Euphytica **140:** 95–106, 2004.

The sexual differentiation of *Cannabis sativa* L.: A morphological and molecular study

V.M. Cristiana Moliterni[1,*], Luigi Cattivelli[2], P. Ranalli[1] & Giuseppe Mandolino[1]
[1]*Istituto Sperimentale per le Colture Industriali, via di Corticella 133, 40128 Bologna, Italy;* [2]*Istituto Sperimentale per la Cerealicoltura, via S. Protaso 308, 29017 Fiorenzuola d'Arda, Italy*
*(*author for correspondence: e-mail: vmc.moliterni@isci.it)*

Key words: Cannabis sativa, cDNA-AFLP, dioecy, sex linked markers, sexual differentiation

Summary

Cannabis sativa L. is a dioecious species with sexual dimorphism occurring in a late stage of plant development. Sex is determined by heteromorphic chromosomes (X and Y): male is the heterogametic sex (XY) and female is the homogametic one (XX). The sexual phenotype of *Cannabis* often shows some flexibility leading to the differentiation of hermaphrodite flowers or bisexual inflorescences (monoecious phenotype). Sex is considered an important trait for hemp genetic improvement; therefore, the study of the mechanism of sexual differentiation is of paramount interest in hemp research. A morphological and molecular study of *Cannabis sativa* sexual differentiation has been carried out in the Italian dioecious cultivar Fibranova.

Microscopic analysis of male and female apices revealed that their reproductive commitment may occur as soon as the leaves of the fourth node emerge; the genetic expression of male and female apices at this stage has been compared by cDNA-AFLP. A rapid method for the early sex discrimination has been developed, based on the PCR amplification of a male-specific SCAR marker directly from a tissue fragment.

Five of the several cDNA-AFLP polymorphic fragments identified have been confirmed to be differentially expressed in male and female apices at the fourth node. Cloning and sequencing revealed that they belong to nine different mRNAs that were all induced in the female apices at this stage. Four out of them showed a high degree of similarity with known sequences: a putative permease, a SMT3-like protein, a putative kinesin and a RAC-GTP binding protein.

Abbreviations: AFLP: Amplified Fragment Length Polymorphism; LINE: Long Interspersed Elements; LTR: Long Terminal Repeat; SCAR: Sequence Characterized Amplified Region

Introduction

Cannabis sativa belongs to the family of Cannabaceae, order Rosales (APGII, 2003). It is a naturally dioecious species with male and female individuals showing unisexual flowers and characterized by sexual dimorphism: male plants are generally taller and slender than female plants and have a shorter life cycle. Unisexual flowers are bored in inflorescences that are terminal at an earlier stage, and terminal or lateral in a later stage.

Male inflorescence consists of hanging panicles sometimes branched, generally with few or no leaves and composed by a variable number of flowers. Male flower has a perianth of five sepals that encloses the androecium, composed by five stamens bored by subtle stalks. The anthers at maturity undergo dehiscence longitudinally, releasing the pollen grains that are mostly wind dispersed (Mohan Ram and Nath, 1964). Female inflorescence is a raceme developing at the apex of the plant or at the axils of leaves or lateral branches. The female flower has a very simple structure as it

is composed by a green bract that completely wraps the rudimental perianth and the ovary. This latter is uniloculate and has a short style that distally differentiates a bifid stigma.

The chromosome set of *Cannabis sativa* is composed by nine pairs of autosomes and one pair of sexual chromosomes: X and Y. The male sex is endowed with an XY pair, and the female one with an XX pair, similarly to what found in other dioecious species such as *Humulus lupulus*, *Silene latifolia*, *Coccinia indica*, *Rumex hastatulus*; however, sex determination in Cannabis has been supposed to be based on a X:autosome dosage rather than on an active-Y mechanism (Westgaard, 1958; Grant et al., 1994). The Y chromosome in Cannabis is subtelocentric and characterized by a satellite at the extremity of the short arm; besides, the long arm is particularly developed and probably responsible for the difference found between the male and the female genome sizes (1683 and 1636 Mbp, respectively; Sakamoto et al., 1998). The X chromosome is submetacentric, and bears a satellite at the end of short arm. There are no specific reports about the chromosome set of monoecious plants.

As already reported for many other plant sexual chromosomes, *Cannabis sativa* Y chromosome is strongly heterochromatic and rich of repetitive sequences that are likely cause of its marked metaphasic condensation. A high percent of the repeated DNA is made of LINE-like sequences (Boecke, 1989), probably representing traces of transposable elements showing a low level of transcription for the presence of still active ORFs, coding for enzymes involved in the transposition mechanism. *Cannabis sativa* LINE elements (LINE-CS) are represented in the X chromosome and in the autosomes too, but their concentration at the end of Y chromosome is particularly high. This observation led to the hypothesis that these sequences might have a role in maintaining the structure of Y chromosome and that they can contribute to the morphological and structural differentiation of the sex chromosomes, by creating heteromorphic regions in which the recombination is prevented (Sakamoto et al., 2000; Peil et al., 2003).

The phenotypic expression of sex in hemp shows some flexibility. Anomalies in flower development are sometimes observed, such as the appearance of hermaphrodite flowers or the development of mixed inflorescences (bearing both male and female flowers), like those occurring in the monoecious phenotypes. Monoecious varieties have been developed from some of these mutations, and need a strict selection to be maintained in the variety during the seed multiplication, due to the recessive nature of the trait.

In some hemp genotypes it is possible to obtain total or partial reversion of the sex. It is known that the treatment with masculinizing or feminizing chemical agents is effective in determining the formation of the opposite sex reproductive organs even in plants that are already sexually well differentiated. Chemicals that inhibit the biosynthesis or the activity of ethylene, such as aminoetoxyvinylglycine, silver thiosulphate and silver nitrate, have a masculinizing effect, while the precursors or activators of the biosynthesis of ethylene, like etephon, have a feminizing effect (Mohan Ram & Sett, 1982a, 1982b). The ability to undergo sexual reversion is thought to have a genetic base: some ecotypes such as the Italian Carmagnola are very resistant to any sex reversion treatment, while plants belonging to Fibranova cv. are quite prone to sex reversion (G. Grassi, E. de Meijer, personal communications).

