



DATA NOTE

In silico discovery of terpenoid metabolism in *Cannabis sativa* [version 1; referees: 2 approved, 1 approved with reservations]

Luca Massimino

Molecular Oncology Unit, San Gerardo Hospital, Monza, Italy

v1 First published: 06 Feb 2017, 6:107 (doi: [10.12688/f1000research.10778.1](https://doi.org/10.12688/f1000research.10778.1))
Latest published: 06 Feb 2017, 6:107 (doi: [10.12688/f1000research.10778.1](https://doi.org/10.12688/f1000research.10778.1))

Abstract

Due to their efficacy, cannabis based therapies are currently being prescribed for the treatment of many different medical conditions. Interestingly, treatments based on the use of cannabis flowers or their derivatives have been shown to be very effective, while therapies based on drugs containing THC alone lack therapeutic value and lead to increased side effects, likely resulting from the absence of other pivotal entourage compounds found in the Phyto-complex. Among these compounds are terpenoids, which are not produced exclusively by cannabis plants, so other plant species must share many of the enzymes involved in their metabolism. In the present work, 23,630 transcripts from the canSat3 reference transcriptome were scanned for evolutionarily conserved protein domains and annotated in accordance with their predicted molecular functions. A total of 215 evolutionarily conserved genes encoding enzymes presumably involved in terpenoid metabolism are described, together with their expression profiles in different cannabis plant tissues at different developmental stages. The resource presented here will aid future investigations on terpenoid metabolism in *Cannabis sativa*.

Open Peer Review

Referee Status:

| | Invited Referees | | |
|--|------------------|------------|------------|
| | 1 | 2 | 3 |
| version 1 published 06 Feb 2017 | report | report | report |

- Akan Das**, University of Science and Technology, Meghalaya (USTM), India
- Jonathan E Page**, University of British Columbia, Canada
Judith K. Booth, University of British Columbia, Canada
- Meirong Jia**, Iowa State University, USA

Discuss this article

Comments (0)

Corresponding author: Luca Massimino (m4x1m1n1o@gmail.com)

Competing interests: No competing interests were disclosed.

How to cite this article: Massimino L. *In silico discovery of terpenoid metabolism in Cannabis sativa* [version 1; referees: 2 approved, 1 approved with reservations] *F1000Research* 2017, 6:107 (doi: [10.12688/f1000research.10778.1](https://doi.org/10.12688/f1000research.10778.1))

Copyright: © 2017 Massimino L. This is an open access article distributed under the terms of the [Creative Commons Attribution Licence](https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Grant information: The author(s) declared that no grants were involved in supporting this work.

First published: 06 Feb 2017, 6:107 (doi: [10.12688/f1000research.10778.1](https://doi.org/10.12688/f1000research.10778.1))

Introduction

Due to its astonishing efficacy¹, nowadays cannabis is prescribed by physicians for the treatment of neurological, psychiatric, immunological, cardiovascular, gastrointestinal, and oncological conditions²⁻⁷. Although therapies based on the use of cannabis flowers or their derivatives are recognized to be very effective, treatments centered on drugs containing Δ^9 -tetrahydrocannabinol (THC) alone lack efficacy and lead to increased side effects^{8,9}. This discrepancy seems to result from the absence of the synergistic effects of additional pivotal compounds found in the Phyto-complex, the so-called entourage effect¹⁰. Among these molecules are other cannabinoids and terpenoids, which are thought to play major roles in the modulation of THC¹¹.

Terpenes are small hydrocarbon (isoprenoid) molecules classified as either as monoterpenes, sesquiterpenes, diterpenes or carotenoids, depending on the number of isoprene units (C_3) used to synthesize them. Terpenoids are small lipids derived from terpenes, often accompanied by a strong odor useful for the plants to protect themselves against possible predators¹². Terpenoids not only have important functions when working in concert with cannabinoids; they have been widely investigated in many different plant species and are being exploited as anti-fungal, anti-bacterial, anti-oxidant, anti-inflammatory, anti-stress, anti-cancer and analgesic agents¹³⁻¹⁸. However, whilst the gene networks controlling the biosynthesis of cannabinoids and their precursors have been extensively studied¹⁹⁻²², the biosynthetic pathway of terpenoid molecules in *Cannabis sativa* is only recently being elucidated. Only two genes have been characterized, one encoding (-)-limonene synthase, the other (+)- α -pinene synthase²³, two enzymes responsible for the conversion of geranyl pyrophosphate into limonene and pinene, respectively^{24,25}. Remarkably, while cannabinoids are only found in cannabis plants, terpenoids are also produced by a variety of other plant species, so they must share many of the enzymes involved in their metabolism.

