

## Review

# The Quest to Understand the Basis and Mechanisms that Control Expression of Introduced Transgenes in Crop Plants

Ajay Kohli<sup>1</sup>

Pablo Gonzalez Melendi<sup>2</sup>

Rita Abranches<sup>3</sup>

Teresa Capell<sup>4</sup>

Eva Stoger<sup>5</sup>

Paul Christou<sup>6,\*</sup>

<sup>1</sup>Institute for Research on Environment & Sustainability (IRES); University of Newcastle upon Tyne; Newcastle, UK

<sup>2</sup>Centro de Investigaciones Biológicas (CSIC); Madrid, Spain

<sup>3</sup>Instituto de Tecnologia Quimica e Biologica; Plant Cell Biology Laboratory; Oeiras, Portugal and Universidade Nova de Lisboa

<sup>4</sup>Universitat de Lleida; <sup>6</sup>ICREA; Departament de Produccio Vegetal I Ciencia Forestal; Lleida, Spain

<sup>5</sup>Biology VII; RWTH Aachen; Aachen, Germany

\*Correspondence to: Paul Christou; ICREA-Universitat de Lleida; Departament de Produccio Vegetal I Ciencia Forestal; Av. Alcalde Rovira Roure 191; Lleida E-25198 Spain; Tel.: +34.973.702693; Fax: +34.973.238264; Email: Paul.Christou@icrea.es

Received 07/07/06; Accepted 07/12/06

Previously published online as a *Plant Signaling & Behavior* E-publication: <http://www.landesbioscience.com/journals/psb/abstract.php?id=3195>

## KEY WORDS

transgene expression, protein trafficking, silencing mechanisms, chromatin remodeling, crop plants

## ACKNOWLEDGEMENTS

Research in the laboratories of P.C. and T.C. is funded in part by the Pharma-Planta FP6 EU project. P.G.-M. and T.C. are funded by the "Ramón y Cajal" program of the Spanish Ministry of Education and Science. R.A. thanks Fundacao para a Ciencia e Tecnologia, Portugal.

## ABSTRACT

We discuss mechanisms and factors that influence levels and stability of expressed heterologous proteins in crop plants. We have seen substantial progress in this field over the past two decades in model experimental organisms such as *Arabidopsis* and tobacco. There is no question such studies have resulted in furthering our understanding of key processes in the plant cell and the elaboration of sophisticated models to explain underlying mechanisms that might influence the fate, levels and stability of expression of recombinant heterologous proteins in plants. However, very often, such information is not applicable outside these laboratory experimental models. In order to generate a knowledge basis that can be used to achieve high levels and stability of heterologous proteins in relevant crop plants it is imperative to perform such studies on the target crops. With this in mind, we discuss key elements of the process at the DNA, RNA and protein levels. We believe it is essential to discuss recombinant protein production in crops in a holistic manner in order to develop a comprehensive knowledge base that will in turn serve plant biotechnology applications well.

## INTRODUCTION

One of the major challenges in plant molecular biology is to generate transgenic plants that express input genes stably over generations. Many investigators have devoted enormous resources to unravel factors that are involved in the control of transgene expression in model plants such as *Arabidopsis*, tobacco, petunia and other species that are easily manipulated in the laboratory. A substantial body of knowledge has thus been generated which allowed researchers to formulate elegant theories and models that explain at least in part why populations of transgenic plants containing the same transgene(s) often exhibit unpredictable and to some extent unstable expression patterns. There is no question that such studies contributed significantly to our understanding of the complex processes involved in transgene expression post integration. A major limitation of these studies, however, stems from the fact that results obtained in model systems often are not congruent with the behavior of the same transgene(s) in crop plants. In our studies which span a period of over two decades, we focused on the behavior of transgenes in commercially important legumes and cereals. On occasion, we observed similar behavior in terms of transgene expression in crop plants to that reported in model species. However, frequently results were contradictory to those reported for *Arabidopsis* and tobacco.

Genetic transformation is a two-stage process: DNA transfer into the cell, followed by integration into the genome. The integration stage is much less efficient than the DNA transfer stage, with the result that only a small proportion of the cells that initially receive DNA actually become stably transformed.<sup>1,2</sup> Transgene integration, mediated by either *A. tumefaciens* or direct DNA transfer, is a random process and the sites of integration are believed to correlate with (1) the positions of naturally occurring chromosome breaks<sup>3</sup> and (2) transcriptionally active regions of the genome particularly the sub-terminal regions of the chromosomes, perhaps because the DNA is more accessible in these areas.<sup>4,5</sup> Once inside the cell, the integration of foreign DNA is largely dependent on host proteins involved in DNA replication, repair and recombination. The exact way in which such proteins interact with exogenous DNA is poorly understood and thus highly unpredictable at this stage. There is evidence that variations in the site of the transgene locus and its structure have a profound effect on the level and stability of transgene expression.<sup>6-9</sup>

