The Inheritance of Chemical Phenotype in *Cannabis sativa* L.

Etienne P. M. de Meijer,*,[1] Manuela Bagatta,[†] Andrea Carboni,[†] Paola Crucitti,[†] V. M. Cristiana Moliterni,[†] Paolo Ranalli[†] and Giuseppe Mandolino[†,2]

**HortaPharm B.V., 1075 VS, Amsterdam, The Netherlands and [†]Istituto Sperimentale per le Colture Industriali, 40128 Bologna, Italy*

Manuscript received April 23, 2002
Accepted for publication October 16, 2002

ABSTRACT

Four crosses were made between inbred *Cannabis sativa* plants with pure cannabidiol (CBD) and pure Δ-9-tetrahydrocannabinol (THC) chemotypes. All the plants belonging to the F_1's were analyzed by gas chromatography for cannabinoid composition and constantly found to have a mixed CBD-THC chemotype. Ten individual F_1 plants were self-fertilized, and 10 inbred F_2 offspring were collected and analyzed. In all cases, a segregation of the three chemotypes (pure CBD, mixed CBD-THC, and pure THC) fitting a 1:2:1 proportion was observed. The CBD/THC ratio was found to be significantly progeny specific and transmitted from each F_1 to the F_2's derived from it. A model involving one locus, *B*, with two alleles, B_D and B_T, is proposed, with the two alleles being codominant. The mixed chemotypes are interpreted as due to the genotype B_D/B_T at the *B* locus, while the pure-chemotype plants are due to homozygosity at the *B* locus (either B_D/B_D or B_T/B_T). It is suggested that such codominance is due to the codification by the two alleles for different isoforms of the same synthase, having different specificity for the conversion of the common precursor cannabigerol into CBD or THC, respectively. The F_2 segregating groups were used in a bulk segregant analysis of the pooled DNAs for screening RAPD primers; three chemotype-associated markers are described, one of which has been transformed in a sequence-characterized amplified region (SCAR) marker and shows tight linkage to the chemotype and codominance.

CHEMOTYPICAL diversity in Cannabis: The class of secondary products unique to the dioecious species *Cannabis sativa* L. (hemp) is the terpenophenolic substances known as cannabinoids, which accumulate mainly in the glandular trichomes of the plant (Mechoulam 1970; Hammond and Mahlberg 1977). Over 60 cannabinoids are known (de Zeeuw *et al.* 1972a), the most abundant being cannabidiol (CBD) and Δ-9-tetrahydrocannabinol (THC). Other common cannabinoids are cannabichromene (CBC; Holley *et al.* 1975) and cannabigerol (CBG). These cannabinoids have a pentyl side chain, but also their propyl homologs occur, which are indicated as cannabidivarin (CBDV), Δ-9-tetrahydrocannabivarin (THCV), cannabichromevarin (CBCV), and cannabigerovarin (CBGV), respectively (de Zeeuw *et al.* 1972b).

Small and Beckstead (1973) were the first to systematically survey a wide number of Cannabis accessions for variability in cannabinoid composition. They recognized, on a population mean basis, three chemical phenotypes (chemotypes): chemotype I, with a THC content >0.3% and a CBD content <0.5% of the inflorescence dry matter; an intermediate chemotype II, with CBD as the prevalent cannabinoid but also THC present at various concentrations; and a chemotype III, with particularly low THC content. These chemotypes were presumed to be associated mainly to geographical provenance. No studies on the respective roles of heredity and environment on the chemotype expression were performed. Tripartite patterns of CBD/THC ratio distributions were recognized within populations by Fournier and Paris (1979) and by Fournier (1981). De Meijer *et al.* (1992), in a survey of a large Cannabis collection, also found that plants belonging to the same population often show distinct CBD/THC ratios. A rare, additional chemotype, characterized by a very low content of both THC and CBD and with CBG as the predominant constituent, was later identified by Fournier *et al.* (1987).

The biosynthesis of cannabinoids: In the Cannabis plant, cannabinoids are synthesized and accumulated as cannabinoid acids [*e.g.*, cannabidiolic acid (CBDA)]. When the herbal product is dried, stored, or heated, the acids decarboxylize gradually or completely into neutral forms (*e.g.*, CBDA → CBD). For convenience, this article indicates all cannabinoids by the abbreviations for their neutral forms.

The first specific step in cannabinoid biosynthesis is the condensation reaction of geranylpyrophosphate (GPP) with the polyketide, olivetolic acid (OA), which is catalyzed by the enzyme geranylpyrophosphate:olivetolate geranyltransferase (GOT; Fellermeier and Zenk 1998; Fellermeier *et al.* 2001). The resulting CBG is the direct precursor for CBD (Taura *et al.* 1996) and

[1]*Present address:* GW Pharmaceuticals plc., Porton Down Science Park, Salisbury, Wiltshire SP4 0JQ, United Kingdom.

[2]*Corresponding author:* Istituto Sperimentale per le Colture Industriali, Via di Corticella 133, 40128 Bologna, Italy.
E-mail: g.mandolino@isci.it

CBC (Gaoni and Mechoulam 1966; Morimoto *et al.*, 1997, 1998). In older references THC was considered a further cyclization product of CBD (Shoyama *et al.* 1974). Later, Fournier and Paris (1980) assumed that this pathway, CBG → CBD → THC, was characteristic for fiber strains only. For drug strains, which are often devoid of even trace amounts of CBD, the direct conversion of CBG into THC was supposed to be typical. Today, THC is considered to be derived directly from CBG in all Cannabis strains (Figure 1b); the existence of the postulated enzyme CBD-cyclase catalyzing the synthesis of THC via CBD has not been experimentally confirmed (Taura *et al.* 1995). The propyl homolog of CBG, *i.e.*, CBGV, is formed if a C_{10}, instead of the common C_{12} version of OA, condenses with GPP (Shoyama *et al.* 1984; Fellermeier and Zenk 1998). The *in vivo* conversions of CBG(V) into the end products THC(V), CBD(V), and CBC(V) are enzymatically catalyzed, and for each reaction an enzyme has been identified: THC synthase (Taura *et al.* 1995), CBD synthase (Taura *et al.* 1996), and CBC synthase (Morimoto *et al.* 1998). Shoyama *et al.* (1984) demonstrated that an enzyme extract from a Cannabis strain containing CBD and THC is able to convert CBGV into THCV and CBDV. This last finding implies that THC, CBD, and probably also CBC synthase, are able to process homologs of CBG regardless of the length of their alkyl side chain.

Chemotype inheritance: There is little doubt that environmental factors have a strong influence in modulating the amount of cannabinoids present in the different parts of the plants at different growth stages, as demonstrated by a number of articles (*e.g.*, Lydon *et al.* 1987; Bócsa *et al.* 1997). However, the tripartite distribution of CBD/THC ratios in most populations is likely to underlie a discrete inheritance of the chemotype trait. Indeed, most authors agree that cannabinoid profiles are under strong genetic control. According to Beutler and der Manderosian (1978), the ratio of CBD/THC is a chemical marker of taxonomic significance. Fournier *et al.* (1987) stated that the cannabinoid profile of each plant—and therefore its CBD/THC ratio—is chiefly dependent on its genetic background and that each individual plant invariably belongs to its distinct chemical group throughout its life cycle.

Quantitative and qualitative aspects of cannabinoid accumulation are often confused as pointed out by Hillig (2002) in a critical comment on Sytnik and Stelmah (1999). To specify the target of the current article it is adequate to express the yield of a certain cannabinoid per crop area unit as a complex trait,

$$CY_n = DM \times P_{flor} \times C_{tot} \times PC_n, \qquad (1)$$

where CY_n is the yield of cannabinoid, n (grams per square meter); DM is the total amount of dry, aboveground biomass (grams per square meter); P_{flor} is the weight proportion of inflorescence leaves and bracts (grams per gram); C_{tot} is the total cannabinoid content in the inflorescence leaves and bract fraction (grams per gram); and PC_n is "purity," the proportion of cannabinoid n in the total cannabinoid fraction (grams per gram).

