Optimization of the expression of a transgene in plants

Subhash C. Minocha
Department of Plant Biology
University of New Hampshire
Durham, NH 03824 USA
Tel: 603-862-3840
FAX: 603-862-3784
E-Mail; sminocha@christa.unh.edu

Chapter contents

1. Introduction
2. Regulation of gene expression
   - 2.1 cDNA vs. genomic DNA as coding sequence
   - 2.2 Promoters and 5'-regulatory sequences
   - 2.3 Post-transcriptional regulation - 5'UTRs
   - 2.4 Post-transcriptional regulation - 3'-end and 3' UTRs
3. Chromatin Organization and Matrix Attachment Regions
4. Transcriptional Factors
5. Gene Silencing and Co-suppression
6. Methods of transformation
7. Targeted transformation
8. Post-translational factors and organelle targeting of proteins
9. What are the limits to protein accumulation?
10. Future Perspectives
11. Acknowledgments
12. References

S.M. Jain and S.C. Minocha (eds.), Molecular Biology of Woody Plants, Vol. 1, 1-30
© 2000. *Kluwer Academic Publishers.*

1. Introduction

Transgenic manipulation of plants is now a well-established technology for crop improvement. In 1998, more than 30 million hectares of land were devoted to the cultivation of genetically modified crops (Moffat, 1998). While the numbers are impressive, plantations are presently limited to only a few herbaceous annual crops. However, for tree species, which potentially have the most to gain from genetic manipulation technologies (because long breeding cycles limit their genetic improvement through conventional breeding), relatively little progress has been made with the production of transgenic plants.

Transgenic manipulation involves the transfer of a recombinant gene into plant cells, selection of the transformed cells, and regeneration of these cells into whole plants (Birch, 1997). The transgenic plants are tested for expression of the transgene and its effectiveness in achieving the desired phenotype. Those plants exhibiting the desired phenotype are eventually used as breeding stocks to incorporate the transgenic trait into commercially valuable genetic backgrounds. Sufficient quantities of seed or vegetatively propagated plantlets are then produced for commercial plantation. The premise is that the expression of a single transgene of known function imparts the desired trait in the target tissues or in the whole plant. Thus far, the most successful examples have involved transgenes whose product (the polypeptide) directly provides a unique phenotype to the plant with minimal impact on cellular metabolism. Some examples of these genes include the Bt (*Cry*) genes for insect resistance, viral coat protein genes for virus resistance, and mutant proteins for herbicide tolerance. Only limited success has been achieved with genes that affect metabolic pathways. Likewise, genes with pleiotropic effects have rarely been used.

During the production of commercially grown varieties of genetically modified plants, and in numerous other studies where transgenic plants were tested in lab/greenhouse environments, we have learned a great deal about the expression of transgenes in plants. At the same time, we have encountered numerous unexpected problems in achieving desired levels of transgene expression in transgenic tissues/plants. The process, seemingly straightforward in design, is by no means so in practice. The fact that there are several steps from the transfer of a gene into a cell to the formation of a functional protein, each one regulated in a complex but precise manner in the natural system, has led us to analyze each step carefully. In a broad sense, these steps involve controlled transcription, production of a functional mRNA by post-transcriptional processing of the primary transcript, translation of the mRNA to produce a polypeptide, and post-translational modifications, processing and targeting of the polypeptide to cellular organelles. It is well known that superabundance of a transcript does not necessarily result in elevated protein levels due to limitations posed by the translational machinery (Bailey-Serres, 1999). Moreover, rapid turnover of mRNA and/or protein may counteract any effort at enhancement of biosynthesis. When attempting to manipulate complex metabolic pathways by genetic engineering, added complexity comes from the existence of multiple forms of regulation of the pathway by metabolite levels and the tight regulation of enzymes affecting the flux of these metabolites. Thus, designing transgenes for optimal expression in plants will require not only a thorough understanding of the regulation of normal (native) gene expression but also a reliable means of gene transfer and regeneration of transformed cells into whole plants.

Two things must be made clear at the outset. First of all, this chapter is not intended

to provide a comprehensive review of all aspects of gene regulation in plants nor of all manipulations of a transgene that could be done to achieve its maximum expression. It is conceivable that all manipulations of transgene constructs will not necessarily have additive or synergistic effects. Detailed reviews on different aspects of gene regulation in wild-type cells and attempts at achieving high expression of transgenes have recently appeared (Meyer, 1995; Koziel et al., 1996; Gallie, 1998 a,b; von Hippel, 1998; Udvardy, 1999). Secondly, there is a scarcity of published information on gene regulation and genetic transformation of woody plants in particular and thus, the chapter contains only limited references on these species. On the other hand, there is little doubt that the information obtained from herbaceous plants, and also from animal and microbial systems, is highly pertinent to future success with the transgenic improvement of woody plants, and therefore, this information is briefly reviewed here. The chapter is aimed at highlighting the contribution of different components of a transgene construct and the host cell as they interact with each other to regulate gene activity.

2. Regulation of gene expression

For a while, most of our concepts concerning regulation of gene activity were based upon work with microbes, yeast and a few animal models. Recent years have seen a tremendous spurt of activity in this area with higher plants, leading to the identification of parallel gene regulation systems and characterization of specific elements in gene regulation. In a broad sense, there are two distinct aspects of regulation of gene activity – a 'global' aspect and a specific 'local' aspect (Doebley and Lukens, 1998). The global aspect refers to the organization of chromatin, the characteristic *cis-* and *trans*-acting elements, and the polymerases that are common to all gene regulation (Zhou, 1999). Local and cellular factors impart spatial and temporal attributes to regulation of gene activity in a cell-specific manner. Although the two share some common features, the distinction between them is what leads to species differences on the one hand and cellular differentiation on the other. It is obvious that most organisms share common physiological functions that are the result of the existence of functionally similar enzymes. Yet, at the morphological level, each species is quite distinct in many ways. Doebley and Lukens (1998) ascribe the evolutionary origin of this variation to changes in *cis*-acting regulatory elements of transcriptional regulators much more than to changes in coding sequences *per se.* Built into this interpretation is the realization that different cells may utilize different *trans*-acting factors to regulate similar genes. This evolutionary change thus affects the spatial, temporal and developmental regulation of similar genes, leading to changes in morphology while maintaining similar overall biochemistry. At the cellular (local) level, we have already identified dozens of general transcription factors and an ever-increasing number of specific transcription factors that interact to control the timing and extent of transcription of a gene or a group of genes. The approach of transgenic expression of foreign genes must, therefore, be adjusted in such a way that it can optimally utilize the *trans*-acting elements present in a given cell type. The identification of gene-specific *cis*-acting elements and cell-specific *trans*-acting elements has already allowed us to experimentally manipulate gene activity (Koziel et al., 1996; Gallie, 1998 a,b), and should potentially aid us in the design of transgenes to achieve optimal expression at the desired time in a target cell type.

Active players in the production of functional proteins in the cell include: the coding

sequence of the gene, a variety of transcribed and untranscribed 5' and 3' regulatory elements, organization of the gene into chromatin, RNA polymerase and associated components of the transcription machinery, characteristics of the mRNA sequence that determine its processing, its turnover and its translational efficiency, the translational machinery of the cell, and numerous factors that determine post-translational modifications and targeting of proteins to specific sites in the cell. The system is thus composed of two classes of active participants: (1) those contained within the transgene sequence, and (2) those contributed by the host cell. Obviously, for transgene manipulation, one has a better control over the former than the latter. The role of some of the components of gene expression system, particularly those that are subject to experimental manipulation, is briefly reviewed here.

2.1 cDNA vs. genomic DNA as coding sequence

Obviously, the most important structural component of a transgene is its coding sequence, (i.e. the transcribed region) which could be derived either from genomic DNA or a cDNA. In studies on transgene expression in plants, cDNA sequences have generally been preferred as the coding sequence. Since cDNAs lack introns, it is assumed that the complications of intron removal and exon splicing will be avoided, thus yielding a functional mRNA as the direct transcript. Removal of introns from the primary transcript is an essential step for maturation of eukaryotic messages. The process involves an interaction of several *cis*-acting elements on the primary transcript with a large ribonucleoprotein complex called the spliceosome. Through a variety of conformational changes, as well as cutting and splicing events, a functional mRNA is put together. The process of splicing precursor mRNAs in higher plants has been reviewed by Simpson and Filipowicz (1996) and Schuler (1998). The formation of the 3' end and the polyadenylation site are a part of the intron removal machinery (Guo and Sherman, 1996). While introns themselves do not appear to possess any unique sequence characteristics, and intron-exon junctions are composed of relatively small recognition sequences, introns are often AU or U-rich sequences in higher plants (Brendel et al., 1998).

Several mutant phenotypes have been recognized as a consequence of faulty cutting and splicing of introns in plants (Fütterer et al., 1994; Luehrsen et al.,1994; Filipowicz et al., 1995; Keller, 1995; Muller et al., 1995). Whereas in mammals and in yeast, heterologous introns can be effectively removed, such is not the case in plants. Not only are animal introns not processed properly in most plants, major problems are also encountered when monocot introns are used for gene expression in dicots (Goodall and Filipowicz, 1991). Interestingly, the reverse, i.e. the removal of dicot introns in monocots, seems to create less of a problem. While many homologous as well as heterologous cDNA sequences have been successfully expressed as transgenes in plants, there are, however, some cases where the presence of introns has been shown to improve transgene expression (Luehrsen et al., 1994; Sinibaldi and Mettler, 1992). Koziel et al. (1996) have recommended the inclusion of some introns in a transgene as a means to optimize its expression in plants. Also, Ku et al (1999) attributed the presence of extremely high levels of maize phosphoenolpyruvate (PEP) carboxylase enzyme in transgenic rice to the use of genomic DNA which retained its original introns. Furthermore, it is known that some introns may themselves code for functional RNAs, e.g. those of ribozymes (Cech, 1993). It has also been proposed that inclusion of introns in the transgene construct may reduce the chances of cosuppression and silencing (Baulcombe, 1996). Since

all heterologous introns are not properly processed in plants, a better understanding of the mechanism of mRNA splicing is needed in order to achieve an improved design of the transgene sequence that would include introns which can be easily processed. Moreover, it is envisioned that it would be easier to design a transgene than to change the properties of host spliceosome because the latter is a very complex structure involving several small nuclear RNAs and at least 8 core proteins (Seraphin, 1995).

Whereas most eukaryotic genes do contain one or more introns, they are not a universal feature of all plant genes. Introns can be located within the translated sequence or they may be present in the 5' and 3' UTRs. The number of introns varies widely within different genes of the same species and within the same gene in different species. Typically, the number of introns may be less than 10, however, some plant genes have been known to contain as many as 31 introns (Roesler et al. 1994). The size of the introns is also quite variable, ranging from about 70 bases to as much as 7 kb. A typical intron is 100-200 bases long (Filipowicz et al., 1995; Simpson and Filipowicz, 1996). While most dicot introns are highly AU-rich (60-70% AU), monocot introns are less so (as low as 30-35% AU).

A typical 5' splice site for a plant intron is GU, which seems to be highly conserved (Schuler, 1998; Brendel et al., 1998). The flanking bases around the GU seem to contain a consensus sequence of AG/GUAAGU, which is quite similar to that found in mammals and yeast. Mutations in the GU pair often lead to faulty or no splicing. Likewise, the AG base pair located at the 3' splice site is also highly conserved in plant introns. Mutations in this pair of bases generally abolish the site specific processing, causing the next AG to become the site of splicing. The consensus sequence around the 3' AG splice site is often UGYAG/GU. There seems little doubt that the secondary structure of a primary transcript plays a crucial role in the processing of introns (Klaff et al., 1996). Thus, insertion of sequences within introns, changing the location of introns within a gene, and any alteration of the transcript that would lead to modification of secondary structure, could lower splicing efficiency or precision. Alternative splicing of the same transcript is highly prevalent in animals, where it plays an important role in the production of different messages from the same transcript in a tissue-specific manner (Moore et al., 1993; Valcarel et al., 1995). In plants, however, alternative splicing is not as common and its developmental significance is not clear.

Introns are not only important for normal gene expression, they also affect the pattern of gene expression in plants. For example, when a 5' UTR intron was removed from a sucrose synthase (*sus 4*) construct, its expression was highly reduced in potato tubers but much less in the roots (Fu et al., 1995a). Removal of a similar intron in another sucrose synthase (*sus3*) gene caused a reduction in its expression in the vascular tissue of tobacco anthers and a substantial increase in its expression in pollen (Fu et al., 1995b). It is further proposed that the regulatory effect of the intron is affected by the promoter and the 3' UTR elements.

