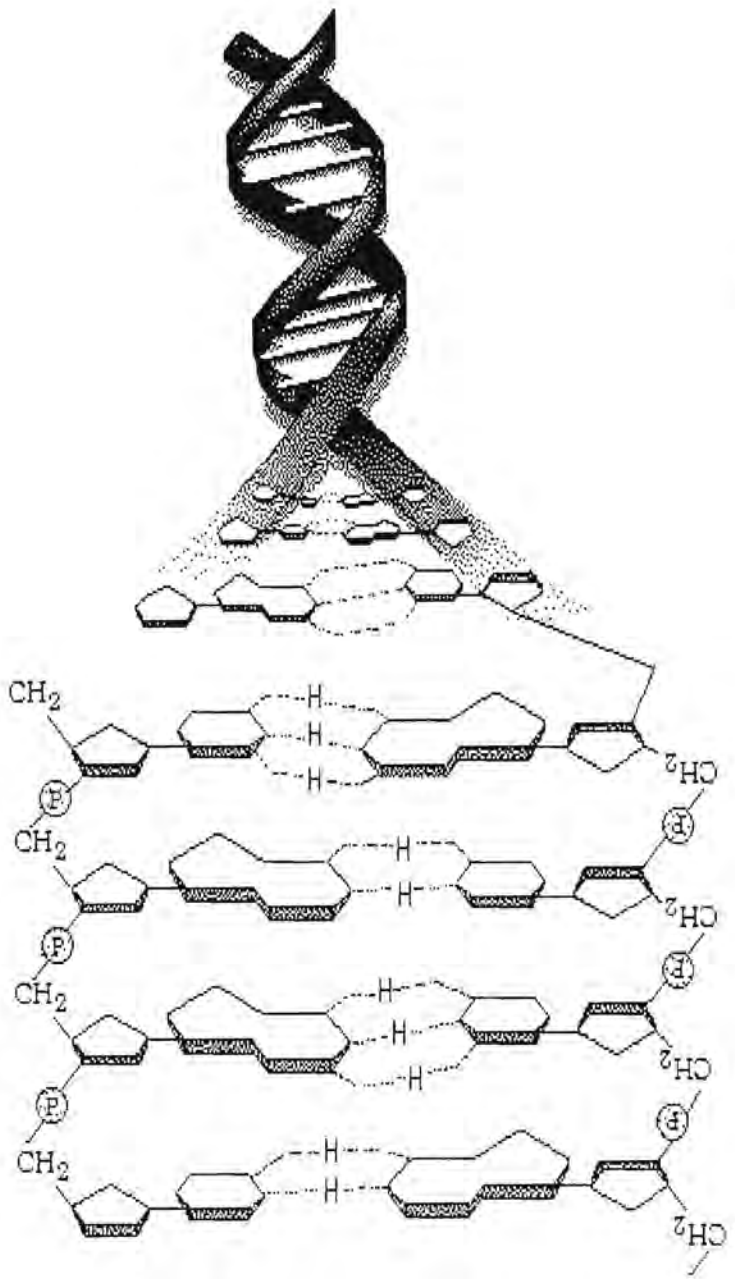
Chapter 5

Detecting methylation changes in the genome



5.1 Abstract

In an extention of the project the RDA technique was applied to see whether changes in genomic methylation between two date palm cultivars and also two types of embryogenic tissues from *Pinus strobus* could be isolated. From date palm a repetitive sequence was isolated as a subtraction product and further products with homology to rRNA genes. These genes also represent repetitive sequences in the plant genome, and as such are a target for methylation. Embryogenic tissue of *Pinus strobes* was used to provide information whether the RDA protocol could also be applied to study large genomes and to isolate subtraction products that could be correlated to differences observed during somatic embryogenesis of tissue grown on different 2,4-D concentrations. Although a difference in the methylation levels of genomic DNA of these cells receiving the different 2,4-D treatments was found, no subtraction products could successfully be isolated, which was very likely due to the size and complexity of the pine genome.

5.2 Objectives

The objectives of this study were to determine whether the RDA protocol could be applied to isolate DNA sequences that are differently methylated between two genomes (MS-RDA). In particular, tissue culture plants from two date palm cultivars were compared ('Barhee' and 'Medjool') as a first objective. They were chosen because of the differences in the occurrence of somaclonal variation that is observed between the two cultivars when regenerated by tissue culture. Further, in a second objective, embryogenic material from *Pinus strobus* (*P. strobus*) that have been grown *in vitro* and exposed to different concentrations of 2,4-D were compared. This experiment was aimed at determining the capability of the standard RDA protocol to compare large genomes and complex genomes such as that of *P. strobus*. *Phoenix dactylifera* (date palm) has a 1C content of 931 Mbp (Olszewska & Osiecka, 1982) whereas *P. strobus* has a IC content of 25137 Mbp (O'Brien *et al.*, 1996)

5.3 Results

5.3.1 Date palm

5.3.1.1 *Comparison of methylation in date palm*

Comparison of methylation was done using the methylation-sensitive restriction enzyme *Hpa*II and the less sensitive enzyme *Msp*I. Using these two enzymes the relative methylation status of the genomic DNA of date palm plants in tissue culture was compared to those that did not derive from a tissue culture process. The results obtained indicate that the date palm DNA obtained from the tissue culture plants were hyper-methylated when compared to that of non-tissue culture plants (Figure 5.1). The genomic DNA of the 'Barhee' cultivar also contained a higher percentage of methylated DNA compared to that of 'Medjool' regardless of whether the plants were exposed to a tissue culture process.

Figure 5.1. Comparison of methylation of tissue culture and non-tissue culture date palm plants. Lanes 1 and 2 represent digested DNA of 'Barhee' non-tissue culture plants digested with with *Hpa*II (1) and *Msp*I (2); lanes 3 and 4 digested DNA of 'Medjool' non-tissue culture plants with *Hpa*II (3) and *Msp*I (4); lanes 5 and 6: digested DNA of 'Barhee' tissue culture plants with HpaII (5) and MspI (6); lanes 7 and 8: digested DNA of 'Medjool' tissue culture plants with HpaII (7) and MspI (8). Lane 9 represents molecular weight marker III.

5.3.1.2 *First MS-RDA subtraction on date palm*

Two MS-RDA subtractions were performed with either 'Medjool' as tester and 'Barhee' as driver or *vice versa*. The subtraction using 'Barhee' DNA as tester yielded no subtraction products. Using 'Medjool' as tester yielded a number of fragments ranging from 200 bp to 500 bp (Figure 5.2 A). Cross hybridisation of the subtraction products revealed a shared homology between all the different cloned fragments representing a common element. This homology was confirmed by subsequent sequence analysis.

Southern blot analysis probing the subtraction products against the first amplicons of both 'Barhee' and 'Medjool' showed that the element was present in both genomes, but a stronger hybridisation signal was observed with the 'Medjool' amplicons. This indicates that a possible higher abundance of this sequence in the 'Medjool' genome. Southern blot analysis probing the subtraction product against genomic DNA gave a smear providing evidence for a dispersed repetitive element. Southern blot hybridisation using a labelled subtraction product hybridised against the other subtraction products revealed homology between all the isolated subtraction products (Figure 5.2 B)

Sequence analysis of the fragments showed a 27 bp repeat within each fragment. The first 24 bp of the repeat was homologous to the 24 bp of the RHpa24 adapter that was used to create the first round amplicons (Figure 5.3). A BlastN homology search revealed some homology of the total sequence with microsattelites from *Hippoglossus hippoglossus* (accession number: AF133244) (E-value: 0.002) and *Phytophthora cinnamomi* (accession number: AF312885) (E-value: 0.12), but no plant-related sequences.

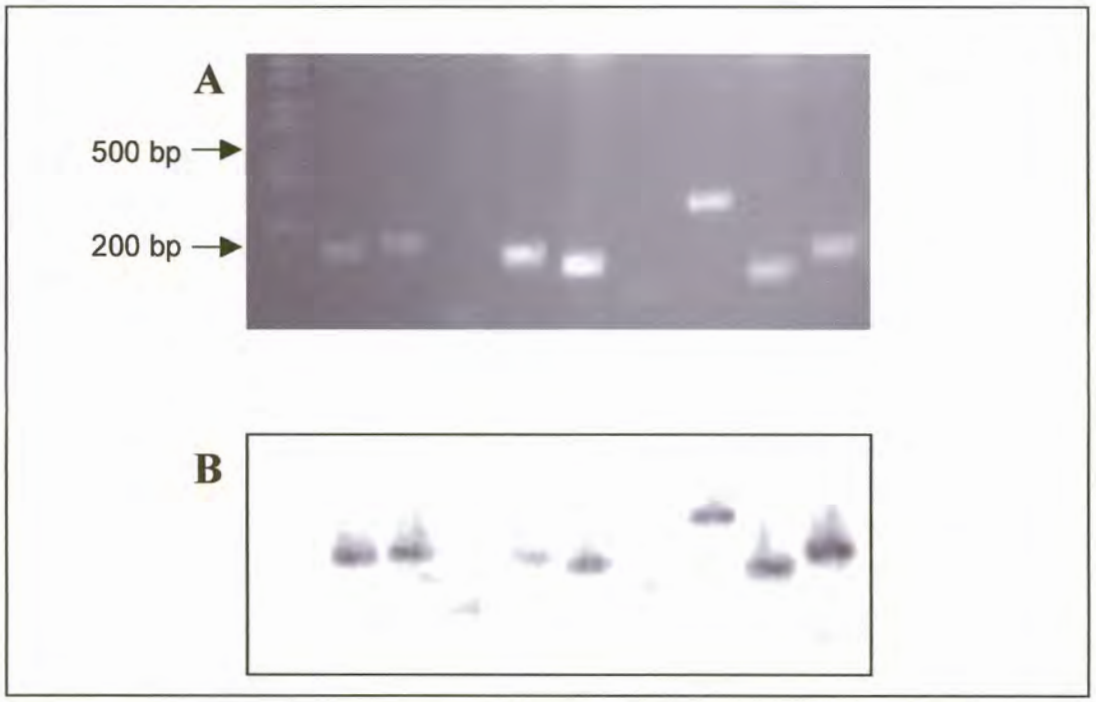


Figure 5.2. Cloned subtraction products (A) cut from plasmids with appropriate restriction enzymes showing various sizes and (B) Southern blot analysis showing homology between the subtraction products.

```
5'GGACCTTGCAAGCCCCTCGCCGAACGCCGTGGTTGT*AGCACTC
TCCAGCCTCTCACCGACCG*AGCACTCTCCAGCCCCTCACCGACCG
C*AGCACTCTCCAGCCTCTCACCGACCGC*GCACTCTCCAGCCTCT
CACCGACCGT*AGCACTCTCCAGCCTCTCACCGACCGT*AGCACTC
CCCAGCCTCTCACCGACCGT*AGCACTCTCCAGCCTCTCACCGACC
GT*AGGCTTCTCCAGCCTCTCACCGACCGT*AGCACTCTTCAGCCT
CTCACCGACCGT*AGCACTCTCCAGCCTCTCA 3'

    RHpa 24   5'AGCACTCTCCAGCCTCTCACCGAC 3'
```

Figure 5.3. Sequence of a 298 bp subtraction product obtained from the first MS-RDA subtraction of date palm. Repeats are separated by a red asterix (*) and the RHpa24 primer sequence is shown for comparison.

5.3.1.3 *Second MS-RDA subtraction on date palm*

Because of the sequence homology found between the first subtraction products and the RHpa24 adapter, the *Hpa*II subtraction was repeated, using the SHpa12 and 24 adapter pair instead of the RHpa adapter pair. The second subtraction resulted in the isolation of three fragments from 'Medjool' as tester (Figure 5.4). The fragment sizes were 444 bp, 330 bp and 215 bp. Sequence analysis and BLAST search of these fragments revealed a 99% homology between the 444 bp fragment and part of the 18S rDNA of tobacco (*Nicotiana tabacum*) (accession number: AJ236016.1) (E-value: 0.0) as well as that of a number of other plant species. The 330 bp fragment showed homology to 23S and 4.5S rDNA sequences including that of tobacco. The 215 bp fragment showed no significant homology to any plant related sequences.

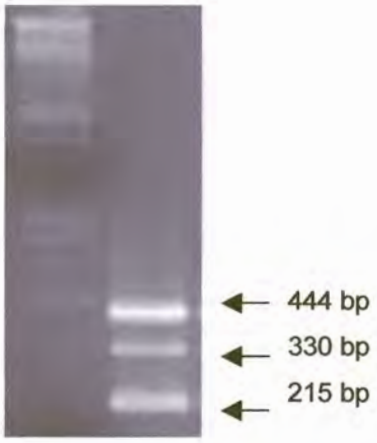


Figure 5.4. Subtraction products obtained from the second MS-RDA subtraction of date palm. Sizes of fragments are indicated.

5.3.2 Pinus strobus

5.3.2.1 *Pinus strobus culture and somatic embryogenesis*

Mega-gametophytes containing zygotic embryos in developmental stages ranging from the cleavage poly-embryony, to the early (pre-cotyledonary) dominance stage were used to initiate embryonal masses (EM) used for somatic embryogenesis. Embryogenic tissue began to grow from explants after the second week of culture, with the majority of explants responding by the 12th week (Figure 5.5A). These

embryonal masses were carefully separated from the mega-gametophytes and transferred to new medium, (Figure 5.5B). Analysis of the tissue under a light microscope confirmed their embryonal nature (Figure 5.5C).

Maintenance of the cultures on the medium used for the initiation of the specific line ensured rapid growth once the line was established after approximately 2 months. However, with a few genotypes an occasional gradual decline in growth was observed. From approximately 300 mg of fresh weight at day 0 the tissue fresh weight increased 6 to 8-fold after a 2-week culture period. This material was used for DNA extraction for MS-RDA, as well as for the maturation experiment to show the differences in the amount of mature embryos produced by the two different 2.4-D treatments.

Embryogenic tissue of *P. strobus* maintained on UL medium produced on average more mature somatic embryos (Figure 5.6) than the ones maintained on Std medium. medium. Figure 5.7 shows differences in the amount of mature embryos produced per plate for two lines (1053 and 1055) when embryogenic tissue maintained on either MLV UL or Std were used for the maturation of embryos. Table 5.1 shows the average number of mature embryos produced per gram of fresh weight for the four lines used.

Table 5.1. Average number of embryos produced per gram of fresh weight used.

Line	Number of embryos per gram of fresh weight MLV UL	Number of embryos per gram of fresh weight MLV Std
1053	319	42
1054	307	13
1055	263	55
1073	282	80

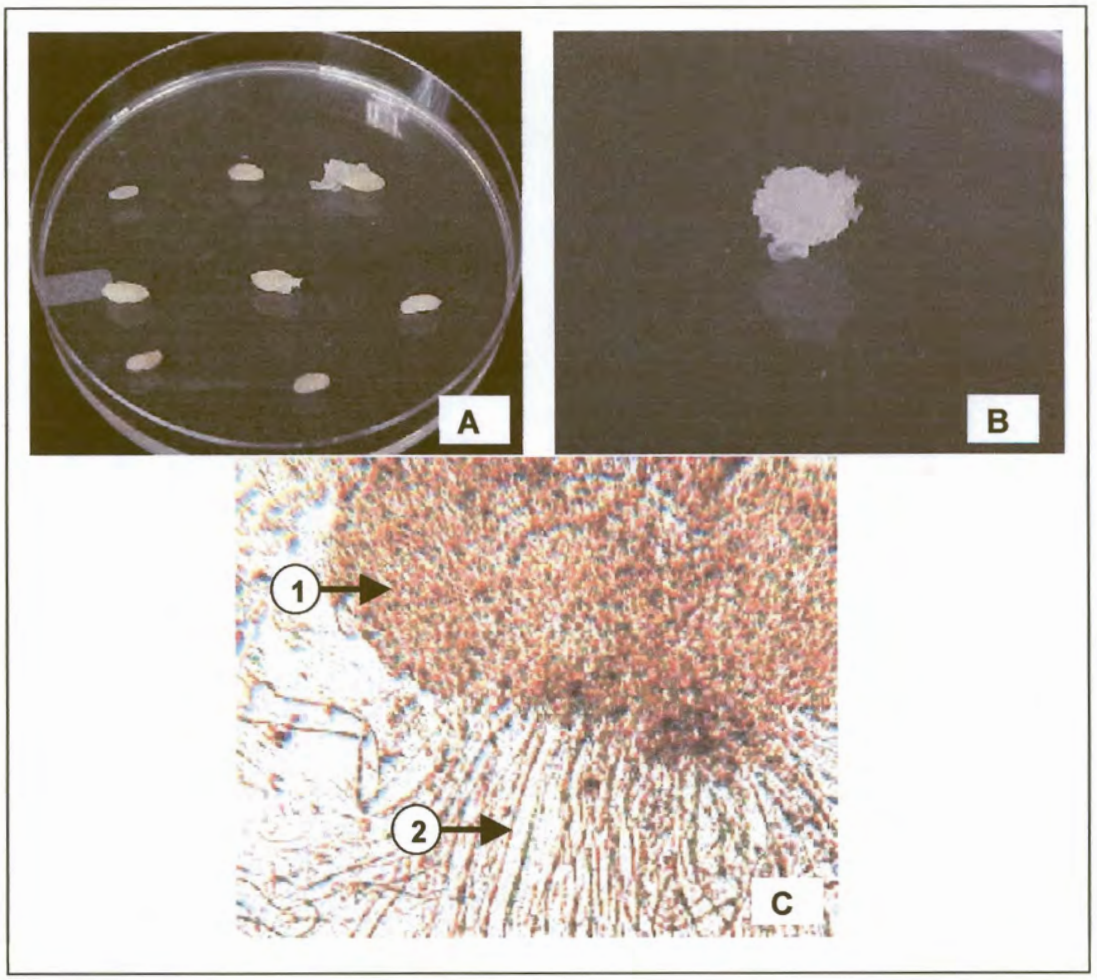


Figure 5.5. (A) Mega-gametophytes on initiation medium. Micropylar extrusion and proliferation of embryonal mass (EM) indicated by arrow. (B) EM on medium and (C) light microscope image of EM showing small embrogenic cells (1) and large suspensor cells (2).

Figure 5.6. Mature somatic embryo of *P. strobus*.

