Molecular Breeding (2006) 17: 257–268
DOI 10.1007/s11032-005-5681-x

© Springer 2006

Genetics and marker-assisted selection of the chemotype in *Cannabis sativa* L.

Daniela Pacifico, Francesca Miselli, Mirta Micheler, Andrea Carboni, Paolo Ranalli and Giuseppe Mandolino*
*Istituto Sperimentale per le Colture Industriali, Via di Corticella 133 – 40128 Bologna, Italy; *Author for correspondence (e-mail: g.mandolino@isci.it; phone: +39-051-6316811; fax +39-051-374857)*

Received 4 August 2005; accepted in revised form 5 December 2005

Key words: Cannabinoids, Cannabis, Chemotype, Hemp, Markers, MAS

Abstract

Cannabis sativa is an interesting crop for several industrial uses, but the legislations in Europe and USA require a tight control of cannabinoid type and content for cultivation and *subsidies release*. Therefore, cannabinoid survey by gas chromatography of materials under selection is an important step in hemp breeding. In this paper, a number of *Cannabis* accessions were examined for their cannabinoid composition. Their absolute and relative content was examined, and results are discussed in the light of both the current genetic model for cannabinoid's inheritance, and the legislation's requirements. In addition, the effectiveness of two different types of markers associated to the locus determining the chemotype in *Cannabis* was evaluated and discussed, as possible tools in marker-assisted selection in hemp, but also for possible applications in the forensic and pharmaceutical fields.

Introduction

Hemp (*Cannabis sativa* L.) is one of the oldest non-food crops in the world, and until World War II, it was one of the most important fibre species, cultivated with high productivity for textile end-uses in both America and Europe. Today, hemp fibres are extracted mainly for non-textile uses: production of pulp for technical applications, injection in press-moulded parts in automotive industry, insulation mats (Karus and Vogt 2004). Hemp seed oil can be extracted for both industrial and nutritional uses (Ranalli et al. 1999; Callaway 2004; Kriese et al. 2004). *Cannabis sativa* cultivation requires little chemical inputs; the plant is a candidate for heavy metals phytoextraction from contaminated soils (Arru et al. 2004), it naturally limits weeds' presence, and it is also used for biodegradable plastic, fuels and medical application (Ranalli et al. 1999; Guy and Stott 2004); finally, hemp is suited for cultivation in marginal soils. For all its potential, hemp cultivation has been supported by UE, becoming for some productive chains an economically interesting option.

It is today generally accepted that *Cannabis* is a monospecific genus (Small et al. 1976; see however Hillig and Mahlberg 2004; Hillig 2005 for a different point of view). The intoxicant/pharmaceutical action of some Cannabis strains is due to the presence of a single secondary product, known as Δ^9-tetrahydrocannabinolic acid (Δ^9-THCA, THCA or simply THC, if indicated by its neutral form); the drug potential, for all analytical and legal purposes, is given by its amount in the extracts from dried

mature inflorescences (de Zeeuw et al. 1972a), which are the portion of the plant to be analysed for accurate quantitation of the cannabinoids content and type. Another intoxicant compound present in *C. sativa* is Δ^8-THC (a THC isomer) much less abundant and potent than Δ^9-THC (Small and Marcus 2003).

Both THC and Δ^8-THC are part of a class of secondary products (cannabinoids) unique to *C. sativa*. Cannabinoids are terpenophenolic substances, composed of a terpenic portion (geranylgeraniol in most cases) condensed with a phenolic moiety, most commonly the olivetolic acid (5-*n*-pentylresorcinol). Besides the common pentyl-cannabinoids, also methyl- (Shoyama et al. 1984), propyl- (de Zeeuw et al. 1972b) and even butyl-cannabinoids (Smith 1997) have been described in extracts of Cannabis samples; this variability in the length of the side chain of the phenolic moiety is certainly genetically determined, though little data are available on the genes controlling it (de Meijer et al. 2003). It is today accepted that cannabigerolic acid (CBGA, CBG in the neutral-form notation) is the first cannabinoid to be synthetised by condensation of the terpenic and phenolic moieties, and it is the precursor of the main cannabinoids found in hemp (Fellermeier et al. 2001). The enzymes involved in this condensation reaction, and in the following synthesis of CBD, CBC and THC, have been characterised from the biochemical point of view, and some of the coding sequences of the genes were obtained (Taura et al. 1995, 1996; Morimoto et al. 1998; Sirikantaramas et al. 2004).

Cannabinoids are accumulated and secreted, by resin exuding, mainly in the glandular trichomes present on the aerial portion of the plant (Dayanandan and Kaufman 1976; Mahlberg and Kim 2004). These organs are concentrated on the floral bracts, anthers and, at a lower density, on the leaves (Pate 1994); they are completely absent in the roots, seeds and cell cultures, where no cannabinoids are indeed detectable (Mandolino and Ranalli 1999).

There are over 60 cannabinoids in Cannabis species (de Zeeuw et al 1972a): CBD, THC, CBG, Δ^8-THC, cannabichromene (CBC) and cannabinol (CBN), are some of the most common. Many of them became the subject of intense research concerning their potential or actual pharmacological properties (Mechoulam 2000; Guy and Stott 2004; Pertwee 2004).

Cannabis sativa exists in different chemical variants, showing chemical but sometimes also morphological differences, known as chemotypes (Small and Beckstead 1973). Three different principal chemotypes were first identified. The first one was defined as the 'drug type' (chemotype I), because its low CBD/THC content ratio, especially due to high THC content. The second chemotype, 'intermediate', has both the two main cannabinoids, THC and CBD, in a content ratio close to the unity (typically ranging from 0.5 to 3.0; chemotype II), but usually with a slight prevalence of CBD; the third one, the so-called 'fibre', or 'non-drug' type (chemotype III), has mainly CBD, an amount of THC lower than 0.3%, and therefore a high CBD/THC ratio, sometimes not calculable due to undetectability of THC (de Meijer et al. 1992). It has been recently demonstrated that all the three main chemotypes can arise simply by segregation at one locus (*B*) within individual F2 progenies of divergent-chemotype parentals (Mandolino et al. 2003; de Meijer et al. 2003). Today a widely accepted view of the inheritance of these three chemotypes, is based upon the occurrence, at *B* locus, of two codominant alleles, B_D and B_T, responsible for the presence of CBD and THC, respectively (de Meijer et al. 2003; Mandolino 2004).