While DNA sequences of most functional genes from any source can be directly expressed in plants, there are situations where this expression is quite low even when a strong promoter is used. Some prominent examples are the genes for jellyfish *gfp*, Bt protein (*CryA*), T4 lysozyme, and cyclodextrin glycosyltransferase gene of *Klebsiella pneumoniae* (Düring et al., 1993). Primary reasons for such low expression of these genes may be high A and U bias in the mRNA and the presence of potential cryptic splicing and polyadenylation sites in the coding sequence. In all cases, improved expression was observed when the gene sequence was modified to increase GC content. The change of certain AU sequences into GC-rich

sequences also results in the removal of potential cryptic introns, making the mRNA more stable and available for proper translation (van Aarssen et al., 1995). For example, for *CryA*(b) toxin gene of *Bacillus thuringiensis*, complete modification of the coding sequence was needed to obtain strong expression in plants (Perlak et al., 1991). A synthetic gene that was devoid of the ATTTA sequence and a potential plant polyadenylation site produced 100-fold higher expression than the wild type gene. Chiu et al. (1996) and Crameri et al. (1996) observed a strong expression of a *gfp* gene which was modified to make it highly G+C rich. Likewise, Rugh et al. (1998) used a modified coding sequence of a bacterial mercury reductase (*mer*) gene to achieve acceptable levels of its expression in poplar cells. However, Rouwendal et al. (1997) demonstrated that increased G and C content of a *gfp* gene only enhanced the frequency of transformants without affecting the level of GFP accumulation in the transformed plants (also see Haseloff et al., 1997).

2.2 Promoters and 5'-regulatory sequences

A promoter is perhaps the most crucial element in the regulation of transgene expression in an optimal fashion (Guilfoyle, 1997). A promoter is a 5' *cis*-acting regulatory element of a gene that determines when, where and how much of the gene will be transcribed. There are essentially three major categories of promoters: constitutive, developmentally-regulated, and inducible. A constitutive promoter is presumed to be active in all cell types at all times. While it is usually assumed that promoters of most house-keeping genes are constitutive, in reality, most of them show a high degree of variation in their regulation and expression. The best that can be said about a constitutive promoter is that it may show unregulated variable expression in different cell types and at different times. Several constitutive promoters have been successfully used to drive the expression of a variety of transgenes in plants. These promoters are very valuable in that their activity is generally high and is distributed in all cell types. Some common examples are CaMV 35S, actin, tubulin, and 16S ribosomal RNA promoters. They all are derived from genes that are abundantly expressed in plants. The most widely used of these promoters in transgenic research is the 35S promoter of CaMV. It has been generally classified as a strong constitutive promoter, however, it also shows different degrees of expression in different cell types (Williamson et al.,1989; Lemmetyinen et al., 1998).

Developmentally-regulated or tissue-specific promoters are those that are active only at certain stages of development and only in specific tissues. These promoters provide excellent fine-tuning of transgene expression in a spatial and temporal fashion. Transgenes that may have deleterious effects in some tissues or are needed to be expressed only at specific stages of development would be ideally regulated by such promoters. A few examples of the successful use of such promoters in transgenic research are: fruit-specific promoters for genes that control ripening, anther-specific promoters for induction of male sterility, LEA protein promoters for gene expression during embryo development and maturation, embryo specific promoters, and a variety of tissue-specific promoters. A subgroup of developmentally-regulated and cell-specific promoters is organelle specific promoters that can be useful for organelle transformation studies without affecting the cytoplasmic protein synthesis system (Daniell et al., 1998). A unique advantage of using homologous cell-specific and developmentally-regulated promoters is that all the transcription machinery for their activity is already present in the cells in which transgene expression may be desired.

The third and the most versatile category of promoters is the inducible promoters, which respond to a variety of physical and chemical factors. Some of the physical factors that induce specific promoter activity are light, high and low temperature, and water-stress. On the other hand, there is a large number of chemically-induced promoters (Gatz and Lenk 1998; Rossi and Blau, 1998) which respond to both internal (plant) factors as well as treatment with synthetic chemicals. The best-studied chemically-inducible promoters are hormone-induced promoters and metabolite-inducible promoters. These promoters continually respond to changing concentrations of the inducer to provide desirable levels of activity in different tissues. Exogenously applied chemicals, e.g. antibiotics, herbicides, ozone, non-plant products, copper and other heavy metals, and a variety of synthetic analogs of naturally occurring biological compounds have been used to activate transgene expression in plants. The induction is a direct positive interaction of the inducer with a cellular protein factor, in most cases, a repressor protein to induce gene expression (Gatz and Lenk, 1998).

Promoters whose expression is regulated by a specific chemical have proven uniquely valuable for microbial systems in aiding commercial fermentation technology as well as helping us understand gene regulation. Common examples include various sugar-induced promoters of bacteria. A series of chemically-inducible promoters have also been used to analyze the effects of ectopic transgene expression in animals. In plants, however, such chemically inducible promoters have only recently become available. Here again, such promoters are valuable for expression of transgenes in crops in the field as well as in delineating the role of specific proteins in normal plant growth and development. They are particularly desirable in situations where overexpression of a transgene may be toxic or may interfere with normal plant development, organogenesis and somatic embryogenesis. Inducible gene expression allows an evaluation of direct vs. indirect effects of the transgene product on plant phenotype. For tree species, such promoters will be highly suitable for regulating conditional male sterility systems, overcoming interference with regeneration of plants from transgenic cells, managed expression of pest resistance genes, and regulation of developmental processes, such as leaf abscission and flowering.

Gatz and Lenk (1998) list some of the properties that inducible promoter systems should have for desirable transgene expression in plants. Some of them are: high specificity of the inducer, lack of toxicity, acceptable environmental compatibility, high efficiency at low concentrations, low cost, ease of application of the inducing chemical in the field, and low or high half-life of the inducer (depending upon the need for short-term or long-term persistence of induction). A few examples of promoter systems that are currently available for use and meet some of these criteria are: tetracyclin-inducible promoters, copper-inducible promoters, ethanol-inducible promoters, and steroid hormone-inducible promoters.

A general approach used to produce chemically-regulated promoters is to combine the *cis*-active regulatory sequences of a highly efficient constitutive or a developmentally-regulated or tissue-specific promoter with a chemically-regulated domain of a microbial promoter region. The best documented of these is the *tet*-inducible promoter, which involves the fusion of a constitutive 35S promoter of CaMV with a *tet*-regulated elements of a bacterial promoter. The latter regulates activity of the 35S promoter in *cis* position through binding of another bacterial gene product (the *tet* repressor) which is constitutively expressed as a separate transgene in the cell. The presence of tetracyclin in low, non-toxic concentrations removes repression by its binding to the *tet*-repressor protein (Gatz et al., 1992; Gatz, 1997).

The value of such a system has been demonstrated by regulation of the overexpression of an isopentenyl transferase gene, whose product causes overproduction of cytokinin and triggers an aberrant phenotype (Motyka et al., 1996; Faiss et al., 1997). Such a system should also be uniquely suitable in studies involving antisense expression of metabolically critical genes. A modification of the *tet*-inducible promoter was described by Weinman et al. (1994) who made it into a *tet*-repressible promoter. Such a promoter will aid in studies aimed at analysis of turnover rates of transgenic messages and proteins. Similar approaches of fusing a constitutive promoter with inducible regulatory domains of bacterial or eukaryotic promoters have also been used to produce steroid hormone-inducible promoters (Lloyd et al., 1994; Simon et al., 1996; Aoyama and Chua, 1997; McNellis et al., 1998), copper-inducible promoters (Mett et al., 1993), and ethanol-inducible promoters (Caddick et al., 1998). Further modifications of such chimeric promoters should open up endless possibilities of achieving organ-specific or developmentally-regulated inducible expression of transgenes in all plants (Mett et al., 1996).

While spraying large acreage of tree plantations with a chemical (inducer) is feasible, its appropriateness will have to be evaluated in terms of the benefits achieved. For some phenotypic traits (early flowering, transgenically-induced male sterility or reversal of constitutive male sterility, etc.), chemically-induced promoters will provide unique advantages for transgenic tree crops. For other traits (e.g. expression of insect resistance genes – Bt or protease inhibitor), it may be desirable that the promoter is responsive to an *in situ* produced chemical, whose production is controlled, for example, by wounding as a result of insect bites. Pathogen-inducible and pest-inducible promoters are additional examples of promoters that may be useful for conditional expression of genes which can impart tolerance to bacterial, fungal and insect pests (for review see Rushton and Somssich, 1998).

An excellent example of creating developmentally regulated promoters with widespread potential use is the *cre-lox* system (Odell et al., 1994). The system involves generation of a functional promoter (may be originally a constitutive promoter) by removal of deliberately inserted sequences between the promoter domains and the TATA box. The system takes advantage of a recombinase that can recognize the flanking regions of the insert, and whose production is controlled by another transgene under the control of a strong developmentally-regulated promoter. The usefulness of such a system has been demonstrated for controlling the development of zygotic embryos in developing seeds by turning on certain suicidal genes within the embryo (a controversial technology termed as 'terminator' technology by some critics - Oliver et al., 1998; Lehmann, 1998; Radin, 1999). One can exploit such a system to control almost any aspect of growth and development in plants by using a chemical inducer to regulate recombinase production or to suppress recombinase production at any time. The use of proprietary chemicals to regulate recombinase production is being developed as a means to protect the patented transgenic germplasm from unauthorized multiplication for commercial use. The system provides unparalleled potential applications, not only in controlling transgene expression but also in removal of unwanted segments of the transgenic DNA from the genome, e.g. removal of an antibiotic resistance gene after selection of transgenic cells.

2.3 Post-transcriptional regulation - 5'UTRs

Although the relative amounts of mRNA present in a tissue are often considered strong

indicators of transgene expression, it cannot be assumed that these mRNAs are being translated with the same efficiency to produce proportionate amounts of proteins. Before translation begins, eukaryotic transcripts undergo a series of structural modifications to produce a functional mRNA. These modifications involve 5' capping, intron removal, 3' truncation, and 3' polyadenylation (Bailey-Serres and Gallie,1998). Most requirements for these processing events are inherent in the transcript sequence, i.e. they are derived from the transcribed DNA sequence and, therefore, can be experimentally manipulated. The process of translation itself, which involves three distinct sets of events (initiation, elongation and termination), is very complex (Browning, 1996; Browning et al., 1998). It is affected by cellular factors (growth regulators, developmental stage of cell, etc.), as well as external environmental stimuli (stress factors, temperature, light, etc.). The most crucial and rate-limiting step for translation is generally ribosomal recruitment (translation initiation). This step involves several interactive events, some that are governed by structural features of the mRNA and others that depend upon the metabolic status of the cell. The presence of numerous cytoplasmic factors is also essential for this step to proceed normally. Among the mRNA characteristics that affect rates of translation initiation, including recycling or re-initiation events, are the presence or absence of the cap, the 5'-UTRs, the 3'-UTRs, the nucleotide sequence context of the starting AUG codon, and 5'-cap and 3'-poly(A) tail interactions.

There is little doubt that selective translation of mRNAs is a common phenomenon (for review, see Bailey-Serres, 1999). Thus transcript abundance is not always an indicator of high translation efficiency. Intrinsic properties that affect the overall rates of specific protein synthesis involve storage of messages as translationally inactive ribonucleoprotein particles and mRNA turnover. Our knowledge of mechanisms by which mRNA structural features, including *cis*-acting elements around the open-reading frame (ORF), determine message stability and its ability to be efficiently translated is, however, rather limited (Abler and Green, 1996; Marcotte, 1998; Johnson et al., 1998).

Selective degradation of mRNA is often used by the cells to regulate cellular polypeptide levels. The half-lives of mRNAs vary from a few minutes to several hours. In the case of eggs, zygotes and seeds, stable ribonucleoprotein particles can last up to several months. The turnover of mRNA is promoted by AU-rich elements (AREs) in the 3' UTR domains and is often initiated in association with translation (Laroia et al., 1999). In animals, a group of ARE-binding proteins are postulated to promote mRNA degradation, whereas, release of stored Ca^{++} or an overexpression of other ARE-binding factors can temporarily stabilize mRNA. The details of how mRNAs intended for selective degradation are recognized and how they are actually degraded are not known. The steady-state levels of different mRNAs are quite different, but often stable in a particular cell type indicating a tight homeostatic regulation of specific mRNA levels.

Unlike the process of transcription initiation of a transgene, which depends upon the presence of numerous *tans*-acting factors in addition to the accessability of the DNA sequence (Zhou, 1999), information for the initiation of translation appears largely to be contained in the final mRNA sequence. Once the primary transcript has been processed to generate a functional mRNA, like most other messages in the cell, it becomes available for translation by ribosomes and the associated translational machinery. Little is presently known about the competition between the transgenic messages and the native messages for factors involved in translation and the turnover of these messages. While some reports have implicated specific

recognition mechanisms that may lead to selective degradation of a message, most of them are based upon mRNA being aberrant, not properly processed, or subject to interaction with antisense messages. It has also been reported that translation initiation can be improved by using a proper consensus sequence surrounding the initiation codon, e.g. (A/G)(A/C)cAUGGC in monocots and A(A/C)aAUGGC in dicots (Fütterer and Hohn, 1996; Joshi et al., 1997).