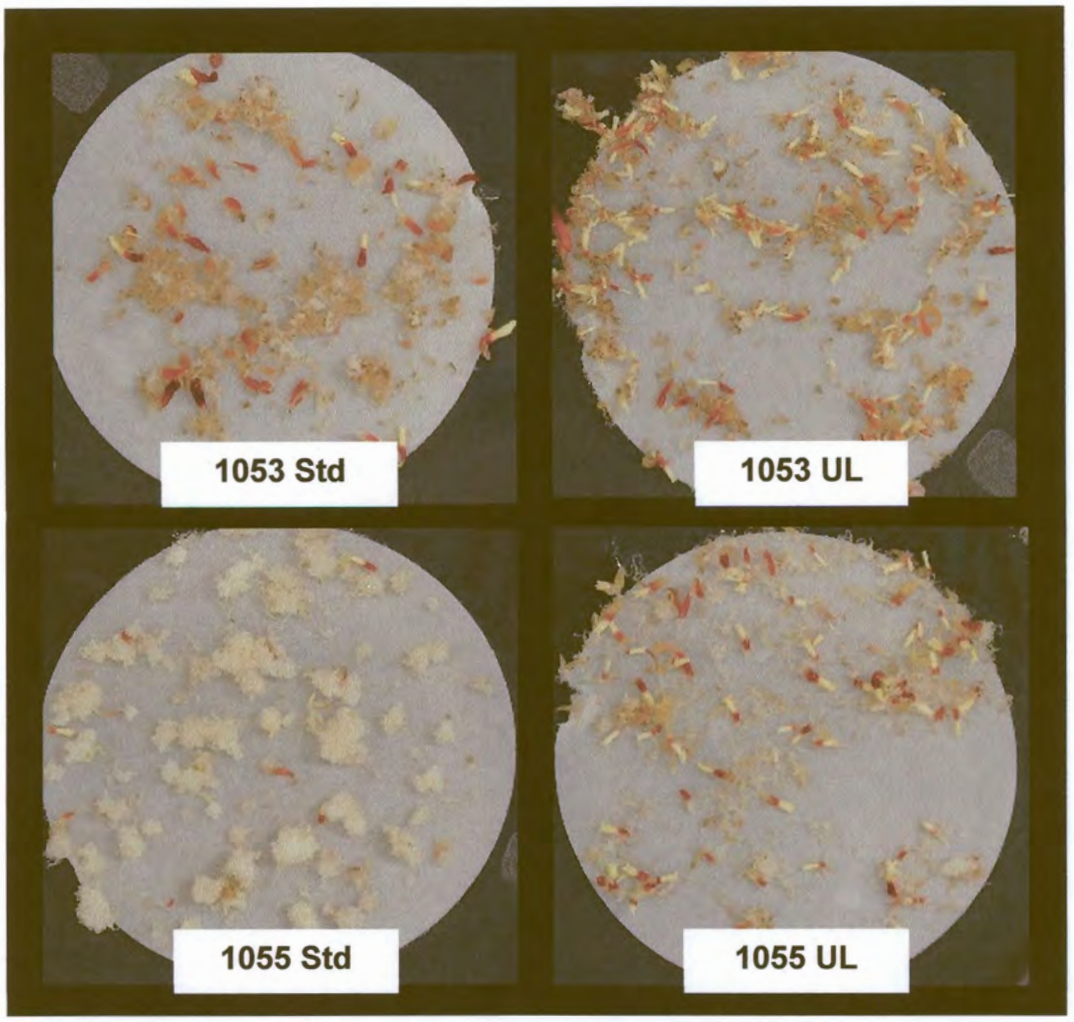


Figure 5.7. Maturation of somatic embryos from tissue maintained on either MLV UL medium (1053 UL and 1055 UL) vs. from tissue maintained on MLV Std medium (1053 Std and 1055 Std).

5.3.2.2 *Comparison of methylation in P. strobus*

Genomic methylation of embryogenic masses of *P. stobus* grown on medium containing different concentrations of 2,4-D was compared. DNA isolated from those cells of *P. strobus* that were grown on high 2,4-D medium also showed a higher level of methylation when compared to those that were grown on a lower 2,4-D medium. This can be seen by the amount of undigested DNA still in the wells of the gel (Figure 5.8; lanes 9-16) and further by the amplification of larger fragments in the *Hpa*II amplicons generated from genomic DNA from cells grown on MLV Std medium *vs.* MLV UL medium (Figure 5.9).

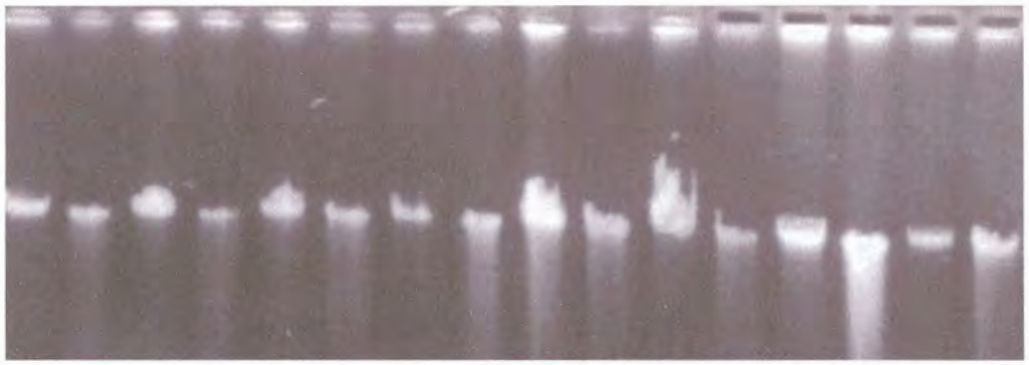


Figure 5.8. Comparison of DNA methylation between cells grown on high and low 2,4-D medium. Lanes 1-8 represent cells grown on low 2,4-D media; lanes 9-16 cells grown on high 2,4-D media. Each DNA sample was cut with *Hpa*II and *Msp*I and the two samples run in adjacent lanes.

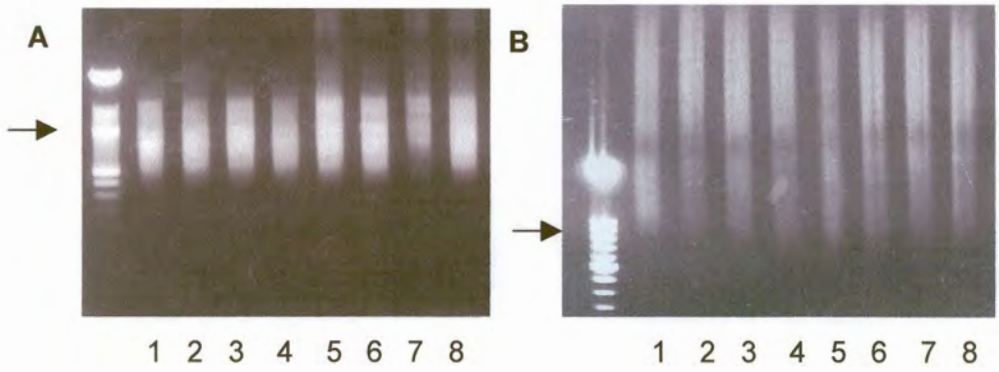


Figure 5.9. Gel profiles of RDA *Hpa*II amplicons generated from genomic DNA of lines grown on (A) MLV UL medium (1.1 μM 2,4-D) and (B) the same lines grown on MLV Std. medium (9.5 μM 2,4-D). The amplification of larger fragments in B indicates higher methylation levels. Arrow indicates the position of the 1000 bp marker.

5.3.2.3 *Pinus strobus MS-RDA subtraction*

Using the tissue grown on MLV UL medium (1.1 μM 2,4-D) as tester and tissue from the same line grown on MLV Std. (9.5 μM 2,4-D) as driver, three rounds of subtractive hybridisation at tester to driver ratios of 1: 1000; 1: 10000 and 1: 100000 were performed. However, the subtractions failed to produce any subtraction

products. The profile obtained after separating the amplicons on a 1.5% agarose gel shows only a smear indicating a high complexity of amplified fragments (Figure 5.10).

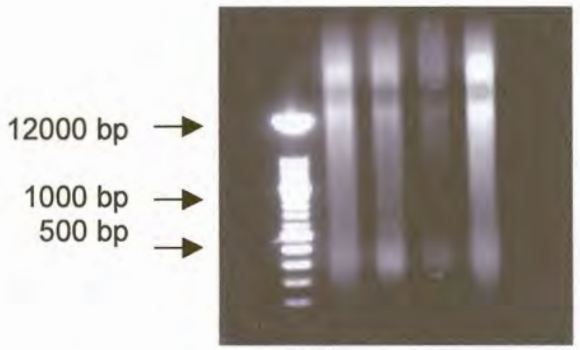


Figure 5.10. Amplification profile of 3rd round subtraction products of 4 *P. strobus* embryogenic lines, 1053 (line 1); 1054 (line 2); 1055 (line 3) and 1073 (line 4).

5.5 Materials and methods

5.5.1 Plant material and DNA extraction

In vitro grown date palm plants from the cultivars 'Medjool' and 'Barhee' and embryogenic tissue of *P. strobus* were used for the experiments. The embryogenic material was obtained from plant material collected from the Nova Scotia Tree Improvement Working Group's (NSTIWG) clonal seed orchard in Debert, Nova Scotia, Canada. Tissue from four cell lines was used (NSTIWG clone numbers 1053, 1054, 1055 and 1073). The embryogenic tissue of each cell line was grown on ML medium containing either 9.5 μM or 1.1 μM of 2,4-D. Total cellular DNA from different types of plant material was isolated from tissue (1g) using a Nucleon Phytopure Plant DNA extraction kit (Amersham Life Science, UK) according to the manufacturer's instructions.

5.5.2 Tissue culture and somatic embryogenesis

5.5.2.1 *Plant material*

For culture initiation, seeds were extracted from developing cones and placed in petri dishes containing moistened filter paper. Within 1 hour of extraction, seeds were surface sterilized in preparation for dissection. Seeds were surface-sterilised by placing them in metal baskets and stirring vigorously for 6 minutes in 200 ml of 3% (v/v) hydrogen peroxide with one drop of Tween 20. Seeds were then rinsed three-times for 2 minutes each with sterile water and transferred to a petri dish containing a moist, sterile filter paper. The seed coat, nucellus, and membrane surrounding the mega-gametophyte were removed under a stereomicroscope, and the mega-gametophyte was placed on initiation medium.

Five to eight mega-gametophytes were cultured in each petri dish containing approximately 25 ml of either MLV UL or Std. medium (see Annex), and dishes were sealed with parafilm. Explants were cultured in darkness at 24°C for 16 weeks and examined every 4 weeks for embryogenic tissue initiation.

5.5.2.2 *Embryogenic tissue proliferation and culture maintenance*

Explants displaying proliferation from the zygotic embryo cells were microscopically selected, then embryogenic tissue was removed, dispersed by vigorous shaking in liquid MLV medium, poured onto a sterile filter paper disc (Whatman No. 2, 5.5 cm) in a Büchner funnel and a short low-pressure pulse (55 kPa) was applied to drain the medium prior to transferring the disc to fresh medium. This approach was also used for maintenance of embryogenic tissue. Maintenance and proliferation cultures were incubated in darkness at 24°C.

5.5.2.3 *Maturation of somatic embryos*

The somatic embryo experiments were performed by combining embryogenic tissue of one line (from several plates), one week after subculture, in a 50 ml test tube, adding liquid MLV media without growth regulators and vigorously shaking the tube to break up the clumps of tissue into a fine suspension. Subsequently 3 ml containing

approximately 0.2 g of suspended embryonal masses were withdrawn with a wide-mouth pipette and placed on a moist filter paper disc (Whatman No. 2 55 mm in diameter) in a Büchner funnel attached to a vacuum pump. A short low-pressure pulse (5 s, 55 kPa) was applied to drain all liquid medium and anchor the embryonal masses to the filter paper as a thin layer. Each disc of filter paper with the embryonal masses were subsequently placed on maturation medium (see Annex) in a 10 X 20 mm petri dish and cultured for up to 10 weeks. The cultures were kept under dim light condition at 1.6 μmol $m^{-2}s^{-1}$ from cool white florescence lamps (Philips F72T12/CW, 65 W) under a 16 hour photoperiod at 24°C.

5.5.3 Comparison of relative methylation

Methylation of total genomic DNA was compared between date palm material derived from tissue culture and non-tissue culture-derived plants (Figure 5.1). For *P. strobus*, DNA methylation was compared between cells grown on high (9.5 μM) and low (1.1 μM) 2,4-D medium (Figure 5.2). The DNA was digested with the restriction enzymes *Hpa*II and *Msp*I both recognizing the same site (C/CGG). The digested DNA was separated by electrophoresis and stained with ethidium bromide and the obtained restriction profiles were compared.

5.5.4 MS-RDA subtraction

5.5.4.1 *Date palm*

RDA was performed following the general outline described by Lisitsyn *et al.* (1993) and as described in the Annex. Total genomic DNA was digested with the methylation sensitive restriction enzyme *Hpa*II. Amplicons were prepared by ligating the adaptor pair RHpa12 and RHpa24 (see Annex) to the digested DNA. After ligation, the DNA was amplified to produce the tester and driver amplicons in eight individual 100 μl volumes using a Perkin-Elmer GeneAmp 9600 thermocycler as described in the Annex.

Driver (150 μg) and tester (10 μg) amplicons were digested with *Hpa*II. Driver and tester DNAs were both re-dissolved at approximately 400 μg/ml. Both the digested

driver and tester amplicons were run on a 1.5% TAE agarose gel alongside an equal aliquot of undigested amplicons to check the completeness of the digestion. On the same gel, standard lambda phage DNA was run so that the final concentrations of driver and tester DNAs could be estimated and adjusted to 400 μg/ml if necessary.

Tester DNA was prepared by adding a second adaptor pair JHpa 12 and 24 to 1 μg of the first round amplicons. An aliquot of tester DNA with ligated adaptors was amplified for 20 cycles in a reaction volume of 20 μl to test that ligated adaptor pair will amplify tester amplicon DNA. A hybridisation reaction was set up by mixing the driver and tester amplicons at a ratio of driver to tester DNA of 100:1. The hybridisation mixture was precipitated and the pellet re-dissolved in 4 μl 3X EE buffer (see Annex). The solution was overlaid with light mineral oil. The DNA was then denatured and sodium chloride solution was added. Hybridisation of DNAs was done overnight at 67°C. Following hybridisation the reaction mixture was amplified by PCR.

The adapters on the amplicons that were obtained after amplification of the first hybridisation reaction, was removed and replaced with a third set of adapters, namely NHpa 12 and 24. A second round of hybridisation was then performed in the same way as the first except that a ratio of driver to tester of 1000:1 was used this time. A second MS-RDA subtraction was also done on the date palm genome that differed from the first in that a different adaptor pair, SHpa 11 and 24 (see Annex), was used instead of the RHpa adapter pair.

5.5.4.2 *P. strobus*

RDA amplicons were prepared as those for date palm using the restriction enzyme *Hpa*II for genomic DNA digestion and the RHpa adaptor pair. The first subtraction was done at a driver to tester ratio of 100:1 followed by a second subtraction of 1000:1 and a third of 10 000:1 using the adaptor pairs (see Annex)

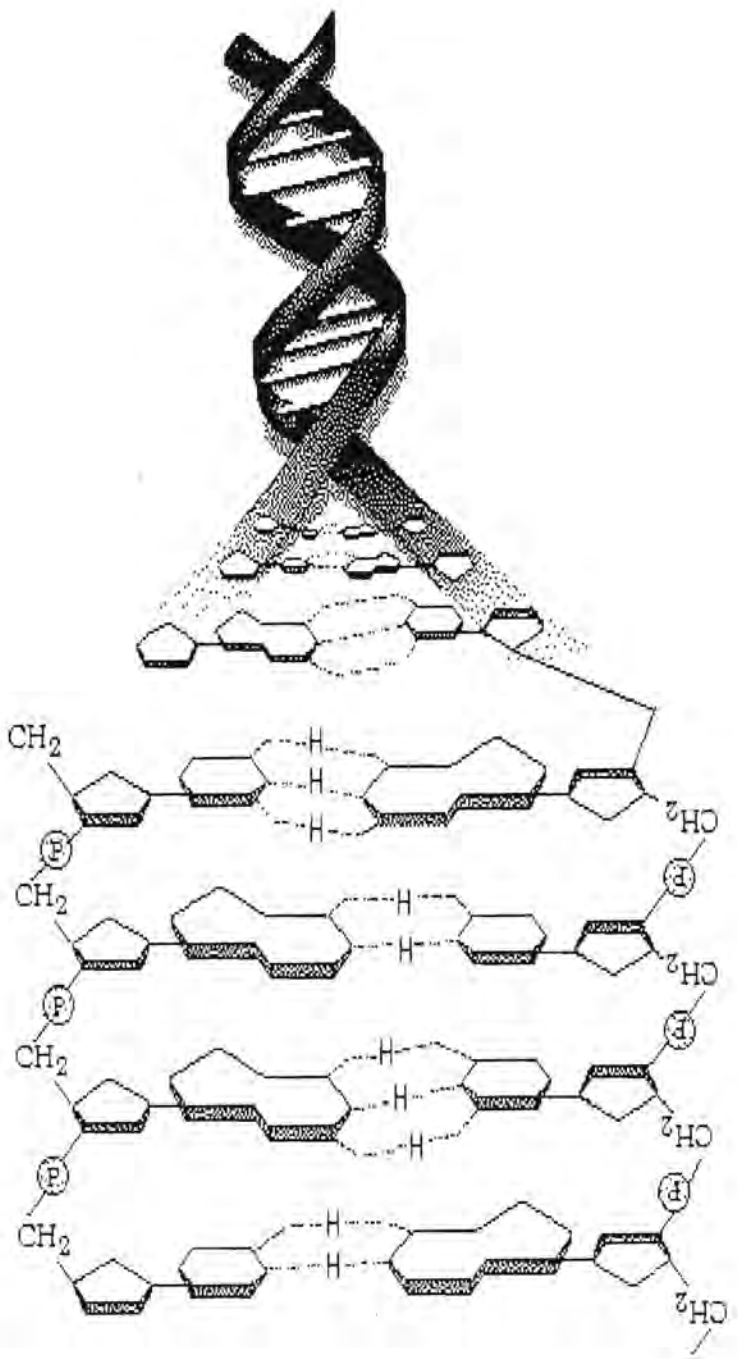
5.4.5 Cloning and sequencing of the difference products

The RDA subtraction products obtained were cloned blunt-ended into the *EcoR*V cloning site of the cloning vector pMos*Blue* (Amersham Life Science, UK). Blue/white selection of *E. coli* colonies and screening for ampicillin-resistant colonies were used to identify recombinant colonies.

Cloned subtraction products were sequenced using an automated sequencer as described in the Annex. Obtained sequences were compared and submitted to the Basic Local Alignment Search Tool (BLAST) (Altschul *et al.*, 1990) to search for homology with known sequences.

Chapter 6

Discussion



6.1 Date palm RAPD analysis

In this study the two date cultivars 'Barhee' and 'Medjool' could be differentiated with polymorphic DNA bands by the RAPD technique as outlined by Corniquel and Mercier (1994) using the RAPD primers OPE-01 and OPE-06. For optimal amplification, it was however vital in this study to optimize the reaction conditions for PCR amplification. This included optimizing the DNA isolation and purification method, the DNA template concentration and the selection of both enzyme and the DNA annealing temperature. The optimal DNA concentration was found to be approximately 25 ng. More or less DNA in the reaction mixture usually resulted in a failure to amplify a DNA fragment. The enzyme used for amplification was also important and best and consistent amplification results were always obtained with Taq polymerase from TaKaRa (Japan). Amplification profiles obtained by other enzymes, such as Taq polymerase from Amersham or Roche, varied and were not useful due to lack of consistency in amplification when test were repeated 3-times. It was also found that profiles were more consistent when the annealing temperature was lowered from the original 36°C to 33°C.

