3

DNA Methylation and Plant Breeding

Athanasios S. Tsaftaris and Alexios N. Polidoros
Department of Genetics and Plant Breeding
Aristotelian University of Thessaloniki
Thessaloniki 54006, Greece

Plant Breeding Reviews, Volume 18, Edited by Jules Janick
ISBN 0-471-35567-4

I. INTRODUCTION

Frequently, when unexpected phenomena are observed they first are ignored, then timidly explored, and only later published and debated with firm conviction in a more coherent framework. We have now entered the third phase for the unexpected phenomena associated with epigenetic mechanisms operating in organisms, particularly in plants. Several epigenetic mechanisms have been described (Adams and Burdon 1985; Jablonka and Lamb 1995). A primary such epigenetic mechanism involves post-replicative covalent modification of DNA by methylation of cytosine bases. Typically, the modification in plants is methylation of cytosine bases in the dinucleotide CpG and the trinucleotide CpNpG (where N could be any of the four nucleotide bases). Because the methylated sequence is palindromic, both strands of DNA can be methylated. The modification is inherited epigenetically because of the existence of a system that recognizes hemimethylated sequences (with one strand modified) and converts them to the fully methylated state (with both strands modified). The epigenetic state can be reversed by removing the methyl group.

It is a paradox of conventional genetics that two alleles can have the same genetic sequence but show different states of inheritance. A methylated sequence is frequently not expressed while the same sequence is expressed when unmethylated. Epigenetic effects of this nature typically have been characterized by their propagation through mitosis, but can also be perpetuated by meiosis. This implies simply that the necessary conditions (including production of the necessary components) occur in both types of division. It means that the determined structures can be perpetuated through the various structural changes that occur to chromosomes in meiosis. This is relatively straightforward for DNA methylation.

In the past, genetic variation was considered to be due to allelic and epistatic combinations that owed their existence to alterations in the primary nucleotide sequence of the respective genes. Nucleotides modified by methylation result from post-replicative events and were usually not considered to be a part of the primary nucleotide sequence of an individual. However, since DNA methylation occurs at defined target sequences (mainly CpG and CpNpG), and not all target sites are methylated, it represents a potentially important form of polymorphism. In this way epigenetic information systems, like DNA methylation, could produce new alleles, called epialleles, and could generate epigenetic variation that had never been considered as cause of phenotypic variation.

Developmental epigenetic and genetic variation could be important for plants. Most structures of a mature plant are formed after embryogenesis by the reiterative action of meristems. Meristems are collections of undifferentiated cells found at the apices of shoots and roots. By delaying the bulk of growth and differentiation until after embryogenesis, plants can fashion their form to the environment, allowing development to substitute for behavior in some measure. The identity of the structures produced by the apical meristems changes through developmental time. This pattern of plant development allows substantive opportunities for embryogenic cells to vary. Such variant cells can subsequently give rise to gametes. Genetic and epigenetic variation arising during ontogeny can therefore be a potentially important source of transmissible variation.

Plant genomes are generally more methylated compared to other eukaryotic genomes. More than 30% of cytosine bases in some plant genomes are methylated in certain tissues and/or certain developmental stages. This is not only due to extra CpG dinucleotide sequences methylated in plants, but also to methylation of cytosine in the trinucleotide CpNpG sequence that is extensively methylated in plants. Does the plasticity of plant described above permit plants to more extensively explore this kind of epigenetic mechanism, like methylation?

Epigenetic phenomena like DNA methylation are associated with all the important steps of a conventional plant breeding program, e.g., the creation of favorable genetic variation that will form the basis for subsequent selection schemes or used directly as F_1 hybrid seed in hybrid breeding; the selection of superior genotypes through their phenotypes in the field; the multi-site multi-year testing of putative new cultivars and estimation of their adaptation; the preservation and stability or even further improvement of pure line cultivars. This is also true for breeding programs using more modern methods like cell and tissue culture and plant genetic engineering. Variation observed in somaclones, the unexpected silencing through methylation of certain transgenes inserted into plants, and sometimes even the silencing through methylation of endogenous plant sequences homologous to the transgene, have raised serious problems for those wishing to exploit transgenic plants. However, research on these areas has also helped us gain an understanding of epigenetic phenomena involved in the regulation of gene action, allelic and epistatic interactions of genes, variation in plant somatic cells, plant virology etc.

The introductory Section II of this review will summarize some basic knowledge on the extent of plant DNA methylation, methods for its detection and the molecular machinery for imposing new methylation, preserving methylated sites or demethylating DNA. Evidence interconnecting this epigenetic modification with gene activity and the effects of methylation in specific aspects of plant breeding will be examined. In Section III the role of methylation in creating favorable genetic variation, selecting and preserving superior genotypes in conventional breeding programs will be examined. The role of methylation in the specific case of heterosis and F_1 hybrid breeding will be examined separately in Section IV. Finally Section V will be devoted to the

interference of methylation in the expression of transgenes in transgenic plants, examining the role of methylation in gene silencing and the stability of performance of newly developed transgenic cultivars.

The present review has three objectives. The first is to present a comprehensive picture of the occurrence and extent of methylation in plant genomes. The second is to show that this epigenetic system of inheritance interferes, in many different ways with plant gene action and interaction with their allelic and non-allelic counterparts. Understanding gene action and interaction, that is understanding the plant genotype and its expression to the phenotype, and the role of external environmental conditions in this transition, are of paramount importance for plant breeding. This is because the choice and efficiency of the different breeding methods are depending upon these factors. The third objective is then to examine the interference and association of this epigenetic system with specific aspects of plant breeding, both conventional and modern.

The philosophy of this article is reflected in the following statements made by Jablonka and Lamb (1995):

> "Not all changes are the result of Darwinian selection of random variations created by the shuffling of genes and rare chance mutations. The nature of different types of heritable variation is now beginning to receive closer attention, and there is growing realization …that there are non-DNA sequence heritable variations that play a crucial part…".

For some topics or sub-topics of this review there are excellent texts and volumes (Adams and Burdon 1985; Jost and Saluz 1993; Jablonka and Lamb 1995; Meyer 1995; Russo et al. 1996; Chadwick and Gardew 1997), special issues of journals (*Trends in Genetics* 13(8): 293-341, 1997) as well as individual recent reviews. However, this is the first integrated examination of the role and interference of epigenetic phenomena, like DNA methylation, in plant breeding. A glossary of terms used in this review is presented in Table 1.

Table 1. Definition of terms used in the text

Buffering: Resistance to environmental changes leading to improved adaptability, increased stability, and reduced genotype by environment interactions.

Chromatin mark: The non-DNA part of a chromosomal locus that affects the nature and stability of gene expression.

Chromosome domain: A region of chromosome, thought to be a unit of function and replication, containing 30-300 kb of DNA anchored at each end to the nuclear matrix.

CpG island: A CG-rich region of DNA, usually at the 5' end of a gene, that is relatively rich in unmethylated CpG dinucleotides.

CpG site: A DNA site at which C (cytosine) is followed by G (guanine); p denotes the phosphate group, so that C is at the 5' position relative to G.

Competition environment: A dense stand of genetically dissimilar plants where genetically induced plant-to-plant variation reduces yield per unit area and allows density and competitive ability to exert their disturbing effects on response to selection. Contrary to the isolation environment where plants are spaced so widely apart as to exclude any plant-to plant interference with the equal use of growth resources.

Degenerative genes: Genes that in homozygous condition have an effect ranging from degenerative to lethal.

DNA modification: The addition of covalently bonded small chemical groups (e.g., the methyl group $-CH_3$) to specific bases of DNA, e.g., the change of cytosine to 5-methyl cytosine.

Epigenetic inheritance system (EIS): A system that enables the phenotypic expression of the information in a cell or individual to be transmitted to the next generation.

Epimutation: A heritable change in phenotype which is not the result of an altered DNA base sequence.

Genomic imprinting: The process that causes the expression of genetic information to depend on the sex of the parent from which it was inherited; also used for the result of this process.

Germ line: The cell lineage that in normal development gives rise to gametes. Germ-line determination refers to irreversible segregation of the germ cell lineage from somatic lineages. It can take place early or late in development.

Homeostasis: The maintenance of relatively steady states in an organism through internal regulatory mechanisms, despite variations in internal and external conditions.

Isoschizomers: Restriction enzymes that recognize the same target sequence but their cutting is methylation specific.

LINEs: *Long interspersed* repeated elements, found throughout the genome, that result from the movement of *transposable elements*.

Methyl-transferase (MT): The enzyme catalyzing the addition of methyl groups to DNA.

Nuclear matrix: Structural elements in the nucleus that are associated with the organization and activity of chromosomes.

Paramutation: A term coined by Brink for the process whereby one allele in a heterozygote alters the heritable properties of the other allele.

Position Effect Variegation (PEV): The mosaic phenotype resulting from the inactivation in some cells of a gene that has been relocated within the genome.

Positive dominant genes: Genes yielding in hemizygous condition an intact multimeric protein in optimal amount. As a result they are capable of masking the defective action of mutant alleles and of complementing themselves additively. Positive dominant loci represent the integrated form of a gene which through qualitative and quantitative degradation gives rise to semidominant, codominant, and negative dominant alleles.

RIP: A process, first identified in fungi, whereby repeated sequences of DNA above a certain length are either excised through rearrangements (*r*earrangements *i*nduced *pre*-meiotically), or are modified and mutated (repeat-induced *p*oint-mutation).

Somatic mutation/selection: A mutation occurring in a somatic cell. Selection among cells within a tissue or organ is referred as somatic selection.

Stability: Resistance to environmental changes that leads to reduced genotype by environment interactions and increased fitness.

Transgene/transgenic organism: Newly integrated DNA within an organism is a transgene, while the transformed organism is a transgenic organism.

II. DNA METHYLATION IN PLANTS

A. Occurrence

While in most vertebrates less than 10% of all cytosines are methylated, plants have up to 30% 5mC residues (Vanyushin et al. 1970). In addition to the methylation of cytosine in some CpG dinucleotides, the only methylation sites of animals and other eukaryotes, plant genomes contain 5mC in the trinucleotide sequence CpNpG, where N is reportedly any of the four common DNA bases (Gruenbaum et al. 1981). 5-Methylcytosine has also been found in non-symmetrical sequences in transgenes (Meyer et al. 1994), but there is no evidence in plants for the N-4 methylcytosine that is found in some prokaryotes. It has been shown that CpG and CpNpG methylations are carried out by different methyltransferases, and one possible consequence of this is that the two reactions could be independently regulated (Pradhan and Adams 1995). This would allow methylation of CpG and CpNpG sequences to have unrelated functions in the plant cell, although to date there are no data to support this suggestion.

CpG and CpNpG methylation was surveyed in a range of non-vascular and vascular plants to determine when CpNpG methylation evolved and whether the two methylation systems found in higher plants were likely to be under common or separate control. Although both systems exist in many vascular plant taxa, the nonvascular plant taxa appear to contain only CpNpG methylation and this in only limited amounts. The data suggest that both systems may have evolved at the same time and that speciation involved the loss of one or the other methylation systems, or involved the evolution of a stage-specific control system operating during differentiation. Thus the extra methylated CpNpG sequence found only in plants is not a recent acquisition of the plant kingdom (Belanger and Hepburn 1990).

The high proportion of methylated C residues in plants compared to animals could be due to the fact that angiosperm genomes contain a higher proportion of CpG dinucleotides, the common recognition sequence for DNA methylation in plants and animals. The higher rate of CpG dinucleotides in plants compared to vertebrates is probably due to differential degrees of CpG depletion. According to the deamination theory, 5mC residues can undergo deamination to thymine, which leads to point mutations (see also Section III F) and to depletion of CpG dinucleotides and subsequently to an increase of TpG and CpA dinucleotides (Coulondre et al. 1978). Depletion levels are determined by comparing the numbers of observed CpG dinucleotides with the level expected from the base composition. In animals this ratio varies between 0.15% and 0.35% (Josse et al. 1961; Russell et al. 1976), while dicot and monocot genomes vary between 0.68% and 0.79%, respectively (Gardiner-Garden et al. 1992).

The degree of DNA methylation varies among individual plant species (Table 2). *Arabidopsis thaliana* contains only 6.3% cytosine methylation, one-fourth of the methylation level of most other angiosperms (Leutwiler et al. 1984). This difference, however, does not reflect a reduced degree of DNA methylation efficiency, but the low proportion of highly methylated repetitive DNA in arabidopsis. CpG depletion rates in arabidopsis do not differ from those of other angiosperm species, suggesting that methylation of genes occurs to a similar extent in all angiosperms.

In tomato 85% of the CpNpG sites were methylated (Messeguer et al. 1991). To estimate the extent of CpNpG methylation in the tomato genome, nuclear DNA was digested with the methylation-sensitive isoschizomeric enzymes *Bst* NI/*Eco* RII. *Bst* NI and *Eco* RII have the same recognition site [CC*(A/T)GG]. However, *Bst* NI is insensitive to cytosine methylation whereas *Eco* RII is sensitive to methylation at the second C residue (see also Section II B). By comparing the observed fragment sizes for *Eco* RII and *Bst* NI, it was

estimated that 85% of the CpNpG sites are methylated. This value is within the range needed to account for the overall level of C methylation in the tomato genome. *Msp* I and *Hpa* II have the same recognition site (C*CGG). However, *Msp* I is methylation sensitive only when the first C is methylated, whereas *Hpa* II is sensitive to methylation at both the first and second C. Any difference in fragment sizes generated by these two enzymes should thus be due to differences in methylation at the CpG site. Average *Msp* I and *Hpa* II fragment sizes, lead to the estimate that 55% of the CpG sites are methylated and the remaining 45% are unmethylated in tomato. These unmethylated islands often correspond to genes, a phenomenon that has been helpful in genome mapping and in locating the position of candidate genes in clones derived from chromosome walking (Antequera and Bird 1993). It has been shown that unmethylated CpG islands also exist in several plant species including maize, tobacco and wheat. In the case of maize, some islands correspond to genes. Unmethylated CpG islands should result in a bimodal distribution of restriction fragments generated by methylation-sensitive enzymes. The low molecular weight fraction represents unmethylated regions (in which the enzyme cuts) whereas the high molecular weight fraction corresponds to intervening stretches of DNA that are either methylated and/or G+C-deficient (Antequera and Bird 1993). In both maize and tobacco, a bimodal distribution could be observed with *Hpa* II (Messeguer et al. 1991). It was concluded from these analyses that tomato nuclear DNA possesses islands of unmethylated CpG and CpNpG sites, but that not all demethylated sites occur in islands and that the islands are not so pronounced, as in maize. From the values determined for CpG and CpNpG site methylation, calculated and expected values for the overall level of C methylation in the tomato genome were compared. The estimated value, 21%, is close to the value of 23% empirically determined by HPLC (Table 2). Because a lower percentage of CpG sites are methylated than are CpNpG sites (55% vs. 85%) and because C residues are less likely to occur in CpG sites, the majority of methylated C residues in the tomato genome (0.14/0.21=0.66=66%) are expected to exist at CpNpG sites.

Table 2. Percent G+C and %5mC content of nuclear DNA from various plant and animal species.

Species	G+C	5mC/(5mC+C)	Reference
Dicotyledonous plants			
Arabidopsis thaliana	41.4	4.6	Leutwiler et al. 1984
Lycopersicon esculentum (leaves)	37.7	25.0	Messeguer et. al. 1991
Solanum tuberosum (leaves)	38.0	24.6	>>
Nicotiana tabacum (leaves)	40.7	27.8	>>
Pisum sativum	41.9	23.2	Wagner et al. 1981
Solanum alba	39.3	12.2	>>
Vicia faba	39.1	30.5	>>
Gossypium hirsutum	34.2	26.6	Ergle et al. 1961
Helianthus annus	37.6	37.2	>>
Monocotyledonous plants			
Zea mays	46.0	26.7	Ergle et. al. 1961
Zea mays	46.0	27.7	Tsaftaris et al. 1993
Triticum aestivum	48.0	22.4	Wagner et al. 1981
Oryza sativa	42.1	18.6	Messeguer et al. 1991

B. Methods of Detection

The modified base 5mC was initially detected using chromatographic techniques. A number of other methods were then developed for estimation of 5mC content. For gross comparative quantitation of genome methylation, isoschizomeric restriction enzymes that recognize and cut similar nucleotide sequences of DNA, which only differ in cytosine methylation, have been used (Steele-Scott et al. 1984). The isoschizomer pairs *Msp* I/*Hpa* II and *Eco* RII/*Bst* NI, which recognize 4bp and 5bp nucleotide stretches, have already been mentioned. Isoschizomers that recognize larger DNA sequences for more infrequent cutting of DNA are also available (Hayashizaki and Watanabe 1997). Efforts to make this kind of analysis more quantitative have been made by densitometric scanning of electrophoregraphs of DNA after cutting but results have not been encouraging (Steele-Scott et al. 1984). Most highly repetitive DNA is constantly methylated in all plant tissues. Therefore, accurate quantitative methods were required to detect subtle changes of methylation in low copy number DNA in different tissues or in different developmental stages.

For more accurate quantitative determinations of DNA methylation, different types of High Performance Liquid Chromatography (HPLC) as well as mass spectroscopy (Wagner and Capesius 1981; Citti et al. 1983) have been employed for amounts as small as 5-10 •g DNA. Vilpo et al. (1986) employed immunological techniques, generating specific antibodies against 5mC, to measure DNA methylation. Unfortunately, none of the above techniques can provide information about the location of methylated nucleotides in the genome.

Southern blot analysis of DNA digested with isoschizomeric restriction enzymes that have different sensitivities to recognition site methylation, has been used to localize methylation in the genome. This technique allows cytosine methylation associated with specific genes or specific regions of DNA such as repeat sequences, to be identified. However, the procedure is laborious and requires a specific probe. Also, Southern analysis does not always identify DNA methylation mutants or polymorphisms because it sometimes cannot discriminate between cytosine methylation at a restriction site and loss of the site due to nucleotide mutation. With the development of PCR novel techniques have been developed to identify DNA methylation and to map methylation polymorphisms. One such technique is based upon the coupled restriction enzyme digestion and random amplification (CRED-RA) of genomic DNA (Cai et al. 1996). Random amplification of genomic DNA by PCR with arbitrary 10-mer oligonucleotide primers is widely used to generate random amplification polymorphic DNA (RAPD) markers for fingerprinting or genetic mapping (Williams et al. 1990). The CRED-RA technique is based on the following hypothesis: a DNA fragment cannot be amplified if it contains a specific restriction site in the region between two primer binding sites and that site is cut by restriction enzyme digestion prior to PCR. If DNA methylation of the restriction site prevents digestion within the genomic fragment, the fragment can be amplified. However the amplified product will then be susceptible to cleavage because the restriction site will not be methylated during DNA amplification. Thus DNA methylation can be identified by comparing the banding patterns of template DNA amplified without restriction, template DNA amplified after restriction and product DNA restricted after amplification. The technique has been used by Bedford and Van Helden (1990) to detect allelic differences in methylation and by Tsaftaris et al. (1997, 1998) to study the variation in patterns of DNA methylation among maize inbreds, and between maize inbreds and hybrids from plants grown in different growth conditions (see also Section IV).

A third method combining PCR with sequencing uses bisulfite treatment to modify cytosine to uracil residues in the DNA. All cytosines are converted to uracil, except those that are methylated, which are resistant to modification and remain as methyl cytosine (Church and Gilbert 1984). Each altered DNA sample must then be amplified, cloned and sequenced. The main disadvantage of this technique is that it is technically difficult and labor intensive, but for a single structural gene and its upstream few thousands bp of regulatory sequences, it can provide a complete map of methylated sites in different tissues and developmental stages. A recent modification of this procedure takes advantage of the bisulfite-mediated chemical conversion of cytosine to uracil, followed by PCR using primers designed to distinguish methylated from unmethylated DNA. The main advantage is that it avoids the use of restriction enzymes and resulting problems associated with incomplete digestion.

For fast analysis of the methylation state of thousands of genes simultaneously, a powerful new technique called Restriction Landmark Genomic Scanning (RLGS) has been recently developed (Hayashizaki and Watanabe 1997). RLGS is a multiplex method that allows simultaneous analysis of more that 3,000 loci. It employs the *Not* I restriction enzyme because its restriction site makes good landmarks for genetic analysis. *Not* I cuts neither GCGG5mCCGC nor GCGGC5mCGC, but cuts GCGGCC5mC and it is blocked by CpG methylation. The technique uses high-resolution, 2-D electrophoresis to visualize radioactive DNA fragments produced by restriction digestion.

In conclusion, a repertoire of methods is available today for bulk estimates or accurate quantitative determination of 5mC in plant DNA and its distribution in individual genes and their regulatory sequences or even its distribution in the whole plant genome.

C. Mechanisms of Maintenance, *De novo* Methylation and Demethylation

Non-random distribution of methylcytosine within the genome varies depending on tissues or developmental state. For example in carrot, a differing content of 5mC was observed among different tissues (Bernacchia et al. 1998). In tomato, the level of methylation of mature tissues was significantly higher than that of immature ones and protoplasts (Messeguer et al. 1991). Moreover, a reversible variation in the methylation pattern was observed during the process of carrot somatic embryogenesis. Thus, as in mammals (Razin and Kafri 1994), a cycle of demethylation and *de novo* methylation appears to take place during plant development and differentiation.

The regulation of programmed changes in methylation patterns during development and differentiation involves highly ordered processes of *de novo* methylation, maintenance methylation and demethylation. The key enzymes involved in the first two processes, the DNA methyltransferases (MTases), catalyze the transfer of the methyl group from the donor S-adenosyl-methionine to the C5 position of cytosine residues within specific DNA sequence contexts. Several eukaryotic MTase genes have been cloned so far. The corresponding enzymes contain the ten conserved amino acid motifs characteristic of the catalytic domain of MTases (Lauster et al.

1989; Kumar et al. 1994). They differ from prokaryotic MTases, however, by the presence of a large N-terminal domain. In contrast to prokaryotic type II MTases, which do not distinguish between unmethylated and hemimethylated substrates eukaryotic MTases strongly prefer hemimethylated DNA as a substrate. Interestingly, the removal of the approximately 1,000 amino acid-long N-terminal domain by proteolytic cleavage stimulates the *de novo* activity of the murine MTase in vitro (Bestor 1992), suggesting that, besides a maintenance methylation activity, this enzyme might also be responsible for some in vivo, *de novo* methylation. At least part of the *de novo* methylation activity in mammalian cells, however, should be catalyzed by a different MTase encoded by a second gene. Evidence for this comes from the analysis of mouse embryonic stem cells homozygous for the disruption of the identified *Dnmt* gene. Despite the lack of DNMT activity, these cells still show a low amount of *de novo* methylation upon differentiation (Lei et al. 1996). Such a second type of MTase, responsible for *de novo* methylation, has been identified in filamentous ascomycetes. In *Ascobolus*, DNA repeats are heavily methylated at cytosines. This *de novo* methylation takes place during sexual reproduction, between fertilization and karyogamy and requires pairing between the homologous repeated DNA sequences. This process has been termed 'methylation induced premeiotically' (MIP) (Rhounim et al. 1992). Once established, the methylation patterns are perpetuated in dividing cells by a maintenance process. Both elements of a duplication are methylated by MIP and, as a result, the genes carried by the duplication are silenced. Towards and understanding of the genetic control and the biological function of *de novo* methylation processes, the cloning and characterization of an *Ascobolus* gene, *masc 1*, encoding a protein with the primary structure of a MTase has been reported (Malagnac et al. 1997). The disruption of *masc1* has no effect on maintenance methylation but prevents *de novo* methylation associated with MIP. Therefore, Masc1, which is involved in *de novo* methylation, is likely to represent a novel type of a eukaryotic MTase.

In plants, MTases have been purified from wheat germ (Theiss et al. 1987), pea shoots (Yesufu et al. 1991), cultured rice cells (Giordano et al. 1991), arabidopsis (Finnegan and Dennis 1993) and carrot (Bernacchia et al. 1998). The molecular masses of wheat and rice enzymes were shown to be 50-55 kDa. In contrast, the molecular mass of purified pea MTase was shown to be 160 kDa; however, the enzyme was unstable and yielded proteolytic products that retained enzymatic activity. Furthermore, two distinct MTase fractions extracted from nuclei of young pea shoots are responsible for the differing specificity of methylation (CpG vs. CpNpG) (Pradhan and Adams 1995). In arabidopsis, a PCR approach using degenerate oligonucleotides (synthesized on the basis of motifs conserved between prokaryotic and mammalian MTases) led to the isolation of a series of overlapping cDNAs that code for a protein of 1,534 amino acids; the C-terminal region of the deduced amino acid sequence shows 50% conservation with the corresponding part of the murine enzyme but only 24% identity was observed for the N-terminal part (Finnegan and Dennis 1993). Sequence divergence between the PCR product and the isolated cDNA and results of Southern analysis suggested that MTase may be encoded by a gene family in arabidopsis.

In a recent study (Bernacchia et al. 1998) isolation and characterization of two distinct cDNAs that code for carrot MTase were reported. The screening of a cDNA library has led to the isolation of clones that belong to two distinct classes of genes (*Met1* and *Met2*) that differ in sequence and size. Amino acid sequences derived from *Met1* and *Met2* were more than 85% identical for most of the polypeptide but completely diverged at the N-terminus. The larger size of *Met2* cDNA was due to the presence of a nearly perfect fivefold repeat of a 171 bp sequence present only once in the *Met1* cDNA. Northern and in situ hybridization analysis of young carrot plants and somatic embryos indicated that both genes were maximally expressed in proliferating cells (suspension cells, meristems and leaf primordia) but differed quantitatively and spatially in their mode of expression. Polyclonal antibodies were raised in rabbit using fusion proteins corresponding to the regulatory and catalytic regions. In nuclear carrot extracts, both antibodies were found to recognize a band of about 200 kDa along with some additional bands of lower size. These results provided the first direct demonstration that MTases of a higher eukaryote are encoded by a gene family. But it still remains unclear whether MTases coded by this gene family have substrate specificity for both CpG and CpNpG methylation or whether different methyl-transferase enzymes exist in plants.

The process of demethylation could be either passive or active. In the first case no specific mechanisms and proteins are required. In dynamic fashion, DNA methylated in both strands will be hemimethylated just after the first cycle of duplication. The unmethylated DNA chain will lose its methylation signals (existing methyl groups in its complementary DNA chain) if the hemimethylated DNA proceeds to a second round of duplication prior to the completion of methylation of both strands by maintenance MTases. Mechanisms of maintenance of DNA methylation patterns during DNA duplication, and of passive demethylation are depicted in Fig. 1.

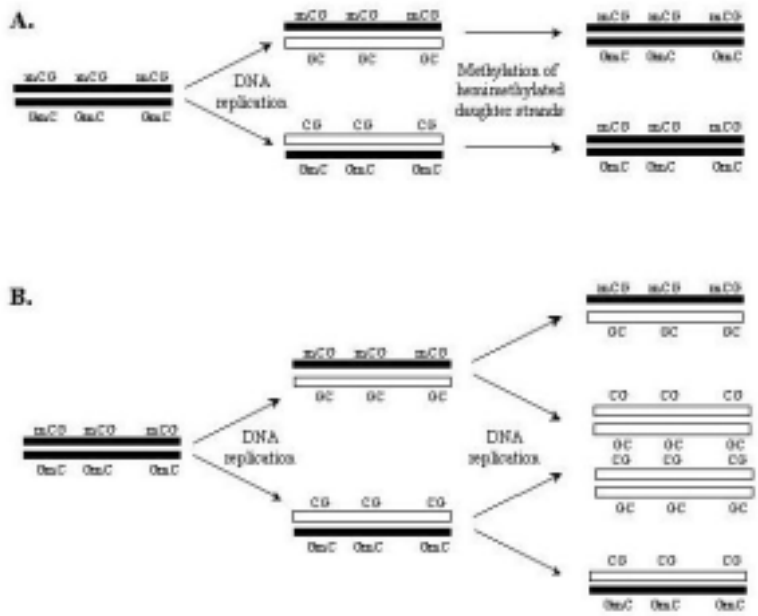

Fig. 1. Possible mechanisms for the inheritance of (A): methylation patterns and (B): passive demethylation.

