Phyton (Horn, Austria) Special issue: "Bioindication ..."	Vol. 36	Fasc. 3	(129)-(138)	15.09.96

Molecular Tools for Population and Ecological Genetics in Coniferous Trees

By

MICHELE MORGANTE[1)], ANTONELLA PFEIFFER[1)], ANTONIA COSTACURTA[1)] & ANGELO M. OLIVIERI[1)]

Key words: Molecular markers, microsatellites, automation, simple sequence repeats, PCR.

Summary

MORGANTE M., PFEIFFER A., COSTACURTA A. & OLIVIERI A. M. 1996. Molecular tools for population and ecological genetics in coniferous trees. - Phyton (Horn, Austria) 36 (3): (129) - (138).

In order for molecular markers to be used in the population genetics of forest trees they have to meet certain requirements such as the ease and speed of genotyping, the codominance of alleles, the reproducibility over time and space, the high information content and the possibility of easily exchanging marker information. None of the many molecular marker systems available fully meets all these requirements. However after taking into account advantages and disadvantages of each of the systems we decided to focus our attention on the use of simple sequence repeats (SSRs or microsatellites) because they are codominant, reproducible, highly informative and easy to exchange. We have been isolating AC/GT and AG/CT SSRs from the Norway spruce (*Picea abies* K.) nuclear genome. We isolated several hundreds positive clones from a small-insert genomic library and following sequence analysis we designed primers for 36 of them, 24 containing AG and 12 AC SSRs. After testing them on a panel of spruce individuals 25% of the primer pairs produced a single-locus hypervariable pattern, with the remaining ones giving either a single monomorphic product (18%) or very poor amplification (19%) or amplification of multiple bands (38%). Segregation in accordance with a simple Mendelian model of inheritance was demonstrated for all the loci amplified with the primer pairs giving a simple variable pattern. We screened a panel of 19 spruce trees at these loci. The average number of alleles per locus was 14 and expected heterozygosity 0.80, with up to 23 alleles per locus and heterozygosities exceeding 0.94. This shows that nuclear SSRs can be very useful markers in the population genetics of trees even though the overall efficiency of the marker identification process is quite low due to the high percentage of primer pairs producing complex or "dirty" patterns. We attribute this phenomenon to the high complexity of the spruce genome. Other methods, including the construction of libraries highly enriched for SSR sequences, that we developed in order to make SSR retrieval and typing easier and faster will be discussed. We recently extended the use of PCR amplified SSR markers to the chloroplast genome. We demonstrated that mononucleotide poly(A/T) stretches are frequent in the

[1)] Dipartimento di Produzione Vegetale e Tecnologie Agrarie, Universita' di Udine, Via delle Scienze 208, I-33100 Udine, Italy.

chloroplast genomes of plants and show high levels of between and within population variation, making them ideal tools for cytoplasmic population genetics overcoming the difficulties in finding within species variation that are frequently encountered when analysing the cpDNA molecule by RFLPs or PCR-RFLPs. We will discuss the possible applications of such markers for studying gene flow and for paternity analysis.

Introduction

Isozyme markers have proven their utility as genetic markers for the analysis of genetic variability, mating system and gene flow in forest trees and have been applied to a variety of problems in population and ecological genetics. However the analysis of isozyme loci variation is limited by the fact that the total number of loci and thus of genomic locations that can be assayed is low and that only coding regions of the genome can be assayed. DNA markers allow analysis of variation at the DNA level in both coding and non coding regions and are almost unlimited in number. DNA markers started to be used in 1980 when the idea that naturally occurring DNA sequence variation could provide a virtually unlimited source of genetic markers was proposed and put into practice by BOTSTEIN and colleagues (BOTSTEIN & al. 1980). The marker system used by BOTSTEIN and his colleagues, Restriction Fragment Length Polymorphism (RFLP) analysis was only the first one of a now fairly large number of technologies available for detecting differences between individuals at the DNA level and thus to obtain molecular markers. Many of the currently used marker systems have taken advantage from the advent of the polymerase chian reaction (PCR) technique (MULLIS & al. 1986). Plant geneticists soon realized the advantages and possibilities offered by the use of molecular markers not only in the traditional field of plant genetics, but also in the more applied field of breeding (BECKMANN & SOLLER 1983, TANSKLEY & al. 1989). Among the marker systems that are most commonly used in plants now (see RAFALSKI & al. 1995 for a detailed technical description of each marker type) are RFLPs, Random Amplified Polymorphic DNAs (RAPDs, WILLIAMS & al. 1990), Simple Sequence Repeats (SSRs or microsatellites, WEBER & MAY 1989, MORGANTE & OLIVIERI 1993), Cleaved Amplified Polymorphic Sequences (CAPS or PCR-RFLP, KONIECZNY & AUSUBEL 1993), Amplified Fragment Length Polymorphisms (AFLPs, ZABEAU & VOS 1993). Each one of these marker systems offers a unique combination of advantages and disadvantages (reviewed in detail in RAFALSKI & TINGEY 1993), and differ in the type of sequence polymorphisms detected (insertion/deletions vs. point mutations), the information content, the dominance relationships between alleles (dominant vs. codominant markers), the amount of DNA required, the need for DNA sequence information in the species under analysis, the development costs, the ease of use and the extent to which they can be automated. The choice of a marker system is to a significant extent dictated by the specific application and there is probably not a single class of markers that can satisfy all the needs encountered by plant geneticists and breeders.

