Available online at www.sciencedirect.com
SCIENCE DIRECT
Forensic Science International 159 (2006) 132–140
Forensic Science International
www.elsevier.com/locate/forsciint

# DNA polymorphisms in the tetrahydrocannabinolic acid (THCA) synthase gene in ''drug-type'' and ''fiber-type'' *Cannabis sativa* L.

Mareshige Kojoma [a,b,*], Hikaru Seki [b], Shigeo Yoshida [b], Toshiya Muranaka [b]

[a] *JYUGEI Institute (The Arboricultural Research Institute), University Forests, Graduate School of Agricultural and Life Sciences, The University of Tokyo, 457 Kano, Minami-izu, Shizuoka 451-0304, Japan*
[b] *Plant Science Center, RIKEN, 1-7-22 Suehirocho, Tsurumi-ku, Yokohama 230-0045, Japan*

Received 20 May 2005; received in revised form 9 July 2005; accepted 18 July 2005
Available online 6 September 2005

**Abstract**

The cannabinoid content of 13 different strains of cannabis plant (*Cannabis sativa* L.) was analyzed. Six strains fell into the ''drug-type'' class, with high Δ-9-tetrahydrocannabinolic acid (THCA) content, and seven strains into the ''fiber-type'' class, with low THCA using HPLC analysis. Genomic DNA sequence polymorphisms in the THCA synthase gene from each strain were studied. A single PCR fragment of the THCA synthase gene was detected from six strains of ''drug-type'' plants. We could also detect the fragment from seven strains of ''fiber-type'' plants, although no or very low content of THCA were detected in these samples. These were 1638 bp from all 13 strains and no intron among the sequences obtained. There were two variants of the THCA synthase gene in the ''drug-type'' and ''fiber-type'' cannabis plants, respectively. Thirty-seven major substitutions were detected in the alignment of the deduced amino acid sequences from these variants. Furthermore, we identified a specific PCR marker for the THCA synthase gene for the ''drug-type'' strains. This PCR marker was not detected in the ''fiber-type'' strains.


*Keywords:* *Cannabis sativa*; Marijuana; Hemp; THCA synthase gene; DNA polymorphism; Cannabinoid

## 1. Introduction

Cannabis (*Cannabis sativa* L.), one of the oldest cultivated plants, has been grown worldwide for thousands of years for its fiber and seed oil. In addition, the dried flowering tops and leaves are used as the marijuana and hashish. The narcotic chemical substances of cannabis plants are cannabinoids (the main substance is Δ-9-tetrahydrocannabinol, THC). Thus, cannabis used for narcotic purposes is called ''drug-type'' and has a high content of THC. On the other hand, cannabis used for fiber is called ''fiber-type'' and has low or no THC. The morphological, geographical, and chemical features of *C. sativa* are highly variable. Several studies have examined the variations in morphology and chemical content of *C. sativa* and, in particular, variations in the cannabinoid content [1–4]. Genetic variations in cannabis populations have been widely distributed throughout the world via commerce and illicit Internet trade.

* Corresponding author. Tel.: +81 558 62 0021; fax: +81 558 62 3170.
*E-mail address:* kojoma@uf.a.u-tokyo.ac.jp (M. Kojoma).

0379-0738/$ – see front matter 
doi:10.1016/j.forsciint.2005.07.005

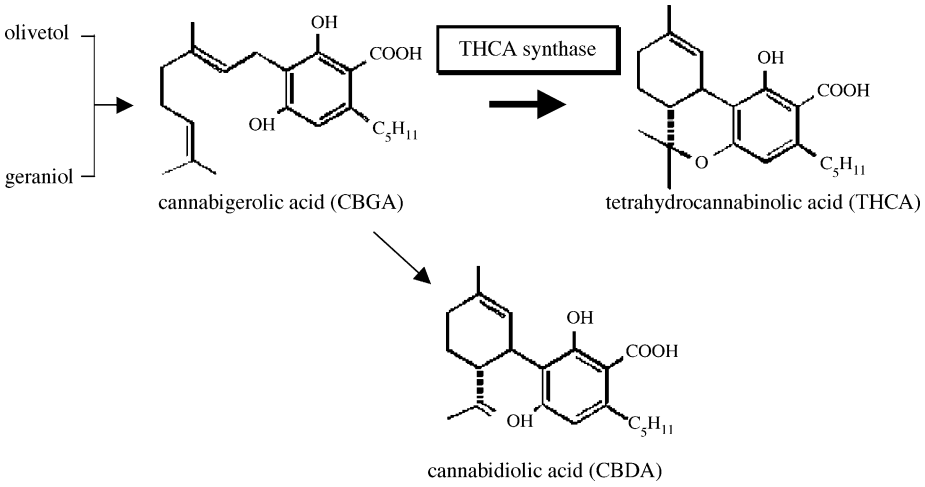


Fig. 1. The biosynthetic pathway of THCA and CBDA in *Cannabis sativa* L.

In many countries, including Japan, possession and cultivation of cannabis are strictly controlled by law enforcement, for which the accurate identification of drug material in seized samples is crucial. Several molecular techniques have been evaluated with respect to their ability to establish genetic relationships between different plants. Recently, *C. sativa* strains have been classified based on genomic DNA markers using random amplified polymorphic DNA (RAPD) [5–8], amplified fragment length polymorphisms (AFLP) [9], and inter-simple sequence repeat amplification (ISSR) [10]. Genomic DNA sequence analysis has been performed for the internal transcribed spacers I and II (ITS I, II) of the nuclear ribosomal DNA [11,12], the 5S-rRNA gene spacer region [13], the intergenic spacer region of the chloroplast DNA (*trnL-trnF* IGS) [14], and the *trnL* intron [15]. Recently, short tandem repeat (STR, microsatellite) markers have been developed as a powerful tool to distinguish and study genetic variations in cannabis samples [16–18].

