Journal of Experimental Botany, Vol. 60, No. 13, pp. 3715–3726, 2009
doi:10.1093/jxb/erp210 Advance Access publication 6 July, 2009
This paper is available online free of all access charges (see http://jxb.oxfordjournals.org/open_access.html for further details)

Journal of Experimental Botany
www.jxb.oxfordjournals.org

RESEARCH PAPER

Identification of candidate genes affecting Δ^9-tetrahydrocannabinol biosynthesis in *Cannabis sativa*

M. David Marks[1,*], Li Tian[2,†], Jonathan P. Wenger[1], Stephanie N. Omburo[1], Wilfredo Soto-Fuentes[1], Ji He[2], David R. Gang[3], George D. Weiblen[1] and Richard A. Dixon[2]

[1] Department of Plant Biology, University of Minnesota, 1445 Gortner Ave, St Paul, MN 55108, USA
[2] Plant Biology Division, Samuel Roberts Noble Foundation, Ardmore, OK 73401, USA
[3] Department of Plant Sciences, University of Arizona, Tucson, AZ 85721-0036, USA

Received 14 May 2009; Revised 3 June 2009; Accepted 5 June 2009

Abstract

RNA isolated from the glands of a Δ^9-tetrahydrocannabinolic acid (THCA)-producing strain of *Cannabis sativa* was used to generate a cDNA library containing over 100 000 expressed sequence tags (ESTs). Sequencing of over 2000 clones from the library resulted in the identification of over 1000 unigenes. Candidate genes for almost every step in the biochemical pathways leading from primary metabolites to THCA were identified. Quantitative PCR analysis suggested that many of the pathway genes are preferentially expressed in the glands. Hexanoyl-CoA, one of the metabolites required for THCA synthesis, could be made via either *de novo* fatty acids synthesis or via the breakdown of existing lipids. qPCR analysis supported the *de novo* pathway. Many of the ESTs encode transcription factors and two putative MYB genes were identified that were preferentially expressed in glands. Given the similarity of the *Cannabis* MYB genes to those in other species with known functions, these *Cannabis* MYBs may play roles in regulating gland development and THCA synthesis. Three candidates for the polyketide synthase (PKS) gene responsible for the first committed step in the pathway to THCA were characterized in more detail. One of these was identical to a previously reported chalcone synthase (CHS) and was found to have CHS activity. All three could use malonyl-CoA and hexanoyl-CoA as substrates, including the CHS, but reaction conditions were not identified that allowed for the production of olivetolic acid (the proposed product of the PKS activity needed for THCA synthesis). One of the PKS candidates was highly and specifically expressed in glands (relative to whole leaves) and, on the basis of these expression data, it is proposed to be the most likely PKS responsible for olivetolic acid synthesis in *Cannabis* glands.

Key words: Chalcone synthase, glandular trichome, hemp, hop, *Humulus lupulus*, marijuana, polyketide synthase, trichomes.

Introduction

Cannabis sativa has a long history of cultivation for a variety of uses including food, fibre, medicine, and recreational drugs (Measham *et al.*, 1994; Ware and Tawfik, 2005; Kostic *et al.*, 2008). *Cannabis* produces many different secondary compounds such as cannabinoids, flavonoids, stilbenoids, alkaloids, lignanamides, and phenolic amides (Flores-Sanchez and Verpoorte, 2008*b*). Δ^9-Tetrahydrocannabinolic acid (THCA), a product of the cannabinoid class, is the primary psychoactive agent. This compound is produced as an acid in the glandular trichomes of inflorescence bracts and undergoes decarboxylation with age or heating to Δ^9-tetrahydrocannabinol (THC) (Mechoulam, 1970; Turner *et al.*, 1980; Pertwee, 2006). *Cannabis* cultivars differ substantially in economic traits that range from marijuana, arguably the most widespread illicit drug, to hemp fibre derived from the stems of the plant. Marijuana consists of the dried female inflorescences in which the quantity of THC exceeds that of cannabidiol (CBD,

* To whom correspondence should be addressed: E-mail: marks004@umn.edu
† Present address: Department of Plant Sciences, University of California Davis, One Shield Avenue, Davis, CA 95616, USA.

produced initially as cannabidiolic acid (CBDA)), and potency varies among cultivars by several orders of magnitude (ElSohly *et al.*, 2000). Marijuana cultivars are known to have THC levels exceeding 2–24% of inflorescence dry weight whereas hemp cultivars produce substantially less THC but rather high levels of CBD (Hillig and Mahlberg, 2004). THCA and CBDA share the same biosynthetic pathway except for the last step in which THCA synthase and CBDA synthase produce THCA or CBDA, respectively (Taura *et al.*, 2007). Recent evidence suggests that the genes encoding the two synthases are allelic (de Meijer *et al.*, 2003; Pacifico *et al.*, 2006). CBD and THC are enatiomers, but only THC elicits psychotropic effects, whereas CBD may mediate anti-psychotropic effects (Long *et al.*, 2006; Zuardi *et al.*, 2006), a difference highlighting the stereo-selectivity of receptors in the human body that bind these compounds.

Although classified as a drug without therapeutic value in the United States, ingestion of THC is widely regarded as having effects including pain relief and appetite stimulation, that may, among other things, increase the tolerance of cancer patients to chemotherapy (Baker *et al.*, 2003). Dronabinol, a synthetic analogue of THC, is approved for use as an appetite stimulant in the United States as a Schedule III drug (marketed as Marinol by Unimed Pharmaceutical, a subsidiary of Solvay Pharmaceuticals, Marietta, GA). Cesamet (Nobiline; Valeant Pharmaceuticals, Costa Mesa, CA), another synthetic analogue, is used as an anti-emetic for patients undergoing cancer therapy. The natural product Sativex (GW Pharmaceuticals, Salisbury, UK) is approved for use in the UK and is derived from *Cannabis* cultivars containing both THC and CBD, and is used to treat pain symptoms associated with multiple sclerosis.

Compounds from *Cannabis sativa* are of undeniable medical interest, and subtle differences in the chemical nature of these compounds can greatly influence their pharmacological properties. For these reasons, a better understanding of the secondary metabolic pathways that lead to the synthesis of bioactive natural products in *Cannabis* is needed (Mechoulam, 2005). Knowledge of genetics underlying cannabinoid biosynthesis is also needed to engineer drug-free and distinctive *Cannabis* varieties capable of supplying hemp fibre and oilseed. In this report, RNA from mature glands isolated from the bracts of female inflorescences was converted into cDNA and cloned to produce a cDNA library. DNA from over 2000 clones has been sequenced and characterized. Candidate genes for almost all of the enzymes required to convert primary metabolites into THCA have been identified. Expression levels of many of the candidate genes for the pathways were compared between isolated glands and intact inflorescence leaves.

Materials and methods

Plant growth and gland isolation

Seeds from the marijuana cultivar Skunk no. 1 were provided by HortaPharm BV (Amsterdam, The Netherlands) and imported under a US Drug Enforcement Administration (DEA) permit to a registered controlled substance research facility. Plants were grown under hydroponic conditions in a secure growth chamber yielding cannabinoid levels in mature plants as reported in Datwyler and Weiblen (2006). Approximately 5 g of tissue was harvested from mature female inflorescences 8 weeks after the onset of flowering. Tissue was equally distributed into four 50 ml tubes (50 ml PP-Tube sterile, Greiner Bio-one, Kremsmünster, Austria) containing 20 ml phosphate buffered saline (PBS) as described by Sambrook *et al.* (1989), but made with all potassium salts and mixed at maximum speed with a Vortex 2 Genie (Scientific Industries, Bohemia, NY) for four repetitions of 30 s mixing followed by 30 s rest on ice, for a total of 2 min of mixing. Material was sieved through four layers of 1×1 mm plastic mesh and the flow-through was split into two 50 ml tubes and spun in a centrifuge for 30 s at 500 rpm. Supernatants were decanted and pellets were resuspended in PBS. The suspensions were combined into one tube and pelleted as before. The resulting pellet was diluted into 100 µl of PBS. Five µl were used for cell counting with a haemocytometer, and the total suspension was estimated to contain 70 000 intact glands. Plant residue was incinerated by a DEA-registered reverse distributor (Return Logistics, Savannah, GA).