Two distinct mechanisms have been proposed for initiation of translation of mRNAs, one for 5'-capped mRNAs and the other for those which lack this cap. The translation of capped messages often begins with the recognition of 5'-^{7m}GpppN cap of mRNA by one or more proteins (cap binding proteins – CBPs) that bind to this site and act as anchors for the small ribosomal subunit binding to message (Browning, 1996; Pain, 1996; Browning et al., 1998). This complex scans the 5' end through 5' UTRs to reach the initiation codon before the large ribosomal subunit associates with it. Insertion of stable hairpins into the 5' UTR reduces translation efficiency. Also insertion of additional AUG triplets upstream of the original start site can cause initiation to begin at the new site (Kozak, 1991). Initiator tRNA and several eukaryotic initiation factors (eIFs) are involved at this step. It has been suggested that CBPs may play an important role in selectivity of mRNA translation through their abundance and activity (Rodriguez et al., 1998). While cap-dependent initiation is most frequent in eukaryotes, cap-independent translation of mRNAs also occurs, e.g. for some viral mRNAs (Fütterer and Hohn, 1996; Hohn et al., 1998). Eukaryotic mRNAs that lack the cap are translated by employing internal-ribosomal-entry-segments (IRES) for initiation.

The length and the sequence of 5' UTRs, and the nucleotides surrounding the initiation codon (AUG) all play important roles in translation initiation. The 5' UTRs of most plant messages are less than 100 nucleotides long, however, unusually long 5' UTRs have been observed in some mammalian mRNAs (e.g. ornithine decarboxylase - ODC and S-adenosylmethionine decarboxylase - SAMDC mRNAs - Kahana and Nathans, 1985; Van Steeg et al., 1991; Shantz et al., 1994). In some cases, one or more functional or non-functional ORFs and complex secondary structures can be found in the 5' UTRs (Kahana and Nathans, 1985). In the case of 5'-capped mRNAs, a strong interaction between the scanning complex on the 5'-terminus and the 3'-terminal poly(A) tail is envisioned through the binding of poly(A)-binding protein. This interaction thus effectively circularizes the mRNA before translation actually commences. Therefore, a proper design of the 5' end of the transgenic mRNA must be ensured to achieve efficient translation. The 5' UTRs are often poor in GC or AU tracts which could form hairpin structures (Pesole et al., 1997). The 5' UTRs have been implicated in various aspects of the regulation of translation of these messages (Shantz et al., 1994; Van Steeg et al., 1991). Most of the transgenic work in plants has utilized 5'-capped mRNA production. Viral mechanisms of translation (some of which are highly efficient - Gallie, 1996) have not been exploited (for a review of translation initiation mechanisms, see Pestova and Hellen, 1999). While the 5' UTRs may provide developmental regulation or response to external stimuli, they can be highly reduced in length without major effects on faithful translation of the message.

It has been shown that extraordinarily long 5' and 3' UTRs of a mouse ODC gene (737 and 342 nucleotides, respectively; Kahana and Nathans, 1985) were not required for the expression of its cDNA in plants. Plant cells transformed with cDNAs lacking over 90% of the 5' UTR, or even the entire 5' UTR, and the entire 3' UTR sequences, strongly expressed

these cDNAs to produce functional enzymes (DeScenzo and Minocha, 1993; Bastola and Minocha, 1995; Glasheen, Bains, Minocha and Minocha, unpublished data). Removal of the coding sequence for 37 of the C-terminus amino acids (which constitute a PEST region responsible for instability of the ODC enzyme - Ghoda et al., 1989; Pegg, 1989) further increased enzyme activity in the transformants, probably due to increased stability of the enzyme. On the other hand, Grens et al., 1990 showed that the addition of the 3' UTR of ODC between the poly(A) site and the coding sequence relieved the 5'UTR-imposed inhibition of a reporter gene. Thus the multipurpose role of the 5' UTRs are quite complex and variable, and affect all aspects of mRNA stability and translation efficiency.

2.4 Post-transcriptional regulation - 3'-end and 3' UTRs

The 3' end of a eukaryotic mRNA plays major roles in both translation initiation (by interaction with the translation initiation complex bound to the 5'-end) and mRNA stability. The 3' region is often made of a 3' UTR which averages about 200-300 nucleotides, and a poly(A) tail of up to 200 nucleotides (Rothnie, 1996; Hunt and Messing, 1998). The 3' UTR as well as the poly(A) tail have been implicated in determining the stability of mRNAs. Whether the transcription termination sequences are recognized by RNA polymerase or by factors that mediate disassembly of the transcription complex is not known. In animals, variable sites for transcription termination of the same transcript are often known to occur (Proudfoot, 1989) while in yeast transcription termination and 3'-end processing are tightly coupled (Guo and Sherman, 1996). In plants, like animals, transcription termination occurs several hundred nucleotides downstream from the actual cleavage and polyadenylation site. Its structure and sequence specificity are not clearly defined. It is believed that the 3' polyadenylation site is created post-transcriptionally by a self-cleavage processing event. A few important features of this processing are: (1) the polyadenylation site is variable in plant mRNAs while it is a single unique site in animal mRNAs, the result being that plants contain a variable size population of functional mRNAs for the same gene, often in the same tissue; (2) poly(A) site cleavage can occur *in vitro* in cell-free transcription systems, showing that *cis*-acting elements are primarily responsible for the process; (3) while in animals the cleavage site is flanked by a conserved AAUAAA motif and U-rich downstream elements, in plants no single conserved motif is involved; and (4) whereas viral or heterologous plant 3' termination and poly(A) signals are effective in transgene translation in plant cells, animal 3'-termination signals are not properly recognized in plants (Hunt et al., 1987).

The observation that the same transcribed sequence can be represented as a highly heterogenous population of mRNAs (as many as 14 distinct 3' cleavage sites were reported for *Nicotiana plumbaginifolia* chloroplast RNA-binding protein - Klahre et al., 1995) is intriguing and raises an interesting question about the significance of this variation in translation efficiency. Are we better off using conserved (highly favored) sites for 3'-end cleavage or should we use heterologous 3' domains which allow variable cleavage sites for expression of a transgene? Detailed comparisons of messages with different sites of polyadenylation have not yet been reported to provide satisfactory answers to this question. However, the importance of *cis*-acting elements near the site of cleavage (called near upstream elements – NUEs) and some further upstream from this site (called further upstream elements – FUEs) has been well documented (see Rothnie, 1996 for review). Another important element is the

cleavage site itself. While the precise roles of the NUEs, the FUEs and the cleavage site in transcript processing are controversial, it seems that downstream sequences do not usually play a major role in determining the site of cleavage and polyadenylation. In most transgenic experiments, *Agrobacterium* T-DNA 3' termination sequences or a CaMV 35S termination sequences have been used with high degrees of success. Both these sequences possess multiple cleavage and polyadenylation sites.

In spite of the diversity observed in the FUEs, NUEs and the cleavage site itself, it is believed that a suitable combination of the three elements is required for efficient and proper processing. The role of secondary structure in this region (demonstrated in yeast poly(A) signals and in some plant pararetrovirus pregenomic RNAs) has not been widely studied in plants (Fütterer et al., 1988, 1994). In mammalian cells, complete processing of the 3' end of an mRNA involves interactions between 'cleavage and polyadenylation specificity factor' (CPSF), 'cleavage stimulation factors' (CstF), two 'cleavage factors' (CFI and CFII), poly(A) polymerase (PAP) and, in some cases, a poly(A) binding protein (PAB). While details of the 3' end processing complex are not clear as yet, it is believed that a similar complex exists in plants. It is further observed that during early stages of embryo development, when several pre-stored mRNAs are translated, the length of poly(A) tail plays a crucial role in their translation efficiency (Richter, 1993; Curtis et al., 1995). Thus, in spite of our poor understanding of the processing of 3'-terminus of mRNA, we must still carefully choose a 3'-termination sequence for construction of a functional transgene. Until detailed comparisons of different 3' termini are made, and the mechanistic role of the 3' end in translation is fully understood, we must rely upon empirical experimentation with known 3' termination sequences of some plant, viral or agrobacterial genes.

3. Chromatin Organization and Matrix Attachment Regions

The structural complexity and functional significance of chromatin have always been topics of hot discussion (Marcand et al., 1996; Björklund, et al., 1999). It is now commonly believed that functionally active domains of chromatin are somehow different from those that are either silent or less active. The activity of chromatin depends upon modifications of both histones and DNA. Histone modifications often include hyperacetylation, phosphorylation and methylation in the regions of active chromatin, while DNA in the inactive regions of chromatin is often associated with hypermethylation of CG pairs (Martienssen and Richards, 1995). What determines the extent and the timing of these modifications is not known. In addition, considerable attention is now being focused on the role of so-called Matrix Attachment Regions (MARs) or Scaffold Attachment Regions (SARs) of DNA in the genome (Spiker and Thompson, 1996; Gallie, 1998a; Holmes-Davis and Comai, 1998). MARs are 300-800 bp long sequences of DNA that are found throughout the genome. They are hypothetically involved in anchoring DNA to a nuclear matrix, thus allowing formation of loops of the unanchored portions of DNA. The size of these loops is variable, averaging about 85 Kb (Benyajati and Worcel, 1976). Each loop consists of two anchor points (fixed boundaries) and a third variable point which functions as the site of replication or transcription. It is postulated that MARs divide the genomic DNA into functional domains that may be regulated differently for gene activity and may also function as potential units for DNA replication and chromatin condensation (Udvardy, 1999). Open chromatin domains within the loop are presumably

active for transcription and condensed or compact domains are not. It is further believed that the anchor points insulate different sections of the genome from adjoining transcription regulatory elements (Udvardy, 1999). In other words, genes within a loop may be influenced by the same group of transcription regulatory elements, thus reducing variation in gene activity within the loop.

It has been observed that highly expressed genes are often closely flanked by MARs while others (less expressed) may be distant from MARs, the former constituting smaller, less compact loops of transcriptionally active DNA. MARs are generally AT-rich domains that often contain conserved sequences. These domains are interchangeable among the genes when used in transformation experiments. Based upon these observations, several workers have utilized cloned regions of MARs to achieve enhanced transgene activity by incorporation of their sequences as flanking regions on the transgene. The results of these studies have been reviewed in Holmes-Davis and Comai (1998).

Breyne et al. (1992) demonstrated that a soybean MAR flanking a transgene resulted in reduced variation in its expression in different transformants. Mlynárová et al. (1994, 1995) also showed reduced variability in transgene expression in tobacco by using a chicken MAR sequence. Further analyses of the role of MARs in transgene expression have revealed that MARs (e.g. Rb7 MAR of tobacco) which bind more strongly to the matrix result in higher levels of expression than weaker-binding MARs (Allen et al., 1996). These authors suggest that the MARs may prevent homologous gene silencing (see discussion below) where multiple copies of a transgene may be integrated into the genome. While some studies have shown that only a 5' MAR may be sufficient for enhanced gene expression (van der Geest and Hall, 1997), others have emphasized the need for both 5' and 3' MARs (Chinn et al., 1996). While assigning specific roles to MARs in transcription (e.g. boundary function for transcription regulation or simply regulating transcription through formation of small functional loops of chromatin) may not be easy at present, their usefulness in the enhancement of transgene activity and their role in reducing gene silencing may still be exploited in transgenic manipulation of all plants, including woody species. Several variations in the design of constructs containing MARs in terms of their location, size, source, spacing vis-a-vis the promoter, etc. must be experimentally tested to take full advantage of their role in regulating transgene expression.

Wolffe (1994) has suggested that the positioning of regulatory domains of genes within the nucleosomes could also play an important role in gene regulation through their accessibility to the transcriptional factors and RNA polymerase. It is argued that in such nucleosomes, contact points for start and finish of DNA-histone binding are at defined sites and are influenced by the DNA sequence in these domains. As a result, the regulatory DNA sequences that are recognized by transcription factors are exposed (facing outward) or are found in linker DNA. It is conceivable that these DNA sequences may even be in contact with specific histones. It is further argued that both DNA sequences and some *trans*-acting factors determine the sites for DNA-histone interactions during nucleosome assembly. In a functional sense, the presence of regulatory sequences in the linker regions that are 160 bp apart could bring two regulatory sequences in close proximity for interaction with other transcription factors. Some of the supporting evidence for such arguments comes from studies with *Drosophila* and yeast where, for example, a nucleosome positioned between an enhancer and the promoter could bring these two elements in close proximity. Likewise, two regulatory

domains separated by 80 bp could be brought together on the nucleosome surface due to one turn of DNA around the histone octamer. Each turn around the octamer takes 80 nucleotides.

A lack of understanding of the mechanism of nucleosome organization prevents us from designing transgene sequences for optimal nucleosome positioning. Moreover, the insertion site of a transgene (which is highly unpredictable) will play a crucial role in determining its positioning in the nucleosome because of the fixed distance between the nucleosomes. If these arguments are indeed valid, this could explain the position effects on transgene activity without involving other complicated mechanisms of variation in transgene activity in different transformants (for review, see Wolffe, 1994).

