

A marker-based strategy for the assessment of epigenetic instability in oil palm

Estelle Jaligot¹, Alain Rival¹, Thierry Beulé¹, James Tregear¹ & E. Jean Finnegan²

¹ CIRAD/IRD, PO Box 64501, 34394 Montpellier, France.

² CSIRO Plant Industry, GPO Box 1600, ACT 2601, Canberra, Australia.

e-mail: Estelle.Jaligot@cirad.fr

Abstract

Plant tissues from the mantled somaclonal variant of oil palm have been found to display both large-scale perturbations of genomic methylation levels and sequence-specific DNA methylation pattern changes when compared to their true-to-type counterparts. Concomitantly, several markers showing phenotype-specific transcription patterns have been identified in our group. Some of them are likely to be indicators of the altered physiological state of abnormal cultures, others are possible targets of epigenetic misregulations of flower development. We now aim to establish the exact link between the expression of these candidate marker genes and the mantled somaclonal variation state, through the exploration of their global epigenetic behaviour. This will be achieved using three complementary approaches applied to normal and abnormal tissues: i) the evaluation of transcription levels through quantitative or semi-quantitative techniques (semi-quantitative RT-PCR, real-time PCR); ii) the determination of sequence-wide DNA methylation patterns (bisulfite sequencing) and iii) the identification of the chromatin regulatory environment of these genes (Chromatin Immuno-Precipitation: ChIP). These combined strategies will ultimately help to elucidate the epigenetic mechanisms which give rise to the mantled somaclonal variant. Also, we will be able to identify which of the candidate markers which are indeed consistently linked with the abnormal state. This work will provide the basis for a molecular detection test of the variant phenotype.

Introduction

Evidence has been accumulating over the past decade that epigenetic mechanisms play a key role in regulating plant development. Naturally-occurring as well as induced differentiation/dedifferentiation processes and various environmental stresses have been found to initiate perturbations in the level and distribution of DNA methylation, in chromatin conformation and in gene expression patterns (Finnegan et al., 2000). These same conditions favoring epigenetic instability can occur during micropropagation and regeneration processes through somatic embryogenesis and

often result in disruptions of clonal fidelity, thus causing a major impediment to the development of large-scale production of clonal material (Kaeppler et al. 2000). In oil palm (*Elaeis guineensis* Jacq.), an average of 5% of the somatic embryo-derived adult palms exhibit an aberrant floral structure, with male parts being transformed into carpel-like structures in flowers of both sexes, hence the name “mantled” for this phenotype (Rival and Parveez, 2005). This unstable somaclonal variation results in partial or complete flower sterility and it is in any case detrimental to the production of palm oil. In accordance with the epigenetic nature postulated for this unstable abnormality, both genome-wide and sequence-specific DNA hypomethylation have been demonstrated in variant tissues when compared to their normal counterparts (Jaligot et al, 2000, 2002 and 2004). Since most of this depletion is likely to be a side-effect of the epigenetic events leading to the “mantled” phenotype, it is of paramount importance to focus on the few sequences which putative functions are consistent with a role in the “mantled” phenotype, and showing a strict phenotype-specific expression pattern.

DNA methyltransferases

In plants, methylation patterns are established *de novo* and maintained by several DNA-methyltransferase enzymes belonging to three families (Figure 1).

The *MET1* gene family was isolated thanks to its extended homologies with mammalian sequences (Finnegan et al., 1996), and since then family members have been found in a number of Dicotyledonous and Monocotyledonous species. MET1 has been identified very early as the main maintenance methyltransferase targeting symmetric sites (CG and CNG).

The CMT (chromomethylase) family, characterized by the presence of a chromatin-association domain (or chromodomain) embedded in the catalytic part of the protein, has a role mainly in the long-term inactivation of repeated sequences such as retrotransposons, both through DNA methylation (mainly at CNG sites) and through the recruitment of histone-methyltransferases (Jackson et al., 2002).

The *de novo* methylation activity is controlled by the DRM (domain-rearranged) family, which shows an inversion in the arrangement of motifs in the catalytic domain, compared to the Dnmt3 *de novo* family found in mammals (Finnegan and Kovac, 2000). This activity is required to establish methylation at previously unmethylated sites and to maintain methylation present at asymmetric (CNN) sites throughout mitotic divisions.

In the model plant *Arabidopsis thaliana*, up to 10 DNA-methyltransferase sequences have been identified with one member from each family being predominantly expressed irrespective of tissue and developmental stage (Finnegan et Kovac, 2000).