RNA isolation, cDNA library construction, EST sequencing and analysis

Total RNA was isolated from the glands using the RNAeasy Plant Mini Kit (Qiagen Inc.-USA, Valencia, CA), and 120 ng of total RNA was used to make a cDNA library with the Creator SMART cDNA Library Construction Kit (Clontech, Mountain View, CA). This kit allows directional cloning of cDNA inserts into the pDNR_Lib plasmid vector. All procedures were followed as described in the kit except that the final cDNA product was size-selected on a 0.8% agarose gel and gel purified with the Qiaex II kit from Qiagen. Over 100 000 clones were generated. Plasmid DNA was isolated from 2112 bacterial clones, and sequenced using an M13 forward primer that reads into the 5′ end of the oriented cDNA inserts. Sequences can be found in GenBank with accession numbers GR220588-222152.

Cannabis *polyketide synthase cloning and analysis*

Full-length cDNAs from CAN24, 383, and 1069 contigs encoding putative polyketide synthases were used as templates for PCR reactions. Sequences identical to CAN24 and 1069 are available in GenBank with accession numbers AB164375 and AAL92879, respectively. The GenBank accession for CAN383 is GQ222379. The PCR reactions were designed to add 5′ *Nco*I and 3′ *Bam*HI restriction enzyme sites to the ends of each sequence. After digestion with *Nco*I and *Bam*HI, the inserts were cloned into the corresponding sites of pHIS8 (Jez *et al.*, 2000), which adds eight histidine residues to the N-terminus of the encoded

protein. Clones corresponding to each PKS-related gene were sequenced. The Lasergene (Madison, WI) MegAlign program using the Clustal W algorithm was used to generate the alignment of PKS genes. PKS gene constructs were transformed into *E. coli* BL21-(DE). For protein isolation, cultures were grown to an OD of approximately 0.6 at 37 °C when transcription of the cDNA inserts was induced with IPTG and cultures were grown for an additional 10–12 h at 28 °C. His-tagged protein was isolated from the bacteria using the MagneHis Protein Purification System (Promega, Madison, WI).

PKS enzyme assays (40 µl) contained 4 µg protein, 100 mM KPO_4 (pH 7.0), 5 mM malonyl-CoA, and either 250 µM hexanoyl-CoA or 67.5 µM 4-coumaroyl-CoA as substrates. Boiled protein was assayed in parallel with all reactions. All negative controls showed a lack of product formation. Reactions were incubated for 1 h at 30 °C, dried in a Speed Vac, resuspended in 40 µl methanol, and applied to Agilent Tech 1200 HPLC (Agilent Technologies, Santa Clara, CA) with a Spherisorb 6 µ ODS2 (Waters, Milano, Italy) separation column. Products and reactants were resolved across a gradient of 1% H_3PO_4 to 100% acetonitrile with molecular weights determined by LC-MS.

Quantitative PCR

Quantitative (qPCR) reactions were performed as described previously (Marks *et al.*, 2007) using primers listed in Supplementary Table 4B at *JXB* online. Equivalent quantities of RNA isolated from glands and inflorescence-associated leaves were used to generate the respective single stranded cDNAs. qPCR reactions containing equal quantities of gland or leaf cDNA were run in duplicate along with reactions containing standards consisting of 100-fold sequential dilutions of isolated target fragments, on a Lightcyler qPCR machine (Roche, Indianapolis, IN). Lightcycler software was used to generate standard curves covering a range of 10^6 to which gland and leaf data were compared. Two biological replicates were used to generate the means and standard deviations shown in Supplementary Table 4A at *JXB* online. These values were used to compute the gland over leaf ratios and *P*-values shown in Supplementary Table 4A at *JXB* online. Raw relative expression data, means, standard deviations, *P*-values from gland versus leaf *t* tests, qPCR primer sequences, and representative real-time qPCR tracings are shown in Supplementary Table 4A at *JXB* online.

Results

Identification of candidate unigenes required for THCA production

Anatomical study revealed that glands located on mature floral bracts of female plants are the site of enhanced secondary metabolism leading to the production of THCA and other compounds in *Cannabis sativa* (Kim and Mahlberg, 1997) (Fig. 1A, C, D). These glands are located on multicellular stalks and typically are composed of eight cells (Hammond and Mahlberg, 1977, 1978). The outer gland surface is composed of a smooth capsule covered by a membrane. The capsule contains exudates derived from the gland cells (Kim and Mahlberg, 2003). The weakly attached glands can easily be separated from the bracts and purified as shown in Fig. 1E and F. An EST library was constructed using RNA isolated from purified glands. Over 100 000 ESTs were cloned. Plasmid DNA was isolated and sequenced from over 2000 clones. Because of the directed orientation of cDNA insertion, sequences are expected to represent the coding strand. After the removal of vector-only, poor quality sequences, and sequences obviously originating from organelles or ribosomal RNA, the remaining sequences were clustered into 1075 unigenes (see Suplementary Tables 1 and 2 at *JXB* online). Overall, 111 of the unigenes were contigs containing two or more closely related ESTs (see Supplementary Table 2 at *JXB* online). Only 14 contigs lacked a similar sequence in the NCBI database. Nine hundred and sixty four of the ESTs were only found once and of these 710 were similar to sequences in the NCBI database (see Supplementary Fig. 2 at *JXB* online).

The top three unigenes representing the greatest number of ESTs encoded proteins related to metallothionein, RD22-like BURP domain-containing proteins, and chitin binding hevein-like proteins (Table 1). All three of these proteins have functions related to biotic or abiotic stress responses (Cobbett and Goldsbrough, 2002; Granger *et al.*, 2002; Van den Bergh *et al.*, 2004). Gene Ontology (GO) analysis was performed on the sequence dataset (Ashburner *et al.*, 2000). An analysis of biological function indicates that 27% of the unigenes encode proteins with metabolic activity. Unigenes with NCBI matches encoding proteins with unknown function comprise 14% of the total and another 28% are predicted to be involved in various cellular processes such as protein synthesis and protein degradation.

Candidate genes in THCA biosynthesis

The specific biochemical steps leading to THCA are proposed to begin with a reaction involving a type III PKS enzyme that catalyses the synthesis of olivetolic acid from hexanoyl-CoA and three molecules of malonyl-CoA (Fig. 2A; Fellermeier *et al.*, 2001). Malonyl-CoA is derived from the carboxylation of acetyl-CoA. ESTs encoding acetyl-CoA carboxylase were identified. Hexanoyl-CoA could be produced by more than one pathway in the trichomes. One route to produce hexanoyl-CoA would involve the early termination of the fatty acid biosynthetic pathway, yielding hexanoyl-ACP (Fig. 2B). The hexanoyl moiety would then be transferred to CoA by the action of an ACP-CoA transacylase or it would be cleaved by the action of a thioesterase, yielding *n*-hexanol, which would then be converted into *n*-hexanoyl-CoA by the action of acyl-CoA synthase. Most of the enzymes needed for this route are represented in the EST database, except for the