4. Transcriptional Factors

Successful transcription of a transgene is determined by two factors: (1) composition of the transgene construct, i.e. the *cis*-acting factors and (2) the host cell components involved in transcription, i.e. the *trans*-acting factors. Elements that are a part of the transgene construct are: the type and the number of transcription factor-binding sites in the target promoter, the potency of activation domains, and the degree of their interaction with other proteins involved in transcription activation. It is believed that a closer match between the global structure of the genes of the host cell and that of a transgene makes it more likely to be successfully transcribed. A variety of *trans*-acting transcription factors work sequentially and cooperatively to initiate and continue transcription of a gene. It has been shown that experimental manipulation of the transcription factors can be used to improve transcription efficiency in plants (Schwechheimer et al., 1998). While many of them are constitutively present in most cells, others are produced in a tissue specific manner to ensure optimal transcription in a spatial and temporal fashion. Since transcription occurs in the nucleus and the transcription factors, like most other proteins, are synthesized in the cytoplasm, their import into the nucleus is subject to regulation by the presence of nuclear localization signals as well as the transporter proteins that shuttle between the nucleus and the cytoplasm (Raikhel, 1992). With our increasing understanding of how different proteins (mostly involved with signal transduction pathways and also transcription) are targeted to the nuclear pore (see Hicks and Raikhel, 1995; Smith and Raikhel, 1999 for reviews), it should be possible to transport modified transcription factors to the nucleus to regulate transcription of a transgene as well as the host genes.

The RNA Pol-II complex is composed of a large number of different subunits which include general transcription factors, TATA-binding proteins (TBP), several TBP-associated factors, and several coactivator proteins (Lewin, 1997; Meisel and Lam, 1997; Zhou, 1999). In addition, several cell/tissue-specific *trans*-acting factors interact with the promoter or enhancer regions to control gene activity. The final interactions between distantly-located enhancers, the 5' regulatory domains, and the *trans*-acting factors that eventually expose the DNA in organized chromatin to the RNA polymerase complex still remain a mystery. Undeniably, DNA sequences in all of the regulatory domains play an important role, not only in transcription initiation, but also in transcription re-initiation.

Some well known transcription factors include: MYB-like proteins, MADS-domain proteins, basic region-leucine zipper (bZIP) proteins, helix-loop-helix proteins, zinc-finger proteins, and homeobox proteins (Pabo and Sauer, 1992; Meshi and Iwabuchi, 1995; Meisel and Lam, 1997). These proteins, in association with several other tissue-specific or inducible

factors, recognize DNA and bind to it in a sequence specific manner (Ranish and Hahn, 1996). The transcription factors include both transcription activators and suppressors. Each one of them is subject to signal-specific phosphorylation and dephosphorylation at specific amino acids by a variety of kinases and phosphatases. Thus, the mere presence of a transcription factor is not sufficient for its function.

The assembly of all transcription factors at the site of transcription along with RNA polymerase is a very complex process that, in spite of a vast amount of data available, is not well understood (Meisel and Lam, 1997; Schwechheimer and Bevan, 1998). Furthermore, the synthesis of some of the inducible and developmentally-regulated transcription factors is itself subject to regulation (Schwechheimer et al., 1998). Obviously, regulated expression of a gene is the result of complex multi-layered regulatory steps that have evolved over time to ensure proper functioning of the system. The desired expression of a transgene then must fit into this hierarchical control mechanism of the cell in order to yield the expected protein at the desired time in the desired tissue. Having little control over the regulation of these transcriptional factors renders the expression of transgenes vulnerable to tremendous variation due to the cellular environment.

5. Gene Silencing and Co-suppression

It has often been observed that transgene expression in many plants is silenced in a developmentally regulated manner or in a stable manner (see reviews by Meins, 1996; Depicker and Van Montagu, 1997; Matzke and Matzke, 1995, 1998a; Wassenegger and Pélissier, 1999). Two types of silencing phenomena are observed: (1) those where silencing is stable and heritable, and (2) those where silencing is less stable and meiotically not inherited. The former generally involves hypermethylation of the transgene or its promoter, and affects transcription. The latter, however, appears to be generally a post-transcriptional phenomenon (called post-transcriptional gene silencing - PTGS), and does not involve DNA hypermethylation; it is related instead to mRNA instability. PTGS is also affected by certain environmental factors, is dependent upon number of transgene copies, and often requires sustained RNA synthesis (English et al., 1997; Vaucheret et al., 1997; Wassenegger and Pélissier, 1998). Altered RNA degradation and the synthesis of aberrant mRNAs are commonly associated with PTGS (Baulcombe, 1996; Meins, 1996; Mette et al., 1999). When two different transgenes are inserted into the cell, they often behave independently with respect to silencing, i.e. one may remain fully functional while the other may be silenced (Leech et al., 1998; Holtorf et al., 1999). Moreover, degradation of mRNA in silenced plants is not related to its translation.

Several models have been proposed to explain PTGS (Baulcombe, 1996; Johnson et al., 1998); two major ones being the 'ectopic pairing' model and the 'RNA threshold' model. The latter seems to have more support in the literature and can explain both stable and unstable PTGS. According to this model, there is a sensing mechanism to monitor the total concentration of similar transcripts. When the total concentration crosses a threshold, a sequence-specific degradation system is activated to degrade this RNA. The process involves the activation of systemic silencing-inducing signals (SSS) which are *trans*-acting and are capable of transport across graft unions (Palauqui et al., 1997; Palauqui and Vaucheret, 1998; Voinnet et al., 1998; Palauqui and Balzergue, 1999; Wassenegger and Pélissier, 1999). The

SSS are presumably aberrant RNAs of a similar sequence which lead to degradation of the specific messages. These aberrant RNAs may even be derived from the 'silenced' genes, are supposedly self-perpetuating, and can be transported to neighboring cells to induce/promote transgene silencing. Although mechanisms involving antisense RNA have also been proposed, no direct evidence for such RNAs has been forthcoming (see Jorgensen et al., 1999 for review). Other than the fact that PTGS has been observed in many studies, no mechanisms or schemes have been formulated to avoid or to reduce the incidence of PTGS. The only recommendation seems to be the preferential selection of transgenic plants which contain a single or only a few copies of the transgene. Whether or not PTGS is a phenomenon that occurs in nature in wild-type plants, many of which contain multiple copies of native genes, is not known (Cogoni and Macino, 1997).

It has been suggested that silencing of the transgenes by selective methylation and other means of chromatin modification is perhaps a normal part of the 'genome surveillance mechanism' of the plant as a defense against intrusive foreign DNAs (Kumpatla et al., 1998). Numerous studies hint at the existence of a variety of such mechanisms. In prokaryotes, these mechanisms exist in the form of restriction-modification systems which cause degradation of nucleic acids that are recognized as non-self. Examples of the existence of surveillance mechanisms in eukaryotes to recognize and inactivate foreign DNA include: chromosome elimination in interspecific somatic cell hybrids in animals, existence of heterochromatization of a chromosomal homologue, extensive heterochromatization of B chromosomes in many plants, silencing of transgenes in plants, selective elimination of duplicated and extragenomic DNA sequences in fungi and *Drosophila*, excision of integrated viral sequences in mammalian cells, and repeated excision of transposable elements in plants and animals. Thus the processes of elimination and silencing of intrusive (foreign) DNA may play an evolutionary function of stabilizing genomes over time. It has been suggested that PTGS may also function as a strong defense mechanism in plants against virus infection and viruses may have evolved mechanisms to counter these defenses (Ratcliff et al., 1997; Matzke and Matzke, 1998b). A better understanding of this mechanism should lead to manipulation of long-term resistance against viruses in tree species. It has also been postulated that these processes are operative, not only against truly invasive DNA, but also against events that cause intrageneric DNA rearrangements (e.g. those resulting from inversions, translocations, insertions, and overamplification of transposons and retroposons), since these events may also pose a threat to genome integrity (Kricker et al., 1992; Yoder et al., 1997). The existence of this phenomenon poses a fundamental question of utmost importance for the future of transgenic manipulation of plants. That is: if these mechanisms are indeed responsible for recognition of 'self' and 'non-self' nucleic acids and, subsequently, for the elimination (or silencing) of foreign DNA at the cellular level, how does a foreign DNA escape this surveillance and destruction? While there is no clear time frame within which foreign DNA is removed or silenced, experimental data from transgenic studies thus far indicate DNA silencing to be a rather short-term response. This should raise serious concerns about the potential stability of transgenes in plants, particularly in tree species which have life spans of several decades.

A number of mechanisms for the recognition of a DNA sequence as foreign have been discussed by Kumpatla et al. (1998). Some of them include: the recognition of aberrant DNA structures formed during DNA integration, sequence repeats, site of integration in the genome leading to 'position effect', and disruption of normal functions of active genes in the

genome. Once a DNA sequence has been recognized as 'non-self', it is either modified (hypermethylated) to render it non-functional and/or eliminated from the genome, particularly during meiosis when it may remain unpaired. Among the mechanisms proposed to achieve this are: (1) targeting the insertion intermediates that are formed during integration of foreign DNA and can be recognized by methyltransferases (Bestor and Tycko, 1996); (2) recognition of GC-rich DNA fragments (foreign) inserted within AT-rich regions of the genome (Bernardi, 1995; Carles et al., 1995; Iglesias et al., 1997); (3) direct DNA-DNA pairing between multiple copy insertions and/or with native homologous genes (Assaad et al., 1992; Bender, 1998); (4) collisions of transcription apparatus moving in opposite directions on a foreign gene inserted in the middle of a functional gene (Liu and Alberts, 1995); and (5) presence of abnormally high levels of mRNA, including aberrant mRNA, in the cytoplasm as discussed above (Palauqui et al., 1997; Voinnet et al., 1998).

Based upon our current understanding of native gene expression, and to some extent upon limited analysis of transgene expression, Kumpatla et al. (1998) suggest a few strategies for avoiding transgene silencing. These are:

- Introduction of base substitutions in the coding regions to increase sequence heterogeneity among multiple copies of transgenes and between transgenes and the native gene.
- Use of native promoters and 3'-termination sequences so that the transgene may function in a normal spatial and temporal fashion within the recipient cell. Substitutions in certain parts of the promoter/enhancer sequence that do not perform a crucial *cis*-function may aid in avoiding promoter silencing due to ectopic pairing with the native promoter sequences.
- Use of introns in the coding region, and possibly also in the 5' and 3' UTRs, to increase sequence diversity among multiple copies of the transgene and between the transgene and the native gene.
- Matching the overall GC richness of the transgene with that of the native genome, particularly in the regions of euchromatin. This can be achieved either by base substitution or addition of GC-rich small (100-300 bp) repeat sequences flanking the transgene, especially sequences isolated from active domains of chromatin.
- Use of MARs flanking the transgenes so that they may form their own independent loop domains in chromatin.
- Avoidance of multiple copy insertions.
- Use of sequences that may be easily recognized by chromatin acetylases and phosphorylases.
- Use of transcription termination sequences flanking a transgene to avoid transcriptional read-through from promoters that may be present around the insertion site.
- Avoiding or reducing the insertion of plasmid sequences that may be easily recognized as foreign DNA by the host genome scanning mechanism.

Razin (1998) has proposed a functional model showing a three-way connection between DNA methylation, chromatin structure and gene activity which may explain the mechanism of gene silencing and provide some guidelines to reduce or even reverse gene silencing. While this model does not explain why certain domains of chromatin are silenced, it sheds some light on how it is achieved and inherited through cell divisions.

6. Methods of transformation

Although several reports have shown a positive correlation between copy number of a transgene and the amount of transgenic product made in the cell, low copy number of integrated genes is generally considered to be a desirable situation. This is largely due to the complications of gene silencing and co-suppression, which are often associated with high copy numbers (Matzke and Matzke, 1998a). Of the different methods of transformation available, some generally result in multiple gene insertions (e.g. electroporation and biolistic bombardment), while others such as *Agrobacterium*-mediated transformation, generally yield lower numbers of integrations in the genome. Depending upon the target tissue used for transformation, one may be able to increase the frequency of single or few gene insertions by choosing an appropriate method of transformation. In all cases, the integration of vector sequences, particularly inverted repeats (e.g. in T-DNA borders) may make the transgene vulnerable to inactivation and silencing. At present, no reliable techniques are available to avoid the transfer and integration of vector DNA. With future modifications of the *Cre-lox* system, which allows precise excision of sequences that are flanked by the *lox* site, it may be possible to remove unwanted transgene sequences that may act as potential sites of methylation and gene silencing (Odell et al., 1990; 1994; Ow, 1996) .

Most of the past experiments on genetic engineering have involved the transfer of only one or two transgenes at a given time. Future demands may include transfer of multiple genes of related function to regulate several steps in a metabolic pathway or a group of genes affecting complex quantitative traits. Regulation of growth rates, biosynthesis of pharmaceuticals and secondary compounds, modulation of cell wall components (e.g. lignin), regulation of uptake and transport of nutrients, and assimilation of carbon and nitrogen, are examples of some of the future needs for plant improvement which will require multiple gene transfers. Similarly, the need to simultaneously transfer a number of diverse and unrelated insect resistance genes into trees is obvious because it will reduce the probability of insect pests developing resistance to a particular gene product over the life of tree plantations.