The biosynthetic pathway that produces Δ-9-tetrahydrocannabinolic acid (THCA) has been studied (Fig. 1) [19]. Recently, the gene encoding THCA synthase (from cannabigerolic acid (CBGA) to THCA) has been identified by Sirikantaramas et al. [20]. Cannabinoids are present as cannabinoid acids (e.g. THCA and CBDA) in plant tissues, although the major narcotic chemical forms of cannabinoids are neutral forms (e.g. tetrahydrocannabinol, THC). For convenience, the cannabinoid acid forms (THCA and CBDA) are referred to here. In the present study, we used HPLC to analyze variations in the cannabinoid content of 13 *C. sativa* strains. The gene for THCA synthase was amplified by PCR and the products were sequenced to examine polymorphisms in the gene.

## 2. Materials and methods

### 2.1. Plant materials

We analyzed 13 different strains of *C. sativa*: #001, #005, #009, #010, #011, #013, #020, #045, #053, #054, #066, #068 and #078 (Table 1). One plant of each strain was used. All plants used in this study were grown in an incubation room (25 °C, constant fluorescent light) to avoid the effects of environmental differences on plant growth and cannabinoid content, as described previously [10,14].

### 2.2. Cannabinoid analysis by HPLC

Quantitative analysis of cannabinoids (THCA and CBDA) in mature leaves was performed by HPLC, as previously described [10].

### 2.3. DNA extraction

Total genomic DNA was extracted from 150 mg of fresh leaves from 40-day-old plants using a modified cetyltrimethylammonium bromide (CTAB) method [14]. Purified DNA was adjusted to a final concentration of 0.5 ng/μl with sterile water.

### 2.4. PCR amplification and detection

A full-length coding region of the THCA synthase gene was amplified by PCR with primers ‘‘a’’ and ‘‘b’’ that were designed in the present study based on the report previously described [20] (primer a, 5′-TGA AGA AAA AAA ATG AAT TGCTCA GCA TTT TCC-3′; primer b, 5′-TCT ATT TAA AGA TAA TTA ATG ATG ATG CGG TGG-3′; Figs. 2 and 5). Amplification was performed in a total volume of

Table 1
List of plant materials used

| Experimental code | Country | Institute or arboretum | Accession number[a] |
|---|---|---|---|
| #001 | Czech | Jaromir Svoboda Orechova | AB212829 |
| #005 | Hungary | Agratudomanyi Egyetem Botanikus Godollo | AB212830 |
| #009 | Poland | The Botanical Garden of Lublin | AB212831 |
| #010 | Iran | Medicinal and Aromatic Plants Research Center, University of Shahied Beheshti, Teheran | AB212832 |
| #011 | Germany | Botaniachwe Garten Heinrich-Heine Universitat, Dusseldorf | AB212833 |
| #013 | Mexico | Kyushu University (Japan), from Mexico | AB212834 |
| #020 | Unknown | Kyushu University (Japan) | AB212835 |
| #045 | Japan | Tochigi Prefectural Institute of Public Health and Envirnmental Science, Tochigi | AB212836 |
| #053 | Japan | Kanto-Shin'etsu Regional Narcotic Control Office, Tokyo | AB212837 |
| #054 | Japan | Kanto-Shin'etsu Regional Narcotic Control Office, Tokyo | AB212838 |
| #066 | France | Jardin Botanique E.M., Marseille | AB212839 |
| #068 | Czech | Agricultural University Prague, Prague | AB212840 |
| #078 | Italy | Civico Orto Botanico, Marchesetti | AB212841 |

[a] The nucleotide sequence data accessioned to the DDBJ/EMBL/GenBank databases.

25 μl, with each reaction containing 200 μM dNTPs, 0.3 μM of primers, 1.25 U of *Pfu Turbo* DNA polymerase (Stratagene), 10× *Pfu Turbo* reaction buffer, and 5 ng of template DNA. *Pfu Turbo* DNA polymerase has 3′- to 5′-exonuclease proofreading activity to avoid misamplification in PCR. The PCR conditions were: preheating at 95 °C for 2 min, 30 cycles at 95 °C for 30 s, 50 °C for 30 s, and 72 °C for 2 min with a final extension at 72 °C for 10 min. AmpliWax™ (TaKaRa) was used according to the manufacturer's instructions to avoid misannealing of primers at the lower temperatures. This system allows annealing of the template DNA and primers at high temperatures (>55 °C). Following amplification of the PCR fragment with *Pfu Turbo* DNA polymerase, 0.5 U of *Tag* DNA polymerase (Roche) was added to the reactions. Then the reactions were incubated at 72 °C for 10 min, which allowed the addition of the 3′ "A" nucleotide overhang necessary for the subsequent cloning procedure. The reactions were carried out in a TaKaRa PCR Thermal Cycle MP (TaKaRa). The amplified products were electrophoresed on 2% agarose gels prepared with Tris-acetate-EDTA (TAE) buffer (pH 8.0). After staining with ethidium bromide, the amplified products were photographed under UV light (254 nm).

### 2.5. Sequence analysis

The amplified PCR products were purified using a QIAquick PCR Purification Kit (QIAGEN). The purified DNA fragments were fused into the pCR II-TOPO vector of the Original TA Cloning Kit (Invitrogen), and plasmid DNA was isolated using a QIAprep Spin Miniprep Kit (QIAGEN). Sequencing was carried out using the BigDye Terminator Cycle Sequencing FS Ready Reaction Kit (Applied Biosystems) using ABI Prism 310, ABI Prism 373A, or ABI Prism 3100 automated DNA sequencing systems (Applied Biosystems). Primers for sequencing used primers "a", "b", "c", "d", "e" and "f" (primer c, 5′-CAA ACT GGT TGC TGT CCC ATC-3′; primer d, 5′-AGC TGG GAA GAA GAC GGC TTT CTC A-3′; primer e, 5′-CGT CTT CTT CCC AGC TGA TCT-3′; and primer f, 5′-CGC CAA CAG TAG GGC AAT ACC-3′, designed in the present study; Figs. 2 and 5). Sequence data were aligned using the computer program GENETYX-MAC ver. 11 (Software Development). All PCR fragments were cloned and sequenced at least three times.