Molecular analysis of transgenic soybean,<sup>10,11</sup> rice,<sup>6,7,12-15</sup> wheat<sup>16,17</sup> and maize<sup>18-20</sup> plants reveals a diverse range of transgene loci structures encompassing intact, truncated and internally rearranged copies and also copies interspersed with genomic DNA. These results are obtained irrespective of the transformation method used, direct DNA or

*Agrobacterium*. Southern blot analysis of long PCR products from the transgenic locus in maize followed by sequencing demonstrated that integration of foreign DNA was often accompanied by extensive transgene rearrangements, including truncations, internal deletions, interspersions of genomic DNA and inversions.<sup>18</sup> Rearrangements often cause loss of transgene expression but on occasion, plants with rearranged transgene copies continue to express the protein product normally. Conversely, while intact transgenes are often expressed, cases of non-expressed intact transgenes have been reported.<sup>21</sup> Rearrangements are rarely given serious consideration as a factor contributing to transgene inactivation and variable transgene expression. Undetected rearrangements have a profound effect on transgene expression with significant implications in studies attempting to correlate transgene copy number and expression levels.<sup>6</sup> The influence of epigenetic silencing mechanisms may be exaggerated when rearrangements are the cause of transgene inactivation. These results highlight some of the difficulties that may cause erroneous conclusions from transgene expression studies.

Multi protein complexes that directly or indirectly affect the structure of the chromatin surrounding an introduced transgene are now recognized as important factors that may influence expression. The latter could be a consequence of the different organization patterns of euchromatin/heterochromatin in the nucleus. This broader view on regulation of transgene expression is gaining importance as we become more aware of additional levels of regulation other than DNA sequences such as promoters or enhancers. Although chromatin-mediated factors are far from being completely understood, the fact that several model and crop species are now fully sequenced, or will be in the near future, along with advances in proteomics, will facilitate the elucidation of these mechanisms. DNA methylation and post-translational modifications of histones are key mechanisms in chromatin-mediated regulation of gene expression. There is also evidence that biochemical changes are associated with physical changes in transgene locus organization. Epigenetic regulation and silencing phenomena are related to such changes in locus organization and therefore influence expression. Intrinsic differences among species in terms of chromatin organization, as a result of genome size and GC content, may also be a factor influencing transgene expression.

Different and often incompatible biochemical reactions occur in a eukaryotic cell simultaneously. This is possible by compartmentalization in specialized membrane-bound organelles where essential enzymes and structural proteins are directed by specific signaling mechanisms.<sup>22</sup> Protein sorting to the correct sub-cellular compartment(s) is dependent on specific signals in their peptide sequence. Trafficking among organelles is mediated by vesicular transport routes, which might vary depending on the physiology of particular cell types. Consequently, protein signaling and the nature of the cell and/or host plant can determine the sub-cellular fate, and in turn, expression levels of heterologous proteins. Such proteins expressed in transgenic plants provide ideal tracers to investigate the complex networks of protein sorting, intracellular trafficking mechanisms and deposition. In situ tracking of recombinant proteins will not only shed light on our understanding of these basic cell processes. It can also aid optimization of strategies to increase expression levels, improve protein recovery and select the best host for particular applications by targeting proteins to optimal destinations in a subcellular or tissue-dependent manner, thus avoiding aberrant localization.

In this review, we discuss key factors that in our collective experience have been shown to exert a profound influence on levels

and stability of transgene expression in crop plants. We also discuss mechanisms of gene regulation that may influence transgene expression, in particular those that are related to chromatin modifications. In our discussion we highlight a number of wide misconceptions that have been propagated in the scientific literature, particularly in respect to the effect of transgene copy number on levels and stability of expression, the purported advantages of transgenes integrating in transcriptionally active euchromatic regions in the host genome and the ability of plants with polyploid genomes to buffer against some of the undesirable effects of transgene integration and chromosome aberrations.

## THE ORGANIZATION OF TRANSGENIC LOCI IN CROP PLANTS IS A MAJOR FACTOR THAT DETERMINES STABILITY AND LEVELS OF EXPRESSION

The simplest transgenic locus comprises a single copy of an introduced transgene with the unit expression construct (promoter-ORF-terminator) intact. Intact single copy integration events resulting in the transgene being flanked by favorable or neutral endogenous sequences in the transcriptionally active euchromatic regions of the host genome have long been proposed as ideal integration events (Fig. 1). However, these are rarely achieved in major crop species in practice and in situations where such a transgenic locus is created this might not always lead to optimal stability and levels of expression.

