2015

Characterisation of cannabinoid composition in a diverse Cannabis sativa L. germplasm collection

Matthew T. Welling
Southern Cross University

Lei Liu
Southern Cross University

Tim Shapter
Southern Cross University

Carolyn A. Raymond
Southern Cross University

Graham J. King
Southern Cross University

Publication details

Postprint of: Welling, MT, Liu, L, Shapter, T, Raymond, CA & King, GJ 2016, 'Characterisation of cannabinoid composition in a diverse Cannabis sativa L. germplasm collection', *Euphytica,* vol. 208. no. 3, pp. 463-475.
Published version available from:
http://dx.doi.org/10.1007/s10681-015-1585-y

Euphytica

Characterisation of cannabinoid composition in a diverse Cannabis sativa L. germplasm collection

--Manuscript Draft--

Manuscript Number:	EUPH-D-15-00581R1
Full Title:	Characterisation of cannabinoid composition in a diverse Cannabis sativa L. germplasm collection
Article Type:	Original Article
Keywords:	Cannabinoids; marijuana; hemp; chemotype; Genetic diversity; LC-MS; DNA genotype
Corresponding Author:	graham John king, Ph.D. Southern Cross University Lismore, AUSTRALIA
Corresponding Author Secondary Information:	
Corresponding Author's Institution:	Southern Cross University
Corresponding Author's Secondary Institution:	
First Author:	Matthew Welling
First Author Secondary Information:	
Order of Authors:	Matthew Welling
	Lei Liu
	Tim Shapter
	Carolyn Raymond
	Graham John King
Order of Authors Secondary Information:	
Funding Information:	Ecofibre Industries Operations — Prof Graham John King
Abstract:	The ability to characterise cannabinoid chemical phenotype (chemotype) accurately is important for the development of Cannabis sativa L. cultivars specific for pharmacological, hemp fibre, or seed end use. Although a number of chemotyping and genotyping methods have previously been developed to predict and characterise cannabinoid composition, only a subset of the gene pool has been examined. A representative survey from a wide range of geographically and genetically diverse C. sativa accessions using liquid chromatography-mass spectrometry (LC-MS) cannabinoid profiling together with dominant and co-dominant DNA marker assays was performed. Overall variability of chemotype across the gene pool was found to be three-fold greater within heterozygote genotypes than previously reported. Interestingly, an individual plant of East Asian origin was found to exhibit a rare propyl alkyl cannabinoid homologue and a chemotype inconsistent with the predicted genotype. We propose that in order to carry out comprehensive screening of genetic resource collections and to identify chemotypic variants specific for end-use pharmacological applications, a strategy which adopts both cannabinoid profiling and the co-dominant DNA marker assay is required. Further research with consideration of propyl alkyl cannabinoid homologues should explore the relationship between chemotype and genotype in greater detail.
Response to Reviewers:	Response to Authors Comments 1. The results of chemical and genetic analysis are not presented. I suggest these should be included either as a supplementary table or as a table in the paper. The cannabinoid composition, the chemotype, and the D589 and B1080/B1192 DNA

	marker genotype score for each plant individual has been recorded in a supplementary table entitled “Online Resource 3” (ESM_3.pdf). Reference to “Online Resource 3” can be found within the main text of the manuscript at line numbers L207/L225. 2. Typo: In figure 3 there are three results presented for each accession with the exception of D (2 results) and O (4 results). I suspect this is a typo. This typo has been corrected and a revised Figure 3 has been uploaded for submission. 3. There may be other ways of presenting the chemometric data. The authors could consider principle component analysis. It would be interesting to see if the data fell into clusters and whether the genotype data or geographic data mapped on to any observed groupings. We have considered principle component analysis. However, there are an insufficient number of independent variables within the data set to support this method of analysis.

Characterisation of cannabinoid composition in a diverse *Cannabis sativa* L. germplasm collection

Matthew T. Welling • Lei Liu • Tim Shapter • Carolyn A. Raymond • Graham J. King

M. T. Welling • L. Liu • T. Shapter • C. A. Raymond • G. J. King (corresponding author)

Southern Cross Plant Science, Southern Cross University, T Block, Lismore, PO Box 157, New South Wales 2480, Australia.

e-mail: graham.king@scu.edu.au

t: +61 2 6620 3410

M. T. Welling • T. Shapter

Ecofibre Industries Operations Pty Ltd, 8 Thynne Court, Maleny, Queensland, 4552, Australia.

Acknowledgments

The authors wish to thank Ecofibre Industries Operations Pty Ltd for allowing us access to the Ecofibre global *Cannabis sativa* L. germplasm collection and for generously providing financial support.

Abstract

The ability to characterise cannabinoid chemical phenotype (chemotype) accurately is important for the development of *Cannabis sativa* L. cultivars specific for pharmacological, hemp fibre, or seed end use. Although a number of chemotyping and genotyping methods have previously been developed to predict and characterise cannabinoid composition, only a subset of the gene pool has been examined. A representative survey from a wide range of geographically and genetically diverse *C. sativa* accessions using liquid chromatography-mass spectrometry (LC-MS) cannabinoid profiling together with dominant and co-dominant DNA marker assays was performed. Overall variability of chemotype across the gene pool was found to be three-fold greater within heterozygote genotypes than previously reported. Interestingly, an individual plant of East Asian origin was found to exhibit a rare propyl alkyl cannabinoid homologue and a chemotype inconsistent with the predicted genotype. We propose that in order to carry out comprehensive screening of genetic resource collections and to identify chemotypic variants specific for end-use pharmacological applications, a strategy which adopts both cannabinoid profiling and the co-dominant DNA marker assay is required. Further research with consideration of propyl alkyl cannabinoid homologues should explore the relationship between chemotype and genotype in greater detail.

Keywords

Cannabinoids • *Cannabis sativa* L. • chemotype • genetic diversity • LC-MS • marker genotype

Introduction

Cannabis sativa L. is an erect, diploid, mostly dieocious (Van Bakel et al. 2011), outcrossing (Forapani et al. 2001) annual herb within the Cannabaceae family (Small and Cronquist 1976). The species is characterised by the production of a large range of biologically active secondary plant metabolites (ElSohly and Slade 2005; Gertsch et al. 2010; Werz et al. 2014), with a subset of over 70 terpenophenolic phytocannabinoid (cannabinoid) compounds (ElSohly and Slade 2005), some of which appear unique to *C. sativa* (Appendino et al. 2011; Gertsch et al. 2010). Cannabinoids are synthesised in plants in their carboxylic acid forms (Swift et al. 2013) and accumulate principally within glandular trichomes occurring on female inflorescences (Happyana et al.

2013). These form neutral cannabinoids, in a non-enzymatic thermal conversion reaction (Dussy et al. 2005), with the most notable conversion being that of delta-9-tetrahydrocannabinolic acid (THCA) to the psychoactive cannabinoid delta-9-tetrahydrocannabinoid (THC).

Cannabinoids have modulating effects on the human endocannabinoid system and are believed to be beneficial in a number of physiopathological processes (Izzo et al. 2009). The pentyl alkyl cannabinoids THCA and cannabidiolic acid (CBDA) are the major constituents found in plants (De Backer et al. 2009; Swift et al. 2013), although a series of propyl alkyl cannabinoid homologues (delta-9-tetrahydrocannabivarinic acid (THCVA) and cannabidivarinic acid (CBDVA)) also occur in plants from specific geographical regions (Baker et al. 1980; Hillig and Mahlberg 2004). Pentyl-cannabinoids are formed from cannabigerolic acid (CBGA) (Taura et al. 2007b), while propyl-cannabinoids are formed from cannabigerovarinic acid (CBGVA) (Shoyama et al. 1984). Oxidocyclization reactions catalysed by THCA and CBDA synthase (Shoyama et al. 1984), form THCA + THCVA (THC(V)A) and CBDA + CBDVA (CBD(V)A) respectively (Sirikantaramas et al. 2004; Taura et al. 2007b) (Fig. 1).

THC(V)A, CBD(V)A and their derivatized forms exert differing actions on the human endocannabinoid system and are considered separate pharmacological entities (Izzo et al. 2009). Three main cannabinoid chemical phenotypes (chemotypes) are recognised based on CBD(V)A and THC(V)A composition (De Meijer et al. 2003); Chemotype I (CBD(V)A: THC(V)A ratio >1), chemotype II (CBD(V)A: THC(V)A ratio close to 1:1), and chemotype III (CBD(V)A: THC(V)A: ratio <1) (De Meijer et al. 1992). Additional chemotypes IV (CBG(V)A: CBD(V)A/THC(V)A >1) and V (total cannabinoid content <0.02%) also occur at low levels within a subgroup of the *C. sativa* gene pool (Pacifico et al. 2006).