As a new aspect of this study, the potential of designed SCAR primers was investigated. RAPD reactions are very sensitive and the slightest variation in either temperature, quality of DNA or in any of the reagents used can easily result in failure of amplification. This is due to the use of a short 10 mer DNA primer in the RAPD technique. Since the ability of such primer to bind correctly can be influenced by very small changes in the reaction conditions, SCAR primers (20 mers) were constructed from an amplified polymorphic 'Medjool' DNA fragment (700 bp fragment). Designed SCAR primers amplified a DNA fragment with the expected size of 128 bp irrespectively which genomic DNA of date palm was used in the PCR reaction and therefore did not differentiate between the two types of genomic DNA. This possibly indicates that the original polymorphic fragment used for SCAR primer design had a base pair change in the OPE-06 primer site preventing efficient binding of the primer in the PCR reaction.

Overall, in this part of the study the results reported by Corniquel and Mercier (1994) for date palm differentiation using the RAPD technique could be confirmed. But, SCAR primers designed to obtain a more robust PCR amplification system than RAPD failed.

6.2 Isolation of RDA subtraction products from date palm

Application of RDA successfully allowed a direct one-step cloning of putative genome differences of two date palm cultivars despite carrying out a more complex technical procedure than the RAPD technique applied in chapter 2 of this study. With RDA, two types of subtraction products from 'Barhee' genomic DNA with a different size were isolated after subtracting the two genomes. However, subtraction products were only identified after *BamH*I digestion of genomic DNA and not after *Hind*III digestion. Also cross-hybridization with 'Medjool' amplicon DNA of the 'Barhee' subtraction product revealed that these products are present in both genomes and not unique to the 'Barhee' genome. Date palm has a relatively small genome ($5X10^8$ bp), which is roughly 4.2-times greater than the genome of *Arabidiopsis thaliana* ($1.17X10^8$ bp) but very similar to that of rice ($4.3X10^8$ bp). Generally, such small genomes tend to be more stable than larger genomes, which are present in many cereals or trees. The relatively small date palm genome might therefore not have sufficient detectable variation specifically when digested with the restriction enzyme *Hind*III, which is required to easily detect significant sequence differences between cultivars.

It is also unclear why subtraction products were formed and could be isolated from 'Barhee' when these products are present in both genomes. One possibility is that these subtraction products are in excess of more than 100-times in the 'Barhee' genome than in the 'Medjool' genome. This might result in elimination of the subtractive potential of the driver DNA when the driver DNA is only applied in a 100-times higher concentration than the tester DNA. A greatly uneven copy number of subtraction products in both genomes might also explain why no difference product was obtained when 'Medjool' instead of 'Barhee' amplicon DNA was used as a tester. Also, methylation differences might exist between the genomic DNAs of the two cultivars, with 'Medjool' DNA being more methylated. The restriction enzyme *Bam*HI is methylation sensitive and any methylation of a restriction site for this enzyme in the driver DNA might have prevented sufficient PCR amplification of the sequence region detected as a subtraction product when the driver amplicon DNA was produced.

6.3 Bioinformatic analysis of isolated subtraction products

In this part of the study several bioinformatic tools, such as the Smith-Waterman algorithm, FASTA algorithm and BLAST, were applied to characterize one subtraction product sequence (Dp41) in more detail. The Smith-Waterman algorithm is the *de facto* standard for searching databases. It employs the Dynamic Programming (DP) algorithm to each and every database sequence, so it is the most sensitive method for finding related proteins in a database. But, because the DP algorithm effectively makes every possible pair-wise comparison between the query (or reference) sequence and the library (or database) sequences, it is also the slowest method for similarity searches of sequence databases. The Smith-Waterman Algorithm produces local alignment(s) between two sequences. Instead of investigating each sequence in it's entirety it compares segments of all possible lengths, and chooses whichever maximize the similarity measure (Smith and Waterman, 1981). This is useful if some parts of the sequences have undergone any processes of evolution that similarity is undetectable; or if one is only a fragment or only partially overlaps the other. These are all common situations.

The FASTA algorithm is a heuristic approximation to the Smith-Waterman algorithm. In Computer science the term 'heuristic' is by definition: Relating to or using a problem solving technique in which the most appropriate of several found by alternative methods is selected at successive stages of a program for use in the next step in the program. The heuristics used by FASTA allows it to run much faster than the Smith-Waterman algorithm, but at the cost of sensitivity (Pearson and Lipman, 1988). Finally BLAST is a much faster heuristic approximation of the Smith-Waterman algorithm. However BLAST cannot guarantee missing a score as sequences not found by BLAST could still have a significant Smith-Waterman score.

In this study the subtraction product sequence (Dp41) was specifically characterized, using molecular tools, with respect to its variability and presence in individual date palm plants of the two different date palm varieties. A 75% homology was found with rice DNA from chromosome 1 using the Smith-Waterman algorithm but without a known gene function. Sequence analysis of the amplified subtraction products from both date palm varieties, 'Barhee' and 'Medjool', identified a variable region and the existence of several variant members of the repeated sequence. This provides strong evidence that this Dp41 family of DNA sequences represents a variable "genetic hotspot" in the date palm genome as found for rye by Linacero et al. (2000). Variation

detected included changes from C→ T and G→A and deletion of single or several base pairs. In general, deletions occurred at a higher frequency in 'Barhee' than in 'Medjool'. Base pair changes represent the most commonly observed point mutations in plants, and can also be a consequence of a plant tissue culture process where plant growth regulators are used (Phillips et al. 1994). The extent of variation observed between the different plants further indicates that Dp41 might be a rapidly evolving/changing sequence. It might be hypothesized that 'Barhee' plants are generally more susceptible to these variations and that these variations have resulted in a heterogenous 'Barhee' population among non-tissue culture-derived plants including variation in the copy number of this sequence in the different date palm cultivars. However, since all tissue culture-derived 'Barhee' plants originating from a single 'Barhee' mother plant, no conclusions about the behaviour of this family of sequences through tissue culture can be drawn at this stage and might be subject of a further detailed investigation.

Furthermore, PCR amplification patterns were more variable between individual plants collected from a genome bank than from tissue-derived plants, as would be expected because of the clonal nature of the tissue culture plants. Higher variability might also be due to the age of tested plants. Whereas all tissue culture-derived plants were young with a uniform genetic composition, plants collected from the genome bank were mature and fruit bearing. The complexity to which this fragment evolved indicate a genetic hot-spot for genome mutation.

6.4 Detection of methylation changes in the genome using RDA

6.4.1 DNA methylation

In this study, comparison of methylation levels in different types of plants showed higher methylation of genomic DNA in tissue culture plants compared to that of non-tissue culture plants. Further, genomic DNA of non-tissue culture plants digested with the restriction enzymes *Hpa*II and *Msp*I gave similar profiles on an agarose gel. The methylation process of genomic DNA in plants and the effect a tissue culture process has on genomic methylation are both well documented (Arnold-Schmitt *et al.* 1995; Phillips *et al.*, 1994 Fredorff, 2000; Martienssen and Coleot, 2001). However in date palm, the profiles of enzyme-digested 'Medjool' and 'Barhee' DNA differed, with 'Barhee' DNA showing a higher degree of methylation compared to that of 'Medjool'.

In plants and filamentous fungi, genomic methylation is restricted mostly to repetitive DNA elements (Rabinowicz et al., 1999; Colot and Rossignol, 1999). Higher level of methylation might therefore be due to a higher copy number of the highly repetitive subtraction product DP41.

In *P. strobus* increased 2,4-D concentrations in the medium caused higher genomic DNA methylation and higher methylation gave different digestion profiles of DNA with the restriction enzymes *Hpa*II vs. *Msp*I. Both enzymes recognise the same DNA sequence but only methylation at the restriction site produces differences in the digestibility of the DNA, with methylation of DNA rendering the DNA indigestible for *Hpa*II. Furthermore, amplification of different, mostly larger, size fragments for amplicons derived from *Hpa*II digested DNA of cells grown on a high 2,4-D containing medium might be explained by the inability of *Hpa*II to digest methylated DNA to smaller amplifiable fragments.

6.4.2 Date palm

The isolation of repetitive DNA sequences with RDA has been previously reported (Ushijima *et al.*, 1997; Chen *et al.*, 1998). In this study, repeated RHpa 24 adapter sequences were identified in one isolated and characterized subtraction product. Very likely this product does not represent an PCR artefact because identified repeats are interrupted by an extra three base pairs unlikely to be present in an PCR artefact. Also, a similar repeated DNA has been found in other organisms, such as *Hippoglossus* and banana (Cullis, unpublished results). Although an artefact cannot be completely excluded.

Methylation of repeated DNA is a common phenomenon in plant genomes and the amplification or deletion of sequences resulting in alteration of copy number is a known phenomenon occurring during tissue culture (Brown PTH, 1989). Differences in the copy number of subtraction product between the two genomes might also be associated with differences in the levels of methylation in the different genomes, allowing the isolation of subtraction product. The variation found in the fragment length might be due to the RHpa primer binding at different positions during the first round amplification reaction due to the sequence homology found between the primer and the fragment. resulting in amplified fragments with different sizes.

In this study, when RDA was repeated, the elimination of the R series adaptor resulted in the isolation of a new set of difference products. Such isolation of a different set of difference products has also been reported by Ushijima *et al.* (1997) using RDA-WEEC (RDA - With Elimination of Excessive Clones) to amplify nine novel fragments after obtaining two repetitive sequences in the first RDA. Two of the three fragments obtained in this study as subtraction products – a 444 bp fragment and a 330 bp fragment – showed a high level of homology with rRNA genes. These genes, in themselves, represent repetitive DNA sequences and a high copy number of rRNA genes with about 570 repeats per haploid genome have been found in plants (Pruitt and Meyerowitz, 1986). These fragments were obtained only from 'Medjool'. Seemingly same size fragments, which can be efficiently amplified by PCR, were not amplified for 'Barhee' amplicons due to methylation at the HpaII restriction site in genomic 'Barhee' DNA.

6.4.3 *P. strobus*

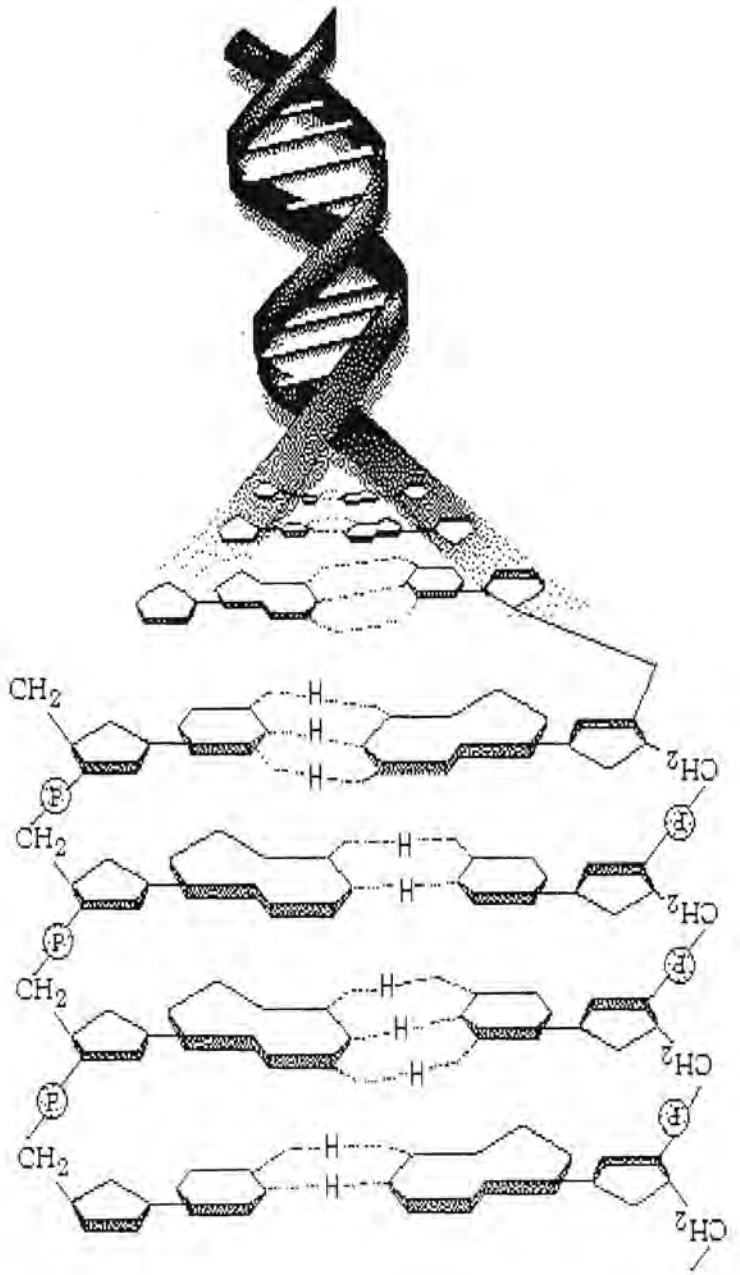
The cleavage polyembryony to the early (precotyledonary) dominance stage represents the most responsive embryo developmental stage for somatic embryogenesis initiation (Percy *et al.*, 2000). Finger *et al.* (1989) showed that high induction rates could be obtained when precoteledonary embryos, enclosed in the megagametophytes, were used as explants. In this study, 40% of explants produced embryonal masses and different levels of 2,4-D applied for cell growth did not resulted in any noticeable differences in growth of the cell cultures. However, in this study noticeable differences were obtained in the number of somatic embryos produced where tissue grown on MLV UL produced more mature embryos. Higher levels of methylation to be present in the genome of tissue grown on high 2,4-D medium might be correlated with loss of gene expression and subsequent loss of embryogenic potential.

The subtractions experiments that were done using genomic DNA of *P. strobus* digested with *Hpa* II failed to reduce the sample size and complexity sufficiently in order to isolate difference products. Even after subtractions with a ratio of 1: 100 000, amplification of the samples resulted in a smear with fragment sizes varying between 200 bp and 2000 bp when visualized on an ethidium bromide stained agarose gel. Whereas the standard RDA protocol is sufficient in reducing the complexity of small

and medium size genomes, large genomes such as that of conifers and specifically *P. strobus* presents an obstacle.

Chapter 7

Achievements and conclusions



There are always advantages and disadvantages associated with every molecular technique and careful consideration should be given before choosing a specific technique. In this study, the applicability of Representational Difference Analysis (RDA) towards plant genome analysis was investigated as one strategy to identify genome differences in two types of plants. Since RDA was originally developed to isolate genetic differences between cancerous and non-cancerous cell lines, one of challenges in this study was to demonstrate if the RDA technique is at all a useful technique to isolate genetic differences on the genome level between closely related plant cultivars or plant cell lines with different phenotypes, and if any isolated difference might ultimately be useful in establishing a genetic marker that can be used either for cultivar identification or for marker assisted selection of superior cell lines and quality control purposes.

In this study, the well-established technique of Random Amplified Polymorphic DNA (RAPD) was first used to demonstrate if two date palm cultivars 'Barhee' and 'Medjool' could be differentiated at all. RAPD results obtained in this study clearly confirmed the findings of Corniquel and Mercier (1994), and RAPD verified the identity of plant material that was used for RDA. Unfortunately, the more robust technique of Sequence Characterised Amplified Region (SCAR) for plant material verification failed in this study indicating that any change in the date palm genome detected by RAPD is possibly due to only a single nucleotide polymorphism in the original RAPD primer-binding site.

For execution of RDA on date palm material three restriction enzymes were used for amplicon production (*Bam*HI, *Hpa*II and *Hind*III) as outlined in the original paper by Lisitsyn *et al.* (1993) and this study achieved as a first outcome to obtain subtraction products from *Bam*HI and *Hpa*II-derived amplicons. *Bam*HI-derived subtraction products were characterized in this study in much more detail and products represented mainly three related sequences. One of these, Dp41, showed more variability and represented a highly repetitive sequence that was up to a hundred-times more abundant in 'Barhee' compared to 'Medjool'. Amplification of this region from a number of individual plants of both cultivars showed further variation in the sequence between cultivars, as well as between individual plants and between different copies amplified from one plant. Several primers were designed from this sequence data and one showed indeed potential for cultivar identification purposes amplifying the expected fragment consistently only from 'Medjool' plants. Further, the sequence variation that was found showed also certain tendencies such as C→ T

and G→ A changes. The sequence variation occurred, however, to such an extent that it could not be explained by applying the rate of natural random mutation to this sequence. This study therefore achieved as a second outcome to identify a possible genetic mutational hot spot in date palm. It might be speculated, that such hot spots act as a genetic buffer to collect mutations, preventing them from occurring in essential genes or other important regions of the genome.

This study also showed as a third outcome, differences in the genomic DNA methylation status between tissue culture and non-tissue culture plants as well as between plants grown on different 2,4-D concentrations. Degree of DNA methylation was clearly higher in tissue culture-derived date palm plants when compared to non-tissue culture plants. Further, DNA methylation of *P.strobus* cells grown on a high 2,4-D containing media were also higher than of cells grown on a lower 2,4-D containing media, which fits to the theory that 2,4-D leads to hyper-methylation.

