

Complete biosynthesis of cannabinoids and their unnatural analogues in yeast

Xiaozhou Luo^{1,15}, Michael A. Reiter^{1,2,15}, Leo d'Espaux^{3,12}, Jeff Wong^{3,12}, Charles M. Denby^{1,13}, Anna Lechner^{4,5,14}, Yunfeng Zhang^{1,6}, Adrian T. Grzybowski¹, Simon Harth³, Wei Yin Lin³, Hyunsu Lee^{3,7}, Changhua Yu^{3,5}, John Shin^{3,4}, Kai Deng^{8,9}, Veronica T. Benites³, George Wang³, Edward E. K. Baidoo³, Yan Chen³, Ishaan Dev^{3,4}, Christopher J. Petzold³ & Jay D. Keasling^{1,3,4,5,10,11*}

Cannabis sativa L. has been cultivated and used around the globe for its medicinal properties for millennia¹. Some cannabinoids, the hallmark constituents of *Cannabis*, and their analogues have been investigated extensively for their potential medical applications². Certain cannabinoid formulations have been approved as prescription drugs in several countries for the treatment of a range of human ailments³. However, the study and medicinal use of cannabinoids has been hampered by the legal scheduling of *Cannabis*, the low in planta abundances of nearly all of the dozens of known cannabinoids⁴, and their structural complexity, which limits bulk chemical synthesis. Here we report the complete biosynthesis of the major cannabinoids cannabigerolic acid, Δ^9 -tetrahydrocannabinolic acid, cannabidiolic acid, Δ^9 -tetrahydrocannabivarinic acid and cannabidivarinic acid in *Saccharomyces cerevisiae*, from the simple sugar galactose. To accomplish this, we engineered the native mevalonate pathway to provide a high flux of geranyl pyrophosphate and introduced a heterologous, multi-organism-derived hexanoyl-CoA biosynthetic pathway⁵. We also introduced the *Cannabis* genes that encode the enzymes involved in the biosynthesis of olivetolic acid⁶, as well as the gene for a previously undiscovered enzyme with geranylpyrophosphate:olivetolate geranyltransferase activity and the genes for corresponding cannabinoid synthases^{7,8}. Furthermore, we established a biosynthetic approach that harnessed the promiscuity of several pathway genes to produce cannabinoid analogues. Feeding different fatty acids to our engineered strains yielded cannabinoid analogues with modifications in the part of the molecule that is known to alter receptor binding affinity and potency⁹. We also demonstrated that our biological system could be complemented by simple synthetic chemistry to further expand the accessible chemical space. Our work presents a platform for the production of natural and unnatural cannabinoids that will allow for more rigorous study of these compounds and could be used in the development of treatments for a variety of human health problems.

We began our generation of the cannabinoid-producing yeast by focusing first on the production of olivetolic acid (Fig. 1), an initial intermediate in the cannabinoid biosynthetic pathway. Two *Cannabis* enzymes, a tetraketide synthase (*C. sativa* TKS; CsTKS)¹⁰ and an olivetolic acid cyclase (CsOAC)⁶, have been reported to produce olivetolic acid from hexanoyl-CoA and malonyl-CoA. To produce olivetolic acid in yeast, we introduced a CsTKS and CsOAC expression cassette into *S. cerevisiae* to generate strain yCAN01 (Extended Data Table 1). The strain produced 0.2 mg l⁻¹ olivetolic acid from galactose (Fig. 2a), consistent with the fact that *S. cerevisiae* maintains low intracellular levels

of hexanoyl-CoA¹¹. To increase the supply of hexanoyl-CoA, we fed the yCAN01 strain with 1 mM hexanoic acid, which can be converted to hexanoyl-CoA by an endogenous acyl activating enzyme (AAE), and observed a sixfold increase in olivetolic acid production (1.3 mg l⁻¹). A known byproduct of TKS, hexanoyl triacetic acid lactone (HTAL)⁶, was also detected (Extended Data Fig. 1).

To optimize the conversion of hexanoic acid to hexanoyl-CoA, we introduced CsAAE1, an AAE from *Cannabis* that is thought to catalyse this step in planta, into yCAN01¹². When fed with 1 mM hexanoic acid, the resulting strain (yCAN02) showed a twofold increase in olivetolic acid titre (3.0 mg l⁻¹) compared with yCAN01 (Fig. 2a). To produce hexanoyl-CoA from galactose and complete the olivetolic acid pathway, we introduced a previously reported hexanoyl-CoA pathway into yCAN01⁵. The resulting strain (yCAN03) produced 1.6 mg l⁻¹ olivetolic acid (Fig. 2a).

Cannabigerolic acid (CBGA)—the precursor to Δ^9 -tetrahydrocannabinolic acid (THCA), cannabidiolic acid (CBDA) and numerous other cannabinoids—is produced from olivetolic acid and the mevalonate-pathway intermediate geranyl pyrophosphate (GPP) by a geranylpyrophosphate:olivetolate geranyltransferase (GOT). GOT activity was detected in *Cannabis* extracts two decades ago¹³, and a *Cannabis* GOT (CsPT1) was patented ten years later¹⁴. To test CsPT1 in vivo, we constructed a GPP-overproducing strain (yCAN10) with an upregulated mevalonate pathway¹⁵ and a mutant version of the endogenous farnesyl pyrophosphate synthase ERG20 (ERG20(F69W/N127W)) that preferentially produces GPP over FPP¹⁶. However, we were unable to observe any GOT activity when we expressed CsPT1, or any truncations of it, in yCAN10.

To identify an enzyme with GOT activity that would function in yeast, we searched for candidate prenyltransferase enzymes from *Cannabis* and other organisms. These included NphB, a soluble prenyltransferase from *Streptomyces* sp.¹⁷ that displayed GOT activity in vitro¹⁸, and HIPT1L and HIPT2, two prenyltransferases involved in bitter acid biosynthesis in *Humulus lupulus* (a close relative of *Cannabis*)¹⁹. In addition, we mined published *Cannabis* transcriptomes²⁰ (<http://medicinalplantgenomics.msu.edu>) for GOT candidates. We searched full-length transcripts using the Basic Local Alignment Search Tool (BLAST) against CsPT1, HIPT1L and HIPT2, and on the basis of this chose six *Cannabis* enzymes (CsPT2–CsPT7).

For functional expression of the six *Cannabis* and two *H. lupulus* prenyltransferases in yeast, we removed predicted N-terminal plastid-targeting sequences^{21,22}, resulting in CsPT2-T–CsPT7-T and HIPT1L-T and HIPT2-T, respectively. Each GOT candidate was then introduced into yCAN10, and the resulting strains (yCAN12–yCAN20)

¹California Institute of Quantitative Biosciences (QB3), University of California, Berkeley, CA, USA. ²Department of Biosystems Science and Engineering, ETH Zurich, Basel, Switzerland. ³Biological Systems and Engineering Division, Lawrence Berkeley National Laboratory, Berkeley, CA, USA. ⁴Department of Chemical and Biomolecular Engineering, University of California, Berkeley, CA, USA. ⁵Department of Bioengineering, University of California, Berkeley, CA, USA. ⁶Key Laboratory of Industrial Biotechnology, Ministry of Education, Jiangnan University, Wuxi, China. ⁷Department of Chemistry, University of California, Berkeley, CA, USA. ⁸Department of Plant and Microbial Biology, University of California, Berkeley, CA, USA. ⁹Biotechnology and Bioengineering Department, Sandia National Laboratories, Livermore, CA, USA. ¹⁰Novo Nordisk Foundation Center for Biosustainability, Technical University of Denmark, Lyngby, Denmark. ¹¹Center for Synthetic Biochemistry, Institute of Synthetic Biology, Shenzhen Institutes of Advanced Technologies, Shenzhen, China. ¹²Present address: Demetrix, Inc., Emeryville, CA, USA. ¹³Present address: Berkeley Brewing Science, Inc., Berkeley, CA, USA. ¹⁴Present address: Genomatica, Inc., San Diego, CA, USA. ¹⁵These authors contributed equally: Xiaozhou Luo, Michael A. Reiter. *e-mail: keasling@berkeley.edu

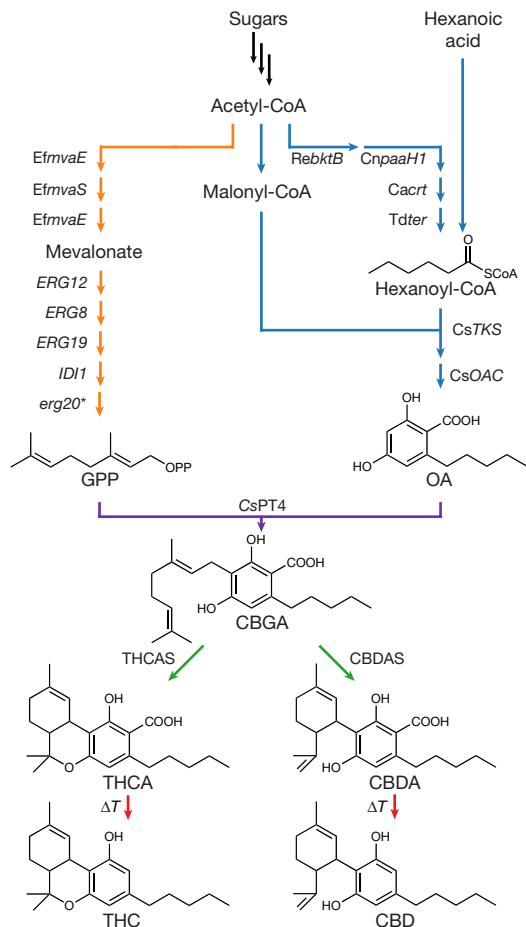


Fig. 1 | Engineered biosynthetic pathway for synthesis of cannabinoids in *S. cerevisiae*. GPP was produced by introducing the *Enterococcus faecalis* genes *EfmvaE* and *EfmvaS*, and by overexpressing the native mevalonate pathway genes (*ERG12*, *ERG8*, *ERG19* and *IDI1*) and a mutated *ERG20*^{F96W/N127W} gene (*erg20**). Hexanoyl-CoA was produced via a heterologous biosynthetic pathway, using genes from *Ralstonia eutropha* (*RebktB*), *Cupriavidus necator* (*CnpaaH1*), *Clostridium acetobutylicum* (*Cacrt*) and *Treponema denticola* (*Tdter*), or by feeding hexanoic acid as a substrate for AAE (encoded by *CsAAE1* from *Cannabis*). Expression of the genes encoding *CsTKS* and *CsOAC* produced olivetolic acid, which was prenylated by *CsPT4*. The resulting CBGA was transformed to the cannabinoid acids THCA and CBDA using the cannabinoid synthases THCAS and CBDAS. After exposure to heat (ΔT), THCA and CBDA decarboxylate to THC and CBD, respectively.

