# 3

# Marker assisted selection and genomics of industrial plants

Giuseppe Mandolino

C.R.A. - Istituto Sperimentale per le Colture Industriali, Via di Corticella 133, 40128 Bologna, Italy (e-mail: g.mandolino@isci.it)

## 1. Introduction

Though by far less popularly publicized than the 50th jubilee of the determination of DNA structure, celebrated three years ago, 2006 will be the 40th anniversary of the introduction of molecular markers in the genetic analysis, by Richard Lewontin in 1966. From the earliest, labour intensive and time-consuming applications to the study of the natural populations, to the massive exploitation in genome mapping of important plant species, it took therefore less than half a century for molecular markers to become a fundamental tool of theoretical and applied genetics. The oldest method for the analysis of the animal and plants' genomes is their mapping, i.e. the ordered positioning of a number of tags, acting as markers, along the entire length of the genome itself or, in the case of the eukaryotic genomes, of each of the chromosomes in which it is fragmented. The concept that two phenotypic traits can be inherited more often together rather than separately, probably dates back to the early breeding experiences. In the early decades of last century, the developments of genetic analysis led to the construction of the first genetic maps, consisting of a few tens of markers – mostly phenotypic, visible or easily scorable traits (Sturtevant 1913). It took about 50 years for a new breakthrough in the linkage mapping technologies and strategies to occur, as it was recognized that molecular tags, for example the isoenzymatic variants coded by different alleles at the



*P. Ranalli (ed.), Improvement of Crop Plants for Industrial End Uses*, 59–82.

same genetic locus, could be treated as markers as well (Lewontin and Hubby 1966), and their association with other traits, or between them, could be studied by standard genetic means.

The previously existing linkage maps, filled with only a limited number of phenotypic markers and genes characterized by segregation ratios, were rapidly enriched by the introduction of isoenzymatic markers. The following revolutions, the use of a wide choice of DNA-based markers for genetic analysis (Botstein 1980), only required 15 years after the introduction of isoenzymes, and led after a few more years to the early highly saturated human, animal and plant linkage maps. At the mid-Nineties, the maps of the major crops, including several important industrial ones (e.g. potato, sugarbeet) were densely studded with several markers, obtained with different molecular technologies, and ordered by detailed genetic analysis. Marker technology was an important tool for either positioning and tagging genes or QTLs on specific chromosome regions, and exploiting them in the marker assisted selection (MAS). It seemed to occur a strict inverse relationship between the time required to develop a new marker type, and the number of markers that each new introduction made available to the geneticists and breeders. The limit such an acceleration pointed at, was soon reached: the collapse in the cost of DNA sequencing and equipment, and the development of automatic, high-throughput technologies made available, in a rapidly increasing number of species, the ultimate marker, i.e. the gene or even the whole genomic sequences, in draft versions constantly refined and annotated; it has been a matter of only a short time from the publication of the human genomic sequence, and the complete sequencing of a model plant, the thale cress (The Arabidopsis Genome Initiative, 2000). Soon the genomes of other plants were sequenced: rice, tomato, poplar. Several international initiatives and research consortia will lead in the very next future to the availability of the entire sequence of other crop plants, though these last years made increasingly evident, for the full exploitation of the data available, the necessity of the functional annotation and characterization of the expressed and regulative sequences available (see Chapter 1 of this book).

Industrial crops were fully part of these development, and the advancement of the research in the area of marker-assisted breeding and genomics, depended on the rapid developments of the technologies and strategies necessary, and on the economic relevance of the crop, leading to more or less extensive gathering of research institutions into consortia engaged in the sequence production and/or characterization.

Some plant species, that were traditionally considered food or forage crops, can today be added to the list of typically industrial crops. Among these there are plants that can now be exploited as biomass or biofuel sources, such as sorghum or rapeseed. These new, industrial uses of

traditional forage, food or oil crops, can take further impulse from the rapidly-growing amount of genomic data accumulated since 2000. Among the several industrial crops for which highly dense genetic maps, and an high amount of genomic and functional genomic data are being accumulated, there are rapeseed, sugarbeet and hemp. These three crops are somehow paradigmatic of the evolution, along with the genomic technologies, of their end-use, and of the possible exploitation, at different levels of development, of the genomic databases. Besides their traditional use (food oil, sugar and fibre), requiring completely different industrial plants, in the latest years these three crop species can all be included in the list of actual or potential crops grown for biofuels, a rapidly expanding field of investments in Europe and America. In the present review, we will highlight some of the main advancements of the genomics and marker-assisted selection in these three industrial crop species.

## 2. Rapeseed

Rapeseed (*Brassica napus*) cultivation experienced a great expansion in the latest years, and it is today one of the leading oil crops in the world, along with soybean and palm. FAO reports a worldwide increase in the last years from 25 to 27 millions ha since 2000, with a 14% increase in production (FAOSTAT). It is likely that these increases were at least partially due to the fact that, along with the traditional uses of rapeseed, for animal feeding and vegetable oil, in the latest years the biodiesel option has been favored by public and private investments, and by specific legislative actions in several countries. Due to its long-lasting economic importance in both developed and developing countries, the state of the art of molecular markers development and genomics in this plant is highly advanced. Already in 1991, a genetic map essentially based on expressed cDNA probes, and covering the 19 linkage groups of the amphidiploid species *Brassica napus* was published, using as parentals of the mapping F2 population two "canola" varieties (Landry et al. 1991). Even in this early map, based on poorly polymorphic probes, it was possible a first assessment of the structure of the rapeseed genome, identifying duplicated regions and rearrangements of the duplicated loci, altering their linear order in different genome areas. Besides rapeseed, other Brassicaceae were thereafter extensively studied by means of molecular markers, revealing a substantially high degree of polymorphisms, detectable especially by DNA markers. Already in the mid-Nineties, it was possible to compare a number of maps for several different species belonging to the "U triangle" (U, 1935), with the only exception of *Brassica carinata* (Cheung and Landry 1996; Lagerkrantz and Lydiate 1996). The availability of the earliest maps in some of

these species, and the extensive use of RFLP and cDNAs deriving from studies on the model species *Arabidopsis thaliana*, started to pay off very soon, driving and accelerating the advancements of genomic research in all Brassicaceae (Teutonico and Osborn 1994). The following studies confirmed that an interesting feature of Brassica maps was the high number of rearrangements found, a feature that allowed to date with great accuracy the divergence of the different members of the family from Arabidopsis (Lagerkrantz 1998). The use of molecular markers in this plant family is therefore a good example of how molecular maps, besides their potential in marker-assisted breeding, have an immediate application to the study of the structure and evolution, natural or through domestication, of the plant families. The colinearity found between the different rapeseed maps and those obtained for other Brassicaceae, was almost complete, and their comparison was fully in support of the postulated relationships between the diplod and amphidiploid members of the family (Paterson et al. 2000; Lydiate et al. 1993; Pradhan et al. 2003). Besides, specific bioinformatics tools were developed for the systematic comparison of their genomes with that of the fully sequenced Arabidopsis: the comparative analysis of the *B. napus* genome led to the identification of 21 conserved units that can be traced back to the Arabidopsis genome, modified by duplication and arrangements (Parkin et al. 2005). The genomes within the Brassicaceae family are comparatively large or even, in the case of the amphidiploid, very large. It is of great importance the availability, for their genetic analysis, of a much smaller, densely mapped and entirely sequenced genome of a close relative with a smaller and simpler genome, such as Arabidopsis. The informations and the functional annotations stored in databases like TAIR (The Arabidopsis Information Resource, www.arabidopsis.org; see also Chapter 1) are useful short cuts for the identification of *B. napus* candidate genes (see for example Brunel et al. 1999, where also polymorphisms discriminating different *B. napus* lines were detected).