### 2.6. Detection of "drug-type" PCR marker

A portion of the THCA synthase gene was amplified by PCR with primers "g" and "h" that were designed in the present study (primer g, 5′-AAT AAC TCC CAT ATC CAA GCA-3′; primer h, 5′-AGG ACT CGC ATG ATT AGT TT-3′; Figs. 2 and 5). PCR was carried out in a total volume of 10 μl using *AmpUTaq Gold* DNA polymerase (Applied Biosystems). PCR conditions were: one cycle at 94 °C for 10 min, followed by 35 cycles at 94 °C for 1 min, 57 °C for 1 min, and 72 °C for 2 min, with a final extension step at 72 °C for 5 min. The amplified products were photographed as above. The ribulose-bisphosphate carboxylase gene (*rbcL*; forward primer, 5′-ATG TCA CCA CAA ACA GAG ACT AAA GC-3′; reverse primer, 5′-GCA GCA GCT AGT TCC GGG CTCCA-3′) [21] was also amplified for the positive control.

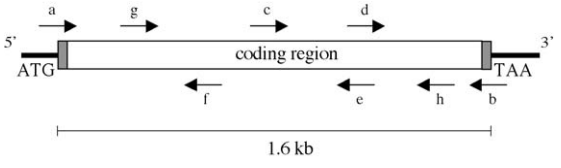


Fig. 2. Schematic representation of the THCA synthase gene. Primers (a–h) are indicated by arrows. ATG, start codon; TAA, stop codon.

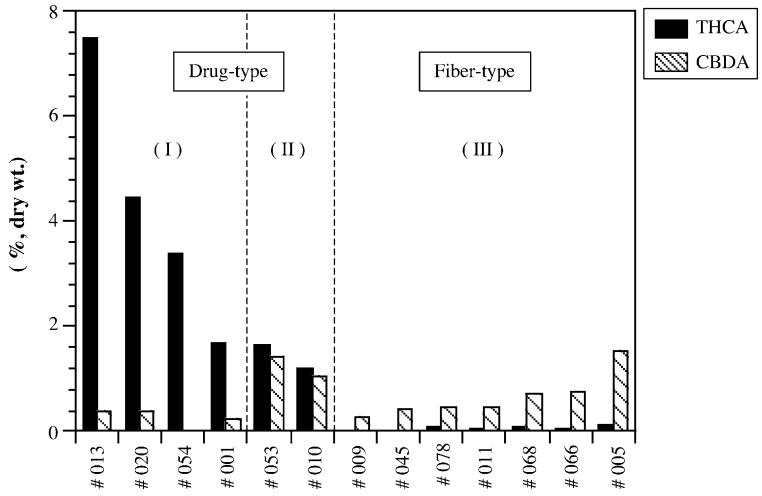


Fig. 3. THCA and CBDA content in the leaves of 13 strains of *Cannabis sativa* L. I–III represents chemo-types.

## 3. Results and discussion

### 3.1. Cannabinoid analysis by HPLC

The THCA and CBDA content in mature leaves of 13 *C. saliva* strains was determined. These strains were separated into two groups, based on the THCA content (Fig. 3). High THCA content (1.19–7.51% dry wt.) was detected in six strains (#013, #020, #054, #001, #053 and #010). These six strains were classified as ''drug-type''. However, low (or not detected) THCA content (0–0.12%) was detected in seven strains (#009, #045, #078, #011, #068, #066 and #005). These seven strains were classified as ''fiber-type''.

Small and Cronquist [2] described three chemical phenotypes (chemo-types I–III), based on the THCA and CBDA content. In the present study, four strains of ''drug-type'', #013, #020, #054 and #001, were recognized chemo-type I with high THCA and low CBDA content. Two ''drug-type'' strains, #053 and #010, were found to be chemo-type II, with high THCA and CBDA. The seven ''fiber-type'' strains were all found to be chemo-type III, with particularly low THCA content.

### 3.2. Sequence of THCA synthase gene

The sequence of CBDA synthase gene [22] is very similar to the one of THCA synthase gene (homology 87.9%). So, we designed primers ''a'' and ''b'' to eliminate producing the CBDA synthase gene fragment (Fig. 5). DNA fragments (approximately 1.6 kb) were obtained from all 13 samples using primers ''a'' and ''b'' (Fig. 4). The nucleotide sequences of these PCR products were successfully determined (Table 1, DDBJ/EMBL/GenBank database accession numbers AB212829, 212830, 212831, 212832, 212833, 212834, 212835, 212836, 212837, 212838, 212839, 212840 and 212841). There were 1638 bp between the 5′-end (the ATG sequence, start codon) and the 3′-end (the TAA sequence, stop codon) in all samples. There was no intron among the sequences obtained. An alignment of the genomic DNA sequences is shown in Table 2. The sequences corre-

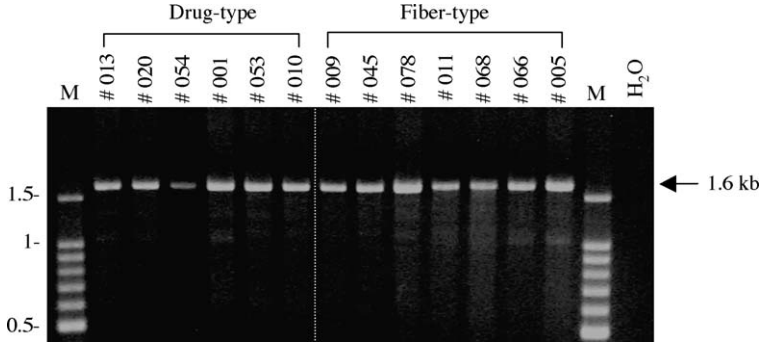


Fig. 4. Detection of PCR products of the THCA synthase gene from 13 strains of *Cannabis sativa* L. Full length of the coding region used primers ''a'' and ''b''. M, 100 bp ladder molecular marker.