Thus, in actively dividing cells (apical meristems, cells in suspension cultures or callus) if the maintenance MTase lags behind the rate of DNA duplication for different reasons (availability of S-Adenosyl Methione, presence of inhibitors etc.), previously methylated genes could under these conditions become unmethylated in certain progeny cells. This could lead to activation of silent genes and/or to their higher expression due to partial demethylation of their regulatory and/or coding sequences.

In active demethylation the participation of regulatory proteins that demethylate specific methylated DNA sequences is anticipated. However, little information is available for such a mechanism. A protein called TnpA, coded by one of the four open reading frames of the *Spm* transposable elements in maize, has been identified to be implicated in such a mechanism (Schlappi et al. 1993; Fedoroff 1995). Like other maize transposable elements, active *Spm* elements promote the reactivation of inactive elements. The element-encoded protein responsible for element reactivation was identified by introducing each of the element's four potential coding sequences into transgenic tobacco plants with a methylated, inactive *Spm* element. One of the most interesting aspects of TnpA's interaction with methylated *Spm* promoter was its ability to promote demethylation of sequences outside its binding sites, suggesting that the mechanism by which it influences promoter demethylation is not simple competition with MTase for binding to its recognition site. Only 12 of the more than 100 methylatable C residues on each DNA strand are within TnpA binding sites. The interaction between TnpA and the methylated promoter results in decreased methylation of the entire region. This may be attributable to the ability of TnpA to maintain DNA in a configuration that is less favorable or accessible to maintenance of methylation during DNA replication. Or its effect may be mediated by components of the transcription machinery. That is, TnpA binding may promote binding of other transcription factors, which in turn decreases the probability of remethylation. Whatever its mechanism of action, it is clear that TnpA action supports the existence of mechanisms of active demethylation of individual methylated sequences (Raina et al. 1998).

In conclusion, extensive information has been obtained in the last few years, especially after isolation and cloning of several MTase genes from plants, for the key mechanisms operating in DNA methylation, namely: *de novo* methylation, methylation maintenance, and demethylation. However, much remains to be learned about their control and their operation during development.

D. DNA Methylation and Gene Action.

Several lines of evidence suggest that DNA methylation in eukaryotes plays a role in gene expression. Studies of numerous tissue-specific genes using different techniques have shown a clear correlation between the methylation status of active and inactive genes. Thus, most genes are undermethylated in the tissues in which they are expressed, while they are heavily methylated in non-expressing tissues (Jost and Saluz 1993). These data suggest that changes in the methylation pattern during differentiation may modulate gene activity. In plants induction of several endogenous genes in certain tissues has been linked to loss of cytosine methylation (Spena et al. 1983; Bianchi and Viotti 1988; Ngernprasirtsiri et al. 1989), correlating hypomethylation of these genes with transcription in the respective tissues.

Although such examples suggest a correlation between gene repression and DNA methylation, other reports can be found in the literature that do not detect any changes in DNA methylation patterns although gene activity is altered. Certainly not all changes in gene activity are based on regulation of gene transcription by DNA methylation. It is most likely that DNA methylation is mainly involved in the regulation of promoter activities, but not in post-transcriptional regulation. If changes in gene activity are due to post-transcriptional regulation, promoter activity would probably not be impaired and no significant changes in DNA methylation should be detectable. Even for transcriptional control it is difficult to exclude the involvement of DNA methylation in changing promoter activity, because most DNA methylation studies have limited accuracy as they frequently use isoschizomers. Very often DNA methylation within a genomic region can be accurately monitored by measuring the methylation sensitivity of a restriction site located within this region. However, the state of DNA methylation at a restriction site might not always correspond to the degree of methylation of a neighboring sequence that is involved in promoter regulation. For example, genomic sequencing analysis provides a precise tool, as the methylation state of every C residue can be analyzed. Genomic sequencing of a 900-bp region upstream from the translation start codon of the maize alcohol dehydrogenase gene did not reveal any cytosine methylation although the gene was silenced (Nick et al. 1986). Apparently the *Adh1* gene provides an example where gene activity is not regulated by DNA methylation. It cannot be excluded, however, that changes in DNA methylation further upstream of the promoter region may have an influence on repression of the gene. For example, the cell-specific transcription of the *PEPCase* gene of a C_4 plant like maize corresponds to demethylation of a region located 3.3 kb upstream of the gene (Langdale et al. 1991).

In vitro DNA methylation of a few specific gene sequences inhibited the activity of these genes when inserted into animal cells in vivo (see Adams and Burdon 1985 for a review and references thereafter). In plants Weber et al. (1990) showed that in vitro hemimethylation of the CaMV 35S promoter inhibited transient gene expression of reporter genes after transfection into protoplasts; the methylated state was maintained and inherited during regeneration of plants and correlated with inhibition of transgene expression.

Evidence for the involvement of DNA methylation in preventing the expression of certain plant genes has been provided by studying the expression of genes within plant transposable elements. Transposable elements of plants, (that could be considered, in a way, special cases of transgenes and the host plants harboring a transposable element as a special type of transgenic plant) were known for many years. The pioneering work of Barbara McClintock (McClintock 1957, 1958) and Peter A. Peterson (Peterson 1965, 1986) with maize revealed that transposons undergo reversible heritable inactivation. Early in the study of the suppresor-mutator (*Spm*) transposable element (identical with Peterson's *En* element), McClintock recognized that certain isolates of the element either cycled between inactive and active phases during development or underwent an inactivation event of longer duration and sufficient stability to be heritable, but which was nonetheless occasionally reversed (McClintock 1957, 1958). In subsequent studies, she developed a deeper understanding of the ways in which the *Spm* element alternated between active and inactive phases (McClintock 1959). She later reported that the *Activator* element was also subject to a similar type of reversible inactivation, although the *Activator* element's inactivation mechanism was not analyzed in detail (McClintock 1964, 1965).

While an inactive *Spm* can neither transpose nor *trans*-activate expression of a *dSpm* element, it is distinguishable from a *dSpm* in two ways: it can return to an active form spontaneously and it can be transiently reactivated by an active element (McClintock 1971). Inactive elements differ considerably in the stability of the inactive state, as well as the developmental pattern of element reactivation. In some cases the element alternates between active and inactive phases frequently during plant development, giving a plant that shows a fine mosaicism for element expression. This type of element has been termed a cycling *Spm* (*Smp-c*). Other elements exhibit heritable developmental patterns of inactivation and reactivation in which the element is active only in certain plant parts or only in certain areas of a given tissue. Still other elements remain largely inactive throughout development, returning to the active phase infrequently and late in development.

Changes in methylation of C residues along the length of the *Spm* element have been investigated using restriction endonucleases (Banks et al. 1988; Banks and Fedoroff 1989). Active and inactive elements alike are methylated throughout most of their sequence, although methylation was not observed to extend into neighboring sequences at the locus studied. Inactive elements could be distinguished from active elements by methylation of sites in the upstream control region (UCR) of the single transcription unit identified in this sequence. Cryptic *Spm* elements could be distinguished from less stable inactive elements by the extent of

methylation in the GC-rich downstream control region (DCR) present in the first untranslated exon of the gene (Fig. 2). These observations imply that methylation of the UCR is correlated with element inactivity, while increasing methylation of sites in the DCR is correlated with the progressive increase in the heritability of the inactive state. This phenomenon is of particular importance, if verified for regulatory and coding parts of other plant genes (see Jacobsen and Meyerowitz (1997) as an example for a gene controlling flower size in arabidopsis). As mentioned earlier, analysis of the *Spm* transposable element and its gene regulation led to the isolation of a novel epigenetic activator protein, TnpA. TnpA interacts specifically with the *Spm* promoter and promotes demethylation of neighboring specific DNA sequences outside its binding sites.

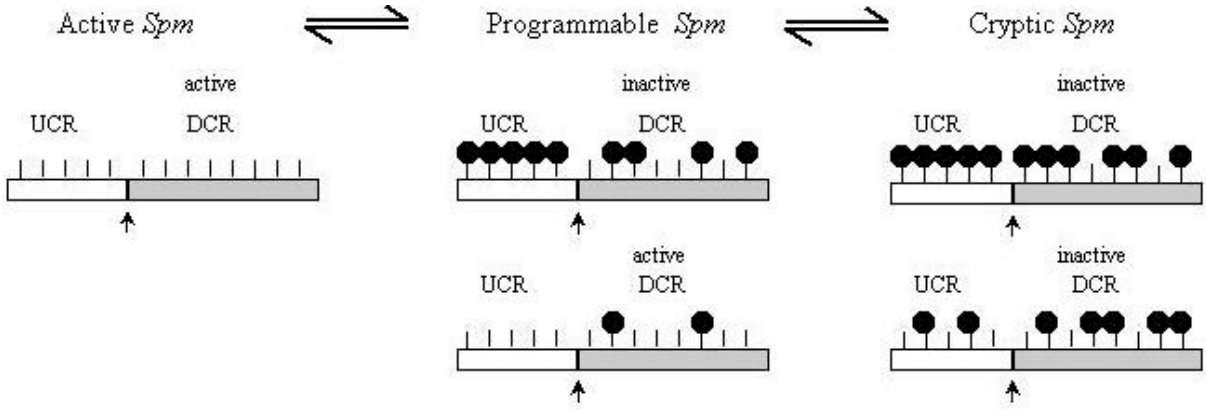

Fig. 2. Fedoroff's model of the different states of the *Spm* element. UCR: upstream control region; DCR: downstream control region; solid lines: methylatable sites, unmethylated; black dots: methyl-groups, methylated sites. The active element autoregulates its own activity. The programmable element is transcribed when an active element provides a gene product, which reduces methylation and activates transcription. The same product also reduces the level of methylation of cryptic elements, but not enough to lead to transcription.

Similar results have been obtained by studying the *Ac* transposable element in maize. *Ac* elements also cycle between active and inactive phases during plant and kernel development (McClintock 1963). As Bennetzen (1987) has shown, these changes in phase were associated with *Ac* methylation. The initial studies of this phase-change were conducted with the *Ac* element in the *waxy* locus. When *Ac* at *wx-m7* is inactive, the *Ac* element can no longer autonomously transpose, *trans*-activate a *Ds* element, or contribute to overall *Ac* dosage effects. When *Ac* activity is regained after a shift in phase, each of these functions is restored. Comparison of active *Ac* elements at several loci and the inactive *Ac* at *wx-m7* by Southern blot analysis revealed that the inactive *Ac* sequence was not susceptible to digestion by the methylation sensitive enzyme *Pvu* II while active elements were susceptible to *Pvu* II digestion. Further analyses with the enzymes *Sst* II and the methylation sensitive and insensitive isoschizomers *Eco* RII and *Bst* NI showed the inactive *Ac* sequence was methylated at these sites, whereas the active *Ac* was hypomethylated. Although the active *Ac* at the *wx-m7* allele in different genetic backgrounds showed differences in the *Ac* DNA modification pattern, at least a fraction of genomic DNA contained *Ac* sequences that were unmethylated at all of the internal sites assayed (Chomet et al. 1987). These data suggested a role for DNA modification in the ability of *Ac* to transpose from the *waxy* locus and to destabilize unlinked *Ds* elements. The findings were the same when Schwartz and Dennis (1986) analyzed another allele of *waxy*. They isolated a spontaneous derivative (*wx-m9Dx-cy*) of the *Ac* element present in the *wx-m9Ac* mutant which does not itself transpose but can be induced to transpose by the presence of an *Ac* element elsewhere in the genome. The *wx-m9Ds-cy* derivative reverts to an active *Ac* form. A comparison of cloned isolates of the three forms of the element shows no differences in restriction enzyme pattern. Southern analysis of the genome organization of the elements shows marked differences in the methylation patterns. The active *Ac* element is methylated at one end of the element while the inactive derivative *wx-m9Ds-cy* is completely methylated at all *Hpa* II sites in the element. Chandler and Walbot (1986) obtained similar results associating DNA methylation with inactivation for the more recently discovered *Mu* element family.

These studies on the regulation of activity of genes belonging to transposable elements strongly support the involvement of DNA methylation in gene regulation (Fedoroff 1996; Martienssen 1996). Indeed, it was extended inactivation of eukaryotic transposable elements through methylation that forced Bestor and colleagues to propose that DNA methylation was restricted almost entirely to transposable elements (Yoder et al. 1997; Bestor 1998b), which constitute more than a third of certain eukaryotic genomes. The primary role of DNA methylation might thus be to prevent transposition, which would otherwise be expected to "lacerate the genome." Every transposition of the element has proven to leave behind some footprint nucleotide(s), mutating the host DNA at that point. Thus, keeping numerous elements in a repressed state could be of benefit to the host plant and

in a sense for the element's benefit, as well (Fedoroff 1996). That could explain why some elements developed specific mechanism such as the Tpn protein to activate themselves in a controllable way. But, as pointed out by Bird (1997), if methylation were to regulate transposon spread, then it has been singularly unsuccessful, as large amplifications of transposons have occurred in recent vertebrate and plant evolution, as well as in insects such as *Drosophila* that do not have methylation. Bird favors the notion that the primary role of methylation is to prevent genome-wide transcription of many genes including transcription of transposons themselves, except those which are required to be active in a certain tissue, and in this way to limit "background" transcriptional noise. In a stimulating discussion, Martienssen (1998) tried to reconcile the two views stressing the importance of DNA methylation in plant gene regulation, presenting also the notion that one way to alleviate the mutagenic effects of transposon movements on genes is to make them jump preferably to its other positions or in areas of a gene where the insertion of a transposon could not be so detrimental (i.e. introns, untranslated regions, etc.). Indeed recent data on the biology of transposon elements support this proposal (Kidwell and Lisch 1998; Waugh O'Neill et al. 1998).

Finally, strong evidence for the role of DNA methylation in modulating plant gene expression has been more recently obtained from studies of transgenic plants. By introducing extra copies of a specific gene, one might expect in many cases to overproduce the corresponding mRNA and protein products. Conversely, attempts at silencing genes have often employed an antisense strategy of expressing single-stranded RNA from the non-coding strand of a gene to bind to the mRNA, thereby preventing accumulation of the corresponding protein. Although these techniques have been successful in numerous applications, a body of literature is emerging that documents cases with unexpected outcomes in organisms as diverse as nematodes and plants. These observations encompass transgene silencing, i.e., failure to express certain transgenes. In some cases not only the transgenes introduced at ectopic positions in plant genomes can be unpredictably silenced, but also if the ectopic sequences are homologous to endogenous plant genes, silencing of the endogenous gene can frequently occur (Flavell 1994; Jorgensen 1995; Matzke and Matzke 1995). The role of DNA methylation in transgene inactivation will be thoroughly analyzed in Section V. Transgene epigenetic inactivation has provided clear-cut evidence for the involvement of DNA methylation in gene action. The involvement of epigenetic phenomena in unpredictable transgene inactivation in transgenic plants has also attracted the attention of the scientific community to epigenetics for practical reasons. Plant cultivars bred for specific characteristics, e.g., herbicide tolerance, in the laboratory, may lose this character when cultivated in farmer's fields. In addition, analysis and understanding of specific cases of allelic gene-gene interactions and inactivation of the transgene through DNA methylation helped our understanding of similar types of naturally occurring phenomena such as paramutation and viral resistance of plants (see also Section V A).

In conclusion, research on structural plant genes, transposon genes, and transgenes, points to a significant role of DNA methylation in gene transcription. But how does DNA methylation repress transcription?

The most direct mechanism by which DNA methylation could interfere with transcription would be to prevent binding of the basal transcriptional machinery to promoters. This is not a generally applicable mechanism because some promoters are transcribed effectively as naked DNA templates independent of DNA methylation. Certain transcription factors bind less well to methylated recognition elements; however, the reduction in affinity is often insufficient to account for the inactivity of promoters in vivo. It seems unlikely that DNA methylation would function to repress transcription globally by modifying the majority of CpGs in a chromosome, if the only sites of action are to be a limited set of recognition elements for individual transcription factors. The second possibility is that specific transcriptional repressors exist, that recognize methyl-CpG and either independently or together with other components of chromatin, turn off transcription. This mechanism would have the advantage of being substantially independent of DNA sequence itself, thereby offering a simple means of global transcriptional control. It would be especially attractive if the methylation-dependent repressors work in a chromatic context because then DNA could maintain the nucleosomal and chromatin fiber architecture necessary to compact DNA. Moreover, because chromatin assembly also represses transcription, methylation-dependent repression mechanisms would add to those already in place. But until recently, how cytosine methylation might affect nucleosome and chromatin organization was missing.

Bird and colleagues identified two repressors, MeCP1 (Meehan et al. 1989) and MeCP2 (Lewis et al. 1992), that bind to methyl-CpG without apparent sequence specificity (Boyes and Bird 1991). Like DNA methylation itself, MeCP2 is dispensable for the viability of embryonic stem cells, however, it is essential for normal development in mammals (Tate et al. 1996). Consistent with the capacity of methylation-dependent repressors to operate in chromatin, recent studies indicate the MeCP2 is a chromosomal protein with the capacity to displace histone III from the nucleosome. Moreover, MeCP2 contains a methyl-CpG DNA-binding domain, which might alter chromatin structure directly and a repressor domain, which might function indirectly to confer long-range repression in vivo. The capacity for MeCP2 to function in chromatin explains several phenomena connected with unique aspects of chromatin assembled on methylated DNA (Kass et al. 1997; Kass and Wolffe 1998; Grunstein 1998)A breakthrough in our understanding of DNA methylation, MeCP2 binding, and chromatin condensation leading to gene inactivation, came recently when Bird and his colleagues (Nan et al. 1998a, 1998b) and Jones et al. (1998) showed that MeCP2 exists as a complex with histone deacetylase. The complex also contains Sin3A, a co-repressor in other deacetylation-dependent silencing processes, as well as

several unidentified proteins. Both papers also show that methylation-dependent transcriptional silencing can be overcome by adding trichostatin A, which specifically inhibits histone deacetylase. The new results indicate that histone deacetylation is guided to specific chromatin domains by the underlying genomic methylation patterns and that methylation-dependent transcriptional silencing relies on histone acetylation (Fig. 3).

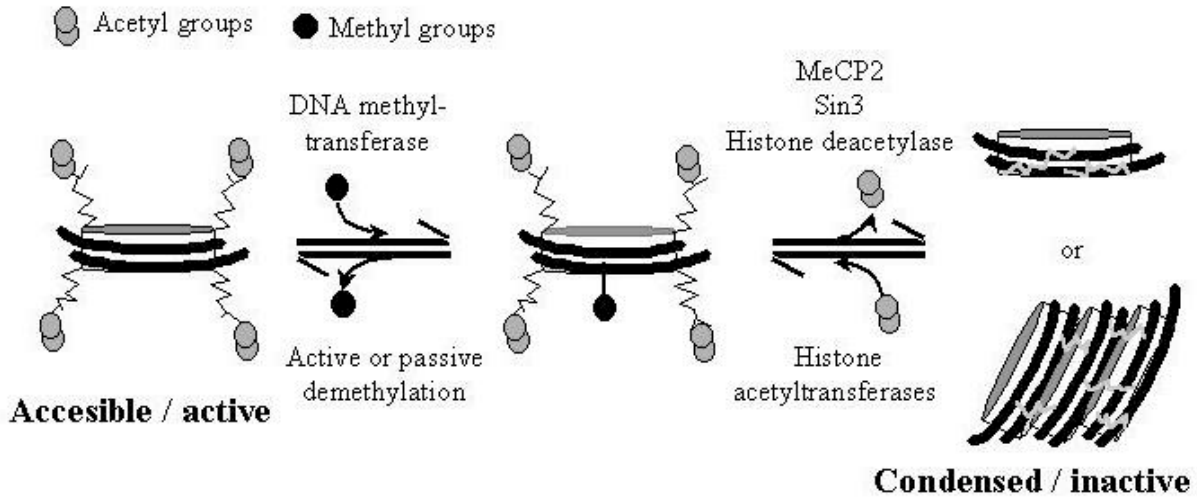

Fig. 3. The effects of cytosine methylation and histone deacetylation on transcription (modified after Bestor 1998a). Transcriptional silencing in many organisms is usually associated with the presence of 5-methylcytosine (5mC) in the DNA. A link between methylation and histone deacetylation has been revealed by the discovery that MeCP2 (a protein that binds methylated DNA) exists in a complex with histone deacetylase.

Thus not only the regulation of gene action by DNA methylation is more clearly established today, but also the details of molecular mechanisms involved are better understood. The data suggest that two global mechanisms of gene regulation: DNA methylation and chromatin structure can by linked by the MeCP2 protein complex. Underlying patterns of methylated cytosine residues are important in guiding chromatin structure, and like other epigenetic systems involving only proteins, can also do without the help of DNA methylation (Bestor 1998a). Organisms like *Drosophila* lacking any DNA methylation largely depend on chromatin structuring proteins to regulate the activity of some of their genes epigenetically. The "lock in" homeotic genes of the polycomb group (Paro et al. 1998; Pirrotta 1998) or the white (w^+) gene conferring red eye color causing a white mottled clonally inherited phenotype when translocated to the vicinity of centromeric heterochromatin, an epigenetic phenomenon called position-effect variegation (Wakimoto 1998), are two examples. Such epigenetic phenomena could also be stably inherited through meiosis (Jablonka and Lamb 1995; Cavalli and Paro 1998; Klar 1998). The work with MeCP2 now proves that in organisms where both epigenetic systems (based on DNA methylation or on chromatin structure) are functioning, MeCP2 coordinates their operation.

E. Implications in Breeding

For characters of major agronomic importance, successful identification of genotypes based on phenotypes requires an understanding of three major questions: 1. How do genes act and interact to control complex characters? 2. What kind of gene action do breeding systems exploit? 3. What conditions optimize heritability? The understanding of gene action is of paramount importance to plant breeders. Alleles with dominant, additive, or deleterious phenotypic effects influence heritability differently depending on whether they are in homozygous or heterozygous condition. An understanding of how heterozygosity and homozygosity affect gene action and interaction will facilitate decisions about whether the end product in breeding programs should be hybrids or inbred lines. Knowledge of the way genes act and interact will also determine which breeding system optimizes gene action more efficiently and will elucidate the role of the breeding systems in the evolution of crop plants.

In discussing the role of gene action (and interaction) in plant breeding in their most recent review in this series Fasoula and Fasoula (1997b) (see also Fasoulas 1993) defined the positive dominant gene as a fully expressed gene that can provide alone the required optimal amount of coded protein. This is also true for a positive epistatic gene in non-allelic interactions. Thus even in the hemizygous condition it could mask the unnecessary action of any kind of allelic counterparts. In homozygous conditions positive dominant genes complement themselves additively by securing their phenotype since whatever happens to either of the two copies the remaining allele could provide an optimum phenotype. Any alterations of these genes that lower quantitatively their expression generates semidominant genes (synonym of partially dominant genes) producing, in hemizygous conditions, sup-optimal amounts of coded proteins. Thus two such semidominant alleles complement each other additively in homozygous or heterozygous conditions. The two reviewers emphasized the importance of these two kind of genes and stressed the characteristic that they produce the best result in homozygous condition only; therefore inbreeding is the system for their exploitation, while codominant alleles are favored by heterozygosity. Several researchers have stressed the importance of quantitative changes in gene activity. Quantitative changes in gene activity, either at the individual RNA level (Tsaftaris and Polidoros 1993), the individual protein level (Damerval et al. 1994), or the enzymatic activity level (Mitchell-Olds and Pedersen 1998) (see also Section IV) have been found to be correlated with better performance of maize F_1 hybrids, thus hybrid vigor. This is in agreement with many classical quantitative genetic studies indicating that additive genetic variation could account for a large portion of vigor (Jones 1917) as well as with modern molecular analysis of hybrids, also stressing the importance of this kind of genes in hybrid vigor (Stuber 1994; Xiao et al. 1995).

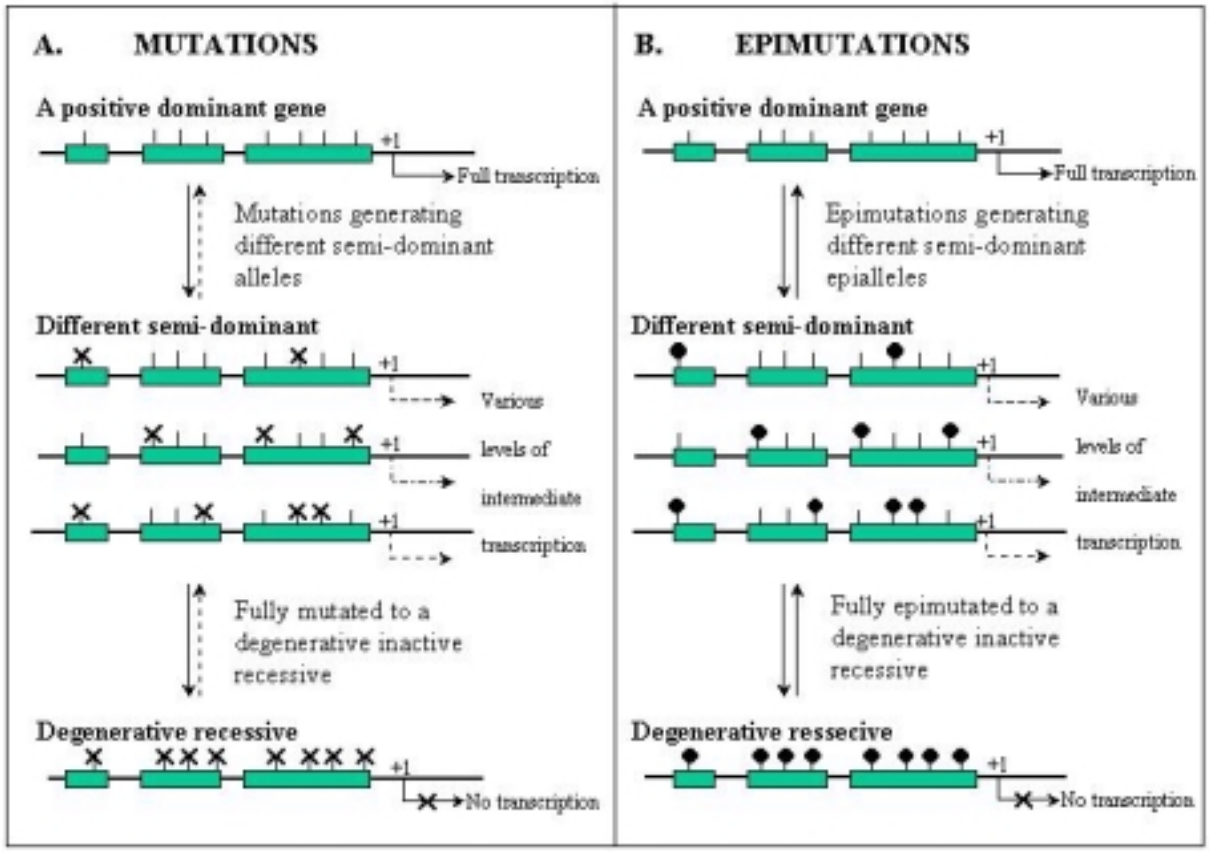

Fig. 4. Parallels between mutational events generating new alleles and epigenetic events generating new epialleles. Events are depicted on a modular promoter area (shown in boxes) that regulates the rate of expression of a gene. Vertical bars on the promoter indicate positions of mutations (x) or epimutations (•).