The size of DNA that can be transferred to a plant cell is determined by the vector and the method of transformation. Whereas electroporation, biolistic bombardment and commonly used *Agrobacterium* strains can transfer relatively small DNA sequences (10-20 kb maximum), microinjection and liposome fusion are capable of delivering much larger pieces of DNA. However, the latter are highly inefficient methods for use with most tree species. Hamilton et al. (1996) reported the development of a binary BAC (bacterial artificial chromosome) vector for transfer of large DNA fragments (as much as 150 kb long) into plant cells. Unfortunately, no follow-up publications have been reported to demonstrate its usefulness in the transfer of functional gene clusters. Nonetheless, such vectors should be extremely valuable in the future in transferring multiple genes or gene families into plant cells.

Another approach to achieving multiple gene transfer is the use of constructs that produce polycistronic mRNAs under the control of a single promoter. Although plants invariably use monocistronic mRNAs as translational units, it has been demonstrated that transgenic bicistronic messages can be successfully translated in plant cells (Iida et al., 1992; Lough et al., 1997). A better understanding of the oraganellar transcription and translation machinery (Barkan and Stern, 1998; Mayfield and Cohen, 1998; Mulligan and Maliga, 1998; Weihe and Börner, 1999; Bruick and Mayfield, 1999) and the availability of organelle

transformation systems (Bilang and Potrykus, 1998; Daniell, 1998) will also pave the way for using polycistronic gene constructs under the control of a single chloroplast-specific promoter for multiple gene transfers.

7. Targeted transformation

There are two ways in which transgenes are believed to integrate into the host genome: (1) via homologous recombination events between similar sequences in the host genome and the transgene, and (2) illegitimate recombination which is independent of any homologies between the transgene and the host genome (Puchta and Hohn, 1996; Puchta, 1998). While the former generally leads to replacement of host genes with the transgene, the latter is a random event and may not cause any loss of host genes.

Targeted insertion of foreign DNA provides a number of unique advantages for transgene expression. These include: regulation of transgene expression by native promoters, replacement of defective genes by functional sequences, gene knockout by replacement of functional genes with mutated non-functional sequences, enhanced expression of foreign genes, and reduction in the frequency of co-suppression when multiple copies of homologous genes are introduced. The technique also provides a unique advantage for analyzing gene function (functional genomics).

While homologous recombination is a common, even predominant event, in prokaryotes and some lower eukaryote (Puchta and Hohn, 1996), in higher eukaryotes it occurs at extremely low frequency (Puchta et al., 1994; Lijegren and Yanofsky, 1998). Even at that, targeted gene insertion by homologous recombination has proven to be highly useful in understanding gene function in animals and is being touted as extremely desirable for human gene therapy to cure genetic ailments (Thomas et al., 1992). With the exception of the single-celled green alga *Chlamydomonas reinhardtii* (Chlorophyta) and the moss *Physcomitrella patens* (Bryophyta), experimental evidence for targeted gene insertion in higher plants is rare (Schaefer and Zryd, 1997; Puchta et al., 1994; Puchta, 1998). Miao and Lam (1995) and Kempin et al. (1997) have presented some preliminary evidence for successful targeted insertion of homologous genes in *Arabidopsis thaliana*. Homologous recombination between the transgene and the host DNA was observed at a frequency of one in 2500 transformed plants. However, Thykjaer et al. (1997) found not a single case of targeted transgene insertion among more than 18,900 transformed plants of *Lotus japonicus* that they analyzed.

While molecular events leading to homologous recombination in bacteria and yeast are becoming better understood, the low frequency or absence of this phenomenon in higher plants is both intriguing and frustrating. Based upon its success with *P. patens* and *C. reinhardtii*, it has been suggested that haploid state of the genome in plants may be more conducive to homologous recombination than the diploid state. Due to limited attempts and the scarcity of success with transformation of haploid tissues in higher plants , the hypothesis has not been duly tested. Thus, while being highly desirable, targeted insertion of the transgene still remains an elusive event in higher plants.

8. Post-translational factors and organelle targeting of proteins

The expression of a transgene to produce a protein is only half of the story to achieve desired

results of genetic engineering. Even when high levels of transcription and translation have been achieved, numerous post-translational modifications of a polypeptide could become limiting factors in the production of a physiologically active protein. Following translation and partial folding of a protein, its targeting to specific sites in the cell (transport to organelles, incorporation into membranes, secretion out of the cell, etc.), its secondary modifications (phosphorylation, glycosylation, acetylation, farnesylation, methylation, proteolytic cleavage, self-cleavage, dimerization, etc.), and its interactions with co-factors and co-enzymes are mostly inherent in the amino acid sequence of the protein. While transgene expression yielding cytoplasmically-localized proteins has resulted in successful production of several varieties of genetically improved crop plants for single gene-regulated traits (e.g. for herbicide tolerance, insect tolerance, and virus resistance), modulation of complex metabolism and organelle or membrane based function will require proper targeting of the transgenic protein to an appropriate site in the cell. In recent years, we have witnessed a tremendous increase in our understanding of the information needed for protein targeting.

The process of protein folding and translocation to the target site is quite complex (Boston et al., 1996; Heins et al., 1998). Information contained in the polypeptide sequences is often both necessary and sufficient for its targeting to a specific site. In most cases, these amino acid sequences - also called signal sequences or transit peptides - are located at the N-terminus of the polypeptide. The structural features of transit peptides for a specific organelles are conserved among plants, thus the transit peptides of heterologous proteins seem to work well for transport of transgene products. The process of translocation and targeting, however, may require several modifications of the transit peptide region by enzymes located in the cytoplasm or in the membrane to be crossed (Boston et al., 1996; Bar-Peled et al., 1996; Ellis et al., 1998; Thieffry and Sarkar, 1998; Vitale and Raikhel, 1999). In addition, the polypeptide is often complexed with additional proteins commonly called chaperones. For example, in case of the chloroplast targeted proteins, an N terminal transit peptide, which must be phosphorylated at a specific site by a cytosolic protein kinase, is cleaved before transfer to chloroplast stroma. A protein phosphatase located in the outer envelope is also required for dephosphorylation of the protein before active transport into the stroma. Once in the stroma, further targeting of the protein into the thylakoids or its maintenance in the soluble fraction of the stroma is a function of additional protein complexes. Proteins that are targeted to the outer envelope use a modification of the system for stromal translocation (Heins et al., 1998). The targeting of proteins into mitochondria and other organelles requires organelle-specific transit peptides (Bar-Peled et al., 1996). Thus, for the targeting of a transgenic protein to a specific organelle, sequences for appropriate transit peptides must be included in the translated coding sequence of the transgene. Information obtained from such studies also allows us to study the effect of targeting certain metabolically active proteins to organelles where the substrate may be localized.

Unlike the targeting signals found on proteins destined for chloroplast and mitochondrial localization, most of which are highly conserved, the proteins targeted for nuclear destination do not possess conserved *cis*-acting amino acid sequences. In the case of the nucleus, protein import as well as export are equally important functions. At present, the import seems to be better understood than the export (Hicks and Raikhel, 1995; Boelens et al., 1995; Nagatani, 1998; Smith and Raikhel, 1999). While each nuclear-targeted protein needs to have an import-export signal (nuclear localization signal - NLS), the process is aided by a

number of carrier proteins called importins. The determination of where the polypeptides end up in the cell is achieved partially by the distribution of their mRNAs into different parts of the ER (Okita and Rogers, 1996). Cellular cytoskeleton (actin microfilaments and microtubules) also plays a crucial role in this process. Whereas properly-folded proteins are translocated to their final site, improperly-folded proteins are recognized by the ER and translocated back to the cytosol to be degraded (Suzuki et al., 1998). Thus the ER performs a reliable quality control function as well as help in protein deposition. Borisjuk et al. (1999) have demonstrated an effective secretion of transgenic proteins from plant roots by designing transgenes using ER signal peptide sequences. This approach will eliminate the need for extensive purification of commercially valuable proteins produced in transgenic plants.

9. What are the limits to protein accumulation?

How much of a single protein can a cell accumulate? This question by no means has a clear answer. The final amounts of specific protein accumulation in a cell are governed not only by the rates of synthesis and degradation of the protein but also by the cell's ability to target, store, translocate, and localize the protein in sub-cellular compartments. It is well known that the enzyme Rubisco can constitute as much as 50% of the total soluble protein in a mesophyll cell. It is, however, a complex enzyme made of several polypeptides, each of which represents a smaller fraction of the total protein. Seed storage proteins, that are often produced from large gene families, can constitute even higher proportions of the total protein content of a cell. It must be emphasized, however, that seed proteins are not found in a soluble form but are stored in membrane-surrounded vesicles (Matsuoka and Bednarek, 1998). For transgenic proteins that are encoded by nuclear genes, cytosolic concentrations typically range between 0.1-1.0% of the total protein (Van Engelen et al., 1994; Verwoerd et al., 1995). However, single proteins representing as much as 5% (Eckes et al., 1989) and 12% (Ku et al., 1999) of the soluble protein fraction have been reported for some transgenic cells/plants. The latter was attributed to a combination of factors, including the presence of introns in the coding sequence, the strength of the promoter, and the similarity between the source of coding sequence and the recipient species (both being monocots). The same gene transferred to tobacco and potato did not produce such high levels of the protein. Often, the progeny of transformants showing high levels of transgenic protein do not show such high levels of accumulation (deNeve et al., 1999). While high levels of expression of the transgene are important in many cases, it is equally important that the expression be stable throughout the life of the plant as well as be transmitted to the progeny.

10. Future Perspectives

It is obvious from the foregoing discussion that successful genetic improvement of a commercially important plant species by genetic engineering depends upon (1) the availability of a gene construct with appropriate coding and regulatory sequences, (2) stable integration of the transgene into the plant genome, and (3) its continued regulated expression through the life of the plant. We have gained significant knowledge about designing gene constructs to achieve regulated expression in desirable tissues, however, at present, we have little control over the host cell and genome factors that determine the integration and expression of a

transgene which ultimately decides its usefulness in the transgenic plant (Tinland, 1996). We must rely upon the selection of appropriate transformants which have been tested for stable expression of the gene, and then incorporate the transgenic trait into desirable crop varieties to produce sufficient seed and/or other planting material, often through a conventional breeding program. The plantation material must also be tested for transgene expression under field conditions. In spite of the potential for a transgene to be silenced through this period of expansion of the transgenic germplasm, remarkable success has been achieved in commercial production of genetically engineered varieties of most of our important dicot and some monocot crops (Dale and Irwin, 1995; Moffat, 1998). In most cases, it appears that if the transgene was not silenced within the first one or two generations, it remains active through successive generations.

While the transgenic approach provides some unique advantages for genetic improvement of trees that are not achievable by conventional methods, at the same time, the technology has some formidable logistical problems for this group of plants. With increased understanding of the structures of plant genomes (Bevan et al., 1998; Bennetzan, 1998; Walbot, 1999), and more information on how the genomes are functionally organized (Bouchez and Hofte, 1998), we should be able to identify regions of the genome that may be better suited for targeted transgene insertion. Even when we have successfully produced transgenic T_o or R_o generations (plants regenerated directly from transgenic cells/callus) of woody plants, mass production of field-plantable material in most cases will have to be achieved by vegetative propagation. Incorporation of the transgenic germplasm into large numbers of individuals for field plantation through breeding programs will take an extraordinarily large amount of time because of the multi-year flowering cycle in most tree species.

Mass vegetative propagation systems are not yet available in most commercially important tree species (particularly the conifers). This fact combined with short-lived regeneration capability of callus cultures in many species requires heavy dependence upon continued generation of new transgenic cell lines. Thus, there will be tremendous variation in expression levels of transgenes in field-planted material. In order to augment this situation, reliable cryopreservation technology must be developed for long term storage of the transgenic tissues with high regeneration capacity.. Furthermore, complications posed by gene silencing will require rigorous selection and monitoring of transgene expression in large populations of plants. Therefore, rapid means of quantifying transgene expression and monitoring of the transgenic phenotype in field-grown plants may have to be developed for evaluating transgene expression on a large scale. Obviously, a better understanding of the control of gene expression and the availability of techniques for targeted transgene integration will lead to better design of constructs for optimal expression of the desired transgenes.

Fussenegger et al. (1999) have discussed several approaches to achieve optimized production of recombinant glycoproteins in mammalian cells, which could provide leads for future experimentation with plants as well. Some of these approaches focus on targeting genes to specific locations on a chromosome, use of bidirectional promoters, coordinated expression of multi-component and multi-subunit proteins using polycistronic expression of several genes, the design of complex regulatory circuits such as positive feedback regulatory systems, and advanced selection and screening procedures.

