Chapter 18

High-Throughput Testing of Terpenoid Biosynthesis Candidate Genes Using Transient Expression in *Nicotiana benthamiana*

Søren Spanner Bach, Jean-Étienne Bassard, Johan Andersen-Ranberg, Morten Emil Møldrup, Henrik Toft Simonsen, and Björn Hamberger

Abstract

To respond to the rapidly growing number of genes putatively involved in terpenoid metabolism, a robust high-throughput platform for functional testing is needed. An *in planta* expression system offers several advantages such as the capacity to produce correctly folded and active enzymes localized to the native compartments, unlike microbial or prokaryotic expression systems. Two inherent drawbacks of plant-based expression systems, time-consuming generation of transgenic plant lines and challenging gene-stacking, can be circumvented by transient expression in *Nicotiana benthamiana*. In this chapter we describe an expression platform for rapid testing of candidate terpenoid biosynthetic genes based on *Agrobacterium* mediated gene expression in *N. benthamiana* leaves. Simultaneous expression of multiple genes is facilitated by co-infiltration of leaves with several engineered *Agrobacterium* strains, possibly making this the fastest and most convenient system for the assembly of plant terpenoid biosynthetic routes. Tools for cloning of expression plasmids, *N. benthamiana* culturing, *Agrobacterium* preparation, leaf infiltration, metabolite extraction, and automated GC-MS data mining are provided. With all steps optimized for high throughput, this *in planta* expression platform is particularly suited for testing large panels of candidate genes in all possible permutations.

Key words *Nicotiana benthamiana*, *In planta*, Transient expression, *Agrobacterium*, High throughput, Enzyme characterization, Terpenoid pathway discovery, Metabolic engineering

1 Introduction

New and increasingly affordable sequencing techniques have become an important tool for discovery of terpenoid pathways. Key biosynthetic steps are cyclization of the terpenoid backbone by terpene synthases (TPS) and oxidative decoration through cytochromes P450 (P450s). TPS are typically encoded by small to medium sized gene families, while genes encoding P450s can reach 1 % of the total genes in vascular plants [1, 2]. The resulting large

Manuel Rodríguez-Concepción (ed.), *Plant Isoprenoids: Methods and Protocols*, Methods in Molecular Biology, vol. 1153, DOI 10.1007/978-1-4939-0606-2_18,

number of candidate genes presents a major challenge in discovery and reconstitution of multi-enzyme terpenoid pathways.

Heterologous expression in bacteria, yeast, or insect cells is the most common approach for functional characterization of terpenoid enzymes from plants. Nevertheless, expression of TPS and P450s in non-plant hosts is no trivial task, which is reflected by the range of strategies employed to optimize expression. These include codon optimization of the candidate genes and the use of host strains overexpressing tRNAs for rare plant codons. Furthermore, monoterpene and diterpene synthases are plastid targeted enzymes and the N-terminal plastid transit peptide can impair the catalytic activity of these enzymes or cause formation of inclusion bodies [3, 4]. Thus, pseudomature proteins, where the plastid transit peptide has been removed, can be characterized instead [4]. Expression and characterization of plant P450s is also challenging in non-plant systems. Bacteria lack the endoplasmic reticulum (ER), where eukaryotic P450s are localized. However, N-terminal amino acid truncations or modifications can lead to expression of functional P450s in bacterial systems [5]. Although yeast has an ER compartment, N-terminal modifications such as membrane anchor swapping has been shown to improve the chance of obtaining functional P450s from vascular plants [6]. Optimization of expression conditions, assay conditions, or precursor concentrations, has also been shown to yield functional enzymes. All the mentioned engineering and optimization steps can be time demanding and are undesirable in a high-throughput context. Moreover, the enzymatic activity can be influenced by the cellular environment of the expression host. For example, the enzymatic products of a Norway spruce diterpene synthase were shown to differ between in vitro and in vivo expression systems [7, 8], whereas the products of a monoterpene synthase from sweet basil expressed in grapevine, *Arabidopsis* or tobacco differed from the products observed when the same enzyme was expressed in microbial systems [9].