Molecular markers have not found so far a widespread application in population and ecological genetics of forest trees because technology hasn't developed so much yet as to be able to satisfy all the requirements for their successful use. The first of these requirements is the capacity to screen a large number of individuals (several individuals per several populations) at a certain number of loci with low costs. So speed, ease of use and low costs (when using the markers, not necessarily when developing them) are primary criteria to select the appropriate marker system. As a consequence automation both of the marker assays and of the marker scoring is also a definite requirement. In addition to this, markers should be codominant because of the outbreeding mode of reproduction of the forest tree species and the frequent occurrence of heterozygous genotypes that would not be recognized when using dominant markers. Estimation of the levels of heterozygosity at the individual, population and species level is one of the key issues in the population and conservation genetics of trees and this cannot simply be achieved on the basis of the frequency of the homozygotes for the recessive allele under the assumption that populations are under Hardy-Weinberg equilibrium because this is frequently not the case. For conifers, where haploid megagametophytes (endosperms) can be analysed and thus the diploid genotype of the mother tree can be inferred by looking at several (6-8) endosperms from the same tree, even dominant markers could be used. However, this implies that the total number of markers assays would also be multiplied by a factor of 6-8 as compared to analysing a diploid tissue using codominant markers. Extraction of good quality DNA from forest trees is often difficult because of the presence of polyphenols and polysaccharides that makes it necessary to use long and tedious extraction protocols. Unless fast and cheap methods to extract high quality DNA are devised, the ideal marker system should therefore not require high quality and high quantities of DNA and not be too sensitive to variations in DNA concentrations in the samples to be analysed. When molecular markers have to be applied to mating sytem, paternity or gene flow analysis, where the markers are used to differentiate individuals rather than to assess their degree of relatedness, they should be highly informative, i.e. show a high number of alleles with similar frequencies. Information content can be measured for each marker by expected heterozygosity H_e (NEI 1973) where

$$H_e = 1 - \text{sum } (p_i^2)$$

or by the within-population differentiation delta_T (GREGORIUS 1987) where

$$\text{delta}_T = [N/(N-1)] \cdot [1 - \text{sum } (p_i^2)]$$

The two measures are very similar but the latter one takes into account the finite sample size. A high information content reduces the number of loci that have to be screened to obtain a sufficient level of resolution between individuals. When

relationships at the above population level or even more at the subspecies or species level have to be determined, i.e. when genetic relatedness has to be estimated and phylogeny has to be inferred, less informative markers may be preferred. So the problems of reversion and the possibility of homoplasy due to the high mutation rates of hypervariable markers can be avoided. Finally, repeatibility over time and space is a key requirement for molecular markers to be used for monitoring variability in forest trees. This is a very important aspect when one considers that studies on genetic variation frequently require to analyse the same population in different years to look at the population dynamics or to compare results obtained by different labs in different populations. The use of a common set of markers that can be easily exchanged would enormously facilitate such tasks. Isozymes have proven so far to be repeatable and to easily allow comparisons between results from different groups both in a qualitative (presence or absence of certain alleles) and quantitative (levels of genetic diversity or differentiation, etc.) way.