were cultured in 1 mM olivetolic acid and assayed for CBGA production using liquid chromatography–mass spectrometry (LC–MS). Of the nine GOT candidates tested, only the strain that expresses *CsPT4-T* (*yCAN14*) produced detectable amounts of CBGA (136 mg l⁻¹, Fig. 2b). *CsPT4-T* is predicted to have eight transmembrane helices²³, and it localized to the purified microsomal fraction when heterologously expressed in yeast (Extended Data Fig. 2a). Subsequent *in vitro* experiments with purified microsomal fractions from *yCAN14* confirmed GOT activity and revealed a Michaelis–Menten constant of $K_m(\text{olivetolic acid}) = 6.73 \pm 0.26 \mu\text{M}$ (mean \pm s.d.), as well as non-Michaelis–Menten type behaviour for GPP (Extended Data Fig. 2b, c). Similar *in vitro* assays with other GOT candidates showed low and non-specific (*NphB*, *HIPT1L-T* and *HIPT2-T*) or no (*CsPT1-T*) GOT activity (Extended Data Fig. 3). *CsPT4* clusters with other *Cannabis* prenyltransferases in a phylogenetic tree (Extended Data Fig. 2d) but shares only 62% homology with *CsPT1*.

Next, we set out to produce CBGA from simpler precursors such as hexanoic acid or galactose by reconstituting the olivetolic acid biosynthesis module in strain *yCAN14*. Adding *CsTKS*, *CsOAC* and *CsAAE1*

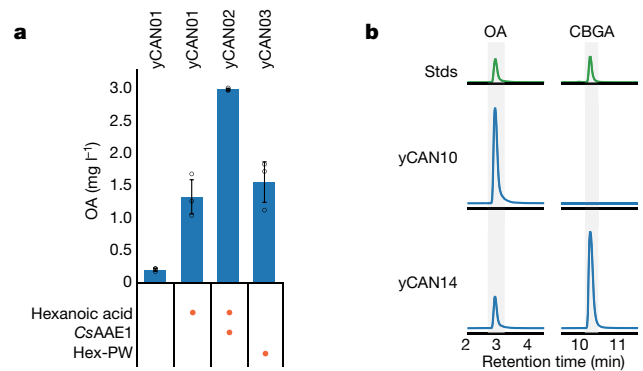


Fig. 2 | Pathway engineering for olivetolic acid and CBGA. **a**, Olivetolic acid (OA) production in *yCAN01*, *yCAN02* and *yCAN03* strains from galactose or 1 mM hexanoic acid 48 h after induction (hex-PW, hexanoyl-CoA pathway). Data are mean \pm s.d.; $n = 3$ biologically independent samples. **b**, *In vivo* production of CBGA by feeding olivetolic acid. *yCAN10* and *yCAN14* cultures fed 1 mM olivetolic acid were sampled 24 h after induction. Extracts were analysed by LC–MS and signals were compared to genuine standards (stds). All LC–MS chromatograms were selected for the theoretical m/z values of the respective compounds of interest (Extended Data Table 3).

generated strain *yCAN31*, which produced 7.2 mg l⁻¹ CBGA from 1 mM hexanoic acid. Finally, integrating the hexanoyl-CoA pathway into *yCAN31* resulted in strain *yCAN32*, which produced 1.4 mg l⁻¹ CBGA from galactose.

Synthases for the conversion of CBGA to THCA²⁴ (THCAS) and to CBDA²⁵ (CBDAS) in *Cannabis* were identified two decades ago. THCAS has been functionally expressed and assayed *in vitro* in insect cells²⁶ and yeast^{7,18}. However, *in vivo* production of THCA, CBDA or other cannabinoids that are produced late in the cannabinoid pathway (after CBGA) from sugars has not been shown so far. To complete the cannabinoid biosynthesis pathway, we replaced the N-terminal secretory signal peptides²⁷ of THCAS and CBDAS with a vacuolar localization tag to enable their functional expression^{7,18}, and integrated the resulting sequences into *yCAN31* (resulting in *yCAN40* and *yCAN41*) and *yCAN32* (resulting in *yCAN42* and *yCAN43*). Culturing *yCAN40* or *yCAN41* with 1 mM hexanoic acid yielded 1.1 mg l⁻¹ THCA or 4.3 $\mu\text{g l}^{-1}$ CBDA, respectively, and culturing *yCAN42* or *yCAN43* with galactose yielded 2.3 mg l⁻¹ THCA or 4.2 $\mu\text{g l}^{-1}$ CBDA, respectively (Fig. 3a, b).

In addition to cannabinoids derived from olivetolic acid, *Cannabis* produces compounds originating from divarinolic acid, an olivetolic acid analogue in which the C3-pentyl side chain is substituted with a propyl moiety²⁸, which suggests promiscuity of at least some pathway enzymes. The precursor to divarinolic acid is presumably butanoyl-CoA¹⁰, an intermediate in our engineered hexanoyl-CoA pathway. Indeed, *yCAN42* and *yCAN43* produced 1.2 mg l⁻¹ Δ^9 -tetrahydrocannabivarinic acid (THCVA) and 6.0 $\mu\text{g l}^{-1}$ cannabidivarinic acid (CBDVA), respectively (Fig. 3). Furthermore, peaks corresponding to the predicted m/z values of divarinolic acid and the propyl-variant of CBGA, cannabigerovarinic acid (CBGVA), were also detected (Extended Data Fig. 4).

To improve titres from galactose and identify bottlenecks in the pathway, we introduced additional single copies of *CsTKS*–*CsOAC*, *CsPT4-T* or *THCAS* to *yCAN42* (resulting in *yCAN50*, *yCAN51* and *yCAN52*). We assumed that the GPP supply was sufficient, as olivetolic acid was converted efficiently to CBGA when the growth medium was supplemented with 1 mM olivetolic acid (Fig. 2b). When assayed, we observed a threefold increase in the concentration of intermediates and in the concentration of THCA and THCVA for the *CsTKS*–*CsOAC* overexpression strain (*yCAN50*) in comparison with the parent strain (*yCAN42*) (Fig. 3c). Cannabinoid production remained essentially unchanged for *yCAN51* and *yCAN52* relative to the parent strain, suggesting that the pathway was primarily limited

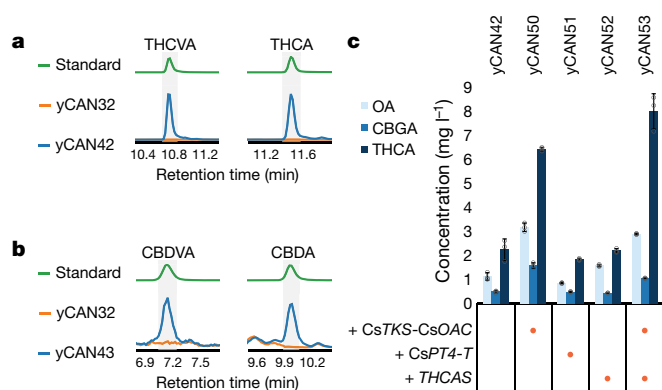


Fig. 3 | In vivo production of THCVA, THCA, CBDVA and CBDA. All LC–MS chromatograms were selected for the theoretical m/z values of the respective compounds of interest (Extended Data Table 3). **a**, yCAN42 produced THCVA and THCA from galactose. **b**, yCAN43 produced CBDVA and CBDA from galactose. yCAN32 is the parent strain to yCAN42 and yCAN43 and is shown as the negative control. Samples were extracted 96 h after induction. **c**, Introduction of additional copies of pathway enzymes revealed bottlenecks and increased overall production. Data are mean \pm s.d.; $n = 3$ biologically independent samples.

by the conversion of malonyl-CoA and hexanoyl-CoA to olivetolic acid. Finally, we added one more copy of *THCAS* to yCAN50 (resulting in yCAN53) to enhance conversion of the now-larger CBGA pool. yCAN53 yielded increased titres of THCA (8.0 mg l^{-1}) and THCVA (4.8 mg l^{-1}) (Fig. 3c). Further gains could be achieved in the future by fine-tuning TKS-OAC and *THCAS* expression, or by upregulating upstream precursor production.

In addition to their natural counterparts, unnatural cannabinoid analogues are actively being investigated for their potentially improved medicinal properties. One of the major pharmacophores of interest in this search is the C3 side chain of Δ^9 -tetrahydrocannabinol (THC) because its length, size, structure and chemistry have been shown to modulate cannabinoid receptor (CB1 or CB2) properties including binding affinity, selectivity and potency⁹. We set out to establish a biosynthetic approach for the production of this class of cannabinoid analogues from different fatty acids, hypothesizing that the observed promiscuity of our pathway towards butanoyl-CoA would translate to other precursors (Fig. 4a). To probe the analogue production capability of our engineered strains, we fed yCAN31 an array of 19 different fatty acids of various chain lengths, branching and degrees of saturation (Extended Data Table 2). LC–MS analysis revealed the production of olivetolic acid and CBGA analogues from pentanoic acid (II), heptanoic acid (III), 4-methylhexanoic acid (IV), 5-hexenoic acid (V) and 6-heptynoic acid (VI) (Fig. 4b, Extended Data Fig. 5). Subsequent supplementation of yCAN40 with this subset of fatty acids yielded the respective THCA analogues (Fig. 4b). Moreover, the functionalization of the pharmacophore with an alkene (V) or alkyne (VI) terminal group enabled simple post-fermentation modification and thus the construction of side chains that were intractable to direct incorporation. As proof of concept, we performed copper-catalysed azide–alkyne cycloaddition on the 6-heptynoic acid CBGA (6hCBGA) and THCA (6hTHCA) analogues with an azide–PEG3–biotin conjugate. The corresponding products were detected by LC–MS (Fig. 4c), demonstrating that the accessible chemical space of our process can be further expanded. Our results illustrate a route towards the production of cannabinoid analogues with tailored C3 side chains.

