Planta (2015) 242:693–708
DOI 10.1007/s00425-015-2367-4

ORIGINAL ARTICLE

Production of mono- and sesquiterpenes in *Camelina sativa* oilseed

Jörg M. Augustin[1] · Yasuhiro Higashi[1,2] · Xiaohong Feng[1] · Toni M. Kutchan[1]

Received: 2 February 2015 / Accepted: 29 June 2015 / Published online: 30 July 2015

Abstract

Main conclusion **Camelina was bioengineered to accumulate (4*S*)-limonene and (+)-δ-cadinene in seed. Plastidic localization of the recombinant enzymes resulted in higher yields than cytosolic localization. Overexpressing 1-deoxy-D-xylulose-5-phosphate synthase (*DXS*) further increased terpene accumulation.**

Many plant-derived compounds of high value for industrial or pharmaceutical applications originate from plant species that are not amenable to cultivation. Biotechnological production in low-input organisms is an attractive alternative. Several microbes are well established as biotechnological production platforms; however, their growth requires fermentation units, energy input, and nutrients. Plant-based production systems potentially allow the generation of high-value compounds on arable land with minimal input. Here we explore whether *Camelina sativa* (camelina), an emerging low-input non-foodstuff Brassicaceae oilseed crop grown on marginal lands or as a rotation crop on fallow land, can successfully be refactored to produce and store novel compounds in seed. As proof-of-concept, we use the cyclic monoterpene hydrocarbon (4*S*)-limonene and the bicyclic sesquiterpene hydrocarbon (+)-δ-cadinene, which have potential biofuel and industrial solvent applications. Post-translational translocation of the recombinant enzymes to the plastid with concurrent overexpression of 1-deoxy-D-xylulose-5-phosphate synthase (*DXS*) resulted in the accumulation of (4*S*)-limonene and (+)-δ-cadinene up to 7 mg g^{-1} seed and 5 mg g^{-1} seed, respectively. This study presents the framework for rapid engineering of camelina oilseed production platforms for terpene-based high-value compounds.

Keywords (+)-δ-Cadinene · Camelina · (4*S*)-Limonene · Metabolic bioengineering · Terpenes

Abbreviations

CDNS	Cadinene synthase
DMAPP	Dimethylallyl pyrophosphate
DXS	1-Deoxy-D-xylulose-5-phosphate synthase
FAME	Fatty acid methyl ester
FDS	Farnesyl diphosphate synthase
FDS CDNS	Plant transformation construct containing *FDS* and *CDNS* ORFs
GDS	Geranyl diphosphate synthase
GDS LS	Plant transformation construct containing *GDS* and *LS* ORFs
GDSLS	Plant transformation construct containing *GDSLS* fusion ORF

J. M. Augustin and Y. Higashi contributed equally to this work.

Electronic supplementary material The online version of this article (doi:10.1007/s00425-015-2367-4) contains supplementary material, which is available to authorized users.

✉ Toni M. Kutchan
tmkutchan@danforthcenter.org

Jörg M. Augustin
jmaug@gmx.de

Yasuhiro Higashi
yasuhiro.higashi@riken.jp

Xiaohong Feng
xiaohong68@gmail.com

[1] Donald Danforth Plant Science Center, 975 N Warson Road, St. Louis, MO 63132, USA

[2] Present Address: Metabolomics Research Group, RIKEN Center for Sustainable Resource Science, Tsurumi, Yokohama, Kanagawa, Japan

GP	Glycinin promotor
GSL	Geranyl diphosphate synthase large subunit
GSS	Geranyl diphosphate synthase small subunit
GT	Glycinin terminator
IPP	Isopentenyl pyrophosphate
LS	Limonene synthase
MEP	2-*C*-Methyl-D-erythritol 4-phosphate
NP	Napin promotor
OP	Oleosin promotor
ORF	Open reading frame
OT	Oleosin terminator
TP	Transit peptide
TPCDNS	Transit peptide cadinene synthase fusion
TPFDS	Transit peptide farnesyl diphosphate synthase fusion
TPFDS TPCDNS	Plant transformation construct containing *TPFDS* and *TPCDNS* ORFs
TPFDS TPCDNS DXS	Plant transformation construct containing *TPFDS*, *TPCDNS*, and *DXS* ORFs
TPGDS	Transit peptide geranyl diphosphate synthase fusion
TPGDS TPLS	Plant transformation construct containing *TPGDS* and *TPLS* ORFs
TPGDS TPLS DXS	Plant transformation construct containing *TPGDS*, *TPLS*, and *DXS* ORFs
TPGDSLS	Plant transformation construct containing *TPGDSLS* fusion ORF
TPLS	Transit peptide limonene synthase fusion
TPLSGDS	Plant transformation construct containing *TPLSGDS* fusion ORF

Introduction

Plants provide mankind with a vast array of phytochemicals that have wide ranging industrial and pharmacological applications. However, low availability of high-value phytochemicals can limit their large-scale use. Microbial production systems such as *Escherichia coli* (Yu et al. 2011) and *Saccharomyces cerevisiae* (Hong and Nielsen 2012) are well established biotechnological platforms that can often be successfully bioengineered to serve as alternative sources for natural compounds. In addition to microbes, plant cell cultures have been exploited as potential biotechnological production platforms for phytochemicals (Kirakosyan et al. 2009). Despite the advantages of such cell-based production systems over native low-yield producers, a common drawback is the requirement for specialized fermentation facilities, energy input, and a continuous supply of macro- and micronutrients. Bioengineering of low-input crop plants to synthesize high-value compounds would allow production of phytochemicals on farmland (Fesenko and Edwards 2014). However, whereas production of pharmacological proteins in plants has recently significantly advanced (Davies 2010), suitable plant feedstocks for the production of small molecules remain under-explored.

Camelina is an emerging low-input non-foodstuff Brassicaceae oilseed crop that grows on marginal land or can be used as a rotation crop on fallow land (Collins-Silva et al. 2011). Historically cultivated in Europe, it is now grown in northwestern regions of the United States and Canada and is considered a practical agronomic oilseed crop. Furthermore, camelina oil has been successfully tested by commercial airlines as a hydrotreated renewable jet fuel (Shonnard et al. 2010). Camelina is genetically similar to the model dicot *Arabidopsis thaliana* (Arabidopsis) and is likewise genetically transformed by the facile floral dip method (Lu and Kang 2008). Previous bioengineering attempts of camelina primarily targeted optimization of oil composition (Lu and Kang 2008; Li et al. 2014; Snapp et al. 2014; Nguyen et al. 2015) and plant morphology (Zhang et al. 2012; Lee et al. 2014; Roy Choudhury et al. 2014). To our knowledge, no study reports refactoring of camelina for the production of terpenes—the largest group of natural compounds.

Terpenes are typical plant natural products used in the flavor and fragrance and medical industries and, moreover, have the potential to be developed as drop-in biofuels (George et al. 2015). Of particular interest in this perspective is the monoterpene limonene that has, based on its physicochemical properties, recently been identified as a suitable bioderived jet fuel substitute (Chuck and Donnelly 2014). Both microbial- as well as plant-based platforms are currently under exploration for producing large-scale quantities of terpenes (Carter et al. 2003; Reiling et al. 2004; Ohara et al. 2010; Lange et al. 2011; Alonso-Gutierrez et al. 2013). Biosynthetically, all terpenes are formed from varying combinations of two five-carbon biosynthetic units, isopentenyl pyrophosphate (IPP) and dimethylallyl pyrophosphate (DMAPP), which in turn are formed from two evolutionarily independent pathways: either from acetyl-CoA via the cytosolic mevalonate pathway or from D-glyceraldehyde-3-phosphate and pyruvate via the 2-*C*-methyl-D-erythritol 4-phosphate (MEP) pathway, which in plants is localized in plastids (Eisenreich et al. 1998). IPP and DMAPP are enzymatically concatemerized by specific prenyltransferases to C_{10} (geranyl pyrophosphate; GPP), C_{15} (farnesyl pyrophosphate; FPP), C_{20} (geranylgeranyl

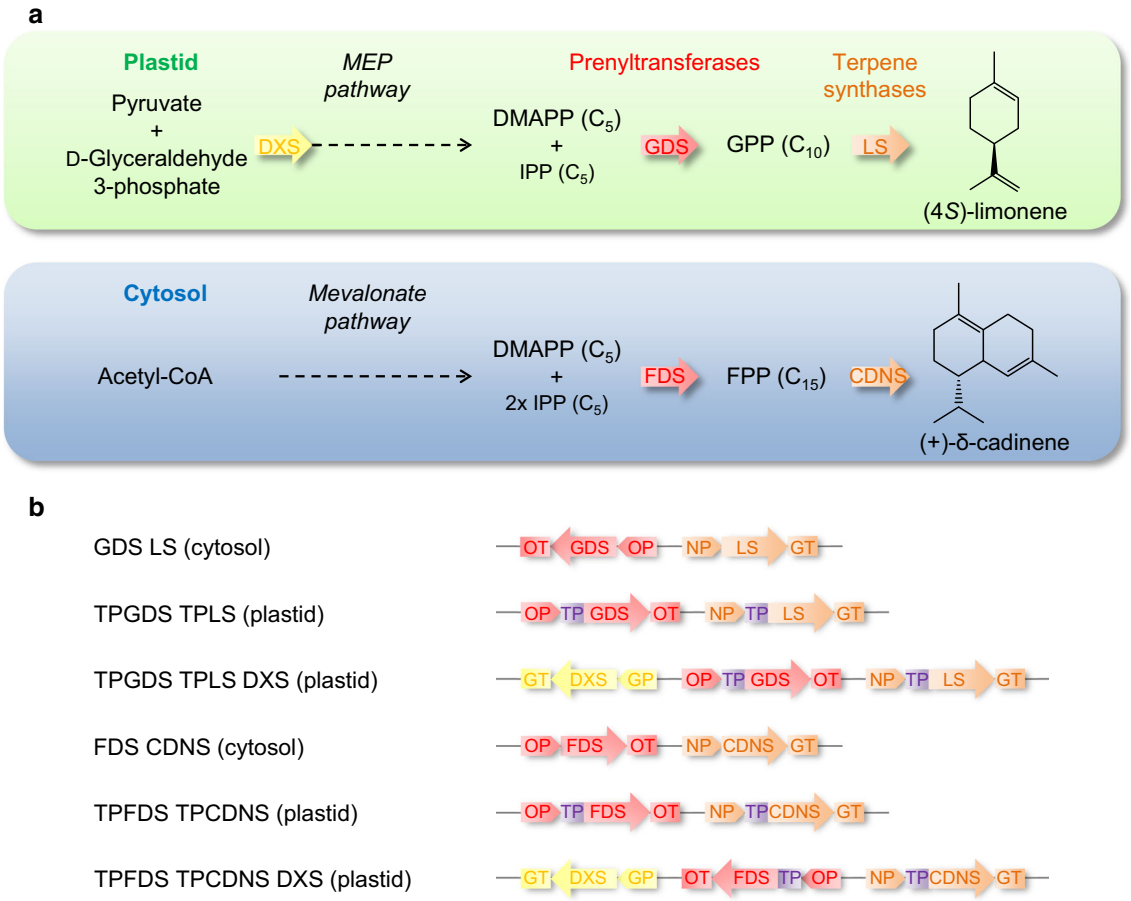

Fig. 1 Graphical representation of terpene metabolic pathways leading to (4*S*)-limonene and (+)-δ-cadinene. The prenyltransferases geranyl diphosphate synthase (GDS) and farnesyl diphosphate synthase (FDS) concatenate dimethylallyl pyrophosphate (DMAPP) and isopentenyl pyrophosphate (IPP) to geranyl pyrophosphate (GPP) and farnesyl pyrophosphate (FPP), which serve as substrates for the terpene synthases (4*S*)-limonene synthase (LS) and (+)-δ-cadinene synthase (CDNS). **a** Peppermint (4*S*)-limonene is produced in the plastid via the MEP pathway. 1-Deoxy-D-xylulose-5-phosphate synthase (DXS) catalyzes the first step of the MEP pathway. Cotton (+)-δ-cadinene is produced in the cytosol via the mevalonate pathway. **b** Plant transformation constructs for bioengineering (4*S*)-limonene and (+)-δ-cadinene production into camelina seed. Expression of the transgenes was controlled by seed-specific promoters. *OP* oleosin promoter, *NP* napin promoter, *GP* glycinin promoter, *OT* oleosin terminator, *GT* glycinin terminator, *TP* Rubisco small subunit transit peptide