The first three components determining the cannabinoid yield are probably polygenic traits not related to specific metabolic pathways and are heavily affected by environment. In contrast, the latter term of Equation 1, the proportion of a certain cannabinoid in the total cannabinoid fraction, depends strictly on the metabolic pathways followed by the plant to convert common precursors into specific end products. The focus in this article is on the proportions of the two most commonly found and abundant cannabinoids, CBD and THC, and restricts the definition of chemotype to the ratio of CBD/THC, with both terms expressed as percentage of the inflorescence dry weight.

Fournier and Paris (1979, 1980) and Fournier (1981) reported a clear-cut segregation for CBD/THC ratios within French fiber cultivars. Two groups could be distinguished in a ratio of 1:4, the first one composed of plants with mixed CBD-THC profiles and the other of plants with fairly pure CBD profiles. Yotoriyama *et al.* (1980) analyzed the F_2 from F_1 hybrids containing both CBD and THC in similar amounts and found segregation of the chemotypes with pure CBD, mixed CBD-THC, and pure THC profiles in a 1:2:1 ratio. The subsequent generations of the pure CBD plants were further investigated and they showed a fixed CBD chemotype. Becu *et al.* (1998) evaluated a segregating F_2 and supposed a monogenic inheritance of THC and CBD ratios. This was not confirmed by other authors (Sytnik and Stelmah 1999).

DNA markers in Cannabis: Today, the concept of Cannabis as a monotypic genus is widely accepted; taxonomical, morphological, and biometrical studies confirm the continuity of its gene pool despite the extremely high variation found within and between populations (Small *et al.* 1976; de Meijer and Keizer 1996). In the last few years, the existence of just a single species within the genus has been confirmed by molecular marker studies that show a limited segregation of the different groups within the genus Cannabis and an extremely high degree of polymorphism, estimated to be of the same magnitude within and between populations (Faeti *et al.* 1996; Forapani *et al.* 2001). Within some of the best-known hemp cultivars, *e.g.*, Carmagnola, the degree of polymorphism was estimated by randomly amplified polymorphic DNA (RAPD) markers to involve ~80% of the markers scored, and the data suggested a huge reservoir of variation within even the most selected Cannabis strains considered during the study. Finally, within the dioecious populations, the presence of a high number of male-specific markers, presumably associated with the Y chromosome, was found by RAPD and amplified fragment length polymorphism

analysis (Mandolino *et al.* 1999, 2002; Flachowsky *et al.* 2001).

Aim of the work: This study aims to clarify the inheritance of cannabinoid chemotype, by isolating pure CBD and pure THC inbred lines. A simple inheritance model was proposed after making crosses between the chemotypically contrasting lines and examining a number of F_1 and F_2 progenies. An RAPD analysis of the parental lines and their offspring was performed and a number of chemotype-associated markers were described.

MATERIALS AND METHODS

Chemotype assessment: Mature floral clusters were collected from each and every individual plant. The flower clusters were air dried, and 50 mg of leafy material was weighed in a filtration tube (Ultrafree-CL, 0.1 μm; Millipore, Bedford, MA). The following steps were then repeated four times: 1 ml of ethanol (99.7%) was added, the sample was sonicated in ethanol for 15 min, and the extract was centrifuged at 4000 rpm for 10 min. Then, the total 4 ml of ethanol containing the extracted cannabinoids was transferred from the filtration tube to a 5-ml volumetric flask; 0.25 ml of a phenanthrene stock solution (10 mg/ml in ethanol) was added as internal standard and the volume was adjusted to 5 ml with ethanol. Finally, extracts were homogenized and transferred to GC vials. Gas-chromatographic analyses were performed on a Hewlett-Packard 6890 GC equipped with an autosampler and a flame ionization detector. Two columns were used: the (slightly polar) HP-5, 320 μm × 30 m, with 0.25-μm film for general quantitative analysis of larger sample loads, and the nonpolar HP-1, 100 μm × 40 m, with 0.20-μm film for an accurate separation of CBD from CBC. Average moisture content per progeny studied was determined by drying samples of the floral material at 105° for 3 hr. Moisture correction factors and a linear calibration equation, obtained with a CBD concentration range, were used to convert GC-derived peak areas to dry weight concentrations. Compound identities were determined by matching retention times with those of pure standards.

Constitution of inbred lines: All parentals used in this study were doubly inbred plants (S_2's) obtained through the self-fertilization of selected female clones from the Cannabis collection of HortaPharm B.V., The Netherlands. The original plants had either CBD or THC as the predominant cannabinoid. The 00.45.1 clone was an exception, having both CBD and THC in similar amounts. The clones were obtained through *in vivo* propagation of lateral branches. An individual from each clone was partially sex reversed according to the procedure described by Mohan Ram and Sett (1982) and allowed to self-pollinate in isolation. In many cases it was possible to collect sufficient viable seed to constitute a first-generation inbred line (S_1), which was completely female and showing the same chemotype as the parental clone. The 00.45.1 S_1, however, segregated into pure CBD, mixed CBD-THC, and pure THC individuals. Here, further inbreeding was restricted to the pure CBD plants. An S_2 generation was produced from some of the S_1 plants, using the same procedure described above. The work focused on six S_2 lines, briefly described in Table 1. A leaf sample was collected for DNA analysis from 10 to 20 plants per S_2.

Production of F_1's and F_2's: Seven individual plants belonging to the six S_2 lines with contrasting chemotypes were chosen to produce hybrid F_1's. The individual female plants used as pollen parents were partially sex reversed and placed in isolation cabinets with the seed parent plants. The crosses performed are summarized in Table 2.

The next season, a variable number of individual plants of the four different F_1's were grown and their chemotype was determined. Again, leaf samples from 5–20 plants belonging to the four F_1's were taken for molecular analysis. Ten F_1 plants (3 plants from cross 99.3, 3 from 99.4, 1 from 99.5, and 3 from 99.6; Table 2) were again treated to obtain partial sex reversion, isolated, and allowed to set F_2 seed. The seeds were sown and a variable number of F_2 mature plants, ranging from 35 to 118, were evaluated for chemotype; leaf samples were again collected for DNA analysis. During three seasons (1998–2000), the complete cycle from the parental S_2 inbred lines to the different F_2 progenies was accomplished under similar greenhouse conditions and strict isolation. Confirmation of the genetic femaleness of all the plants was based on the absence of any male-specific marker.

Molecular analysis: From each leaf sample taken from S_2, F_1, and F_2 individual plants, genomic DNA was prepared using the Nucleon Phytopure kit (Amersham Pharmacia Biotech, Buckinghamshire, UK). The DNA concentration of each sample was adjusted to 1 μg/μl after 260-nm readings, and 20 ng were used for amplification reactions. RAPD analysis using decamer primers of random sequence (purchased from Operon Technologies, Alameda, CA, and the Nucleic Acid-Protein Service Unit, Biotechnology Laboratory, University of British Columbia, Canada) and gel electrophoresis were conducted as described elsewhere (Faeti *et al.* 1996). For the analysis of S_2 and F_1 individuals, a matrix composed of 1's and 0's was obtained after scoring the RAPD bands reproducible after at least two rounds of amplifications of the same template. The number of loci and the polymorphism percentage were calculated on the basis of these matrices. F_2 samples were analyzed by bulk segregant analysis (Michelmore *et al.* 1991); seven pairs of bulks were composed from the contrasting (CBD and THC) chemotypes of the F_2's by mixing equimolar amounts of DNA from 6–14 individual plants per chemotype for each of the seven progenies. The RAPD analysis of the bulked DNA samples was performed as described above for individual analysis. Single DNA bands differentiating the F_2 bulks were eluted from the gel using the QIAquick gel extraction kit (QIAGEN, Valencia, CA) and cloned in the pGEM-T vector system II (Promega, Madison, WI); *Escherichia coli* cells (strain JM109) were then transformed with the recombinant plasmid using the Gene Pulser electroporator (Bio-Rad, Richmond, CA) and plated on suitable media. The positive clones were separately cultured, and the inserts were excised from the plasmid vector, labeled with [α-^{32}P]dCTP according to Feinberg and Vogelstein (1983), and used as probes in Southern blot hybridization experiments of RAPD amplified products as described elsewhere (Mandolino *et al.* 1999), to check the identity of the cloned fragment with the original RAPD marker. Once the identity was confirmed, the DNA insert was sequenced using the BigDye Terminator cycle sequencing kit (Applied Biosystems, Foster City, CA) and the automatic sequencer AbiPrism 310 Genetic Analyzer (Applied Biosystems). From the sequences obtained, specific 20-mer primers were constructed and tested in PCR reactions carried out under the same conditions described elsewhere (Mandolino *et al.* 1999).