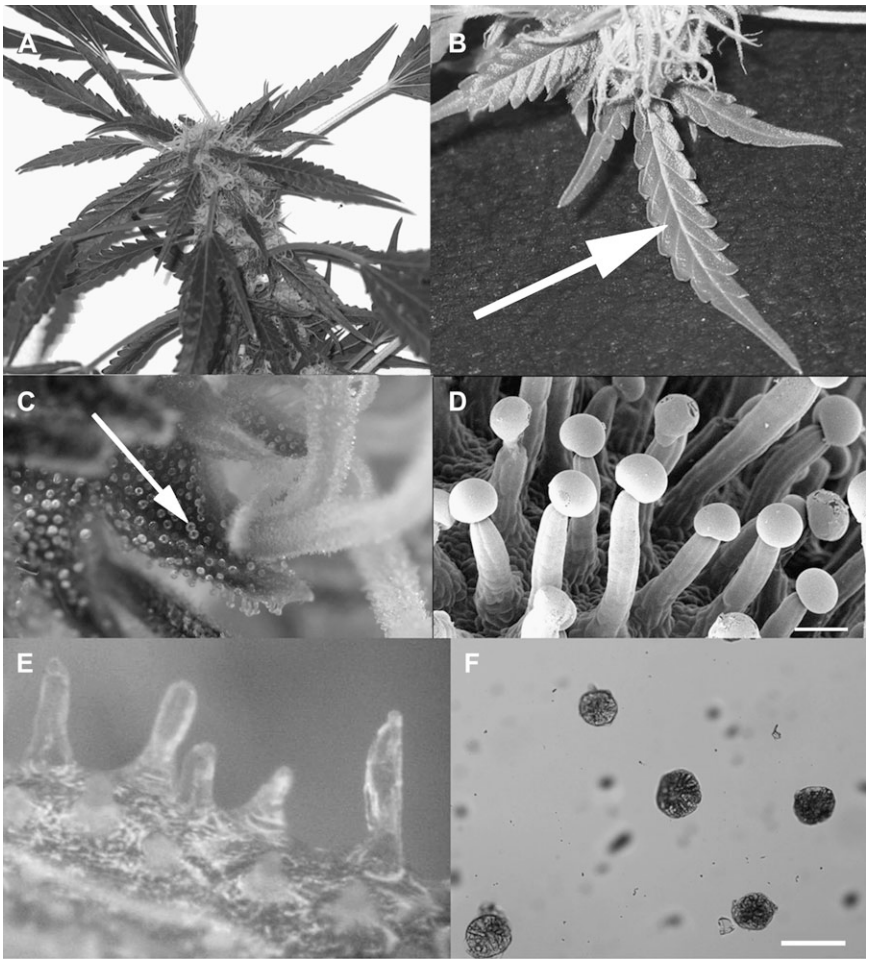

Fig. 1. Isolation of glandular trichomes from mature female bracts. (A) Female inflorescence 8 weeks after germination. (B) Leaves associated with the female inflorescence (note arrow). (C) Glands on the bracts of pistillate florets, each bearing a pale bifid style (note arrow indicating glands. (D) Scanning electron micrograph of capitate glands. (E) Micrograph of bract after removal of glands. (F) Micrograph of isolated glands. Bar in (D) and (F)=80 μm.

transacylase and 2,3-*trans*-enoyl-ACP reductase (Fig. 2B). A second route to hexanoyl-CoA would involve the production of hexanol from the breakdown of the fatty acid linoleic acid via the lipoxygenase (LOX) pathway (Fig. 2C; Hatanaka, 1999). A survey of the sequenced ESTs revealed candidate genes encoding the enzymes needed to synthesize linoleic acid from acetyl-CoA by the typical fatty acid biosynthetic pathway in plastids followed by the production of hexanol from linoleic acid via the LOX pathway. An third pathway related to the biosynthesis of branched chain amino acids has been proposed to be involved in the production of short-chain and medium-chain fatty acids (Kroumova *et al.*, 1994). However, the enzymes in this pathway [2-isopropylmalate synthase, 3-isopropylmalate dehydratase, 3-isopropylmalate dehydrogenase, and 2-oxoisovalerate dehydrogenase (acylating)] were not represented in the *Cannabis* trichome EST library.

After the formation of olivetolic acid, a prenyltransferase is proposed to add a prenyl group derived from geranyl diphosphate (GPP) to create cannabigerolic acid. GPP is derived from the fusion of two isoprene units (Fig. 2D). Two different biochemical pathways support the synthesis of isoprenoids in plants (Eisenreich *et al.*, 1998; Rohmer, 1999). Within the list of unigenes all but one of the enzymatic activities needed to convert pyruvate and glyceraldehyde-3-phosphate into isopentenyl and dimethylallyl diphosphate via the methylerythritol 4-phosphate (MEP) pathway were represented (Fig. 2D). This finding is consistent with isotopic studies showing that the GPP cannabinoid precursors are synthesized via this pathway (Fellermeier *et al.*, 2001). The formation of GPP is mediated by GPP synthase. Several unigenes related to GPP synthase were identified (with an average of 10% identity), however, they were more closely related to other terpene synthases. In particular, CAN36 and CAN55, which possibly were derived from the same gene, and the closely related CAN37, are most similar to hop (*Humulus lupulus*) sesquiterpene synthases HISTS1 and HISTS2 (NCBI accessions EU760350 and EU760351, respectively), with an average identity of 56% over the first 160 amino acid residues (Wang *et al.*, 2008). CAN41 is most similar to hop monoterpene synthase HIMTS2 (NCBI accession EU760349; 66% identity over the first 160 amino acid residues; Wang *et al.*, 2008).

Table 1. NCBI hits to most abundant *Cannabis* gland unigenes

NCBI accession	NCBI description	Unigene	No. of ESTs
ABM21763.1	Metallothionein-like protein MT1A	CAN4	31
BAB60848.1	RD22-like BURP domain-containing protein	CAN5	22
AAL30422.1	Hevein-like protein	CAN7	19
AAF73006.1	NADP-dependent malic protein	CAN6	18
1Q53	Hypothetical *Arabidopsis thaliana* protein At3g17210	CAN14	9
ABB29926.1	Fructose-*bis*phosphate aldolase-like	CAN19	9
NP_194153.1	VEP1 (VEIN PATTERNING)	CAN15	8
AAX11454.1	Chalcone synthase	CAN24	8
O24248	Microsomal oleic acid desaturase	CAN20	8
AAD22104.1	Major allergen Pru av 1 (Pru a 1)	CAN21	6
ABA12220.1	Translation elongation factor 1A-4	CAN26	6
AAC50014.1	B12D protein	CAN22	5
AAO33357.1	Non-specific lipid transfer protein 1	CAN25	5
AAO37754.1	Delta-12 oleate desaturase	CAN27	5
ABA27052.1	TO71-3	CAN30	5
AAG21984.1	LYTB-like protein precursor	CAN31	5
AAD56020.1	Elongation factor-1 alpha 3	CAN33	4
O22342	ADP,ATP carrier protein 1, mitochondrial precursor	CAN38	4
AAF85975.1	Cytosolic phosphoglycerate kinase	CAN40	4
AAS79351.1	Terpenoid synthase	CAN41	4
AAF35186.1	Lipid transfer protein precursor	CAN42	4
AAA33401.1	Ubiquitin	CAN44	4
BAB47196.1	Aspartate aminotransferase glyoxysomal isozyme AAT1 precursor	CAN23	3
Q96423	*Trans*-cinnamate 4-mono-oxygenase (cytochrome P450 73)	CAN32	3
BAB09969.1	Cytochrome P450	CAN35	3
XP_476547.1	Putative succinate dehydrogenase flavoprotein alpha subunit	CAN47	3
NP_172092.1	Kinase/ribokinase	CAN54	3
AAS66357.1	Terpene synthase	CAN55	3

The nature of the prenyltransferase is unknown. However, previous studies identified a soluble aromatic geranylpyrophosphate:olivetolate geranyltransferase in the extract of young leaves with the appropriate activity (Fellermeier and Zenk, 1998). The only EST encoding a predicted prenyltransferase was CAN121. However, the encoded protein is more closely related to members of the membrane-bound chloroplast-localized family of prenyltransferases than to soluble prenyltransferases (Heide, 2009). The final step in the pathway is mediated by THCA synthase, which mediates the conversion of cannabigerolic acid to THCA (Fig. 2A). Two ESTs (CAN296 and 720) with sequences identical to the previous reportedly THCA synthase were identified (Sirikantaramas *et al.*, 2004).

Characterization of Cannabis PKS *genes*

Whereas the nature of the prenyltransferase responsible for the synthesis of cannabigerolic acid is unknown, three unigenes, CAN24, CAN383, and CAN1069, comprising eight, one, and two ESTs, respectively, could encode the PKS activity needed to synthesize olivetolic acid. These were therefore characterized in more detail. All three unigenes were represented by individual ESTs encoding complete PKS polypetides. These were sequenced and compared to related PKS sequences (Fig. 3; see Supplementary Table 3 at *JXB* online). CAN1069 was identical to a previously identified *Cannabis* gene encoding a chalcone synthase, and is the most closely related of the PKS sequences to other known chalcone synthases from hop and *Arabidopsis* (Raharjo *et al.*, 2004*b*). The relationships of hop phlorisovalerophenone synthase (VPS), which mediates the conversion of malonyl-CoA and isovaleryl-CoA to phlorisovalerophenone, to CAN24 and CAN383 is less clear (Okada and Ito, 2001). CAN24 and CAN383 show 64.6% identity and are nearly equally similar to hop VPS at 71.2% and 72.0%, respectively.