As for the exploitation of molecular markers, a particularly useful approach was developed by the Brassica Microsatellite initiative, a public-private consortium that developed over 2,000 microsatellite markers for *B. napus* genetic analysis. Several genome-specific microsatellites (i.e. specific for the A or C genome of the amphidiploid) were identified; besides, a number of highly polymorphic and reliable SSR loci for each linkage group were made publicy available, allowing the research community to exploit them for placing in "frame" any given new marker or trait in its unambiguously defined linkage group (Lydiate and Sharpe 2003). SSR markers were also largely used in rapeseed for studying their genetic diversity and the relationships between the different gene pools (Hasan et al. 2006), and high-throughput systems were devised, involving multiplex PCR

reactions and fluorescence-based detection of the fragments, for genotyping and pedigree analysis (Mitchell et al. 1997).

Finally, it should be pointed out that some members of the Brassica family were used to develop SRAP (Sequence Related Amplified Polymorphisms) markers; these markers are especially designed to target coding sequence in the genome. They are based on two 17 or 18 bp primers, one consisting of a 14 bp "core" sequence composed of a 11 bp long "filler" and a CCGG short tail, while the other primer has an AATT sequence. In both primers, three selective nucleotides are at the end of the sequence. In Brassica species, this type of multi-locus marker proved to be an effective strategy to amplify sequences particularly rich in coding regions; in fact, the CCGG sequence "targets" the exons of ORFs, particularly rich in G+C, while, at the opposite end, AATT sequence targets the AT-rich regions at the 3' ends of the genes. Upon sequencing of the polymorphic bands, it was shown that for most of them, many highly similar sequences in the Gene Bank were found following BLAST analysis. SRAP markers can be relatively easily converted to codominant ones, and they were found to have a good genome coverage in *B. oleracea*, where a combined AFLP-SRAP map was developed (Li and Quiros 2001). SRAP markers are also being tested, along with other more traditional marker systems, for marker-assisted selection of important traits in the rapeseed genetic improvement. Among the traits under selection with the aid of these markers, there is the identification of resistance genes; particularly relevant are the attempts of marker-assisted pyramiding of the different sources of blackleg resistance; white rust resistance, another resistance trait important to introgress into Indian mustard (*B. juncea*; Somers et al. 2002). As for partial resistance to *Sclerotinia sclerotiorum*, a mixed RFLP, AFLP, SSR and RAPD map was developed, and QTLs accounting for a large proportion of the variation were identified and mapped (Zhao and Meng 2003).

Finally, markers have also been exploited for the tagging of the genes involved in the biosynthesis of glucosinolates, the main secondary metabolite produced by all members of the Brassicaceae family (Li and Quiros 2001; Mahmood et al. 2003); also for the putative identification of the genes involved in their biosynthesis, it was extremely useful the candidate gene approach, based on the knowledge of the homologous genes of *A. thaliana*. By this approach, several genes involved in the synthesis and degradation of glucosinolates in different Brassicaceae, have been isolated and characterized (Barth and Jander 2006; Gao et al. 2004; Mahmood et al. 2003).

## 3. Sugarbeet

On February 2006, the European Council of the Ministers gave green light for the reform of the OCM sugar. The immediate consequence was a cut of the sugarbeet production in all European countries, ranging from 5% for Germany, up to 50% for Italy. One of the consequences of the reduction of the production of sugar from beets has been the necessity, for many European governments, to provide incentives for alternative uses of both the sugar factories (e.g. converted to the production of biodiesel or energy from biomasses), while for the sugar beet crop, one of the possible pathways of utilization of the sucrose was the conversion to ethanol, for biofuel production. However, for this relatively new end-use, sugar beet does not appear competitive with other crops, and the sugar beet research community will probably focus its attention on the exploitation of the biotechnology and breeding strategies for increasing sugar beet yield and quality, through the introduction of resistances to abiotic and biotic stresses, bolting and retrogradation, still a major cause of economic losses in many areas.

Sugarbeet belongs to the Chenopodiaceae family, and all the cultivated forms (sugar, chard, fodder and red beet) belong to the same species, *Beta vulgaris* ssp. *vulgaris*, included in the Section *Beta* of genus Beta (Lange et al. 1999). In the same species is also included a distinct subspecies, perfectly intermating with all the cultivated forms, *Beta vulgaris* ssp. *maritima* (the sea beet). Populations of this wild subspecies exist along most of the coastal areas of Europe, and have been repeatedly exploited in sugarbeet breeding, as they often are source of resistances to cercospora leaf spot and rhizomania, two of the major diseases causing severe production losses (Stevanato et al. 2001; Biancardi et al. 2002). The three remaining Sections of the genus Beta: *Nanae*, *Corollinae* and *Procumbentes*, only include wild species (e.g. *B. nana*, *B. patellaris*, *B. procumbens*, *B. webbiana*). In these Sections, other important sources of resistance to cercospora and to the nematode *Heterodera schachtii* exist (Luterbacher et al. 2005; Speckmann and de Bock 1982; Heijbroek et al. 1983), though their introgression in the cultivated gene pool is more difficult, due to high incompatibility with the *B. vulgaris* species. Among the earliest applications of molecular markers to Beta studies, there has been the comparison of the cultivated and wild germplasm, and the characterization of the ancestral or actual gene flow occurring between them. Isoenzymes, RFLP and cDNA probes were used to detect polymorphisms, to establish phylogenetic relationships between the different Beta taxa, and to fingerprint interspecific hybrids or addition or translocation lines obtained from crosses between incompatible *B. vulgaris* and *B. procumbens* or *B. patellaris* (Oléo et al. 1986; Nagamine et al. 1989). In these early works, genomic RFLPs were

mainly used, but the detection of polymorphisms between sugar beet varieties was quite low, at least until the use of hypervariable "minisatellite" probes, targeting several different loci in one single hybridization step, was introduced (Jung et al. 1993). These works confirmed the strict proximity of the cultivated beet gene pool with the sea beet, while other species, belonging to different Sections of the genus Beta, were much more distantly related, based on Nei and Li similarity coefficients: it was reported that only 34% of the RFLP probes used cross-hybridized between these species (Jung et al. 1993). Molecular markers were extensively used for detailed studies of the gene flow within wild *B. vulgaris* ssp. *maritima* populations (Raybould et al. 1996; Tufto et al. 1998), and particular relevance assumed the studies of the gene flow (as estimated by AFLP or RFLP markers) between the bolting or flowering transgenic sugar beets, and the wild or weedy populations. In these studies, the risk of the transgene escape from cultivated fields was assessed (Desplanque et al. 1999), and the genetic erosion of the wild beet populations was evaluated (Bartsch et al. 2002). Gene flow studies, and their impact on the risk assessment protocols and on the legislations regulating the introduction of transgenic varieties, were indeed one of the main driving force for the discovery and application of several microsatellite markers in the Beta genus (Cureton et al. 2002; Richards et al. 2004). Although, strictly speaking, this research field is not real marker-assisted selection, it can be considered as a new field of application, a "marker-assisted risk assessment" strategy, that will probably play an increasingly important role in a number of cases of decision-making processes related with the introduction of transgenic varieties and their coexistence with traditional crops.