In the present work, evolutionarily conserved genes encoding enzymes predicted to be involved in terpenoid metabolism have been identified within the transcripts of the canSat3 reference transcriptome of *Cannabis sativa*²¹. Moreover, by taking advantage of available gene expression data²⁶, gene expression profiling of these enzymes was performed in cannabis plant tissue at different developmental stages. The data note presented here will provide researchers with a corollary of candidate genes that will considerably accelerate future investigations on terpenoid metabolism in *Cannabis sativa*.

Material and methods

Analysis of evolutionarily conserved transcripts

Cannabis sativa transcript sequences (n=23,630) taken from the canSat3 genome assembly (<http://genome.cabr.utoronto.ca/>)²¹ were annotated with Blast2GO 4.0.7²⁷ using NCBI blastx and InterProScan databases. Terpene metabolism related genes were selected if found to be present in datasets downstream of the “terpene metabolic process” gene ontology category from the AmiGO 2 repository (GO:0042214), including the carotene metabolic process (GO:0016119), ent-kaurene metabolic process (GO:0033331), ent-pimara-8(14),15-diene metabolic process (GO:1901539), isoprene metabolic process (GO:0043611), miltiradiene metabolic

process (GO:1901944), monoterpene metabolic process (GO:0043692), sesquiterpene metabolic process (GO:1903192), sesquiterpene metabolic process (GO:0051761), terpene biosynthetic process (GO:0046246), and terpene catabolic process (GO:0046247)²⁸.

Gene expression analysis

Gene expression profiles from cannabis plant tissue at different developmental stages were downloaded from the NCBI GEO repository (<https://www.ncbi.nlm.nih.gov/geo/>). Gene expression heatmaps and unsupervised hierarchical clustering were performed with GENE-E 3.0.213²⁹.

Results

Although the *Cannabis sativa* reference genome and transcriptome has been publicly released²¹, only a few genes have been characterized and surveyed for their molecular functions. To define the possible roles of these genes, 40,197 canSat3 transcript sequences were downloaded from the cannabis genome browser (<http://genome.cabr.utoronto.ca/>), translated *in silico*, and scanned for evolutionarily conserved protein domains for functional anno-

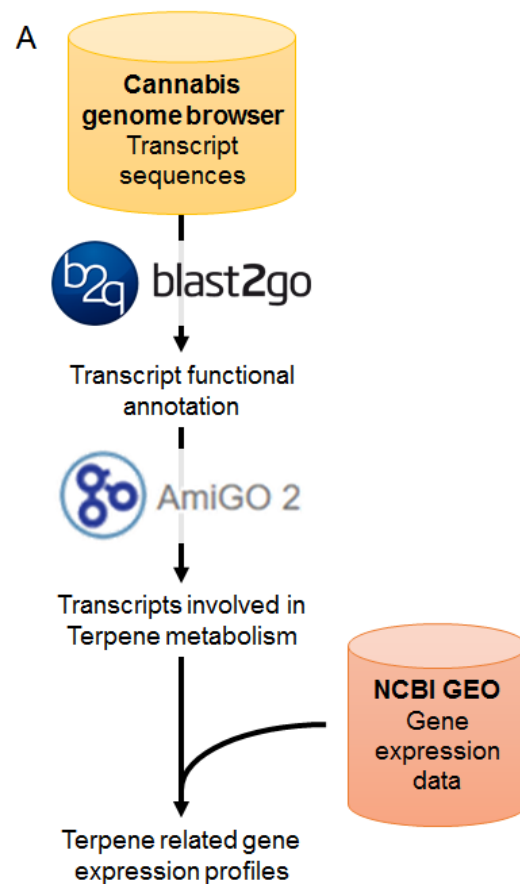


Figure 1. Workflow analysis. Schematic of bioinformatics pipeline utilized in this work. *Cannabis sativa* transcript sequences were taken from the canSat3 reference genome, functionally annotated with Blast2GO 4.0.7, filtered for terpenoid metabolism categories (AmiGO 2), and integrated with gene expression data downloaded from NCBI.

tation (Figure 1). To identify the genes presumably playing a role in terpenoid metabolism, annotated transcripts were filtered for gene ontology (GO) categories involved in terpene biosynthesis and catabolism using the AmiGO 2 reference database. A total of 288 transcripts representing 215 different genes were predicted to be involved in the metabolism of bisabolene, cadinene, carotene, copaene, ent-kaurene, farnesol, geraniol, germacrene,

lycopene, limonene, myrcene, phytoene, pinene, squalene, and others (Supplementary table 1). Functional characterization of this subset confirmed an enrichment for GO categories involved in different terpene biosynthetic and catabolic processes (Figure 2).