Two tentative additional chemotypes were occasionally reported. Chemotype IV had CBG as the predominant cannabinoid, but also CBD was present (Fournier et al. 1987); chemotype V was proposed by Mandolino and Carboni (2004) to classify material with undetectable amounts of any cannabinoids, qualitatively described by some authors (Virovets 1998). The inheritance mode of these additional chemotypes cannot be explained in terms of the two alleles postulated for the first three chemotypes, and required completion of the current genetic model with further elements (de Meijer and Hammond 2005, and the present paper).

Marker-assisted selection (MAS) in hemp produced, in recent years, molecular tools useful for the breeding and for forensic applications (reviewed in Mandolino and Carboni 2004). Today, tightly associated markers are available for early identification of the male sex (Mandolino et al. 1999, 2002); sequence characterised amplified region (SCAR) markers associated with the first three chemotypes were developed, by RAPD screening of F2 groups segregating for chemotype

(bulk segregant analysis), followed by isolation and sequencing of the markers identified, and design of specific primers (Mandolino et al. 2003; de Meijer et al. 2003).

In this work, a number of *C. sativa* accessions were surveyed for their cannabinoid type, content and ratio. The chemical phenotypes of the plants were examined in relation to the current genetic model, and an F1 group of plants with chemotype IV was obtained. Besides, a number of plants were checked for the maintenance of association of the chemotype with the previously developed SCAR marker. This marker is compared with another one, developed on the basis of the published gene sequences of B_D and B_T alleles (entries E55107 and E33090 in the NCBI gene bank). The effectiveness in chemotype identification of these markers is discussed in relation to their utilisation in hemp-breeding strategies for the development of new varieties, and to forensic and legislation issues.

Materials and methods

Three hundred and twelve Cannabis plants belonging to 12 hemp varieties and accessions of different origin were grown in pots containing a 1:1 mixture of sand and peat, in a greenhouse under natural spring–summer photoperiod. Because some of the plant material used were presumed to have a high THC content, it was necessary to raise all the plants in greenhouse rather than in open field conditions. The temperature ranged from about 24 °C at the beginning of the life cycle (April) up to 32 °C, when the plants flowered and the gas chromatographic analyses were carried out (September).

The hemp accessions used were the following: Carmagnola, an Italian dioecious fibre variety, and CS, a selection deriving from it; Fibranova, an Italian dioecious cross-bred cultivar with high fibre content; Fibrimon 56 and Epsilon 68, two French monoecious cultivars; Dioica 88, a French dioecious cultivar; USO31, an Ukrainian monoecious variety; Panorama, an Hungarian dioecious and dwarf variety (Bócsa 1994); (21R×15R)×NL1, an experimental hybrid obtained in our Institute by the cross between a drug strain, Northern Lights, and the fibre cv. Fibranova; the breeding line Bernabeo, under selection at ISCI, and an F1 Bernabeo obtained from intercross of 5 plants (3 males and 2 females); a group of plants from a seed sample seized by the Italian Police in 1996 and conferred to ISCI for the gas chromatographic analysis necessary to determine the liceity of the possession of such seeds. Italian varieties belong to a 1998–2001 seed lot deposited at ISCI gene bank; the French varieties were obtained in 1999 from O. Beherec (Fédération Nationale des Producteurs de Chanvre, Le Mans, France); seeds of the Ukrainian variety USO31 was kindly given in 2001 by Dr. V.G. Virovets (Institute of Bast Crops, Glukhov, Ukraine); Panorama seeds were obtained in 2000 from Dr. L. Frese, Bundesanstalt für Züchtungsforschung an Kulturpflanzen, Quedlinburg, Germany); Bernabeo breeding line was kindly provided in 2003 by Dr. M. Di Candilo, ISCI, Bologna, Italy. Cannabinoid profiles published in the present paper refer to plants grown from seed samples specific of the reported year and provenance.

In order to extract and quantify cannabinoids, the official method for *C. sativa* sampling for chemical analysis (Official Journal of the European Union, December 22nd, 2002) was carried out, taking 30 cm of the upper flowering part, when staminate and carpellate inflorescence were completely developed, but before the seed formation. The excised inflorescences were dried in an oven at 65 °C for 48 h, and crushed and sieved to remove the stems and the petioles. The powdered samples were stored in glass jars in the dark at 5 °C until the moment of analysis. The extraction solvent (hexane), the internal standard (squalene) and the reference standard (squalane) were added to 100 mg of sample powder, and the mixture treated as in de Meijer et al. (2003).

GC analyses were carried out with a gas chromatograph 8000 Serie Top (ThermoQuest Italia) equipped with autosampler and flame ionisation detector. The column used was a 30 m×0.32 mm fused silica capillary column (Labocest, Italy), 0.5 μm thick. The oven temperature varied between 220 °C and 300 °C, and the injector and detector temperatures were 300 °C. The pressure was fixed for carrier gas (He) at 59 KPa and for make up gas at 100 Kpa (air) and 50 Kpa (H). In pressure mode the flow rates were variable. Compound identities were initially determined by GC-MS, and by comparing their retention times (RT) with those of commercial CBD and THC, used as standards. For quantitative determination of each

compound, a calibration was carried out using cannabinoid standards (CBD and THC) at various concentrations added to hexane, to squalene and to squalane. Chrom-Card for Windows (version 1.21, CE Instruments, Italy) determined a linear calibration equation using a fixed value of Response Factor (RF) for squalene (RF = 1) and worked out the RFs for any cannabinoid of interest. RF for CBG was estimated according to de Meijer and Hammond (2005). These RFs allowed to convert GC-derived peak areas to dry weight concentrations; cannabinoid amounts were calculated as mg/g of dry weight, taking into account the weight of the powder (100 mg), the extraction volume (5 ml), and the average residual humidity of the sample after drying (5%).