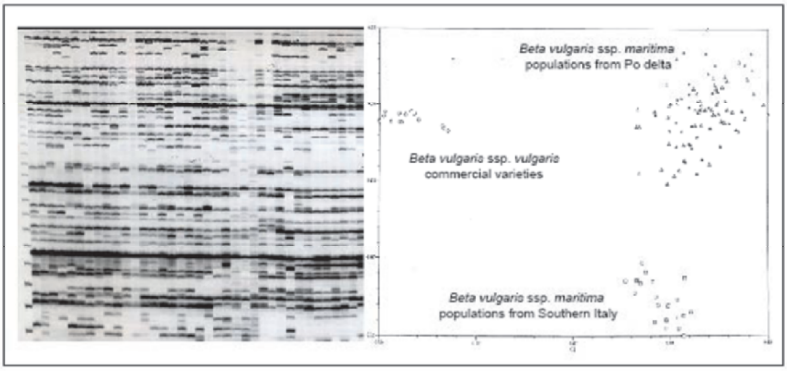


**Fig. 3.1.** AFLP patterns (left) and principal coordinate analysis of AFLP data (right) for varieties and wild populations of *Beta vulgaris* (ssp. *vulgaris* and *maritima*).

The existence in Beta of some Section-specific genomic probes, corresponding to repeated, satellite sequences, had far-reaching consequences. In fact, since 1990, probes hybridizing specifically to satellite DNA of members of the Procumbentes Section were used in the marker-assisted identification of the sugarbeet lines carrying the monosomic addition from *B. procumbens*, enclosing the resistance gene to *Heterodera schachtii* (Schmidt et al. 1990; Salentijn et al. 1992). These Section-specific markers were used to "bracket" the resistance gene, and then to clone it (Klein-Lankhorst et al. 1994; Cai et al. 1997). Marker-assisted tagging of specific chromosome segments was also carried out in screenings of monosomic additions and translocation lines from interspecific crosses (*B. vulgaris* x *B. procumbens* or *B. vulgaris* x *B. corolliflora*), for the isolation of other sources of nematode resistance (Heller et al. 1996; Kleine et al. 1998; Gao et al. 2001).

Rhizomania is another important disease for which sugar beet breeding has reached, in the last 15 years, several successes. The introgression of the resistance trait appears the main strategy to obtain high yields and root quality in the rhizomania-infected soils (Asher 1993; Biancardi et al. 2002). The breeding work took advantage of the availability of greenhouse tests based on a positive correlation between rhizomania resistance and virus content, as evaluated by ELISA assay of the roots of artificially-infected plantlets. This test was therefore the basis of many breeding programs, but was also useful for establishing an association between the various resistance sources, and the different linkage groups and markers in different mapping populations. The "Holly type" resistance was mapped in the linkage group IV, with an RFLP marker about 7 cM from *Rr1* gene (Barzen et al. 1992); bulk segregant analysis (BSA; Michelmore et al. 1991) was used to screen for markers associated to this same locus, and RAPD and SCAR markers were obtained (Barzen et al. 1997). Interestingly, these markers also were able to identify an apparently different source of resistance, originating from *B. vulgaris* ssp. *maritima*, and introgressed in the sugar beet line R104 (Biancardi et al. 2002). Further research, also based on map comparison and the construction of consensus maps (Schumacher et al. 1997), revealed that another associated resistance, deriving from a further *B. vulgaris* ssp. *maritima* accession, WB42, was associated but distinct, mapping at a different locus called *Rz2*, to keep it distinct from the Holly/R104 one (called *Rz1*; Scholten et al. 1999). In the case of rhizomania, molecular markers, therefore, were not only useful for marker-assisted selection (with the limit that most of the available markers are protected by proprietary rights, as they have been developed by private companies), but also contributed important insights on the long-debated issue of the relationships between the different resistance sources (Biancardi et al. 2002). Efforts were also made to combine different linkage

maps from different populations, each contributing specific resistance or quality traits, and segregating for QTL for yield-related traits (Weber et al. 1999; 2000). However, more recently the potential for marker-assisted selection for yield and quality has also been explored by a functional genomic approach (see below).

As for the markers for disease resistances, a different approach recently applied to sugar beet illustrates well the spreading of non-mapping strategies, made possible by the increasingly wide number of expressed sequences available in public repositories. Sequencing of resistance genes against several pathogens and in several plant species, made clear that shared domains exist. Therefore, it has been possible to design primers and to isolate resistance gene analogues, determining their functional characteristics, mapping them on the existing linkage maps, and associating them to mapped resistance loci, as it has been the case for rhizomania and cercospora leaf spot resistance (Hunger et al. 2003). This approach can be particularly promising for identification of candidate genes for the resistance to *Cercospora beticola*, the causal agent of cercospora leaf spot; this resistance trait showed in selection programs a slow rate of gains, as it is typical of quantitative traits, and it is inversely correlated with the yield parameters (Skaracis and Biancardi 2000). Difficulties in the development of effective markers for this trait are due to problems in exploiting a reliable *in vitro* method for resistance evaluation, and on the number and chromosomal location of QTLs associated, that are different in different studies, and even different depending upon the inoculation conditions (artificial conditions or naturally infested fields) (Weiland and Koch 2004).

Among the several sugar beet linkage maps published, a special mention deserves the one from Pillen et al. (1993), as this map includes evidences of close association (and therefore, possible exploitability in the breeding) with the fertility restoration gene, a very important trait in the development of O-type lines. These latter are plants capable of "maintaining" the male sterility, and therefore essential for the multiplication of the male sterile lines, in turn used as parental for hybrid seed production. The "restoration of fertility" locus, at which maintainer lines are endowed with the recessive allele x, was mapped terminally on linkage group III, about 9 cM from an RFLP marker; however, it should be pointed out that current genetical models suggest the existence of two distinct loci for fertility restoration (X and Z), on two different linkage groups. The importance of this trait in the sugar beet industry is confirmed by the amount of studies carried out in the last ten years, often with markedly different approaches. QTLs were described on linkage group III and IV in segregating populations screened by RFLP markers, and explaining from 72 to 79% of the variation observed; this analysis suggested the existence of minor alleles in all mapping populations (Hjerdin-Panagopoulos et al. 2002). In another study, BSA was

used, and several AFLP and RAPD markers were identified, three of which were confirmed to co-segregate with the restoration of fertility (Hagihara et al. 2005). Despite these successes, the gene(s) responsible of the maintaining phenotype have not yet been cloned, their functional analysis is still to come, and details of the interactions of nuclear gene products with the cytoplasmic-encoded male sterility determinants have not yet been clarified for sugar beet.

Despite the transition from genetic to genomic analysis has been in sugar beet less advanced than in other crop species, huge collections of expressed genes have been deposited in the gene banks (e.g. TIGR, www.tigr.org; The Sugarbeet EST Database http://genomics.msu.edu/sugarbeet/) and are presently being screened for functional analysis and for genes possibly correlated to agronomically valuable traits. The progressive accumulation of sequence informations on sugar beet was exploited for the enrichment of the existing genetic maps with SNPs, or for the fingerprinting of sugar beet genotypes. Particularly interesting is the approach linking fragments of expressed genes, often of some relevance from the industrial point of view, and the SNPs found in sugar beet germplasm within these sequences, and anchored to the existing molecular maps (Schneider et al. 1999). SNPs were also identified in sugar beet by gel-based assays, such as single strand conformation polymorphisms (SSCP) and heteroduplex analysis (HA) (Schneider et al. 2001); the markers developed in this way proved to have a polymorphism information content (PIC) around 0.47, not as high as SSR, but superior to AFLP or RFLP markers. Sugar beet SNPs identified by these approaches could also be located in the different linkage group; multiplex assays were developed to fingerprint genotypes simultaneously at different loci in single PCR reactions for each linkage group, and to detect these polymorphisms by fluorescence in capillary electrophoresis, in a high-throughput, semi-automated system (Möhring et al. 2004).