pyrophosphate; GGPP), etc. units that are converted to a plethora of terpenes through catalysis by terpene synthases (Degenhardt et al. 2009). In plants, monoterpenes (C_{10} terpenes) are formed in plastids, whereas sesquiterpenes (C_{15} terpenes) are synthesized in the cytosol (Fig. 1a). Specialized cell types are often required for terpene biosynthesis and storage, as exemplified by glandular trichomes of herbaceous plant species (Gershenzon et al. 1992) and resin ducts of conifers (Abbott et al. 2010). Although often oxygenated, terpene formation can be limited to the formation of cyclic hydrocarbons in a heterologous host through the choice of appropriate terpene synthases.

Here we report the biosynthesis and accumulation of the cyclic monoterpene hydrocarbon (4*S*)-limonene and the bicyclic sesquiterpene hydrocarbon (+)-δ-cadinene in bioengineered camelina seed. Comparison of cytosolic and plastidic localization of the recombinant enzymes revealed higher (4*S*)-limonene and (+)-δ-cadinene yields for the latter. Production rates of both terpenes were found further increased by concurrent overexpression of DXS, the first enzyme of the plastidic MEP pathway, which provides the terpene precursors DMAPP and IPP. Emission measurements showed minimal evaporation losses of (4*S*)-limonene during seed development and no detectable losses from matured seeds. Finally, determination of calorific values of camelina seed oils revealed predicted increase in the energy density of seed oil derived from (4*S*)-limonene and (+)-δ-cadinene producing lines.

Materials and methods

Cloning

Due to unawareness of any plant that concurrently produces both (4*S*)-limonene and (+)-δ-cadinene, the limonene synthase (LS) open reading frame (ORF) was cloned from *Mentha* × *piperita* L. (peppermint), whose LS has been previously extensively studied, whereas the (+)-δ-cadinene synthase (CDNS) ORF was obtained from *Gossypium hirsutum* L. cv. Coker 312 (cotton), as cotton is to our knowledge to date the only source of enzymatically characterized CDNSs. *FDS* and *DXS* were cloned from *A. thaliana* (L.) Heyn., ecotype Columbia (Col-0) (Arabidopsis) as due to its close relationship to camelina both encoded enzymes were likely to also be functionally expressed in camelina tissue. In contrast, as no *GDS* gene has been identified in Arabidopsis the previously exhaustively studied peppermint geranyl diphosphate synthase small subunit (GSS) and geranyl diphosphate synthase large subunit (GSL) and their fusion, GDS, respectively, were used instead. The *Pisum sativum* L. var. *saccharatum* (snow pea) Rubisco small subunit transit peptide ORF (TP) was applied for achieving post-translational localization to plastids due to its verified functionality in a wide range of plant species.

Total RNA from peppermint leaves (obtained from a domestic vegetable garden, St. Louis, MO, USA), Arabidopsis leaves (seeds obtained from Donald Danforth Plant Science Center greenhouse seed stock, St. Louis, MO, USA) and cotton embryos (seeds obtained from Donald Danforth Plant Science Center greenhouse seed stock) was extracted applying the RNeasy plant mini kit (Qiagen). First-strand cDNAs were synthesized using SuperScript III (Invitrogen). Genomic DNA of snow pea siliques (purchased from Global Foods Market, St. Louis, MO, USA) was extracted using the DNeasy plant mini kit (Qiagen).

The ORFs for GSS and GSL were cloned from peppermint cDNA with primers GSSfC/GSSr4 and GSLfC/GSLr2, respectively (see Table S1 for primer sequences). GDS is comprised of GSS, a linker sequence encoding for a 10 amino acid (SSNNLGIEGR) peptide sequence and GSL. The GDS ORF encoding for the geranyl diphosphate synthase fusion protein (GDS) was generated by a 2-stage polymerase chain reaction (PCR) method (Burke et al. 2004) using *GSS* and *GSL* clones as templates and primers GSSfC/GSSr10aa/10aaGSLfC/GSLr2. The LS ORF was cloned from peppermint cDNA using primers LSfC3/LSr. Fusions of the *GDS* and *LS* ORFs, connected by a 9 amino acid peptide sequence (9aa, SGGSGGSGG) encoding linker sequence, were generated using 2-stage PCRs with primer combinations LSfC3/9AA-LSr/9AA-GSSfc/GSLr2 (*LS-9aa-GDS*) and GSSfc/9AA-GSLr2/9AA-LSfC3/LSr (*GDS-9aa-LS*). The ORF of farnesyl diphosphate synthase (FDS) was cloned from Arabidopsis cDNA with primers FDSf/FDSr. The *CDNS* sequence was amplified from cotton cDNA using primers BamHI_CDNS-C5_NT_for/NotI_CDNS-C5_CT_rev. The pea TP ORF was cloned from genomic pea DNA with primers RuSfwd/BamHI,RuTP. All sequences were amplified in PCRs using Phusion polymerase (New England BioLabs).

Transformation constructs

All promotor and terminator sequences applied in this study were obtained from collaborators (Drs. Jan Jaworski ((Donald Danforth Plant Science Center, St. Louis, Missouri, USA) and Ed Cahoon (University of Nebraska-Lincoln, Lincoln, Nebraska, USA)), who predetermined their functionality in camelina seed.

The pNaMluIOleosin entry vector (OP-OT) contains an expression cassette consisting of the *Glycine max* (soybean) oleosin promoter (OP) and soybean oleosin terminator (OT) sequences. Similarly, the pNaAscINapin entry vector (NP-GT) harbors a *Brassica napus* (rapeseed) napin promoter (NP) and soybean glycinin terminator (GT) expression cassette. Insertion of the *TP* ORF into OP-OT and NP-GT yielded entry vectors pNaMluIOleosinTP (OP-TP-OT) and pNaAscINapinTP (NP-TP-GT), respectively. The binary vector pRS, received as a kind gift from Dr. Jan Jaworski, contains a soybean glycinin promotor (GP) and soybean glycinin terminator expression cassette (GP-GT) and the Discosoma red fluorescent protein ORF (*DsRed2*) as a selection marker between the left and right border T-DNA repeat sequences. Removal of the GP-GT expression cassette from pRS yielded the binary vector pRSe2.

Insertion of the cloned *GDS*, *LS*, *FDS*, and *CDNS* sequences into the respective entry vectors yielded the expression cassettes OP-GDS-OT, OP-TP-GDS-OT, NP-LS-GT, NP-TP-LS-GT, OP-FDS-OT, OP-TP-FDS-OT, NP-CDNS-GT, and NP-TP-CDNS-GT. Combinations of OP-GDS-OT/NP-LS-GT, OP-TP-GDS-OT/NP-TP-LS-GT, OP-FDS-OT/NP-CDNS-GT, and OP-TP-FDS-OT/NP-TP-CDNS-GT were stepwise inserted into pRSe2 yielding camelina transformation constructs GDS LS (cytosol), TPGDS TPLS (plastid), FDS CDNS (cytosol), and TPFDS TPCDNS (plastid), respectively. Similarly, the fusion ORFs *LS-9aa-GDS* and *GDS-9aa-LS* were first inserted into OP-OT and OP-TP-OT yielding the expression cassettes OP-GDS-9aa-LS-OT, OP-TP-GDS-9aa-LS-OT, and OP-TP-LS-9aa-GDS-OT. Insertion of these expression cassettes into pRSe2 gave rise to the plant transformation constructs GDSLS fusion (cytosol), TPGDSLS fusion (plastid), and TPLSGDS fusion

(plastid). The Arabidopsis *DXS* ORF, which contains its own transit peptide, was a kind gift from Dr. Ed Cahoon. Ligation of *DXS* into the GP-GT cassette of pRS yielded the binary vector pRSDXS. Stepwise insertion of the expression cassettes OP-TP-GDS-OT/NP-TP-LS-GT and OP-TP-FDS-OT/NP-TP-CDNS-GT into pRSDXS led to the TPGDS TPLS DXS (plastid) and TPFDS TPCDNS DXS (plastid) plant transformation constructs. All constructs were verified by Sanger sequencing.

Camelina plant transformation

The plant transformation constructs were transformed into *Agrobacterium tumefaciens* strain GV3101 (pMP90) using a freeze–thaw method (Weigel and Glazebrook 2006) (GDS LS, TPGDS TPLS, TPGDS TPLS DXS, GDSLS, TPGDSLS, and TPLSGDS) or electroporation (2 mm gap cuvettes, 2 kV, 400 Ω, and 25 μF) (FDS CDNS, TPFDS TPCDNS, and TPFDS TPCDNS DXS). Selection of transformed bacteria was carried out on YEP medium containing 10 g l^{-1} bacto-peptone, 10 g l^{-1} yeast extract and 5 g l^{-1} NaCl at pH 7.0 with 25 mg l^{-1} rifampicin, 40 mg l^{-1} gentamicin, and 50 mg l^{-1} kanamycin. Plasmid insertion was confirmed by PCR and diagnostic digests with appropriate restriction enzymes. Overnight precultures of the transformed bacteria were transferred into 2-l flasks containing 400 ml YEP medium with 50 mg l^{-1} kanamycin to a start $OD_{600\ nm}$ of 0.0125 and incubated at 28 °C, 200 rpm for 15 h. Cells were harvested by centrifugation at 4 °C at 5000*g* for 10 min, and then resuspended in an infiltration medium consisting of 0.5× Murashige and Skoog medium with vitamins, 50 g l^{-1} sucrose, and 0.05 % (v/v) Silwet L-77 (Lehle Seeds, Round Rock, TX, USA) to a final $OD_{600\ nm}$ of 1.0 to 1.5.

