, only two pairs were found to be expressed in glandular trichomes (LOC107828762/ LOC107771747 and LOC107821041/LOC107786596). The total expression level of endogenous *SQE* genes is in the same range as the one of *ScERG1* (the *ScERG1/ endogenous SQE* ratio ranged between 0.21 and 1.70 depending on the analyzed line; supplemental Fig. 3). It is also worth noting that expression levels of endogenous SQS and SQE are similar in WT and transgenic lines. Altogether, *ScERG9* and *ScERG1* expression levels were induced in trichomes but limited at roughly the same expression level as endogenous orthologs whose expression was not affected in the transgenic lines.

By specifically expressing these three enzymes in the glandular heads of the trichomes, we did not observe any detectable accumulation of oxidosqualene in leaves, whatever the subcellular location of these enzymes. However, we observed a significant accumulation of squalene, the precursor of oxidosqualene (Fig. 4B). The concentration of squalene reached up to 17.7 μg/g dry weight (DW) of leaves in transgenic lines while WT plants had 9.7 $\mu g/g_{DW}$ of leaves which corresponds to a 1.8-fold increase in squalene accumulation compared to WT (Fig. 4B). Interestingly, we did not observe any significant difference between the three configurations tested neither any linear correlation between the expression level of the *AtFPS2* transgene and squalene accumulation in the transgenic lines (Supplemental Fig. 4). No modified phenotypes were observed in transgenic plants.

As FPP can be transported across the plastid envelope (Bick and Lange, 2003), it is actually quite likely that plastid-derived FPP, synthesized by the overexpressed AtFPS2, could partly be used by endogenous ER membrane-associated SQS and SQE and in consequence also contribute to squalene and oxidosqualene formation via the endogenous (cytosolic) pathway. It is indeed worth noting that endogenous SQS and SQE genes being expressed in trichomes, endogenous pathways probably contribute to routing products into tobacco sterols/native triterpenoids and provide part of the (oxido)squalene pool for further cyclization, oxidation and conjugation events.

3.3. From IPP/DMAPP to triterpenic acids in trichomes

The subcellular localizations of MdOSCI and MdCYP90-like were confirmed by transient expression in *N. benthamiana*. Colocalization with HDEL-CFP (ER marker) confirmed the ER localization of these recombinant proteins (Fig. 5).

After testing the subcellular localization of AtFPS2, ScERG9, ScERG1 in addition to the one of MdOSC1 and MdCYP90-like, plants expressing the five enzymes specifically in the head of the glandular trichomes were generated and selected. Given that the *GUS-VENUS* coding sequence is on the same T-DNA as *AtFPS2*, *ScERG9* and *ScERG1* (Supplemental Fig. 2A-C) and that the *mCherry* coding sequence is on the same T-DNA as *MdOSC1* and *MdCYP90-like* (Supplemental Fig. 2D) and that these seven coding sequences were individually placed under control of the same *pMALD1* trichome-specific promoter (Pottier et al., 2020), confocal microscopy was used to monitor the expression pattern of mVENUS and mCherry and to select appropriate lines (Supplemental Fig. 2E). Only lines presenting trichome-specific Venus and mCherry signals were selected. Our assumption was that, given the fact the same promoter was used to drive each individual expression cassette, the cellular expression pattern of the two reporters (GUS-VENUS and mCherry) should mirror the cellular expression pattern of the different transgenes.

Concerning the subcellular localizations, we kept the constructions where the AtFPS2 and the ScERG9 were targeted to the plastids, ScERG1 being either in the plastids or in the ER (Fig. 2A and 2B). Regarding OSCs and CYP450, these proteins are known to be localized in the ER