RESULTS

Variation in cannabinoid composition: Examples of Cannabis gas-chromatographic profiles are shown in Figure 1a. The variation found within accessions is usually at both levels of cannabinoid type (different reten-

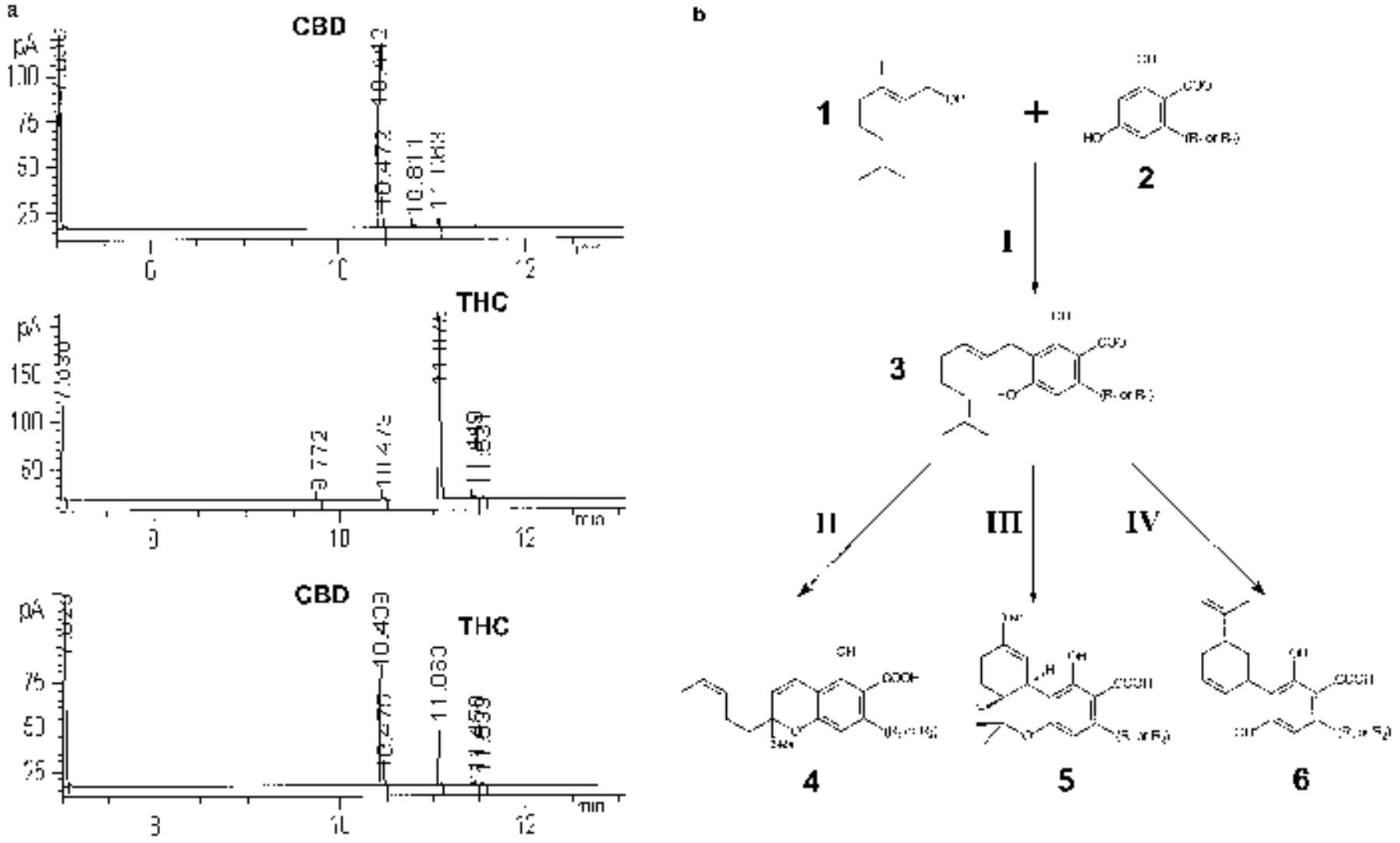

FIGURE 1.—(a) GC chromatograms of female individuals belonging to three different chemotypes (pure-CBD, pure-THC, and mixed chemotype) from different breeding lines are shown. The horizontal axis indicates the retention time and the vertical axis the detector signal (picoamps). CBD and THC peaks appear in the interval 10.439–10.442 and 11.063–11.075 min, respectively. (b) The biosynthetic pathway of cannabinoids is shown (modified after FELLERMEIER *et al.* 2001). 1, geranylpyrophosphate; 2, olivetolic acid; 3, CBG(V); 4, CBC(V); 5, THC(V); 6, CBD(V); I, geranylpyrophosphate:olivetolate geranyltransferase (GOT); II, CBC(V) synthase; III, THC(V) synthase; IV, CBD(V) synthase. R_1 (= $-C_3H_7$) and R_2 (= $-C_5H_{11}$) indicate the propyl and pentyl forms of the different metabolites.

TABLE 1

Characteristics of the parental inbred lines

S_2 inbred line	Source population	Predominant cannabinoid	Purity[a] (%)	Content of major cannabinoid (%): avg. ± SD	Polymorphism[b] (%)
00.45.1.2	Afghani-Skunk hybrid	CBD	96.0	2.94 ± 1.58	33.3
94.4.12.63	German fiber landrace	CBD	97.7	1.33 ± 0.33	45.8
94.5.2.30	Turkish fiber landrace	CBD	84.2	4.00 ± 2.01	17.3
55.22.7.10	Thai marijuana landrace	THC	93.9	5.45 ± 0.90	46.4
55.24.4.34	South Indian marijuana landrace	THC	87.5	6.69 ± 1.15	34.3
55.28.1.4	U.S. strain "Duckfoot"	THC	86.8	2.58 ± 1.55	36.0

[a] The proportion of the major cannabinoid in the total cannabinoid fraction.
[b] As evaluated by RAPD markers (see MATERIALS AND METHODS).

tion times) and amount (different peak areas), confirming other authors' observations (SMALL and BECKSTEAD 1973; DE MEIJER *et al.* 1992). Plants with only a single dominant cannabinoid (THC or CBD in this study) are present as 95–98% of the total cannabinoid content (Figure 1a, top and middle chromatograms). Such plants are here termed as pure chemotypes and can be found naturally occurring within accessions or can be produced by crossing or self-fertilizing plants showing a mixed chemotype (Figure 1a, bottom chromatogram; see below).

S_2 inbred lines: If the clone originally used to produce the S_2 was of a pure CBD or THC chemotype, this chemotype is preserved throughout all the subsequent inbred generations, although the absolute amount of the dominant cannabinoid still shows considerable variation, as demonstrated by the standard deviations found (Table 1).

The molecular analysis performed on the S_2 lines suggested a narrowing of the genetic variation within these materials, especially if compared with noninbred populations as examined previously (FORAPANI *et al.* 2001). The percentage of polymorphisms detectable by RAPD analysis (three primers) ranged from 17.3 to 46.4% (Table 1), while it was previously reported to be from 65 to 80% within the most common Italian and French fiber cultivars (FORAPANI *et al.* 2001).

F_1 hybrids: When the S_2 plants were mutually crossed as indicated in Table 2, all F_1 plants examined contained both CBD and THC in considerable amounts, and no pure chemotypes were found. An example of the distribution of the F_1 plants (F_1 99.3) in a CBD *vs.* THC plot is presented in Figure 2, where the values for the parental plants are also indicated. Figure 2 shows a heterotic effect for the total cannabinoid content (CBD + THC) and in particular a strong increase of the CBD content of the F_1's as compared with the parental CBD content. Similar patterns were found in two of the other F_1's. However, there was no such effect on the total cannabinoid content in the 99.4 F_1.