Wild-type camelina plants were grown in the Donald Danforth Plant Science Center green house (day: 21 °C, night: 20 °C, humidity: 40–90 %). Camelina transformation was performed using a floral dip procedure (Lu and Kang 2008). Flowering camelina plants were inoculated with the *Agrobacterium* suspension prepared as described above and filled into a 500 ml beaker. One to two camelina plants were placed into a vacuum desiccator with their inflorescences immersed in the *Agrobacterium* suspension and degassed under vacuum for 5 min. The inoculated plants were covered with plastic trays for 24 h and kept in the dark before returning them to the greenhouse. Mature seeds of the transformed plants were illuminated with a green LED flashlight, and transgenic seeds identified based on their fluorescence visualized by a red lens (Wratten #25) screen.

RT-PCR of transgenes in developing seeds

RNA was extracted from developing camelina seeds using the RNeasy plant mini kit (Qiagen). First-strand cDNAs were synthesized with M-MLV reverse transcriptase (Invitrogen) using an oligo-dT primer. Primer combinations used for the RT-PCR were fwdCsTub3/revCsTub3, GDS_rof/GDS_RTrev, LS_inf/LS_ror, FDS_rof/FDS_ror, CDNS-C5_inf/CDNS-C1.2_ror, and DXS_inf2/DXS_RTrev.

GC–MS analysis of the transgenic seeds

Per replicate ten mature, homozygous T3 seeds were pooled and ground with a glass rod, soaked overnight in diethyl ether at room temperature, and then finally shook every 15 min for 2 h. iso-Butylbenzene and hexadecane were added as internal standards for (4*S*)-limonene and (+)-δ-cadinene analyses, respectively. T3 seeds were chosen as they represent the first generation that allows unequivocally determination of homozygosity for the introduced trait. The extracts were concentrated under nitrogen and analyzed by GC–MS. GC–MS analyses were performed on an Agilent 7890A GC system equipped with a Phenomenex ZB-5MSi column (32.5 m × 250 μm × 0.25 μm) connected to a 5975C inert XL MSD mass spectrometer. Helium was used as carrier gas. Oven temperature for (4*S*)-limonene analyses was 30 °C for 3 min, raised to 85 °C at a rate of 55 °C min^{-1}, raised to 120 °C at a rate of 10 °C min^{-1}, raised to 300 °C at 120 °C min^{-1}, and held for 3 min. Oven temperature for (+)-δ-cadinene analyses was 50 °C for 3 min, raised to 160 °C at a rate of 80 °C min^{-1}, raised to 170 °C at a rate of 1.4 °C min^{-1}, raised to 300 °C at 120 °C min^{-1}, and held for 3 min. (4*S*)-limonene and (+)-δ-cadinene concentrations were calculated using standard curves of commercial (4*S*)-limonene (Sigma) and, due to the commercial unavailability of (+)-δ-cadinene, the sesquiterpene valencene (Sigma), respectively. Other monoterpenes and sesquiterpenes were assigned by comparison of their EI-MS spectra with those of the NIST library.

Heterologous expression and purification of His-tagged LS and CDNS in *E. coli*

300 ml LB medium containing 50 mg l^{-1} kanamycin and 32 mg l^{-1} chloramphenicol in a 2-l flask were inoculated with a single *E. coli* colony (strain: BL21(DE3) Rosetta) harboring either the pET28a-LS or the pET28a-CDNS plasmid. The inoculated medium was incubated at 30 °C, 180 rpm for 12 h, followed by addition of 600 ml terrific broth containing 50 mg l^{-1} kanamycin and 32 mg l^{-1} chloramphenicol, and adjusted to a final isopropyl β-D-1-thiogalactopyranoside concentration of 0.1 mM before it was incubated for an additional 24 h at 15 °C, 200 rpm. Bacterial cells were harvested by centrifugation at 4 °C, 2500*g* and resuspended in 14 ml equilibration buffer (50 mM NaP_i pH 7.0, 300 mM NaCl, and 10 % (v/v)

glycerol). 1.75 ml of a 10 g l^{-1} lysozyme solution in equilibration buffer were added to the resuspended cells followed by 30 min incubation on ice to aid lysis, which was subsequently performed by five 15 s sonication pulses with a Sonic Dismembrator Model 100 equipped with an Ultrasonic Convertor FS2952 (Fisher Scientific). The sonication pulse power level was set to 5 and the cell suspension was chilled on ice for 45 s between individual pulses. Insoluble fragments of the lysate were removed by 20 min centrifugation at 4 °C, 50,000*g*, and the cleared supernatant incubated for 1 h with 1.5 ml equilibrated TALON Metal Affinity Resin (Clonetech) under continuous rotation at 4 °C. After settling of the resin, the supernatant was discarded and the resin washed three times as follows: 30 ml equilibration buffer were added to the recovered resin followed by 10 min incubation under continuous rotation at 4 °C before the resin was collected by 5 min centrifugation at 4 °C, 700*g* and the supernatant subsequently discarded. The washed resin was transferred to a TALON gravity flow column (Clonetech), washed once with 7.5 ml equilibration buffer, and once with 7.5 ml equilibration buffer containing 10 mM imidazole before the bound protein was eluted with 2.5 ml equilibration buffer containing 150 mM imidazole. The buffer of the eluate was exchanged with 1x phosphate buffered saline (137 mM NaCl, 2.7 mM KCl, 10 mM Na_2HPO_4, and 2 mM KH_2PO_4) using a PD-10 desalting column (GE Healthcare) according to manufacturer's instructions. Final protein concentrations were determined using the Bio-Rad Protein Assay (Bio-Rad) and the purified protein stored at −80 °C until further use.

Western blot analysis of camelina seed protein extracts

Per extract ten pooled mature seeds were ground on ice in 1.5 ml tubes filled with 200 µl extraction buffer using plastic homogenizers. Extraction buffer conditions were 100 mM Tris–HCl pH 7.5, 250 mM sucrose, 50 mM NaCl, 2 mM EDTA, 5 mM DTT, and 5 % (w/v) polyvinylpolypyrrolidone. Insoluble particles were removed by two-fold centrifugation for 20 min at 16,100*g*, 4 °C. 20 µl of each protein extract were further purified by methanol–chloroform precipitation. For this 75 µl methanol, 25 µl chloroform and 75 µl water were sequentially added to the extracts. The denatured proteins were collected in the interphase by centrifugation for 1 min at 16,100*g*, room temperature. The upper phase was discarded, 75 µl methanol added to lower phase and protein pellet, and the newly formed supernatant discarded again. The dried protein pellet was dissolved in 75 µl 5 % (w/v) SDS solution. 5 mg total protein of each purified extract, quantified with a nanodrop 2000 spectrophotometer using the Protein A280 module, was loaded per lane for SDS-PAGE. After protein transfer to polyvinylidene fluoride membranes, the membranes were blocked with 5 % (w/v) bovine serum albumin, 0.1 % (v/v) Tween-20 in Tris buffered saline (TBS) pH 7.6. All wash procedures of the membranes with 0.1 % (v/v) Tween-20 TBS pH 7.6 were performed according to manufacturer's instructions. The membranes were first incubated with 1:1000 dilutions of rabbit antisera (Rockland Immunochemicals) against His-tagged versions of (4*S*)-limonene synthase and (+)-δ-cadinene synthase in 0.1 % (v/v) Tween-20 TBS pH 7.6. 1:4000 dilutions of an alkaline phosphatase conjugated anti-rabbit IgG antibody (Rockland Immunochemicals) in 0.1 % (v/v) Tween-20 TBS pH 7.6 were used for secondary antibody binding. Visualization of secondary antibody binding was achieved by incubation of the membranes with a 250 µM nitro-blue tetrazolium, 250 µM 5-bromo-4-chloro-3′-indolyphosphate, and 50 mM $MgCl_2$ TBS pH 9.5 solution.

Enzyme assays with camelina seed protein extracts

Total protein extracts were prepared from dry mature seeds. *Ca.* 22 pooled seeds (corresponding to ∼20 mg) were ground in a 1.5 ml tube with plastic homogenizers on ice for 3 min in 20 µl mg^{-1} of extraction buffer containing 50 mM Tris–HCl pH 7.5, 100 mM NaCl, 10 % (v/v) glycerol, 5 mM 2-mercaptoethanol, and a protease inhibitor cocktail (Sigma). The protein extract was twice centrifuged at 15,000*g* for 30 min at 4 °C. 45 µl aliquots of the resulting supernatants were frozen in liquid nitrogen and stored at −80 °C until further usage.

GDS enzyme activity was analyzed as follows: reactions were performed in a total volume of 1 ml adjusted to 50 mM Hepes pH 7.2, 10 % (v/v) glycerol, 20 mM $MgCl_2$, 0.5 mM DTT, and with 50 µg *E. coli* recombinant (4*S*)-limonene synthase purified from heterologous expression in *E. coli* BL21(DE3)RIL with the pET28a-LS plasmid, 2 nmol IPP, 2 nmol DMAPP, and 45 µl camelina seed protein extract. The enzyme reaction was initiated by the addition of IPP and DMAPP, overlaid with hexane and incubated at 30 °C for 30 min to 4 h. LS enzyme activity was analyzed as follows: reactions were performed in a total volume of 1 ml adjusted to 50 mM Hepes pH 7.2, 10 % (v/v) glycerol, 20 mM $MgCl_2$, 0.5 mM DTT, 500 mM KCl, and with 2 nmol geranyl pyrophosphate (GPP) and 45 µl camelina seed protein extract. The enzyme reaction was initiated by the addition of GPP, overlaid with hexane and incubated at 30 °C for 1 h to 8 h. Reactions were stopped by chilling on ice followed by vigorous mixing. After addition of isobutylbenzene as internal standard, enzymatically produced

(4*S*)-limonene was three times extracted with hexane. The combined hexane extract was dehydrated by Na_2SO_4, concentrated, and analyzed by GC–MS as described above.

Head-space measurement of volatile (4*S*)-limonene

For 7 consecutive weeks, the same individual TPGDS TPLS (plastid, T4 generation) and wild-type plants were analyzed once per week for terpene emission. T4 plants were chosen as they derived from previously unequivocally determined homozygous seeds, hence ensuring homozygosity of the investigated developing plants, which otherwise could have only been determined post-harvest. Whole potted plants were enclosed in a plastic vacuum desiccator under fluorescent light. Air was aspirated by a small bench top vacuum pump at 2.5 in Hg for 8 h. Activated charcoal (Sigma, 100–400 mesh) was placed both in the air-in and air-out valves in order to bind volatile compounds. Following the incubation period, the charcoal plugs were washed 5 times with diethyl ether. After addition of isobutylbenzene as internal standard, the combined diethyl ether extracts were concentrated under nitrogen gas and analyzed by GC–MS. (4*S*)-limonene content was determined based on the peak height of an *m/z* 136 ion. Total (4*S*)-limonene emission during seed development was estimated based on the results of biological replicates derived from two independent cultivation periods. Volatile emissions of mature seeds were investigated in a similar manner.

