TRICHOME SPECIFIC PROMOTERS FOR THE MANIPULATION OF CANNABINOIDS AND OTHER COMPOUNDS IN GLANDULAR TRICHOMES

CROSS-REFERENCE TO RELATED APPLICATION

[0001] The present application claims priority to U.S. Provisional Patent Application No. 62/397,212, filed on September 20, 2016, the contents of which are hereby incorporated by reference in their entirety.

TECHNICAL FIELD

[0002] The present technology relates generally to trichome specific promoters of cannabinoid biosynthesis enzyme genes from *Cannabis*, nucleotide sequences of the trichome specific promoters, and uses of the promoters for modulating cannabinoid production or for modulating other trichome specific production of biochemicals in organisms. The present technology also relates to transgenic cells and organisms, including plant cells and plants, comprising the trichome specific promoters.

BACKGROUND

[0003] The following description is provided to assist the understanding of the reader. None of the information provided or references cited is admitted to be prior art.

[0004] Plant trichomes are epidermal protuberances, including branched and unbranched hairs, vesicles, hooks, spines, and stinging hairs covering the leaves, bracts, and stems. There are two major classes of trichomes, which may be distinguished on the basis of their capacity to produce and secrete or store secondary metabolites, namely glandular trichomes and non-glandular trichomes. Non-glandular trichomes exhibit low metabolic activity and provide protection to the plant mainly through physical means. By contrast, glandular trichomes, which are present on the foliage of many plant species including some solanaceous species (*e.g.*, tobacco, tomato) and also cannabis, are highly metabolically active and accumulate metabolites, which can represent up to 10-15% of the leaf dry weight (Wagner et al., *Ann. Bot.* 93:3-11 (2004)). Glandular trichomes are capable of secreting (or storing) secondary metabolites as a defense mechanism.

[0005] *Cannabis sativa* L. (cannabis, hemp, marijuana), an annual herb that has been cultivated for thousands of years, contains a unique set of secondary metabolites called cannabinoids, which constitute a group of terpenophenolics. Cannabinoids are primarily synthesized and accumulate in glandular trichomes that are present at high densities on female flowers and at lower densities on male flowers of *C. sativa* plants. The accumulation of cannabinoids in the storage cavity of trichomes is beneficial to the plant as cannabinoids are known to be cytotoxic to other plant cells and have been shown to induce apoptosis in both hemp and tobacco cell suspension cultures (Sirikantaramas et al., *Plant Cell Physiol.* 46:1578-1582 (2005)). Cannabinoids are formed by a three-step biosynthetic process: polyketide formation, aromatic prenylation, and cyclization (**FIG. 1**). The cannabinoid pathway is supplied by hexanoyl-CoA, the formation of which is catalyzed by hexanoyl-CoA synthetase. The first enzymatic step in cannabinoid biosynthesis is the formation of 3,5,7-trioxododecanoyl-CoA by a tetraketide synthase enzyme (TKS), termed olivetolic acid synthase (OLS1). The second enzymatic step in cannabinoid biosynthesis is the formation of olivetolic acid by olivetolic acid cyclase (OAC). The next step is the prenylation of olivetolic acid to form cannabigerolic acid (CBGA) by the aromatic prenyltransferase. CBGA is a central branch-point intermediate for the biosynthesis of the different major classes of cannabinoids. Alternative cyclization of the prenyl side-chain of CBGA yields Δ^9tetrahydrocannabinolic acid (THCA) or its isomers cannabidiolic acid (CBDA) or cannabichromenic acid (CBCA). THCA and CBDA are later decarboxylated by a non-enzymatic reaction during storage or smoking to yield Δ^9-tetrahydrocannabinol (THC) or cannabidiol (CBD), respectively (**FIG. 1**).

[0006] Cannabinoids are valuable plant-derived natural products. Cannabis preparations, such as marijuana and hashish, have been used for centuries for their well-known psychoactive effects. Cannabinoids have attracted a renewed interest for medical applications due to their ability to act through mammalian cannabinoid receptors. Major cannabinoids include Δ^9-tetrahydrocannabinol (THC), the compound responsible for the psychoactive and therapeutic effects of marijuana consumption, and cannabidiol (CBD), which has neuroprotective properties. (Gaoni & Mechoulam, *J. Am. Chem. Soc.* 86:1646-1647 (1964); Mechoulam et al., *J. Clin. Pharmacol.* 42:11S-19S (2002)). THCA is the major cannabinoid in drug strains of cannabis while CBDA is the predominant cannabinoid in hemp forms grown for fiber or seed. Cannabinoids are currently being explored for therapeutic

purposes, including the treatment of chronic pain, nausea, the control of spasticity and tremor in patients suffering from multiple sclerosis or epilepsy, as well as a therapy for arthritis. The possibility to direct cannabinoid production in trichomes avoids interference in the plants' metabolic pathways and performance. Accordingly, there is a need to identify trichome specific promoters to modulate the synthesis of cannabinoids in organisms including transgenic plants, transgenic cells, and derivatives thereof, which allow for targeting of gene expression specifically in trichomes.

SUMMARY

[0007] Disclosed herein are trichome specific promoters and uses of these promoters for directing the expression of coding nucleic acid sequences in plant trichomes.

[0008] In one aspect, the present disclosure provides a synthetic DNA molecule comprising a nucleotide sequence selected from the group consisting of: (a) a nucleotide sequence set forth in any one of SEQ ID NOs: 1-3, 5, 6, 8, 9, 11-14, 16-19, 21-26, 28, 29, or 31-33; and (b) a nucleotide sequence that is at least about 80% identical to the nucleotide sequence of any one of SEQ ID NOs: 1-3, 5, 6, 8, 9, 11-14, 16-19, 21-26, 28, 29, or 31-33, and which encodes a promoter having plant trichome gland specific transcriptional activity, wherein the nucleotide sequence is operably linked to a heterologous nucleic acid.

[0009] In some embodiments, the present disclosure provides an expression vector comprising the synthetic DNA molecule operably linked to one or more nucleic acid sequences encoding a polypeptide.

[0010] In some embodiments, the present disclosure provides a genetically engineered host cell comprising the expression vector. In some embodiments, the genetically engineered host cell is a *Cannabis sativa* cell. In some embodiments, the genetically engineered host cell is a *Nicotiana tabacum* cell.

[0011] In some embodiments, the present disclosure provides a genetically engineered plant comprising a cell comprising a chimeric nucleic acid construct comprising the synthetic DNA molecule. In some embodiments, the genetically engineered plant belongs to the family *Solanacea*. In some embodiments, the engineered *Solanacea* plant is an *N. tabacum* plant. In some embodiments, the genetically engineered plant belongs to the family *Cannabaceae*. In

some embodiments, the engineered *Cannabaceae* plant is a *C. sativa* plant. In some embodiments, the present disclosure provides seeds from the genetically engineered plant, wherein the seeds comprise the chimeric nucleic acid construct.

[0012] In some embodiments, the present disclosure provides a synthetic DNA molecule with the nucleotide sequence as set forth in SEQ ID NO: 1. In some embodiments, the present disclosure provides a synthetic DNA molecule having a nucleotide sequence that is at least about 80% identical to the nucleotide sequence of SEQ ID NO: 1, and which encodes a promoter that has plant trichome gland specific transcriptional activity.

[0013] In some embodiments, the present disclosure provides a synthetic DNA molecule with the nucleotide sequence as set forth in SEQ ID NO: 2. In some embodiments, the present disclosure provides a synthetic DNA molecule having a nucleotide sequence that is at least about 80% identical to the nucleotide sequence of SEQ ID NO: 2, and which encodes a promoter that has plant trichome gland specific transcriptional activity.

[0014] In some embodiments, the present disclosure provides a synthetic DNA molecule with the nucleotide sequence as set forth in SEQ ID NO: 3. In some embodiments, the present disclosure provides a synthetic DNA molecule having a nucleotide sequence that is at least about 80% identical to the nucleotide sequence of SEQ ID NO: 3, and which encodes a promoter that has plant trichome gland specific transcriptional activity.

[0015] In some embodiments, the present disclosure provides a synthetic DNA molecule with the nucleotide sequence as set forth in SEQ ID NO: 5. In some embodiments, the present disclosure provides a synthetic DNA molecule having a nucleotide sequence that is at least about 80% identical to the nucleotide sequence of SEQ ID NO: 5, and which encodes a promoter that has plant trichome gland specific transcriptional activity.

[0016] In some embodiments, the present disclosure provides a synthetic DNA molecule with the nucleotide sequence as set forth in SEQ ID NO: 6. In some embodiments, the present disclosure provides a synthetic DNA molecule having a nucleotide sequence that is at least about 80% identical to the nucleotide sequence of SEQ ID NO: 6, and which encodes a promoter that has plant trichome gland specific transcriptional activity.

[0017] In some embodiments, the present disclosure provides a synthetic DNA molecule with the nucleotide sequence as set forth in SEQ ID NO: 8. In some embodiments, the present disclosure provides a synthetic DNA molecule having a nucleotide sequence that is at least about 80% identical to the nucleotide sequence of SEQ ID NO: 8, and which encodes a promoter that has plant trichome gland specific transcriptional activity.

[0018] In some embodiments, the present disclosure provides a synthetic DNA molecule with the nucleotide sequence as set forth in SEQ ID NO: 9. In some embodiments, the present disclosure provides a synthetic DNA molecule having a nucleotide sequence that is at least about 80% identical to the nucleotide sequence of SEQ ID NO: 9, and which encodes a promoter that has plant trichome gland specific transcriptional activity.

[0019] In some embodiments, the present disclosure provides a synthetic DNA molecule with the nucleotide sequence as set forth in SEQ ID NO: 11. In some embodiments, the present disclosure provides a synthetic DNA molecule having a nucleotide sequence that is at least about 80% identical to the nucleotide sequence of SEQ ID NO: 11, and which encodes a promoter that has plant trichome gland specific transcriptional activity.

[0020] In some embodiments, the present disclosure provides a synthetic DNA molecule with the nucleotide sequence as set forth in SEQ ID NO: 12. In some embodiments, the present disclosure provides a synthetic DNA molecule having a nucleotide sequence that is at least about 80% identical to the nucleotide sequence of SEQ ID NO: 12, and which encodes a promoter that has plant trichome gland specific transcriptional activity.

[0021] In some embodiments, the present disclosure provides a synthetic DNA molecule with the nucleotide sequence as set forth in SEQ ID NO: 13. In some embodiments, the present disclosure provides a synthetic DNA molecule having a nucleotide sequence that is at least about 80% identical to the nucleotide sequence of SEQ ID NO: 13, and which encodes a promoter that has plant trichome gland specific transcriptional activity.

[0022] In some embodiments, the present disclosure provides a synthetic DNA molecule with the nucleotide sequence as set forth in SEQ ID NO: 14. In some embodiments, the present disclosure provides a synthetic DNA molecule having a nucleotide sequence that is at least about 80% identical to the nucleotide sequence of SEQ ID NO: 14, and which encodes a promoter that has plant trichome gland specific transcriptional activity.

[0023] In some embodiments, the present disclosure provides a synthetic DNA molecule with the nucleotide sequence as set forth in SEQ ID NO: 16. In some embodiments, the present disclosure provides a synthetic DNA molecule having a nucleotide sequence that is at least about 80% identical to the nucleotide sequence of SEQ ID NO: 16, and which encodes a promoter that has plant trichome gland specific transcriptional activity.

[0024] In some embodiments, the present disclosure provides a synthetic DNA molecule with the nucleotide sequence as set forth in SEQ ID NO: 17. In some embodiments, the present disclosure provides a synthetic DNA molecule having a nucleotide sequence that is at least about 80% identical to the nucleotide sequence of SEQ ID NO: 17, and which encodes a promoter that has plant trichome gland specific transcriptional activity.

[0025] In some embodiments, the present disclosure provides a synthetic DNA molecule with the nucleotide sequence as set forth in SEQ ID NO: 18. In some embodiments, the present disclosure provides a synthetic DNA molecule having a nucleotide sequence that is at least about 80% identical to the nucleotide sequence of SEQ ID NO: 18, and which encodes a promoter that has plant trichome gland specific transcriptional activity.

[0026] In some embodiments, the present disclosure provides a synthetic DNA molecule with the nucleotide sequence as set forth in SEQ ID NO: 19. In some embodiments, the present disclosure provides a synthetic DNA molecule having a nucleotide sequence that is at least about 80% identical to the nucleotide sequence of SEQ ID NO: 19, and which encodes a promoter that has plant trichome gland specific transcriptional activity.

[0027] In some embodiments, the present disclosure provides a synthetic DNA molecule with the nucleotide sequence as set forth in SEQ ID NO: 21. In some embodiments, the present disclosure provides a synthetic DNA molecule having a nucleotide sequence that is at least about 80% identical to the nucleotide sequence of SEQ ID NO: 21, and which encodes a promoter that has plant trichome gland specific transcriptional activity.

[0028] In some embodiments, the present disclosure provides a synthetic DNA molecule with the nucleotide sequence as set forth in SEQ ID NO: 22. In some embodiments, the present disclosure provides a synthetic DNA molecule having a nucleotide sequence that is at least about 80% identical to the nucleotide sequence of SEQ ID NO: 22, and which encodes a promoter that has plant trichome gland specific transcriptional activity.

[0029] In some embodiments, the present disclosure provides a synthetic DNA molecule with the nucleotide sequence as set forth in SEQ ID NO: 23. In some embodiments, the present disclosure provides a synthetic DNA molecule having a nucleotide sequence that is at least about 80% identical to the nucleotide sequence of SEQ ID NO: 23, and which encodes a promoter that has plant trichome gland specific transcriptional activity.

[0030] In some embodiments, the present disclosure provides a synthetic DNA molecule with the nucleotide sequence as set forth in SEQ ID NO: 24. In some embodiments, the present disclosure provides a synthetic DNA molecule having a nucleotide sequence that is at least about 80% identical to the nucleotide sequence of SEQ ID NO: 24, and which encodes a promoter that has plant trichome gland specific transcriptional activity.

[0031] In some embodiments, the present disclosure provides a synthetic DNA molecule with the nucleotide sequence as set forth in SEQ ID NO: 25. In some embodiments, the present disclosure provides a synthetic DNA molecule having a nucleotide sequence that is at least about 80% identical to the nucleotide sequence of SEQ ID NO: 25, and which encodes a promoter that has plant trichome gland specific transcriptional activity.

[0032] In some embodiments, the present disclosure provides a synthetic DNA molecule with the nucleotide sequence as set forth in SEQ ID NO: 26. In some embodiments, the present disclosure provides a synthetic DNA molecule having a nucleotide sequence that is at least about 80% identical to the nucleotide sequence of SEQ ID NO: 26, and which encodes a promoter that has plant trichome gland specific transcriptional activity.

[0033] In some embodiments, the present disclosure provides a synthetic DNA molecule with the nucleotide sequence as set forth in SEQ ID NO: 28. In some embodiments, the present disclosure provides a synthetic DNA molecule having a nucleotide sequence that is at least about 80% identical to the nucleotide sequence of SEQ ID NO: 28, and which encodes a promoter that has plant trichome gland specific transcriptional activity.

[0034] In some embodiments, the present disclosure provides a synthetic DNA molecule with the nucleotide sequence as set forth in SEQ ID NO: 29. In some embodiments, the present disclosure provides a synthetic DNA molecule having a nucleotide sequence that is at least about 80% identical to the nucleotide sequence of SEQ ID NO: 29, and which encodes a promoter that has plant trichome gland specific transcriptional activity.

[0035] In some embodiments, the present disclosure provides a synthetic DNA molecule with the nucleotide sequence as set forth in SEQ ID NO: 31. In some embodiments, the present disclosure provides a synthetic DNA molecule having a nucleotide sequence that is at least about 80% identical to the nucleotide sequence of SEQ ID NO: 31, and which encodes a promoter that has plant trichome gland specific transcriptional activity.

[0036] In some embodiments, the present disclosure provides a synthetic DNA molecule with the nucleotide sequence as set forth in SEQ ID NO: 32. In some embodiments, the present disclosure provides a synthetic DNA molecule having a nucleotide sequence that is at least about 80% identical to the nucleotide sequence of SEQ ID NO: 32, and which encodes a promoter that has plant trichome gland specific transcriptional activity.

[0037] In some embodiments, the present disclosure provides a synthetic DNA molecule with the nucleotide sequence as set forth in SEQ ID NO: 33. In some embodiments, the present disclosure provides a synthetic DNA molecule having a nucleotide sequence that is at least about 80% identical to the nucleotide sequence of SEQ ID NO: 33, and which encodes a promoter that has plant trichome gland specific transcriptional activity.

[0038] In one aspect, the present disclosure provides a genetically engineered plant or plant cell comprising a chimeric gene integrated into its genome, the chimeric gene comprising a trichome specific promoter operably linked to a homologous or heterologous nucleic acid sequence, wherein the promoter is selected from the group consisting of: (a) a nucleotide sequence of any one of SEQ ID NOs: 1-3, 5, 6, 8, 9, 11-14, 16-19, 21-26, 28, 29, or 31-33; and (b) a nucleotide sequence that is at least about 80% identical to the nucleotide sequence of any one of SEQ ID NOs: 1-3, 5, 6, 8, 9, 11-14, 16-19, 21-26, 28, 29, or 31-33, and which encodes a promoter that has plant trichome gland specific transcriptional activity.

[0039] In some embodiments, the genetically engineered plant or plant cell belongs to the family *Solanacea*. In some embodiments, the *Solanacea* plant is *N. tabacum*. In some embodiments, the genetically engineered plant belongs to the family *Cannabaceae*. In some embodiments, the engineered *Cannabaceae* plant is *C. sativa*.

[0040] In one aspect, the present disclosure provides a method for expressing a polypeptide in plant trichomes, comprising: (a) introducing into a host cell an expression vector comprising a nucleotide sequence selected from the group consisting of: (i) a nucleotide

sequence set forth in any one of SEQ ID NOs: 1-3, 5, 6, 8, 9, 11-14, 16-19, 21-26, 28, 29, or 31-33; and (ii) a nucleotide sequence that is at least about 80% identical to the nucleotide sequence of any one of SEQ ID NOs: 1-3, 5, 6, 8, 9, 11-14, 16-19, 21-26, 28, 29, or 31-33, and which encodes a promoter that has plant trichome gland specific transcriptional activity; wherein the nucleic acid sequence of (i) or (ii) is operably linked to one or more nucleic acid sequences encoding a polypeptide; and (b) growing the plant under conditions which allow for the expression of the polypeptide.