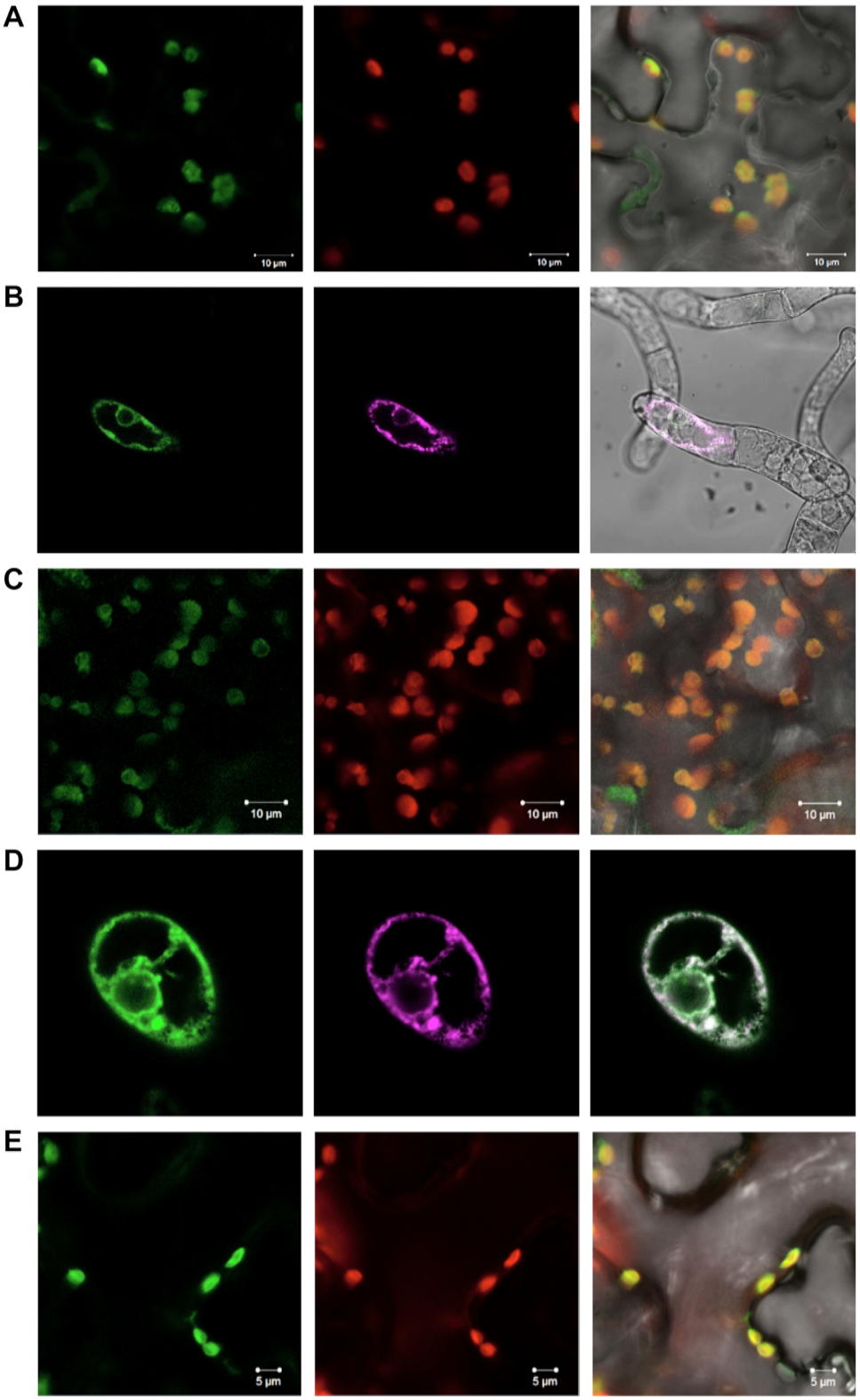

Fig. 3. Subcellular localization of AtFPS2 from *Arabidopsis thaliana*, ScERG1, and ScERG9 from *Saccharomyces cerevisiae* with or without a plastid targeting sequence. (A) *N. benthamiana* plant cells transiently expressing a fusion protein of GFP (green fluorescent protein) and AtFPS2 with a plastid targeting sequence (green). The red signal corresponds to chlorophyll autofluorescence (chloroplasts). Both images are merged in the right-hand panel. (B) *N. tabacum* BY-2 cells transiently expressing a fusion protein of GFP and the yeast squalene synthase (ScERG9) (green) and a fusion protein of CFP (cyan fluorescent protein) and HDEL, an endoplasmic reticulum marker (magenta). Both images are merged in the right-hand panel. (C) *N. benthamiana* plant cells transiently expressing a fusion protein of GFP and ScERG9 with a plastid targeting sequence (green). Red signals correspond to chlorophyll autofluorescence (chloroplasts). Both images are merged in the right-hand panel. (D) *N. tabacum* BY-2 cells transiently expressing a fusion protein of GFP and the yeast squalene epoxidase (ScERG1) (green) and a fusion protein of CFP and HDEL (magenta). Both images are merged in the right-hand panel. (E) *N. benthamiana* plant cells transiently expressing a fusion protein of GFP and ScERG1 with a plastid targeting sequence (green). The red signal corresponds to chlorophyll autofluorescence. Both images are merged in the right-hand panel.

membrane (Thoma et al., 2004, Hasemann et al., 1995, Jozwiak et al., 2020). Because these enzymes use electron transporters (NADPH) localized in the ER membrane for the formation of the acid function, among other things (Fukushima et al., 2011, Upadhyay et al., 2015), they were not targeted to the plastids.

One of our objectives was to determine the potential of *pMALD1* for metabolic engineering in *N. tabacum* trichomes. A previous study showed specific activity of this promoter in trichomes (Pottier et al., 2020), but the authors did not determine the expression level of transgenes using this promoter. To get a better idea of the activity of the *pMALD1* promoter, we measured the RNA expression levels of the five transgenes specifically in isolated glandular trichomes (Fig. 6 A). Although the transgene expression level conferred by the *MALD1* promoter is gene-specific, all five transgenes were found to be expressed,

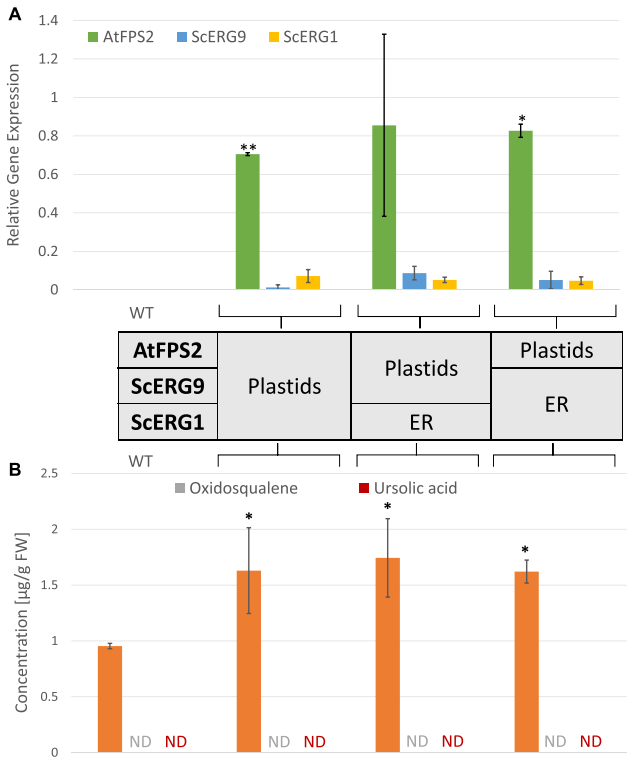

Fig. 4. (A) RT-qPCR for T1 transgenic lines expressing *AtFPS2, ScERG9* and *ScERG1* with different subcellular localizations in isolated trichomes. For each combination of constructs, two independent lines were analyzed (trichomes from three plants per line were collected and pooled prior to RNA extraction). The subcellular localization of the different recombinant proteins (AtFPS2, ScERG9, ScERG1) is indicated at the bottom, either in the endoplasmic reticulum (ER) or in plastids. Gene expression levels are expressed relative to that corresponding to the geometric mean of three different internal control genes (*NtEF1a, NtTac9, NtUbc2*). (B) Squalene and oxidosqualene accumulation in leaves from T0 plants expressing *AtFPS2, ScERG9* and *ScERG1*. AtFPS2, ScERG9 and ScERG1 are localized in plastids or in the ER, under the trichome-specific promoter pMALD1. For each combination of construct, means were obtained from five randomly chosen independent T0 lines. Results obtained by GC-MS/MS. ND: not detected (below the GC-MS detection limit). Results are displayed as mean± SD. Asterisks indicate significant differences determined using Student's t-test: *$p < 0.05$, * *$p < 0.01$.

with a relative gene expression between 0.1 and 0.8 compared to the one of the reference genes (Fig. 4 A and Fig. 6 A). However, these expression levels are relatively low compared to the one of endogenous *NtMALD1*, which has a relative gene expression level up to 12 (Supplemental Figure 6).

Regarding the metabolites, we observed a significant 2.2-fold decrease in the quantity of squalene in three lines which expressed the five transgenes specifically in the trichomes, regardless of the subcellular location of ScERG1 (down to 1.2 µg/g $_{FW}$, while WT plants have 2.6 µg/g $_{FW}$; Fig. 6B). Given that the transgenes were specifically expressed in the glandular heads of the trichomes, while extracts for GC-MS analysis were prepared from whole leaf discs, this is a big difference.

Concerning the triterpenic acids, as *MdOSCI* is a multifunctional enzyme that produces a 5:1 ratio of α-amyrin to β-amyrin (Andre et al., 2016), we looked at the concentrations of ursolic and oleanolic acids. These triterpenic acids correspond to the triterpene acids of α- and β-amyrin respectively. Of the eighteen plants analyzed, we detected a small quantity of ursolic acid (~ 2.1 µg/g $_{FW}$) in three plants (2 plants when ScERG1 was in plastids and one plant when in the ER; Fig. 6B and Fig. 6 C). The concentration was above the limit of detection defined as the mean+ 3 SD of the signal baseline (Knoll, 1985, Fig. 6C). Ursolic acid concentration was quantified based on the standard calibration curve (Supplemental Fig. 7A). No olanolic acid was detected as it was below the limit of detection (Fig. 6C). Besides squalene, 2,3-oxidosqualene, oleanolic and ursolic acid (as shown in Fig. 6), we also attempted to measure the levels of campesterol, stigmasterol, β-sitosterol, β-amyrin and α-amyrin on whole leaf extracts from T1 transformants, using GC-MS/MS. No significant difference was observed in term of sterol profile/accumulation in the different T1 lines we tested: the levels of campesterol, stigmasterol and b-sitosterol were similar to those of WT lines (Supplemental Table 3). No accumulation of oxidosqualene, β-amyrin or α-amyrin was detected in WT or T1 transgenic lines as the concentration of these terpenes was below the limit of detection in our GC-MS/MS assays (Supplemental Table 3).

No modified phenotypes were observed in these transgenic plant lines. No ursolic acid was detected in WT lines (Fig. 6C) or in lines co-expressing only the first three transgenes (*AtFPS2*, *ScERG9* and *ScERG1*). As a result, ursolic acid production was specifically due to the activity of the recombinant MdOSCI and MdCYP90-like enzymes when expressed in addition to recombinant AtFPS2, ScERG9 and ScERG1. We did not test the presence of ursolic acid upon co-expression of *MdOSC1* and *MdCYP90-like* in a wild type background and may not exclude ursolic acid could be formed in this case (probably at an even lower amount than the one we observed in this study).