Terpenoids are produced by several plants species and in several types of plant tissue as defense against predators³⁰. Similar to other

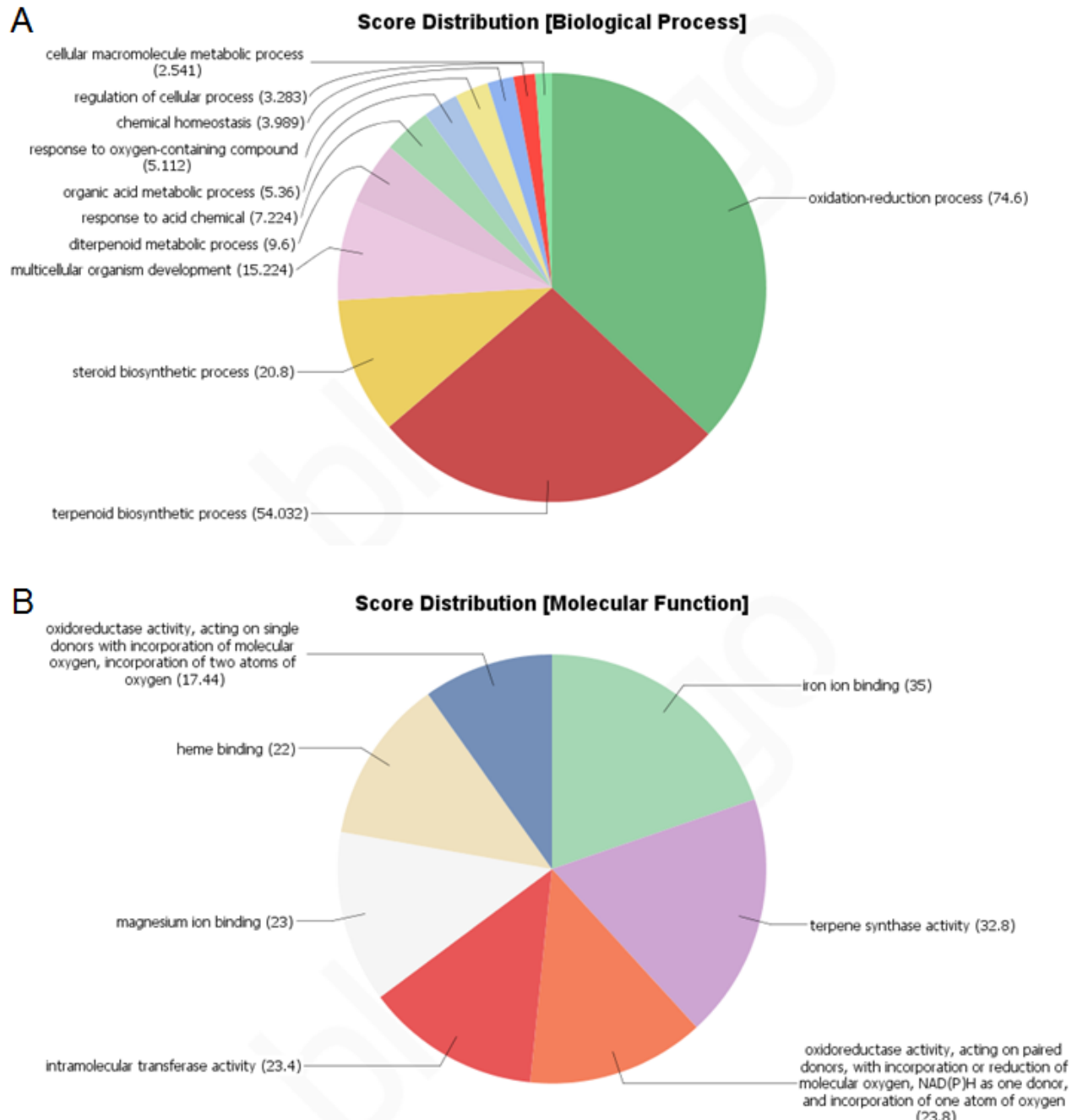


Figure 2. Gene ontology analysis of putative terpenoid metabolism related transcripts. Functional enrichment analysis of putative terpenoid metabolism related transcripts taken from the canSat3 reference genome. Enrichment for terpene biosynthesis and catabolism is shown for Biological Process (A) and Molecular function (B) Gene ontology categories.

