COMPLETE BIOSYNTHESIS OF CANNABINOIDS AND THEIR UNNATURAL ANALOGUES IN YEAST

Xiaozhou Luo[1]*, Michael A. Reiter[1,2]*, Leo d'Espaux[3]†, Jeff Wong[3]†, Charles M. Denby[1]†, Anna Lechner[5,6]†, Yunfeng Zhang[1,7], Adrian T. Grzybowski[1], Simon Harth[3], Weiyin Lin[3], Hyunsu Lee[3,8], Changhua Yu[3,6], John Shin[3,5], Kai Deng[4,9], Veronica T. Benites[3], George Wang[3], Edward E. K. Baidoo[3], Yan Chen[3], Ishaan Dev[3,5], Christopher J. Petzold[3], Jay D. Keasling[1,3,5,6,10,11]

1. California Institute of Quantitative Biosciences (QB3), University of California, Berkeley, California 94720, United States
2. Department of Biosystems Science and Engineering, ETH Zurich, Mattenstrasse 26, 4058 Basel, Switzerland
3. Biological Systems and Engineering Division, Lawrence Berkeley National Laboratory, Berkeley, California 94720, United States
4. Department of Plant and Microbial Biology, University of California, Berkeley, CA 94720, United States
5. Department of Chemical & Biomolecular Engineering, University of California, Berkeley, California 94720, United States
6. Department of Bioengineering, University of California, Berkeley, California 94720, United States
7. Key Laboratory of Industrial Biotechnology, Ministry of Education, Jiangnan University, 1800 Lihu Road, Wuxi, Jiangsu 214122, China
8. Department of Chemistry, University of California, Berkeley, California 94720, United States
9. Biotechnology and Bioengineering Department, Sandia National Laboratories, Livermore, California 94551, United States
10. Novo Nordisk Foundation Center for Biosustainability, Technical University of Denmark, Building 220, Kemitorvet, DK-2800 Kgs, Lyngby, Denmark
11. Center for Synthetic Biochemistry, Institute of Synthetic Biology, Shenzhen Institutes of Advanced Technologies, Shenzhen, Guangdong 518055, China

†Present addresses: Demetrix, Inc., 5858 Horton Street, Emeryville, California 94608, USA (L.d'E., J.W.); Genomatica, Inc., 4755 Nexus Center Dr, San Diego, CA 92121, USA (A.L.); Berkeley Brewing Science, Inc., 2332 4th St, Suite E, Berkeley, California 94710, USA (C.M.D)

*These authors contributed equally to this work.

Correspondence and requests for materials should be addressed to J.D.K. (keasling@berkeley.edu).

Cannabis sativa L. (*Cannabis*) has been cultivated and used around the globe for its medicinal properties for millennia.[1] Some cannabinoids, the hallmark constituents of *Cannabis*, as well as analogues thereof have been investigated extensively for their potential medical applications.[2] Certain cannabinoid formulations have been approved as prescription drugs in several countries for the treatment of a variety of human ailments.[3] However, greater study and medicinal use of cannabinoids has been hampered by the legal scheduling of *Cannabis*, low abundances of nearly all of the several dozens of known cannabinoids *in planta*,[4] and their structural complexity, which limits bulk chemical synthesis. Here, we report the complete biosynthesis of the major cannabinoids cannabigerolic acid (CBGA), Δ^9-tetrahydrocannabinolic acid (THCA), cannabidiolic acid (CBDA), Δ^9-tetrahydrocannabivarinic acid (THCVA), and cannabidivarinic acid (CBDVA) in *Saccharomyces cerevisiae* from the simple sugar galactose. To accomplish this, we engineered the native mevalonate pathway to provide high flux of geranyl pyrophosphate (GPP) and introduced a heterologous, multi-organism-derived hexanoyl-CoA biosynthetic pathway[5] as well as the *Cannabis* genes encoding the enzymes involved in olivetolic acid (OA) biosynthesis,[6] a previously undiscovered enzyme with geranylpyrophosphate:olivetolate geranyltransferase (GOT) activity, and corresponding cannabinoid synthases.[7,8] Furthermore, we established a biosynthetic approach, harnessing the promiscuity of several pathway genes, for the production of cannabinoid analogues. Feeding different fatty acids to our engineered strains yielded cannabinoid analogues with modifications in the part of the molecule known to alter receptor binding affinity and potency.[9] We also demonstrated that our biological system can 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, which will allow for more rigorous study of cannabinoids and ultimately the development of new treatments for a variety of human ailments.

We initiated construction of the cannabinoid-producing yeast by focusing first on production of OA (Fig. 1), an initial intermediate in the cannabinoid biosynthetic pathway. Two *Cannabis* enzymes, a tetraketide synthase (CsTKS)[10] and an olivetolic acid cyclase (CsOAC),[6] have been reported to produce

1

OA from hexanoyl-CoA and malonyl-CoA. To produce OA 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 OA from galactose (Fig. 2a), consistent with the fact that *S. cerevisiae* maintains low intracellular levels of hexanoyl-CoA.[11] To increase the supply of hexanoyl-CoA, we fed 1 mM hexanoic acid, which can be converted to hexanoyl-CoA by an endogenous acyl activating enzyme (AAE), and observed six-fold higher OA production (1.3 mg/L) compared to no exogenous hexanoic acid feeding. 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 into yCAN01 an AAE from *Cannabis* (CsAAE1), which is thought to catalyze this step *in planta*.[12] The resulting strain (yCAN02) showed a two-fold increase in OA titer (3.0 mg/L) when fed 1 mM hexanoic acid (Fig. 2a). To produce hexanoyl-CoA from galactose and complete the OA pathway, we introduced into yCAN01 a previously reported hexanoyl-CoA pathway[5]. The resulting strain (yCAN03) produced 1.6 mg/L OA (Fig. 2a).