Based on the needs that we have previously defined we came to the conclusion that the marker assays that do not require species-specific sequence information, such as RAPDs and AFLPs are not suited for being used in the population genetics of forest trees because they produce mostly if not only dominant markers. Additionally RAPDs have been shown not to be very repeatable not only between different labs but also within the same lab. For AFLPs where a complex pattern is produced with as many as 100 bands per lane it could be difficult to identify exactly the band (i.e. marker) that has been looked at previously or by someone else. RFLPs, that require Southern blotting and hybridisation, are not a viable alternative because they are time consuming, technically demanding and cannot be easily automated. We therefore focused our attention on PCR-based markers that require species-specific sequence information, i.e. Sequence Tagged Sites (OLSON & al. 1989), where the primer sequences unequivocally define the locus under analysis and markers can be exchanged simply by exchanging sequence information. We especially focused on SSRs or microsatellites because on the one hand they fully share the advantages of STSs and on the other hand their hypervariability makes them extremely informative codominant markers (BECKMANN & SOLLER 1990). They are thus almost ideal markers for population genetics studies. The effort required for developing the markers should be paid back through the greater utility of such markers and the possibility to efficiently automate the assay. Presence and hypervariability of SSRs has been demonstrated in nuclear genomes of plants (MORGANTE & OLIVIERI 1993) and more recently also in the chloroplast genome (POWELL & al. 1995a, 1995b). Nuclear SSRs have been shown to be definitely more informative than the other markers systems available in soybean (MORGANTE & al. 1994, POWELL & al. 1996) as well as in other crop species. Very little information is available on SSRs in conifers (SMITH & DEVEY 1994). We will here present results of our work aimed at identifying and using SSRs for the nuclear and chloroplast genome in conifers, mainly in Norway spruce (*Picea abies* K.).

Nuclear SSRs in Norway spruce

Our work was focused into two directions: one is the identification and testing of SSRs markers in Norway spruce, the other is the development and/or adaptation of the enabling technology for fast and efficient identification and use of SSR markers in plants.

We have been isolating AC/GT and AG/CT SSRs from the Norway spruce (*Picea abies* K.) nuclear genome. Based on hybridization data we estimate that there is one AC SSR every 500-600 kb and one AG SSR every 200-300 kb. When the genome size of spruce is considered (30-40 pg per haploid genome) this corresponds to an extremely large number of SSRs available for analysis. We have isolated more than 200 positive clones containing AC/GT or AG/CT repeats from two unenriched small insert lambda phage genomic libraries and sequenced over 50 of them. We designed primers for 36 of them so far and tested all of them on a panel of individual trees representing the diversity of the species. Amplification conditions were optimised for each primer pair to give the best possible amplification pattern by testing different $MgCl_2$ concentrations and different annealing temperatures. We also tested the use of touch-down PCR profile and the digestion of genomic DNA prior to amplification even though neither of these two approaches improved the amplification results. After the optimisation work primers were divided into four classes:

1) No or unsatisfactory amplification	19 %
2) Good amplification, no variability	18 %
3) Good amplification of multiple loci, variability	38 %
4) Good amplification of a single locus, variability	25 %

In light of our previous experience with SSRs in other plant species (soybean, corn, barley) the proportion of primers falling into class 4 is low. We are trying to determine what is the reason for either poor amplification (class 1) or for the appearance of multiple bands corresponding to the coamplification of different loci (class 3) and whether or not the monomorphic fragments (class 2) contain SSRs. Our working hypothesis is that the main cause may be the very high complexity of the spruce genome (30-40 pg DNA per haploid genome) and that many of the SSRs may be embedded in repetitive rather than single copy DNA. If primers are taken from the repetitive sequences, then this would result in either poor amplification or amplification of multiple fragments. Based on Southern blotting results, all of the monomorphic products (class 2) do contain the expected SSR motif and have a size that is similar to that expected on the basis of the sequence data. They therefore seem to correspond to SSR regions that do not show size variation. We are currrently working on the determination of the sequence content of the multiple fragments that are coamplified when using the primers falling into class 3 by Southern blotting and/or sequencing. It looks as if many of the amplified fragments obtained with primer pairs falling into class 3 do not

contain SSRs and thus do not correspond to the sequences that were used for primer design. A procedure for prescreening and classifying positive clones according to their repetitive DNA content is being set up based on hybridisation to labelled total genomic DNA. Preliminary data show that primers falling into class 1 are usually obtained from clones giving weak signals on such blots. The signal intensity in such an experiment is proportional to the abundance of the sequences contained in each clone in the genome. Weak signals therefore correspond to clones present in single or low-copy number.