The enzymatic activities encoded by CAN24 and CAN1069 were explored in detail. The coding regions of the two genes were inserted into the pHis8 vector in frame with a His8 tag. The tagged proteins were purified on a nickel-containing magnetic bead matrix and were assayed for chalcone and olivetol/olivetolic acid synthase activities (Fig. 4). Recombinant protein from CAN1069, but not CAN24, produced reaction products when incubated with coumaroyl-CoA and malonyl-CoA (Fig. 4C, peak 2). The reaction products were analysed by LC-MS and peak 2 was found to have a molecular mass and absorption spectrum consistent with naringenin (Fig. 4E), the major product of chalcone synthases. Both CAN24 and CAN1069 were capable of using malonyl-CoA and hexanoyl-CoA as reaction substrates (Fig. 4C, peak 1; D, peak 3) and LC-MS indicated that products of these enzymes were the same, but neither molecular mass nor the absorption spectrum of this product matched olivetol or olivetolic acid (Fig. 4F). Results similar to CAN24 were obtained using protein purified from CAN383 (data not shown).

qPCR of candidate genes in THCA biosynthesis

Genes required for THCA production are probably more highly expressed in glands of pistillate inflorescences because this is where THCA is most highly concentrated. To test this hypothesis, the relative expression levels in isolated glands versus young inflorescence-associated leaves of selected unigenes were compared using real-time qPCR. The identity of the genes assayed and the differences in relative expression levels are listed in Table 2 and in Supplementary Table 4A at *JXB* online. Consistent with this hypothesis, THCA synthase (CAN720) expression was 437 times higher in isolated glands than in leaves. CAN24 was expressed 1600 times higher in glands of the inflorescence than in associated leaves. CAN1069 encoding CHS was also more highly expressed in glands than leaves (220-fold). The expression of a third PKS, CAN383, was expressed at similar levels in glands and leaves ($Pval$=0.564). These results are not explained by poor

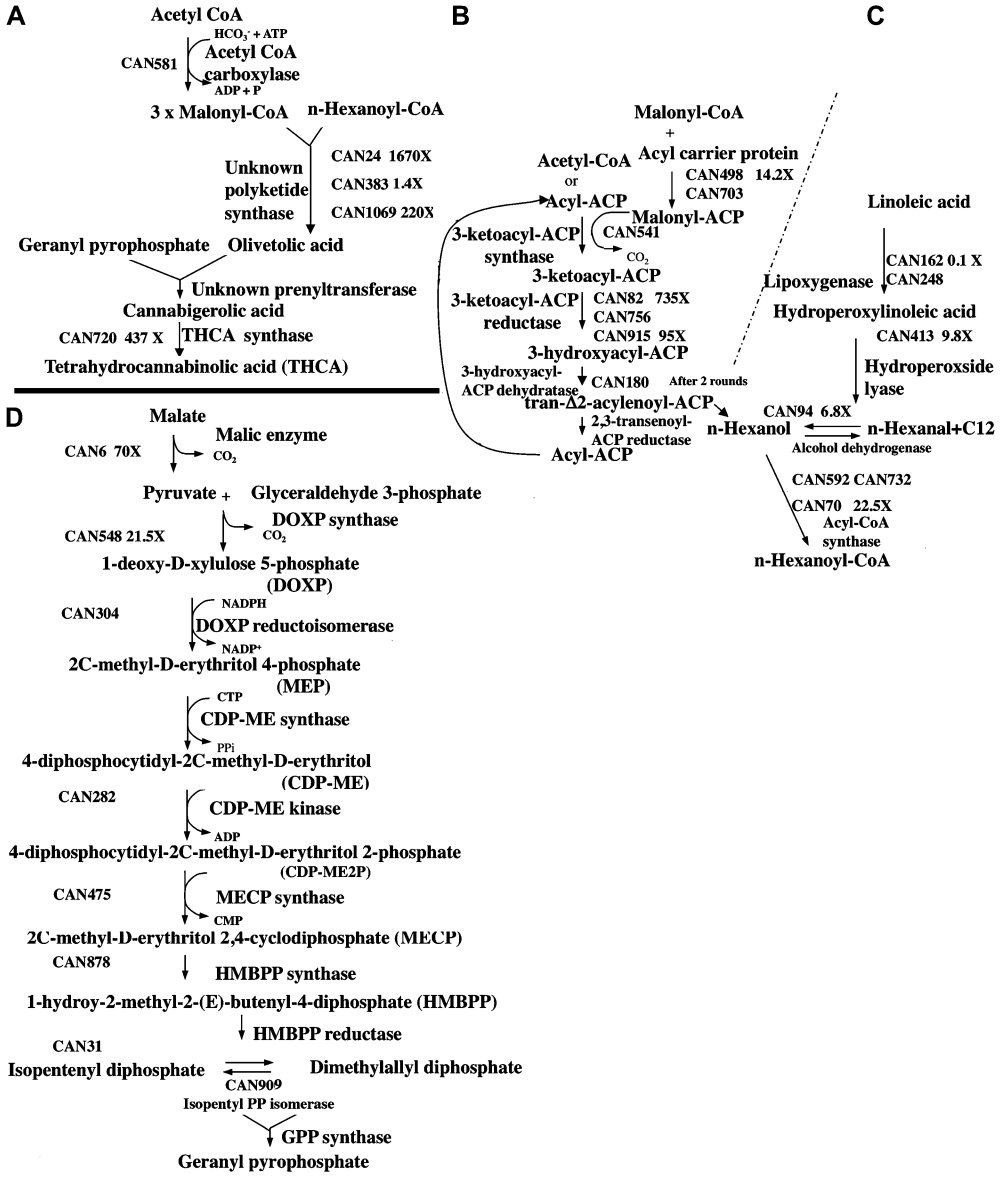

Fig. 2. Biochemical pathways leading from primary metabolites to THCA showing candidate *Cannabis* unigenes encoding enzymes in the pathway. (A) Production of THCA. (B) *De novo* fatty acid pathway leading to the formation of hexanol. (C) Breakdown of fatty acids leading to the formation of hexanol. (D) MEP pathway leading to geranyl pyrophosphate. Candidate *Cannabis* unigenes are shown, as well as the relative gland over leaf expression ratios for a subset of the candidates (see Table 2).

RNA isolation from leaves as unigene CAN219 encoding chlorophyll A/B binding protein showed elevated leaf expression levels (Table 2; See Supplementary Table 4A, C at *JXB* online). The activities of several housekeeping genes were also tested. A relatively modest increase in levels of histone H2A (CAN986) and beta tubulin (CAN1084) expression in glands compared to leaves was detected. The increase in expression levels of these latter two genes might reflect a combination of the heightened metabolic activity and the unique cellular structure of glandular trichomes.

Two different pathways could provide the hexanol required for olivetolic acid synthesis, as shown in Fig. 2. Expression levels provide support for the *de novo* pathway as a primary source, given that CAN498, CAN82, and CAN915 were much more highly expressed in glands than

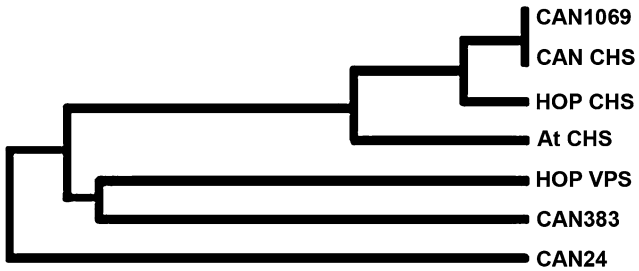

Fig. 3. Comparison of PKS-related unigenes from *Cannabis sativa* glandular trichomes. Accession numbers: *Arabidopsis thaliana* CHS, NP_196897; hop (*Humulus lupulus*) VPS, BAB121202; hop CHS, CAK19319; *Cannabis* CHS, AAL92879. The alignment was generated as described in the Materials and methods. The distance relationships are described in the main text.