When it comes to successful regulation of sustained gene expression, there is a lot

more that we can learn from viruses (Maia et al., 1996; Gallie, 1996; Wang et al., 1997). They employ a set of unique strategies that allow them to not only replicate their genome but to also produce the necessary proteins in a sequential manner. The polypeptide production is regulated more by the nucleic acid sequence than the cellular (host) factors. At the same time, the viruses utilize (exploit) the host machinery in the most successful manner, with relatively small contribution from their own genome or its products. The mechanisms of RNA amplification, RNA stability, selective translation from a polycistronic message, selective stop signals to produce different proteins from the same genome, post-translational modifications of their proteins, and assembly of stable nucleo-protein complexes in the cell are some of the unique features of the viral genome and its organization. Functional analysis of DNA or RNA domains of viruses should yield extremely useful information for designing transgene constructs that would function in a plant cell in optimal fashion.

11. Acknowledgments

I would like to express sincere thanks to Dr. Eric Schaller, Dr. John Wallace, Dr. Rakesh Minocha, Ms. Laura Morgenthau and Ms. Pratiksha Bhatnagar for their valuable suggestions for improvement of the manuscript, and to Ms. Diane Lavalliere for word processing. This is a scientific contribution number 2018 from the NH Agricultural Experiment Station.

12. References

Abler, M.L. and Green, P.J. (1996) Control of mRNA stability in higher plants, *Plant Mol Biol* **32**, 63-78.

Allen, G.C., Hall Jr., G., Michalowski, S., Newman, W., Spiker, S., Weissinger, A.K., and Thompson, W.F. (1996) High-level transgene expression in plant cells: effects of a strong scaffold attachment region from tobacco, *Plant Cell* **8**, 899-913.

Aoyama, T. and Chua, N.-H. (1997) A glucocorticoid-mediated transcriptional induction system in transgenic plants, *Plant J* **11**, 605-612.

Assaad, F.F., Tucker, K.L., and Signer, E.R. (1993) Epigenetic repeat-induced gene silencing (RIGS) in *Arabidopsis*, *Plant Mol Biol* **22**, 1067-1085.

Bailey-Serres, J. (1999) Selective translation of cytoplasmic mRNAs in plants, *TIPS* **4**, 142-148.

Bailey-Serres, J. and Gallie, D.R. (eds.) (1998) *A Look Beyond Transcription: Mechanisms Determining mRNA Stability and Translation in Plants*, American Society of Plant Physiologists, Rockville, MD, 183 pp.

Bar-Peled, M., Bassham, D.C., and Raikhel, N.V. (1996) Transport of proteins in eukaryotic cells: more questions ahead, *Plant Mol Biol* **32**, 223-249.

Barkan, A. and Stern, D. (1998) Chloroplast mRNA processing: intron splicing and 3'-end metabolism, in J. Bailey-Serres and D.R. Gallie (eds.), *A Look Beyond Transcription: Mechanisms Determining mRNA Stability and Translation in Plants*, American Society of Plant Physiologists, Rockville, MD, pp. 162-173.

Bastola, D.R. and Minocha, S.C. (1995) Increased putrescine biosynthesis through transfer of mouse ornithine decarboxylase cDNA in carrot promotes somatic embryogenesis, *Plant Physiol* **109**, 63-71.

Baulcombe, D.C. (1996) RNA as a target and an initiator of post-transcriptional gene silencing in transgenic plants, *Plant Mol Biol* **32**, 79-88.

Bender, J. (1998) Cytosine methylation of repeated sequences in eukaryotes: the role of DNA pairing, *TIBS* **23**, 252-256.

Bennetzen, J.L. (1998) The structure and evolution of angiosperm nuclear genomes, *Curr Opin Plant Biol* **1**, 103-108.

Benyajati, C. and Worcel, A. (1976) Isolation, characterization, and structure of the folded interphase genome of *Drosophila melanogaster*, *Cell* **9**, 393-407.

Bernardi, G. (1995) The human genome: organization and evolutionary history, *Annu Rev Genet* **29**, 445-476.

Bestor, T.H. and Tycko, B. (1996) Creation of genomic methylation patterns, *Nature Genet* **12**, 363-367.

Bevan, M., Bennetzen, J.L., and Martienssen, R. (1998) Genome studies and molecular evolution commonalities, contrasts, continuity and change in plant genomes, *Curr Opin Plant Biol* **1**, 101-102.

Bilang, R. and Potrykus, I. (1998) Containing excitement over transplastomic plants, *Nature Biotech* **16**, 333-334.

Birch, R.G. (1997) Plant transformation: problems and strategies for practical application, *Ann Rev Plant Physiol Plant Mol Biol* **48**, 297-326.

Björklund, S., Almouzni, G., Davidson, I., Nightingale, K.P., and Weiss, K. (1999) Global transcription regulators of eukaryotes, *Cell* **96**, 759-767.

Boelens, W.C., Dargemont, C., and Mattaj, I.W. (1995) Export of mRNA through the nuclear pore complex, in A. Lamond (ed.), *Pre-mRNA Processing*, R.G. Landes Publishers, Georgetown, TX, pp. 173-186.

Borisjuk, N.V., Borisjuk, L.G., Logendra, S., Petersen, F., Gleba, Y., and Raskin, I. (1999) Production of recombinant proteins in plant root exudates, *Nature Biotech* **17**, 466-469.

Boston, R.S., Viitanen, P.V., and Vierling, E. (1996) Molecular chaperones and protein folding in plants, *Plant Mol Biol* **32**, 191-222.

Bouchez, D. and Hofte, H. (1998) Functional genomics in plants, *Plant Physiol* **118**, 725-732.

Brendel, V., Carle-Urioste, J.C., and Walbot, V. (1998) Intron recognition in plants, in J. Bailey-Serres and D.R. Gallie (eds.), *A Look Beyond Transcription: Mechanisms Determining mRNA Stability and Translation in Plants*, American Society of Plant Physiologists, Rockville, MD, pp. 20-28.

Breyne, P., van Montagu, M., Depicker, N., Gheysen, G. (1992) Characterization of a plant scaffold attachment region in a DNA fragment that normalizes transgene expression in tobacco, *Plant Cell* **4**, 463-471.

Browning, K.S. (1996) The plant translational apparatus, *Plant Mol Biol* **32**, 107-144.

Browning, K.S., Goss, D.J., Roth, D.A., and Gallie, D.R. (1998) The translational machinery of plants, in J. Bailey-Serres and D.R. Gallie (eds.), *A Look Beyond Transcription: Mechanisms Determining mRNA Stability and Translation in Plants*, American Society of Plant Physiologists, Rockville, MD, pp. 68-83.

Bruick, R.K. and Mayfield, S.P. (1999) Light-activated translation of chloroplast mRNAs, *TIPS* **4**, 190-195.

Caddick, M.X., Greenland, A.J., Jepson, I., Krause, K.P., Qu, N., Riddell, K.V., Salter, M.G., Schuch, W., Sonnewald, U., and Tomsett, A.B. (1998) An ethanol inducible gene switch for plants used to manipulate carbon metabolism, *Nature Biotech* **16**, 177-180.

Carles, N., Barakat, A., and Bernardi, G. (1995) The gene distribution of the maize genome, *Proc Natl Acad Sci USA* **92**, 11057-11060.

Cech, T.R. (1993) Ribozymes, in R.F. Gesteland and J.F. Atkin, (eds) *The RNA World*, Cold Spring Harbor Press, Cold Spring Harbor, NY, pp 239-269.

Chinn, A.M., Payne, S.R., and Comai, L. (1996) Variegation and silencing of the *Heat Shock Cognate 80* gene are relieved by a bipartite downstream regulatory element, *Plant J* **9**, 325-339.

Chiu, W., Niwa, Y., Zeng, W., Hirano, T., Kobayashi, H., and Sheen, J. (1996) Engineered GFP as a vital reporter in plants, *Curr Biol* **6**, 325-330.

Cogoni, C. and Macino, G. (1997) Conservation of transgene-induced post-transcriptional gene silencing in plants and fungi, *TIPS* **2**, 438-443.

Crameri, A., Whitehorn, E.A., Tate, E., and Stemmer, W.P.C. (1996) Improved green fluorescent protein by molecular evolution using DNA shuffling, *Nature Biotech* **14**, 315-319.

Curtis, D., Lehmann, R., and Zamore, P.D. (1995) Translational regulation in development, *Cell* **81**, 171-178.

Dale, P.J. and Irwin, J.A. (1995) The release of transgenic plants from containment, and the move towards their widespread use in agriculture, *Euphytica* **85**, 425-431.

Daniell, H., Datta, R., Varma, S., Gray, S., and Lee, S.-B. (1998) Containment of herbicide resistance through genetic engineering of the chloroplast genome, *Nature Biotech* **16**, 345-348.

De Neve, M., De Buck, S., De Wilde, C., Van Houdt, H., Strobbe, I., Jacobs, A., Van Montagu, M., and Depicker, A. (1999) Gene silencing results in instability of antibody production in transgenic plants, *Mol Gen Genet* **260**, 582-592.

Depicker, A. and van Montagu, M. (1997) Post-transcriptional gene silencing in plants, *Curr Opin Cell Biol* **9**, 373-382.

DeScenzo, R.A. and Minocha, S.C. (1993) Modulation of cellular polyamines in tobacco by transfer and expression of mouse ornithine decarboxylase cDNA, *Plant Mol Biol* **22**, 113-127.

Doebley, J. and Lukens, L. (1998) Transcriptional regulators and the evolution of plant form, *Plant Cell* **10**, 1075-1082.

Düring, K., Porsch, P., Fladung, M., and Lörz, H. (1993) Transgenic potato plants resistant to the phytopathogenic bacterium *Erwinia carotovora*, *Plant J* **3**, 587-598.

Eckes, P., Schmitt, P., Daub, W., and Wengenmayer, F. (1989) Overexpression of alfalfa glutamine synthetase in transgenic tobacco plants, *Mol Gen Genet* **217**, 263-268.

Ellis, R.J., Dobson, C., and Hartl, U. (1998) Sequence does specify protein conformation, *TIBS* **23**, 468

English, J.J., Davenport, G.F., Elmayan, T., Vaucheret, H., and Baulcombe, D.C. (1997) Requirement of sense

transcription for homology-dependent virus resistance and *trans*-inactivation, *Plant J* **12**, 597-604.

Faiss, M., Zalubilova, J., Strnad, M., Schmülling, T. (1997) Conditional transgenic expression of the *ipt* gene indicates a function for cytokinins in paracrine signaling in whole tobacco plants, *Plant J* **12**, 401-415.

Filipowicz, W., Gniadkowski, M., Klahre, U., and Liu, H.-X. (1995) Pre-mRNA splicing in plants, in A. Lamond (ed.), *Pre-mRNA Processing*, R.G. Landes Publishers, Georgetown, TX, pp. 65-78.

Fu, H., Kim, S.Y., and Park, W.D. (1995a) High-level tuber expression and sucrose inducibility of a potato *Sus4* sucrose synthase gene require 5' and 3' flanking sequences and the leader intron, *Plant Cell* **7**, 1387-1394.

Fu, H., Kim, S.Y., and Park, W.D. (1995b) A potato *Sus3* sucrose synthase gene contains a context-dependent 3' element and a leader intron with both positive and negative tissue-specific effects, *Plant Cell* **7**, 1395-1403.

Fussenegger, M., Bailey, J.E., Hauser, H., and Mueller, P.P. (1999) Genetic optimization of recombinant glycoprotein production by mammalian cells, *Trends Biotech* **17**, 35-42.

Fütterer, J., Gordon, K., Bonneville, J.M., Sanfacon, H., Pisan, B., Penswick, J., and Hohn, T. (1988) The leading sequence of caulimovirus large RNA can be folded into a large stem-loop structure, *Nucleic Acids Res* **16**, 8377-8390.

Fütterer, J. and Hohn, T. (1996) Translation in plants – rules and exceptions, *Plant Mol Biol* **32**, 159-189.

Fütterer, J., Potrykus, I., Valles-Brau, M.P., Dasgupta, I., Hull, R., and Hohn, T. (1994) Splicing in a plant pararetrovirus, *Virology* **198**, 663-670.

Gallie, D.R. (1996) Translational control of cellular and viral mRNAs, *Plant Mol Biol* **32**, 145-158.

Gallie, D.R. (1998a) Controlling gene expression in transgenics, *Curr Opin Plant Biol* **1**, 166-172.

Gallie, D.R. (1998b) A tale of two termini: a functional interaction between the termini of an mRNA is a prerequisite for efficient translation initiation, *Gene* **216**, 1-11.

Gatz, C. (1997) Chemical control of gene expression, *Annu Rev Plant Physiol Plant Mol Biol* **48**, 89-108.

Gatz, C., Frohberg, C., and Wendenburg, R. (1992) Stringent repression and homogeneous de-repression by tetracycline of a modified CaMV 35S promoter in intact transgenic tobacco plants, *Plant J* **2**, 397-404.

Gatz, C. and Lenk, I. (1998) Promoters that respond to chemical inducers, *TIPS* **3**, 352-358.

Ghoda, L., van Daalen, W.T., Macrae, M., Ascherman, D., and Coffino, P. (1989) Prevention of rapid intracellular degradation of ODC by a carboxyl-terminal truncation, *Science* **243**, 1493-1495.