In Italian open field conditions, the life cycle of a typical dioecious variety has a 5–6 months duration, and sexual maturity is attained after 3–4 months, when the earliest unisexual flowers appear. Sexual dimorphism of dioecious hemp is generally apparent only in a much later stage of development, just before the onset of flowering, when a marked elongation of the last internodes occurs in male plants causing them to become taller and slender than female plants.

Sex is considered an important trait for hemp genetic improvement. The Bredemann's strategy of selection for fibre quality implies a relatively early quali-quantitative analysis of fibre in male plants before pollen dispersion (Bredemann, 1938). This analysis is followed by the elimination of lower-quality male plants, not intended for pollination. Therefore, the possibility of early sex identification and the study of the mechanism of sexual differentiation in dioecious varieties are of paramount interest in hemp research. Attempts have been made in the pre-genomic era by multivariate analysis of morphological traits followed by correlation to the sex expression (Lacombe, 1980). Since the Nineties, DNA markers were developed, capable of discriminating the male plants from the female and the monoecious ones (Mandolino et al., 1998, 1999). Such markers (see also the paper by G. Mandolino and A. Carboni in this special issue) can be fruitfully used in the selection schemes for hemp breeding, and in the assessment of the number of male plants in monoecious seed lots.

Sex linked markers, provided that they are tightly and reliably associated to the sexual phenotype, can

be of great importance in the study of the earliest stages of Cannabis sexual differentiation. Beyond the relationship between the presence of the Y chromosome and the male sex, very little is known about the molecular mechanisms underlying sexual differentiation of hemp. We have investigated the onset of sexual differentiation both at the histological and molecular level. The characterization of different developmental stages in male and female plants was carried out by optical microscopy; differentially expressed sequences at the same developmental stage in the two sexes were also identified by cDNA-AFLP analysis. The possible role of these differently expressed sequences in male and female plants at the onset of sexual differentiation is discussed.

Materials and methods

Plant material

All the experiments were carried out using cv. Fibranova, an Italian cross-bred variety, characterized by high fiber content; Fibranova is dioecious and shows a sex ratio close to 1. Seeds were sown in peat paper pots, and plants grown in the greenhouse at a 16 h photoperiod, with a daily thermo period of 11–27 °C, throughout the plant's life cycle.

Sex identification

Early sex identification was carried out according to Klimyuk et al. (1993) with minor modifications (Mandolino and Ranalli, 2002). Tissue fragments 3–4 mm long were picked up from young leaves at the 2–3° node and collected in sterile 1.5 ml vials; 40 μl of a 0.25 M NaOH solution were added to each vial that was then placed in boiling water (100 °C) for 50 s. An amount of 40 μl of a 0.25 M NaCl solution and 20 μl of a detergent-buffer solution (0.5 M Tris pH 8, 0.25% Triton X-100) were added; the vials were then centrifuged for few seconds at maximum speed and placed in boiling water (100 °C) for 2 min. Each tissue fragment was transferred in a PCR vial and 25 μl of the PCR reaction mixture for the male specific marker amplification were added (Mandolino et al., 1999). Amplicons were run in a 1% agarose gel in TAE 1X buffer and visualized by U.V. exposition (305 nm), after ethidium bromide staining. All the buffers and disposables used in this procedure were sterile.

Histological methods

Apices from male and female plants grown in the above described conditions were collected at the emergence of the leaves of the second, fourth, sixth and eight nodes. The collected samples were submerged in the fixing solution (glutaraldehyde 3% in KH_2PO4/Na_2PO4 0.1 M buffer, pH 7); this step sometimes needed a brief treatment in a vacuum jar, to allow the air to flow out of the stem vessels. Fixing was obtained after a 1–3 weeks incubation at 4 °C. Samples were then rinsed several times in phosphate buffer 0.1 M, pH 7.0 at 4 °C, dehydrated in a graded ethanol series (ethanol/phosphate buffer, ethanol/water), and pre-infiltrated in a mixed solution of ethanol and Technovit 7100 resin (Heraeus Kulzer) for 4 h at 4 °C. The infiltration and embedding steps were carried out for 36 h at room temperature. Longitudinal sections 2–4 μm thick, obtained with an ultra microtome (Reichert Jung) fitted with a glass blade, were then stained with toluidine blue. A toluidine blue 0.1%, borax 1% aqueous water solution was dropped on the sections adhering on a glass microscope slide, and was then dried on a hot plate. Differential staining was obtained by adding few drops of an acid–alcohol solution (ethanol 70%, HCl 1.5%) and rinsing with water; a drop of water allowed the cover glass to adhere to the stained sections that were then observed and photographed with an optical microscope (Leitz Orthoplan).

RNA extraction and cDNA synthesis

Pooled apices, respectively from undifferentiated plants at the second node, male plants at fourth node and female plants at fourth node, were frozen and grinded under liquid nitrogen. Extraction Buffer (50 mM Tris-HCl, pH 9.0, 100 mM NaCl, 10 mM EDTA, 2% SDS) was warmed at 37 °C and added at a 5:1 ratio (v/w) to the powdered sample. Homogenate was transferred in a sterile polypropylene tube, extracted twice with an equal volume of a phenol/chloroform/isoamyl alcohol mixture (25:24:1), and once with two volumes of a chloroform/isoamyl alcohol mixture (100:1). Oligo-dT cellulose (Roche) 50 mg/ml and NaCl 0.4 M (final concentrations) were added to the supernatant. Samples were shaken (30–50 rpm) for 30 min and then centrifuged for few minutes at low speed. The cellulose was rinsed twice with 10 ml of washing Buffer 1 (10 mM Tris-HCl, pH 7.5, 400 mM NaCl, 0.2% SDS) and twice with 10 ml of washing Buffer 2 (20 mM Tris-HCl, pH 7.5, 100 mM

NaCl); the tubes were centrifuged at 800 × *g* for 5 min. after each washing. Cellulose was transferred to a 10 ml sterile chromatography column, and rinsed three times with 10 ml of washing Buffer 2. The mRNA was eluted from cellulose by applying to the column 10 mM Tris-HCl, pH 7.5, pre-heated at 65 °C. The eluted mRNA was precipitated with two volumes of ethanol, and resuspended in DEPC treated water (diethylpyrocarbonate 0.1%). mRNA (1 μg) was used as template for the synthesis of double stranded cDNA. First strand was synthesized by using the Superscript II Reverse Transcriptase (400 U, Life Technologies) in a final volume of 30 μl. RNAase H (1U, Life Technologies) and DNA Polymerase I (40U, Life Technologies) were added to the first strand reaction mixture, for the synthesis of the double stranded cDNA in a final volume of 150 μl. The mRNA template was finally digested by RNAase A (90 μg, Life Technologies) at 37 °C for 15 min. The reaction mixture was extracted with chlorophorm/isoamyl alcohol (100:1) and the double stranded cDNA purified by Microcon YM-100 columns (Amicon). One tenth of the recovered volume was quantified by gel electrophoresis, using lambda DNA as standard.