LC–MS/MS glycoside analysis

Approximately two grams of homozygous T5 camelina seeds were homogenized using an Ultra-Turrax T25 high-speed homogenizer and three times extracted with 100 % methanol. T5 seeds were chosen as larger quantities of genetically identical seeds were required as were provided by the individual plant giving rise to T3 seeds. The extracts were combined and evaporated under reduced pressure to dryness. Dried residues were dissolved in 50 % methanol and analyzed by a LC (LC-20AD, Shimadzu) in tandem with a 4000 QTRAP mass spectrometer (Applied Biosystems). LC separation was carried out on a Phenomenex Gemini-NX c18 column (150 × 2 mm, 5 μm) using a linear gradient with a flow rate of 0.3 ml min^{-1}; solvent A was 5 mM ammonium acetate in water, and solvent B was 5 mM ammonium acetate in methanol. Putative glycosides were identified by precursor ion scan (*m/z* 161.0) and product ion scan (*m/z* 391.2) in negative ionization mode. Glycoside quantitation was performed by multiple reactions monitoring (MRM) scan using phenyl-β-D-glucopyranoside as an internal standard.

Seed oil quantity and fatty acid composition analysis

Acyl composition of camelina seed oil was determined by hydrolyzing triacylglycerols and transferring the released fatty acids into their respective fatty acid methyl esters (FAMEs). 50 μg mg^{-1} Triheptadecanoin (Nu-Chek) was added as an internal standard to ten mature, homozygous T3 seeds in screw top glass tubes. T3 seeds were chosen as they represent the first generation that allows unequivocally determination of homozygosity for the introduced trait. The seeds were ground in 98 % methanol/2 % sulfuric acid using a glass rod and incubated at 85 °C for 1.5 h. After addition of half the volume 1 M NaCl, FAMEs were twice extracted with two-third the volume of hexanes. FAME extracts were analyzed by GC-FID on a Thermoquest Trace GC 2000 system equipped with an Agilent HP-INNOWax column (30 m × 250 μm × 0.25 μm) using Helium as carrier gas. GC conditions were as follows: 60 °C for 1 min, raised to 185 °C at a rate of 40 °C min^{-1}, raised to 235 °C at a rate of 5 °C min^{-1}, and held for 5 min. Individual FAME species were identified by retention time comparison to known standards and quantified relative to the amount of detected 17:0 FAME. *t* tests were performed to test for statistical significance.

Determination of seed oil calorific values

Cold-pressed camelina seed oil was obtained using a French press equipped with a 3/8-in diameter piston pressure cell. 10 g seeds of each investigated plant line were pressed in five portions (2 g each) with a gage pressure of 900 psi and an environmental temperature of 4 °C. T4 and T5 seeds were chosen as larger quantities of genetically identical seeds were required as were provided by the individual plant giving rise to T3 seeds. Parallel growth under identical greenhouse conditions was given priority over identical generation. Solid seed particles were removed by centrifugation at 16,100*g*. Terpene contents were determined using the same GC–MS protocol described above on 1:20 dilutions of the oils with hexane. Calorific values were determined with an IKA C2000 calorimeter equipped with a C5010 decomposition vessel using 0.35 ml oil per run. Every oil sample was investigated in triplicates. *t* tests were performed to test for statistical significance.

Results

Generation of terpene accumulating camelina lines

To bioengineer (4*S*)-limonene biosynthesis into camelina seed, cDNAs encoding for the peppermint prenyltransferase

geranyl diphosphate synthase (GDS) (Burke et al. 1999) and the peppermint monoterpene synthase (4*S*)-limonene synthase (LS) (Colby et al. 1993) were combined into plant transformation vectors for either cytosolic (transformation vector GDS LS) or plastidic (transformation vector TPGDS TPLS) localization of the recombinant enzymes (Fig. 1b). Native peppermint GDS is a heterodimer, consisting of a small and a large GDS subunit. For this study, the two cDNAs for the small and large subunits were combined into one continuous ORF as described by Burke et al. (2004) and expressed as a fusion protein in camelina seed. To achieve cytosolic localization of GDS and LS both were cloned without their native transit peptides into the GDS LS construct. To unify post-translational transport efficiency, plastidic localization was restored by subsequently adding the pea RuBisCO small subunit transit peptide (TP) (Van den Broeck et al. 1985) to the truncated *GDS* and *LS* sequences for generating the TPGDS TPLS construct. Constructs to bioengineer (+)-δ-cadinene biosynthesis, consisted of the Arabidopsis prenyltransferase farnesyl diphosphate synthase (FDS) (Cunillera et al. 1996) and the cotton sesquiterpene synthase (+)-δ-cadinene synthase (CDNS) (Chen et al. 1995) cDNAs (transformation vectors FDS CDNS and TPFDS TPCDNS). As FDS and CDNS are naturally localized to the cytosol, no modification was applied to the two native sequences for the FDS CDNS construct. Both sequences were extended by the TP sequence to achieve translocation into plastids with the TPFDS TPCDNS construct. For testing the effect of overexpressing DXS (Estévez et al. 2000), the two transformation constructs for plastidic localization were extended by a synthetic Arabidopsis *DXS* ORF (transformation vectors TPGDS TPLS DXS and TPFDS TPCDNS DXS). Seed-specific expression of all transgenes was ensured by placing each reading frame under control of either the oleosin (Rowley and Herman 1997) (*GDS*/*FDS*), napin (Josefsson et al. 1987) (*LS*/*CDNS*), or glycinin (Nielsen et al. 1989) (*DXS*) promoters; any given promoter was used only once in a transformation vector to avoid potential gene silencing. All plant transformation vectors further contained a gene encoding the red-emitting fluorescent protein (DsRed2), to facilitate transgenic seed identification under green light and were introduced into camelina inflorescences by floral dip (Lu and Kang 2008). At least fifteen plants were transformed with each transformation vector construct. Six to seven weeks post-floral dip, mature camelina seeds were harvested; red fluorescing, putative transgenic seeds were propagated in the greenhouse to subsequent generations for further analysis. Typically, a total of ca. 70–140 red fluorescing seeds were obtained, representing ca. 0.2–0.8 % of total seeds produced in fifteen plants. Transgenic plants generally developed indistinguishable from wild-type camelina. Occasionally, transgenic seedlings showed chlorotic cotyledons, leading to delayed development. Affected seedlings developed normally, however, after formation of the first true leaves.

GC–MS analysis of seed oil of transgenic plants showed that camelina transformed with *GDS* and *LS* cDNAs accumulated (4*S*)-limonene and minor amounts of six additional monoterpenes and seven monoterpene oxides in seed oil, whereas wild-type camelina seed oil was devoid of monoterpenes (Fig. 2a, Fig. S1a). Likewise, whereas no sesquiterpenes were detected in wild-type camelina seed oil, camelina transformed with FDS and CDNS encoding cDNAs accumulated (+)-δ-cadinene and in addition trace amounts of five other sesquiterpenes, five sesquiterpene oxides, and three sesquiterpene dioxides in seed oil (Fig. 2b, Fig. S1b). Expression of the transgenes in developing seeds was confirmed by RT-PCR (Fig. S2).

To determine the effect of the different transformation construct designs on mono- and sesquiterpene yields, T3 seeds of multiple homozygous transgenic plants transformed with each construct were subjected to quantitative GC–MS analysis. Monoterpene lines with the (4*S*)-limonene biosynthetic enzymes directed to the plastid of camelina seed (TPGDS TPLS) yielded 1.5–3 mg (4*S*)-limonene per gram seed (Fig. 3a). Notably, when the biosynthetic enzymes were instead directed to the cytosol (GDS LS), substantially lower (4*S*)-limonene levels were observed in camelina seed (up to 0.065 mg g^{-1} seed). Camelina plants concurrently overexpressing DXS, the enzyme catalyzing the first step of the plastidic MEP pathway, accumulated 1.5–7 mg g^{-1} (4*S*)-limonene in seed. Thus, DXS overexpression more than doubled the amount of (4*S*)-limonene in the most productive plant lines. Similar to (4*S*)-limonene, plastidic accumulation of the (+)-δ-cadinene biosynthetic enzymes (TPFDS TPCDNS) resulted in higher accumulation of (+)-δ-cadinene (2.4–3.0 mg g^{-1} seed) than cytosolic localization of the recombinant enzymes (FDS CDNS, 0.4–0.5 mg g^{-1} seed) (Fig. 3b), though the difference was less marked than for (4*S*)-limonene production. Moreover, enhancing the (+)-δ-cadinene transformation construct with *DXS* also increased yields of (+)-δ-cadinene to 4.5–5 mg g^{-1} seed.

To determine the stability of the newly introduced terpene production traits in camelina, TPGDS TPLS plants were propagated and monitored through to the T5 generation. Although some variation was observed among the offspring derived from the same transformation events, no substantial changes in the (4*S*)-limonene yields occurred throughout the generations monitored (Fig. S3).

With an average of 90 days between sowing of camelina seeds and harvesting of the subsequent seed generation, overall terpene production for the most productive (4*S*)-limonene accumulating line was about 0.078 mg g^{-1} seed day^{-1} and about 0.056 mg g^{-1} seed day^{-1} for the best (+)-δ-cadinene producing line.

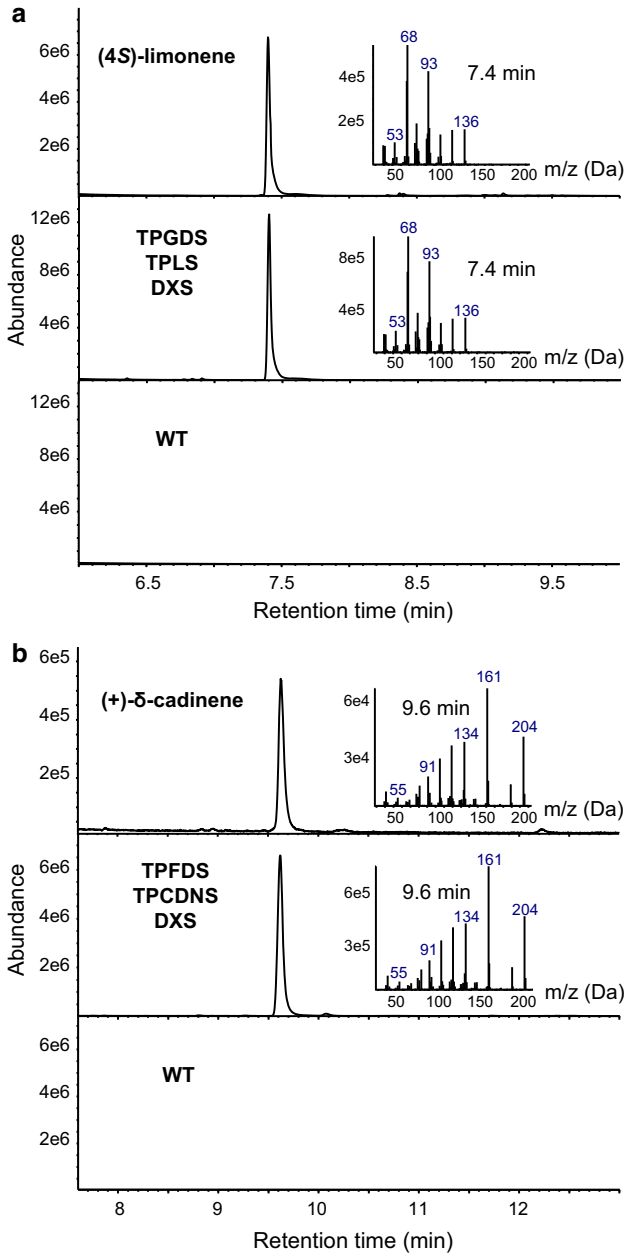

Fig. 2 GC-MS chromatograms of (4*S*)-limonene and (+)-δ-cadinene in cold-pressed oil from transgenic camelina. **a** Total ion chromatograms of commercial (4*S*)-limonene as well as oils of TPGDS TPLS DXS (plastid; T5 generation) and wild-type camelina seed. For the (4*S*)-limonene reference and TPGDS TPLS DXS seed oil mass spectra at 7.4 min are shown. **b** Total ion chromatograms of (+)-δ-cadinene generated by recombinant CDNS expressed in *E. coli*, TPFDS TPCDNS DXS (plastid; T4 generation) seed oil and wild-type camelina seed oil. The mass spectra at 9.6 min are shown for the (+)-δ-cadinene reference and TPFDS TPCDNS DXS seed oil