biological compounds, their abundance directly correlates with the expression levels of the enzymes involved in their metabolism. To this end, gene expression analysis of genes likely to be involved in terpenoid metabolism was performed using previously published

datasets (Figure 3; Supplementary table 2)²⁶. Notably, unsupervised hierarchical clustering identified four gene clusters. Cluster 1 genes display high expression in roots and stems; cluster 3 genes in hemp flowers; cluster 4 genes in leaves and flowers. Cluster 2 genes

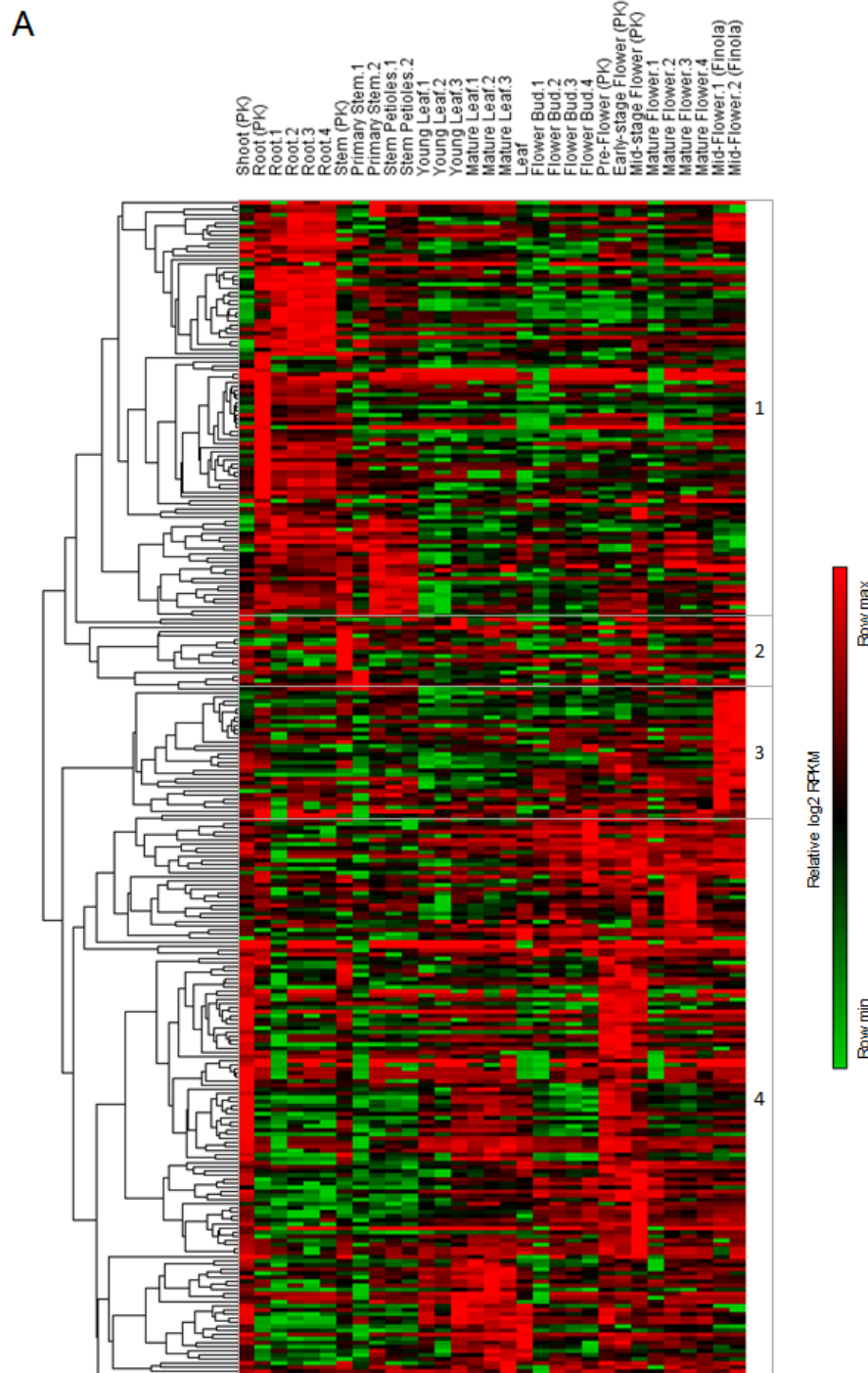


Figure 3. Gene expression analysis of terpenoid metabolism related genes. Heatmap showing relative expression values (log₂ RPKM) of putative terpenoid metabolism related genes from cannabis plant tissue taken at different developmental stages (shoot, root, stem, young and mature leaf, early-, mid- and mature-stage flower). Five gene clusters were defined in accordance to unsupervised hierarchical clustering.