In general, these examples show that marker-assisted selection is at a quite mature stage in sugar beet, for different industrially important traits. Besides, in the last few years, sugar beet functional genomics took advantage of several specific international projects and cooperation's, and it is likely that in the next future the number of disease resistance, quality and yield traits, for which the gene(s) and alleles responsible for the variation are completely characterized, will substantially increase.

## 4. Hemp

Hemp, *Cannabis sativa*, is probably one of the most ancient non food crops cultivated by mankind. Hemp is endowed of a particularly oil-rich

seed, and of species-specific secondary compounds accumulating in the inflorescences (cannabinoids); because of these characteristics, hemp has been considered a candidate for biofuel production and pharmaceutical applications respectively; besides, a number of novel uses for its fibre have been developed (e.g. in paper and automotive industry). Several initiatives and industrial enterprises, aimed at demonstrating the feasibility and sustainability of such uses, and the consequent worth of the re-discovery of this ancient crop also for more modern industrial uses, have been carried out, often with good commercial success (Guy and Stott 2004).

The first concept that is important to stress, when dealing with this typically industrial species, is that most taxonomists recognize the existence of a single species in the Cannabaceae family, namely *C. sativa*. Therefore, the distinction between *C. sativa* as a fibre crop -deserving sustain for its beneficial added values such as low chemical inputs, eradication of some soil-borne pathogens, potentially complete utilization of all its parts, etc.- and *C. indica* as an illegal species, genetically different from the first, but phenotypically undistinguishable from it, and only suited for drug abuse, should be overcome. The two forms of Cannabis, that can be at most considered different subspecies, but more correctly different races or varieties (chemotypes: see below), are perfectly interfertile, and they should be considered belonging to a single gene pool (Small et al. 1976). However, the recent history of hemp cannot be made without considering two facts: first, forms of *Cannabis sativa* L., endowed with levels of a single secondary compound, $\Delta^9$-tetrahydrocannabinol (THC), above 0.30% of the inflorescence dry weight, are considered illegal (and not eligible for EU subsidies) and cannot be cultivated; and second, that the close relatedness of the tradition-revivifying, environmentally sustainable and agronomically beneficial fibre hemp with the marihuana, has been used by opinion groups as an argument in favor of its liberalization. As an obvious consequence, far from limiting itself to industrial and technical issues, the question of hemp cultivation in Europe and America has been strongly conditioned by the analytical and legislative issues, in most cases hampering the re-discovery and diffusion of this plant for a number of industrial exploitations (Ranalli et al. 1999; Karus and Vogt 2004; Guy and Stott 2004). A further consequence of this paradox was that forensic scientists, rather than breeders, set up the basic knowledges leading to the exploitation of molecular markers in Cannabis research (see Mandolino and Carboni 2004, for a review of forensic applications). Basically, these early applications aimed at the development of methods recognizing the presence of the illicit "cannabis" plant material, distinguishing it from other plant sources, and at the analysis of the variability within illegal plant material, to reconstruct phylogenies of the different drug strains, and the routes of their diffusion. For example, a short intergenic sequence located between the chloroplast genes for the

LEU and PHE transport RNAs can act, upon amplification, as a marker for the recognition of *Cannabis sativa* DNA (Wilkinson and Linacre 2000). The internal transcribed spacers I and II (ITS1 and ITS2) of the nuclear ribosomal genes were used in forensic applications to univocally identify *Cannabis sativa* (Siniscalco Gigliano et al. 1997).

These examples made confirm the fact that forensic research had a major role in the development of molecular markers in Cannabis research: AFLP analysis specific for marihuana samples (Coyle et al. 2003), and RAPD (Gillan et al. 1995; Jagadish et al. 1996) and ISSR (Kojoma et al. 2002) markers were used in attempts to establish a direct relationship between the cannabinoid type of the plant, and the markers identified. However, no correlations between any specific markers and the gas-chromatographic or HPLC cannabinoid profiles of different Cannabis chemotypes were identified in these studies, and the development of chemotype-specific markers has only recently been achieved (see below).

Microsatellites markers were described and applied to the study of Cannabis population and their variability only very recently (Alghanim and Almirall 2003; Gilmore and Peakall 2003; Hsieh et al. 2003). In hemp, the most common repeated motif is the dinucleotide GA/CT; di- and trinucleotide repeats were the most frequent, with an allele number ranging from 2 up to 28 (Hsieh et al. 2003). These markers were useful in describing the genetic relatedness of the Cannabis accessions considered, but no marker associated with the chemotype, and suitable for marker-assisted marijuana identification were found (Mandolino and Carboni 2004).

Variation is a key feature of *Cannabis sativa*, and this makes its fingerprinting and mapping essentially a statistical task. SSR and RAPD markers were used for genotyping, and the variability found analyzing varieties, populations and accessions was very high; in one study, out of 93 plants examined by SSRs, only four (belonging to the same drug accession) were not distinguishable from each other (Gilmore et al. 2003), and also RAPD markers showed reproducibly high levels of variation (Faeti et al. 1996). Therefore, a first level of analysis is variation assessment within varieties, populations or accessions, and between them; other levels of analysis are then possible (between and within chemotypes, sexes, etc.). Analysis of Molecular Variance (AMOVA, Excoffier et al. 1992) has been a widely used tool for analysis of marker variation, especially in natural and allogamous populations. In Cannabis, the majority of the observed marker variation (from 50 to 73%) was attributable to individual differences within accessions (Forapani et al. 2001; Gilmore and Peakall 2003). If the accessions are roughly divided into the two groups “drug” and “non drug” (chemotype I and III, see below), only 6% of the variation was attributable to the chemotype: it was concluded that there was "no clear split or defined boundary between drug and non drug materials". Conceptually similar

results had been obtained also by Forapani et al. (2001), that confirmed the very high degree of variation within the cultivars or accessions previously identified, ranging from 31% of polymorphic loci in an inbred female Cannabis strain, up to 78% in Fibranova, a cross-bred fibre variety. Forapani et al. (2001) also found that the majority (66%) of the markers had a calculated $F_{ST}$ value below the average (0.48); the authors interpreted this finding as a confirmation of the existence of a widely shared gene pool in Cannabis, with limited cultivar boundaries and relatively poor loci segregation between different populations.

The marker variation observed was exploited to construct the few available molecular maps of *Cannabis sativa*, including RAPD or AFLP markers (Carboni et al. 2000; Flachowski et al. 2001; Mandolino and Ranalli 2002), but lacking until now agronomically relevant traits, like fibre content or quality, or monoecy; the genetic basis of these traits are indeed still poorly understood. The available maps still have limited markers density, and no attempts to define anchor markers, or to establish relationships between the molecular maps and the Cannabis chromosomes has been done yet, with the only exception of the Y chromosome and of the male-associated markers (see below). The extremely high variability found is a general problem in many allogamous crop species, but it might reveal particularly important in hemp.