**Why are stably expressing-single copy transgene loci formed at a relatively low frequency in crop plants?** Enzymes dictating DNA structure and function in processes such as DNA synthesis, replication, transcription, repair and degradation are active in the cell during most phases of the cell cycle. Hence, functional enzymes such as topoisomerases, gyrases, helicases, polymerases, ligases and importantly, endo- and exo-nucleases are readily available in the cell. Whatever the method of delivery, transgenes are directly exposed to this enzymatic machinery. Additionally, transgene delivery elicits a wound response in the host cell. Wounding was shown to be necessary for *Agrobacterium*-mediated T-DNA transfer.<sup>23</sup> Bombardment-mediated transgene delivery was also shown to elicit a wound response.<sup>24</sup> Upregulation of DNA repair and degradation enzymes and additional DNA manipulation enzymes such as (retro)transposases<sup>25</sup> and integrases<sup>26</sup> are one of the consequences of this wound response. The latter two enzymes facilitate DNA integration into the host genome. The probability of transgenic DNA escaping any modification by these enzymes is extremely low. Higher recombination frequency under wound-induced stress was put forward as a means of DNA repair<sup>27</sup> and this might serve as an alternative route to integration. Most stress-mediated integration occurs through illegitimate recombination, hallmarks of which such as microhomology and filler DNA at recombination junctions are characteristic of transgene-host genome junctions through *Agrobacterium*<sup>28</sup> or bombardment-mediated transformation.<sup>8</sup>

Transformation frequency was demonstrated to be higher for cells in the M or G<sub>2</sub> phase compared to the S phase of the cell cycle in one study<sup>29</sup> but in other later studies it was shown to be higher in the S and G<sub>2</sub><sup>30</sup> or S and M<sup>31</sup> phase of the cell cycle. This suggests that open chromatin structure may not always be a pre-requisite for integration. However, it also implies that integration is likely to occur not only in the euchromatic but also in the heterochromatic regions. For these reasons, it is not difficult to recognize that the probability of achieving intact, single copy transgene integration at a transcriptionally favorable site of the host genome is not very high.

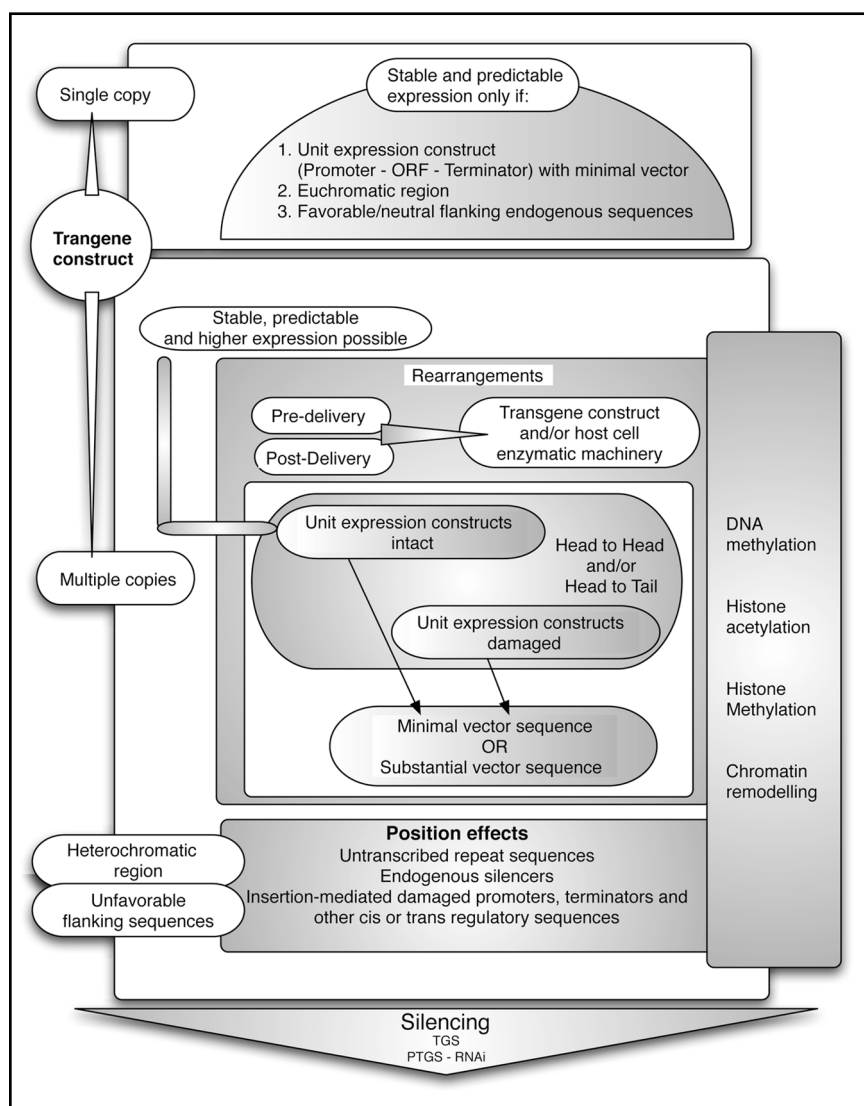


Figure 1. Fate of a transgene construct in a plant transformation experiment. The construct may integrate as a single copy or as multiple copies. In either case, a number of factors as discussed in the text influence stability and predictability of expression in the primary transformant and progeny. If multiple copy integration events do not contain all copies as intact expression units, factors such as DNA methylation, histone methylation and acetylation and chromatin remodeling, may lead to transgene silencing through mechanisms of transcriptional or post-transcriptional gene silencing (TGS/PTGS).