were constitutively expressed in all tissues. These results highlight which enzymes are expressed by specific tissues and will provide a strong rationale for further investigations on the molecular basis of terpenoid metabolism.

Discussion

The active principles inside plants have been exploited by humans for centuries, with *Cannabis sativa* being one of the oldest ever used for medicinal purposes³¹. Surprisingly, contrary to whole plant extracts, medicinal products containing exclusively THC have been found to lack efficacy and lead to unbearable side effects^{8,9}. These results arise from the fact that these products lack other important co-factors typically found in the Phyto-complex, such as terpenoids and other cannabinoids¹⁰ that contribute to the synergistic effects seen with whole plant extracts.

While genes involved in cannabinoid biosynthesis have been widely investigated^{19–22}, the gene network controlling terpenoid metabolism is only recently being elucidated, with genes encoding (-)-limonene synthase and (+)- α -pinene synthase being the only two characterized²³. To this end, *Cannabis sativa* transcripts²¹ have been scanned for evolutionarily conserved protein domains and annotated according to their presumptive molecular function. As a result, 215 evolutionarily conserved genes were predicted to be involved in terpenoid metabolism. Furthermore, *in silico* gene expression profiling²⁶ of these enzymes in cannabis plant tissue at different developmental stages highlighted different gene clusters with peculiar expression patterns. For instance, cluster 3 genes (Figure 3) displayed high expression specifically in hemp flowers, which could be of great interest as different cannabis strains harbor different entourage effects.

Since the current cannabis reference transcriptome is still at preliminary stages²¹, it is very likely that false negatives have caused important transcripts to still be missing. For example, the two genes

encoding for (-)-limonene synthase and (+)- α -pinene synthase²³ align on the same transcript predicted to encode for Myrcene synthase (PK25781.1 in Supplementary table 1), and therefore cannot be discriminated. Unfortunately, to overcome this issue at whole genome level we need the complete version of the reference transcriptome to be available. Until that time, researchers are forced to validate single transcripts with classic low-throughput technology, such as molecular cloning followed by Sanger sequencing.

Nevertheless, the data presented here will ease future investigations on terpenoid metabolism in *Cannabis sativa* by providing researchers with a collection of candidate genes. For instance, one of these genes was predicted to encode for β -bisabolene synthase (PK05069.1 in Supplementary table 1). Bisabolene is being used as an antimicrobial agent³², as well as a biofuel³³. However, prior to this report nothing was known about the gene network controlling its metabolism in *Cannabis sativa*. As soon as future studies will integrate gene expression data with chemical analysis, the complete molecular scenario underlying terpenoid metabolism will be revealed.

Data availability

Processed gene expression data can be found in the NCBI GEO repository (<https://www.ncbi.nlm.nih.gov/geo/>) with accession number GSE93201.

Competing interests

No competing interests were disclosed.

Grant information

The author(s) declared that no grants were involved in supporting this work.

Supplementary material

Supplementary table 1. Evolutionarily conserved terpenoid metabolism transcripts

List of putative terpenoid metabolism genes obtained with Blast2GO.

[Click here to access the data.](#)

Supplementary table 2. Terpenoid metabolism gene profiling in different tissues and developmental stages

Gene expression matrix of predicted terpenoid metabolism genes. Expression units are expressed in RPKM.