[0041] In another aspect, the present disclosure provides a method for increasing a cannabinoid in a host plant trichome, comprising: (a) introducing into a host cell an expression vector comprising a nucleotide sequence selected from the group consisting of: (i) a nucleotide sequence set forth in any one of SEQ ID NOs: 1-3, 5, 6, 8, 9, 11-14, 16-19, 21-26, 28, 29, or 31-33; and (ii) a nucleotide sequence that is at least about 80% identical to the nucleotide sequence of any one of SEQ ID NOs: 1-3, 5, 6, 8, 9, 11-14, 16-19, 21-26, 28, 29, or 31-33, and which encodes a promoter that has plant trichome gland specific transcriptional activity; wherein the nucleic acid sequence of (i) or (ii) is operably linked to one or more nucleic acid sequences encoding an enzyme of the cannabinoid biosynthetic pathway; and (b) growing the plant under conditions which allow for the expression of the cannabinoid biosynthetic pathway enzyme; wherein expression of the cannabinoid biosynthetic pathway enzyme results in the plant having an increased cannabinoid content relative to a control plant grown under similar conditions.

[0042] In some embodiments of the method, the cannabinoid biosynthetic pathway enzyme is cannabidiolic acid (CBDA) synthase, cannabichromenic acid (CBCA) synthase or Δ^9tetrahydrocannabinolic acid (THCA) synthase.

[0043] In some embodiments, the method further comprises providing the plant with cannabigerolic acid (CBGA).

[0044] In some embodiments, the present disclosure provides a method for producing a genetically-engineered plant having increased Δ^9tetrahydrocannabinol (THC), cannabichromene (CBC), and/or cannabidiol (CBD) content relative to a control plant.

[0045] The technologies described and claimed herein have many attributes and embodiments including, but not limited to, those set forth or described or referenced in this

brief summary. It is not intended to be all-inclusive and the inventions described and claimed herein are not limited to or by the features or embodiments identified in this brief summary, which is included for purposes of illustration only and not restriction. Additional embodiments may be disclosed in the detailed description below.

[0046] In one aspect, the present disclosure provides a synthetic DNA molecule comprising a nucleotide sequence selected from the group consisting of: (a) a nucleotide sequence set forth in any one of SEQ ID NOs: 31, 32, or 33; and (b) a nucleotide sequence that is at least about 80% identical to the nucleotide sequence of any one of SEQ ID NOs: 31, 32, or 33, and which encodes a promoter having plant trichome gland specific transcriptional activity, wherein the nucleotide sequence is operably linked to a heterologous nucleic acid.

[0047] In some embodiments, the present disclosure provides an expression vector comprising the synthetic DNA molecule operably linked to one or more nucleic acid sequences encoding a polypeptide.

[0048] In some embodiments, the present disclosure provides a genetically engineered host cell comprising the expression vector. In some embodiments, the genetically engineered host cell is a *Cannabis sativa* cell. In some embodiments, the genetically engineered host cell is a *Nicotiana tabacum* cell.

[0049] In some embodiments, the present disclosure provides a genetically engineered plant comprising a cell comprising a chimeric nucleic acid construct comprising the synthetic DNA molecule. In some embodiments, the engineered plant is an *N. tabacum* plant. In some embodiments, the engineered plant is a *C. sativa* plant.

[0050] In some embodiments, the present disclosure provides seeds from the engineered plant of any one of 5, wherein the seeds comprise the chimeric nucleic acid construct.

[0051] In one aspect, the present disclosure provides a synthetic DNA molecule comprising a nucleotide sequence set forth in SEQ ID NO: 33.

[0052] In some embodiments, the present disclosure provides an expression vector comprising the synthetic DNA molecule operably linked to one or more nucleic acid sequences encoding a polypeptide.

[0053] In some embodiments, the present disclosure provides a genetically engineered host cell comprising the expression vector. In some embodiments, the genetically engineered host cell is a *Cannabis sativa* cell. In some embodiments, the genetically engineered host cell is a *Nicotiana tabacum* cell.

[0054] In some embodiments, the present disclosure provides a genetically engineered plant comprising a cell comprising a chimeric nucleic acid construct comprising the synthetic DNA molecule. In some embodiments, the engineered plant is an *N. tabacum* plant. In some embodiments, the engineered plant is a *C. sativa* plant.

[0055] In some embodiments, the present disclosure provides seeds from the engineered plant, wherein the seeds comprise the chimeric nucleic acid construct.

BRIEF DESCRIPTION OF THE DRAWINGS

[0056] **FIG. 1** depicts the cannabinoid biosynthetic pathway that leads to the formation of the major cannabinoids in *Cannabis sativa*, Δ^9-tetrahydrocannabinolic acid (THCA) and cannabidiolic acid (CBDA).

[0057] **FIGS. 2A-2B** show the promoter:reporter gene design utilized in the experiments described herein. Each promoter is fused to the GUS reporter gene and promoter activity is indicated by a blue color upon staining (**FIG. 2A**). **FIG. 2B** is a whole leaf view of trichome specific expression of cannabinoid biosynthetic enzyme gene promoters in tobacco trichomes. Results from a representative promoter (CBDA synthase 20800 gene promoter) are shown. Results from the other promoters in the cannabinoid biosynthetic pathway are qualitatively identical (not shown). **FIG. 2B** shows that expression of the reporter gene is restricted to the trichomes.

[0058] **FIG. 3** shows trichome specific expression of promoters from the complete cannabinoid biosynthetic pathway in tobacco trichomes. Promoter activity is indicated by a blue color caused by activity of the GUS reporter gene. A representative promoter from every gene in the pathway shows trichome specific expression and therefore all promoters from genes in the pathway will direct expression in the trichomes.

[0059] **FIG. 4A** shows the design of the 4 x Cannabinoid On (CANON) fragment synthetic promoter of the present technology. **FIG. 4B** shows trichome specific expression of the 4 x

CANON fragment synthetic promoter of the present technology. The activity of the synthetic promoter is indicated by the blue color caused by activity of the GUS reporter gene.

DETAILED DESCRIPTION

I. INTRODUCTION

[0060] The present technology relates to the discovery of nucleic acid sequences for twenty-three trichome specific promoters of enzymes involved in the cannabinoid biosynthetic pathway: (1) olivetol synthase (OLS; also referred to as tetraketide synthase) promoter; (2) OLS1 promoter; (3) OLS2 promoter; (4) olivetolic acid cyclase (OAC) promoter; (5) OAC1 promoter; (6) aromatic prenyltransferase (PT) promoter; (7) PT1 promoter; (8) hexanoyl-CoA synthetase (AAE1-1) promoter; (9) hexanoyl-CoA synthetase (AAE1-1′) promoter; (10) hexanoyl-CoA synthetase (AAE3) promoter; (11) hexanoyl-CoA synthetase (AAE12) promoter; (12) CBDA synthase (CBDAS) promoter; (13) CBDA synthase 1 (CBDAS1) promoter; (14) CBDA synthase (CBDAS) 20800 promoter; (15) CBDA synthase (CBDAS) 20800′ promoter; (16) THCA synthase (THCAS) 19603 promoter; (17) THCA synthase (THCAS) 19603′ promoter; (18) THCA synthase (THCAS) 50320 promoter; (19) THCA synthase (THCAS) 50320′ promoter; (20) THCA synthase (THCAS) 1330 promoter; (21) THCA synthase (THCAS) 1330′ promoter; (22) CBDA synthase (CBDAS) 3498 promoter; and (23) CBDA synthase (CBDAS) 3498′ promoter.

[0061] The nucleic acid sequences for each promoter have been determined. The nucleic acid sequences of the (i) olivetol synthase (OLS) promoter, (ii) OLS1 promoter, and (iii) OLS2 promoter are set forth in SEQ ID NOs: 1, 2, and 3, respectively, and the open reading frame (ORF) of OLS is set forth in SEQ ID NO: 4. The nucleic acid sequence of the olivetolic acid cyclase (OAC) promoter is set forth in SEQ ID NO: 5, the nucleic acid sequence of the OAC1 promoter is set forth in SEQ ID NO: 6, and the ORF of OAC is set forth in SEQ ID NO: 7. The nucleic acid sequence of aromatic prenyltransferase (PT) promoter is set forth in SEQ ID NO: 8, the nucleic acid sequence of the PT1 promoter is set forth in SEQ ID NO: 9, and the ORF of PT is set forth in SEQ ID NO: 10. The nucleic acid sequences of the (i) hexanoyl-CoA synthetase (AAE1-1) promoter, (ii) hexanoyl-CoA synthetase (AAE1-1′) promoter; (iii) hexanoyl-CoA synthetase (AAE3) promoter, and (iv) hexanoyl-CoA synthetase (AAE12) promoter are set forth in SEQ ID NOs: 11, 12, 13, and 14, respectively, and the ORF of hexanoyl-CoA (AAE-1) is set forth in SEQ ID NO: 15. The

nucleic acid sequences of the (i) CBDA synthase (CBDAS) promoter, (ii) CBDAS synthase 1 (CBDAS1) promoter, (iii) CBDA synthase (CBDAS) 20800 promoter, and (iv) CBDA synthase (CBDAS) 20800′ promoter are set forth in SEQ ID NOs: 16, 17, 18, and 19, respectively, and the nucleic acid sequence of the ORF of CBDA synthase (CBDAS) is set forth in SEQ ID NO: 20. The nucleic acid sequences of (i) THCA synthase (THCAS) 19603 promoter, (ii) THCA synthase (THCAS) 19603′ promoter, (iii) THCA synthase (THCAS) 50320 promoter, (iv) THCA synthase (THCAS) 50320′ promoter, (v) THCA synthase (THCAS) 1330 promoter, and (vi) THCA synthase (THCAS) 1330′ promoter are set forth in SEQ ID NOs: 21, 22, 23, 24, 25, and 26, respectively, and the ORF of THCAS is set forth in SEQ ID NO: 27. The nucleic acid sequence of the CBCA synthase (CBCAS) 3498 promoter is set forth in SEQ ID NO: 28, the nucleic acid sequence of the CBCA synthase (CBCAS) 3498′ promoter is set forth in SEQ ID NO: 29, and the ORF of CBCA synthase (CBCAS) is set forth in SEQ ID NO: 30.

[0062] The present technology also relates to the discovery of nucleic acid sequences for a "cannabinoid on" or "CANON" promoter fragment that is sufficient to direct trichome specific expression. The nucleic acid sequence for the CANON fragment that is sufficient to direct trichome specific expression in glandular trichomes is set forth in SEQ ID NO: 31. The nucleic acid sequence for the 4 x CANON fragment synthetic promoter that comprises four copies of the consensus CANON fragment is set forth in SEQ ID NO: 33.

[0063] Given the known cytotoxic effects of cannabinoids such as THC on plant cells, the expression of genes driving their production under a strong ubiquitous promoter, like the Cauliflower Mosaic Virus (CaMV) 35S, may lead to the perturbation of metabolic pathways in the whole plant and may have deleterious consequences on plant development and physiology. Thus, trichomes, as distinct entities with restricted communication to the rest of plant, represent a potential target for metabolic engineering.

[0064] Accordingly, in some embodiments, the present technology provides previously undiscovered trichome specific promoters from cannabinoid biosynthesis genes or biologically active fragments thereof that may be used to genetically manipulate the synthesis of cannabinoids (*e.g.*, THC, CBD, CBC, CBG) in host plants, such as *C. sativa*, plants of the family *Solanaceae*, and other plant families and species that do not naturally produce cannabinoids.

II. DEFINITIONS

[0065] All technical terms employed in this specification are commonly used in biochemistry, molecular biology and agriculture; hence, they are understood by those skilled in the field to which the present technology belongs. Those technical terms can be found, for example in: *Molecular Cloning: A Laboratory Manual* 3rd ed., vol. 1-3, ed. Sambrook and Russel (Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N.Y., 2001); *Current Protocols In Molecular Biology*, ed. Ausubel et al. (Greene Publishing Associates and Wiley-Interscience, New York, 1988) (including periodic updates); *Short Protocols In Molecular Biology: A Compendium Of Methods From Current Protocols In Molecular Biology* 5th ed., vol. 1-2, ed. Ausubel et al. (John Wiley & Sons, Inc., 2002); *Genome Analysis: A Laboratory Manual*, vol. 1-2, ed. Green et al. (Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N.Y., 1997). Methodology involving plant biology techniques are described here and also are described in detail in treatises such as *Methods In Plant Molecular Biology: A Laboratory Course Manual*, ed. Maliga et al. (Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N.Y., 1995).

[0029] A **"chimeric nucleic acid"** comprises a coding sequence or fragment thereof linked to a nucleotide sequence that is different from the nucleotide sequence with which it is associated in cells in which the coding sequence occurs naturally.

[0066] The terms **"encoding"** and **"coding"** refer to the process by which a gene, through the mechanisms of transcription and translation, provides information to a cell from which a series of amino acids can be assembled into a specific amino acid sequence to produce an active enzyme. Because of the degeneracy of the genetic code, certain base changes in DNA sequence do not change the amino acid sequence of a protein.

[0067] **"Endogenous nucleic acid"** or **"endogenous sequence"** is **"native"** to, *i.e.*, indigenous to, the plant or organism that is to be genetically engineered. It refers to a nucleic acid, gene, polynucleotide, DNA, RNA, mRNA, or cDNA molecule that is present in the genome of a plant or organism that is to be genetically engineered.

[0068] **"Exogenous nucleic acid"** refers to a nucleic acid, DNA or RNA, which has been introduced into a cell (or the cell's ancestor) through the efforts of humans. Such exogenous

nucleic acid may be a copy of a sequence which is naturally found in the cell into which it was introduced, or fragments thereof.

[0069] As used herein, **"expression"** denotes the production of an RNA product through transcription of a gene or the production of the polypeptide product encoded by a nucleotide sequence. **"Overexpression"** or **"up-regulation"** is used to indicate that expression of a particular gene sequence or variant thereof, in a cell or plant, including all progeny plants derived thereof, has been increased by genetic engineering, relative to a control cell or plant.

[0070] **"Genetic engineering"** encompasses any methodology for introducing a nucleic acid or specific mutation into a host organism. For example, a plant is genetically engineered when it is transformed with a polynucleotide sequence that suppresses expression of a gene, such that expression of a target gene is reduced compared to a control plant. In the present context, **"genetically engineered"** includes transgenic plants and plant cells. A genetically engineered plant or plant cell may be the product of any native approach (*i.e.*, involving no foreign nucleotide sequences), implemented by introducing only nucleic acid sequences derived from the host plant species or from a sexually compatible plant species. *See, e.g.*, U.S. Patent Application No. 2004/0107455.

[0071] **"Heterologous nucleic acid"** or **"homologous nucleic acid"** refer to the relationship between a nucleic acid or amino acid sequence and its host cell or organism, especially in the context of transgenic organisms. A homologous sequence is naturally found in the host species (*e.g.*, a cannabis plant transformed with a cannabis gene), while a heterologous sequence is not naturally found in the host cell (*e.g.*, a tobacco plant transformed with a sequence from cannabis plants). Such heterologous nucleic acids may comprise segments that are a copy of a sequence that is naturally found in the cell into which it has been introduced, or fragments thereof. Depending on the context, the term "homolog" or "homologous" may alternatively refer to sequences which are descendent from a common ancestral sequence (*e.g.*, they may be orthologs).

[0072] **"Increasing," "decreasing," "modulating," "altering,"** or the like refer to comparison to a similar variety, strain, or cell grown under similar conditions but without the modification resulting in the increase, decrease, modulation, or alteration. In some cases, this

may be a non-transformed control, a mock transformed control, or a vector-transformed control.

[0073] By **"isolated nucleic acid molecule"** is intended a nucleic acid molecule, DNA, or RNA, which has been removed from its native environment. For example, recombinant DNA molecules contained in a DNA construct are considered isolated for the purposes of the present technology. Further examples of isolated DNA molecules include recombinant DNA molecules maintained in heterologous host cells or DNA molecules that are purified, partially or substantially, in solution. Isolated RNA molecules include *in vitro* RNA transcripts of the DNA molecules of the present technology. Isolated nucleic acid molecules, according to the present technology, further include such molecules produced synthetically.

[0074] **"Plant"** is a term that encompasses whole plants, plant organs (*e.g.*, leaves, stems, roots, *etc.*), seeds, differentiated or undifferentiated plant cells, and progeny of the same. Plant material includes without limitation seeds, suspension cultures, embryos, meristematic regions, callus tissues, leaves, roots, shoots, stems, fruit, gametophytes, sporophytes, pollen, and microspores.

[0075] **"Plant cell culture"** means cultures of plant units such as, for example, protoplasts, cell culture cells, cells in plant tissues, pollen, pollen tubes, ovules, embryo sacs, zygotes, and embryos at various stages of development. In some embodiments of the present technology, a transgenic tissue culture or transgenic plant cell culture is provided, wherein the transgenic tissue or cell culture comprises a nucleic acid molecule of the present technology.

[0076] **"Promoter"** connotes a region of DNA upstream from the start of transcription that is involved in recognition and binding of RNA polymerase and other proteins to initiate transcription. A **"constitutive promoter"** is one that is active throughout the life of the plant and under most environmental conditions. Tissue-specific, tissue-preferred, cell type-specific, and inducible promoters constitute the class of **"non-constitutive promoters."** A **"trichome specific promoter"** is a promoter that preferentially directs expression of an operably linked gene in trichome tissue, as compared to expression in the root, leaf, stem, or other tissues of the plant. **"Operably linked"** refers to a functional linkage between a promoter and a second sequence, where the promoter sequence initiates and mediates

transcription of the DNA sequence corresponding to the second sequence. In general, "operably linked" means that the nucleic acid sequences being linked are contiguous.

[0077] **"Sequence identity"** or **"identity"** in the context of two polynucleotide (nucleic acid) or polypeptide sequences includes reference to the residues in the two sequences that are the same when aligned for maximum correspondence over a specified region. When percentage of sequence identity is used in reference to proteins it is recognized that residue positions which are not identical often differ by conservative amino acid substitutions, where amino acid residues are substituted for other amino acid residues with similar chemical properties, such as charge and hydrophobicity, and therefore do not change the functional properties of the molecule. Where sequences differ in conservative substitutions, the percent sequence identity may be adjusted upwards to correct for the conservative nature of the substitution. Sequences which differ by such conservative substitutions are said to have **"sequence similarity"** or **"similarity."** Means for making this adjustment are well-known to those of skill in the art. Typically this involves scoring a conservative substitution as a partial rather than a full mismatch, thereby increasing the percentage sequence identity. Thus, for example, where an identical amino acid is given a score of 1 and a non-conservative substitution is given a score of zero, a conservative substitution is given a score between zero and 1. The scoring of conservative substitutions is calculated, for example, according to the algorithm of Meyers & Miller, *Computer Applic. Biol. Sci.* 4: 11-17 (1988), as implemented in the program PC/GENE (Intelligenetics, Mountain View, California, USA).

