



Engineering modular diterpene biosynthetic pathways in *Physcomitrella patens*

Aparajita Banerjee^{1,2} · Jonathan A. Arnesen¹ · Daniel Moser^{1,3} · Balindile B. Motsa¹ · Sean R. Johnson¹ · Bjoern Hamberger^{1,2}

Received: 3 August 2018 / Accepted: 14 November 2018 / Published online: 23 November 2018
© The Author(s) 2018

Abstract

Main conclusion Modular assembly and heterologous expression in the moss *Physcomitrella patens* of pairs of diterpene synthases results in accumulation of modern land plant diterpenoids.

Physcomitrella patens is a representative of the ancient bryophyte plant lineage with a genome size of 511 Mb, dominant haploid life cycle and limited chemical and metabolic complexity. For these plants, exceptional capacity for genome editing through homologous recombination is met with recently demonstrated in vivo assembly of multiple heterologous DNA fragments. These traits earlier made *P. patens* an attractive choice as a biotechnological chassis for photosynthesis-driven production of recombinant peptides. The lack of diterpene gibberellic acid phytohormones in *P. patens* combined with the recent targeted disruption of the single bifunctional diterpene synthase yielded lines devoid of endogenous diterpenoid metabolites and well-suited for engineering of terpenoid production. Here, we mimicked the modular nature of diterpene biosynthetic pathways found in modern land plants by developing a flexible pipeline to install three combinations of class II and class I diterpene synthases in *P. patens* to access industrially relevant diterpene biomaterials. In addition to a well-established neutral locus for targeted integration, we also explored loci created by a class of Long Terminal Repeat Retrotransposon present at moderate number in the genome of *P. patens*. Assembly of the pathways and production of the enzymes from the neutral locus led to accumulation of diterpenes matching the reported activities in the angiosperm sources. In contrast, insights gained with the retrotransposon loci indicate their suitability for targeting, but reveal potentially inherent complications which may require adaptation of the experimental design.

Keywords Diterpenoids · Synthetic biology · Homologous recombination · Photosynthetic production · Neutral locus · Retrotransposon

Aparajita Banerjee and Jonathan A. Arnesen contributed equally to this work.

Electronic supplementary material The online version of this article (<https://doi.org/10.1007/s00425-018-3053-0>) contains supplementary material, which is available to authorized users.

✉ Bjoern Hamberger
hamberge@msu.edu

¹ Biochemistry and Molecular Biology, Michigan State University, East Lansing, MI 48824, USA

² Wisconsin Energy Institute, 1552 University Ave, Madison, WI 53726, USA

³ Institute of Biochemistry, Heinrich-Heine-Universität Düsseldorf, Universitätsstraße 1, 40225 Düsseldorf, Germany

Introduction

Terpenoids constitute the oldest and structurally most diverse group of specialized metabolites in plants. The most recent count reports 58,091 terpenoids from the plant kingdom, with 15,804 compounds of the diterpenoid class, constituted by predominantly heavily oxidized and further functionalized multicyclic scaffolds of 20 carbon atoms (Dictionary of Natural Products 27.1; <http://dnp.chemnetbas.com/faces/chemical/ChemicalSearch.xhtml>). This diversity reflects important and broad roles in development, adaptation and interaction with the environment. For example, in general (i.e., primary) metabolism, the group of phytohormone gibberellins is critical for the growth and development of higher land plants (Helliwell et al. 1999, 2001), while

specialized (i.e., secondary) metabolites such as resin acids play roles in the defense of conifers against herbivore attacks (Hamberger et al. 2011; Keeling and Bohlmann 2006; Zerbe and Bohlmann 2015).

Industrial interest in diterpenoids stems from equally broad applications such as renewable feedstocks, inks, tackifiers, flavors and nutraceuticals, fragrances, and on the far end of the value-spectrum, therapeutics. However, access to these biomaterials is limited by the native sources, where they often accumulate as part of complex mixtures of related but undesired compounds, or because formal chemical synthesis is economically challenging. Similarly, the development of the simpler diterpene scaffolds into lead compounds for semi-synthesis of analog libraries is challenged by low accumulation in the plant source, which may drop below the detection limit of modern analytics due to highly active routes of downstream functionalization.

While not yet as well established as the conventional workhorses of biotechnology, such as *Escherichia coli* or *Saccharomyces cerevisiae*, photosynthetic systems for plant diterpenoid production may offer advantages in terms of sustainability, and similarity to plant systems. Perceived advantages of *E. coli* or *S. cerevisiae* include a short doubling time and available advanced genetic tools (Ajikumar et al. 2010; David and Siewers 2015; Huang et al. 2001; Krivoruchko et al. 2011). In photosynthetic systems, these may be compensated by an abundance of (free) photosynthetically fixed carbon and reducing equivalents, next to native membrane systems, subcellular compartments amenable to targeting of biosynthetic steps and dedicated storage organelles for the products (Englund et al. 2015; Kempinski et al. 2015). The moss *Physcomitrella patens*, representative of a non-vascular ancient lineage of land plants, appears as a promising choice for stable heterologous production of diterpenes. Considerations include its ability to undergo homologous recombination, allowing direct genome editing without the requirement of CRISPR-systems, known genome sequence, presence of the carbon-efficient 2-C-methyl-D-erythritol 4-phosphate (MEP) pathway providing C5-building blocks in the plastids and industrial scalability (King et al. 2016; Reski et al. 2015, 2018). *Physcomitrella patens* has been used as a platform for the biosynthesis of commercially valuable products including the small molecule sesquiterpenes

patchoulol, α/β -santalene, the diterpenes sclareol and taxadiene to complex glycoprotein-based recombinant therapeutic peptides (King et al. 2016; Pan et al. 2015; Reski et al. 2015, 2018; Zhan et al. 2014). However, each of those platforms required the labor intensive generation of highly specialized lines and lacked the flexibility required for formation of multiple diterpenoid targets. Building on the most recent demonstration of efficient transformation of *P. patens* and in vivo assembly of multiple linear DNA parts for production of amorphadiene, and Artemisinin (Khairul Ikram et al. 2017; King et al. 2016), we set out to develop a modular transformation system for diterpenes, imitating their natural formation by pairs of diterpene synthases (diTPSs), and origin of structural diversity through different combinations of diTPSs.

To generate stable lines of engineered *P. patens*, we initially used targeted integration into a neutral locus, defined by the lack of phenotypic change upon insertion of transgenes. Since their discovery over two decades ago, *Pp108*, *Pp213*, and *Pp420* remain the only known such integration sites (Schaefer and Zryd 1997). Of these, *Pp108* is by far the most popular target because it has the highest reported transformation efficiency (*P. patens* v3.3, Chr20, 520,941–523,718) (Fig. 1) (Goodstein et al. 2012; Lang et al. 2018; Schaefer and Zryd 1997).

To overcome the limitation imposed by single loci on assembly of multi-gene pathways with increased complexity, and to investigate if gene dosage by integration in multiple loci can improve production levels, we explored potential additional integration sites in the *P. patens* genome. In the yeast *S. cerevisiae*, it has been demonstrated that retrotransposons can be exploited as potential integration sites for homologous recombination of target genes and that independent integration of multiple copies of the target gene can occur (Cho et al. 1999; Fang et al. 2011; Juretzek et al. 2001; Ohgiya et al. 1997). In *P. patens*, a large number of long-terminal-repeat retrotransposons (LTR-Rs) make up about half of the genome. They fall into four different types of chromodomain-containing Gypsy repeatable elements, which are members of an ancient clade predating the speciation of plants and fungi and which disappeared in seed plants (Lang et al. 2008). Among these, the only intact sequence was reported for *PpatensLTR2* (Novikova et al. 2008), which we used here

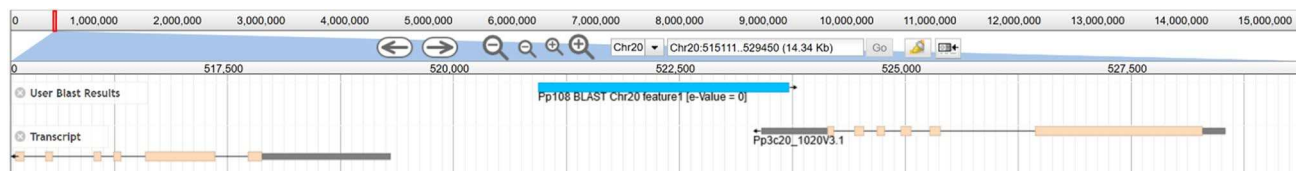


Fig. 1 *Pp108* is located on chromosome 20 between coordinates 520,000 and 525,000

to design constructs to test if the corresponding genomic loci are tolerant to integration of transgenes by homologous recombination and their expression.