Moreover, dysfunctions of all three enzyme families have been involved in the formation of aberrant DNA methylation patterns, and, to varying extent, to the emergence of abnormal developmental phenotypes (Finnegan et al., 2000).

The complementarities between the three DNA-methyltransferase activities make them all likely candidates in our search for molecular markers of the “mantled” somaclonal variant, and therefore we undertook the isolation of all three gene families in oil palm. Our aim was to identify the respective role of each METase family in the hypomethylation of genomic DNA which has been measured in “mantled” material (Jaligot et al., 2000). The considerable sequence conservation detected between sequences identified in different plant species was used to implement two complementary homology-based approaches: i) a PCR-based strategy built upon the design of degenerated primers (Rose et al., 1998 and 2003); ii) a screening strategy involving oil palm cDNA libraries and heterologous (*AtMET1* and *PsMET1* cDNA clones) or homologous (*Zmet3*-like oil palm EST) probes. The resulting partial cDNA clones were used in turn to isolate the corresponding full-length coding sequences using the RACE (Rapid Amplification of cDNA Ends) technique. The *EgMET1* coding sequence is approximately 5 kb in length and is mostly similar to rice and maize sequences. The *EgCMT* sequence (ca. 3 kb) displays extensive homologies with rice, maize and barley sequences. The *EgDRM* sequence (2,4 kb) is mostly similar to *DRM*-like sequences found in tobacco, rice and maize, as well as two ubiquitin-associated proteins from *Medicago truncatula*.

In parallel to the isolation of the coding sequences, preliminary expression analyses by semi-quantitative reverse-transcription PCR (sqRT-PCR) have been performed on embryogenic calli (Figure 2). A slight overexpression of both *EgMET1* and *EgCMT* genes was found in Fast-Growing Calli (FGC, generating 100% of “mantled” palms) compared to Nodular Compact Calli (which yield on average 5% of variant palms). This might be considered surprising, since the genome of FGC shows a 4.5% decrease in global methylation rate compared to that of NCC (Jaligot et al., 2000). Conversely, no significant difference was found in the transcription level of *EgDRM* in this material. These results are somehow reminiscent of the cancer paradox, where hypermethylated sequences are embedded in a globally hypomethylated genome and DNA methyltransferase expression is unchanged or increased. This preliminary result will need to be confirmed on different tissues and genotypes.

MADS-box genes

Transcription factors of the MADS-box family have been implicated in the regulation of floral organ formation through what is now called the “ABCDE model”. According to this model, the identity of each whorl of the flower is governed by the expression

of one or more homeotic genes of function A, B or C. Expression of the A-class function alone specifies sepal formation. The combination of A- and B-class functions specifies the development of petals, and the combination of B- and C-class functions results in the formation of stamens. The expression of the C function alone determines the development of carpels. Since its initial conception by Coen and Meyerowitz (1991) this model has been modified to take account of newer data, revealing a D-type activity involved in the specification of ovules (Angenent and Colombo, 1996) and an E function necessary for the determination of the corolla, androecium and gynoecium (Pelaz et al., 2000).

The misregulation of such homeotic genes results in either the transformation or the ectopic formation of floral organs. In particular, mutation of B-type MADS-box genes responsible for petal and stamen formation leads in model plants to a phenotype in which stamens are homeotically transformed to carpel-like structures, reminiscent of the “mantled” phenotype (Jacobsen and Meyerowitz, 1997; Rohde et al., 1999; Jacobsen et al., 2000). The fact that the “mantled” abnormality involves a characteristic homeotic modification of oil palm floral architecture implies that the activity of a specific subset of genes has been altered in somaclonal variant plants, at least within the flower and fruit tissues. Floral homeotic genes of the MADS-box transcription factor family, which might be affected by the chain of events resulting in the “mantled” abnormality, have been identified. Genes of this type are probably expressed only in floral tissues but may well bear differential epigenetic marks (DNA methylation and/or chromatin modifications) at the vegetative stages of growth. At the vegetative stages of development, no distinct morphological criteria are available for screening out abnormal palms, thus the identification of “fingerprint” genes displaying “mantled”-related expression at pre-planting stages would thus be of great benefit by providing us with the means to formulate molecular tests for clonal conformity. Such a test should also allow us to amend the micropropagation protocol while monitoring any environmental factor that could increase the epigenetic vulnerability of the embryogenic cell culture.