One hundred and fortyeight plants belonging to the 12 accessions were screened with *B190/B200* (de Meijer et al. 2003) and *B1080/B1192* SCAR markers, in order to verify their degree of association with the chemotype. A small amount of tissue was picked up, and genomic DNA was extracted using the Nucleon Phytopure kit (Amersham Pharmacia Biotec, U.K.). DNA concentrations were determined by 260 nm readings. DNA was amplified using the two 20-mer primers described elsewhere (de Meijer et al. 2003), producing one 190-bp and/or a 200-bp fragment. PCR reactions were performed in a PCR Express thermal cycler (Hybaid, U.K.) using the conditions reported elsewhere (de Meijer et al. 2003). The same DNA samples were also amplified with a multiplex system, developed in our Institute on the basis of the gene sequences of THC- and CBD-synthases deposited in NCBI gene bank (E55108 and E33091; Sirikantaramas et al. 2004). This primer combination is designed to amplify segments of the THC and CBD synthase gene sequences. The amplicons obtained by both the marker systems tested were run on a 3% (*B190/B200*) and 1% (*B1080/B1192*) agarose gel, in 1× TAE buffer, stained with ethidium bromide, and visualised by UV radiation. For more informations on the sequences and conditions of the multiplex marker system, inquire to GM.

Results

Cannabinoids content in different Cannabis *accessions*

The results of GC analysis are reported in Table 1, as average % of inflorescence dry weight, for each of the varieties or lines considered. The two cannabinoids, THC and CBD, accounted together from a minimum of 77% up to close to 100% of the total detectable cannabinoids, confirming they are the two most common and abundant cannabinoids in Cannabis germplasm (with the exception of a few Bernabeo plants, see below).

The mean THC contents were significantly higher in the seized material (2.08%), in (21R×15R)×NL (0.80%) and in Panorama (0.49%) than in all other plants; the threshold level

Table 1. Content of the main cannabinoids found in Cannabis accessions, number of plants exceeding the 0.20% THC threshold, and CBD/THC ratios averaged over the entire accession. See Introduction for chemotype definitions.

Accession	Sample size	Av. CBD (% d.w.) ± s.d.	Av. THC (% d.w.) ± s.d.	Plants with THC > 0.2% d.w.	Av. CBG (% d.w.) ± s.d.	Av. CBD/THC[a] ± s.d.	Total cannabinoid (% d.w.) ± s.d.
Carmagnola	43	0.85 ± 0.43	0.09 ± 0.17	4	0.02 ± 0.04	17.09 ± 6.48	0.98 ± 0.49
C.S.	4	1.09 ± 0.72	0.14 ± 0.17	1	0.04 ± 0.08	16.85 ± 11.59	1.27 ± 0.73
Fibranova	10	0.43 ± 0.19	0.02 ± 0.03	0	0.01 ± 0.01	33.20 ± 16.80	0.46 ± 0.18
Fibrimon 56	7	1.28 ± 0.38	0.04 ± 0.01	0	0.05 ± 0.04	30.23 ± 1.50	1.37 ± 0.42
Epsilon 68	35	0.51 ± 0.42	0.02 ± 0.02	0	0.00	18.92 ± 3.90	0.53 ± 0.44
Dioica 88	35	1.04 ± 0.48	0.04 ± 0.02	0	0.06 ± 0.09	25.59 ± 7.54	1.15 ± 0.53
USO31	48	0.16 ± 0.23	0.01 ± 0.01	0	0.01 ± 0.03	22.93 ± 10.13	0.18 ± 0.27
Seized material	39	0.28 ± 0.18	2.08 ± 0.72	39	0.10 ± 0.09	0.14 ± 0.07	2.46 ± 0.84
Panorama	3	0.86 ± 0.46	0.49 ± 0.31	3	0.09 ± 0.08	1.84 ± 0.46	1.44 ± 0.83
(21R×15R)×NL	45	1.44 ± 0.86	0.80 ± 0.49	45	0.10 ± 0.14	1.82 ± 0.30	2.36 ± 1.43
Bernabeo	9	0.14 ± 0.15	0.003	0	0.09 ± 0.09	25.75 ± 3.18	0.23 ± 0.11
F1 Bernabeo	34	0.21 ± 0.19	0.00	0	0.53 ± 0.12	n.c.	0.74 ± 0.21

[a]Calculated on the basis of the plants showing detectable amounts of THC.

admitted by EU for granting subsidies, 0.20% THC, was exceeded for each and every of the 87 single plants analysed for these three groups (Table 1). However, while in the seized material CBD content (and therefore CBD/THC ratio) was found to be low, as expected from chemotype I plants, the hybrid (21R×15R)×NL and the cv. Panorama had comparable contents of THC and CBD, leading to ratios of 1.82 and 1.84, respectively, typical of chemotype II plants. In the other accessions, average THC content ranged from less than 0.01% (Bernabeo and USO31) to 0.14% (CS), and only in Italian cvs. Carmagnola, Fibranova and CS, a few plants (4 out of 43, 1 out of 10 and 1 out of 4, respectively; Table 1) showed THC concentrations high enough to give CBD/THC ratios typical of chemotype II; however, among these, only the 5 Carmagnola and CS plants showed THC absolute values exceeding 0.20%. In these particular plants, as in all Panorama and (21R×15R)×NL plants, significant levels of the non-psychoactive cannabinoid CBD were also present, and CBD/THC ratios close to 1 were found (Table 2); therefore, Carmagnola, Fibranova and CS were found still quite heterogeneous as far chemotype is concerned, with most plants belonging to chemotype III but some to chemotype II (Table 2).

CBD contents varied from an average of 0.14% in the breeding line Bernabeo, up to 1.44% in the (21R×15R)×NL hybrid; the standard deviations, however, were much higher for this cannabinoid than for THC average levels (Table 1).