Here, we report the pathway engineering for a range of diterpenes in *P. patens* using a modular approach involving linear DNA building blocks and in vivo assembly. Three class II and two class I diTPS genes were chosen based on previous work, and combined to generate stereospecific diterpene scaffolds, either as single compounds, or as mixtures of several high-value biomaterials of industrial relevance (Andersen-Ranberg et al. 2016). Specifically, the structurally complex diterpenes, manoyl oxide in both *ent*- and normal configuration and manool, are lead compounds for diterpenes with important therapeutic activities while sclareol and (+)-abienol are precursors for semisynthetic production of ambroxoides relevant for the fragrance industry. We demonstrate integration and activity of all three diterpene modules into the neutral locus *Pp108*. While integration of a fluorescent reporter in *PpatensLTR2* was successful, stably transformed lines for the significantly larger diterpene modules could not be recovered. Strong expression differences of the reporter in loci of *PpatensLTR2* over *Pp108* may indicate potential future advantages for engineering, if the practical challenges here identified can be solved.

Materials and methods

General maintenance and propagation of *Physcomitrella patens*

The general maintenance and propagation of moss tissue cultures were performed in a laminar flow hood under sterile conditions. Propagation of moss involved adding tissue from one plate into 5 mL of autoclaved water, before tissue disruption into a homogeneous suspension. One-fifth of this mixture was spread over each cellophane-covered BCD-media plate (composition: 45 μ M iron(II) sulfate heptahydrate ($\text{FeSO}_4 \cdot 7\text{H}_2\text{O}$), 1 mM magnesium sulfate (MgSO_4), 1.84 mM monopotassium phosphate (KH_2PO_4), 10 mM potassium nitrate (KNO_3), trace element solution (1000 $\times$ dilution), 1 mM calcium chloride (CaCl_2), 5 mM diammonium tartrate ($(\text{NH}_4)_2\text{C}_4\text{H}_4\text{O}_6$), agar (0.7% (w/v)). Trace element solution, $\text{Al}_2(\text{SO}_4)_3 \cdot \text{K}_2\text{SO}_4 \cdot 24\text{H}_2\text{O}$, $\text{CoCl}_2 \cdot 6\text{H}_2\text{O}$, $\text{CuSO}_4 \cdot 5\text{H}_2\text{O}$, H_3BO_3 , KBr, KI, LiCl, $\text{MnCl}_2 \cdot 4\text{H}_2\text{O}$, $\text{SnCl}_2 \cdot 2\text{H}_2\text{O}$, $\text{ZnSO}_4 \cdot 7\text{H}_2\text{O}$). Calcium chloride and diammonium tartrate (autoclaved separately) were added immediately before use. After addition of the cell suspension, plates were dried for up to 1 h with the lids half opened before sealing with micropore tape (3 M, VWR) and incubation at 18 h light and 6 h dark cycles under 100–150 $\mu\text{mol m}^{-2} \text{s}^{-1}$ of light.

PCR-based generation of promoter and terminator fragments

Both the 5' and 3' termini of the LTR retrotransposons are composed of direct repeats as result of the integration. *PpatensLTR2* is 6686 bp long (GenBank: GQ294565.1) and contains a 474 bp 5' and 452 bp 3' LTR region. The first open reading frame, ORF1, encodes for a protein with homology to retroviral GAG protein. ORF2 contains the *POL* gene, consisting of the characteristic domains of aspartyl protease (PR), reverse transcriptase (RT), ribonuclease H (Rnase H), integrase (Int), chromodomains (chromo), and zinc finger motif (H2C2) (Novikova et al. 2008).

To generate flanking fragments for homologous integration into retrotransposon loci, the *GAG*- and the *POL* regions were amplified from genomic DNA of *P. patens*, subcloned and sequence-verified (all oligonucleotide sequences are given in supplemental table S1). The *GAG*-region was fused by PCR with the neomycin phosphotransferase resistance cassette (*NPTII*) and the *Zea mays* ubiquitin promoter (*ZmUbi*) to form the GNZ-fragment. The *POL*-region was fused with the OCS-terminator (OCS-T) to form the OPO fragment using a two-step overlap extension PCR based assembly. The *GAG* region was amplified with primer pair P5 and P6, where P6 has 20 bp overlap with the 5'-end of the *NPTII* cassette. The *NPTII* cassette/*ZmUbi* region was amplified with the primer pair P7 and P2 introducing a 20 bp overlap with the 3'-end of the *GAG* region and resulting in an overall 40 bp overlap between the *GAG*- and *NPTII/ZmUbi* region. *GAG* and the *NPTII/ZmUbi* PCR products were used as templates with the primer pairs P5 and P2 to assemble the *GAG/NPTII/ZmUbi* (GNZ) fragment. The *POL* region was amplified using primer pair P8 and P9 generating a 20 bp overlap with the 3'-end of the *OCS-T*. The *OCS-T* region was amplified using the primer pairs P3 and P10. *OCS-T* and the *POL* PCR products were used as templates with the primer pairs P3 and P9 to assemble the *OCS-T/POL* (OPO) fragment. All products were subcloned and sequence-verified. The promoter (termed PNZ) and the terminator (OP) fragment for the targeted integration at the *Pp108* locus were amplified with the primer pairs P1, P2 and P3, P4, respectively.

PCR-based generation of the diterpene module and the EYFP fragment

Each of the diterpene modules (Table 1) tested in this work consists of a class II and class I diTPS connected through the LP4-2A linker. Those and the Enhanced Yellow Fluorescent Protein (EYFP) were designed with 20 bp overlap at the 5' and 3' end with the *ZmUbi*-region of the PNZ- or GNZ-fragment and the *OCS-T* region of the OP- or OPO-fragments, respectively. Except *CfTPS2* which was cloned

Table 1 Different modules used for transformation of line pBK3 and the corresponding diterpene products

Module	Class II diTPS	Class I diTPS	Product	
			Name	Number
1	<i>Cf</i> TPS2	<i>Ss</i> SCS	13 <i>R</i> -(+)-manoyl oxide	(i)
			13 <i>S</i> -(+)-manoyl oxide	(iia)
			(+)-manool	(iii)
			(+)-abienol	(iv)
			(+)-sclareol	(v)
2	<i>Tw</i> TPS21	<i>Ep</i> TPS1	13 <i>R</i> - <i>ent</i> -manoyl oxide	(iib)
3	<i>Cf</i> TPS1	<i>Ss</i> SCS	(+)-manool	(iii)

Tw *Tripterygium wilfordii*, *Ep* *Euphorbia peplus*, *Cf* *Coleus forskohlii*, *Ss* *Salvia sclarea*; *TPS* terpene synthase, *SCS* sclareol synthase

from *Coleus forskohlii* cDNA, diTPS genes were synthesized by GenScript. Products were ligated into pJET1.2 and sequence-verified. Diterpene modules were assembled from the single diTPS-coding sequences and the LP4-2A linker sequence using a PCR-based strategy. diTPS sequences were connected to the linker LP4-2A and a two-step assembly PCR was used to generate the final diTPS-modules. In brief, the two-step PCR assembly reaction was adapted with minor modifications from a previously described procedure (Shevchuk et al. 2004). The initial PCR reaction of the assembly included the purified products of class II diTPS gene/linker and linker/class I diTPS and was carried out with ten cycles of denaturation, annealing and extension, followed by a second reaction, using 1 μ L of the product as template.

Isolation of the linear DNA fragment for moss transformation

To obtain sufficient amounts of the linear DNA fragment for the transformation, up to sixteen 50 μ L PCR reactions were performed for each fragment or module. Reactions were pooled, precipitated, resuspended in 50 μ L of autoclaved water and quantified [NanoDrop Lite Spectrophotometer instrument (Thermo Scientific)] after confirming the formation of a single product by gel electrophoresis.