Table 2
Aligned sequences of polymorphic sites for the THCA synthase gene of *Cannabis sativa* L.

| | | Position (bp) | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Strain | Type | 4 | 6 | 7 | 9 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 2 | 2 | 2 | 2 | 2 | 2 | 2 | 2 | 2 | 3 | 3 | 3 | 3 | 3 | 3 | 3 | 4 | 4 | 4 | 4 | 4 | 5 | 6 |
| | | 6 | 4 | 9 | 2 | 1 | 1 | 2 | 2 | 3 | 3 | 5 | 5 | 8 | 9 | 2 | 2 | 6 | 6 | 8 | 8 | 8 | 9 | 9 | 0 | 1 | 5 | 6 | 8 | 8 | 9 | 0 | 1 | 1 | 2 | 9 | 0 | 1 |
| | | | | | | 7 | 8 | 1 | 9 | 6 | 7 | 1 | 4 | 7 | 8 | 1 | 9 | 8 | 9 | 2 | 5 | 7 | 2 | 4 | 0 | 2 | 5 | 6 | 3 | 5 | 9 | 9 | 2 | 8 | 4 | 4 | 5 | 2 |
| #013 | Drug-type | A | C | A | G | A | A | C | C | G | T | C | G | A | G | T | A | A | A | A | A | C | T | A | T | C | A | A | T | G | A | T | A | A | G | A | C | T |
| #020 | | * | * | * | * | * | * | * | * | * | * | * | * | * | * | * | * | * | * | * | * | * | * | * | * | * | * | * | * | * | * | * | * | * | * | * | * | * |
| #054 | | * | * | * | * | * | * | * | * | * | * | * | * | * | * | * | * | * | * | * | * | * | * | * | * | * | * | * | * | * | * | * | * | * | * | * | * | * |
| #001 | | * | * | * | * | * | * | * | * | * | * | * | * | C | * | * | * | * | * | * | * | * | * | * | * | * | * | T | * | * | G | * | * | * | * | * | * | * |
| #010 | | * | * | * | * | * | * | * | * | * | * | * | * | * | * | * | * | * | * | * | * | * | * | * | * | * | * | * | * | * | * | * | * | * | * | * | * | * |
| #053 | | * | * | * | * | * | * | * | * | * | * | * | * | * | * | * | * | * | * | * | * | * | * | * | * | * | * | * | * | * | * | * | * | * | * | * | * | * |
| #009 | Fiber-type | * | A | * | A | G | G | T | T | C | C | T | A | G | A | C | * | G | T | G | C | G | C | C | C | T | T | * | C | A | * | A | G | G | A | T | T | C |
| #045 | | * | A | T | A | G | G | T | T | C | C | T | A | G | * | C | * | G | T | G | C | G | C | C | C | T | T | * | C | A | * | A | G | G | A | T | T | C |
| #078 | | * | A | T | A | G | G | T | T | C | C | T | A | G | * | C | * | G | T | G | C | G | C | C | C | T | T | * | C | A | * | A | G | G | A | T | T | C |
| #011 | | * | A | * | A | G | G | T | T | C | C | T | A | G | * | C | * | G | T | G | C | G | C | C | C | T | T | * | C | A | * | A | G | G | A | T | T | C |
| #068 | | * | A | * | A | G | G | T | T | C | C | T | A | G | * | C | G | G | T | G | C | G | C | C | C | T | T | * | C | A | * | A | G | G | A | T | T | C |
| #066 | | * | A | * | A | G | G | T | T | C | C | T | A | G | * | C | * | G | T | G | C | G | C | C | C | T | T | * | C | A | * | A | G | G | A | T | T | C |
| #005 | | T | A | * | A | G | G | T | T | C | C | T | A | G | * | C | * | G | T | G | C | G | C | C | C | T | T | * | C | A | * | A | G | G | A | T | T | C |