MADS-box genes of oil palm were isolated and characterized in our group in order to address their role in the formation of male and female flower structure, and possibly in the onset of the “mantled” phenotype. Thirteen genes belonging to this family have been isolated and their respective expression patterns have been characterized (Adam et al., 2006 and 2007a). We show on [Figure 3](#) the putative functions of several of these genes, with respect to floral organ morphogenesis.

Results from preliminary expression studies by RT-PCR have shown that the oil palm MADS orthologs of the B-type display differential transcription between normal and abnormal tissues (Adam et al., 2007a). Indeed, both *EgDEF1* and *EgGLO2* show a

decrease in mRNA accumulation in mantled flower tissues, compared to their normal counterpart.

Our future projects include a global investigation of the epigenetic behaviour of MADS-box genes in the context of the “mantled” somaclonal variation in oil palm.

1. Studying methylation changes in and around MADS-box genes

As a preliminary step, we will identify the genomic sequences which might play a regulatory role in the expression of these genes. Upstream and downstream regulatory sequences will be identified using RACE (Rapid Amplification of cDNA Ends) and chromosome walking approaches, which will also allow the localization of the introns.

Methylation studies will be performed using two complementary methods: the MS-RFLP method (Methylation Sensitive-Restriction Fragment Length Polymorphism), associated with the Southern blot technique, gives an insight into the methylation status of one specific restriction site, targeted alternatively by a methylation-sensitive enzyme or by its methylation-insensitive isoschizomer (Jaligot et al., 2002 and 2004); on the other hand, the bisulfite sequencing method (Grunau et al., 2001) makes it possible to localize each methylated C and to evaluate the average occurrence of methylation at a given site, irrespective of the sequence context. These investigations will be undertaken simultaneously on both the regulatory and coding sequences of each oil palm MADS-box family member, since the epigenetic regulations can affect these different regions independently (reviewed in Bender, 2004).

2. Elucidating chromatin-based regulation of the MADS-box genes

We will follow two different approaches to monitor changes in the chromatin structure surrounding MADS-box genes in oil palm, in connection with the “mantled” variant phenotype.

The first approach will consist in investigating nuclease accessibility of chromatin around candidate genes. We will obtain a rough estimate of the accessibility by subjecting crude nuclei preparations to a time-course micrococcal nuclease digestion. This rough analysis will be refined using a Chromatin Accessibility Real-Time (ChART) assay, which has been used successfully to characterize changes in chromatin structure around the IL-2 gene (Rao et al., 2001).

The second approach, chromatin immunoprecipitation (ChIP), will be used to study the methylation and acetylation status of histones around target sequences. The “ChIP on chip” technique (Lippman and Martienssen 2004) will provide a complementary overview of large-scale conformational changes underlying the regulation of genes governing flower organ formation.

3. Analyzing the expression of MADS-box genes: relationship with the susceptibility to the “mantled” somaclonal variation

We need a method that shows both sensitivity and accuracy in order to discriminate subtle differences in mRNA accumulation. For that purpose, the use of real-time quantitative PCR (rtQ-PCR) (Heid et al., 1996) is recommended, since it is really efficient for the sensitive quantitation of low gene expression differences, while being considerably more reliable and reproducible than other techniques (Wong and Medrano, 2005).

Small non-coding species of RNA are now considered as widely used mediators in the sequence-specific epigenetic regulation of gene expression (Kanno et al., 2005a and 2005b). We will investigate their role in the transcriptional and post-transcriptional regulation of MADS-box genes. The search for micro-RNAs (miRNAs) will be undertaken by analyzing the overall miRNA content of floral tissues through the new 454 pyrosequencing technology (Kasschau et al., 2007) and identifying the molecules of interest by RNA blotting with MADS-box probe. We plan to follow the production of regulatory miRNAs in the course of flower organ formation, in parallel with the quantitation of MADS-box genes transcripts accumulation. Their involvement in the “mantled” variation will be investigated by comparing their abundance between normal and variant tissues, and also by evaluating the influence of the reversion phenomenon. The specificity of the regulation will be assessed in rice transformants expressing either a mutated copy of the oil palm MADS-box gene, or a mutated micro-RNA.

This ambitious project is aimed at linking directly the phenotype to the changing epigenetic status of one or several sequence(s). This epigenetic status will be evaluated simultaneously at three levels (i.e. DNA methylation, chromatin configuration and RNA accumulation) thereby providing us with an accurate picture of the regulations governing MADS genes expression.