Transformation of *Physcomitrella patens*

Physcomitrella patens has a very simple diterpene background of *ent*-kaurene and 16-hydroxy-*ent*-kaurane as major products and *ent*-beyerene and *ent*-sandaracopimaradiene as minor products. Those are formed by a bifunctional copalyl diphosphate synthase/kaurene synthase (*PpCPS/KS*) from geranylgeranyl diphosphate (GGDP). An earlier reported *Ppcps/ks* knock-out line (pBK3), lacking detectable

diterpenoids (King et al. 2016; Zhan et al. 2015) was used in this work. *Physcomitrella patens* was transformed with different combinations of fragments according to an established protocol with some modifications (Liu and Vidali 2011). Briefly, protoplasts were isolated from 5 to 7 days old *P. patens* tissues using 8% D-mannitol and 2% driselase. 30 μ g of total DNA with equimolar ratios of the individual fragments were added to the protoplasts before addition of the PEG/Calcium solution. The protoplasts were plated on PRMB plates and left for 5 days for regeneration. All centrifugation steps in the protocol were performed at 200 $\times$ *g* under slow acceleration and breaking to limit damage to the protoplasts. After regeneration, the cells were subjected to two consecutive rounds of selection on kanamycin and non-selection, each for 2 weeks, on BCD media plates with 18 h light and 6 h dark cycle under reduced light at 50 μ mol m⁻² s⁻¹, before genotyping to test the assembly of the entire expression cassette at the targeted locus.

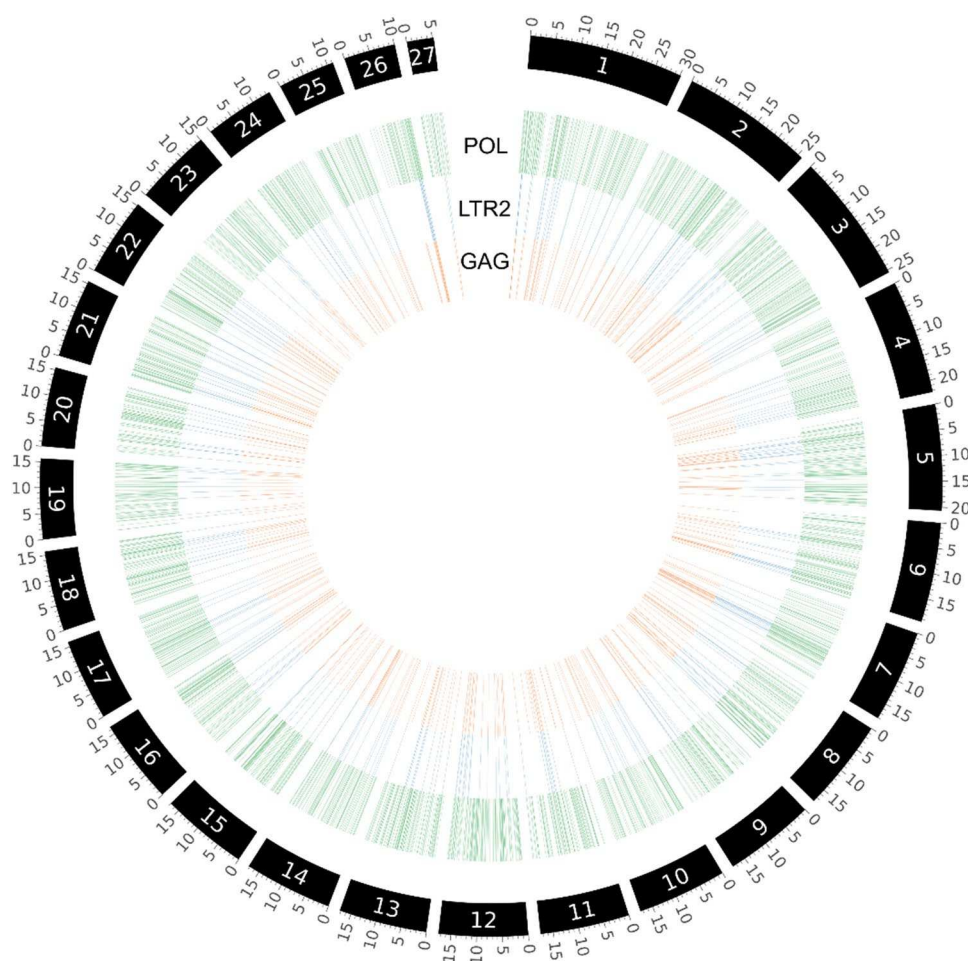
Physcomitrella patens sampling and genotyping

Genotyping PCR was performed using the Phire Plant Direct PCR Kit according to the manufacturer's protocol (Thermo Scientific). In brief, plant tissue was harvested from isolated colonies and suspended in 20 μ L of dilution buffer solution by pipetting and vortexing before direct use as template for the genotyping according to the kit specifications. A 297 bp fragment of a highly conserved region of chloroplast DNA gene was used as a positive control. For the transformants at the *Pp108* locus, the presence of outer-most junctions of the cassette with the *P. patens* genome was verified in a first round of genotyping PCR (Fig. S1a). Positive clones were further tested for the correct assembly of the internal junctions of the PNZ-fragment with the target genes as well as of the target gene with the OP-fragment (Fig. S1b). For the transformants at the *PpatensLTR2* loci, the presence of internal junctions between different fragments were used for screening (Fig. S2a). Due to the significantly higher specificity, integration of the transgene at the *PpatensLTR2* loci was verified through amplification of the 5' LTR region upstream of GAG and the integration junction (Fig. S2b). PCR products were sequence verified (Fig. S8). The reaction targeting the 3' LTR region downstream of POL yielded no specific product, plausibly due to an abundance in other, non-*PpatensLTR2* linked regions of the genome (Fig. 2).

Microscopy

Fluorescence of enhanced yellow fluorescent protein (EYFP) and chlorophyll for transformant lines of *P. patens* was measured with a confocal laser-scanning microscope, Fluoview FV10i (Olympus), at excitation 480 nm/emission 527 nm and excitation 559 nm/emission 570–670 nm, respectively.

Fig. 2 Locations of *Ppatens*LTR2-GAG, *Ppatens*LTR2-POL, and *Ppatens*LTR2 sites with at least 80% completeness in the genome of *P. patens*. Chromosomes in solid black; coordinate units are megabases



Freshly harvested tissues for each line were used for the fluorescent imaging. Images were processed using the FV10-ASW 2.1 microscopy software (Olympus).

Gas-chromatography mass-spectrometry (GC-MS)

Approx. 100 mg (fresh weight) of tissue for each line of *P. patens* was harvested and extracted with 1 mL of hexane containing 1 $\mu\text{g mL}^{-1}$ 1-eicosene as an internal standard. For line *Pp108/M1*, this amount was increased to 1 g of fresh tissue to compensate for relatively lower accumulation of the bioproduct. Extractions were incubated overnight at room temperature on a rocking shaker before centrifugation at 3000 $\times g$ for 5 min and reduction of the supernatant under a stream of nitrogen to 200 μL . GC-MS analysis was performed on an Agilent Technologies 7890A GC system coupled with an Agilent Technologies 5975C inert XL MSD with Triple Axis Detector on an Agilent VF-5 ms (40 m $\times$ 250 μm $\times$ 0.25 μm) column with a purge flow of 3 mL min $^{-1}$ using helium as a carrier gas and a flow of sample at a rate of 1 mL min $^{-1}$. One μL of the sample was injected at an inlet temperature of 250 $^{\circ}\text{C}$ in splitless mode.

The GC oven temperature program used for the separation was comprised of an initial temperature held at 100 $^{\circ}\text{C}$ for 1 min, then a raise of temperature to 250 $^{\circ}\text{C}$ at a rate of 10 $^{\circ}\text{C min}^{-1}$, followed by a raise to 310 $^{\circ}\text{C}$ at a rate of 20 $^{\circ}\text{C min}^{-1}$, and hold for 2 min. The MS was operated in electron-ionization mode (70 eV) at a source temperature of 230 $^{\circ}\text{C}$ and MS quadrupole temperature of 150 $^{\circ}\text{C}$. A solvent delay of 3 min was used and the MS data was collected from 30 to 700 m/z . The diterpenes produced in the engineered moss lines were identified by comparison to the reference MS data earlier reported (Andersen-Ranberg et al. 2016). The diterpenes accumulating in transgenic lines expressing diterpene modules M1 and M2 were quantified against an authentic standard of 13R-(+)-manoyl oxide. In brief, 13R-(+)-manoyl oxide was biosynthesized using the transient *Agrobacterium/Nicotiana benthamiana* system, purified and quantified by NMR as previously described (Andersen-Ranberg et al. 2016). Relative quantification of the diterpenes accumulating in *P. patens* was performed on an Agilent Technologies 7890A GC system equipped with a flame ionization detector (FID). The chromatography was done using an Agilent HP-5MS

(30 m × 0.250 mm × 0.25 µm) column with helium as a carrier gas and a flow of sample at a rate of 1 mL min⁻¹. One µL of the sample was injected at an inlet temperature of 275 °C in splitless mode. The GC oven temperature program used for the separation was comprised of an initial temperature held at 80 °C for 0.5 min, then a raise of temperature to 250 °C at a rate of 50 °C min⁻¹, followed by a raise to 280 °C at a rate of 10 °C min⁻¹, and hold for 0.1 min, followed by a raise to 320 °C at a rate of 50 °C min⁻¹, and hold for 5 min. The FID was set at 300 °C, with H₂ flow at 30 mL min⁻¹ and air flow at 400 mL min⁻¹. Data sampling rate was 50 Hz.