In summary, we engineered yeast strains capable of producing the major cannabinoids found in *Cannabis* from galactose. Pending the identification of novel cannabinoid synthases, we expect to be able to produce a large fraction of this class of natural molecules. Furthermore, we expanded the chemical space of cannabinoids by establishing and then harnessing the intrinsic promiscuity of the cannabinoid pathway

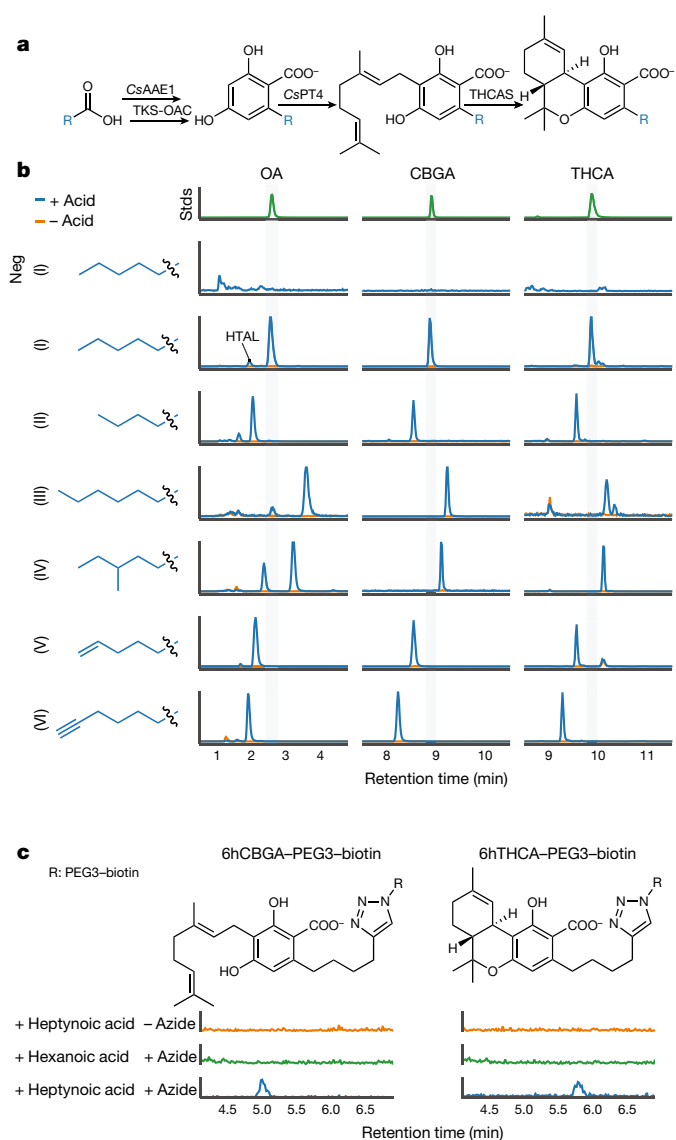


Fig. 4 | Production of cannabinoid analogues and post-fermentation derivatization. **a**, Pathway assayed for incorporation of different fatty acid precursors into the corresponding olivetolic acid, CBGA and THCA analogues. **b**, Ion-extracted chromatograms of the engineered strains (yCAN31 for olivetolic acid and CBGA columns; yCAN40 for THCA column) grown in the presence of hexanoic acid (I), pentanoic acid (II), heptanoic acid (III), 4-methylhexanoic acid (IV), 5-hexenoic acid (V) or 6-heptynoic acid (VI) showed the production of olivetolic acid, CBGA and THCA, or their respective analogues. Detected analogue peaks shifted in retention time depending on their change in hydrophobicity relative to the olivetolic acid, CBGA and THCA standards. Feeding 1 mM hexanoic acid to the parent strain of yCAN31 (yCAN30) did not yield olivetolic acid or CBGA, and correspondingly, no THCA was produced by the parent strain of yCAN40 (yCAN31) (neg (I)). All LC–MS chromatograms were selected for the theoretical m/z values of the respective compounds of interest (Extended Data Fig. 6). Additional peaks were observed for the byproducts (HTAL and its analogues) of the *Cs*TKS- and *Cs*OAC-catalysed reaction step. **c**, yCAN40 was cultured in the presence of 1 mM 6-heptynoic acid or 1 mM hexanoic acid. The corresponding extracts were reacted with an azide–PEG3–biotin conjugate (+ azide) or water (– azide) under copper-catalysed azide–alkyne cycloaddition conditions. The production of 6hCBGA–PEG3–biotin and 6hTHCA–PEG3–biotin was observed only in the presence of both 6-heptynoic acid and azide–PEG3–biotin conjugate.

to produce unnatural cannabinoids, including molecules with side groups amenable to further chemical derivatization. This work lays the foundation for the large-scale fermentation of cannabinoids, independent of *Cannabis* cultivation, which will enable the pharmacological

study of these highly promising compounds and could ultimately lead to new and better medicines.

Online content

Any methods, additional references, Nature Research reporting summaries, source data, statements of data availability and associated accession codes are available at <https://doi.org/10.1038/s41586-019-0978-9>.

Received: 6 May 2018; Accepted: 24 January 2019;

Published online: 27 February 2019

- Pacher, P., B tkai, S. & Kunos, G. The endocannabinoid system as an emerging target of pharmacotherapy. *Pharmacol. Rev.* **58**, 389–462 (2006).
- Whiting, P. F. et al. Cannabinoids for medical use: a systematic review and meta-analysis. *J. Am. Med. Assoc.* **313**, 2456–2473 (2015).
- Hazekamp, A., Ware, M. A., M ller-Vahl, K. R., Abrams, D. & Grotenhermen, F. The medicinal use of cannabis and cannabinoids—an international cross-sectional survey on administration forms. *J. Psychoactive Drugs* **45**, 199–210 (2013).
- ElSolhly, M. A. & Gul, W. in *Handbook of Cannabis* (ed. Pertwee, R. G.) Ch. 1 (Oxford Univ. Press, Oxford, 2014).
- Dekishima, Y., Lan, E. I., Shen, C. R., Cho, K. M. & Liao, J. C. Extending carbon chain length of 1-butanol pathway for 1-hexanol synthesis from glucose by engineered *Escherichia coli*. *J. Am. Chem. Soc.* **133**, 11399–11401 (2011).
- Gagne, S. J. et al. Identification of olivetolic acid cyclase from *Cannabis sativa* reveals a unique catalytic route to plant polyketides. *Proc. Natl Acad. Sci. USA* **109**, 12811–12816 (2012).
- Zirpel, B., Stehle, F. & Kayser, O. Production of Δ^9 -tetrahydrocannabinolic acid from cannabigerolic acid by whole cells of *Pichia (Komagataella) pastoris* expressing Δ^9 -tetrahydrocannabinolic acid synthase from *Cannabis sativa* L. *Biotechnol. Lett.* **37**, 1869–1875 (2015).
- Taura, F. et al. Cannabidiolic-acid synthase, the chemotype-determining enzyme in the fiber-type *Cannabis sativa*. *FEBS Lett.* **581**, 2929–2934 (2007).
- Razdan, R. K. in *The Cannabinoid Receptors* (ed. Reggio, P. H.) 3–19 (Humana Press, New York, 2009).
- Taura, F. et al. Characterization of olivetol synthase, a polyketide synthase putatively involved in cannabinoid biosynthetic pathway. *FEBS Lett.* **583**, 2061–2066 (2009).
- Saerens, S. M. G., Delvaux, F. R., Verstrepen, K. J. & Thevelein, J. M. Production and biological function of volatile esters in *Saccharomyces cerevisiae*. *Microb. Biotechnol.* **3**, 165–177 (2010).
- Stout, J. M., Boubakir, Z., Ambrose, S. J., Purves, R. W. & Page, J. E. The hexanoyl-CoA precursor for cannabinoid biosynthesis is formed by an acyl-activating enzyme in *Cannabis sativa* trichomes. *Plant J.* **71**, 353–365 (2012).
- Fellermeier, M. & Zenk, M. H. Prenylation of olivetolate by a hemp transferase yields cannabigerolic acid, the precursor of tetrahydrocannabinol. *FEBS Lett.* **427**, 283–285 (1998).
- Page, J. E. & Boubakir, Z. Aromatic prenyltransferase from Cannabis. US patent 2012/0144523 A1 (2012).
- Reider Apel, A. et al. A Cas9-based toolkit to program gene expression in *Saccharomyces cerevisiae*. *Nucleic Acids Res.* **45**, 496–508 (2017).
- Ignea, C., Pontini, M., Maffei, M. E., Makris, A. M. & Kampranis, S. C. Engineering monoterpene production in yeast using a synthetic dominant negative geranyl diphosphate synthase. *ACS Synth. Biol.* **3**, 298–306 (2014).
- Kuzuyama, T., Noel, J. P. & Richard, S. B. Structural basis for the promiscuous biosynthetic prenylation of aromatic natural products. *Nature* **435**, 983–987 (2005).
- Zirpel, B., Degenhardt, F., Martin, C., Kayser, O. & Stehle, F. Engineering yeasts as platform organisms for cannabinoid biosynthesis. *J. Biotechnol.* **259**, 204–212 (2017).
- Li, H. et al. A heteromeric membrane-bound prenyltransferase complex from hop catalyzes three sequential aromatic prenylations in the bitter acid pathway. *Plant Physiol.* **167**, 650–659 (2015).
- van Bakel, H. et al. The draft genome and transcriptome of *Cannabis sativa*. *Genome Biol.* **12**, R102 (2011).
- Emanuelsson, O., Brunak, S., von Heijne, G. & Nielsen, H. Locating proteins in the cell using TargetP, SignalP and related tools. *Nat. Protocols* **2**, 953–971 (2007).
- Emanuelsson, O., Nielsen, H. & von Heijne, G. ChloroP, a neural network-based method for predicting chloroplast transit peptides and their cleavage sites. *Protein Sci.* **8**, 978–984 (1999).
- Krogh, A., Larsson, B., von Heijne, G. & Sonnhammer, E. L. Predicting transmembrane protein topology with a hidden Markov model: application to complete genomes. *J. Mol. Biol.* **305**, 567–580 (2001).
- Taura, F., Morimoto, S., Shoyama, Y. & Mechoulam, R. First direct evidence for the mechanism of Δ^1 -tetrahydrocannabinolic acid biosynthesis. *J. Am. Chem. Soc.* **117**, 9766–9767 (1995).
- Taura, F., Morimoto, S. & Shoyama, Y. Purification and characterization of cannabidiolic-acid synthase from *Cannabis sativa* L. Biochemical analysis of a novel enzyme that catalyzes the oxidocyclization of cannabigerolic acid to cannabidiolic acid. *J. Biol. Chem.* **271**, 17411–17416 (1996).
- Sirikantaramas, S. et al. The gene controlling marijuana psychoactivity: molecular cloning and heterologous expression of Δ^1 -tetrahydrocannabinolic acid synthase from *Cannabis sativa* L. *J. Biol. Chem.* **279**, 39767–39774 (2004).
- Sirikantaramas, S. et al. Tetrahydrocannabinolic acid synthase, the enzyme controlling marijuana psychoactivity, is secreted into the storage cavity of the glandular trichomes. *Plant Cell Physiol.* **46**, 1578–1582 (2005).
- de Meijer, E. P. M. & Hammond, K. M. The inheritance of chemical phenotype in *Cannabis sativa* L. (V): regulation of the propyl-/pentyl cannabinoid ratio, completion of a genetic model. *Euphytica* **210**, 291–307 (2016).