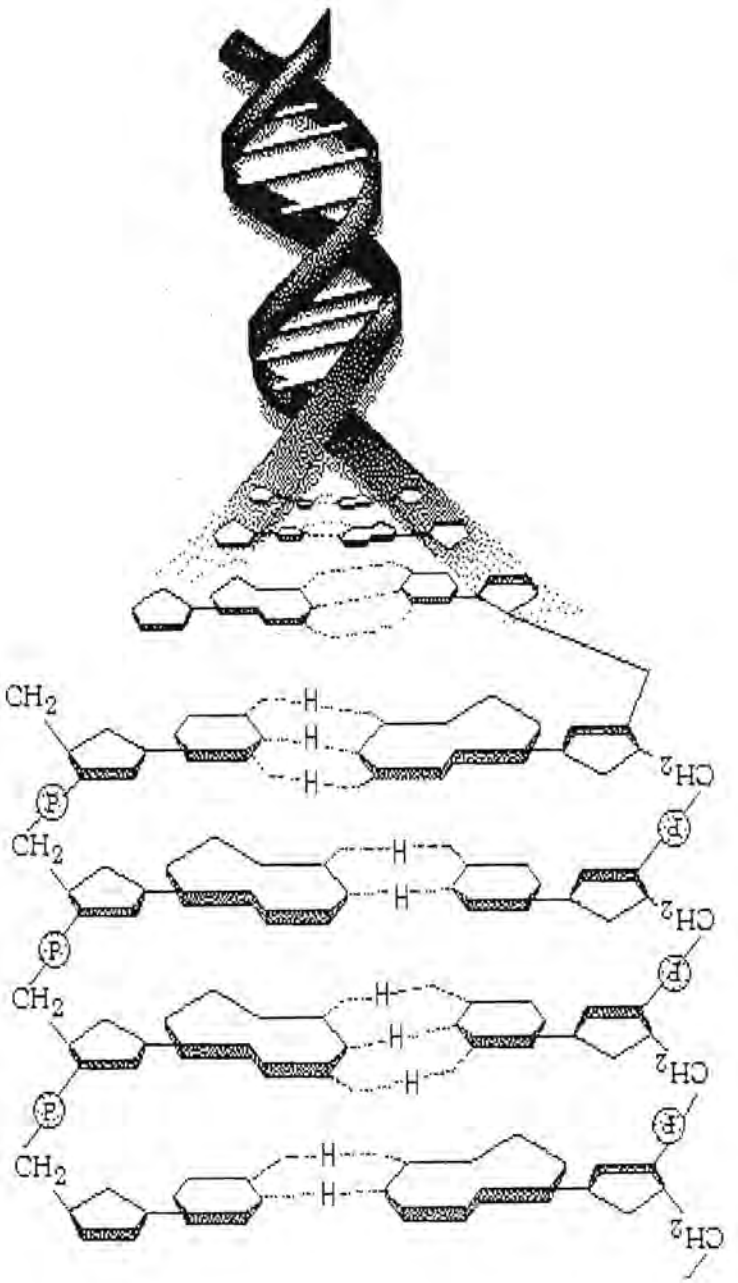
This study also focused on application of the RDA technology on large plant genome using cell lines of *Pinus strobus* as a model and testing if possible DNA methylation differences in a genome due to 2,4-D treatment could be detected by the RDA technology. Unfortunately, RDA was inefficient to identify any subtraction products after three rounds of subtractions of *Hpa*II-derived amplicons and subtraction product profile was still too complex representing a mixture of a great number of fragments too difficult to analyze.

In conclusion, the RDA technique could be extended in this study to two further plant species and has shown potential to isolate variable DNA regions, which differ in plant genomes. Using date palm several repetitive elements were isolated. This would be by far more difficult with, for example, RAPD only based on mutations in short DNA primer sites. However, RDA results in plants have still to be interpreted with caution because a subtractive cloning approach is always subject to several sources of bias. First, the representation of the genome in RDA is based on digestion of the genomic DNA with a single restriction enzyme and therefore depends on the sequence of the restriction site. Secondly, the tester to driver ratios used for subtractive hybridization are critical for the selection of specific sequences, so that a bias introduced by a badly chosen ratio could be further amplified during the PCR step (Zoldos *et al.*, 2001).

Although it was demonstrated in this study that RDA could be successfully applied to analyse small plant genomes, more work remains to be done in the future. To fully understand the significance of these products analysis of the flanking sequences need to be done and the position of the sequences in the genome needs to be determined. Using, RDA-WEEC as previously described by Ushijima *et al.* (1997) can be useful to eliminate the repetitive sequences found commonly in plant genomes and isolate other low-copy regions. In terms of larger genomes such as that of *P. stobus* the size and complexity of the genome seems to be an obstacle for the successful isolation of subtraction products, therefore a greater reduction in the genome complexity by the introduction of a second restriction enzyme such as *Msp*I that cuts the sequence A/ATT commonly found within repetitive sequences might be useful. Otherwise cDNA-RDA that is currently been used to determine differences in gene expression between cell lines grown on high and low 2,4-D containing media might be a better approach when dealing with large plant genomes.

The results obtained from this study resulted in two publications Kunert *et al.*, 2000 and Vorster *et al.*, 2002.

References



Ait-Chitt M., Ainsworth CC. and Thangavelu M. 1993. A rapid and efficient method for extraction of total DNA from mature leaves of date palm (*Phoenix dactylifera* l.) Plant Molecular Biology Reports 11: 317-319

Altschul SF., Gish W., Miller W., Myers EW. and Lipman DJ. 1990. Basic local alignment search tool. Journal of Molecular Biology 215: 403-410

Arnault C. and Dufourmel I. 1994. Genome and stresses: reactions against aggressions, behaviour of transposable elements. Genetica 93: 149-160

Arnold-Schmitt B., Herterich S. and Neumann K-H. 1995. Physiological aspects of genome variability in tissue culture. I. Growth phase-dependant differential DNA methylation of the carrot genome (Daucus carota L.) during primary culture. Theoretical and Applied Genetics 91: 809–815

Ayala F. 1970. Teleological explanations in evolutionary biology. Philosophy of Science (March) 1-15

Barakat A., Han DT., Banslimane A-A., Rode A. and Bernardi G. 1999. The gene distribution in the genomes of pea, tomato and date palm. FEBS Letters 463: 139-142

Barral V., This P., Imbert-Establet D., Combes C. and Delseny M. 1993. Genetic variability and evolution of the Schistosoma genome analised by using random amplified polymorphic DNA markers. Molecular Biochemistry and Parasitology 59: 211- 222

Bender CM., Pao MM. and Jones PA. 1998. Inhibition of DNA methylation by 5-aza-2'-deoxycytidine suppresses the growth of human tumor cell lines. Cancer Research 58: 95- 101

Bennett MD. and Smith JB. (1991). Nuclear DNA amounts in angiosperms. *Philosophical Transactions of the Royal Society of London* B 334: 309-345

Bennetzen JL. and Kelogg EA. 1997. Do plants have a one-way ticket to genome obesity? Plant Cell 9: 1509-1514

Bennetzen JL., Schrick K., Springer PS., Brown WE. And SanMiguel P. 1994. Active maize genes are unmodified and flanked by diverse classes of modified, highly repetitive DNA. Genome 37: 565– 576.

Bloom KS. and Carbon J. 1982. Yeast centromere DNA is in a unique and highly ordered structure in chromosomes and small circular minichromosomes. Cell 29: 305-317

Blundy KS., Cullis CA. and Hepburn A.G. 1987. Ribosomal DNA methylation in a flax genotroph and a crown gall tumor. Plant Molecular Biology 8: 217-225

Bodens C., Joandet S., Laigret F. and Kremer A. 1997. Detection of genomic regions differentiating two closely related oak species *Quercus petreae* (Matt.) and *Q. robur* (L.). Heredity 78: 433-444

Bouman H. and de Klerk GJ. 1997. Somaclonal variation. In: Biotechnology of Ornamental Plants p. 165-183. Oxford University Press, NY.

Brown PTH. 1989. DNA methylation in plants and its role in tissue culture. Genome 31: 717- 729

Brown TA. 1992. Genetics a molecular approach. Second edition. Chapman and Hall London.

Buchanan BB., Gruissem W. and Jones RL. 2000. Biochemistry and molecular biology of plants. American Society of Plant Physiologists, Rochville, Maryland.

Burns TD., White TJ. and Taylor JW. 1991. Fungal molecular systematics. Annual Review of Ecological Systematics 22: 524- 564

Buscot F., Wipf D., Di Battista C., Munch J-C, Botten B. and Martin F. 1996. DNA polymorphisim in morels: PCR/ RFLP analysis of the ribosomal DNA spacer and microsatellite primed PCR. Mycology Research 100: 63- 71

Capy P. 1998. A plastic genome. Nature 396: 522-523

Casacuberta JM., Vernhettes S. and Grandbastien M-A. 1995. Sequence variability within the tobacco retrotransposon *Tnt*1 population. EMBO Journal 14: 2670-2678

Chandler V., Rivin C. and Walbot W. 1986. Stable non-mutator stocks of maize have sequences homologous to the Mu1 transposable element. Genetics 114: 1007-1021

Chang Y., Cesarman E., Pessin M.S., Lee F., Culpepper J., Knowles DM. and Moore PS. 1994. Identification of herpes virus-like DNA sequences in AIDS-associated Kaposi's sarcoma. Science 266: 1865-1869

Chen M., SanMiguel P. and Bennetzen JL. 1998. Sequence organizatin and conservation in sh2/ a1- homologous regions of sorgum and rice. Genetics 148: 435-443.

Colot V. and Rossignol. 1999. Eukaryotic DNA methylation as a evolutionary device. BioEssays 21: 402- 411

Copenhaver GP. and Preuss D. 1999. Centromeres in the genomic era: unravelling paradoxes. Current Opinions in Plant Biology 2: 104- 108

Corniquel B. and Mercier L. 1994. Date palm (*Phoenix dactylifera* L.) cultivar identification by RFLP and RAPD. Plant Science 101: 163-172

Crow JF. 1997. The high spontaneous mutation rate: Is it a health risk? Proceedings of the National Academy of Science USA 94: 8380-8386

Cullis CA. 1990. DNA rearrangements in response to environmental stress. Advances in Genetics 28: 73-97

Cullis CA. 1999. Environmental Stress - a generator of adaptive variation? In: H.R. Lerner (ed.). Plant Adaptations to Stress Environments. Marcel Dekker, New York, in press.

Cullis CA. and Kunert KJ. 2000. Isolation of tissue culture-induced polymorphisms in bananas by representational difference analysis. Acta Horticulturae 530: 421-428

Cullis CA., Rademan S. and Kunert KJ. 1999. Method for finding genetic markers of somaclonal variation. International publication number WO 99/53100.

Curtis CA. and Lukaszewski AJ. 1991. Genetic linkage between C- bands and storage protein genes in chromosome 1B of tetraploid wheat. Theoretical and Applied Genetics 81: 245- 252

Deng Z., Huang S., Xiao S. and Gmitter JR. FG. 1997. Development and characterization of SCAR markers linked to the *Citrus tristeza*. Genome 40: 697-740

Donnison IS., Siroky J., Vyskot B., Saedler H. and Grant SR. 1996. Isolation of Y chromosome-specific sequences from *Silene latifolia* and mapping of male sex-determining genes using representational difference analysis. Genetics 144: 1893-1901

Eggleston WB., Alleman M. and Kermicle JL. 1995. Molecular organization and germinal instability of *R- stippled* maize. Genetics 141: 347- 360

Evans DA. 1988. Applications of somaclonal variation. Biotechnology in Agriculture p. 203-223

Fedoroff N. 2000. Transposons and genome evolution in plants. Proceedings of the National Academy Of Science USA 97: 7002-7007

Fedoroff N., Furtek D. and Nelson O. 1984. Cloning of the Bronze locus in maize by a simple and generalizable procedure using the transposable controlling element activator (AC). Proceedings of the National Academy Science USA 81: 3825-3829

Fedoroff N., Wessler S. and Shure M. 1983. Isolation of the transposable maize controlling elements Ac and Ds. Cell 35: 243- 251

Feldman M., Liu B., Segal G., Abbo S., Levy AA. and Vega JM. 1997. Rapid elimination of low- copy DNA sequences in polyploid wheat: a possible mechanism for differentiation of Homoeologous chromosomes. Genetics 147: 1381- 1387.

Fitzgerald-Hayes M., Clarke L. and Carbon J. 1982. Nucleotide sequence comparisons and functional analysis of yeast centromere DNAs. Cell 29: 235-244

Flavell AJ., Dunbar E., Anderson R., Pearce SR., Hartly R. and Kumar A. 1992b. Ty1-*copia* group retrotransposons are ubiquitous and heterogeneous in higher plants. Nucleic Acids Research 20: 3639-44

Flavell AJ., Smith DB. and Kumar A. 1992a. Extreme heterogeneity of Ty1-*copia* group retrotransposons in plants. Molecular and General Genetics. 231:233-42

Flavell RB., Bennett MD., Smith JB. and Smith D. 1974. Genome size and the proportion of repeated nucleotide sequence DNA in plants. Biochemical Genetics 12: 257-269

Finer JJ., Krebel HB. and Becwar MR. 1989. Initiation of embryogenic callus and suspension cultures of eastern white pine (*Pinus strobus* L.) Plant Cell Reports 8: 203- 206

Foster LM., Kozac KR., Loftus MG., Stevens JJ. and Ross IK. 1993. The polymerase chain reaction and its application to the filamentous fungi. Mycological Research 97: 769-781

Gill BS., Fiebe B. and Endo TR. 1991. Standard karyotype and nomenclature system for description of chromasome bands and structural aberrations in wheat (*Triticum aestivum*). Genome 34 : 830- 839

Grandbastien M-A. 1998. Activation of plant retrotransposons under stress conditions. Trends Plant Science 3: 181-187

Graur D., Shuali Y. and Li WH. 1989. Deletions in processed pseudogenes accumulate faster in rodents than in humans. Journal of Molecular Evolution 28: 279-285

Gu WK., Weeden NF., Yu J. and Wallace DH. 1995. Large-scale, cost-effective screening of PCR products in marker assisted selection application. Theoretical and Applied Genetics 91: 465- 470

Haberer G. and Fisher TC. 1996. Mapping of the nucleolus organizer on chromosome 4 in *Arabidopsis thaliana*. Molecular and General Genetics 250: 123-128.

Hadrys H., Balick M. and Schierwater B. 1992. Applications of RAPD fingerprinting in molecular ecology. Molecular Ecology

Heldt HW. 1997. Plant Biochemistry & Molecular Biology. Oxford University Press, New York.

Henry RJ. 1997. Practical Applications of Plant Molecular Biology. Chapman & Hall, London.

Hirochika H. 1993. Activation of tobacco retrotransposons during tissue-culture. EMBO Journal 12: 2521-2528

Hirochika H., Sugimoto K., Otsuki Y., Tsugawa H. and Kanda M. 1996. Retrotransposons of rice involved in mutations induced by tissue culture. Proceedings of the National Academy of Science USA 93: 7783-7788

Hu W., Das OP. and Messing J. 1995. *Zeon*-1, a member of a new maize retrotransposon family. Molecular and General Genetics 248: 471-480

Jiang JM., Nasuda S., Dong FG., Scherrer CW. and Woo SS. *et al.*, 1996 A conserved repetitive DNA element located in the centromeres of cereal chromosomes. Proceedings of the National Academy of Science USA 93:14210-14213

Jin Y-K. and Bennetzen JL. 1989. Structure and coding properties of *Bs*1, a maize retrovirus-like transposon. Proceedings of the National Academy of Science USA 86:6235-6239

Johnston JS., Jensen A., Czeschin DG. and Price HJ. 1996. Environmentally induced nuclear DNA content instability in *Helianthus annus* (Asteraceae). American Journal of Botany 83: 1113-1120

Jones N., Ougham H. and Thomas H. 1997. Markers and mapping: we are all geneticists now. New Phytologist 137: 165-177

Judson HF. 1996. The Eighth Day of Creation. Cold Spring Harbor Press, New York

Karp A. 1993. Are your plants normal? - Genetic instability in regenerated and transgenic plants. Agro-Food-Industry Hi-Tech: May/June 7-12

Kunert KJ., Vorster J., Bey E. and Cullis C. 2000. Introduction of DNA-based quality assurance procedure for commercially micropropagated date palms in South Africa. Proceedings of the International Date Palm Symposium,Windhoek/Namibia, pp 74-80.

Larkin PJ. and Scowcroft WR. 1981. Somaclonal variation - a novel source of variability from cell cultures for plant improvement. Theoretical and Applied Genetics 60: 197-214

Laszlo E. 1996. The systems view of the world: A holistic vision for our time. (Advances in systems theory, complexity, and the human sciences). Hampton Press, Australia.

Lee M. and Phillips RL. 1986. Genomic rearrangements in maize induced by tissue culture. Genome 29: 122-128

Leroy XJ., Hily JM. and Chaumeil P. 2001. Detection of in vitro culture-induced instability through inter-simple sequence repeat analysis. Theoretical and Applied Genetics 102: 885-891

Leutwiler LS., Hough- Evans BR. and Meyerowitz EM. 1984. The DNA of *Arabidopsis thaliana*. Genetics 194: 15- 23

Levin I., Crittenden LB. and Dogson JB. 1993. Genetic map of the Z chromosome using random amplified polymorphic DNA (RAPD) markers. Genomics 16: 224- 230

Levitt J. 1972. Responses of plants to environmental stresses. Academic Press, New York and London

Lieb M. and Rehmat S. 1997. 5-Methylcytosine is not a mutation hot spot in nondividing *Escherichia coli*. Proceedings of the National Academy of Science USA 94: 940- 945

Linacero R., Alves EF. and Vazquez AM. 2000. Hot spots of DNA instability revealed through the study of somaclonal variation in rye. Theoretical and Applied Genetics 100: 506-511

Lisitsyn N., Lsistsyn N. and Wigler M. 1993. Cloning the differences between two complex genomes. Science 259: 946-950

Lu Z., Kunnimalaiyaan M. and Nielsen BL. 1996. Characterization of replication origins flanking the 23S rRNA gene in tobacco chloroplast DNA. Plant Molecular Biology 32: 693-706

MacLean DJ., Braithwaite KS., Manners JM. and Irwin JAG. 1993. How do we identify and classify fungal pathogens in the era of DNA analysis? In Advances in Plant Pathology Volume 10. Andrews JH. and Tommerup IC. (Ed.) Academic Press London

Marillonnet S. and Wessler SR. 1998. Extreme structural heterogeneity among the members of a maize retrotransposon family. Genetics 150: 1245- 1256

Martiensenn RA. and Colot V. 2001. DNA methylation and epigenetic inheritance in plants and filamentous fungi. Science 293: 1070- 1074

Masson P., Surosky R., Kingsbury JA. and Fedoroff NV. 1987. Genetics 177: 177-137

McClintock B. 1984. The significance of responses of the genome to challenge. Science 226: 792-801

McCouch SR. 2001. Genomics and synteny. Plant Physiology 125: 152-155

Mhiri C., De Wit PJGM. and Grandbastien M-A. 1999. Activation of the promoter of the *Tnt*1 retrotransposon in tomato after inoculation with the fungal pathogen *Cladosporium fulvum*. Molecular Plant-Microbe Interactions 12: 592-603

Mhiri C., Morel JB., Vernhettes S., Casacuberta JM., Lucas H. and Grandbastien MA. 1997. The promoter of the tobacco *Tnt*1 retrotransposon is induced by wounding and by abiotic stress. Plant Molecular Biology 33: 257-266

Miller JT., Dong F., Jackson SA., Song J. and Jiang J. 1998. Retrotransposon-related DNA sequences in the centromeres of grass chromosomes. Genetics 150: 1615- 1623.

Moor G., Lucas H., Batty N. and Flavell R. 1991. A family of retrotransposons and associated genomic variation in wheat. Genomics 10: 461- 468.

Moreau-Mhiri C., Morel JB., Audeon C., Ferault M., Grandbastien MA. and Lucas H. 1996. Regulation of expression of the tobacco *Tnt*1 retrotransposon in heterologous species following pathogen related stresses. Plant Journal 9: 409-419

Murray B. (1998). Nuclear DNA amounts in gymnosperms. *Annals of Botany* 82 (Supplement A): 3-15.