Surprisingly, the CBD/THC ratio appeared to vary in a clear progeny-specific way. The average ratio varied significantly ($P < 0.001$) from 0.50 (F_1 99.4) up to 1.57 (F_1 99.5), where only four plants could be analyzed (Table 3).

TABLE 2

Pedigrees and codes of the studied progenies

CBD parent (seed parent)	THC parent (pollen parent)	F_1 code	Code of the F_1 plants selfed	F_2 code
94.4.12.63.5	55.22.7.10.6	99.3	99.3.10	99.3.10
			99.3.34	99.3.34
			99.3.49	99.3.49
94.5.2.30.1	55.24.4.34.8	99.4	99.4.2	99.4.2
			99.4.6	99.4.6
			99.4.10	99.4.10
00.45.1.2.6	55.28.1.4.9	99.5	99.5.5	99.5.5
94.4.12.63.8	55.28.1.4.9	99.6	99.6.2	99.6.2
			99.6.14	99.6.14
			99.6.25	99.6.25

The pedigrees of the F_1 and F_2 progenies examined for chemotype segregation are shown. The underlined ciphers of the codes denote the specific individuals that have been crossed (S_2's) or self-fertilized (F_1's) to obtain the subsequent generation.

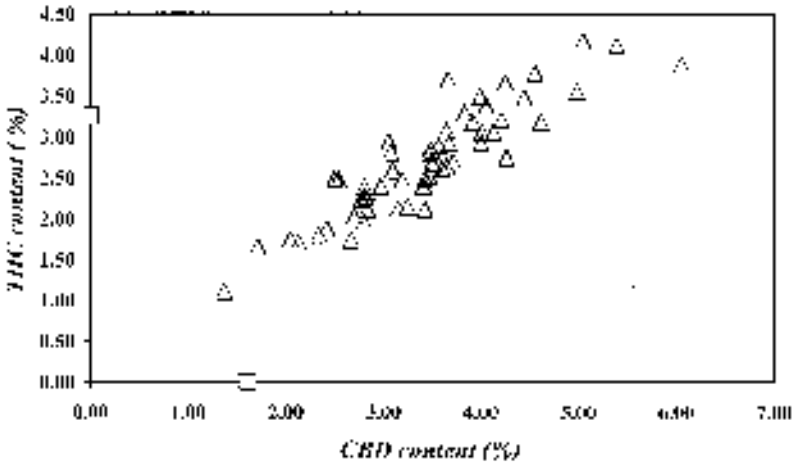

FIGURE 2.—CBD *vs.* THC content for the 99.3 F_1 progeny. The positions of the F_1 plants are indicated by open triangles; their parents (belonging to the CBD line 94.4.12.63 and to the THC line 55.22.7.10, respectively) are indicated by open squares. The cannabinoid content is expressed as weight percentage of the dry, mature inflorescence.

The molecular analysis of the F_1 plants was again carried out by using three RAPD primers and comparing the loci identified in the parental plants with those present in the F_1 progenies. In the four progenies, from 5.6% (99.3) to 37.5% (99.4) of the RAPD loci scored in the parental plants segregated in the F_1 (data not shown). This indicates that a still significant number of loci were in the heterozygous state in the S_2 parentals.

F_2 inbreds: The CBD *vs.* THC contents of one of the F_2's examined are plotted in Figure 3. Within each F_2, the individuals could unmistakably be assigned to three different segregant groups on the basis of large discontinuities in the calculated CBD/THC ratios. Data on the segregation of chemotypes in the different F_2's are shown in Table 4. For all F_2's, the results of the χ^2 test accepted the model of a single locus with two codominant alleles. In those F_2's where the segregation ratios deviated most from 1:2:1 (largest χ^2 values), this was consistently due to an underrepresentation of pure CBD chemotypes.

As in the F_1's, the CBD/THC ratio for heterozygous plants in the F_2's again appeared to be strongly progeny dependent. ANOVA and a multiple-comparison least-squared difference (LSD) test performed on heterozygous CBD/THC ratios of all evaluated F_1 and F_2 progenies showed highly significant differences among progenies ($P < 0.001$; Table 3). Especially, the 99.4 progenies form a very distinct cluster as they are composed of heterozygotes having a much higher proportion of THC than of CBD. In all other progenies the proportion of CBD in heterozygotes exceeds, to a varying extent, the THC proportion. Small, though significant, differences in mean CBD/THC ratio do occur among progenies sharing the same pedigree. However, the F_1 CBD/THC ratios (assessed in a different year from the F_2 ratios) are transmitted with very little change to the heterozygotes of the corresponding inbred F_2's.

TABLE 3

CBD/THC ratios of heterozygous plants

Progeny	CBD/THC ratio (avg. ± SD)	$P = 0.05$	$P = 0.01$
99.4 F_1	0.50 ± 0.07	a	a
99.4.2 F_2	0.59 ± 0.06	ab	a
99.4.10 F_2	0.60 ± 0.08	ab	a
99.4.6 F_2	0.68 ± 0.11	b	a
99.3.49 F_2	1.09 ± 0.24	c	bc
99.3.10 F_2	1.15 ± 0.18	c	bcd
99.3 F_1	1.18 ± 0.13	c	bcd
99.6.14 F_2	1.26 ± 0.13	d	cde
99.3.34 F_2	1.27 ± 0.13	de	cde
99.6.2 F_2	1.34 ± 0.22	e	def
99.6 F_1	1.43 ± 0.28	f	efg
99.6.25 F_2	1.47 ± 0.25	f	efg
99.5.5 F_2	1.47 ± 0.20	f	fg
99.5 F_1	1.57 ± 0.22	f	efg

The F_1 and F_2 progenies were ranked according to the CBD/THC ratio measured in the heterozygotes (all F_1 plants and ~50% of the F_2 plants). Means with different letters are significantly different at the given *P* value (LSD test).

Molecular markers associated to chemotype: The clear-cut segregation observed in all the F_2's considered allowed the application of the bulk segregant analysis (BSA; MICHELMORE *et al.* 1991) strategy to find molecular markers linked to chemotype. Seven pairs of DNA bulks were made, corresponding to the F_2 progenies 99.3.10, 99.3.49, 99.4.2, 99.4.6, 99.4.10, 99.5.5, and 99.6.25, and each was composed of 8–10 DNAs from the contrasting chemotype groups. Fifty RAPD primers were used to screen the seven bulks, and ~400 bands were scored. In several cases, bands discriminating one or more of the bulks were observed, but only three primers, OPA07 (5′-GAAACGGGTG-3′) and OPB06 (5′-TGCTCTGCCC-3′) from Operon Technologies and UBC109 (5′-TGTACGTGAC-3′) from the University of British Columbia, produced three bands, two THC and one CBD associated, discriminating six or seven pairs of DNA bulks. The CBD-associated band ($UBC109_{620}$) was ~620 bp, while the two THC-associated markers ($OPB06_{1000}$ and $OPA07_{2100}$) were ~1000 and 2100 bp.

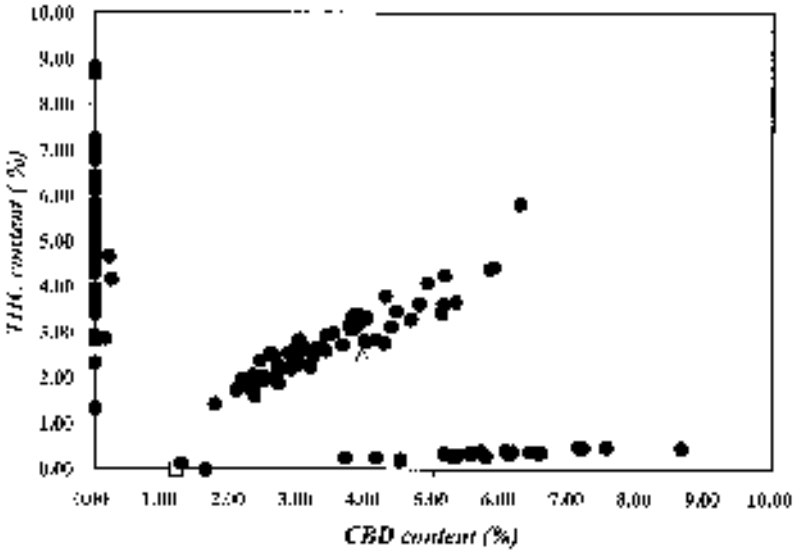

FIGURE 3.—CBD *vs.* THC content plot of the 99.6.14 F_2 progeny (circles). The positions of the single F_1 plant (open triangle), self-fertilized to obtain this F_2, and of the initial pure-CBD and pure-THC parents (open squares) are indicated. Cannabinoid content is expressed as weight percentage of the dry, mature inflorescences.