Acknowledgements We thank G. Wang for the pESC-HIPT1L, pESC-HIPT2 and pESC-HIPT1L-HIPT2 plasmids; T. Laursen for help with microsomal preparations; C. Eiben, T. de Rond, R. Lee, C. Joshua, T. Tian, P. Shih, M. Wehrs and A. Zargar for discussion; L. Davis, T. Lease and B. Sandmann for help on controlled substance regulatory processes; E. Mendez and T. Dunn for strain archiving support; and O. Diaz and E. Coyne for administrative support. This work was supported by a National Science Foundation grant 1330914 to J.D.K.; Y.Z. acknowledges support from a Jiangnan University Graduate Scholarship for Oversea Study.

Reviewer information *Nature* thanks James Collins, Bernd Lange, Jonathan Mielenz and Jordan Zager for their contribution to the peer review of this work.

Author contributions X.L., M.A.R., L.d'E., J.W., C.M.D., A.L. and J.D.K. conceived the study. X.L., M.A.R., L.d'E., J.W., A.L., C.M.D., Y.Z., A.T.G., S.H., W.L., H.L., C.Y., J.S. and I.D. constructed the plasmids and yeast strains and performed microbiological manipulations and extractions. X.L., M.A.R. and K.D. performed synthetic organic chemistry. X.L., M.A.R., V.T.B., G.W., E.E.K.B., Y.C. and C.J.P. performed mass spectrometry. M.A.R. performed enzyme kinetic study and in vitro studies. J.W. and L.d'E. identified and functionally expressed CsPT4. X.L., M.A.R., A.L. and J.W. performed bioinformatic analysis. M.A.R. developed data analysis scripts. All authors contributed to the manuscript.

Competing interests X.L., M.A.R., L.d'E., J.W., C.M.D., E.E.K.B. and C.J.P. have a financial interest in Demetrix. J.D.K. has a financial interest in Amyris, Lygos, Demetrix, Constructive Biology, Napigen, and Maple Bio.

Additional information

Extended data is available for this paper at <https://doi.org/10.1038/s41586-019-0978-9>.

Supplementary information is available for this paper at <https://doi.org/10.1038/s41586-019-0978-9>.

Reprints and permissions information is available at <http://www.nature.com/reprints>.

Correspondence and requests for materials should be addressed to J.D.K.

Publisher's note: Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

  The Author(s), under exclusive licence to Springer Nature Limited 2019

METHODS

General. A detailed description of the methods used for the generation of *S. cerevisiae* yCAN strains is provided in the Supplementary Information.

Cannabinoid standards (CBGA, THCA, CBDA, CBGVA, THCVA and CBDVA) were purchased from Cerilliant. Olivetolic acid was purchased from A1 Biochem Labs. All other chemicals were obtained from Fisher Scientific and Sigma-Aldrich. Oligonucleotides and codon-optimized gBlock gene fragments were obtained from Integrated DNA Technologies. Yeast culture media were purchased from Becton, Dickinson and Company, and all agar plates were obtained from Teknova. The project was registered with, and under the regulation of, the University of California, Berkeley and Lawrence Berkeley National Laboratory controlled substances programmes, the Research Advisory Panel of California and the US Drug Enforcement Administration (DEA). All researchers who had direct access to controlled substances, or yeast strains capable of making such compounds, passed background screening. All work was conducted in a DEA-certified laboratory, and all protocols were reviewed by the authorities, including the production log, product destruction and disposal. To prevent illicit use, standards, strains and samples were stored in a TL-15 safe with group 1R lock.

In vivo production, purification and LC-TOF-MS/LC-QTOF-MS analysis of cannabinoids. Strains were pre-grown in yeast extract peptone dextrose (YPD) medium overnight and then back-diluted to $OD_{600\text{ nm}} = 0.2$ into yeast extract peptone galactose (YPG). Medium was supplemented with 1 mM olivetolic acid or corresponding fatty acid as indicated (Extended Data Table 3). Strains were incubated for 24 h (yCAN10–yCAN20), 48 h (yCAN01–yCAN03, yCAN30–yCAN33) or 96 h (yCAN40–yCAN53) in 24-deep-well plates (800 r.p.m.) at 30°C while supplementing with 2% (w/v) galactose every 24 h. Subsequently, samples were treated with 2 U/OD zymolyase (Zymo Research) (2 h, 30°C, 800 r.p.m.) (step was skipped for extraction of analogues), followed by ethyl acetate/formic acid (0.05% (v/v)) extraction in a 2:1 ratio and bead-beating (30 s^{-1} , 3 min). Organic and inorganic layers were separated by centrifugation at 12,000g for 1 min. Samples were extracted three times. The combined organic layers were evaporated in a vacuum oven (50°C) and the remainders were resuspended in acetonitrile/H₂O/formic acid (80%/20%/0.05% (v/v/v)). Finally, samples were filtered with Ultrafree-MC columns (0.22 µm pore size, polyvinylidene difluoride (PVDF) membrane).

Products were analysed using either high-performance liquid chromatography with UV detection (HPLC-UV, Agilent 1200 series) or LC-MS (Agilent 6210 for time-of-flight (TOF) and Agilent 6545 for quadrupole time-of-flight (QTOF); Agilent Technologies) equipped with a reverse phase C18 column (Kinetex 2.6 µm, 100×2.1 mm, XB-C18, Phenomenex). The mobile phase was composed of 0.05% (v/v) formic acid in water (solvent A) and 0.05% (v/v) formic acid in acetonitrile (solvent B). Cannabinoids were separated via gradient elution as follows: linearly increased from 45% B to 62.5% B in 3.5 min, held at 62.5% B for 4.5 min, increased from 62.5% B to 97% B in 0.5 min, held at 97% B for 4 min, decreased from 97% B to 45% B in 0.2 min, and held at 45% B for 2.8 min. The flow rate was held at 0.2 ml min^{-1} for 12.5 min, increased from 0.2 ml min^{-1} to 0.4 ml min^{-1} in 0.2 min, and held at 0.4 ml min^{-1} for 2.8 min. The total liquid chromatography run time was 15.5 min. Cannabinoid analogues from different acids were separated via gradient elution as follows: linearly increased from 45% B to 73% B in 5.6 min, increased from 73% B to 97% B in 0.6 min, held at 97% B for 5.1 min, decreased from 97% B to 45% B in 1.4 min, and held at 45% B for 2.8 min. The flow rate was held at 0.2 ml min^{-1} for 5.6 min, increased from 0.2 ml min^{-1} to 0.3 ml min^{-1} in 0.6 min, and held at 0.3 ml min^{-1} for 9.3 min. The total liquid chromatography run time was 15.5 min. Sample injection volumes of 1 µl and 2 µl were used for QTOF-MS and TOF-MS, respectively. The sample tray and column compartment were set to 6°C and 40°C, respectively. For HPLC-UV, cannabinoids were detected by diode array detection at 270 nm. For TOF-MS, electrospray ionization was conducted in the negative ion mode and a capillary voltage of 3,500 V was used. The fragmentor, skimmer and OCT 1 RF Vpp voltages were set to 150 V, 50 V and 170 V, respectively. Drying and nebulizing gases were set to 11 l min^{-1} and 206.8 kPa (30 lb inch²), respectively, and a drying-gas temperature of 330°C was used throughout. For QTOF-MS, electrospray ionization was conducted via the Agilent Jet Stream thermal gradient focusing technology, in which the sheath gas flow rate and temperature were set to 12.1 ml min^{-1} and 350°C, respectively. Drying and nebulizing gases were set to 10.1 ml min^{-1} and 172.4 kPa (25 lb inch²), respectively, and a drying-gas temperature of 300°C was used throughout. All other conditions were the same as that of TOF-MS. Data files were processed by Agilent MassHunter Qualitative Analysis software via extraction of the corresponding $[M-H]^{-}$ ion counts (extraction window ± 20 p.p.m.; mass accuracy < 5 p.p.m.) and analysed by Openchrom 1.2.0²⁹, custom Python 3.6 scripts using Scipy (v.1.0.0) (<http://www.scipy.org/>) and Pandas (v.0.22.0)³⁰ packages. All plots were generated with Plotly (Plotly Technologies).

In vitro enzyme assays of CsPT4-T. yCAN10 and yCAN14 were inoculated to $OD_{600\text{ nm}} = 0.2$ in 200 ml YPG medium and incubated for 20 h in baffled shake flasks (200 r.p.m.). Collected cells were resuspended in 50 ml buffer (50 mM Tris-HCl, 1 mM EDTA, 0.1 M KCl, pH 7.4, 125 units benzonase), then lysed (60 bar, 10 min; Emulsiflex C3, Avestin). Cell debris was removed by centrifugation (10,000g, 10 min, 4°C). Subsequently, the supernatant was subjected to ultracentrifugation (150,000g, 1 h, 4°C; Beckman Coulter L-90K, TI-70). The resulting membrane fractions of yCAN10 and yCAN14 were resuspended in 3.3 ml buffer (10 mM Tris-HCl, 10 mM MgCl₂, pH 8.0, 10% glycerol) and homogenized with a tissue grinder. Then 4% (v/v) of the respective membrane preparations were dissolved in reaction buffer (50 mM Tris-HCl, 10 mM MgCl₂, pH 8.5) and substrate was added (500 µM olivetolic acid, 500 µM GPP) to a total volume of 50 µl. Samples were incubated for 1 h at 30°C. Assays were extracted three times by adding two reaction volumes of ethyl acetate followed by vortexing and centrifugation. The organic layer was evaporated as described above for 30 min, resuspended in acetonitrile/H₂O/formic acid and filtered.