Fasoula and Fasoula (1997b) presented a number of classical mutational events that, by taking place intralocusly, could change quantitatively positive dominant genes to semidominant or even cause qualitative changes that mutate a positive dominant allele to different degenerated codominants or negative dominant alleles. In addition to postulated classical mutational events underlying different allelic phenotypes, particularly of quantitative nature, epigenetic changes of genes could generate different epialleles. The data presented in this Section supported the view that changes in the methylation status of a locus interfere in modulating its expression. In quantitative terms, complete demethylation of methylation sites critical for expression (i.e.

promoter and enhancer areas) will optimize expression, while gradual addition of methyl groups especially to these critical sites will quantitatively decrease expression. If a fully expressed unmethylated gene provides alone the optimal amount of an encoded protein, it will look like a positive dominant. After gradual addition of methyl groups will resemble semi-dominant and when fully methylated, resulting in complete inactivation, will appear to be a degenerative recessive.

In Fig. 4 such a parallel is depicted between the two mechanisms of generating new alleles with classical mutational events and new epialleles with changes in DNA methylation patterns (epimutations). For simplicity, changes are indicated in the modular promoter area, quantitatively regulating the rate of expression of the encoded protein. Thus, because this kind of epigenetic change is inherited it could cause significant variation that was unaccounted in the past and confounded with genetic variation (see also Section III).

Further comparisons of these two mechanisms generating variants (mutations and epimutations) are given in Table 3. DNA methylation as a generator of epimutations, could have other important implications for the breeder. Epialleles could emerge at high frequency in a single generation, by far exceeding the rate of mutational events giving rise to new alleles. Their reversion rate is by far higher and that will interfere in heritability estimation. Their emergence is highly affected by the conditions of plant growth (see also Sections IV and V) while random mutational events are largely considered independent of growth conditions. DNA methylation, by its mutational role (see also Section III F), gives rise to more permanent mutant alleles at a locus, while mutations only rarely lead to new epialleles (when, by chance, critical C-residues in methylation sites, are eliminated or generated). A better and clearer view of the role and significance of this new source of variants in plants will be obtained only when DNA methylation is systematically studied, taking into consideration the evolution of plants, the mode of their reproduction, their genotype (inbred line, hybrid, clone, their ploidy level), the degree of isolation during domestication, as well as the time and intensity of breeding effort.

Table 3. A comparison of heritable epigenetic and genetic generation of variants.

Property	Epigenetic	Genetic
Type of variation	Usually does not involve change in DNA sequence; involves change in DNA methylation, chromatin structure, or architectural organization of cellular structures	Involves change in DNA sequence
Origin of variation	Random changes due to imperfections of the DNA methylation pattern copying system or to non-directed effects of environmental factors	Usually random
Unit of variation	The activity of the gene	DNA bases, sequences
Frequency of "forward" variation	Very wide range: up to 100% per locus	More limited range: $<10^{-4}$ per locus
Frequency of "backward" variation	Very wide range: up to 100% per locus	More limited range: usually extremely low
Locus and tissue specificity	May be highly specific; the probability of a specific change could be 100% for a particular gene in a specific cell type and developmental stage; coordinated changes in several loci probably common	The probability of a particular change varies, but is always extremely low; several genes may change but only rarely in a coordinated way
Heritability	Varies	100%
Interconnection	Very frequently methylation leads to mutations	Only in rare cases mutations could affect methylation

III. DNA METHYLATION, GENETIC, AND EPIGENETIC VARIATION

A. Genetic and Epigenetic Variation

Creation of favorable variation that will enable the selection of superior genotypes is at the core of a successful breeding program. Kuckuck et. al. (1991), classified breeding methods into three groups: **selection breeding** where the breeder relies on existing variation in natural populations, genotype mixtures, etc., **combination breeding** where combinatorial crosses are made and the F_1 is not used directly but only to generate subsequent segregating material that will form the basis for selection, and **hybrid breeding** where combinatorial crosses are made to create new genotypes that will be used as F_1 hybrid seed.

In breeding programs of the past, genetic variation was considered to be due to allelic and epistatic combinations that owed their existence to alterations in the primary nucleotide sequences of the respective genes. Rare or induced chance mutations and their shuffling through meiosis and fertilization were considered as the only source of variation and formed the basis for selection. But, epigenetic information systems, like DNA methylation, could generate epigenetic variation that had never been considered in plant breeding as a source of phenotypic variation. As stated by Jablonka and Lamb (1995) this heritable epigenetic variation is now beginning to receive closer attention and there is a growing realization that it plays a critical role in plant development (the transition from genotype to phenotype) and consequently in plant breeding.

Polymorphism for DNA methylation is extensive not only among species but among genotypes of a species, individuals belonging to the same genotype, between different organs and tissues of an individual and even among mitotically derived somatic cells of a certain tissue. This polymorphism in DNA methylation interferes with gene expression and consequently becomes associated with phenotypic variation. It has been shown that epigenetic variation is not only mitotically stable but also could persist through meiosis in the next generation, mimicking genetic variation (Jablonka and Lamb 1995; Klar 1998). This situation is reflected in the so-called somaclonal variation, in asexually propagated copies of a plant generated in vitro from somatic cells. Populations of somatic cells are mosaics of different epialleles, thus containing much epigenetic variation, which is not only stable in mitotic somatic division but could be inherited through meiosis, if some of these somatic cells generate gametes (Buss 1987, see Section III D).

In Section III B data are presented showing that epigenetic variation is under developmental control. Reproducible and systematic differences are observed in different tissues and developmental stages. In addition, and perhaps more significantly, epigenetic changes are highly affected by plant-environmental interactions. This indicates that epigenetic phenomena will interfere not only with the creation of favorable variation, but also in the later steps of the breeding program, for the evaluation and selection of superior genotypes through expression of phenotypes.

DNA methylation in addition to being the cause of epigenetic variation is also the cause of genetic variation due to its role in generating new mutations. Methylated cytosines frequently deaminate to T, thus 5mC are hot spots for mutations, providing an interconnection between epigenetic and genetic variation. This relationship between DNA methylation and mutation will be examined in Section III D.

Elaboration of some unique aspects of plant development provides a context in which the role of genetic and epigenetic changes in the plant genome as a source of variation can be understood (Richards 1997). Structures in a mature plant develop after embryogenesis by the reiterative action of meristems, which are collections of undifferentiated cells poised at the apices of growing shoots and roots. Delaying the bulk of growth and differentiation until after embryogenesis, is one way plants can fashion their form to the environmental conditions. The identity of structures produced by apical meristems changes throughout development. For example, early in development the shoot apical meristem produces vegetative leaf structures, while later in development, a transition to production of reproductive floral structures occurs. The relatively late divergence of vegetative (somatic) and floral (reproductive) lineages allows hereditary alterations that occur through somatic development, either of genetic or epigenetic nature, to be transmitted to subsequent generations.

The need for developmental plasticity and environmental interaction suggests that plants extensively employ epigenetic regulatory strategies. Such strategies rely on heritable and often reversible changes to access the underlying genetic information, and not only to mutation. Epigenetic strategies might also be elaborated in plants because of the opportunity to select and transmit to the next generation, metastable epigenetic states established throughout development in somatic tissues. Consequently, one should expect that alteration of plant DNA methylation affects plant development and plant breeding significantly, provided that differential DNA methylation is an important component of epigenetic regulation of plant genes.

B. The Extent of Variation in DNA Methylation and its Inheritance.

Significant variation exists for percentage of total genomic methylation among plant species (Table 2). Significant variation also exists in the pattern of DNA methylation. This variation occurs among relative species, among cultivars or lines belonging to the same species, among individual plants of a certain line, among different organs and tissues (especially during successive developmental stages), and even among individual somatic cells.

For example, analyzing the tomato genome, Messeguer et. al. (1991) screened DNA from two different cultivars with methylation sensitive isoschizomers for polymorphism. Both *Lycopersicon esculentum* cv. VF36 and *L. pennellii* (LA716) have homosequential chromosomes and constitute parents of an F_2 population from which an RFLP map has been constructed. Of the clones tested with *Bst* NI/*Eco* RII, 86% detected polymorphism between the two parents with one or both enzymes; of this, 19% were detected by *Eco* RII only and thus attributable to methylation. Since *Eco* RII is sensitive to methylation at the second C in the restriction site CC*(A/T)GG, the methylation polymorphism can be attributed to variation in methylation at the CpNpG target site. Similar comparisons were made for the enzymes *Msp* I/*Hpa* II; 83% of the clones detected polymorphism with one or both enzymes, and of this 52% were detected with either *Msp* I or *Hpa* II, but not with both, and probably represent methylation polymorphism at the CpG site. Twenty-nine percent of all polymorphisms fell into this category and in all cases, the polymorphic fragment detected by *Hpa* II was larger than that generated by *Msp* I, consistent with the interpretation that polymorphism is due to a difference in methylation at the CpG site. Thus, methylation polymorphisms are not uncommon. In the case of *Bst* NI/*Eco* RII, CpNpG methylation polymorphisms account for approximately 25% of all polymorphism detected. For *Msp* I/*Hpa* II, in which methylation polymorphism can be detected at either CpG or CpNpG sites, nearly 50% of the polymorphisms are due to differential methylation of allelic DNA and occur with nearly equal frequency at both the CpG and CpNpG target sites.

Since for any polymorphism to be useful, it must be stably inherited in an allele-specific manner, Messeguer et al.(1991) tested this by segregation analysis in the F_2 in the same cross used to create a map composed of more than 250 RFLP markers. Selected probes, determined to uncover methylation polymorphism with either *Bst* NI/*Eco* RII or *Msp* I/*Hpa* II, were placed onto the RFLP map using methylation-insensitive enzymes to determine the map position of the primary target site. The same probes were then mapped (using DNA from the same F_2 plants) with the appropriate enzyme (either *Eco* RII or *Hpa* II) known to detect the methylation polymorphism. In all cases, the methylation polymorphism co-segregated perfectly with the non-methylation genetic polymorphism. The methylation polymorphism observed in the parents was highly heritable, since it is passed to the offspring in a normal Mendelian fashion. Because the presence or absence of methylation segregates in perfect accordance with the primary target sequence (as determined by the methylation-insensitive enzymes), the basis for the methylation polymorphism resides at or near the methylated sequence and cannot be attributed to polymorphism for a *trans*-acting factor required for methylation. While allele-specific methylation is apparently inherited, the high level of methylation polymorphism observed in this study suggests that this is a dynamic process and that new alleles with heritable methylation patterns are likely to occur with a relatively high frequency in plant populations.

Similar results were obtained by Tsaftaris et. al. (1997, 1998) in analyzing the polymorphism in the pattern of DNA methylation among different maize inbred lines and hybrids using the CRED-RA technique. By changing different methylation sensitive restriction enzymes for targeting the CpG or CpNpG sequences and/or changing the sequence of primers for the PCR step, significant variation was recorded between different inbred lines of maize, even among related second-cycle lines originated from the F_2 generation of a single F_1 hybrid used as starting material for the selection program. Accordingly, variation was recorded among F_1 hybrids (see also Section IV for more details). Also working with maize inbreds, Bedford and Van Helden (1990) detected significant variation for DNA methylation patterns. While working with orange trees and using the CRED-RA technique, Cai et al. (1996) detected extensive variation in DNA methylation patterns among orange tree cultivars.

For total percentage methylation, considerable variation has been observed between different organs and tissues as well as in different developmental stages. In tomato for example several stages/tissues were distinguished with different levels of methylation. At the low end were protoplasts from leaves as well as the immature stem, leaf and root tissues with an average value of 20% 5mC. At the high end were mature leaves and fruit as well as green fruit and seeds (average value of approximately 25%). Pollen had a value (22%) intermediate between immature and mature tissues (Messeguer et al. 1991). Similarly DNA in young maize seedlings is less methylated and there is a steady increase of total DNA methylation as the plant passes through different developmental stages (Tsaftaris and Polidoros 1993; Tsaftaris et al. 1997). A statistically significant increase in DNA methylation was observed, from an average of 27% for leaves of 10-day-old seedlings, to 29% for leaves of 1-month-old plants in different maize genotypes, inbreds and hybrids.

Lower levels of methylation in immature tissues may be due, at least in part, to the fact that these tissues are likely experiencing higher levels of cell division and thus contain higher proportions of hemi-methylated DNA (Messeguer et al 1991). Increased levels of methylation in mature tissues may also relate to

gene expression. As tissues differentiate and cease dividing, it is likely that they need only a subset of the genes required by actively growing tissues. Higher methylation levels in more mature tissues may result from the methylation of genes whose products are no longer required. Indeed studying genome expression through three successive stages in maize development, using dot or slot blot hybridization to quantitate individual amounts of many different mRNAs, Tsaftaris and Polidoros (1993) and Tsaftaris (1995) found a steady decrease shortly after germination (Fig. 5).

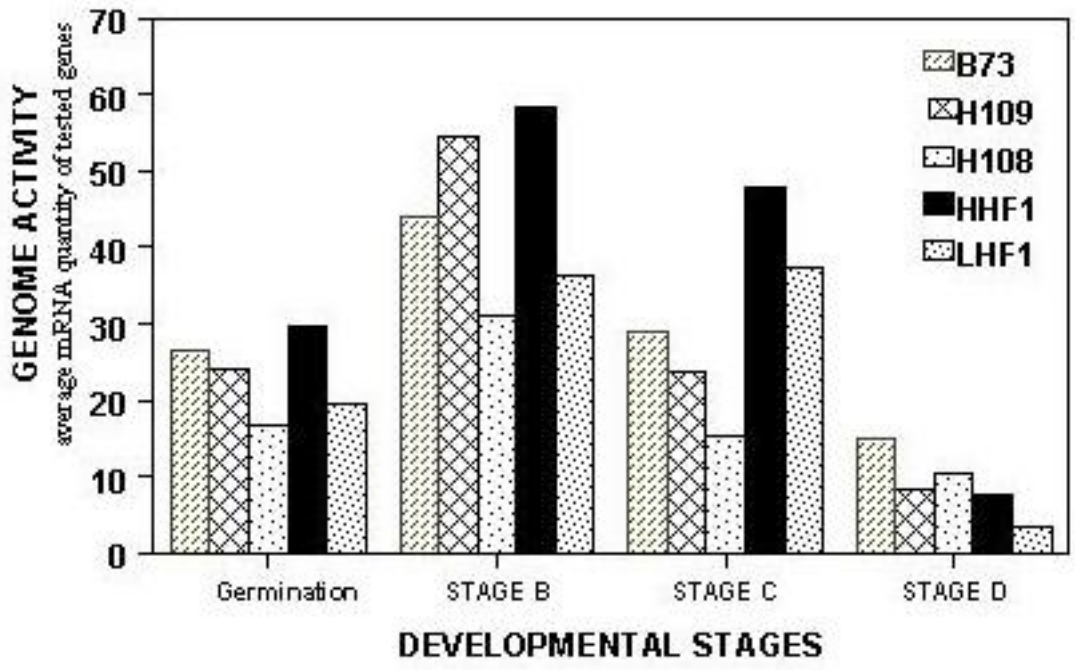

Fig. 5. The genome activity of parental lines B73, H109, H108, and two of their hybrids, the highly heterotic HHF_1 and the poorly heterotic LHF_1, in scutellum during germination and in three developmental stages of the leaves (Stage B: 15 days, Stage C: 30 days, Stage D: 66 days), calculated after pooling the quantitative measurements of the individual mRNAs for the 35 tested gene loci for every genotype in each of four developmental stages.

Pollen represents a possible exception to the correlation between level of methylation and potential gene expression. At the time of maturation, many, if not most of the gene products for germination and fertilization are already present in the pollen grain and thus there is not a requirement for transcription of larger numbers of genes during pollen germination (Mascarenhas 1990). Yet, the observed level of methylation in mature tomato pollen (22%) is not as high as that observed in several of the mature tissues/stages. Seeds have significantly higher levels of methylation than pollen. This suggests that, after fertilization and during development and maturation, embryonic DNA could be subjected to *de novo* methylation. Germination and subsequent cell division and growth result in a dramatic drop in methylation, until it gradually increases as cells become terminally differentiated and the plant ages.

Inbred lines have reduced genetic variability; yet, as indicated earlier, they show substantial heritable variation. Even when individuals have identical DNA sequences at a particular locus, the chromatin associated with these sequences is not always identical. In an inbred line of wheat, Flavell and O'Dell (1990) found that genes coding for a high molecular weight glutenin protein had seven methylation variants (seven epialleles). Some of these patterns of methylation were stably inherited both somatically and between generations. In crosses between epiallelic variants, the F_1 individuals showed the patterns of both parents and, in the F_2 generation, the patterns segregated so that both the F_1 and parental patterns were found. In other words, the methylation patterns showed the same behaviour as Mendelian alleles, and segregated with the DNA sequences on which they were imposed. Occasionally a new methylation pattern appeared, showing that the gene had high epimutabilitv.

Perhaps the most clear-cut demonstration of a relationship between heritable differences in methylation and heritable differences in gene activity has come from studies of maize transposable elements (see also Section II D). The action of transposable elements during development of a plant often results in characteristic patterns of variegation reflecting the tissue, the timing, and the frequency of transposition events among individual somatic cells during development. Unstable changes in variegation patterns, termed "changes in phase" associated with changes in methylation of the element's DNA, have been described for all transposon families. Suitable genetic

marker loci were developed to monitor transposon activities in different tissues, like the *P* locus of maize, for instance, providing a somatic marker conditioning pigmentation of pericarp and glume tissue of the ear (Styles and Ceska 1977). It is now clear that many of these transitions are associated with changes in the methylation level of the element: a genetic property, the transposability of an element, is correlated with its level of methylation as has been shown clearly by Fedoroff and her co-workers (Fedoroff 1996) in their studies of the *Spm* elements (Fig. 2).

The likelihood that a programmable inactive element will become heritably active or inactive depends on several factors, all of which have been shown to affect methylation levels. First it depends on position in the plant. When transmitted through tiller gametes, the likelihood of reactivation in the following generation is increased, whereas transmission through gametes produced on the main stalk tends to stabilise the inactive state. In general a plant's ability to inactivate an active element increases with the distance from the base of the plant. The second factor influencing the activity of programmable elements is their parental origin: activation is most likely if the element is transmitted through female gametes. Third, change in a programmable element depends on the presence or absence of other active *Spm* elements (see Section V). An active element can reactivate a programmable inactive element and increase the probability that the new activity will be transmitted to the next generation. The presence of an active element also increases by two to three orders of magnitude the likelihood of reactivation of a cryptic element in the same genome. Thus, stages in the development of the plant exist during which changes in element methylation is reproducibly altered, suggesting also the existence of a genetic mechanism that modulates DNA methylation in certain tissue and/or developmental stages. Likewise, environmental conditions of plant growth play a role.

The switch from an active to an inactive state can be gradual. Breeding plants in which there was a low frequency of reactivation of a programmable element produced progeny in which the element was even more inactive. Eventually, after several generations of selection, the element no longer showed differential expression during development. It had become a cryptic element. This progressive change to a stably inactive element was paralleled by a progressive increase in methylation.

Studies of other transposable elements in maize and other plants have led to similar conclusions (Brutnell and Dellaporta 1994; Peterson 1995). Fedoroff et al. (1989) concluded an analysis of *Spm* regulation and transmission by stating:

> "Perhaps the most striking observation that has emerged from the analysis of the *Spm* element's developmental control mechanism is that epigenetic changes in the present generation can influence the expression pattern of the element in the next generation."

Another system useful for studying plant DNA methylation in transgenic plants utilizes the maize *A1* as a marker gene (Meyer 1995). Introduction of the *A1* marker gene triggers the accumulation of bric-red pelargonidin pigments in floral cells of transgenic *Petunia* . The *A1* gene encodes a dihydroflavonol reductase (DFR), an enzyme of the anthocyanin pigmentation pathway. The receptor line RI01 produces only low amounts of cyanidin and delphinidin, the two pigments that determine floral pigmentation in petunia. The almost white flowers of RI01 accumulate dihydrokaempferol, a substrate of the maize DFR. In transgenic plants expressing the *A1* gene, dihydrokaemoferol is converted into leucopelargonidin, which is further processed by endogenous enzymes into a red pelargonidin pigment. Petunia plants are especially suitable for long-term studies on transgene stability because individual plants, regenerated from leaf tissue or single protoplasts, can be preserved in tissue culture or in the greenhouse over several years. Defined genotypes are readily propagated and amplified via stem cuttings or protoplast culture. New flowers continuously emerge from plants grown in the greenhouse, which allows a constant monitoring of transgene expression. Transcription instabilities are detected by changes in floral pigmentation patterns, and epigenetic variants of isogenic material can be selected and propagated as individual plants, thus providing sufficient plant tissue for molecular analysis. The system has proven to be particularly useful for the identification of multiple parameters that influence DNA methylation and its involvement in different gene silencing mechanisms.

In transgenic petunia harboring a chimeric *A1* gene from maize, the number of plants producing flowers with white or variegated petals and with weakly pigmented blooms was dependent on the developmental stage of the progeny plant producing gametes and seeds. Plants grown from seed that was produced from aged parental plants late in the previous season showed higher inactivation rates through methylation, again supporting the view that endogenous developmental signals affect the level and pattern of DNA methylation and that variation in DNA methylation is stably inherited.

In addition to internal (genetic) mechanisms affecting the extent of DNA methylation in different tissues and in different developmental stage, external (environmental) signals have been shown to affect the extent of DNA methylation by their interaction with the plant. This is true for both: the total level and the pattern of DNA methylation.

It was the pioneering work of Cullis (1981, 1983) and Schneeberger and Cullis (1991) that provided strong evidence that the environmental conditions during plant growth (in this case the amount of phosphoric fertilizer applied to flax) could lead to variant phenotypes. This new epigenetically induced variation did not only give rise to different phenotypes but also was stably inherited in the next generation, suggesting an environmentally imposed heritable epigenetic change.

The role of environmental conditions of plant growth on the extent of variation of DNA methylation in maize was also studied by Tsaftaris et al. (1997, 1998). Maize parental inbreds of different origin and different F_1 hybrids were grown under different plant densities. Wide spacing of 1.5 m was considered a favorable, less stressful condition of growth (isolation environment) in comparison to a high density (0.3 m) stress (competition environment). Significant differences were found for total methylation or patterns of DNA methylation between the two conditions of growth. On average stressful conditions of growth lead to more methylated sequences in comparison to the wide spacing with nil competition (see also Section IV).

Similar results were obtained by monitoring the status of methylation of the *Ac* transposable element in maize plants grown under competitive stress (0.3 m) versus nil competition (1.5 m) (Kafka 1996; Tsaftaris et al. 1998; Kafka and Tsaftaris in preparation). The genotype used in this study was an F_1 hybrid harboring the *wx-m7* allele (*wx-m7* X *wx*) that was shown by Schwartz and Dennis (1986) to contain an inactive methylated *Ac* element at the *wx* locus. Demethylation, thus reactivation of the *Ac* element was monitored by movement of defective *Ds* element from the a_1m_3 gene controlling anthocyanin pigmentation and was measured either as percentage of seeds with demethylation events or average frequency of demethylation events. Analysis of a few thousands selfed F_2 seeds in three consecutive years indicated that demethylation occurred more extensively in the isolation environment, with no stress (Fig. 6).

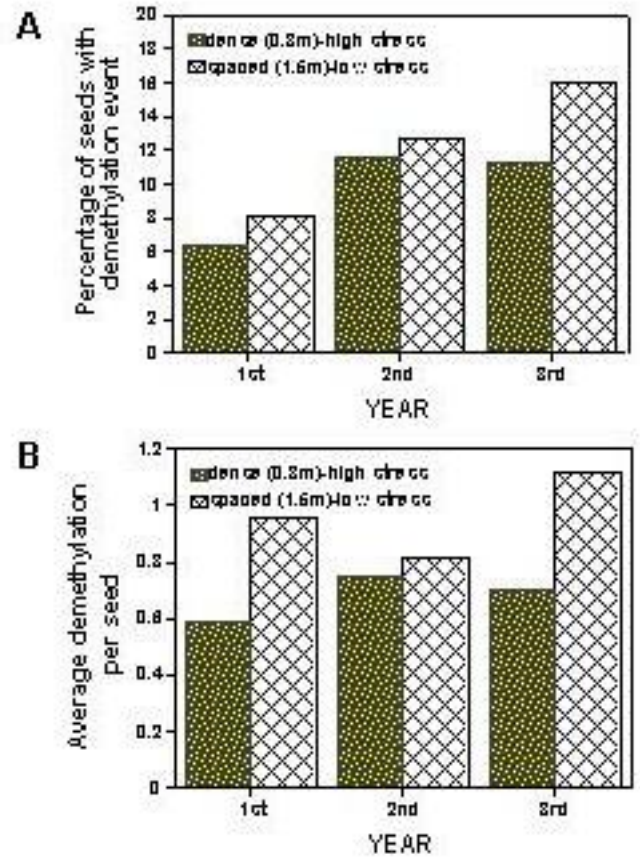

Fig. 6. Demethylation study of the methylated *Ac* element for three consecutive years in the field in F_1 plants harboring a methylated *Ac* element in the *wx* locus (*wx-m7* X *wx*) and growing under stressful competitive conditions or spaced plants with nil competition and low stress (modified after Kafka, 1996)

Several researchers who investigated the occurrence and pattern of DNA methylation of transposable elements also reported that environmental conditions of plant growth strongly affect the extent of methylation thus the activity of transposable elements (Pan and Peterson 1988; Fedoroff 1996).

For methylation involved in the paramutagenicity of the paramutant *r*-gene in maize (see Section V) Mikula (1995) provided data that environmental conditions, applied at a specific stage of development, cause specific epigenetic changes and affect the expression of a specific allele. At maturity, *R*-allele expression in test-crosses of male gametes derived from *R/R*-lst seedlings raised 15 days at 32^0C and continuous light differed significantly from those of sib seedlings raised for 15 days at 22^0C and continuous light and shifted to six 12-h light/12-h dark cycles, for days 16-21. Thus, environmental conditions of seedlings strongly affected the level of epigenetic change. In addition, these epigenetic changes were stably inherited through the gametes.

Additional data supporting the view that the environmental conditions of plant growth affect the extent and pattern of DNA methylation come from studies of transgenic plants. In the first such report on silencing of

the *A1* transgene in petunia, silencing of the transgene through methylation appeared to require high light intensity and was strikingly greater in field-grown plants than in greenhouse-grown plants (Meyer et al. 1992). Also populations of transgenic *N. sylvestris* plants harboring a chitinase transgene, germinated in closed vessels, transferred to soil, and then raised to maturity in a greenhouse showed a high incidence of silencing, but when the same seeds were germinated in a greenhouse in open vessels containing soil, then raised to maturity, no methylation was observed (Dorlhac de Borne et al. 1994).

That environmental factors affect the stability of transgene expression was also supported by a study of transgenic cell suspension cultures of *M. sativa* carrying a single copy of an introduced herbicide resistance gene (Walter et al. 1992). A 10-day heat treatment at 37^0C resulted in loss of phosphinothrycin (PPT) resistance in 95% of cells grown in non-selective medium. In contrast, only 12% of the same cultures maintained for up to 150 days under non-selective conditions, at a constant 25^0C, showed loss of PPT resistance. Loss of resistance in response to high temperatures was not due to loss of the introduced DNA, but rather to inactivation of the transgene.

Important data for the role of environmental conditions transgene inactivation were obtained by Brandle (1995) who studied in tobacco the role of transplanting in inactivation of transgenes through methylation. Young transgenic tobacco plants maintained transgene activity when grew under control conditions in the nursery, whereas high percentage of those that were transplanted in the field lost the activity of their transgene.