Ultimately, the shifting of the studied material from adult palms to embryogenic *in vitro* cultures will pave the way to the elaboration of a molecular detection test able to discriminate the material prone to generate “mantled” palms as early as possible in the course of the micropropagation process.

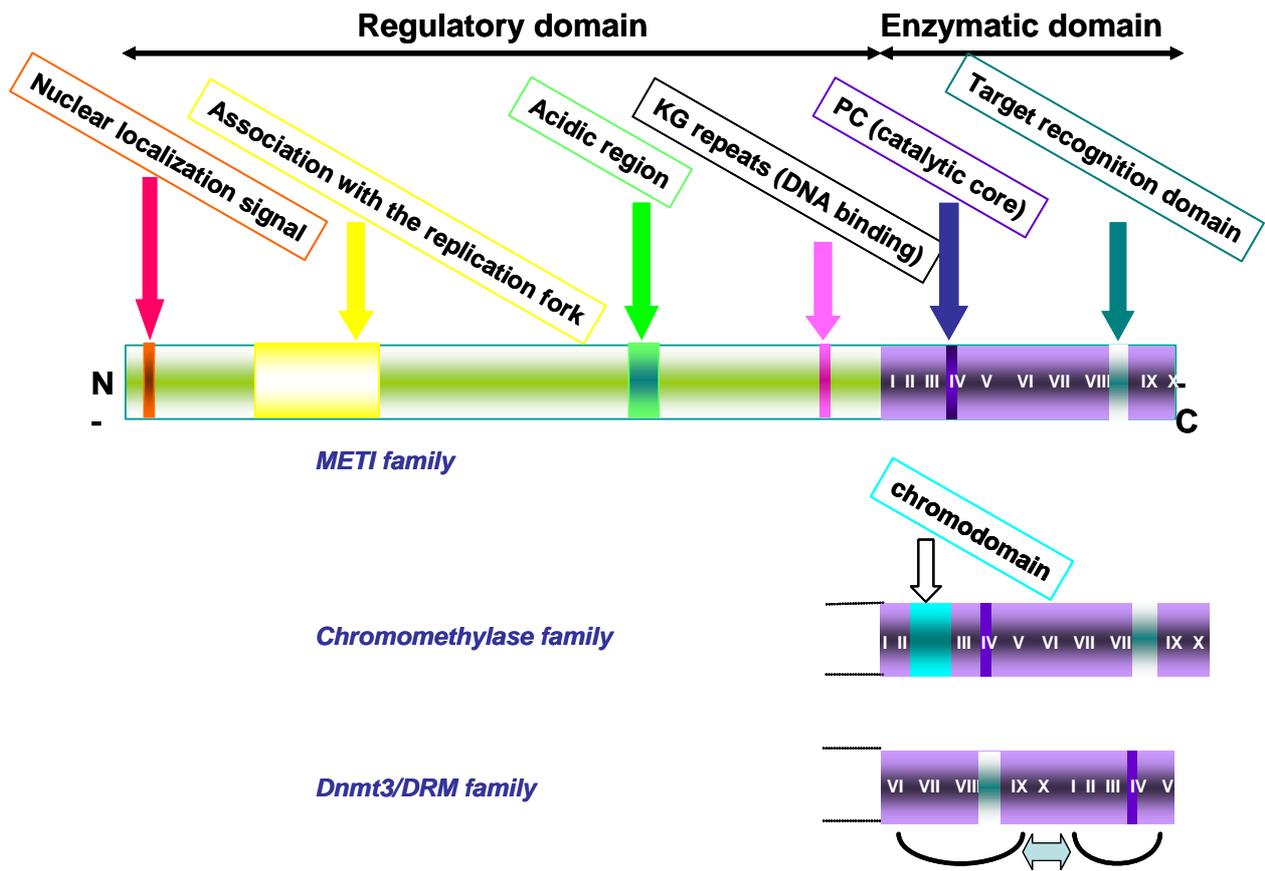


Figure 1: Structure of the three families of plant DNA-methyltransferases. The roman numerals stand for the ten signature motifs of DNA-methyltransferase proteins (adapted from Finnegan and Kovac, 2000).