[Click here to access the data.](#)

References

1. Committee on the Health Effects of Marijuana: **The Health Effects of Cannabis and Cannabinoids**. (National Academies Press). 2017. [Publisher Full Text](#)
2. Hosking R, Zajicek J: **Pharmacology: Cannabis in neurology—a potted review**. *Nat Rev Neurol*. 2014; **10**(8): 429–30. [PubMed Abstract](#) | [Publisher Full Text](#)
3. Curran HV, Freeman TP, Mokrysz C, *et al.*: **Keep off the grass? Cannabis, cognition and addiction**. *Nat Rev Neurosci*. 2016; **17**(5): 293–306. [PubMed Abstract](#) | [Publisher Full Text](#)
4. Klein TW: **Cannabinoid-based drugs as anti-inflammatory therapeutics**. *Nat Rev Immunol*. 2005; **5**(5): 400–11. [PubMed Abstract](#) | [Publisher Full Text](#)

5. Di Marzo V, Després JP: **CB1 antagonists for obesity—what lessons have we learned from rimonabant?** *Nat Rev Endocrinol.* 2009; 5(11): 633–8.
[PubMed Abstract](#) | [Publisher Full Text](#)
6. Gerich ME, Isfort RW, Brimhall B, *et al.*: **Medical marijuana for digestive disorders: high time to prescribe?** *Am J Gastroenterol.* 2015; 110(2): 208–14.
[PubMed Abstract](#) | [Publisher Full Text](#)
7. Swami M: **Cannabis and cancer link.** *Nat Rev Cancer.* 2009; 9: 148.
[Publisher Full Text](#)
8. Ben Amar M: **Cannabinoids in medicine: A review of their therapeutic potential.** *J Ethnopharmacol.* 2006; 105(1–2): 1–25.
[PubMed Abstract](#) | [Publisher Full Text](#)
9. Kowal MA, Hazekamp A, Grotenhermen F: **Review on clinical studies with cannabis and cannabinoids 2010–2014.** *Cannabinoids.* 2016; 11(special issue): 1–18.
[Reference Source](#)
10. Russo EB: **Taming THC: potential cannabis synergy and phytocannabinoid-terpenoid entourage effects.** *Br J Pharmacol.* 2011; 163(7): 1344–1364.
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
11. Andre CM, Hausman JF, Guerriero G: **Cannabis sativa: The Plant of the Thousand and One Molecules.** *Front Plant Sci.* 2016; 7: 19.
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
12. Pichersky E, Noel JP, Dudareva N: **Biosynthesis of Plant Volatiles: Nature's Diversity and Ingenuity.** *Science.* 2006; 311(5762): 808–811.
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
13. Khosla C, Keasling JD: **Metabolic engineering for drug discovery and development.** *Nat Rev Drug Discov.* 2003; 2(12): 1019–1025.
[PubMed Abstract](#) | [Publisher Full Text](#)
14. Hakkim FL, Al-Buloshi M, Al-Sabahi J: **Frankincense derived heavy terpene cocktail boosting breast cancer cell (MDA-MB-231) death in vitro.** *Asian Pac J Trop Biomed.* 2015; 5(10): 824–828.
[Publisher Full Text](#)
15. Smanski MJ, Zhou H, Claesen J, *et al.*: **Synthetic biology to access and expand nature's chemical diversity.** *Nat Rev Microbiol.* 2016; 14(3): 135–149.
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
16. d'Alessio PA, Bisson JF, Béné MC: **Anti-stress effects of d-limonene and its metabolite perillyl alcohol.** *Rejuvenation Res.* 2014; 17(2): 145–149.
[PubMed Abstract](#) | [Publisher Full Text](#)
17. Klauke AL, Racz I, Pradier B, *et al.*: **The cannabinoid CB₁ receptor-selective phytocannabinoid beta-caryophyllene exerts analgesic effects in mouse models of inflammatory and neuropathic pain.** *Eur Neuropsychopharmacol.* 2014; 24(4): 608–620.
[PubMed Abstract](#) | [Publisher Full Text](#)
18. Bonamin F, Moraes TM, Dos Santos RC, *et al.*: **The effect of a minor constituent of essential oil from Citrus aurantium: the role of β-myrcene in preventing peptic ulcer disease.** *Chem Biol Interact.* 2014; 212: 11–19.
[PubMed Abstract](#) | [Publisher Full Text](#)
19. Sirikantaramas S, Morimoto S, Shoyama Y, *et al.*: **The gene controlling marijuana psychoactivity: molecular cloning and heterologous expression of Delta¹-tetrahydrocannabinolic acid synthase from Cannabis sativa L.** *J Biol Chem.* 2004; 279(38): 39767–74.
[PubMed Abstract](#) | [Publisher Full Text](#)
20. Taura F, Tanaka S, Taguchi C, *et al.*: **Characterization of olivetol synthase, a polyketide synthase putatively involved in cannabinoid biosynthetic pathway.** *FEBS Lett.* 2009; 583(12): 2061–2066.
[PubMed Abstract](#) | [Publisher Full Text](#)
21. van Bakel H, Stout JM, Cote AG, *et al.*: **The draft genome and transcriptome of Cannabis sativa.** *Genome Biol.* 2011; 12(10): R102.
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
22. Gagne SJ, Stout JM, Liu E, *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.* 2012; 109(31): 12811–6.
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
23. Günnewich N, Page JE, Köllner TG, *et al.*: **Functional expression and characterization of trichome-specific (-)-limonene synthase and (+)-α-pinene synthase from Cannabis sativa.** *Nat Prod Commun.* 2007; 2: 223–232.
24. Colby SM, Alonso WR, Katahira EJ, *et al.*: **4S-limonene synthase from the oil glands of spearmint (Mentha spicata). cDNA isolation, characterization, and bacterial expression of the catalytically active monoterpene cyclase.** *J Biol Chem.* 1993; 268(31): 23016–24.
[PubMed Abstract](#)
25. Bohlmann J, Meyer-Gauen G, Croteau R: **Plant terpenoid synthases: molecular biology and phylogenetic analysis.** *Proc Natl Acad Sci U S A.* 1998; 95(8): 4126–33.
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
26. Massimino L: **In silico gene expression profiling in Cannabis sativa.** *F1000Research.* 2017; 6: 69.
[Publisher Full Text](#)
27. Conesa A, Götz S, García-Gómez JM, *et al.*: **Blast2GO: A universal tool for annotation, visualization and analysis in functional genomics research.** *Bioinformatics.* 2005; 21(18): 3674–3676.
[PubMed Abstract](#) | [Publisher Full Text](#)
28. Carbon S, Ireland A, Mungall CJ, *et al.*: **AmiGO: Online access to ontology and annotation data.** *Bioinformatics.* 2009; 25(2): 288–289.
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
29. **GENE-E.** Cambridge (MA): The Broad Institute of MIT and Harvard.
[Reference Source](#)
30. Martin DM, Gershenzon J, Bohlmann J: **Induction of volatile terpene biosynthesis and diurnal emission by methyl jasmonate in foliage of Norway spruce.** *Plant Physiol.* 2003; 132(3): 1586–1599.
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
31. Pain S: **A potted history.** *Nature.* 2015; 525(7570): S10–1.
[PubMed Abstract](#) | [Publisher Full Text](#)
32. Satyal P, Setzer WN: **Chemotyping and Determination of Antimicrobial, Insecticidal, and Cytotoxic Properties of Wild-Grown Cannabis sativa from Nepal.** 2014; 3(1–4): 9–16.
[Reference Source](#)
33. Peralta-Yahya PP, Ouellet M, Chan R, *et al.*: **Identification and microbial production of a terpene-based advanced biofuel.** *Nat Commun.* 2011; 2: 483.
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)