Goodall, G.J. and Filipowicz, W. (1991) Different effects of intron nucleotide composition and secondary structure on pre-mRNA splicing in monocot and dicot plants, *EMBO J* **10**, 2635-2644.

Grens, A. and Scheffler, I.E. (1990) The 5'- and 3'-untranslated regions of ornithine decarboxylase mRNA affect the translational efficiency, *J Biol Chem* **265**, 11810-11816.

Guilfoyle, T.J. (1997) The structure of plant gene promoters, *Genet Eng* **19**, 15-47.

Guo, Z. and Sherman, F. (1996) 3'-end-forming signals of yeast mRNA, *TIBS* **21**, 477-481.

Hamilton, C.M., Frary, A., Lewis, C., and Tanksley, S.D. (1996) Stable transfer of intact high molecular weight DNA into plant chromosomes, *Proc Natl Acad Sci USA* **93**, 9975-9979.

Haseloff, J., Siemering, K.R., Prasher, D.C., and Hodge, S. (1997) Removal of cryptic intron and subcellular localization of green fluorescent protein are required to mark transgenic *Arabidopsis* plants brightly, *Proc Natl Acad Sci USA* **94**, 2122-2127.

Heins, L., Collerison, I., and Soll, J. (1998) The protein translocation apparatus of chloroplast envelopes, *TIPS* **3**, 56-61.

Hicks, G.R. and Raikhel, N.V. (1995) Protein import into the nucleus: an integrated view, *Annu Rev Cell Dev Biol* **11**, 155-188.

Hohn, T., Dominguez, D.I., Schärer-Hernández, Pooggin, M.M., Schmidt-Puchta, W., Hemmings-Mieszczak, M., and Fütterer, J. (1998) Ribosome shunting in eukaryotes: what the viruses tell me, in J. Bailey-Serres and D.R. Gallie (eds.), *A Look Beyond Transcription: Mechanisms Determining mRNA Stability and Translation in Plants*, American Society of Plant Physiologists, Rockville, MD, pp. 84-95.

Holmes-Davis, R. and Comai, L. (1998) Nuclear matrix attachment regions and plant gene expression, *TIPS* **3**, 91-97.

Holtorf, H., Schöb, H., Kunz, C., Waldvogel, R., and Meins Jr., F. (1999) Stochastic and nonstochastic post-transcriptional silencing of chitinase and β-1,3-glucanase genes involves increased RNA turnover – possible role for ribosome-independent RNA degradation, *Plant Cell* **11**, 471-483.

Hunt, A.G., Chu, N.M., Odell, J.T., Nagy, F., and Chua, N. (1987) Plant cells do not properly recognize animal gene polyadenylation signals, *Plant Mol Biol* **8**, 23-35.

Hunt, A.G. and Messing, J. (1998) mRNA polyadenylation in plants, in J. Bailey-Serres and D.R. Gallie (eds.), *A Look Beyond Transcription: Mechanisms Determining mRNA Stability and Translation in Plants*, American Society of Plant Physiologists, Rockville, MD, pp. 29-39.

Iglesias, V.A., Moscone, E.A., Papp, I., Neuhuber, F., Michalowski, S., Phelan, T., Spiker, S., Matzke, M., and

Matzke, A.J.M. (1997) Molecular and cytogenetic analyses of stably and unstably expressed transgene loci in tobacco, *Plant Cell* **9**, 1251-1264.

Iida, S., Scheid, O.M., Saul, M.W., Seipel, K., Miyazaki, C., and Potrykus, I. (1992) Expression of a downstream gene from a bicistronic transcription unit in transgenic tobacco plants, *Gene* **119**, 199-205.

Johnson, M.A., Baker, E.J., Colbert, J.T., and Green, P.J. (1998) Determinants of mRNA stability in plants, in J. Bailey-Serres and D.R. Gallie (eds.), *A Look Beyond Transcription: Mechanisms Determining mRNA Stability and Translation in Plants*, American Society of Plant Physiologists, Rockville, MD, pp. 40-53.

Jorgensen, R.A., Que, Q., and Stam, M. (1999) Do unintended antisense transcripts contribute to sense cosuppression in plants? *Trends Genet* **15**, 11-12.

Joshi, C.P., Zhou, H., Huang, X., and Chiang, V.L. (1997) Context sequences of translation initiation codon in plants, *Plant Mol Biol* **35**, 993-1001.

Kahana, C. and Nathans, D. (1985) Nucleotide sequence of murine ornithine decarboxylase mRNA, *Proc Natl Acad Sci USA* **82**, 1673-1677.

Keller, W. (1995) 3' end cleavage and polyadenylation of nuclear messenger RNA precursors, in A. Lamond (ed.), *Pre-mRNA Processing*, R.G. Landes Publishers, Georgetown, TX, pp. 113-128.

Kempin, S.A. Liljegren, L.J., Block, L.M., Rounsly, S.D. and Yanofsky, M.F. (1997) Targeted disruption in *Arabidopsis*, *Nature* **389**, 802-803.

Klaff, P., Riesner, D., and Steger, G. (1996) RNA structure and the regulation of gene expression, *Plant Mol Biol* **32**, 89-106.

Klahre, U., Hemmings-Mieszczak, M., and Filipowicz, W. (1995) Extreme heterogeneity of polyadenylation sites in mRNAs encoding chloroplast RNA-binding proteins in *Nicotiana plumbaginifolia, Plant Mol Biol* **28**, 569-574.

Kozak, M. (1991) Structural features in eukaryotic mRNAs that modulate the initiation of translation, *J Biol Chem* **266**, 19867-19870.

Koziel, M.G., Carozzi, N.B., and Desai, N. (1996) Optimizing expression of transgenes with an emphasis on post-transcriptional events, *Plant Mol Biol* **32**, 393-405.

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

Ku, M.S.B., Agarie, S., Nomura, M., Fukayama, H., Tsuchida, H., Ono, K., Hirose, S., Toki, S., Miyao, M., and Matsuoka, M. (1999) High-level expression of maize phosphoenolpyruvate carboxylase in transgenic rice plants, *Nature Biotech* **17**, 76-81.

Kumpatla, S.P., Chandrasekharan, M.B., Iyer, L.M., Li, G., and Hall, T.C. (1998) Genome intruder scanning and modulation systems and transgene silencing, *TIPS* **3**, 97-104.

Laroia, G., Cuesta, R., Brewer, G., and Schneider, R.J. (1999) Control of mRNA decay by heat shock-ubiquitin-proteasome pathway, *Science* **284**, 499-502.

Leech, M.J., May, K., Hallard, D., Verpoorte, R., De Luca, V., and Christou, P. (1998) Expression of two consecutive genes of a secondary metabolic pathway in transgenic tobacco: molecular diversity influences levels of expression and product accumulation, *Plant Mol Biol* **38**, 765-774.

Lehmann, V. (1998) Patent on seed sterility threatens seed saving, *Biotech Devel Monitor* **35**, 6-8.

Lemmetyinen, J., Keinonen-Mettälä, Lännenpää, M., von Weissenberg, K., and Sopanen, T. (1998) Activity of the CaMV 35S promoter in various parts of transgenic early flowering birch clones, *Plant Cell Rep* **18**, 243-248.

Lewin, B. (1997) *Genes VI*, Oxford University Press, New York, NY, pp. 809-883.

Lijegren, S.J. and Yanofsky, M.F. (1998) ...Response: targeting *Arabidopsis*, *TIPS* **3**, 79-80.

Liu, B. and Alberts, B.M. (1995) Head-on collision between a DNA replication apparatus and RNA polymerase transcription complex, *Science* **267**, 1131-1137.

Lloyd, A.M., Schena, M., Walbot, V., and Davis, R.W. (1994) Epidermal cell fate determination in *Arabidopsis*: patterns defined by a steroid-inducible regulator, *Science* **266**, 436-439.

Lough, T., Tourneur, C., Masson, J., and Robaglia, C. (1997) Expression of genes in transgenic plants from bicistronic transcriptional units, *Plant Sci* **129**, 91-99.

Luehrsen, K.R., Taha, S., and Walbot, V. (1994) Nuclear pre-mRNA processing in higher plants, *Prog Nucl Acid Res Mol Biol* **47**, 149-193.

Maia, I.G., Séron, K., Haenni, A.-L., and Bernardi, F. (1996) Gene expression from viral RNA genomes, *Plant Mol Biol* **32**, 367-391.

Marcand, S., Gasser, S.M., and Gilson, E. (1996) Chromatin: a sticky silence, *Curr Biol* **6**, 1222-1225.

Marcotte Jr., W.R. (1998) Developmental regulation of translation and mRNA stability, in J. Bailey-Serres and D.R. Gallie (eds.), *A Look Beyond Transcription: Mechanisms Determining mRNA Stability and Translation*

in Plants, American Society of Plant Physiologists, Rockville, MD, pp. 64-67.

Martienssen, R.A. and Richards, E.J. (1995) DNA methylation in eukaryotes, *Curr Opin Genet Dev* **5**, 234-242.

Matsuoka, K. and Bednarek, S.Y. (1998) Protein transport within the plant cell endomembrane system: an update, *Curr Opin Plant Biol* **1**, 463-469.

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

Matzke, A.J.M. and Matzke, M.A. (1998a) Position effects and epigenetic silencing of plant transgenes, *Curr Opin Plant Biol* **1**, 142-148.

Matzke, M.A. and Matzke, A.J.M. (1998b) Epigenetic silencing of plant transgenes as a consequence of diverse cellular defense responses, *Cell Mol Life Sci* **54**, 94-103.

Mayfield, S.P. and Cohen, A. (1998) Translational regulation in the chloroplast, in J. Bailey-Serres and D.R. Gallie (eds.), *A Look Beyond Transcription: Mechanisms Determining mRNA Stability and Translation in Plants*, American Society of Plant Physiologists, Rockville, MD, pp. 174-179.

McNellis, T.W., Mudgett, M.B., Li, K., Aoyama, T., Horvath, D., Chua, N.-H., and Staskawicz, B.J. (1998) Glucocorticoid-inducible expression of a bacterial avirulence gene in transgenic *Arabidopsis* induces hypersensitive cell death, *Plant J* **14**, 247-257.

Meins Jr., F. (1996) Epigenetic modifications and gene silencing in plants, in V.E.A. Russo, R.A. Martienssen, and A.D. Riggs (eds.), *Epigenetic Mechanisms of Gene Regulation*, Cold Spring Harbor Press, Cold Spring Harbor, NY, pp. 415-442.

Meisel, L., and Lam, E. (1997) Switching on gene expression: analysis of the factors that spatially and temporally regulate plant gene expression, *Genet Eng* **19**, 183-195.

Meshi, T. and Iwabuchi, M. (1995) Plant transcription factors, *Plant Cell Physiol* **36**, 1405-1420.

Mett, V.L., Lochhead, L.B., and Reynolds, P.H.S. (1993) Copper controllable gene expression system for whole plants, *Proc Natl Acad Sci USA* **90**, 4567-4571.

Mett, V.L., Podivinsky, E., Tennant, A.M., Lochhead, L.P., Jones, W.T., and Reynolds, P.H. (1996) A system for tissue-specific copper-controllable gene expression in transgenic plants: nodule-specific antisense expression of aspartate aminotransferase-P2, *Transgen Res* **5**, 105-113.

Mette, M.F., van der Winden, J., Matzke, M.A., and Matzke, A.J.M. (1999) Production of aberrant promoter transcripts contributes to methylation and silencing of unlinked homologous promoters *in trans*, *EMBO J* **18**, 241-248.

Meyer, P. (1995) Variation of transgene expression in plants, *Euphytica* **85**, 359-366.

Miao, Z.H. and Lam, E. (1995) Targeted disruption of the TGA3 locus in *Arabidopsis thaliana*, *Plant J* 7, 359-365.

Mlynárová, L., Loonen, A., Heldens, J. Jansen, R.C., Keizer, P., Stiekema, W.J., and Nap, J.-P. (1994) Reduced position effect in mature transgenic plants conferred by the chicken lysozyme matrix-associated region, *Plant Cell* **6**, 417-426.

Mlynárová, L., Jansen, R.E., Conner, A.J., Stiekema, W.J., and Nap, J.-P. (1995) The MAR-mediated reduction in position effect can be uncoupled from copy number-dependent expression in transgenic plants, *Plant Cell* 7, 599-609.

Moffatt, A.S. (1998) Toting up the early harvest of transgenic plants, *Science* **282**, 2176-2178.

Moore, M.J., Query, C.C., and Sharp, P.A. (1993) Splicing of precursors to messenger RNAs by the spliceosome, in R.F. Gestland and J. F. Atkins (eds.), *The RNA World*, Cold Spring Harbor Press, Cold Spring Harbor, NY, pp. 303-358.

Motyka, V. Faiss, M., Strnad, M., Kaminek, M., and Schmülling (1996) Changes in cytokinin content and cytokinin oxidase activity in response to derepression of *ipt* gene transcription in transgenic tobacco calli and plants, *Plant Physiol* **112**, 1035-1043.