| | | Position (bp) | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Strain | Type | 6 | 6 | 6 | 7 | 7 | 7 | 8 | 8 | 8 | 8 | 8 | 8 | 8 | 9 | 9 | 9 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 |
| | | 1 | 7 | 9 | 4 | 4 | 6 | 0 | 6 | 6 | 6 | 8 | 8 | 8 | 4 | 5 | 7 | 0 | 0 | 0 | 0 | 1 | 1 | 1 | 2 | 2 | 2 | 3 | 3 | 3 | 4 | 4 | 4 | 4 | 4 | 5 | 5 | 5 |
| | | 9 | 8 | 9 | 4 | 9 | 3 | 0 | 2 | 4 | 9 | 1 | 5 | 7 | 5 | 3 | 5 | 3 | 7 | 7 | 8 | 7 | 7 | 7 | 2 | 3 | 7 | 4 | 6 | 9 | 0 | 0 | 1 | 8 | 8 | 6 | 6 | 7 |
| | | | | | | | | | | | | | | | | | | 5 | 4 | 9 | 0 | 2 | 6 | 9 | 4 | 2 | 2 | 2 | 5 | 8 | 4 | 7 | 0 | 1 | 4 | 0 | 9 | 4 |
| #013 | Drug-type | G | G | T | G | C | T | A | G | A | T | T | A | A | A | A | C | T | A | T | T | C | T | A | T | C | G | T | T | T | T | G | T | A | C | G | T | T |
| #020 | | * | * | * | * | * | * | * | * | * | * | * | * | * | * | * | * | * | * | * | * | * | * | * | * | * | * | * | * | * | * | * | * | * | * | * | * | * |
| #054 | | * | * | * | * | * | * | * | * | * | * | * | * | * | * | * | * | * | * | * | * | * | * | * | * | * | * | * | * | * | * | * | * | * | * | * | * | * |
| #001 | | * | * | * | * | * | * | * | * | * | * | * | * | * | * | * | * | * | * | * | * | * | * | T | * | * | * | * | * | * | * | * | * | * | * | * | * | * |
| #010 | | * | * | * | * | * | * | * | * | * | * | * | * | * | * | * | * | * | * | * | * | * | * | * | * | * | * | * | * | * | * | * | * | * | * | * | * | * |
| #053 | | * | * | * | * | * | * | * | * | * | * | * | * | * | * | * | * | * | * | * | * | * | * | * | * | * | * | * | * | * | * | * | * | * | * | * | * | * |
| #009 | Fiber-type | * | A | A | T | T | G | * | A | G | C | G | T | G | C | T | T | G | A | A | C | T | A | T | G | T | T | A | C | * | C | A | * | C | A | * | C | C |
| #045 | | * | A | A | T | T | G | G | A | G | C | G | T | G | C | T | T | G | A | A | C | T | A | T | G | T | T | A | C | C | C | A | * | C | A | * | C | C |
| #078 | | * | A | A | T | T | G | * | A | G | C | G | T | G | C | T | T | G | A | A | C | T | A | T | G | T | T | A | C | * | C | A | G | C | A | * | C | C |
| #011 | | T | A | A | T | T | G | * | A | G | C | G | T | G | C | T | T | G | C | A | C | T | A | T | G | T | T | A | C | * | C | A | * | C | A | A | C | C |
| #068 | | * | A | A | T | T | G | * | A | G | C | G | T | G | C | T | T | G | A | A | C | T | A | T | G | T | T | A | C | * | C | A | * | C | A | * | C | C |
| #066 | | * | A | A | T | T | G | * | A | G | C | G | T | G | C | T | T | G | A | A | C | T | A | T | G | T | T | A | C | * | C | A | * | C | A | * | C | C |
| #005 | | * | A | A | T | T | G | * | A | G | C | G | T | G | C | T | T | G | A | A | C | T | A | T | G | T | T | A | C | * | C | A | * | C | A | * | C | C |

Identical nucleotides for strain #013 are shown in ‘‘*’’. Nucleotide substitutions for strain #013 strain are shown in gray.