CDNA-AFLP analysis and band elution

The cDNA-AFLP procedure (Bachem et al., 1996) was adapted for *Cannabis sativa* following Hartings (1999), with minor modifications. Double stranded cDNA (20 ng) was digested by *Mse I* (5U) and *BstY I* (0.5U) (Life technologies) in a final volume of 40 μl, for 1 h at 37 °C. *Mse I* adapter (50 pmoles) and *BstY I* adapter (20 pmoles) were added to the digested cDNA and the ligation was carried out using theT4 DNA Ligase (5U, Amersham Pharmacia) for 3 h at 37 °C in a final volume of 50 μl. The ligation mixture was diluted to 100 μl, and 5 μl were used as template in the pre-amplification step. The pre-amplification (30 pmoles of the *BstY I* and *Mse I* anchor primers, 1.5 mM $MgCl_2$, 0.2 mM each dNTPs, 1 U Taq Polymerase and 1X Buffer, Life Technologies) was carried out for 20 cycles using the following profile: 94 °C, 30 s; 56 °C, 60 s; 72 °C, 60 s. One tenth of the secondary template produced was checked on a 0.8% agarose gel in TBE 0.5 X. Selective amplifications were carried out with 60 combinations of two anchor primers for *BstY I*, bearing an extension of one selective base (C or T), with 30 anchor primers for *Mse I* bearing a three selective bases extension (sequences of adapters and primers used are reported in Table 1). The *BstY I* selective primers were labelled with Redivue-[γ-^{33}P]dATP (Amersham Pharmacia) using the T4 Polynucleotide Kinase (Life Technologies), following the protocol suggested by the manufacturer. PCR amplification was carried out as follows: 94 °C, 30 s; 65 °C, 30 s (−0.7 °C/cycle) for 13 cycles, and 94 °C, 30 s; 56 °C, 30 s; 72 °C, 60 s for 13 cycles. Amplicons were size-fractionated on 6%, 0.35 mm thick polyacrylamide sequencing gels (urea 8 M) at 40 W for 2.5 h in the Sequigen GT apparatus (Bio-Rad). The gels were dried at 60 °C for 1 h, coated with transparent film and exposed to autoradiography film (Biomax MR, Kodak) for 24–48 h at −20 °C. The bands of interest were excised from the gel and collected in sterile vials; 50 μl of TE buffer 0.1 X were added, and vials incubated at 65 °C for 15 min for band recovery.

Table 1. Sequences of the adapters and primers used for cDNA-AFLP fingerprinting (Hartings, 1999)

Mse I adapter; top strand	5′-GACGATGAGTCCTGAG-3′
Mse I adapter; bottom strand	5′-TACTCAGGACTCAT-3′
BstY I adapter; top strand	5′-CTCGTAGACTGCGTACC-3′
BstY I adapter; bottom strand	5′-GATCGGTACGCAGTCTAC-3′
Mse I anchor primer	5′-GATGAGTCCTGAGTA-3′
BstY I anchor primer	5′-GTAGACTGCGTACCGATC-3′
Mse I selective primer	5′-GATGAGTCCTGAGTANNN-3′
BstY I selective primer 1	5′-GTAGACTGCGTACCGATCT-3′
BstY I selective primer 2	5′-GTAGACTGCGTACCGATCC-3′

The variable selective nucleotides are represented by N.

Reverse Northern hybridization

The recovered bands (10 μl of the eluted volume) were re-amplified with the same primer combinations and PCR profile used for their selective amplification. PCR products were checked on 2% agarose gels in TBE 1X and blotted on nylon membranes to obtain two identical blots of all the polymorphic fragments eluted. mRNA (1 μg) from males and females apices was transcribed to labelled cDNA in a mixture containing 10 mM DTT, 0.3 M each dNTP mix (with no dCTP), 40 μCi Redivue- [α-^{32}P] dCTP (Amersham Pharmacia), 1 μg oligo-dT_{18}, 400 U Superscript II RT enzyme in the 1X first strand RT Buffer (Life Technologies), in a final volume of 30 μl. Labelled cDNA was purified on Sephadex G-50 packed columns (Sigma-Aldrich) and quantified by the Cerenkov method. The two identical blots were hybridized with the same quantity (total cpm) of labelled cDNA from males and from females apices, for 8 h at 65 °C. The membranes were rinsed with SSC 2X and 1X, 0.1% SDS solutions for

20 min at 65 °C and then exposed to autoradiography film (Biomax MR, Kodak) for 7–10 days.

Northern hybridization

Messenger RNA belonging to males and females apices was separated on 1% agarose gel (40 mM MOPS, 10 mM sodium acetate, 1 mM EDTA, 0.1% DEPC, 2.2 M formaldehyde) and blotted on nylon membranes according to Sambrook et al. (1989). AFLP fragments were re-amplified from the eluted cDNA band as described above, and purified on a 1.15% agarose gel in TAE modified buffer 1X (40 mM TRIS-Acetate, 1 mM disodium-EDTA pH 8.0), with the Montage DNA gel Extraction Kit (Millipore). Eighty nanograms of the purified fragments were radioactively labelled with the random oligonucleotide priming method, according to Sambrook et al. (1989), using 40 μCi of Redivue- [$\alpha-^{32}$P]dCTP (Amersham Pharmacia), in a final volume of 25 μl. Labelled probes were purified on Sephadex G-50 (Sigma Aldrich) packed columns and quantified. Northern hybridization was carried out according to Sambrook et al. (1989); hybridized blots were then exposed to autoradiography film (Biomax MR, Kodak) for 1–3 days. Quantitation of mRNA immobilized on nylon membranes was made by radio labelled oligo-dT_{18} probes hybridization (Harley, 1997). Eighty nanograms of Oligo-dT_{18} were labelled in a mixture containing 50 μCi of Redivue- [γ-^{32}P] dATP (Amersham Pharmacia), 6U of T4 Polynucleotide Kinase and 1X enzyme buffer (Life Technologies), in a final volume of 50 μl, at 37 °C for 1 h. The labelled probe was precipitated in absolute ethanol (2.5 volumes) and ammonium acetate 10 M (2 volumes) for 45 min at −20 °C, centrifuged at 10.000 rpm and the pellet resuspended in 500 μl of Oligo Buffer (5X Denhardt, 0.8 M NaCl). Filters were pre-hybridized with Oligo Buffer at 30 °C for 6 h and hybridization was carried out over night. Filters were rinsed in SSC, 0.1% SDS solutions, at low but increasing stringency (SSC 6X, 5X, 4X) at 30 °C, and autoradiographed.