The RAPD bands discriminating the chemotypes within each bulk are shown in Figure 4, a–c. The chemotype-associated bands were then examined in the single-plant DNAs comprising the bulks. The results are summarized in Table 5. Marker $OPA07_{2100}$ appeared to be the most effective as it is present in all THC plants and in only 2 out of 61 CBD plants examined. Marker $OPB06_{1000}$ is equally effective for THC plants, but it was detected in 6 out of 61 CBD plants as well. Marker $UBC109_{620}$, CBD associated, is present in all CBD F_2's except 99.5.5, where only 1 plant out of 9 consistently showed the marker. The same marker is also present in 12 out of 61 THC plants examined. In general, if lack of association can be attributed to genetic recombination between the chemotype locus and the marker, then the marker $OPA07_{2100}$ had 1.3% recombination, marker $OPB06_{1000}$ had 5.3%, and marker $UBC109_{620}$ had ~10.3%, calculated as average of all the seven F_2's.

The three markers were sequenced and different combinations of 20-mer specific primers were constructed on the basis of the DNA sequence and tested. The best primers were found to be derived from the marker $OPB06_{1000}$, originally THC associated. This sequence-characterized amplified region (SCAR) marker, termed B190/B200, was tested on all single plants of the F_2's examined, pure THC, pure CBD, or heterozygous CBD/THC, and on all the S_2 plants from the original cross parents. The result of DNA amplification primed by B190/B200 primers (forward, 5′-TGCTCTGCCCA AAGTATCAA-3′; reverse, 5′-CCACTCACCACTCCACC TTT-3′) is shown in Figure 4d. The THC phenotype is associated with the amplification of a band of approximate molecular weight of 190 bp, whereas pure CBD plants show a band of ~200 bp. Heterozygous plants showed both fragments in most cases, indicating that this marker had the same codominance characteristics

TABLE 4

Chemotype segregation data

F_2	No. plants analyzed	CBD	CBD-THC	THC	χ^2 value	1:2:1 accepted
99.3.10	40	10	21	9	0.15	Yes
99.3.34	35	4	17	14	5.74	Yes
99.3.49	41	10	23	8	0.80	Yes
99.4.2	58	17	27	14	0.59	Yes
99.4.6	38	10	17	11	0.47	Yes
99.4.10	58	8	35	15	4.17	Yes
99.5.5	66	10	34	22	4.42	Yes
99.6.2	55	7	32	16	4.42	Yes
99.6.14	118	22	56	40	5.80	Yes
99.6.25	37	11	15	11	1.32	Yes

The number of plants belonging to the different chemotypes is shown; the data are relative to the 10 F_2's analyzed (see Table 2 for the pedigrees). The χ^2 values were calculated to test the goodness of fit of the data to the model of one Mendelian locus for chemotype with two codominant alleles responsible for CBD and THC accumulation, respectively. The χ^2 threshold for acceptance at d.f. = 2 and $P = 0.05$ is 5.99.

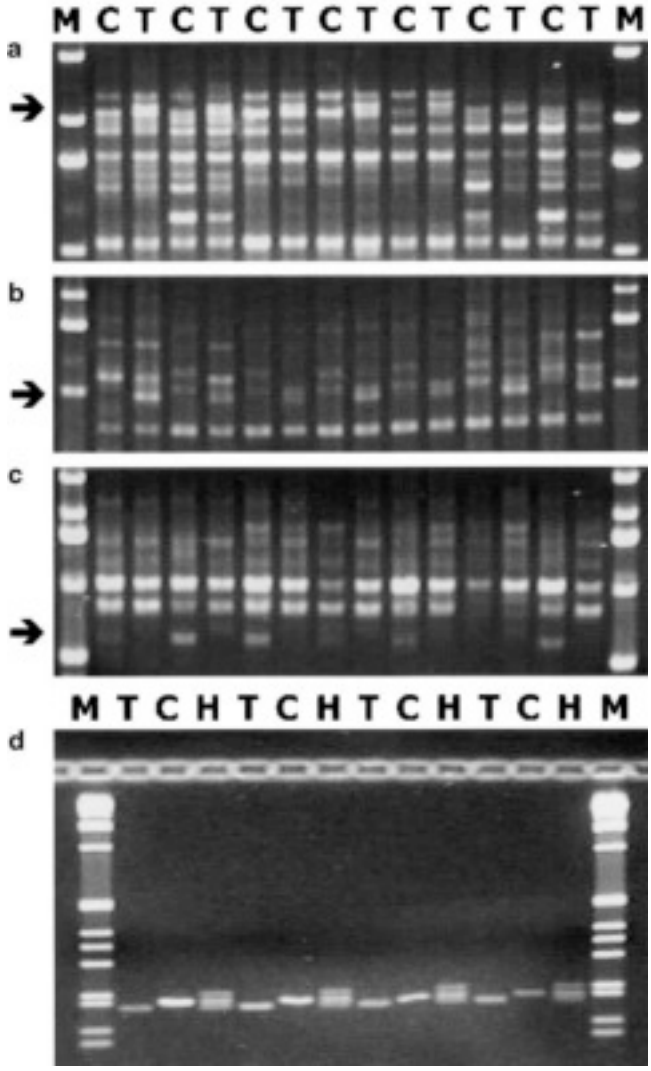

FIGURE 4.—Result of the amplification mediated by the three RAPD primers OPA07, OPB06, and UBC109 (a–c) on the bulked DNAs from THC (T) and CBD (C) segregants of the seven different F_2's listed in Table 5. The arrows indicate the two THC- (a and b) and the CBD- (c) specific DNA fragments. (d) The amplification produced by the SCAR marker obtained on the basis of the sequence of the THC-specific fragment shown in b; each lane shows the result of the amplification of DNA from single plants, pure THC (T), pure CBD (C), or mixed (H). M indicates the molecular weight marker (1-kb ladder; Life Technologies).

as the chemotype locus; these plants also showed an additional band at higher molecular weight (~250 bp). The efficiency of the B190/B200 marker in predicting the chemotype of the plants examined in this work is summarized in Table 6. Remarkably, the efficiency with which the concurrent appearance of both markers identified the heterozygous chemotypes is 95.3%. These values were calculated on the basis of the results of the markers in 63 plants belonging to the S_2 inbred lines, in 39 plants from the four different F_1's, and in 246 plants of the different F_2's. The sequenced portions of the DNA fragments generated by the RAPD primers were found to share no significant homology with the published DNA sequences for the THC and CBD synthases (GenBank nos. E55108/GI 18629739 and E33091/GI 18623981; patent pending).

DISCUSSION

Within Cannabis populations, large variations in cannabinoid composition and content can be found among individual plants. Therefore, to study the inheritance of the chemotype trait, we chose to use inbred, female lines with fixed, pure chemotype. RAPD data confirmed the relative narrowing of the genetic basis of the parental plants due to this strategy (Table 1). In the parental S_2 lines, the percentage of the major cannabinoid (PC_n) ranged from 84 to 98% of the total cannabinoid fraction C_{tot}. For the F_1 99.3, 99.5, and 99.6 and for their F_2 offspring, the sum of CBD and THC reached proportions of 95% of C_{tot}. The remaining fractions were composed of varying mixtures of CBG, CBC, THCV, and CBDV. The 99.4 F_1's and their F_2 descendants had a lower proportion of CBD + THC (91%). This was due to a consistent presence of higher amounts of CBG, a feature inherited from both parents of this progeny.