3.4. Boosting the precursors: overexpression of HMGR in trichomes

As one of the possibilities to increase the overall yield is to increase the availability of IPP/DMAPP precursors, overexpression of rate limiting enzymes for IPP/DMAPP biosynthesis, such as *AtHMG1/HMGR* (Leivar et al., 2005), in the glandular heads of trichomes for the MVA pathway was attempted. *HMGR* overexpression in trichome glandular heads caused a modest but significant 1.2-fold accumulation of squalene (up to 3.7 µg/g $_{FW}$, while WT plants have 3.2 µg/g $_{FW}$; Supplemental Fig. 5). No modified phenotypes were observed in transgenic lines.

4. Discussion

In the present effort, our objectives were threefold. The first objective was to test the feasibility of redirecting the metabolic flux of terpene biosynthesis to produce triterpenic acids in *N. tabacum*. Our second objective was to evaluate the impact of various subcellular localizations (plastidial or cytosolic) of the enzymes to produce triterpenic acids most efficiently from the existing pools of IPP/DMAPP precursors. Our last objective was to evaluate the potential of a newly identified trichome-specific promoter, *pMALD1*, for metabolic engineering purposes (Pottier et al., 2020). Given these objectives, our study initially establishes the first three biosynthetic steps to the intermediate oxidosqualene in a transient assay, re-targeting the enzymes to the plastids and with that demonstrating the principle. Moving the route into stably engineered transgenic plants with the trichome specific promoter and adding the last two enzymes resulted in detectable accumulation of the oxidized triterpenoid, albeit at low level.

4.1. Redirecting the metabolic flux of terpenes

A *de novo* biosynthesis of oxidosqualene, the last common intermediate of triterpenic acids, was attempted in *N. tabacum* glandular trichomes. When AtFPS2, truncated ScERG9 and ScERG1 were transiently expressed in plastids of *N. benthamiana* leaves, an 82-fold increase in oxidosqualene, compared to non-infiltrated plants, was observed (Fig. 1).

However, when the three transgenes were stably expressed in *N. tabacum* trichomes, we did not detect accumulation of oxidosqualene. This result was expected since SQE is a rate-limiting enzyme in the sterol synthesis pathway (Asadollahi et al., 2010; Satoh et al., 1993; Jennings et al., 1991). Another possibility could be that oxidosqualene was

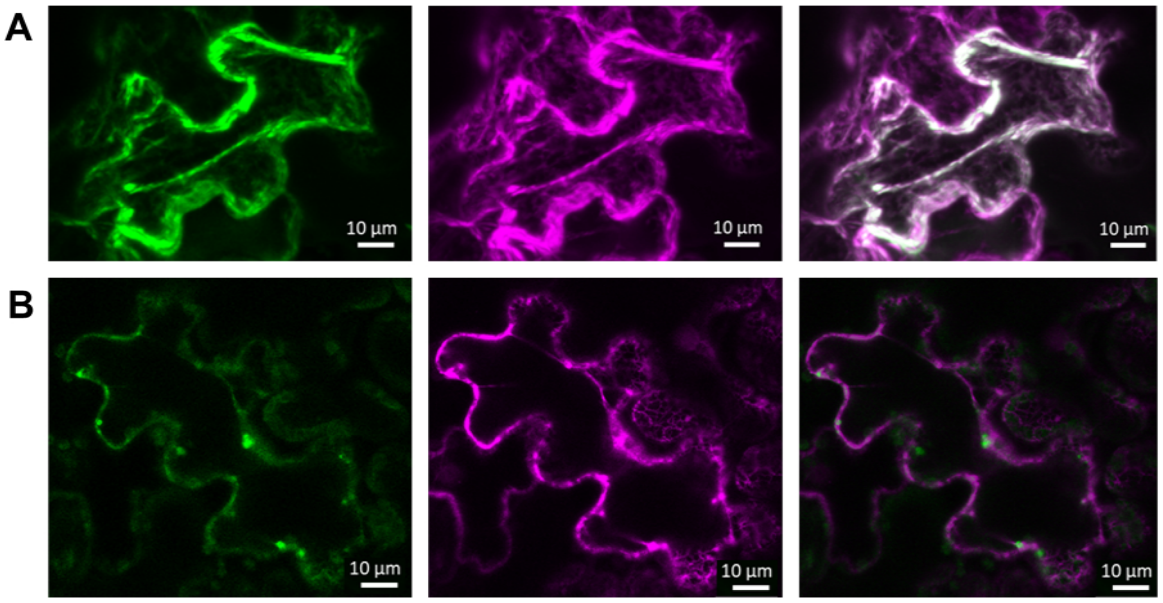

Fig. 5. Confocal fluorescence microscopy images of leaf tissue of *N. benthamiana* expressing transiently the fusion proteins (A) p35S:MdCYP90-like-eGFP (left) and p35S:HDEL-CFP (middle) or (B) p35S:MdOSCI-eGFP (left) and p35S:HDEL-CFP (middle)after *Agrobacterium*-mediated infiltration. On the right panels, superposition of fluorescence signals allowing the visualization of the colocalization of the p35S:MdCYP90-like-eGFP or p35S:MdOSCI-eGFP protein with HDEL-CFP, a marker of the endoplasmic reticulum.

converted to triterpenes and/or sterols by endogenous enzymes.

Therefore, we decided to look at squalene, synthesized upstream of oxidosqualene in the biosynthesis pathway. Whatever the subcellular localization of the three enzymes, a 1.5–2-fold increase in squalene concentration was observed compared to WT controls (Fig. 4B). Compared to published results (up to 1760 µg/g $_{FW}$) by locating AtFPS2 and ScERG9 in plastids of *N. tabacum* trichomes (Wu et al., 2012), the observed accumulations seem relatively low. However, in Wuś study, the plants had many deleterious phenotypes (chlorosis, dwarfism, …) probably due to the lack of tissue specificity of the promoter used. Indeed, the authors used the *CBTS* promoter with additional enhancers of the 35 S promoter of cauliflower mosaic virus. In another study (Tissier et al., 2012), the same promoter was used with and without these enhancers and these latter indeed seemed to be responsible for the deleterious phenotypes. Therefore, one cannot rule out that the accumulation observed by Wu et al. (2012) did not only originate from specific trichome-born enzymatic activity, but also from activity in the entire leaf tissue considering that the activity of this *CBTS* promoter may not be restricted to trichomes (Jiang et al., 2016, Pottier et al., 2020).

The second part of the project consisted of expressing AtFPS2, ScERG9, ScERG1, MdOSC1 and MdCYP90-like necessary for the synthesis of triterpenic acids from the pool of IPP/DMAPP. First, we looked at the squalene accumulation in these transgenic plants. Interestingly, we observed a 2.2-fold decrease in the amount of squalene when all five transgenes were present (down to 1.2 µg/g $_{FW}$, while WT plants have 2.6 µg/g $_{FW}$), whereas an increase in squalene accumulation was observed when only the first three enzymes were overexpressed in trichomes. This decrease in squalene accumulation upon overexpression of the five transgenes in trichomes suggests squalene did not accumulate and was further metabolized into triterpenes upon MdOSCI and MdCYP90-like activity. Indeed, out of the eighteen plant lines analyzed by GC-MS, we detected a modest but significant accumulation of ursolic acid (a triterpenic acid of interest) in three lines, but no oleanolic acid. This indicates that our genetic constructs indeed trigger a redirection of the metabolic flux in glandular trichomes. As *MdOSCI* is a multifunctional enzyme that produces a 5:1 ratio of α-amyrin to β-amyrin (Andre et al., 2016), oleanolic acid, synthetized from β-amyrin, is probably produced, but below the detection limit. Ursolic acid production seemed to be specifically due to the activity of the recombinant MdOSCI and MdCYP90-like enzymes as no ursolic acid was detected in WT lines or in lines co-expressing only the first three transgenes (*AtFPS2*, *ScERG9* and *ScERG1*).