Cloning, reverse Northern hybridization and sequencing of the subclones

Positive, differentially hybridizating probes were ligated to the pCR 2.1-TOPO vector and cloned in the TOP 10 F′ cells One Shot chemically competent, by the TOPO TA Cloning kit (Invitrogen). Twenty mini preps per transformation were performed by the CONCERT Rapid Plasmid Miniprep System (Life Technologies). Cloned fragments were excised from vectors by *Eco RI* digestion and separated onto agarose gel. Two identical gels of the excised fragments were blotted onto nylon filters and hybridized with radio labelled cDNA from males and females, following the reverse northern procedure described above.

Recombinant subclones obtained from each transformation reaction were sequenced by ABI PRISM 310 Genetic Analyzer (PE Applied Biosystem). The sequences obtained were analyzed by the *BLAST* algorithm (Altschul et al., 1990).

Results

Microscopic analysis of apex differentiation

Microscopic analysis of male and female apices was performed at various developmental stages (emergence of the leaves of the second, fourth, sixth and eight nodes). At the last stage examined, the plants had started flowering, and the inflorescence buds were already macroscopically visible. The microscopic analysis of the inflorescence buds and mature flowers was also carried out. In most of the male and female apices analyzed at the fourth node stage, meristematic buds were observed at the axils of the leaves of the earlier nodes (third and/or second node; Figure 1a). No floral primordia were visible at this stage, but in the apices observed at the subsequent stages of differentiation (sixth and eighth node stages), meristematic buds were clearly more developed and showed a stronger mitotic activity (Figure 1b). From these observations, we argued that the meristem primordia produced at the fourth node stage could in most cases develop into an inflorescence bud, in response to external or internal cues. No meristem buds were observed at the axils of the leaves when apices at the second node were analyzed (Figure 1c), therefore it seems possible that the earliest step of apex sexual commitment could occur at the emergence of the fourth node's leaves in Fibranova cv., under the environmental conditions used. This stage was therefore chosen for the subsequent study of differential gene expression in male and female plants; as a control, apices picked up at the second node were used, as they represent a fully vegetative and undifferentiated meristem tissue.

Differential gene expression analysis

Sixty different primers combinations were used to screen double stranded cDNA belonging to male and

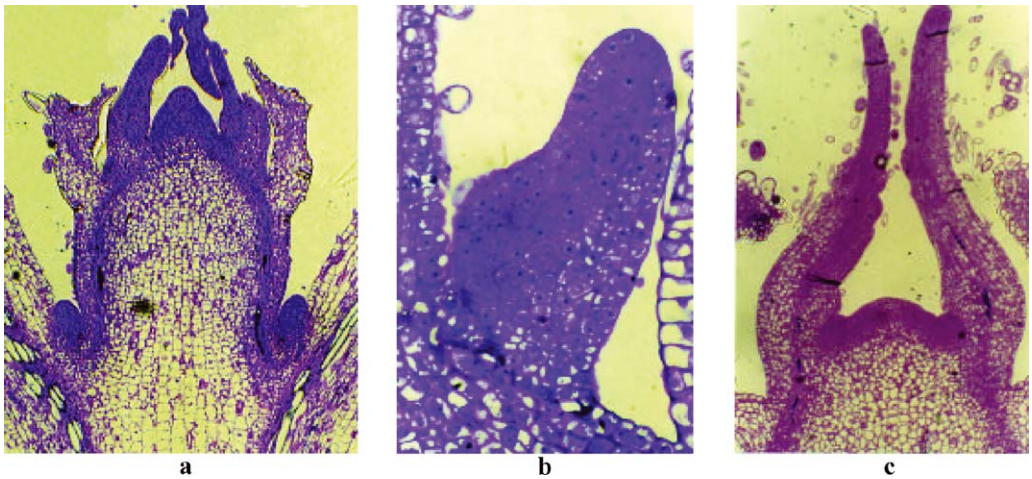

Figure 1. (a) Longitudinal section of the apex at the fourth node (25×). (b) Detail of a meristematic bud at the axil of a sub apical leaf (40×). (c) Longitudinal section of the apex at the second node (25×).

female apices at the fourth node. Four thousands eight hundred bands were obtained, with an average of 80 bands per primer combination. Nine hundreds out of the 4800 fragments were polymorphic, and therefore putatively belonging to differentially expressed mRNAs, as they were present only in male or in female cDNAs and were absent, or present below detectability levels, in the control cDNA sample (second node stage; Figure 2). Among the different combinations tested, those obtained using the B0+T primer, produced a higher number of polymorphic bands (509) compared to those obtained by the B0+C primer (384). In general, more polymorphic fragments were produced in the male cDNAs (average value of 9.6 bands per primer combination) than in the female ones (average 6.3 bands).

All the polymorphic fragments were eluted from the gel and blotted onto nylon membranes, in order to check by reverse northern hybridisation their true differential expression. Most of the isolated AFLP fragments resulted to be either equally expressed or not detectable when probed with labelled cDNA from males and female's fourth node apices. Only 22 fragments were confirmed to be actually differentially expressed. Five of these fragments, showing a clear differential expression and a high hybridization signal, were used as probes in northern analysis.