Muller, K.J., Romano, N., Gerstner, O., Garcia-Maroto, F., Pozzi, C., and Salamini, R.W. (1995) The barley *Hooded* mutation is caused by a duplication in a homebox gene intron, *Nature* **374**, 727-730.

Mulligan, R.M. and Maliga, P. (1998) RNA editing in mitochondria and plastids, in J. Bailey-Serres and D.R. Gallie (eds.), *A Look Beyond Transcription: Mechanisms Determining mRNA Stability and Translation in Plants*, American Society of Plant Physiologists, Rockville, MD pp. 153-161.

Nagatani, A. (1998) Regulated nuclear targeting, *Curr Opin Plant Biol* **1**, 470-474.

Oakes, J.V., Shewmaker, C.K., and Stalker, D.M. (1991) Production of cyclodextrins, a novel carbohydrate, in the tubers of transgenic potato plants, *Biotechnology* **9**, 982-986.

Odell, J.T., Hoopes, J.L., and Vermerris, W. (1994) Seed-specific gene activation mediated by the *cre/lox* site-specific recombination system, *Plant Physiol* **106**, 447-458.

Okita, T. and Rogers, J. (1996) Compartmentation of proteins in the endomembrane system of plant cells, *Annu Rev Plant Physiol Plant Mol Biol* **47**, 327-350.

Oliver, M.J., Quisenberry, J.E., Trolinder, N.L.G., and Keim, D.L. (1998) Control of plant gene expression, Patent # 5,723,765, Delta and Pine Land Co (Scott, MS).

Ow, D.W. (1996) Recombinase-directed chromosome engineering in plants, *Curr Opin Biotechn* 7, 181-186.

Pabo, C.O. and Sauer, R.T. (1992) Transcription factors: structural families and principles of DNA recognition, *Annu Rev Biochem* **61**, 1053-1095.

Pain, V.M. (1996) Initiation of protein synthesis in eukaryotic cells, *Eur J Biochem* **236**, 747-771.

Palauqui, J.-C. and Balzergue, S. (1999) Activation of systemic acquired silencing by localised introduction of DNA, *Curr Biol* **9**, 59-66.

Palauqui, J.-C., Elmayan, T., Pollien, J., and Vaucheret, H. (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.

Palauqui, J.-C. and Vaucheret, H. (1998) Transgenes are dispensable for the RNA degradation step of cosuppression, *Proc Natl Acad Sci USA* **95**, 9675-9680.

Pegg, A.E. (1989) Characteristics of ornithine decarboxylase from various sources, in S.-I. Hayashi (ed.), *Ornithine Decarboxylase: Biology, Enzymology, and Molecular Genetics*, Pergamon Press, New York, pp. 21-34.

Perlak, F.J., Fuchs, R.L., Dean, D.A., McPherson, S.L., and Fischhoff, D.A. (1991) Modification of the coding sequence enhances plant expression of insect control protein genes, *Proc Natl Acad Sci USA* **88**, 3324-3328.

Pesole, G., Liuni, S., Grillo, G., and Saccone, C. (1997) Structural and compositional features of untranslated regions of eukaryotic mRNAs, *Gene* **205**, 95-102.

Pestova, T.V. and Hellen, C.U.T. (1999) Ribosome recruitment and scanning: what's new?, *TIBS* **24**, 85-87.

Proudfoot, N. (1989) How RNA polymerase II terminates transcription in higher eukaryotes, *Trends Biochem Sci* **14**, 105-110.

Puchta, H. (1998) Towards targeted transformation in plants, *TIPS* **3**, 77-78.

Puchta, H. and Hohn, B. (1996) From centiMorgans to basepairs: homologous recombination in plants, *TIPS* **1**, 340-348.

Puchta, H., Swoboda, P. and Hohn, B. (1994) Homologous recombination in plants. *Experientia,* **50**, 277-284.

Radin, J.W. (1999) The technology protection system: revolutionary or evolutionary?, *Biotech Devel Monitor* **37**, 24.

Raikhel, N. (1992) Nuclear targeting in plants, *Plant Physiol* **100**, 1627-1632.

Ranish, J.A. and Hahn, S. (1996) Transcription: basal factors and activation, *Curr Opin Genet Devel* **6**, 151-158.

Ratcliff, F., Harrison, B. and Baulcombe, D. (1997) A similarity between viral defense and gene silencing in plants, *Science* **276**,1558-1560.

Razin, A. (1998) CpG methylation, chromatin structure and gene silencing – a three-way connection, *EMBO J* **17**, 4905-4908.

Richter, J.D. (1993) Translational control in development: a perspective, *Devel Genet* **14**, 407-411.

Rodriguez, C.M., Freire, M.A., Camilleri, C., and Robaglia, C. (1998) The *Arabidopsis thaliana* cDNAs coding for eIF4E and eIF(iso)4E are not functionally equivalent for yeast complementation and are differentially expressed during plant development, *Plant J* **13**, 465-473.

Roesler, K.R., Shorrosh, B.S., and Ohlrogge, J.B. (1994) Structure and expression of an *Arabidopsis* acetyl-coenzyme A carboxylase gene, *Plant Physiol* **105**, 611-617.

Rossi, F.M.V. and Blau, H.M. (1998) Recent advances in inducible gene expression systems, *Curr Opin Biotech* **9**, 451-456.

Rothnie, H.M. (1996) Plant mRNA 3'-end formation, *Plant Mol Biol* **32**, 43-61.

Rouwendal, G.J.A., Mendes, O., Wolbert, E.J.H., and deBoer, A.D. (1997) Enhanced expression in tobacco of the gene encoding green fluorescent protein by modification of its codon usage, *Plant Mol Biol* **33**, 989-999.

Rugh, C.L., Senecoff, J.F., Meagher, R.B., and Merkle, S.A. (1998) Development of transgenic yellow poplar for mercury phytoremediation, *Nature Biotech* **16**, 925-928.

Rushton, P.J. and Somssich, I.E. (1998) Transcriptional control of plant genes responsive to pathogens, *Curr Opin Plant Biol* **1**, 311-315.

Schaefer, D.G. and Zryd, J.-P. (1997) Efficient gene targeting in the moss *Physcomitrella patens*, *Plant J.* **11**, 1195-1206.

Schuler, M.A. (1998) Plant pre-mRNA splicing, in J. Bailey-Serres and D.R. Gallie (eds.), *A Look Beyond Transcription: Mechanisms Determining mRNA Stability and Translation in Plants*, American Society of Plant Physiologists, Rockville, MD, pp. 1-19.

Schwechheimer, C. and Bevan, M. (1998) The regulation of transcription factor activity in plants, *TIPS* **3**, 378-383.

Schwechheimer, C., Smith, C., and Bevan, M.W. (1998) The activities of acidic and glutamine-rich transcriptional activation domains in plant cells: design of modular transcription factors for high-level expression, *Plant*

Mol Biol **36**, 195-204.

Seraphin, B. (1995) Sm and Sm-like proteins belong to a large family: identification of proteins of U6 as well as the U1, U2 U4 and U5snRNPs, *EMBO J* **14**, 2089-2098.

Shantz, L.M., Viswanath, R., and Pegg, A.E. (1994) Role of the 5'-untranslated region of mRNA in the sysnthesis of S-adenosylmethionine decarboxylase and its regulation by spermine, *Biochem J* **302**, 765-772.

Simpson, G.G. and Filipowicz, W. (1996) Splicing of precursors to messenger RNA in higher plants: mechanism, regulation and sub-nuclear organization of the spliceosomal machinery, *Plant Mol Biol* **32**, 1-41.

Sinibaldi, R.M. and Mettler, I.J. (1992) Intron splicing and intron-mediated enhanced gene expression in monocots, *Progr Nucl Acid Res Mol Biol* **42**, 229-257.

Smith, H.M.S. and Raikhel, N.V. (1999) Protein targeting to the nuclear pore. What can we learn from plants?, *Plant Physiol* **119**, 1157-1163.

Spiker, S. and Thompson, W.F. (1996) Nuclear matrix attachment regions and transgene expression in plants, *Plant Physiol* **110**, 15-21.

Suzuki, T., Yan, Q., and Lennarz, W.J. (1998) Complex, two-way traffic of molecules across the membrane of the endoplasmic reticulum, *J Biol Chem* **273**, 10083-10086.

Thieffry, D. and Sarkar, S. (1998) Forty years under the central dogma, *TIBS* **23**, 312-316.

Thomas, K.R., Deng, C., and Capecchi, M.R. (1992) High-fidelity gene targeting in embryonic stem cells by using sequence replacement vectors, *Mol Cell Biol* **12**, 2919-2923.

Thykjaer, T., Finnemann, J., Schauser, L., Christensen, L., Poulsen, C., and Stougaard, J. (1997) Gene targeting approaches using positive-negative selection and large flanking regions, *Plant Mol Biol* **35**, 523-530.

Tinland, B. (1996) The integration of T-DNA in plant genomes, *TIPS* **1**, 178-184.

Udvardy, A. (1999) Dividing the empire: boundary chromatin elements delimit the territory of enhancers, *EMBO J*, **18**, 1-8.

Valcarel, J., Singh, R., and Green, M.R. (1995) Mechanisms of regulated pre-mRNA splicing, in A. Lamond (ed.), *Pre-mRNA Processing*, R.G. Landes Publishers, Georgetown, TX, pp. 97-112.

van Aarssen, R., Soetaert, P., Stam, M., Dockx, J., Gosselé, V., Seurinck, J., Reynaerts, A., and Cornelissen, M. (1995) *cry*IA(b) transcript formation in tobacco is inefficient, *Plant Mol Biol* **28**, 513-524.

van der Geest, A.H.M. and Hall, T.C. (1997) The *β-phaseolin* 5' matrix attachment region acts as an enhancer facilitator, *Plant Mol Biol* **33**, 553-557.

van Engelen, F.A., Schouten, A., Molthoff, J.W., Roosien, J., Salinas, J., Dirkse, W.G., Schots, A., Bakker, J., Gommers, F.J., Jongsma, M.A., Bosch, D., and Stiekema, W.J. (1994) Coordinate expression of antibody subunit genes yields high levels of functional antibodies in roots of transgenic tobacco, *Plant Mol Biol* **26**, 1701-1710.

van Steeg, H., Van Oostrom, C.T., Hodemaekers, H.M., Peters, L., and Thomas, A.A. (1991) The translation *in vitro* of rat ornithine decarboxylase mRNA is blocked by its 5' untranslated region in a polyamine-independent way, *Biochem J* **274**, 521-526.

Vaucheret, H., Nussaume, L., Palauqui, J.-C., Quilleré, I., and Elmayan, T. (1997) A transcriptionally active state is required for post-transcriptional silencing (co-suppression) of nitrate reductase host genes and transgenes, *Plant Cell* **9**, 1495-1504.

Verwoerd, T.C., van Paridon, P.A., van Ooyen, A.J.J., van Lent, J.W.M., Hoekema, A., and Pen, J. (1995) Stable accumulation of *Aspergillus niger* phytase in transgenic tobacco leaves, *Plant Physiol* **109**, 1199-1205.

Vitale, A. and Raikhel, N.V. (1999) What do proteins need to reach different vacuoles?, *TIPS* **4**, 149-155..

Voinnet, O., Vain, P., Angell, S., and Baulcombe, D.C. (1998) Systemic spread of sequence-specific transgene RNA degradation in plants is initiated by localized introduction of ectopic promoter-less DNA, *Cell* **177**, 187.

von Hippel, P.H. (1998) An integrated model of the transcription complex in elongation, termination, and editing, *Science* **281**, 660-665.

Walbot, V. (1999) Genes, genomes, genomics. What can plant biologists expect from the 1998 National Science Foundation plant genome program? *Plant Physiol* **119**, 1151-1155.

Wang, S., Browning, K.S., and Miller, W.A. (1997) A viral sequence in the 3'-untranslated region mimics a 5' cap in facilitating translation of uncapped mRNA, *EMBO J* **13**, 4107-4116.

Wassenegger, M. and Pélissier, T. (1998) A model for RNA-mediated gene silencing in higher plants, *Plant Mol Biol* **37**, 349-362.

Wassenegger, M. and Pélissier, T. (1999) Signalling in gene silencing, *TIPS* **4**, 207-209.

Weihe, A. and Börner (1999) Transcription and the architecture of promoters in chloroplasts, *TIPS* **4**, 169-170.

Weinmann, P., Gossen, M., Hillen, W., Bujard, H. and Gatz, C. (1994) A chimeric transactivator allows tetracycline-response gene expression in whole plants, *Plant J* **5**, 559-569.

Williamson, J.D., Hirsch-Wyncott, M.E., Larkins, B.A., and Gelvin, S.B. (1989) Differential accumulation of a transcript driven by the CaMV 35S promoter in transgenic tobacco, *Plant Physiol* **90**, 1570-1576.

Wolffe, A.P. (1994) Nucleosome positioning and modification: chromatin structures that potentiate transcription, *TIBS* **19**, 240-244.

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

Zhou, D.-X. (1999) Regulatory mechanism of plant gene transcription by GT-elements and GT-factors, *TIPS* **4**, 210-214.