CsPT4-T kinetics. To determine the kinetics of CsPT4, 4% (v/v) of yCAN14 microsomal preparation was incubated in 50 µl reaction buffer for 10 min at 30°C with a range of olivetolic acid concentrations (0.25 µM–0.56 mM) and constant GPP (1.67 mM), or different GPP concentrations (0.25 µM–1.67 mM) and constant olivetolic acid (1.67 mM). Assays were quenched with 50 µl ice-cold 32% HCl and extracted with 100 µl ice-cold ethyl acetate. Product formation was detected by HPLC-UV as described above. Kinetic parameters were determined by nonlinear regression using Scipy (<http://www.scipy.org/>). Kinetic studies were carried out in triplicate or quadruplicate.

Click chemistry. yCAN40 was pre-grown in YPD, subsequently diluted to $OD_{600\text{ nm}} = 0.2$ in YPG with 1 mM 6-heptynoic acid or 1 mM hexanoic acid, cultured for 96 h, and extracted as described above. After evaporation, samples were resuspended in 25 µl dimethyl sulfoxide (DMSO)/H₂O (55%:45% (v/v)) and filtered. An aliquot of 8 µl of freshly prepared copper-catalysed azide-alkyne cycloaddition master mix (1.25 mM Tris[(1-benzyl-1H-1,2,3-triazol-4-yl)methyl]amine, 1.25 mM CuSO₄, 2.5 mM ascorbic acid, 2.5 mM azide-PEG3-biotin conjugate in DMSO/H₂O (38.75%/61.25% (v/v)) was added to 20 µl of sample and incubated for 24 h at 37°C.

Identification of candidate Cannabis GOTs. Previously reported assembled transcripts and expression abundance estimations were retrieved from the *Cannabis* genome browser gateway (<http://genome.ccb.utoronto.ca/cgi-bin/hgGateway>) and the medicinal plant genomics resource (<http://medicinalplantgenomics.msu.edu>). To select prenyltransferases for functional testing, published transcriptomes of *Cannabis* were mined using a BLAST search against CsPT1, HIPT1L and HIPT2. **Phylogenetic analysis.** Phylogenetic analysis was performed using the predicted amino acid sequences of *Cannabis* candidate prenyltransferases and related prenyltransferases (Extended Data Table 4). Sequence alignments were generated using MUSCLE with default parameters³¹. Maximum likelihood analysis was conducted with the IQ-TREE web server (<http://iqtree.cibiv.univie.ac.at/>) and default parameters³². The resulting tree was visualized using ETE3³³.

Statistics and reproducibility. No statistical methods were used to predetermine sample size. The experiments were not randomized and the investigators were not blinded to allocation during experiments and outcome assessment. If not indicated otherwise, all experiments were conducted with $n = 3$ biological replicates. In cases for which representative traces are shown in the display items, all other traces were similar. For bar charts, the mean of the individual measurements was used as the measure of centre and the s.d. was used for error bars.

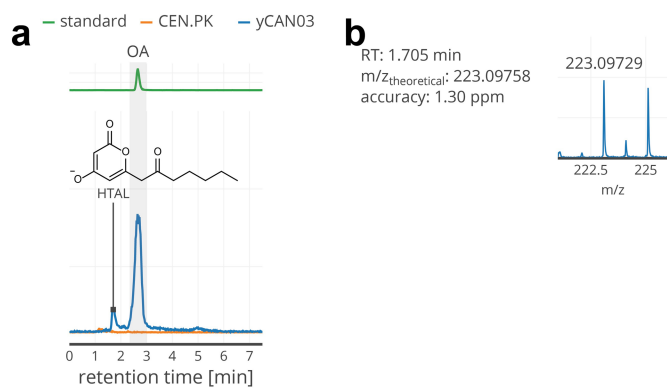
Reporting summary. Further information on research design is available in the Nature Research Reporting Summary linked to this paper.

Data availability

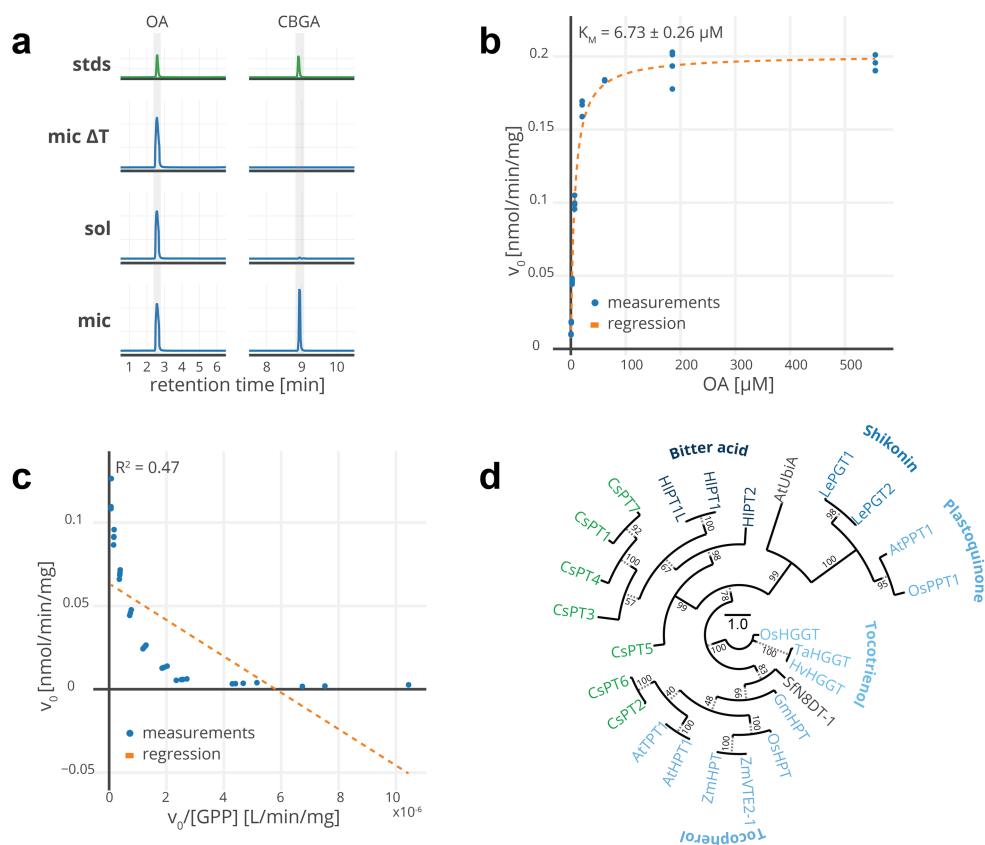
Nucleotide sequence data of *Cannabis* candidate prenyltransferases are available in the third-party annotation section of the DDBJ/ENA/GenBank databases (Extended Data Table 4). Strains and plasmids developed for this study (Extended Data Table 1), along with annotated sequences, have been deposited in the Synthetic Biology Engineering Research Center (Synberc) Registry (<https://synberc-registry.jbei.org/>) and are physically available from the authors upon reasonable request. Strains producing controlled substances or its direct precursors can only be provided to laboratories and institutions with appropriate approvals and licenses (for example, DEA permits). Custom Python 3.6 scripts for data analysis are available from the authors upon reasonable request.

- Wenig, P. & Odermatt, J. OpenChrom: a cross-platform open source software for the mass spectrometric analysis of chromatographic data. *BMC Bioinformatics* **11**, 405 (2010).
- McKinney, W. Data structures for statistical computing in Python. In *Proc. 9th Python in Science Conference* (eds van der Walt, S. & Millman, J.) 51–56 (SciPy, 2010).

31. Edgar, R. C. MUSCLE: multiple sequence alignment with high accuracy and high throughput. *Nucleic Acids Res.* **32**, 1792–1797 (2004).
32. Trifinopoulos, J., Nguyen, L.-T., von Haeseler, A. & Minh, B. Q. W-IQ-TREE: a fast online phylogenetic tool for maximum likelihood analysis. *Nucleic Acids Res.* **44**, W232–W235 (2016).
33. Huerta-Cepas, J., Serra, F. & Bork, P. ETE 3: reconstruction, analysis, and visualization of phylogenomic data. *Mol. Biol. Evol.* **33**, 1635–1638 (2016).
34. Lee, T. S. et al. BglBrick vectors and datasheets: a synthetic biology platform for gene expression. *J. Biol. Eng.* **5**, 12 (2011).

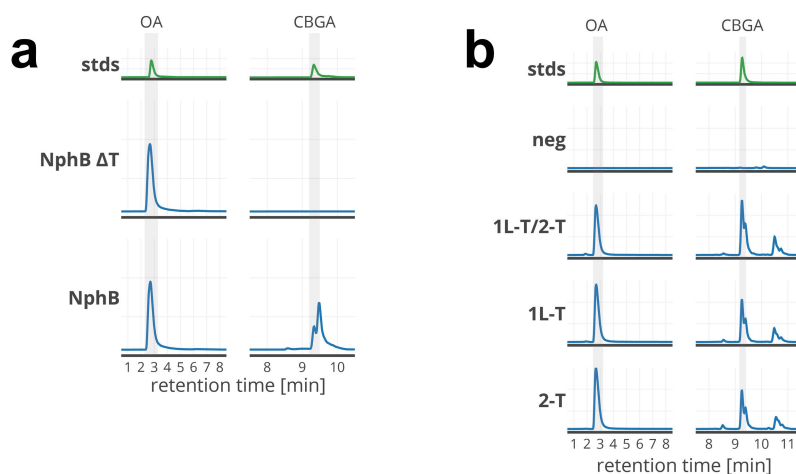


Extended Data Fig. 1 | Olivetolic acid production of yCAN03. All LC-MS chromatograms were selected for the theoretical m/z values of the respective compounds of interest (Extended Data Table 3). **a**, yCAN03 produced olivetolic acid. An additional peak was observed for the byproduct HTAL (same m/z as olivetolic acid), which has previously been reported. No zymolyase was added during extraction. Chromatography gradient for cannabinoid analogues was used. **b**, Mass spectrum of HTAL. Mass accuracy for observed m/z at a given retention time (RT) is reported in parts per million (p.p.m.).