The uniformity of F_1 chemotypes and the F_2 segregation ratios demonstrate the presence of a single locus, which is referred to as *B*, showing simple Mendelian inheritance of the two alleles, B_D and B_T, evidenced by this study. The model proposes that a pure CBD plant has a B_D/B_D genotype at the *B* locus, while a pure THC plant has a B_T/B_T genotype. F_1 and ~50% of the F_2 plants are therefore heterozygous B_D/B_T, with the two alleles being codominant and therefore simultaneously expressed in the hybrids. The hypothesis of two alleles at one locus was accepted by χ^2 tests for all the F_2's examined (Table 4). This model agrees with the assumption of a monogenic inheritance as expressed by BECU *et al.* (1998).

It should be acknowledged that these results may also be explained with the hypothesis of two duplicated loci, one encoding for a CBD synthase and the other for a THC synthase, mapping so closely that observation of linkage rupture was impossible in the progenies examined. Such a situation was found in different cases of secondary metabolism genes where duplicated members encoded for enzymes catalyzing either consecutive metabolic steps or alternative reactions from a common precursor. In maize, a family of four duplicated genes (BX2–5) was shown to encode for cytochrome P450-dependent monooxygenases, each catalyzing one of the consecutive steps from indole to 2,4-dihydroxy-7-methoxy-1,4-benzoxazin-3-one (DIMBOA) synthesis (FREY *et al.* 1995; GLAWISCHNIG *et al.* 1999). In Arabidopsis, genes encoding for different 2-oxoglutarate-dependent dioxygenases (AOP1–3) were identified as responsible for the synthesis of different glucosinolates; these genes map to the same position and code for alternative reac-

TABLE 5

Chemotype-associated markers

	Pure THC plants				Pure CBD plants			
F_2	No. of plants	Marker $OPA07_{2100}$	Marker $OPB06_{1000}$	Marker $UBC109_{620}$	No. of plants	Marker $OPA07_{2100}$	Marker $OPB06_{1000}$	Marker $UBC109_{620}$
99.3.10	9	9	9	3	9	1	1	9
99.3.49	6	6	6	1	6	0	1	6
99.4.2	14	14	14	2	14	1	1	14
99.4.6	10	10	10	3	10	0	0	10
99.4.10	6	6	6	1	6	0	1	6
99.5.5	9	9	9	2	9	0	2	1
99.6.25	7	7	7	0	7	0	0	7

Presence of the three RAPD markers, two THC associated ($OPA07_{2100}$ and $OPB06_{1000}$) and one CBD associated ($UBC109_{620}$), in the individual plants composing the contrasting bulks used for bulk segregant analysis is shown.

tions from a common precursor, leading to 3-hydroxypropyl and 3-butenyl glucosinolate (Kliebenstein *et al.* 2001). These authors could also identify a recombinant inbred (RI) line endowed with null alleles, accumulating the precursor 4-methylsulfinylbutyl glucosinolate. AOP genes were considered as duplicated, rather than as allelic, essentially on the basis of the presence of a cluster of candidate genes in the sequenced Arabidopsis genome. Besides, little homology was found between AOP and other genes, though a 60–70% sequence homology was found among the cluster components. In the case of cannabinoid genes, however, although the possibility of the presence of duplicated genes cannot be ruled out on the basis of the presented experiments, the consequences of a model with duplicated loci should be examined. Had parental CBD lines carried defective alleles at the *THC* locus (*thc/thc-CBD/CBD*) and THC lines at the *CBD* locus (*THC/THC-cbd/cbd*), a similar, chemotypically uniform, F_1 would have been found, as well as the same segregation of chemotypes in the F_2. Theoretically the screening of wide populations should reveal the existence of fixed doubly dominant homozygous (*THC/THC-CBD/CBD*), showing both cannabinoids: these plants should not segregate on selfing. Such a situation has never been observed during several years of germplasm screening and selfing in the breeding programs. Further consideration suggests that if there were any chance of a cross that separated the two duplicated loci, then the CBG chemotype (*thc/thc-cbd/cbd*) should be found more frequently than has actually been observed. In fact, in populations where a high frequency of all the three chemotypes is found, as in hashish landraces, the recessive alleles *cbd* and *thc* should occur with significant frequency and, because Cannabis of necessity outbreeds, should have a good chance to occur in the homozygous state. Instead, CBG plants have been detected in fiber cultivars that, according to a two-locus model, should have a very high frequency of *CBD* and *thc* alleles. Therefore, even extremely low frequencies of the *cbd* allele should lead to frequent CBG chemotypes, due to the virtual absence of *THC* alleles in fiber hemp. Conversely, as yet only a few reports of single plants show the CBG chemotype (Fournier *et al.* 1987; G. Grassi and V. G. Virovets, personal communications), suggesting a very low frequency for this chemotype, despite the observation that the plants carrying defective alleles suffered no loss of vitality (G. Fournier, personal communication).

These facts are more convincingly explained by the rare occurrence of a mutated B_0 allele for a defective enzyme at a single locus. Horkay (1986) estimated the degree of self-fertilization in monoecious populations, like the French cultivars in which the CBG plants were found, at 20–26%, and therefore it is conceivable that a mutated, inactive allele at the locus *B* could have thrived through repeated and frequent inbreeding until becoming fixed in a few plants.

Although based on negative evidence, it is our opinion that the model of a single allelic locus governing the synthesis of CBD and THC better explains the chemotype distribution in Cannabis populations.

In those F_2's where the segregation ratios deviated most from 1:2:1 (largest χ^2 values), this was consistently due to an underrepresentation of CBD homozygotes, possibly an effect of a recessive semilethal factor loosely associated with the B_D allele. For the 99.3.34 F_2, differential seed viability and seedling survival could explain the lower proportion of CBD homozygotes. The other F_2's with higher χ^2 values showed an insignificant loss of seeds and seedlings. The assumed semilethal factor in these progenies must therefore already be effective during embryogenesis. An additional indication of reduced viability of the B_D/B_D genotype is provided by the dramatic drop of fertility in pure CBD plants during line selection, a phenomenon that is usually absent in THC inbreds (data not shown). Although one can expect most fiber cultivars to have the B_D/B_D genotype, to the best of our knowledge there are no reports on reduced viability and fertility of such strains, as compared with high-THC populations. However, only a chemotypically

TABLE 6

Linkage of chemotypes with the marker B190/B200

Chemotype	No. of plants	Marker B190	Marker B200	Both markers
THC	89	78	3	8
CBD-THC	169	3	5	161
CBD	90	0	88	2

Degree of linkage of the different forms of the SCAR marker B190/B200 with the chemotypes is shown. The data are relative to all the plants belonging to the original S_2 lines listed in Table 1, to the four F_1's listed in Table 2, and to the seven F_2's of Table 5.

segregating population could provide proper evidence for such a phenomenon, as all other genetic traits contributing to viability and fertility need to be randomized among chemotype groups.

In three out of the four F_1 progenies, the CBD content was higher than in the CBD parental (Figure 2). This can be explained by the fact that CBD parental lines, usually derived from fiber strains, have low values for P_{flor} and C_{tot} (see Equation 1). These components, polygenic in nature, show a strong heterotic effect; therefore, the B_D allele of F_1 plants is active in a much more productive genetic environment than that in the parental lines. This does not hold true for the B_T allele, which already comes from drug strains with high P_{flor} and C_{tot} values.

When working with young, vegetative plant materials, the molecular markers described may be more effective than GC chromatograms in genotyping plants for their chemotype. The RAPD markers originally identified were completely dominant, as expected from a PCR marker; however, the marker B190/B200, one of the SCAR markers developed on the basis of sequence information, behaves codominantly (Figure 4d). Both types of marker appear tightly linked to the chemotype in the pedigrees so far examined (Tables 5 and 6). The B190/B200 marker could be profitably employed in the breeding work, though at present we have no data on its utility beyond the specific crosses made in this study. The marker seems particularly suitable to distinguish pure CBD and heterozygous plants, which can be valuable when counterselecting for THC chemotypes in fiber hemp breeding.