Nagl W. and Rucker W. 1976. Effect of phytohormones on thermal denaturation profiles of *Cymbidium* DNA: Indication of differential DNA replication. Nucleic Acids Research 3: 2033-2039

Navin A, Prekeris R, Lisitsyn NA, Sonti MM, Grieco DA, Narayanswami S, Lander ES, Simpson EM. 1996. Mouse Y-specific repeats isolated by whole chromosome representational difference analysis. Genomics 36: 349-353

Nekrutenko A, Makova KD, Baker RJ. 2000. Isolation of binary species-specific PCR-based markers and their value for diagnostic applications. Gene 249: 47-51

Nelke M., Nowak J., Wright J.M. and Mclean N.L. 1993. DNA fingerprinting of red clover (*Trifolium pratens* L.) with Jeffrey's probes: detection of somaclonal variation and other applications. Plant Cell Report 13: 72-78

Nierras CR., Liebman SW. and Warner JR. 1997. Does Saccharomyces need and organized nucleolus? Chromosoma 105: 444-451

Nixon RW. (1950) Imported varieties of dates in the United States. US Department of Agriculture Circular No. 834

O'Brien IEW., Smith DR., Gardner RC. and Murray BG. 1996. Flow cytometric determination of genome size in Pinus. Plant Science 115: 91-99

O'Hanlon PC., Peakall R. and Briese DT. 2000. A review of new PCR- based genetic markers and their utility to weed ecology. Weed Research 40: 239- 254

O'Niell RJW., O'Niell, MJ., and Graves, MA. 1998. Undermethylation associated with retroelements activation and chromosomes remodelling in an interspecific mammalian hybrid. Nature 393: 68-72

Olszewska MJ. and Osiecka R. 1982. The relationship between 2C DNA content, life-cycle type, systematic position, and the level of DNA endoreplication in nuclei of parenchyma cells during growth and differentiation of roots in some monocotyledonous species. Biochemie Und Physiologie Der Pflanzen 177: 319-336

Paran I. and Michelmore, RW. 1993. Development of reliable PCR-based markers linked to downy mildew resistance genes in lettuce. Theoretical and Applied Genetics 85: 985-993

Pardue ML., Danilevskaya ON., Lowenhaupt K., Slot F. and Traverse KL. 1996. Drosophila telomeres: new views on chromosome evolution. Trends in Genetics 12: 48-52

Pearce SR., Harrison G., Li D., Heslop-Harrison JS., Kumar A. and Flavell AJ. 1996. The Ty1-copia group retrotransposons in Vica species: copy number, sequence heterogeneity and chromosomal location. Molecular and General Genetics 250: 305-315

Pearson WR. and Lipman DJ. 1988. Improved tools for biological sequence comparison. Proceedings of the National Academy of Science USA 85: 2444-2448

Percy RE.,m Klimaszewska K. and Cyr DR. 2000. Evaluation of somatic embryogenesis for clonal propagation of eastern white pine. Canadian Journal of Forest Reasearch 30: 1867- 1876.

Peschke VM., Phillips RL. and Gengenbach BG. 1987. Discovery of transposable element activity among progeny of tissue culture-derived maize plants. Science 238: 804-807

Petrov DA., Lozovskaya ER. and Hartl DL. 1996. High intrinsic rate of DNA loss in Drosophila. Nature 384: 346-349

Phillips RL., Kaeppler SM. and Olhoft P. 1994. Genetic instability of plant tissue cultures: Breakdown of normal controls. Proceedings of the National Academy of Science USA 91: 5222-5226

Poulsen GB., Kahl G. and Weising K. 1993. Abundance and polymorphism of simple repetitive DNA sequences in *Brassic napus* L. Theoretical and Applied Genetics 85: 994-1000

Pouteau S., Gandbasien MA., Boccara M. 1994. Microbial elicitors of plant defense response activate transcription of a retrotransposon. Plant Journal 5: 535-542

Pruitt RE., and Meyerowitz EM. 1986. Characterization of the genome of *Arabidopsis thaliana*. Journal of Molecular Biology 187: 169-183

Rabinowicz PD., Schutz K., Dedhia N., Yordan C., Parnell LD., Stein L., McCombie WR. and Martienssen RA. 1999. Differential methylation of genes and retrotransposons facilitates shotgun sequencing of the maize genome. Nature Genetics 23: 305- 308

Rothnie HM, McCurrach KJ., Clover LA. and Hardman N. 1991. Retrotransposon-like nature of Tp1 elements: implications for the organisation of highly repetitive, hypermethylated DNA in the genome of *Physarum polycephalum*. Nucleic Acids Research 19: 279- 286

Russell JR., Fuller JD., Macaulay M., Hatz BG., Jahoor A., Powell W. and Waugh R. 1997. Direct comparison of levels of genetic variation among barley accessions detected by RFLPs, AFLPs, SSRs and RAPDs. Theoretical and Applied Genetics 95: 714–722

Sabir A., Newbury HJ., Todd G., Catty J. and Fors-Lloud BV. 1992. Detection of genetic stability using isozymes and RFLPs in beet plants regenerated in vitro. Theoretical and Applied Genetics 84: 113-114

Sambrook J., Fritsch EF. and Maniatis T. 1989. Molecular cloning: a laboratory manual, 2nd edn. Cold Spring Harbor Laboratory Press, New York

Sandhu D. and Gill KS. 2002. Gene- containing regions of wheat and the other grass genomes. Plant Physiology 128: 803- 811

Sandhu D., Champoux JA., Bondareva SN. and Gill KS. 2001. Identification and physical localization of useful genes and markers to a major gene- rich region on wheat group *1S* chromosomes. Genetics 157: 1735- 1747

SanMiguel P., Gaut, BS., Tikhonov, A., Nakajima, Y. and Bennetzen, JL. 1998. The palaeontology of intergene retrotransposons of maize. Nature Genetics 20: 43-45

SanMiguel P., Tikhonov A., Jin YK., Motchoulskaia N. and Zakharov D, et al., 1996. Nested retrotransposons in the intergenic regions of the maize genome. Science 274: 765-68

Schilling AG. 1996. Polymerase chain reaction-based assays for species-specific detection of *Fusarium culmorum*, *F.graminearum*, and *F.avenaceum*. Phytopathology 86: 515 522

Schwarz-Sommer Z., Gieri A., Cuypers H., Peterson PA. and Saedler H. 1985. Plant transposable elements generate the DNA sequence diversity needed for evolution. EMBO Journal 4: 591-597

Scribner KT. and Pearce JM. 2000. Microsatellites: evolutionary and methodological background and empirical applications at an individual, population and phylogenetic levels. In Molecular methods in ecology, Baker A (ed), Blackwell London pp. 235-271

Sedra MH., Lashermes P., Trouslot P., Combes MC. and Hamons S. 1998. Identification and genetic diversity analysis of date palm (*Phoenix dactilyfera* L.) varieties from Morocco using RAPD markers. Euphytica 103: 75- 82.

Selker EU. and Garrett PW. 1988. DNA sequence duplications trigger gene inactivation in *Neurospora crassa*. Proceedings of the National Academy of Science USA 85: 6870- 6874.

Selker EU., Cambareri EB., Jenesen BC. and Haack KR. 1987. Rearrangement of duplicated DNA in specialized Neurospora.Cell 51: 741- 752.

Skirvin RM. and Janick J. 1976. Tissue culture-induced variation in scented *Pelargonium* spp. Journal of the American Society for Horticultural Science 101: 281-290

Skirvin RM., McPheeters KD. and Norton M. 1994. Sources and frequency of somaclonal variation. HortScience 29: 1232-1237

Smith TF. and Waterman MS. 1981. Identification of common molecular subsequences. Journal of Molecular Biology 25; 147: 195-197

Smyth DR., Kalitsis P., Joseph JL. and Sentry JW. 1989. Plant retrotransposon from *Lilium henryi* is related to Ty3 of yeast and the *gypsy* group of *Drosophila*. Proceedings of the National Academy of Science USA 86: 5015- 5019.

Soller M. and Beckmann JS. 1986. Restriction fragment length polymorphisims in poultry breeding. Poultry Science 65: 1474- 1488

Strathdee G. and Brown R. 2002. Aberrand DNA methylation in cancer: potential clinical interventions. Expert Reviews in Molecular Medicine. 4 March pp. 1- 17

Sun X., Wahlstrom J. and Karpen G. 1997. Molecular structure of a functional Drosophila centromere. Cell 91: 1007-1019

Suoniemi A., Anamthawat-Jonsson K., Arna T. and Schulman AH. 1996. Retrotransposon BARE-1 is a major, dispersed component of the barley (Hordeum vulgare L.) genome. Plant Molecular Biology 30: 1321- 29

Takeda S., Sugimoto K., Otsuki H. and Hirochika H. 1998. Transcriptional activation of the tobacco retrotransposon *Tto1* by wounding and methyl jasmonate. Plant Molecular Biology 36: 365–376.

Takeda S., Sugimoto K., Otsuki H. and Hirochika H. 1999. A 13-bp *cis*-regulatory element in the LTR promoter of the tobacco retrotransposon *Tto1* is involved in responsiveness to tissue culture, wounding, methyl jasmonate and fungal elicitors. Plant Journal 18: 1- 11.

Thomas CA. 1971. The genetic organization of chromosomes. Annual Review of Genetics 5: 237-256

Toder R., Grutzner F., Haaf T. and Bausch E. 2001. Species-specific evolution of repeated DNA sequences in great apes. Chromosome Research 9: 431-435

Twyman RM. 1998. Advances in molecular biology: a concise reference. Bios Scientific Publishers, Springer- Verlag, New York.

Ushijima T., Morimura K., Hosoya Y., Okonogi H., Tatematsu M., Sugimura T. and Nagao M. 1997. Establishment of methylation-sensitive-representational difference analysis and isolation of hypo- and hypermethylated genomic fragments in mouse liver tumors. Proceedings of the National Academy of Science USA. 94: 2284-2289

Vos P., Hogers R., Bleeker M., Reijans M., Lee T., Van de Hornes M., Frijters A., Pot J., Peleman J., Kuiper M. and Zabeau M. 1995. AFLP: a new technique for DNA fingerprinting. Nucleic Acids Research 23: 4407- 4414.

Vongs A., Kakutani T., Martienssen RA. and Richards EJ. 1993. *Arabidopsis thaliana* DNA methylation mutants. Science 260: 1926- 1928.

Vorster BJ., Kunert KJ. and Cullis CA. 2002. Use of representational difference analysis for the characterization of sequence differences between date palm varieties. Plant Cell Reports 21:271–275

Voytas DF., Cummings MP., Konieczny A., Ausubel FM. and Rodermel SR. 1992. *Copia*-like retrotransposons are ubiquitous among plants. Proceedings of the National Academy of Science USA 89: 7124-7128

Wang DG., Fan JB., Siao CJ., Berno A., Young P., Sapolsky R., Ghandour G., Perkins N., Winchester E., Spencer J., Kruglyak L., Stein L., Hsie L., Topaloglou T., Hubbell E., Robinson E., Mittmann M., Morris MS., Shen N., Kilburn D., Rioux J., Nusbaum C., Rozen S., Hudson TJ., Lander ES., et al. 1998. Large-scale identification, mapping, and genotyping of single-nucleotide polymorphisms in the human genome. Science 280: 1077-1082

Welsh J. and McClelland M. 1990. Fingerprinting genomes using PCR with arbitrary primers. Nucleic Acids Research 18: 7213- 7218

White SE., Habera LF. and Wessler SR. 1994. Retrotransposons in the flanking regions of normal plant genes – a role for *copia*-like elements in the evolution of gene structure and expression. Proceedings of the National Academy of Science USA 91: 11792- 96

Williams JGK., Kubelik AR., Livak KJ., Rafalski JA. and Tingey SV. 1990. DNA polymorphisms amplified by arbitrary primers are useful as genetic markers. Nucleic Acids Research 18: 6531- 6535.

Wright WE., Brasiskyte D., Piatyszek MA. and Shay JW. 1996. Experimental elongation of telomeres extends the lifespan of immortal x normal cell hybrids. EMBO Journal 15: 1734-1741

Yoder JA., Walsh CP. and Bestor TH. 1997. Cytosine methylation and the ecology of intagenomic parasites. Trends in Genetics 13: 335-40

Zakian VA. 1996. Structure, function and replication of Saccharomyces cerevisiae telomeres. Annual Review of Genetics 30: 141-172

Zheng KL., Castiglione S., Biasini MG., Biroli A., Morandi C. and Sala F. 1987. Nuclear DNA amplification in cultured cells of *Oryza sativa* L. Theoretical and Applied Genetics 74: 65-70

Zoldos V., Siljak-Yakovlev S., Papes D., Sarr A. and Panaud O. 2001. Representational difference analysis reveals genomic differences between *Q. robur* and *Q. suber*: implications for the study of genome evolution in the genus *Quercus*. Molecular Genetics and Genomics 265: 234-241

Annex

I : Molecular methods

II : Tissue culture

III : General solutions and buffers

IV : Plasmid maps

V : Primer sequences

I) Molecular Methods

I.I Plant material and DNA extraction

'Medjool' and 'Barhee' plant material originated from 14 plants grown in California that have been propagated solely via off-shoots from mother plant material imported in the beginning of the last century from Northern Africa (Nixon 1950). The *in vitro* plants used were 'Medjool' derived from explant material collected in California, and 'Barhee' derived from explant material collected in the United Arabic Emirates.

Total cellular DNA was isolated from leaves (1 g) using the Nucleon Phytopure Plant DNA extraction kit (Amersham Life Science, UK) according to the manufacturer's instructions. To test the quality and amount of isolated DNA, samples of isolated DNA (1 µl) were run on a 1% agarose gel in TAE buffer (0.04 M Tris-acetate; 1 mM EDTA, pH 8) as outlined by Sambrook et al. (1989). After staining of the gel with ethidium bromide for 15 min, DNA quality was determined on a white/UV-transilluminator, photographed using a Grab-IT system (Vacutec, USA) and the DNA concentration of the samples was visually determined using 4 different λ-phage DNA amounts as standards (25 ng, 50 ng, 100 ng and 250 ng) for comparison.

I.II Polymerase chain reaction (PCR)

Standard DNA amplifications by PCR were carried out in 25 µl reaction volumes containing 500 mM KCl, 25 mM $MgCl_2$, 100 mM Tris-HCl (pH 8.3), 25 mM of each dNTP and 5 units TaKaRa DNA polymerase (TaKaRa, Japan) in a GeneAmp PCR 9600 system (Perking Elmer, USA). Primers for PCR were obtained either from Operon Technologies (Operon Technologies, USA) or were designed using the online primer design tool of Molecular Biology Shortcuts (MBS), program 'Oligos and Primers' (www.mbshortcuts.com/biotools/index.htlm). The designed primers were then manufactured and purchased from MWG-Biotech AG (Germany). The standard PCR program consisted of 94°C (5 min) to denature double stranded DNA. This was followed by 35 cycles of amplification consisting of denaturing DNA at 94°C (1 min), primer annealing at 55°C or higher depending on primer pair (1 min), and extension of the DNA chain at 72°C (2 min). This was followed by another extension cycle at 72°C for 5 min.

I.III Representational difference analysis (RDA)

I.III.I *Preparation of RDA amplicons*

RDA was performed following the general outline described by Lisitsyn et al. (1993). Two micrograms of each of the DNAs ('Barhee' and 'Medjool') were digested with 80 units of the enzyme *Bam*HI or *Hind*III (Amersham Life Science, UK) 100 μg of yeast tRNA (Sigma-Aldrich, USA) was added to the digest to limit non-specific binding of DNA against the sides of eppendorf tubes and pipet tips. The digests were then extracted with an equal volume of phenol/chloroform. The DNA containing upper phase was removed and the DNA precipitated using 1/10 volume 3 M sodium acetate and 2.5 volumes 100% ethanol. The precipitate was washed in 70% ethanol and dried. The digested DNA was dissolved in 18 μl of TE buffer (10 mM Tris-Cl, 1 mM EDTA), pH 8, at a concentration of 100 μg/ml.

I.III.II Ligation of adaptor sequences

One μg of the *Bam*HI and *Hind*III digested DNA were ligated to the adaptor pair RBam 12 and RBam 24 or RHind 12 and RHind 24, respectively. The DNA, primer set (0.6 μM), 10X ligase buffer and water to a final volume of 30 μl were mixed in a microcentrifuge tube. The tubes were placed in a heating block at 55°C and the block was the placed at 4°C until the temperature had fallen to 12°C (60 to 75 min). The tubes were then placed on ice for 3 min, after which 1 μl of T4 DNA ligase (1 unit) (Amersham Life Science, UK) was added and the reaction incubated at 16°C overnight. The ligation reaction was then diluted with 970 μl TE buffer, pH 8, and the amplification reaction was set up as follows:

80 μl 10X *Taq* Buffer (100 mM Tris-HCl, pH 8.3, 500 mM KCl, 15 mM $MgCl_2$)
64 μl dNTP mix (160 mM of each dNTP)
16 μl (1.0 μM) of one primer of the appropriate pair (J Bam 24 or R Hind 42)
80 μl of the diluted ligation mixture
560 μl water

The tube was placed in a heating block at 72°C for 5 min to melt of the 12-mer oligonucleotide. Six µl (30 units) of *Taq* polymerase (TaKaRa, Japan) was added and mixed by pipetting. The mixture was then aliquoted into eighth PCR tubes in a Perkin-Elmer GeneAmp 9600 thermocycler, the block of which was being held at 72°C. After 5 minutes, the thermocycler program run was: 20 cycles: 95°C for 30 sec; 72°C for 2 min, followed by 10 min at 72°C, followed by holding at 4°C.

The 8 tubes were combined an a 10 µl aliquot was run on a 1.5% TAE agarose gel at 100 V for 30 minutes to check the amplification and estimate the concentration by comparison to lambda phage DNA standards. The remainder was extracted with 600 µl phenol/chloroform. The upper phase was removed and the DNA precipitated with 3 M sodium acetate and an equal volume propanol. The precipitate was collected by centrifugation, washed twice with 70% ethanol and dried. The amplicons was redissolved in 80µl TE buffer, pH 8.

I.III.III Removal of the adaptors from amplicons

Of the amplicons that was to be used as driver 150 µg and 10 µg of the tester amplicons were digested with the appropriate enzyme in 800 µl and 200 µl, respectively, at an enzyme concentration of 20 units/µg DNA, at 37°C for one hour. Ten µg Yeast tRNA was added and each digest extracted with phenol/chloroform. The DNA was precipitated with an equal volume of isopropanol and 1/10 volume 3 M sodium acetate, as described above. Isopropanol is used instead of ethanol in order to eliminate very small fragments (e.g. adaptors) from the precipitate. The driver and tester DNAs were both redissolved at approximately 400 µg/ml. Both the digested driver and tester amplicons were run on a 1.5% TAE agarose gel alongside an equal aliquot of undigested amplicons to check the completeness of the digestion. On the same gel, standard lambda phage DNA was run so that the final concentrations of driver and tester DNAs could be estimated and adjusted if necessary.