We determined information content and mode of inheritance for the SSRs markers that fall into class 4. Variability was checked on a panel of 18 trees by using fluorescent labelling and a Pharmacia A.L.F. automated sequencer (MANSFIELD & al. 1994). The average number of alleles per locus was 14 and expected heterozygosity 0.80, with up to 23 alleles per locus and heterozygosities exceeding 0.94. This is in contrast with what we previously found using isozyme markers on the same populations (GIANNINI & al. 1991, MORGANTE & VENDRAMIN 1991) where the average number of alleles per locus was 1.78 and expected heterozygosity 0.16. SSR markers therefore appear to be five times more informative than isozyme markers, if we look at expected heterozygosity. Segregation was tested on meiotic products (megagametophytes) from heterozygous individuals and was always in accordance with a simple Mendelian mode of inheritance. SSR markers hence appear to be perfectly fit for the screening of variation in natural populations. We are going to screen 30 to 40 individuals from 4 italian populations with these primers for better assessing levels and distribution of genetic variation for SSR markers and thus validate the use of the markers. We have also used the same primers on other species belonging to the genus *Picea* to determine how transferrable the markers are to other species. We had succesful amplification in a significant proportion of cases and are now determining if the amplified fragments are homologous to the ones from Norway spruce and if they contain a SSR sequence.

As far as the technology development is concerned we initially developed procedures for the non radioactive hybridisation screening of genomic libraries and for the PCR screening of clones (RAFALSKI & al. 1995). We use a robotic pipetting workstation (Beckmann Biomek 1000) to speed up the set up of PCR reactions. We more recently worked on two other steps, the construction of enriched genomic libraries and the automated separation and scoring of the amplified fragments. The availability of libraries where the proportion of clones containing SSRs is significantly increased compared to their real frequency in the genome is extremely helpful for rapid identification of SSR markers, because it can reduce the amount of work required for the screening of the libraries to a minimum. Several methods have been recently proposed based on different principles (ITO & al. 1992, OSTRANDER & al. 1992, KARAGYOZOV & al. 1993, KANDPAL & al. 1994). We have developed a very efficient method for the enrichment of libraries based on liquid hybridisation of genomic fragments to solid phase captured oligonucleotide probes which is basically a modification of the methods described by KARAGYOZOV & al.

1993 and KANDPAL & al. 1994 (MORGANTE & al., in preparation). The method has been thoroughly tested and we have made a spruce library enriched for AC/GT repeats which shows a 200-fold enrichment, with one positive every four clones. We are currently isolating and sequencing new SSRs from this library. Enriched libraries can be of even greater advantage when rare SSR motifs, such as trinucleotide repeats, have to be isolated. Trinucleotide repeats are attractive because they offer the advantage of an easier resolution between alleles that will differ by at least 3 bp instead of 2. We have performed the separation of the SSR markers on a Pharmacia A.L.F. automated sequencer. The advantages of this kind of analysis are the non radioactive detection, the automated sizing of fragments and allele identification and the high throughput achievable. Our work has focused on testing different gel matrices, reloading the same gel several times, using different labelling strategies for making the fragments fluorescent, multiplexing by size and verifying the repeatibility and accuracy of the system. We are now capable of separating up to 800 different SSRs in one day with a 0.2 bp accuracy.

Chloroplast SSRs

In plants chloroplast and mitochondrial genomes exhibit different patterns of genetic differentiation compared to nuclear alleles due to their generally uniparental mode of transmission. The analysis of the chloroplast organelle provides information on the population dynamics of plants which is complementary to that obtained from the nuclear genome. The chloroplast genome is highly conserved and has a much lower mutation rate than plant nuclear genomes. Restriction site analysis of chloroplast DNA (cpDNA) has been widely used for inter-specific studies and in some cases the magnitude of intra-specific variation has been sufficient to allow population based studies. It has been observed, at least in Angiosperms, that the majority of the intraspecific variation is found between populations with little or no intrapopulational variation possibly due to the maternal mode of inheritance. The availability of a high resolution chloroplast specific polymorphic assay would facilitate the analysis of population differentiation and gene flow. Furthermore, in Gymnosperms where chloroplasts are primarily paternally transmitted such an assay would allow pollen movement to be monitored.