The ability to characterise cannabinoid composition is important for the comprehensive screening of germplasm collections, and is especially relevant in planning targeted breeding within obligate outcrossing dioecious species such as *C. sativa* (Forapani et al. 2001). Accurately characterising chemotypic variation offers the potential to identify allelic variation that can contribute to novel cultivars capable of meeting the demands of quality, safety, and efficacy necessary for the manufacture of cannabinoid-based pharmaceutical-grade botanical extracts (Potter 2014), or indeed for eliminating psychoactive constituents from industrial hemp (Mandolino and Carboni 2004). Genetic metabolic engineering through the exploiting of natural occurring allelic and chemotypic diversity within the gene pool (Barker et al. 2007), and subsequent generation of novel recombined breeding lines using marker assisted selection (Mandolino and Carboni 2004; Potter 2014), may lead to the development of standardised multi-targeting botanical drug products from a single plant line. Rigorous

characterisation of chemotype may also aid in the development of fibre hemp and seed cultivars which can maintain levels of THC within legal regulatory thresholds of 0.2% dry weight (w/w) (DW) (Kojoma et al. 2006; Mandolino and Carboni 2004; Pacifico et al. 2006).

A number of chemotyping and genotyping methods have been developed to characterise and predict cannabinoid composition. Gas chromatography is commonly used to characterise chemotype and relies on the thermal conversion of acidic to neutral cannabinoids, although this reaction can vary between laboratories and is only partial (Dussy et al. 2005). Liquid chromatography methods developed more recently can detect both acidic and neutral cannabinoids, therefore providing a more precise characterisation of chemotype (De Backer et al. 2009). Several DNA markers associated with the genes encoding THCA and/or CBDA synthase have been found beneficial in predicting chemotype during early stages of plant development (Kojoma et al. 2006; Pacifico et al. 2006; Rotherham and Harbison 2011; Staginnus et al. 2014), with the most comprehensively studied in terms of genetic linkage and sample population screening being the dominant D589 (Staginnus et al. 2014) and co-dominant B1080/B1192 (Pacifico et al. 2006) DNA sequence characterised amplified region (SCAR) markers respectively.

To date, the available chemotyping and genotyping methods have only been applied to a subset of the *C. sativa* gene pool (Pacifico et al. 2006; Rotherham and Harbison 2011; Staginnus et al. 2014). Given the extensive genetic (Faeti et al. 1996; Gao et al. 2014; Gilmore et al. 2007; Hillig 2005) and chemotypic variability (Baker et al. 1980; Hillig and Mahlberg 2004) which appears to exist, use of each approach in isolation may not be sufficient to account for the full extent of variation in cannabinoid composition within the species. In the present study, we carried out a representative survey from a wide range of geographically and genetically diverse accessions with differing domestication histories. Using a combination of liquid chromatography-mass spectrometry (LC-MS) cannabinoid profiling and DNA SCAR markers, we explored the relationship between chemotype and genotype in order to develop a comprehensive strategy for characterising chemotypic diversity in *C. sativa*.

Materials and methods

Genetic resources

All experimental work, acquisition and storage of the prohibited plant was performed in accordance with Section 23(4)(b) of the Drug Misuse and Trafficking Act 1985 and Section 41(b) of the Drug Misuse and

Trafficking Act 1985, and under the appropriate licences issued by the New South Wales Ministry of Health, Australia. Accessions were sourced from a global germplasm collection owned by Ecofibre Industries Operations (EIO) and managed by Southern Cross University (Table 1). Three genetically diverse *C. sativa* groupings were selected (Hillig 2005), including seven European fibre-type accessions, seven East Asian fibre-type accessions and eight mixed origin drug-type accessions.

Growth parameters

Seeds were planted in seedling trays at a depth of 1 cm in a soil mix comprising of one part vermiculite, one part peat moss, one part perlite and dolomite (110g/100L), and supplemented with CANNA® Aqua Vega nutrient solution. Seeds were grown at temperatures between 26-30°C. 500 mL of water was applied to13 seedling trays (40 cells of 5.5 x 4 cm) three times daily for 14 days and were subject to 11 hours of 600 w HPS/metal halide light (luminous flux equal to 72,000 lumens) per day. Three female progeny from each dioecious accession were selected, while three hermaphrodite plants with mainly female reproductive tissues from monoecious accessions A (TS1300610), B (TS1300026), D (TS1300041) and G (TS1300287) were selected.

Seedlings were transferred into 8 L pots with 1 part vermiculite, 1 part peat moss, 1 part perlite soil mix and dolomite (110 g/100 L). Each 8 L pot contained 100 g Osmocote® Exact nutrient mix and 8 g of Micromax® micronutrient formula. Plants were grown in bespoke pollen secure environmentally controlled closed loop filtered air system growth chambers fitted with 'smart valves' for optimal water regimes. Plants were subject to 11 hours of 600 w HPS/metal halide light (luminous flux equal to 72,000 lumens) per day and kept at a temperature between 28-32°C to promote flowering. Plants were harvested at maturation, which was determined when approximately 95% of the stigma present on the apical inflorescence were browned and shrivelled (Staginnus et al. 2014) and before the onset of seed formation (Pacifico et al. 2006).

DNA marker assays

DNA extraction

Plant leaf material was removed from the apical node of 14 day-old plants. DNA was extracted using a DNeasy® plant mini kit (Qiagen #69104). Tissue disruption was completed manually using a pestle and mortar and ground under liquid nitrogen. Purification of DNA was determined using the Nanodrop 2000 UV-vis

Spectrophotometer (Thermo Scientific). Ratio of absorbance at 260/280 nm was determined for all samples, with ratios of 1.7–1.9 and symmetric peaks at 260 nm used to infer optimal DNA purity.

DNA amplification

Amplification of the D589 marker fragment was achieved using the primers F: 5′ CCTGAATTCGACAATACAAAATCTTAGATTCAT 3′; R: 5` ACTGAATATAGTAGACTTTGATGGGACAGCAACC 3′ (Staginnus et al. 2014). These primers included four SNPs associated with the functional variant of *THCAS* (Kojoma et al. 2006) (B_T allele). Each reaction contained 1.5 mM $MgCl_2$, 0.2 mM dNTPs, 0.2 µM of each of the forward and reverse primers, and 2 U Platinum® Taq DNA Polymerase (Life Technologies #10966-034). Thermocycling parameters used during PCR were 96°C for 2 minutes, then 35 cycles of 94°C for 20 seconds, 64°C for 30 seconds, 72°C for 1minute 30 seconds, followed by a final extension of 72°C for 5 minutes.

Amplification of B1080/B1192 marker fragments were achieved using a primer common to both synthases F: 5′ AAGAAAGTTGGCTTGCAG 3′, and a *THCAS*-specific primer R: 5′ TTAGGACTCGCATGATTAGTTTTTC 3′, and a *CBDAS*-specific primer R: 5′ ATCCAGTTTAGATGCTTTTCGT 3′ (Pacifico et al. 2006). PCR parameters have not been previously reported for the co-dominant DNA marker (Onofri et al. 2015; Pacifico et al. 2006) and these were optimised as follows: Each reaction contained 1.5 mM $MgCl_2$, 0.2 mM dNTPs, 0.4 µM for the forward primer and 0.2 µM for *THCAS*-specific and *CBDAS*-specific reverse primers, and 2 U Platinum® Taq DNA Polymerase (Life Technologies #10966-034). Thermocycling parameters were 94°C for 2 minutes, then 25 cycles of 94°C for 30 seconds, 58°C for 30 seconds, 72°C for 1 minute 15 seconds.

PCR reactions were performed in 0.2 mL 96 well PCR plates (Thermo Scientific #AB-0600) sealed with flat cap strips (Thermo Scientific #AB-0786) using a Gradient Palm–Cycler™ (Corbett Life Science) and occurred in a total volume of 50 µL. D589 and B1080/B1192 amplification products were separated by electrophoresis on a 1.5% and 1% SeaKem® LE agarose gel (Cambrex #50004) stained with GelRed™ (Biotium #41003) respectively. Amplification products were then visualized under UV illumination using the Bio-Rad Molecular Imager® Gel Doc™ XR+ system using Image Lab™ software.

LC-MS Cannabinoid profiling

Sample extraction

Bracts, bracteoles and surrounding leaf tissues were collected from the upper 30 cm of female inflorescences (International Union for the Protection of New Varieties of Plants (UPOV) (www.upov.int)). Sample preparation and extraction followed those of De Backer et al. (2009). Leaf material was dried at 35 ºC in a forced ventilation oven for 48 hours. Seeds and stalks were separated and discarded manually and 1 g of dried leaf material was ground using a Mixer Mill MM 301 (Retsch GmbH) at 30 rotations /second for 30 seconds.