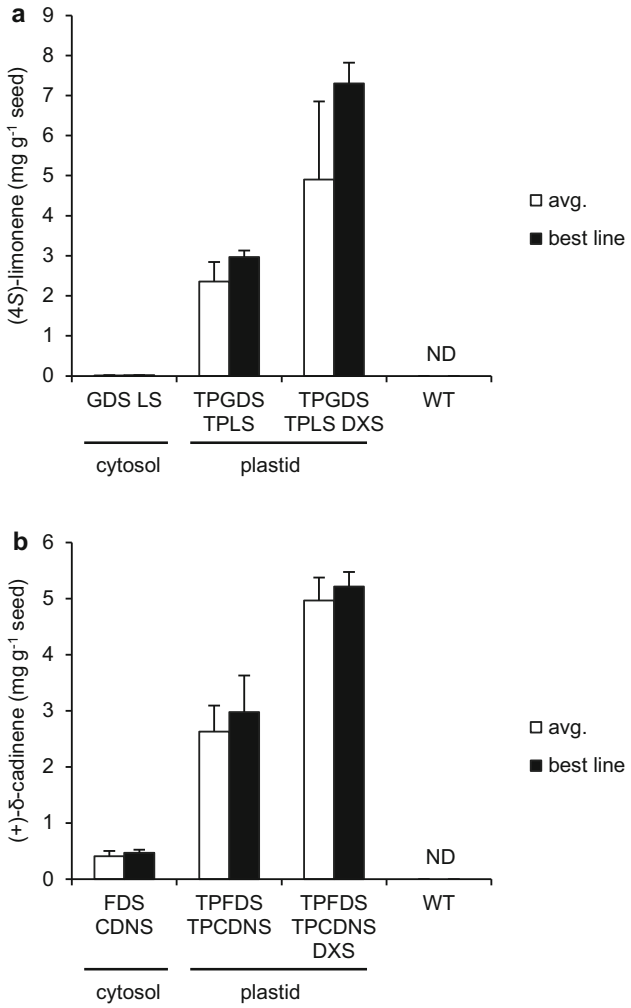

Fig. 3 (4*S*)-limonene and (+)-δ-cadinene accumulation in seed of transgenic camelina lines. **a** (4*S*)-limonene yield in seed of T3 generation plants transformed with the GDS LS (cytosol; $n = 4$), TPGDS TPLS (plastid; $n = 13$), and TPGDS TPLS DXS (plastid; $n = 30$) constructs. **b** (+)-δ-cadinene yield in seed of T3 generation plants transformed with the FDS CDNS (cytosol; $n = 12$), TPFDS TPCDNS (plastid; $n = 12$) and TPFDS TPCDNS DXS (plastid; $n = 10$) constructs. *ND* not detected. *WT* wild-type plants. For each construct are the average ± SD of all investigated plants as well as the mean ± SD of the plant with the highest observed terpene accumulation shown

Verification of functional terpene synthase expression

Polyclonal antibodies raised against either LS or CDNS were used for Western blotting to verify expression of the terpene synthases in seed of transgenic camelina lines. Similar amounts of LS were detected in seed of plants transformed with both plastid and cytosolic localization constructs (Fig. 4a). The Western blotting further showed that increased (4*S*)-limonene accumulation in lines overexpressing DXS does not result from higher LS accumulation in seed as compared to lines without DXS overexpression. In contrast to the situation observed for the (4*S*)-limonene accumulating lines, in camelina lines with FDS and CDNS targeted to the cytosol, CDNS accumulation appears lowered as compared to plants with plastid

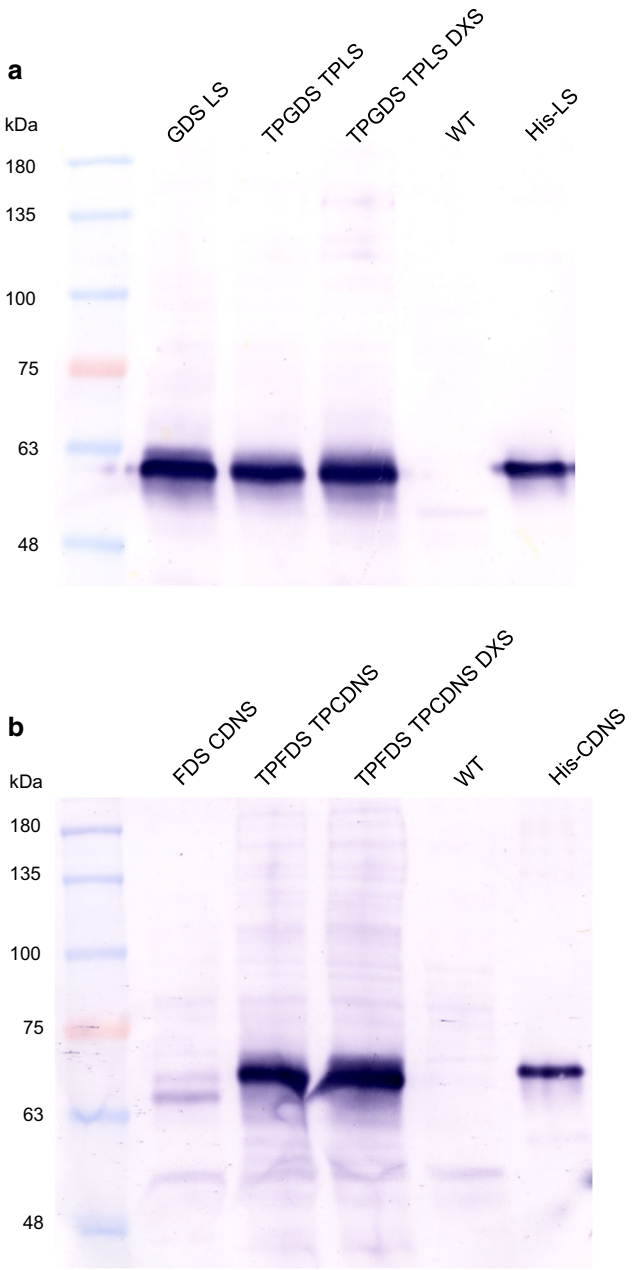

Fig. 4 Western blot of crude protein extracts of (4*S*)-limonene and (+)-δ-cadinene accumulating camelina seed. **a** Protein extracts from plants transformed with the GDS LS (cytosol), TPGDS TPLS (plastid), and TPGDS TPLS DXS (plastid) construct using an anti-LS antiserum. Applied protein extracts derive from those T3 generation plants referred to as 'best line' in Fig. 3a. The reference lane contains 5 ng His-tagged LS (His-LS) that was used to raise the antiserum. **b** Protein extracts from camelina plants transformed with the FDS CDNS (cytosol), TPFDS TPCDNS (plastid), and TPFDS TPCDNS DXS (plastid) constructs using an anti-CDNS antiserum. Protein extracts were obtained from the T3 generation plants with the highest (+)-δ-cadinene yield for each construct (Fig. 3b). 10 ng His-tagged CDNS (His-CDNS), used to raise the antiserum, was loaded as reference

localized CDNS (Fig. 4b). Moreover, the observation of two distinct CDNS protein bands in plants transformed with the construct for cytosolic localization of CDNS, indicates partial protein degradation of CDNS in this cellular compartment. As for LS, CDNS accumulation appears unaffected by overexpression of DXS, thus ruling out higher CDNS amounts as cause for higher (+)-δ-cadinene levels in DXS overexpression lines.

The specific activities of GDS and LS were further tested in vitro in crude protein extracts prepared from camelina seed in order to examine whether the differences in terpene accumulation in the plastid and cytosol experiments were due to variations in the ability to form functional enzymes. Both the plastidic and cytosolic accumulation experiments yielded similar ranges of GDS specific activity (TPGDS TPLS, 7–13.5 pmol min^{-1} mg^{-1} protein; GDS LS, 6.5–15 pmol min^{-1} mg^{-1} protein; Fig. 5a) and LS specific activity (TPGDS TPLS, 0.2–0.75 pmol min^{-1} mg^{-1} protein; GDS LS, 0.05–1.55 pmol min^{-1} mg^{-1} protein; Fig. 5b), which indicates that GDS and LS are functionally equivalent, independent of subcellular localization.

Effect of GDSLS fusion proteins on (4*S*)-limonene accumulation

To determine whether monoterpene production in camelina seed can be increased by fusing the prenyltransferase and monoterpene synthase, three different fusion constructs of GDS and LS (Fig. 6a) were introduced into camelina (transformation vectors GDSLS, TPGDSLS, and TPLSGDS). In all three cases, resulting (4*S*)-limonene amounts (Fig. 6b) were lower compared to experiments in which GDS and LS were produced as discrete enzymes (Fig. 3a). Western blotting indicated lower expression of the recombinant fusion proteins (Fig. 6c) compared to lines expressing discrete GDS and LS (Fig. 4a), which may result from generally lowered transcription or translation efficiency of large ORFs in camelina seed. The substantial drop of GDS and LS specific activities in camelina seed expressing TPGDSLS and TPLSGDS fusion proteins (Fig. 6d, e) further indicated lowered enzymatic activities of the fusion proteins compared to the discrete enzymes (Fig. 5).