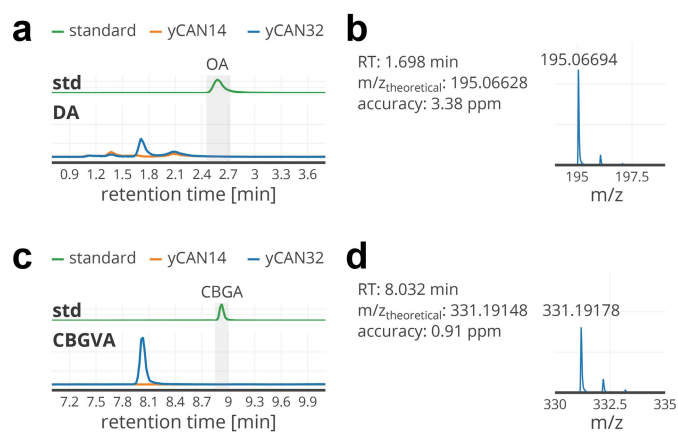
Extended Data Fig. 2 | CsPT4-T characterization. **a**, CsPT4-T localized to the purified microsomal fraction. All LC-MS chromatograms were selected for the theoretical m/z values of the respective compounds of interest (Extended Data Table 3). Chromatography gradient for cannabinoid analogues was used. Signals were compared to genuine standards. We incubated boiled microsomal fraction (mic ΔT), soluble fraction (sol) and microsomal fraction (mic) of γ CAN14 in the presence of 500 μ M olivetolic acid and 500 μ M GPP for 1 h at 30 °C and observed GOT activity only in the microsomal fraction. **b**, CsPT4-T olivetolic acid kinetics. Using nonlinear regression to fit the Michaelis-Menten kinetic model for varied olivetolic acid (0.25 μ M–0.56 mM) and constant GPP (1.67 mM) concentrations revealed a K_m (olivetolic acid) = $6.73 \pm 0.26 \mu$ M (mean \pm s.d.) ($n = 4$ technically independent

samples; measurements were plotted individually). **c**, CsPT4-T GPP kinetics. Eadie-Hofstee linearization of Michaelis-Menten model showed non-Michaelis-Menten type behaviour for CsPT4-T for varied GPP (0.25 μ M–1.67 mM) and constant olivetolic acid (1.67 mM) concentrations. Measurements do not fall on a line as would be expected (R^2 , coefficient of determination; $n = 3$ technically independent samples; measurements were plotted individually). **d**, Phylogenetic tree of *Cannabis* and related prenyltransferases. The numbers indicate the bootstrap value (%) from 1,000 replications. Grey dashed lines depict branch offsets to accommodate labels. The scale bar shows the amino acid substitution ratio. Prenyltransferases cluster by biosynthetic pathway. CsPT4 catalyses CBGA production. The function of the other *Cannabis* prenyltransferases remains unknown.



Extended Data Fig. 3 | In vitro activity of NphB and HIPT. All LC-MS chromatograms were selected for the theoretical m/z values of the respective compounds of interest (Extended Data Table 3). Chromatography gradient for cannabinoid analogues was used. Signals were compared to genuine standards. **a**, Purified NphB catalysed the condensation of GPP and olivetolic acid to CBGA when incubated with 5 mM olivetolic acid and 5 mM GPP at room temperature for 24 h. The enzyme produced at least one other isomer of CBGA, which is

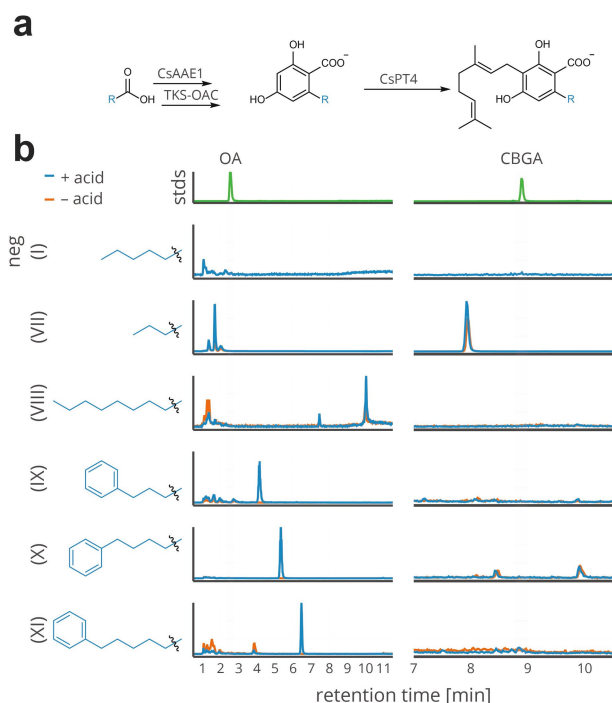
consistent with previous reports¹⁸. Boiling NphB (NphB Δ T) abolished activity. **b**, Microsomal fractions were prepared from γ CAN21, γ CAN22, and γ CAN23 that expressed HIPT1L-T (1L-T), HIPT1L-T and HIPT2-T (1L-T/2-T), and HIPT2-T (2-T), respectively. Incubation with 5 mM olivetolic acid and 5 mM GPP for 24 h at 30 °C yielded CBGA as well as several isomers. CBGA production was not observed when incubating HIPT1L-T and HIPT2-T with their native substrates phlorisovalerophenone and dimethylallyl diphosphate (neg).



Extended Data Fig. 4 | Divarinolic acid and CBGVA production.

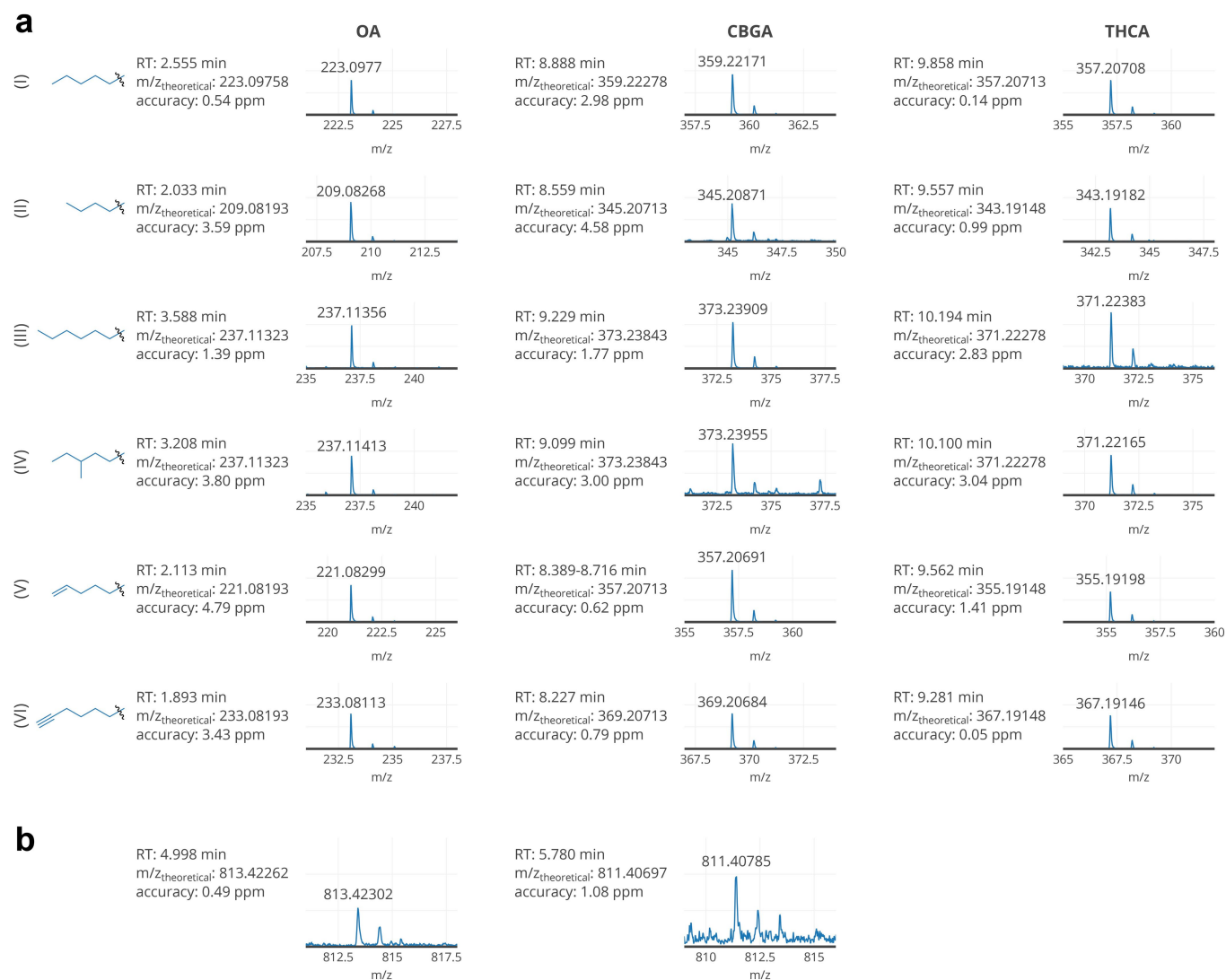
All LC–MS chromatograms were selected for the theoretical m/z values of the respective compounds of interest (Extended Data Table 3).

a, Compared with a strain that lacks the olivetolic acid synthesis pathway (yCAN14), yCAN32 produced an additional compound with an m/z value corresponding to divarinolic acid (DA). The retention time of divarinolic acid is reduced relative to olivetolic acid, owing to the reduced hydrophobicity that is conferred by the shorter 5-propyl side chain. No zymolyase was added during extraction. Chromatography gradient for cannabinoid analogues was used. **b**, Mass spectrum of divarinolic acid. Mass accuracy for observed m/z at a given retention time is reported in parts per million. **c**, As expected, the prenylated variant of divarinolic acid, CBGVA, was also detected. **d**, Mass spectrum of CBGVA, with mass accuracy reported in parts per million.