By overexpressing these five enzymes specifically in the glandular head of trichomes, we were able to redirect the metabolic flux of terpenes but only observed low yields of triterpenic acids. Further research is needed to investigate and resolve the bottlenecks impacting such biosynthesis.

However, while transgenes were specifically expressed in the glandular heads of trichomes, all chromatographic analysis were carried out on extracts from entire leaf disks where the bulk of the plant mass is not contributing to the biosynthesis of the compounds of interest. Therefore, although the accumulation may seem modest, these results show we were actually able to redirect the flux of terpenes in glandular trichomes.

4.2. Impact of subcellular localization

The bioavailability of existing pools of natural precursor molecules is controlled by so-called cross-talks between different biosynthetic pathways. As highlighted in Huchelmann et al. (2017), the exact nature and extent of such cross-talks are only partially understood. The second objective was therefore to test whether sinking precursors localized in plastids, by exploiting potential cross-talks allowing the transport of terpenoid precursors from the chloroplast to the cytosol, could be a way to improve overall yield. To do so, we decided to test the impact of different subcellular localizations of the enzymes on the yield in terpenoids of interest. Studies using constitutive promoters showed that yields were higher when enzymes were localized in plastids, which are metabolically active in leaves (Wu et al., 2006, Jiang et al., 2016). However, in stable transgenic lines we did not detect, a significant impact of the different subcellular localization of the enzymes (cytosolic vs plastidial) on the production of squalene or of triterpenic acids whether upon trichome-specific co-expression of AtFPS2, ScERG9 and ScERG1 only (Fig. 4B) or in conjunction with MdOSC1 and MdCYP90-like (Fig. 6B).

In all the transgenic lines we analyzed, the *AtFPS2* expression level was significantly higher than the one of *ScERG9* and *ScERG1* (Fig. 4A and Fig. 6A). It is hard to explain how the activity of the *pMALD1* promoter could be regulated by feedback mechanisms, especially given the

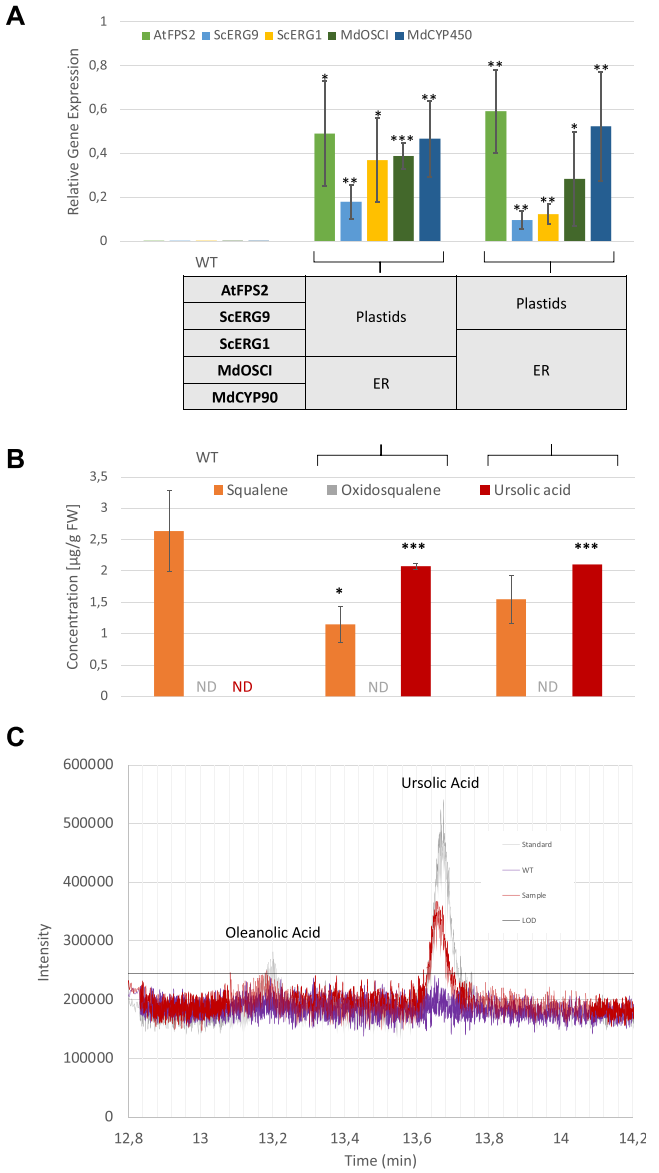

Fig. 6. (A) RT-qPCR for T1 transgenic lines co-expressing *AtFPS2*, *ScERG9*, *ScERG1*, *MdOSC1* and *MdCYP90-like* with different subcellular localizations in isolated glandular trichomes. For each combination of constructs, five independent lines were analyzed (trichomes from three plants per line were collected and pooled prior to RNA extraction). The subcellular localization of the different recombinant proteins (AtFPS2, ScERG9, ScERG1, MdOSC1 and MdCYP90-like) is indicated at the bottom, either in the endoplasmic reticulum (ER) or in plastids. Gene expression levels are expressed relative to that corresponding to the geometric mean of three different internal control (*NtEF1a, NtTac9, NtUbc2*). (B) Squalene and ursolic acid accumulation in T1 lines expressing *AtFPS2*, *ScERG9*, *ScERG1*, *MdOSC1* and *MdCYP90-Like* in trichomes. GC-MS analysis of squalene, oxydosqualene and ursolic acid accumulation (for each combination of constructs, three plants per line and three independent lines were analyzed). ND: below the limit of detection in GC-MS. (C) Chromatogram of authentic standards (500 µg/mL) of oleanolic acid (peak:13,22 min) and ursolic acid (peak:13,67 min) (grey trace), a sample prepared from a transgenic line expressing all five transgenes (red trace) and a wild type line (violet trace). The limit of detection (LOD, defined as the mean+3 SD of the signal baseline) for the sample is shown (black line). The chromatogram shows selected ions at *m/z*: 203.140 + /- 0.001. Results are displayed as mean ± SD. For ursolic acid, means were calculated only from plants where ursolic acid was detected. Asterisks indicate significant differences determined using Student's t-test: *p < 0.05, * *p < 0.01, * **p < 0.001.

fact that the expression levels of *ScERG9* and *ScERG1* were much more affected than the one of the *AtFPS2* transgene (despite the use of the same transcriptional promoter in the construct). The lower expression levels of *ScERG1* and *ScERG9* could be caused by the architecture of the construct used in this study and thus reflect an experimental artifact rather than actual feedback regulation mechanisms. Readthrough from the first transcript (AtFPS2) could cause reduced expression of downstream transgenes (*ScERG9* and *ScERG1*), especially since *tNOS* is not a very strong terminator.

Actually, given that FPP can be transported across the plastid envelope (Bick and Lange, 2003), plastid-derived FPP, synthesized by the plastid-targeted AtFPS2 could be used, thanks to existing cross-talks, by endogenous ER membrane-associated SQS and SQE and thereby accounts for the observed squalene accumulation via the endogenous cytosolic pathway. No linear correlation could be observed between the expression level of the *AtFPS2* transgene and squalene accumulation in the different transgenic lines we analyzed, squalene accumulation appearing to plateau around 1.5 $\mu g/g_{FW}$ in transgenic lines (Supplemental Fig. 4), a concentration much lower than the > 600 $\mu g/g_{FW}$ observed by Wu et al., 2012.