Sex and monoecy are particularly important traits to score during the selection programs for industrial hemp. Dioecious varieties, made of male and female plants, are generally high-yielding, more disease-tolerant, but less uniform (a strong sexual dimorphism is present) and less amenable to mechanical harvesting than monoecious varieties. This latter type of varieties (presently the most diffused and cultivated in Europe) are made of variable proportions of female and monoecious plants, have high seed yields (because all plants contribute to seed formation), high habitus uniformity (favoring mechanical harvesting), but a narrower genetic base (due to the selective pressure necessary to maintain the monoecious trait in a significant proportion of the plants), a lower biomass yield, and the necessity of strict isolation and seed batch control during seed multiplication, due to the lower competitiveness of monoecious pollen compared with pollen from contaminating male (dioecious) plants. The practice of hemp breeding sometimes requires the identification of the sexes (male, female, monoecious), with the male sex determination being especially important for the different strategies of improvement in dioecious hemp, and for seed quality controls in monoecious hemp. Early identification of the male plants is today an example of completely effective and 100% reliable marker-assisted selection. In fact, since 1999, strictly male-specific (probably Y-chromosome located) molecular markers were described by several groups, obtained by RAPD markers (Mandolino et al. 1999), then

transformed in SCAR markers upon sequencing of the discriminating fragments (Mandolino et al. 1999; Mandolino et al. 2002); also AFLP and SSR sex-specific markers were later described (Peil et al. 2003; Rode et al. 2005). In most cases, the strict association to the male phenotype was attributable to the localization of the marker on the region of the Y chromosome excluded from recombination with homologous areas of the X chromosome (Mandolino et al. 2002), while in another case the marker was located on the Y chromosome by direct FISH analysis (Sakamoto et al. 2000). In all cases reported where these male-specific markers have been sequenced, a high similarity between them and with LINE-like retrotransposons was reported (Mandolino et al. 1999; Sakamoto et al. 2000; Sakamoto et al. 2005).

The marker named $SCAR_{400}$, described in Mandolino et al. (1999), was routinely used in selection programs for both dioecious and monoecious hemp, demonstrating its full potential and association; direct PCR amplification protocols of tiny tissue fragments were devised, simplifying the scoring of the male sex in a huge number of samples (Mandolino and Ranalli 2002); this marker was tested and used also by commercial seed companies with excellent results.

One peculiar exploitation of sex-linked markers, came from the possibility to score at an extremely early stage of differentiation the genetically determined male plants, well before any sex-related structure was visible. This opportunity allowed the study of genome-wide gene expression at very early stages of the sexual determination in male and female sexes. By this "marker-assisted" gene expression analysis, it was possible to identify by cDNA-AFLP and clone some sequences differentially expressed in male and female plants (Moliterni et al. 2004).

Monoecy is another very important character for which a marker system would be useful. However, the above described male-specific marker is also suitable for keeping under strict control the “contaminating” male plants in a monoecious stand or seed lot. The exploitation of this marker could allow the introduction, by the seed companies commercializing monoecious seed, of the “elite” or “superelite” terms for seed batches, depending on the number of contaminating males expected (Mandolino and Carboni 2004). This marker can be therefore already be exploited both in MAS and in seed quality certifications of monoecious varieties. Monoecy and flowering time are today among the main targets for genetic analysis and marker-assisted breeding; the availability of male- and of monoecious-specific markers, possibly combined in a single PCR assay, would allow the complete identification of the sexual phenotype of plants belonging to any variety or breeding lines.

As anticipated at the beginning of this section, one of the main targets of modern industrial hemp breeding is the early identification of the

chemotype, i.e. of the cannabinoid(s) a given genotype is endowed with. Today, the determination of the chemotype and of the absolute amounts of each cannabinoid is carried out by labor-intensive gas-chromatography or HPLC; such analysis is mandatory to get the EU subsidies by industrial hemp growers: THC levels in a sample of plants of the stand must be below 0.30% of the inflorescence dry weight. Though the necessary techniques are susceptible of being speeded up and automated, the development of specific molecular markers predictive of the type and (ideally) amount of each of the cannabinoids a plant can manufacture, would be very useful for hemp breeders. Five main chemotypes (I through V) are recognized and/or proposed today (Small and Beckstead 1973; Mandolino and Carboni 2004); the mode of inheritance of the first three (I: prevalent cannabidiol, CBD; III: prevalent tetrahydrocannabinol, THC; and II: mixed CBD and THC), has been recently extensively studied and clarified (de Meijer et al. 2003; Mandolino et al. 2003; Mandolino 2004). CBD prevalence in the plant (as observed, for example, in most modern fibre cvs.) is determined by homozygosity at a single locus, *B*, of the allele $B_D$, while prevalence of THC in the plant (e.g. in the drug varieties) is largely determined by homozygosity at the same locus for the $B_T$ allele. Therefore, the chemotype is simply determined by the allelic constitution at one single locus, while the total amount of cannabinoids can be considered a quantitative trait; the combination of these two determination types, yields the specific cannabinoid's profile of a genotype (Mandolino 2004). The other two chemotypes proposed (IV: prevalent CBG, and V: zero cannabinoids) are still poorly characterized, though recently genetical models have been proposed also for chemotype IV (de Meijer and Hammond 2005). Markers associated to the allelic status at the *B* locus ($B_D/B_D$, $B_D/B_T$ or $B_T/B_T$) have been developed by bulk segregant analysis of segregating F2 progenies (de Meijer et al. 2003; Mandolino et al. 2003).

The genes responsible for the synthesis of the two main cannabinoids, the non-psychoactive CBD and the psychoactive THC, have been characterized, and several sequences related with the enzyme activities involved are deposited in the Gene Bank (Sirikantaramas et al. 2004). The availability of the gene sequences involved in the synthesis of the two main cannabinoids, and the knowledge of their characteristics, has also led to the development of highly predictive sequence-based PCR markers for the chemotype (Kojoma et al. 2005); besides, marker systems based on multiplex