Northern data confirmed the differential expression of the five fragments at the fourth node stage; the mRNAs corresponding to the isolated fragments resulted all more expressed in female apices compared to the male ones (Figure 3). The five AFLP differentially displayed fragments, named C1, C2, C3, T1 and T2, were sub-cloned in pCR2.1 vector. About 20 preps from each subcloned fragment were further controlled by reverse Northern analysis. Reverse Northern revealed that only a subset of the subclones obtained from each AFLP polymorphic fragment corresponded to differentially expressed sequences in male and female apices at the fourth node, suggesting that in some cases the original amplified AFLP fragment could contained a mixture of different sequences.

Sequence analysis

All the subclones corresponding to male/female differentially expressed mRNAs were sequenced. Sequence analysis further confirmed the heterogeneous composition of some of the AFLP fragment isolated: the five differentially expressed AFLP fragments were in fact separated into nine different clones belonging to nine different mRNAs, all induced in the female apices at the fourth node. Clones derived from the same AFLP fragment have been designated by different letters (e.g. T1A, T1B, etc.); their main characteristics are summarised in Table 2. The clones showing the highest level of similarity with known sequences were: the *C2.A* clone (283 bp), within a 205 bp region of the cDNA for an SMT3-like protein (86% similarity); the *C3.A* clone (435 bp), within a 130 bp region of the cDNA for a kinesin nine heavy chain (84% similarity); the *T2.A* clone (282 bp), within a 231 bp long region of

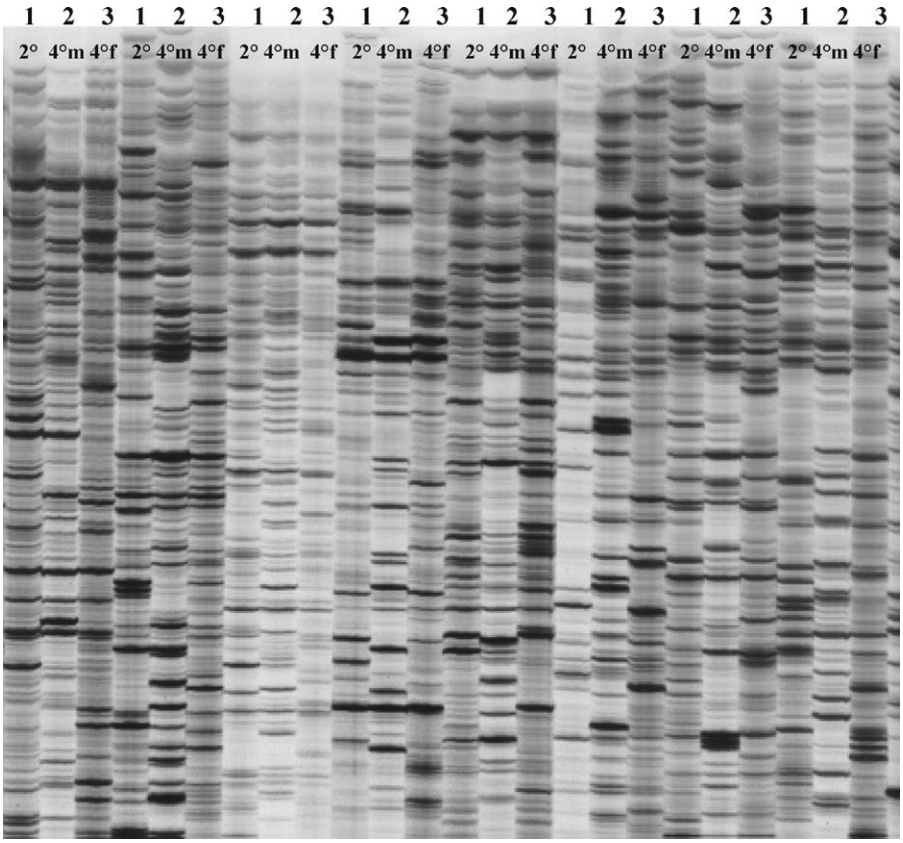

Figure 2. AFLP pattern produced by eight BstY I+1/Mse I+3 primer combination on cDNAs from apices at the second node stage (lane 1), male apices (lane 2) and female apices (lane 3) at the fourth node stage.

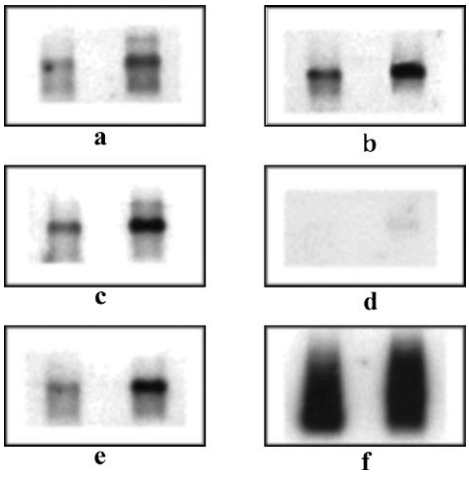

Figure 3. Northern analysis of five polymorphic cDNA AFLP fragments: (a) C1; (b) C2; (c) C3; (d) T1; (e) T2. Male and female mRNA from apices at the fourth node are blotted respectively on the left and right lane, and quantified in (f).

the cDNA for a putative permease (80% similarity); the *T2.B* clone (284 bp), that produced significant matches (86% similarity) within a 203 bp region of a Rac-GTP binding protein-like. Two examples of best matches are graphically represented in Figure 4.

Discussion

Under the condition described, the beginning of the sexual differentiation of *Cannabis sativa* (cv Fibranova), occurs 50–60 days after seed germination; at this stage the plant reaches a height of 160–180 cm. However, the reduction of the photoperiod length, or the exposition to low temperatures, causes the lifecycle to shorten, and the development of mature flowers to occur even 15–20 days after the plant emergence. These observations confirmed that the sexual commitment in hemp plants may take place in a very early phase of the vegetative development,

Table 2. Results of *BLAST* analysis for the nine clones differentially expressed at the fourth node (see text for details)