I.III.IV Change of adaptors on tester amplicons

Tester DNA was prepared by adding a second adaptor pair JBam 12 and 24 for *Bam*HI digested DNA or JHind 12 and 24 for *Hind*III digested DNA to 1 µg of the first

round amplicons in the same way as described for the ligation of the first set. The ligation reaction was then diluted to a final volume of 100 µl with water. An aliquot of the ligate was amplified for 20 cycles in a reaction volume of 20 µl to check that the newly ligated adaptors would support amplification with the new primer.

I.III.V Subtractive hybridisation and kinetic enrichment

The hybridisation reaction was set up by mixing 80 µl of the driver (~ 40 µg) with 40 µl tester (~ 0.4 µg) amplicons (a ratio of driver : tester of 100 : 1). To this was added 30 µl of 10 M ammonium acetate and 380 µl 100% ethanol. After incubation on ice for 10 min the DNA precipitate was collected by centrefugation for 10 min at full speed in a microcentrifuge. The pellet was washed twice with 70% ethanol and dried after which it was redissolved in 4 µl 3X EE buffer [30 mM N-(2-hydroxyethyl piperizine)-N'-(3-propene sulfonic acid) (EPPS), pH 8, 3 mM EDTA] followed by repeated vortexing interspersed with centrifugation. The solution was collected at the bottom of the tube and overlaid with light mineral oil so that the spherical droplet could be seen to be completely covered by oil. The DNA was denatured by placing it in a heating block at 98°C for 5 minutes. One µl of 5 M sodium chloride solution was added and the tube briefly centrifuged to mix the aqueous phases. The DNA was then incubated at 67°C overnight. Ten µl TE buffer, pH 8, was added and mixed with the aqueous phase, which was then carefully removed from under the oil by pipetting and added to clean microcentrifuge tube containing 380 µl TE buffer and 40 µg tRNA. The appropriately annealed fragments were amplified in the following reaction:

80 µl 10X *Taq* buffer
64 µl dNTP mix (160 mM of each dNTP)
80 µl of the diluted hybridisation mixture
560 µl ds H_2O
6 µl *Taq* polymerase (30 units)

The tube was placed in a heating block at 72°C for 5 min to fill in the ends. Sixteen µl of the appropriate primer (the 24mer used in the adapter ligated to the tester amplicons) was added, mixed by pipetting, and the mixture aliquoted into eighth PCR tubes. The tubes were placed in a Perkin-Elmer thermocycler, the block of which, were being held at 72°C.

The following thermocycler program was then run: 10 cycles of 95°C for 30 sec; 72°C for 2 min followed by 10 min at 72°C, followed by holding at 4°C. The tubes were combined and 20 µl of the solution subjected to a further 20 cycles of amplification under the same conditions.

Ten µg of tRNA was added to the remainder of the amplified reaction and then extracted with 600 µl phenol/chloroform. An aliqoute (750 µl) was removed and 75 µl (1/10 volume) of 3 M sodium acetate and 825 µl (one volume) propanol added. The precipitate was collected by centrifugation, after incubation on ice for 10 minutes. The pellet was washed twice with 70% ethanol, dried and dissolved in 40 µl TE buffer, pH 8. At this stage the aliquot subjected to the additional amplification cycles was run on a 1.5% TAE agarose gel to check that amplification occurred. This first round difference product (20 µl) was digested with 20 units of mung bean nuclease (Amersham Life Science, UK) at 30°C for 30 minutes to remove all single- stranded DNA. The reaction was stopped by the addition of 160 µl of TE buffer, pH 8. The digested product was amplified in a reaction mixture consisting of:

80 µl 10X *Taq* buffer
64 µl dNTP mix (160 mM of each dNTP)
16 µl of the same primer used for the 10-cycle amplification
80 µl of the diluted nuclease treated mixture
560 µl ds H_2O
6 µl *Taq* polymerase (30 units)

The mixture was aliquoted into eighth PCR tubes in a Perkin-Elmer thermocycler, the block of which was being held at 72°C. The thermocycler program run was 20 cycles: 95°C for 30 seconds; 72°C for 2 minutes followed by 10 minutes at 72°C, followed by holding at 4°C.

The 8 tubes were combined and a 10 µl aliquot was run on a 1.5% TAE agarose gel at 100 V for 30 min to check the amplification. The remainder was extracted with phenol/chloroform. The upper phase was pipetted into a clean tube and a 1/10 volume of 3 M sodium acetate and 1 volume of propanol added. The tubes were

mixed by inversion and placed on ice for 15 min. The precipitate was collected by centrifugation for 15 min at full speed in a microsentrifuge, washed twice with 70% ethanol and dried. The amplicons were redissolved in 80 µl of TE buffer, pH 8. The concentration was estimated by electrophoresis on a 1.5% TAE agarose gel with lambda phage DNA standards and adjusted to 100 µg/ml.

I.IV Cloning of the subtraction products

Subtraction products (2 µg) were digested with the 50 units of the appropriate restriction enzyme for 30 minutes at 37°C. One hundred ng of this digest was mixed with 10 ng of pBluescriptII vector (Strategene, USA) digested with the same enzyme (a ratio of insert to vector of 10 : 1). These were extracted with phenol/chloroform and precipitated with sodium acetate and ethanol. The precipitate was collected by microcentrifugation and the pellet washed twice with 70% ethanol before being dried. The dried pellet was redissolved in 8 µl water. To this was added 1 µl 10X ligase buffer and 1 µl T4 ligase, and the mixture incubated overnight at 16°C. Two µl of the ligate was used to transform competent XL1Blue cells and 50 plasmid-containing colonies carrying an insert were selected and probed with either the 'Barhee' or 'Medjool' labeled driver amplicons using the *Gene Images* random prime-labeling module (Amersham Life Sciences UK). Ten colonies that showed a much stronger signal after hybridization with the 'Barhee' amplicons than with the 'Medjool' amplicons were selected for plasmid isolation and determination of the insert sequence and size.

I.V Southern-blot hybridization

Total genomic DNA of 'Barhee' and 'Medjool' (1 µg) were digested with two units of each of the restriction enzymes *Bam*HI and *Eco*RI at 37°C for two hours. The restricted DNA was then separated on a 1.5% TAE agarose gel using electrophoresis, and transferred to a positively charged nylon membrane (Roche, Switzerland), as described by Sambrook et al. (1989). Cloned fragments of Dp41 and Dp36 were used as probes, which were labeled using the *Gene Images* random prime, labeling module. Membranes were pre-hybridized and hybridized at 65°C in a hybridization buffer containing 5% SSC, 0.1% SDS and 20-fold dilution of the liquid

block provided. Hybridization was carried out overnight and the membranes washed at 60°C using a 1% SSC and 0.1% SDS solution. Detection was preformed using the *Gene Images CDP-Star* detection module (Amersham Life Sciences UK), according to the manufacturer's instructions. Membranes were exposed to Hyperfilm ECL (Amersham Life Sciences UK) and the films developed.

I.VI Sequence analysis

DNA sequence analysis was carried out with the dideoxy chain terminator method developed by Sanger et al. (1997). Recombinant plasmids were sequenced using fluorescent dye terminators and AmliTaq from the ABI PRISM™ Big DYE Terminator Cycle Sequencing Ready Reaction Kit (Perkin-Elmer Applied Biosystems, Inc., Foster City, Calif.), in a cycle sequence protocol according to the recommendations of the manufacturer. Sequence reactions were run on PAGE (polyacrylimide gel electrophoresis) using an ABI PRISM™ 377 Autosequencer (Perkin-Elmer Applied Biosystems). Sequence electropherograms were analysed using Sequence Navigator version 1.0.1 (Perkin-Elmer).

II) Tissue culture and somatic embryogenesis

II.I Tissue culture initiation

The most responsive embryo developmental stages for somatic embryogenesis initiation range from the cleavage polyembryony, to the early (precotyledonary) dominance stage (Percy *et al.*, 2000). For culture initiation, seeds were extracted from developing cones and placed in Petri dishes containing moistened filter paper. Within 1 hour of extraction, seeds were surface sterilized in preparation for dissection. Seeds were surface sterilised by placing them in metal baskets and stirring vigorously for 6 min in 200 ml of 3% (v/v) hydrogen peroxide with one drop of Tween 20. Seeds were then rinsed tree times for 2 min each with sterile water and transferred to a petridish containing a moist, sterile filter paper. The seed coat, nucellus, and membrane surrounding the metagametophyte were removed under a stereomicroscope, and the megagametophyte was placed on initiation medium. Five to eight megagametophytes were cultured in each Petri dish containing approximately 25 ml of medium, and dishes were sealed with parafilm. Explants were cultured in darkness at 24°C for 16 weeks and examined every 4 weeks for embryogenic tissue initiation.

II.II Embryogenic tissue proliferation and culture maintenance

Explants displaying proliferation from the zygotic embryo cells were microscopically selected, then embryogenic tissue was removed, dispersed by vigorous shaking in liquid MLV maintenance medium, poured onto a sterile filter paper disk (Whatman No. 2, 5.5 cm) in a Büchner funnel and a short low-pressure pulse (55 kPa) was applied to drain the medium prior to transferring the disc to fresh medium. This approach was also used for maintenance of embryogenic tissue. Maintenance and proliferation cultures were incubated in darkness at 24°C.

II.III Maturation of somatic embryos

The somatic embryo experiments were performed by combining embryogenic tissue of one line (from several plates), one week after subculture, in a 50 ml test tube,

adding liquid media without growth regulators and vigorously shaking the tube to break up the clumps of tissue into a fine suspention. Subsequently 3 ml containing approximately 0.2 g of suspended embryonal masses were withdrawn with a wide-mouth pipette and placed on a moist filter paper disk (Whatman No. 2, 5.5 cm in diameter) in a Büchner funnel attached to a vacuum pump. A short low-pressure pulse (5 sec, 55 kPa) was applied to drain all liquid medium and anchor the embryonal masses to the filter paper as a thin layer. Each disk of filter paper with the embryonal masses was subsequently placed on maturation medium in a 10 X 20 mm Petri dish and cultured for up to 10 weeks. The cultures were kept under dim light condition at 1.6 µmol $m^{-2}s^{-1}$ from cool white florescence lamps (Philips F72T12/CW, 65 W) under a 16 h photoperiod at 24°C.

II.IV Tissue culture media

II.IV.I MLM 10X stock

	2 Liters	Final concentration in medium [g/L]
NH_4NO_3	16.5 g	0.825
KNO_3	19.0 g	0.95
$MgSO_4.7H_2O$	18.5 g	0.925
KH_2PO_4 (monobasic)	3.4 g	0.17
$CaCl_2.2H_2O$	0.22 g	0.011
LM Minor stock (100X)	200 ml	
LM Vitamin stock (100X)	200 ml	
Myo-Inisitol	2.0 g	0.1
Iron Solution	100 ml	
dH_2O to	2 L	
Stored at -20°C or 2 weeks at +4°C		

II.IV.II MLM Minor Stock 100X

	1 Liter	Final concentration in medium [g/L]
$Na_2MoO_4.2H_2O$	0.125 g	0.00125
KI	0.415 g	0.00415
H_3BO_3	3.1 g	0.031
$MnSO_4.H_2O$ (or $MnSO_4.4H_2O$)	2.1 g (2.7 g)	0.021 (0.027)
$ZnSO_4.7H_2O$	4.3 g	0.043
$CuSO_4.5H_2O$	0.05 g	0.0005
$CoCl_2.6H_2O$	0.013	0.00013
DH_2O to	1 L	
Stored at -20°C		

II.IV.III MLM Vitamin stock 100X

	1 Liter	Final concentration in medium [g/L]
Nicotinic acid	0.05 g	0.0005
Pyridoxine.HCL	0.01 g	0.0001
Thiamine.HCL	0.01 g	0.0001
dH_2O to	1 L	
Stored at -20°C		

II.IV.IV MLM Iron solution

	100 ml	Final concentration in medium [g/L]
Na_2EDTA	0.746 g	0.0373
$FeSO_4.7H_2O$	0.556 g	0.0287
dH_2O to	100 ml	
Made fresh		

II.V Growth regulators

II.V.I 2,4-Dichlorophenoxyacetic Acid (2,4-D) stock (1 mg/ml)

2,4-D	100 mg
95% Ethanol	50 ml
DH_2O to	100 ml
Stored at +4°C	

II.V.II 6-Benzyl-Aminopurine (BA) stock (2.5 mg/ml)

BA	50. mg
Warmed 0.5 N HCL (or NaOH)	5 ml to dissolve
Warmed dH_2O to	100 ml
Stored at +4°C	

II.V.III (±) cis-trans Abscic (ABA) stock (10mM)

ABA	0.02643 g
1 N NaOH	drops to dissolve
dH_2O to	10 ml
Filter sterilized and wrapped in foil. Made fresh. 1 ml stock/ L = 10 μM	

II.V.IV Glutamine stock (25 mg/ml) (Amino acids)

Glutamine	25.0 g
dH_2O to	1 L
pH 5.8. Filter sterilized. Stored frozen	

II.VI MLV Media

II.VI.I MLV Ultra low medium (MLV UL)

		1 Liter
MLM 10X stock		100 ml
Casein Hydrolosate (Casamino Acids)	1 g/L	1.0 g
Sucrose	2 %	20.0 g
2,4-D (1 mg/ml stock)	1.1 μM	0.25 ml
BA (0.5 mg/ml stock)	1.1 μM	0.5 ml
dH_2O to		1 L
Phytagel	0.4 %	4 g
Set pH to 5.7. Autoclaved		
Glutamine (25 mg/ml sterile stock)	0.5 g/L	20 ml

II.VI.II MLV Standard medium (MLV Std)

		1 Liter
MLM 10X stock		100 ml
Casein Hydrolosate (Casamino Acids)	1 g/L	1.0 g
Sucrose	2 %	20.0 g
2,4-D (1 mg/ml stock)	9.5 μM	2 ml
BA (0.5 mg/ml stock)	4.4 μM	2 ml
dH_2O to		1 L
Phytagel	0.4 %	4 g
Set pH to 5.7. Autoclaved		
Glutamine (25 mg/ml sterile stock)	0.5 g/L	20 ml

II.VI.III Maturation medium (MM)

		1 Liter
MLM 10X stock		100 ml
Casein Hydrolosate (Casamino Acids)	1 g/L	1.0 g
Sucrose	6 %	60.0 g
dH_2O to		1 L
Phytagel	1 %	10 g
Set pH to 5.7. Autoclaved		
Glutamine (25 mg/ml sterile stock)	0.5 g/L	20 ml
ABA (10 mM sterile stock)	120 µM	12 ml

III General solutions and buffers

III.I Antibiotics

III.I.I Ampicillin stock solution (50 mg/ml)

Dissolve 2 g Ampicillin (D(-)-a-Aminobenzylpenicillin sodium salt) (Sigma-Aldrich, Germany) powder in 40 ml of sterile distilled water. Aliquote in 1.5 ml tubes and store at -20°C.

III.I.II Tetracycline stock solution (10 mg/ml)

Dissolve 0.1 g of tetracycline powder (Sigma-Aldrich, Germany) in 5 ml 100% ethanol, adjust volume to 10 ml with sterile distilled water. Cover tube with foil and store at -20°C.

III.II Southern blotting solutions

III.II.I Neutralization buffer

Combine 121.1 g of Tris base (Sigma-Aldrich, Germany) and 87.7 g NaCl. Add 900 ml distilled water. Adjust pH to 8.0 with concentrated HCl. Adjust volume with distilled water to 1 L and autoclave.

III.II.II Denaturing buffer

Combine 87.7 g NaCl and 20 g NaOH. Add distilled water to 1 L. Stir to dissolve and autoclave.

III.II.III 20X Standard Saline-Citrate (SSC)

Combine 174.4 g NaCl, 88.3 g $C_6H_5Na_3O_7.2H_2O$ (Sodium-citrate) in 800 ml distilled water. Stir to dissolve and adjust volume to 1 L. Autoclave.

III.III Buffers and salt solutions

III.III.I Agarose gel buffer (TAE) (50X stock)

Combine 968 g Tris base, 228.4 ml glacial acetic acid and 400 ml 0.5 M EDTA, pH 8.0. Add 3 L distilled water and stir until solids dissolve. Adjust volume to 4 L with distilled water. Dilute to 1X before use.

III.III.II Agarose gel loading buffer (6X)

Combine 0.063 g bromophenol blue (BPB), 0.063 g xylene cyanol FF (XC) and 2.5 ml glycerol. Add distilled water to 25 ml. Aliquot into 1.5 ml tubes and store at -20°C. Add 1 µl to sample before loading onto agarose gel.

III.III.III Ethidium Bromide solution

Dissolve 0.2 g ethidium bromide powder in 20 ml distilled water. Vortex to dissolve and cover container with foil.

III.III.IV 5 M Ammonium acetate (NH_4OAc)

Dissolve 385.4 g ammonium acetate in 500 ml distilled water by slowly adding the ammonium acetate powder into the stirring water. Adjust volume to 1 L. Autoclave and store at RT.

III.III.VI 3 M Sodium acetate (NaOAc) pH 4.8

Dissolve 40.82 g $NaOAc.3H_2O$ in 80 ml distilled water and stir on a hot plate until dissolved. Adjust pH to 4.8 with glacial acetic acid. Adjust volume to 100 ml with distilled water. Autoclave and store at RT.

III.III.VII 5 M Sodium chloride

Dissolve 29.22 g NaCl in 80 ml distilled water. Adjust volume to 100 ml, autoclave and store at RT.

III.III.VIII 1 M Tris-HCl pH 8.0

Dissolve 12.11 g Tris base in 80 ml of distilled water. Adjust pH to 8.0 using concentrated HCl. Adjust volume to 100 ml with distilled water. Autoclave and store at RT.

III.III.IX 0.5 M EDTA pH 8.0

Dissolve 8 g NaOH pelles in 400 ml distilled water. Add 93.05 g $Na_2EDTA.2H_2O$ and allow to dissolve. Adjust pH to 8.0 using more NaOH pellets. Adjust volume to 500 ml with distilled water. Autoclave and store at RT.

III.III.X TE buffer (10 mM Tris-HCL, 1 mM EDTA)

Combine 1 ml 1M Tris-HCl pH 8.0 and 0.2 ml 0.5 M EDTA pH 8.0. Adjust volume to 100 ml with distilled water. Autoclave and store at RT.