Genomic DNA isolation

Genomic DNA was isolated from different *P. patens* lines using the illustra™ DNA Extraction Kit Phytopure™ (GE Healthcare). Freshly harvested tissue was dried by pressing between filter papers, before snap-freezing in liquid nitrogen and grinding to a fine powder using mortar and pestle. Approx. 100 mg of the ground tissue was transferred to a pre-chilled tube. DNA extraction was performed according to the kit's protocol using Nucleon Phyto Pure DNA extraction resin. After quantification of the genomic DNA, the quality was checked by agarose gel electrophoresis (Fig. S3).

Quantitative realtime PCR

Quantitative Realtime PCR was performed using the Luna® Universal Probe qPCR Master Mix kit (New England Biolabs Inc) according to the manufacturer's instructions. Primers used for the RT-qPCR reactions were tested by end-point PCR and the PCR products were sequence verified. Approx. 10 ng of the genomic DNA from transformant lines was used as template. A 195 bp fragment of the EYFP gene was amplified using the primers P11 and P12 with the condition of denaturation at 95 °C for 15 s, and 40 cycles of annealing and extension at 60 °C for 30 s. The single-copy *PpCYP701B1* gene was used as an internal reference and a 189 bp fragment was amplified using primers P13 and P14. The RT-qPCR was performed with three technical repeats and two biological replicates for each line. The relative EYFP gene dosage (N) of the RT/EYFP transformant lines were determined relative to that of the Pp108/EYFP transformant line following the equation below (Pfaffl 2004).

$$N = \frac{E_{\text{target}}^{\Delta Ct_{\text{target}}(\text{control-sample})}}{E_{\text{reference}}^{\Delta Ct_{\text{reference}}(\text{control-sample})}}$$

where *E* is the efficiency of the qPCR reaction for that particular primer pairs; control and sample represent the Pp108/EYFP line R and RT/EYFP line, respectively.

Results

Integration loci

To identify suitable retrotransposon regions that could be used as potential integration sites, Blast searches against the most recent release of the *P. patens* genome v3.3 (Lang et al. 2018) revealed 371, 499, and 2283 putative *PpatensLTR2*, *PpatensLTR2-GAG*, and *PpatensLTR2-POL* sites, respectively, of at least 80% completeness (Fig. 2). The data indicated that most GAG sites are colocalized with *PpatensLTR2* over the 27 chromosomes, while the POL sites are more abundant and not exclusively specific to *PpatensLTR2*. Considering the abundance of the GAG and POL regions, they were chosen as potential homologous recombination sites for integration of transgenes into the retrotransposon regions.

Construction of the promoter and the terminator fragments

For targeted integration into the *Pp108* locus, a modular approach was devised where promoter and terminator fragments can be combined with any target gene (e.g., different diterpene modules or the *EYFP* gene). The flanking regions for homologous recombination with the promoter and terminator fragments were inserted into the target gene fragments. The promoter fragment (PNZ) consists of an 879 bp region homologous to the 5' end of the *Pp108* locus (*P. patens* v3.3, Chr20, 520,941–521,819) followed by the resistance cassette (*NPTII*) and the *Zea mays* ubiquitin promoter (*ZmUbi*) (Fig. 3a). The terminator fragment (OP) consists of the OCS-terminator followed by the 1247 bp region homologous to the 3' end of the *Pp108* locus (*P. patens* v3.3, Chr20, 522,472–523,718) (Fig. 3b). Correct in vivo assembly of the PNZ-fragment, the diterpene module, and the OP-fragment by homologous recombination inserts the target genes into the *Pp108* locus of the *P. patens* genome (Fig. 3d).

To target *PpatensLTR2* loci in the *P. patens* genome, the flanking integration sites of the promoter (GNZ) and the terminator (OPO) fragment were designed to carry the 5' region (518 bp) of the GAG-gene (*P. patens* v3.3, Chr20, 15,443,889–15,444,406) and 3' region (531 bp) of the POL-gene (*P. patens* v3.3, Chr20, 2,120,160–2,119,630). With this design, successful integration will replace the structural elements of *PpatensLTR2* with the transgenic DNA construct (Fig. 3e). Consistent with the single integration approach, for *PpatensLTR2* the promoter fragment (GNZ) carries the *NPTII* resistance cassette and *ZmUbi* promoter (Fig. 3a) while the terminator fragment (OPO) is

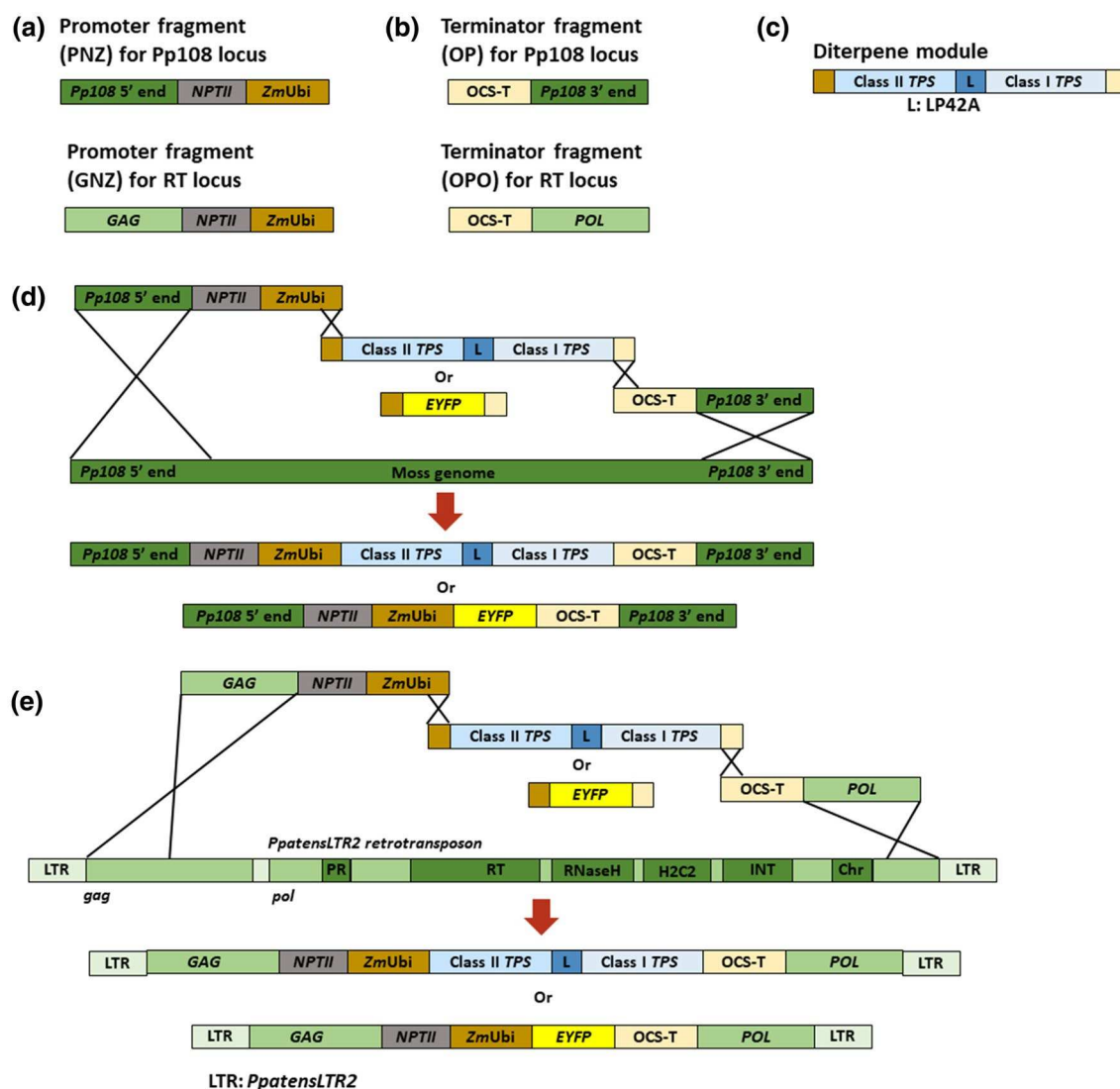


Fig. 3 Schematic of different fragments for transformation into *P. patens* (a–c). Schematic of genome editing at the (d) *Pp108* locus and (e) *PpatensLTR2* loci using homologous recombination

composed of OCS-terminator followed by the *POL* region (Fig. 2b). Sequence verification confirmed that the isolated *GAG*-sequence from *P. patens* was 518 bp long and shared 97% identity with the *PpatensLTR2* sequence (Fig. S4). The isolated *POL*-sequence from *P. patens* was 531 bp long and shared a 452 bp stretch with 97% identity with the original sequence. The 3' *POL*-sequence contained a 39 bp insert, followed by 40 bp with 100% identity to the original sequence of *PpatensLTR2* (Fig. S5).