In conclusion, extensive variation exists for DNA methylation among plant species, cultivars, tissues and cells within tissue. Both internal (genetic) and external (environmental) mechanisms have been described that modulate genome methylation. Environmental conditions of plant growth affect the methylation status of its genes. In general, imposing different biotic and abiotic stresses to the plant, leads to increased gene methylation, thus, to degeneration of genome activity. In contrast, favorable growth conditions and absence of stress is associated with lower genome methylation and optimum expression.

C. Variation of DNA Methylation as a Source of Heritable Epigenetic Phenotypic Variation.

Epigenetic changes mimic the way allelic and epistatic interactions modulate gene activity leading to phenotypic variation. For example, transgenic plants bred for a specific character may loose their specific phenotype when the cloned DNA has been epigenetically silenced through methylation. Similarly, some transgenes (like other internal plant genes) can *trans*-inactivate their allelic or non-allelic homologous loci leading to further phenotypic changes. Thus polymorphism in DNA methylation due to *de novo* methylation of certain loci could lead to modified expression with heritable variants. In addition changes to transgene methylation have been shown to be imposed by environmental conditions of plant growth.

Another line of research has provided data indicating that gross changes in total methylation of plant genomes lead to phenotypic variation. Until recently, alteration of DNA methylation in eukaryotic organisms was principally achieved by application of DNA methylation inhibitors, such as 5-azacytidine or 5-azadeoxycytidine (Jones 1985). Treatment with methylation inhibitors induced developmental changes in several plant species, including rice (Sano et al. 1990), *Triticale* (Heslop-Harrison and Bennett 1990), flax (Fieldes 1994), and tobacco (Vyskot et al. 1995). Although similar changes have been observed in several studies, the effects were variable. The phenotypes induced in the first generation of exposure to 5-azadeoxycytidine were inherited in subsequent generations in certain cases (Sano et al. 1990; Fieldes 1994), suggesting heritable changes. Similar inhibitor studies also demonstrated inheritance of genomic hypomethylation in progeny (Heslop-Harrison and Bennett 1990; Sano et al. 1990; Vyskot et al. 1995). Finally, in other cases, the induced changes were transitory and were not transmitted in a recognizable manner to progeny. These inhibitor studies suggest that DNA methylation plays a role in formulating the phenotype of a plant and its progenies.

More recently, genetic tools have been developed that allow more specific and controlled manipulation of gross genomic DNA methylation in plants. A reverse genetic approach has been taken by two different groups (Finnegan et al. 1996; Ronemus et al. 1996) who made transgenic arabidopsis plants that express antisense constructs of a native cytosine-DNA methyltransferase gene (*MET1*). Genomic hypomethylation varies greatly between independent antisense transgenic lines, but the most severely affected lines have 5-methylcytosine levels ranging from 10 to 30% that of wild-type. Both single-copy and repetitive DNA sequences were hypomethylated in *MET1* antisense lines. Taking a traditional genetic approach, Richards and his colleagues (Vongs et al. 1993) isolated arabidopsis mutations (*ddm*: decrease in DNA methylation) that reduced nuclear 5-methylcytosine levels. The best-characterized mutations defined the *DDM1* gene. Homozygotes carrying recessive *ddm1* alleles contain 30% of the wild-type levels of 5-methylcytosine. The *ddm1* mutations do not map to the two known cytosine-DNA methyltransferase genes of arabidopsis (Finnegan and Dennis 1993), nor do they affect DNA methyltransferase activity detectable in nuclear extracts. In addition, *ddm1* mutations do not appear to affect the metabolism of the active methyl group donor, S-adenosylmethionine (Kakutani et al. 1995). Consequently, the *DDM1* gene product is a candidate for a novel component of the DNA methylation system, or is involved in determining the cellular context (e.g., chromatin structure, subnuclear localization) of the

methylation reaction. Recently, several groups have isolated additional mutants that reduce methylation in arabidopsis. The *hog1* mutant (Furner et al. 1998) and the *som2* mutant (Mittelstein-Scheid et al. 1998) impair gene silencing and homology dependent gene silencing (HDG), reduce genomic DNA methylation, and are not alleles of the *DDM1* locus. In parallel, Furner et al. (1998) described two monogenic, *trans*-acting, recessive mutations, *sil1* and *sil2*, that specifically reduce gene silencing but have little or no effect on HDG. Further work is needed to define how many genes and biochemical processes affect DNA methylation in plants.

The phenotypic consequences of lowering genomic methylation in arabidopsis through genetic means are varied, but striking. A variety of morphological phenotypes arose in the *MET1* antisense lines and different lines had different combinations of phenotypic characters. Similarly, independent inbred *ddm1* mutant lines had different constellations of developmental changes. Related phenotypic characters occurred in the different *ddm1* and *MET* antisense lines, indicating that DNA methylation played a role in the regulation of processes controlling these characters (Kakutani et al. 1996).

In the antisense expression studies, the extent of morphological alterations was correlated with the extent of genome-wide hypomethylation. In addition, removal of the *MET1* antisense construct by genetic segregation lead to genomic methylation and a reduction in the changes of the morphological phenotypes (Ronemus et al. 1996). In the *ddm1* mutant studies, progressively more severe developmental changes were seen upon inbreeding the mutants, which coincided with the progressive loss of DNA methylation in the single-copy genomic component. Isogenic *DDM1* wild-type control lines failed to show phenotypic changes or methylation changes, demonstrating that inbreeding depression was dependent on the DNA hypomethylation background. The parallels between the effects caused by *ddm1* and *MET1* antisense lines indicated that DNA hypomethylation was directly involved in the onset of phenotypic changes in *ddm1* lines.

In contrast to the *MET1* antisense results, removal of the hypomethylation locus did not lead to reversion of the morphological phenotypes that developed in *ddm1* inbred lines. Genetic analysis indicated that different phenotypic characters in *ddm1* inbred lines could be separated from the potentiating *ddm1* mutation and mapped to heritable variants at dispersed sites in the genome. In the cases examined to date, the variants behaved as stable Mendelian factors, some of which act as dominant alleles and other as recessive alleles. Variant formation appears to occur at particular loci of the arabidopsis genome as evidenced by the recovery of several independent dominant lesions that conditioned late flowering and mapped to the same genetic location.

Data with *MET1* antisense and *ddm1* should be repeated with other higher plants. Arabidopsis has a small genome with few copies of transposable elements (Bevan et al. 1998). According to Martienssen (1998), in plants like maize, with thousands of cryptic methylated transposons, genome-wide demethylation perhaps would reveal numerous mutations that may even contribute to early lethality, contrary to viability of the *ddm1* and the *Met1* of arabidopsis.

Additional evidence that epigenetic changes in gene activity could lead to variation comes from the phenomenon of parental imprinting (Surani 1998; Martienssen and Richards 1995). Parental imprinting is defined as the differential expression of alleles dependent upon their parental origin. Genomic or gametic imprinting is a unique manifestation of epigenetic inheritance in which expression of certain genes is governed by their parental origin, from generation to generation. Some of these genes show expression after paternal inheritance while others are expressed only when inherited from the maternal germ line. The most striking feature of imprinted genes therefore is that the active and inactive parental alleles of the same DNA sequence coexist within individual cells. In animals, over 20 imprinted genes have so far been identified (Surani 1998). Parental alleles of imprinted genes display differential DNA methylation capable of regulating gene expression responsible for regulating and transmitting the repressed and active states of imprinted genes to the gametes. While parental imprinting has been widely observed in animals, (especially mammals) it occurs more rarely in plants (Lin 1975; 1982; 1984; Heslop-Harrison and Bennett 1990). The best-documented case is for a maize gene(s) controlling endosperm development in the kernel and located on the long arm of chromosome 10 (Lin 1975). The evolutionary significance of the endosperm is due to its role in seed development, while due to the role of endosperm as storage tissue is the target of many breeding programs. In cereals, for example, breeding for yield is largely breeding for increased endosperm weight. Endosperm is a triploid tissue that develops normally only when the maternal:paternal genome ratio is 2:1. Lin (1982; 1984) not only verified the importance of such a balance, but also verified that the imprinting of paternal origin gene(s) controlled endosperm size and mapped the gene(s) to the long arm of chromosome 10. Thus, epigenetic changes by imprinting genes can cause significant variation in a character, such as the size of the maize endosperm, which is closely tied to grain yield.

D. DNA Methylation and Somaclonal Variation.

Tissue culture is used extensively in plant biotechnology. Its direct applications include clonal multiplication of plants by micropropagation, elimination of viruses by meristem-tip culture, the production of doubled haploids by anther culture and the use of cell cultures for the production of secondary products. Tissue culture also has crucial indirect application as a means of providing target cells for genetic engineering, particularly where *Agrobacterium*-mediated gene transfer is not possible and approaches such as direct DNA uptake into protoplasts or particle bombardment are required. In these applications it is of paramount importance that the

plants derived from culture are true-to-type. However, uncontrolled instability can occur when plant cells are cultured in vitro. This unexpected source of variation or somaclonal variation was once hailed as a "novel source of variation for crop improvement," but, due largely to its unpredictability as a breeding tool, enthusiasm for this application has diminished and somaclonal variation has lost much popularity in recent years (Karp, 1993). This should not be interpreted to mean that somaclonal variation is no longer a problem in plant tissue culture. The production of plants that are not true-to-type still causes embarrassment and sometimes serious financial losses in the micropropagation industry and reduces the commercial value and breeding potential of somatic hybrid and transgenic plants (Karp, 1993). What is the source of the phenotypic variability observed in regenerated plants when, according to the rules of asexual reproduction, "constancy should prevail" (Karp and Bright, 1985)? Despite numerous studies, the molecular basis of somaclonal variation is poorly understood. Fundamental questions, such as the roles of tissue-culture-induced versus preexisting variability in somatic cells and genetic versus epigenetic changes, remain unresolved. Other peculiar aspects of somaclonal variation, such as the high frequency of occurrence obtained (which routinely exceed what has been achieved using conventional mutagens), the appearance of homozygous changes, and the generation of variants in tissue culture identical to those observed among homozygous lines of inbreeding species as soybean (Roth et al. 1989), require an explanation.

Somaclonal variation is generally attributed to tissue-culture induced, heritable genetic changes (Kuckuck et al. 1991; Phillips et al. 1994) rather than preexisting genetic or even epigenetic variation in the somatic cells of the explant. Distinguishing between tissue-culture induced and preexisting genetic or epigenetic variability is difficult unless the exact genetic or epigenetic makeup of individual cells in the explant was known prior to its culture.

A somatic cell or small group of cells, having a unique genetic constitution or epigenetic modification, may reproduce preferentially in culture. Any plant eventually regenerated from this cell(s) would probably be scored as a variant acquired during the culture period, although in reality it comprised a minuscule, possibly even undetectable, proportion of the cells in the original explant (Matzke and Matzke, 1995). As it is becoming more apparent that somatically acquired epigenetic modifications in plants can be mitotically stable and meiotically heritable, more emphasis is given to variation in DNA methylation as a source of somaclonal variation.

The first such evidence involving DNA methylation originated from studies of regenerants from crown gall tumor lines where changes in T-DNA methylation were associated with phenotypic variation (John and Amasino 1989) as well as from demethylation, thus reactivation of silent *Ac* elements following tissue culture (Brettell and Dennis 1991; Peschke et al. 1991). But the most compelling evidence suggesting that a substantial proportion of somaclonal variation might be due to diverse, preexisting epigenetic states that are mitotically and meiotically stable in plants regenerated from individual somatic cells, was provided by Neuhuber et al. (1994). Their system involved double transformation experiments in which two transgene constructs, containing different antibiotic selection markers Hyg^R and Kan^R, were introduced into tobacco leaf cells by using sequential transformation steps. Regenerated double transformants (DTs) were grown to maturity. Seeds obtained through self-fertilization and backcrosses to untransformed tobacco were sown on medium containing either kanamycin or hygromycin, or a combination of both antibiotics. With respect to sexual transmission of hygromycin resistance, there were several categories of DT lines. Some behaved like the homozygous parent, producing 100% Hyg^R progeny in selfs and backcrosses. Other DTs, however, did not segregate for the Hyg^R marker like the parental explant source. Among the variations observed, there were lines which did not produce any Hyg^R progeny at all, and one line whose DT backcrosses produced progeny segregated primarily in 1 Hyg^R:1 Hyg^S. Finally, one line, H11, spawned DTs that clearly behaved as if they were homozygous, hemizygous, or recessive for the *Hyg* locus. Further examination specifically of H11 proved that the *Hyg* locus was present in all DTs. One explanation for the results with H11 is that the *Hyg* transgene complex in H11 homozygotes existed as an epigenetic mosaic prior to tissue culture and transformation. In individual cells, one or the other of the two *Hyg* transgene alleles might have been inactivated, so that regenerants from single cells behaved as homozygous resistant, hemizygous, or homozygous sensitive, according to their "epiallelic" constitution at the transgene locus. These "epigenotypes" would thus correspond to the apparent genotypes of the DTs. Any plant regenerated from a single cell of line H11 maintained its specific epigenotype, which was also meiotically heritable through several selfed and backcross generations. Examination of methylation of the *Hyg* transgene complex in the parental H11 line revealed that its promoter was partially methylated and became even more methylated in the DTs in which the Hyg^R marker gene activity was decreased. Furthermore, other lines that produced DTs with impaired Hyg^R were also found to have partially methylated transgene promoters. These results are consistent with the presence of the transgene complex in the original line as an epigenetic mosaic resulting from partial methylation.

To the extent that these data with transgenes will be verified with endogenous plant genes (newly developed technology should allow analysis of individual somatic cells (Schindler 1998; Schutze and Lahr 1998)) this would reveal considerable variation among individual somatic cells. This new source of variation, when generated in the meristematic cells of a plant could be of importance, since in plants, some of these meristematic cells have the chance to become gametes.

E. Implications in Breeding.

Recognition that the concept of heredity has to be extended to incorporate epigenetic inheritance systems (EIS) is likely to have major impacts on plant breeding. The theory of selection is based on the existence of heritable variation that affects performance. A theory of variation is therefore a fundamental part of a theory of selection and will determine its efficiency, its limits, and the end result. The present theory is based largely on the assumption that heritable variation is random and involves changes in DNA sequences. If some variation is not based on sequence change but rather is epigenetic, which in addition, is affected by the environmental conditions of plant growth, this must modify and complement breeding theory. By inclusion of inherited epigenetic information as a source of variation, the interpretation of some breeding results could be different, and often simpler.

Conventional breeding is a Mendelian approach in the sense that the genotypic merit is assessed from the phenotypic expression. In other words, genotypic values have to be assessed from phenotypic values, and this requires knowledge of the conditions that ensure the best correspondence between genotype and phenotype.

The degree of correspondence between genotypic and phenotypic values is measured by the coefficient of heredity (h^2), usually expressed as the ratio of genetic over phenotypic variance ($h^2=\sigma_g^2/\sigma_p^2$). The more the coefficient of heritability approaches unit, the closer the correspondence between genotypic and phenotypic values and, therefore, the greater the expected response through phenotypic selection.

For those who are concerned with measuring heritability, the existence of epigenetic inheritance will produce an unwelcome complication. Measures of heritability are notoriously difficult to interpret, particularly because of genotype-environment interactions. Epigenetic inheritance of environmentally induced variation obviously contributes to these interactions, making the distinction between the genetic and environmental components of variance impossible to disentangle, even in theory, and rendering the interpretation of heritability even more obscure.

Heritability values represent average estimates of the population as a whole, whereas breeders are principally interested in knowing the heritability of individual plants and more particularly heritability of the few plants selected to be parents. The problem with single-plant heritabilities is that they require progeny testing to be assessed, meaning that they are not available in the critical generation but only in the generation that follows (Fasoulas, 1993).

Variation of DNA methylation may affect these estimates, since part of this epigenetic variation is inherited. The three principals assumed to interfere with single plant heritabilities, namely: competition, heterozygosity and genotype by environment interaction (Fasoula and Fasoula, 1997a), should be complemented by epigenetic effects on a character. Since epigenetic modifications are affected by environmental conditions of plant growth, they should confound genotype by environment estimates as well (for further discussion on this see Section IV).

The finding that somatic tissue could be mosaic of cells of different epigenetic stages, varying in their level and pattern of DNA methylation, is relevant to plant breeding. The events taking place within somatic cells of plant are specific and concern individual plants rather than populations. Thus selection conditions and/or experimental design that could help these epigenetic states: (a) to be expressed in the phenotype, (b) to be identified and selected, and (c) to be inherited, should increase selection efficiency.

Somatic selection of epialleles may also play a role in creating new variation in clonally propagated plants, homozygous inbred lines, etc. Many plants show clonal variation, but usually the causes of variation are poorly understood: the variation does not behave like classical genetic mutations, yet some is heritable. Breese and colleagues studied somatic variation in various characters in populations of the perennial rye-grass *Lolium perenne* (Breese et al. 1965; Hayward and Breese 1968). They compared sexual and asexual populations, and found that somatic variation occurred in clones of this species; the amount was related to genotype, age of the clone, and environmental conditions. The variation was selectable, with different characters having different heritabilities and transmissibilities. One of the most intriguing results of these studies was the marked differences found within asexual populations. Epigenetic variation followed by somatic selection could explain variability in these asexual populations. Selectable, seemingly non-genetic, somatic variability has been found in other clonal populations. It is likely that epigenetic variation is important in clonal plants. It may enhance the speed of adaptive evolution, especially of specialization. The evolutionary potential of heritable variation that can be selected somatically is enormous. As Silander pointed out:

> " Somatic variation in clonal plants may provide an opportunity for rapid evolution. Somatic effects that are deleterious may impose little or no genetic load on the population or on clonal ramets (Whitham and Slobodchikoff, 1981). The defective organs (or ramets) could simply be shed. Adaptive changes, whether small or saltational, could be perpetuated clonally even if they are not immediately incorporated into the germ line. Many of the somatic effects described for agricultural species represent dramatic changes in form or structure, literally saltational.... Analogous changes in natural populations that are advantageous may spread rapidly and be transmitted in the germ line as well as clonally. The general phenomenon of somatic variation represents an area that has been seriously neglected..." (Silander, 1985).

Similarly, with homozygous inbreds, even for obligate inbreeders, significant variation is frequently detected that could not be explained with the usual low mutation rates. This creates a particular problem especially for preserving for years the homogeneity of a newly developed inbred line after its circulation. The involvement of EIS in generating epigenetic variation as well the role of environmental conditions of growth in inducing such variation in inbreds was mentioned already in describing the data of Cullis with flax inbred lines (Cullis, 1981).

A comparable restricted repertoire of genetic variation has been found in the soybean, *Glycine max* L., an obligate inbreeder (Roth et al. 1989). Although cultivars are inbred and generally homozygous, many different phenotypic and genotypic variants are present. For some genetic markers, there are usually only two alleles and the difference between them is the result of mutation. What is particularly interesting is that in tissue cultures that have been maintained for many generations, the frequent newly generated variation for the markers is always the same as this found in intact plants. This suggests that the same specific, reproducible changes occur in both somatic and germ-line cells. The workers who made this fascinating discovery suggested that the repeated generation of particular alleles may mean that obligate inbreeders such as soybean have evolved mechanisms that are internal generators of genetic variation, possibly as a response to stress. Whether EIS like DNA methylation mediates these changes requires further investigation.

F. DNA Methylation and Genetic Variation.

DNA methylation, in addition to being the cause of epigenetic variation, could be the cause of mutation and generation of genetic variation. Methylated cytosines are hot spots for mutations since 5mC frequently deaminates to T (Coulondre et al. 1978; Jones et al. 1992). In this Section we address the heavy mutational burden induced by methylation of C that could be seen as either a mutagenesis system not requiring the use of exogenous mutagens and occurring in non replicating DNA, or as the price that must be paid for employing a 5mC epigenetic system. The mutability of 5mC was first demonstrated in *E. coli* (Coulondre et al. 1978). Cytosine bases that were methylated in the *E. coli lacI* gene were found to be hot spots for spontaneous base substitution mutations, and the hot spots disappeared when the same sites were unmethylated. It was speculated that the reason for this increase was that, whereas C deaminates to uracil (U), 5mC deaminates to T, which is a normal DNA base and therefore inherently more difficult to repair (Duncan and Miller 1980; Shenoy et al. 1987). In vertebrates, the presence of high levels of CpG methylation was associated with significant deamination of methylcytosine to thymine, a change that was incompletely or inefficiently repaired (Bird 1986; Shen et al. 1994). Thus, where an 5mCpG dinucleotide pair was initially present in a gene, the deamination process would convert this into a TG/CA dinucleotide pair. Presently, mutation at CpG sites continues to play a significant role in the formation of new germ-line mutations contributing to genetic disease. Cooper and Krawczak (1990), in a survey of a wide variety of genetic diseases, found 44 of 139 (32%) point mutations were C to T or G to A transitions occurring at CpG dinucleotides. The isolation of tumor suppressor genes and the detection of mutations within them in somatic cells, has led to the realization that 5mC is a frequent contributor to mutations relevant to human carcinogenesis (Jones et al. 1992).

Over evolutionary time, this deamination reaction may have resulted in the depletion of CpG dinucleotides which, in vertebrates, now occur at only one-fifth their expected frequency (Adams and Eason 1984), a phenomenon known as CpG suppression.

Evidence of the role of methylation in causing CpG depletion is seen by comparison of the genomes of organisms containing 5mC with organisms such as *Drosophila*, which do not methylate their DNA and are not CpG-depleted (Bird, 1986). The entire vertebrate genome, however, is not CpG-depleted, and both 5mC and CpG sites are not randomly distributed. Unmethylated CpG-rich clusters, which make up 1-2% of the vertebrate genome, have been described (McClelland and Ivarie 1982; Cooper et al. 1983). These CpG-rich regions, referred to as "CpG islands," do not show the strong suppression of CpG seen in the rest of the genome and are usually associated with genes or the promoter region of genes (Bird 1986). Importantly, CpG islands are not methylated in the germ line (Bird 1986).

The fact that CpG islands are not methylated in the germ line implies that two forces play a role in their generation: selective pressure, which protects against mutations of CpG sites in vital genes, and lack of methylation in the germ line, which relieves CpG islands of the mutagenicity of 5mC. Thus, CpG islands are not formed by the creation of CpG-rich regions, but rather are formed by the loss of CpG sites in the bulk of the genome, leaving a non-CpG-depleted region. The key to maintenance of CpG islands is their protection from methylation in the germ line, which can be carried out in two possible ways. Either CpG islands are protected from *de novo* methylation, or CpG islands are actively demethylated. Evidence exists for both mechanisms (Frank et al. 1990; Szyf et al. 1990).

Plant DNA is considerably more methylated than vertebrate DNA. Thus, in the absence of selection pressure, plant genomes with high levels of C-methylation are thought to shift to lower G+C contents. But, despite the evidence that indeed such a CpG suppression is operating in plants, the considerably higher amount of methylcytosine in higher plant genomes has not been accompanied by such an analogous high level of CpG or CpNpG suppression (Gardiner-Garden et al. 1992; Montero et al. 1992). In tomato for example, if the lower

G+C content is the result of previous transitions from 5mC to T one would predict a lower G+C content in non-coding regions compared with coding regions, since the former should be under relaxed selection pressure. To test this hypothesis, Messeguer et al. (1991) analyzed the base composition of published sequences for coding as well as non-coding regions and repetitive sequences from tomato. Averaged over all sequences, the non-coding regions were found to have a G+C content of 32%, which is lower than the overall G+C content of the tomato genome (37%) and much lower than that found in coding regions (46%). The hypothesis that deamination of 5mC leads to a lower G+C content yields another prediction, namely that there should be excesses and deficiencies of certain di- and tri-nucleotides caused by 5mC to T transitions. Specifically TpG, CpA, TpApG, CpTpA, TpTpG and CpApA should be in excess and CpG, CpApG and CpTpG should be deficient. These predictions were also tested based on published sequences. Many of these prediction were borne out and add general support to the hypothesis. These results with tomato fit predictions that 5mC has been subjected to deamination and transitions in both mCpG and mCpNpG sites in regions of the genome under relaxed selection pressure (i.e., non-coding regions). But for coding regions, the results were different. Although CpG was also deficient in these regions, neither CpApG nor CpTpG were deficient. The use of these tri-nucleotides as codons (CpApG and CpTpG code for glutamine and leucine, respectively) may explain the higher frequency in coding versus non-coding regions. McClelland and Ivarie (1982) analyzed published gene sequences from legumes and they also observed a reduction of CpG, but not CpNpG sites. Based on these results, they suggested that there is some mechanism either preventing or significantly reducing CpNpG transitions, or that CpNpG methylation was so new to plants (evolutionarily) that not enough time had elapsed to allow transitions to accumulate. But, Belanger and Hepburn (1990) presented evidence discussed already in Section II A, that CpNpG methylation is as old as CpG methylation, eliminating in this way the argument that plants had not enough time to allow transitions in the CpNpG sites to accumulate, and pointing to the existence of a much more effective mechanism of repairing these sequences in plants. Such a mechanism could be a more effective mismatch repair system in plants. Employing a more efficient mismatch system by plants is important because plants do not have the well-defined germ-line tissues characteristic of vertebrates, where essential sequences can be protected from the consequences of deamination by being maintained in an unmethylated state (Walbot and Cullis 1985; Bird 1986).

In discussing the role of plant DNA methylation in relation to somaclonal variation, we have made note of the data indicating that the high frequency of somaclonal variation can be mainly accounted for by pervasive, preexisting epigenetic variability in somatic tissues used as explant sources, and stable maintenance of a given epimutation in the regenerated plants. However, there is still a role for tissue-culture-induced changes, both genetic and epigenetic. Phillips et al. (1994) attempted to explain the unusually high incidence of somaclonal variation by assigning a central mutagenic role to tissue-culture-induced changes in DNA methylation, considering a role for methylation-induced genetic changes in somaclonal variation because 5-methylcytosine can spontaneously deaminate to produce T. Moreover, as stressed by Matzke and Matzke (1996) the C methylation pathway is inherently mutagenic, particularly under conditions in which the methyl donor S-adenosyl methionine (SAM) is limiting. Such an environment permits the accumulation of an intermediate in the C methylation pathway (5,6-dihydrocytosine), which has a 10^4-fold higher rate deamination than 5-methylcytosine (Matzke and Matzke 1996). In vitro culture might subject plant cells to SAM-deficient conditions, leading to increased frequencies of C-G to A-T transitions via rapid deamination of 5,6-dihydrocytosine. Thus a possible cost of maintaining the high content of 5-methylcytosine in plant genomes during plant cell culture might be an elevated rate of mutation, which then contributes to somaclonal variation of regenerated plants.

In conclusion, deamination of 5mC leading to point mutation is taking place in plants despite its lower rates. Its contribution in generating new genetic variation require further experimentation, especially in cases like clonally propagated plant species, homozygous obligate inbreeders, etc.

Thus, on the one hand, a methylated cytosine is readily deaminated to thymine, so methylated CpG sites are hot spots for point mutations caused by C to T transitions. On the other hand, CpG methylation participates in the organization of chromatin into an inactive "closed" conformation, which is less transcribed, less accessible to mutagenic processes than active chromatin, relatively protected from rearrangements and transpositions. Hence, in organisms that methylate their DNA, the quantity and distribution of methylated CpGs may be the result of a compromise between the mutability of methylated cytosines, and the protection afforded by methylated DNA from change in sequence organization through recombination and transposition.

The way that chromatin structure can influence the probability of changes in DNA sequence highlights the complex role of the environment in evolutionary change. The environment is not just the agent of selection in plants. Through its effects on the gene's phenotype, it also biases the direction, rate, and type of DNA changes at a locus. Consequently, the frequency of mutation and the frequency of recombination are not independent.