[0078] Use in this description of a percentage of sequence identity denotes a value determined by comparing two optimally aligned sequences over a comparison window, wherein the portion of the polynucleotide sequence in the comparison window may comprise additions or deletions (*i.e.*, gaps) as compared to the reference sequence (which does not comprise additions or deletions) for optimal alignment of the two sequences. The percentage is calculated by determining the number of positions at which the identical nucleic acid base or amino acid residue occurs in both sequences to yield the number of matched positions, dividing the number of matched positions by the total number of positions in the window of comparison, and multiplying the result by 100 to yield the percentage of sequence identity.

[0079] The terms **"suppression"** or **"down-regulation"** are used synonymously to indicate that expression of a particular gene sequence variant thereof, in a cell or plant, including all

progeny plants derived thereof, has been reduced by genetic engineering, relative to a control cell or plant.

[0080] **"Trichome"** encompasses herein different types of trichomes, both glandular trichomes and/or non-glandular trichomes.

[0081] **"Trichome cells"** refers to the cells making up the trichome structure, such as the gland, or secretory cells, base cells and stalk, or stripe cells, extra-cellular cavity and cuticle cells. Trichomes can also consist of one single cell.

[0082] **"Cannabis"** or **"cannabis plant"** refers to any species in the *Cannabis* genus that produces cannabinoids, such as *Cannabis sativa* and interspecific hybrids thereof.

[0083] A **"variant"** is a nucleotide or amino acid sequence that deviates from the standard, or given, nucleotide or amino acid sequence of a particular gene or polypeptide. The terms **"isoform," "isotype,"** and **"analog"** also refer to "variant" forms of a nucleotide or an amino acid sequence. An amino acid sequence that is altered by the addition, removal, or substitution of one or more amino acids, or a change in nucleotide sequence, may be considered a variant sequence. A polypeptide variant may have "**conservative**" changes, wherein a substituted amino acid has similar structural or chemical properties, *e.g.*, replacement of leucine with isoleucine. A polypeptide variant may have **"nonconservative"** changes, *e.g.*, replacement of a glycine with a tryptophan. Analogous minor variations may also include amino acid deletions or insertions, or both. Guidance in determining which amino acid residues may be substituted, inserted, or deleted may be found using computer programs well known in the art such as Vector NTI Suite (InforMax, MD) software. Variant may also refer to a **"shuffled gene"** such as those described in Maxygen-assigned patents (*see, e.g.*, U. S. Patent No. 6,602,986).

[0084] As used herein, the term "**about**" will be understood by persons of ordinary skill in the art and will vary to some extent depending upon the context in which it is used. If there are uses of the term which are not clear to persons of ordinary skill in the art given the context in which it is used, "about" will mean up to plus or minus 10% of the particular term.

[0085] The term "**biologically active fragments"** or **"functional fragments"** or **"fragments having promoter activity"** refer to nucleic acid fragments which are capable of

conferring transcription in one or more trichome types and/or one or more trichome cells found on one or more different types of plant tissues and organs. Biologically active fragments confer trichome specific and/or at least trichome preferred expression, and they preferably have at least a similar strength (or higher strength) as the promoter of SEQ ID NOs: 1, 3, 5, 7, 9, or 11-15. This can be tested by transforming a plant with such a fragment, preferably operably linked to a reporter gene, and assaying the promoter activity qualitatively (spatio-temporal transcription) and/or quantitatively in trichomes. In some embodiments, the strength of the promoter and/or promoter fragments of the present technology is quantitatively identical to, or higher than, that of the CaMV 35S promoter when measured in the glandular trichome. In some embodiments, a biologically active fragment of a trichome promoter described herein can be about 5%, about 10%, about 15%, about 20%, about 25%, about 30%, about 35%, about 40%, about 45%, about 50%, about 55%, about 60%, about 65%, about 70%, about 75%, about 80%, about 85%, about 90%, about 91%, about 92%, about 93%, about 94%, about 95%, about 96%, about 97%, about 98%, or about 99% of the full length sequence nucleic acid sequence for the promoter. In other embodiments, a biologically active nucleic acid fragment of a trichome promoter described herein can be, for example, at least about 10 contiguous nucleic acids. In yet other embodiments, the biologically active nucleic acid fragment of a trichome promoter described herein can be (1) about 10 contiguous nucleic acids up to about 554 contiguous nucleic acids for the OLS promoter (*e.g.*, SEQ ID NO: 1); (2) about 10 contiguous nucleic acids up to about 550 contiguous nucleic acids for the OLS1 promoter (SEQ ID NO: 2); (3) about 10 contiguous nucleic acids up to about 558 contiguous nucleic acids for the OLS2 promoter (SEQ ID NO: 3); (4) about 10 contiguous nucleic acids up to about 996 contiguous nucleic acids for the OAC promoter (*e.g.*, SEQ ID NO: 5); (5) about 10 contiguous nucleic acids up to about 992 contiguous nucleic acids for the OAC1 promoter (*e.g.*, SEQ ID NO: 6); (6) about 10 contiguous nucleic acids up to about 1361 contiguous nucleic acids for the PT promoter (*e.g.*, SEQ ID NO: 8); (7) about 10 contiguous nucleic acids up to about 1357 contiguous nucleic acids for the PT1 promoter (*e.g.*, SEQ ID NO: 9); (8) about 10 contiguous nucleic acids up to about 805 contiguous nucleic acids for the AAE1-1 promoter (*e.g.*, SEQ ID NO: 11); (9) about 10 contiguous nucleic acids up to about 800 contiguous nucleic acids for the AAE1-1′ promoter (*e.g.*, SEQ ID NO: 12); (10) about 10 contiguous nucleic acids up to about 1000 contiguous nucleic acids for the AAE3 promoter (*e.g.*, SEQ ID NO: 13); (11) about 10 contiguous nucleic acids up to about 869 contiguous nucleic acids for the AAE12 promoter

(*e.g.*, SEQ ID NO: 14); (12) about 10 contiguous nucleic acids up to about 420 contiguous nucleic acids for the CBDA synthase promoter (*e.g.*, SEQ ID NO: 16); (13) about 10 contiguous nucleic acids up to about 416 contiguous nucleic acids for the CBDAS1 promoter (*e.g.*, SEQ ID NO: 17); (14) about 10 contiguous nucleic acids up to about 535 contiguous nucleic acids for the CBDAS 20800 promoter (*e.g.*, SEQ ID NO: 18); (15) about 10 contiguous nucleic acids up to about 531 contiguous nucleic acids for the CBDAS 20800'promoter (*e.g.*, SEQ ID NO: 19); (16) about 10 contiguous nucleic acids up to about 800 contiguous nucleic acids for the THCAS 19603 promoter (*e.g.*, SEQ ID NO: 21); (17) about 10 contiguous nucleic acids up to about 796 contiguous nucleic acids for the THCAS 19603' promoter (*e.g.*, SEQ ID NO: 22); (18) about 10 contiguous nucleic acids up to about 796 contiguous nucleic acids for the THCAS 50320 promoter (*e.g.*, SEQ ID NO: 23); (19) about 10 contiguous nucleic acids up to about 792 contiguous nucleic acids for the THCAS 50320' promoter (*e.g.*, SEQ ID NO: 24); (20) about 10 contiguous nucleic acids up to about 720 contiguous nucleic acids for the THCAS 1330 promoter (*e.g.*, SEQ ID NO: 25); (21) about 10 contiguous nucleic acids up to about 716 contiguous nucleic acids for the THCAS 1330' promoter (*e.g.*, SEQ ID NO: 26); (22) about 10 contiguous nucleic acids up to about 804 contiguous nucleic acids for the CBCAS 3498 promoter (*e.g.*, SEQ ID NO: 28); or (23) about 10 contiguous nucleic acids up to about 800 contiguous nucleic acids for the CBCAS 3498' promoter (*e.g.*, SEQ ID NO: 29). In yet other embodiments, the biologically active fragment of the trichome promoter can be any value of contiguous nucleic acids in between these two amounts, such as but not limited to about 20, about 30, about 40, about 50, about 60, about 70, about 80, about 90, about 100, about 110, about 120, about 130, about 140, about 150, about 160, about 170, about 180, about 190, about 200, about 250, about 300, about 350, about 400, about 450, about 500, about 550, about 600, about 650, about 700, about 750, about 800, about 850, about 900, about 950, about 1000, about 1050, about 1100, about 1150, about 1200, about 1250, or about 1300 contiguous nucleic acids. .

III. GENETIC ENGINEERING OF HOST CELLS AND ORGANISMS USING TRICHOME SPECIFIC PROMOTERS

A. Trichome Specific Promoters

[0086] The disclosure of the present technology relates to the identification of twenty-three promoters, which are capable of regulating transcription of coding nucleic acid sequences operably linked thereto in trichome cells.

[0087] Accordingly, the present technology provides an isolated polynucleotide having a nucleic acid sequence that is at least about 50%, about 55%, about 60%, about 65%, about 70%, about 75%, about 80%, about 85%, about 90%, about 91%, about 92%, about 93%, about 94%, about 95%, about 96%, about 97%, about 98%, about 99%, or 100% identical to a nucleic acid sequence described in any of SEQ ID NOs: 1-3, 5, 6, 8, 9, 11-14, 16-19, 21-26, 28, 29, and 31-33, wherein the nucleic acid sequence is capable of regulating transcription of coding nucleic acid sequences operably linked thereto in trichome cells. Differences between two nucleic acid sequences may occur at the 5' or 3' terminal positions of the reference nucleotide sequence or anywhere between those terminal positions, interspersed either individually among nucleotides in the reference sequence or in one or more contiguous groups within the reference sequence.

[0088] The present technology also includes biologically active "variants" of SEQ ID NOs: 1-3, 5, 6, 8, 9, 11-14, 16-19, 21-26, 28, 29, and 31-33, with one or more bases deleted, substituted, inserted, or added, wherein the nucleic acid sequence is capable of regulating transcription of coding nucleic acid sequences operably linked thereto in trichome cells. Variants of SEQ ID NOs: 1-3, 5, 6, 8, 9, 11-14, 16-19, 21-26, 28, 29, and 31-33, include nucleic acid sequences comprising at least about 50%, about 55%, about 60%, about 65%, about 70%, about 75%, about 80%, about 85%, about 90%, about 91%, about 92%, about 93%, about 94%, about 95%, about 96%, about 97%, about 98%, about 99% or more nucleic acid sequence identity to SEQ ID NOs: 1-3, 5, 6, 8, 9, 11-14, 16-19, 21-26, 28, 29, and 31-33, and which are trichome specific in their activity.

[0089] In some embodiments of the present technology, the polynucleotides (promoters) are modified to create variations in the molecule sequences such as to enhance their promoting activities, using methods known in the art, such as PCR-based DNA modification, or standard mutagenesis techniques, or by chemically synthesizing the modified polynucleotides.

[0090] Accordingly, the sequences set forth in SEQ ID NOs: 1-3, 5, 6, 8, 9, 11-14, 16-19, 21-26, 28, 29, and 31-33 may be truncated or deleted and still retain the capacity of directing the transcription of an operably linked nucleic acid sequence in trichomes. The minimal length of a promoter region can be determined by systematically removing sequences from the 5' and 3'-ends of the isolated polynucleotide by standard techniques known in the art,

including but not limited to removal of restriction enzyme fragments or digestion with nucleases.

[0091] Trichome specific promoters of the present technology may also be used for expressing a nucleic acid that will decrease or inhibit expression of a native gene in the plant. Such nucleic acids may encode antisense nucleic acids, ribozymes, sense suppression agents, or other products that inhibit expression of a native gene.

[0092] The trichome specific promoters of the present technology may also be used to express proteins or peptides in "molecular farming" applications. Such proteins or peptides include but are not limited to industrial enzymes, antibodies, therapeutic agents, and nutritional products.

[0093] In some embodiments, novel hybrid promoters can be designed or engineered by a number of methods. Many promoters contain upstream sequences which activate, enhance, or define the strength and/or specificity of the promoter. *See, e.g.*, Atchison, *Ann. Rev. Cell Biol.* 4:127 (1988). T-DNA genes, for example contain "TATA" boxes defining the site of transcription initiation and other upstream elements located upstream of the transcription initiation site modulate transcription levels.

B. Cannabinoid On (CANON) Fragment for Trichome Specific Expression

[0094] In some embodiments, the disclosure of the present technology also relates to the identification of a nucleic acid molecule termed the "Cannabinoid On" or "CANON" fragment that is sufficient for directing trichome specific expression of coding nucleic acid sequences operably linked thereto.

[0095] The 171-base pair CANON fragment (SEQ ID NO: 31) is shown below in **Table 1**. The consensus CANON fragment (highlighted) is shown together with the putative TATA box (bold underline), 5' UTR, and start codon ("atg" in bold underline) as SEQ ID NO: 32 in **Table 1**. The consensus is derived from the trichome specific promoters from THCA synthases 19603, 1330, and 50320, CBCA synthase 3498, and CBDA synthase 20800. The 171-base pair CANON fragment (SEQ ID NO: 31) is sufficient to direct trichome specific expression in glandular trichomes of tobacco (and cannabis).

[0096] A nucleic acid sequence comprising four copies of the CANON fragment in front of one copy of the minimal promoter (*i.e.*, TATA Box, start of transcription, and first ATG) termed "4 x CANON fragment synthetic promoter" is also shown as SEQ ID NO: 33 in **Table 1**. The first CANON fragment of SEQ ID NO: 33 is shown in bold, followed by a second fragment shown in underline, followed by a third fragment in bold, and a fourth fragment in underline.

Table 1. CANON fragment sequences.

Cannabinoid On ("CANON") fragment (171 bp) atgatgccaaactattcaatgtacaatgtacatttatttttaataagggcttcacctaacaa aggtgcctaatttttgtgaactttttttttaccacatgtgactatttaatgactatcaaatta taaaatatttaagtcaatttctttgcccccactccaatatataatgt (SEQ ID NO: 31)
CANON fragment with putative TATA Box, 5' UTR, and start codon (232 bp) atgatgccaaactattcaatgtacaatgtacatttatttttaataagggcttcacctaacaa aggtgcctaatttttgtgaactttttttttaccacatgtgactatttaatgactatcaaatta taaaatatttaagtcaatttctttgcccccactccaatatataatgt**tataaata**ggataat tctcaattcatagtaattcaaaaatcattaggactaaagaaaa**atg** (SEQ ID NO: 32)
4 x CANON fragment synthetic promoter (709 bp) **atgatgccaaactattcaatgtacaatgtacatttatttttaataagggcttcacctaacaa** **aggtgcctaatttttgtgaactttttttttaccacatgtgactatttaatgactatcaaatta** **taaaatatttaagtcaatttctttgcccccactcc**atgatgccaaactattcaatgtacaat gtacatttatttttaataagggcttcacctaacaaaggtgcctaatttttgtgaacttttttt ttaccacatgtgactatttaatgactatcaaattataaaatatttaagtcaatttctttgcc cccactccatgatgccaaactattcaatgtacaatgtacatttatttttaataagggcttca **cctaacaaaggtgcctaatttttgtgaacttttttttaccacatgtgactatttaatgacta** **tcaaattataaaatatttaagtcaatttctttgcccccactccatgatgccaaactattcaa** tgtacaatgtacatttatttttaataagggcttcacctaacaaaggtgcctaatttttgtga acttttttttaccacatgtgactatttaatgactatcaaattataaaatatttaagtcaatt tctttgcccccactccaatatataatgttataaataggataattctcaattcatagtaattc aaaaatcattaggactaaagaaaaatg (SEQ ID NO: 33)

[0097] The CANON fragment is sufficient to direct trichome specific expression in a gain-of-function promoter. Without wishing to be bound by theory, it is believed that the CANON fragment is responsible for the trichome specific expression of a number of cannabinoid biosynthetic enzyme genes including THCA, CBDA, and CBCA synthases.

C. Nucleic Acid Constructs

[0098] In some embodiments, the trichome specific promoter sequences and CANON fragments of the present technology, or biologically active fragments thereof, can be incorporated into nucleic acid constructs, such as expression constructs (*i.e.*, expression vectors), which can be introduced and replicate in a host cell, such as plant trichome cell.

Such nucleic acid constructs may include a heterologous nucleic acid operably linked to any of the promoter sequences or CANON fragments of the present technology. Thus, in some embodiments, the present technology provides the use of any of the promoters or CANON fragments set forth in SEQ ID NOs: 1-3, 5, 6, 8, 9, 11-14, 16-19, 21-26, 28, 29, and 31-33, or biologically active fragments thereof, for the expression of homologous or heterologous nucleic acid sequences in a recombinant cell or organism, such as a plant cell or plant. In some embodiments, this use comprises operably linking any of the promoters or CANON fragments set forth in SEQ ID NOs: 1-3, 5, 6, 8, 9, 11-14, 16-19, 21-26, 28, 29, and 31-33, or biologically active fragments thereof, to a homologous or heterologous nucleic acid sequence to form a nucleic acid construct and transforming a host, such as a plant or plant cell. In some embodiments, various genes that encode enzymes involved in biosynthetic pathways for the production of cannabinoids (*e.g.*, at least one of the nucleic acid sequences set forth in SEQ ID NOs: 4, 7, 10, 15, 20, 27, or 30) can be suitable as transgenes that can be operably linked to a trichome specific promoter or CANON fragment of the present technology. In some embodiments, the nucleic acid constructs of the present technology modulate the expression of one or more proteins that regulate cannabinoid biosynthesis. In some embodiments, the nucleic acid constructs of the present technology can be used to modulate the expression of cannabinoids or other compounds (*e.g.*, terpenes) in trichome cells.