Clone	Primer combination	Best match (blastn)	Score	*E*-value	Best match (blastx)	Score	*E*-value	Putative function
T2.A (282bp)	*B0+T/M0+CGT*	AF466198.1 putative permease	97.6	$7e^{-18}$	AF466198 putative permease	159	$1e^{-38}$	Permease
C2.A (283bp)	*B0+C/M0+AGT*	AF451278.1 SMT3 ubiquitin like protein	184	$4e^{-44}$	AF451278 SMT3 protein	149	$1e^{-35}$	Ubiquitin like protein
C3.A (435bp)	*B0+C/M0+CAC*	AF272756 kinesin 9 heavy chain	101	$8e^{-19}$	AF272756 kinesin heavy chain	156	$4e^{-38}$	Kinesin heavy chain
T2.B (435bp)	*B0+T/M0+CGT*	–	–	–	AL163816 Rac-GTP binding protein like	100	$3e^{-21}$	Rac-GTP binding protein-like
C3.B (476bp)	*B0+C/M0+CAC*	DCU47095 putative ribosomal protein	135	$5e^{-29}$	AP003241 putative ribosomal protein 60S	94	$2e^{-23}$	Ribosomal protein
C1.A (218bp)	*B0+C/M0+ACG*	NM119070.1 acid phosphatase like protein	42	0.24	NM119070 acid phosphatase like protein	103	$6e^{-22}$	Acid phosphatase protein
C1.B (212bp)	*B0+C/M0+ACG*	AF108891 ADP ribosilation factor	98	$5e^{-18}$	AY062539 Calcium dependent protein kinase	50	$2e^{-09}$	–
T1.A (106bp)	*B0+T/M0+AAC*	–	–	–	PPA1 LYCES acid phosphatase precursor	40	0.006	Acid phosphatase precursor
T2.C (280bp)	*B0+T/M0+CGT*	ZMRNARPP2 ribosomal protein	91	$5e^{-16}$	–	–	–	Ribosomal protein

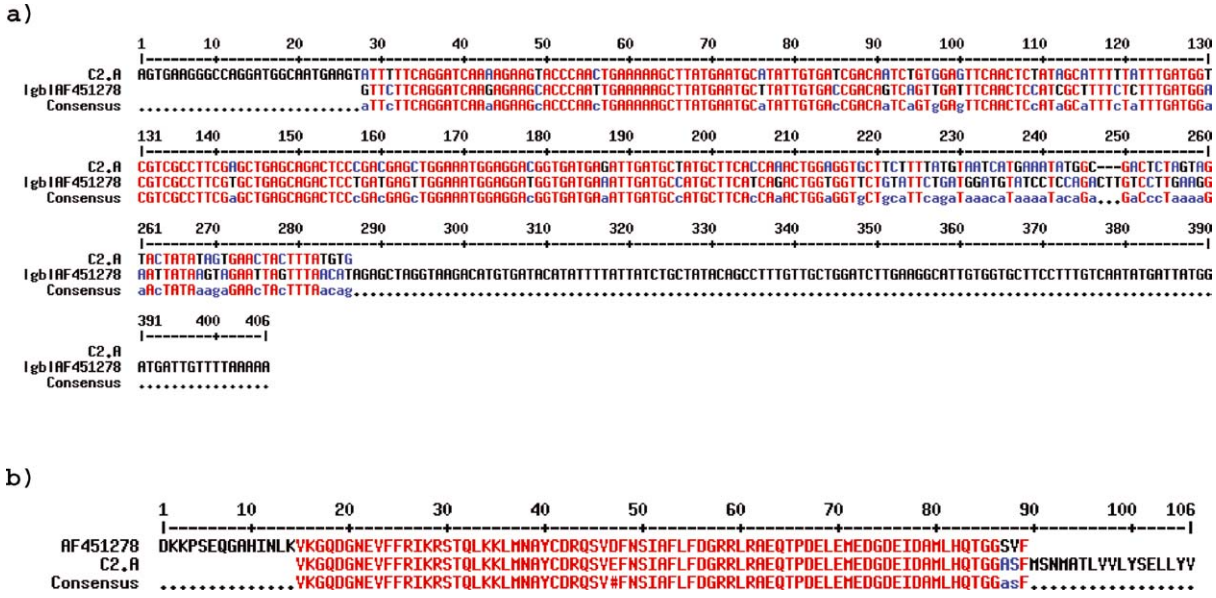

Figure 4. (a) Alignment between the C2.A clone and the AF451278 (ubiquitin-like protein SMT3 of *Phaseolus vulgaris*, partial coding sequence) nucleotide sequences and (b) aminoacid sequences (C2.A frame +2), performed by the MultAlin tool (Corpet, 1988).

and suggested to undertake a microscopic study of the changes in apex morphology during development, in support of gene expression analysis. Microscopic analysis was carried out in male and female apices picked up at various stages of development, from the emergence of the leaves of the second node, until the production of unisexual flowers. The choice to use the node number to define the plant developmental stage, instead of other morphologic parameters as the plant height or the time after seed germination, was suggested by the observation that these parameters are more flexible and more conditioned by environmental factors and genetic variability (Mediavilla et al., 1998). The microscopic analysis of male and female apices under non-inductive conditions (long photoperiod) revealed the formation of meristematic buds at the axils of the stem leaves as soon as the fourth node stage is reached, at a plant height of 15–20 cm. These undifferentiated meristematic buds could develop in an inflorescence bud under the opportune endogenous stimulus, as pointed out by the results of the microscopic analysis of the apices at the subsequent stage of development (data not shown). Apices then could be committed to reproductive development as early as the leaves of the fourth node emerge. Our interest was therefore focused on the male and female apices at the fourth node, and in particular on the genes that are expressed at this early stage of apex differentiation, and possibly responsible of its sexual commitment.

Sexual dimorphism in Cannabis naturally occurs late during its lifetime, immediately before the production of the unisexual flowers, when the vegetative development is almost completed. Therefore, the investigation of the early stages of sexual differentiation requires a method for the precise identification of the sex. A rapid method for sex determination in Cannabis has been developed, based on the PCR amplification of the 391 bp male-specific SCAR marker directly from a leaf tissue fragment. This method is suitable for a precise, early and rapid identification of male plants and was of great importance in the setting of the experimental design for the analysis of differential gene expression in the early phase of sexual differentiation; it can be also useful during the fiber hemp MAS programs.