To what extent is mutagenesis of 5mC regulated? Deamination of 5mC is considered to be a randomly occurring event but deeper examination of this phenomenon reveals otherwise. In *Neurospora*, for example, a phenomenon called RIP (Repeat Induced Point mutation) refers specifically to a process that riddles duplicated sequences with C-G to A-T transition mutations in haploid genomes of special dikaryotic cells formed during the sexual cycle of the fungus (Selker 1997). Sequences altered by RIP greatly accelerate *de novo* cytosine

methylation (Selker, 1997). RIP greatly accelerates the mutation rate in a regulated way, explaining the high incidence of somaclonal variation by the unleashing of a RIP-like process during the stress of tissue culture (Phillips et al. 1994). There has been some discussion, though, about whether RIP or a related process operates in vertebrates (Kricker et al. 1992) or plants (Matzke et al. 1994).

IV. DNA METHYLATION , HETEROSIS AND F_1 HYBRID BREEDING

A. Genome Function and Hybrid Vigor

In breeding the terms hybrid vigor and heterosis are frequently used inter-changeably, sometime resulting in confusion. Hybrid vigor refers to a generally better performance of F_1 plants, growing more vigorously than their parents, without the necessity to express their vigor quantitatively as deviation from the parents. Heterosis is a quantitative term with parental performance (mid-parental value or best-parent value) always taken into consideration. A contemporary F_1 hybrid for instance could grow vigorously giving high yield but simultaneously could be of low heterosis, if its parents are also high yielding inbred lines. Most of the today's hybrids, even for maize, are of this type. They have high yield but their heterosis is low, on a percentage basis in comparison with past hybrids. Conversely, a hybrid with moderate vigor and yield could be quantified as highly heterotic if its parents were bad performers. F_1 hybrids of maize in the past were of this type. Constant breeding in recent years resulted in more vigorous, high yielding hybrids surpassing the old ones, while their heterosis has steadily decreased (Duvick 1992).

Breeders have made wide use of hybrid vigor. Evidence for the manifestation of hybrid vigor appears in hundreds of millions of hectares of field and vegetable crops throughout the world. The magnitude of hybrid vigor in F_1 hybrids is one of the primary reasons for the success of the commercial maize industry.

Production of hybrid seed-propagated crops is a successful breeding scheme not only because it exploits vigor but also because it promotes uniformity and homogeneity in cross-pollinating species and is a way for commercial breeders to control their product (Duvick 1984; Janick 1998). Uniformity is one of the principal benefits of hybrids for three reasons: product uniformity is essential in marketing, particularly in horticultural species, uniformity in maturity permits crop scheduling, and uniformity in plant structure and maturation permits efficient mechanical harvest. With the increasing importance of urban markets, product uniformity is an essential feature of crop quality. Genetic homogeneity refers to uniform populations as a result of identical genotypes. In the typical crop environment, in farmer's fields, a monogenotypic cultivar like F_1 hybrids (the same is also true for other monogenotypic cultivars like inbred lines and clones) is grown in self-competition also called iso-competition (Fasoulas and Tsaftaris 1975; Donald and Hamblin 1976; Fasoulas 1993). This results in plant-to-plant interference among the genetically identical plants for exclusive use of the environmental resources, that are available in shorter and shorter supply, as the density increases, and the growth advances. Yield per unit area is maximized under zero mutual interference for equal use of resources among plants. This is the ideal condition where resources are shared equally by all plants and consequently, the yield of all plants is evenly suppressed. A first prerequisite for approximating this condition is the use of monogenotypic cultivars with ample and evenly distributed growth resources, and secondly elimination of acquired differences among plants. Failure to approximate this ideal condition supports unequal sharing of resources, uneven suppression of yield, undercompensation by the plants with high yield for the yield loss of other plants and, finally, reduction of yield per unit area (Fasoula and Fasoula 1997a).

In seed propagated, cross-pollinated species where open-pollinated populations consist of a mixture of genotypes, genetic homogeneity can be achieved by inbreeding to create inbred or pure lines followed by intercrossing divergent inbred genotypes with high homozygosity while selecting within and between inbred lines for adaptability, vigor, and combining ability (Janick 1998).

In addition, farmers prefer F_1 hybrids for their high but most importantly stable yield while growing in different fields and in different years. Genetic stability (homeostasis) refers to reduced genotype x environment interaction, particularly under stress. The selection of adapted single crosses with high yield and high stability has been responsible for the consistent increase in maize yields in the United States since their introduction in the early 1960s (Duvick 1992, 1997 ; Janick 1998).

However, while agronomists and subsequently farmers have been utilizing hybrid vigor, its biological basis remains unknown. Our understanding of hybrid vigor will enhance our ability to form new genotypes that may be used directly as F_1 hybrids or form the basis for future selection programs for possibly fixing in homozygous conditions part of the additively acting interlocusly dominant alleles that contribute to F_1 hybrid vigor (Fasoulas 1993; Tsaftaris 1995; Xiao et al. 1995).

The genetic theories for heterosis (which have changed little since 1952) include: (1) dominance, and (2) true overdominance, which is nearly impossible to distinguish from pseudo-overdominance (i.e., nearby loci at which alleles having dominant or partially dominant advantageous effects are in repulsion phase linkage). Certain types of epistatic interactions between interacting loci were included, too. Elucidation of these theories

remains a major challenge. Biometrical approaches can only evaluate average (or net) genetic effects on heterosis while studies that have used biochemical, physiological or molecular approaches may provide some limited insight into a better understanding of this phenomenon. Even though the physiological and/or biochemical responses reported may only represent manifestations of heterosis at a level other than at the overall mature plant stage (as measured in the field), these responses should be closer to the gene level and may ultimately help to reach the genetic basis of heterosis.

Several investigators have estimated the correlation between isozyme allelic diversity and hybrid vigor (Tsaftaris 1995) but the number of isozyme markers has been too small for effectively marking the plant genome. With molecular markers, the possibility of drastically increasing the number of markers has been ralized (Stuber 1995; Tsaftaris 1995; Coors 1998).

The morphological, isozyme and molecular markers provided a repertory of neutral markers distributed throughout the genome. They can substitute for kinship coefficients to study the correlation between parental divergence and F_1 performance, but they provide no information about the genes and the molecular and physiological mechanisms involved for the phenotypic expression of complicated characters, such as yield or yield heterosis. Only indirectly, through their linkage to relevant QTLs, such as those controlling yield or yield heterosis, can these markers be associated with quantitative characters.

Parameters derived from the variability of genome expression could provide a source of non-neutral markers for different phenotypes including hybrid vigor. It has been suggested that molecular foundations of phenotypic diversity could reside in the variability of genome expression. Some authors have even proposed that the quantitative difference in patterns of gene expression during development lies at the heart of organismal evolution (Atchley 1989). The question then arises whether or not quantitative variation in gene expression also plays a role in the phenotypic diversity of the different genotypes of a plant species, its relation to variation in different biochemical/physiological processes, and finally its relation to hybrid vigor. Variability in gene expression can be assessed through polymorphism of individual RNA amount (RAP), which is consistently detected when comparing genotypes by northern analysis of RNA or by dot or slot blot hybridization of RNA, and through the polymorphism of individual protein amount (PAP) which is consistently detected when comparing genotypes by two-dimensional PAGE of denatured proteins or enzyme assays.

The advantages of using PAP have been enumerated by Damerval et al. (1987), and Leonardi et al. (1991) who showed that among five maize lines PAP was related to a Mahalanobis distance computed from general combining abilities of 14 morphological and developmental traits. They also reported non-additive inheritance for protein amounts in hybrids of maize (Leonardi et al. 1987, 1988). Later, in a study by their group, 28 maize single-cross hybrids of a diallel design between eight parental lines were characterized for agromorphological performances in four different environments (Leonardi et al. 1991). Numerous significant correlations between PAP indices and hybrid vigor for agronomic traits were found.

The significant correlation between PAP and hybrid values led these researchers to conclude that loci controlling protein amounts would themselves be QTLs. Moreover, the PAP index is correlated to the number of non-additive situations, which appeared also to be related to hybrid values. This correlation cannot, in any case, be explained in terms of linkage disequilibrium. The authors suggested that the correlation between PAP and performances of the hybrids observed in three different experiments meant that loci for protein level are themselves QTLs. These results suggest that genes controlling protein amounts and enzyme activities, and particularly those with multiple effects, directly affect the expression of hybrid vigor. Basically, such a conclusion is almost a necessary consequence of the role of the regulation of gene expression in development. Genetic variation in the profile of gene expression can hardly be without any phenotypic effect and certainly constitutes a non-neutral polymorphism. Recently a methodology was devised (Damerval et al. 1994) to dissect the genetic architecture of quantitative variation. Damerval et al. (1994) used the procedures for QTL mapping to analyze the genetic determinism of the quantitative variation of 72 anonymous proteins separated by high resolution two-dimensional polyacrylamide gel electrophoresis in an F_2 population of maize. Recently Mitchell-Olds and Pederson (1998) working with arabidopsis have identified a number of such QTLs controlling ten enzymes in central and secondary metabolism.

Results obtained for polymorphism of individual RNA amounts (RAP) were similar to those of PAP. In one such RAP analysis (Tsaftaris and Polidoros 1993), expression of many genes in different maize tissues and developmental stages was measured. As genetic material three maize parental inbreds (B73, H108 and H109) and two of their hybrids the highly heterotic F_1 H109 X B73 (HHF_1) and the lowly heterotic F_1 H109 X H108 (LHF_1) were used. In accordance with the protein data, the qualitative analysis with RAP gave similar results. Significant quantitative differences in the amounts of the individual RNAs for the tested genes were found among the five genotypes in every developmental stage (Fig. 5). This quantitative variation in the amounts of individual RNAs was significantly correlated with agronomic performance. The HHF_1 exceeded the LHF_1 as well as the inbreds in genome activity (average of the mRNA quantities for the tested genes) at three developmental stages. In a more detailed analysis of the data, the deviation in quantity of each RNA of every one of the two hybrids from the quantity of the same mRNA in the two parental hybrids from the quantity of the same developmental stage was examined. The heterotic hybrid had a significant number of genes expressed over the quantity of the better parent at three stages. Genes belonging to this category were the 30% of the tested

genes at the first stage and reached 63% at the third stage. The non-heterotic hybrid had a minor number of genes expressed over the better parent at the first and second stages, representing no more than 15% of the tested genes, and reached 57% at the third stage. In addition to the above, also for the LHF_1 hybrid, a significant number of genes (28% of the tested genes) expressed below the level of the worst parent. Thus, genes of the two hybrids are differentially expressed during development. These data are in accordance with those of Romagnoli et al. (1990), who also conducted a RAP analysis in one heterotic hybrid in comparison with it's parental inbreds. They showed differential expression of many genes, since approximately 33% of the major proteins of the hybrid, translated in vitro from isolated mRNA, were more abundant or possibly new, in relation to parental mRNA. Thus, they also suggested that increased expression of certain loci may be important in manifestation of vigor.

The above data with PAP, enzyme amounts, and RAP indicate that quantitative variation in the expression of certain loci may be important in vigor manifestation, and underline again the significance of regulatory mechanisms involved in the quantitative modulation of gene expression in manifesting vigor. It remains then for the regulatory mechanisms responsible for this extended variation in the amounts of mRNAs or proteins to be analyzed, as well as their mode of action to be defined.

B. Genome Methylation and Hybrid Vigor

DNA methylation could be considered as a genome-wide general regulatory mechanism that affects the expression of many genes important for the manifestation of heterosis (Tsaftaris and Kafka 1998; Tsaftaris et al. 1998). A model implicating DNA methylation as a molecular regulator of heterosis was presented by R. L. Phillips (at the Workshop on Molecular Basis of Heterosis in Plants, March 19-20, 1991, University of Minnesota) and cited by Stuber (1998).

To obtain more evidence for the possible involvement of methylation, Tsaftaris and Polidoros (1993) used the RAP analysis described above to obtain a gross estimation of genome activity, and HPLC chromatography to estimate DNA methylation. They found a significant negative correlation between genome activity and total DNA methylation (Fig. 7).

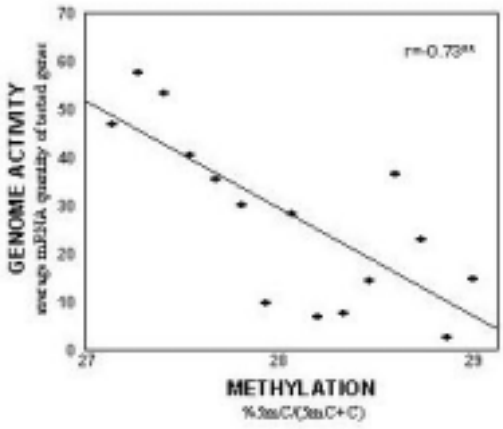

Fig. 7. Linear regression of genome activity estimated as average mRNA quantity for 35 cloned genes on total DNA methylation, for the five genotypes of the first group in three developmental stages.

Isoschizomeric restriction enzymes sensitive and insensitive to methylation also were used to analyze DNA methylation in different maize inbreds of different pedigrees, and hybrids, revealing differences in total methylation between inbreds and hybrids (Tsaftaris and Kafka 1997; Tsaftaris et al. 1998). Estimations of DNA methylation of ten S_2 second cycle lines examined in conjunction with DNA methylation of the original progenitor parental lines and their F_1 hybrids are shown in Table 4. The FS68 parental line (a low yielding inbred) showed 31.4% methylation, one of the highest values observed. The second parental line showed 28.3% methylation and their F_1 hybrid 27.4%. Thus the hybrid had a lower percentage of methylation than either parent did. The ten S_2 high yielding lines had an average methylation of 24.9%. Their methylation rate varied from 23.8% (line 2B) to 25.8% (line 4A). These values are significantly lower, not only from percent methylation of their original progenitor parental lines FS68 and NE2, but also from methylation of the progenitor F_1, indicating that breeding the original F_2 population for higher yield for five continuous generations (Koutsika-Sotiriou et al. 1990), using the mass honeycomb selection design, was accompanied with concomitant decrease of the methylation level in the derived S_2 second cycle lines.

These studies paralleled studies using the CRED-RA technique for monitoring the pattern of methylation in certain sites of DNA in maize parental inbreds and their hybrids (Tsaftaris et al. 1997, 1998). Like the results of isoschizomeric enzymes and HPLC (concerning total methylation), the CRED-RA analysis

indicated that: (1) hybrids in general are less methylated than their parental inbreds; (2) heterotic hybrids are less methylated than related non-heterotic hybrids; (3) old, low-yielding inbreds are highly methylated; (4) more modern inbreds, especially those selected for high and stable yield under spacing in the isolation environment, have lower percentages of methylation in comparison with old progenitor lines. These findings support the hypotheses made by several researchers (reviewed by Stuber, 1998) that systematic selfing for isolation of inbreds, with emphasis only on combining ability of inbreds (leading to line inbreeding depression), is also leading to a gradual accumulation of more methylated sites, which then could be released and/or repatterned when the selfed lines are crossed to generate hybrids. Shifting emphasis to line performance per se during selection, though, could moderate line inbreeding depression. While approaching homozygosity, degenerative alleles, (possibly methylated and/or even mutated genes) in homozygous inbreds could be constantly selected out. The proven higher efficiency of this kind of selection, in inbreds approaching homozygosity, growing in the isolation environment could be correlated with the lower methylation levels of plants grown under this condition (Tsaftaris et al. 1997, 1998).

Table 4. %5mC content measured by HPLC analysis in leaves of two old low yielding parental lines, their F_1 hybrid and ten new second cycle S_2 selected maize inbred lines of similar pedigree derived from the hybrid.

Genotype	5mC content
FS68	31.4
NE2	28.3
FS68 X NE2	27.4
Selected S_2 lines	
2B	23.8
4E	24.1
2D	24.3
5D	24.3
3A	24.7
2F	25.4
3E	25.4
5E	25.6
6AT	25.6
4A	25.8

C. Genome Function, DNA Methylation and Performance of F_1 Hybrids.

F_1 hybrids, in addition to their homogeneity and high yield, are preferred for their stability of performance. Stability of performance refers to reduced genotype X environmental interaction. This stability of performance of hybrid plants has been particularly obvious when the plants are facing different kinds of stress. Duvick (1997) compared 36 widely grown successful hybrids released at intervals from 1934 to 1991. During this period, genetic yielding ability of these hybrids has increased at an annual linear rate of about 74 kg/ha. The hybrid series showed continuing improvement in tolerance to stresses, particularly being more efficient in grain production under high density planting. More recent data implicating heterosis as a form of stress tolerance were presented by Elings (1998) and Kang (1998). These results agree with the hypothesis that increased grain yielding ability and stability of performance of widely successful maize hybrids is due primarily to improved tolerance to stresses, particularly high density planting.

Production of contemporary hybrids with high and stable yield was made possible by efficient multiyear and multisite testing of the genetic material. Different years and different sites imposed different kind of stresses (water, nutritional, light, pH, temperature, etc.) on the tested material. The concomitant tolerance of newer hybrids to higher and higher plant densities in the crop conditions of farmers indicates that density induced stress in the crop environment consists of an integrated form of all the above partial stresses for different environmental parameters. G X E interactions (that are considered a significant factor interfering with heritability of selected plants, thus efficiency of selection) actually reflect the reaction of genotypes to different

stresses imposed by growing the genotype in varying sites for years. Thus multiyear and multisite selection of the genetic material was found indispensable for consideration of G x E interaction and breeding for stability of performance.

What is the role of methylation in a hybrid's genomic DNA, especially in tolerating stress, leading to more stable yield? Data have been presented showing that F_1 hybrids are in general less methylated than their parental inbreds. The possible role of methylation in the expression of maize genes and performance of hybrids under different growth conditions has been examined in experiments with maize inbreds and hybrids grown under different plant densities (Tsaftaris and Kafka 1998; Tsaftaris et al. 1998). Genotypes growing under wide spacing (1.5 m) in the isolation environment are considered to be growing under less stress in comparison to those of high density planting (0.3 m). Total genome methylation using HPLC was measured. The data concern a set of inbreds and two of their most productive hybrids at two developmental stages, an early stage when the plants were small and the density-induced stress non-existent and a later stage when the density induced stress was greatest. In late stages when the effect of density was fully imposed (Table 5), methylation of inbreds was significantly higher (29.2%) in the dense conditions in comparison to spaced plants (27.5%). The situation was different in the hybrids. Hybrids showed exactly the same methylation in dense and in spaced plantings (26.5% vs. 26.2%). This indicates that, hybrids are more tolerant than inbreds to density-induced increases in methylation. Similar results were obtained using the CRED-RA analysis targeted to methylation of specific sites. Hybrids were less methylated than inbreds, genotypes were found to be more methylated under the high-density stressful conditions, and, most importantly, hybrids remained less methylated even under higher density in comparison to parental inbreds.

Table 5. 5mC content measured in stage B (60 days after planting) under two different conditions of growth in the field in five S_4 lines selected from the F_2 of the commercial hybrid 'LORENA' and two of their hybrids. Two more lines: the old NE_2 and an S_2 line (see Table 4) were also included for comparison (Tsaftaris et al.1997; 1998).

	%5mC content at STAGE B	
GENOTYPES	**SPACED**	**DENSE**
Inbreds		
22M	26.2	22.4
30B	25.9	27.3
102A	24.2	28.7
5C	29.8	30.1
NE2	29.3	30.6
4D	29.5	32.4
60M	27.9	32.9
Average	27.5	29.2
Hybrids		
22M X 102A	26.2	26.4
60M X 30B	26.1	26.5
Average	26.1	26.4

We can conclude that the genotype, the developmental stage, and conditions of growth affect the methylation status of genomic DNA. Selection could change the level of methylation. The uniform low methylation of all S_2 high yielding second cycle lines tested in this study suggests that improvement of line performance per se is related to a decrease of methylation after five continuous generations of selection in wide spacing of individual plants. On the other hand, methylation of the ten S_2 lines of similar pedigree was lower and varied less than that of the seven commercial inbreds of different pedigrees used for years in the past for hybrid seed production. A systematic evaluation of methylation in inbred lines of different eras in the history of maize, particularly today's inbreds bred for line yield per se, always in conjunction with their yield and combining ability, will be highly informative.

Stressful growth conditions result in more methylated DNA (less expressed) and, in general, vigorous hybrids are more resistant to such density induced methylation and suppression of genome activity in their genomic DNA. This resistance of the hybrid genome to genome methylation under different stresses, and consequently, avoidance of suppression of many of its genes could be at the core of high F_1 yield and, maybe

more importantly, F_1 stable yield. Taking advantage of existing criteria for evaluating stability of performance of different hybrids (Fasoula and Fasoula 1997b) experiments are underway for a systematic study of possible correlation between stability of performance and methylation changes in the genotype (A. S. Tsaftaris and colleagues, unpublished data). If such correlation exists then time and money could be saved from multiyear and multisite evaluation of the genetic material, since criteria and tests could be devised for fast evaluation of genotypes for tolerance to density induced integrated stress, securing later stability of performance. As emphasis in line selection has been shifting to more productive lines that could generate high yielding hybrids with stable yield in higher densities (substituting double and triple hybrids with single F_1 seeds), there was a concomitant shift in parental lines giving less heterosis, less inbreeding depression, with less methylation. What will be the result if this trend with crossbreeding species such as maize will continue (or even be intensified by selecting for line performance per se at the isolation environment that promotes demethylation of many genes)?

To what extent epigenetic phenomena contribute to hybrid vigor will determine the answer to the above question. Bringing for instance in homozygous conditions, in an inbred, many gene pairs currently in repulsion phase linkage (which is apparently mostly responsible for pseudo-overdominant effects on hybrid vigor) will require time. Thus, in the meantime, exploiting them in heterozygous conditions in F_1 hybrids will continue to be a fast solution to the pressing and competitive demands for more productive crops. On the contrary, if epigenetic changes like DNA methylation also contribute to hybrid vigor, transition to more productive inbreds will be faster since demethylation of different genes is not generally related to their linkage. Selection for line per se, is what nature and breeders are doing with crops that are obligate inbreeders. Despite their homozygosity, high yielding inbred lines developed in these species are widely cultivated.

V. DNA METHYLATION AND TRANSGENE INACTIVATION IN TRANSGENIC PLANTS

Directed genetic manipulation for breeding plants, leading to production of transgenic plants, is fast becoming routine for a number of crops. However, such transgenic crop plants will only be of value if their phenotype is stable in the field and transmitted faithfully in subsequent generations. Gene silencing in transgenic plants by methylation has been mentioned many times in this review. Transgene inactivation generated a more intense interest for this important basic biological phenomenon. Gene silencing has come as an unwelcome surprise. Early reviews of the prospects for plant genetic engineering did not identify this as a potential obstacle. Rather, the anticipated challenge was to identify tissue and stage-specific promoters that could be used to obtain regulated transgene expression. Yet, silencing of transgenes is turning out to be a substantial problem as according to Finnegan and McElroy (1994) of 30 companies polled, nearly all reported some problems with unwanted silencing of transgenes.

It has been shown that this unwelcome inactivation of the transgene mediated by methylation, is triggered by stresses including the common agronomic practice of seedling transplantation in the field (Brandle et al. 1995). Thus, it requires closer attention since many crops need transplantation in the field and this imposes a severe stress for the young plantlet. It is clear that some steps that were taken for granted may need to be further investigated for successful commercialization of transgenic crops.

In this Section, we focus on transgene-induced gene silencing of both, the transgene, and certain cases of other endogenous sequences, in plants. The different molecular mechanisms involved will be described separately. Considering the step in gene expression where inactivation occurs, it can be transcriptional or post-transcriptional.

Transcriptional silencing will be divided in transcriptional inactivation of single transgenes through methylation and, unidirectional *trans*-inactivation of other homologous allelic genes, either (allelic unidirectional *trans*-inactivation) or non-allelic (epistatic unidirectional *trans*-inactivation). Epistatic *trans*-inactivation could refer to cis-interaction of linked homologous sequences or to *trans*-interaction of homologous sequences located on different chromosomes (Vaucheret et al. 1998).

The different transgene silencing systems share striking similarities with basic biological phenomena such as paramutation, position effects, changes in chromatin structure, repeat induced gene silencing, and may reflect natural mechanisms of plant protection against transposable elements or viral pathogens (Matzke and Matzke 1998). Understanding the role of DNA methylation in different transgene silencing systems will help comprehension of these, important for plant breeding, phenomena.

A. Transcriptional Inactivation of Single Transgenes.

Transgene silencing through DNA methylation has been thought to involve multiple copies of a transgene, or somehow homologous DNA sequences in the same genome. However, there are cases where individual transgenes, with no apparent copies and/or homologous sequences, become inactivated when inserted into plant chromosomes (Meyer et al. 1992). This could be expected if the transgene "landed" in a chromosomal area

where the chromatin is heavily methylated, condensed and thus inactive. In other cases, transgene silencing occurs specifically and only in the cloned DNA of transgenes that "landed" in actively transcribed chromatin.

Early studies on the activity of T-DNA genes in crown gall tumor lines revealed a correlation between DNA methylation and inactivity of T-DNA genes in certain lines (John and Amasino, 1989). Apparently, T-DNA that integrates randomly into the genome becomes methylated at certain integration sites. The fact that certain copies of a transgene became hypermethylated while others remained hypomethylated and transcriptionally active, strongly suggested that the degree of DNA methylation may be influenced by the integration region. Occasionally this position-specific inactivation was found to revert either spontaneously or after treatment of the cell line with the demethylation agent 5-azacytidine (Van Slogteren et al. 1984; John and Amasino 1989).

The best-studied case of silencing of single transgenes is the inactivation of petunia transgenic plants carrying the *A1* gene from maize. To study position-dependent effects for *de novo* methylation of the *A1* gene, Prols and Meyer (1992) analyzed single-copy transformants with different *A1* expression levels. They found that when the transgene had been integrated into a highly repetitive and hypermethylated region, it was inactive. Apparently the methylation state of the integration region was, at least to some extent, responsible for position-dependent differences in gene expression.

In another line the transgene remained hypomethylated and active in the majority of plants that were propagated in the greenhouse. Hypermethylation was significantly enhanced when plants were grown in the field, which allowed selection of epigenetic variants containing the same transgene in different methylation states (Meyer et al. 1992). A detailed analysis of the transgene and its chromosomal integration region in these epigenetic variants, and also, in a derivative of the original line that carries a deletion in the *A1* coding region, showed that hypermethylation was limited to the transgene DNA only, while the hypomethylation state of the integration region remained unaltered (Meyer and Heidmann 1994). This might suggest the presence of a DNA methylation mechanism that specifically recognizes foreign DNA. This mechanism, which has already been proposed for animal systems (Bestor 1990; Doerfler 1991) might identify foreign DNA by its nucleotide composition. Nuclear genomes of angiosperms are mosaics of long, compositionally homogeneous DNA segments called isochores (Salinas et al. 1988). Isochores contain very defined GC content of functional genes and their chromosomal environment (Matassi et al. 1989). The *A1* transgene differs significantly in AT content from the chromosomal environment of the host. The integrated DNA has an average AT content of 47.5%, while the neighboring 269 bp at the 5' end and 196 bp at the 3' end show a much higher average AT content of 74% and 77%, respectively. It is, therefore, conceivable that the isochore composition of a transgene has to match that of its integration region to avoid specific methylation of the transgene.

Alternatively, it might not be the nucleotide composition but rather the transcriptional activity, which makes the transgene a specific target for DNA methylation. Promoter activity might open the local chromatin structure, thus providing access in this region for cellular proteins involved in DNA methylation or heterochromatin formation. So far, there is no data about DNA methylation of promoterless transgenes. For the future it will be interesting to compare DNA methylation patterns in transgenes embedded into different isochore regions and differing in nucleotide composition, secondary structure and promoter content.