[0099] In some embodiments, an expression vector comprises a promoter or CANON fragment comprising a nucleic acid sequence selected from the group consisting of SEQ ID NOs: 1-3, 5, 6, 8, 9, 11-14, 16-19, 21-26, 28, 29, and 31-33, or a biologically active fragment thereof, operably linked to the cDNA encoding a polypeptide, such as one or more of olivetol synthase (OLS), olivetolic acid cyclase (OAC), aromatic pyrenyltransferase (PT), hexanoyl-CoA synthetase (AEE1-1), CBDA synthase, CBCA synthase, and THCA synthase. In another embodiment, a plant cell line comprises an expression vector comprising a promoter or CANON fragment comprising a nucleic acid sequence selected from the group consisting of SEQ ID NOs: 1-3, 5, 6, 8, 9, 11-14, 16-19, 21-26, 28, 29, and 31-33, or a biologically active fragment thereof, operably linked to the cDNA encoding a polypeptide, such as one or more of olivetol synthase (OLS), olivetolic acid cyclase (OAC), aromatic pyrenyltransferase (PT), hexanoyl-CoA synthetase (AEE1-1), CBDA synthase, CBCA synthase, and THCA synthase. In another embodiment, a transgenic plant comprises an expression vector comprising a promoter or CANON fragment comprising a nucleic acid sequence selected

from the group consisting of SEQ ID NOs: 1-3, 5, 6, 8, 9, 11-14, 16-19, 21-26, 28, 29, and 31-33, or a biologically active fragment thereof, operably linked to the cDNA encoding a polypeptide, such as one or more of olivetol synthase (OLS), olivetolic acid cyclase (OAC), aromatic pyrenyltransferase (PT), hexanoyl-CoA synthetase (AEE1-1), CBDA synthase, CBCA synthase, and THCA synthase. In another embodiment, methods for genetically modulating the production of cannabinoids are provided, comprising: introducing an expression vector comprising a promoter or CANON fragment comprising a nucleic acid sequence selected from the group consisting of SEQ ID NOs: 1-3, 5, 6, 8, 9, 11-14, 16-19, 21-26, 28, 29, and 31-33, or a biologically active fragment thereof, operably linked to the cDNA encoding a polypeptide, such as one or more of olivetol synthase (OLS), olivetolic acid cyclase (OAC), aromatic pyrenyltransferase (PT), hexanoyl-CoA synthetase (AEE1-1), CBDA synthase, CBCA synthase, and THCA synthase.

[0100] In another embodiment, an expression vector comprises one or more promoters or CANON fragments comprising a nucleic acid sequence selected from the group consisting of SEQ ID NOs: 1-3, 5, 6, 8, 9, 11-14, 16-19, 21-26, 28, 29, and 31-33, or a biologically active fragment thereof, operably linked to cDNA encoding a polypeptide, such as one or more of olivetol synthase (OLS), olivetolic acid cyclase (OAC), aromatic pyrenyltransferase (PT), hexanoyl-CoA synthetase (AEE1-1), CBDA synthase, CBCA synthase, and THCA synthase. In another embodiment, a plant cell line comprises one or more promoters comprising a nucleic acid sequence selected from the group consisting of SEQ ID NOs: 1-3, 5, 6, 8, 9, 11-14, 16-19, 21-26, 28, 29, and 31-33, or a biologically active fragment thereof, operably linked to cDNA encoding a polypeptide, such as one or more of olivetol synthase (OLS), olivetolic acid cyclase (OAC), aromatic pyrenyltransferase (PT), hexanoyl-CoA synthetase (AEE1-1), CBDA synthase, CBCA synthase, and THCA synthase. In another embodiment, a transgenic plant comprises one or more promoters or CANON fragments comprising a nucleic acid sequence selected from the group consisting of SEQ ID NOs: 1-3, 5, 6, 8, 9, 11-14, 16-19, 21-26, 28, 29, and 31-33, or a biologically active fragment thereof, operably linked to cDNA encoding a polypeptide, such as one or more of olivetol synthase (OLS), olivetolic acid cyclase (OAC), aromatic pyrenyltransferase (PT), hexanoyl-CoA synthetase (AEE1-1), CBDA synthase, CBCA synthase, and THCA synthase. In another embodiment, methods for genetically modulating the production level of cannabinoids are provided, comprising introducing into a host cell an expression vector comprising one or more promoters or

CANON fragments, comprising a nucleic acid sequence selected from the group consisting of SEQ ID NOs: 1-3, 5, 6, 8, 9, 11-14, 16-19, 21-26, 28, 29, and 31-33, or a biologically active fragment thereof, operably linked to cDNA encoding a polypeptide, such as one or more of olivetol synthase (OLS), olivetolic acid cyclase (OAC), aromatic pyrenyltransferase (PT), hexanoyl-CoA synthetase (AEE1-1), CBDA synthase, CBCA synthase, and THCA synthase.

[0101] Constructs may be comprised within a vector, such as an expression vector adapted for expression in an appropriate host (plant) cell. It will be appreciated that any vector which is capable of producing a plant comprising the introduced DNA sequence will be sufficient.

[0102] Suitable vectors are well known to those skilled in the art and are described in general technical references such as Pouwels et al., *Cloning Vectors, A Laboratory Manual*, Elsevier, Amsterdam (1986). Vectors for plant transformation have been described (*see, e.g.*, Schardl et al., *Gene* 61:1-14 (1987)). In some embodiments, the nucleic acid construct is a plasmid vector, or a binary vector. Examples of suitable vectors include the Ti plasmid vectors.

[0103] Recombinant nucleic acid constructs (*e.g.*, expression vectors) capable of introducing nucleotide sequences or chimeric genes under the control of a trichome specific regulatory sequence (*e.g.*, promoter, CANON fragment) may be made using standard techniques generally known in the art. To generate a chimeric gene, an expression vector generally comprises, operably linked in the 5' to 3' direction, a trichome specific promoter sequence or CANON sequence that directs the transcription of a downstream homologous or heterologous nucleic acid sequence, and optionally followed by a 3' untranslated nucleic acid region (3'-UTR) that encodes a polyadenylation signal which functions in plant cells to cause the termination of transcription and the addition of polyadenylate nucleotides to the 3' end of the mRNA encoding the protein. The homologous or heterologous nucleic acid sequence may be a sequence encoding a protein or peptide or it may be a sequence that is transcribed into an active RNA molecule, such as a sense and/or antisense RNA suitable for silencing a gene or gene family in the host cell or organism. Expression vectors also generally contain a selectable marker. Typical 5'to 3' regulatory sequences include a transcription initiation site, a ribosome binding site, an RNA processing signal, a transcription termination site, and/or polyadenylation signal.

[0104] In some embodiments, the expression vectors of the present technology may contain termination sequences, which are positioned downstream of the nucleic acid molecules of the present technology, such that transcription of mRNA is terminated, and polyA sequences added. Exemplary terminators include *Agrobacterium tumefaciens* nopaline synthase terminator (Tnos), *Agrobacterium tumefaciens* mannopine synthase terminator (Tmas), and the CaMV 35S terminator (T35S). Termination regions include the pea ribulose bisphosphate carboxylase small subunit termination region (TrbcS) or the Tnos termination region. The expression vector also may contain enhancers, start codons, splicing signal sequences, and targeting sequences.

[0105] In some embodiments, the expression vectors of the present technology may contain a selection marker by which transformed cells can be identified in culture. The marker may be associated with the heterologous nucleic acid molecule, *i.e.*, the gene operably linked to a promoter. As used herein, the term "marker" refers to a gene encoding a trait or a phenotype that permits the selection of, or the screening for, a plant or cell containing the marker. In plants, for example, the marker gene will encode antibiotic or herbicide resistance. This allows for selection of transformed cells from among cells that are not transformed or transfected.

[0106] Examples of suitable selectable markers include but are not limited to adenosine deaminase, dihydrofolate reductase, hygromycin-B-phosphotransferase, thymidine kinase, xanthine-guanine phospho-ribosyltransferase, glyphosate and glufosinate resistance, and amino-glycoside 3'-O-phosphotransferase (kanamycin, neomycin and G418 resistance). These markers may include resistance to G418, hygromycin, bleomycin, kanamycin, and gentamicin. The construct may also contain the selectable marker gene *bar* that confers resistance to herbicidal phosphinothricin analogs like ammonium gluphosinate. *See, e.g.*, Thompson et al., *EMBO J.* 9:2519-23 (1987)). Other suitable selection markers known in the art may also be used.

[0107] Visible markers such as green florescent protein (GFP) may be used. Methods for identifying or selecting transformed plants based on the control of cell division have also been described. *See, e.g.*, WO 2000/052168 and WO 2001/059086.

[0108] Replication sequences, of bacterial or viral origin, may also be included to allow the vector to be cloned in a bacterial or phage host. Preferably, a broad host range prokaryotic origin of replication is used. A selectable marker for bacteria may be included to allow selection of bacterial cells bearing the desired construct. Suitable prokaryotic selectable markers also include resistance to antibiotics such as kanamycin or tetracycline.

[0109] Other nucleic acid sequences encoding additional functions may also be present in the vector, as is known in the art. For example, when *Agrobacterium* is the host, T-DNA sequences may be included to facilitate the subsequent transfer to and incorporation into plant chromosomes.

[0110] Whether a nucleic acid sequence of present technology or biologically active fragment thereof is capable of conferring transcription specifically in trichomes and whether the activity is "strong," can be determined using various methods. Qualitative methods (*e.g.*, histological GUS (β-glucuronidase) staining) are used to determine the spatio-temporal activity of the promoter or CANON fragment (*i.e.*, whether the promoter or CANON fragment is active in a certain tissue or organ (*e.g.*, trichomes, or under certain environmental/developmental conditions). Quantitative methods (*e.g.*, fluorometric GUS assays) also quantify the level of activity compared to controls. Suitable controls include, but are not limited to, plants transformed with empty vectors (negative controls) or transformed with constructs comprising other promoters, such as the *Arabidopsis* CER6 promoter, which is active in the epidermis and trichomes of *Nicotiana tabacum*.

[0111] To test or quantify the activity of a promoter or CANON fragment of the present technology, a nucleic acid sequence as set forth in SEQ ID NOs: 1-3, 5, 6, 8, 9, 11-14, 16-19, 21-26, 28, 29, or 31-33, or biologically active fragments thereof, may be operably linked to a known nucleic acid sequence (*e.g.*, a reporter gene such as *gusA*, or any gene encoding a specific protein) and may be used to transform a plant cell using known methods. The activity of the promoter or CANON fragment can, for example, be assayed (and optionally quantified) by detecting the level of RNA transcripts of the downstream nucleic acid sequence in trichome cells by quantitative RT-PCR or other PCR-based methods. Alternatively, the reporter protein or activity of the reporter protein may be assayed and quantified, by, for example a fluorometric GUS assay if the reporter gene is the *gus* gene.

[0112] In some embodiments, the promoters of the present technology can be used to drive expression of a heterologous nucleic acid of interest in trichome cells. The heterologous nucleic acid can encode any man-made recombinant or naturally occurring or protein, such as the cannabinoid biosynthetic pathway enzymes olivetol synthase (OLS), olivetolic acid cyclase (OAC), aromatic prenyltransferase (PT), hexanoyl-CoA synthetase (AAE1-1), CBDA synthase, THCA synthase, or CBCA synthase as set forth in SEQ ID NOs: 4, 7, 10, 15, 20, 27, and 30, respectively.

D. Host Plants and Cells and Plant Regeneration

[0113] The nucleic acid construct of the present technology can be utilized to transform a host cell, such as a plant cell. In some embodiments, the nucleic acid construct of the present technology is used to transform at least a portion of the cells of a plant. These expression vectors can be transiently introduced into host plant cells or stably integrated into the genomes of host plant cells to generate transgenic plants by various methods known to persons skilled in the art.

[0114] Methods for introducing nucleic acid constructs into a cell or plant are well known in the art. Suitable methods for introducing nucleic acid constructs (*e.g.*, expression vectors) into plant trichomes to generate transgenic plants include, but are not limited to, *Agrobacterium*-mediated transformation, particle gun delivery, microinjection, electroporation, polyethylene glycol-assisted protoplast transformation, and liposome-mediated transformation. Methods for transforming dicots primarily use *Agrobacterium tumefaciens*.

[0115] *Agrobacterium rhizogenes* may be used to produce transgenic hairy roots cultures of plants, including cannabis and tobacco, as described, for example, by Guillon et al., *Curr. Opin. Plant Biol.* 9:341-6 (2006). "Tobacco hairy roots" refers to tobacco roots that have T-DNA from an Ri plasmid of *Agrobacterium rhizogenes* integrated in the genome and grow in culture without supplementation of auxin and other phytohormones.

[0116] Additionally, plants may be transformed by *Rhizobium*, *Sinorhizobium*, or *Mesorhizobium* transformation. (Broothaerts et al., *Nature* 433: 629-633 (2005)).

[0117] After transformation of the plant cells or plant, those plant cells or plants into which the desired DNA has been incorporated may be selected by such methods as antibiotic resistance, herbicide resistance, tolerance to amino-acid analogues or using phenotypic markers.

[0118] The transgenic plants can be used in a conventional plant breeding scheme, such as crossing, selfing, or backcrossing, to produce additional transgenic plants containing the transgene.

[0119] Suitable host cells include plant cells. Any plant may be a suitable host, including monocotyledonous plants or dicotyledonous plants, such as, for example, maize/corn (*Zea* species, *e.g., Z. mays, Z. diploperennis* (chapule), *Zea luxurians* (Guatemalan teosinte), *Zea mays* subsp. *huehuetenangensis* (San Antonio Huista teosinte), *Z. mays* subsp. *mexicana* (Mexican teosinte), *Z. mays* subsp. *parviglumis* (Balsas teosinte), *Z. perennis* (perennial teosinte) and *Z. ramosa*, wheat (*Triticum* species), barley (*e.g., Hordeum vulgare*), oat (*e.g., Avena sativa*), *sorghum* (*Sorghum bicolor*), rye (*Secale cereale*), soybean (*Glycine* spp, *e.g., G. max*), cotton (*Gossypium* species, *e.g., G. hirsutum, G. barbadense*), *Brassica* spp. (*e.g., B. napus, B. juncea, B. oleracea, B. rapa, etc.*), sunflower (*Helianthus annus*), tobacco (*Nicotiana species*), alfalfa (*Medicago sativa*), rice (*Oryza* species, *e.g., O. sativa* indica cultivar-group or *japonica* cultivar-group), forage grasses, pearl millet (*Pennisetum* species. *e.g., P. glaucum*), tree species, vegetable species, such as *Lycopersicon* ssp (recently reclassified as belonging to the genus *Solanum*), *e.g.*, tomato (*L. esculentum*, syn. *Solanum lycopersicum*) such as *e.g.*, cherry tomato, var. *cerasiforme* or current tomato, var. *pimpinellifolium*) or tree tomato (*S. betaceum*, syn. *Cyphomandra betaceae*), potato (*Solanum tuberosum*) and other *Solanum* species, such as eggplant (*Solanum melongena*), pepino (*S. muricatum*), cocona (*S. sessiliflorum*) and naranjilla (*S. quitoense*); peppers (*Capsicum annuum, Capsicum frutescens*), pea (*e.g., Pisum sativum*), bean (*e.g., Phaseolus* species), carrot (*Daucus carona*), *Lactuca* species (such as *Lactuca sativa, Lactuca indica, Lactuca perennis*), cucumber (*Cucumis sativus*), melon (*Cucumis melo*), zucchini (*Cucurbita pepo*), squash (*Cucurbita maxima, Cucurbita pepo, Cucurbita mixta*), pumpkin (*Cucurbita pepo*), watermelon (*Citrullus lanatus* syn. *Citrullus vulgaris*), fleshy fruit species (grapes, peaches, plums, strawberry, mango, melon), ornamental species (*e.g.*, Rose, *Petunia, Chrysanthemum*, Lily, Tulip, *Gerbera* species), woody trees (*e.g.*, species of *Populus, Salix, Quercus, Eucalyptus*), fibre species *e.g.*, flax (*Linum usitatissimum*), and hemp (*Cannabis sativa*). In

some embodiments, the plant is *Cannabis sativa*. In some embodiments, the plant is *Nicotiana tabacum*.

[0120] Thus, in some embodiments, the present technology contemplates the use of the trichome specific promoters and/or CANON fragments comprising the nucleic acid sequences set forth SEQ ID NOs: 1-3, 5, 6, 8, 9, 11-14, 16-19, 21-26, 28, 29, and 31-33, or biologically active fragments thereof, to genetically manipulate the synthesis of cannabinoids (*e.g.*, THC, CBD, CBC) or other molecules in host plants, such as *C. sativa*, plants of the family *Solanaceae*, such as *N. tabacum*, and other plant families and species that do not naturally produce cannabinoids.

[0121] The present technology also contemplates cell culture systems (*e.g.*, plant cell cultures, bacterial or fungal cell cultures, human or mammalian cell cultures, insect cell cultures) comprising genetically engineered cells transformed with the nucleic acid molecules described herein. In some embodiments, a cell culture comprising cells comprising a promoter or CANON fragment of the present technology is provided.

[0122] Various assays may be used to determine whether a plant cell shows a change in gene expression, for example, Northern blotting or quantitative reverse transcriptase PCR (RT-PCR). Whole transgenic plants may be regenerated from the transformed cell by conventional methods. Such transgenic plants may be propagated and self-pollinated to produce homozygous lines. Such plants produce seeds containing the genes for the introduced trait and can be grown to produce plants that will produce the selected phenotype.

[0123] To enhance the expression and/or accumulation of a molecule of interest in trichome cells and/or to facilitate purification of the molecule from trichome cells, methods to down-regulate at least one molecule endogenous to the plant trichomes can be employed. Trichomes are known to contain a number of compounds and metabolites that interfere with the production of other molecules in the trichome cells. These compounds and metabolites include, for example, proteases, polyphenol oxidase (PPO), polyphenols, ketones, terpenoids, and alkaloids. The down-regulation of such trichome components has been described. *See, e.g.*, U.S. Patent No. 7,498,428.

EXAMPLES

[0124] The following examples are provided by way of illustration only and not by way of limitation. Those of skill in the art will readily recognize a variety of non-critical parameters that could be changed or modified to yield essentially the same or similar results. The examples should in no way be construed as limiting the scope of the present technology, as defined by the appended claims.