III.III.XI Low TE buffer (10 mM Tris-HCL, 0.1 mM EDTA)

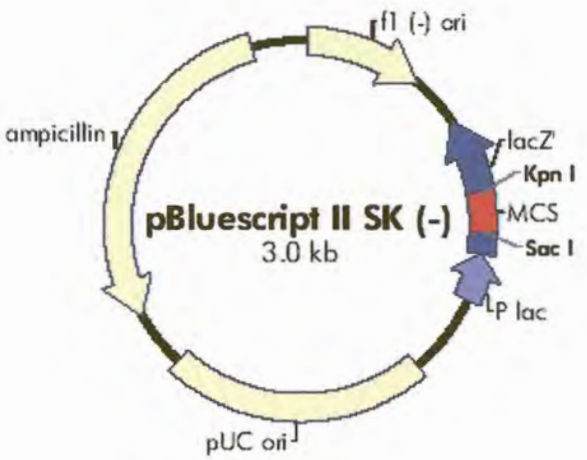
Combine 1 ml 1M Tris-HCl pH 8.0 and 20 µl 0.5 M EDTA pH 8.0. Adjust volume to 100 ml with distilled water. Autoclave and store at RT.

III.III.XII 10 N Sodium hydroxide (NaOH)

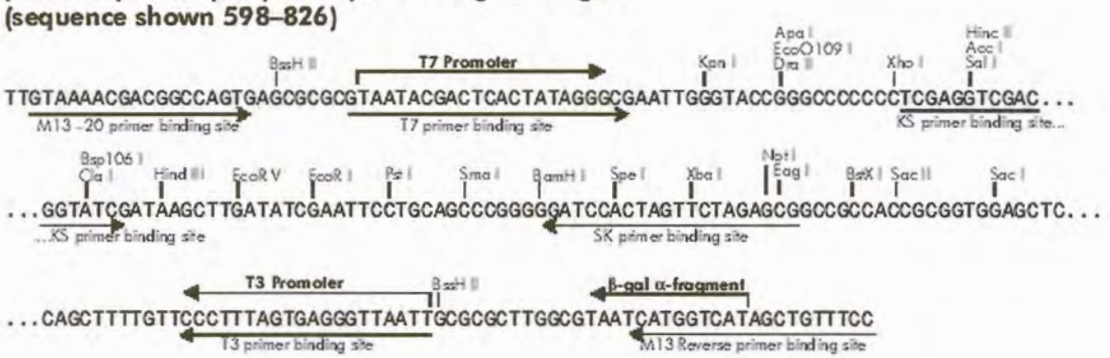
Dissolve 40 g NaOH pellets in 70 ml distilled water. Adjust volume to 100 ml. Store at RT.

IV. Plasmid maps

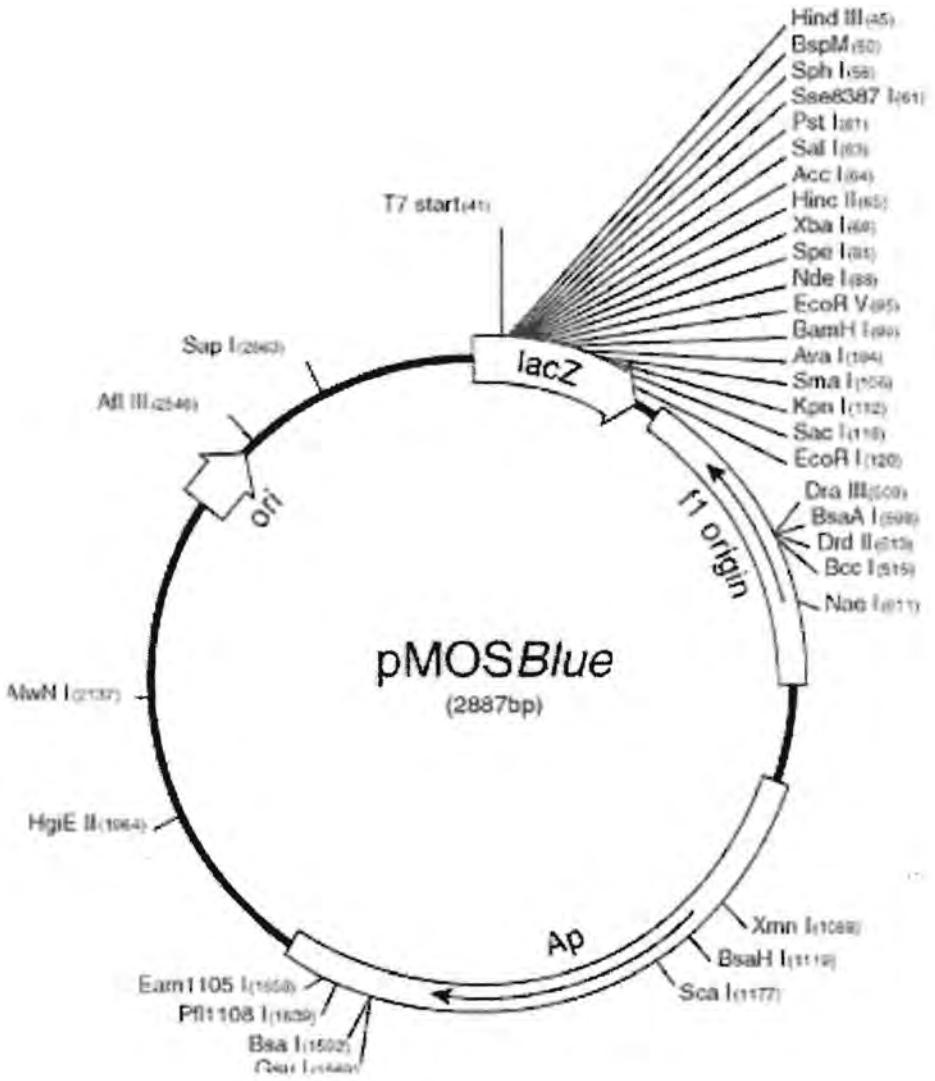
IV.I pBluescript II (Stratagene USA)



pBluescript II SK (+/–) Multiple Cloning Site Region (sequence shown 598–826)



IV.II pMosBlue (Amersham UK)



V. Primer sequences

RAPD and SCAR primers	
OPE-01	5'-CCCAAGGTCC-3'
OPE-06	5'-AAGACCCCTC-3'
DpSL	5'- GTGTTAGGGGCAAAATGGAA–3'
DpSR	5'- TTGTCCGTCTGAGACTCCCT–3'
Date palm subtraction product specific primers	
DP36L	5'-CTATCGACGACAGGCTGACA -3'
DP36R	5'-GACCCGGACTTGTTGGAGTA-3'
DP41L	5'-CCTTCTCCCCGTAGTAACCG-3'
DP41R	5'-AGGAAAGGCAACCTACCGAG-3'
DP50L	5'-TACACGATGTCCCTCAACCA-3'
DP50R	5'-GGAACATTTCCTCGGTATCC-3'
PLM 1	5'-TTACAGAGGGGAAAGGAGGA-3'
PLM 4	5'-GGAAGGAGGTGGCTCCG-3'
PLB11	5'-CGCAATCTTGCAAGTATCAGT-3'
RDA adapters	
R Bam 24	5'-AGCACTCTCCAGCCTCTCACCGAG-3'
R Bam 12	5'-GATCCTCGGTGA-3'
J Bam 24	5'-ACCGACGTCGACTATCCATGAACG-3'
J Bam 12	5'-GATCCGTTCATG-3'
N Bam 24	5'-AGGCAACTGTGCTATCCGAGGGAG-3'
N Bam 12	5'-GATCCTCCCTCG-3'
R Hind 24	5'-AGCACTCTCCAGCCTCTCACCGCA-3'
R Hind 12	5'-AGCTTGCGGTGA-3'
J Hind 24	5'-ACCGACGTCGACTATCCATGAACA-3'
J Hind 12	5'-AGC TTGTTCATG-3'
N Hind 24	5'-AGGCAGCTGTGGTATCGAGGGAGA-3'
N Hind 24	5'-AGCTTCTCCCTC-3'

R Hpa 24	5'-AGCACTCTCCAGCCTCTCACCGAC-3'
R Hpa 11	5'-CGGTCGGTGAG-3'
J Hpa 24	5'-ACCGACGTCGACTATCCATGAAAC-3'
J Hpa 11	5'-CGGTTTCATGG-3'
N Hpa 24	5'-AGGCAACTGTGCTATCCGAGGGAC-3'
N Hpa 11	5'-CGGTCCCTCGG-3'
S Hpa 24	5'-ACTTCTACGGCTGAATTCCGACAC-3'
S Hpa 12	5'-CGGTGTCGGAAT-3'
Sequencing primers	
T7	5'-TAATACGACTCACTATAGGG-3'
Sp6	5'-CATACGATTTAGGTGACACTATAG-3'

UNIVERSITEIT VAN PRETORIA
UNIVERSITY OF PRETORIA
YUNIBESITHI YA PRETORIA

Plant Cell Rep (2002) 21:271–275
DOI 10.1007/s00299-002-0501-9

GENETICS AND GENOMICS

B.J. Vorster · K.J. Kunert · C.A. Cullis

Use of representational difference analysis for the characterization of sequence differences between date palm varieties

Received: 12 December 2001 / Revised: 11 June 2002 / Accepted: 12 June 2002 / Published online: 21 August 2002
© Springer-Verlag 2002

Abstract Representational difference analysis was applied to subtract the genomes of the two date palm varieties, Barhee and Medjool, for identification and characterization of unique genome differences suitable for discriminating between individual plants and the two varieties. Three different DNA difference products were isolated from Barhee representing families of dispersed, repeated variable sequences present in the genome of both varieties. Several variant members of repeated DNA were detected by sequence analysis, containing base changes from C to T and G to A and short deletions. Mutated DNA sequences could be amplified in a polymerase chain reaction-based test from a much smaller number of Barhee plants than from Medjool plants allowing the differentiation between individual plants and partial discrimination between varieties.

Keywords Date palm · Representational difference analysis · Genome analysis · Hyper-variable DNA · Plant variety identification

Introduction

Genomes of closely related plants or varieties can differ by only a few coding genes or in minor genome re-organizations and a range of different approaches is available for detection of such genetic differences. Among these are the analyses of r-DNA intergenic regions (Scribner and Pearce 2000) and of simple sequence repeats (SSRs), which are also known as microsatellites. SSRs have been used for example for identification of varieties of *Chrysanthemum* and *Citrus* (Wolff et al. 1995; Fang and Roose 1997) as well as individual plants of oilseed rape cultivars (Charters et al. 1996). However, the two most widely used genetic techniques to detect plant variation are random amplified polymorphic DNA (RAPD) analysis, which detects DNA polymorphisms amplified by arbitrary primers (Williams et al. 1990; Welsh and McClelland 1990) and amplified fragment length polymorphisms (AFLPs) (Vos et al. 1995; O'Hanlon et al. 2000). Recently, the RAPD technique has also been applied to identify date palm on the variety level (Corniquel and Mercier 1994; Sedra et al. 1998). All these techniques are useful for the differentiation of plants by identifying random polymorphisms. However, the comparisons are usually made on the basis of the presence or absence of a band rather than directly on any DNA sequence variation.

In contrast, representational difference analysis (RDA) allows the cloning and sequencing of fine genome differences between two highly similar genomes and further provides exact sequence information about these differences. RDA detects any kind of labile DNA region in two genomes to be compared and can be used to derive probes for genomic losses, rearrangements, amplifications, point mutations and pathogenic organisms found within any of the genomes compared (Lisitsyn et al. 1993, 1994; Michiels et al. 1998). RDA has been applied in a variety of contexts including the isolation of repetitive sequences present in only one of the compared genomes (Navin 1996). Nekrutenko et al. (2000) used RDA to create a species-specific marker for voles and Toder et al. (2001) have applied RDA in evolutionary genomics to search for overall genome differences between humans and the great apes. RDA has also been used to determine differences between two distantly related oak species where similarities of isolated RDA fragments with known retrotransposons were found (Zoldos et al. 2001). In addition, Donnison et al. (1996)

Communicated by J.C. Register III

C.A. Cullis (✉)
Biology Department, Case Western Reserve University,
10900 Euclid Avenue, Cleveland, OH 4404, USA
e-mail: cac5@po.cwru.edu
Tel.: +1-216-3685362, Fax: +1-216-3684672

B.J. Vorster · K.J. Kunert
Botany Department,
Forestry and Agricultural Biotechnology Institute,
74 Lunnon Road, University of Pretoria, Hillcrest, Pretoria 0002,
South Africa

UNIVERSITEIT VAN PRETORIA
UNIVERSITY OF PRETORIA
YUNIBESITHI YA PRETORIA

applied RDA to identify male-specific restriction fragments in the dioecious plant *Silene latifolia*. RDA has also been used to identify polymorphisms in banana lines that are a result of genomic rearrangements during in vitro propagation resulting in markers useful for the detection of early variation in the initiation of tissue culture plants (Cullis and Kunert 2000). One of the specific advantages of RDA is that subtractions between pooled DNA samples can be performed in order to identify specific polymorphisms that are only present in either a particular individual or a particular variety rather than relying on identification based on a pattern of polymorphic bands.

The aim of this study was to apply the RDA technology to date palm and to investigate if it is a useful technique for generation of markers suitable for identification and characterization of variable regions in the date palm genome. Such variable regions might ultimately be applicable for either individual date palm plant or variety identification/discrimination. In general, there is a need for DNA-based markers in date palm that will facilitate variety identification as well as confirmation of the genetic fidelity of individual propagated plants. The current identification based on morphological characteristics is often difficult. For example, in the Kingdom of Saudi Arabia alone almost 400 date palm cultivars have been classified (Hussain and El-Zeid 1978), based mainly on their fruit characteristics that are expressed in the mature stage of the plant. In this study, plants of the two date palm varieties Barhee and Medjool, which are commonly used in commercial tissue culture for the asexual production of date palm via somatic embryogenesis, have been investigated. A repetitive labile DNA region has been identified allowing the differentiation between individual plants and partial discrimination between varieties.

Materials and methods

Plant material and DNA extraction

The date palm varieties Medjool and Barhee were used as the source of DNA. The in vitro plants used were Medjool derived from explant material collected in California, and Barhee derived from explant material collected in the United Arabic Emirates. Non-tissue culture-derived Medjool and Barhee plant material originated from 14 plants grown in California that have been propagated solely via off-shoots from mother plant material imported in the beginning of the last century from Northern Africa (Nixon 1950). Total cellular DNA was isolated from leaves (1 g) using either the technique outlined by Aitchitt et al. (1993) or the Nucleon Phytopure plant DNA extraction kit (Amersham Life Sciences, UK) according to the manufacturer's instructions. Both techniques gave similar results.

Preparation of RDA amplicons

RDA was performed following the general outline described by Lisitsyn et al. (1993). Two micrograms of each of the DNAs (Barhee and Medjool) were digested with 80 units of the enzyme *Bam*HI or *Hin*dIII. The digests were then extracted with phenol/chloroform, precipitated and resuspended at a concentration of 100 µg/ml. The *Bam*HI and *Hin*dIII digests were ligated to the adaptor pair JBam 12 (5′-GATCCGTTCATG-3′) and JBam 24 (5′-ACCGACGTCGACTATCCATGAACG-3′) or RHind 12 (5′-AGCTTGCGGTGA-3′) and RHind 24 (5′-AGCACTCTCCA-GCCTCTCACCGCA-3′), respectively. The ligation products were amplified by polymerase chain reaction (PCR) by using the primer JBam 24 or RHind 24 to generate the first-round amplicons, followed by digestion with *Bam*HI or *Hin*dIII to remove the adaptors. Tester DNA was prepared by adding a second adaptor pair NBam 12 and 24 (5′-GATCCTCCCTCG-3′ and 5′-AGGCA-ACTGTGCTATCCGAGGGAG-3′) for *Bam*HI-digested DNA or JHind 12 and 24 (5′-AGCTTGTTCATG-3′ and 5′-ACCGAC-GTCGACTATCCATGAACA-3′) for *Hin*dIII-digested DNA to the ends of the first-round amplicons.

Subtractive RDA hybridization, kinetic enrichment and cloning of the difference products

The hybridization reaction was set up using 40 µg driver DNA (Medjool) and 0.4 µg tester DNA (Barhee) (100:1 driver/tester ratio) in a final volume of 4 µl hybridization buffer consisting of 30 m*M* EPPS [(2-hydroxyethyl piperizine)-*N*′-(3-propene sulfonic acid)], pH 8, and 3 m*M* EDTA. The DNA was denatured at 100°C for 10 min, 1 µl of sodium chloride (5 *M*) was added to a final concentration of 1 *M* and the reaction incubated at 67°C for 16 h. The hybridization mix was then diluted and an aliquot amplified using NBam 24 or JHind 24. The first round of amplification was for ten cycles, followed by digestion of the products by mung bean nuclease. The nuclease-treated product was then amplified for an additional 20 cycles. The resulting amplicons, which are called the first difference product, were used in this study. These subtraction products were digested with the appropriate restriction enzyme and ligated into the appropriately digested pBluescript II (Stratagene, USA). The ligation products were transformed into XL1Blue-competent cells and 50 plasmid-containing colonies carrying an insert were selected and probed with either the Barhee- or Medjool-labeled driver amplicons using the Gene Images random prime-labeling module (Amersham Life Sciences). Colonies that showed a much stronger signal after hybridization with the Barhee amplicons than with the Medjool amplicons were selected for plasmid isolation and determination of the insert sequence and size.

Primer design and testing

Pairs of primers were designed using a standard design program (Expasy, Switzerland). The primer pairs were used in a PCR reaction using Barhee and Medjool DNA as template at various annealing temperatures to optimize the PCR reaction. The PCR reactions were carried out in 25-µl volumes containing 25 ng total genomic DNA, 15 ng primer, 100 m*M* of each dNTP, 10 m*M* TRIS-HCl, pH 8.3, 2 m*M* $MgCl_2$ and 0.5 units Taq polymerase (Takara, Japan). Amplification was performed using a Perkin Elmer GeneAmp PCR system 9600 with the following program: (1) 94°C for 5 min×1 cycle; (2) 94°C for 1 min, 65°C or 60°C for 1 min depending on the primer pair, 72°C for 1 min×35 cycles; (3) 72°C for 5 min×1 cycle, and optional soak period at 4°C. The products were separated on a 1.5% agarose gel, stained with ethidium bromide and visualized under UV light.

Southern blot hybridization

Genomic DNA was digested with different restriction enzymes, separated on a 1.5% TAE agarose gel using electrophoresis, and transferred to a positively charged nylon membrane (Roche, Switzerland), as described by Sambrook et al. (1989). Cloned bands were used as probes and were labeled using the Gene Images random prime labeling module (Amersham Life Sciences). Mem-

UNIVERSITEIT VAN PRETORIA
UNIVERSITY OF PRETORIA
YUNIBESITHI YA PRETORIA

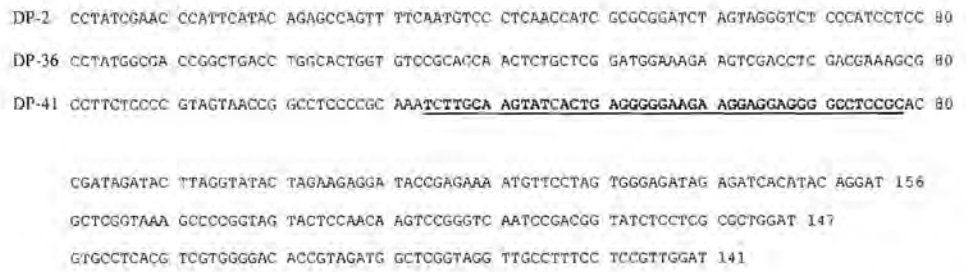


Fig. 1 Sequence analysis of isolated RDA clones DP2, DP36 and DP41 and position of the most variable DNA region of the DP41 clone

branes were pre-hybridized and hybridized at 65°C in a buffer containing 5× SSC, 0.1% SDS and a 20-fold dilution of the liquid block provided overnight and washed at 60°C using a 1× SSC and 0.1% SDS solution. Detection was performed using the Gene Images CDP-Star detection module, according to the manufacturer's instructions. Membranes were exposed to Hyperfilm ECL (Amersham Life Sciences) and the films developed.