Despite attack by host enzymes and the influence of host cell processes on the transgene, various transgenic plants generated through different transformation methods do contain single copy integration events. Most single copy integration events express the transgenic protein in a stable and predictable manner. However, some express the transgene in primary transformants but lose expression in subsequent generations.<sup>21,32</sup> Precise reasons for this are not known but it was demonstrated that even subtle changes in the integrated transgene construct can cause unpredictable expression patterns associated with altered DNA cytosine methylation patterns in the transgene and the flanking endogenous sequences.<sup>32</sup> DNA methylation patterns are reset during the transition from the sporophytic to the gametophytic phases of the plant life cycle. Evidently a single such transition is sufficient for the cellular machinery to initiate the gene silencing processes. It is still an open question if associated changes in methylation patterns are the cause or the effect

of silencing.<sup>33</sup> However, the low frequency of transgenic crop plants containing a single copy of the transgene that also exhibit stable and predictable expression patterns over generations invariably contain no changes at all in comparison to the delivered construct, contain no filler DNA at the junctions, occupy a transcriptionally active site on the genome and maintain stable methylation patterns.<sup>34,35</sup>

**Why single copy transgene loci might not always be desirable or ideal in transgenic crops?** As described earlier the probability of generating functional single copy integration events is small. Those single copy integration events that express stably are useful if relatively low levels of the gene product are adequate to confer the desired phenotype. The single copy transgenic cotton event MON826, containing the Bt cryIA gene expresses the protein stably over a number of generations to render the plant insecticidal in the field.<sup>36</sup> However, in most cases higher expression levels are desirable in transgenic plants than those achieved by single copy transgenes. The promoter and enhancer cis-elements combined with RNA and protein stability and/or turnover factors facilitate higher expression of the transgene product. However, these have limited value in cases where substantial amounts of transgenic protein are required. It has been demonstrated that multi-copy transgenic plants can produce higher levels of the desired gene product, compared to their single transgene counterparts. Transgenic cotton containing two Bt genes is a useful example.<sup>37</sup> Tang et al.<sup>38</sup> demonstrated a positive correlation between copy number and expression levels for acyl-acyl carrier protein (Acyl-ACP) thioesterase transgenes in *Brassica napus* after analyzing three different populations generated after (1) self pollination of the transgenic plants; (2) cross pollination with non-transgenic parental lines; and (3) doubled haploid lines from F1 of (2) above. In addition, higher amounts of the transgenic protein(s) may be required in applications such as molecular pharming. Huang and colleagues<sup>39-41</sup> provide useful data in support of multi-copy transgenic plants. When using rice plants to produce pharmaceuticals and nutraceuticals, the expression level of the recombinant protein is absolutely critical since this dictates the economics of downstream processing. For example, Nandi et al.<sup>42</sup> have achieved levels of approximately 5 mg human lactoferrin per gram of flour and have estimated production costs at approximately \$6 per gram of pharmaceutical-grade protein. If the expression level dropped to 0.5 mg or 0.05 mg g<sup>-1</sup>, the cost would increase to \$60 and \$400 per gram, respectively. Our experience shows that, regardless of the transformation method, lines selected for high recombinant protein expression levels tend to contain three or more transgene copies.

In nature most plant genes exist in more than one copy or as gene families. For example, the redox genes such as peroxidase, catalase, superoxide dismutase, NADPH-P450-reductase ferredoxin-NADP reductase exist in 4–20 copies in *Arabidopsis* and rice as evident from their genome sequence at the TIGR genome database website.<sup>43,44</sup> A number of other plants e.g., wheat, barley, maize, potato and

tobacco also contain multiple copies of the redox genes per haploid genome. Interestingly, most biotic and abiotic stress resistance genes in plants exist as gene families. For example in rice the resistance gene against *Xanthomonas oryzae* belongs to a family of Xa21 resistance genes.<sup>45</sup> Indeed, members of the gene families may exhibit specificity of spatio-temporal expression patterns but multiple copies often contribute to higher amounts of the product per se, e.g., histones,<sup>46</sup> rubisco,<sup>47</sup> ribosomal proteins<sup>48</sup> etc. Hence, single copy integration of the transgene may not always be ideal or desirable.