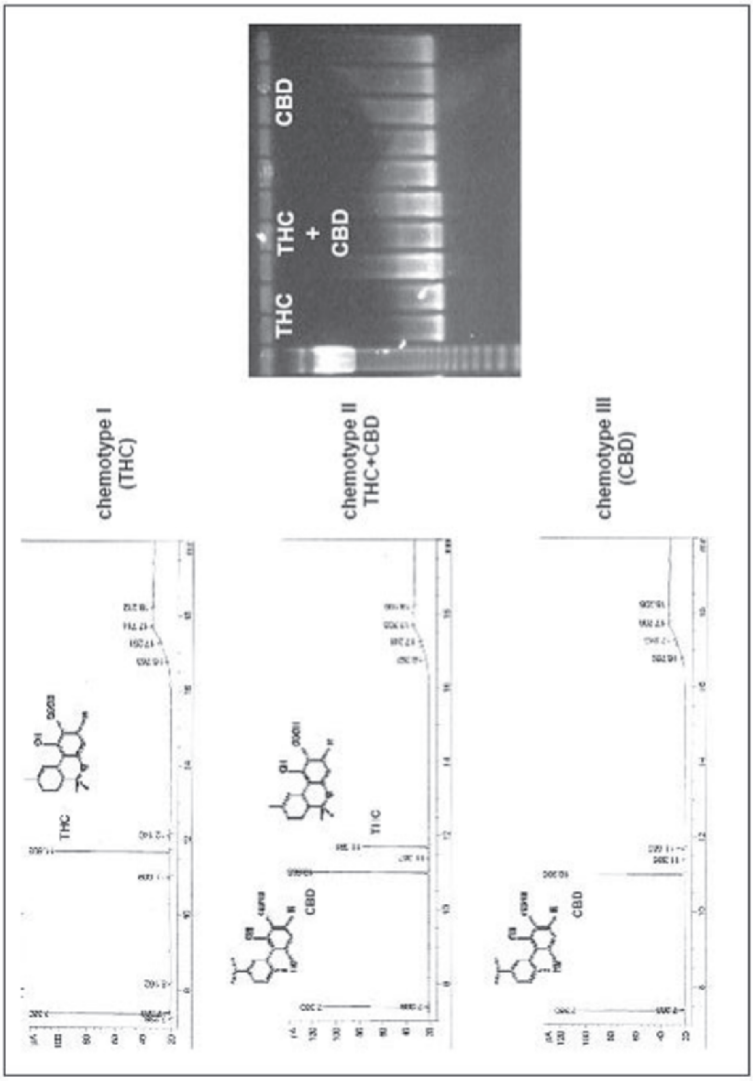


**Fig. 3.2.** The marker-assisted selection for chemotype in hemp.

PCR amplification, capable of identifying both alleles (and therefore either homozygous or heterozygous genotypes at the *B* locus) have been developed; this particularly effective marker system has been extensively used in the marker-assisted elimination of THC-containing plants (genotyped $B_T/B_T$) from newly-developed hemp varieties. As expected, these markers, very much like the male-specific markers described above, have a 100% efficiency of correct identification of the genotype at the *B* locus, and also gives indirect evaluation of the absolute THC content of any Cannabis plant (Pacifico et al. 2006). Besides, the same marker can be exploited for selecting high-THC producing plants, today used as a source of this important active principle by pharmaceutical industry (Guy and Stott 2004).

Traditionally, the main product of industrial hemp cultivation was its fibre; the genetics and functional genetics of fibre production is still in its infancy, not only in hemp, but also in other fibre crops. It is likely that the study of fibre characteristics and of their development will play a major role in marker assisted selection for this trait; the functional genomics approach is at its beginning, but is already generating a number of interesting results, both from the point of view of the basic research and of the development of diagnostic tools (see chapter 6 of this book for a review of fibre genomics).

## 5. Conclusions

In this review, the main advancements achieved in marker-assisted selection and genomics of three industrial crops have been highlighted. These three crops represent somehow three different stages of maturity of the markers' exploitation for genetic improvement. Rapeseed has a well defined and consolidated body of knowledges about the structure and organization of its genome and the several genes relevant from the point of view of industrial utilization, and it can exploit the knowledge of the complete sequence of *Arabidopsis thaliana*; sugar beet is less characterized, especially from the functional genomics point of view, but recent research efforts, and the fact that sugar beet is becoming a model species for root crops, are rapidly filling the gap; hemp genomics still is in its infancy, but a number of important traits, such as sex and chemotype are already characterized, and effective tools for marker assisted breeding have already been developed. It can be foreseen that in the next few years, pushed by

the recent new opportunities of industrial utilization of these crops, important breakthrough in genomics and genetic mapping of other important traits will occur.

## References

Alghanim HJ & Almirall JR (2003) Development of microsatellite markers in *Cannabis sativa* for DNA typing and genetic relatedness analysis. Anal Bioanal Chem 376: 1225-1233

Asher MJC (1993). Rhizomania. In: Cooke DA and Scott RK (eds) The sugar beet crop. Chapman and Hall, London, pp 311-346

Barth C and Jander G (2006) Arabidopsis myrosinases TGG1 and TGG2 have redundant function in glucosinolate breakdown and insect defence. Plant J 46: 549-562

Bartsch D, Stevanato P, Lehnen M, Mainolfi A, Mücher T, Moschella A, Driessen S, Mandolino G, Hoffmann A, De Biaggi M, Wehres U, Saeglitz C and Biancardi E (2002) Biodiversity of sea beet in Northern Italy. In: Proc. of the 65th IIRB Congress, Brussels, pp 171-180.

Barzen E, Mechelke W, Ritter E, Seitzer JF and Salamini F (1992) RFLP markers for sugar beet breeding: chromosomal linkage maps and location of major genes for rhizomania resistance, monogermy and hypocotyl colour. Plant J 2: 601-611

Barzen E, Stahl R, Fuchs E, Borchardt DC and Salamini F (1997) Development of coupling-repulsion phase SCAR markers diagnostic for the sugar beet *Rr1* allele conferring resistance to rhizomania. Mol Breeding 3: 231-238

Biancardi E, Lewellen RT, De Biaggi M, Erichsen AW and Stevanato P (2002). The origin of rhizomania resistance in sugar beet. Euphytica 127: 383-397

Botstein D, White RL, Skolnik M and Davis RW (1980) Construction of a genetic linkage map in man using restriction fragment length polymorphism. Am J Hum Genet 32: 314-331

Brunel D, Froger N and Pelletier G (1999) Development of amplified consensus genetic markers (ACGM) in *Brassica napus* from *Arabidopsis thaliana* sequences of known biological function. Genome 42: 387- 402

Cai D, Kleine M, Kifle S, Harloff HJ, Sandal NN, Marcker KA, Klein-Lankhorst RM, Salentijn EMJ, Lange W, Stiekema WJ, Wyss U, Grundler FMW and Jung C (1997) Positional cloning of a gene for nematode resistance in sugar beet. Science 275: 832-834

Carboni A, Paoletti C, Moliterni VMC, Ranalli P and Mandolino G (2000) Molecular markers as genetic tools for hemp characterization. In: Proc. Bioresource Hemp, Wolfsburg

Cheung WY and Landry BS (1996) Current status of genome mapping in the Cruciferae. In: Paterson AH (ed) Genome mapping in plant. Academic Press, 193-210

Coyle HM, Shutler G, Abrams S, Hanniman J, Neylon S, Ladd C, Palmbach T and Lee HC (2003) A simple DNA extraction method for marijuana samples used in amplified fragment length polymorphism (AFLP) analysis. J Forensic Sci 48: 343-347

Cureton AN, Burns MJ, Ford-Lloyd BV and Newbury HJ (2002) Development of simple sequence repeat (SSR) markers for the assessment of gene flow between sea beet (*Beta vulgaris* ssp. *maritima*) populations. Mol Ecology Notes 2: 402-403

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

de Meijer EPM and Hammond KM (2005) The inheritance of chemical phenotype in *Cannabis sativa* L. (II): Cannabigerol predominant plants. Euphytica 145: 189-198

Desplanque B, Boudry P, Broomberg K, Samitou-Laprade P, Cuguen J and van Dijk H (1999) Genetic diversity and gene flow between wild, cultivated and weedy forms of *Beta vulgaris* L. (Chenopodiaceae), assessed by RFLP and microsatellite markers. Theor Appl Genet 98: 1194-1201

Excoffier L, Smouse PE and Quattro JM (1992) Analysis of molecular variance inferred from metric distances among DNA haplotypes: application to human mitochondrial DNA restriction data. Genetics 131: 479-491

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

Flachowsky H, Schumann E, Weber WE and Peil A (2001) Application of AFLP for the detection of sex-specific markers in hemp. Plant Breeding 120: 305-309

Forapani S, Carboni A, Paoletti C, Moliterni VMC, Ranalli P and Mandolino G (2001) Comparison of hemp (*Cannabis sativa* L.) varieties using Random Amplified Polymorphic DNA markers. Crop Sci 41: 1682-1689

Gao D, Guo D and Jung C. (2001) Monosomic addition lines of Beta corolliflora Zoss in sugar beet: cytological and molecular-marker analysis. Theor Appl Genet 103: 240-247

Gao M, Li G, Yang B, McCombie WR and Quiros CF (2004) Comparative analysis of a Brassica BAC clone containing several major aliphatic glucosinolate genes with its corresponding Arabidopsis sequence. Genome 47: 666-679