Extractions were performed in duplicate for each plant. 250 mg of dried ground leaf tissue was weighed into a 25 mL volumetric flask and extracted by agitation in 25 mL mixture of methanol/chloroform (v/v: 9/1) for 30 minutes. 1 mL of the extract was evaporated under a gentle stream of nitrogen (N_2) and resuspended in 1 mL of water/methanol (v/v: 5/5). Samples were sonicated for 30 seconds and centrifuged using a Compact centrifuge 2-5 (Sigma) at 3000 rpm for 7 minutes to remove particulate material. 500 µL of the supernatant was transferred into 2 mL screw cap glass amber vial (Agilent Technologies #5182-0716). All solvents used for extractions and HPLC analysis were HPLC grade.

LC-MS parameters

LC-MS chromatographic runs were performed using an Agilent 1290 Infinity analytical HPLC instrument (Agilent Technologies, Palo Alto, CA, USA) equipped with a vacuum degasser, binary pump, autoinjector, and diode array detector (DAD, 1260), coupled with an Agilent 6120 Quadrupole mass detector (MSD). The LC-MS system was controlled using Agilent ChemStation software (Rev. B.04.03). Absorbance was monitored at 210 nm, 272 nm, 280 nm, 330 nm, and 360 nm. The column used was an Agilent Eclipse plus rapid resolution high definition (RRHD) C18 column (1.8 µm; 50 × 2.1 mm internal diameter). Column temperature was set at 30 ºC.

A linear gradient elution program was applied with the mobile phase containing acetonitrile with 0.005% trifluoroacetic acid (TFA) and Milli-Q water with 0.005% TFA. The solvent gradient was programmed from 10% to 99% acetonitrile with 0.005% TFA in 11.5 minutes. 99% acetonitrile with 0.005% TFA was maintained for 1.5 minutes, recalibrated to 10% in 2 minutes and then held at 10% for a further 1 minute. The total run time was 15 minutes. Flow rate and injection volume were set at 0.3 mL/minute and 3 µL respectively.

MSD parameters and settings followed those of a previously validated method (Liu et al. 2014). The MSD was operated in atmospheric pressure ionization–electrospray mode using the following parameters: scan mass

range, 100−1200; fragmentor, 150; capillary voltage, 3000 V (positive); drying gas flow, 12.0L/min (N_2); vaporizer temperature, 350°C; nebuliser pressure, 35 psi; drying gas temperature, 350°C. THCA, CBDA, CBGA, cannabigerol (CBG), delta 9-tetrahydrocannabivarin (THCV) (Weesp, The Netherlands), THC, CBD, and cannabinol (CBN) (Capalaba, Australia) cannabinoid reference standards were scanned to determine the most abundant and representative signal. Cannabinoids were quantified using selected-ion monitoring. Cannabinoids were allocated to one of four available MSD signal channels (Online Resource 1).

Linearity and reproducibility

Concentrations of cannabinoids were determined from the linear regression equation of calibration curves of individual reference standards. Calibration curves were obtained from five standard solutions comprising all eight cannabinoid standards, with standard solution concentrations at 0.08 µg/mL, 0.4 µg/mL, 2 µg/mL, 10 µg/mL, and 50 µg/mL. Calibration curves were calculated using unweighted linear regression analysis and expressed using R^2. Calibration curves for all eight cannabinoids were linear within the concentration range R^2 >0.99 (Online Resource 2). Calibration curves were conducted every 24 hours to allow for changes in response factor. Sample cannabinoid concentrations >50 µg/mL were diluted in water/methanol (v/v: 5/5) in a 1:20 dilution to ensure signals were within calibration range.

To determine the precision of the LC-MS instrument, standard solutions were injected six times. Relative standard deviation (RSD) for each cannabinoid peak area was found to be <2%. THC_{total} (= THC + THCA), CBD_{total} (= CBD + CBDA), and CBG_{total} (= CBG + CBGA) (Swift et al. 2013; Taura et al. 2007a) between sample replicates were typically <5% RSD. One way analysis of variance (ANOVA) calculated using GenStat 64-bit Release 16.1 (VSN International 2014) indicated that there was no significant difference in THC_{total} $F_{1, 130}$ = 0.01, P = 0.936, CBD_{total} $F_{1, 130}$ = 0.00, P = 0.985, and CBG_{total} $F_{1, 122}$ = 0.02, P = 0.898 between replicates.

Results

We were able to classify individual plants of each accession according to their chemotype and DNA marker genotypes. A combination of the D589 and B1080/B1192 DNA SCAR marker assays generated a haplotype for each plant which enabled us unequivocally to assign haplotype to chemotype. Associations between chemotype and haplotype were determined specifically at the plant level, as variation in chemotype and haplotype occurred between accessions and segregated within accessions.

Chemotype

Individual plants from each accession were allocated a chemotype on the basis of their THC_{total}, CBD_{total}, and CBG_{total} % DW concentrations (Online Resource 3). A histogram of $\log_{10}$ THC_{total} and CBD_{total} ratios of individual plants (*N*=66) shows three discrete distributions (Fig. 2). Based on these distributions, plants with a $\log_{10}$ CBD_{total}: THC_{total} ratio -2.64 to -1.88 were assigned to chemotype I, plants with a $\log_{10}$ CBD_{total}: THC_{total} ratio between -1.15 to 0.46 were assigned to chemotype II, while plants with a $\log_{10}$ CBD_{total}: THC_{total} ratio of 0.95 to 1.51 were assigned to chemotype III.

No plants were found to have levels of CBG_{total} that exceeded 10% of the cannabinoid fraction or with levels greater than either THC_{total} or CBD_{total}, although variability in the accumulation of CBG_{total} was observed across chemotypes (Fig. 3). Therefore, no plants were classified as chemotype IV. Individual plants within 8 out of 22 accessions had more than one chemotype. Chemotype I, II, III were evenly distributed within the sample population. CBN contributed negligibly to chemotype and was either <LOQ, or at levels no greater than 0.004% DW (data not shown).

Five individual plants from the East Asian accessions J (TS1300283) and H (TS1300317) had THCV: THC ratio >1. HPLC chromatograms at 272 nm showed two unknown peaks at 7.5 and 9.1 minutes in all five plants (Table 2). Using UV maximum data and peak fragment ions determined from a positive MSD scan, these compounds were tentatively identify as THCVA and CBDVA.

Marker genotypes

Individual plants from each accession were classified according to their marker genotypes, generating a haplotype from a combination of the D589 and B1080/B1192 DNA SCAR marker assays. Plants were scored separately for each marker (Online Resource 3). Conformity between the D589 and B1080/B1192 marker genotypes was observed (Table 3). The D589 marker assay scored plants as either $B_{Tpresent}$ or $B_{Tabsent}$, while the B1080/B1192 marker assay scored plants as either B_{THCAS}, $B_{THCAS}B_{CBDAS}$, or B_{CBDAS}. All plants with a $B_{Tpresent}$ D589 marker genotype had either a B_{THCAS} or $B_{CBDAS/THCAS}$ B1080/B1192 marker genotype, while all plants with $B_{Tabsent}$ D589 marker genotype were found to have a B_{CBDAS} B1080/B1192 marker genotype.

Chemotype and marker genotype

The chemotype of 65 out of 66 plants were correctly predicted by both D589 and B1080/B1192 DNA SCAR marker assays (Table 3). However, the dominant D589 marker assays was less specific in predicting chemotype then the co-dominant B1080/B1192 marker assay, and was unable to differentiate between chemotype I and II. 45 out of 46 chemotype I and II plants were correctly predicted by the presence of the D589 B_{Tpresent} marker genotype (sensitivity of 97.8%); while 20 out of 20 chemotype III plants were correctly predicted by the absence of the B_{Tabsent} genotype (specificity of 100.00%). Similar results were observed with B1080/B1192 marker assay, with 27 out of 27 chemotype I and 18 out of 19 chemotype II plants correctly predicted by B_{THCAS} and $B_{\text{THCAS}}B_{\text{CBDAS}}$ marker genotypes respectively, while 20 out of 20 chemotype III plants were correctly predicted by the B_{CBDAS} marker genotype. As expected there were highly significant associations between chemotype I-III and marker genotype, Pearson's $\chi^2_{2,\,66} = 61.63$, $P < 0.001$ (GenStat 64-bit Release 16.1).