Construction of target modules

Three different modules have been chosen to engineer diterpene biosynthetic pathway in *P. patens*. Modules consist of pairs of class II and class I diTPSs, connected through

a hybrid LP4/2A linker (L) (Fig. 3c). LP4/2A is a hybrid linker peptide containing the first 9 amino acid residues of LP4 linker peptide from the naturally occurring polypeptide precursor from *Impatiens balsamina* and 20 amino acid residues of 2A linker peptide from the foot-and-mouth disease virus (François et al. 2004). Module 1 consists of the pair of class II and class I diTPSs from *Coleus forskohlii* (synonym *Plectranthus barbatus*) CfTPS2 and *Salvia sclarea* SsSCS, shown earlier to yield stereospecific mixture of diterpenes 13R-(+)-manoyl oxide (i), 13S-(+)-manoyl oxide (iia), (+)-manool (iii), (+)-abienol (iv), and (+)-sclareol (v) (Table 1) (Andersen-Ranberg et al. 2016). Module 2 consists of *Tripterygium wilfordii* TwTPS21 and *Euphorbia peplus* EpTPS1, while *C. forskohlii* CfTPS1 and *SsSCS* constitute module 3. Both module 2 and 3 were shown to result

in a single stereospecific product, 13*R*-*ent*-manoyl oxide (iib) from module 2 and (+)-manool (iii) from module 3 (Andersen-Ranberg et al. 2016). The 20 bp overlap region to the PNZ- or GNZ- and the OP- or OPO-fragments at the 5' and 3' end on the target (diterpene module/*EYFP*) gives full flexibility to combine each target with the PNZ- and OP-fragments. This modular approach allows convenient engineering of *P. patens* with any pair of diTPSs.

Moss transformation and genotyping of the transformants

For the stable expression of the target genes, PCR-amplified linear DNA fragments of diterpene modules/*EYFP* gene along with that of PNZ- or GNZ- and OP- or OPO-fragments were used to transform the diterpene-free *Ppcps/eks* knock-out line pBK3 (King et al. 2016). Genotyping of kanamycin resistant lines indicated that for the transformation targeting the *Pp108* locus, 25% of M1, 6.1% of M2, 5% of M3 module transformants and 2.5% of *EYFP* transformants were completely assembled (Table S2). For the transformation targeting *PpatensLTR2* loci, 16.7% of the *EYFP* transformants were found positive (Table S2). For the diterpene modules M2 or M3 targeting *PpatensLTR2*, no transformant with correct assembly of all fragments could be identified, after screening of 320 of kanamycin resistant lines.

Metabolite analysis

The transformants of the diterpene modules at *Pp108* locus were tested for accumulation of diterpenes by GC–MS analysis. Lines of module M1 (*CfTPS2/SsSCS*) were found to accumulate multiple products, 13*R*-(+)-manoyl oxide (i) and 13*S*-(+)-manoyl oxide (iia), (+)-manool (iii), (+)-abienol (iv), and (+)-sclareol (v) (Fig. 4, Table 1), consistent with earlier findings using a transient expression system (Andersen-Ranberg et al. 2016). Transformants of diterpene modules M2 (*TwTPS21/EpTPS1*) and M3 (*CfTPS1/SsSCS*) accumulated single products, 13*R*-*ent*-manoyl oxide (iib) and (+)-manool (iii), respectively (Fig. 4, Table 1), also matching earlier reported profiles. The accumulation of diterpenes in these transgenic lines ranges from 0.13 ± 0.02 ng of 13*R*-*ent*-manoyl oxide/mg of fresh weight for the diterpene module M2 to 0.74 ± 0.14 ng of total diterpene/mg of fresh weight for diterpene module M1.

Analysis of the *Pp108/EYFP* and *PpatensLTR2/EYFP* lines by microscopy

Expression of *EYFP* was analyzed for the representative lines of *Pp108/EYFP* and *PpatensLTR2/EYFP* by confocal microscopy. The fluorescence signal of lines with *EYFP* integrated at *Pp108* (representative line R) was consistently of an intermediate brightness. In contrast, the signal detected in different lines with *EYFP* integrated in *PpatensLTR2* loci varied greatly in intensity (representative lines F and H) (Fig. 5), i.e., line F showed a significantly brighter fluorescence signal compared to the intermediate *Pp108/EYFP* line R, while the fluorescence signal from line H was barely detectable.

To test whether the observed differences of fluorescence in the representative lines are caused by gene dosage effects, i.e., possible integration of targets in multiple *PpatensLTR2* loci, we compared the relative quantity of the *EYFP* gene by qPCR on genomic DNA.

Genomic analysis of the *Pp108/EYFP* and selected *PpatensLTR2/EYFP* lines

The distinctly different signal intensity observed among the *Pp108/EYFP* line R and the *PpatensLTR2/EYFP* lines F and H (Fig. 4) prompted us to test if this variation is caused by gene dosage effect using quantitative real time PCR analysis of the genomic DNA.

Our analysis showed no significant difference in the *EYFP* gene dosage between *Pp108/EYFP* line R and *PpatensLTR2/EYFP* line F (Fig. 6), whereas significantly lower gene dosage was observed in *RT/EYFP* line H (~ 10-fold lower). It has been reported that transformation of *P. patens* can generate antibiotic resistant lines containing episomally replicating DNA. This episomal DNA was found to undergo formation of concatemers consisting of 3–40 copies, which can also chromosomally integrate (Ashton et al. 2000; Muren et al. 2009). To test for concatemer formation, genomic DNA of representative *Pp108/EYFP* line R was subjected to PCR. A single product was formed with a forward primer on the 3' end of the OP-fragment and a reverse primer on the 5' end of the PNZ-fragment, which supports concatemer formation, specifically in a head-to-tail orientation of individual monomeric constructs (Fig. S6). Sequence verification corroborated the finding and indicated a 20-bp deletion at the junction between 3' end of the OP-fragment (last 6 bp at the 3' end of the OP-fragment) and 5' end of the PNZ-fragment (first 14 bp at the 5' end of the PNZ-fragment). Other combinations considering head-to-head, or tail-to-tail orientations did not yield detectable products by PCR.