4.3. *pMALD1 for metabolic engineering purposes*

The last objective was to assess the potential of a newly identified trichome-specific promoter, pMALD1(Pottier et al., 2020), for metabolic engineering purposes.

We could confirm that pMALD1 promoter sequence used is appropriate to drive gene expression specifically in secretory cells of glandular trichomes (Supplemental Fig. 2) as reported in Pfigottier et al. (2020). This 2 kb promoter sequence was also sufficient to produce clearly detectable amounts of GUS-VENUS or mCherry proteins using confocal microscopy (Supplemental Fig. 2E) in the different selected lines.

The expression levels of the different transgenes are in the same range or slightly lower than the one of the control genes used to normalize the data (Fig. 4A and Fig. 6A). However, these expression levels are relatively low compared to the expression of *MALD1*, which has a relative gene expression up to 12 compared to the control genes (Supplemental Fig. 6). This could suggest that the 2 kb promoter sequence probably does not fully reflect the transcriptional activity of the endogenous *MALD1* promoter (due to a different genomic context (presence of enhancers) or the fact the 2 kb sequence cloned did not include all the necessary regulatory regions). In addition, we cannot exclude the existence of retroinhibition mechanisms (or even RNA silencing events) in our transgenic lines that would limit the expression levels of the transgenes, in a gene-specific way.

From a metabolic point of view, we have shown that it is possible, using this promoter, to accumulate squalene by overexpressing *AtFPS2*, *ScERG9* and *ScERG1* specifically in trichomes (up to 17.7 $\mu g/g_{DW}$; Fig. 4B) and some ursolic acid by overexpressing five transgenes. However, the concentrations obtained in this project (~ 2.1 $\mu g/g_{FW}$ of ursolic acid) remain low compared to the ones of endogenous diterpenes (~ 300 $\mu g/g_{FW}$ of diterpenes).

During the development of this project, we decided to express proteins from different species (*S. cerevisiae*, *M. domestica*, *A. thaliana*) whose activity had already been characterized and to try to avoid/limit endogenous negative feedback regulation mechanisms. Therefore, genes previously studied in the literature (Wu et al., 2012) or whose enzymatic activity had already been characterized (Andre et al., 2016) were used. This choice of enzymes could probably have been optimized in term of enzymatic activity as there is a possibility that more efficient enzymes (in term of enzymatic activity or selectivity profile) could be found.

As boosting the MVA pathway via overexpression of full length HMGR had only a modest impact on squalene accumulation (Supplemental Fig. 5), it might be interesting to test the effect of a truncated version of this protein (Chappell et al., 1995, Harker et al., 2003). N-terminally truncated HMGRs, lacking the membrane binding domain,

are soluble and are not subjected to retroinhibition making them significantly stronger in supporting the engineered isoprenoid pathways. Overexpression of either a truncated hamster HMGR or a truncated rubber tree HMGR in transgenic tobacco indeed resulted in sterol overaccumulation due to enhanced HMGR activities (Chappell et al., 1995; Harker et al., 2003).

As improperly terminated transcripts can serve as a basis for gene silencing, optimization of the choice of terminator used in the different constructs could be a way to increase the gene expression levels (Luo and Chen, 2007). In this project, the terminator of *nopaline synthase* from *A. tumefaciens* (tNOS) was used, which is less efficient than other terminators (de Felippes et al., 2020). Optimizing the choice of terminator might be another option.

Diterpenes are synthesized and continuously secreted via ABC transporters (Jasiński et al., 2001, Crouzet et al., 2013). In this project, we did not express any ABC transporter. Therefore, we cannot exclude that the metabolites were accumulated in the head of the glandular trichomes without getting secreted. When metabolites from leaf surface extracts prepared from T0 lines expressing the five transgenes were analyzed by GC-MS/MS, no significant difference compared to WT was detected regarding the terpenoid profile (Supplemental Figure 8). This suggests that the metabolites synthesized by the engineered pathway were not properly secreted. This could negatively impact their biosynthesis (e.g., cytotoxicity, activation of negative feedback regulation mechanisms) in comparison to a situation where, as for diterpenes, they would get secreted at the leaf surface.

Altogether those results suggest that using *pMALD1,* a trichome-specific promoter to drive the expression of *AtFPS2*, *ScERG9*, *ScERG1*, *MdOSC1* and *MdCYP90-like* we were able to partly redirect the metabolic flux of terpene biosynthesis towards the production of low amounts of ursolic acid, a triterpenic acid, in secretory cells of glandular trichomes in *Nicotiana tabacum*. However, further research is needed to investigate and resolve the bottlenecks impacting such biosynthesis to significantly improve the yield.

Funding

This work was supported by the Belgian National Fund for Scientific Research (PDR/PGY T.0116.20 to CH) and by the fund for research training in industry and agriculture (FRIA). NG also benefitted from funding from UCLouvain Special Research Funds (FSR 2017). IA was a recipient of a BELSPO postdoctoral fellowship.

Author contributions

NG, CA, MB and CH designed the project. NG, IA, KS, CA, AB performed all the experiments described in this work. NG, AB, KS and CH analyzed the data and created the figures, NG and CH wrote the manuscript with contributions from all other authors.

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Data Availability

Data will be made available on request.

Acknowledgments

We thank Emmanuelle Cocco (LIST), Sophie Charton (LIST) and Denis Pittois (LIST) for their contribution in chromatographic and mass spectrometric analyses, Hélène Laurent, Alexandre Huchelmann and Charles Claeys for their contributions to the initiation of this project as well as Joseph Nader for his technical advice. IA is grateful to Dr. Md. Salimullah for approving the deputation to avail the BELSPO postdoctoral fellowship.

Appendix A. Supporting information

Supplementary data associated with this article can be found in the online version at doi:10.1016/j.plantsci.2022.111573.