CBGA, the precursor to THCA, CBDA, and numerous other cannabinoids, is produced from the mevalonate pathway intermediate GPP and OA by GOT. GOT activity was detected in *Cannabis* extracts over two decades ago,[13] and a *Cannabis* GOT (CsPT1) was patented ten years later.[14] To enable *in vivo* testing of CsPT1, we constructed a GPP-overproducing strain (yCAN10) with an upregulated mevalonate pathway[15] and a mutant gene of the endogenous farnesyl pyrophosphate synthase ERG20(F96W-N127W), which preferentially produces GPP over FPP.[16] However, we were unable to observe any GOT activity when we expressed CsPT1 or truncations thereof in yCAN10.

To identify an enzyme with GOT activity that would function in yeast, we searched for candidate prenyltransferase (PT) enzymes from *Cannabis* and other organisms. These included NphB, a soluble PT from *Streptomyces* sp.[17] with GOT activity *in vitro*,[18] as well as HlPT1L and HlPT2, two PTs involved in bitter acid biosynthesis in *Humulus lupulus*, a close relative of *Cannabis*.[19] In addition, we mined published

2

Cannabis transcriptomes[20,21] for GOT candidates. We searched full-length transcripts using Basic Local Alignment Search Tool (BLAST) against CsPT1, HlPT1L, and HlPT2, ultimately settling on six enzymes (CsPT2–7).

For functional expression in yeast, we removed predicted N-terminal plastid-targeting sequences[22,23] from all *Cannabis* and *Humulus lupulus* PTs resulting in CsPT2-T–CsPT7-T and HlPT1L-T, HlPT2-T, respectively. Each GOT candidate was introduced into yCAN10, and the resulting strains (yCAN12–yCAN20) were cultured in 1 mM OA and assayed for CBGA production by liquid chromatography-mass spectrometry (LC-MS). Of the nine GOT candidates tested, only the strain expressing CsPT4-T (yCAN14) produced detectable amounts of CBGA (136 mg/L, Fig. 2b). CsPT4-T is predicted to have eight transmembrane helices[24] and 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 (OA) = 6.73 ± 0.26 µM and non-Michaelis-Menten type behavior for GPP (Extended Data Fig. 2b,c). Similar *in vitro* assays with the other GOT candidates (NphB, HlPT1L-T, and HlPT2-T) showed low and non-specific GOT activity or no activity (CsPT1-T) (Extended Data Fig. 3). CsPT4 clusters with other *Cannabis* PTs in a phylogenetic tree (Extended Data Fig. 2d) but bears only 62% homology with CsPT1.

Next, we set out to produce CBGA from simpler precursors such as hexanoic acid or galactose by reconstituting the OA biosynthesis module in strain yCAN14. Adding *CsTKS*, *CsOAC* and *CsAEE1* generated strain yCAN31, which produced 7.2 mg/L of CBGA from 1 mM hexanoic acid. Finally, integrating the hexanoyl-CoA pathway into yCAN31 resulted in strain yCAN32, which produced 1.4 mg/L CBGA from galactose.

Synthases for the conversion of CBGA to THCA[25] (THCAS) and CBDA[26] (CBDAS) in *Cannabis* were identified two decades ago. THCAS has been functionally expressed and assayed *in vitro* in insect cells[27] and yeast.[7,28] Yet, *in vivo* production of THCA, CBDA or other late cannabinoids from sugars has not been shown to date. To complete the cannabinoid biosynthesis pathway, we replaced the N-terminal

secretory signal peptides[29] of THCAS and CBDAS with a vacuolar localization tag to enable their functional expression[7,28] and integrated the resulting sequences into yCAN31 (resulting in yCAN40 and yCAN41) and into yCAN32 (resulting in yCAN42 and yCAN43). Culturing yCAN40 or yCAN41 with 1 mM hexanoic acid yielded 1.1 mg/L THCA or 4.3 µg/L CBDA, respectively. Culturing yCAN42 or yCAN43 with galactose yielded 2.3 mg/L THCA or 4.2 µg/L of CBDA, respectively (Fig. 3a,b).

Besides cannabinoids derived from OA, *Cannabis* also produces compounds originating from divarinolic acid (DA), an OA analogue wherein the C3-pentyl sidechain is substituted with a propyl moiety,[30] suggesting promiscuity of at least some pathway enzymes. The precursor to DA presumably is butanoyl-CoA,[10] an intermediate in our engineered hexanoyl-CoA pathway. Indeed, yCAN42 and yCAN43 produced 1.2 mg/L of tetrahydrocannabivarinic acid (THCVA) and 6.0 µg/L cannabidivarinic acid (CBDVA), respectively (Fig. 3). Furthermore, peaks corresponding to the predicted m/z-values of DA and the propyl-variant of CBGA, cannabigerovarinic acid (CBGVA), were also detected (Extended Data Fig. 4).