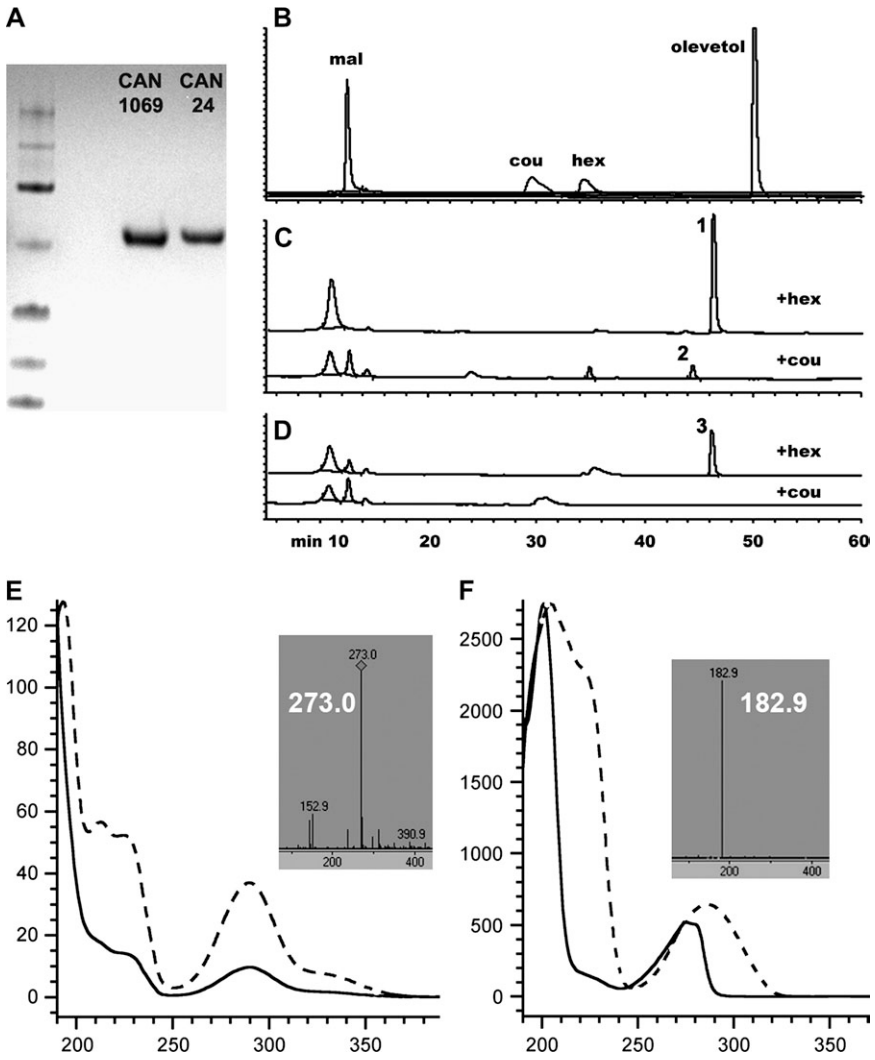

Fig. 4. Analysis of PKS activity for two candidate gene products expressed in *Cannabis* glands. (A) SDS-PAGE analysis of PKS proteins encoded by CAN1069 and CAN24 expressed in *E. coli* harbouring the corresponding gene expression plasmids. (B) HPLC separation of the reactants malonyl-CoA (mal), coumaroyl-CoA (cou), and hexanol-CoA (hex), and the potential product olivetol. Separation by HPLC of reaction products obtained using CAN1069 protein (C) or CAN24 protein (D) with either hex or cou as substrates. Numbered peaks represent products not seen in control reactions containing boiled protein. (E) Absorption spectrum of product 2 (solid line) compared to naringenin (dashed line). The molecular mass of the most abundant ion species is shown in the insert. (F) Absorption spectrum of product 3 (solid line) compared to olivetol (dashed line). The molecular mass of the most abundant ion species is shown in the insert. Product 1 yielded the same mass and spectrum as product 3 shown in (F).

leaves (Fig. 2; Table 2; see Supplementary Table 4A at *JXB* online), whereas the relative expression of genes encoding enzymes in the lipid breakdown pathway were depressed or modestly elevated in glands.

Identification of potential transcription factors expressed in Cannabis *glands*

Cannabis unigenes were compared to known *Arabidopsis thaliana* transcription factors (Guo *et al.*, 2005). Eighty *Cannabis* unigenes were similar to transcription factors found in *Arabidopsis* and 11 contain MYB DNA binding domains (see Supplementary Table 5 at *JXB* online for *P*-scores). Expression of four MYB genes in isolated glands and leaves was compared by real-time qPCR (Table 2; see Supplementary Table 4A at *JXB* online). CAN833 and CAN738 exhibited 954-fold and 586-fold higher expression in glands, respectively, whereas CAN483 and 792 showed more modest induction in glands. None of the other putative transcription factors that were assayed showed the same degree of differential expression as CAN833 and CAN738 (Table 2).

Discussion

Characterization of Cannabis *gland ESTs*

The identities of the most abundant ESTs derived from the glandular trichomes of *Cannabis sativa* are consistent with the protective function of plant glands. For example, the

Table 2. Real-time qPCR analysis of selected unigenes

See Supplementary Table 4 at *JXB* online for raw data, means, and standard deviations and qPCR primers. Means and standard deviations were used in *t* tests to calculate the *P* values shown in the last column.

Unigene	Predicted protein	Gland versus leaf	*P*-value
MEP pathway			
CAN548	DOXP synthase	21.5	0.0002
CAN6	Malic enzyme	70.5	0.0001
***de novo* fatty acid pathway**			
CAN498	Acyl carrier protein	148	0.0001
CAN82	3-ketoacyl-ACP reductase	735	0.0005
CAN915	3-ketoacyl-ACP reductase	95.3	0.0001
Fatty acid breakdown pathway			
CAN162	Lipoxygenase	0.0754	0.0001
CAN295	Lipoxygenase	1.03	0.87
CAN413	Hydroperoxide lyase	9.15	0.0001
CAN94	3-ketoacyl-CoA thiolase; acetyl-CoA acyltransferase	6.79	0.0001
THCA synthesis			
CAN70	Acyl-CoA synthetase	22.5	0.0001
CAN24	Putative olivetolic acid synthase	1670	0.0003
CAN1069	CAN CHS	220	0.0001
CAN383	CHS-like protein	1.4	0.564
CAN720	THCA synthase	437	0.0009
Control genes			
CAN1084	Beta tubulin	3.45	0.0397
CAN986	Histone H2A	4.24	0.0002
CAN219	Chlorophyll *a/b* binding protein	0.0246	0.0019
Putative transcription factors			
CAN833	AtMYB78&112	954	0.0001
CAN738	AtMYB12&111	586	0.0001
CAN483	MYB related	7.09	0.0004
CAN792	MYB related	8.51	0.0012
CAN747	COP9 complex subunit	11.9	0.0002
CAN227	Homeodomain *HAT4*	3.48	0.0024
CAN618	*CONSTANS-LIKE3*	0.00531	0.1903

most abundant ESTs encoded a protein closely related to type II metallothioneins. These proteins bind heavy metals such as Cd, Zn, and Cu, and their proposed primary function is the maintenance of Cu tolerance (Cobbett and Goldsbrough, 2002). The second most abundant class of ESTs encoded an RD22-like BURP domain containing protein. This class of proteins contains a hydrophobic N-terminal signal peptide, and an N-terminal conserved region followed by a series of small repeats (Granger *et al.*, 2002). The BURP domain of approximately 230 amino acids is located in the C-terminal region. The function of RD22-like proteins is unknown but some members of this class of genes are induced by dehydration (Yamaguchi-Shinozaki and Shinozaki, 1993). The third most abundant ESTs encoded a protein containing a hevein domain. Hevein domains contain a conserved 43-amino acid motif that binds chitin and members of this protein class are known for antifungal activity (Van den Bergh *et al.*, 2004). The unique secondary metabolism in *Cannabis* may also play a role in plant defence. Synthesis of THCA is extracellular and results in hydrogen peroxide production, which has general antimicrobial properties (Sirikantaramas *et al.*, 2005), and a recent report further indicates that THCA may directly inhibit microbial growth (Appendino *et al.*, 2008).

The analysis of gland-derived ESTs has identified nearly all the candidate genes required for THCA synthesis from primary metabolic products. These findings differ from a proteomic study that aimed to identify genes expressed in *Cannabis* glands but failed to associate any highly expressed proteins with THCA synthesis (Raharjo *et al.*, 2004*c*). This difference reflects the much greater volume of genomic data enabling more robust identification of DNA sequences when compared to proteomics approaches based on the molecular weights of fragmented polypeptides. This is especially true for species such as *Cannabis sativa* for which there is little amino acid sequence data available to compare with peptide profiles.

The present study highlights the utility of using isolated glands as starting material for making EST libraries to study gland metabolism, as was the case in other plant species (Lange *et al.*, 2000; Gang *et al.*, 2001; Fridman *et al.*, 2005). In this study more than 50% of the ESTs with NCBI matches were involved in metabolism or cellular activities such as transport and protein translation. Many other cannabinoids, in addition to THCA, have been identified in *Cannabis* (Mechoulam, 1970), and it is likely that many of the genes identified in Supplementary Table 2 at *JXB* online are involved in the production of these other compounds. In addition to cannabinoids, many other classes of secondary compounds have been found in *Cannabis* (Flores-Sanchez and Verpoorte, 2008*b*). For example, both monoterpenes and sesquiterpenes have been identified and candidate ESTs encoding activities to produce these compounds have been identified.