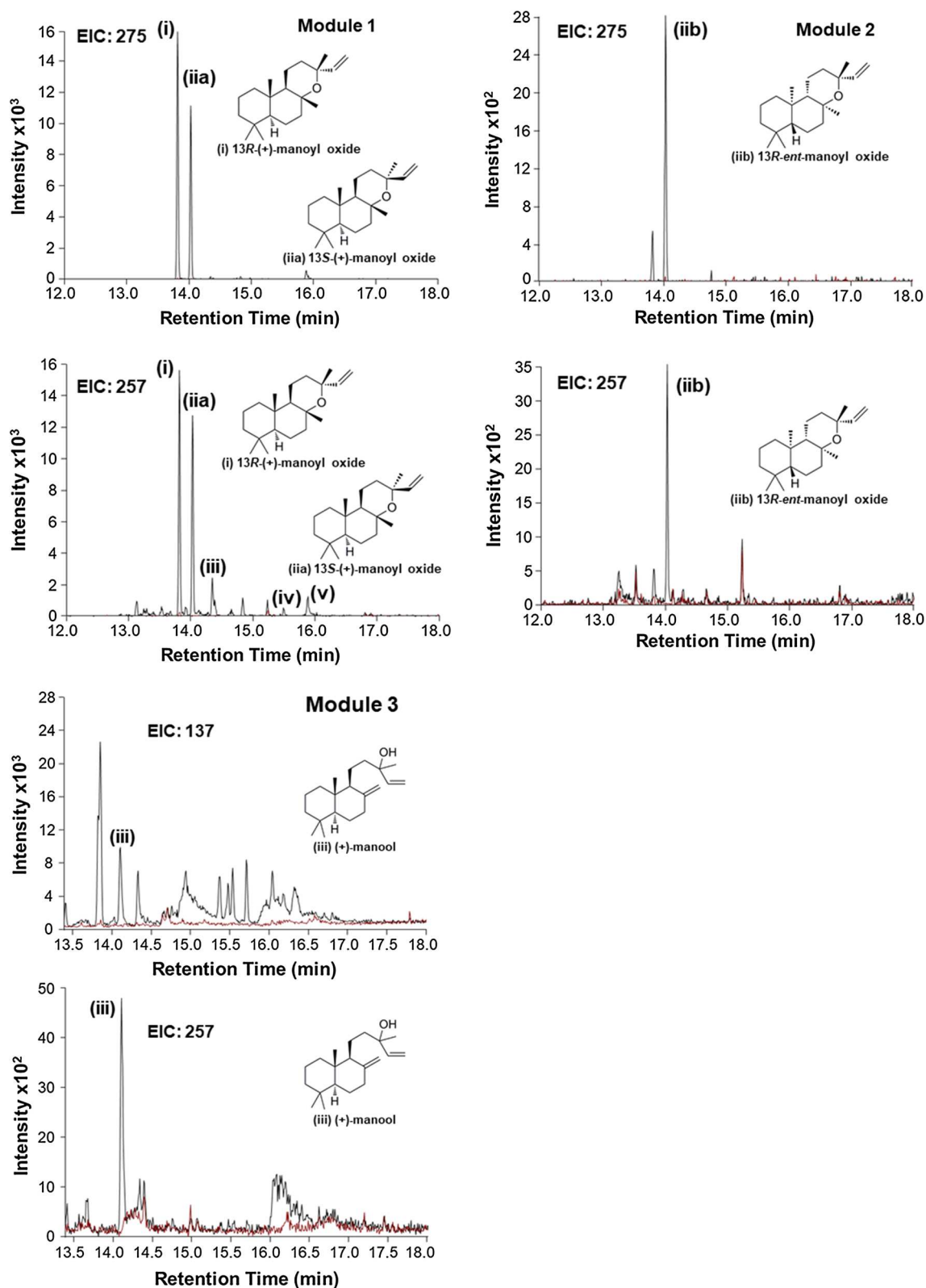
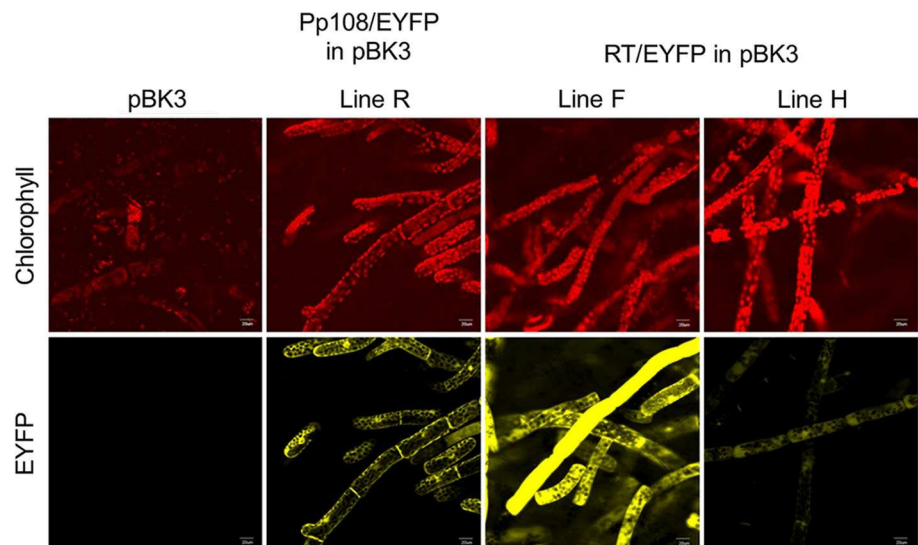


Fig. 4 Extracted ion chromatograms of pBK3 line transformed with different diterpene modules in the *Pp108* locus. Black and red chromatograms represent pBK3 line harboring diterpene modules and non-transformed pBK3 control (negative control) respectively

Fig. 5 Confocal microscopy images of representative lines transformed with *EYFP* at different loci



Evidence of assembly of the individual fragments (GNZ, *EYFP*, and OPO) and amplification and sequencing of the GNZ fragment with the insertion junction with the 5'LTR region upstream of GAG support integration of the construct in *PpatensLTR2* (Fig. S2a and S2b).

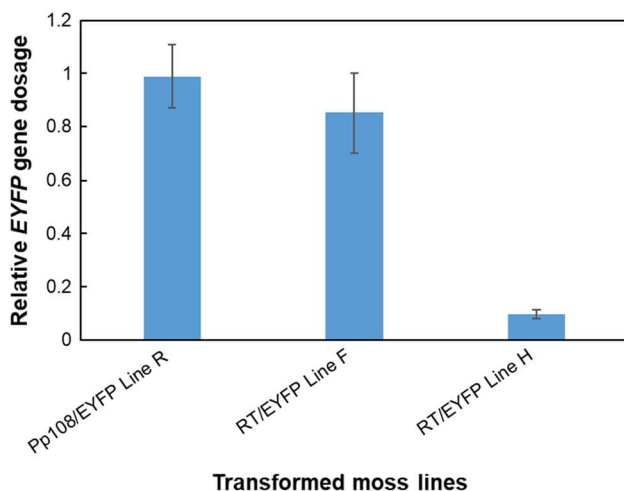


Fig. 6 Relative *EYFP* gene dosage of the various transformed moss lines with *EYFP* in the *Pp108* locus and *PpatensLTR2* loci. Quantitative real time PCR analysis was done with the genomic DNA isolated from the different *EYFP* lines. Relative *EYFP* gene dosage was analyzed using the *Pp108/EYFP* line R as reference. The single copy gene encoding CYP701B1 was used as the internal reference gene. Results are shown as mean $\pm$ SD ($n=6$, three technical repeats of two biological replicates for each line)

Discussion

Physcomitrella patens can be engineered for biosynthesis of modern diterpenes

In this work, we demonstrated a modular approach using linear DNA fragments for homologous recombination and engineering of diterpene biosynthetic pathways in *P. patens*. The regions of the genomic loci for homologous recombination were linked to the genes for diterpene biosynthesis and the entire expression cassette was assembled in vivo and integrated into the genome by homologous recombination. Typical biosynthetic pathways to diterpene scaffolds in higher plants consist of pairs of class II and class I diterpene synthases. The modular approach allows convenient recombination of diTPSs to generate chemical diversity. It has been shown earlier that *P. patens* can be engineered for the production of the diterpenes taxa-4(5), 11(12)-diene and sclareol (Anterola et al. 2009; Pan et al. 2015). In both cases, the required building blocks were assembled in a vector, which was subsequently linearized to obtain the expression cassette for genome targeting by homologous recombination. Here, we simplified the procedure and adapted it for combinatorial assembly of the essential parts required, namely the promoter fragment, the fragment containing the target gene/s, and the terminator fragment, all synthesized using PCR. Linear DNA fragments were then assembled in vivo during the transformation to generate the expression cassette.

Genome editing exploiting homologous recombination is common in yeast (*Saccharomyces cerevisiae*). Therefore, it has been widely used as production host for biosynthesis of various commercial products using synthetic biology and metabolic engineering (David and Siewers 2015; Krivoruchko et al. 2011; Xu et al. 2013). In contrast,

photosynthetic hosts offer a range of advantages over microbial production of terpenoids, including abundance of reduction equivalents and compartmentation. Nevertheless, disturbance through metabolic engineering of the tightly regulated biosynthesis of the diterpene phytohormones, derived from gibberellic acid, was suggested as factor limiting growth and development of the host plants (Besumbes et al. 2004; Kovacs et al. 2007). While *P. patens* lacks gibberellic acid, early diterpenoid intermediates accumulate. We started from a *P. patens* genetic background, pBK3, with disrupted biosynthesis of endogenous diterpenoids. Building and expanding on earlier findings, we demonstrated formation of a range of diterpenes from higher land plants by combinations of class II and class I diTPSs. No growth impairment was observed for any of our diterpene producing transgenic *P. patens* lines compared to the WT (Fig. S7). An accumulation of diterpenes ranging from 0.13 ± 0.02 ng of 13*R*-ent-manoyl oxide/mg of fresh weight for the diterpene module M2 to 0.74 ± 0.14 ng of total diterpene/mg of fresh weight for diterpene module M1 was observed in this work which is lower compared to that of taxa-4(5), 11(12)-diene ($0.5 \mu\text{g}$ taxadiene/gm of fresh weight) and sclareol ($2.84 \mu\text{g}$ sclareol/gm dry weight) reported earlier (Anterola et al. 2009; Pan et al. 2015). Potential reasons for those differences could be the single enzyme pathway to taxadiene, or the exceptionally efficient biosynthesis of sclareol. Our results indicate that *P. patens* has potential as production host for metabolic engineering of structurally diverse diterpenes evolved in modern land plants. Boosting the levels of diterpene products may be achieved by co-expression of DXS and GGDPS, enzymatic steps in the biosynthesis of the acyclic C-20 precursor GGDP, and as found highly successful in the transient *Nicotiana benthamiana* system (Bruckner and Tissier 2013; Gnanasekaran et al. 2015).