In order to improve titers 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, yCAN52). We assumed that the GPP supply was sufficient, since OA was converted efficiently to CBGA when supplementing the growth medium with 1 mM OA (Fig. 2b). When assayed, we observed a 3-fold increase in the concentration of intermediates as well as of THCA and THCVA for the TKS-OAC overexpression strain (yCAN50) in comparison to 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 by malonyl-CoA and hexanoyl-CoA conversion to OA. Lastly, we added one more copy of THCAS to yCAN50 (resulting in yCAN53) to enhance conversion of the now higher CBGA pool. yCAN53 yielded increased titers of THCA (8.0 mg/L) and THCVA (4.8 mg/L) (Fig. 3c). Further gains could be achieved by fine-tuning TKS-OAC and THCAS expression as well as by upregulating upstream precursor production.

In addition to their natural counterparts, unnatural cannabinoid analogues are actively being investigated for improved medicinal properties. One of the major pharmacophores of interest in this search is the C3 sidechain of THC since its length, size, structure and chemistry have been shown to modulate cannabinoid receptor (CB1 or CB2) binding affinity, CB1/CB2 selectivity and potency, amongst others.[9] 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 with various chain lengths, branching and degrees of saturation (Extended Data Table 2). LC-MS analysis revealed the production of OA 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). Furthermore, 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 intractable to direct incorporation. As proof of concept, we performed copper-catalyzed azide-alkyne cycloaddition on the respective 6-heptynoic acid CBGA (6hCBGA) as well as 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 novel avenue towards the production of cannabinoid analogues with tailored C3 sidechains.

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. Additionally, we further expanded the chemical space of cannabinoids by establishing and harnessing the intrinsic promiscuity of the cannabinoid pathway 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 ultimately new and better medicines.

ACKNOWLEDGEMENTS

AUTHOR CONTRIBUTIONS

X.L., M.A.R., L.d'E., J.W., C.M.D., A.L., and J.D.K. conceived of 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 FINANCIAL INTERESTS

X. Luo, M. A. Reiter, L. d'Espaux, J. Wong, C. M. Denby, E. E. K. Baidoo, and C. J. Petzold have a financial interest in Demetrix. J. D. Keasling has a financial interest in Amyris, Lygos, Demetrix, Constructive Biology, Napigen, and Maple Bio.

REFERENCES MAIN TEXT

1. Pacher, P., Bátkai, S. & Kunos, G. The endocannabinoid system as an emerging target of pharmacotherapy. *Pharmacol. Rev.* **58,** 389–462 (2006).

2. Whiting, P. F. *et al.* Cannabinoids for Medical Use: A Systematic Review and Meta-analysis. *JAMA* **313,** 2456–2473 (2015).

3. Hazekamp, A., Ware, M. A., Muller-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).

4. Mahmoud ElSolhly, W. G. Constituents of Cannabis Sativa. in *Handbook of Cannabis* (ed. Pertwee, R. G.) 3– 22 (Oxford University Press, 2014).

5. 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).

6. 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. U. S. A.* **109,** 12811–12816 (2012).

7. Zirpel, B., Stehle, F. & Kayser, O. Production of Δ9-tetrahydrocannabinolic acid from cannabigerolic acid by whole cells of Pichia (Komagataella) pastoris expressing Δ9-tetrahydrocannabinolic acid synthase from Cannabis sativa L. *Biotechnol. Lett.* **37,** 1869–1875 (2015).

8. Taura, F. *et al.* Cannabidiolic-acid synthase, the chemotype-determining enzyme in the fiber-type Cannabis sativa. *FEBS Lett.* **581,** 2929–2934 (2007).

9. Razdan, R. K. Structure–Activity Relationships of Classical Cannabinoids. in *The Cannabinoid Receptors* 3– 19 (2009).

10. Taura, F. *et al.* Characterization of olivetol synthase, a polyketide synthase putatively involved in cannabinoid biosynthetic pathway. *FEBS Lett.* **583,** 2061–2066 (2009).

11. Saerens, S. M. G., Delvaux, F. R., Verstrepen, K. J. & Thevelein, J. M. Production and biological function of volatile esters inSaccharomyces cerevisiae. *Microb. Biotechnol.* **3,** 165–177 (2010).

12. 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).

13. 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).

14. Page, J. E. & Boubakir, Z. Aromatic Prenyltransferase from Cannabis. *US Patent* (2012).

15. Reider Apel, A. *et al.* A Cas9-based toolkit to program gene expression in Saccharomyces cerevisiae. *Nucleic Acids Res.* **45,** 496–508 (2017).

16. 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).

17. Kuzuyama, T., Noel, J. P. & Richard, S. B. Structural basis for the promiscuous biosynthetic prenylation of aromatic natural products. *Nature* **435,** 983–987 (2005).

18. Zirpel, B., Degenhardt, F., Martin, C., Kayser, O. & Stehle, F. Engineering yeasts as platform organisms for cannabinoid biosynthesis. *J. Biotechnol.* **259,** 204–212 (2017).

19. 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).

20. Resource, M. P. G. Medicinal Plant Genomics Resource. Available at: http://medicinalplantgenomics.msu.edu. (Accessed: 4th April 2018)

21. van Bakel, H. *et al.* The draft genome and transcriptome of Cannabis sativa. *Genome Biol.* **12,** R102 (2011).