We have examined the polymorphism of such sequences in the chloroplast genomes of plants, using a PCR-based assay. GenBank searches identified the presence of several $(dA)_n \cdot (dT)_n$ mononucleotide stretches in chloroplast genomes. A chloroplast simple sequence repeat (cpSSR) was identified in three pine species (*Pinus contorta*, *Pinus sylvestris* and *Pinus thunbergii*), 312 bp upstream of the psbA gene. DNA amplification of this repeated region from 11 pine species identified 9 length variants (POWELL & al. 1995a). The polymorphic amplified fragments were isolated and the DNA sequence was determined, confirming that the length polymorphism was caused by variation in the length of the repeated region. Analysis of 305 individuals from 7 populations of *Pinus leucodermis* Ant., an endangered conifer with a restricted and discontinuous distribution in Southern Europe, revealed the presence of 4 variants with intrapopulational diversities

ranging from 0.000 to 0.629 and an average of 0.320. Restriction fragment length polymorphism (RFLP) analysis of chloroplast DNA (cpDNA) on the same populations previously failed to detect any variation (BOSCHERINI & al. 1994). Population subdivision based on cpSSR was higher (G_{st}=0.22) than that revealed in a previous isozyme study (G_{st}=0.05) (BOSCHERINI & al. 1994). These data show that chloroplast SSRs exhibit high levels of between and within population variation, making them ideal tools for cytoplasmic population genetics. They allow to overcome difficulties in finding within species variation that are frequently encountered when analysing the cpDNA molecule by RFLPs or PCR-RFLPs. Similar findings were also obtained in an Angiosperm species such as soybean (*Glycine max*) (POWELL & al. 1995b).

The high sequence conservation of the chloroplast genome makes it easier to design primers that work even in relatively distant species. Taking advantage of the recent availability of the complete chloroplast genome sequence of *Pinus thunbergii* we designed primers that amplify 20 different SSR regions distributed over the whole cpDNA molecule. We amplified succesfully the corresponding regions from three mediterranean pine species, *Pinus leucodermis*, *Pinus halepensis* and *Pinus brutia* and are currently screening different populations for variation. The use of hypervariable cpSSRs will be of particular importance for studying partitioning of genetic variation, gene flow via pollen and for paternity analysis in conifers.

Conclusions

Nuclear and chloroplast SSRs hold the promise to fulfill many of the requirements for molecular markers to be applied to the population genetics of conifers. Retrieval of useful nuclear SSRs seems to be more time consuming than in other species possibly as a consequence of the extremely high complexity of conifer genomes and the high amount of repeated DNA they possess. More efficient ways to retrieve the desired sequences are, however, being worked out as well as techniques to speed up and automate the assays and the scoring of the individual genotypes. Dissemination of these results and methodologies to the tree population genetics community will be a key issue in determining a widespread utilization of these molecular tools. The use of such tools holds the promise of providing us with a much deeper insight into many of the population and ecological genetics processes of trees that have a profound impact on the dynamics and survival of forests.

Acknowledgments

Financial support from the European Commission (BIOTECH project BI02-CT93-0373) is gratefully acknowledged. We would like to thank the Colleagues of the Molecular Breeding group at DuPont Agricultural Products, Biotechnology Research, Wilmington DE, and W. POWELL from the Scottish Crop Research Institute, Invergowrie, Scotland for many helpful discussions and their contributions to the methods and ideas discussed in this paper.

References

BECKMANN J.S. & SOLLER M. 1983. Restriction fragment length polymorphisms in genetic improvement: methodologies, mapping and costs. - Theor. Appl. Genet. 67: 35-43.

-- & -- 1990. Toward a unified approach to genetic mapping of eukaryotes based on sequence tagged microsatellite sites. - Bio Technology 8: 930-932.

BOSCHERINI G., MORGANTE M., ROSSI P. & VENDRAMIN G.G. 1994. Allozyme and chloroplast DNA variation in Italian and Greek populations of *Pinus leucodermis.* - Ant. Heredity, 73: 284-290.

BOTSTEIN D., WHITE R.L., SKOLNICK M.H. & DAVIS R.W. 1980. Construction of a genetic map in man using restriction fragment length polymorphism. - Am. J. Hum. Genet. 32: 314-331.

GIANNINI R., MORGANTE M. & VENDRAMIN G.G. 1991. Allozyme variation in Italian populations of *Picea abies* K.. - Silvae Genetica 40: 160-165.

GREGORIUS H.R. 1987. The relationship between the concepts of genetic diversity and differentiation. - Theor. Appl. Genet. 74: 397-401.

ITO T., SMITH C.L. & CANTOR C.R. 1992. Sequence-specific DNA purification by triplex affinity capture. - Proc. Natl. Acad. Sci. USA 89: 495-498.

KANDPAL R.P., KANDPAL G. & WEISSMAN S.M. 1994. Construction of libraries enriched for sequence repeats and jumping clones, and hybridization selection for region-specific markers. - Proc. Natl. Acad. Sci. USA 91: 88-92.