The protein synthesis machinery, posttranslational modifications, cellular compartments, and the function of transit peptides for protein targeting are conserved among higher plants. Thus, plant expression systems have the capacity to express native plant terpenoid genes without any modifications of the coding sequence. Additionally, *in planta* expression systems contain the machineries for posttranslational modification and protein targeting, the relevant cellular compartments (plastids and endoplasmic reticulum), and the universal C_5 building blocks for isoprenoid biosynthesis. However, most plant based systems are not suitable for high-throughput approaches due to time consuming generation of transgenic lines and challenges of gene stacking. In this chapter we present an *in planta* expression platform suitable for high-throughput screening of terpenoid candidate genes using *Agrobacterium tumefaciens* mediated transient expression in

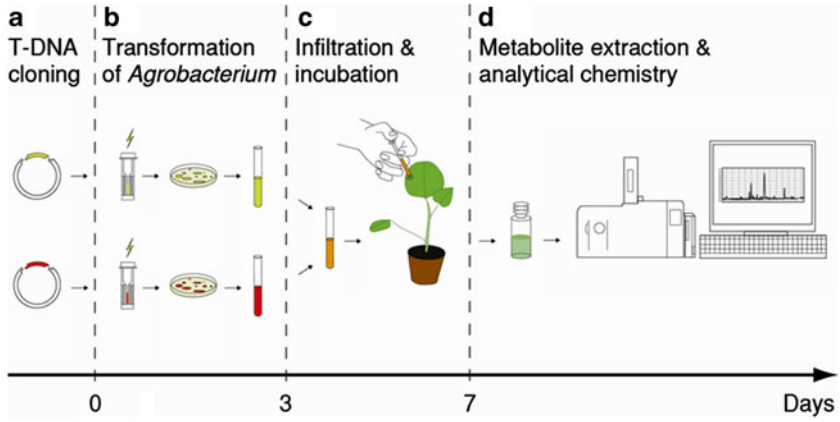

Fig. 1 Workflow illustrating co-expression of two candidate genes. (**a**) Transfer DNA (T-DNA) expression plasmids are created by cloning of native and full-length candidate gene PCR products. (**b**) Expression plasmids are transformed into *Agrobacterium* by electroporation and overnight cultures are subsequently inoculated. (**c**) Different *Agrobacterium* strains, carrying different genes of interest are mixed and co-infiltrated into leaves of *N. benthamiana*. This facilitates co-expression. (**d**) Following 4 days of incubation, metabolites are extracted and analyzed by GC-MS. After the T-DNA expression plasmids have been created, transformation, infiltration, and metabolite analysis can be performed within 7 days

Nicotiana benthamiana leaves. This system facilitates testing of hundreds of native full-length gene combinations in 7 days with minimal hands-on time (Fig. 1). In terms of workload and experimental time-line, this is comparable with microbial based heterologous systems (*see* Chapter 17) but with the advantages of a plant-based expression system. In our laboratory we have successfully expressed mono-, sesqui-, di-, and tri-terpene synthases as native full-length genes. Detection of the corresponding terpenoid products did not require supplementation of substrates, indicating availability of endogenous precursor molecules. In addition, a simple terpenoid biosynthetic pathway has been reconstituted by co-expression of a TPS and a P450. The "plug and play" nature of this transient expression platform is suitable for high-throughput screening of candidate genes, and the option of easy co-expression of multiple genes makes it a very valuable tool in discovery and reconstitution of multi-enzyme terpenoid pathways from vascular plants. The protocol presented here provides all necessary information for establishing this transient *in planta* expression platform in any molecular biological laboratory with access to plant growth and analytical facilities.

2 Materials

2.1 Growth of Biological Materials

1. *N. benthamiana* seeds can be obtained from the Tobacco Institute of Bergerac in France (http://www.imperial-tobacco-bergerac.com).

2. 5.5 cm plastic pots and a tray.
3. Greenhouse soil (e.g., Pindstrup substrate no. 2, blonde sphagnum peat with granulated clay, medium fertilizer level, pH 6.0).
4. Biological larvicide (e.g., Vectobac®).
5. Toothpicks.
6. Transparent film and Parafilm.
7. Common garden NPK 5-1-4 fertilizer.
8. Greenhouse with 16 h light (24 °C) and 8 h dark (17 °C) period, 80 % humidity (*see* **Note 1**).
9. *Agrobacterium* strain AGL-1—GV3850 [10] (*see* **Note 2**).
10. T-DNA expression plasmid encoding the anti-posttranscriptional gene silencing protein p19 (35S:p19) [11].
11. LB Medium: LB Broth low salt, adjusted to pH 7.2 with NaOH, with and without 50 μg/mL kanamycin, 34 μg/mL rifampicin, or 50 μg/mL carbenicillin. For solid media add 1.5 % agar.
12. TE buffer: 10 mM Tris–HCl pH 8.0, 1 mM EDTA.
13. 10 % (v/v) sterile glycerol solution.