22. Emanuelsson, O., Brunak, S., von Heijne, G. & Nielsen, H. Locating proteins in the cell using TargetP, SignalP and related tools. *Nat. Protoc.* **2,** 953–971 (2007).

23. 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).

24. 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).

25. Taura, F., Morimoto, S., Shoyama, Y. & Mechoulam, R. First direct evidence for the mechanism of .DELTA.1-tetrahydrocannabinolic acid biosynthesis. *J. Am. Chem. Soc.* **117,** 9766–9767 (1995).

26. 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).

27. Sirikantaramas, S. *et al.* The gene controlling marijuana psychoactivity: molecular cloning and heterologous expression of Delta1-tetrahydrocannabinolic acid synthase from Cannabis sativa L. *J. Biol. Chem.* **279,** 39767–39774 (2004).

28. Zirpel, B., Degenhardt, F., Martin, C., Kayser, O. & Stehle, F. Engineering yeasts as platform organisms for cannabinoid biosynthesis. *J. Biotechnol.* **259,** 204–212 (2017).

29. 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).

30. 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).

LEGENDS

Figure 1 | Engineered biosynthetic pathway for synthesis of cannabinoids in *S. cerevisiae*. GPP was produced by introducing the *E. faecalis* genes *EfMvaE* and *EfMvaS,* overexpressing the native mevalonate pathway genes (*ERG12, ERG8, ERG19, IDI1*) and a mutated *ERG20(F96W-N127W)* (*Erg20**). Hexanoyl-CoA was produced using 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 genes encoding CsTKS and CsOAC from *Cannabis* produced OA, which was prenylated by CsPT4 from *Cannabis*. The resulting CBGA was transformed to the cannabinoid acids THCA and CBDA using the cannabinoid synthases THCAS and CBDAS. Upon exposure to heat THCA and CBDA decarboxylate to THC and CBD, respectively.

Figure 2 | Pathway engineering for OA and CBGA. a, OA production in yCAN01, yCAN02 and yCAN03 from galactose or 1 mM hexanoic acid 48 h after induction (hex-PW: hexanoyl-CoA pathway, error bars represent standard deviation around the mean, n = 3 biologically independent samples). **b,** *In vivo* production of CBGA by feeding OA. yCAN10 and yCAN14 cultures fed 1 mM OA were sampled 24 h after induction and extracts analyzed by LC-MS and signals 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).

Figure 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

11

pathway enzymes revealed bottlenecks and increased overall production (error bars represent standard deviation around the mean, n = 3 biologically independent samples).

Figure 4 | Production of cannabinoid analogues and post-fermentation derivatization. a, Pathway assayed for incorporation of different fatty acid precursors into the corresponding OA, CBGA and THCA analogues. **b,** Ion-extracted chromatograms of the engineered strains (yCAN31 for OA 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 OA, CBGA and THCA or their respective analogues. Detected analogue peaks shifted in retention time depending on their change in hydrophobicity relative to the OA, CBGA and THCA standards (stds). Feeding 1 mM hexanoic acid to the parent strain of yCAN31 (yCAN30) did not yield OA 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). **c,** yCAN40 was cultured in the presence of 1 mM 6-heptynoic acid (+ heptynoic acid) or 1 mM hexanoic acid (+ hexanoic acid). The corresponding extracts were reacted with an azide-PEG3-biotin conjugate (+ azide) or water (− azide) under copper-catalyzed 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.

METHODS

Detailed description of the methods used for the generation of *S. cerevisiae* yCAN strains are provided in the Supplementary Information.

General. Cannabinoid standards (CBGA, THCA, CBDA, CBGVA, THCVA and CBDVA) were purchased from Cerilliant Corporation (Round Rock, TX). OA was purchased from A1 Biochem Labs (Syracuse, NY). All other chemicals were obtained from Fisher Scientific and Sigma-Aldrich. Oligonucleotides and codon-optimized gBlock gene fragments were obtained from Integrated DNA Technologies (Skokie, IL). Yeast culture media were purchased from Becton, Dickinson and Company (Franklin Lakes, NJ), and all agar plates were obtained from Teknova (Hollister, CA). The project was registered with and under the regulation of University of California, Berkeley and Lawrence Berkeley National Laboratory controlled substances programs, Research Advisory Panel of California, and U.S. Drug Enforcement Administration. All researchers who have direct access to controlled substances or yeast strains capable of making such compounds have passed background screening. All work was conducted in a DEA-certified laboratory, and all protocols were reviewed by 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_{600nm} 0.2 into yeast extract peptone galactose (YPG). When indicated, medium was supplemented with 1 mM OA or corresponding fatty acid (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 rpm) at 30°C while supplementing 2% (w/v) galactose every 24 h. Subsequently, samples were treated with 2U/OD Zymolyase (Zymo Research) (2 h, 30°C, 800 rpm) (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 12,000 × g

for 1 min. Samples were extracted three times. The combined organic layers were evaporated in a vacuum oven (50°C) and the remainders resuspended in acetonitrile/H_2O/formic acid (80%:20%:0.05% (v/v)) (AHF). Lastly, samples were filtered with Ultrafree®-MC columns (0.22 μm pore size, PVDF membrane material).