Retrotransposon region as potential integration sites

Exploiting highly repetitive retrotransposon loci as potential integration sites for metabolic engineering was reported in *S. cerevisiae* (Cho et al. 1999; Fang et al. 2011; Juretzek et al. 2001; Ohgiya et al. 1997). We report targeting of simple constructs to a retrotransposon region in *P. patens*. The reporter gene *EYFP* was targeted to loci of *PpatensLTR2* or the neutral locus *Pp108*. While reporter activity was stable across lines targeting the neutral locus, substantial variation in fluorescence signal was observed in lines targeting the retrotransposon loci. We investigated whether difference in gene dosage could contribute to the observed variation. Even though presence of a concatemer was detected, the gene dosage did not correlate with the difference in fluorescence. Alternatively, it is tempting to speculate that positional effects, caused by the integration of transgenes

into differentially expressed regions of the genome, may influence the expression level. These positional effects may include epigenetic regulation of gene expression at the retrotransposon loci as well as modulation of gene expression by sequences acting as cryptic promoters and/or terminators at those loci (Lisch and Bennetzen 2011; Lisch and Slotkin 2011).

Retrotransposon regions were not successful for generating stable production lines

Our attempts to target two diterpene modules to *PpatensLTR2* loci were not successful. 244 and 76 kanamycin resistant transformants for M2 and M3 were genotyped. We found for M2 that the fragments were not properly assembled in 14 lines, i.e., we detected the junction between class I diterpene synthase of the M2 module and OPO-fragment but not between the GNZ-fragment and the class II diterpene synthase of the M2 module, indicating only partial assembly. The remaining kanamycin resistant lines were also found incomplete as only the selection marker containing fragment could be detected. None of the kanamycin resistant M3 lines were completely assembled, as shown by genotyping. It is possible that this incomplete assembly is due to the markedly increased size of the diterpene modules (approx. 4800 bp) compared to *EYFP* (717 bp). Considering the divergence in the recombination sites across the copies scattered over the genome, i.e., the 5' end of the *GAG* and 3' end of the *POL* region with our cloned single representative region of *PpatensLTR2* (Fig. S4 and S5), the larger diterpene modules may reduce integration and homologous recombination efficiency. This observation, combined with the strong variation in reporter activity from those loci successfully targeted, indicate that the potential these new integration sites may offer should be carefully weighed against features of the limited number of neutral sites.

Conclusion

We have shown here that diterpene biosynthetic pathways involving novel combinations of class II and class I diTPSs can be successfully engineered at the *Pp108* locus in *P. patens*. We built a modular strategy where the flanking regions required for homologous recombination with the linear fragments of promoter and terminator fragments were fused to target genes. This allowed convenient shuffling of the target genes, while keeping the design of the promoter and terminator fragments unaltered.

Detection of concatemers indicates that genome editing, even at the neutral *Pp108* locus, is more complicated than expected. Exploring redundant retrotransposon regions of *PpatensLTR2* as integration sites with a reporter construct

was successful, but installation of the bigger diterpene modules remained a challenge. We found that one of the *PpatensLTR2/EYFP* lines had higher expression of *EYFP* compared to that of lines for *Pp108/EYFP*, possibly due to positional effect. Exploiting this for metabolic engineering in generation of non-neutral loci carrying diterpene modules will require adjustments of the experimental strategy due to the decreased efficiency in multi-piece in vivo assembly.

Author contribution statement BH and JAA conceived and designed research. AB and JAA conducted experiments. DM and BBM contributed critical technical assistance to carry out the experiments. SRJ performed the genome-wide mapping of *PpatensLTR2* as well as *GAG* and *POL*. BH, JAA and AB analyzed data and created the figures. AB and BH wrote the manuscript. All authors read and approved the manuscript.

Acknowledgements We thank Dr. Henrik Toft Simonsen (Department of Systems Biology, Technical University of Denmark) for providing the pBK3 strain. We would also like to thank Britta Hamberger, Jiaxing Chen, Amanda Agosto, and Peiyen Kuo at Michigan State University for excellent technical support. D.M. acknowledges the Bachelor PLUS/International Biochemistry exchange program between Heinrich-Heine-University Düsseldorf and Michigan State University. This work was primarily supported by the U.S. Department of Energy-Great Lakes Bioenergy Research Center Cooperative Agreement DE-FC02-07ER64494 and DE-SC0018409. B.H. gratefully acknowledges the Michigan State University Strategic Partnership Grant program (“Evolutionary-Driven Genome Mining of Plant Biosynthetic Pathways” and “Plant-inspired Chemical Diversity”), startup funding from the Department of Molecular Biology and Biochemistry, Michigan State University and support by the USDA National Institute of Food and Agriculture, HATCH project MICL02454. Mass spectra were collected using instruments at the Michigan State University Mass Spectrometry and Metabolomics core. B.H. is in part supported by the National Science Foundation under Grant Number 1737898. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the author(s) and do not necessarily reflect the views of the National Science Foundation.

Open Access This article is distributed under the terms of the Creative Commons Attribution 4.0 International License (<http://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made.

References

- Ajikumar PK et al (2010) Isoprenoid pathway optimization for taxol precursor overproduction in *Escherichia coli*. *Science* 330:70–74. <https://doi.org/10.1126/science.1191652>
- Andersen-Ranberg J et al (2016) Expanding the landscape of diterpene structural diversity through stereochemically controlled combinatorial biosynthesis. *Angew Chem Int Ed* 55:2142–2146. <https://doi.org/10.1002/anie.201510650>
- Anterola A, Shanle E, Perroud PF, Quatrano R (2009) Production of taxa-4(5),11(12)-diene by transgenic *Physcomitrella patens*. *Transgenic Res* 18:655–660. <https://doi.org/10.1007/s11248-009-9252-5>
- Ashton NW, Champagne CEM, Weiler T, Verkoczy LK (2000) The bryophyte *Physcomitrella patens* replicates extrachromosomal transgenic elements. *New Phytol* 146:391–402. <https://doi.org/10.1046/j.1469-8137.2000.00671.x>
- Besumbes O, Sauret-Gueto S, Phillips MA, Imperial S, Rodriguez-Concepcion M, Boronat A (2004) Metabolic engineering of isoprenoid biosynthesis in *Arabidopsis* for the production of taxadiene, the first committed precursor of taxol. *Biotechnol Bioeng* 88:168–175. <https://doi.org/10.1002/bit.20237>
- Bruckner K, Tissier A (2013) High-level diterpene production by transient expression in *Nicotiana benthamiana*. *Plant Methods* 9:46. <https://doi.org/10.1186/1746-4811-9-46>
- Cho KM, Yoo YJ, Kang HS (1999) δ -Integration of endo/exo-glucanase and β -glucosidase genes into the yeast chromosomes for direct conversion of cellulose to ethanol. *Enzyme Microb Technol* 25:23–30. [https://doi.org/10.1016/S0141-0229\(99\)00011-3](https://doi.org/10.1016/S0141-0229(99)00011-3)
- David F, Siewers V (2015) Advances in yeast genome engineering. *FEMS Yeast Res* 15:1–14. <https://doi.org/10.1111/1567-1364.12200>
- Englund E, Andersen-Ranberg J, Miao R, Hamberger B, Lindberg P (2015) Metabolic engineering of *Synechocystis* sp. PCC 6803 for production of the plant diterpenoid manoyl oxide. *ACS Synth Biol* 4:1270–1278. <https://doi.org/10.1021/acssynbio.5b00070>
- Fang F et al (2011) A vector set for systematic metabolic engineering in *Saccharomyces cerevisiae*. *Yeast* 28:123–136. <https://doi.org/10.1002/yea.1824>
- François IEJA, Van Hemelrijck W, Aerts AM, Wouters PFJ, Proost P, Broekaert WF, Cammue BPA (2004) Processing in *Arabidopsis thaliana* of a heterologous polypeptide resulting in differential targeting of the individual plant defensins. *Plant Sci* 166:113–121. <https://doi.org/10.1016/j.plantsci.2003.09.001>
- Gnanasekaran T, Vavitsas K, Andersen-Ranberg J, Nielsen AZ, Olsen CE, Hamberger B, Jensen PE (2015) Heterologous expression of the isopimaric acid pathway in *Nicotiana benthamiana* and the effect of N-terminal modifications of the involved cytochrome P450 enzyme. *J Biol Eng* 9:24. <https://doi.org/10.1186/s13036-015-0022-z>
- Goodstein DM et al (2012) Phytozome: a comparative platform for green plant genomics. *Nucleic Acids Res* 40:D1178–D1186. <https://doi.org/10.1093/nar/gkr944>
- Hamberger B, Ohnishi T, Hamberger B, Seguin A, Bohlmann J (2011) Evolution of diterpene metabolism: Sitka spruce CYP720B4 catalyzes multiple oxidations in resin acid biosynthesis of conifer defense against insects. *Plant Physiol* 157:1677–1695. <https://doi.org/10.1104/pp.111.185843>
- Helliwell CA, Poole A, Peacock WJ, Dennis ES (1999) *Arabidopsis* ent-kaurene oxidase catalyzes three steps of gibberellin biosynthesis. *Plant Physiol* 119:507–510
- Helliwell CA, Sullivan JA, Mould RM, Gray JC, Peacock WJ, Dennis ES (2001) A plastid envelope location of *Arabidopsis* ent-kaurene oxidase links the plastid and endoplasmic reticulum steps of the gibberellin biosynthesis pathway. *Plant J* 201:8
- Huang Q, Roessner CA, Croteau R, Scott AI (2001) Engineering *Escherichia coli* for the synthesis of taxadiene, a key intermediate in the biosynthesis of taxol. *Bioorg Med Chem* 9:2237–2242
- Juretzek T, Le Dall M-T, Mauersberger S, Gaillardin C, Barth G, Nicaud J (2001) Vectors for gene expression and amplification in the yeast *Yarrowia lipolytica*. *Yeast* 18:97–113. [https://doi.org/10.1002/1097-0061\(20010130\)18:2%3c97::AID-YEA652%3e3.0.CO;2-U](https://doi.org/10.1002/1097-0061(20010130)18:2%3c97::AID-YEA652%3e3.0.CO;2-U)