A higher level of variability in cannabinoid composition occurred in chemotype II $B_{\text{THCAS}}B_{\text{CBDAS}}$ genotypes comparatively with chemotype I B_{THCAS} genotypes and chemotype III B_{CBDAS} (B_{Tabsent}) genotypes (Fig. 4). Chemotype II $B_{\text{THCAS}}B_{\text{CBDAS}}$ genotypes ranged from 25.1% $\text{THC}_{\text{total}}$ (72.4% $\text{CBD}_{\text{total}}$) /total cannabinoid fraction to 86.3% $\text{THC}_{\text{total}}$ (13.8% $\text{CBD}_{\text{total}}$). In contrast, chemotype I B_{THCAS} genotypes and chemotype III B_{CBDAS} (B_{Tabsent}) genotypes ranged from 91.3% $\text{THC}_{\text{total}}$ (0.3% $\text{CBD}_{\text{total}}$) to 99.5% $\text{THC}_{\text{total}}$ (0.5% $\text{CBD}_{\text{total}}$) and 6.3% $\text{THC}_{\text{total}}$ (93.7% $\text{CBD}_{\text{total}}$) to10.1% $\text{THC}_{\text{total}}$ (89.9% $\text{CBD}_{\text{total}}$) /total cannabinoid fraction respectively.

Cannabinoid homologues

A lack of correspondence between chemotype and genotype was observed in a single individual, H-3, from the East Asian accessions H (TS1300317). This individual had B_{Tabsent} and B_{CBDAS} genotypes associated with chemotype III (Table 3), despite exhibiting a $\text{THC}_{\text{total}}$: $\text{CBD}_{\text{total}}$ ratio associated with chemotype II (Fig. 3). CBDVA: THCVA $\log_{10}$ ratios, determined from THCA and CBDA standards, corresponded to chemotype in four/five plants which demonstrated a THCV: THC>1. H-3 was found to have CBDVA at levels exceeding other cannabinoids, and produced a CBDVA: THCVA $\log_{10}$ ratio of 1.17 and so was associated with chemotype III. Pentyl- and propyl-chemotypes and $B_{\text{Tabsent}} + B_{\text{CBDAS}}$ haplotype for H-3 were confirmed through duplicated LC-MS sample replicates, and through repeat DNA extraction and amplification using D589 and B1080/B1192 DNA SCAR markers.

Discussion

Chemotype II variability

Previously it was shown that chemotype II $B_{THCAS}B_{CBDAS}$ genotypes exhibited a CBD_{total}: THC_{total} log_{10} ratio ranging from 0.36 to 0.01 (Pacifico et al. 2006). In the present study, chemotype II $B_{THCAS}B_{CBDAS}$ genotypes exhibited a CBD_{total}: THC_{total} log_{10} ratio ranging from 0.46 to -1.15, a greater than three-fold increase in variability (Fig. 4). GC analysis of chemotype II has typically demonstrated a more narrow range skewed towards a high CBD_{total} to THC_{total} ratio (De Meijer et al. 1992; Hillig and Mahlberg 2004; Weiblen et al. 2015), although a low CBD_{total} to THC_{total} ratio similar to those reported here have been observed in Southeast Asian landraces (Tipparat et al. 2012). Nevertheless, this is the first time that such a wide range in chemotype II variability has been unequivocally assigned to genotype and heterozygosity of *THCAS* and *CBDAS*.

In a single (*B*) locus model, chemotype II individuals are assumed to be heterozygote for the co-dominant B_T (*THCAS*) and B_D (*CBDAS*) alleles (De Meijer et al. 2003). These alleles encode functional sequence variants of THCA and CBDA synthase, and therefore the efficiency (Pacifico et al. 2006) with which CBG(V)A is converted to THC(V)A (B_T) and CBD(V)A (B_D) (Sirikantaramas et al. 2004; Taura et al. 2007b) (Fig. 1). All things being equal chemotype II individuals would be expected to have a CBD_{total}: THC_{total} ratio close to 1:1. Deviation from this ratio within chemotype II has previously been proposed to be due to either an innate catalytic superiority of CBDA synthase over THCA synthase (Weiblen et al. 2015), or from genetic vs environment interactions (Potter 2009). Considering that ratios of chemotype II plants deviated both towards CBD_{total} and THC_{total}, and that all plants were grown in environmentally controlled conditions, it could be suggested that chemotype II variability is largely influenced by genetic control.

Chemotype II $B_{THCAS}B_{CBDAS}$ genotype variability may be generated by the presence of alternative *B* locus alleles. Four additional alleles encoding reduced or non-functional variants of *THCAS* (B_{T0}) and *CBDAS* (B_{DW}; B_{D01}; B_{D02}) have been observed within chemotype IV (Onofri et al. 2015) (Fig. 1). Individual's homozygote for these alleles have a reduced capacity to form THCA/CBDA, resulting in the accumulation CBGA. These alleles are associated with nonsynonymous substitutions (Onofri et al. 2015), some of which are non-specific to the D589 and B1080/B1192 marker genotypes. It is possible these and potentially other alleles contribute to chemotype II variability. However, if these alleles were present in sufficient frequency to account for chemotype II variability, one would expect a high frequency of chemotype IV throughout the gene pool, although similar to

the tripartite distributions of chemotype I, II, and III observed within the sample collection (Fig. 3), this has not been found (De Meijer et al. 2003; Hillig and Mahlberg 2004; Pacifico et al. 2006).

Variability within chemotype II $B_{THCAS}B_{CBDAS}$ genotypes may be more comprehensively explained by a multi-locus model. The identification of multiple sequence variants of *THCAS* or *CBDAS* within individuals (Kojoma et al. 2006; Onofri et al. 2015; Van Bakel et al. 2011; Weiblen et al. 2015) and segregation in mapping populations, suggests that at least two closely linked loci govern THC(V)A and CBD(V)A composition. Recent evidence indicates that *THCAS* and *CBDAS* are positioned 1.1 cM apart (Weiblen et al. 2015). The presence of contrasting functional allelic variants of *THCAS* and *CBDAS* occurring at multiple loci may be a possible explanation as to the range and variability in CBD_{total}: THC_{total} ratios occurring within chemotype II, and is also consistent with the reduced representation of chemotype IV among the gene pool. Analysis of *THCAS* and *CBDAS* sequence variants and comparative genomic approaches using next generation sequencing technologies may help to determine the genomic organisation of chemotype, and if alternative alleles contribute to chemotype II variability.

Characterisation of chemotype

The large range of CBD_{total}: THC_{total} ratios observed within chemotype II $B_{THCAS}B_{CBDAS}$ genotypes would appear to obscure the distinction between chemotypes. Chemotypes have previously been classified on the basis of CBD_{total}: THC_{total} $\log_{10}$ histogram frequency distributions (Hillig and Mahlberg 2004; Pacifico et al. 2008; Staginnus et al. 2014; Tipparat et al. 2012), with CBD_{total}: THC_{total} $\log_{10}$ ratio of 0.0 (Pacifico et al. 2008) and -1.0 being used as an arbitrary division between chemotype I and II (Hillig and Mahlberg 2004; Tipparat et al. 2012). In the present study chemotype II $B_{THCAS}B_{CBDAS}$ genotypes were found to exceed these $\log_{10}$ ratios, albeit the latter $\log_{10}$ ratio of -1.0 was only exceeded by a single individual I-3 (-1.15). Nevertheless, this highlights a potential limitation of characterising chemotype exclusively from cannabinoid profiling, and questions the utility of this approach in determining genotype.

Characterisation and selection for chemotype I B_{THCAS} and chemotype III B_{CBDAS} genotypes may aid in the compositional-selection of uniform plant lines, with B_{THCAS} and B_{CBDAS} genotypes found to vary by only 8.2% THC_{total} (0.2% CBD_{total}) and 3.2% THC_{total} (3.8% CBD_{total}) within the cannabinoid fraction respectively (Fig. 4). Selecting for chemotype III B_{CBDAS} genotypes may prove especially beneficial for development of uniform plant lines for hemp fibre, seed, and pharmacological production, given the strong association between chemotype III and THC content <0.2 % DW (Pacifico et al. 2008; Pacifico et al. 2006), and growing interest in CBD(V)A and

CBD(V) derivatives as pharmacological entities (De Petrocellis et al. 2011; Gallily et al. 2015; Hill et al. 2013; Iseger and Bossong 2015). However, additional cannabinoid profiling is required in order to differentiate cannabinoid homologue compositions and to characterise CBD_{total}: THC_{total} chemotype II variability accurately.