**Loci with multiple copies of integrated transgenes in crops.** Integration of multiple copies of the transgene has mostly been seen as a disadvantage and closely associated with transgene silencing.<sup>49,50</sup> However, many examples demonstrate that multi-copy transgenic plants can express the gene product at a higher level stably compared to their single copy counterparts.<sup>40</sup> Most multi-copy transgenic plants are single locus events, consequently transgenes, either multiple copies or different transgenes integrated in one locus, do not segregate in subsequent generations. Such loci may frequently contain rearranged transgenes but what is important in terms of expression is the nature of the rearrangements. Head-to-tail arrays of multiple copies are useful and frequently result in stable expression, whereas head-to-head or tail-to-tail arrangements generally result in silencing.

Within the context of transgenic plants, rearrangement is a generic term that may encompass any unintended changes to the construct in the amplifying/transforming host and/or in the plant host prior to and/or during integration in the genome.<sup>7,8</sup> Figure 1 shows that numerous factors and processes concordant with rearrangements that lead to multi-copy integration might also lead to silencing. It is important to reiterate that only the head-to-tail rearrangement may be useful in terms of expression, provided that the multiple expression units involved are all intact and preferably contain minimal vector sequences. The head-to-head or tail-to-tail rearrangements can lead to formation of four-way Holliday junctions that on one hand elicit recombination and DNA break repair, while at the same time may create physical steric hindrance to limit processes such as replication and transcription. Such structures when resolved may lead to deletions that can scramble local endogenous sequences, or they may rearrange further as a result of recombination to yield scrambled transgenic sequences. Such scrambled sequences most likely will not yield any protein product at all. However, occasionally such sequences may still retain the capacity to initiate transcription and result in aberrant RNA synthesis. Aberrant RNA-mediated silencing associated with DNA methylation may lead to the silencing of an otherwise expressing locus at a different position in the genome.<sup>6</sup> Rearranged sequences may also provide templates for RNAi. The snap-back hairpin structures are classic templates for generating siRNA which will in turn lead to silencing an expressing locus at a different position on the genome.<sup>50</sup> This 'trans-silencing' effect of an aberrant 'silencing locus' can even extend to an expressing locus of the same construct in an independent transformant, when the 'silencing locus' is moved into the latter through crossing.<sup>52</sup>

Rearranged multiple copy constructs in transgenic lines generated irrespective of the transformation method used encompass substantial portions of the vector sequence. In fact most rearrangement junctions are situated in highly recombinogenic sequences of the vector such as the origin of replication and promoters for bacterial antibiotic resistance genes.<sup>8</sup> Plant transformation using 'vectorless' minimal cassettes through bombardment discourages such rearrangements and results in a higher frequency of transgenic plants expressing the transgenes stably and predictably.<sup>53-55</sup>

A variety of transgenic loci created either through *Agrobacterium* or bombardment in rice,<sup>7,8</sup> oats,<sup>56</sup> maize,<sup>18</sup> wheat,<sup>16,17,57</sup> soybean<sup>58,59</sup> and *Arabidopsis*<sup>60</sup> all exhibit common features suggesting that the host cell plays a major role in transgene integration. The integration and rearrangement junctions are mostly formed through micro-homology, occasionally with filler DNA or through blunt end ligation.<sup>8</sup> Both of these are characteristic of illegitimate recombination. Multiple copies are more often than not clustered together and do not segregate. Integration mechanisms facilitating such loci have been proposed<sup>7</sup> and though these have not been completely confirmed there are increasing reports describing similar transgenic loci in a variety of plants generated through different transformation methods.<sup>61,62</sup> Loci created through different transformation methods e.g., PEG<sup>61</sup> *Agrobacterium*<sup>60,62</sup> and bombardment<sup>8</sup> share common molecular features such as presence of topoisomerase I sites on the genomic sequences immediately flanking the transgene. Such common features suggest a conserved mechanism largely operative through the host cell machinery. The importance of the host cell machinery in generating the particular structure and organization of transgenic loci as proposed by Kohli et al.<sup>7</sup> and Pawlowski and Somers<sup>55</sup> was recently validated through elegant experiments involving movement of DNA from chloroplast to the nucleus.<sup>63</sup> Interestingly the genomic loci created by integration of the chloroplast gene(s) moving into the nucleus were similar to those created through *Agrobacterium* or bombardment-mediated transformation.