References

K.P. Adam, R. Thiel, J. Zapp, Incorporation of 1-[1-^{13}C]deoxy-D- xylulose in chamomile sesquiterpenes, Arch. Biochem. Biophys. 369 (1999) 127–132.
C.M. Andre, S. Legay, A. Deleruelle, N. Nieuwenhuizen, M. Punter, C. Brendolise, J. M. Cooney, M. Lateur, J.F. Hausman, Y. Larondelle, Multifunctional oxidosqualene cyclases and cytochrome P450 involved in the biosynthesis of apple fruit triterpenic acids, N. Phytol. 211 (2016) 1279–1294.
M.A. Asadollahi, J. Maury, M. Schalk, A. Clark, J. Nielsen, Enhancement of farnesyl diphosphate pool as direct precursor of sesquiterpenes through metabolic engineering of the mevalonate pathway in *Saccharomyces cerevisiae*, Biotechnol. Bioeng. 106 (2010) 86–96.
E. Babiychuk, P. Bouvier-Navé, V. Compagnon, M. Suzuki, T. Muranaka, V.A. N. Montagu, M. Kushnir, S, H. Schaller, Allelic mutant series reveal distinct functions for *Arabidopsis* cycloartenol synthase 1 in cell viability and plastid biogenesis, Proc. Natl. Acad. Sci. 105 (2008) 3163–3168.
K. Back, J. Chappell, Identifying functional domains within terpene cyclases using a domain-swapping strategy, Proc. Natl. Acad. Sci. 93 (1996) 6841–6845.
S. Bartram, A. Jux, G. Gleixner, W. Boland, Dynamic pathway allocation in early terpenoid biosynthesis of stress-induced lima bean leaves, Phytochemistry 67 (2006) 1661–1672.
V. Bhatia, J. Maisnam, A. Jain, K.K. Sharma, R. Bhattacharya, Aphid-repellent pheromone E-β-farnesene is generated in transgenic *Arabidopsis thaliana* over-expressing farnesyl diphosphate synthase2, Ann. Bot. 115 (2015) 581–591.
J.A. Bick, B.M. Lange, Metabolic cross talk between cytosolic and plastidial pathways of isoprenoid biosynthesis: unidirectional transport of intermediates across the chloroplast envelope membrane, Arch. Biochem. Biophys. 415 (2) (2003) 146–154.
M.C. Chang, R.A. Eachus, W. Trieu, D.K. Ro, J.D. Keasling, Engineering *Escherichia coli* for production of functionalized terpenoids using plant P450s, Nat. Chem. Biol. 3 (2007) 274–277.
J. Chappell, F. Wolf, J. Prouix, R. Cueller, C. Saunders, Is the reaction catalyzed by 3-hydroxy-3-methylglutaryl coenzyme A reductase a rate limiting step for isoprenoid biosynthesis in plants? Plant Physiol. 109 (1995) 1337–1343.
M. Closa, E. Vranová, C. Bortolotti, L. Bigler, M. Arró, A. Ferrer, W. Gruissem, The *Arabidopsis thaliana* FPP synthase isozymes have overlapping and specific functions in isoprenoid biosynthesis, and complete loss of FPP synthase activity causes early developmental arrest, Plant J. 63 (2010) 512–525.
R. Croteau, T.M. Kutchan, N.G. Lewis, Natural products (secondary metabolites), Biochem. Mol. Biol. Plants 24 (2000) 1250–1319.
J. Crouzet, J. Roland, E. Peeters, T. Trombik, E. Ducos, J. Nader, M. Boutry, NtPDR1, a plasma membrane ABC transporter from Nicotiana tabacum, is involved in diterpene transport, Plant Mol. Biol. 82 (1–2) (2013) 181–192.
E.M. Davis, R. Croteau, Cyclization enzymes in the biosynthesis of monoterpenes, sesquiterpenes, and diterpenes. Biosynthesis, Springer, 2000.
L. Dokládal, M. Oborił, K. Stejskal, Z. Zdráhal, N. Ptáčková, R. Chaloupková, J. Damborský, T. Kašparovský, S. Jeandroz, M. Žďárská, Physiological and proteomic approaches to evaluate the role of sterol binding in elicitin-induced resistance, J. Exp. Bot. 63 (2012) 2203–2215.
N. Dudareva, S. Andersson, I. Orlova, N. Gatto, M. Reichelt, D. Rhodes, W. Boland, J. Gershenzon, The nonmevalonate pathway supports both monoterpene and sesquiterpene formation in snapdragon flowers, Proc. Natl. Acad. Sci. 102 (2005) 933–938.
M. Farhi, E. Marhevka, J. Ben-Ari, A. Algamas-Dimantov, Z. Liang, V. Zeevi, O. Edelbaum, B. Spitzer-Rimon, H. Abeliovich, B. Schwartz, Generation of the potent anti-malarial drug artemisinin in tobacco, Nat. Biotechnol. 29 (2011) 1072.
D.E. Felippes, F. Mchale, M. Doran, R.L. Roden, S. Eamens, A.L. Finnegan, E. J, P. M. Waterhouse, The key role of terminators on the expression and post-transcriptional gene silencing of transgenes, Plant J. 104 (2020) 96–112.
E.O. Fukushima, H. Seki, K. Ohyama, E. Ono, N. Umemoto, M. Mizutani, K. Saito, T. Muranaka, CYP716A subfamily members are multifunctional oxidases in triterpenoid biosynthesis, Plant Cell Physiol. 52 (2011) 2050–2061.
E. Gerber, A. Hemmerlin, M. Hartmann, D. Heintz, M.A. Hartmann, J. Mutterer, M. Rodríguez-Concepción, A. Boronat, A. Van Dorsselaer, M. Rohmer, The plastidial 2-C-methyl-D-erythritol 4-phosphate pathway provides the isoprenyl moiety for protein geranylgeranylation in tobacco BY-2 cells, *Plant Cell* 21 (2009) 285–300.
J. Gershenzon, N. Dudareva, The function of terpene natural products in the natural world, Nat. Chem. Biol. 3 (2007) 408–414.
J. Gershenzon, D. Mccaskill, J.I.M. Rajaonarivony, C. Mihaliak, F. Karp, R. Croteau, Isolation of secretory cells from plant glandular trichomes and their use in biosynthetic studies of monoterpenes and other gland products, Anal. Biochem. 200 (1992) 130–138.
I.J. Goderis, D.E. Bolle, M.F. François, I.E. Wouters, P.F. Broekaert, W. F, B.P. Cammue, A set of modular plant transformation vectors allowing flexible insertion of up to six expression units, Plant Mol. Biol. 50 (2002) 17–27.