(4*S*)-Limonene emission from developing camelina plants

To measure the loss of (4S)-limonene due to evaporation into the atmosphere, head-space analysis of whole camelina plants in different development stages and of harvested

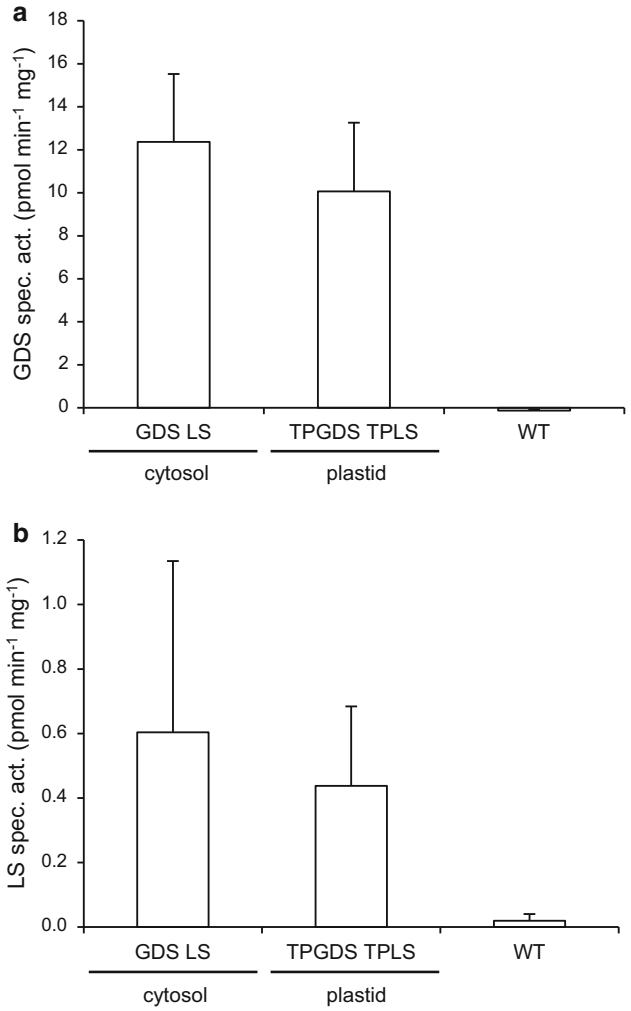

Fig. 5 GDS and LS specific activities in crude camelina seed protein extracts. **a** Specific GDS activity in protein extracts of seed of camelina plants transformed with the GDS LS (cytosol; T3 generation; $n = 7$) and TPGDS TPLS (plastid; T3 generation; $n = 3$) constructs in comparison to a seed protein extract from wild-type camelina. Presented results are the average ± SD of all plants of the corresponding transgenic lines, each of which investigated in triplicates. **b** Specific LS activity in protein extracts of camelina seed harboring the GDS LS (cytosol; T3 generation; $n = 7$) and TPGDS TPLS (plastid; T3 generation; $n = 3$) constructs as well as a wild-type camelina seed protein extract. Results are average ± SD of all plants of the corresponding transgenic lines. Each plant was investigated in triplicates

mature seeds was carried out. (4*S*)-Limonene emissions in developing plants peaked at week 9 and 10, subsequently decreasing below detection limits. No (4*S*)-limonene emissions were detected from mature seeds. The total amount of volatile (4*S*)-limonene emitted from developing plants during weeks 7 to 13 was calculated as 7.0 and 24.8 μg plant^{-1}, which corresponded to 0.24 and 0.84 % of accumulated (4*S*)-limonene in mature seed, respectively (Table 1).

Stable, long-term storage of terpenes in camelina seed was further corroborated by a reanalysis of seeds of mono- and sesquiterpene producing camelina lines after storage up to 3 years at room temperature, resulting in detection of similar terpene amounts as in the initial analysis immediately after harvest (data not shown).

Glycosylation of (4*S*)-limonene in camelina seed

Glycosylation of the mono- and sesquiterpenes *in planta* was investigated using LC–MS/MS. Several putative monoterpene glycosides were identified in T5 TPGDS TPLS seed, which, when using phenyl-β-D-glucopyranoside as standard, were estimated to add up to a total amount of 1.54 ± 0.54 mg g^{-1} (ca. 1/3 of total (4*S*)-limonene accumulated). Although the exact glycosylation positions were not determined, high-resolution mass spectra indicated that both mono- and dihydroxylation of the monoterpene skeleton had occurred in camelina seed and facilitated *O*-glycosylation. Deglycosylation with 1 N HCl at 100 °C for 2 h resulted in formation of monohydroxylated (4*S*)-limonene and 4-isopropyltoluene (also called cymene, a constituent of oil of cumin and thyme), thereby confirming dihydroxylation of the hexene ring of (4*S*)-limonene. Similarly, various sesquiterpene mono- and dioxide glycosides were detected in seeds of (+)-δ-cadinene accumulating camelina lines. However, quantification of those putative (+)-δ-cadinene *O*-glycosides was hampered due to rather distinct fractionation properties to phenyl-β-D-glucopyranoside and the lack of any authentic (+)-δ-cadinene glycoside standards.

Effects on seed oil quantity and quality

For determining effects of mono- and sesquiterpene production in camelina seed on natural seed total fatty acid amounts and fatty acid composition, fatty acid methyl esters (FAMEs) from mature seeds of homozygous T3 generation TPGDS TPLS DXS and TPFDS TPCDNS DXS plants, respectively, and wild-type plants were extracted and measured. Total oil amounts in the seed of TPGDS TPLS DXS plants (29.03 ± 1.78 wt%) compared favorably with wild-type plants grown in parallel (30.55 ± 2.76 wt%) (Fig. S4a). TPFDS TPCDNS DXS seed accumulated significantly ($P < 0.005$) less oil (31.80 ± 2.35 wt%) than seed of wild-type plants grown in parallel (34.66 ± 1.52 wt%) (Fig. S4b). Fatty acid composition of TPGDS TPLS DXS and TPFDS TPCDNS DXS seed differed only marginally from that of wild-type seed (Fig. S4e,f). Both plant lines showed a small reduction in unsaturated long-chain fatty acids (20:X and 22:X) and a corresponding increase in saturated long-chain fatty acids (20:0 and 22:0).

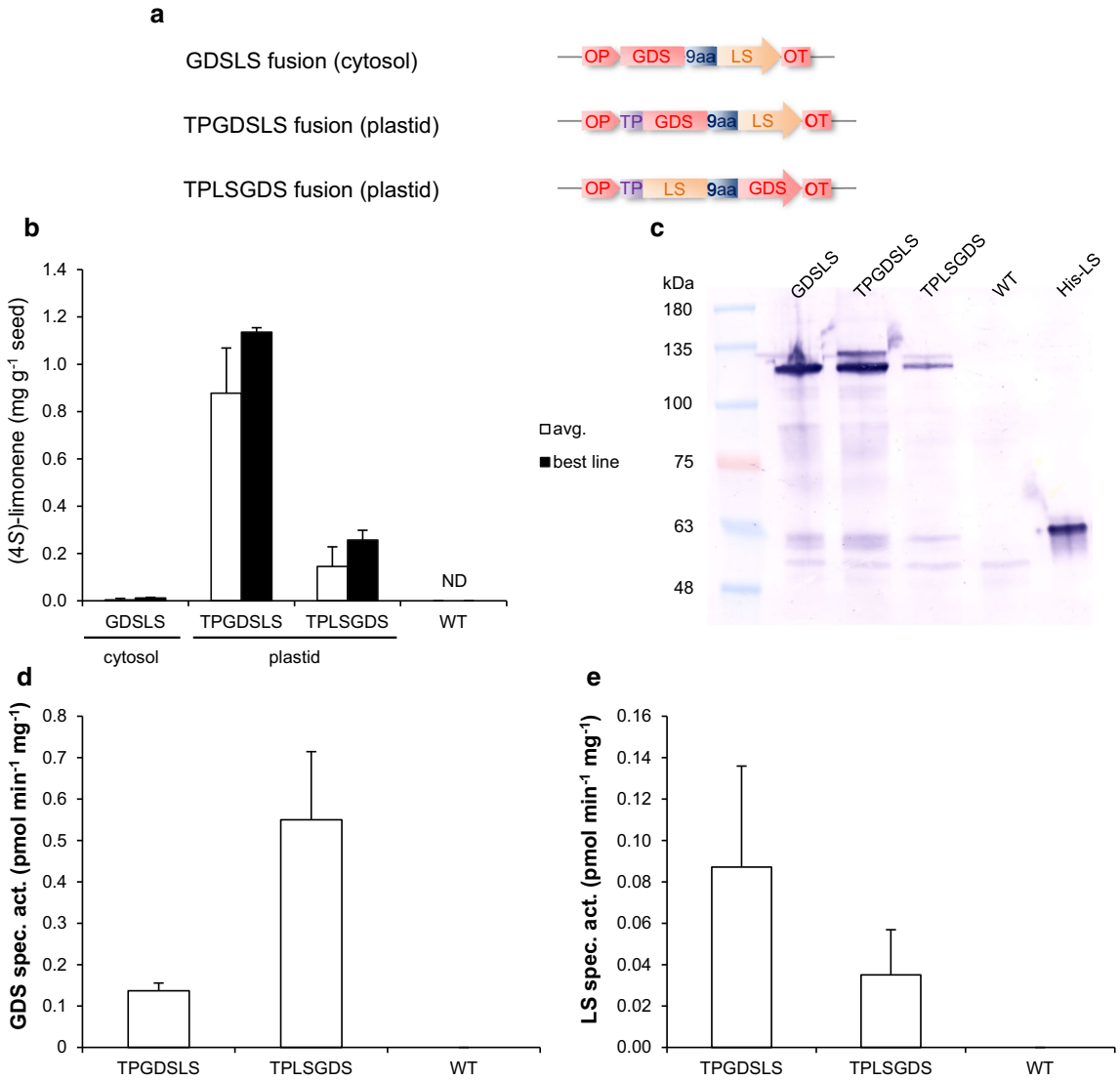

Fig. 6 Fusion constructs of GDS and LS. **a** Overview about generated GDS/LS fusion constructs. *OP* oleosin promotor, *OT* oleosin terminator, *TP* Rubisco small subunit transit peptide, *9aa* 9 amino acid linker sequence. **b** (4*S*)-limonene accumulation in seed of T3 generation plants ($n = 3$ for each group) transformed with the fusion constructs shown in **a**. For each construct are the average ± SD of all investigated plants as well as the mean ± SD of the plant with the highest observed terpene accumulation shown. All samples were analyzed in triplicates. *ND* not detected, *wt* wild-type plant. **c** Western blot of protein extracts from camelina plants transformed with the constructs shown in **a** using an anti-LS antiserum. Protein extracts used for the Western blot analysis derive from those plants, whose (4*S*)-Iimonene accumulation data are shown as best line in panel **b**. 5 ng His-tagged LS (His-LS), used to raise the antiserum, was added as reference. **d** Specific GDS activity in protein extracts of camelina seed transformed with the TPGDSLS (plastid, T2 generation, $n = 3$) and TPLSGDS (plastid; T2 generation; $n = 3$) constructs. Shown results are average ± SD of all plants of the corresponding transgenic lines, each of which investigated once. **e** Specific LS activity in protein extracts harboring the TPGDSLS (plastid; T2 generation; $n = 3$) and TPLSGDS (plastid; T2 generation; $n = 3$) constructs. Results are average ± SD of all plants of the corresponding transgenic lines. Each plant was analyzed once

To further measure changes in seed oil content, cold-pressed oil was obtained from pooled T5 seeds of the most productive TPGDS TPLS DXS line and pooled T4 seeds of the most productive TPFDS TPCDNS DXS line. The (4*S*)-limonene content in the TPGDS TPLS DXS oil was found to be 29.06 ± 0.61 mg ml^{-1} (ca. 3.2 wt%) and the (+)-δ-cadinene content in the TPFDS TPCDNS DXS oil 21.99 ± 0.14 mg ml^{-1} (ca. 2.4 wt%). The calorific value of the two terpene containing oils was compared to oils from wild-type plants grown in previous years (2012, 2013) and in parallel (2014) in the greenhouse, as well as to commercial camelina oil. With 39.59 ± 0.03 kJ g^{-1} commercial oil had the lowest calorific value. Oil from wild-type camelina seeds grown in the green house had

Table 1 (4*S*)-Limonene emission of developing TPGDS TPLS (plastid) camelina plants

Plant age (weeks)	TPGDS TPLS (plastid, T4 generation)		Wild-type
	Experiment 1[a]	Experiment 2[a]	
	In ng (4*S*)-limonene emitted plant^{-1} h^{-1}		
7	0.0	0.0	0.0
8	2.6	0.0	0.0
9	17.2	78.9	0.0
10	10.6	68.6	0.0
11	7.5	0.0	0.0
12	4.0	0.0	0.0
13	0.0	0.0	0.0
Mature seed	0.0	0.0	0.0

[a] The amount of total volatile (4*S*)-limonene emitted during weeks 7 to 13 was calculated as 7.0 and 24.8 μg plant^{-1}, which corresponded to 0.24 and 0.84 % of accumulated (4*S*)-limonene in mature seed, respectively

calorific values of 39.61 ± 0.01 kJ g^{-1} (2012), 39.62 ± 0.02 kJ g^{-1} (2013), and 39.64 ± 0.01 kJ g^{-1} (2014), whereas oils from the TPGDS TPLS DXS and TPFDS TPCDNS DXS lines had significantly increased calorific values of 39.78 ± 0.02 kJ g^{-1} ($P < 1e^{-7}$) and 39.76 ± 0.03 kJ g^{-1} ($P < 5e^{-6}$), respectively.