Small amounts of CBG were detected in all accessions (except the French fibre cv. Epsilon 68), with average contents ranging from 0.01 up to 0.10% of inflorescence dry weight. However, individual plants reached values as high as 0.57%, but with a share on the total cannabinoids never exceeding 23%. The breeding line Bernabeo was the only exception (Table 1); it had a mean CBG content of 0.09 ± 0.09%, but when the data for the single plants belonging to this line were examined, it was observed that, out of nine plants analysed, four had no detectable amounts of CBG, while the remaining 5 had an average content of 0.17%, the value of 0.09% being therefore the resultant of these two types of plants. These five peculiar plants had a CBG share on the total amount of cannabinoids ranging from 71 to 84% (data not shown). Therefore, the Bernabeo line was composed of plants that can be assigned to different chemotypes (III and IV; Table 2). Interestingly, though all Bernabeo plants contained CBD, the five CBG-containing plants had a low CBD content (average 0.05%), while the four plants devoid of any detectable CBG had much higher CBD amounts (average 0.26%); as a final result, the

Table 2. Chemotypes and CBD/THC ratios found within each of the accessions tested.

Accession	No. plants	Chemotype	CBD/THC	CBG/CBD
Carmagnola	39	III	18.96 ± 3.60	0.03 ± 0.05
	4	II	1.21 ± 0.18	
Fibranova	9	III	37.13 ± 12.93	0.02 ± 0.02
	1	II	2.10	
CS	3	III	22.01 ± 6.46	0.02 ± 0.04
	1	II	1.37	
Fibrimon 56	7	III	30.23 ± 1.50	0.04 ± 0.02
Epsilon 68	32	III	18.92 ± 3.90	0.01 ± 0.02
	3	V	n.d.[a]	
Dioica 88	35	III	25.59 ± 7.54	0.06 ± 0.07
USO31	32	III	22.93 ± 10.13	0.02 ± 0.04
	16	V	n.d.[a]	
Seized material	39	I	0.14 ± 0.07	0.46 ± 0.42
Panorama	3	II	1.84 ± 0.46	0.08 ± 0.07
(21R×15R)×NL	45	II	1.82 ± 0.30	0.06 ± 0.06
Bernabeo	4	III	25.75 ± 3.18	0.00
	5	IV	n.d.[a]	3.69 ± 1.09
F1 Bernabeo	34	IV	n.d.[a]	5.31 ± 3.96

[a]Not determined, due to undetectable THC amount.

two type of plants within Bernabeo had similar amounts of total cannabinoids.

Another cannabinoid present at very small levels by GC analysis was Δ^8-THC. It was detected only in Carmagnola (average 0.01%) and (21R×15R)×NL (average 0.02%; data not shown); this cannabinoid never exceeded 1% of the total amount of cannabinoids present in the plant.

The monoecious Ukrainian cultivar USO31 showed very low average CBD, THC and total cannabinoid contents (Table 1); indeed, 16 out of 48 plants (33%) had no detectable cannabinoids, confirming previous reports from USO31 breeders (Virovets, personal communication) of the existence of zero-cannabinoids plants within this variety. When the data relative to these plants were ignored, the average CBD content of the remaining USO31 plants raised from 0.16 to 0.53%. The same holded for 3 out of 35 Epsilon 68 plants; these plants are assigned therefore to chemotype V in Table 2.

The total amount of cannabinoids, showed in Table 1, was found to be quite different among the varieties and within them. The lowest values were found for USO31 and Bernabeo; the variability in the overall cannabinoid content was very high for these materials. The groups with the highest cannabinoid content also had the highest THC content, i.e. the seized material, the (21R×15R)×NL hybrid, and Panorama variety. Comparing the individual plants belonging to these groups, variation was found very high too, around 50% of the content (Table 1). The average cannabinoid content of Italian (Carmagnola, Fibranova and CS) and French (Fibrimon 56, Epsilon 68 and Dioica 88) varieties, all belonging to the EU list of cvs. for which subsidies are issued, was very similar (0.90 and 1.01%, respectively). However, French varieties had a lower THC content (on average, 0.03%) than Italian ones (0.08%). Both groups were well below the threshold of 0.20% (averaged), though only in Italian varieties some individual plants with THC above 0.20% were found.

The THC amount of the seized material was 2.08% ± 0.72 (Table 1). These values places this material among illicit one, and the same holds true for Panorama cv., at least for the seed lot examined, and for the experimental hybrid (21R×15R)×NL, with contents of 0.49 ± 0.31 and 0.80 ± 0.49. Growing these plants would be considered illicit by most European legislations.

Chemotype distribution

The data relative to all plants are shown in Figure 1a, b. The eight accessions with a cannabinoid content mostly determined by CBD and THC, distribute in a tripartite way in CBD vs. THC scatter plots (Figure 1a). The plants with CBG-predominant or undetectable cannabinoid content, are described by a CBG vs. CBD + THC scatter plot in Figure 1b, where chemotype IV plants were

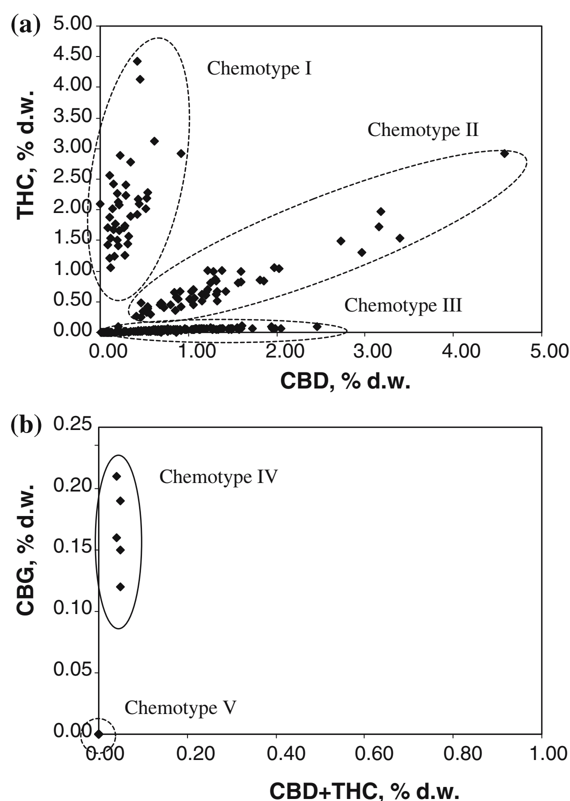

Figure 1. Chemotype distributions in the Cannabis germplasm analysed. Each point represents the analytical GC data for a single plant. (a) the tripartite distribution of the THC and CBD contents. Chemotype I cluster is made of the seized plants; chemotype II cluster gathers the (21R×15R)×NL hybrids, Panorama plants and a few plants from Carmagnola, CS and Fibranova. The chemotype III cluster is made of plants from the French and Italian varieties and from about 2/3 of USO31 and some Bernabeo plants. (b) distribution of CBG vs. CBD + THC content. Chemotype IV cluster is made of the Bernabeo plants predominant in CBG; in chemotype V cluster are gathered all the zero-cannabinoids plants from USO31 and Epsilon 68 cvs. Data from F1 Bernabeo are not shown in this graph.

distributed along the Y-axis, while chemotype V plants were gathered in a single point on the origin of any scatter plot.