The organization of the different types of transgenic loci, generated predominantly through the host cell machinery as described above, can be viewed in a hierarchical manner at the genome level. Even single or multi-copy integration events paint a complex picture of genome function in relation to transgene function. Transgenic loci far apart may come to lie close together at different stages of the cell cycle and transgene activity.<sup>5</sup> The nature (euchromatic or heterochromatic) and influence of flanking host sequences on the transgenes at the site of integration is known as position effect. Integration in heterochromatic repeat sequences or adjacent to endogenous silencers and insulators renders the event useless in terms of expression. This and related chromatin remodeling aspects are addressed in more detail in the subsequent section. An element of position effects that is often neglected in studies focusing on transgene expression is the influence of deletions, inversions and/or filler sequences at the site of insertion.<sup>8,60</sup> T-DNA or transposon insertion strategies to study gene function are possible because of changes in the host genome due to transgene insertion. Insertion of T-DNA or transposons renders a gene non-functional by interruption in the coding region or through structural changes as discussed earlier. Such structural alterations also constitute an element of position effects directing transgene expression.

It is thus clear that both, integrating transgene sequences and host sequences at the site of integration are prone to changes and that such changes affect transgene expression. Hence a thorough analysis of the position, structure and organization of the transgenic locus is needed to ascertain the usefulness of particular transgenic plants.

**Variation in levels of expression amongst plants from genetically identical lineages.** Wide variation in the expression of transgenes in progeny derived from one primary transformant has been reported in different plants. Jones et al.<sup>64</sup> described variability in transgene expression in petunia transformed with *Agrobacterium*. Kohli et al.<sup>6</sup> quantified levels of marker genes such as *gusA* and *bar* in transgenic rice plants. Within a particular lineage (R0 and R1) GUS expression levels ranged from 0.1–40 pmol MU per mol<sup>-1</sup> per hour<sup>-1</sup>. Bar

expression levels in progeny were also highly variable. Rice plants with enhanced  $\beta$ -carotene (vitamin A) content also showed similar behavior observed as different shades of yellow color in individual rice grains.<sup>65</sup> Additional examples of such variation amongst levels of expression in individual plants within a lineage derived from the same primary transformant (R0) have also been documented in rice transformed with the *Arabidopsis GAI* gene.<sup>66</sup> Reduction in the height of different rice transformants expressing this gene<sup>66,67</sup> was shown to be proportional to the levels of *GAI* mRNA and also *GAI* protein, with different plants within each lineage showing a wide variation in levels of *GAI* expression and consequently different heights. Such wide variation in end product accumulation for different genes cannot be accounted for by the zygosity status of transgenic plants (i.e., hemi- versus homo-zygous sub-populations<sup>68</sup>).

Careful evaluation of end-product and enzymatic activities in progeny derived from individual primary transformants expressing genes of the polyamine biosynthetic pathway varied widely.<sup>69-73</sup> Expression of the *Datura samdc* cDNA under the control of the maize Ubi-1 promoter and first intron in transgenic rice plants was studied in order to elucidate underlying factors contributing to different levels of expression of transgenes within a particular lineage derived from the same primary transformant.<sup>74</sup> Three particular transgenic plant lineages (R0 and R1) harboring one, two or three integration sites for the transgene, in all cases linked at one genetic locus, respectively were characterized fully and studied in detail (Fig. 2). By selecting the population in this manner, the possibility of position effects influencing expression levels within one particular plant lineage was excluded. Steady state levels of the *Datura samdc* mRNA varied considerably among members of the same lineage. Such variation also manifested at the enzyme activity and end-product levels (Fig. 2). A detailed analysis of the methylation status of the transgene and also of the host genome revealed profound changes in their degree of methylation. These results demonstrated that genomic DNA and the *Datura samdc* transgene within a lineage were methylated to different degrees and this correlated to transgene expression levels in each individual plant.

Progeny from tobacco plants harboring the *nptII* reporter gene showed highly variable levels of *nptII* expression.<sup>75</sup> Nuclear run-on transcription analysis confirmed that the reduced steady-state *nptII* mRNA levels were not a result of reduced transcription initiation. Low expressing lines were associated with increased methylation of the *nptII* sequence. Fu et al., demonstrated three alternative silencing mechanisms in a three transgene single copy locus in rice.<sup>32</sup> Three distinct types of silencing effects associated with different methylation patterns were observed in a line of rice plants carrying a single copy 6.6 Kb locus comprising the *bar*, *hpt* and *gusA* transgenes. This includes a novel form of transcriptional silencing involving methylation of cytosine residues only at non-conventional acceptor sites in the transgene(s) coding region. They demonstrated that silencing in the three transgenes arose independently, in different plants derived from the same primary transformant (R1, R2, R3) despite the stability of the transgenic locus, even though the basic structure of the locus, transgene dosage and position effects remained constant. Furthermore, they demonstrated that different silencing effects could occur concurrently in adjacent heterologous transgenes in the same plant, with no evidence for spreading of silenced states or methylation patterns from one transgene to another.