Due to variability within chemotype II the uniformity of botanical drug product CBD_{total}: THC_{total} ratios has been achieved by combining cannabinoid extracts from chemotype I and III plants (Potter 2014). However, this approach requires the growth of chemotype I drug-type plants (Swift et al. 2013) and limits compositional control over other potentially pharmacologically relevant cannabinoids (Izzo et al. 2009). Identification and recombination of novel chemotype-determining alleles (Barker et al. 2007) may facilitate the selection of chemotype variants with CBD(V)A: THC(V)A-specific ratios. This approach to genetic metabolic engineering may result in the development of intermediate chemotypes tailed for specific pharmaceutical applications.

Propyl- pentyl-cannabinoid chemotypes

A single individual from the East Asian accessions H (H-3) was characterised as having a propyl-chemotype III and pentyl-chemotype II profile, while exhibiting a chemotype III $B_{CBDAS}B_{CBDAS}$ + $B_{Tabsent}$ haplotype. THCA and CBDA synthase are considered isoforms of the same enzyme (De Meijer et al. 2003; Taura et al. 2007a), sharing 83.9% amino acid identity (Taura et al. 2007b). Residues of THCA in chemotype III are thought to be due to an imperfect capacity of CBDA synthase to form CBDA (De Meijer et al. 2003). Chemotype III individuals have been observed to transitionally deviate from III to II during early stages of plant development (Pacifico et al. 2008). THCA residue formation may have occurred in H-3 during this period and the remaining CBGVA substrate catalysed to CBDVA during later developmental stages. Plants which exhibit propyl-chemotype III profiles are not frequently distributed among the *C. sativa* genepool (Baker et al. 1980; Hillig and Mahlberg 2004). It is possible that two alleles encoding CBGA-specific or CBGVA-specific CBDA synthases exist, although this would fail to explain why THC_{total} and CBD_{total} were found at an almost equal ratio in an individual presumably carrying a functional variant of *CBDAS*. Further allelism tests on progenies segregating for propyl- and pentyl-cannabinoids may provide insight into the genetic regulation of homologue variability.

Conclusions

We carried out a representative survey from a wide range of genetically diverse accessions to explore the relationship between chemotype and genotype, and to identify a suitable strategy for characterising cannabinoid

compositional variation. A high level of variability in CBD_{total}: THC_{total} ratio was observed within chemotype II $B_{THCAS}B_{CBDAS}$ genotypes. However, the genetic regulation underlying this variability remains uncertain. More detailed exploration of the relationship between chemotype and genotype using next generation sequencing technologies offers the potential of characterising chemotype with greater accuracy, and may lead to advances in the genetic metabolic engineering of *C. sativa* for pharmacological development. In the interim, a strategy which adopts the use of both the B1080/B1192 DNA SCAR marker genotype and careful cannabinoid profiling may provide an effective approach for classifying chemotypic variability within diverse germplasm collections, and may also contribute to the development of cultivars with cannabinoid profiles specific for end-use applications.

References

Appendino G, Chianese G, Taglialatela-Scafati O (2011) Cannabinoids: occurrence and medicinal chemistry. Curr Med Chem 18:1085-1099

Baker P, Gough T, Taylor B (1980) Illicitly imported *Cannabis* products: some physical and chemical features indicative of their origin. B Narcotics 32:31-40

Barker GC, Larson TR, Graham IA, Lynn JR, King GJ (2007) Novel insights into seed fatty acid synthesis and modification pathways from genetic diversity and quantitative trait loci analysis of the *Brassica* C genome. Plant Physiol 144:1827-1842

De Backer B et al. (2009) Innovative development and validation of an HPLC/DAD method for the qualitative and quantitative determination of major cannabinoids in cannabis plant material. J Chromatogr B 877:4115-4124

De Meijer EPM, Bagatta M, Carboni A, Crucitti P, Moliterni VMC, Ranalli P, Mandolino G (2003) The inheritance of chemical phenotype in *Cannabis sativa* L. Genetics 163:335-346

De Meijer EPM, Hammond KM, Sutton A (2009) The inheritance of chemical phenotype in *Cannabis sativa* L. (IV): cannabinoid-free plants. Euphytica 168:95-112

De Meijer EPM, Van der Kamp HJ, Van Eeuwijk FA (1992) Characterisation of *Cannabis* accessions with regard to cannabinoid content in relation to other plant characters. Euphytica 62:187-200

De Petrocellis L et al. (2011) Effects of cannabinoids and cannabinoid-enriched *Cannabis* extracts on TRP channels and endocannabinoid metabolic enzymes. Brit J Pharmacol 163:1479-1494

Dussy FE, Hamberg C, Luginbühl M, Schwerzmann T, Briellmann TA (2005) Isolation of Δ9-THCA-A from hemp and analytical aspects concerning the determination of Δ9-THC in *Cannabis* products. Forensic Sci Int 149:3-10

ElSohly MA, Slade D (2005) Chemical constituents of marijuana: the complex mixture of natural cannabinoids. Life Sci 78:539-548

Faeti V, Mandolino G, Ranalli P (1996) Genetic diversity of *Cannabis sativa* germplasm based on RAPD markers. Plant Breeding 115:367-370

Forapani S, Carboni A, Paoletti C, Moliterni VMC, Ranalli P, Mandolino G (2001) Comparison of hemp varieties using random amplified polymorphic DNA markers. Crop Sci 41:1682-1689

Gallily R, Yekhtin Z, Hanuš LO (2015) Overcoming the bell-shaped dose-response of cannabidiol by using *Cannabis* extract enriched in cannabidiol. Pharmacology & Pharmacy 6:75

Gao C et al. (2014) Diversity analysis in *Cannabis sativa* based on large-scale development of expressed sequence tag-derived simple sequence repeat markers. PLoS One 9:e110638

Gertsch J, Pertwee RG, Di Marzo V (2010) Phytocannabinoids beyond the *Cannabis* plant–do they exist? Brit J Pharmacol 160:523-529

Gilmore S, Peakall R, Robertson J (2007) Organelle DNA haplotypes reflect crop-use characteristics and geographic origins of *Cannabis sativa.* Forensic Sci Int 172:179-190

Happyana N, Agnolet S, Muntendam R, Van Dam A, Schneider B, Kayser O (2013) Analysis of cannabinoids in laser-microdissected trichomes of medicinal *Cannabis sativa* using LCMS and cryogenic NMR. Phytochemistry 87:51-59

Hill TDM et al. (2013) Cannabidivarin-rich cannabis extracts are anticonvulsant in mouse and rat via a CB1 receptor-independent mechanism. Brit J Pharmacol 170:679-692

Hillig KW (2005) Genetic evidence for speciation in *Cannabis* (*Cannabaceae*). Genet Resour Crop Ev 52:161-180

Hillig KW, Mahlberg PG (2004) A chemotaxonomic analysis of cannabinoid variation in *Cannabis* (*Cannabaceae*). Am J Bot 91:966-975

Iseger TA, Bossong MG (2015) A systematic review of the antipsychotic properties of cannabidiol in humans. Schizophr Res 162:153-161

Izzo AA, Borrelli F, Capasso R, Di Marzo V, Mechoulam R (2009) Non-psychotropic plant cannabinoids: new therapeutic opportunities from an ancient herb. Trends Pharmacol Sci 30:515-527

399 Kojoma M, Seki H, Yoshida S, Muranaka T (2006) DNA polymorphisms in the tetrahydrocannabinolic acid
400 (THCA) synthase gene in “drug-type” and “fiber-type” *Cannabis sativa* L. Forensic Sci Int 159:132-
401 140
402 Liu L, Tong C, Bao J, Waters DLE, Rose TJ, King GJ (2014) Determination of starch lysophospholipids in rice
403 using liquid chromatography–mass spectrometry (LC-MS). J Agr Food Chem 62:6600-6607
404 Mandolino G, Carboni A (2004) Potential of marker-assisted selection in hemp genetic improvement. Euphytica
405 140:107-120
406 Onofri C, de Meijer EPM, Mandolino G (2015) Sequence heterogeneity of cannabidiolic- and
407 tetrahydrocannabinolic acid-synthase in *Cannabis sativa* L. and its relationship with chemical
408 phenotype. Phytochemistry. doi:10.1016/j.phytochem.2015.03.006
409 Pacifico D, Miselli F, Carboni A, Moschella A, Mandolino G (2008) Time course of cannabinoid accumulation
410 and chemotype development during the growth of *Cannabis sativa* L. Euphytica 160:231-240
411 Pacifico D, Miselli F, Micheler M, Carboni A, Ranalli P, Mandolino G (2006) Genetics and marker-assisted
412 selection of the chemotype in *Cannabis sativa* L. Mol Breeding 17:257-268
413 Potter DJ (2009) The propagation, characterisation and optimisation of *Cannabis Sativa* L. as a
414 phytopharmaceutical. Dissertation, King's College London
415 Potter DJ (2014) A review of the cultivation and processing of cannabis (*Cannabis sativa* L.) for production of
416 prescription medicines in the UK. Drug Testing and Analysis 6:31-38
417 Rotherham D, Harbison SA (2011) Differentiation of drug and non-drug *Cannabis* using a single nucleotide
418 polymorphism (SNP) assay. Forensic Sci Int 207:193-197
419 Shoyama Y, Hirano H, Nishioka I (1984) Biosynthesis of propyl cannabinoid acid and its biosynthetic
420 relationship with pentyl and methyl cannabinoid acids. Phytochemistry 23:1909-1912
421 Sirikantaramas S, Morimoto S, Shoyama Y, Ishikawa Y, Wada Y, Shoyama Y, Taura F (2004) The gene
422 controlling marijuana psychoactivity molecular cloning and heterologous expression of Δ1-
423 tetrahydrocannabinolic acid synthase from *Cannabis sativa* L. J Biol Chem 279:39767-39774
424 Small E, Cronquist A (1976) A practical and natural taxonomy for *Cannabis.* Taxon 25:405-435
425 Staginnus C, Zörntlein S, de Meijer E (2014) A PCR marker linked to a THCA synthase polymorphism is a
426 reliable tool to discriminate potentially THC-rich plants of *Cannabis sativa* L. J Forensic Sci 59:919-
427 926