B. Transgenes Inactivate Transcription of Other Homologous Genes.

1. Transgene Inactivates Unidirectonally the Transcription of its Allelic Gene. Several phenomena have been described in higher eucaryotes that represent an exception to the common observation that two alleles are expressed independently of each other. In *Drosophila* transvection effects imply the ability of regulatory elements on one chromosome to affect the expression of the homologous gene in a somatically paired chromosome (Wu and Goldberg 1989). The brown gene and some extreme derivatives of the white gene behave like dominant mutations in *Drosophila*, causing *trans*-inactivation of its homologous alleles. Both genes are known to undergo position-effect variegation (Heinkoff and Dreasen 1989).

In plants, the term paramutation has been coined (Brink, 1956) to describe a heritable change in gene function directed by its allele (Kermicle 1996; Martienssen 1996; Richards 1997). Paramutation is a regularly occurring, directed, and heritable alteration of gene expression resulting from the interaction of two alleles. The expression of the one allele, referred to as paramutable, changes following paramutation. The other allele, referred to as paramutagenic, causes the change in the paramutable allele. When the paramutable/paramutagenic heterozygote is crossed to allow segregation of the two alleles, virtually 100% of the paramutable alleles transmitted display a decrease in expression. This decreased expression level (the paramutant phenotype) persists through many generation. Paramutation was first described in 1956, but mechanisms to explain how one allele can heritably affect the expression of another have remained obscure.

One of the best studied examples of paramutation is the *r1* locus of maize (Brink 1973; Kermicle 1996; Martienssen 1996; Richards 1997). The *r1* genes encode helix-loop-helix proteins that are capable of directing transcription of the structural genes in the anthocyanin biosynthetic pathway. Thus, dominant *R1* genes conferred a purple anthocyanin pigmentation to the tissues in which they were expressed. The paramutant *R-r* complex is referred to as *R-r'*, to distinguish it from nonparamutant *R-r*. The effect on aleurone pigmentation is only

observed if *R-r'* is transmitted through male gametes. When *R-r'* is transmitted through the female, the normal dark uniform pigmentation pattern is observed. This dependence on the mode of transmission is not due to a simple dosage effect; two paternally transmitted copies of *R-r'* still confer a paramutant phenotype in aleurone (Kermicle 1970). Nonparamutant *R-r* also exhibits mild silencing on male transmission. *R-r'* is characterized as being metastable; under certain conditions, it reverts toward the standard full color phenotype. The paramutable *r1* complex, *r'r*, consists of four components of the *rl* transcription unit (Walker et al. 1995). These four copies, referred to as components, are designated *p*, *q*, *S1* and *S2*. The *S1* and *S2* components are complete genes and are responsible for pigmentation of the aleurons layer of the seed. The *q* component is nonfunctional because it lacks downstream coding sequences. The *p* component is the third complete *r1* gene, and is active in several tissues in the plant including the coleoptile, the roots, and the anther walls. The three complete genes of the complex are not equally subject to paramutation: the *S* genes (*S1* and *S2*) show a large decrease in the amount of pigment they confer following paramutation, whole the *P* gene shows a relatively small decrease following heterozygosity with a paramutagenic *rl* complex (Brink and Mikula 1958; Brown 1966).

The *S* genes are arranged in an inverted head-to-head orientation and are separated by a rearranged remnant of a transposable element of the CACTA family (Cone et al. 1993) called *doppia* (Walker et al. 1995). At its ends, a sequence called *σ* contains partial copies of the terminal inverted repeats of the *doppia* element. Internal to these inverted repeats are multiple copies of a subterminal repeated element of *doppia* and a second region consisting of sequences that have undergone extensive rearrangement and that may or may not have been derived from *doppia*. The • region functions as the promoter for both the *S1* and *S2* genes and *R-r*. The finding that a mutant without • has lost most of its ability to acquire secondary paramutagenicity implies that the • region has a role in causing paramutability of *R-r* (Kermicle 1996).

Finally, evidence has been recently obtained that *r1* paramutation is associated with changes in the level of C-methylation of the paramutable *R-r* complex (Martienssen 1996; Richards, 1997). A mutant *r-r*: N13-1 allele that has a deletion of most of the *doppia*-containing • region, but retains the multiple homologous genetic elements of the *R-r* complex (i.e., the *p*, *q*, *S1* and *S2* genes) failed to become methylated (in a context that elicits strong C-methylation of intact *R-r*). These data suggests that the • region is needed for induction of C-methylation during paramutation while the duplicated gene segments at *R-r* do not themselves trigger C-methylation.

A variety of paramutation and paramutation-like phenomena has been observed that involve other endogenous genes. In maize, two other genes, *b1* and *pl1*, are known to be subject to paramutation, but since their expression was not found to involve DNA methylation, are not discussed further in this review (Coe 1966; Hollick et al. 1994; Patterson and Chandler 1995).

A striking similarity has been observed between paramutation of the *R* locus of maize and the behavior of *A1* allele in certain transgenic petunia plants (Meyer, 1996). The transgenic petunia line 17-R contains one copy of the maize *A1* gene, which mediates brick-red palargonidin pigmentation of the flower. A white derivative, 17-W, was isolated from homozygous progeny of this line in which no pelargonidin pigmentation was observed. In 17-W the 35S promoter driving the *A1* gene was hypermethylated, in contrast to its hypomethylated state in 17-R. Progeny plants carrying both the 17-R and 17-W allele did not show the expected *A1* phenotype. Predominantly white progeny and variable plants were observed which showed a continuous change in pattern and intensity of pelargonidin pigmentation. This reduction of *A1* activity argues for a semidominant effect of the 17-W allele, which inhibits the activity of its homologue, 17-R. Most of the F_1 plants that had received a hypomethylated *A1* allele from one parent and a hypermethylated allele from the other parent showed a significant reduction in *A1* expression. Flowers of these F_1 plants were either white or highly variable with white and colored spots or sectors within individual flowers. Again, inhibition of *A1* expression correlated with hypermethylation of the promoter region. Apparently the hypermethylated *A1* allele has paramutagenic potential, inducing methylation in the paramutable *A1* allele that had previously been hypomethylated and that remained hypomethylated in control crosses when it was not combined with the hypermethylated allele. After both *A1* alleles had segregated from each other, the paramutated *A1* allele remained hypermethylated over two generations, strongly supporting again the heritable nature of epigenetic changes. Only occasionally did a weak reactivation of the hypermethylated *A1* allele occur in a few cells. The molecular mechanisms postulated for this interaction and the models presented will be discussed below.

2. Transgenes Inactivate Unidirectionally the Transcription on Non-allelic (Epistatic) Genes. Although single transgenes can be inactivated and sometimes can inactivate their allele, transgene inactivation generally occurs more frequently when multiple copies of a gene are integrated in the genome. The multi-copy integration could take place at a single insertion site or dispersed throughout the genome. In this sub-Section we will focus on this epistatic *trans*-inactivation of homologous genes. Epistatic *trans*-inactivation is defined as unidirectional interaction that occurs between homologous or partially homologous transgenes present in nonallelic chromosomal locations and the outcome is that one hypermethylated transgene inactivates the transcription of another, promoting its methylation (Matzke and Matzke 1996; Vaucheret et al. 1998).

Pyramiding of single-copy transgenes that have sequences in common, whether by sequential transformation or by crossbreeding of transgenic plants can result in the inactivation of one or more of the

introduced genes. *Trans*-inactivation, i.e., the inactivation of genes encoded on one T-DNA following transformation with a second T-DNA encoding a different marker gene, has been initially reported (Matzke et al. 1989; Matzke and Matzke 1990, 1991). Since then several epistatic transgene inactivation events have been described in different plant species, such as petunia, tomato, rice and arabidopsis (Finnegan and McElroy 1994; Finnegan et al. 1998).

In many aspects, epistatic *trans*-inactivation of homologous genes resembles allelic *trans*-inactivation. Another class of epistatic homology dependent *trans*-inactivation, also called co-suppression, will be described in the next sub-Section since it operates post-transcriptionally. Even though the consequences and the final phenotypes of co-suppression and epistatic *trans*-inactivation are the same, i.e., silencing of homologous genes, they can be distinguished by several features: co-suppression is bidirectional and does not block transcription thus indicating a fundamentally different mechanism at the post-transcriptional level.

The importance of repeated sequences in gene inactivation was clearly demonstrated in a study of transgene silencing in arabidopsis, where the copy number of a transgene, which remained at the same genomic location, varied (Davies et al. 1997). Inactivation occurred more frequently when the gene was present in multiple copies leading to the suggestion that this phenomenon be termed RIGS (repeat induced gene silencing) in keeping with similar phenomena, such as repeat-induced point mutation (RIP) of *Neurospora* and methylation induced premeiotically (MIP) in *Ascobolus*.

The clearest indication that multiple copies of homologous transgenes could lead to silencing of transgene expression in tobacco came in sequential transformation experiments in which a previously active transgene locus became inactive and methylated following the introduction, by retransformation, of a second, partially homologous construct, (Matzke et al. 1989). An important aspect that distinguished these experiments from previous ones was that both the methylation and the inactivation were gradually reversible when the second "silencing" transgene locus segregated from the first "target" locus. This established that the silencing of the target locus was directly dependent on the presence of the silencing locus, and not just due to progressive inactivation and methylation, as would be expected of a foreign DNA response (Bestor 1990; Doerfler 1991). A second novel feature was that, unlike previous studies involving multiple transgene copies at a single locus, the target locus and silencing locus in this experiment were clearly shown to be unlinked. This introduced the notion that copies of homologous transgenes on non-homologous chromosomes could somehow interact to modify gene expression and induce epigenetic alterations. Two more studies in transgenic tobacco confirmed the basic observation that homologous, or partially homologous, unlinked transgenes could interact in plant nuclei, leading to reduced expression of a target transgene locus. Hobbs et al. (1990) identified transgenic tobacco lines that either strongly or weakly expressed a transgene construct that encoded resistance to kanamycin and glucuronidase activity. Strongly expressing lines normally contained single, unmethylated copies of the construct, whereas weakly expressing lines contained more than one copy and these were methylated. When one weakly expressing line, containing two methylated copies of the construct arranged as an inverted repeat, was crossed with a strongly expressing line, the activity of the latter was reduced. Vaucheret (1993) crossed a transgenic tobacco line that was homozygous for a silencing locus with more than 20 other tobacco lines that contained a variety of different target transgene constructs. The primary region of homology present between the silencing locus and various potential target loci consisted of the 35S promoter and/or the 19S promoter of cauliflower mosaic virus. Inactivation of all target transgenes was observed, even when the region of homology was reduced to 90 bp of the 35S promoter.

In conclusion, these studies using transgenic tobacco established that a previously active, unmethylated target transgene locus could be reversibly inactivated and methylated when combined with a silencing locus that shared regions of homology with the target. In the silenced target transgenes examined, increased methylation was detected. In all studies the imposed methylation was stably inherited but the observed methylation and inactivation of target transgene loci was reversed when the epistatic silencing locus (which is itself non expressed due to hypermethylation) segregated in progeny from the target locus.

Matzke and Matzke (1996) stressed that in certain cases homology in promoter regions alone may be sufficient to provoke epistatic *trans*-inactivation. This constitutes another significant difference between epistatic *trans*-inactivation and co-suppression where homology in the coding region and expression of both coding regions is required. Table 6 summarizes all the known differences between these two fundamentally different mechanisms of transinactivation, despite their common result.

The epistatic silencing loci are hypermethylated, not expressed at all or weekly expressed and, in many cases, contain multiple copies of the transferred construct. However, methylated single copy silencing loci also exist. Multiple copy silencing loci are sensitive to hypermethylation and, the more methylated they are, the more effective they become in silencing their target.

Table 6. **C**haracteristic differences between the two mechanisms of transinactivation of homologous sequences, namely: epistatic transcriptional silencing and post-transcriptional cosuppression

Transcriptional silencing	**Cosuppression**
Unidirectional	Operates in both directions
Transcription is blocked	Transcription is required
Heavy methylation of homologous sequences is involved	In certain cases partial methylation of homologous sequences could be imposed in a second stage, as an additional contribution to more permanent silencing
Homology in the promoter is adequate	Homology only in the promoter is not adequate; homology in coding areas transcribed in mRNA is required
Run-off test, negative	Run-off test, positive

Potential target transgene loci to *trans*-inactivation vary to their susceptibility to silencing effects. Targets of *trans*-inactivation loci are inherently unstable, with a tendency to become methylated, and less active, even in the absence of a silencing locus. Matzke and Matzke (1996) observed that these intrinsic characteristics were particularly pronounced when the target locus was homozygous, suggesting that allelic interactions between susceptible alleles contribute to the instability. They conclude that a silencing *trans*-inactivating epistatic locus could substitute for an allele, in what was otherwise an interaction between susceptible alleles. This conclusion resembles that for paramutation (Brink and Mikula 1958) in that paramutable sensitive alleles of *R* are inherently unstable with a propensity to spontaneously decrease in activity even in the absence of a paramutagenic silencing *R* alleles, the presence of which might only serve to potentiate an intrinsic instability of the paramutable target locus. This could also be related to inbreeding depression associated with increasing homozygosity in inbred line development when breeding is only restricted to combining ability tests. In discussing hybrid vigor we presented a number of proposals and data, indicating that progressive methylation of inbreds is associated with homozygosity and inbreeding depression. This increase in methylation could be accelerated as methylated and unmethylated homologous genes become homozygous, a state that could be reversed after hybridization in F_1 hybrid as it was proposed by R. L. Phillips (at the Workshop on Molecular Basis of Heterosis in Plants, March 19-20, 1991, University of Minnesota) and cited by Stuber (1998).

Homology restricted even to the non-transcribed promoter area has been shown sufficient to induce silencing in allelic and epistatic *trans*-inactivation. This excludes the involvement of mRNA transcripts and leads to mechanisms involving some short of "communications" between homologous loci. Direct DNA-DNA pairing of long enough homologous loci could be the mode of such communication. Rossignal and Faugeron (1994) estimated that more than 300 bp of homology is required for such pairing to occur. In cases where the homology is shorter, like in constructs whith homology restricted to only 90 bp in the promoter area, an indirect pairing mediated by multimeric DNA-binding proteins (mostly regulatory proteins like multimeric transcription factors) has been postulated. Identical promoters could be recognized by an aggregate protein, leading to clustering in the same nuclear compartment where methylation and/or chromatin changes could be imposed coordinately. Such a model might also be relevant for *trans*-silencing in a non-transgenic system in which paramutagenicity has been mapped to the 5' flanking region of the *B*' allele in maize and no association of paramutagenicity with methylation was found (Patterson et al. 1995). This indicated an epigenetic system of paramutation involving alterations in chromatin structure via DNA-protein interactions without clear involvement of methylation.

C. Bidirectional Inactivation of Homologous Sequences Post-Transcriptionally

A common goal of many efforts in plant biotechnology is the over-expression of an endogenous or introduced trait. Introduction of extra copies of an endogenous gene to boost expression may, however, result in the coordinate silencing not only of the introduced transgenes but also of the endogenous gene. This phenomenon, termed co-suppression, was first described by Jorgensen and his colleagues (Napoli et al. 1990). A petunia chalcone synthase (*chs*) gene, encoding one step in floral pigment biosynthesis, was introduced into a deep violet line of *Petunia hybrida*. Surprisingly up to 42% of transgenic plants produced flowers that were completely white or had variegated flowers with white or pale sectors on a normally colored background. While over-expression of *chs* mRNA occurred in normally pigmented floral tissue with one or more copies of the transgene, the steady state level of *chs* mRNA was reduced in white or pale floral sectors. In contrast, nascent transcription,

as measured by nuclear run-on experiments, was comparable in white and pigmented tissue, indicating that co-suppression is a post-transcriptional event (Flavell 1994; Flavell et al. 1998).

In another case of transgene silencing, transgene mediated silencing of the endogenous gene encoding •-1,3-glucanase in tobacco, de Carvalho Niebel et al. (1995) demonstrated that the suppressed genes are actively transcribed. Subsequently, it was shown that gene silencing in this line correlate with an increased turnover of both the transgenic and the endogenous transcripts of •-1,3-glucanase.

The most notable examples of co-suppression are those related to resistance to some RNA viral infections, especially the mechanism of resistance of some transgenic plants bred for resistance to RNA viral infections. Attempts to engineer virus resistance are forging a connection between post-transcriptional silencing phenomena and cytoplasmic activities that help protect plants from viral infection.

The idea emerging is that plants can somehow sense elevated levels of an RNA sequence arising from a replicating viral genome or a highly transcribed nuclear transgene, or both acting additively, and target that RNA for destruction (Wassenegger et al. 1994; Kumagai et al. 1995; Angell and Baulcombe 1997). The developmental of such a strategy from the plants was of extreme importance for their survival, as many of them can be asexually propagated using propagules that could have been infected by viruses and 90% of plant viruses have an RNA genome.

The so-called biochemical switch model, suggested by Meins and Kunz (1994), best explains most of the co-suppression cases observed. According to this model, a product of gene expression, e.g., RNA, would accumulate until a threshold has been reached, at which point RNA degradation would be initiated. In transgenic plants, overproduction of a given RNA, resulting from transcription of a transgene in addition to transcription of a homologous endogenous gene, would provoke its own turnover. The same could result when extremely active promoters in combination with strong enhancers have been used to express a transgene.

Unlike the *trans*-inactivation phenomena described in the previous paragraphs, co-suppression does not appear to be associated with DNA methylation of the suppressed genes and is reversible both somatically and upon segregation of the transgene, indicating that the repressed state is unstable. For this reason, co-suppression should be out of the scope of this review. But, there was evidence that elevated levels of certain RNA transcripts could trigger co-suppression and, could also induce methylation of the coding genes (thus more permanent inactivation of transcription). This interconnects co-suppression with DNA methylation. By immediate destruction of RNA paralleled with inactivation, through methylation of the coding gene, a dual mechanism is secured for the final objective: first, quick inactivation of translatable or available viral RNAs, and second, more permanent inhibition of their synthesis by methylating existing copies of DNA in the nucleus.

Strong evidence for the involvement of RNA in this type of silencing comes from experiments in which RNAs have been introduced in the absence of a DNA template (using RNA viruses as vectors). The ability of viral RNAs to interfere with a homologous gene in the plant genome is one of the strongest arguments for the existence of RNA-mediated silencing mechanisms (Wassenegger et al. 1994; Kumagai et al. 1995; Angell and Baulcombe 1997). Other evidence in favor of the involvement of RNA in co-suppression is the finding that the phenomenon can spread from the site of interfering RNA synthesis or application. Two groups of researchers have demonstrated the systemic spread of co-suppression in plants (Palauqui et al. 1997; Voinnet and Baulcombe 1997). An RNA molecule, spreading throughout the plant via phloem, has been proposed as the mobile agent responsible for transmitting the co-suppression state (Voinnet and Baulcombe 1997). Recently, Montgomery and Fire (1998) discussed the possibility that double-stranded RNA (dsRNA), rather than sense or antisense single-stranded RNA alone, is the effector molecule responsible for RNA-mediated silencing and co-suppression, and presented different ways as to how these dsRNAs are formed from complementary RNA molecules. Extensive transcription of a sense chain from a strong promoter leaving a chromatin area in a relaxed form that allows formation of antisense transcripts, from a hidden promoter of the opposite complementary side, is one case. Arguments against this hypothesis have been presented by Jorgensen et al. (1999) who favor a role of copy RNA (cRNA) molecules produced post-transcriptionally, in co-supression.

The original observation for another type of interaction between the coding sequence and its RNA product was made by Wassenegger et al. (1994) who showed that viral or transgene-generated RNA could direct *de novo* methylation of a homologous sequence in the plant genome. Could there be a mechanistic link between (1) RNA-mediated degradation of RNA and (2) RNA-mediated methylation of DNA? One possibility is that these are two separate processes mediated by similar RNA molecules; alternatively, there could be a causal relationship, perhaps from an ability of RNA decay products to trigger methylation of homologous sequences in replicating DNA. Crucial data towards elucidation of this matter were presented by Ruiz et al. (1998) supporting the presence of two separate steps: (1) initiation of the degradation reaction and (2) maintenance of the silenced stage not requiring any further the presence of RNA.

D. Avoiding Transgene Silencing

If the inactivation of transgenes reflects normal aspects of plant development, can transgenes be engineered in such a way as to evade this response with a view towards improving their stability under field conditions? Strategies to minimize silencing could be tailored to avoid different kind of silencing mechanisms. As noted

earlier, transgene inactivation is frequently associated with the integration of multiple copies of the introduced DNA, either dispersed throughout the genome or at a single insertion site. Thus, the approach towards stabilizing expression that is most commonly used is to select and/or screen for those transgenic plants which contain a single intact copy of the transgene. As direct gene transfer methods (Klein et al. 1988; Koziel et al. 1993; Thompson et al. 1995) with some species, especially monocots, generally result in multiple insertion events, a method to reduce the copy number of the integrated transgene would be desirable. This might be achieved most readily by extending *Agrobacterium*-mediated transformation technologies to monocotyledonous plants, and encouraging results towards this goal have been published already (Gould et al. 1991; Konstantinov et al. 1991; Cheng et al. 1997; Heath et al. 1997; Shen et al. 1998).

Eliminating repeated elements from transgene constructs should alleviate problems with DNA-DNA pairing and *de novo* methylation. If a particular isochore composition of a transferred gene is important for its stable activity, the source of the transgene should be carefully selected for the AT/GC content to mimic, as close as possible, the composition of the targeted isochore of the host plant. If necessary the AT/GC content of the transgene could be changed with in vitro methods making use of redundancy of the genetic code.

The co-transfer of neighboring chromosomal regions significantly stabilizes the expression of the human globin gene (Grosveld et al. 1987). A similar result was obtained from work with the matrix attachment regions of the chicken lysozyme gene which co-maps with the boundaries of the chromatin domain of this locus. A reporter gene that was flanked on both sides by a 3-kb attachment element from the 5' region of the chicken lysozyme gene mediated elevated and position-independent transgene activity (Stief et al. 1989). In plants similar stabilizing effects have been achieved with transgenes flanked by scaffold attachment sequences (Breyne et al. 1992; Allen et al. 1993; Schoffl et al. 1993).

Finally since the integration region of the host genome contributes to stable transgene expression, the use of methods that would allow the selection of a suitable integration region into which the transgenes might be targeted should contribute to their stability. The present low efficiency of methods for gene targeting into the plant genome is constantly being improved so that in plants, as it is already possible for other eukaryotes, site specific integration of a transgene will be achievable (Paszkowski et al. 1988).

E. Implications in Breeding

The significance of transgene silencing in breeding new cultivars with genetic engineering techniques has been discussed, followed by measures for avoiding such undesirable transgene silencing. But to the extent that the kind of phenomena observed in transgenes are simply the outcome of basic biological mechanisms operating in plants, thus affecting internal genes, too (i.e. paramutation, imprinting, transposon silencing etc.) they have implications in conventional plant breeding.

From the examples presented here it is obvious that the presence of 2-3 extra copies of homologous genes in a genome frequently leads to their inactivation. Yet, many of today's crops are polyploids harboring extra copies of homologous sequences. For many enzyme encoding plant genes, the existence of small families of 2-3 and sometimes up to 8-10 copies of homologous genes and pseudogenes in the genome is a rule rather than an exception. Also, according to Stelly et al. (1997) duplication and expansion of plant genomes frequently leads to imbalances and inactivation of many genes. Evolutionary time may be needed for the duplicated genomes to harmonize. Epigenetic inactivation could provide quick solutions to this kind of problems, allowing for the genome to explore new opportunities, for an appropriate reversion of extra gene copies during development, and offering new opportunities to the plant for adaptation to any new environment.

Potato, for example, is an autotetraploid arose from the doubling of the chromosome number within a species. Because the chromosome sets are not distinct and pair at meiosis, plants will exhibit tetrasomic inheritance and duplicate loci will fail to diverge. Thus similar homologous sequences should exist. The potato genome due to its polyploidy could "hide" a number of degenerative alleles and due to its asexual reproduction could "hide" them further from being exposed to selection, and elimination. Genomic allotetraploids, such as wheat or cotton, arose via the fusion of the genomes of two distinct species. Because the genomes are distinct, chromosomes do not pair but evolve separately within the tetraploid. However, homologous sequences should exist in multiple copies. In this respect, even maize has long been thought to be an ancient tetraploid whose genome has reverted over time to functional diploidy and thus lacks two clear sets of duplicated chromosomes. Part of the evidence for this interpretation is that while maize has ten gametic chromosomes, several members of the *Andropogonae* tribe to which maize belongs have only five. In addition, maize has two unlinked copies of many genes, each of which is found on duplicate chromosomal segments (Helentjaris et al. 1988). The same is true for the "diploid" tomato (Messeguer et al. 1991).

Does the presence of extra copies of homologous sequences promote silencing of certain loci and what was the role of methylation in shaping the evolution of these genomes? Data are only rarely available for answering this question. But one way to study the load of defective genes is to bring some of them into homozygous conditions, thereby exposing their phenotype. The variation in single plant performance, grown in isolation (as quantified by the coefficient of variation - CV) as well as the distribution in plotting single plant performances, are strong indicators of this load (Fasoulas, 1993). As shown in Fig. 8, an F_2 derived by selfing

the tomato hybrid 'Dombo' grown in the isolation environment (1.25X1.25 m apart in the field) shows a low CV of 0.26 and a normal distribution in plotting single plant performance, indicating a low load of degenerated alleles. On the contrary, F_2 plants derived by selfing the asexual potato hybrid 'Jaerla' grown in the same isolation environment show a larger CV of 0.94 and strongly positive skew (L-shaped distribution) in plotting. Thus, the two species have significant different genetic loads. The asexually propagated polyploid potato significantly surpasses the self fertilized diploid tomato (Fasoula and Fasoula 1997a).

Part of this load is due to mutated alleles, but another part could be due to epialleles derived from differential methylation. The extent and size of each part needs to be considered. Crops should be ranked according to both kinds of genetic loads always taking into consideration different parameters such as their mode of reproduction, polyploidization, time of domestication, degree of isolation during domestication, and breeding efforts.

Doebley and colleagues (White and Doebley 1998) studying the evolution of the maize genome especially in comparison with the progenitor relative *Teosinte* found minor role of transposable elements in the evolution of this most unusual plant, despite their widespread occurrence in the genome. On the contrary, of the five genes (or chromosomal areas with multiple tightly linked genes) postulated to have played a critical role for the divergence of maize from *Teonsite*, two genes isolated and studied in detail were regulatory and for one of them (tb_1 - teosinte branched 1) a quantitative change in the level of expression (that may be associated with changes in its DNA methylation) was found to play critical role.