Example 1: Identifying trichome specific promoters

[0125] The nucleic acid sequences of the (i) olivetol synthase (OLS) promoter, (ii) OLS1 promoter, and (iii) OLS2 promoter are set forth in SEQ ID NOs: 1, 2, and 3, respectively, and the open reading frame (ORF) of OLS is set forth in SEQ ID NO: 4. The nucleic acid sequence of the olivetolic acid cyclase (OAC) promoter is set forth in SEQ ID NO: 5, the nucleic acid sequence of the OAC1 promoter is set forth in SEQ ID NO: 6, and the ORF of OAC is set forth in SEQ ID NO: 7. The nucleic acid sequence of aromatic prenyltransferase (PT) promoter is set forth in SEQ ID NO: 8, the nucleic acid sequence of the PT1 promoter is set forth in SEQ ID NO: 9, and the ORF of PT is set forth in SEQ ID NO: 10. The nucleic acid sequences of the (i) hexanoyl-CoA synthetase (AAE1-1) promoter, (ii) hexanoyl-CoA synthetase (AAE1-1′) promoter; (iii) hexanoyl-CoA synthetase (AAE3) promoter, and (iv) hexanoyl-CoA synthetase (AAE12) promoter are set forth in SEQ ID NOs: 11, 12, 13, and 14, respectively, and the ORF of hexanoyl-CoA (AAE-1) is set forth in SEQ ID NO: 15. The nucleic acid sequences of the (i) CBDA synthase (CBDAS) promoter, (ii) CBDAS synthase 1 (CBDAS1) promoter, (iii) CBDA synthase (CBDAS) 20800 promoter, and (iv) CBDA synthase (CBDAS) 20800′ promoter are set forth in SEQ ID NOs: 16, 17, 18, and 19, respectively, and the nucleic acid sequence of the ORF of CBDA synthase (CBDAS) is set forth in SEQ ID NO: 20. The nucleic acid sequences of (i) THCA synthase (THCAS) 19603 promoter, (ii) THCA synthase (THCAS) 19603′ promoter, (iii) THCA synthase (THCAS) 50320 promoter, (iv) THCA synthase (THCAS) 50320′ promoter, (v) THCA synthase (THCAS) 1330 promoter, and (vi) THCA synthase (THCAS) 1330′ promoter are set forth in SEQ ID NOs: 21, 22, 23, 24, 25, and 26, respectively, and the ORF of THCAS is set forth in SEQ ID NO: 27. The nucleic acid sequence of the CBCA synthase (CBCAS) 3498 promoter is set forth in SEQ ID NO: 28, the nucleic acid sequence of the CBCA synthase (CBCAS) 3498′ promoter is set forth in SEQ ID NO: 29, and the ORF of CBCA synthase (CBCAS) is set forth in SEQ ID NO: 30.

[0126] Trichome specific promoters were identified by searching the draft genome sequence of *cannabis sativa* using the BLAT search facility with the coding regions of the biosynthetic enzyme genes in the cannabinoid biosynthetic pathway. For each genomic sequence hit, a gene prediction program was run (using the FGENESH program) to establish the first ATG. The start of transcription was then established by comparing the genomic sequence to the longest available cDNA sequences in the NCBI NR database. Both multiple sequence alignments (for the CBDA and THCA synthase genes) and querying the PLACE database established TATA Box regions. With the start codon, start of transcription, and TATA Box established, the location and sequence of the promoters was verified.

Example 2: Trichome specific promoters for directing cannabinoid production in *Nicotiana tabacum*

[0127] Cannabinoids are synthesized and accumulated in cannabis trichomes. Olivetol synthase (OLS), olivetolic acid cyclase (OAC), aromatic prenyltransferase (PT), hexanoyl-CoA synthetase (AAE1-1), CBDA synthase, CBCA synthase, and THCA synthase are enzymes of the cannabinoid biosynthetic pathway. Accordingly, it is expected that the promoters for each of these enzymes will direct the expression of coding nucleic acids in trichome cells. This example demonstrates the use of the trichome specific promoters of the present technology, or biologically active fragments thereof, to express cannabinoid biosynthetic enzymes in plants and plant cells that do not naturally produce cannabinoids.

Methods

[0128] *Vector constructs.* Promoter sequences (SEQ ID NOs: 1-3, 5, 6, 8, 9, 11-14, 16-19, 21-26, 28, 29, or 31-33) are placed in front of a GUS-A marker in a vector adapted for expression in a *Nicotiana tabacum* cell, such as a Ti plasmid vector. The constructs are incorporated into *Agrobacterium tumafaciens* and used to transform *N. tabacum* according to methods known in the art. Constructs are transformed and regenerated under kanamycin selection and primary regenerants (T_0) are grown to seed.

[0129] As a control, a construct containing the tobacco *NtCPS2* promoter is transformed into tobacco. The *NtCPS2* promoter has been shown to be highly effective in directing trichome-specific expression in *N. tabacum* (Sallaud et al., *The Plant Journal* 72:1-17 (2012)).

[0130] *Expression analysis.* Quantitative and qualitative β-glucuronidase (GUS) activity analyses are performed on T_1 plants. Qualitative analysis of promoter activity is carrier out using histological GUS assays and by visualization of the Green Fluorescent Protein (GFP) using a fluorescence microscope. For GUS assays, various plant parts are incubated overnight at 37 °C in the presence of atmospheric oxygen with Xglue (5-Bromo-4-chloro-3-indolyl-β-D-glucuronide cyclohexylamine salt) substrate in phosphate buffer (1 mg/mL, K_2HPO_4, 10 μM, pH 7.2, 0.2% Triton X-100). The samples are de-stained by repeated washing with ethanol. Non-transgenic plants are used as negative controls. It is anticipated that trichomes of transgenic plants with OLS:GUS, OAC:GUS, PT:GUS, AAEE1-1:GUS, CBDAS:GUS, CBCAS:GUS, and THCAS:GUS will show bright blue trichomes whereas the non-trichome tissues of these transgenic plants and the trichomes of non-transgenic control plants will not be colored.

[0131] Quantitative analysis of promoter activity is carrier out using a fluorometric GUS assay. Total protein samples are prepared from young leaf material; samples are prepared from pooled leaf pieces. Fresh leaf material is ground in PBS using metal beads followed by centrifugation and collection of the supernatant.

Results

[0132] Plants and plant cells genetically engineered with expression vectors comprising the promoters of the present technology, or biologically active fragments thereof, exhibit trichome specific transcriptional activity. As shown in **FIG. 2B**, trichomes of a transgenic plant transformed with CBDAS 20800:GUS show bright blue trichomes whereas the non-trichome tissues of the plant are not colored. Trichome specific expression for the other promoters in the cannabinoid biosynthetic pathway are qualitatively similar (**FIG. 3**). Results shown in **FIG. 3** demonstrate the trichome specific expression of promoters from the complete cannabinoid biosynthetic pathway in tobacco trichomes. Accordingly, these results demonstrate that the promoters from the complete cannabinoid biosynthetic pathway as described herein are useful for preferentially directing expression of an operably linked gene in trichome tissue, as compared to expression in the root, leaf, stem, or other tissues of a plant. This trichome-specific expression will be a crucial tool for the manipulation of the biosynthesis of trichome-specific biochemical compounds such as the cytotoxic cannabinoids. In addition, these promoters will be crucial to strategies aimed at using

tobacco or cannabis trichomes as biofactories for the controlled production of specific biochemical compounds.

Example 3: Identifying "Cannabinoid On" ("CANON") fragments for directing trichome specific expression in *Nicotiana tabacum*

[0133] The nucleic acid sequence of the "Cannabinoid On" ("CANON") fragment is set forth in SEQ ID NO: 31. The nucleic acid sequence of the CANON fragment together with the putative TATA Box, 5' UTR, and start codon is set forth in SEQ ID NO: 32. The nucleic acid sequence of the 4 x CANON fragment synthetic promoter is set forth in SEQ ID NO: 33. The consensus CANON fragment is derived from the trichome specific promoters from THCA synthases 19603, 1330, 50320, CBCA synthase 3498, and CBDA synthase 20800.

Methods

[0134] CANON fragments were identified by searching the purple Kush genome sequence using the BLAT search facility. A BLAT search of the purple Kush genome sequence using a THCA synthase gene yielded the following results:

BLAT Search Results

ACTIONS	QUERY	SCORE	START	END	QSIZE	IDENTITY	CHRO	STRAND	START	END	SPAN
browser details	THCA	979	1	979	979	100.0%	scaffold19603	+	6901	7879	979
browser details	THCA	489	257	979	979	89.4%	scaffold50320	-	8585	9462	878
browser details	THCA	466	257	979	979	89.1%	scaffold1330	-	2572	3369	798
browser details	THCA	332	593	979	979	94.2%	scaffold3498	+	9881	10270	390
browser details	THCA	275	590	979	979	85.6%	scaffold6274	+	24419	24803	385
browser details	THCA	271	590	979	979	86.5%	scaffold39155	-	4079	4464	386
browser details	THCA	251	1	258	979	97.7%	scaffold17297	+	9858	10113	256
browser details	THCA	222	590	979	979	86.0%	scaffold46100	-	1795	2180	386
browser details	THCA	206	3	271	979	89.0%	scaffold33063	-	10241	10479	239
browser details	THCA	196	575	979	979	91.2%	scaffold20800	-	4646	5042	397
browser details	THCA	167	69	267	979	94.7%	scaffold12036	+	133813	134227	415
browser details	THCA	156	89	258	979	96.5%	scaffold17775	-	18105	18274	170
browser details	THCA	142	260	514	979	87.5%	scaffold59317	+	863	1300	438
browser details	THCA	136	120	268	979	96.6%	scaffold12751	-	46584	46732	149
browser details	THCA	128	120	257	979	97.1%	scaffold87266	-	12096	12233	138
browser details	THCA	128	120	257	979	97.1%	scaffold23985	+	18853	18990	138
browser details	THCA	127	120	258	979	96.4%	scaffold14882	+	59722	59860	139
browser details	THCA	123	98	269	979	85.9%	scaffold4283	-	23256	23416	161

[0135] The highlighted promoters were selected for further analysis and the sequences of the five promoters were compared. The five promoters showed high sequence similarity from the start of translation to about 160 base pairs upstream of the putative TATA Box, but little sequence similarity upstream of the TATA Box.

[0136] A consensus sequence was derived from the five promoters and the sequence was used to test for trichome specific promoter activity in tobacco trichomes using a synthetic

promoter termed the "4 x CANON fragment synthetic promoter." The 4 x CANON fragment synthetic promoter comprises four copies of the consensus CANON fragment, which is not found in nature, in front of one copy of the minimal promoter (TATA Box, start of transcription and first ATG). *N. tabacum* was transformed with the 4 x CANON fragment synthetic promoter placed in front of a GUS-A marker as described in Example 2 and GUS expression analyses were performed.

Results

[0137] As shown in **FIG. 4B**, the 4 x CANON fragment synthetic promoter directs expression in the trichome. Accordingly, these results demonstrate that the CANON fragment is sufficient to direct trichome specific expression in a gain-of-function promoter and is therefore responsible for the trichome specific expression of a number of cannabinoid biosynthetic enzyme genes including THCA, CBDA, and CBCA synthases.

EQUIVALENTS

[0138] The present technology is not to be limited in terms of the particular embodiments described in this application, which are intended as single illustrations of individual aspects of the present technology. Many modifications and variations of this present technology can be made without departing from its spirit and scope, as will be apparent to those skilled in the art. Functionally equivalent methods and apparatuses within the scope of the present technology, in addition to those enumerated herein, will be apparent to those skilled in the art from the foregoing descriptions. Such modifications and variations are intended to fall within the scope of the appended claims. The present technology is to be limited only by the terms of the appended claims, along with the full scope of equivalents to which such claims are entitled. It is to be understood that this present technology is not limited to particular methods, reagents, compounds compositions or biological systems, which can, of course, vary. It is also to be understood that the terminology used herein is for the purpose of describing particular embodiments only, and is not intended to be limiting.

[0139] In addition, where features or aspects of the disclosure are described in terms of Markush groups, those skilled in the art will recognize that the disclosure is also thereby described in terms of any individual member or subgroup of members of the Markush group.

[0140] As will be understood by one skilled in the art, for any and all purposes, particularly in terms of providing a written description, all ranges disclosed herein also encompass any and all possible subranges and combinations of subranges thereof. Any listed range can be easily recognized as sufficiently describing and enabling the same range being broken down into at least equal halves, thirds, quarters, fifths, tenths, *etc.* As a non-limiting example, each range discussed herein can be readily broken down into a lower third, middle third and upper third, *etc.* As will also be understood by one skilled in the art all language such as "up to," "at least," "greater than," "less than," and the like, include the number recited and refer to ranges which can be subsequently broken down into subranges as discussed above. Finally, as will be understood by one skilled in the art, a range includes each individual member. Thus, for example, a group having 1-3 cells refers to groups having 1, 2, or 3 cells. Similarly, a group having 1-5 cells refers to groups having 1, 2, 3, 4, or 5 cells, and so forth.

[0141] All publicly available documents referenced or cited to herein, such as patents, patent applications, provisional applications, and publications, including GenBank Accession Numbers, are incorporated by reference in their entirety, including all figures and tables, to the extent they are not inconsistent with the explicit teachings of this specification.

[0142] Other embodiments are set forth within the following claims.

SEQUENCE LISTING

SEQ ID NO: 1 (554 bp)

Tetraketide synthase/olivetol synthase (OLS) promoter

AAAAATAAAATTAATAAATTTTTAATTATTAATATCATTTTATTTTTTAA
TAAATAAATATCATTTTTATATTATTATAATATGTATAAAGTTTTAATTG
TATACAAGAAGTCTTATAGTAAGAGTATACACCTTACATCATAATAACTA
CTCGATCTGAAATCAATGGTCAAGAAAAGTTCCCTACCGGTAGGAAACTT
TTGCTAGATCCTACCATAGTCTTCCCTTATATTTATTATGTAGAATCTAT
TATTATATCTAATAATTAACAAATATTAACAAATCATTTTGAAAAATTAT
ATTAAAAAAAAAACTTGAAAAGTCAAAGATTAACCATCAATTTGCCAAAT
CAAATTAGTGAGAAAGTAGGTATTATATACCTAACACTCATTTTACATAT
GTCTAGACGTTTATATGTATAGCGTTGTTGTGTGTAATAAGTTCACTTGT
AGTATCTTGTACATACATAATATATATATATAGGTGTGTTTGTGAACAAT
TATATTATCACACATACACAACTCATTATTATCATAATAATAATAATACC
ATGG

SEQ ID NO: 2 (550 bp)

Tetraketide synthase/olivetol synthase 1 (OLS1) promoter

AAAAATAAAATTAATAAATTTTTAATTATTAATATCATTTTATTTTTTAA
TAAATAAATATCATTTTTATATTATTATAATATGTATAAAGTTTTAATTG
TATACAAGAAGTCTTATAGTAAGAGTATACACCTTACATCATAATAACTA
CTCGATCTGAAATCAATGGTCAAGAAAAGTTCCCTACCGGTAGGAAACTT
TTGCTAGATCCTACCATAGTCTTCCCTTATATTTATTATGTAGAATCTAT
TATTATATCTAATAATTAACAAATATTAACAAATCATTTTGAAAAATTAT
ATTAAAAAAAAAACTTGAAAAGTCAAAGATTAACCATCAATTTGCCAAAT
CAAATTAGTGAGAAAGTAGGTATTATATACCTAACACTCATTTTACATAT
GTCTAGACGTTTATATGTATAGCGTTGTTGTGTGTAATAAGTTCACTTGT
AGTATCTTGTACATACATAATATATATATATAGGTGTGTTTGTGAACAAT
TATATTATCACACATACACAACTCATTATTATCATAATAATAATAATAAT

SEQ ID NO: 3 (558 bp)

Tetraketide synthase/olivetol synthase 2 (OLS2) promoter

AAAAATAAAATTAATAAATTTTTAATTATTAATATCATTTTATTTTTTAA
TAAATAAATATCATTTTTATATTATTATAATATGTATAAAGTTTTAATTG
TATACAAGAAGTCTTATAGTAAGAGTATACACCTTACATCATAATAACTA
CTCGATCTGAAATCAATGGTCAAGAAAAGTTCCCTACCGGTAGGAAACTT
TTGCTAGATCCTACCATAGTCTTCCCTTATATTTATTATGTAGAATCTAT
TATTATATCTAATAATTAACAAATATTAACAAATCATTTTGAAAAATTAT
ATTAAAAAAAAAACTTGAAAAGTCAAAGATTAACCATCAATTTGCCAAAT
CAAATTAGTGAGAAAGTAGGTATTATATACCTAACACTCATTTTACATAT
GTCTAGACGTTTATATGTATAGCGTTGTTGTGTGTAATAAGTTCACTTGT
AGTATCTTGTACATACATAATATATATATATATAGGTGTGTTTGTGAACA
ATTATATTATCACACATACACAACTCATTATTATTAGCATAATAATAATA
ATAATAAT

SEQ ID NO: 4 (1158 bp)

Olivetol synthase (OLS) ORF

```
   1 atgaatcatc ttcgtgctga gggtccggcc tccgttctcg ccattggcac cgccaatccg
  61 gagaacattt tattacaaga tgagtttcct gactactatt ttcgcgtcac caaaagtgaa
 121 cacatgactc aactcaaaga aaagtttcga aaaatatgtg acaaaagtat gataaggaaa
 181 cgtaactgtt tcttaaatga agaacaccta aagcaaaacc caagattggt ggagcacgag
 241 atgcaaactc tggatgcacg tcaagacatg ttggtagttg aggttccaaa acttgggaag
 301 gatgcttgtg caaaggccat caaagaatgg ggtcaaccca agtctaaaat cactcattta
 361 atcttcacta gcgcatcaac cactgacatg cccggtgcag actaccattg cgctaagctt
 421 ctcggactga gtccctcagt gaagcgtgtg atgatgtatc aactaggctg ttatggtggt
 481 ggaaccgttc tacgcattgc caaggacata gcagagaata acaaaggcgc acgagttctc
 541 gccgtgtgtt gtgacataat ggcttgcttg tttcgtgggc cttcagagtc tgacctcgaa
 601 ttactagtgg gacaagctat ctttggtgat ggggctgctg cggtgattgt tggagctgaa
 661 cccgatgagt cagttgggga aaggccgata tttgagttgg tgtcaactgg gcaaacaatc
 721 ttaccaaact cggaaggaac tattgggggа catataaggg aagcaggact gatatttgat
 781 ttacataagg atgtgcctat gttgatctct aataatattg agaaatgttt gattgaggca
 841 tttactccta ttgggattag tgattggaac tccatatttt ggattacaca cccaggtggg
 901 aaagctattt tggacaaagt ggaggagaag ttgcatctaa agagtgataa gtttgtggat
 961 tcacgtcatg tgctgagtga gcatgggaat atgtctagct caactgtctt gtttgttatg
1021 gatgagttga ggaagaggtc gttggaggaa gggaagtcta ccactggaga tggatttgag
1081 tggggtgttc tttttgggtt tggaccaggt ttgactgtcg aaagagtggt cgtgcgtagt
1141 gttcccatca aatattaa
```

SEQ ID NO: 5 (996 bp)