2.2 General Laboratory Equipment

1. Incubators at 28 °C with and without agitation.
2. Tabletop centrifuge.
3. Spectrophotometer and UV cuvette.
4. Liquid nitrogen.
5. –80 °C freezer.
6. Electroporation cuvette (Bio-Rad; 2 mm) and Gene Pulser (Bio-Rad; Capacity 25 μF; 2.5 kV; 400 Ω).
7. Water spray atomizer.
8. Razor blade.

2.3 Metabolite Analysis

1. Silanized 2 and 20 mL glass vials with screw top including screw caps with Teflon-coated silicone septa.
2. Solid-phase micro-extraction (SPME) fiber made of polydimethylsiloxane (PDMS). Grey fiber.
3. GC-MS: Shimadzu GC-2010GC, with a 30 m Agilent HP-5MS column (250 μm i.d., 0.25 μm film) and helium as carrier gas. CTC AOC-5000 auto injector and a GCMS-QP2010 PlusMS Detector using electrical ionization mode. Similar GC-MS systems can be applied.
4. Internal standards (e.g., limonene, valencene or 1-eicosene). C7-C30 alkane standard for calculation of retention index.

5. Solvent for extraction e.g., GC-grade hexane.
6. Methanol (GC-grade) and trimethylsilyl (TMS) diazomethane for derivatization.
7. N_2 gas (grade 2 or higher).

3 Methods

Transient expression in *N. benthamiana* has been used for numerous research applications; however this protocol is optimized for rapid characterization of terpene genes.

3.1 Growth of N. benthamiana Plants

1. Fill the tray with pots containing greenhouse soil. Use water supplemented with a biological larvicide (e.g., Vectobac®) to wet the soil and a toothpick to sow a single seed in each pot.
2. Cover the tray with transparent film, to keep the seeds moist, and move the tray into the greenhouse under 16 h light (24 °C) and 8 h dark (17 °C) regime and 80 % humidity (*see* **Note 1**).
3. When seedlings start to emerge (normally after 3–4 days), remove the film and cover the bottom of the tray with water three times per week using general NPK fertilized water once per week.
4. Plants are ready for infiltration when they have developed 5–6 leaves. This usually takes 4–6 weeks.

3.2 Construction of T-DNA Expression Plasmids

T-DNA vectors for *Agrobacterium* infiltration can be created using many different cloning strategies. We find the Uracil-Specific Excision Reagent (USER™) cloning technique robust and suitable for high-throughput cloning. However, other rapid-cloning techniques can be used (e.g., In-Fusion HD® cloning or Gibson Assembly™). A subset of T-DNA expression vectors can conveniently be obtained from pCAMBIA (http://www.cambia.org/daisy/cambia/585). This protocol is intended for T-DNA expression plasmids encoding a bacterial kanamycin resistance marker (e.g., pCAMBIA130035Su) [12].

3.3 Preparation of Electrocompetent Agrobacterium

Electrocompetent A*grobacterium* cells can also be acquired from commercial providers or prepared using standard procedures (*see* **Note 2**).

1. *Agrobacterium* is stored as a 50 % glycerol stock at −80 °C.
2. Inoculate a petri dish containing LB medium and appropriate antibiotics with the *Agrobacterium* strain AGL-1 (rifampicin resistant) harboring the pGV3850 Ti plasmid (providing carbenicillin resistance) and incubate at 28 °C for 1–2 days until colonies are visible.

3. Inoculate 5 mL of LB liquid medium, including the antibiotics and incubate overnight at 28 °C under 200 rpm agitation.
4. Use the starter culture to inoculate 500 mL of LB medium including antibiotics, and incubate approximately 15–18 h at 28 °C with agitation (170 rpm) until the OD_{600} reaches 0.5–0.8.
5. Cool the culture to 4 °C (30 min) and centrifuge for 10 min at 3,000 × *g* at 4 °C.
6. Discard the supernatant and resuspend the pellet in 5 mL of ice-cold water. Adjust the volume to 500 mL with ice-cold water.
7. This washing step is repeated twice by resuspending in 250 and 50 mL of ice-cold water, respectively.
8. Finally resuspend in 5 mL of 10 % (v/v) ice-cold glycerol solution and freeze 50 μL aliquots in liquid nitrogen.
9. The electrocompetent *Agrobacterium* cells are stored at −80 °C.