Chemotype of CBG-containing F1 plants

When the five CBG-containing Bernabeo plants were intercrossed to generate an F1, the chemotype IV resulted fixed (Table 2), as all 34 plants examined showed prevalence of CBG and smaller CBD peaks, a typical chemotype IV feature. The average cannabinoid content of the F1 plants obtained from the 5 CBG-containing Bernabeo plants are shown in the last row of Table 1, for comparison with its Bernabeo parental group. The CBD/THC ratio found for these latter plants was 25.75 (Table 1), though this actually was an average of only the few plants showing some THC traces, and for which this ratio could be calculated, the other plants and the entire F1 showing no THC (Table 2).

Markers associated to chemotypes

DNA extracted from leaf samples of 148 plants chosen among the 312 analysed by GC, was amplified using the 20-mer primers yielding the SCAR marker *B190/B200* described elsewhere (de Meijer et al. 2003); the expected DNA fragments of either 190 or 200 bp were obtained (Figure 2a). However, only in 1 out of 13 chemotype II plants, both DNA fragments were detected, contrarily to previous observations in segregating F2s (Mandolino et al. 2003; de Meijer et al. 2003). In the present analyses, the marker was indeed polymorphic, as it produced either the 190 or the 200 bp fragment, but unable to discriminate heterozygous, chemotype II plants. The degree of association of this marker with the chemotype, determined gas chromatographically for each plant, is shown in Table 3. The overall association was 93% for correct identification of chemotype III, but only 20% for the identification of chemotype I plants.

For comparison, a multiplex PCR assay was developed; the three primers mediating the amplification reaction were designed on the basis of the gene sequences of CBD- and THC-synthases

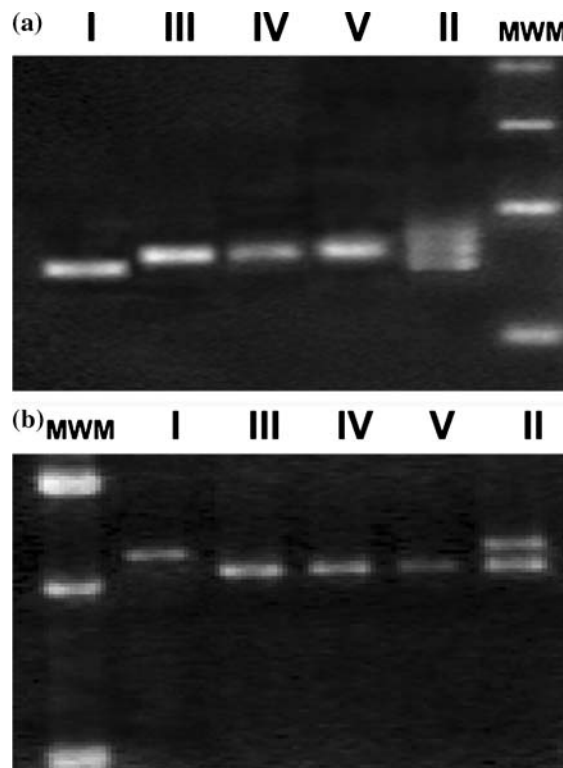

Figure 2. Amplification of Cannabis DNA by *B190/B200* (a) and *B1080/1192* (b) markers. MWM, molecular weight markers (panel a: 123 bp ladder; panel b: 1 kb, Invitrogen). I, II, III, IV, and V indicate amplicons from the different chemotypes.

deposited in GeneBank. These primers produced the fragments of the expected size, 1080 and 1190 bp approximatively (Figure 2b). The association of such markers was complete for the three main chemotypes (100% correct identification; Table 3), and the marker was codominant, being able to fully typify the DNA at the B locus. However, all chemotype IV and V plants were typified as endowed with the DNA fragment associated to the B_D allele, i.e. they were not distinguished from chemotype III (CBD) plants (Figure 2b; Table 3).

Discussion

The data presented here support the introduction of new chemotypes, here termed chemotype IV (prevalent CBG) and V (undetectable cannabinoids), as already proposed elsewhere (Mandolino and Carboni 2004). The frequency of the plants

Table 3. Association between the number of plants belonging to the different chemotypes as determined by GC analysis, and the occurrence of *B190*/*B200* and *B1080*/*B1192* SCAR markers.

Accession	No. plants	Chemotype	SCAR *B190*/*B200*			SCAR *B1080*/*B1192*		
			B190	*B200*	*B190*/*B200*	*B1080*	*B1192*	*B1080*/*B1192*
Carmagnola	26	III	–	26	–	26	–	–
	2	II	–	1	1	–	–	2
Fibranova	9	III	1	3	5	9	–	–
	1	II	1	–	–	–	–	1
CS	3	III	–	3	–	3	–	–
	1	II	–	1	–	–	–	1
Fibrimon 56	7	III	–	7	–	7	–	–
Epsilon 68	12	III	–	12	–	12	–	–
Dioica 88	9	III	–	9	–	9	–	–
USO31	15	III	–	15	–	15	–	–
	5	V	–	5	–	5	–	–
Seized material	5	I	1	4	–	–	5	–
Panorama	3	II	–	3	–	–	–	3
(21R×15R)×NL	6	II	–	6	–	–	–	6
Bernabeo	4	III	–	4	–	4	–	–
	5	IV	–	5	–	5	–	–
F1 Bernabeo	35	IV	–	35	–	35	–	–

belonging to these chemotypes is low but not negligible, and tends to increase within the populations subjected to extensive inbreeding work to fix specific traits, such as monoecy (USO31) or lack of THC (Bernabeo, Epsilon 68).