The synthesis of THC and CBD in Cannabis plants has been described as an oxidoreduction coupled to a cyclization of CBG, catalyzed by a THC and a CBD synthase, respectively. A CBC synthase has also been described, catalyzing a similar reaction leading to CBC. These enzymes were isolated from different drug or fiber strains, and many of their characteristics were elucidated. Most of the properties of CBD and THC synthase were very similar, like the mass (75 kD), the existence as a monomer localized in the cytosol, the pI, the optimum pH, the rate constant k_{cat}, the V_{max}, and the K_m for their substrate. Also, the NH_2-terminal sequence of the two synthases shared 87% of identity (Taura *et al.* 1995, 1996). The properties described for CBC synthase (Morimoto *et al.* 1998) were quite different; this enzyme showed a lower K_m (23 μM instead of 134 and 137 μM of the other two synthases), a lower turnover number ($k_{cat} = 0.04\ s^{-1}$ against 0.19 and 0.20 of CBD and THC synthase). It can be hypothesized that each of the alleles identified in this work, *i.e.*, B_D and B_T, codes for an isoform of the same enzyme, showing specificity for the conversion of CBG to CBD or THC, respectively. In the heterozygous state, both isoforms would be present and therefore both conversions would occur, in accordance with the mixed chemotypes observed in the F_1's and in one-half of the F_2 plants. This hypothesis, presented here for the first time, is also supported by the recent publication of the cDNA sequence of the CBD and THC synthases; the two sequences share 89% identity, and the longest nonmatching stretch is four nucleotides. The fact that, according to Taura *et al.* (1995, 1996), the two synthases have very similar affinities for CBG would theoretically result in CBD/THC ratios close to 1.0 in B_D/B_T genotypes. It is intriguing that the different parental combinations examined in our experiments show different CBD/THC ratios in the resulting F_1 hybrid, often strongly deviating from 1.0 and fairly stably inherited by the F_2 heterozygotes (Table 3). Some heritable factor seems to affect the balance between CBD and THC synthase in their competition to convert the CBG precursor. Had this factor been at a different locus, segregation for the CBD/THC ratio in the F_2's should be observed. However, as shown in Figure 3, within an F_2 progeny there is no evidence of several heterozygous clusters with distinct slopes. It is therefore possible that B_D and B_T are part of a wider allelic series, coding for several isoenzymatic forms of CBD synthase and THC synthase, respectively, with differential affinities for the CBG substrate, resulting in significantly different CBD/THC ratios in heterozygotes. When two homozygous parents are crossed, one with a certain isoform of CBD synthase, the other with a certain isoform of THC synthase, the CBD/THC ratio in the F_1's will depend on the balance between the efficiencies of the two synthases and will remain fixed in any further heterozygous descendant obtained through self-fertilization.

The data obtained in this work do not take into account the inheritance of the conversion of CBG into CBC. In principle, it is possible to suppose the existence of a further allele, B_C, at the *B* locus, coding for a CBC synthase (Morimoto *et al.* 1998), but no direct evidence is available and there is the possibility that a different locus could be involved. CBC synthase was originally isolated from a juvenile stage of a CBD strain, and as yet it has proved very difficult to obtain pure CBC plants, although plants reaching CBC proportions up to 64%

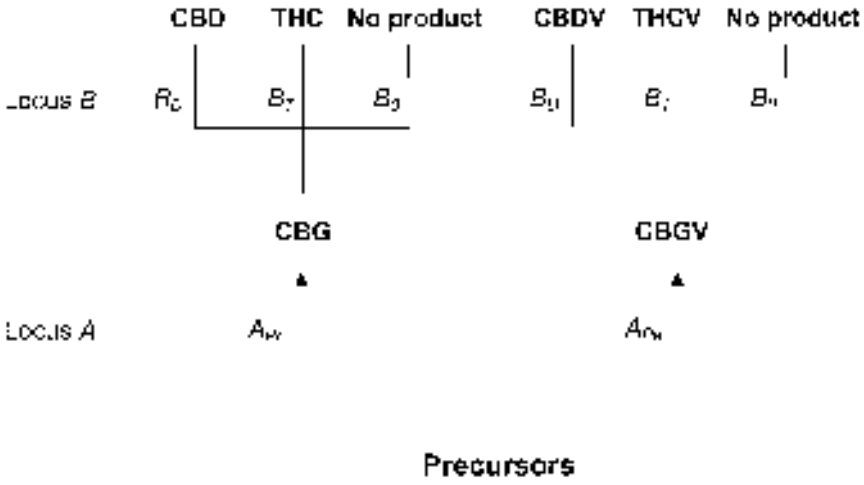

FIGURE 5.—The genetic model presented postulates the existence of two separate loci, *A* and *B*. Locus *A* has two different alleles A_{Pr} and A_{Pe}, responsible for the synthesis of either the propyl or the pentyl form of cannabigerol (CBG or CBGV, with $R_1 = -C_3H_7$ and $R_2 = -C_5H_{11}$). Locus *B* has different alleles, each responsible for the conversion of CBG(V) into the end products CBD(V) and THC(V) accumulating in the inflorescences. The B_0 allele codes for a nonfunctional enzyme, leading to CBG(V) accumulation. All cannabinoids are drawn in their neutral form.

of the total cannabinoid fractions have been reported (HOLLEY *et al.* 1975).

The model for the biogenesis of cannabinoids with its relations between alleles and resulting chemotypes is illustrated in Figure 5. In this scheme, it is assumed that the pathway leading to CBG or CBGV is governed by at least one allelic locus, called *A*, on which the experiments presented here provide no information. However, there is evidence for the existence of "null" genotypes at the *A* locus, leading to plants devoid of any cannabinoids; such phenotypes have indeed been observed (V. G. VIROVETS and G. GRASSI, personal communication). The pathways shown are consistent with the assumption of SHOYAMA *et al.* (1984) that the synthases characterized so far can convert equally well both CBG and CBGV into the end products THC(V) and CBD(V). This assumption is also supported by our observation that CBDV traces were detected only in CBD homozygotes (B_D/B_D) and heterozygotes (B_D/B_T), whereas THCV traces occurred exclusively in THC homozygotes (B_T/B_T) and heterozygotes (data not shown).

In the materials studied, the proportions of both CBD and THC reached at best ~96–98% of the total cannabinoid fraction. Generally, even after five cycles of inbreeding selection aimed at one target cannabinoid, at least a 2–4% impurity consisting of other cannabinoids remains. Therefore, the alleles postulated here, even in homozygous genotypes, seem to have an imperfect control over the biosynthetic events. Apparently, any of the postulated isoenzymatic forms encoded by the alleles at the *B* locus show a residual ability to convert the precursor CBG(V) into cannabinoids other than the major one.

The existence of a single locus determining the chemotype, with at least two alleles, gives a clear genetic meaning to the tripartite distribution of the chemotypes within populations, as observed by several authors when CBD *vs.* THC content plots are considered (FOURNIER and PARIS 1979; FOURNIER 1981; DE MEIJER *et al.* 1992). According to our model, these plots do not merely show phenotypic distributions, but rather they visualize the allele frequency within a population. It should be possible to study the frequencies of the B_D and B_T alleles and their changes during time as a function of the population structure, the action of environmental conditions, and the different fitness values carried by them. A further consequence of the fact that CBD *vs.* THC plots actually are to be considered allele distribution plots is that no barriers between the different chemotypes of Cannabis can be postulated. The plants that are differently distributed in a CBD *vs.* THC plot have no large genetic differences, only different alleles at one single locus. Therefore, the commonly practiced application of chemotype as a taxonomic criterion is very disputable. Probably a polygenic character, such as the total cannabinoid content, is better balanced and preserved in populations and hence is a more robust criterion with which to discriminate subspecific taxa.

In the work presented here, the strategy of partial sex reversion of female plants was used to obtain S_2, F_1's, and F_2's; yet the results obtained are also expected to hold true if dioecious plants had been used as parentals or F_1 plants had been intercrossed to obtain segregating F_2's. The model proposed here is therefore highly predictive and intended to stimulate further research: it provides a tool to elucidate the possible existence of other genetic loci regulating the cannabinoid composition.

We gratefully acknowledge the plant breeding and chemical-analytical assistance of Tina Ent (Horta Pharm B.V.) and Kathy Hammond (GW Pharmaceuticals plc.) for text revision. This work was funded in part by the Italian Ministry of Agriculture and Forestry in the framework of the project "Induction of phenotypic markers and improvement of common hemp."