Gillan K, Cole MD, Linacre A, Thorpe JW and Watson ND (1995) Comparison of *Cannabis sativa* by Random Amplified Amplified of Polymorphic DNA (RAPD) and HPLC of cannabinoids: a preliminary study. Sci Justice 35: 169-177

Gilmore S and Peakall R (2003) Isolation of microsatellite markers in *Cannabis sativa* L. (marijuana). Mol Ecol Notes 3: 105-108

Gilmore S, Peakall R and Robertson J (2003) Short Tandem Repeat (STR) DNA markers are hypervariable and informative in *Cannabis sativa*: implications for forensic investigations. Forensic Sci Intl 131: 65-74

Guy GW and Stott CG (2004) Cannabinoids for the pharmaceutical industry. Euphytica 140: 83-93

Hagihara E, Itchoda N, Habu Y, Iida S, Mikami T and Kubo T (2005) Molecular mapping of a fertility restorer gene for Owen cytoplasmic male sterility in sugar beet. Theor Appl Genet 111: 250-255

Hasan M, Seyis F, Badani AG, Pons-Kühnemann J, Friedt W, Lühs W and Snowdon RJ (2006) Analysis of genetic diversity in the *Brassica napus* L. gene pool using SSR markers. Genet Res Crop Evol 53: 793-802

Heijbroek W, Roelands AJ and de Jong JH (1983) Transfer of resistance to beet cyst nematode from *Beta patellaris* to sugar beet. Euphytica 32: 287-298

Heller R, Schondelmaier J, Steinrücken G and Jung C (1996) Genetic localization of four genes for nematode (*Heterodera schachtii* Schm.) resistance in sugar beet (*Beta vulgaris* L.). Theor Appl Genet 92: 991-997

Hjerdin-Panagopoulos A, Kraft T, Rading IM, Tuvesson S and Nilsson NO (2002) Three QTL regions for restoration of Owen CMS in sugar beet. Crop Sci 42: 540-544

Hsieh HM, Hou RJ, Tsai LC, Wei CS, Liu SW, Huang LH, Kuo YC, Linacre A and Lee JCI (2003) A highly polymorphic STR locus in *Cannabis sativa*. Forensic Sci Intl 131: 53-58

Hunger S, Di Gaspero G, Möhring S, Bellin D, Schäfer-Pregl R, Borchardt DC, Durel CE, Werber M, Weisshaar B, Salamini F and Schneider K (2003) Isolation and linkage analysis of expressed disease-resistance gene analogues of sugar beet (*Beta vulgaris* L.). Genome 46: 70-82

Jagadish V, Robertson J and Gibbs J (1996) A RAPD analysis distinguished *Cannabis sativa* samples from different sources. Forensic Sci Int 79: 113-121

Jung C, Pillen K, Frese L, Fähr S and Melchinger AE (1993) Phylogenetic relationships between cultivated and wild species of the genus Beta revealed by DNA "fingerprinting". Theor Appl Genet 86: 449-457

Karus M and Vogt D (2004) European hemp industry: cultivation, processing and product lines. Euphytica 140: 7-12

Kleine M, Voss H, Cai D and Jung C (1998) Evaluation of nematode-resistant sugar beet (*Beta vulgaris* L.) lines by molecular analysis. Theor Appl Genet 97: 896-904

Klein-Lankhorst RM, Salentijn EMJ, Dirkse WG, Arens-de Reuver M and Stiekema WJ (1994) Construction of a YAC library from a *Beta vulgaris* fragment addition and isolation of a major satellite DNA cluster linked to the beet cyst nematode resistance locus $Hs1^{pat-1}$. Theor Appl Genet 89: 426-434

Kojoma M, Iida O, Makino Y, Sekita S and Satake M (2002) DNA fingerprinting of *Cannabis sativa* using Inter-Simple Sequence Repeat (ISSR) amplification. Planta Medica 68: 60-63

Kojoma M, Seki H, Yoshida S and Muranaka T (2005) DNA polymorphisms in the tetrahydrocannabinolic acid (THCA) synthase gene in "drug-type" and "fiber-type" *Cannabis sativa* L. For Sci Intl, in press.

Lagerkrantz U and Lydiate DJ (1996) Comparative genome mapping in Brassica. Genetics 144: 1903-1910

Lagerkratz U (1998) Comparative mapping between *Arabidopsis thaliana* and *Brassica nigra* indicates that Brassica genomes have evolved through extensive genome replication accompanied by chromosome fusion and frequent rearrangements. Genetics 150: 1217-1228

Landry, BS, Hubert N and Etoh T (1991) A genetic map for *Brassica napus* based on restriction fragment length polymorphisms detected with expressed DNA sequences. Genome 34: 543-552

Lange W, Brandenburg WA and de Bock TSM (1999). Taxonomy and cultonomy of beet (*Beta vulgaris* L.). Bot J Linn Soc 130: 81-96.

Lewontin RC and Hubby JL (1966) A molecular approach to the study of genetic heterozygosity in natural populations. II. Amount of variation and degree of heterozygosity in natural populations of *Drosophila pseudoobscura*. Genetics 54: 577-594

Li G and Quiros CF (2001) Sequence-related amplified polymorphism (SRAP), a new marker system based on a simple PCR reaction: its application to mapping and gene tagging in Brassica. Theor Appl. Genet 103: 455-461

Luterbacher MC, Asher MJC, Beyer W, Mandolino G, Scholten OE, Biancardi E, Mechelke W, Slyvchenko O, and Frese L (2005) Sources of resistance to diseases of sugar beet in related Beta germplasm: II. Soil-borne diseases. Euphytica 141: 49-63

Lydiate D, Sharpe A, Lagerkrantz U and Parkin I (1993) Mapping the Brassica genome. Outlook Agric 2. 85-92

Lydiate D and Sharpe A (2003) Aligning genetic maps of *Brassica napus* using microsatellite markers. In: Proc. of the PAG-XI, Plant and Animal Genomes Conference, San Diego, p 473.

Mahmood T, Ekuere U, yeh F, Good AG and Stringam GR (2003) Molecular mapping of seed aliphatic glucosinolates in *Brassica juncea*. Genome 46: 753-760

Mandolino G (2004) Again on the nature of inheritance of chemotype. Letter to the editor. J Ind Hemp 9: 5-7

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

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

Mandolino G and Ranalli P (2002) The application of molecular markers in genetics and breeding of hemp. J Ind Hemp 7: 7-23

Mandolino G, Bagatta M, Carboni A, Ranalli P and de Meijer E (2003) Qualitative and quantitative aspects of the inheritance of chemical phenotype in Cannabis. J Ind Hemp 8: 51-72

Mandolino G and Carboni A (2004) Potential of marker assisted selection in hemp genetic improvement. Euphytica 140: 107-120

Michelmore RW, Paran I and Kesseli RV (1991). Identification of markers linked to isease-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.