- Keeling CI, Bohlmann J (2006) Diterpene resin acids in conifers. *Phytochemistry* 67:2415–2423. <https://doi.org/10.1016/j.phytochem.2006.08.019>
- Kempinski C, Jiang Z, Bell S, Chappell J (2015) Metabolic engineering of higher plants and algae for isoprenoid production. *Adv Biochem Eng Biotechnol* 148:161–199. https://doi.org/10.1007/10_2014_290
- Khairul Ikram NKB, Beyraghdar Kashkooli A, Peramuna AV, van der Krol AR, Bouwmeester H, Simonsen HT (2017) Stable production of the antimalarial drug artemisinin in the moss *Physcomitrella patens*. *Front Bioeng Biotechnol* 5:47. <https://doi.org/10.3389/fbioe.2017.00047>
- King BC et al (2016) In vivo assembly of DNA-fragments in the moss, *Physcomitrella patens*. *Sci Rep* 6:25030. <https://doi.org/10.1038/srep25030>
- Kovacs K, Zhang L, Linforth RST, Whittaker B, Hayes CJ, Fray RG (2007) Redirection of carotenoid metabolism for the efficient production of taxadiene [taxa-4(5),11(12)-diene] in transgenic tomato fruit. *Transgenic Res* 16:121–126. <https://doi.org/10.1007/s11248-006-9039-x>
- Krivoruchko A, Siewers V, Nielsen J (2011) Opportunities for yeast metabolic engineering: lessons from synthetic biology. *Biotechnol J* 6:262–276. <https://doi.org/10.1002/biot.201000308>
- Lang D, Zimmer AD, Rensing SA, Reski R (2008) Exploring plant biodiversity: the *Physcomitrella* genome and beyond. *Trends Plant Sci* 13:542–549. <https://doi.org/10.1016/j.tplants.2008.07.002>
- Lang D et al (2018) The *Physcomitrella patens* chromosome-scale assembly reveals moss genome structure and evolution. *Plant J* 93:515–533. <https://doi.org/10.1111/tpj.13801>
- Lisch D, Bennetzen JL (2011) Transposable element origins of epigenetic gene regulation. *Curr Opin Plant Biol* 14:156–161. <https://doi.org/10.1016/j.pbi.2011.01.003>
- Lisch D, Slotkin RK (2011) Strategies for silencing and escape: the ancient struggle between transposable elements and their hosts. *Int Rev Cell Mol Biol* 292:119–152. <https://doi.org/10.1016/B978-0-12-386033-0.00003-7>
- Liu YC, Vidali L (2011) Efficient polyethylene glycol (PEG) mediated transformation of the moss *Physcomitrella patens*. *J Vis Exp* 50:e2560. <https://doi.org/10.3791/2560>
- Muren E, Nilsson A, Ulfstedt M, Johansson M, Ronne H (2009) Rescue and characterization of episomally replicating DNA from the moss *Physcomitrella*. *Proc Natl Acad Sci USA* 106:19444–19449. <https://doi.org/10.1073/pnas.0908037106>
- Novikova O, Mayorov V, Smyshlyaev G, Fursov M, Adkinson L, Pisarenko O, Blinov A (2008) Novel clades of chromodomain-containing Gypsy LTR retrotransposons from mosses (Bryophyta). *Plant J* 56:562–574. <https://doi.org/10.1111/j.1365-3113X.2008.03621.x>
- Ohgiya S, Goda T, Hoshino T, Kamataki T, Ishizaki K (1997) Establishment of a novel host, high-red yeast that stably expresses hamster NADPH–cytochrome P450 oxidoreductase: usefulness for examination of the function of mammalian cytochrome P450. *Arch Biochem Biophys* 343:215–224. <https://doi.org/10.1006/abbi.1997.0148>
- Pan X-W, Han L, Zhang Y-H, Chen D-F, Simonsen HT (2015) Sclareol production in the moss *Physcomitrella patens* and observations on growth and terpenoid biosynthesis. *Plant Biotechnol Rep* 9:149–159. <https://doi.org/10.1007/s11816-015-0353-8>
- Pfaffl MW (2004) Quantification strategies in real-time PCR. In: Bustin SA (ed) A-Z of quantitative PCR. International University Line, La Jolla, pp 87–112
- Reski R, Parsons J, Decker EL (2015) Moss-made pharmaceuticals: from bench to bedside. *Plant Biotechnol J* 13:1191–1198. <https://doi.org/10.1111/pbi.12401>
- Reski R, Bae H, Simonsen HT (2018) *Physcomitrella patens*, a versatile synthetic biology chassis. *Plant Cell Rep*. <https://doi.org/10.1007/s00299-018-2293-6>
- Schaefer DG, Zryd JP (1997) Efficient gene targeting in the moss *Physcomitrella patens*. *Plant J* 11:1195–1206
- Shevchuk NA, Bryksin AV, Nusinovich YA, Cabello FC, Sutherland M, Ladisch S (2004) Construction of long DNA molecules using long PCR-based fusion of several fragments simultaneously. *Nucleic Acids Res* 32:1–12. <https://doi.org/10.1093/nar/gnh014>
- Xu P, Bhan N, Koffas MAG (2013) Engineering plant metabolism into microbes: from systems biology to synthetic biology. *Curr Opin Biotechnol* 24:291–299. <https://doi.org/10.1016/j.copbio.2012.08.010>
- Zerbe P, Bohlmann J (2015) Plant diterpene synthases: exploring modularity and metabolic diversity for bioengineering. *Trends Biotechnol* 33:419–428. <https://doi.org/10.1016/j.tibtech.2015.04.006>
- Zhan X, Zhang Y-H, Chen D-F, Simonsen HT (2014) Metabolic engineering of the moss *Physcomitrella patens* to produce the sesquiterpenoids patchoulol and α/β -santalene. *Front Plant Sci* 5:636. <https://doi.org/10.3389/fpls.2014.00636>
- Zhan X, Bach SS, Hansen NL, Lunde C, Simonsen HT (2015) Additional diterpenes from *Physcomitrella patens* synthesized by copalyl diphosphate/kaurene synthase (PpCPS/KS). *Plant Physiol Biochem* 96:110–114. <https://doi.org/10.1016/j.plaphy.2015.07.011>