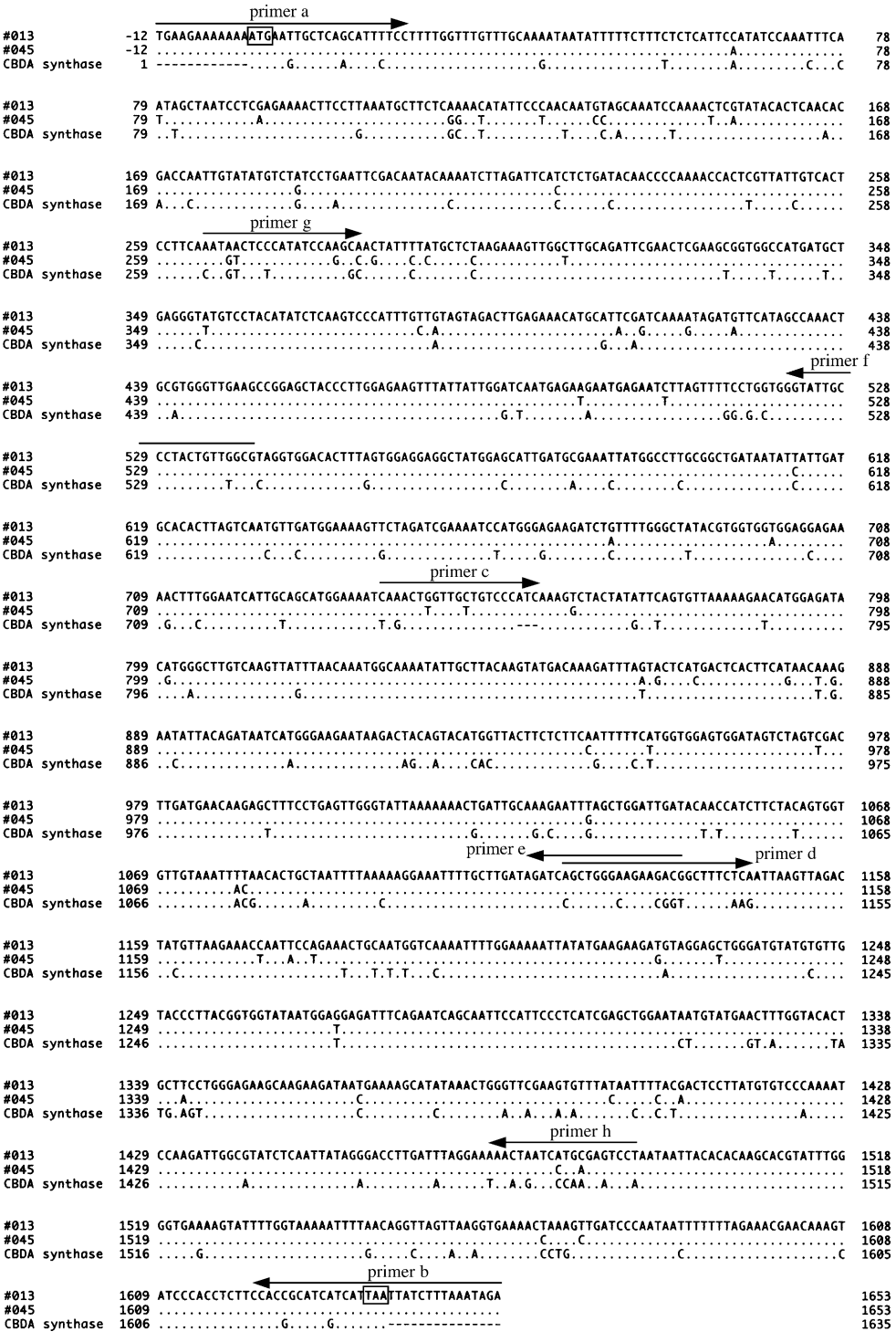


Fig. 5. Nucleotide sequences of the THCA synthase gene of *Cannabis sativa* L. "#013", "drug-type" strain; "#045", "drug-type" strain. "CBDA synthase" is the sequence of the CBDA synthase gene. Identical nucleotides for strain #013 are shown in "*". Primers used in the present study are indicated by arrows.

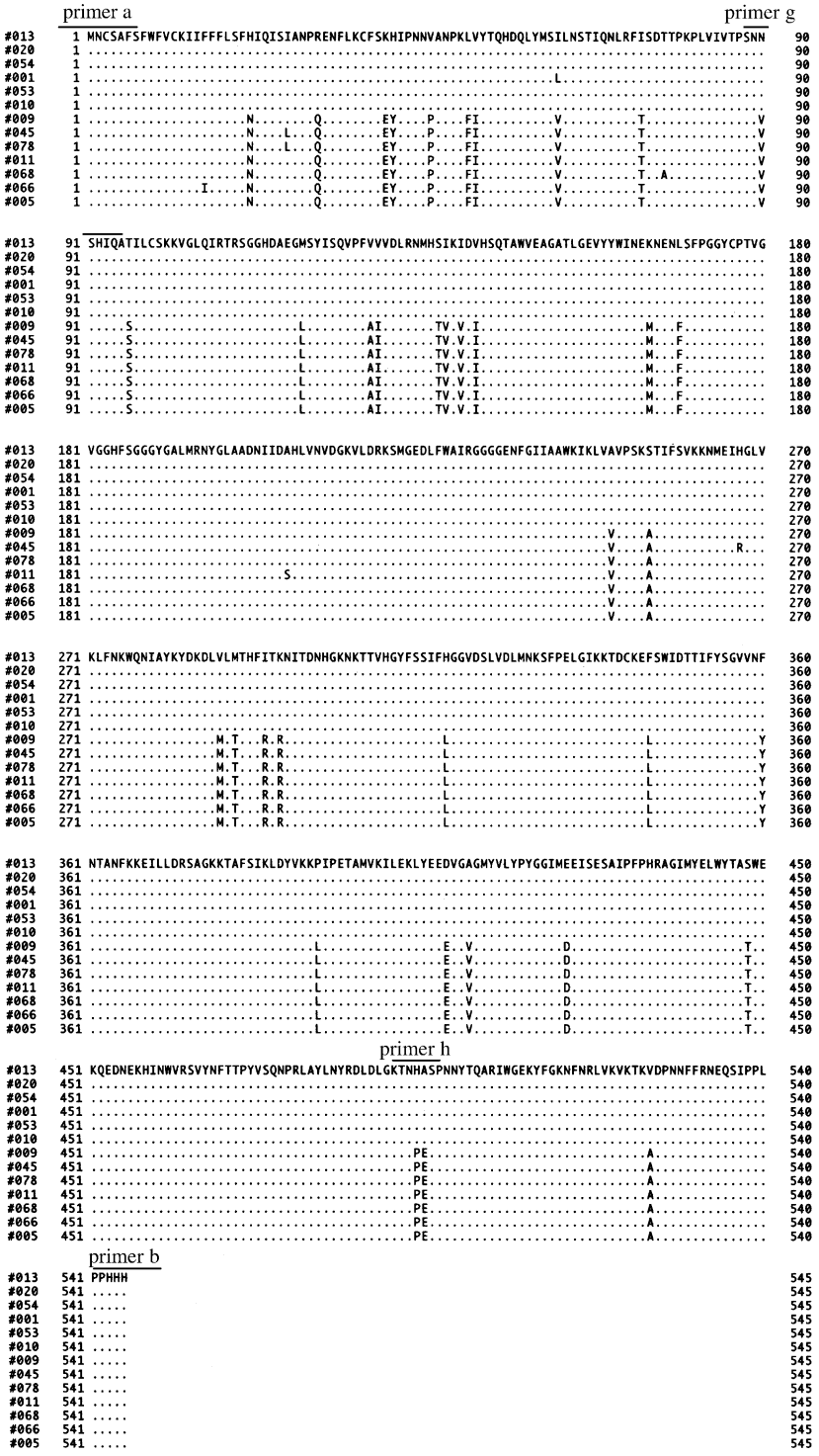


Fig. 6. Amino acid sequence alignment of the THCA synthase gene from 13 strains of *Cannabis sativa* L. Primer "a", "b", "g" and "h" are indicated lines.

sponding to the THCA synthase gene were separated into two groups, "drug-type" and "fiber-type". The 62 nucleotides from the six "drug-type" strains (#013, #020, #054, #001, #053 and #010) were substituted with different nucleotides from the seven "fiber-type" strains (#009, #045, #078, #011, #068, #066 and #005); the nucleotide substitutions for strain #013 are shown in gray in Table 2. The sequences of strain #013 ("drug-type") and #045 ("fiber-type") were shown in Fig. 5. Both of the genomic DNA sequences of the two groups encoded complete amino acid sequences (545 amino acids, Fig. 6), although many substitutions existed in the genomic DNA. A total of 37 major amino acid substitutions were detected in the alignment of the sequences from "drug-type" and "fiber-type" strains. It seemed that the amino acid substitutions induced a decrease in THCA synthase activity in the "fiber-type" cannabis plants. In the present study, we clarified that there were two variants of the THCA synthase gene in the "drug-type" and "fiber-type" cannabis plants.