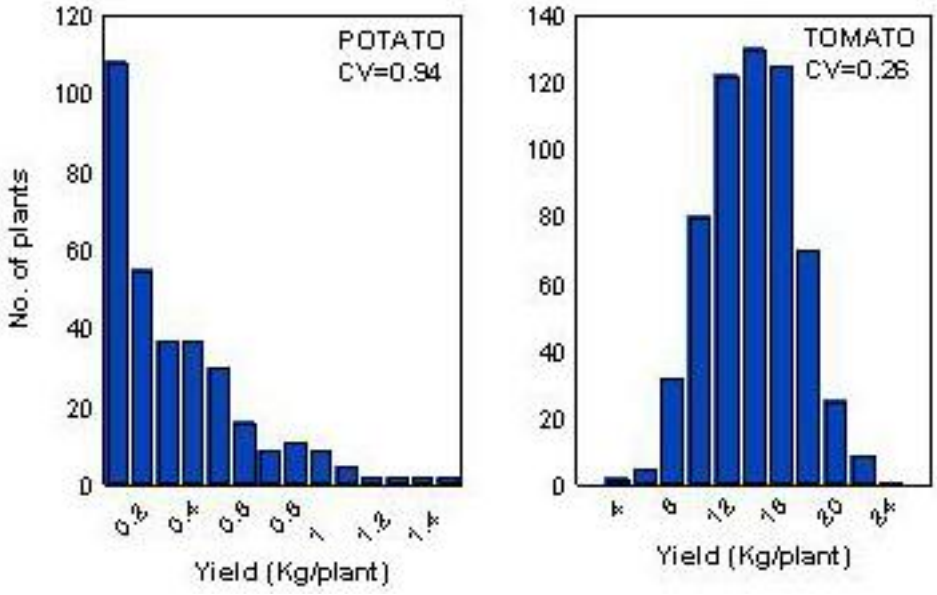

Fig. 8. Single-plant yield histograms of F_2s grown in an isolation environment. The distribution of seedlings after self pollination of the potato cultivar has a high CV and is skewed toward the low-yield end, indicative of high load of degenerative genes, in contrast with the distribution of tomato F_2s which is normal with low CV, indicative of low load of degenerative genes (see Fasoula and Fasoula 1997a).

The tb_1 locus controls tillering and prolificacy in maize. During maize evolution, domestication, and subsequent breeding there was selection for increased apical dominance and repression of growth of branches and prolificacy. Indian maize, as described in Leonard Fuchs' herbal *De historia stirpium*, published in 1542 (Janick et al. 1974) was multiculmed and multi-eared. Farmers and breeders, by selecting for hundreds of years for uniculm and uni-ear plants, gave maize its uniculm trait. This was accomplished by a quantitative shift in the amount of tb_1 expression.

Recent studies have shown that prolificacy is indispensable for increasing hybrid yield and stability (Hallauer 1973; Prior and Russell 1975). Motto and Moll (1983) reviewed the available information on the prolific habit of maize and indicated that prolificacy is a relevant component of grain yield, which strongly affects stability over a wide range of environmental conditions. In fact, Tollenaar (1991) has indicated that at the density of 2 plants/m^2, new hybrids tillered more profusely than old hybrids. Fasoulas (1993) presented data supporting the view that selecting maize plants for a few generations for high yield in the isolation environment (1.5 m apart, nil competition) very frequently leads to more profusely tillering plants and prolificacy. Lonnquist (1967) was able to increase the productivity of `Hays Golden,' an open-pollinated maize cultivar, 6.3% per generation for the first five generations by selecting for ear prolificacy. Russell and Machado (1978) suggested that the advantage of the prolific lines in increasing hybrid yield and stability was because they were better able to produce ears in stressful environments, whereas single-eared lines had a greater incidence of barren plants. Obviously, prolific plants are better buffered against environmental stress. Thus, tb_1 activity could possibly be

changed by a few generations of selection after relaxing selection for uniculm habit, arguing for reversible epigenetic control of tb_1 activity. To what extent this is so should soon be clarified since the tb_1 gene has been cloned using *Mutator* transposon tagging (Doebley et al. 1997); thus its sequence could be used to analyse recent maize genetic material with prolificacy.

IV. CONCLUSIONS AND FUTURE PROSPECTS

It seems clear that genes, the unit of inheritance, are more complex than mere DNA nucleotide sequence. Epigenetic inheritance systems, like DNA methylation, are receiving attention and complement classical genetics in many fields including molecular and developmental biology, epidemiology, cancer biology and evolution. In this review, experimental data and hypotheses have been presented supporting the view that epigenesis should be an integral part of conventional plant breeding. DNA methylation was found to be associated with action and interaction of plant genes during development. By this means DNA methylation affects important parameters of a conventional plant breeding program including creation of usable variation, single plant heritability, hybrid vigor, plant-environment interactions, stress tolerance and stability of performance. DNA methylation was found to be associated with the induction of mutation and new genetic variation. DNA methylation was proven equally important for transgenic technology because of the interference with the phenotypic stability in new transgenic cultivars.

Because the development of germ cells occurs late in the life cycle of annual plants genetic and epigenetic variants arising in somatic cells during development may yield variant gametes. Thus, epigenetic variation could be of particular importance for many plants. Plant breeders should follow new strategies and methods for further improvements of crops through this newly opened path provided by epigenesis. The initial of this review was to clear this path and to broaden horizons. We have attemted to integrate the existing principles of plant breeding with the emerging information on DNA methylation. Our rational is based on the puissant comment of Zirkle, (1952):

> "Great steps forward generally are made not by the discovery of new facts, important as they are, or by new ideas, brilliant as they may be, but by the organization of existing data in such a way that hitherto unperceived relationships are revealed, and by incorporating the pertinent data into the general body of knowledge so that new, basic principles emerge".

As more data accumulates we anticipate a clearer role for the exploitation of epigenetic phenomena such as DNA methylation in plant breeding. Better understanding and use of plant breeding methods will widen our views and ideas for plant evolution and domestication.

LITERATURE CITED

Adams, R. L. and R. Eason. 1984. Increased G + C content of DNA stabilizes methyl CpG dinucleotides. Nucleic Acids Res. 12: 5869-77.

Adams, R. L. P. and R. H. Burdon. 1985. Molecular biology of DNA methylation. Springer-Verlag, New York.

Allen, G. C., G. E. Hall, L. C. Childs, A. K. Weissinger, S. Spiker and W. F. Thompson. 1993. Scaffold attachment regions increase reporter gene expression in stably transformed plant cells. Plant Cell 5: 603-613.

Angell, S. M. and D. C. Baulcombe. 1997. Consistent gene silencing in transgenic plants expressing a replicating potato virus X RNA. EMBO J. 16: 3675-3684.

Antequera, F. and A. Bird. 1993. Number of CpG islands and genes in human and mouse. Proc. Nat. Acad. Sci. (USA) 90: 11995-11999.

Atchley, W. R. 1989. Introduction to the symposium: Evolution of developmental systems. Am. Naturalist 134: 437-512.

Banks, J. A. and N. Fedoroff. 1989. Patterns of developmental and heritable change in methylation of the *Suppressor-mutator* transposable element. Dev. Genet. 10: 425-437.

Banks, J. A., P. Masson and N. Fedoroff. 1988. Molecular mechanisms in the developmental regulation of maize *Suppressor-mutator* transposable element. Genes Dev. 2: 1364-1380.

Bedford, M. T. and P. P. Van Helden. 1990. Analysis of allele-specific methylation. Trends Genet. 6: 140-141.

Belanger, F. C. and A. G. Hepburn. 1990. The evolution of CpNpG methylation in plants. J. Mol. Evol. 30: 26-35.

Bennetzen, J. L. 1987. Covalent DNA modification and the regulation of *Mutator* element transposition in maize. Mol. Gen. Genet. 208: 45-51.

Bernacchia, G., A. Primo, L. Giorgetti, L. Pitto and R. Cella. 1998. Carrot DNA-methyltransferase is encoded by two classes of genes with differing patterns of expression. Plant J. 13: 317-329.

Bestor, T. H. 1990. DNA methylation: evolution of a bacterial immune function into a regulator of gene expression and genome structure in higher eukaryotes. Phil. Trans. Royal Soc. Lond. (Biol.) 326: 179-187.

Bestor, T. H. 1992. Activation of mammalian DNA methyltransferase by cleavage of Zn binding regulatory domain. EMBO J. 11: 2611-2617.

Bestor, T.H. 1998a. Methylation meets acetylation. Nature 393: 311-312.

Bestor, T. H. 1998b. The host defence function of genomic methylation patterns. p. 187-195. In D. J. Chadwick and G. Gardew (eds): Epigenetics. John Wiley and Sons Ltd, Chichester, England.

Bevan, M., I. Bancroft, E. Bent, K. Love, H. Goodman, et al. 1998. Analysis of 1.9 Mb of contiguous sequence from chromosome 4 of *Arabidopsis thaliana*. Nature 391: 485-488.

Bianchi, M. W. and A. Viotti. 1988. DNA methylation and tissue-specific transcription of the storage proteins of maize. Plant Mol. Biol. 11: 203-214.

Bird, A. 1997. Does DNA methylation control transposition of selfish elements in the germline? Trends Genet. 13: 469-470.

Bird, A. P. 1986. CpG-rich islands and the function of DNA methylation. Nature 321: 209-13.

Boyes, J. and A. Bird. 1991. DNA methylation inhibits transcription indirectly via a methyl-CpG binding protein. Cell 64: 1123-34.
Brandle, J. E., S. G. McHugh, L. James, H. Labbe and B. L. Miki. 1995. Instability of transgene expression in field grown tobacco carrying the *csr1-1* gene for sulfonylurea herbicide resistance. Biotechnology 13: 994-998.
Breese, E. L., M. D. Hayward and A. C. Thomas. 1965. Somatic selection in perennial ryegrass. Heredity 20: 367-379.
Brettell, R. I. S. and E. S. Dennis. 1991. Reactivation of a silent Ac following tissue culture is associated with heritable alterations in its methylation pattern. Mol. Gen. Genet. 229: 365-372.
Breyne, P., M. Van Montague, A. Depicker and G. Gheysen. 1992. Characterization of a plant scaffold attachment region in a DNA fragment that normalizes transgene expression in tobacco. Plant Cell 4: 463-471.
Brink, R. A. 1956. A genetic change associated with the *R* locus in maize which is directed and potentially reverible. Genetics 41: 872-889.
Brink, R. A. and B. Mikula. 1958. Plant color effects of certain anomalous forms of the *R-r* allele in maize. Z. ind. Abstamm. Vererb-lehre. 89: 94-102.
Brown, D. F. 1966. Paramutability of *R-g* and *r-r* mutant genes derived from an *R-r* allele in maize. Genetics 54: 899-910.
Brutnell, T. P. and S. L. Dellaporta. 1994. Somatic inactivation and reactivation of Ac associated with changes in cytosine methylation and transposase expression. Genetics 138: 213-225.
Buss, L. W. 1987. The evolution of individuality. Princeton Univ. Press, Princeton NJ.
Cai, Q., C. L. Guy and G. A. Moore. 1996. Detection of cytosine methylation and mapping of a gene influencing cytosine methylation in the genome of *Citrus*. Genome 39: 235-242.
Cavalli, G. and R. Paro. 1998. The *Drosophila* Fab-7 chromosomal element conveys epigenetic inheritance during mitosis and meiosis. Cell 93: 505-18.
Chadwick D. J. and G. Gardew (eds). 1997. Epigenetics. John Wiley and Sons Ltd, Chichester, England.
Chandler, V. and V. Walbot. 1986. DNA modification of a maize transposable element correlates with loss of activity. Proc. Nat. Acad. Sci. (USA) 83: 1767-1771.
Cheng, M., J. E. Fry, S. Z. Pang, H. P. Zhou, C. M. Hironaka, D. R. Duncan, T. W. Conner, Y. C. Wan. 1997. Genetic transformation of wheat mediated by *Agrobacterium tumefaciens*. Plant Physiol. 115: 971-980.
Chomet, P. S., S. Wessler and S. L. Dellaporta. 1987. Inactivation of the maize transposable element activator (Ac) is associated with its DNA modification. EMBO J. 6: 295-302.
Church, G. M. and W. Gilbert. 1984. Genomic sequencing. Proc. Nat. Acad. Sci. (USA) 81: 1991-1995.
Citti, L., P. G. Gervasi, G. Turchi and L. Mariani. 1983. Improved method for determination of 5-methylcytosine by ion-pair reversed-phase high-performance liquid chromatography. J. Chromatog. 261: 315-319.
Coe, E. H. 1966. The properties, origin, and mechanism of conversion-type inheritance at the *B* locus in maize. Genetics 53: 1035-1063.
Cone, K. C., S. M. Cacciolone, C. A. Moehlenkamp, T. Weber, B. J. Drummond, L. A. Tagliani, B. A. Bowen and G. H. Perrot. 1993. Role of the regulatory gene *pl* in the photocontrol of maize antocyanin pigmentation. Plant Cell 5: 1807-1816.
Cooper, D. N., L. H. Errington and R. M. Clayton. 1983. Variation in the DNA methylation pattern of expressed and nonexpressed genes in chicken. DNA 2: 131-140.
Cooper, D. N. and M. Krawczak. 1990. The mutational spectrum of single base-pair substitutions causing human genetic disease: patterns and predictions. Human Genet. 85: 55-74.
Coors, J. (ed.). 1998. The genetics and exploitation of heterosis in crops. Am. Soc. Agronomy, Madison, WI.
Coulondre, C., J. H. Miller, P. J. Farabaugh and W. Gilbert. 1978. Molecular basis of base substitution hotspots in *Escherichia coli*. Nature 274: 775-780.
Cullis, C. A. 1981. Environmental induction of heritable changes in flax: defined environments inducing changes in rDNA and peroxidase isozyme band pattern. Heredity 47: 87-94.
Cullis, C. A. 1983. Environmentally induced DNA changes in plants [Interdependence of the genotype and of the environment]. Crit. Rev. Plant Sci. 1: 117-131.
Damerval, C., Y. Hebert and D. de Vienne. 1987. Is the polymorphism of protein amounts related to phenotypic variability? A comparison of two-dimensional electrophoresis data with morphological traits in maize. Theor. Appl. Genet . 74: 194-202.
Damerval, C., A. Maurice, J. M. Josse and D. De Vienne. 1994. Quantitative trait loci underlying gene product variation: A novel perspective for analyzing regulation of genome expression. Genetics 137: 289-301.
Davies, G. J., M. A. Sheikh, O. J. Ratcliffe, G. Coupland and I. J. Furner. 1997. Genetics of homology-dependent gene silencing in *Arabidopsis*; a role for methylation. Plant J. 12: 791-804.
De Carvalho Niebel, F., P. Frendo, M. Van Montagu and M. Cornelissen. 1995. Post-transcriptional cosuppression of beta-1,3-glucanase genes does not affect accumulation of transgene nuclear mRNA. Plant Cell 7: 347-358.
Doebley, J., A. Stec and L. Hubbard. 1997. The evolution of apical dominance in maize. Nature 386: 485-488.
Doerfler, W. 1991. Patterns of DNA methylation-evolutionary vestiges of foreign DNA inactivation as a host defense mechanism-a proposal. Biol. Chem. Hoppe-Seyler 372: 557-564.
Donald, C. M. and J. Hamblin. 1976. The biological yield and harvest index of yield as agronomic and plant breeding criteria. Adv. Agron. 28: 361-405.
Dorlhac de Borne, F., M. Vincentz, Y. Chupeau and H. Vaucheret. 1994. Co-suppression of nitrate reductase host genes and transgenes in transgenic tobacco plants. Mol Gen. Genet. 243: 613-621.
Duncan, B. K. and J. H. Miller. 1980. Mutagenic deamination of cytosine residues in DNA. Nature 287: 560-561.
Duvick, D. N. 1984. Genetic contributions to yield gains of U.S. hybrid maize, 1930 to 1980. p. 15-47. In: W. Fehr (ed.), Genetic contributions to yield gains of five major crop plants. CSSA Special Pub. 7, Crop Sci. Soc. Am., Madison, WI.
Duvick, D. N. 1992. Genetic contributions to advances in yield of U.S. maize. Maydica 37: 69-79.
Duvick, D. N. 1997. What is yield? In: G. O. Edmeades, B. Banziger, H. R. Mickelson and C. B. Pena-Valdivia (eds.), Developing drought and low N-tolerant maize. Centro International de Mejoramiento de Maiz y Trigo, El Batan, Mexico, D.F.
Elings, A., G. O. Edmeades and M. Banziger. 1998. Is heterosis a form of stress tolerance? in press. In: J. Coors (eds.), The genetics and exploitation of heterosis in crops. Am. Soc Agronomy, Madison, WI.
Fasoula, D. A. and V. A. Fasoula. 1997a. Competitive ability and plant breeding. Plant Breed. Rev. 14: 89-138.
Fasoula, D. A. and V. A. Fasoula. 1997b. Gene action and plant breeding. Plant Breed. Rev. 15: 315-374.
Fasoulas, A. C. 1993. Principles of crop breeding. Fasoulas, A. C. , P.O. Box 1555, GR-54006, Thessaloniki, Greece.
Fasoulas, A. C. and D. A. Fasoula. 1995. Honeycomb selection designs. Plant Breed. Rev. 13: 87-139.
Fasoulas, A. C. and A. S. Tsaftaris. 1975. An integraded approach to plant breeding and field experimentation. Dept. Gen. Plant Breed., Aristotelian Univ., Thessaloniki, Greece, Publ. 5.
Fedoroff, N. 1989. Maize transposable elements. p. 375-411. In: D. Berg and M. Howe (eds.), Mobile DNA. Am. Soc. Microbiol., Washington DC.
Fedoroff, N. 1996. Epigenetic regulation of the maize *Spm* transposable element. p. 575-592. In: V. E. A. Russo, R. A. Martienssen and A. D. Riggs (eds.), Epigenetic mechanisms of gene regulation. Cold Spring Harbor Laboratory Press, New York.
Fedoroff, N. V. 1995. DNA methylation and activity of the maize *Spm* transposable element. p. 143-164. In: P. Meyer (eds.), Gene silencing in higher plants and related phenomena in other eukaryotes. Springer-Verlag, New York.

Fieldes, M. A. 1994. Heritable effects of 5-azacytidine treatments on the growth and development of flax (*Linum usitatissimum*) genotrophs and genotypes. Genome 37: 1-11.
Finnegan, E. J. and E. S. Dennis. 1993. Isolation and identification by sequence homology of a putative cytosine methyltransferase from *Arabidopsis thaliana*. Nucl. Acids Res. 21: 2383-2388.
Finnegan, E. J., W. J. Peacock and E. S. Dennis. 1996. Reduced DNA methylation in *Arabidopsis thaliana* results in abnormal plant development. Proc. Nat. Acad. Sci. U. S. A. 93: 8449-8454.
Finnegan, J. and D. McElroy. 1994. Transgene inactivation: Plants fight back! Biotechnology 12: 883-888.
Flavell, R. B. 1994. Inactivation of gene expression in plants as a consequence of novel sequence duplications. Proc. Nat. Acad. Sci. (USA) 91: 3490-3496.
Flavell, R. B., M. O'Dell and M. Metzlaff. 1998. Transgene-promoted epigenetic switches of chalcone synthase in petunia plants. p. 144-154. In D. J. Chadwick and G. Gardew (eds): Epigenetics. John Wiley and Sons Ltd, Chichester, England.
Frank, D., M. Lichtenstein, Z. Paroush, Y. Bergman, M. Shani, A. Razin and H. Cedar. 1990. Demethylation of genes in animal cells. Phil. Trans. Royal Soc. London B. 326: 241-252.
Furner, I. J., M. A. Sheikh and C. E. Collett. 1998. Gene silencing and homology-depented gene silencing in *Arabidopsis*: genetic modifiers and DNA methylation. Genetics 149: 651-662.
Gardiner-Garden, M., J. A. Sved and M. Frommer. 1992. Methylation sites in angiosperm genes. J. Mol. Evol. 34: 219-230.
Giordano, M., M. E. Mattachini, R. Cella and G. Pedrali-Noy. 1991. Purification and properties of a novel DNA methyltransferase from cultured rice cells. Biochem. Biophys. Res. Commun. 177: 711-719.
Gould, J., M. Devey, O. Hasegawa, E. C. Ulian, G. Peterson and R. H. Smith. 1991. Transformation of *Zea mays* L. using *Agrobacterium tumefaciens* and the shoot apex. Plant Physiol. 95: 426-434.
Grosveld, F., G. Blom van Assendelft, D. R. Greaves and G. Kollias. 1987. Position-independed, high-level expression of the human beta-globin gene in transgenic mice. Cell 51: 975-985.
Gruenbaum, Y., T. Naveh-Many, H. Cedar and A. Razin. 1981. Sequence specificity of methylation in higher plant DNA. Nature 292: 860-862.
Grunstein, M. 1998. Yeast heterochromatin: regulation of its assembly and inheritance by histones. Cell 93: 325-328.
Hallauer, A. R. 1973. Hybrid development and population improvement by reciprocal full-sib selection. Egypt. J. Genet. Cytol. 1: 84-101.
Hayashizaki, Y. and S. Watanabe (eds.). 1997. Restriction landmark genomic scanning (RLGS). Springer-Verlag, New York.
Hayward, M. D. and E. L. Breese. 1968. The genetic organization of natural populations of *Lolium perenne* L. III. Productivity. Heredity 23: 357-368.
Heath, J. D., M. I. Boulton, D. M. Raineri, S. L. Doty, A. R. Mushegian, T. C. Charles, J. W. Davies, E. W. Nester. 1997. Discrete regions of the sensor protein VirA determine the strain-specific ability of *Agrobacterium* to agroinfect maize. Mol. Plant-Microb. interact. 10: 221-227.
Heinkoff, S. and T. D. Dreasen. 1989. *Trans*-inactivation of the *Drosophila* brown gene: evidence for transcriptional repression and somatic pairing dependence. Proc. Nat. Acad. Sci. (USA) 86: 6704-6708.
Helentjaris, T., D. Weber and S. Wright. 1988. Identification of genomic locations of duplicate nucleotide sequences in maize by analysis of restriction fragment length polymorphisms. Genetics 118: 353-363.
Heslop-Harrison, J. S. and M. D. Bennett. 1990. Nuclear architecture in plants. Trends Genet. 6: 401-405.
Hobbs, S. L. A., P. Kpodar and C. M. O. Delong. 1990. The effect of T-DNA copy number, position and methylation on reporter gene expression in tobacco transformants. Plant. Mol. Biol. 15: 851-864.
Hollick, J. B., G. I. Patterson, E. H. J. Coe, K. C. Cone and V. L. Chandler. 1994. Allelic interactions heritably alter the activity of a metastable maize *pl* allele. Genetics 141: 709-719.
Jablonka, E. and M. J. Lamb. 1995. Epigenetic inheritance and evolution. The Lamarkian dimension. Oxford University Press, New York.
Jacobsen, S. E. and E. M. Meyerowitz. 1997. Hypermethylated SUPERMAN epigenetic alleles in *Arabidopsis*. Science 277: 1100-1103.
Janick, J. 1998. Exploitation of heterosis: uniformity and stability considerations. in press. In: J. Coors (eds.), The genetics and exploitation of heterosis in crops. Am. Soc. Agronomy, Madison, WI.
Janick, J., R. W. Schery, F. M. Woods and V. W. Ruttan. 1974. Plant science: an introduction to world crops. W. H Freeman and Co., San Francisco, CA.
John, M. C. and R. M. Amasino. 1989. Extensive changes in DNA methylation patterns accompany activation of a silent T-DNA *ipt* gene in *Agrobacterium tumefaciens*- transformed plant cells. Mol. Cell Biol. 9: 4298-4303.
Jones, D. F. 1917. Dominance of linked factors as a means of accounting for heterosis. Genetics 2: 466-479.
Jones, P. L., G. C. Jan Veenstra, P. A. Wade, D. Vermaak, S. U. Kass, N. Landsberger, J. Strouboulis and A. P. Wolffe. 1998. Methylated DNA and MeCP2 recruit histone deacetylase to repress transcription. Nature Genet. 19: 187-191.
Jones, P. A. 1985. Altering gene expression with 5-azacytidine. Cell 40: 485-486.
Jones, P. A., W. M. Rideout, J.-C. Shen, C. H. Spruck and Y. C. Tsai. 1992. Methylation, mutation and cancer. BioEssays 14: 33-36.
Jorgensen, R. A. 1995. Cosuppression, flower color patterns, and metastable gene expression states. Science 268: 686-691.
Jorgensen, R. A., Q. Que, M. Stam. 1999. Do unintended antisense transcripts contribute to sense cosuppression in plants? Trends Genet. 15: 11-12.
Josse, J., A. D. Kaiser and A. Kornberg. 1961. Enzymatic synthesis of deoxyribonucleic acid VIII. Frequencies of nearest neighbor base sequences in deoxyribonucleic acid. J. Biol. Chem. 236: 864-875.
Jost, J.-P. and H.-P. Saluz. 1993. DNA methylation: molecular biolgy and biological significance. Brinkhauser Verlag, Boston.
Kafka, M. 1996. Control of gene expression in lines and hybrids of maize (*Zea mays*). Ph.D. Thesis. Aristotelian Univ., Thessaloniki, Greece.
Kakutani, T., J. A. Jeddeloh, S. K. Flowers, K. Munakata and E. J. Richards. 1996. Developmental abnormalities and epimutations associated with DNA hypomethylation mutations. Proc. Nat. Acad. Sci. U. S. A. 93: 12406-12411.
Kakutani, T., J. A. Jeddeloh and E. J. Richards. 1995. Characterization of an *Arabidopsis thaliana* DNA hypomethylation mutant. Nucleic Acids Res. 23: 130-137.
Kang, M. S. 1998. Phenotypic plasticity, heterosis, and environmental stress: a concise review. in press. In: J. Coors (eds.), The genetics and exploitation of heterosis in crops. Am. Soc. Agronomy, Madison, WI.
Karp, A. 1993. Are your plants normal?-Genetic instability in regenerated and transgenic plants. Agro-Food-Industry Hi-Tech. May/June: 7-12.
Karp, A. and S. W. J. Bright. 1985. On the causes and origin of somaclonal variation. Oxf. Surv. Plant Mol. Cell Biol. 2: 199-234.
Kass, S. U., D. Pruss and A. P. Wolffe. 1997. How does DNA methylation repress transcription? Trends Genet. 13: 444-449.
Kass, S. U. and A. P. Wolffe. 1998. DNA methylation, nucleosomes and the inheritance of chromatin structure and function. p. 22-35. In D. J. Chadwick and G. Gardew (eds): Epigenetics. John Wiley and Sons Ltd, Chichester, England.
Kermicle, J. L. 1970. Dependence of the *R*-mottled phenotype in maize on the mode of sexual transmission. Genetics 66: 69-85.
Kermicle, J. L. 1996. Epigenetic silencing and activation of a maize *r* gene. p. 267-287. In: V. E. A. Russo, R. A. Martienssen and A. D. Riggs (eds.), Epigenetic mechanisms of gene regulation. Cold Spring Harbor Laboratory Press, New York.
Kidwell, M. G. and D. R. Lisch. 1998. Transposons unbound. Nature 393: 22-23.

Klar, A. J. S. 1998. Propagating epigenetic states through meiosis: where Mendel's gene is more than a DNA moiety. Trends Genet. 14: 299-301.

Klein, T. M., M. Fromm, A. Weissinger, D. Tomes, S. Schaaf, M. Sletten and J. C. Sanford. 1988. Transfer of foreign genes into intact maize cells using high velocity microprojectiles. Proc. Nat. Acad. Sci. (USA) 85: 4305-4309.

Konstantinov, K., S. Mladenovic and M. Denic. 1991. Recombinand DNA technology in maize breeding. V. Plant and indigenous bacteria strains transformed by the gene controlling kanamycin resistance. Genetika (Beograd). 23: 121-135.

Koutsika-Sotiriou, M., I. Boss and A. C. Fasoulas. 1990. Hybrid reconstruction in maize. Euphytica 45: 257-266.

Koziel, M. G., G. L. Beland, C. Bowman, N. B. Carozzi, R. Crenshaw, L. Crossland, J. Dawson, N. Desai, M. Hill, S. Kadwell, K. Launis, K. Lewis, D. Maddox, K. McPherson, M. R. Meghji, E. Merlin, R. Rhodes, G. W. Warren, M. Wright and S. V. Evola. 1993. Field performance of elite transgenic maize plants expressing an insecticidal protein derived from *Bacillus thuringiensis*. Bio/Technology 11: 194-200.