Extended Data Fig. 5 | Production of cannabinoid analogues.

a, Pathway assayed for incorporation of different fatty acid precursors into the corresponding olivetolic acid and CBGA analogues. **b**, Ion-extracted chromatograms of the engineered strain (*yCAN31*) grown in the presence of butanoic acid (VII), nonanoic acid (VIII), 5-phenylpentanoic acid (IX), 6-phenylhexanoic acid (X) and 7-phenylheptanoic acid (XI); peaks corresponding to their respective olivetolic acid and HTAL analogues were detected. Detected analogue peaks shifted in retention time depending on their change in hydrophobicity relative to olivetolic acid and CBGA standards. Feeding 1 mM hexanoic acid to *yCAN30* yielded neither olivetolic acid nor CBGA (neg (I)). Supplementation of the medium with butanoic acid did not increase production of divarinolic acid or CBGVA over baseline levels. Feeding phenyl-substituted acids yielded only either the corresponding HTAL analogues or olivetolic acid analogues. All LC-MS chromatograms were selected for the theoretical m/z values of the respective compounds of interest (Extended Data Table 3).



Extended Data Fig. 6 | LC-MS spectra of cannabinoid analogues. Mass accuracies for observed m/z values at a given retention time are reported in parts per million. **a**, Mass spectra of cannabinoid analogues produced from fatty acids (I) to (VI). **b**, Mass spectra of 6hCBGA and 6hTHCA.

Extended Data Table 1 | Strains and plasmids used in this study

Strain	Genotype/Description	Source	SYNBERC Part ID
CEN.PK2-1C	MATa; his3Δ1; leu2-3_112; ura3-52; trp1-289; MAL2-8c; SUC2	Euroscarf	-
yCAN01	CEN.PK2-1C {ACC1::GAL1p-CsTKS-ENO1t/GAL10p-CsOAC-SSA1t}	This work	SBa_001907
yCAN02	yCAN01 {URA3::HHF1p-CsAAE1-ADH1t}	This work	SBa_001908
yCAN03	yCAN01 {URA3::GAL1p-CnPaaH1-ENO1t/GAL10-CaCrt-SSA1t/GAL7p-TdTer-ADH1t/TEF2p-ReBktB-PGK1t}	This work	SBa_001909
yCAN10	CEN.PK2-1C {erg9::KanMX/CTR3p-ERG9; leu2-3_112::His3MX6/GAL1p-ERG19/GAL1p-ERG8; ura3-52::GAL1p-EfmVaS(A110G)-CYC1t/GAL10p-EfmVaE-ADH1t; his3Δ1::hphMX4/GAL1p-ERG12/GAL10p-IDI1; 308a::GAL1p-ERG20(F96W-N127W)-TDH1t}	This work	SBa_001910
yCAN11	yCAN10 {1114a::GAL1p-CsPT1-T-TDH1t}	This work	SBa_001911
yCAN12	yCAN10 {1114a::GAL1p-CsPT2-T-TDH1t}	This work	SBa_001912
yCAN13	yCAN10 {1114a::GAL1p-CsPT3-T-TDH1t}	This work	SBa_001913
yCAN14	yCAN10 {1114a::GAL1p-CsPT4-T-TDH1t}	This work	SBa_001938
yCAN15	yCAN10 {1114a::GAL1p-CsPT5-T-TDH1t}	This work	SBa_001939
yCAN16	yCAN10 {1114a::GAL1p-CsPT6-T-TDH1t}	This work	SBa_001940
yCAN17	yCAN10 {1114a::GAL1p-CsPT7-T-TDH1t}	This work	SBa_001941
yCAN18	yCAN10 {1114a::GAL1p-HIPT1L-T-TDH1t}	This work	SBa_001942
yCAN19	yCAN10 {1114a::GAL1p-HIPT2-T-TDH1t}	This work	SBa_001943
yCAN20	yCAN10 {1114a::GAL1p-NphB-TDH1t}	This work	SBa_001944
yCAN21	CEN.PK2-1C pCAN03	This work	SBa_001945
yCAN22	CEN.PK2-1C pCAN04	This work	SBa_001946
yCAN23	CEN.PK2-1C pCAN05	This work	SBa_001947
yCAN30	yCAN14 {607b::HHF1p-CsAAE1-ADH1t}	This work	SBa_001948
yCAN31	yCAN30 {911b::GAL1p-CsTKS-CsOAC-ENO1t}	This work	SBa_001949
yCAN32	yCAN31 {1014a::GAL1p-CnPaaH1-ENO1t/GAL10p-CaCrt-SSA1t/GAL7p-TdTer-ADH1t/TEF2p-ReBktB-PGK1t}	This work	SBa_001950
yCAN40	yCAN31 {416d::GAL1p-proA-CsTHCAS-ADH1t}	This work	SBa_001951
yCAN41	yCAN31 {416d::GAL1p-proA-CsCBDAS-CYC1t}	This work	SBa_001952
yCAN42	yCAN32 {416d::GAL1p-proA-CsTHCAS-ADH1t}	This work	SBa_001953
yCAN43	yCAN32 {416d::GAL1p-proA-CsCBDAS-CYC1t}	This work	SBa_001954
yCAN50	yCAN42 {YPRCd15c::GAL1p-CsTKS-CsOAC-ENO1t}	This work	SBa_002026
yCAN51	yCAN42 {YPRCd15c::GAL1p-CsPT4-T-TDH1t}	This work	SBa_002027
yCAN52	yCAN42 {YPRCd15c::GAL1p-proA-CsTHCAS-ADH1t}	This work	SBa_002028
yCAN53	yCAN52 {720a::GAL1p-CsTKS-CsOAC-ENO1t}	This work	SBa_002029
BL21 (DE3)	<i>fhuA2</i> [<i>lon</i>] <i>ompT gal</i> (λ DE3) [<i>dcm</i>] λ DE3 = λ sBamHlo Δ EcoRI-B int::(<i>lacI::PlacUV5::T7 gene1</i>) <i>i21</i> Δ <i>nin5</i>	Δ <i>hsdS</i> New England Biolabs, Inc.	
eCAN01	BL21 pCAN01	This work	SBa_001956
pCAN01	pBbE7a Amp ^r T7p-NphB-6xHis-DBLt	This work	SBa_001955
pCAN02	E. coli plasmid with p15A origin and Amp selection marker	Ref. ³⁴	-
pCAN03	pESC-URA with GAL1p-HIPT1L-T-CYC1t	Ref. ¹⁹	-
pCAN04	pESC-URA with GAL1p-HIPT1L-T-CYC1t-GAL10p-HIPT2-T-ADH1t	Ref. ¹⁹	-
pCAN05	pESC-URA with GAL10p-HIPT2-T-ADH1t	This work	SBa_002030

Extended Data Table 2 | Fatty acids tested for cannabinoid analogue production

group	fatty acid	final concentration in medium [mM]	OA analogue	CBGA analogue	THCA analogue
C4	butanoic acid	1.0	y*	y*	-
	4-phenylbutanoic acid	0.1	n	n	-
C5	pentanoic acid	1.0	y	y	y
	4-pentynoic acid	1.0	n	n	-
	trans-2-pentenoic acid	1.0	n	n	-
	5-phenylpentanoic acid	0.2	?	n	-
C6	hexanoic acid	1.0	y	y	y
	5-hexynoic acid	1.0	n	n	-
	4-methylhexanoic acid	1.0	y	y	y
	trans-2-hexenoic acid	1.0	?	n	-
	5-hexenoic acid	1.0	y	y	y
	6-phenylhexanoic acid	0.1	?	n	-
C7	heptanoic acid	1.0	y	y	y
	6-heptynoic acid	1.0	y	y	y
	7-phenylheptanoic acid	0.1	?	n	-
C8	octanoic acid	0.1	n	n	-
	trans-2-octenoic acid	0.1	n	n	-
C9	nonanoic acid	0.1	?	n	-
	trans-2-nonenoic acid	0.1	n	n	-
C10	decanoic acid	0.1	n	n	-

n, no detection of olivetolic acid/CBGA/THCA analogue; y, detected olivetolic acid/CBGA/THCA analogue; -, not tested; *, feeding fatty acid did not increase olivetolic acid/CBGA/THCA baseline levels; ?, possible olivetolic acid or HTAL analogue.