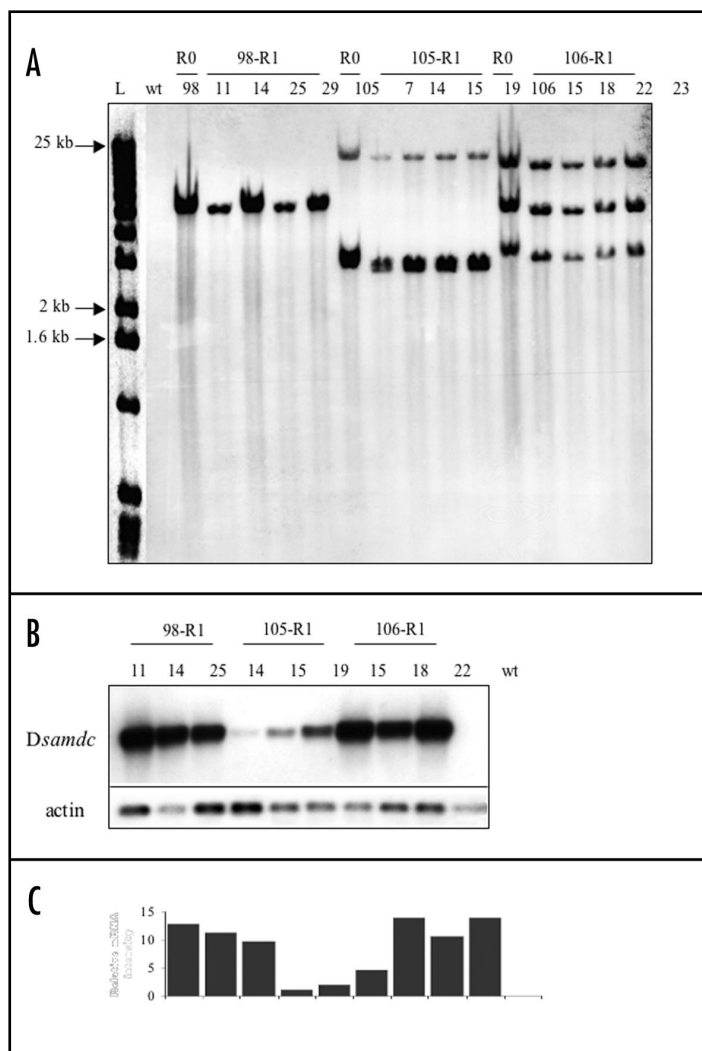


Figure 2. Variable levels of expression in specific transgenic rice lineages (primary transformants and R1 progeny) expressing the *Datura samdc* cDNA. (A) Three independent primary transformants and progeny, lines 98, 105 and 106, exhibiting integration of *Samdc* in 1, 2 or 3 sites, respectively (Southern blot). (B) Quantitative RT-PCR analysis of *Datura samdc* transcript levels from total RNA extracted from wild type (wt) and R1 progeny from plants in A above; numbers in gel (bottom number) represent siblings from same parental line (top number). Actin amplification (25 cycles) used as internal control.

## CHROMATIN MEDIATED (TRANS)GENE REGULATION

Our views of eukaryotic chromatin have changed from the traditional notion of a static to a highly dynamic functional entity that is essential for regulating cellular functions. The activity state of a gene is determined not only by sequence specific regulatory factors but also by multi-protein complexes that directly or indirectly affect the structure of the chromatin surrounding the gene. An increasing number of studies on histone and DNA modifications have strengthened the idea of chromatin environment as key to the establishment and maintenance of transcriptional activity or repression, and these need to be considered when we picture a transgene integrated into the host genome.

Chromatin remodeling complexes were initially studied in the context of gene activation, but they have later become also associated with transcriptional repression.<sup>76</sup> It has also become evident that

remodeling complexes are important regulators of both local and global transcriptional repression, and it is possible that different genomic regions, e.g., highly repetitive regions versus low copy sequences, are silenced by different sets of chromatin- and DNA-modifying activities.<sup>77</sup>

Epigenetic information can be introduced by cytosine methylation and by marking nucleosomal histones. Post-translational modifications of histones include acetylation and methylation among others. Interestingly and in contrast to the universal nature of the DNA code, the histone language differs among animals, plants and fungi (reviewed in ref. 78). Enzymes that modify the histones and chromatin remodeling machines that utilize ATP (ref. 78 and refs. therein) affect chromatin structure and function. There is accumulating evidence for an interrelation between DNA methylation and histone modifications.<sup>79</sup> Four major mechanisms govern these interactions: DNA and histone methylation, histone acetylation and chromatin remodeling/heterochromatin. Three marks are recognized for heterochromatin: cytosine methylation, hypo-acetylation of lysine residues in histones, and histone H3 methylated at lysine-9. In contrast, euchromatin is characterized by methylation of histone H3 at lysine 4, high levels of histone acetylation, and low levels of cytosine methylation. Histone acetylation at the level of lysine residues requires the activity of two groups of enzymes, histone acetyltransferases (HATs) and histone deacetylases (HDACs). The unraveling of these mechanisms will contribute to devise ways to increase transgene stability and levels of expression. We can no longer set our goals solely on the design of better plant expression vectors; we need to consider the environment where the transgene is embedded and also the effect(s) that the transgenic sequence will have on the surrounding chromatin.