Characterization of Cannabis *PKSs and the difficulty in the identification of olivetolic acid synthase*

Synthesis of olivetolic acid from malonyl-CoA and hexanoyl-CoA represents the first committed step toward the synthesis of THCA. Olivetolic acid synthesis is predicted to be mediated by a member of the type III PKS family through a series of three condensation reactions producing a triketide (Fig. 7B; Fellermeier *et al.*, 2001). CAN24, represented by eight ESTs and one of the most highly expressed unigenes in our analysis, encodes a member of the PKS family. This gene was expressed 1600-fold higher in glands than in leaves. CAN1068, another PKS member represented by two ESTs, corresponds to a previously identified *Cannabis* CHS gene (Raharjo *et al.*, 2004*b*). A third PKS represented by a single EST, CAN383, was also identified. Analyses of PKS crystal structures indicate that

the type III PKS enzymes are composed of a dimer with conserved reaction centres and a hollow reaction cavity (Austin and Noel, 2003). All three *Cannabis* PKS genes encoded polypeptides containing the conserved amino acids, Cys 167, His 307, and Asn 340, that are believed to constitute the reaction centre (see Supplementary Table 3 at *JXB* online; Jez *et al.*, 2000). In addition, two of the three amino acids that are important for defining the size of the reaction cavity in chalcone synthases (Jez *et al.*, 2000; Abe *et al.*, 2006) are conserved. The third amino acid, Thr at position 300 that is conserved in all chalcone synthases, was missing in CAN24 and CAN383 (see Supplementary Table 3 at *JXB* online). Instead CAN24 and CAN383 contained Leu and Iso, respectively, at position 300. Such differences might alter substrate specificity.

It has been proposed that either olivetol or olivetolic acid are products of polyketide synthase in the THCA pathway (Raharjo *et al.*, 2004*b*). However assays of plant extracts found that olivetol, the decarboxylation product of olivetolic acid, was not a substrate in the pathway (Fellermeier and Zenk, 1998). Products of the three PKS genes identified in this study were tested for olivetolic acid synthesis *in vitro*. CAN24 and CAN383 yielded identical products according to HPLC analysis (data not shown). Because CAN24 was more abundant, this PKS gene was analysed in detail, along with CAN1069, which had CHS activity as shown in Fig. 4. The size of the product produced by the CAN24-encoded enzyme was smaller than olivetolic acid (Fig. 4F). Further, the absorption spectrum did not match olivetol. A sequence identical to CAN24 has been deposited in the NCBI database (accession AB164375) and was annotated as having olivetol synthase activity but without supporting data. The product generated in our analyses possibly represents a derailment product in which the enzyme catalyses two decarboxylative condensations instead of three (Fig. 5). The failure of *in vitro* conditions to support the complete synthesis of native PKS products is well-documented (Springob *et al.*, 2007). The artificial nature of *in vitro* reaction products is reflected by the failure to find these products *in vivo* (Austin and Noel, 2003). This may indicate that reaction conditions were insufficient or that accessory proteins may be needed to produce olivetolic acid. Similar results were obtained in a study that characterized a gene identical to CAN1069 where the enzyme could use hexanoyl-CoA as a starter molecule, but only yielded a possible derailment product (Raharjo *et al.*, 2004*b*). That plant extracts used in *in vitro* assays have also not yielded olivetolic acid suggests that *in vivo*-like reaction conditions have yet to be imitated (Raharjo *et al.*, 2004*a*; Flores-Sanchez and Verpoorte, 2008*a*).

The assay results with CAN1069 also highlight the permissiveness of substrate use by PKSs. CAN1069 clearly had CHS activity in that it could use coumaroyl-CoA as a substrate to produce naringenin. Given that CAN1069 was preferentially expressed in the glands and can act on hexanoyl-CoA, this PKS also may contribute to THCA production.

Insights on THCA biosynthesis may be gained by comparison with the secondary metabolism of glandular trichomes in hop, the closest relative to *Cannabis*. Production of the bitter acid humulone in hop inflorescence glands requires a PKS called VPS. In two independent studies, ESTs representing VPS were among the most abundant in the collections and were at least 5-fold more abundant than other PKS encoding ESTs (Nagel *et al.*, 2008; Wang *et al.*, 2008). Accordingly, the role of CAN24 as olivetolic acid synthase remains tentative pending further biochemical support, however, it is the best candidate based on expression data.

Identification of MYBs with potential roles in gland chemistry

A large number of *Cannabis* gland unigenes encoded proteins similar to transcription factors (see Supplementary Table 5 at *JXB* online). Several of these were analysed by qPCR and two were found to be preferentially expressed in glands. This is potentially significant as studies have shown that all but one of the 12 transcription factors required for *Arabidopsis* trichome formation are preferentially expressed in trichomes compared to whole leaves (Marks *et al.*, 2009). In this study it was found that two R2R3-type *Cannabis* MYBs, encoded by CAN833 and CAN738, were preferentially expressed in isolated glands compared to leaves by 954-fold and 586-fold, respectively. CAN833 and CAN738 are most similar to the *Arabidopsis* MYBs related to AtMYB112 and AtMYB12, respectively (see Supplementary Table 5 at *JXB* online; Stracke *et al.*, 2001). AtMYB112 corresponds to the *BOTRYTIS SUSCEPTIBLE1* (*BOS1*) gene. *bos1* mutants are more susceptible to pathogens such as *Botrytis cinerea* and *Alternaria brassicicola*, and have impaired tolerance to oxidative stress (Mengiste *et al.*, 2003). A role for CAN883 in tolerance to oxidative stress in *Cannabis* glandular trichomes is logical, as the last enzymatic reaction in THCA synthesis releases hydrogen peroxide (Sirikantaramas *et al.*, 2005). AtMYB12 controls the synthesis of flavonol secondary metabolites in *Arabidopsis* and can induce the synthesis of similar compounds in tobacco (Mehrtens *et al.*, 2005; Luo *et al.*, 2008). Flavonoids have been isolated from *Cannabis* leaves and flowers, but evidence is lacking for gland flavonoid production (Flores-Sanchez and Verpoorte, 2008*b*). Since flavonols are not predominant in *Cannabis* glands, it is possible that CAN738 instead plays a role in controlling the expression of genes required for other secondary metabolites in *Cannabis* such as THCA.

Potential utility of comparing gland ESTs from hop and Cannabis

The PKSs and many other genes identified in this study are closely related to those from hop (*Humulus lupulus*). *Humulus* and *Cannabis* are monotypic sister genera in the family Cannabaceae (Datwyler and Weiblen, 2004).