Olivetolic acid cyclase (OAC) promoter

AATTTTTGACAATTTTTTTTTAATATAACTAACTTGAAGATAATTCCTAAT
ACGAATAAATACAGAAAATATAAACAGTTTTGTTATAACACTTTTAGATC
AGATTATAATTAAATTTTATATTTTTGATAAAAAAATCAATTGAGGGTCC
TATTTGTACCATTTTAGAAAATATAGTGTCCATTTTTATTATTTAGTAAA
ACAGAGAGTCTAATGCGTAACTTTTACAAAAATACATGGTCCAAAATAGT
ATTTACCCTTTAATTTTATTAGCAACTGTCCCAAAAAAAATATGTTTTGA
TGACTGACGTACGGGAATGTAAAGTTTTGTAATAGTGTATTGATTCAAAG
ACAACATAAAGACATCAATCTGAAATCGATTTCAATGTCCCAAAAACACA
TTAATGAGCCTTAATTGCATCCGATAATCATTTTCCATTGGTTTTATTAC
TTCTCATATATATAGACATCACATATATGATAGGATTTCTTGAGAATAAT
GTTAACATTGAGATTTTTATTACAACTGATATATTGATTATGCTTAGATG
GCTTGAATTTGAGCGACATATATAGAAAGAGTATAGAATGATATATATGC
ACATCCAAAATATGTACCAAAATATTAGGTGACAATTTAGATAAATGGAA
GAGAAAATTATAGAAAAATGAGGGTTTATATTTTTGTGGTTATTTAATTT
TAGACATAGTGTGATCCAATAATTTTAGGATGTATAATTGTTAGGCACAT
GGATTACTTGTTTTTTTATTAAGTATAACCTTACAAAGTAGATGGTAGTA
ATTAATGTAGAAGGTTCCAATAATGTATTTATATAAATTGTTAGGCATGC
AAAGCCTAATTAATTAATAAATGAGGGTGGCCAATGGCCACTATATATAT
CAAGGCATCGACTGTATGTAGCATAATGTGATTTATATAATTATCAAAAA
AAAAAAATAAAAATAAGAAGAAGAAGAAAGTTGAGAAAGACCATGG

SEQ ID NO: 6 (992 bp)

Olivetolic acid cyclase 1 (OAC1) promoter

AATTTTTGACAATTTTTTTTTAATATAACTAACTTGAAGATAATTCCTAAT
ACGAATAAATACAGAAAATATAAACAGTTTTGTTATAACACTTTTAGATC
AGATTATAATTAAATTTTATATTTTTGATAAAAAAATCAATTGAGGGTCC
TATTTGTACCATTTTAGAAAATATAGTGTCCATTTTTATTATTTAGTAAA
ACAGAGAGTCTAATGCGTAACTTTTACAAAAATACATGGTCCAAAATAGT
ATTTACCCTTTAATTTTATTAGCAACTGTCCCAAAAAAAATATGTTTTGA
TGACTGACGTACGGGAATGTAAAGTTTTGTAATAGTGTATTGATTCAAAG
ACAACATAAAGACATCAATCTGAAATCGATTTCAATGTCCCAAAAACACA
TTAATGAGCCTTAATTGCATCCGATAATCATTTTCCATTGGTTTTATTAC
TTCTCATATATATAGACATCACATATATGATAGGATTTCTTGAGAATAAT
GTTAACATTGAGATTTTTATTACAACTGATATATTGATTATGCTTAGATG
GCTTGAATTTGAGCGACATATATAGAAAGAGTATAGAATGATATATATGC
ACATCCAAAATATGTACCAAAATATTAGGTGACAATTTAGATAAATGGAA
GAGAAAATTATAGAAAAATGAGGGTTTATATTTTTGTGGTTATTTAATTT
TAGACATAGTGTGATCCAATAATTTTAGGATGTATAATTGTTAGGCACAT
GGATTACTTGTTTTTTTATTAAGTATAACCTTACAAAGTAGATGGTAGTA
ATTAATGTAGAAGGTTCCAATAATGTATTTATATAAATTGTTAGGCATGC
AAAGCCTAATTAATTAATAAATGAGGGTGGCCAATGGCCACTATATATAT
CAAGGCATCGACTGTATGTAGCATAATGTGATTTATATAATTATCAAAAA
AAAAAAATAAAAATAAGAAGAAGAAGAAAGTTGAGAAAGAGA

SEQ ID NO: 7 (485 bp)

Olivetolic acid cyclase (OAC) ORF

```
   1 aaaaaagaag aagaagaaga aagttgagaa agagaatggc agtgaagcat ttgattgtat
  61 tgaagttcaa agatgaaatc acagaagccc aaaaggaaga atttttcaag acgtatgtga
 121 atcttgtgaa tatcatccca gccatgaaag atgtatactg gggtaaagat gtgactcaaa
 181 agaataagga agaagggtac actcacatag ttgaggtaac atttgagagt gtggagacta
 241 ttcaggacta cattattcat cctgcccatg ttggatttgg agatgtctat cgttctttct
 301 gggaaaaact tctcattttt gactacacac cacgaaagta gactatatat agtagccgac
 361 caagctgcct tcatcttcat cttctcaaat aatatatcta atatctaatt atataataat
 421 aactacttaa taaaagactg tgtttataac attaaataat aataataata aagtcttttg
 481 tagct
```

SEQ ID NO: 8 (1361 bp)

Aromatic prenyltransferase (PT) promoter

TCAAACCAGATTTTTTATAATAATTT
ACACACCAATTAACTAAGAAATTCTAACTAGAGTTAGTGTTGAATTAGAA
GGAGTATCAAGGTAACCTATCTCCTTTGTCTCTTTAAATTTTTAAAAGAA
AAGCTaTTGTTGTAATTAGTTTGGATTTTATGGGTTAGGTTTTGTAGAGT
TTGTTTGAGCTTAAAATTGTTGTTGGTTTGATTTAAGGTTGAATTCTGTG
TTGGTTTGTAAGATTTTAGCTGATGTAGTTCGTGTTGGAGGAGTTCGAGT
TTTTGAACTCAAAGTCCACTCAACAATATGGTGAAATTAGTTTTAGGTGG
TTTTAGATGCTTTTGAAGTGTATGTTTTGTTTAGATTACCCTCTTTTTGG
TAAGTTTGATAGCCTTTGGTCAAGGAATGACTTAATTTTGAAGGTCTGAA
AGTTGAGTAGTTTTGAACTCTTTTCTGCCCAAAATGGTTTCGAATGAGTT
GTACTTGTCATATGTGATTTGAGTGGTAAGGTTCGGAAGGGTAGTCCTAG
ACAGTGTTTTCAATGATTAGGTTTGGGGTTGGTGAATCTAGAGTTAAGAG
CCTAGAGTTTTGGTTGCTTGGATTGGTGTGACCTGGTTGTTAGTATCCTT
ATCAAGGGATCCCGAATAGGTGATGAGTGAGGTTTGGATGGTATATTCTG
GATCAAAATTTTTAGATGGGGTCCAAACTGATAGTTTCAAAAATGTGTGT
GTGTGTCTGGTTAAGGATTTATTTTGGGAAGGCTCTAGTTGGGAGGTCCT
GGATTGGAAGTTCTAGATGTGAGGCTCCGGGTAGGAGGGTCCTGTAGGAA
GATTCTAGATGGTATTTGGTGGTCGAGATGAGGGATTTCCGAACATGTTT
TAATAGTTACAATAATTCATAAATTTTCAACACTAAATCAAAATATTAAT
TTTGATATGATTGTATAAATTTAATCTCATCGTAGTGTAGAGCTCTACAT
GCATGCATGGTTTGTATAGCCATAATTAATTAATTAGGCAAGAAAGATAG
ATGTGCCTACCATGCATGACTTCTAAAACATGGATGACTTTAAATTTTGA
AACCAAGCAAAAATTGTTTTACTATAAATCTTGTTAATTTTTCATTATTG
ATATCACTATGTTATTCTAACCCTTTCGTAGCCATAATTAGTTAATTAGG
CAAGAAAGATAGATGTGCCTACCATGCATGACTTCTAAAACATGGATGAC
TTTAAATTTTGAAACCAAGCAAAAATTATTTTACTATAAATCTTGTTATT
TTTTCATTATTGATATCACTATGTTATTCTAACCCTTTCTTTCCCTCATT
TTTTCTTAATATTCAATCAATAATAATCTCCATGG

SEQ ID NO: 9 (1357 bp)

Aromatic prenyltransferase 1 (PT1) promoter

TCAAACCAGATTTTTTATAATAATTT
ACACACCAATTAACTAAGAAATTCTAACTAGAGTTAGTGTTGAATTAGAA
GGAGTATCAAGGTAACCTATCTCCTTTGTCTCTTTAAATTTTTAAAAGAA
AAGCTaTTGTTGTAATTAGTTTGGATTTTATGGGTTAGGTTTTGTAGAGT
TTGTTTGAGCTTAAAATTGTTGTTGGTTTGATTTAAGGTTGAATTCTGTG
TTGGTTTGTAAGATTTTAGCTGATGTAGTTCGTGTTGGAGGAGTTCGAGT
TTTTGAACTCAAAGTCCACTCAACAATATGGTGAAATTAGTTTTAGGTGG
TTTTAGATGCTTTTGAAGTGTATGTTTTGTTTAGATTACCCTCTTTTTGG
TAAGTTTGATAGCCTTTGGTCAAGGAATGACTTAATTTTGAAGGTCTGAA
AGTTGAGTAGTTTTGAACTCTTTTCTGCCCAAAATGGTTTCGAATGAGTT
GTACTTGTCATATGTGATTTGAGTGGTAAGGTTCGGAAGGGTAGTCCTAG
ACAGTGTTTTCAATGATTAGGTTTGGGGTTGGTGAATCTAGAGTTAAGAG
CCTAGAGTTTTGGTTGCTTGGATTGGTGTGACCTGGTTGTTAGTATCCTT
ATCAAGGGATCCCGAATAGGTGATGAGTGAGGTTTGGATGGTATATTCTG
GATCAAAATTTTTAGATGGGGTCCAAACTGATAGTTTCAAAAATGTGTGT

GTGTGTCTGGTTAAGGATTTATTTTGGGAAGGCTCTAGTTGGGAGGTCCT
GGATTGGAAGTTCTAGATGTGAGGCTCCGGGTAGGAGGGTCCTGTAGGAA
GATTCTAGATGGTATTTGGTGGTCGAGATGAGGGATTTCCGAACATGTTT
TAATAGTTACAATAATTCATAAATTTTCAACACTAAATCAAAATATTAAT
TTTGATATGATTGTATAAATTTAATCTCATCGTAGTGTAGAGCTCTACAT
GCATGCATGGTTTGTATAGCCATAATTAATTAATTAGGCAAGAAAGATAG
ATGTGCCTACCATGCATGACTTCTAAAACATGGATGACTTTAAATTTTGA
AACCAAGCAAAAATTGTTTTACTATAAATCTTGTTAATTTTTCATTATTG
ATATCACTATGTTATTCTAACCCTTTCGTAGCCATAATTAGTTAATTAGG
CAAGAAAGATAGATGTGCCTACCATGCATGACTTCTAAAACATGGATGAC
TTTAAATTTTGAAACCAAGCAAAAATTATTTTACTATAAATCTTGTTATT
TTTTCATTATTGATATCACTATGTTATTCTAACCCTTTCTTTCCCTCATT
TTTTCTTAATATTCAATCAATAATAATCTCC

SEQ ID NO: 10 (1188 bp)

Aromatic prenyltransferase (PT) ORF

```
   1 atgggactct catcagtttg tacctttca tttcaaacta attaccatac tttattaaat
  61 cctcacaata ataatcccaa aacctcatta ttatgttatc gacaccccaa aacaccaatt
 121 aaatactctt acaataattt tccctctaaa cattgctcca ccaagagttt tcatctacaa
 181 aacaaatgct cagaatcatt atcaatcgca aaaaattcca ttagggcagc tactacaaat
 241 caaactgagc ctccagaatc tgataatcat tcagtagcaa ctaaaatttt aaactttggg
 301 aaggcatgtt ggaaacttca aagaccatat acaatcatag catttacttc atgcgcttgt
 361 ggattgtttg ggaaagagtt gttgcataac acaaatttaa taagttggtc tctgatgttc
 421 aaggcattct ttttttggt ggctatatta tgcattgctt cttttacaac taccatcaat
 481 cagatttacg atcttcacat tgacagaata aacaagcctg atctaccact agcttcaggg
 541 gaaatatcag taaacacagc ttggattatg agcataattg tggcactgtt tggattgata
 601 ataactataa aaatgaaggg tggaccactc tatatatttg gctactgttt tggtattttt
 661 ggtgggattg tctattctgt tccaccattt agatggaagc aaaatccttc cactgcattt
 721 cttctcaatt tcctggccca tattattaca aatttcacat tttattatgc cagcagagca
 781 gctcttggcc taccatttga gttgaggcct tcttttactt tcctgctagc atttatgaaa
 841 tcaatgggtt cagctttggc tttaatcaaa gatgcttcag acgttgaagg cgacactaaa
 901 tttggcatat caaccttggc aagtaaatat ggttccagaa acttgacatt attttgttct
 961 ggaattgttc tcctatccta tgtggctgct atacttgctg ggattatctg gccccaggct
1021 ttcaacagta acgtaatgtt actttctcat gcaatcttag cattttggtt aatcctccag
1081 actcgagatt ttgcgttaac aaattacgac ccggaagcag gcagaagatt ttacgagttc
1141 atgtggaagc tttattatgc tgaatattta gtatatgttt tcatataa
```

SEQ ID NO: 11 (805 bp)

Hexanoyl-CoA synthetase (AAE1-1) promoter

ATAATGTTATTTATTATTTAATACAATATGACATTTATACAGCGTACCAA
ACGAGCCCATAATGTATTGCAATTGTATTGTTGTACGAATTAAAATAATA
AAATACTAATTAAGTTAATATTGTAAGATAGATAATATGGCAAACAAATA
AATAATTTTCAGTTGAGAATATGTACAAAATTAAAGTTGATTTTACAAGA
AACAAGTAGTACTAATTAAAATGGTAGATTTACATATAGTGTTCGTTAAA
TTTGAAAAAAGAAAACGACTAATTCATTGGAAAATTGATTGTGGAAGAAA
TCAAAATTTGGTGGCTGAACCGCCCTAGTTACTTAACACAATTCATAAAA
CAACCCTACACTACTTACCCTACACTATATACGTATATTTTTTTTTTATA
TTCCATTTGTTAATAAAATAAATATGCATTGTTATTATGTGTACCAAGAT
GCTAATACGGTGACGTTTTATATAAGTTTATAATTTTTTATTAAAGTAAA
ACAATTGTCCCTCAATATTAAATTAAAGAAAATAAGTAGGTTTTAGCATT
TCTCATGATAAATAAGTAGGAGTAGTAAGTATATTAATTTATATGCATAT
GTTAAAAGGCAACACAGTGATAGGTTGTGTAGTACGTACCCAAACCAAAT
TCCAGCCTAATCCATTATTAATGAGCAAGTCCTTGATAAAATAAGGTGCA
TGTTGTTAGATCAAATAACCCGTCCCGAAATCCAATATATATATATATAT
AATATTCAAACTCTCTCTTTCTATCTTCGTACAATTTAATAGAAGTAACC
ATGGG

SEQ ID NO: 12 (800 bp)

Hexanoyl-CoA synthetase (AAE1-1′) promoter

ATAATGTTATTTATTATTTAATACAATATGACATTTATACAGCGTACCAA
ACGAGCCCATAATGTATTGCAATTGTATTGTTGTACGAATTAAAATAATA
AAATACTAATTAAGTTAATATTGTAAGATAGATAATATGGCAAACAAATA
AATAATTTTCAGTTGAGAATATGTACAAAATTAAAGTTGATTTTACAAGA
AACAAGTAGTACTAATTAAAATGGTAGATTTACATATAGTGTTCGTTAAA
TTTGAAAAAAGAAAACGACTAATTCATTGGAAAATTGATTGTGGAAGAAA
TCAAAATTTGGTGGCTGAACCGCCCTAGTTACTTAACACAATTCATAAAA
CAACCCTACACTACTTACCCTACACTATATACGTATATTTTTTTTTTATA
TTCCATTTGTTAATAAAATAAATATGCATTGTTATTATGTGTACCAAGAT
GCTAATACGGTGACGTTTTATATAAGTTTATAATTTTTTATTAAAGTAAA
ACAATTGTCCCTCAATATTAAATTAAAGAAAATAAGTAGGTTTTAGCATT
TCTCATGATAAATAAGTAGGAGTAGTAAGTATATTAATTTATATGCATAT
GTTAAAAGGCAACACAGTGATAGGTTGTGTAGTACGTACCCAAACCAAAT
TCCAGCCTAATCCATTATTAATGAGCAAGTCCTTGATAAAATAAGGTGCA
TGTTGTTAGATCAAATAACCCGTCCCGAAATCCAATATATATATATATAT
AATATTCAAACTCTCTCTTTCTATCTTCGTACAATTTAATAGAAGTAATA

SEQ ID NO: 13 (1000 bp)

Hexanoyl-CoA synthetase (AAE3) promoter

TAAAAATTAGATTATATTTTTATCAGAATTTTTTATTTATATTTTTCATC
GAGTTACGTTGATTTAACGTATGTTGACGGTGCTTATAATTGTCTATTAT
TTTTAAGTTATATACACGCACTACTTACGTAAATTTAGTTGGCAGATTTT
CTGTCTCTATGGGCTATATTGTTGGACATCAGCTATCTCATTAATTAATA
ATACCATACTTTTTCTTAAAAGTTAAATAAAACAAAATTATACACGCACT
GATTCGAATGGTGTAGCCCACGCACTTGCTAAGTCTGTTACTAGTTCTGT
AGGAGTTAATGTGTGGACTGATGCTTGTCCTAGATTTTATCAGCTATCTC

ATTAATTAATAATACCATACTTTTTCTTAAAAGTTAAATAAAACAAAATT
ATACACGCACTACGATAAGGATTATATAAATGATTAACTAAAAAATATCT
ATTTAACATGCCATACCGCGAAAGGTTTTGCGCCATTCGATGGTGTCCGG
GATCTCGACGCTCTCCCTTATGCGACTCCTGCATTAGGAAGCAGCCCAGT
AGTAGGTTGAGGCCGTTGAGCACCGCCGCCGCAAGGAATGGTGCATGCAA
GGAGATGGCGCCCAACAGTCCCCCGGCCACGGGGCCTGCCACCATACCCA
CGCCGAAACAAGCGCTCATGAGCCCGAAGTGGCGAGCCCGATCTTCCCCA
TCGGTGATGTCGGCGATATAGGCGCCAGCAACCGCACCTGTGGCGCCGGT
GATGCCGGCCACGATGCGTCCGGCGTAGAGGATCGAGATCTCGATCCCGC
GAAATTAATACGACTCACTATAGGGGAATTGTGAGCGGATAACAATTCCC
CTCTAGAAATAATTTTGTTTAACTTTAAGAAGGAGATATACCATGAAACA
CCACCACCACCACCACCACCACGGTGGTCTGGTTCCGCGTGGTTCCCATG
ATCAAACAAGTTTGTACAAAAAAGCAGGCTCCGAATTCGCCCTTATGGAG