Minor substitutions were also detected in the sequences of the genomic DNA. In the "drug-type" groups, four nucleotide substitutions (positions 187, 366, 399 and 1179) were detected in strain #001. Three nucleotides (366, 399 and 1179) in strain #001 encoded the same amino acids (isoleucine, arginine and proline) as those in the other "drug-type" strains. The substitution at 187 bp was non-synonymous: the substituted nucleotide C coded for leucine in #001 and for isoleucine in another "drug-type" strain. However, these two amino acids have the same chemical character, and therefore it seems that the substitutions at position 187 bp should not seriously affect the activity of THCA synthase. Furthermore, we analyzed more three different individuals of strain #013 (#013-2, THCA-4.04%, CBDA-not detected; #013-3, T-4.05%, C-not detected; #013-4, T-5.08%, C-not detected). The sequences of these three samples were same as to the one of #013-1, which was shown above as "#013". In the "fiber-type" group, nine nucleotide substitutions were detected. The nucleotide T at position 46 bp in #005 codes for phenylalanine, which was same amino acid in the other strains. Likewise, substitution of A at position 198 bp in #009, C at position 1398 bp in #045, G at position 1410 bp in #078, and A at position 1560 bp in #011 were all synonymous substitutions, so amino acid substitutions did not occur. The substituted nucleotides T at 79 bp in #045 and #078, G at 229 bp in #068, T at 619 bp in #011, and G at 800 bp in #045, encoded different amino acids from the other strains.

### 3.3. Detection of "drug-type" PCR marker

It is difficult to distinguish between "drug-type" and "fiber-type" plants from morphology and chemical analysis in the early vegetative stage. To simplify the detection of "drug-type" cannabis plants, we tried to identify a PCR marker. We could detect a specific PCR fragment of the THCA synthase gene using *Taq* DNA polymerase in the "drug-type" strains (#013, #020, #054, #001, #053 and #010) using primers "g" and "h" (Fig. 7). These primers were also designed to eliminate producing the CBDA synthase gene fragment (Fig. 5). Using these primers, no PCR fragment was detected from the "fiber-type" strains, although fragment of *rbcL* was detected from all strains. The sequences of the "drug-type" specific PCR fragments were the same as described using primers "a" and "b" (data not shown).

It is very important to remove high-THCA plant in the cultivation of cannabis for fiver production. Thus, THCA check of the samples from the cannabis field is necessary. But young vegetative stage "drug-type" cannabis plant contains no or very low content of THCA. Therefore, the detection of THCA using chemical analysis is not easy. It would be possible to check "drug-type" plant easily and rapidly in early vegetative stage using the PCR marker. Moreover, in the forensic situation, the PCR marker would provide useful information.

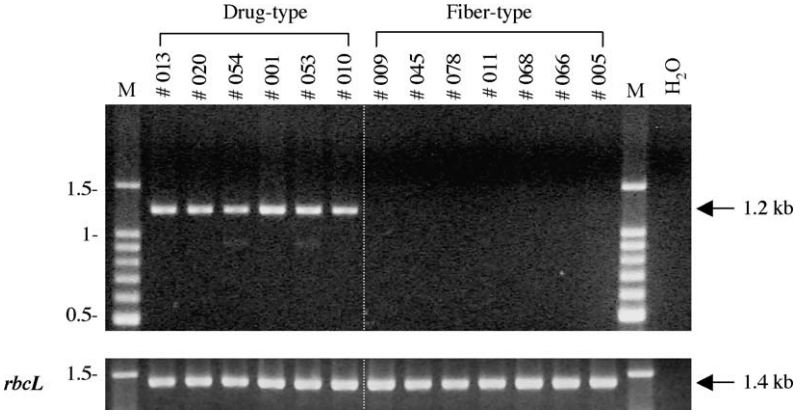


Fig. 7. Detection of the "drug-type" specific PCR marker. A portion of the coding region of the THCA synthase gene was amplified using primers "g" and "h" in "drug-type" strains, #013, #020, #054, #001, #053 and #010 of *Cannabis sativa* L. M, 100 bp ladder molecular marker; *rbcL*, positive control.

## 4. Conclusion

In this study, we identified the difference between ''drug-type'' and ''fiber-type'' cannabis plants by the THCA synthase gene analysis. The PCR fragments of the THCA synthase gene were detected from six strains of ''drug-type'' plants. We could also detect the PCR fragments from seven strains of ''fiber-type'' plants, although no or very low content of THCA were detected in these samples using HPLC analysis. The sequences of the THCA synthase gene were separated into two groups, ''drug-type'' and ''fiber-type''. There were 37 major substitutions of the deduced amino acid sequences between ''drug-type'' and ''fiber-type''. It seemed that the substitutions of the amino acid induced the change of enzymatic activity of the THCA synthase in ''fiber-type''. We also identified a specific PCR marker in ''drug-type'' cannabis plants. This PCR marker would be useful for identifying cannabis types in the fields of forensic science and crop breeding.