Sequence analysis

For plasmid sequencing, PCR products were cloned into the plasmid pMOS*Blue* (Amersham Pharmacia Biotech, UK) and plasmids were recovered from transformed MOS*Blue* cells selected on an appropriate antibiotic. Sequencing was performed using Sequenase (Perkin Elmer, USA) according to the manufacturer's instructions on an automated DNA sequencer (Applied Biosystems, USA).

Table 1 Sequences of primers used to amplify different parts of clone DP41 with DP41R 5′-GCAACCTACCGAGCCATCT-ACGGT-3′ as the right-hand primer and the difference product DP36 with DP36R 5′-GGAACATTTTCTCGGTATCCTC-3′ as the right-hand primer, and a total number of seven non-tissue culture-derived plants amplifying a DNA fragment with the predicted size. *Numbers in parentheses* indicate the total number of plants tested

Primer	Sequence (5′–3′)	Amplification product	
		Barhee	Medjool
DP41L	CCTTCTCCCCGTAGTAACCG	5 (7)	7 (7)
PLM1	TTACAGAGGGGAAAGGAGGA	5 (7)	7 (7)
PLM4	GGAAGGAGGTGGCTCCG	1 (7)	7 (7)
PLB11	CGCAATCTTGCAAGTATCAGT	2 (7)	6 (7)
DP36L	TCGAACCCATTCATACAGAGC	4 (7)	7 (7)

Results

The subtractions between Barhee and Medjool were performed with either Barhee as tester and Medjool as driver or with Medjool as tester and Barhee as driver for amplicons derived from *Bam*HI- or *Hin*dIII-digested DNAs. Following a single round of subtraction using a tester to driver ratio of 1:100 only one of the four subtractions, that with Barhee *BamH*I-digested DNA as tester and Medjool as driver, produced a DNA difference product, which is approximately 150 bp in length. The difference product was cloned and 50 *Escherichia coli* colonies containing the cloned difference product (data not shown) hybridized separately with labeled Barhee tester and Medjool driver amplicons. Both sets of amplicons hybridized to the colonies indicating that the isolated difference product or a closely related sequence was present in both varieties.

Ten *E. coli* colonies with the greatest differential signal between tester and driver hybridization were selected, plasmid DNA isolated and the insert sequence determined. This analysis revealed that the cloned difference product consisted of at least three sequences with lengths of 141 bp (DP41), 147 bp (DP36), and 156 bp (DP2) (Fig. 1) indicating a complex mixture of fragments in the difference product. Seven of the sequenced inserts were identical to DP2 and two identical to DP36. The DP41 sequence contained an AGG motif repeated in tandem 3 times. A search of these three sequences using different databases and DNA sequence analysis tools, such as Blast, FastA and the Smith-Waterman algorithm, resulted in a 75% homology to *Oryza sativa* genomic DNA, chromosome 1 (accession no. AP002902) when a local alignment was done with DP41 but no homology was found for DP2 or DP36.

Since most of the sequenced clones were identical to DP2, we selected only the clones DP36 and DP41, the latter of which was unique, for further characterization. Hybridization of labeled DP41 to *Eco*RI- or *Bam*HI-digested DNAs from tissue culture-derived Medjool and Barhee gave a smear pattern, which is characteristic of a dispersed, repetitive sequence (data not shown). Similar hybridization profiles were obtained when either DP36 or DP2 were used as probes.

Four primers, DP41L and DP41R and DP36L and DP36 R (Table 1), were designed from the DP41 and DP36 sequences. These primer pairs amplified the predicted 120-bp or 110-bp PCR product, respectively, from genomic DNA of six different tissue culture-derived Medjool and Barhee plants but failed to amplify a fragment in two non-tissue culture-derived Medjool and three non-tissue culture-derived Barhee plants (Fig. 2: Table 1). The DP41 amplification product was then characterized in more detail. From each of the six tissue-culture-derived Barhee and Medjool plants, three independent clones of the DP41 amplification product were sequenced (Fig. 3). These 36 clones from different Medjool and Barhee plants revealed a high degree of homology with DP41, the differences being minor base pair changes or single base deletions occurring mainly in a

UNIVERSITEIT VAN PRETORIA
UNIVERSITY OF PRETORIA
YUNIBESITHI YA PRETORIA

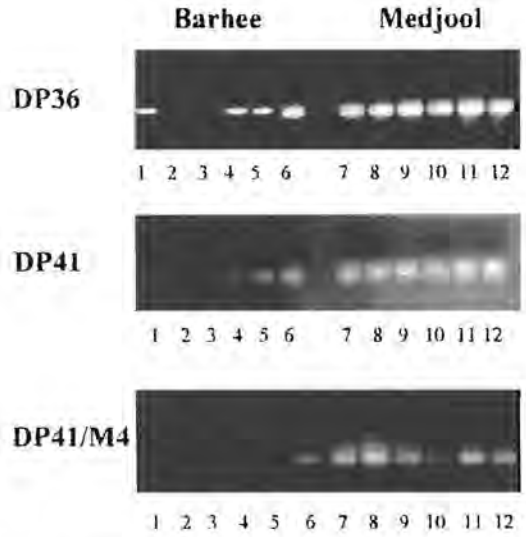


Fig. 2 PCR products of genomic DNA from six out of seven Barhee (B) and Medjool (M) plants amplified with primers for DP36, DP41 and DP41/M4 with the following template DNAs: *lanes 1–6* non-tissue culture-derived B plants; *lanes 7–12* non-tissue culture-derived M plants

```
DP41   TCTTGCAAGTATCACTGAGGGGGAAGAAGGAGGAGGGGCCTCCGC
M1     ----------T--A------A---G----------------T
M2     ----------------------G----------C-------
M3     ----------T--A----------G-----------------
M4     -T---------A-----------G---      T-G------
M5     -G----------T--------T-------------------
M6     ---A-----A-------------------------------
M7     ---------------------------------C-------
M10    ---------------------------------G-------
M11    ----------------------------------AA--
M12    ----------------------------------T--
M13    ----------------------------------T--
M16    ----------TT-----------------------------

DP41   TCTTGCAAGTATCACTGAGGGGGAAGAAGGAGGAGGGGCCTCCGC
B1     ---- -------A-----------------------G------
B2     ----------A----------------------T-G------
B5     ----------A----------------------T-G------
B6     ---------A--A-----A-----------------------
B9     ---A-----A-------------------------------
B10    ----------------T-------------------------
B11    ------CGCA---TTGC-A-TATC--T----------------
B12    ---------------------- ----------------------
B14    ---- ---------------- ---- -- -- ---- -------
B15    ---C---------------- ---------------- -------
```

Fig. 3 Sequence alignment of DP41 with amplified DNA products from genome of M and B plants showing differences to DP41. *Empty space* indicates a base pair deletion. For abbreviations, see Fig. 2

variable 45-bp region of the fragment (Fig. 1). However, in general more changes were observed in the Barhee sequences (specifically single base deletions) than in the Medjool sequences. Only those sequences showing differences in comparison to DP41 are shown in Fig. 3. From the 18 sequences analyzed for each variety, six of the Medjool sequences and eight of the Barhee sequences were identical to DP41. This indicates that the region of the DP41 sequence used to design the primers is common to both genomes. Among the variants found within the different genomes, two, M6 and B9, were identical. Two variants, M1 and M4, the latter with a six base deletion in its sequence, were unique to Medjool and one variant (B11) was unique to Barhee.

Primers PLM1, PLM4 and PLB11 (Table 1) were designed from M1, M4 and B11, to cover the variable portion of these sequences (Fig. 3) when used in conjunction with DP41R. All six tissue culture-derived Medjool plants and also all tissue culture-derived Barhee plants, which originated from a single mother plant, amplified a PCR product with the expected size with all three primers (data not shown). However, primer PLB11 only amplified a PCR product with the expected size of about 110 bp from two of seven non-tissue culture-derived Barhee and six of seven Medjool plants at an optimal annealing temperature of 65°C (Table 1). An identical result was observed with primer PLM1 at an optimal annealing temperature of 60°C (Table 1). Primer PLM4 (at 65°C annealing temperature), which covered a unique 6-bp deletion, amplified a PCR product from DNA of all seven non-tissue culture-derived Medjool plants (Fig. 3) but only from one Barhee plant (Table 1).

Discussion

RDA has been successfully applied in this study for date palm and has resulted in the isolation of three repetitive DNA sequences from the Barhee genome. A subset of each of these families of sequences appears either to be restricted to Barhee, or present at different multiplicity in Barhee, since the reverse subtraction with Medjool as tester yielded no products. The presence of these sequences in the DNA from both varieties could also be explained by methylation differences between the two varieties, with Medjool DNA being more methylated in this region, since *Bam*HI, a methylation-sensitive restriction enzyme, was used to generate the initial amplicons. However, the lack of difference products when *Bam*HI-derived amplicons of genomic Barhee DNA were used as a driver, or when the subtractions were performed with *Hin*dIII-derived amplicons, is consistent with a general high degree of genomic similarity between plants of the two varieties.

One sequence, which is highly variable, has been further characterized with respect to its presence in individual date palm plants of two date palm varieties. Sequence analysis of the amplified difference products from both date palm varieties identified the existence of several variant members of the repeated sequence, consistent with this family representing a variable "genetic hotspot" in the genome (Linacero et al. 2000). Variation included changes from C→T and G→A, deletion of single base pairs (which occurred at a higher frequency in Barhee) and deletion of several base pairs. Base pair changes represent the most commonly observed point mutations in plants, and can also be a consequence of plant tissue culture (Phillips et al. 1994). The extent of variation observed between plants indicates that this might be a rapidly evolv-

UNIVERSITEIT VAN PRETORIA
UNIVERSITY OF PRETORIA
YUNIBESITHI YA PRETORIA

ing/changing sequence. We currently hypothesize that Barhee plants are generally more susceptible to these variations and that these variations have resulted in a heterogeneous Barhee population among non-tissue culture-derived plants including variations in copy number. Since all tissue culture-derived Barhee plants originated from a single Barhee mother plant, conclusions about the behavior of this family of sequences through tissue culture cannot be drawn. However, Pluhar et al. (2001) found an unequal copy number of repeated DNA among callus samples of alfalfa, and speculated that genomic stress induced by tissue culture may have caused that unequal copy number.

In the results reported here, RDA has proved to be useful in identifying a particular repetitive class of sequences that is highly variable in date palms, which is consistent with earlier observations that RDA can be used to isolate families of repetitive sequences (Cullis and Kunert 2000; Nekrutenko et al. 2000; Zoldos et al. 2001). These types of sequences are more difficult to identify with either AFLPs or RAPDs since they either result in many related polymorphisms or generate no size polymorphisms. An additional advantage of RDA is that it can also be performed using bulked amplicons, and can thus be used to identify polymorphisms that are restricted to a particular group of individuals. Therefore, bulking a series of Barhee and Medjool samples and then doing the subtraction will identify variety-specific polymorphisms, rather than individual specific polymorphisms. The types of sequences identified in these experiments would likely be identified as polymorphisms using other techniques, but due to the hyper-variability, each individual would have a unique pattern, or there would be a number of different patterns, none of which would be variety specific. The ideal probes for identifying genomes are those which have a unique location in a given genotype that can be identified directly. Application of RDA therefore offers the opportunity to generate such useful probes.

Acknowledgements This work was funded by a grant of AECI in South Africa to K. J. K. The authors thank Dr E. Bey of Highveld Biological in South Africa for tissue culture material and Dr R. Krueger of the National Clonal Germplasm Repository for Citrus and Dates in Riverside, California for supply of date palm material and NovoMark Technologies LLC in Cleveland, Ohio for technical support to carry out the RDA work

References

Aitchitt M, Ainsworth CC, Thangavelu M (1993) A rapid and efficient method for the extraction of total DNA from mature leaves of the date palm (*Phoenix dactylifera* L.). Plant Mol Biol Rep 11:317–319

Charters YM, Robertson A, Wilkinson MJ, Ramsay G (1996) PCR analysis of oilseed rape cultivars (*Brassica napus* L. ssp. *oleifera*) using 5′-anchored simple sequence repeat (SSR) primers. Theor Appl Genet 92:442–447

Corniquel B, Mercier L (1994) Date palm (*Phoenix dactylifera* L.) cultivar identification by RFLP and RAPD. Plant Sci 101:163–172

Cullis C, Kunert KJ (2000) Isolation of tissue culture-induced polymorphisms in bananas by representational difference analysis. Acta Hortic 530:421–428

Donnison IS, Siroky J, Vyskot B, Saedler H, Grant SR (1996) Isolation of Y chromosome-specific sequences from *Silene latifolia* and mapping of male sex-determining genes using representational difference analysis. Genetics 144:1893–1901

Fang DQ, Roose ML (1997) Identification of closely related citrus cultivars with inter-simple sequence repeat markers. Theor Appl Genet 95:408–417

Hussain F, El-Zeid A (1978) Studies on physical and chemical characteristics of date varieties of Saudi Arabia. Ministry of Agriculture and Water, Saudi Arabia

Linacero R, Alves EF, Vazquez AM (2000) Hot spots of DNA instability revealed through the study of somaclonal variation in rye. Theor Appl Genet 100:506–511

Lisitsyn N, Lisitsyn N, Wigler M (1993) Cloning the differences between two complex genomes. Science 259:946–950

Lisitsyn NA, Segre JA, Kusumi K, Lisitsyn NM, Nadeau JH, Frankel WN, Wigler MH, Lander ES (1994) Direct isolation of polymorphic markers linked to a trait by genetically directed representational difference analysis. Nat Genet 6:57–63

Michiels L, Van Leuven F, Van den Oord JJ, De Wolf-Peeters C, Delabie J (1998) Representational difference analysis using minute quantities of DNA. Nucleic Acids Res 26:3608–3610

Navin A, Prekeris R, Lisitsyn NA, Sonti MM, Grieco DA, Narayanswami S, Lander ES, Simpson EM (1996) Mouse Y-specific repeats isolated by whole chromosome representational difference analysis. Genomics 36:349–353

Nekrutenko A, Makova KD, Baker RJ (2000) Isolation of binary species-specific PCR-based markers and their value for diagnostic applications. Gene 249:47–51

Nixon RW (1950) Imported varieties of dates in the United States. US Department of Agriculture circular no. 834. US Department of Agriculture, Washington, D.C.

O'Hanlon PC, Peakall R, Briese DT (2000) A review of new PCR-based genetic markers and their utility to weed ecology. Weed Res 40:239–254

Phillips RL, Kaeppler SM, Olhoff P (1994) Genetic instability of plant tissue cultures: breakdown of normal controls. Proc Natl Acad Sci USA 91:5222–5226

Pluhar SA, Erickson L, Pauls KP (2001) Effects of tissue culture on a highly repetitive DNA sequence (E180 satellite) in *Medicago sativa*. Plant Cell Tissue Organ Cult 67:195–199

Sambrook J, Fritsch EF, Maniatis T (1989) Molecular cloning: a laboratory manual, 2nd edn. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N.Y.

Scribner KT, Pearce JM (2000) Microsatellites: evolutionary and methodological background and empirical applications at individual, population, and phylogenetic levels. In: Baker A (ed) Molecular methods in ecology. Blackwell, London, pp 235–271

Sedra MH, Lashermes P, Trouslot P, Combes MC, Hamon S (1998) Identification and genetic diversity analysis of date palm (*Phoenix dactylifera* L.) varieties from Morocco using RAPD markers. Euphytica 103:75–82

Toder R, Grutzner F, Haaf T, Bausch E (2001) Species-specific evolution of repeated DNA sequences in great apes. Chromosome Res 9:431–435

Vos P, Hogers R, Bleeker M, Reijans M, Lee T van de, Hornes M, Frijters A, Pot J, Peleman J, Kuiper M, Zabeau M (1995) AFLP: a new technique for DNA fingerprinting. Nucleic Acids Res 23:4407–4414

Welsh J, McClelland M (1990) Fingerprinting genomes using PCR with arbitrary primers. Nucleic Acids Res 18:7213–7218

Williams JGK, Kubelik AR, Livak KJ, Rafalski JA, Tingey SV (1990) DNA polymorphisims amplified by arbitrary primers are useful as genetic markers. Nucleic Acids Res 18:6531–6535

Wolff K, Zietkiewicz E, Hofstra H (1995) Identification of *Chrysanthemum* cultivars and stability of DNA fingerprint patterns. Theor Appl Genet 91:439–447

Zoldos V, Siljak-Yakovlev S, Papes D, Sarr A, Panaud O (2001) Representational difference analysis reveals genomic differences between *Q. robur* and *Q. suber*: implications for the study of genome evolution in the genus *Quercus*. Mol Genet Genom 265:234–41

Copy number the number of copies of a given gene in a set of chromosomes, see multigene family.

Cross-hybridization the binding of a probe to a DNA sequence other than the intended target sequence. This occurs because of homology between the probe and the sequence and because low stringency hybridization wash coinditions are followed.

Cultivar a variety of plant developed through selective breeding programs.

Cytosine (C) pyrimidine base that pairs with guanosine in DNA.

Deletion removal of gene region or base pair from chromosome.

Denaturation for DNA or RNA, describes separation of double-stranded molecule to a single-stranded state, usually by heating; for protein, describes change in physical shape, which is usually renders it inactive.

Digested DNA DNA cleaved by the action of restriction enzymes.

Diploid the normal number of chromosomes (two copies of each – 2n) in virtually all eukaryotes.

Direct repeats multiple identical (or closely related) nucleotide sequences in the same orientation in a DNA molecule.

DNA (deoxyribonucleic acid) the molecular basis of heredity. DNA consists of a polysugar-phosphate backbone from which the bases (nucleotides) project. DNA forms a double helix that is held together by hydrogen bonds between specific pairs of bases (thymine to adenine, guanine to cytosine). Each strand in the double helix is complementary to its partner strand in terms of its base sequence.

DNA ligase enzyme that joins two double-stranded DNAs together, end to end, by catalyzing 3'OH and 5'P termini bond formation.

DNA polymerase an enzyme that catalyzes synthesis of DNA under direction of a single-stranded DNA template.