Mitchell SE, Kresovich S, Jester CA, Hernandez CJ and Szewc-McFadden AK (1997) Application of multiplex PCR and fluorescence-based, semi-automated allele sizing technology for genotyping plant genetic resources. Crop Sci 37: 617-624

Möhring S, Salamini F and Schneider K (2004) Multiplexed, linkage group-specific SNP marker sets for rapid genetic mapping and fingerprinting of sugar beet (*Beta vulgaris* L.). Mol Breeding 14: 475- 488

Moliterni VMC, cattivelli L, Ranalli P and Mandolino G (2004) The sexual differentiation of Cannabis sativa L.:a morphological and molecular study. Euphytica 140: 95-106

Nagamine T, Todd GA, McCann KP, Newbury HJ and Ford-Lloyd BV (1989) Use of restriction fragment length polymorphism to fingerprint beets at the genotype and species levels. Theor Appl Genet 78: 847-851

Oléo M, van Geyt JPC, Lange W and de Bock TSM (1986) Investigations on an intespecific hybrid involving three species of the genus Beta, with special reference to isozyme polymorphism. Theor Appl Genet 73: 261-266

Pacifico D, Miselli F, Micheler M, Carboni A, Ranalli P and Mandolino G (2006) Genetics and marker-assisted selection of the chemotype in *Cannabis sativa* L. Mol Breeding 17: 257-268

Parkin IAP, Gulden SM, Sharpe AG, Lukens L, Trick M, Osborn TC and Lydiate DJ (2005) Segmental structure of the *Brassica napus* genome based on comparative analysis with *Arabidopsis thaliana*. Genetics 171: 763-781

Paterson AH, Bowers JE, Burow MD, Draye X, Elsik CG, Jiang CX, Katsar CS, Lan TH, Lin YR, Ming R and Wright RJ (2000) Comparative genomics of plant chromosomes. Plant Cell 12: 1523-1540

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

Pillen K, Steinrücken G, Herrmann RG and Jung C (1993) An extended linkage map of sugar beet (*Beta vulgaris* L.) including nine putative lethal genes and the restorer gene X. Plant Breeding 111: 265-272

Pradhan AK, Gupta V, Mukhopadhyay A, Arumugam N, Sodhi YS and Pental D (2003 A high-density linkage map in *Brassica juncea* (Indian mustard) using AFLP and RFLP markers. Theor Appl Genet 106: 607-614

Ranalli P, Di Candilo M, Mandolino G, Grassi G and Carboni A (1999) Hemp for sustainable agricultural systems. Agro-Food-Industry Hi-Tech 2: 33-38

Raybould AF, Mogg RJ and Clarke RT (1996) The genetic structure of *Beta vulgaris* ssp. *maritima* (sea beet) populations: RFLPs and isozymes show different patterns of gene flow. Heredity 77: 245-250

Richards CM, Brownson M, Mitchell SE, Kresovich S and Panella L (2004) Polymorphic microsatellite markers for inferring diversity in wild and domesticated sugar beet (*Beta vulgaris*). Mol Ecology Notes 4: 243-245

Rode J, In-Chol K, Saal B, Flachowski H, Kriese U and Weber WE (2005) Sex-linked SSR markers in hemp. Plant Breeding 124: 167-170

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

Sakamoto K, Abe T, Matsuyama T, Yoshida S, Ohmido N, Fukui K and Satoh S (2005) RAPD markers encoding retrotransposable elements are linked to the male sex in Cannabis sativa L. Genome 48: 931-936

Salentijn EMJ, Sandal NN, Lange W, de Bock TSM, Krens FA, Marcker KA and Stiekema WJ (1992) Isolation of DNA markers linked to a beet cyst nematode resistance locus in *Beta patellaris* and *Beta procumbens*. Mol Gen Genet 235: 432-440

Schmidt T, Junghans H and Metzlaff M (1990) Construction of *Beta procumbens*-specific DNA probes and their application for the screening of *B. vulgaris* x *B. procumbens* (2n=19) addition lines. Theor Appl Genet 79: 177-181

Schneider K, Borchardt D, Schäfer-Pregl R, Nagl N, Glass C, Jeppson A, Gebhardt C and Salamini F (1999) PCR-based cloning and segregation analysis of functional gene homologues in *Beta vulgaris*. Mol Gen Genet 262: 515-524

Schneider K, Weisshaar B, Borchardt DC and Salamini F (2001) SNP frequency and allelic haplotype structure of *Beta vulgaris* expressed genes. Mol Breeding 8: 63-74

Scholten OE, de Bock TSM, Klein-Lankhorst RM and Lange W (1999) Inheritance of resistance to beet necrotic yellow vein virus in *Beta vulgaris* conferred by a second gene for resistance. Theor Appl Genet 99: 740-746

Schumacher K, Schondelmaier J, Barzen E, Steinrücken G, Borchardt DC, Weber WE, Jung C and Salamini F (1997) Combining different linkage maps in sugar beet (*Beta vulgaris* L.) to make one map. Plant Breeding 116: 23-38

Siniscalco Gigliano G, Caputo P and Cozzolino S (1997) Ribosomal DNA analysis as a tool for the identification of *Cannabis sativa* L. specimens of forensic interest. Sci Justice 37: 171-174

Sirikantaramas S, Morimoto S, Shoyama Y, Ishikawa Y, Wada Y, Shoyama Y and Taura F (2004) The gene controlling marijuana psichoactivity. J Biol Chem 279: 39767-39774

Skaracis GN and Biancardi E (2000) Breeding for Cercospora resistance in sugar beet. Adv Sugar Beet Res 2: 177-196

Small E and Beckstead HD (1973) Common cannabinoid phenotypes in 350 stocks of *Cannabis*. Lloydia 36: 144-165

Somers J, Rakow G and Rimmer R (2002) *Brassica napus* DNA markers linked to white rust resistance in *Brassica juncea*. Theor Appl Genet 104: 1121-1124

Speckmann GJ and de Bock TSM (1982) The production of alien monosomic additions in *Beta vulgaris* as a source for the introgression of resistance to beet root nematode (*Heterodera schachtii*) from Beta species of the Section *Patellares*. Euphytica 31: 313-323

Stevanato P, De Biaggi M, Skaracis GN, Colombo M, Mandolino G, and Biancardi E (2001) The sea beet (*Beta vulgaris* L. ssp. *maritima*) of the Adriatic coast as source of resistance for sugar beet. Sugar Tech 3: 77-82

Sturtevant AH (1913) The linear arrangement of six sex-linked factors in Drosophila, as shown by their mode of association. J Exp Zoology 14: 43-59

Teutonico RA and Osborn TC (1994) Mapping of RFLP and qualitative trait loci in *Brassica rapa* and comparison to the linkage maps of *B. napus*, *B. oleracea*, and *Arabidopsis thaliana*. Theor Appl. Genet 89: 885-894

The Arabidopsis Genome Initiative (2000) Analysis of the genome sequence of the flowering plant *Arabidopsis thaliana*. Nature 408: 796-815

Tufto J, Raybould AF, Hindar K and Engen S (1998) Analysis of genetic structure and dispersal patterns in a population of sea beet. Genetics 149: 1975-1985

UN (1935) Genome analysis in Brassica with special reference to the experimental formation of *B. napus* and its peculiar mode of fertilization. Japan J Bot 7: 389-452

Weber WE, Borchardt DC and Koch G (1999) Combined linkage maps and QTLs in sugar beet (*Beta vulgaris* L.) from different populations. Plant Breeding 118: 193-204

Weber WE, Borchardt DC and Koch G (2000) Marker analysis for quantitative traits in sugar beet. Plant Breeding 119: 97-106

Weiland J and Koch G (2004) Sugarbeet leaf spot disease (*Cercospora beticola* Sacc.). Mol Plant Pathology 5: 157-166

Wilkinson M and Linacre A (2000) The detection and persistence of *Cannabis sativa* DNA on skin. Sci Justice 40: 11-14

Zhao J and Meng J (2003) Genetic analysis of loci associated with partial resistance to *Sclerotinia sclerotiorum* in rapeseed (*Brassica napus*). Theor Appl Genet 106: 759-764