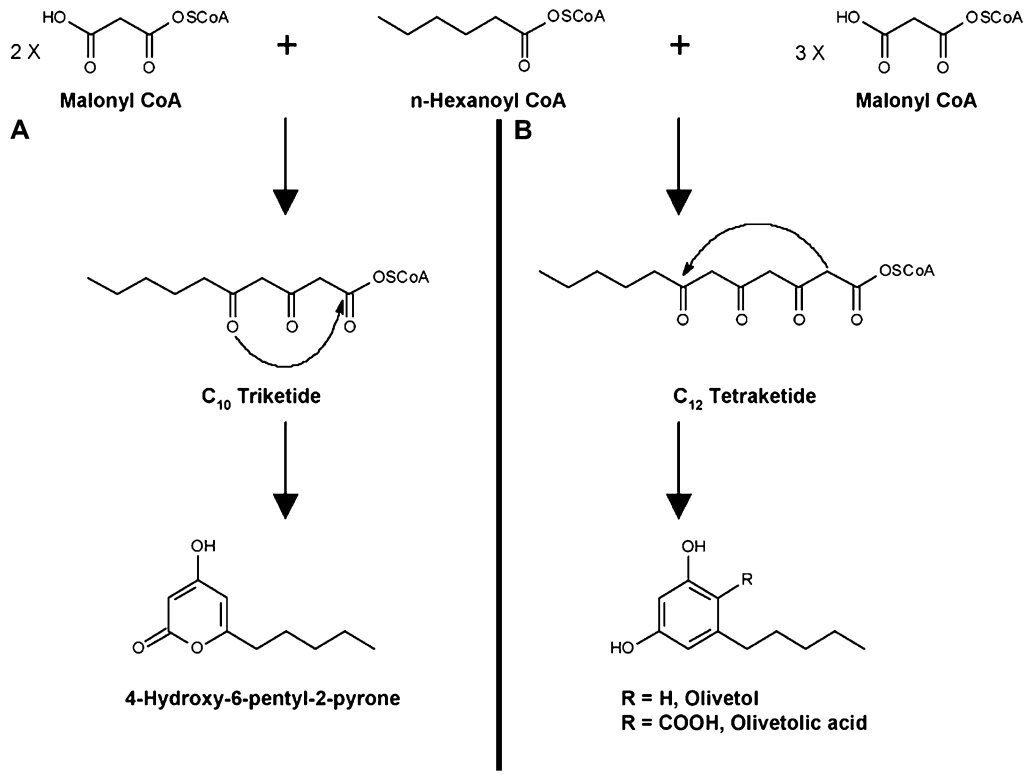

Fig. 5. Comparison of potential products produced by two or three decarboxylative condensation reactions. (A) Prediction of pyrone production from two decarboxylative condensations. (B) Expected product from three decarboxylative condensations.

Glandular trichomes located on the inflorescence bracts of both *Humulus* and *Cannabis* are the location of unique PKS-derived secondary metabolism (Nagel *et al.*, 2008; Wang *et al.*, 2008). Hop glands produce the bitter acid humulone, which is important for beer flavour, and the prenylated chalcone xanthohumol, which has several potential health beneficial properties (Stevens and Page, 2004). The biochemical pathways leading to THCA, xanthohumol, and humulone have common steps that include polyketide synthases and prenyltransferases. It is probable that these plants share other homologous biochemical pathways given their close ancestry. Information from *Cannabis* ESTs has the potential to improve the understanding of hop biochemical pathways as well.

Supplementary data

Supplementary data are available at *JXB* online.

Supplementary Table 1. Sequences of the derived unigenes along with the names of the ESTs used to generate the sequences.

Supplementary Table 2. The BLAST analysis of the unigene dataset.

Supplementary Table 3. A comparison of the polyketide synthase amino acid sequences analysed in this study.

Supplementary Table 4. The raw relative qPCR expression data with means, standard deviations, *t* test *P*-values, and representative real-time qPCR tracings.

Supplementary Table 5. A BLAST comparison of the *Cannabis* gland unigenes against a list of transcription factors found in *Arabidopsis thaliana*.

Acknowledgements

This work was funded by a sabbatical supplement from the University of Minnesota College of Biological Sciences (MDM), by the University of Minnesota Life Sciences Summer Undergraduate Research Programs (SNO and WS-F), the Samuel Roberts Noble Foundation (LT and RAD), the Minnesota Agricultural Experiment Station (GDW), and a Packard Fellowship in Science and Engineering (GDW). We acknowledge Mahmoud El Sohly and Zlatko Mehmedic at the University of Mississippi for previously published analyses of cannabinoid content under the National Institute on Drug Abuse contract number N01DA-5-7746.

References

Abe I, Watanabe T, Lou WW, Noguchi H. 2006. Active site residues governing substrate selectivity and polyketide chain length in aloesone synthase. *FEBS Journal* **273,** 208–218.

Appendino G, Gibbons S, Giana A, Pagani A, Grassi G, Stavri M, Smith E, Rahman MM. 2008. Antibacterial cannabinoids from *Cannabis sativa*: a structure–activity study. *Journal of Natural Products* **71,** 1427–1430.

Ashburner M, Ball CA, Blake JA, *et al*. 2000. Gene ontology: tool for the unification of biology. The gene ontology consortium. *Nature Genetics* **25,** 25–29.

Austin MB, Noel AJP. 2003. The chalcone synthase superfamily of type III polyketide synthases. *Natural Product Reports* **20,** 79–110.

Baker D, Pryce G, Giovannoni G, Thompson AJ. 2003. The therapeutic potential of cannabis. *Lancet Neurology* **2,** 291–298.

Cobbett C, Goldsbrough P. 2002. Phytochelatins and metallothioneins: roles in heavy metal detoxification and homeostasis. *Annual Review of Plant Biology* **53,** 159–182.

Datwyler SL, Weiblen GD. 2004. On the origin of the fig: phylogenetic relationships of Moraceae from ndhF sequences. *American Journal of Botany* **91,** 767–777.

Datwyler SL, Weiblen GD. 2006. Genetic variation in hemp and marijuana (*Cannabis sativa* L.) according to amplified fragment length polymorphisms. *Journal of Forensic Sciences* **51,** 371–375.

de Meijer EP, Bagatta M, Carboni A, Crucitti P, Moliterni VM, Ranalli P, Mandolino G. 2003. The inheritance of chemical phenotype in *Cannabis sativa* L. *Genetics* **163,** 335–346.

Eisenreich W, Schwarz M, Cartayrade A, Arigoni D, Zenk MH, Bacher A. 1998. The deoxyxylulose phosphate pathway of terpenoid biosynthesis in plants and microorganisms. *Chemical Biology* **5,** R221–R233.

ElSohly MA, Ross SA, Mehmedic Z, Arafat R, Yi B, Banahan BF. 2000. Potency trends of Delta(9)-THC and other cannabinoids in confiscated marijuana from 1980–1997. *Journal of Forensic Sciences* **45,** 24–30.

Fellermeier M, Eisenreich W, Bacher A, Zenk MH. 2001. Biosynthesis of cannabinoids. Incorporation experiments with (13)C-labeled glucoses. *European Journal of Biochemistry* **268,** 1596–1604.

Fellermeier M, Zenk MH. 1998. Prenylation of olivetolate by a hemp transferase yields cannabigerolic acid, the precursor of tetrahydrocannabinol. *FEBS Letters* **427,** 283–285.

Flores-Sanchez IJ, Verpoorte R. 2008*a*. PKS activities and biosynthesis of cannabinoids and flavonoids in *Cannabis sativa* L. plants. *Plant and Cell Physiology* **49,** 1767–1782.

Flores-Sanchez IJ, Verpoorte R. 2008*b*. Secondary metabolism in cannabis. *Phytochemistry Review* **7,** 615–639.

Fridman E, Wang J, Iijima Y, Froehlich JE, Gang DR, Ohlrogge J, Pichersky E. 2005. Metabolic, genomic, and biochemical analyses of glandular trichomes from the wild tomato species *Lycopersicon hirsutum* identify a key enzyme in the biosynthesis of methylketones. *The Plant Cell* **17,** 1252–1267.

Gang DR, Wang J, Dudareva N, Nam KH, Simon JE, Lewinsohn E, Pichersky E. 2001. An investigation of the storage and biosynthesis of phenylpropenes in sweet basil. *Plant Physiology* **125,** 539–555.

Granger C, Coryell V, Khanna A, Keim P, Vodkin L, Shoemaker RC. 2002. Identification, structure, and differential expression of members of a BURP domain containing protein family in soybean. *Genome* **45,** 693–701.

Guo A, He K, Liu D, Bai S, Gu X, Wei L, Luo J. 2005. DATF: a database of *Arabidopsis* transcription factors. *Bioinformatics* **21,** 2568–2569.

Hammond CT, Mahlberg PG. 1977. Morphogenesis of capitate glandular hairs of *Cannabis sativa* (Cannabaceae). *American Journal of Botany* **64,** 1023–1031.

Hammond CT, Mahlberg PG. 1978. Ultrastructural development of capitate glandular hairs of *Cannabis sativa* L. (Cannabaceae). *American Journal of Botany* **65,** 140–151.

Hatanaka A. 1999. Biosynthesis of so-called 'green odor' emitted by green leaves. In: Sankawa U, ed. *Comprehensive natural products chemistry*, Vol. 1. *Polyketides and other secondary metabolites including fatty acids and their derivatives*. Oxford: Elsevier, 83–115.

Heide L. 2009. Prenyl transfer to aromatic substrates: genetics and enzymology. *Current Opinion in Chemical Biology* **13,** 171–179.

Hillig KW, Mahlberg PG. 2004. A chemotaxonomic analysis of cannabinoid variation in *Cannabis* (Cannabaceae). *American Journal of Botany* **91,** 966–975.