SEQ ID NO: 14 (869 bp)

Hexanoyl-CoA synthetase (AAE12) promoter

GTGACAAACTGAAGAATTCTGTTTATTTTATTTTATACATT
AGTGATGTTTTTTACTTTTTTAGAGTTATTTTTATGCCTATGTAATTTGC
ATATTGTAGTGGAGTTTATGCAGTTATTTTTTTGTTGTTATTTAGTTGTT
TAACTAAGAGAAATAAATATATATATTTTTTATTTGTGCGTTGTTTTTAT
TTTATTTTTTAAGTTGACGTAAGTGTTGATGCTCAATTAATTTTTATATT
AACATAAAATTGTTATTGTGTTGCTTATTCGTTGCTATTGAGTTATTTCA
ATATTGTACTTTTTGAAGTTTTTGTCTTATAAATTTATTTTTATTTGCAT
AACATATGATAAACTTTAATGGATTTATATGTTACGGTTAAGTTGTTTTC
TAGTTGCTTTTGTGTTGTTGTTTAGTTGTTTTGCGCTGAGAAAAGAAGCC
CTATACGTTAAAATTCACTTTGTCCGAATTCGCCACGTGGCAACTACTGC
TACGTTCTGACGCTACAGATTGTGAGCGTCAGGCAAGTCTGAATGCTTTG
TCCCGCACTCACGTGTGAGTCTCACTCGCCCACATAGCCATATGTGTGCC
ACACATCTCCTTGCGCATCTAGGGGAGTTTTTCTACTGTATTTTTGTAAA
TAAGAAACTTTGAAAATAGTATATTTGAAATTAATAAATAATATAAATGT
TTATTTTTACTTATAGTAGTATAAATATTTATTTAAGATAATTTTAACTT
AAAATAGAAAAAATCTCAAAAAAAAAATAAAATAAAATAAACTATACCTTT
AATTTTTATATTTATGAAAATTTCCCGTATAAAAGGGAAGATTTTTTTTT
TTAAAATTAGTAGGTAAGTAATAGGATG

SEQ ID NO: 15 (2351 bp)

Hexanoyl-CoA synthetase (AAE1-1) ORF

```
   1 tcttggagtg tatatataga tatatatata catttctcca tttttggcca cccaaattac
  61 ctaaattgaa ttcgtcgtgt tcgtctttct ccgccataaa taacaattga tttccttcgt
 121 tctttcaagt ttttgtcaac tgagtactga gatccagtct gagtcgacac ttgtagaggt
 181 tcgaaaatgg cgcagaagaa gtttattttt gaggttgaaa aggctaagga agctaaggat
 241 ggaaggcctt ccatgggacc tgtttatcga agtatgtacg ccaaggatgg ctttcctcct
 301 ccgattgatg gattggattg ttgttgggat gtgtttaggc tttctgtgga gagaaatcct
 361 ggaaacaaaa tgctcggtcg ccgggaggtt gtggatggga aggctggaaa gtatgtttgg
 421 caaacttata aggaggttta tgattgtgtg ataaaacttg gaaatgccat gcgcagttgt
 481 ggtttcggag aaggagacaa atgtggtatc tatggtgcca attgctctga atggattatg
 541 agcatgcagg cttgcaatgc tcatggactc tattgtgttc cgttgtatga caccttaggt
 601 gctaatgcaa ttgagtttgt cataaaacat gccgaggtta caattgcttt tgcagaagag
 661 aaaaagcttc ctgagctgtt gaaaacactt cctaacacca cagactattt aaaaacactt
 721 gtgagctttg ggaaggtcac accggaacaa agggaatctg ttgaaaagtt tggattgaag
 781 atctattcct gggaagagtt cttgcaattg ggtgattgta aagatttcga tcttccagtg
 841 aaaaagaaaa ctgatatatg cactataatg tatactagtg gaactactgg tgacccaaag
 901 ggagtactga tatccaatga gagcatcatt tctcttttgg ctggggtgag gcgattgctt
 961 gagagtgtaa atgaaaagtt gaccgagaaa gatgtatttc tttcatatct tccccttgca
1021 catatcttcg atcgggtgat tgaggagtta tttattttgc atggtgcctc aatcgggttc
1081 tggcgtggcg atgtcaaatt attacttgaa gacattgggg agctaaaacc aactattttc
1141 tgtgccgttc cccgtgtgct ggatagagtc tatgctggtt tgacgcagaa gatttctgct
1201 gggggattct taaagaagac tctcttcaat ttcgcatact cacgcaagta taataacatg
1261 cgaaaggggc ttaaacacgg ggacgcagct ccattatgtg acaaaattgt atttagtaag
1321 gtgaagcaag gtttgggagg taatgtgcgg cttattctgt ctggagctgc gcctctagct
1381 cctcatgttg aagcttactt acgagttgtg acatgtgctc atgtttgcca aggatatggg
1441 ctaactgaaa cttgtgcggg gacatttgtc tcgataccaa atgaactacc aatgctcggt
1501 acagttggtc ctccagtacc caatgttgat gtatgtctgg aatcagttcc ggaaatggga
1561 tatgatgccc ttgcagaagt accgcgtgga gaagtatgtg tgaagggaag caccttattc
1621 tctgggtact acaaacgtga agacttaacc aaggaggtca tgattgatgg gtggttccat
1681 acaggtgatg ttggtgagtg gcaagcagat ggaagcttga aaatcattga ccgtaagaaa
1741 aatattttca aactttctca aggagaatat gtggcagttg agaacttaga gaacatttat
1801 ggtcttgttt ctgacattga ctcaatatgg gtttacggga acagcttcga gtccttcctc
1861 gttgctgtga ttaacccgaa ccagcaagca cttgaacgtt gggccggaga gaataatata
1921 tctggtgact tcaagtccat ttgtgaaaat ccaagagcaa aggaatacat tttgggagag
1981 ctcaataaga ttggcaaaga gaaaaagttg aaaggttttg agttcataag agctgttcac
2041 cttgacccag aggcatttga catagaacgt gaccttatta ctccaacata taagaaaaag
2101 agaccccagt tgctcaaata ctaccagagt gttattgata acatgtacaa ggaaggaaac
2161 aagcccagga cctgaaagat ataatagaca tttagagctt cactttctaa tatttcttca
2221 cacacttccc attttcatct ctctcttact tattacacta taaataatta acaaggttta
2281 ctgtgtttgta aatgacattt taatattttc atagtctcaa tgtttcaatt cttagaggac
2341 gttttgattg g
```

SEQ ID NO: 16 (420 bp)

CBDA Synthase (CBDAS) promoter

GTTAATAAACATACTTGTGAGTAGATCTAAGATCCTGGTAAAATAATTCCCAACATATT
TCAATCGTAAATTCATGAGTGATTCTTATCACTTCTTAAAAAAAAGATAT
AAAGATCCATAGATAAATAT TTGCTAGTCC TTTCTTTCCA TAGTCAACTT
CAATATGGCA TCTTAACTTC TTTCCATCAA CAAATGATGC TCAACTACTT
AATGTACAAT TTATATTTAT TTTTAGTAAG GGATACACCT AACAATGATG
CCTAATTTTG GTGAATTTTT TTTTACCACA TGTGACTTAA TGATATCAAA
TTTGGAAATA TTTAGTTAAT TTATTTGCCC CTGCTCCATT ATATAAAGCT
ATAAATAGGA TAGTTCTTAA TTCATAGTAA TTCAAAAATT ATTAGAACTA
AAGAACCATG G

SEQ ID NO: 17 (416 bp)

CBDA Synthase 1 (CBDAS1) promoter

GTTAATAAACATACTTGTGAGTAGATCTAAGATCCTGGTAAAATAATTCCCAACATATT
TCAATCGTAAATTCATGAGTGATTCTTATCACTTCTTAAAAAAAAGATAT
AAAGATCCATAGATAAATAT TTGCTAGTCC TTTCTTTCCA TAGTCAACTT
CAATATGGCA TCTTAACTTC TTTCCATCAA CAAATGATGC TCAACTACTT
AATGTACAAT TTATATTTAT TTTTAGTAAG GGATACACCT AACAATGATG
CCTAATTTTG GTGAATTTTT TTTTACCACA TGTGACTTAA TGATATCAAA
TTTGGAAATA TTTAGTTAAT TTATTTGCCC CTGCTCCATT ATATAAAGCT
ATAAATAGGA TAGTTCTTAA TTCATAGTAA TTCAAAAATT ATTAGAACTA
AAGAAAA

SEQ ID NO: 18 (535 bp)

CBDA Synthase (CBDAS) 20800 promoter

AGACGTT ATAGCTTTATTGTCTAAATT TCTTGGGGTA GTTTTGTCCC ATGTTGCTCG
CTTACTTTAGAAGTTCAAAG TTTGACAAAA CATGCTATTC GGTTATACAA
TGAGCTATCC
TAGTTCAAGG AGATTCCTGT GCTATTTGTG GATGTCTACA TTGTAAATTC
ATGAGTGATT CTTATAACTT TATTAAAAAA AAGACAAATA GATCCATAGA
TAAATATTTG CAAGTGCTTT CTTGTCATAG CAATTTTTTT TTACATAGGT
AATTTAAAAA ATTCATCTTA ACTTTTTTTC ATCAATAAAT GATGCTCAAT
TATTCAATGC ACCATGTACA TTTATTTTTA ATAAGGGCTG CACCTAACAA
AGGTGCCTAA TTTTGGTAAA AAAGAAATTA CGGCATGTGA ATATTTAATG
AGCATCAAAT TACAAAATAT TTAGCTAATT TCTTTACCCC CACTCCAATC
TATAAAGCTA TAAATAGGAT AATTTTCCAT TCATAGTAAT TTCCAACATT
ACGACTAAAG AACCATGG

SEQ ID NO: 19 (531 bp)

CBDA Synthase (CBDAS) 20800′ promoter

AGACGTT ATAGCTTTATTGTCTAAATT TCTTGGGGTA GTTTTGTCCC ATGTTGCTCG
CTTACTTTAGAAGTTCAAAG TTTGACAAAA CATGCTATTC GGTTATACAA
TGAGCTATCC
TAGTTCAAGG AGATTCCTGT GCTATTTGTG GATGTCTACA TTGTAAATTC

ATGAGTGATT CTTATAACTT TATTAAAAAA AAGACAAATA GATCCATAGA
TAAATATTTG CAAGTGCTTT CTTGTCATAG CAATTTTTTT TTACATAGGT
AATTTAAAAA ATTCATCTTA ACTTTTTTTC ATCAATAAAT GATGCTCAAT
TATTCAATGC ACCATGTACA TTTATTTTTA ATAAGGGCTG CACCTAACAA
AGGTGCCTAA TTTTGGTAAA AAAGAAATTA CGGCATGTGA ATATTTAATG
AGCATCAAAT TACAAAATAT TTAGCTAATT TCTTTACCCC CACTCCAATC
TATAAAGCTA TAAATAGGAT AATTTTCCAT TCATAGTAAT TTCCAACATT
ACGACTAAAG AAAA

SEQ ID NO: 20 (1635 bp)

CBDA Synthase (CBDAS) ORF

```
   1 atgaagtgct caacattctc cttttggttt gtttgcaaga taatattttt ctttttctca
  61 ttcaatatcc aaacttccat tgctaatcct cgagaaaact tccttaaatg cttctcgcaa
 121 tatattccca ataatgcaac aaatctaaaa ctcgtataca ctcaaaacaa cccattgtat
 181 atgtctgtcc taaattcgac aatacacaat cttagattca cctctgacac aaccccaaaa
 241 ccacttgtta tcgtcactcc ttcacatgtc tctcatatcc aaggcactat tctatgctcc
 301 aagaaagttg gcttgcagat tcgaactcga agtggtggtc atgattctga gggcatgtcc
 361 tacatatctc aagtcccatt tgttatagta gacttgagaa acatgcgttc aatcaaaata
 421 gatgttcata gccaaactgc atgggttgaa gccggagcta cccttggaga agtttattat
 481 tgggttaatg agaaaaatga gaatcttagt ttggcggctg ggtattgccc tactgtttgc
 541 gcaggtggac actttggtgg aggaggctat ggaccattga tgagaaacta tggcctcgcg
 601 gctgataata tcattgatgc acacttagtc aacgttcatg gaaaagtgct agatcgaaaa
 661 tctatggggg aagatctctt ttgggcttta cgtggtggtg gagcagaaag cttcggaatc
 721 attgtagcat ggaaaattag actggttgct gtcccaaagt ctactatgtt tagtgttaaa
 781 aagatcatgg agatacatga gcttgtcaag ttagttaaca aatggcaaaa tattgcttac
 841 aagtatgaca aagatttatt actcatgact cacttcataa ctaggaacat tacagataat
 901 caagggaaga ataagacagc aatacacact tacttctctt cagttttcct tggtggagtg
 961 gatagtctag tcgacttgat gaacaagagt tttcctgagt tgggtattaa aaaaacggat
1021 tgcagacaat tgagctggat tgatactatc atcttctata gtggtgttgt aaattacgac
1081 actgataatt ttaacaagga aattttgctt gatagatccg ctgggcagaa cggtgctttc
1141 aagattaagt tagactacgt taagaaacca attccagaat ctgtatttgt ccaaatttttg
1201 gaaaaattat atgaagaaga tataggagct gggatgtatg cgttgtaccc ttacggtggt
1261 ataatggatg agatttcaga atcagcaatt ccattccctc atcgagctgg aatcttgtat
1321 gagttatggt acatatgtag ttgggagaag caagaagata acgaaaagca tctaaactgg
1381 attagaaata tttataactt catgactcct tatgtgtcca aaaatccaag attggcatat
1441 ctcaattata gagaccttga tataggaata aatgatccca agaatccaaa taattacaca
1501 caagcacgta tttggggtga gaagtatttt ggtaaaaatt ttgacaggct agtaaaagtg
1561 aaaaccctgg ttgatcccaa taacttttttt agaaacgaac aaagcatccc acctcttcca
1621 cggcatcgtc attaa
```

SEQ ID NO: 21 (800 bp)

THCA Synthase (THCAS) 19603 promoter

CTTATCACTC ATTCGGGAGC AACTTAAAAA AATTACAATT ATTATGAAAG
ACGGCACAAA AAGCGAATTA ATTACCTAAT CGTAAATAAA AGAAATTTTA
CATAAATATA TTATATAATT TTATATTATT CATAAAAAAT ATAAAATAGT
ATTAGACTGA ATTGAGATTT GATTCGTAAC TAAAAATTAA TTTCTTTAAT
GTAAAATTTT AAAAACAAAT TATTAAAACC GCGTGAAGCG CGGATCTATT
CCCTAGTTGA ATAATAAAGT AGATAGTAGA GGAGGTTAGA ATTTATAATT
TTTCATTTAA ATATTTGAAT TTACTTTACT GTTCTTTGTT ATTCTTCATT
TAATTTTGCT ATTTGTTATT TTACTTTTCA AAAATTATAA TTTTAATTAC
CAAATAAAAA GTAAAATAGA TATTGGTACT TGATATTCAC TCTTTATGGG
AACCATAATA AACTATAAAA GTCATTATGT GTACTTGCTA CCATAGGCAC
CTATATCCCA CAAACTAGCT ACCATAGCCA ATTTCTTGTT TTTTGTTTCC
AATATCCAAT TTTTATTGAT GCCAAACTAT TCAATGTACA ATGTACATTT
ATTTTCAATA AGGGCTTCAC CTAACAAAGG TGCCTAATTT TAGTTGATTT
ATTTTTTATC ACATGTGACT ATTTAATGAC TATCAAATTA TAAAATATTT
AAGTCAATTT ATTTGCCCCA ACTCCAATAT ATAATATTAT AAATAGGATA
GTTCTCAATT CCTAATAATT CAAAAAATCA TTAGGACTGA AGAACCATGG

SEQ ID NO: 22 (796 bp)

THCA Synthase (THCAS) 19603′ promoter

CTTATCACTC ATTCGGGAGC AACTTAAAAA AATTACAATT ATTATGAAAG
ACGGCACAAA AAGCGAATTA ATTACCTAAT CGTAAATAAA AGAAATTTTA
CATAAATATA TTATATAATT TTATATTATT CATAAAAAAT ATAAAATAGT
ATTAGACTGA ATTGAGATTT GATTCGTAAC TAAAAATTAA TTTCTTTAAT
GTAAAATTTT AAAAACAAAT TATTAAAACC GCGTGAAGCG CGGATCTATT
CCCTAGTTGA ATAATAAAGT AGATAGTAGA GGAGGTTAGA ATTTATAATT
TTTCATTTAA ATATTTGAAT TTACTTTACT GTTCTTTGTT ATTCTTCATT
TAATTTTGCT ATTTGTTATT TTACTTTTCA AAAATTATAA TTTTAATTAC
CAAATAAAAA GTAAAATAGA TATTGGTACT TGATATTCAC TCTTTATGGG
AACCATAATA AACTATAAAA GTCATTATGT GTACTTGCTA CCATAGGCAC
CTATATCCCA CAAACTAGCT ACCATAGCCA ATTTCTTGTT TTTTGTTTCC
AATATCCAAT TTTTATTGAT GCCAAACTAT TCAATGTACA ATGTACATTT
ATTTTCAATA AGGGCTTCAC CTAACAAAGG TGCCTAATTT TAGTTGATTT
ATTTTTTATC ACATGTGACT ATTTAATGAC TATCAAATTA TAAAATATTT
AAGTCAATTT ATTTGCCCCA ACTCCAATAT ATAATATTAT AAATAGGATA
GTTCTCAATT CCTAATAATT CAAAAAATCA TTAGGACTGA AGAAAA

SEQ ID NO: 23 (796 bp)