Open Peer Review

Current Referee Status:



Version 1

Referee Report 19 June 2017

doi:10.5256/f1000research.11622.r23344



Meirong Jia

Roy J. Carver Department of Biochemistry, Biophysics & Molecular Biology, Iowa State University, Ames, IA, USA

The bioinformatics study here has reported 215 genes that are putatively involved in terpenoid biosynthesis in *Cannabis sativa* using available genome and transcriptome resources. This knowledge would predictably accelerate the terpenoid investigations in this specific species, or even broader. The research approaches are clear and sound. The writing is generally logical and organized. So I would recommend the publication of this work as a data note.

I would suggest the author to make the following minor changes.

1. The author might wish to clarify throughout the text as to the original “transcriptome” (or “genome”) resources they have used for initial study. For instance, adding the specific link of that file name.
2. In the caption of figure 3, the author claimed “five gene clusters were defined” while in the text, it was “four gene clusters”, the author needs to correct this. Also, it is confused here that some genes clustered are expressed in “Finola” because from the previous text in “introduction” “para 3”, it was said the study was conducted “within the transcripts of the canSat3 reference transcriptome of Cannabis sativa²¹”. Please clarify this.
3. The 215 gene candidates for terpenoids have been further predicted to express distinctively in tissues and developmental stages. While it is not easy to specifically check the expression pattern of each gene from the currently provided data. It would be great if possible that the author could add this piece of information, which would be valuable for future characterization of these genes. An example study showing the expression pattern of terpene genes would be found in the paper by Wang, *et al.* 2016¹.

References

1. Wang Q, Jia M, Huh JH, Muchlinski A, Peters RJ, Tholl D: Identification of a Dolabellane Type Diterpene Synthase and other Root-Expressed Diterpene Synthases in Arabidopsis. *Front Plant Sci.* 2016; 7: 1761 [PubMed Abstract](#) | [Publisher Full Text](#)

Is the rationale for creating the dataset(s) clearly described?