Products were analyzed using either high-performance liquid chromatography with UV detection (HPLC-UV, Agilent 1200 series) or LC-MS (Agilent 6210 TOF, time-of-flight, and Agilent 6545 QTOF, quadrupole time-of-flight, Agilent Technologies, Santa Clara, CA, USA) equipped with a reverse phase C18 column (Kinetex® 2.6 μM, 100×2.1 mm, XB-C18, Phenomenex, Inc., Torrance, CA, USA). 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 for 12.5 min, increased from 0.2 mL/min to 0.4 mL/min in 0.2 min, and held at 0.4 mL/min for 2.8 min. The total LC 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 for 5.6 min, increased from 0.2 mL/min to 0.3 mL/min in 0.6 min, and held at 0.3 mL/min for 9.3 min. The total LC run time was 15.5 min. Sample injection volumes of 1 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 (ESI) was conducted in the negative ion mode and a capillary voltage of 3,500 V was utilized. 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 and 30 lb/in^2, respectively, and a drying-gas temperature of 330°C was used throughout. For QTOF-MS, ESI was conducted via the Agilent Jet Stream thermal gradient

focusing technology, where the sheath gas flow rate and temperature were set to 12 L/min and 350°C, respectively. Drying and nebulizing gases were set to 10 L/min and 25 lb/in^2, 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 ppm, mass accuracy: <5 ppm) and analyzed by Openchrom 1.2.0,[31] custom Python 3.6 scripts utilizing Scipy(v.1.0.0),[32] and Pandas(v.0.22.0)[33] packages. All plots were generated with Plotly (Plotly Technologies Inc.).

***In vitro* enzyme assays of CsPT4-T.** yCAN10 and yCAN14 were inoculated to OD$_{600nm}$ 0.2 in 200 mL YPG medium and incubated for 20 h in baffled shake flasks (200 rpm). Harvested 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 (Emulsiflex C3, Avestin, INC., 60 bar, 10 min). Cell debris was removed by centrifugation (10,000 × g, 10 min, 4°C). Subsequently, the supernatant was subjected to ultracentrifugation (150,000 × g, 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 MgCl$_2$, 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 MgCl$_2$, pH 8.5) and substrate was added (500 µM OA, 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 AHF 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 OA 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 OA (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 non-linear regression using Scipy.[32] Kinetic studies were carried out in triplicate or quadruplicate.

Click-chemistry. yCAN40 was pre-grown in YPD, subsequently diluted to OD_{600nm} 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)/H_2O (55%:45% (v/v)) and filtered. An aliquot of 8 µL of freshly prepared copper-catalyzed azide-alkyne cycloaddition master mix (1.25 mM Tris[(1-benzyl-1H-1,2,3-triazol-4-yl)methyl]amine, 1.25 mM $CuSO_4$, 2.5 mM ascorbic acid, 2.5 mM azide-PEG3-biotin conjugate in DMSO/H_2O (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.ccbr.utoronto.ca/cgi-bin/hgGateway) and the Medicinal Plant Genomics Resource (http://medicinalplantgenomics.msu.edu). To select PTs for functional testing, published transcriptomes of *Cannabis* were mined using a BLAST search against CsPT1, HlPT1L and HlPT2.

Phylogenetic analysis. Phylogenetic analysis was performed using the predicted amino acid sequences of *Cannabis* candidate PTs and related PTs (Extended Data Table 4). Sequence alignments were generated using MUSCLE with default parameters.[34] Maximum likelihood analysis was conducted with the IQ-TREE web server (http://iqtree.cibiv.univie.ac.at/) and default parameters.[35] The resulting tree was visualized using ETE3.[36]

Statistics and Reproducibility. If not indicated otherwise, all experiments were conducted with n = 3 biological replicates. When representative traces are shown in the display items, all other traces were similar. In bar charts the mean of the individual measurements was used as measure of center and the standard deviation for error bars.

METHODS REFERENCES

31. Wenig, P. & Odermatt, J. OpenChrom: a cross-platform open source software for the mass spectrometric analysis of chromatographic data. *BMC Bioinformatics* **11,** 405 (2010).

32. Jones E, Oliphant E, Peterson P, et al. SciPy: Open Source Scientific Tools for Python. Available at: http://www.scipy.org/. (Accessed: 20th February 2018)

33. McKinney, W. Data Structures for Statistical Computing in Python,. *Proceedings of the 9th Python in Science Conference* 51–56 (2010).

34. Edgar, R. C. MUSCLE: multiple sequence alignment with high accuracy and high throughput. *Nucleic Acids Res.* **32,** 1792–1797 (2004).

35. 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,** W232–5 (2016).

36. Huerta-Cepas, J., Serra, F. & Bork, P. ETE 3: Reconstruction, Analysis, and Visualization of Phylogenomic Data. *Mol. Biol. Evol.* **33,** 1635–1638 (2016)

37. Lee, T. S. et al. BglBrick vectors and datasheets: A synthetic biology platform for gene expression. J. Biol. Eng. 5, 12 (2011).

EXTENDED DATA LEGENDS

Extended Data Figure 1 | OA 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 OA. An additional peak was observed for the byproduct HTAL (same m/z as OA), which has been reported previously. No Zymolase was added during extraction. Chromatography gradient for cannabinoid analogues was used. **b,** MS-spectrum of HTAL. Mass accuracy for observed m/z at a given retention time (RT) is reported in parts per million (ppm).