Swift W, Wong A, Li KM, Arnold JC, McGregor IS (2013) Analysis of cannabis seizures in NSW, Australia: cannabis potency and cannabinoid profile. PloS One 8:e70052

Taura F, Sirikantaramas S, Shoyama Y, Shoyama Y, Morimoto S (2007a) Chemistry & biodiversity. WILEY-VCH Verlag, Zurich

Taura F, Sirikantaramas S, Shoyama Y, Yoshikai K, Shoyama Y, Morimoto S (2007b) Cannabidiolic-acid synthase, the chemotype-determining enzyme in the fiber-type *Cannabis sativa.* FEBS Lett 581:2929-2934

Tipparat P, Natakankitkul S, Chamnivikaipong P, Chutiwat S (2012) Characteristics of cannabinoids composition of *Cannabis* plants grown in Northern Thailand and its forensic application. Forensic Sci Int 215:164-170

Van Bakel H, Stout J, Cote A, Tallon C, Sharpe A, Hughes T, Page J (2011) The draft genome and transcriptome of *Cannabis sativa.* Genome Biol 12:R102

Weiblen GD, Wenger JP, Craft KJ, ElSohly MA, Mehmedic Z, Treiber EL, Marks MD (2015) Gene duplication and divergence affecting drug content in *Cannabis sativa*. New Phytol. doi:10.1111/nph.13562

Werz O et al. (2014) Cannflavins from hemp sprouts, a novel cannabinoid-free hemp food product, target microsomal prostaglandin E2 synthase-1 and 5-lipoxygenase. PharmaNutrition 2:53-60

Figure legends

Fig. 1 A single *A* and *B* locus genetic model governing cannabinoid composition in *Cannabis sativa* L. Adapted from (De Meijer et al. 2009). Black lines indicate alternative cannabinoid biosynthetic pathways. Arrows correspond to associated cannabinoid accumulation and chemotype. The A_{pentyl} allele is associated with the accumulation of pentyl-cannabinoid intermediate CBGA, while the A_{propyl} allele is associated with the accumulation of propyl-cannabinoid intermediate CBGVA (Shoyama et al. 1984). The B_T and B_D alleles encode functional THCA synthase and CBDA synthase homologs respectively (De Meijer et al. 2003). The B_{T0} allele encodes a non-functional THCA synthase homolog, while B_{Dw}, B_{D01}, and B_{D02} encode non-functional CBDA synthase homologs (Onofri et al. 2015). The B_T allele is associated with THCA + THCVA (THC(V)A) and chemotype I (B_T) and chemotype II (B_TB_D). The B_D allele is associated with CBDA + CBDVA (CBD(V)A) accumulation and chemotype III (B_D) and chemotype II (B_TB_D). The B_{T0}, B_{Dw}, B_{D01}, and B_{D02} alleles are associated with CBGA + CBGVA (CBG(V)A) accumulation and chemotype IV; cannabidiolic acid (CBDA); cannabidivarinic acid (CBDVA); cannabigerolic acid (CBGA); cannabigerovarinic acid (CBGVA); delta-9-tetrahydrocannabinolic acid (THCA); delta-9-tetrahydrocannabivarinic acid (THCVA)

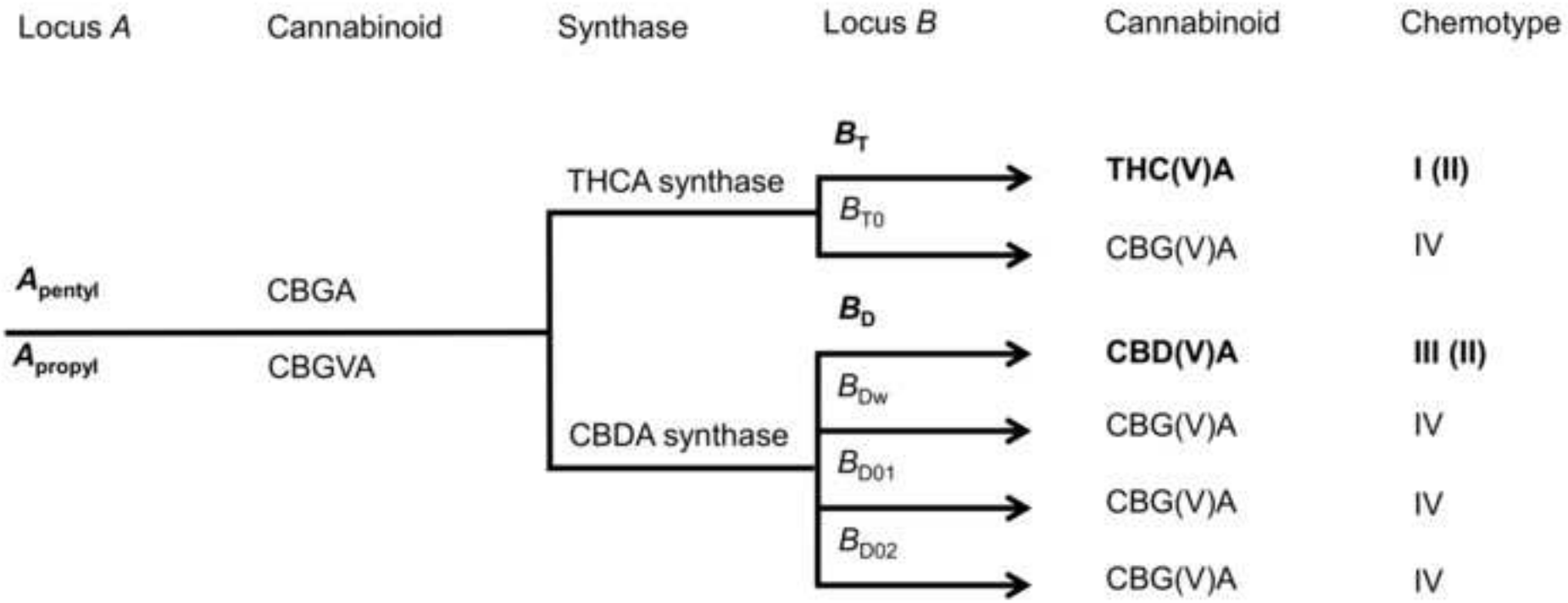
Fig. 2 Frequency distribution of CBD_{total}: THC_{total} DW $\log_{10}$ ratios of individual *Cannabis sativa* L. plants derived from 22 accessions. Three discrete distributions are associated with chemotypes I-III (Hillig and Mahlberg 2004). Dotted lines indicate $\log_{10}$ ratio divisions between chemotype I, II, and III; cannabidiol (CBD); cannabidiolic acid (CBDA); CBD + CBDA (CBD_{total}); dry weight (w/w) (DW); delta-9-tetrahydrocannabinol (THC); delta-9-tetrahydrocannabinolic acid (THCA); THC + THCA (THC_{total})

Fig. 3 Penyl-cannabinoid composition within the cannabinoid fraction of individual *Cannabis sativa* L. plants derived from 22 accessions indicating variability both within and between accessions and chemotypes. Letters indicate accession ID (Table 1). Numbers correspond to each of three plant individuals per accession. Chemotypes determined from the frequency distributions of CBD + CBDA (CBD_{total}): THC + THCA (THC_{total}) $\log_{10}$ ratios of individual *C sativa* plants; cannabidiol (CBD); cannabidiolic acid (CBDA); cannabigerol (CBG); cannabigerolic acid (CBGA); CBG + CBGA (CBG_{total}); delta-9-tetrahydrocannabinol (THC); delta-9-tetrahydrocannabinolic acid (THCA)