LITERATURE CITED

BECU, D. M. S., H. D. MASTEBROEK and H. J. P. MARVIN, 1998 Breeding for root knot nematode resistance in hemp. Proceedings of Bast Fibrous Plants Today and Tomorrow, St. Petersburg, Russia, p. 149.

Beutler, J. A., and A. H. der Manderosian, 1978 Chemotaxonomy of *Cannabis* I. Crossbreeding between *Cannabis sativa* and *C. ruderalis*, with analysis of cannabinoid content. Econ. Bot. **32:** 387–394.

Bócsa, I., P. Mathé and L. Hangyel, 1997 Effect of nitrogen on tetrahydrocannabinol (THC) content in hemp (*Cannabis sativa* L.) leaves at different positions. J. Int. Hemp Assoc. **4:** 80–81.

de Meijer, E. P. M., and L. C. P. Keizer, 1996 Patterns of diversity in *Cannabis*. Genet. Res. Crop Evol. **43:** 41–52.

de Meijer, E. P. M., H. J. van der Kamp and F. A. van Eeuwijk, 1992 Characterisation of *Cannabis* accessions with regard to cannabinoid content in relation to other plant characters. Euphytica **62:** 187–200.

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

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

Faeti, V., G. Mandolino and P. Ranalli, 1996 Genetic diversity of *Cannabis sativa* germplasm based on RAPD markers. Plant Breed. **115:** 367–370.

Feinberg, A. P., and B. Vogelstein, 1983 A technique for radiolabeling DNA restriction endonuclease fragments to high specific activities. Anal. Biochem. **132:** 6–13.

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

Fellermeier, M., W. Eisenreich, A. Bacher and M. H. Zenk, 2001 Biosynthesis of cannabinoids. Incorporation experiments with ^{13}C-labeled glucoses. Eur. J. Biochem. **268:** 1596–1604.

Flachowsky, H., E. Schumann, W. E. Weber and A. Peil, 2001 Application of AFLP for the detection of sex-specific markers in hemp. Plant Breed. **120:** 305–309.

Forapani, S., A. Carboni, C. Paoletti, V. M. C. Moliterni, P. Ranalli *et al.*, 2001 Comparison of hemp (*Cannabis sativa* L.) varieties using RAPD markers. Crop Sci. **41:** 1682–1689.

Fournier, G., 1981 Les chemiotypes du chanvre (*Cannabis sativa* L.). Intérêt pour un programme de sélection. Agronomie **1:** 679–688.

Fournier, G., and M. R. Paris, 1979 Le chanvre papetier (*Cannabis sativa* L.) cultivé en France: le point sur ses constituants. Plant. Med. Phytother. **13:** 116–121.

Fournier, G., and M. Paris, 1980 Détermination de chimiotypes à partir des cannabinoïdes chez le chanvre à fibres monoïque (*Cannabis sativa* L.). Possibilités de sélection. Physiol. Veg. **18:** 349–356.

Fournier, G., C. Richez-Dumanois, J. Duvezin, J.-P. Mathieu and M. Paris, 1987 Identification of a new chemotype in *Cannabis sativa*: cannabigerol-dominant plants, biogenetic and agronomic prospects. Plant. Med. **53:** 277–280.

Frey, M., R. Kleim, H. Saedler and A. Gierl, 1995 Expression of a cytochrome P450 gene family in maize. Mol. Gen. Genet. **246:** 100–109.

Gaoni, Y., and R. Mechoulam, 1966 Cannabichromene, a new active principle in hashish. Chem. Commun. **1:** 20–21.

Glawischnig, E., S. Grün, M. Frey and A. Gierl, 1999 Cytochrome P450 monooxygenases of DIBOA biosynthesis: specificity and conservation among grasses. Phytochemistry **50:** 925–930.

Hammond, C. T., and P. G. Mahlberg, 1977 Morphogenesis of capitate glandular hairs of *Cannabis sativa* (Cannabaceae). Am. J. Bot. **64:** 1023–1031.

Hillig, K., 2002 Letter to the editor. J. Ind. Hemp **7:** 5–6.

Holley, J. H., K. W. Hadley and C. E. Turner, 1975 Constituents of *Cannabis sativa* L. XI. Cannabidiol and cannabichromene in samples of known geographical origin. J. Pharm. Sci. **64:** 892–894.

Horkay, E., 1986 Establishing the share of self- and cross fertilisation by means of population genetics in a monoecious hemp stand. Növénytermelés **35:** 177–182 (in Hungarian).

Kliebenstein, D. J., V. M. Lambrix, M. Reichelt, J. Gershenzon and T. Mitchell-Olds, 2001 Gene duplication in the diversification of secondary metabolism: tandem 2-oxoglutarate-dependent dioxygenases control glucosinolate biosynthesis in *Arabidopsis*. Plant Cell **13:** 681–693.

Lydon, J., A. H. Teramura and C. B. Coffman, 1987 UV-B radiation effects on photosynthesis, growth and cannabinoid production of two *Cannabis sativa* chemotypes. Photochem. Photobiol. **46:** 201–206.

Mandolino, G., A. Carboni, S. Forapani, V. Faeti and P. Ranalli, 1999 Identification of DNA markers linked to the male sex in dioecious hemp (*Cannabis sativa* L.). Theor. Appl. Genet. **98:** 86–92.

Mandolino, G., A. Carboni, M. Bagatta, V. M. C. Moliterni and P. Ranalli, 2002 Occurrence and frequency of putatively Y chromosome linked DNA markers in *Cannabis sativa* L. Euphytica **126:** 211–218.

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

Michelmore, R. W., I. Paran and R. V. Kesseli, 1991 Identification of markers linked to disease-resistance genes by bulked segregant analysis: a rapid method to detect markers in specific genomic regions by using segregating populations. Proc. Natl. Acad. Sci. USA **88:** 9828–9832.

Mohan Ram, H. Y., and R. Sett, 1982 Induction of fertile male flowers in genetically female *Cannabis sativa* plants by silver nitrate and silver thiosulphate anionic complex. Theor. Appl. Genet. **62:** 369–375.

Morimoto, S., K. Komatsu, F. Taura and Y. Shoyama, 1997 Enzymological evidence for cannabichromenic acid biosynthesis. J. Nat. Prod. **60:** 854–857.

Morimoto, S., K. Komatsu, F. Taura and Y. Shoyama, 1998 Purification and characterization of cannabichromenic acid synthase from *Cannabis sativa*. Phytochemistry **49:** 1525–1529.

Shoyama, Y., M. Yagi and I. Nishioka, 1974 Biosynthesis of cannabinoid acids. Phytochemistry **14:** 2189–2192.

Shoyama, Y., H. Hirano and I. Nishioka, 1984 Biosynthesis of propyl cannabinoid acid and its biosynthetic relationship with pentyl and methyl cannabinoid acids. Phytochemistry **23:** 1909–1912.

Small, E., and H. D. Beckstead, 1973 Common cannabinoid phenotypes in 350 stocks of *Cannabis*. Lloydia **36:** 144–165.

Small, E., P. Y. Jui and L. P. Lefkovitch, 1976 A numerical taxonomy analysis of *Cannabis* with special reference to species delimitation. Syst. Bot. **1:** 67–84.

Sytnik, V. P., and A. F. Stelmah, 1999 The character of inheritance of differences in cannabinoid content in hemp (*Cannabis sativa* L.). J. Ind. Hemp Assoc. **6:** 8–9.

Taura, F., S. Morimoto and Y. Shoyama, 1995 First direct evidence for the mechanism of delta-1-tetrahydrocannabinolic acid biosynthesis. J. Am. Chem. Soc. **38:** 9766–9767.

Taura, F., S. Morimoto and Y. Shoyama, 1996 Purification and characterization of cannabidiolic-acid synthase from *Cannabis sativa* L. J. Biol. Chem. **271:** 17411–17416.

Yotoriyama, M., I. Ito, D. Takashima, Y. Shoyama and I. Nishioka, 1980 Breeding of Cannabis. Determination of cannabinoids by high-pressure liquid chromatography. Yakugaku Zasshi **100:** 611–614.

Communicating editor: C. S. Gasser