Gene expression in male and female apices at the fourth node stage was compared by cDNA-AFLP, carried out testing 60 primer combinations. Nine hundreds polymorphic fragments were produced and tested by reverse Northern and Northern hybridization; however, only five fragments were confirmed as truly differentially expressed. CDNA-AFLP is a high-throughput technique: the restriction with two enzymes allows to reduce the size of the fragments to be analyzed in a

range of length (50–1000 bp) that could be resolved in a sequencing gel; the use of specific primers increases the reproducibility, and the two PCR amplifications are likely to increase the detectability of the genes with a low expression rate. The best enzymes to be used as rare cutter in the cDNA-AFLP should have one restriction site per transcript, in order to obtain as a product of the double digestion with rare and frequent cutter, one fragment per transcript. The most frequently used rare cutters are *Eco RI*, *BamH I* and *Pst I* enzymes, although it has been demonstrated that all of them have a restriction site only in half of the transcripts (Bachem et al., 1996), causing the reduction of the transcripts set to be analysed. *BstY I*, the rare cutter enzyme used in this work, has two variable positions in its restriction site, this allows to increase the probability for a given cDNA to be digested, but also the number of the fragments produced. In fact, we obtained an average of 80 bands per amplification, in comparison to a mean of 50–70 bands obtained with standard combination of rare and frequent cutter enzymes used in plant cDNA-AFLP analysis (Bachem et al., 1998). The high number of false positives obtained might be explained considering the high degree of intra-specific genetic variability, characteristic of a dioecious species with an obligate allogamous reproduction system. The degree of genetic polymorphism in *Cannabis sativa*, assessed for cv. Fibranova by the occurrence and frequency of RAPD marker, resulted close to 85% (Faeti et al., 1996; Forapani et al., 2001). As this study represents the first report on gene expression analysis in *Cannabis sativa*, there are no useful works to assess the impact of genetic variability and plant heterogeneity on gene expression analysis. It is however conceivable that a high degree of genetic polymorphism might affects the efficiency of techniques designed to identify differentially expressed genes, based on restriction analysis, such as the cDNA-AFLP.

Four of the identified clones putatively belong to the mRNA for a permease, for an SMT3-like protein, for the heavy chain of a kinesin 9 protein and for a Rac-GTP binding protein. Rac proteins are generally involved in the signal transduction pathways that could be activated by external or internal cues; in plants they are more represented in the meristematic tissues (Valster et al., 2000). Recently, it has been demonstrated for Rac-GTP binding protein a signalling role in auxin-regulated gene expression in Arabidopsis (Tao et al., 2002). There is some evidence of the auxin role in the differentiation of the female sex in *Cannabis sativa*, as they accumulate in female plant during development and reach high levels just before the transition to flowering (Galoch, 1980). The SMT3 protein belongs to the family of ubiquitins that are involved in the post-translational modification of the most plant's proteins. The family of kinesins comprises proteins that mediate the cytoskeleton movements. In mammals and plants, kinesins are involved in the vescicles trafficking, and in the formation of mitotic spindle (Baskin, 2000). The permeases are a heterogeneous family of membrane's proteins that mediate the ions or metabolites exchange between cells or within cell. The high expression of these genes in the female apices at the fourth node suggests that, at this stage, some metabolic processes are more active in the females than in males, either due to their specific activation in the female plants, or to their repression in the male plants. The induction of Rac-GTP binding proteins could be functional to the activation of auxin-induced genes, probably involved in the female sex differentiation. (Galoch, 1980).

It has been proposed that, from the evolutionary point of view, dioecy derived from the hermafroditism through an intermediate state of ginodioecy, and that the male genotype derived from the hermaphrodite one by the repression of the female characters; the male and female sexes would then differ by the presence of female-suppressing factors in the male plants (Charlesworth and Guttman, 1999). This theory would be confirmed by the actual transition from an intermediate state of hermafroditism to a definitive condition of dioecy, observed in many dioecious species during the ontogenesis of unisexual flowers (*Silene latifolia*, *Fragaria* spp., *Asparagus officinalis* and *Vitis* spp.; Dellaporta, 1993); another evidence supportive of this hypothesis is the possibility, in some dioecious species, of partial or complete sex reversion (Grant et al., 1994). It is therefore conceivable that in *Cannabis sativa*, the repression of female characteristics in the male plants apices implies the down-regulation of the genes coding for enzymes involved in metabolic pathways more strictly related to the differentiation of the female sex, as suggested by the results presented here.

The study of sexual differentiation in dioecious species has been often approached by the identification of sex-specific DNA markers that could be often mapped on sexual chromosomes (Donnison and Grant, 1999; Peil et al., 2003). These markers often belong to regions rich in repetitive DNA, and to LINE-like or non-LTR retrotransposons repetitive sequences (Scutt et al., 1999; Sakamoto et al., 2000; Mandolino et al., 2002). In *Silene latifolia*, the analysis of subtracted libraries from different stages of male or female flower

differentiation, allowed the isolation of *Men* 1–10 and *MROS* 1–4 genes, all male specific (except *MROS 3*), and generally coding for functions related to the male reproductive organs development or to pollen maturation (Scutt et al., 1999). Our work represents an alternative approach since gene expression in male and female apices was analyzed when there were no floral buds visible yet, and therefore differences in gene expression between the two sexes could be related to the onset of apex sexual differentiation. Though it is difficult to make a direct correlation between the induction of the differentially expressed clones and sexual differentiation, the strategy used allowed the isolation of gene sequences differentiating male and female plants at an early stage of development, and represents the first step on identification of sex-associated gene expression in *Cannabis sativa*.

Finally, this research reported for the first time a combined morphological and molecular description of a critical stage of *Cannabis* sexual differentiation. The sequences identified so far will be checked at various stage of sexual development and their full length sequences will be tracked down, allowing the cloning and the characterization of sex-related genes in hemp.

Acknowledgments

The authors would like to thank Prof. Piera Medeghini Bonatti (University of Modena and Reggio Emilia) for passing on the theoretical and practical knowledge about plant sample microscopic observation, and for the technical supplies.

This research was realized in the framework of the project "Hemp for textiles: from the production to end uses," supported by the Italian Ministry of Agriculture.

References

Altschul, S.F., W. Gish, W. Miller, E.W. Myers & D.J. Lipman, 1990. Basic local alignment search tool. J Mol Biol 215: 403–410.

Bachem, C.W.B., R.S. van der Hoeven, S.M. de Bruijn, D. Vreugdenhil, M. Zabeau & R.G.F. Visser, 1996. Visualization of differential gene expression using a novel method of RNA finger-printing based on AFLP: Analysis of gene expression during potato tuber development. Plant J 9: 745–753.