Z. Guo, G.J. Wagner, Biosynthesis of labdenediol and sclareol in cell-free extracts from trichomes of *Nicotiana glutinosa*, Planta 197 (1995) 627–632.
Y.S. Gwak, J.Y. Han, P.B. Adhikari, C.H. Ahn, Y.E. Choi, Heterologous production of a ginsenoside saponin (compound K) and its precursors in transgenic tobacco impairs the vegetative and reproductive growth, Planta 245 (2017) 1105–1119.
C. Hachez, T. Laloux, H. Reinhardt, D. Cavez, H. Degand, C. Grefen, D.E. Rycke, R. Inzé, D. Blatt, M. R, E. Russinova, Arabidopsis SNAREs SYP61 and SYP121 coordinate the trafficking of plasma membrane aquaporin PIP2; 7 to modulate the cell membrane water permeability, Plant Cell 26 (2014) 3132–3147.
D. Hampel, A. Mosandl, M. Wust, Biosynthesis of mono- and sesquiterpenes in carrot roots and leaves (Daucus carota L.): metabolic cross talk of cytosolic mevalonate and plastidial methylerythritol phosphate pathways, Phytochemistry 66 (2005) 305–311.
D. Hampel, A. Swatski, A. Mosandl, M. Wüst, Biosynthesis of monoterpenes and norisoprenoids in raspberry fruits (*Rubus idaeus L.*): the role of cytosolic mevalonate and plastidial methylerythritol phosphate pathway, J. Agric. Food Chem. 55 (2007) 9296–9304.
M. Harker, N. Holmberg, J.C. Clayton, C.G. Gibbard, A.W. Wallace, S. Rawlins, S. A. Hellyer, A. Lanot, R. Safford, Enhancement of seed phytosterol levels by expression of an N-terminal truncated *Hevea brasiliensis* (rubber tree) 3-hydroxy-3-methylglutaryl-CoA reductase, Plant Biotechnol. J. 1 (2003) 113–121.
C.A. Hasemann, R.G. Kurumbail, S.S. Boddupalli, J.A. Peterson, J. Deisenhofer, Structure and function of cytochromes P450: a comparative analysis of three crystal structures, Structure 3 (1995) 41–62.
A. Hemmerlin, J.-F. Hoeffler, O. Meyer, D. Tritsch, I.A. Kagan, C. Grosdemange-Billiard, M. Rohmer, T.J. Bach, Cross-talk between the cytosolic mevalonate and the plastidial methylerythritol phosphate pathways in tobacco bright yellow-2 cells, J. Biol. Chem. 278 (2003) 26666–26676.
A. Hemmerlin, J.L. Harwood, T.J. Bach, A raison d'ĕtre for two distinct pathways in the early steps of plant isoprenoid biosynthesis? Prog. Lipid Res 51 (2012) 95–148.
R.B. Horsch, J. Fry, N. Hoffmann, J. Neidermeyer, S.G. Rogers, R.T. Fraley, Leaf disc transformation. Plant Molecular Biology Manual, Springer, 1989.
A. Huchelmann, M.S. Brahim, E. Gerber, D. Tritsch, T.J. Bach, A. Hemmerlin, Farnesol-mediated shift in the metabolic origin of prenyl groups used for protein prenylation in plants, Biochimie 127 (2016) 95–102.
A. Huchelmann, M. Boutry, C. Hachez, Plant glandular trichomes: natural cell factories of high biotechnological interest, Plant Physiol. 175 (2017) 6–22.
D. Itoh, K. Kawano, K. Nabeta, Biosynthesis of chloroplastidic and extrachloroplastidic terpenoids in liverwort cultured cells: 13C serine as a probe of terpene biosynthesis via mevalonate and non-mevalonate pathways, J. Nat. Prod. 66 (2003) 332–336.
M. Jasiński, Y. Stukkens, H. Degand, B. Purnelle, J. Marchand-Brynaert, M. Boutry, A plant plasma membrane ATP binding cassette-type transporter is involved in antifungal terpenoid secretion, Plant Cell 13 (2001) 1095–1107.
S.M. Jennings, Y.H. Tsay, T.M. Fisch, G.W. Robinson, Molecular cloning and characterization of the yeast gene for squalene synthetase, Proc. Natl. Acad. Sci. 88 (1991) 6038–6042.
Z. Jiang, C. Kempinski, C.J. Bush, S.E. Nybo, J. Chappell, Engineering triterpene and methylated triterpene production in plants provides biochemical and physiological insights into terpene metabolism, Plant Physiol. 170 (2016) 702–716.
A. Jozwiak, P.D. Sonawane, S. Panda, C. Garagounis, K.K. Papadopoulou, B. Abebie, H. Massalha, E. Almekias-Siegl, T. Scherf, A. Aharoni, Plant terpenoid metabolism co-opts a component of the cell wall biosynthesis machinery, Nat. Chem. Biol. 16 (2020) 740–748.
L. Kandra, G.J. Wagner, Studies of the site and mode of biosynthesis of tobacco trichome exudate components, Arch. Biochem. Biophys. 265 (1988) 425–432.
C.K. Keene, G.J. Wagner, Direct demonstration of duvatrienediol biosynthesis in glandular heads of tobacco trichomes, Plant Physiol. 79 (1985) 1026–1032.
J.E. Knoll, Estimation of the limit of detection in chromatography, J. Chromatogr. Sci. 23 (1985) 422–425.
R.W. Kortbeek, J. Xu, A. Ramirez, E. Spyropoulou, P. Diergaarde, I. Otten-Bruggeman, D. E. Both, M. Nagel, R. Schmidt, A, R. Schuurink, Engineering of tomato glandular trichomes for the production of specialized metabolites. Methods in Enzymology, Elsevier, 2016.
A.B. Kroumova, G.J. Wagner, Different elongation pathways in the biosynthesis of acyl groups of trichome exudate sugar esters from various solanaceous plants, Planta 216 (2003) 1013–1021.
B.M. Lange, A. Ahkami, Metabolic engineering of plant monoterpenes, sesquiterpenes and diterpenes—current status and future opportunities, Plant Biotechnol. J. 11 (2013) 169–196.
P. Leivar, V.M. González, S. Castel, R.N. Trelease, C. López-Iglesias, M. Arró, A. Boronat, N. Campos, A. Ferrer, X. Fernàndez-Busquets, Subcellular localization of Arabidopsis 3-hydroxy-3-methylglutaryl-coenzyme A reductase, Plant Physiol. 137 (1) (2005) 57–69.
H.K. Lichtenthaler, The 1-deoxy-D-xylulose-5-phosphate pathway of isoprenoid biosynthesis in plants, Annu. Rev. Plant Biol. 50 (1999) 47–65.
K.B. Linscott, T.D. Niehaus, X. Zhuang, S.A. Bell, J. Chappell, Mapping a kingdom-specific functional domain of squalene synthase, Biochim. Et. Biophys. Acta BBA Mol. Cell Biol. Lipids 1861 (2016) 1049–1057.
J. Liu, Oleanolic acid and ursolic acid: research perspectives, J. Ethnopharmacol. 100 (2005) 92–94.
K.J. Livak, T.D. Schmittgen, Analysis of relative gene expression data using real-time quantitative PCR and the 2− ΔΔCT method, Methods 25 (2001) 402–408.
S. Lodeiro, Q. Xiong, W.K. Wilson, M.D. Kolesnikova, C.S. Onak, S.P. Matsuda, An oxidosqualene cyclase makes numerous products by diverse mechanisms: a challenge to prevailing concepts of triterpene biosynthesis, J. Am. Chem. Soc. 129 (2007) 11213–11222.
P.V. Lograsso, D.A. Soltis, B.R. Boettcher, Overexpression, purification, and kinetic characterization of a carboxyl-terminal-truncated yeast squalene synthetase, Arch. Biochem. Biophys. 307 (1993) 193–199.
J. Lücker, W. Schwab, M.C. Franssen, L.H.W. Van Der Plas, H.J. Bouwmeester, H. A. Verhoeven, Metabolic engineering of monoterpene biosynthesis: two-step production of (+)-trans-isopiperitenol by tobacco, Plant J 39 (2004) 135–145.
Z. Luo, Z. Chen, Improperly terminated, unpolyadenylated mRNA of sense transgenes is targeted by RDR6-mediated RNA silencing in Arabidopsis, Plant Cell 19 (2007) 943–958.
P. Maliga, Á.S.- Breznovits, L. Marton, Streptomycin-resistant plants from callus culture of haploid tobacco, Nat. New Biol. 244 (1973) 29–30.
V.J. Martin, D.J. Pitera, S.T. Withers, J.D. Newman, J.D. Keasling, Engineering a mevalonate pathway in *Escherichia coli* for production of terpenoids, Nat. Biotechnol. 21 (2003) 796.
I. Mendoza-Poudereux, E. Kutzner, C. Huber, J. Segura, W. Eisenreich, I. Arrillaga, Metabolic cross-talk between pathways of terpenoid backbone biosynthesis in spike lavender, Plant Physiol. Biochem 95 (2015) 113–120.
T. Nakashima, T. Inoue, A. Oka, T. Nishino, T. Osumi, S. Hata, Cloning, expression, and characterization of cDNAs encoding *Arabidopsis thaliana* squalene synthase, Proc. Natl. Acad. Sci. 92 (1995) 2328–2332.
R.M. Napier, L.C. Fowke, C. Hawes, M. Lewis, H. Pelham, Immunological evidence that plants use both HDEL and KDEL for targeting proteins to the endoplasmic reticulum, J. Cell Sci. 102 (1992) 261–271.
C. Navarre, A. Sallets, E. Gauthy, M. Maîtrejean, B. Magy, J. Nader, D.E. Thozée, C. P. Crouzet, J. Batoko, H, M. Boutry, Isolation of heat shock-induced *Nicotiana tabacum* transcription promoters and their potential as a tool for plant research and biotechnology, Transgenic Res. 20 (2011) 799–810.
V. Ninkuu, L. Zhang, J. Yan, Z. Fu, T. Yang, H. Zeng, Biochemistry of Terpenes and Recent Advances in Plant Protection, Int. J. Mol. Sci. 22 (2021) 5710.
C.J. Paddon, P.J. Westfall, D.J. Pitera, K. Benjamin, K. Fisher, D. Mcphee, M. Leavell, A. Tai, A. Main, D. Eng, High-level semi-synthetic production of the potent antimalarial artemisinin, Nature 496 (2013) 528–532.
M. Pottier, R. Laterre, V.A.N. Wessem, A. Ramirez, A.M. Herman, X. Boutry, M, C. Hachez, Identification of two new trichome-specific promoters of *Nicotiana tabacum*, Planta 251 (2020) 58.
L. Ralston, S.T. Kwon, M. Schoenbeck, J. Ralston, D.J. Schenk, R.M. Coates, J. Chappell, Cloning, heterologous expression, and functional characterization of 5-epi-aristolochene-1, 3-dihydroxylase from tobacco (*Nicotiana tabacum*), Arch. Biochem. Biophys. 393 (2001) 222–235.
D.-K. Ro, E.M. Paradise, M. Ouellet, K.J. Fisher, K.L. Newman, J.M. Ndungu, K.A. Ho, R. A. Eachus, T.S. Ham, J. Kirby, Production of the antimalarial drug precursor artemisinic acid in engineered yeast, Nature 440 (2006) 940.
C. Sallaud, C. Giacalone, R. Töpfer, S. Goepfert, N. Bakaher, S. Rösti, A. Tissier, Characterization of two genes for the biosynthesis of the labdane diterpene Z-abienol in tobacco (*Nicotiana tabacum*) glandular trichomes, Plant J. 72 (2012) 1–17.
T. Satoh, M. Horie, H. Watanabe, Y. Tsuchiya, T. Kamei, Enzymatic properties of squalene epoxidase from *Saccharomyces cerevisiae*, Biol. Pharm. Bull. 16 (1993) 349–352.
B. Saxena, M. Subramaniyan, K. Malhotra, N.S. Bhavesh, S.D. Potlakayala, S. Kumar, Metabolic engineering of chloroplasts for artemisinic acid biosynthesis and impact on plant growth, J. Biosci. 39 (2014) 33–41.
A. Schilmiller, F. Shi, J. Kim, A.L. Charbonneau, D. Holmes, A. Daniel Jones, R.L. Last, Mass spectrometry screening reveals widespread diversity in trichome specialized metabolites of tomato chromosomal substitution lines, Plant J. 62 (2010) 391–403.
G.W. Schmidt, S.K. Delaney, Stable internal reference genes for normalization of real-time RT-PCR in tobacco (*Nicotiana tabacum*) during development and abiotic stress, Mol. Genet. Genom. 283 (2010) 233–241.
C. Schmidt-Dannert, D. Umeno, F.H. Arnold, Molecular breeding of carotenoid biosynthetic pathways, Nat. Biotechnol. 18 (2000) 750.
L. Somova, A. Nadar, P. Rammanan, F. Shode, Cardiovascular, antihyperlipidemic and antioxidant effects of oleanolic and ursolic acids in experimental hypertension, Phytomedicine 10 (2003) 115–121.
C.M. Starks, K. Back, J. Chappell, J.P. Noel, Structural basis for cyclic terpene biosynthesis by tobacco 5-epi-aristolochene synthase, Science 277 (1997) 1815–1820.
S. Takahashi, Y. Yeo, B.T. Greenhagen, T. Mcmullin, L. Song, J. Maurina-Brunker, R. Rosson, J.P. Noel, J. Chappell, Metabolic engineering of sesquiterpene metabolism in yeast, Biotechnol. Bioeng. 97 (2007) 170–181.
R. Thimmappa, K. Geisler, T. Louveau, P. O'maille, A. Osbourn, Triterpene biosynthesis in plants, Annu. Rev. Plant Biol. 65 (2014) 225–257.
R. Thoma, T. Schulz-Gasch, B. D'arcy, J. Benz, J. Aebi, H. Dehmlow, M. Hennig, M. Stihle, A. Ruf, Insight into steroid scaffold formation from the structure of human oxidosqualene cyclase, Nature 432 (2004) 118–122.
B. Thornton, C. Basu, Real-time PCR (qPCR) primer design using free online software, Biochem. Mol. Biol. Educ. 39 (2011) 145–154.
A. Tissier, C. Sallaud, D. Rontein, Tobacco trichomes as a platform for terpenoid biosynthesis engineering. Isoprenoid Synthesis in Plants and Microorganisms, Springer, 2012.
A.K. Upadhyay, A.R. Chacko, A. Gandhimathi, P. Ghosh, K. Harini, A.P. Joseph, A. G. Joshi, S.D. Karpe, S. Kaushik, N. Kuravadi, Genome sequencing of herb Tulsi (*Ocimum tenuiflorum*) unravels key genes behind its strong medicinal properties, BMC Plant Biol. 15 (2015) 1–20.
J. Vandesompele, D.E. Preter, K. Pattyn, F. Poppe B, N. Van Roy, A. De Paepe, F. Speleman, Accurate normalization of real-time quantitative RT-PCR data by geometric averaging of multiple internal control genes, Genome Biol. (2002), 18;3 (7):RESEARCH0034.