Kricker, M. C., J. W. Drake and M. Radman. 1992. Duplication-targeted DNA methylation and mutagenesis in the evolution of eukaryotic chromosomes. Proc. Nat. Acad. Sci. (USA) 89: 1075-1079.

Kuckuck, H., G. Kobabe and G. Wenzel. 1991. Fundamentals of Plant Breeding. Springer-Verlag, Berlin.

Kumagai, M. H., J. Donson, G. Della Cioppa, D. Harvey, K. Hanley and L. K. Grill. 1995. Cytoplasmic inhibition of carotenoid biosynthesis with virus-derived RNA. Proc. Nat. Acad. Sci. (USA) 92: 1679-1683.

Kumar, S., X. Cheng, S. Klimasauskas, S. Mi, J. Posfai, R. J. Roberts and G. G. Wilson. 1994. The DNA (cytosine-5) methyltransferases. Nucleic Acids Res. 22: 1-10.

Langdale, J. A., W. C. Taylor and T. Nelson. 1991. Cell-specific accumulation of maize phosphoenolpyruvate carboxylase is correlated with demethylation at a specific site >3 kb upstream of the gene. Mol. Gen. Genet. 225: 49-55.

Lauster, R., T. A. Trautner and M. Noyer-Weidner. 1989. Cytosine-specific type II DNA methyltransferases - a conserved enzyme core with variable target-recognizing domains. J. Mol. Biol. 206: 305-312.

Lei, H., S. Oh, M. Okano, R. Jutterman, K. A. Goss, R. Jaenisch and E. Li. 1996. De novo DNA cytosine methyltransferase activities in mouse embryonic stem cells. Development 122: 3195-3205.

Leonardi, A., C. Damarval and D. de Vienne. 1987. Inheritance of protein amounts: comparison of two-dimensional electrophoresis patterns of leaf sheaths of two maize lines (*Zea mays* L.) and their hybrids. Genet. Res. 50: 1-5.

Leonardi, A., C. Damarval and D. de Vienne. 1988. Organ-specific variability and inheritance of maize proteins revealed by two-dimensional electrophoresis. Genet. Res. 52: 97-103.

Leonardi, A., C. Damerval, Y. Hebert, A. Gallais and D. de Vienne. 1991. Association of protein amount polymorphism (PAP) among maize lines with performances of their hybrids. Theor. Appl. Genet. 82: 552-560.

Leutwiler, L. S., B. R. Haigh-Evans and E. M. Meyerowitz. 1984. The DNA of *Arabidopsis thaliana*. Mol. Gen. Genet. 194: 15-23.

Lewis, J. D., R. R. Meehan, W. J. Henzel, I. Maurer Fogy, P. Jeppesen, F. Klein and A. Bird. 1992. Purification, sequence, and cellular localization of a novel chromosomal protein that binds to methylated DNA. Cell 69: 905-14.

Lin, B.-Y. 1975. Parental effects on gene expression in maize endosperm development. Ph.D. thesis.p. 190. Univ. Wisconsin, Madison.

Lin, B.-Y. 1982. Association of endosperm reduction with parental imprinting in maize. Genetics 100: 475-486.

Lin, B.-Y. 1984. Ploidy barrier to endosperm development in maize. Genetics 107: 103-115.

Lonnquist, J. H. 1967. Mass selection for prolificacy in maize. Zuchter 37: 185-187.

Malagnac, F., B. Wendel, C. Goyon, G. Faugeron, D. Zickler, J.-L. Rossignol, M. Noyer-Weidner, P. Vollmayr, T. A. Trautner and J. Walter. 1997. A gene essential for de novo methylation and development in *Ascobolus* reveals a novel type of eucatyotic DNA methyltransferase structure. Cell 91: 281-290.

Martienssen, B. 1998. Transposons, DNA methylation and gene control. Trends Genet. 14: 263-264.

Martienssen, R. A. 1996. Epigenetic silencing of *Mu* transposable elements in maize. p. 593-610. In: V. E. A. Russo, R. A. Martienssen and A. D. Riggs (eds.), Epigenetic mechanisms of gene regulation. Cold Spring Harbor Laboratory Press, New York.

Mascarenhas, J. P. 1990. Gene activity during pollen development. Annu. Rev. Plant Physiol. Plant Mol. Biol. 41: 317-338.

Matassi, G., L. M. Montero, J. Salinas and G. Bernardi. 1989. The isochore organization and the compositional distribution of homologous coding sequences in the nuclear genome of plants. Nucleic Acids Res. 17: 5273-5290.

Matzke, A. J. M., F. Neuhuber, Y. D. Park, P. F. Ambros and M. A. Matzke. 1994. Homology-dependent gene silencing in transgenic plants: epistatic silencing loci contain multiple copies of methylated transgenes. Mol. Gen. Genet. 224: 219-229.

Matzke, M. A. and A. J. M. Matzke. 1990. Gene interactions and epigenetic variation in transgenic plants. Dev. Genet. 11: 214-223.

Matzke, M. A. and A. J. M. Matzke. 1991. Differential inactivation and methylation of a transgene in plants by two suppressor loci containing homologous sequences. Plant Mol. Biol. 16: 821-830.

Matzke, M. A. and A. J. M. Matzke. 1995. How and why do plants inactivate homologous (trans)genes? Plant Physiol. 107: 679-685.

Matzke, M. A. and A. J. M. Matzke. 1996. Stable epigenetic states in differentiated plant cells: implications for somaclonal variation and gene silencing in transgenic plants. p. 377-392. In: V. E. A. Russo, R. A. Martienssen and A. D. Riggs (eds.), Epigenetic mechanisms of gene regulation. Cold Spring Harbor Laboratory Press, New York.

Matzke, M. A. and A. J. M. Matzke. 1998. Gene silencing in plants: relevance for genome evolution and the acquisition of genomic methylation patterns. p. 168-180. In D. J. Chadwick and G. Gardew (eds): Epigenetics. John Wiley and Sons Ltd, Chichester, England.

Matzke, M. A., M. Priming, J. Trnovsky and A. J. M. Matzke. 1989. Reversible methylation and inactivation of marker genes in sequentially transformed tobacco plants. EMBO J. 8: 643-649.

McClelland, M. and R. Ivarie. 1982. Asymmetrical distribution of CpG in an "average" mammalian genome. Nucleic Acids Res. 10: 7865-7877.

McClintock, B. 1957. Genetic and cytological studies of maize. Carnegie Inst. Wash. Yearb. 56: 393-401.

McClintock, B. 1958. The suppressor-mutator system of control of gene action in maize. Carnegie Inst. Wash. Yearb. 57: 415-452.

McClintock, B. 1959. Genetic and cytological studies of maize. Carnegie Inst. Wash. Yearb. 58: 415-429.

McClintock, B. 1963. Further studies of gene-control systems in maize. Carnegie Inst. Wash. Yearb. 62: 486-493.

McClintock, B. 1964. Aspects of gene regulation in maize. Carnegie Inst. Wash. Yearb. 63: 592-602.

McClintock, B. 1965. Components of action of the regulators Spm and Ac. Carnegie Inst. Wash. Yearb. 64: 527-534.

McClintock, B. 1971. The contribution of one component of a control system to versatility of gene expression. Carnegie Inst. Wash. Yearb. 70: 5-17.

Meehan, R. R., J. D. Lewis, S. McKay, E. L. Kleiner and A. P. Bird. 1989. Identification of a mammalian protein that binds specifically to DNA containing methylated CpGs. Cell 58: 499-507.

Meins, F. J. and C. Kunz. 1994. Silencing of chitinase expression in transgenic plants: an autoregulated model. p. 335-348. In: J. Paszkowski (eds.), Homologous recombination and gene silencing. Kluwer Academic Publishers, Dordrecht.

Messeguer, R., M. W. Ganal, J. C. Steffens and S. D. Tanskley. 1991. Characterization of the level, target sites and inheritance of cytosine methylation in tomato nuclear DNA. Plant Mol. Biol. 16: 753-770.

Meyer, P. (ed.). 1995. Gene silencing in higher plants and related phenomena in other eukaryotes. Springer-Verlag, New York.

Meyer, P. 1995. DNA methylation and transgene silencing in *Petunia hybrida*. p. 15-28. In: P. Meyer (ed.), Gene silencing in higher plants and related phenomena in other eukaryotes. Springer-Verlag, New York.

Meyer, P. and I. Heidmann. 1994. Epigenetic variants of a transgenic petunia line show hypermethylation in transgene-DNA: an indication for specific recognition of foreign DNA in transgenic plants. Mol. Gen. Genet. 243: 390-399.

Meyer, P., F. Linn, I. Heidmann, Z. A. H. Meyer, I. Niedenhof and H. Saedler. 1992. Endogenous and environmental factors influence 35S promoter methylation of a maize *A1* gene construct in transgenic petunia and its colour phenotype. Mol. Gen. Genet. 231: 345-352.

Meyer, P., I. Niedenhof and M. T. Lohuis. 1994. Evidence for cytosine methylation of non-symmetrical sequences in transgenic *Petunia hybrida*. EMBO J. 13: 2084-2088.

Mikula, B. C. 1995. Environmental programming of heritable epigenetic changes in paramutant r-gene expression using temperature and light at a specific stage of early development in maize seedlings. Genetics 140: 1379-1387.

Mitchell-Olds, T. and D. Pedersen. 1998. The molecular basis of quantitative genetic variation in central and secondary metabolism in *Arabidopsis*. Genetics 149: 739-747.

Mittelstein-Scheid, O., K. Afsar and J. Paszkowski. 1998. Release of epigenetic gene silencing by *trans*-acting mutations in *Arabidopsis*. Proc. Nat. Acad. Sci. (USA) 95: 632-637.

Montero, L. M., J. Filipski, P. Gil, J. Capel, J. M. Martinez Zapater and J. Salinas. 1992. The distribution of 5-methylcytosine in the nuclear genome of plants. Nucleic Acids Res. 20: 3207-3210.

Montgomery, M. K. and A. Fire. 1998. Double-stranded RNA as a mediator in sequence-specific genetic silencing and co-suppression. Trends Genet. 14: 255-258.

Motto, M. and R. H. Moll. 1983. Prolificacy in maize: a review. Maydica 28: 53-76.

Nan, X., H.-H. Ng, C. A. Johnson, C. D. Laherty, B. M. Turner, R. N. Eiseman and A. Bird. 1998a. Transcriptional repression by methyl-CpG-binding protein MeCP2 involves a histone deacetylase complex. Nature 393: 386-389.

Nan, X., S. Cross and A. Bird. 1998b. Gene silencing by methyl-CpG-binding proteins. p. 6-16. In D. J. Chadwick and G. Gardew (eds): Epigenetics. John Wiley and Sons Ltd, Chichester, England.

Napoli, C., C. Lemieux and R. Jorgensen. 1990. Introduction of a chimeric chalcone synthase gene into petunia results in reversible co-suppression of homolgous genes *in trans*. Plant Cell 2: 279-289.

Neuhuber, F., Y. D. Park, A. J. M. Matzke and M. A. Matske. 1994. Susceptibility of transgene loci to homology-dependent gene silencing. Mol. Gen. Genet. 224: 230-241.

Ngernprasirtsiri, J., R. Chollet, H. Kobayashi, T. Sugiyama and T. Akazawa. 1989. DNA methylation and the differential expression of C4 photosynthesis genes in mesophyll and bundle sheath cells of greening maize leaves. J. Biol. Chem. 264: 8241-8248.

Nick, H., B. Bowen, R. J. Ferl and W. Gilbert. 1986. Detection of cytosine methylation in the maize alcohol dehydrogenase gene by genomic sequencing. Nature 319: 243-246.

Palauqui, J. C., T. Elmayan, J. M. Pollien and H. Vaucheret. 1997. Systemic acquired silencing: Transgene-specific post-transcriptional silencing is transmitted by grafting from silenced stocks to non-silenced scions. EMBO J. 16: 4738-4745.

Pan, Y.-B. and P. A. Peterson. 1988. Spontaneous activation of quiescent *Uq* transposable element during endosperm development in *Zea mays*. Genetics 119: 457-464.

Paro, R., H. Strutt and G. Cavalli. 1998. Heritable chromatin states induced by the Polycomb and trithorax group genes. p. 51-61. In D. J. Chadwick and G. Gardew (eds): Epigenetics. John Wiley and Sons Ltd, Chichester, England.

Paszkowski, J., M. Baur, A. Bogucki and I. Potrykus. 1988. Gene targeting in plants. EMBO J. 7: 4021-4026.

Patterson, G. I. and V. L. Chandler. 1995. Paramutation in maize and related allelic interactions. p. 121-141. In: P. Meyer (eds.), Gene silencing in higher plants and related phenomena in other eukaryotes. Springer-Verlag, Berlin.

Patterson, G. I., K. M. Kubo, T. Shroyer and V. L. Chandler. 1995. Sequences required for paramutation of the maize *b* gene map to a region containing the promoter and upstream sequences. Genetics 140: 1389-1406.

Peschke, V. M., R. L. Phillips and B. G. Gengenbach. 1991. Genetic and molecular analysis of tissue-culture derived *Ac* elements. Theor. Appl. Genet. 82: 121-129.

Peterson, P. A.1965. A relationship between the *Spm* and *En* control systems in maize. Am. Nat. 44: 391-398.

Peterson, P. A. 1986. Mobile elements in maize: a force in evolutionary and plant breeding processes. Stadler Genet. Symp. 17: 47-78

Peterson, P. A. 1995. The development of transposon biology: from variegation to molecular confirmation. Maydica 40: 117-124.

Phillips, R. L., S. M. Kaeppler and P. Olhoft. 1994. Genetic instability of plant tissue cultures: Breakdown of normal controls. Proc. Nat. Acad. Sci. (USA) 91: 5222-5226.

Pirrotta, V. 1998. Polycombing the genome: PcG, trxG, and chromatin structure. Cell 93: 333-336.

Pradhan, S. and R. L. P. Adams. 1995. Distinct CG and CNG DNA methyltransferases in *Pisum sativum*. Plant J. 7: 471-481.

Prior, C. L. and W. A. Russell. 1975. Yield performance of nonprolific and prolific maize hybrids at six plant densities. Crop Sci. 15: 482-486.

Prols, F. and P. Meyer. 1992. The methylation patterns of chromosomal integration regions influence gene activity of transferred DNA in *Petunia hybrida*. Plant J. 2: 465-475.

Raina, R., M. Schlappi and N. Fedoroff. 1998. Epigenetic mechanisms in the regulation of the maize *Suppressor-mutator* transposon. p. 133-140. In D. J. Chadwick and G. Gardew (eds): Epigenetics. John Wiley and Sons Ltd, Chichester, England.

Razin, A. and T. Kafri. 1994. DNA methylation from embryo to adult. Prog. Nucl. Acid Res. Mol. Biol. 48: 53-81.

Rhounim, L., J.-L. Rossignol and G. Faugeron. 1992. Epimutation of repeated genes in *Ascobolus immersus*. EMBO J. 11: 4451-4457.

Richards, E. J. 1997. DNA methylation and plant development. Trends Genet. 13: 319-323.

Romagnoli, S., M. Maddaloni, C. Livini and M. Motto. 1990. Relationship between gene expression and hybrid vigor in primary root tips of young maize (*Zea mays* L.) plantlets. Theor. Appl. Genet. 80: 767-775.

Ronemus, M. J., M. Galbiati, C. Ticknor, J. Chen and S. L. Dellaporta. 1996. Demethylation-induced developmental pleiotropy in *Arabidopsis*. Science 273: 654-657.

Rossignol, J.-L. and G. Faugeron. 1994. Gene inactivation triggered by recognition between DNA repeats. Experimentia 50: 307-317.

Roth, E. J., B. L. Frazier, N. R. Apuya and K. G. Lark. 1989. Genetic variation in an inbred plant: variation in tissue cultures of soybean [*Glycine max* (L.) Merrill]. Genetics 121: 359-368.

Ruiz, M. T., O. Voinnet and D. C. Baulcombe. 1998. Initiation and maintenance of virus-induced gene silencing. Plant Cell 10: 937-946.

Russell, G. J., P. M. B. Walker, R. A. Elton and J. H. Subak-Sharpe. 1976. Doublet frequency analysis of fractionated vertebrate nuclear DNA. J. Mol. Biol. 108: 1-23.

Russell, W. A. and V. Machado. 1978. Selection procedures in the development of maize inbred lines and the effects of plant densities on the relationships between inbred traits and hybrid yields. Iowa Agr. Home Econ. Exp. Sta. Res. Bul. 585: 902-932.

Russo, V. E. A., R. A. Martienssen and A. D. Riggs. 1996. Epigenetic mechanisms of gene regulation. Cold Spring Harbor Laboratory Press, New York.

Salinas, J., G. Matassi, L. M. Montero and G. Bernardi. 1988. Compositional compartmentalization and compositional patterns in the nuclear genomes of plants. Nucleic Acids Res. 16: 4269-4285.

Sano, H., I. Kamada, S. Youssefian, M. Katsumi and H. Wabiko. 1990. A single treatment of rice seedlings with 5-azacytidine induces heritable dwarfism and undermethylation of genomic DNA. Mol. Gen. Genet. 220: 441-447.
Schindler, M. 1998. Select, microdissect, and eject. Nature Biotechnology 16: 719-720.
Schlappi, M., D. Smith and N. Fedoroff. 1993. TnpA *trans*-activates methylated maize Suppressor-mutator transposable elements in transgenic tobacco. Genetics 133: 1009-1021.
Schneeberger, R. G. and C. A. Cullis. 1991. Specific DNA alterations associated with the environmental induction of heritable changes in flax. Genetics 128: 619-630.
Schoffl, F., G. Schroder, M. Kliem and M. Rieping. 1993. An SAR sequence containing 395 bp fragment mediated enhanced, gene-dosage correlated expression of a chimeric heat shock gene in transgenic tobacco plants. Transgen. Res. 2: 93-100.
Schutze, K. and G. Lahr. 1998. Identification of expressed genes by laser-mediated manipulation of single cells. Nature Biotechnology 16: 373.
Schwartz, D. and E. Dennis. 1986. Transposase activity of the *Ac* controlling element in maize is regulated by its degree of methylation. Mol. Gen. Genet. 205: 476-482.
Selker, E. U. 1997. Epigenetic phenomena in filamentous fungi: useful paradigms or repeated induced confusion? Trends Genet. 13: 296-301.
Shen, J. C., W. M. Rideout, 3rd and P. A. Jones. 1994. The rate of hydrolytic deamination of 5-methylcytosine in double-stranded DNA. Nucleic Acids Res. 22: 972-6.
Shen, W. H., C. Ramos, B. Hohn. 1998. Excision of Ds1 from the genome of maize streak virus in response to different transposase-encoding genes. Plant Mol. Biol. 36: 387-392.
Shenoy, S., K. C. Ehrlich and M. Ehrlich. 1987. Repair of thymine cntdot guanine and uracil cntdot guanine mismatched base-pairs in bacteriophage M13mp18 DNA heteroduplexes. J. Mol. Biol. 197: 617-626.
Silander, J. A. 1985. Microevolution in clonal plants. p. 107-152. In: J. B. C. Jackson, L. W. Buss and R. E. Cook (eds.), Population biology and evolution of clonal organisms. Yale Univ. Press, New Haven.
Spena, A., A. Viotti and V. Pirrotta. 1983. Two adjacent genomic zein sequences: Structure, organization and tissue specific restriction pattern. J. Mol. Biol. 169: 799-811.
Steele-Scott, N., T. A. Kavanagh and J. N. Timmis. 1984. Methylation of rRNA genes in some higher plants. Plant Sci. Letters 35: 213-217.
Stelly, D. M., H. J. Price, R. E. Hanson, M. Zwick, M. Gomez, M. N. Islam-Faridi and C. F. Crane. 1997. Angiosperm polyploidization: incidence, mechanisms and ramifications. Book of abstracts. 9th Int. Cong. on Genes, Gene Families, and Isozymes., San Antonio, TX.
Stief, A., D. M. Winter, W. H. Straetling and A. E. Sippel. 1989. A nuclear DNA attachment element mediates elevated and position-independent gene activity. Nature 341: 343-345.
Stuber, C. W. 1994. Heterosis in plant breeding. Plant Breed. Rev. 12: 227-251.

Stuber, C. W. 1995. Mapping and manipulating quantitative traits in maize. Trends Genet. 11: 447-481.
Stuber, C. W. 1998. Biochemistry, molecular biology and physiology of heterosis. in press. In: J. Coors (eds.), The genetics and exploitation of heterosis in crops. Am. Soc. Agronomy, Madison, WI.
Styles, E. D. and O. Ceska. 1977. The genetic control of flavonoid synthesis in maize. Can. J. Genet. Cytol. 19: 289-302.
Surani, M. A. 1998. Imprinting and the initiation of gene silencing in the germ line. Cell 93: 309-312.
Szyf, M., G. Tanigawa and P. L. J. McCarthy. 1990. A DNA signal from the *thy*-1 gene defines *de novo* methylation patterns in embryonic stem cells. Mol. Cel. Biol. 10: 4396-4400.
Tate, P., W. Skarnes and A. Bird. 1996. The methyl-CpG binding protein MeCP2 is essential for embryonic development in the mouse. Nature Genet. 12: 205-8.
Theiss, G., R. Schleicher, R. Schimpff-Weiland and H. Follmann. 1987. DNA methylation in wheat. Purification and properties of DNA methyltranferase. Eur. J. Biochem. 167: 89-96.
Thompson, J. A., P. R. Drayton, B. R. Frame, K. Wang and J. M. Dunwell. 1995. Maize transformation utilizing silicon carbide whiskers: a review. Euphytica 85: 75-80.
Tollenaar, M. 1991. Physiological basis of genetic improvement of maize hybrids in Ontario from 1959 to1988. Crop Sci. 31: 119-124.
Tsaftaris, A. S. 1995. Molecular aspects of heterosis in plants. Physiol. Plant. 94: 362-370.
Tsaftaris, A. S. and M. Kafka. 1998. Mechanisms of heterosis in crop plants. J. Crop Prod. 1: 95-111.
Tsaftaris, A. S., M. Kafka, A. Polidoros and E. Tani. 1998. Epigenetic changes in maize DNA and heterosis. in press. In: J. Coors (eds.), The genetics and exploitation of heterosis in crops. Am. Soc. Agronomy, Madison, WI.
Tsaftaris, A. S., M. Kafka and A. N. Polidoros. 1997. Epigenetic modifications of total genomic maize DNA: The role of growth conditions. p. 125-130. In: A. S. Tsaftaris (eds.), Genetics, Biotechnology and Breeding of Maize and Sorghum. Royal Society of Chemistry, Cambridge, UK.
Tsaftaris, A. S. and A. N. Polidoros. 1993. Studying the expression of genes in maize parental inbreds and their heterotic and non-heterotic hybrids. p. 283-292. In: A. Bianci, E. Lupotto and M. Motto (eds.), Proc. XVI Eucarpia Maize and Sorghum Conference. Bergamo, Italy.
Van Slogteren, G. M. S., P. J. J. Hooykaas and R. A. Schilperoot. 1984. Silent T-DNA genes in plant lines transformed by *Agrobacterium tumefaciens* are activated by grafting and 5-azacytidine treatment. Plant Mol. Biol. 3: 333-336.
Vanyushin, B. F., S. G. Tkacheva and A. N. Belozerski. 1970. Rare bases in animal DNA. Nature 225: 948-949.
Vaucheret, H. 1993. Identification of a general silencer for 19S and 35S promoters in a transgenic tobacco plant: 90 bp of homology in the promoter sequence are sufficient for *trans*-inactivation. C. R. Acad. Sci. Paris. 316: 1471-1483.
Vaucheret, H., C. Beclin, T. Elmayan, F. Feuerbach, C. Godon, J-B. Morel, P. Mourrain, J-C. Palauqui and S. Vernhettes. 1998. Transgene-induced gene silencing in plants. Plant J. 16: 651-659.
Vilpo, J. A., S. Rasi, E. Suvato and L. M. Vilpo. 1986. An improved radioimmunoassay for quantitation of DNA methylation. J. Immun. Methods. 87: 179-183.
Voinnet, O. and D. C. Baulcombe. 1997. Systemic signaling in gene silencing. Nature 389: 553.
Vongs, A., T. Kakutani, R. A. Martienssen and E. J. Richards. 1993. *Arabidopsis thaliana* DNA methylation mutants. Science 260: 1926-1928.
Vyskot, B., B. Koukalova, A. Kovarik, L. Sachambula, D. Reynolds and M. Bezdek. 1995. Meiotic transmission of a hypomethylated repetitive DNA family in tobacco. Theor. Appl. Genet. 91: 659-660.
Wagner, I. and I. Capesius. 1981. Determination of 5-methylcytosine from plant DNA by high performance liquid chromatography. Biochim. Biophys. Acta 654: 52-56.
Wakimoto, B. T. 1998. Beyond the nucleosome: epigenetic aspects of position-effect variegation in *Drosophila*. Cell 93: 321-324.
Walbot, V. and C. A. Cullis. 1985. Rapid genomic change in higher plants. Annu. Rev. Plant Physiol. 36: 367-396.
Walker, E. L., T. P. Robbins, T. E. Bureau, J. Kermicle and S. L. Dellaporta. 1995. Transposon-mediated chromosomal rearrangements and gene duplications in the formation of the maize *R-r* complex. EMBO J. 14: 2350-2363.

Walter, C., I. Broer, D. Hillemann and A. Puhler. 1992. High frequency, heat treatment induced inactivation of a phospinothricin resistance gene in transgenic single cell suspension cultures of *Medicago sativa*. Mol. Gen. Genet. 235: 189-196.

Wassenegger, M., S. Heimes, L. Riedel and H. L. Saenger. 1994. RNA-directed *de novo* methylation of genomic sequences in plants. Cell 76: 567-576.

Waugh O'Neill, R. J., M. J. O'Neill and J. A. Marshall Graves. 1998. Undermethylation associated with retroelement activation and chromosome remodelling in an interspecific mammalian hybrid. Nature 393: 68-72.

Weber, H., C. Ziechmann and A. Graessmann. 1990. In vitro DNA methylation inhibits gene expression in transgenic tobacco. EMBO J. 9: 4409-4415.

White, S. and J. Doebley. 1998. Of genes and genomes and the origin of maize. Trends Genet. 14: 327-332.

Whitham, T. G. and C. N. Slobodchikoff. 1981. Evolution by individuals, plant-herbivore interactions, and mosaics of genetic variability: the adaptive significance of somatic mutations in plants. Oecologia 49: 287-292.

Williams, J. G. K., A. R. Kubelik, K. J. Livak, J. A. Rafalski and S. V. Tingey. 1990. DNA polymorphisms amplified by arbitrary primers are useful as genetic markers. Nucleic Acids Res. 18: 6531-6535

Wu, C. and M. L. Goldberg. 1989. The *Drosophila zeste* gene and transvection. Trends Genet. 5: 189-194.

Xiao, J., L. L. Yuan and S. D. Tanskley. 1995. Dominance is the major genetic basis of heterosis in rice as revealed by QTL analysis using molecular markers. Genetics 140: 745-754.

Yesufu, H. M. I., A. Honley, A. Rinaldi and R. L. P. Adams. 1991. DNA methylase from *Pisum sativum*. Biochem. J. 273: 469-475.

Yoder, J. A., C. P. Walsh and T. H. Bestor. 1997. Cytosine methylation and the ecology of intragenomic parasites. Trends Genet. 13: 335-340.

Zirkle, C. 1952. Early ideas on inbreeding and crossbreeding. p. 1-13. In: J. W. Gowen (ed), Heterosis. Iowa State College Press, Ames.