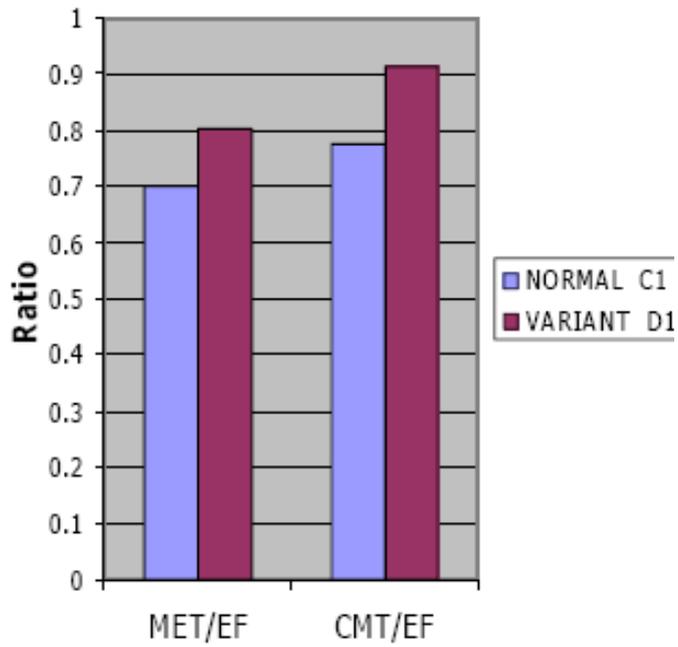


Figure 2: Expression of oil palm *MET1* and *CMT* methyltransferase homologues in embryogenic calli, as estimated by semi-quantitative RT-PCR. C1: Nodular Compact Calli (NCC); D1: Fast-Growing Calli (FGC) (see text for details). EF: protein elongation factor EF1 α (transcription standard).

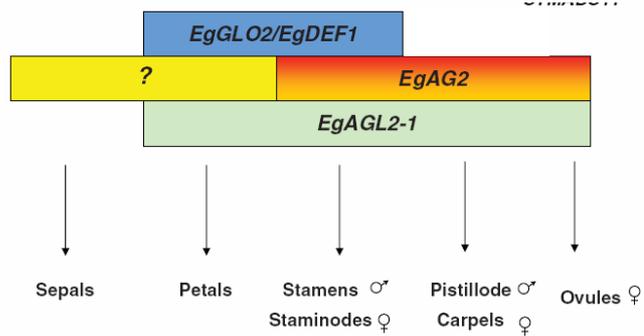


Figure 3: A model to explain the possible roles of various oil palm genes in the determination of flower structure in oil palm, as based upon the generic Eudicot ABC model (Coen and Meyerowitz, 1991; Angenent and Colombo, 1996; Pelaz et al., 2000). Boxes are colour shaded according to homeotic functions as follows: yellow, A function; blue, B function; orange, C and D functions; green, E functions. (From Adam et al., 2007a).

References

- Adam H., Jouannic S., Morcillo F., Richaud F., Duval Y. & Tregear J.W. (2006). MADS-box genes in oil palm (*Elaeis guineensis*): Patterns in the evolution of the *SQUAMOSA*, *DEFICIENS*, *GLOBOSA*, *AGAMOUS*, and *SEPALLATA* subfamilies. *J Mol Evol* 62, 15-31.
- Adam H., Jouannic S., Morcillo F., Verdeil J.-L., Duval Y. & Tregear J.W. (2007a). Determination of Flower Structure in *Elaeis guineensis*: Do palms use the same homeotic genes as other species? *Ann Bot (Lond.)* 1–12.
- Adam H., Jouannic S., Orioux Y., Morcillo F., Richaud F., Duval Y. & Tregear J.W. (2007b). Functional characterization of MADS box genes involved in the determination of oil palm flower structure. *J Exp Bot* 58, 1245-59.
- Angenent G.C. & Colombo L. (1996). Molecular control of ovule development. *Trends in Plant Science* 1, 228–232.
- Bender J. (2004). DNA methylation and epigenetics. *Annual Rev Plant Biol* 55 41-68.
- Coen E.S. & Meyerowitz E.M. (1991). The war of the whorls: genetic interactions controlling flower development. *Nature* 353, 31–37.
- Finnegan E.J. & Kovac K.A. (2000). Plant DNA methyltransferases. *Plant Mol Biol* 43, 189-201.
- Finnegan E.J., Peacock W.J. & Dennis E.S. (1996). Reduced DNA methylation in *Arabidopsis thaliana* results in abnormal plant development. *Proc Natl Acad Sci USA* 93, 8449-8454.
- Finnegan E.J., Peacock W.J. & Dennis E.S. (2000). DNA methylation, a key regulator of plant development and other processes. *Curr Opin Genet Dev* 10, 217-223.
- Grunau C., Clark S. J. & Rosenthal A. (2001). Bisulfite genomic sequencing: systematic investigation of critical experimental parameters. *Nucleic Acids Res*, 29: e65.
- Heid C.A., Stevens J., Livak K.J. & Williams P.M. (1996). Real-time quantitative PCR. *Genome Res* 6, 986-994.
- Jackson J. P., Lindroth A. M., Cao X. F. & Jacobsen S. E. (2002). Control of CpNpG DNA methylation by the KRYPTONITE histone H3 methyltransferase. *Nature* 416, 556-560.
- Jacobsen S.E. & Meyerowitz E.M. (1997). Hypermethylated *SUPERMAN* epigenetic alleles in *Arabidopsis*. *Science* 277, 1100-1103.
- Jacobsen S.E., Sakai H., Finnegan E.J., Cao X. & Meyerowitz E.M. (2000). Ectopic hypermethylation of flower-specific genes in *Arabidopsis*. *Curr Biol* 10, 179-186.
- Jaligot E., Beulé T., Baurens F.-C., Billote N. & Rival A. (2004). Search for methylation-sensitive amplification polymorphisms (MSAPs) associated with the «