KARAGYOZOV L., KALCHEVA I. & CHAPMAN V.M. 1993. Construction of random small-insert genomic libraries highly enriched for simple sequence repeats. - Nucleic Acids Res. 21: 3911-3912.

KONIECZNY A. & AUSUBEL F.M. 1993. A procedure for quick mapping of *Arabidopsis* mutants using ecotype-specific markers. - Plant J. 4: 403-410.

MANSFIELD D.C., BROWN A.F., GREEN D.K., CAROTHERS A.D., MORRIS S.W., EVANS H.J. & WRIGHT A.F. 1994. Automation of genetic linkage analysis using fluorescent microsatellite markers. - Genomics 24: 225-233.

MORGANTE M. & VENDRAMIN G.G. 1991. Genetic variation in Italian populations of *Picea abies* (L.) Karst. and *Pinus leucodermis* Ant. - In: MUELLER-STARCK G. & ZIEHE M. (eds.), Genetic variation in European populations of forest trees, pp. 205-227. JD Sauerlaender's Verlag, Frankfurt am Main, Germany.

-- & OLIVIERI A.M. 1993. PCR-amplified microsatellites as markers in plant genetics. - Plant J. 3: 175-182.

--, RAFALSKI J.A., BIDDLE P., TINGEY S.V. & OLIVIERI A.M. 1994. Genetic mapping and variability of seven soybean microsatellite loci. - Genome 37: 763-769.

MULLIS K., FALOONA S., SCHARF S., SAIKI R., HORN G. & ERLICH H. 1986. Specific enzymatic amplification of DNA in vitro: the polymerase chain reaction. - Cold Spring Harbor Symp. Quant. Biol. 51: 263-273.

NEI M. 1973. Analysis of gene diversity in subdivided populations. - Proc. Natl. Acad. Sci. USA 70: 3321-3323.

OLSON M., HOOD L., CANTOR C. & BOTSTEIN D. 1989. A common language for physical mapping of the human genome. - Science 245: 1434-1435.

OSTRANDER E.A., JONG P.M., RINE J. & DUYK G. 1992. Construction of small-insert genomic DNA libraries highly enriched for microsatellite repeat sequences. - Proc. Natl. Acad. Sci. USA 89: 3419-3423.

POWELL W., MORGANTE M., MCDEVITT R., VENDRAMIN G.G. & RAFALSKI J.A. 1995a. Polymorphic simple sequence repeat regions in chloroplast genomes: applications to the population genetics of pines. - Proc. Natl. Acad. Sci. USA 92: 7759-7763.

--, --, ANDRE C., MCNICOL G., DOYLE J.J. & RAFALSKI J.A. 1995b. Hypervariable chloroplast simple sequence repeats provide a general source of polymorphic DNA markers for the chloroplast genome. - Current Biol. 5: 1023-1029.

--, --, --, HANAFEY M., VOGEL J., TINGEY S.V. & RAFALSKI J.A. 1996. The utility of RFLP, RAPD, AFLP and SSR (microsatellite) markers for germplasm analysis. - Mol. Breed., accepted.

RAFALSKI J.A. & TINGEY S.V. 1993. Genetic diagnostics in plant breeding: RAPDs, microsatellites and machines. - Trends Genet. 9: 275-280.

--, MORGANTE M., POWELL W., VOGEL J.M. & TINGEY S.V. 1995. Generating and using DNA markers in plants. - In: BIRREN B. & LAI E. (eds.), Analysis of non-mammalian genomes - A practical guide. Academic Press, New York, in press.

SMITH D.N. & DEVEY M.E. 1994. Occurrence and inheritance of microsatellites in *Pinus radiata*. - Genome 37: 977-983.

TANSKLEY S.D., YOUNG N.D., PATERSON A.H. & BONIERBALE M.W. 1989. RFLP mapping in plant breeding: new tools for an old science. - Bio Technology 7: 257-264.

WEBER J. & MAY P.E. 1989. Abundant class of human DNA polymorphisms which can be typed using the polymerase chain reaction. - Am. J. Hum. Genet. 44: 388-396.

WILLIAMS J.G.K., KUBELIK A.R., LIVAK K.J., RAFALSKI J.A. & TINGEY S.V. 1990. DNA polymorphisms amplified by arbitrary primers are useful as genetic markers. - Nucleic Acids Res. 18: 6531-6535.

ZABEAU M. & VOSS P. 1993. Selective restriction fragment amplification: a general method for DNA fingerprinting. - European Patent Application 92402629.7 (Publ. Number 0 534 858 A1).