According to previous studies (Small and Beckstead 1973; de Meijer et al. 1992; Mandolino et al. 2003), the CBD/THC content ratio is a discriminant between the different chemotypes. However, the comparison of different varieties solely on the basis of their average CBD/THC values, as listed in Table 1, might mask the existence of different chemotypes within each accession. When the individual plants were examined, it became apparent that not all the materials analysed were homogeneous for the CBD/THC ratio, but some included distinct groups, with markedly different CBD/THC ratios (Table 2). When the data relative to the different chemotypes within each variety, line or strain are considered, the standard deviation of the CBD/THC ratio values considerably lowered for Fibranova, Carmagnola, and CS (Table 2). In this way, it is possible to classify all the plants examined in five different chemotypes, which appear to cover all those described up to now in the literature. The only exception is the CBC-prevalent chemotype, never found to be the main cannabinoid in any plant, but only as a minor component (Holley et al. 1975). However, specific breeding programs recently achieved in some materials proportions of this cannabinoid up to 50% (E. de Meijer, personal communication); in this survey, no data about any other cannabinoids present in significant amounts were obtained besides those described, and the integration of a putative allele specific for CBC biosynthesis in the current genetic model, originally postulated in de Meijer et al. (2003), remains a pure speculation.

The observed discrete distribution of the chemotypes (Figure 1a, b) matched closely the tripartite distribution of THC vs. CBD scatter plots observed in Cannabis segregant F2s in de Meijer et al. (2003); this fact supports the idea that this distribution actually reflects that of the B_D and B_T alleles. The present work, therefore, by introducing evidences confirming the existence of a new allele at the *B* locus (B_0, responsible for CBG chemotype), already postulated by de Meijer et al. (2003), and revealed by direct genetic analysis by de Meijer and Hammond (2005), extends the validity of the model previously described (de Meijer et al. 2003), and confirms the interpretation in terms of allele frequencies of the chemotypes' distribution in Cannabis germplasm.

The ratios between the different cannabinoids present in a Cannabis plant, reflect the relative

efficiency with which the different cannabinoid synthases convert the common precursor, CBG, to either THC or CBD or other minor cannabinoids. Because the biochemical properties of THC- and CBD-synthases are extremely similar (Taura et al. 1995, 1996), heterozygous B_D/B_T plants, supposed to possess both synthases, have CBD/THC ratios around the unity, while homozygous B_D or B_T plants have ratios strongly shifted towards very high or exceedingly low values, sometimes not calculable (e.g. undetectable THC; Tables 1 and 2). These predictions are completely supported by the data presented here. The genotype at the *B* locus was identified by using a SCAR marker designed on the basis of the sequences of CBD- and THC-synthase deposited in GeneBank, and it was always in agreement with the observed gas chromatographic profile, with some remarkable exceptions: chemotype V plants, characterised by a flat chromatogram, and belonging to USO31 and Epsilon 68 varieties, showed a marker pattern, using both the marker systems (*B190*/*B200* and the *B1080*/*B1192*) not distinguishable from chemotype III plants (Table 3); also chemotype IV plants (belonging to Bernabeo line and their F1), when screened by both marker types, could not be discriminated from chemotype III plants (Figure 2a, b).

CBG-prevalent plants should be defined, rather than by their CBD/THC ratio, by high CBG/CBD values (Figure 1b and Table 2). The chemotype IV was confirmed to occur in ‘fibre’ germplasm, having a very high frequency of B_D alleles (Fournier et al. 1987; de Meijer and Hammond 2005). The data presented demonstrate that chemotype IV plants belonging to Bernabeo line, bred true, i.e. their F1 offspring was composed exclusively of CBG-prevalent plants (Table 2). In all chemotype IV plants (both parental or F1), an inverse relationships between CBG and CBD was observed (see Results); such relationship was also observed by other authors in crosses involving CBG plants (de Meijer and Hammond 2005). Finally, the markers designed on the gene sequence of CBD synthase failed to distinguish chemotype III and chemotype IV plants, a further hint supporting the idea that a new allele, called B_0, deriving from B_D allele and coding for a largely non-functional CBD-synthase, and with consistent sequence homology with it, is present in CBG-prevalent plants at the homozygous state. This hypothesis is currently the simplest able to explain the data. However, it needs to be confirmed by sequencing this new proposed allele, and characterising the biochemical properties of the relative synthase.

As for chemotype V, the lack of any detectable cannabinoid in one third of USO31 and in a minority of Epsilon 68 plants, suggests a metabolic block, possibly involving a gene upstream of the *B* locus, preventing the formation of the precursor CBG. However, it is also possible that chemotype V plants were counter-selected for some quantitative trait indirectly influencing cannabinoid’s amount, e.g. for a particularly low trichome density. Whatever the reason of the undetectability of cannabinoids in chemotype V plants, the cannabinoids produced would go undetected under standard analytical conditions, irrespective of the alleles at the locus *B* the plants would be endowed with; the marker analysis would not therefore be able to predict the chemotype of these plants. Currently, it is not possible to distinguish between the different explanations for this particular chemical phenotype. Further investigations will require a functional genomic approach, to establish whether the absence of cannabinoids is due to transcriptional, translational or organogenetic blocks.

To identify the chemotypes and for validation of the markers available, two molecular marker systems were used in the present work: a marker (*B190*/*B200*), developed by bulk segregant analysis (de Meijer et al. 2003), and a marker (*B1080*/*B1192*), based on the sequences of CBD- and THC-synthases. This latter is a multiplex system, based upon the use of one primer common to both sequences, and two specific primers, designed on the variant regions of the gene sequences. The *B190*/*B200* marker revealed only limited polymorphism in the materials tested, and it was not associated to chemotype as tightly as it was in the pedigrees it was originally developed for (de Meijer et al. 2003); besides, often only one single DNA fragment was amplified, and consequently only occasionally it was possible to identify the heterozygous plants. Therefore, *B190*/*B200* it is not very suitable for a mass screening during improvement or forensic identification of chemotypes in *C. sativa* in the materials tested, though a high capacity of identification of the ‘fibre’ chemotype has been confirmed.