Discussion

The present study shows that co-expression of the prenyltransferase GDS and the monoterpene synthase LS in camelina seed results in production and accumulation of (4*S*)-limonene. Similarly, expression of a second prenyltransferase FDS and the sesquiterpene synthase CDNS leads to accumulation of (+)-δ-cadinene in camelina seed. Overexpression of *LS* in *Eucalyptus camaldulensis* (eucalyptus) and peppermint has previously been demonstrated to result in elevated limonene levels *in planta* (Ohara et al. 2010; Lange et al. 2011). In contrast to the work presented here, however, the previous *LS* overexpression studies in eucalyptus and peppermint focused on plant tissues that naturally produce limonene as a precursor to their respective monoterpene bouquets, whereas camelina seed is, by our analysis, devoid of mono- and sesquiterpenes.

Typically, monoterpene biosynthesis in plants is located in plastids utilizing DMAPP and IPP derived from the MEP pathway, whereas sesquiterpenes are formed in the cytosol using DMAPP and IPP originating from the mevalonate pathway. In this study, however, both the monoterpene (4*S*)-limonene and the sesquiterpene (+)-δ-cadinene achieved higher accumulation levels when the transgenic enzymes were post-translationally targeted to the plastids as compared to their cytosolic localization. Western blot analysis and in vitro enzyme activity assays with crude seed protein extracts showed that the observed differences in (4*S*)-limonene accumulation do not derive from differential *LS* gene expression efficiency or LS folding properties in the two cellular compartments. This indicates that GDS and/or LS are either less active under cytosolic conditions as compared to their native localization in plastids or are outcompeted for their access to DMAPP and IPP by cytosolic enzymes leading, e.g., to phytosterols. In contrast, Western blotting showed CDNS to accumulate in lower amounts in transgenic camelina seed when retained in the cytosol compared to post-translational redirection to plastids. The observation of two distinct CDNS bands in Western blots of protein extracts from plants with cytosolic CDNS localization further indicates that the lower CDNS amount, and subsequently also the lower (+)-δ-cadinene yield in those plants, results from proteolytic digest of CDNS in the camelina seed cytosol. Both (4*S*)-limonene and (+)-δ-cadinene accumulation further increased when the post-translational direction of the two transgenic enzymes was combined with overexpression of *DXS*, the first enzyme of the plastidic MEP pathway. This result is in agreement with previous reports by Estévez et al. (2001), Botella-Pavía et al. (2004), Carretero-Paulet et al. (2006), and Lange et al. (2011), who observed increased terpene accumulation in response to the overexpression of MEP pathway enzymes in multiple plant species. Moreover, this observation suggests that one of the current limiting factors for terpene production in camelina seed is availability of the terpene precursors DMAPP and IPP.

Fusions between prenyltransferases and terpene synthase occur in nature, at least for formation of the fusicoccins, which are diterpenes in the plant pathogenic fungus *Phomopsis amygdali* (Toyomasu et al. 2007). Fusion of FDS from *Artemisia annua* and 5-epi-aristolochene synthase from tobacco produced a functional chimera in *E. coli*. The K_m values of the fusion protein were unchanged as compared to the individual enzymes;

however, a more efficient conversion of IPP to 5-epi-aristolochene was achieved with the fusion protein (Brodelius et al. 2002). In the present study, three different fusions of *GDS* and *LS* were constructed and expressed in camelina seed. Similar to expression of the discrete enzymes, higher (4*S*)-limonene yields resulted from plastidic localization of the fusion proteins. The overall (4*S*)-limonene accumulation of all GDSLS fusions, however, was considerably lower than observed for the discrete enzymes. Western blotting and enzyme activity assays indicate this to be the combined effect of lowered transcription or translation efficiency and lowered specific activity of the fusion proteins. It cannot be excluded that alternative linker sequences, gene sequences, or gene sequence orientations could lead to fusion proteins that result in higher (4*S*)-limonene accumulation than observed here. The strongly decreased terpene yields observed with the fusion proteins, however, discouraged pursuing this path.

In plants, terpene carbon skeletons are often oxygenated and further modified by e.g., addition of sugar moieties (Lücker et al. 2001). Similarly, in the present study, various mono- and sesquiterpenes oxides were found in seed of (4*S*)-limonene and (+)-δ-cadinene producing plants that are absent in wild-type camelina. LC–MS/MS analysis further suggested that substantial amounts of terpene oxides were subsequently glycosylated. The high number of structurally distinct terpene oxides and glycosides indicates either multiple functionally related enzymes to be involved in their formation or their origination from enzymes that are promiscuous not only with respect to substrate specificity, but also regiospecificity. Likely candidates for introducing the observed modifications are enzymes belonging to the families of cytochrome P450 monooxygenases and uridine diphosphate glycosyltransferases, which are known to be involved in the detoxification of xenobiotics (Wang 2009; Mizutani and Ohta 2010).

Specialized cellular compartments have evolved to store terpenes in plants, such as the subcuticular space between trichome head cells and the cuticle that encloses them in herbaceous plant species (Gershenzon et al. 1992). Due to the high volatility of monoterpenes, in particular, expression of terpene production pathways in plant tissues lacking similar specialized storage compartments could potentially lead to loss of terpenes due to release into the atmosphere. Head-space analysis on whole (4*S*)-limonene producing camelina plants showed only marginal losses of (4*S*)-limonene by evaporation during seed development, which further dropped to undetectable levels upon seed maturation. The natural camelina seed architecture, thus, seems suitable to store mono- and sesquiterpenes. This may be due to the existence of oil bodies in seed, which provide a lipophilic environment allowing the partitioning of mono- and sesquiterpenes. Additionally, the seed coat in the closely related *A. thaliana* is known to contain a suberin layer (Radchuck and Borisjuk 2014). This seed coat may provide an impermeable barrier for mono- and sesquiterpenes, thereby preventing evaporation into the atmosphere.

Triacylglycerols constitute the major energy storage form in camelina seed. They are the source material for FAMEs applied as biodiesel as well as the paraffin proportion used in jet biofuel. Current (4*S*)-limonene and (+)-δ-cadinene accumulation in camelina seed affects both total fatty acid amount and fatty acid composition only marginally. However, the slight reduction in the total fatty acid amount of plants accumulating (+)-δ-cadinene indicates that further increases in terpene production could potentially result in lowered triacylglycerol accumulation. The energy content of terpenes is higher than that of triacylglycerols. Accordingly, proportionately increased energy contents were observed for (4*S*)-limonene and (+)-δ-cadinene containing camelina seed oils. Both measured values are in range of the expected 39.76 kJ g^{-1} for either 3.2 wt% (4*S*)-limonene or 2.4 wt% (+)-δ-cadinene containing oil, which were predicted using the measured calorific value of the 2014 wild type camelina oil and, based on bond enthalpies calculated, standard combustion enthalpies of (4*S*)-limonene (43.57 kJ g^{-1}) and (+)-δ-cadinene (44.84 kJ g^{-1}). Thus, from an economic point of view, further increase of terpene accumulation at the cost of triacylglycerols is preferable. However, since triacylglycerols are an essential carbon and energy source for the germinating seedling, maximal achievable terpene to triacylglycerol ratios in camelina seed will be, in addition to biochemical restrictions, limited by seedling fitness.

Conclusions

Taken together, the results presented in this study demonstrate that camelina seed, naturally devoid of mono- and sesquiterpenes, is a suitable synthetic biology platform for the production of (4*S*)-limonene and (+)-δ-cadinene. The plant is genetically tractable by floral dip, selection of transgenic seed is facilitated by fluorescence resulting from expression of the gene encoding DsRed2 in the transformation vector and transgene expression is stable over at least five generations. Importantly, loss of volatile terpenes during seed development is minimal and virtually undetectable for mature seeds.

Cyclic terpenes are currently being considered as alternatives to Jet A-1 aviation kerosene (Chuck and Donnelly 2014) and diesel (Peralta-Yahya et al. 2011). Accordingly, we observed that accumulation of terpenes in camelina seed increases the calorific value of the cold-pressed oil. Current

terpene producing camelina lines, however, do not yet meet industry requirements for biofuel applications, as both total terpene amounts and terpene to triacylglycerol ratio are still low. Additional interest in terpenes derives from the use of terpenes as drugs and pharmacological agents (Wang et al. 2005; Paduch et al. 2007; Sun 2007). Although limited to three transgenes in this study, the molecular tool kit used here allows for more genes in the transformation vector. Alternatively, using other selection markers such as herbicide resistance, extended numbers of transgenes can be stepwise transferred into the camelina genome. To achieve industrial feasibility, we expect to further elevate terpene accumulation by increasing flux through the MEP pathway and by altering gene dosage of prenyltransferases vs. terpene synthases to balance the difference in steady-state kinetics observed for these two enzyme classes.

Author contribution statement T.M.K. conceived the project and T.M.K., J.M.A. and Y.H. planned experiments. J.M.A. and Y.H. performed all molecular, genetic and enzymatic work. X.F. and J.M.A. performed terpene extraction and mass spectrometric analyses. J.M.A. performed FAME and oil extraction procedures as well as following mass spectrometric and calorimetric analyses. J.M.A., T.M.K. and Y.H. wrote the manuscript. All authors commented on the manuscript.

Acknowledgments The authors thank Jan Jaworski and Ed Cahoon for camelina seeds, transformation vectors, and transformation protocols. The results presented herein are based upon work supported as part of the Center for Advanced Biofuels Systems (CABS), an Energy Frontier Research Center funded by the U.S. Department of Energy, Office of Science, Office of Basic Energy Sciences under Award Number DE-SC0001295.