3.4 Transformation of Agrobacterium Cells

1. Mix 50 ng of expression vector with 50 μL of electrocompetent *Agrobacterium* cells thawed on ice.
2. Transfer the mixture to a pre-cooled 2 mm electroporation cuvette and electroporate using a Gene Pulser (Capacity 25 μF; 2.5 kV; 400 Ω).
3. Let the transformed bacteria recover in 950 μL of LB media for 4 h at 28 °C under agitation (170 rpm).
4. Spin the cells 3 min at 3,000 × *g* and remove 900 μL of the supernatant. Resuspend the cells in the residual liquid and spread them on LB solid medium supplemented with kanamycin, rifampicin, and carbenicillin. If necessary, leave the lid open for 10 min to let excess water evaporate.
5. Seal the petri dish with parafilm and incubate at 28 °C. Colonies should appear after 48 h.
6. The plate can be stored at 4 °C (*see* **Note 3**).

3.5 Agroinfiltration of N. Benthamiana Leaves

1. Inoculate 10 mL of LB medium containing kanamycin, rifampicin, and carbenicillin with several *Agrobacterium* colonies from the same plate and grow for 20 h at 28 °C under vigorous shaking (200 rpm). In order to ensure sufficient airflow use a 50 mL culture tube with a loose lid taped on. This is performed for each of the constructs to be included in the experiment. Cultures of an *Agrobacterium* strain harboring a plasmid encoding the anti-posttranscriptional gene silencing protein p19 should be included to ensure high levels of transient expression [11] (*see* **Note 4**).

2. After 20 h incubation at 28 °C the bacteria should have reached an OD_{600} of 1.0 or higher. Pellet the bacteria by spinning at 4,000 × *g* for 10 min at 16 °C and resuspend in 5 mL of water.
3. Repeat twice to remove residual antibiotics.
4. Measure OD_{600} of each *Agrobacterium* culture and adjust the OD_{600} to 0.4 with water. Mix each *Agrobacterium* strain harboring the plasmids encoding your gene of interest and the *Agrobacterium* strain encoding the p19 protein in a 1:1 ratio. For co-infiltration of two *Agrobacterium* strains harboring different plasmids as well as the p19 strain, strains are mixed in a 1:1:1 ratio (*see* **Notes 5** and **6**). Three leaves can be infiltrated with approximately 5 mL of infiltration solution. This will generate experimental triplicates.
5. Spray the plants with water using a water spray atomizer 1–10 min before infiltration.
6. *N. benthamiana* leaves are infiltrated by placing a 1 mL syringe onto the abaxial side of a leaf. Support with a finger the adaxial side of the leaf while pressing the piston gently (*see* **Note 7**).
7. Avoid bubbles in the infiltration solution. Infiltrate 3 leaves per plant with the same constructs to yield experimental triplicates.
8. Return the plants to the greenhouse (*see* **Note 1**).
9. Metabolites are extracted 3–4 days after infiltration.

3.6 Metabolite Extraction

3.6.1 Analysis of Volatile Terpenoid Products

This method allows an automated high-throughput analysis of headspace for monoterpenoids and sesquiterpenoids. By using SPME fibers and auto sampling it is possible to screen several samples per hour with minimal hands-on time.

1. Harvest 1–2 leaves by cutting the petiole with the razor blade. Place leaves with the petiole down in a 20 mL GC vial containing 3 mL of sterile water.
2. Cap the vials and incubate the leaf material in the water for 24–48 h (*see* **Note 8**).
3. Inject a SPME fiber through the silicone septum of the vial and incubate the fiber, without touching the leaf, for 20 min at room temperature to adsorb volatiles and analyze by GC-MS.

3.6.2 Analysis of Nonvolatile Terpenoid Products

This method allows an automated high-throughput analysis of leaf extracts for diterpenoids and triterpenoids.

1. Transfer two leaf disks (approximately 3 cm in diameter) to a 2 mL amber GC vial containing 1 mL of hexane spiked with an internal standard (e.g., 1 mg/L 1-eicosene) (*see* **Note 9**).
2. Vortex vigorously for 15 s and leave for 1 h on an orbital shaker (150 rpm) at room temperature.