THCA Synthase (THCAS) 50320 promoter

TTATTTAAAT TTCTCAGAGA GATATAGAGA ATTTACAATA GGTCAAGTTG
TATATTTTTT AATCCGAATA ATTAATGCTT TTGAATATCC ATCATCAATA
TTGAATAATA AAGTAGATAG CAGAGGAGAT TAGAACTCAT AATTTTTCAT
TTATATTTTT GAATTTAGTT TACTGTTCTT TATTATTCTT CATTTAATTT
TGCTATATGT TATTTTACTT TTCAAAAATT CTAATTTTGA TTACCAAATA
AAAAATAAAA TATATATTGG TACTTCATAT TTAGTTTTTA TGGGAATGAT
AATGAATGAT ATTGTAATAA ACTATAAAAG TAATTTTGTA TACTTGCATA
GTTCTTAATT TTTACCCAAA AAAATATATA TTAATGAAAA AAAAAGGTGG
AAAGTGCCAT AGGCACCTAT ATCCCACAAA CTAGCTATAA GATATTGAAA
GACAAATAGA TCCATAGATA AATATTTGCG TGTCCTTTCT TTTCATATCT
AATTTTTTGT TTTTTTATTT TTTTCCAATA GTCAATTTTG GCAGCTTAAC
TTCTTTCCAT AAATAATGAT GCCAAACTAT TCAATGTAAA ATTTAGATTT
ATTTTCATTA AGGGCTTCAC CTAACAAATG TGCCTAATTT TTGTGGATTT
TTTTACCATA TGTCGCTATT TAATGACTAC CAAATTATAG AATTATTTAA
GTCAATTTCT CAGTCCCCGC TCCAATATAT AAAGTTATAG AAAGGACAAT
TCTTAATTCA TAGTAATTCA AAAATCATTA AGACTAAAAA CCATGG

SEQ ID NO: 24 (792 bp)

THCA Synthase (THCAS) 50320′ promoter

TTATTTAAAT TTCTCAGAGA GATATAGAGA ATTTACAATA GGTCAAGTTG
TATATTTTTT AATCCGAATA ATTAATGCTT TTGAATATCC ATCATCAATA
TTGAATAATA AAGTAGATAG CAGAGGAGAT TAGAACTCAT AATTTTTCAT
TTATATTTTT GAATTTAGTT TACTGTTCTT TATTATTCTT CATTTAATTT
TGCTATATGT TATTTTACTT TTCAAAAATT CTAATTTTGA TTACCAAATA
AAAAATAAAA TATATATTGG TACTTCATAT TTAGTTTTTA TGGGAATGAT
AATGAATGAT ATTGTAATAA ACTATAAAAG TAATTTTGTA TACTTGCATA
GTTCTTAATT TTTACCCAAA AAAATATATA TTAATGAAAA AAAAAGGTGG
AAAGTGCCAT AGGCACCTAT ATCCCACAAA CTAGCTATAA GATATTGAAA
GACAAATAGA TCCATAGATA AATATTTGCG TGTCCTTTCT TTTCATATCT
AATTTTTTGT TTTTTTATTT TTTTCCAATA GTCAATTTTG GCAGCTTAAC
TTCTTTCCAT AAATAATGAT GCCAAACTAT TCAATGTAAA ATTTAGATTT
ATTTTCATTA AGGGCTTCAC CTAACAAATG TGCCTAATTT TTGTGGATTT
TTTTACCATA TGTCGCTATT TAATGACTAC CAAATTATAG AATTATTTAA
GTCAATTTCT CAGTCCCCGC TCCAATATAT AAAGTTATAG AAAGGACAAT
TCTTAATTCA TAGTAATTCA AAAATCATTA AGACTAAAAA AA

SEQ ID NO: 25 (720 bp)

THCA Synthase (THCAS) 1330 promoter

TTTTCAATTG ATTTAATTTC TTATATTGAT ATAAAGAATT TGCAATACAT
TGAGTTTCTT AACCCGAATA TTAAACGGTT TTGAATATCT TCATCATTGA
TTGAATAATA AAGTGGATAG TAGAGGGAAT TAGAATCCAT AATTTTTTAT
TTATATATTT GAAGTTAGTT TATTATTCTT TGCTATCCTT CATTTAATTT
TGCTATTTGT TATTTTAGTT TCAAAAATTT ATTTTTCATT ACAAAATAAA
AAATAAGATA GGATATTGGT ACTTGATAAG TCTTCTTTGT GGAAACGATA
ATCGGTATTA TTAGGTATAT TTGCATATTT CAATATTATT GCAAAAAATA
AATATATTAA TAAATAAAGT GGAAGGTGCC ATAGGTACCT ATATCCACAA
ACTAGCATAT TGAAAGAAAA TGGATCCAAG GATAAATATT TGCAAGTCTT
TTTTTTTTTG CATAGCCAAT TTCTTTTTTT ATTATTATTT TAATAGAATA
TTTCAAAAGG GCATCTAACA TTTATTTTTA ATAAGGACTG CACCTAACAA
AGGTGCCTAA TTTTTGTGAA CTTTTTTTTA CCACATGTGA CTATTTAATG
AGTACTAAAT TATGAAATAT TTAGTTAATT TCTTTGCCCC CGCTCCAATA
TATAATGCTA TAAATAGCAT AATTTTCTAT TCATAGTAAT TCAAAAATCA
TTAGGACTAA AGAACCATGG

SEQ ID NO: 26 (716 bp)

THCA Synthase (THCAS) 1330′ promoter

TTTTCAATTG ATTTAATTTC TTATATTGAT ATAAAGAATT TGCAATACAT
TGAGTTTCTT AACCCGAATA TTAAACGGTT TTGAATATCT TCATCATTGA
TTGAATAATA AAGTGGATAG TAGAGGGAAT TAGAATCCAT AATTTTTTAT
TTATATATTT GAAGTTAGTT TATTATTCTT TGCTATCCTT CATTTAATTT
TGCTATTTGT TATTTTAGTT TCAAAAATTT ATTTTTCATT ACAAAATAAA
AAATAAGATA GGATATTGGT ACTTGATAAG TCTTCTTTGT GGAAACGATA
ATCGGTATTA TTAGGTATAT TTGCATATTT CAATATTATT GCAAAAAATA
AATATATTAA TAAATAAAGT GGAAGGTGCC ATAGGTACCT ATATCCACAA
ACTAGCATAT TGAAAGAAAA TGGATCCAAG GATAAATATT TGCAAGTCTT
TTTTTTTTTG CATAGCCAAT TTCTTTTTTT ATTATTATTT TAATAGAATA
TTTCAAAAGG GCATCTAACA TTTATTTTTA ATAAGGACTG CACCTAACAA
AGGTGCCTAA TTTTTGTGAA CTTTTTTTTA CCACATGTGA CTATTTAATG
AGTACTAAAT TATGAAATAT TTAGTTAATT TCTTTGCCCC CGCTCCAATA
TATAATGCTA TAAATAGCAT AATTTTCTAT TCATAGTAAT TCAAAAATCA
TTAGGACTAA AGAAAG

SEQ ID NO: 27 (1635 bp)

THCA Synthase (THCAS) ORF

```
   1 atgaattgct cagcattttc cttttggttt gtttgcaaaa taatattttt ctttctctca
  61 ttcaatatcc aaatttcaat agctaatcct caagaaaact tccttaaatg cttctcggaa
 121 tatattccta acaatccagc aaatccaaaa ttcatataca ctcaacacga ccaattgtat
 181 atgtctgtcc tgaattcgac aatacaaaat cttagattca cctctgatac aaccccaaaa
 241 ccactcgtta ttgtcactcc ttcaaatgtc tcccatatcc aggccagtat tctctgctcc
 301 aagaaagttg gtttgcagat tcgaactcga agcggtggcc atgatgctga gggtatgtcc
 361 tacatatctc aagtcccatt tgttgtagta gacttgagaa acatgcattc gatcaaaata
 421 gatgttcata gccaaactgc gtgggttgaa gccggagcta cccttggaga agtttattat
```

```
 481 tggatcaatg agaagaatga gaattttagt tttcctggtg ggtattgccc tactgttggc
 541 gtaggtggac actttagtgg aggaggctat ggagcattga tgcgaaatta tggccttgcg
 601 gctgataata tcattgatgc acacttagtc aatgttgatg gaaaagttct agatcgaaaa
 661 tccatgggag aagatctatt ttgggctata cgtggtggag gaggagaaaa ctttggaatc
 721 attgcagcat ggaaaatcaa actggttgtt gtcccatcaa agtctactat attcagtgtt
 781 aaaaagaaca tggagataca tgggcttgtc aagttattta acaaatggca aaatattgct
 841 tacaagtatg acaaagattt agtactcatg actcacttca taacaaagaa tattacagat
 901 aatcatggga agaataagac tacagtacat ggttacttct cttcaatttt tcatggtgga
 961 gtggatagtc tagtcgactt gatgaacaag agctttcctg agttgggtat taaaaaaact
1021 gattgcaaag aatttagctg gattgataca accatcttct acagtggtgt tgtaaatttt
1081 aacactgcta attttaaaaa ggaaattttg cttgatagat cagctgggaa gaagacggct
1141 ttctcaatta agttagacta tgttaagaaa ccaattccag aaactgcaat ggtcaaaatt
1201 ttggaaaaat tatatgaaga agatgtagga gctgggatgt atgtgttgta cccttacggt
1261 ggtataatgg aggagatttc agaatcagca attccattcc ctcatcgagc tggaataatg
1321 tatgaacttt ggtacactgc ttcctgggag aagcaagaag ataatgaaaa gcatataaac
1381 tgggttcgaa gtgtttataa ttttacgact ccttatgtgt cccaaaatcc aagattggcg
1441 tatctcaatt atagggacct tgatttagga aaaactaatc ctgagagtcc taataattac
1501 acacaagcac gtatttgggg tgaaaagtat tttggtaaaa attttaacag gttagttaag
1561 gtgaaaacca aagctgatcc caataatttt tttagaaacg aacaaagtat cccacctctt
1621 ccaccgcatc atcat
```

SEQ ID NO: 28 (804 bp)

CBCA Synthase (CBCAS) 3498 promoter

GATAATTTTT CTTCCAAATT TAATAGTATC TTTTATATAA GTTCAGTCAT
GTTTTTCTTA TGTAAGTTAG GCTATTAATT ATTTTGATGT ATTTATTATA
TTGTATGTCG TTTACTGCAT TCTGCCAACA TTGGTTTTTG TCGATTTATG
ATATATAATA TGATATGCTG TTAAGTTAGA GGTAAGATTC AATTTATTTA
AGATTATTTA ATTTGCGCTT TTATTATATA CATCTTCTAG TTGTAATTAA
TGGGGTAGAA TTAAGACTGT GTTTTGAATT AATATCGGTA TTAGAATAGG
ATTTGCATGA CTATATTTCT ACCTTACCAA GTTTTTTATC TGATAACACC
CTTTCACTTA TTATTTTCTT ATTTTTAATT TATTTAT
ATTTTTACCATATTCTATTTATACTAATCTTAAGTGGTACATTACCTCCCTGTG
GATACGACATATA
ATCTGTTTACTATCGTGACCGAAGTATAACTAATTGGGCGACATCACACCTATA
TCCCACAAACTAGCTACTATAGTCAATTTCTTGTTTTTTTTCCAATAGCCAATT
TTAAATGATGCCAAACTATTCAATGTATAATGTACATTTATTTTAAATAAGGGC
TTCACCTAACAAATGTGCCTAATTTTAGTTAATTTATTTTTTTATCGCATCTTA
CTATTTAAAGGAACTATCAAATTATAAAATATTTATGTCAATTCATTTGCCCCA
ACTCCAATATATAATATTATAAATAGGATAGTTCTCTATTCATAATAATTCAAA
ATATCATTAGGACTGAAGAACCATGG

SEQ ID NO: 29 (800 bp)

CBCA Synthase (CBCAS) 3498′ promoter

GATAATTTTT CTTCCAAATT TAATAGTATC TTTTATATAA GTTCAGTCAT
GTTTTTCTTA TGTAAGTTAG GCTATTAATT ATTTTGATGT ATTTATTATA
TTGTATGTCG TTTACTGCAT TCTGCCAACA TTGGTTTTTG TCGATTTATG
ATATATAATA TGATATGCTG TTAAGTTAGA GGTAAGATTC AATTTATTTA
AGATTATTTA ATTTGCGCTT TTATTATATA CATCTTCTAG TTGTAATTAA
TGGGGTAGAA TTAAGACTGT GTTTTGAATT AATATCGGTA TTAGAATAGG
ATTTGCATGA CTATATTTCT ACCTTACCAA GTTTTTTATC TGATAACACC
CTTTCACTTA TTATTTTCTT ATTTTTAATT TATTTAT
ATTTTTACCATATTCTATTTATACTAATCTTAAGTGGTACATTACCTCCCTGTG
GATACGACATATA
ATCTGTTTACTATCGTGACCGAAGTATAACTAATTGGGCGACATCACACCTATA
TCCCACAAACTAGCTACTATAGTCAATTTCTTGTTTTTTTTCCAATAGCCAATT
TTAAATGATGCCAAACTATTCAATGTATAATGTACATTTATTTTAAATAAGGGC
TTCACCTAACAAATGTGCCTAATTTTAGTTAATTTATTTTTTTATCGCATCTTA
CTATTTAAAGGAACTATCAAATTATAAAATATTTATGTCAATTCATTTGCCCCA
ACTCCAATATATAATATTATAAATAGGATAGTTCTCTATTCATAATAATTCAAA
ATATCATTAGGACTGAAGAAAA

SEQ ID NO: 30 (1635 bp)

CBCA Synthase (CBCAS) ORF

```
   1 atgaattgct caacattctc cttttggttt gtttgcaaaa taatattttt ctttctctca
  61 ttcaatatcc aaatttcaat agctaatcct caagaaaact tccttaaatg cttctcggaa
 121 tatattccta acaatccagc aaatccaaaa ttcatataca ctcaacacga ccaattgtat
 181 atgtctgtcc tgaattcgac aatacaaaat cttagattca cctctgatac aaccccaaaa
```

241 ccactcgtta ttgtcactcc ttcaaatgtc tcccatatcc aggccagtat tctctgctcc
301 aagaaagttg gtttgcagat tcgaactcga agcggtggcc atgatgctga gggtttgtcc
361 tacatatctc aagtcccatt tgctatagta gacttgagaa acatgcatac ggtcaaagta
421 gatattcata gccaaactgc gtgggttgaa gccggagcta cccttggaga agtttattat
481 tggatcaatg agatgaatga gaattttagt tttcctggtg ggtattgccc tactgttggc
541 gtaggtggac actttagtgg aggaggctat ggagcattga tgcgaaatta tggccttgcg
601 gctgataata tcattgatgc acacttagtc aatgttgatg gaaaagttct agatcgaaaa
661 tccatgggag aagatctatt ttgggctata cgtggtggag gaggagaaaa ctttggaatc
721 attgcagcat gtaaaatcaa acttgttgtt gtcccatcaa aggctactat attcagtgtt
781 aaaaagaaca tggagataca tgggcttgtc aagttattta acaaatggca aaatattgct
841 tacaagtatg acaaagattt aatgctcacg actcacttca gaactaggaa tattacagat
901 aatcatggga agaataagac tacagtacat ggttacttct cttccatttt tcttggtgga
961 gtggatagtc tagttgactt gatgaacaag agctttcctg agttgggtat taaaaaaact
1021 gattgcaaag aattgagctg gattgataca accatcttct acagtggtgt tgtaaattac
1081 aacactgcta attttaaaaa ggaaattttg cttgatagat cagctgggaa gaagacggct
1141 ttctcaatta agttagacta tgttaagaaa ctaatacctg aaactgcaat ggtcaaaatt
1201 ttggaaaaat tatatgaaga agaggtagga gttgggatgt atgtgttgta cccttacggt
1261 ggtataatgg atgagatttc agaatcagca attccattcc ctcatcgagc tggaataatg
1321 tatgaacttt ggtacactgc tacctgggag aagcaagaag ataacgaaaa gcatataaac
1381 tgggttcgaa gtgtttataa tttcacaact ccttatgtgt cccaaaatcc aagattggcg
1441 tatctcaatt atagggacct tgatttagga aaaactaatc ctgagagtcc taataattac
1501 acacaagcac gtatttgggg tgaaaagtat tttggtaaaa attttaacag gttagttaag
1561 gtgaaaacca aagctgatcc caataatttt tttagaaacg aacaaagtat cccacctctt
1621 ccaccgcgtc atcat

SEQ ID NO: 31 (171 bp)
CANON fragment

atgatgccaaactattcaatgtacaatgtacatttatttttaataagggcttcacctaacaa
aggtgcctaatttttgtgaacttttttttaccacatgtgactatttaatgactatcaaatta
taaaatatttaagtcaatttctttgcccccactccaatatataatgt

SEQ ID NO: 32 (232 bp)
CANON fragment with putative TATA Box, 5'UTR, and start codon

atgatgccaaactattcaatgtacaatgtacatttatttttaataagggcttcacctaacaa
aggtgcctaatttttgtgaacttttttttaccacatgtgactatttaatgactatcaaatta
taaaatatttaagtcaatttctttgcccccactccaatatataatgt**tataaata**ggataat
tctcaattcatagtaattcaaaaatcattaggactaaagaaaa**atg**

SEQ ID NO: 33 (709 bp)
4 x CANON fragment synthetic promoter

atgatgccaaactattcaatgtacaatgtacatttatttttaataagggcttcacctaacaa
aggtgcctaatttttgtgaacttttttttaccacatgtgactatttaatgactatcaaatta
taaaatatttaagtcaatttctttgcccccactccatgatgccaaactattcaatgtacaat
gtacatttatttttaataagggcttcacctaacaaaggtgcctaatttttgtgaacttttttt
ttaccacatgtgactatttaatgactatcaaattataaaatatttaagtcaatttctttgcc
cccactcc**atgatgccaaactattcaatgtacaatgtacatttatttttaataagggcttca**
cctaacaaaggtgcctaatttttgtgaacttttttttaccacatgtgactatttaatgacta
tcaaattataaaatatttaagtcaatttctttgcccccactccatgatgccaaactattcaa
tgtacaatgtacatttatttttaataagggcttcacctaacaaaggtgcctaatttttgtga
acttttttttaccacatgtgactatttaatgactatcaaattataaaatatttaagtcaatt
tctttgcccccactcc**aatatataatgttataaataggataattctcaattcatagtaattc**
aaaaatcattaggactaaagaaaaatg

CLAIMS

1. A synthetic DNA molecule comprising a nucleotide sequence set forth in SEQ ID NO: 33.

2. An expression vector comprising the DNA molecule of claim 1, operably linked to one or more nucleic acid sequences encoding a polypeptide.

3. A genetically engineered host cell comprising the expression vector of claim 2.

4. The genetically engineered host cell of claim 3, wherein the cell is a *Cannabis sativa* cell.

5. The genetically engineered host cell of claim 3, wherein the cell is a *Nicotiana tabacum* cell.

6. A genetically engineered plant comprising a cell comprising a chimeric nucleic acid construct comprising the synthetic DNA molecule of claim 1.

7. The engineered plant of claim 6, wherein the plant is an *N. tabacum* plant.

8. The engineered plant of claim 6, wherein the plant is a *C. sativa* plant.

9. Seeds from the engineered plant of any one of claims 6-8, wherein the seeds comprise the chimeric nucleic acid construct.