DNA amplification mediated by *B1080*/*B1192*, yielded bands that are in perfect association with

the chemotype, showing 100% of correct identification of the three main chemotypes. Chemotype III and chemotype I were found associated with two bands of different molecular weight (1080 and 1192 bp, respectively), and the hybrid chemotype II is identified by the simultaneous presence of both bands (Figure 2b). Such marker is therefore very effective in ruling out the presence of the B_T allele in the varieties under selection or examination, and can therefore be fruitfully employed in both the breeding programs aimed to eradicate THC-containing plants from a breeding line, and to identify illegal crops in the forensic field; however, as already stated above, CBG- and zero-cannabinoid plants cannot be identified by this marker system.

Today, the upper THC content limit for accessing to EU subsidies and for the cultivation to be considered licit is (quite arbitrarily) fixed at 0.20% of the plant inflorescence's dry weight.

The present survey of different fibre cultivars has shown that they are all, on average, below such limit, at least under the growth conditions chosen. In these conditions, the response of the *B1080/B1192* SCAR marker is adequate also for the screening of plants with THC levels below the threshold; when classified by the marker as B_D/B_D, the THC amounts of a Cannabis plant apparently cannot overcome this quantitative threshold, at least in the environmental conditions used. The marker can therefore also be tentatively considered as a 'threshold' indicator, suitable for checking the quantitative limits for cultivation as well. However, it should be pointed out that some variations, and occasional exceeding the 0.20% d.w. threshold, was observed by growing the same Cannabis genotype under different environmental or agronomical conditions (de Meijer et al. 1992; Bócsa et al. 1997; Small and Marcus 2003; for an in-depth discussion of the quantitative and qualitative factors influencing the chemotype in Cannabis, see Hillig 2002; de Meijer et al. 2003; Mandolino 2004).

There are no specific reports that the chemotype can change in different growth stages or plant organs and tissues, with the only possible exception of CBC, reported to be a juvenile stage cannabinoid (Vogelmann et al. 1988). However, some authors reported that the fluctuations in the quantitative levels of the different cannabinoids can, to some extent, reflect on the cannabinoid content ratios (Hemphill et al. 1980), depending on the growth stages, the sex, and even the portion of the plant analysed. However, it is our opinion that these variations are not of such an entity to induce a shift in chemotype, and that this latter is correctly scored also by determining the CBD/THC ratio of young leaves at the vegetative stage, rather than in the mature inflorescences. The recognition of this uniformity throughout the development is at the basis of proposals to simplify and speed up the GC determinations, mandatory for the farmers to get the EU subsidies, by analysing a limited number of leaves before flowering and determining their CBD/THC ratio rather than the THC content alone.

Variation within hemp varieties is high, as expected by an obligate outbreeder (Forapani et al. 2001), and cannabinoids content is no exception. In our analyses, 6 plants out of 57 belonging to the fibre germplasm, were found to have a CBD/THC ratio around the unity, which places them within chemotype II; in all these cases, except one, the 0.20% d.w. THC content threshold was also exceeded. Therefore, the use of the CBD/THC ratio threshold instead of the total THC amount may be even too conservative, bringing to the rejection of plants that have very low THC contents (0.10% of d.w. in the case of the single Fibranova plant with a 2.1 CBD/THC ratio). Recently, Hillig and Mahlberg (2004) reported that 20% of the cultivars they analysed exceeded the 0.30% d.w. limit fixed in USA for licit commercialisation of hemp products, and suggested that an upper limit of 0.80% d.w. would be more indicative to distinguish chemotypes. There is currently a great debate about the limits for THC in cultivated hemp; the use of markers for the genes involved in the different steps of cannabinoid biosynthesis like the SCAR marker described in this paper should make easier and safer the correct identification of the plants to be eliminated from multiplication stands, allowing in a limited time significant progresses in the reduction of frequency of undesired alleles, and a consequent lowering of the amount of the cannabinoids synthetised.

References

Arru L., Rognoni S., Baroncini M., Medeghini Bonatti P. and Perata P. 2004. Copper localization in *Cannabis sativa* L. grown in a copper-rich solution. Euphytica 140: 33–38.