mantled » variant phenotype of oil palm (*Elaeis guineensis* Jacq.). Genome 47, 224-228.

Jaligot E., Beulé T. & Rival A. (2002). Methylation-sensitive RFLPs: characterisation of two oil palm markers showing somaclonal variation-associated polymorphisms. Theor Appl Genet 104, 1263-1269.

Jaligot E., Rival A., Beulé T., Dussert S. & Verdeil, J.-L. (2000). Somaclonal variation in oil palm (*Elaeis guineensis* Jacq.): the DNA methylation hypothesis. Plant Cell Rep 19, 684-690.

Kaeppler S.M., Kaeppler H.F. & Rhee Y. (2000). Epigenetic aspects of somaclonal variation in plants. Plant Mol Biol 43, 179-188.

Kanno T., Aufsatz W., Jaligot E., Mette M.F., Matzke M. & Matzke A.J.M. (2005a) A SNF2-like protein facilitates dynamic control of DNA methylation. EMBO Reports 6, 649-655.

Kanno T., Huettel B., Mette M.F., Aufsatz W., Jaligot E., Daxinger L., Kreil D.P., Matzke M. & Matzke A.J.M. (2005b). Atypical RNA polymerase subunits required for RNA-directed DNA methylation. Nature Genetics 37, 761-765.

Kasschau K.D., Fahlgren N., Chapman E.J., Sullivan C.M., Cumbie J.S., Givan S.A. & Carrington J.C. (2007). Genome-wide profiling and analysis of *Arabidopsis* siRNAs. PLoS Biology 5, e57.

Lippman Z. & Martienssen R (2004). The role of RNA interference in heterochromatic silencing. Nature 431, 364-370.

Pelaz S., Ditta G.S., Baumann E., Wisman E., Yanofsky M.F. (2000). B and C floral organ identity functions require *SEPALLATA* MADS-box genes. Nature 405,200–203.

Rao S., Procko E. & Shannon M. F. (2001). Chromatin remodeling, measured by a real-time polymerase chain reaction assay, across the proximal promoter region of the *IL-2* gene. J. Immunol. 167, 4494-4503.

Rival, A. & Parveez M. (2005). *Elaeis guineensis*, Oil Palm. In: Biotechnology of Fruit and Nut Crops. R. Litz (ed.), Biotechnology in Agriculture Series no 29, CABI Publishing, Wallingford, UK, pp: 113-143.

Rohde A., Grunau C., De Beck L., Van Montagu M., Rosenthal A. & Boerjan W. (1999). *Carpel*, a new *Arabidopsis* epi-mutant of the *SUPERMAN* gene: phenotypic analysis and DNA methylation status. Plant Cell Physiol 40, 961-972.

Rose T.M., Henikoff J.G. & Henikoff S. (2003). CODEHOP (CONsensus-DEgenerate Hybrid Oligonucleotide Primer) PCR primer design. Nucleic Acids Research, 31: 3763-3766.

Rose T.M., Schultz E.R., Henikoff J.G., Pietrokovski S., McCallum C.M. & Henikoff, S. (1998). CONsensus-DEgenerate Hybrid Oligonucleotide Primers for amplification of distantly-related sequences. *Nucl Acids Res* 26, 1628-1635.

Wong M.-L. & Medrano J.F. (2005). Real-time PCR for mRNA quantitation. *BioTechniques* 39, 75-85.