Extended Data Figure 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 (stds). We incubated boiled microsomal fraction (mic ΔT), soluble fraction (sol) and microsomal fraction (mic) of yCAN14 in the presence of 500 µM OA and 500 µM GPP for 1 h at 30°C and observed GOT activity only in the microsomal fraction. **b,** CsPT4-T OA kinetics. Using non-linear regression to fit the Michaelis-Menten kinetic model for varied OA (0.25 µM–0.56 mM) and constant GPP (1.67 mM) concentrations revealed a K_M (OA) = 6.73 ± 0.26 µM. (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 behavior for CsPT4-T for varied GPP (0.25 µM–1.67 mM) and constant OA (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 PT's. The numbers indicate the bootstrap value (%) from 1000 replications. Grey dashed lines depict branch offsets to accommodate labels. The scale bar shows the amino acid substitution ratio. PTs cluster by biosynthetic pathway. CsPT4 catalyzes CBGA production. The function of the other Cannabis PTs remains unknown.

Extended Data Figure 3 | *In vitro* activity of NphB and HlPT. 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 (stds). **a,** Purified NphB catalyzed the condensation of GPP and OA to CBGA when incubated with 5 mM OA and 5 mM GPP at room temperature for 24 h. The enzyme produced at least one other isomer of CBGA consistent with previous reports.[28] Boiling NphB (NphB ΔT) abolished activity. **b,** Microsomal fractions were prepared from yCAN21, yCAN22 and yCAN23 which expressed HlPT1L-T (1L-T), HlPT1L-T and HlPT2-T (1L-T/2-T), and HlPT2-T (2-T), respectively. Incubation with 5 mM OA and 5 mM GPP for 24 h at 30°C yielded CBGA as well as several isomers. CBGA production was not observed when incubating HlPT1L-T and HlPT2-T with their native substrates phlorisovalerophenone and dimethylallyl diphosphate (neg).

Extended Data Figure 4 | DA 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,** yCAN32 produced an additional compound of m/z-value corresponding to DA compared to a strain lacking the OA synthesis pathway (yCAN14). The retention time of DA is reduced relative to OA due to its reduced hydrophobicity conferred by the shorter 5-propyl sidechain. No Zymolase was added during extraction. Chromatography gradient for cannabinoid analogues was used. **b,** MS-spectrum of DA. Mass accuracy for observed m/z at a given retention time (RT) is reported in parts per million (ppm). **c,** Accordingly, the prenylated variant of DA, CBGVA, was also detected. **d,** MS-spectrum of CBGVA with mass accuracy reported in parts per million.

Extended Data Figure 5 | Production of cannabinoid analogues. All LC-MS chromatograms were selected for the theoretical m/z-values of the respective compounds of interest (Extended Data Table

3). When yCAN31 was grown in the presence of butanoic acid (II), nonanoic acid (III), 5-phenylpentanoic acid (IV), 6-phenylhexanoic acid (V) and 7-phenylheptanoic acid (VI) peaks corresponding to their respective OA and HTAL analogues were detected. Detected analogue peaks shifted in retention time depending on their change in hydrophobicity relative to OA and CBGA standards (stds). Feeding 1 mM hexanoic acid to yCAN30 yielded neither OA nor CBGA (neg (I)). Medium supplementation with butanoic acid did not increase production of DA or CBGVA over baseline levels. Feeding phenyl substituted acids yielded either only the corresponding HTAL analogues or OA analogues.

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

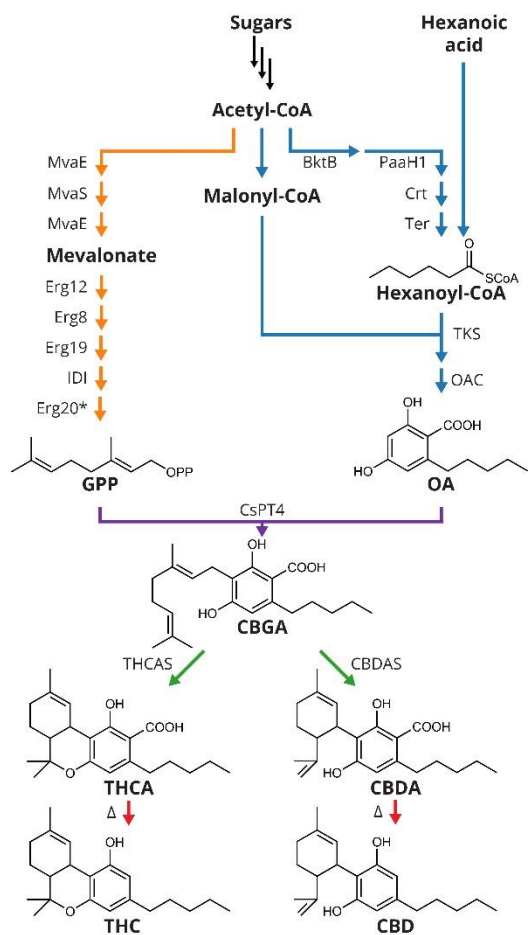**Extended Data Table 1 | Strains and plasmids used in this study.**

Extended Data Table 2 | Fatty acids tested for cannabinoid analogue production. y: detected OA/CBGA/THCA analogue. n: no detection of OA/CBGA/THCA analogue. *: feeding fatty acid did not increase OA/CBGA/THCA baseline levels. ?: possibly OA or HTAL analogue. -: not tested.
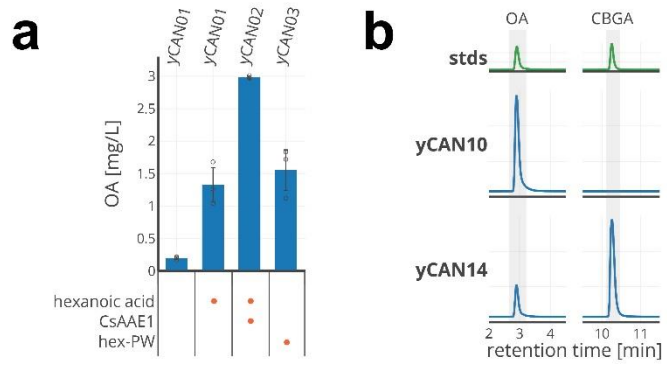
Extended Data Table 3 | m/z-values of compounds analyzed in this study. All values are for the $[M-H]^-$ ion.