G. Wagner, E. Wang, R. Shepherd, New approaches for studying and exploiting an old protuberance, the plant trichome, Ann. Bot. 93 (2004) 3.
E. Wang, G.J. Wagner, Elucidation of the functions of genes central to diterpene metabolism in tobacco trichomes using posttranscriptional gene silencing, Planta 216 (2003) 686–691.
E. Wang, R. Wang, J. Deparasis, J.H. Loughrin, S. Gan, G.J. Wagner, Suppression of a P450 hydroxylase gene in plant trichome glands enhances natural-product-based aphid resistance, Nat. Biotechnol. 19 (2001) 371–374.
Y. Wang, J. Wen, L. Liu, J. Chen, C. Wang, Z. Li, G. Wang, E. Pichersky, H. Xu, Engineering of tomato type VI glandular trichomes for trans-chrysanthemic acid biosynthesis, the acid moiety of natural pyrethrin insecticides, Metab. Eng. 72 (2022) 188–199.
S. Wei, I. Marton, M. Dekel, D. Shalitin, E. Lewinsohn, B.A. Bravdo, O. Shoseyov, Manipulating volatile emission in tobacco leaves by expressing *Aspergillus niger* β-glucosidase in different subcellular compartments, Plant Biotechnol. J. 2 (2004) 341–350.
S. Wu, M. Schalk, A. Clark, R.B. Miles, R. Coates, J. Chappell, Redirection of cytosolic or plastidic isoprenoid precursors elevates terpene production in plants, Nat. Biotechnol. 24 (2006) 1441.
S. Wu, Z. Jiang, C. Kempinski, E.R.I.C. Nybo, S. Husodo, S. Williams, R, J. Chappell, Engineering triterpene metabolism in tobacco, Planta 236 (2012) 867–877.
J. Wungsintaweekul, W. De-Eknamkul, Biosynthesis of plaunotol in *Croton stellatopilosus* proceeds via the deoxyxylulose phosphate pathway, Tetrahedron Lett. 46 (2005) 2125–2128.
H. Zhang, S. Zhang, Y. Yang, H. Jia, H. Cui, Metabolic flux engineering of cembratrien-ol production in both the glandular trichome and leaf mesophyll in *Nicotiana tabacum*, Plant Cell Physiol. 59 (2018) 566–574.
C. Zhao, Y. Kim, Y. Zeng, M. Li, X. Wang, C. Hu, C. Gorman, S.Y. Dai, S.Y. Ding, J.S. Yuan, Co-Compartmentation of Terpene Biosynthesis and Storage via Synthetic Droplet, ACS Synth Biol 7 (3) (2018) 774–781.
Y. Zhao, J. Yang, B. Qin, Y. Li, Y. Sun, S. Su, M. Xian, Biosynthesis of isoprene in *Escherichia coli* via methylerythritol phosphate (MEP) pathway, Appl. Microbiol. Biotechnol. 90 (2011) 1915–1922.
F. Zhou, E. Pichersky, More is better: the diversity of terpene metabolism in plants, Curr. Opin. Plant Biol. 55 (2020) 1–10.