Fig. 4 CBD_{total}: THC_{total} compositional variability within and between *Cannabis sativa* L. chemotypes and its relationship with B1080/B1192 DNA SCAR marker genotype. Black arrow indicates previous range in CBD_{total}: THC_{total} compositional variability within chemotype II $B_{THCAS}B_{CBDAS}$ genotype (Pacifico et al. 2006). Data points correspond to CBD_{total}: THC_{total} composition within the total cannabinoid fraction. Square, circular, and triangular data points correspond to chemotype III, II, and I respectively; sequence characterised amplified region (SCAR); cannabidiol (CBD); cannabidiolic acid (CBDA); CBD + CBDA (CBD_{total}); delta-9-tetrahydrocannabinol (THC); delta-9-tetrahydrocannabinolic acid (THCA); THC + THCA (THC_{total})

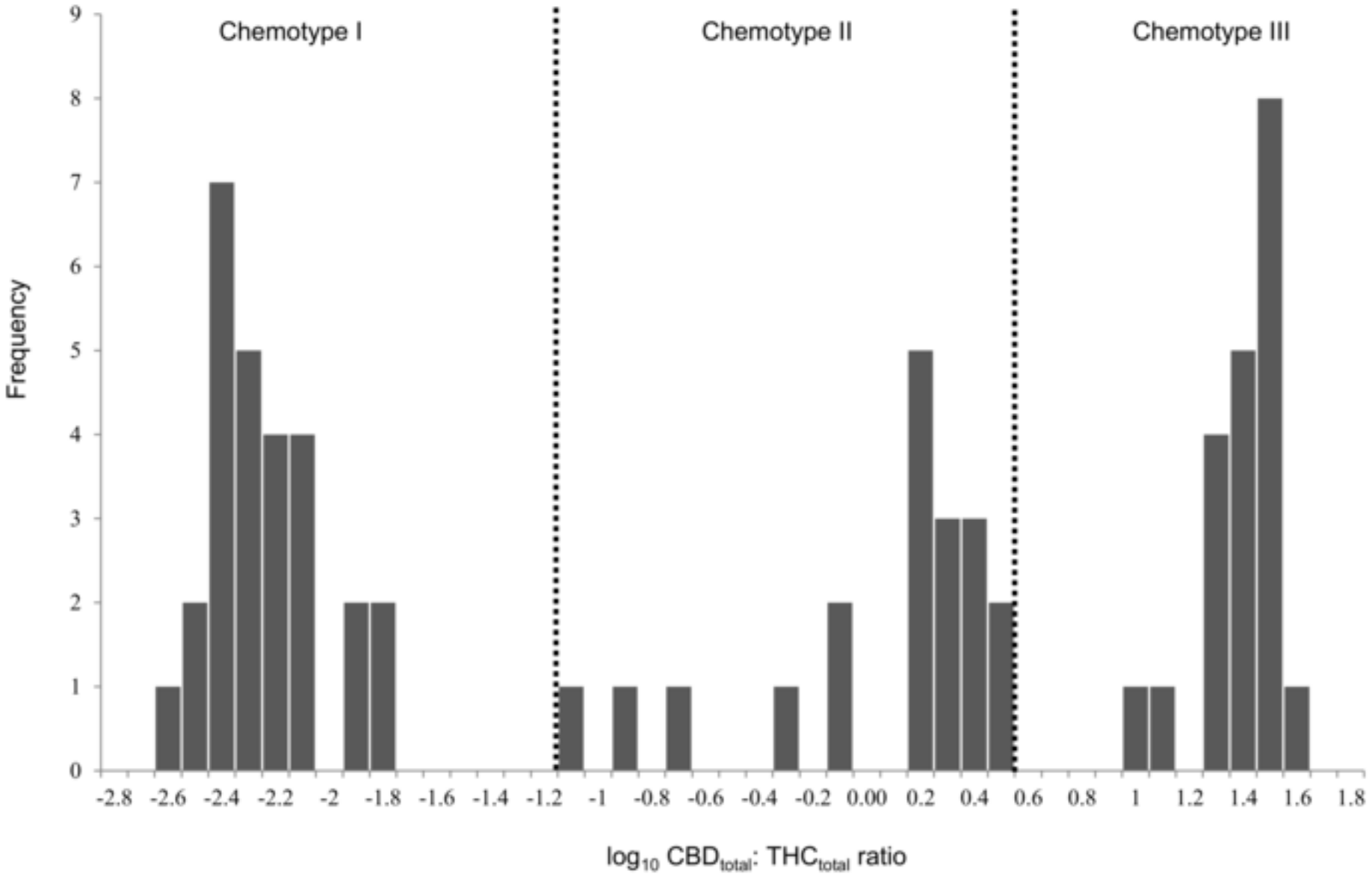
colour figure

Click here to download colour figure Fig1.tif



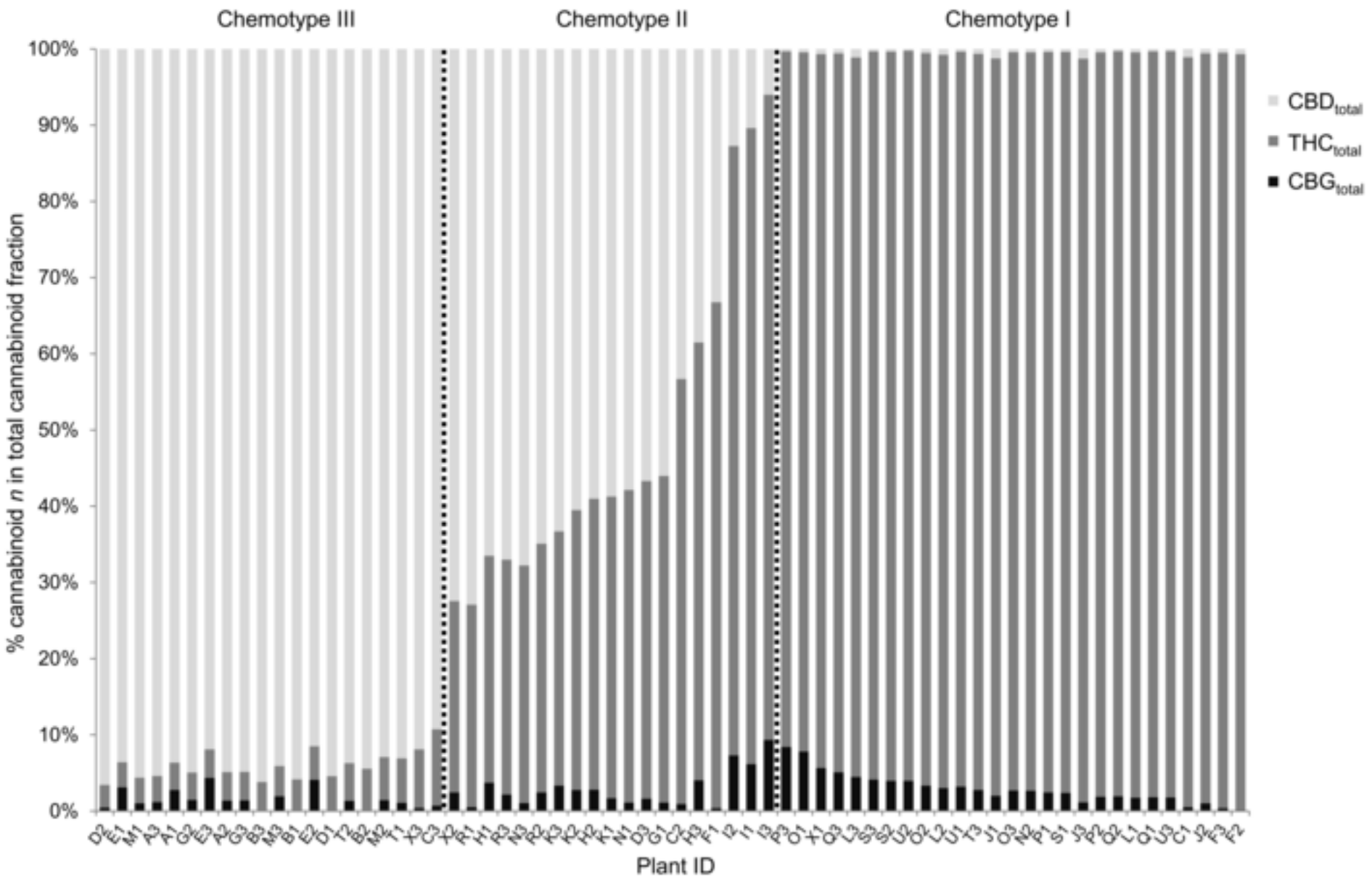
colour figure

Click here to download colour figure Fig2.tif



colour figure

Click here to download colour figure Fig 3.tif



colour figure

Click here to download colour figure Fig 4.tif

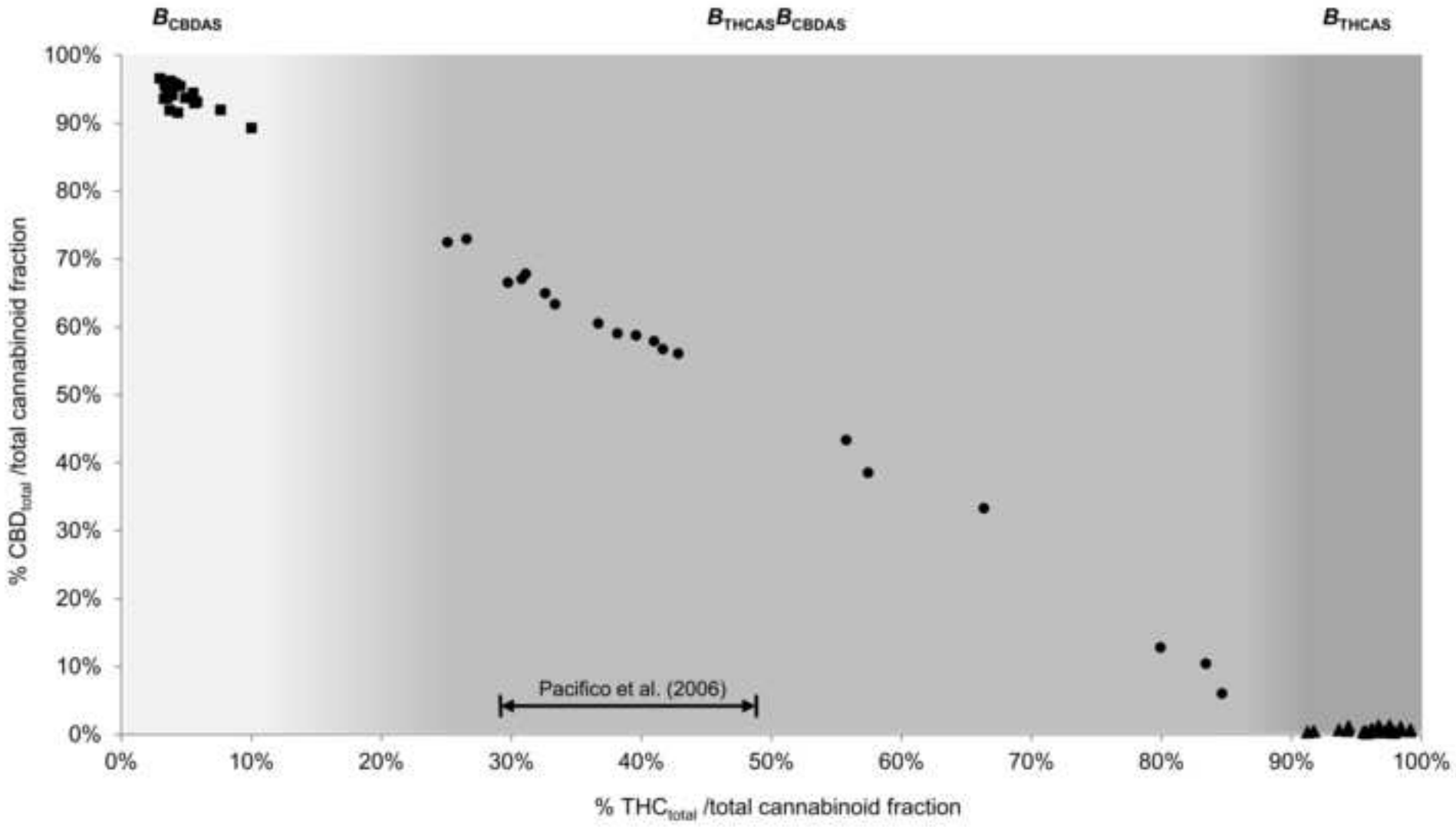


Table 1 Description of 22 *Cannabis sativa* L. accessions used to examine the relationship between chemotype and genotype[a]

Accession[b]	ID	Domestication status	Cultivar	Crop-type	Grouping	Region of origin	Country of origin
TS1300610	A	Cultivar	Futura 75	Fibre	EF	Europe	France
TS1300026	B	Cultivar	USO11	Fibre	EF	Europe	Ukraine
TS1300037	C	Cultivar	Kompolti	Fibre	EF	Europe	Hungary
TS1300041	D	Cultivar	Futura 77	Fibre	EF	Europe	France
TS1300285	E	Cultivar	Rastislavicka	Fibre	EF	Europe	Slovakia
TS1300289	F	Cultivar	Krasnodarskaya	Fibre	EF	Europe	Russia
TS1300287	G	Cultivar	Fedrina 74	Fibre	EF	Europe	France
TS1300317	H	Landrace	-	Fibre	EAF	East Asia	China
TS1300592	I	Landrace	-	Fibre	EAF	East Asia	China
TS1300283	J	Landrace	-	Fibre	EAF	East Asia	China
TS1300567	K	Cultivar	MS-77	Fibre	EAF	East Asia	China
TS1300477	L	Landrace	-	Fibre	EAF	East Asia	China
TS1300594	M	Cultivar	CHG SSL#12	Fibre	EAF	East Asia	China
TS1300394	N	Landrace	-	Fibre	EAF	East Asia	China
TS1300301	O	Cultivar	Thai Skunk	Drug	MOD	Mixed	Netherlands
TS1300246	P	Cultivar	Skunk #1	Drug	MOD	Mixed	North America