Extended Data Table 4 | Accession numbers of (putative) PTs tested for GOT activity and PTs used for phylogenetic analysis.
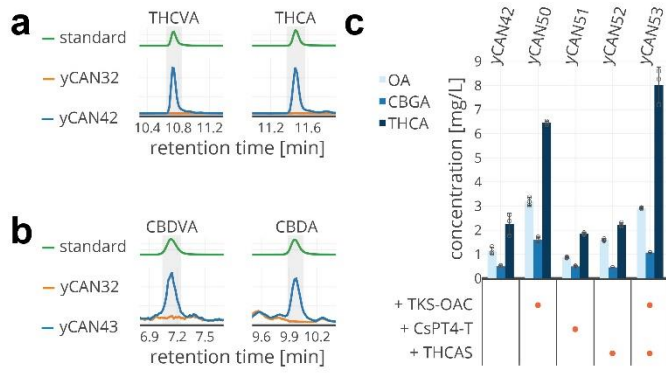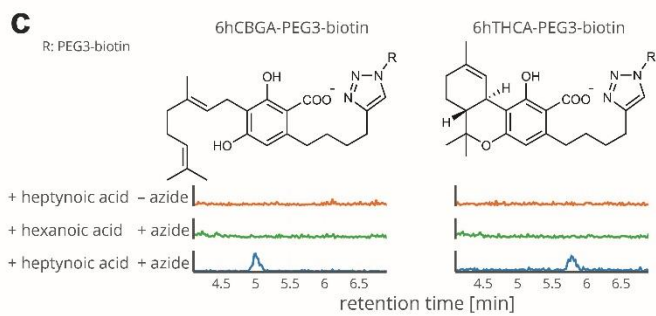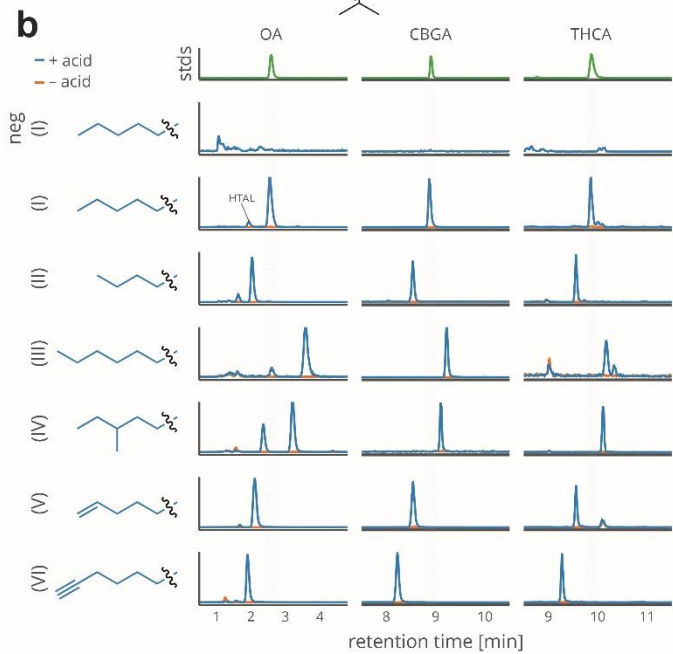
DATA AVAILABILITY STATEMENT

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). Custom Python 3.6 scripts for data analysis are available from the authors upon reasonable request.

a

standard — CEN.PK — yCAN03

OA

HTAL

retention time [min]
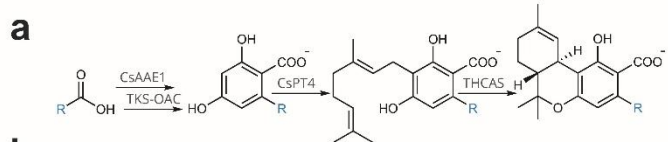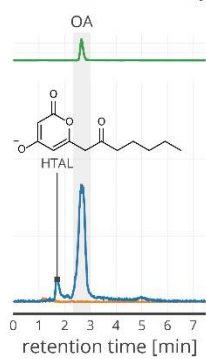
b