Jez JM, Ferrer JL, Bowman ME, Dixon RA, Noel JP. 2000. Dissection of malonyl-coenzyme A decarboxylation from polyketide formation in the reaction mechanism of a plant polyketide synthase. *Biochemistry* **39,** 890–902.

Kim ES, Mahlberg PG. 1997. Immunochemical localization of tetrahydrocannabinol (THC) in cryofixed glandular trichomes of *Cannabis* (Cannabaceae). *American Journal of Botany* **84,** 336–342.

Kim ES, Mahlberg PG. 2003. Secretory vesicle formation in the secretory cavity of glandular trichomes of *Cannabis sativa* L. (Cannabaceae). *Molecules and Cells* **15,** 387–395.

Kostic M, Pejic B, Skundric P. 2008. Quality of chemically modified hemp fibres. *Bioresource Technology* **99,** 94–99.

Kroumova AB, Xie Z, Wagner GJ. 1994. A pathway for the biosynthesis of straight and branched, odd- and even-length, medium-chain fatty acids in plants. *Proceedings of the National Academy of Sciences, USA* **91,** 11437–11441.

Lange BM, Wildung MR, Stauber EJ, Sanchez C, Pouchnik D, Croteau R. 2000. Probing essential oil biosynthesis and secretion by functional evaluation of expressed sequence tags from mint glandular trichomes. *Proceedings of the National Academy of Sciences, USA* **97,** 2934–2939.

Long LE, Malone DT, Taylor DA. 2006. Cannabidiol reverses MK-801-induced disruption of prepulse inhibition in mice. *Neuropsychopharmacology* **31,** 795–803.

Luo J, Butelli E, Hill L, Parr A, Niggeweg R, Bailey P, Weisshaar B, Martin C. 2008. AtMYB12 regulates caffeoyl quinic acid and flavonol synthesis in tomato: expression in fruit results in very high levels of both types of polyphenol. *The Plant Journal* **56,** 316–326.

Marks MD, Gilding E, Wenger JP. 2007. Genetic interaction between glabra3-shapeshifter and siamese in *Arabidopsis thaliana* converts trichome precursors into cells with meristematic activity. *The Plant Journal* **522,** 352–361.

Marks MD, Wenger JP, Gilding E, Jilk R, Dixon RA. 2009. Transcriptome analysis of Arabidopsis wild type and gl3-sst sim trichomes identifies four additional genes required for trichome development. *Molecular Plant* **2,** (in press).

Measham F, Newcombe R, Parker H. 1994. The normalization of recreational drug use amongst young people in north-west England. *British Journal of Sociology* **45,** 287–312.

Mechoulam R. 1970. Marihuana chemistry. *Science* **168,** 1159–1166.

Mechoulam R. 2005. Plant cannabinoids: a neglected pharmacological treasure trove. *British Journal of Pharmacology* **146,** 913–915.

Mehrtens F, Kranz H, Bednarek P, Weisshaar B. 2005. The Arabidopsis transcription factor MYB12 is a flavonol-specific regulator of phenylpropanoid biosynthesis. *Plant Physiology* **138,** 1083–1096.

Mengiste T, Chen X, Salmeron J, Dietrich R. 2003. The *BOTRYTIS SUSCEPTIBLE1* gene encodes an R2R3MYB transcription factor protein that is required for biotic and abiotic stress responses in Arabidopsis. *The Plant Cell* **15,** 2551–2565.

Nagel J, Culley LK, Lu Y, Liu E, Matthews PD, Stevens JF, Page JE. 2008. EST analysis of hop glandular trichomes identifies an *O*-methyltransferase that catalyses the biosynthesis of xanthohumol. *The Plant Cell* **20,** 186–200.

Okada Y, Ito K. 2001. Cloning and analysis of valerophenone synthase gene expressed specifically in lupulin gland of hop (*Humulus lupulus* L.). *Bioscience, Biotechnology and Biochemistry* **65,** 150–155.

Pacifico D, Miselli F, Micheler M, Carboni A, Ranalli P, Mandolino G. 2006. Genetics and marker-assisted selection of the chemotype in *Cannabis sativa* L. *Molecular Breeding* **17,** 257–268.

Pertwee RG. 2006. Cannabinoid pharmacology: the first 66 years. *British Journal of Pharmacology* **147,** S163–S171.

Raharjo TJ, Chang WT, Choi YH, Peltenburg-Looman AMG, Verpoorte R. 2004*a*. Olivetol as product of a polyketide synthase in *Cannabis sativa* L. *Plant Science* **166,** 381–385.

Raharjo TJ, Chang WT, Verberne MC, Peltenburg-Looman AM, Linthorst HJ, Verpoorte R. 2004*b*. Cloning and over-expression of a cDNA encoding a polyketide synthase from *Cannabis sativa*. *Plant Physiology and Biochemistry* **42,** 291–297.

Raharjo TJ, Widjaja I, Roytrakul S, Verpoorte R. 2004*c*. Comparative proteomics of *Cannabis sativa* plant tissues. *Journal of Biomolecular Technology* **15,** 97–106.

Rohmer M. 1999. The discovery of a mevalonate-independent pathway for isoprenoid biosynthesis in bacteria, algae and higher plants. *Natural Products Reports* **16,** 565–574.

Sambrook J, Fritsch EF, Maniatis T. 1989. *Molecular cloning: a laboratory manual*. Cold Spring Harbor, New York: Cold Spring Harbor Laboratory Press.

Sirikantaramas S, Morimoto S, Shoyama Y, Ishikawa Y, Wada Y, Taura F. 2004. The gene controlling marijuana psychoactivity: molecular cloning and heterologous expression of Delta1-tetrahydrocannabinolic acid synthase from *Cannabis sativa* L. *Journal of Biological Chemistry* **279,** 39767–39774.

Sirikantaramas S, Taura F, Tanaka Y, Ishikawa Y, Morimoto S, Shoyama Y. 2005. Tetrahydrocannabinolic acid synthase, the enzyme controlling marijuana psychoactivity, is secreted into the storage cavity of the glandular trichomes. *Plant and Cell Physiology* **46,** 1578–1582.

Springob K, Samappito S, Jindaprasert A, Schmidt J, Page JE, De-Eknamkul W, Kutchan TM. 2007. A polyketide synthase of *Plumbago indica* that catalyses the formation of hexaketide pyrones. *FEBS Journal* **274,** 406–417.

Stevens JF, Page JE. 2004. Xanthohumol and related prenylflavonoids from hops and beer: to your good health! *Phytochemistry* **65,** 1317–1330.

Stracke R, Werber M, Weisshaar B. 2001. The R2R3-MYB gene family in *Arabidopsis thaliana*. *Current Opinion in Plant Biology* **4,** 447–456.

Taura F, Sirikantaramas S, Shoyama Y, Yoshikai K, Morimoto S. 2007. Cannabidiolic-acid synthase, the chemotype-determining enzyme in the fibre-type *Cannabis sativa*. *FEBS Letters* **581,** 2929–2934.

Turner CE, Bouwsma OJ, Billets S, ElSohly MA. 1980. Constituents of *Cannabis sativa* L. XVIII–Electron voltage selected ion monitoring study of cannabinoids. *Biomedical Mass Spectrometry* **7,** 247–256.

Van den Bergh KP, Rouge P, Proost P, Coosemans J, Krouglova T, Engelborghs Y, Peumans WJ, Van Damme EJ. 2004. Synergistic antifungal activity of two chitin-binding proteins from spindle tree (*Euonymus europaeus* L.). *Planta* **219,** 221–232.

Wang G, Tian L, Aziz N, Broun P, Dai X, He J, King A, Zhao PX, Dixon RA. 2008. Terpene biosynthesis in glandular trichomes of hop. *Plant Physiology* **148,** 1254–1266.

Ware MA, Tawfik VL. 2005. Safety issues concerning the medical use of cannabis and cannabinoids. *Pain Research Management* **10,** Supplement A, 31A–37A.

Yamaguchi-Shinozaki K, Shinozaki K. 1993. The plant hormone abscisic acid mediates the drought-induced expression but not the seed-specific expression of *rd22*, a gene responsive to dehydration stress in *Arabidopsis thaliana*. *Molecular and General Genetics* **238,** 17–25.

Zuardi AW, Crippa JA, Hallak JE, Moreira FA, Guimaraes FS. 2006. Cannabidiol, a *Cannabis sativa* constituent, as an antipsychotic drug. *Brazilian Journal of Medical and Biological Research* **39,** 421–429.