University of Pretoria etd – Van der Vyver C 2002

Chapter 6: Summary and Perspective

One of the most challenging tasks in plant science is the understanding of the genome changes in plants that occur as a response to stress. There is accumulating evidence that introduction of foreign DNA molecules into the plant genome via an *in vitro* plant tissue culture process might be a stressful event carrying the risk of genome alterations. At the onset of this PhD study the basic assumption was therefore made, which was based on the detection of a genetically modified plant off-type, that the plant genome possesses plasticity and might reorganize in response to a stress induced by a plant tissue culture/gene insertion process. For this study, it was therefore considered important to obtain more detailed information about susceptible DNA regions that may change due to stress. Such regions might well have a hierarchy, in which they vary, and stress-induced changes might occur only in certain labile genome regions ultimately affecting plant performance and also compromise the bio-safety of plants.

As a first new result this PhD study allowed the identification of an unusual phenotype following a plant tissue culture process/gene insertion process using the Agrobacterium method and a tobacco cultivar, which has been previously widely selected for gene insertion approaches. The observed phenotype in a plant expressing a cystatin transgene, namely reduced stem elongation under growth conditions of low light intensities, has so far not been reported. However, this part of the study, which preceded the genetic analysis of genetically modified tobacco plants, did not provide sufficient evidence for the exact cause of the phenotypic change. Beside transgene interaction with metabolic processes, transgene gene insertion at a locus responsible for elongation, genetic variation present in the original inbreds, somaclonal variation due to the plant tissue culture or gene insertion process and changes in DNA methylation could not be ruled out.

University of Pretoria etd – Van der Vyver C 2002

This study therefore accomplished as a second new result the application of the RDA technique to identify and isolate possible putative genome changes from a complex tetraploid tobacco genome. RDA also allowed the detailed characterization of these changes with bio-informatic tools and extending the application of the technique to a further plant species. In general, execution of the RDA technique required only basic laboratory equipment and was relatively inexpensive and resulted in the isolation of three putative altered DNA sequences from tobacco plants derived from a tissue culture/gene insertion process. This study provided therefore first evidence that variability in these regions might be a direct result of a tissue culture/gene insertion process. Two of these variable regions were successfully identified as similar to part of the tobacco chloroplast genome and tobacco ribosomal RNA. In this regard, RDA has proved to be useful in identifying a particular repetitive class of sequences in tobacco. This is consistent with earlier observations that RDA can be used to isolate such families of repetitive sequences. The second chloroplast variable sequence might further confirm possible interchange of chloroplastic DNA with nuclear DNA. However, the possibility still exist that this DNA fragment originated from nuclear DNA that subsequently changed possibly due to a tissue culture/gene insertion process. Future research might therefore focus on the possibility that a plant tissue culture/gene insertion process stimulate the acceleration of chloroplast / nuclear transfer normally considered as a long-term evolutionary process. The third isolated sequence could not be matched to any previously reported genomic sequence and maybe unique to the tobacco genome.

A third new result of this study was the successful identification of adjacent sequence fragments of these variable DNA sequences in the plant genome. The plastid origin of two of the isolated DNA sequences was confirmed by matching the isolated library clones to known DNA sequences. By applying the Tail PCR method, DNA flanking sequences for the third RDA-derived sequence could also be determined. This sequence was partially homologous to sequences of

University of Pretoria etd – Van der Vyver C 2002

ribosomal RNA and to general cloning vectors. However, vector homology could not be linked to an Agrobacterium-derived sequence.

This study could further achieve as a fourth new result to provide first evidence that both the copy number and DNA sequence changes in certain variable regions of the tobacco genome. There is a trend to more genome variation in plants derived from a plant tissue culture/gene insertion process. Unclear is still why this trend was also found in a selection of tobacco cultivars. The question if any stressful event might result in similar genome changes found in this study has to be answered in a further future study.

None of the RDA-derived DNA sequences could be clearly linked in this study to detection of plants either derived from the plant tissue culture/gene insertion process or to plants with a phenotypic change expressing an exogenous cystatin transgene. Failure might be due to the very high degree of genomic identity between plants. Genetically modified plants might differ, if at all, only in a very small portion of the genome. Possible genome differences might be consequently too small, possibly consisting only of point mutations, to be easily detected by RDA. One should be also aware that a subtractive technology, such as RDA, is inherently subject to several sources of bias. The representation of the genome is based on digestion of the genomic DNA with single restriction enzymes. The genomic subset obtained depends, therefore, on the sequence of the restriction site and particularly its GC content. Further, tester/driver ratios used for subtractive hybridization are critical for the elimination of common regions and enrichment of specific sequences. Also, the initial representation is influenced by the size of the restriction enzyme-digested fragments from total genomic DNA, where larger fragments amplify less efficient by PCR than smaller fragments. These factors might ultimately have also accounted for the inability to isolate, for example, the inserted exogenous cystatin coding sequence by the RDA procedure although plants selected for the experiments clearly showed the expression and insertion of the transgene. In addition, a single or very low copy

University of Pretoria etd – Van der Vyver C 2002

number DNA difference sequence might also not have been efficiently amplified and enriched by the RDA protocol applied in this study.

Actions to overcome current failure of a clear identification, if possible at all, of plants from a stressful event, such as plant tissue culture/gene insertion process, might involve the usage of a greater range of different restriction enzymes for genome digestion. This will allow limiting the genome bias and the selectivity in the genome digestion step. By using different restriction enzymes several representations of the same genome can be scanned in each subtraction. So far the RDA technology has been developed only for four different restriction enzymes namely *Hind*III, *Bgl*II, *Bam*HI and *Hpa*II. However, many more restriction enzymes could possibly be used for genomic digestion, if they generate the same staggered ends, thereby allowing the use of already developed adaptors. Further, elimination of repetitive Further, elimination of repetitive DNA sequences with unequal copy numbers, which seemingly are controlled by stresses, in plants might improve the discovery and enrichment of stress-related very unique induced genome changes. Certain restriction enzymes, such as *Mse*I, are known to digest DNA quite frequently in retrotransposons (personal communication, M. van der Merwe) and should be evaluated for their potential in eliminating repeated DNA sequences before the production of RDA representations. A further future approach might also include using bulked amplicons for RDA. Although two different plants were bulked in this study, bulking of a greater number of plants of stressed and non-stressed plants might identify polymorphisms that are restricted to a particular group of individuals. Therefore, bulking a series of samples of the two different types of plants and then executing the subtraction might preferably identify group specific polymorphisms, rather than individual specific polymorphisms.