Bachem, C.W.B., R.J.F.J. Oomen & R.G.F. Visser, 1998. Transcript imaging with cDNA-AFLP: A step-by-step protocol. Plant Mol Biol Rep 16: 157–173.

Baskin, T.I., 2000. The cytoskeleton. In: B. Buchanan, W. Gruissem & R. Jones (Eds.), Biochemisty and Molecular Biology of Plants, pp. 219–221. American Society of Plant Physiologists, Rockville, Maryland.

Boecke, J.D., 1989. Transposable elements in *Saccharomyces cerevisiae*. In: D.E. Berg & M.M. Howe (Ed.), Mobile DNA, pp. 335–374. American Society of Microbiology, Washington DC.

Bredemann, G., 1938. Züchtung des Hanfes auf Fasergehaltes. Die Ergebnisse des Jahres 1937. Faserforschung 4: 239–258.

Charlesworth, D. & D.S. Guttman, 1999. The evolution of dioecy and plant sex chromosome systems. In: C.C. Ainsworth (Ed.), Sex Determination in Flowering Plants, pp. 25–33. Bios Scientific Publishers, Oxford, Washington DC.

Corpet, F., 1988. Multiple sequence alignment with hierarchical clustering. Nucl Acids Res 16(22): 10881–10890.

Donnison, I.S. & S.G. Grant, 1999. Male sex-specific DNA in *Silene latifolia* and other dioecious plant species. In: C.C. Ainsworth (Ed.), Sex Determination in Flowering Plants, pp. 73–88. Bios Scientific Publishers, Oxford–Washington DC.

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

Forapani, S., A. Carboni, C. Paoletti, V.M.C. Moliterni, P. Ranalli & G. Mandolino, 2001. Comparison of hemp varieties using random amplified polymorphic DNA markers. Crop Sci 41: 1682–1688.

Galoch, E., 1980. The hormonal control of sex differentiation in dioecious plants of hemp (*Cannabis sativa*). Acta Physiol Plant vol II (n.1): 31–39.

Grant, S., A. Houben, B. Vyskot, J. Siroky, W.H. Pan, J. Macas & H. Saedler, 1994. Genetics of sex determination in flowering plants. Dev Genet 15: 214–230.

Harley, C.B., 1997. Hybridization of Oligo(dT) to RNA on nitrocellulose. Gene Anal Tech 4: 17–22.

Hartings, H., 1999. High resolution fingerprinting of transcribed genes by means of a modified cDNA-AFLP method. Maydica 44: 179–186.

Klimyuk, V.I., B.J. Carrolm, C.M. Thomas & J.D. Jones, 1993. Alkali treatment for rapid preparation of plant material for reliable PCR analysis. Plant J 3: 493–494.

Lacombe, J.-P., 1980. Discrimination des sexes en fonction de caractères végétatifs précoces chez le Chanvre dioïque (*Cannabis sativa* L.). Physiol Vég 18: 419–430.

Mandolino, G., A. Carboni, S. Forapani & P. Ranalli, 1998. DNA markers associated with sex phenotype in hemp (*Cannabis sativa* L.). In: Proc Bast Fibrous Plants Today and Tomorrow, St Petersburg, September 28–30, pp. 197–201.

Mandolino, G., A. Carboni, S. Forapani, V. Faeti & 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 & P. Ranalli, 2002. Occurrence and frequency of putatively Y chromosome linked DNA markers in *Cannabis sativa* L. Euphytica 126: 211–216.

Mandolino, G. & P. Ranalli, 2002. The applications of molecular markers in genetics and breeding of hemp. J Ind Hemp 7: 7–24.

Mediavilla, V., M. Jonquera, I. Schmid-Slembrouck & A. Soldati, 1998. A decimal code for growth stages of hemp (*Cannabis sativa* L.). J Ind Hemp Ass 5: 65–74.

Mohan Ram, H.Y. & R. Nath, 1964. The morphology and embryology of *Cannabis sativa* L. Phytomorphology 14: 414–429.

Mohan Ram, H.Y. & R. Sett, 1982a. Modification of growth and sex expression in *Cannabis sativa* by aminoethoxyvinylglycine and etephon. Z Planzenphysiol Bd 105: 165–172.

Mohan Ram, H.Y. & R. Sett, 1982b. 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.

Peil, A., H. Flachowsky, E. Schumann & W.E. Weber, 2003. Sex-linked AFLP markers indicate a pseudoautosomal region in hemp (*Cannabis sativa* L.). Theor Appl Genet 107: 102–109.

Sakamoto, K., Y. Akiyama, K. Fukui, H. Kamada & S. Satoh, 1998. Characterization, genome size and morphology of sex chromosomes in hemp (*Cannabis sativa* L.). Cytologia 63: 459–464.

Sakamoto, K., N. Ohmido, K. Fukui, H. Kamada & S. Satoh, 2000. Site-specific accumulation of a LINE-like retrotrasposon in a sex chromosome of the dioecious plant *Cannabis sativa*. Plant Mol Biol 44: 723–732.

Sambrook, J., E.F. Fritisch & T. Maniatis, 1989. Molecular Cloning: A Laboratory Manual. Cold Spring Harbor Laboratory Press.

Scutt, C.P., S.E. Robertson, M.E. Willis, K. Kamisugi, Y. Li, M.R. Shenton, R.H. Smith, H. Martin & P.M. Gilmartin, 1999. Molecular approaches to the study of sex determination in dioecious *Silene latifolia*. In: C.C. Ainsworth (Ed.), Sex Determination in Flowering Plants, pp. 51–72. Bios Scientific Publishers, Oxford, Washington DC.

Valster, A.H., P.K. Hepler & J. Chernoff, 2000. Plant GTP-ases: The Rhos in bloom. Trends Cell Biol 10: 141–146.

Tao, L., A.Y. Cheung & H. Wu, 2002. Plant Rac-Like GTPases are activated by auxin and mediate auxin-responsive gene expression. Plant Cell 14: 2745–2760.

The Angiosperm Phylogeny Group, 2003. An update of the Angiosperm Phylogeny Group classification for the orders and families of flowering plants: APGII. Bot J Linn Soc 141: 399–436.

Westgaard, M., 1958. The mechanism of sex determination in dioecious plants. Adv Genet 9: 217–281.