RT: 1.705 min
m/z$_{theoretical}$: 223.09758
accuracy: 1.30 ppm

223.09729

m/z

a

— standard — yCAN14 — yCAN32

std
DA

OA

retention time [min]

b

RT: 1.698 min
m/z$_{theoretical}$: 195.06628
accuracy: 3.38 ppm

195.06694

m/z

c

— standard — yCAN14 — yCAN32

std
CBGVA

CBGA

retention time [min]

d

RT: 8.032 min
m/z$_{theoretical}$: 331.19148
accuracy: 0.91 ppm

331.19178

m/z

a

			OA		CBGA		THCA

(I) RT: 2.555 min
m/z_theoretical: 223.09758
accuracy: 0.54 ppm
223.0977

RT: 8.888 min
m/z_theoretical: 359.22278
accuracy: 2.98 ppm
359.22171

RT: 9.858 min
m/z_theoretical: 357.20713
accuracy: 0.14 ppm
357.20708

(II) RT: 2.033 min
m/z_theoretical: 209.08193
accuracy: 3.59 ppm
209.08268

RT: 8.559 min
m/z_theoretical: 345.20713
accuracy: 4.58 ppm
345.20871

RT: 9.557 min
m/z_theoretical: 343.19148
accuracy: 0.99 ppm
343.19182

(III) RT: 3.588 min
m/z_theoretical: 237.11323
accuracy: 1.39 ppm
237.11356

RT: 9.229 min
m/z_theoretical: 373.23843
accuracy: 1.77 ppm
373.23909

RT: 10.194 min
m/z_theoretical: 371.22278
accuracy: 2.83 ppm
371.22383

(IV) RT: 3.208 min
m/z_theoretical: 237.11323
accuracy: 3.80 ppm
237.11413

RT: 9.099 min
m/z_theoretical: 373.23843
accuracy: 3.00 ppm
373.23955

RT: 10.100 min
m/z_theoretical: 371.22278
accuracy: 3.04 ppm
371.22165

(V) RT: 2.113 min
m/z_theoretical: 221.08193
accuracy: 4.79 ppm
221.08299

RT: 8.389-8.716 min
m/z_theoretical: 357.20713
accuracy: 0.62 ppm
357.20691

RT: 9.562 min
m/z_theoretical: 355.19148
accuracy: 1.41 ppm
355.19198

(VI) RT: 1.893 min
m/z_theoretical: 233.08193
accuracy: 3.43 ppm
233.08113

RT: 8.227 min
m/z_theoretical: 369.20713
accuracy: 0.79 ppm
369.20684

RT: 9.281 min
m/z_theoretical: 367.19148
accuracy: 0.05 ppm
367.19146

b

RT: 4.998 min
m/z_theoretical: 813.42262
accuracy: 0.49 ppm
813.42302

RT: 5.780 min
m/z_theoretical: 811.40697
accuracy: 1.08 ppm
811.40785

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-*HlPT1L-T*-TDH1t}	This work	SBa_001942
yCAN19	yCAN10 {1114a::GAL1p-*HlPT2-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-Cs*TKS*-*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-Cs*TKS*-*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)	*fhuA2 [lon] ompT gal (λ DE3) [dcm] ΔhsdS λ DE3 = λ sBamHlo ΔEcoRI-B int::(lacI::PlacUV5::T7 gene1) i21 Δnin5*	New England Biolabs, Inc.	
eCAN01	BL21 pCAN01	This work	SBa_001956
pCAN01	pBbE7a Amp[R] T7p-*NphB*-6xHis-DBLt	This work	SBa_001954
pCAN02	E. coli plasmid with p15A origin and Amp selection marker	Ref. [37]	-
pCAN03	pESC-URA with GAL1p-*HlPT1L-T*-CYC1t	Ref. [21]	-
pCAN04	pESC-URA with GAL1p-*HlPT1L-T*-CYC1t-GAL10p-*HlPT2-T*-ADH1t	Ref. [21]	-
pCAN05	pESC-URA with GAL10p-*HlPT2-T*-ADH1t	This work	SBa_002030

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	-

Category	Name	Abbreviation	m/z (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	cannabigerovarinic 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-pentynyl)-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-oxtyl-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	Δ⁹-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
	Δ⁹-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 triacetic acid lactone	HTAL	223.09758	hexanoic acid

Enzyme	Organism of origin	GenBank Accession No.
CsPT1	*Cannabis*	BK010678
CsPT2	*Cannabis*	BK010679
CsPT3	*Cannabis*	BK010680
CsPT4	*Cannabis*	BK010648
CSPT5	*Cannabis*	BK010681
CsPT6	*Cannabis*	BK010682
CsPT7	*Cannabis*	BK010683
LePGT1	*Lithospermum erythrorhizon*	BAB84122
LePGT2	*Lithospermum erythrorhizon*	Q8W404.1
AtUbiA	*Arabidopsis thaliana*	NP_190750.1
AtVTE2	*Arabidopsis thaliana*	OAP10166.1
AtHPT1	*Arabidopsis thaliana*	NP849984
AtPPT1	*Arabidopsis thaliana*	Q93YP7.1
ZmVTE2-1	*Zea mays*	XP_008659042.1
ZmHPT	*Zea mays*	CAC18910.1
OsPPT1	*Oryza sativa*	BAE96574.1
OsHGGT	*Oryza sativa*	AAP43913.1
OsHPT	*Oryza sativa*	CAC18914.1
SfN8DT-1	*Sophora flavescens*	BAG12671
HlPT1L	*Humulus lupulus*	A0A0B5A051.1
HlPT1	*Humulus lupulus*	E5RP65.1
HlPT2	*Humulus lupulus*	A0A0B4ZTQ2.1
TaHGGT	*Triticum aestivum*	AAP43912.1
HvHGGT	*Hordeum vulgare*	AAP43911.1
GmHPT	*Glycine max*	CAC18917.1