Bócsa I. 1994. Interview. J. Int. Hemp Assoc. 1(2): 61–63.
Bócsa I., Mathé P. and Hangyel L. 1997. Effect of nitrogen on tetrahydrocannabinol (THC) content in hemp (*Cannabis sativa* L.) leaves at different positions. J. Int. Hemp Assoc. 4: 80–81.
Callaway J. 2004. Hempseed as a nutritional resource. An overview. Euphytica 140: 65–72.
Dayanandan P. and Kaufman P.B. 1976. Trichomes of *Cannabis sativa* L. (Cannabaceae). Am. J. Bot. 63: 578–591.
Fellermeier M., Eisenreich W., Bacher A. and Zenk M.H. 2001. Biosynthesis of cannabinoids. Incorporation experiments with ^{13}C-labeled glucoses. Eur. J. Biochem. 268: 1596–1604.
Forapani S., Carboni A., Paoletti C., Moliterni V.M.C., Ranalli P. and Mandolino G. 2001. Comparison of hemp (*Cannabis sativa* L.) varieties using RAPD markers. Crop Sci. 41: 1682–1689.
Fournier G., Richez-Dumanois C., Duvezin J., Mathieu J.-P. and Paris M. 1987. Identification of a new chemotype in *Cannabis sativa*: cannabigerol-dominant plants, biogenetic and agronomic prospects. Plant. Med. 53: 277–280.
Guy G.W. and Stott C.G. 2004. Cannabinoids for the pharmaceutical industry. Euphytica 140: 83–93.
Hemphill J.K., Turner J.C. and Mahlberg P.G. 1980. Cannabinoid content of individual plant organs from different geographical strains of *Cannabis sativa* L. J. Nat. Prod. 43: 112–122.
Hillig K. 2002. Letter to the Editor. J. Industrial Hemp 7: 5–6.
Hillig K. 2005. Genetic evidence for speciation in Cannabis (Cannabaceae). Genet. Res. Crop Evol. 52: 161–180.
Hillig K.W. and Mahlberg P.G. 2004. A chemotaxonomical analysis of cannabinoid variation in Cannabis (Cannabaceae). Am. J. Bot. 91: 966–975.
Holley J.H., Hadley K.W. and Turner C.E. 1975. Constituents of *Cannabis sativa* L. XI Cannabidiol and cannabichromene in samples of known geographical origin. J. Pharm. Sci. 64: 892–894.
Karus M. and Vogt D. 2004. European hemp industry: cultivation, processing and product lines. Euphytica 140: 7–12.
Kriese U., Schumann E., Weber W.E., Beyer M., Brühl L. and Matthäus B. 2004. Oil content, tocopherol composition and fatty acid patterns of the seeds of 51 *Cannabis sativa* L. genotypes. Euphytica 137: 339–351.
Mahlberg P.G. and Kim E.S. 2004. Accumulation of cannabinoids in glandular trichomes of Cannabis (Cannabaceae). J. Industrial Hemp 9: 15–36.
Mandolino G. 2004. Again on the nature of inheritance of chemotype. Letter to the editor. J. Industrial Hemp 9: 5–7.
Mandolino G., Bagatta M., Carboni A., Ranalli P. and de Meijer E.P.M. 2003. Qualitative and quantitative aspects of the inheritance of chemical phenotype in Cannabis. J. Industrial Hemp 8: 51–72.
Mandolino G. and Carboni A. 2004. Potential of marker assisted selection in hemp genetic improvement. Euphytica 140: 107–120.
Mandolino G., Carboni A., Bagatta M., Moliterni V.M.C. and Ranalli P. 2002. Occurrence and frequency of putatively Y chromosome linked DNA markers in *Cannabis sativa* L. Euphytica 126: 211–216.
Mandolino G., Carboni A., Forapani S., Faeti V. and Ranalli P. 1999. Identification of DNA markers linked to the male sex in dioecious hemp (*Cannabis sativa* L.). Theor. Appl. Genet. 98: 86–92.
Mandolino G. and Ranalli P. 1999. Advances in biotechnological approaches for hemp breeding and industry. In: Ranalli P. (ed.), Advances in Hemp Research. The Haworth Press, U.S.A, pp. 185–212.
Mechoulam R. 2000. Looking back at Cannabis research. Curr. Pharm. Des. 6: 1313–1322.
Meijer de E.P.M., van der Kamp H.J. and van Eeuwijk F.A. 1992. Characterization of Cannabis accessions with regard to cannabinoid content in relation to other plant characters. Euphytica 62: 187–200.
Meijer de E.P.M., Bagatta M., Carboni A., Crucitti P., Moliterni V.M.C., Ranalli P. and Mandolino G. 2003. The inheritance of chemical phenotype in *Cannabis sativa* L. Genetics 163: 335–346.
Meijer de E.P.M. and Hammond K.M. 2005. The inheritance of chemical phenotype in *Cannabis sativa* L. (II). Cannabigerol predominant plants. Euphytica 145: 189–198.
Morimoto S., Komatsu K., Taura F. and Shoyama Y. 1998. Purification and characterization of cannabichromenic acid synthase from *Cannabis sativa*. Phytochemistry 49: 1525–1529.
Pate D.W. 1994. Chemical ecology of Cannabis. J. Int. Hemp Assoc. 1: 29–37.
Pertwee R.G. 2004. Pharmacological and therapeutic targets for Δ^9-tetrahydrocannabinol and cannabidiol. Euphytica 140: 73–82.
Ranalli P., Di Candilo M., Mandolino G., Grassi G. and Carboni A. 1999. Hemp for sustainable agricultural systems. Agro-Food-Industry Hi-Tech. 2: 33–38.
Shoyama Y., Hirano H. and Nishioka I. 1984. Biosynthesis of propyl cannabinoid acid and its biosynthetic relationship with pentyl and methyl cannabinoid acids. Phytochemistry 23: 1909–1912.
Sirikantaramas S., Morimoto S., Shoyama Y., Ishikawa Y., Wada Y., Shoyama Y. and Taura F. 2004. The gene controlling marijuana psychoactivity. J. Biol. Chem. 279: 39767–39774.
Small E. and Beckstead H.D. 1973. Common cannabinoid phenotypes in 350 stocks of Cannabis. Lloydia 36: 144–165.
Small E., Jui P.Y. and Lefkovitch L.P. 1976. A numerical taxonomic analysis of Cannabis with special reference to species delimitation. Syst. Bot. 1: 67–84.
Small E. and Marcus D. 2003. Tetrahydrocannabinol levels in hemp (*Cannabis sativa*) germplasm resources. Econ. Bot. 57: 545–558.
Smith R.M. 1997. Identification of butyl cannabinoids in marijuana. J. Forensic Sci. 42: 610–618.
Taura F., Morimoto S. and Shoyama Y. 1995. First direct evidence for the mechanism of delta-1-tetrahydrocannabinolic acid biosynthesis. J. Am. Chem. Soc. 38: 9766–9767.
Taura F., Morimoto S. and Shoyama Y. 1996. Purification and characterization of cannabidiolic-acid synthase from *Cannabis sativa* L. J. Biol. Chem. 271: 17411–17416.
Virovets V.G. 1998. Interview. J. Int. Hemp Assoc. 5: 32–34.
Vogelmann A.F., Turner J.C. and Mahlberg P.G. 1988. Cannabinoid composition in seedlings compared to adult plants of *Cannabis sativa*. J. Nat. Prod. 51: 1075–1079.

Zeeuw de R.A., Malingre Th.M. and Merkus F.W.H.M. 1972a. Tetrahydrocannabinolic acid, an important component in the evaluation of *Cannabis* products. J. Pharm. Pharmacol. 24: 1–6.

Zeeuw de R.A., Wijsbek J., Breimer D.D., Vree T.B., van Ginneken C.A. and van Rossum J.M. 1972b. Cannabinoids with a propyl side chain in *Cannabis*. Occurrence and chromatographic behaviour. Science 175: 778–779.