3. Spin the vials at 2,500 × *g* for 30 s to sediment leaf material and transfer the solvent into new GC vials avoiding transfer of leaf material.
4. Capped samples can be stored at –20 °C until GC-MS analysis.

3.6.3 Derivatization of Oxidized Terpenoid Acid Products

1. Transfer 400 μL of solvent to a new 2 mL GC vial and add 200 μL of methanol and 100 μL of TMS diazomethane. In the presence of methanol the TMS diazomethane reacts with the carboxylic acid and yields a more volatile methyl ester.
2. Vortex briefly and leave at room temperature for 20 min.
3. Evaporate the solvent under a gentle nitrogen-stream and resuspend the residue in 400 μL of hexane by vortexing.
4. Samples can be stored at –20 °C until GC-MS analysis (*see* **Note 10**).

3.6.4 Metabolite Analysis Using GC-MS

In this section a general GC-MS program suitable for detection of most terpenoids is provided (mono- to diterpenoids). These general conditions can be further modified to shorten runtime, improve resolution or sensitivity for specific compounds. Injection volume 1 μL, splitless injector temperature 250 °C, 30 m HP-5MS column, 2 mL/min He carrier gas, GC temperature program (50 °C for 3 min, 15 °C/min to 310 °C and hold 5 min), scan range from *m/z* 50 to *m/z* 350 with 70 eV electrical ionization.

3.7 High-Throughput Data Mining

When performing large combinatorial co-expression experiments in *N. benthamiana*, detection and identification of products can be challenging and time consuming. A general strategy on how this process can be automated in order to ease the task of elucidating patterns in compounds detected from numerous samples is provided here. Furthermore, tips are given how the quantification of peak-area for each chromatogram can be exported into for example Microsoft Excel, avoiding excessive manual cut and paste. This automated data mining is especially suitable for large experimental setups containing complex mixtures of known and unknown terpenoid products.

First a custom library of compounds is built. Starting with the empty vector control and all available authentic standards, each compound is added to the custom compound library together with the corresponding mass fingerprint and retention index. The compound is assigned a unique identifier (e.g., name or number). In the analysis of the following biological sample the newly built library is used to identify known compounds and compounds constituting the phytochemical background of *N. benthamiana*. New compounds, not found in the custom library, are added to the library. As the number of previously detected compounds in the custom library grows, the chromatogram analysis becomes faster.

Most modern GC-MS software can export information of detected compounds including their unique identifier, peak-area and other relevant information. Our software (Shimadzu GCMS Postrun Analysis Program) has the option of exporting this information as text files. Using basic data processing software (Matlab, C++, or macros in Microsoft Excel) the data files can rapidly be imported into Microsoft Excel.

4 Notes

1. Since transient expression in *N. benthamiana* is very robust, growth conditions can be varied.
2. Various *Agrobacterium* strains can be used for transient *N. benthamiana* infiltration. In our laboratory we have successfully used the strains AGL-1, GV3850, and LB4404-virGN54D [14]. Competent ElectroMAX™ *Agrobacterium* LBA4404 cells can conveniently be purchased from Invitrogen or prepared using standard procedures [13].
3. In our experience freshly transformed *Agrobacterium* is best for transient expression; however, *Agrobacterium* can be kept on plates for a few weeks, without losing the ability to transform *N. benthamiana*.
4. In order to prevent posttranslational gene silencing we recommend co-infiltrating with an *Agrobacterium* strain harboring an expression plasmid encoding the p19 protein [11]. The *Agrobacterium* strain harboring p19 is included in each leaf infiltration and, thus, several 10 mL cultures of p19 *Agrobacterium* strain should be inoculated in parallel for each experiment.
5. It is possible to leave the *Agrobacterium* solution for several hours at this stage.
6. To optimize transformation efficiency, we recommend testing different densities of the bacterial suspension when using new *Agrobacterium* strains and new constructs. Furthermore, we recommend monitoring the metabolite accumulation from 3 to 10 days to determine the optimal time for metabolite extraction.
7. Choosing leaf material and conditions before infiltration is critical. Infiltration of medium sized leaves is easier and shows better reproducibility. By following guidelines the leaf stomata will be open which will ease the infiltration procedure and give best reproducibility in metabolite yield: Do not water the plants the last 24 h before infiltration. Avoid wind, temperature shifts, direct sunlight, and hygrometric shocks before and during infiltration. If possible, move the plants more than 2 h before infiltration for acclimation.