Yes

Are the protocols appropriate and is the work technically sound?

Yes

Are sufficient details of methods and materials provided to allow replication by others?

Yes

Are the datasets clearly presented in a useable and accessible format?

Partly

Competing Interests: No competing interests were disclosed.**I have read this submission. I believe that I have an appropriate level of expertise to confirm that it is of an acceptable scientific standard.**

Referee Report 09 June 2017

doi:10.5256/f1000research.11622.r23373

**Jonathan E Page**¹, **Judith K. Booth**²¹ Botany Department, University of British Columbia, Vancouver, BC, Canada² Michael Smith Laboratories, University of British Columbia, Vancouver, BC, Canada

The present study is appropriate as a data note, and may provide a useful shortcut for future studies on terpenoid biosynthesis in cannabis. The author identifies transcripts putatively involved in terpenoid biosynthesis, including processes of both primary and specialized metabolism. The rationale is well explained, and the methods are technically sound.

Major changes:

- More information is required in the Materials and Methods for the study to be replicable. Two versions of the canSat3 transcriptome are available, and the authors should specify whether they used the full or representative transcript set. The authors must also provide the NCBI accession numbers for the GEO dataset they used.
- There is an unsupported statement in the results section, which reads: "Similar to other biological compounds, their abundance directly correlates with the expression levels of the enzymes involved in their metabolism." In a complex system, this is not obvious and requires textual support.

Minor changes:

- The Introduction generally provides a good rationale for the study. However, the definitions provided for 'terpene' and 'terpenoid' in the second sentence of the second paragraph do not reflect the definitions used in reference 12, nor are they the traditional definitions for those terms.
- References 24 and 25 address enzymes characterized in other organisms, and so do not support the statement regarding their biosynthetic activities in cannabis. That information is provided in reference 23.
- The final statement of the second introductory paragraph is unsupported and requires citation.
- Figure 2 caption states that the transcripts were taken from the reference genome, but the Materials and Methods indicate that a transcriptome was used.
- Figure 3 caption insufficiently explains x axis labels. The caption should define 'PK' and 'Finola'. Otherwise, this figure shows an interesting and valuable result.

Readers should note the specific products of terpene synthases and other enzymes in specialized metabolism are currently difficult to predict using *in silico* methods. A recent paper (Booth et al., 2017¹) includes biochemical and phylogenetic analysis of some of the candidate genes highlighted here, and may be of interest to readers.

References

1. Booth JK, Page JE, Bohlmann J: Terpene synthases from *Cannabis sativa*. *PLoS One*. 2017; **12** (3): e0173911 [PubMed Abstract](#) | [Publisher Full Text](#)

Is the rationale for creating the dataset(s) clearly described?

Partly

Are the protocols appropriate and is the work technically sound?

Yes

Are sufficient details of methods and materials provided to allow replication by others?

Partly

Are the datasets clearly presented in a useable and accessible format?

Yes

Competing Interests: No competing interests were disclosed.

Referee Expertise: Plant metabolism

We have read this submission. We believe that we have an appropriate level of expertise to confirm that it is of an acceptable scientific standard, however we have significant reservations, as outlined above.

Referee Report 10 March 2017

doi:10.5256/f1000research.11622.r19961



Akan Das

Department of Applied Biology, University of Science and Technology, Meghalaya (USTM), Ri-Bhoi, Meghalaya, India

In the present work, 23,630 transcripts were scanned for evolutionarily conserved protein domains and annotated using Gene Ontology analysis. A total of 215 evolutionarily conserved genes encoding enzymes presumably involved in terpenoid metabolism are described on the basis of Gene Expression profiles of NCBI GEO repository. The identification of candidate genes on terpenoid metabolism in *Cannabis sativa* will ease the future investigations on terpenoid metabolism pathways.

It is a small piece of work on bioinformatics analysis of transcripts which present useful information to the plant researcher. Hence, the article can be recommended to be published as a Data Note. I suggest the author make a minor change in the title to be "*In silico* discovery of terpenoid metabolism associated transcripts in *Cannabis sativa*" and that the discussion, rather than describing general things, should be modified to discuss the findings in the work.

Competing Interests: No competing interests were disclosed.

I have read this submission. I believe that I have an appropriate level of expertise to confirm that it is of an acceptable scientific standard.