TS1300248	Q	Cultivar	Super Skunk	Drug	MOD	Mixed	Netherlands
TS1300536	R	Landrace	-	Drug	MOD	South Asia	Nepal
TS1300585	S	Landrace	-	Drug	MOD	South Asia	India
TS1300308	T	Landrace	-	Drug	MOD	Central-Southern Asia	Afghanistan
TS1300493	U	Landrace	-	Drug	MOD	East Asia	China
TS1300609	X	Cultivar	Pan 3	Drug	MOD	Mixed	Unknown

[a]Accessions sourced from the Ecofibre global *Cannabis sativa.* L. germplasm collection. [b]Accessions were selected from three genetically and chemotypically diverse *C. sativa* groupings (Hillig 2005; Hillig and Mahlberg 2004); European fibre (EF); East Asian fibre (EAF); mixed origin drug-type (MOD)

Table 2 HPLC-DAD spectrum and MSD fragment ion of acidic cannabinoids

Compound	Retention time (min)	UV maximum (nm)	Fragment ion [M +1] (*m/z*)	Molecular weight (g/mol)
CBDA	8.2	223, 269, 307	341, 359	358.47
CBDVA	7.5	223, 269, 307	313, 331	330.42
THCA	10.0	222, 271, 306	341, 359	358.47
THCVA	9.1	222, 271, 306	313, 331	330.42

High performance liquid chromatography-diode array detector (HPLC-DAD); mass detector (MSD); cannabidiolic acid (CBDA); cannabidivarinic acid (CBDVA); delta-9-tetrahydrocannabinolic acid (THCA); delta-9-tetrahydrocannabivarinic acid (THCVA); ultraviolet (UV)

Table 3 D589 and B1080/B1192 DNA SCAR marker genotypes and associated chemotype of individual *Cannabis sativa* L. plants derived from 22 accessions

Marker Genotype	Chemotype I *N*=27	Chemotype II *N*=19	Chemotype III *N*=20
***B*Tpresent**	27	18	0
B_{THCAS}	27	0	0
$B_{THCAS}B_{CBDA}$	0	18	0
***B*Tabsent**	0	1[a]	20
B_{CBDAS}	0	1[a]	20

Bold indicates the dominant D589 DNA sequence characterised amplified region (SCAR) marker genotype (Rows 2 and 5). Rows 3, 4, and 6 indicate the co-dominant B1080/B1192 DNA SCAR marker genotypes. Columns 2-4 indicate the number of DNA SCAR marker genotypes which predicted chemotype. [a]Non-correspondence between DNA SCAR marker genotype and chemotype of plant individual H-3

Supplementary Files

Click here to view linked References

Click here to access/download
Supplementary Files
ESM_1.pdf

Supplementary Files

Click here to view linked References

Supplementary Files

Click here to view linked References