8. The incubation time in the GC vials can be extended if only minor levels of products are detected. A detached leaf can be incubated up to 1 week in a GC vial.
9. For normalization of the biomass a 50 mL falcon tube, cork borer or wad punch can be used to generate uniform leaf disks.
10. Glycosylation of transiently produced terpenoids at a hydroxyl-, or carboxyl-group has been reported, probably as a part of a detoxification metabolism [15]. However, in our lab glycosylation of terpenoid acids has not yet been observed. If terpenoid acids are glycosylated by *N. benthamiana* the glucose moieties can easily be hydrolyzed and the terpenoid can be detected after derivatization by GC-MS [15]. Alternatively one can employ LC-MS methods [16].

Acknowledgements

This work was supported by the UNIK [Universitetsforskningens Investeringskapital (Investment Capital for University Research)] Center for Synthetic Biology at the University of Copenhagen, funded by the Research Initiative of the Danish Ministry of Science, Technology, and Innovation and the Novo Nordisk Foundation, grant to BH.

References

1. Hamberger B, Bak S (2013) Plant P450s as versatile drivers for evolution of species-specific chemical diversity. Phil Trans R Soc B 368:20120426
2. Chen F, Tholl D, Bohlmann J et al (2011) The family of terpene synthases in plants: a mid-size family of genes for specialized metabolism that is highly diversified throughout the kingdom. Plant J 66:212–229
3. Williams DC, McGarvey DJ, Katahira EJ et al (1998) Truncation of limonene synthase preprotein provides a fully active "pseudomature" form of this monoterpene cyclase and reveals the function of the amino-terminal arginine pair. Biochem 37:12213–12220
4. Peters RJ, Flory JE, Jetter R et al (2000) Abietadiene synthase from grand fir (*Abies grandis*): characterization and mechanism of action of the "pseudomature" recombinant enzyme. Biochemistry 39:15592–15602
5. Barnes HJ, Arlotto MP, Waterman MR (1991) Expression and enzymatic activity of recombinant cytochrome P450 17-alpha-hydroxylase in *Escherichia coli*. Proc Natl Acad Sci U S A 88:5597–5601
6. Nedelkina S, Jupe SC, Blee KA et al (1999) Novel characteristics and regulation of a divergent cinnamate 4-hydroxylase (CYP73A15) from French bean: engineering expression in yeast. Plant Mol Biol 39:1079–1090
7. Hamberger B, Ohnishi T, Hamberger B et al (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
8. Martin DM, Fäldt J, Bohlmann J (2004) Functional characterization of nine Norway spruce TPS genes and evolution of gymnosperm terpene synthases of the TPS-d subfamily. Plant Physiol 135:1908–1927
9. Fischer MJ, Meyer S, Claudel P et al (2012) Specificity of *Ocimum basilicum* geraniol synthase modified by its expression in different heterologous systems. J Biotechnol 163:24–29
10. Zambryski P, Joos H, Genetello C et al (1983) Ti-plasmid vector for the introduction of

DNA into plant-cells without alteration of their normal regeneration capacity. EMBO J 2: 2143–2150

11. Voinnet O, Rivas S, Mestre P et al (2003) An enhanced transient expression system in plants based on suppression of gene silencing by the p19 protein of tomato bushy stunt virus. Plant J 33:949–956
12. Nour-Eldin HH, Hansen BG, Nørholm MHH et al (2006) Advancing uracil-excision based cloning towards an ideal technique for cloning PCR fragments. Nucleic Acids Res 34(18):e122
13. Weigel D, Glazebrook J (2006) Transformation of agrobacterium using electroporation. Cold Spring Harb Protoc. doi:10.1101/pdb.prot4665.
14. van der Fits L, Deakin EA, Hoge JHC et al (2000) The ternary transformation system: constitutive *virG* on a compatible plasmid dramatically increases Agrobacterium-mediated plant transformation. Plant Mol Biol 43:495–502
15. van Herpen TW, Cankar K, Nogueira M et al (2010) *Nicotiana benthamiana* as a production platform for artemisinin precursors. PLoS One 5:e14222
16. De Vos RCH, Moco S, Lommen A et al (2007) Untargeted large-scale plant metabolomics using liquid chromatography coupled to mass spectrometry. Nat Protoc 2:778–791