Extended Data Table 3 | *m/z* values of compounds analysed in this study

Category	Name	Abbreviation	<i>m/z</i> (negative mode)	Fatty acid precursor, substrate
OA and analogues	divarinolic acid	DA	195.06628	butanoic acid
	3-butyl-resorcylic acid		209.08193	pentanoic acid
	3-(3-butynyl)-resorcylic acid		205.05063	4-pentynoic acid
	3-(trans-1-butenyl)-resorcylic acid		207.06628	trans-2-pentenoic acid
	olivetolic acid	OA	223.09758	hexanoic acid
	3-(3-methylpentyl)-resorcylic acid		237.11323	4-methylhexanoic acid
	3-(3-pentynyl)-resorcylic acid		219.06628	5-hexynoic acid
	3-(trans-1-pentenyl)-resorcylic acid		221.08193	trans-2-hexenoic acid
	3-(4-pentenyl)-resorcylic acid		221.08193	5-hexenoic acid
	3-hexyl-resorcylic acid		237.11323	heptanoic acid
	3-(5-hexynyl)-resorcylic acid		233.08193	6-heptynoic acid
	3-heptyl-resorcylic acid		251.12888	octanoic acid
	3-(trans-1-hexyl)-resorcylic acid		249.11323	trans-2-octenoic acid
	3-octyl-resorcylic acid		265.14453	nonanoic acid
	3-(trans-1-octenyl)-resorcylic acid		263.12888	trans-2-nonenoic acid
	3-nonyl-resorcylic acid		279.16018	decanoic acid
	3-(3-Phenylpropyl)-resorcylic acid		257.08193	4-Phenylbutanoic acid
3-(4-Phenylbutyl)-resorcylic acid		271.09758	5-Phenylpentanoic acid	
3-(5-Phenylpentyl)-resorcylic acid		285.11323	6-Phenylhexanoic acid	
3-(6-Phenylhexyl)-resorcylic acid		299.12888	7-Phenylheptanoic acid	
CBGA and analogues	cannabigerovanilic acid	CBGVA	331.19148	butanoic acid
	3-butyl-6-geranyl-resorcylic acid		345.20713	pentanoic acid
	3-(3-butynyl)-6-geranyl-resorcylic acid		341.17583	4-pentynoic acid
	3-(trans-1-butenyl)-6-geranyl-resorcylic acid		343.19148	trans-2-pentenoic acid
	cannabigerolic acid	CBGA	359.22278	hexanoic acid
	3-(3-methylpentyl)-6-geranyl-resorcylic acid		373.23843	4-methylhexanoic acid
	3-(4-pentenyl)-6-geranyl-resorcylic acid		355.19148	5-hexynoic acid
	3-(trans-1-pentenyl)-6-geranyl-resorcylic acid		357.20713	trans-2-hexenoic acid
	3-(4-pentenyl)-6-geranyl-resorcylic acid		357.20713	5-hexenoic acid
	3-hexyl-6-geranyl-resorcylic acid		373.23843	heptanoic acid
	3-(5-hexynyl)-6-geranyl-resorcylic acid	6hCBGA	369.20713	6-heptynoic acid
	3-heptyl-6-geranyl-resorcylic acid		387.25408	octanoic acid
	3-(trans-1-heptenyl)-6-geranyl-resorcylic acid		385.23843	trans-2-octenoic acid
	3-octyl-6-geranyl-resorcylic acid		401.26973	nonanoic acid
	3-(trans-1-octenyl)-6-geranyl-resorcylic acid		399.25408	trans-2-nonenoic acid
3-nonyl-6-geranyl-resorcylic acid		415.28538	decanoic acid	
3-(3-Phenylpropyl)-6-geranyl-resorcylic acid		393.20713	4-Phenylbutanoic acid	
3-(4-Phenylbutyl)-6-geranyl-resorcylic acid		407.22278	5-Phenylpentanoic acid	
3-(5-Phenylpentyl)-6-geranyl-resorcylic acid		421.23843	6-Phenylhexanoic acid	
3-(5-Phenylhexyl)-6-geranyl-resorcylic acid		435.25408	7-Phenylheptanoic acid	
THCA and analogues, CBDA	Δ^8 -tetrahydrocannabivarinic acid/ cannabidivarinic acid	THCVA/CBDVA	329.17583	butanoic acid
	(6aR,10aR)-1-Hydroxy-6,6,9-trimethyl-3-butyl-6a,7,8,10a-tetrahydro-6H-dibenzo[b,d]pyran-2-carboxylic acid		343.19148	pentanoic acid
	Δ^8 -tetrahydrocannabinolic acid/ cannabidiolic acid	THCA/CBDA	357.20713	hexanoic acid
	(6aR,10aR)-1-Hydroxy-6,6,9-trimethyl-3-(3-methylpentyl)-6a,7,8,10a-tetrahydro-6H-dibenzo[b,d]pyran-2-carboxylic acid		371.22278	4-methylhexanoic acid
	(6aR,10aR)-1-Hydroxy-6,6,9-trimethyl-3-(4-butenyl)-6a,7,8,10a-tetrahydro-6H-dibenzo[b,d]pyran-2-carboxylic acid		355.19148	5-hexenoic acid
	(6aR,10aR)-1-Hydroxy-6,6,9-trimethyl-3-hexyl-6a,7,8,10a-tetrahydro-6H-dibenzo[b,d]pyran-2-carboxylic acid		371.22278	heptanoic acid
(6aR,10aR)-1-Hydroxy-6,6,9-trimethyl-3-(5-hexynyl)-6a,7,8,10a-tetrahydro-6H-dibenzo[b,d]pyran-2-carboxylic acid	6hTHCA	367.19148	6-heptynoic acid	
Click chemistry products		biotin-PEG3-6hCBGA	813.42262	6hCBGA
		biotin-PEG3-6hTHCA	811.40697	6hTHCA
Miscellaneous	hexanoyl triacetate lactone	HTAL	223.09758	hexanoic acid

All values are for the $[M-H]^-$ ion.

Extended Data Table 4 | Accession numbers of (putative) prenyltransferases tested for GOT activity and prenyltransferases used for phylogenetic analysis

Enzyme	Organism of origin	GenBank Accession No.
CsPT1	<i>Cannabis</i>	BK010678
CsPT2	<i>Cannabis</i>	BK010679
CsPT3	<i>Cannabis</i>	BK010680
CsPT4	<i>Cannabis</i>	BK010648
CSPT5	<i>Cannabis</i>	BK010681
CsPT6	<i>Cannabis</i>	BK010682
CsPT7	<i>Cannabis</i>	BK010683
LePGT1	<i>Lithospermum erythrorhizon</i>	BAB84122
LePGT2	<i>Lithospermum erythrorhizon</i>	Q8W404.1
AtUbiA	<i>Arabidopsis thaliana</i>	NP_190750.1
AtVTE2	<i>Arabidopsis thaliana</i>	OAP10166.1
AtHPT1	<i>Arabidopsis thaliana</i>	NP849984
AtPPT1	<i>Arabidopsis thaliana</i>	Q93YP7.1
ZmVTE2-1	<i>Zea mays</i>	XP_008659042.1
ZmHPT	<i>Zea mays</i>	CAC18910.1
OsPPT1	<i>Oryza sativa</i>	BAE96574.1
OsHGGT	<i>Oryza sativa</i>	AAP43913.1
OsHPT	<i>Oryza sativa</i>	CAC18914.1
SfN8DT-1	<i>Sophora flavescens</i>	BAG12671
HIPT1L	<i>Humulus lupulus</i>	A0A0B5A051.1
HIPT1	<i>Humulus lupulus</i>	E5RP65.1
HIPT2	<i>Humulus lupulus</i>	A0A0B4ZTQ2.1
TaHGGT	<i>Triticum aestivum</i>	AAP43912.1
HvHGGT	<i>Hordeum vulgare</i>	AAP43911.1
GmHPT	<i>Glycine max</i>	CAC18917.1

Reporting Summary

Nature Research wishes to improve the reproducibility of the work that we publish. This form provides structure for consistency and transparency in reporting. For further information on Nature Research policies, see [Authors & Referees](#) and the [Editorial Policy Checklist](#).

Statistical parameters

When statistical analyses are reported, confirm that the following items are present in the relevant location (e.g. figure legend, table legend, main text, or Methods section).

n/a | Confirmed

- The exact sample size (n) for each experimental group/condition, given as a discrete number and unit of measurement
- An indication of whether measurements were taken from distinct samples or whether the same sample was measured repeatedly
- The statistical test(s) used AND whether they are one- or two-sided
Only common tests should be described solely by name; describe more complex techniques in the Methods section.
- A description of all covariates tested
- A description of any assumptions or corrections, such as tests of normality and adjustment for multiple comparisons
- A full description of the statistics including central tendency (e.g. means) or other basic estimates (e.g. regression coefficient) AND variation (e.g. standard deviation) or associated estimates of uncertainty (e.g. confidence intervals)
- For null hypothesis testing, the test statistic (e.g. F , t , r) with confidence intervals, effect sizes, degrees of freedom and P value noted
Give P values as exact values whenever suitable.
- For Bayesian analysis, information on the choice of priors and Markov chain Monte Carlo settings
- For hierarchical and complex designs, identification of the appropriate level for tests and full reporting of outcomes
- Estimates of effect sizes (e.g. Cohen's d , Pearson's r), indicating how they were calculated
- Clearly defined error bars
State explicitly what error bars represent (e.g. SD, SE, CI)

Our web collection on [statistics for biologists](#) may be useful.

Software and code

Policy information about [availability of computer code](#)

Data collection

No software was used

Data analysis

Agilent MassHunter Qualitative Analysis B.06.00
MAFFT 7.394
RAxML v.7.2.8
FigTree v1.4.3
Openchrom 1.2.0
custom Python 3.6 scripts with Scipy(v.1.0.0), and Pandas(v.0.22.0) packages

For manuscripts utilizing custom algorithms or software that are central to the research but not yet described in published literature, software must be made available to editors/reviewers upon request. We strongly encourage code deposition in a community repository (e.g. GitHub). See the Nature Research [guidelines for submitting code & software](#) for further information.

Data

Policy information about [availability of data](#)

All manuscripts must include a [data availability statement](#). This statement should provide the following information, where applicable:

- Accession codes, unique identifiers, or web links for publicly available datasets
- A list of figures that have associated raw data
- A description of any restrictions on data availability

Nucleotide sequence data of Cannabis candidate prenyltransferases are available in the Third Party Annotation Section of the DDBJ/ENA/GenBank databases (Extended Data Table 4). Strains and plasmids developed for this study (Extended Data Table 1), along with annotated sequences, have been deposited in the Synthetic Biology Engineering Research Center (Synberc) Registry (<https://synberc-registry.jbei.org/>) and are physically available from the authors upon reasonable request. Strains producing controlled substances or its direct precursors can only be provided to laboratories/institutions with appropriate approvals and licenses (e.g., DEA permits). The custom Python 3.6 scripts for data analysis are available from the authors upon reasonable request.

Field-specific reporting

Please select the best fit for your research. If you are not sure, read the appropriate sections before making your selection.

Life sciences Behavioural & social sciences Ecological, evolutionary & environmental sciences

For a reference copy of the document with all sections, see [nature.com/authors/policies/ReportingSummary-flat.pdf](https://www.nature.com/authors/policies/ReportingSummary-flat.pdf)

Life sciences study design

All studies must disclose on these points even when the disclosure is negative.

Sample size	Sample size of three or four was chosen following all previous publications in similar field.
Data exclusions	No data were excluded from the analysis.
Replication	All in vivo productions and in vitro experiments were performed in triplicates or quadruplicates, all attempts at replication were successful.
Randomization	All strains for in vivo study were randomly picked from the corresponding agar plate.
Blinding	For all experiments, the analytical team who analyzed the amounts of substrates and products were blind to the samples they were analyzing. All data were analyzed and checked by multiple authors and reviewed by the corresponding author.

Reporting for specific materials, systems and methods

Materials & experimental systems

n/a	Included in the study
<input checked="" type="checkbox"/>	<input type="checkbox"/> Unique biological materials
<input checked="" type="checkbox"/>	<input type="checkbox"/> Antibodies
<input checked="" type="checkbox"/>	<input type="checkbox"/> Eukaryotic cell lines
<input checked="" type="checkbox"/>	<input type="checkbox"/> Palaeontology
<input checked="" type="checkbox"/>	<input type="checkbox"/> Animals and other organisms
<input checked="" type="checkbox"/>	<input type="checkbox"/> Human research participants

Methods

n/a	Included in the study
<input checked="" type="checkbox"/>	<input type="checkbox"/> ChIP-seq
<input checked="" type="checkbox"/>	<input type="checkbox"/> Flow cytometry
<input checked="" type="checkbox"/>	<input type="checkbox"/> MRI-based neuroimaging