## Acknowledgements

The authors would like to thank O. Iida (National Institute of Health Sciences, Japan), S. Sekita (Tokushima Bunri University), M. Satake (Ocyanomizu University), and Y. Makino (Kanto-Shin'etsu Regional Bureau of Health and Welfare Sciences, Japan) for assistance with this study. This research was supported in part by a Health Sciences Research Grant from the Ministry of Health, Labour and Welfare of Japan.

## References

[1] M.W. Quimby, N.J. Doorenbos, C.E. Turner, A. Masoud, Mississippi-grown marihuana-*Cannabis sativa* cultivation and observed morphological variations, Econ. Bot. 27 (1973) 117–127.

[2] E. Small, A. Cronquist, A practical and natural taxonomy for Cannabis, Taxon 25 (1976) 405–435.

[3] M.G. Rowan, J.W. Fairbairn, Cannabinoid patterns in seedling of *Cannabis sativa* L. and their use in the determination of chemical race, J. Pharm. Pharmacol. 29 (1977) 491–494.

[4] P.B. Baker, T.A. Gough, B.J. Taylor, The physical and chemical features of cannabis plants grown in the United Kingdom of Great Britain and Northern Ireland from seeds of known origin, Bull. Narcotics 34 (1982) 27–36.

[5] R. Gillan, M.D. Cole, A. Linacre, J.W. Thorpe, N.D. Watson, Comparison of *Cannabis sativa* by random amplification of polymorphic DNA (RAPD) and HPLC of cannabinoid: a preliminary study, Sci. Justice 35 (1995) 169–177.

[6] V. Jagadish, J. Robertson, A. Gibbs, RAPD analysis distinguished *Cannabis sativa* samples from different sources, Foren. Sci. Int. 79 (1996) 113–121.

[7] G.S. Gigliano, A.D. Finizio, P. Caputo, S. Cozzolino, Cannabis fingerprints by using random amplified polymorphic DNA (RAPD), Delpinoa 37 (1995-1996) 35–47.

[8] E.P.M. de Meijer, M. Bagatta, A. Carboni, P. Crucitti, V.M.C. Moliterni, P. Ranalli, G. Mandolino, The inheritance of chemical phenotype in *Cannabis sativa* L., Genetics 163 (2003) 335–346.

[9] C.H. Miller, G. Shutler, S. Abrams, J. Hanniman, S. Neylon, C. Ladd, T. Palmbach, H.C. Lee, A simple DNA extraction method for marijuana samples used in amplified fragment length polymorphism (AFLP) analysis, J. Foren. Sci. 48 (2003) 343–347.

[10] M. Kojoma, O. Iida, Y. Makino, S. Sekita, M. Satake, DNA fingerprinting of *Cannabis sativa* using inter-simple sequence repeat (ISSR) amplification, Planta Med. 68 (2002) 60–63.

[11] G.S. Gigliano, Preliminary data on the usefulness of internal transcribed spacer 1 (ITS1) sequence in *Cannabis sativa* L. identification, J. Foren. Sci. 44 (1999) 457–475.

[12] G.S. Gigliano, P. Caputo, Ribosomal DNA analysis as a tool for the identification of *Cannabis sativa* L. specimens of forensic interest, Sci. Justice 37 (1997) 171–174.

[13] M. Miyahara, K. Sugaya, A. Tanimura, M. Satake, Nucleotide sequences of 5S-rRNA gene spacer region from Moraceae and Cannabaceae, Nat. Med. 52 (1998) 209–212.

[14] M. Kohjyouma, I. Lee, O. Iida, K. Kurihara, K. Yamada, Y. Makino, S. Sekita, M. Satake, Intraspecific variation in *Cannabis sativa* L. based on intergenic spacer region of chloroplast DNA, Biol. Pharm. Bull. 23 (2000) 727–730.

[15] G.S. Gigliano, Restriction profiles of *trnL* (UAA) intron as a tool in *Cannabis sativa* L. identification, Delpinoa 37 (1995–1996), pp. 37–38, 85–95.

[16] H. Hsieh, R. Hou, L. Tsai, C. Wei, S. Liu, L. Huang, Y. Kuo, A. Linacre, J. Lee, A highly polymorphic STR locus in *Cannabis sativa*, Foren. Sci. Int. 131 (2003) 53–58.

[17] S.R. Gilmore, J. Robertson, Short tandem repeat (STR) DNA markers are hypervariable and informative in *Cannabis sativa*: implications for forensic investigations, Foren. Sci. Int. 131 (2003) 65–74.

[18] H.J. Alghanim, J.R. Almirall, Development of micro satellite markers in *Cannabis sativa* for DNA typing and genetic relatedness analyses, Anal. Bioanal. Chem. 376 (2003) 1225–1233.

[19] F. Taura, S. Morimoto, Y. Shoyama, First direct evidence for the mechanism of Δ-1-tetrahydrocannabinolic acid biosynthesis, J. Am. Chem. Soc. 117 (1995) 9766–9767.

[20] S. Sirikantaramas, S. Morimoto, Y. Shoyama, Y. Ishikawa, Y. Wada, Y. Shoyama, F. Taura, Molecular cloning and heterologous expression of $\Delta^1$-tetrahydrocannabinolic acid synthase from *Cannabis sativa* L., J. Biol. Chem. 279 (2004) 39767–39774.

[21] M. Hasebe, T. Omori, M. Nakazawa, T. Sano, M. Kato, K. Iwatsuki, *rbcL* gene sequences provide evidence for the evolutionary lineages of leptosporangiate ferns, in: Proceedings of the National Academy of Sciences of the United States of America, vol. 91, 1994, pp. 5730–5734.

[22] K. Yoshikai, T. Taura, S. Morimoto, Y. Shoyama, DNA encoding cannabidiolate synthase, Patent in Japan JP 2000-78979A, 2001, DDBJ/EMBL/GenBank database accession numbers E55107.