References

Abbott E, Hall D, Hamberger B, Bohlmann J (2010) Laser microdissection of conifer stem tissues: isolation and analysis of high quality RNA, terpene synthase enzyme activity and terpenoid metabolites from resin ducts and cambial zone tissue of white spruce (*Picea glauca*). BMC Plant Biol 10:106

Alonso-Gutierrez J, Chan R, Batth TS, Adams PD, Keasling JD, Petzold CJ, Lee TS (2013) Metabolic engineering of *Escherichia coli* for limonene and perillyl alcohol production. Metab Eng 19:33–41

Botella-Pavía P, Besumbes Ó, Phillips MA, Carretero-Paulet L, Boronat A, Rodríguez-Concepción M (2004) Regulation of carotenoid biosynthesis in plants: evidence for a key role of hydroxymethylbutenyl diphosphate reductase in controlling the supply of plastidial isoprenoid precursors. Plant J 40:188–199

Brodelius M, Lundgren A, Mercke P, Brodelius PE (2002) Fusion of farnesyldiphosphate synthase and *epi*-aristolochene synthase, a sesquiterpene cyclase involved in capsidiol biosynthesis in *Nicotiana tabacum*. Eur J Biochem 269:3570–3577

Burke CC, Wildung MR, Croteau R (1999) Geranyl diphosphate synthase: cloning, expression, and characterization of this prenyltransferase as a heterodimer. Proc Natl Acad Sci USA 96:13062–13067

Burke C, Klettke K, Croteau R (2004) Heteromeric geranyl diphosphate synthase from mint: construction of a functional fusion protein and inhibition by bisphosphonate substrate analogs. Arch Biochem Biophys 422:52–60

Carretero-Paulet L, Cairó A, Botella-Pavía P, Besumbes O, Campos N, Boronat A, Rodríguez-Concepción M (2006) Enhanced flux through the methylerythritol 4-phosphate pathway in *Arabidopsis* plants overexpressing deoxyxylulose 5-phosphate reductoisomerase. Plant Mol Biol 62:683–695

Carter OA, Peters RJ, Croteau R (2003) Monoterpene biosynthesis pathway construction in *Escherichia coli*. Phytochemistry 64:425–433

Chen XY, Chen Y, Heinstein P, Davisson VJ (1995) Cloning, expression, and characterization of (+)-δ-cadinene synthase: a catalyst for cotton phytoalexin biosynthesis. Arch Biochem Biophys 324:255–266

Chuck CJ, Donnelly J (2014) The compatibility of potential bioderived fuels with jet A-1 aviation kerosene. Appl Energy 118:83–91

Colby SM, Alonso WR, Katahira EJ, McGarvey DJ, Croteau R (1993) 4*S*-Limonene synthase from the oil glands of spearmint (*Mentha spicata*). cDNA isolation, characterization, and bacterial expression of the catalytically active monoterpene cyclase. J Biol Chem 268:23016–23024

Collins-Silva JE, Lu C, Cahoon EB (2011) Camelina: a designer biotech oilseed crop. Inform 22:610–613

Cunillera N, Arró M, Delourme D, Karst F, Boronat A, Ferrer A (1996) *Arabidopsis thaliana* contains two differentially expressed farnesyl-diphosphate synthase genes. J Biol Chem 271:7774–7780

Davies HM (2010) Commercialization of whole-plant systems for biomanufacturing of protein products: evolution and prospects. Plant Biotechnol J 8:845–861

Degenhardt J, Köllner TG, Gershenzon J (2009) Monoterpene and sesquiterpene synthases and the origin of terpene skeletal diversity in plants. Phytochemistry 70:1621–1637

Eisenreich W, Schwarz M, Cartayrade A, Arigoni D, Zenk MH, Bacher A (1998) The deoxyxylulose phosphate pathway of terpenoid biosynthesis in plants and microorganisms. Chem Biol 5:221–233

Estévez JM, Cantero A, Romero C, Kawaide H, Jiménez LF, Kuzuyama T, Seto H, Kamiya Y, León P (2000) Analysis of the expression of *CLA1*, a gene that encodes the 1-deoxyxylulose 5-phosphate synthase of the 2-*C*-methyl-D-erythritol-4-phosphate pathway in Arabidopsis. Plant Physiol 124:95–104

Estévez JM, Cantero A, Reindl A, Reichler S, León P (2001) 1-Deoxy-D-xylulose-5-phosphate synthase, a limiting enzyme for plastidic isoprenoid biosynthesis in plants. J Biol Chem 276:22901–22909

Fesenko E, Edwards R (2014) Plant synthetic biology: a new platform for industrial biotechnology. J Exp Bot 65:1927–1937

George KW, Alonso-Gutierrez J, Keasling JD, Lee TS (2015) Isoprenoid drugs, biofuels, and chemicals-artemisinin, farnesene, and beyond. Adv Biochem Eng Biotechnol. doi:10.1007/10_2014_288

Gershenzon J, McCaskill D, Rajaonarivony JIM, Mihaliak C, Karp F, Croteau R (1992) Isolation of secretory cells from plant glandular trichomes and their use in biosynthetic studies of monoterpenes and other gland products. Anal Biochem 200:130–138

Hong K-K, Nielsen J (2012) Metabolic engineering of *Saccharomyces cerevisiae*: a key cell factory platform for future biorefineries. Cell Mol Life Sci 69:2671–2690

Josefsson LG, Lenman M, Ericson ML, Rask L (1987) Structure of a gene encoding the 1.7 S storage protein, napin, from *Brassica napus*. J Biol Chem 262:12196–12201

Kirakosyan A, Cseke LJ, Kaufman PB (2009) The use of plant cell biotechnology for the production of phytochemicals. In: Kirakosyan A, Kaufman PB (eds) Recent advances in plant biotechnology. Springer, Dordrecht, pp 15–33

Lange BM, Mahmoud SS, Wildung MR, Turner GW, Davis EM, Lange I, Baker RC, Boydston RA, Croteau RB (2011) Improving peppermint essential oil yield and composition by metabolic engineering. Proc Natl Acad Sci USA 108:16944–16949

Lee SB, Kim H, Kim RJ, Suh MC (2014) Overexpression of Arabidopsis *MYB96* confers drought resistance in *Camelina sativa* via cuticular wax accumulation. Plant Cell Rep 33:1535–1546

Li M, Wei F, Tawfall A, Tang M, Saettele A, Wang X (2014) Overexpression of patatin-related phospholipase AIIIδ altered plant growth and increased seed oil content in camelina. Plant Biotechnol J. doi:10.1111/pbi.12304

Lu C, Kang J (2008) Generation of transgenic plants of a potential oilseed crop *Camelina sativa* by *Agrobacterium*-mediated transformation. Plant Cell Rep 27:273–278

Lücker J, Bouwmeester HJ, Schwab W, Blaas J, Van der Plas LHW, Verhoeven HA (2001) Expression of *Clarkia S*-linalool synthase in transgenic petunia plants results in the accumulation of *S*-linalyl-β-D-glucopyranoside. Plant J 27:315–324

Mizutani M, Ohta D (2010) Diversification of P450 genes during land plant evolution. Annu Rev Plant Biol 61:291–315

Nguyen HT, Park H, Koster KL, Cahoon RE, Nguyen HT, Shanklin J, Clemente TE, Cahoon EB (2015) Redirection of metabolic flux for high levels of omega-7 monounsaturated fatty acid accumulation in camelina seeds. Plant Biotechnol J 13:38–50

Nielsen NC, Dickinson CD, Cho T-J, Thanh VH, Scallon BJ, Fischer RL, Sims TL, Drews GN, Goldberg RB (1989) Characterization of the glycinin gene family in soybean. Plant Cell 1:313–328

Ohara K, Matsunaga E, Nanto K, Yamamoto K, Sasaki K, Ebinuma H, Yazaki K (2010) Monoterpene engineering in a woody plant *Eucalyptus camaldulensis* using a limonene synthase cDNA. Plant Biotechnol J 8:28–37

Paduch R, Kandefer-Szerszeń M, Trytek M, Fiedurek J (2007) Terpenes: substances useful in human healthcare. Arch Immunol Ther Exp 55:315–327

Peralta-Yahya PP, Ouellet M, Chan R, Mukhopadhyay A, Keasling JD, Lee TS (2011) Identification and microbial production of a terpene-based advanced biofuel. Nat Commun 2:483. doi:10.1038/ncomms1494

Radchuck V, Borisjuk L (2014) Physical, metabolic and developmental functions of the seed coat. Front Plant Sci 5:510

Reiling KK, Yoshikuni Y, Martin VJJ, Newman J, Bohlmann J, Keasling JD (2004) Mono and diterpene production in *Escherichia coli*. Biotechnol Bioeng 87:200–212

Rowley DL, Herman EM (1997) The upstream domain of soybean oleosin genes contains regulatory elements similar to those of legume storage proteins. Biochim Biophys Acta 1345:1–4

Roy Choudhury S, Riesselman AJ, Pandey S (2014) Constitutive or seed-specific overexpression of Arabidopsis *G-protein γ subunit 3* (*AGG3*) results in increased seed and oil production and improved stress tolerance in *Camelina sativa*. Plant Biotechnol J 12:49–59

Shonnard DR, Williams L, Kalnes TN (2010) Camelina-derived jet fuel and diesel: sustainable advanced biofuels. Environ Prog Sustain Energy 29:382–392

Snapp AR, Kang J, Qi X, Lu C (2014) A fatty acid condensing enzyme from *Physaria fendleri* increases hydroxy fatty acid accumulation in transgenic oilseeds of *Camelina sativa*. Planta 240:599–610

Sun J (2007) D-Limonene: safety and clinical applications. Altern Med Rev 12:259–264

Toyomasu T, Tsukahara M, Kaneko A, Niida R, Mitsuhashi W, Dairi T, Kato N, Sassa T (2007) Fusicoccins are biosynthesized by an unusual chimera diterpene synthase in fungi. Proc Natl Acad Sci USA 104:3084–3088

Van den Broeck G, Timko MP, Kausch AP, Cashmore AR, Van Montagu M, Herrera-Estrella L (1985) Targeting of a foreign protein to chloroplasts by fusion to the transit peptide from the small subunit of ribulose 1,5-bisphosphate carboxylase. Nature 313:358–363

Wang X (2009) Structure, mechanism and engineering of plant natural product glycosyltransferases. FEBS Lett 583:3303–3309

Wang G, Tang W, Bidigare RR (2005) Terpenoids as therapeutic drugs and pharmaceutical agents. In: Zhang L, Demain AL (eds) Natural products: drug discovery and therapeutic medicine. Humana Press Inc, Totowa NJ, pp 197–227

Weigel D, Glazebrook J (2006) Transformation of *Agrobacterium* using the freeze-thaw method. Cold Spring Harb Protoc. doi:10.1101/pdb.prot4666

Yu C, Cao Y, Zou H, Xian M (2011) Metabolic engineering of *Escherichia coli* for biotechnological production of high-value organic acids and alcohols. Appl Microbiol Biotechnol 89:573–583

Zhang Y, Yu L, Yung KF, Leung DY, Sun F, Lim BL (2012) Overexpression of AtPAP2 in *Camelina sativa* leads to faster plant growth and higher seed yield. Biotechnol Biofuels 5:19