Evidence has indicated that two classes of Cis-acting elements—transcriptional enhancers and matrix-attachment regions (MARs)—might regulate transcription through the acetylation of histones.<sup>80,81</sup> In plants, it has been shown that the pea PetE enhancer associates with the nuclear matrix, suggesting that MARs may function as entry sites for HATs during transcriptional activation.<sup>82</sup> This enhancer seems to activate transcription by associating with the nuclear matrix, mediating the acetylation of histones on the promoter and the nearby coding region, and resulting in an altered chromatin structure. It has also become clear that some of the biochemical modifications important for packaging heterochromatic domains are used to silence genes in euchromatin.<sup>83</sup> The initiation of heterochromatin formation, or similar silencing of euchromatic domains, requires acquisition of at least one epigenetic mark. In some organisms, entry into this cycle, i.e. the initial step towards heterochromatin formation, could be triggered by the repetitive nature of the locus, rather than the primary DNA sequence, and this could apply also to transgenes in tandem. One can indeed view silencing of individual genes residing in euchromatin as a process that involves establishment of a small micro-cosmos of heterochromatin.<sup>84</sup> However, there are important differences between small and large plant genomes. In *Arabidopsis*, euchromatin is characterized by high levels of dimethylated lysine-4 in H3, whereas high levels of dimethylated lysine-9 in H3 were observed for heterochromatic chromocenters.<sup>85</sup> A comparative study over the whole genome of different plants revealed that, in small genome plants, high levels of H3 dimethylated lysine-9 were confined to constitutive heterochromatin, whereas, in large genome plants, dimethylated lysine-9 was uniformly distributed.<sup>86</sup> This suggests that large genome plants have to silence dispersed repetitive sequences also within euchromatic regions with a high level of methylated H3

lysine-4. Indeed, the genome size and the way heterochromatin is arranged in the nucleus are interrelated, and the methylation of histone H3 in euchromatin of plant chromosomes depends on basic nuclear DNA content. The packaging of certain genes into a chromosomal matrix is regulated via sophisticated chromatin remodeling mechanisms that define whether and when a gene becomes accessible to the transcription machinery. In plants the analysis of transgenes and transposable elements exhibits parallels to epigenetic control mechanisms, which can affect the stable expression of transgenes. The levels of epigenetic silencing in plants are multiple and highly complex and it is clear that RNA interference, chromatin modification and silencing are interdependent mechanisms. At this point, we can contemplate the origins of epigenetic regulation (see ref. 87 and accompanying articles in ref. 87). For RNA-mediated silencing and DNA methylation there is evidence that they have evolved as part of a host defense mechanism against viruses and parasitic DNA. The substrate for post-transcriptional gene silencing (PTGS) (or RNAi) and transcriptional gene silencing (TGS) seen in plants—double stranded RNA—is a common intermediate in the life cycle of many viruses and (retro)transposons.

Mechanisms to coordinate gene regulation are also important. A large number of studies in animal cells have provided evidence that chromatin structure plays a role in regulating the expression of clustered genes.<sup>88</sup> Sharing regulatory elements could be one reason to keep genes clustered together on a chromosome. Similarly, multi-copy transgenes could be considered gene clusters since they share promoters and other regulatory sequences or even because of homology of the coding sequence. In fact, there is indication that in interphase nuclei there is co-localization of transgene copies that are located far apart in the metaphase chromosomes.<sup>89</sup> It is possible to accept a concerted activation of different genes, or transgenes, that will be transcribed synchronously. The effect of transgene locus structure on expression is mostly evident in transgenic organisms where independent transformants carrying the same transgene show different levels of expression. There is now much support that these differences are at least in part a consequence of biochemical changes in chromatin. There is also evidence that the biochemical changes are associated with physical changes in transgene locus organization. Multiple copy transgene insertions tend to more likely form a local heterochromatin-like structure that favors repression of transgene transcription. Moreover, interactions between several copies of transgenes alter chromatin organization in the host nuclei<sup>90</sup> that can also affect transgene expression. Recent evidence suggests that organization, not duplication, triggers silencing in a complex transgene locus in rice.<sup>91</sup> These authors claim that two or even multiple copies of a given promoter may be used for transformation without inducing silencing, if their genomic integration does not result in aberrant transcription of the promoters. Finally, it has also been postulated that the host genome size may be related to the genomic reaction when receiving a foreign gene. This could be related to the different organization patterns of euchromatin/heterochromatin in the nucleus as described above.

## RECOMBINANT PROTEIN TRAFFICKING: HOST CELL-SPECIFIC AND POST-TRANSLATIONAL FACTORS AFFECT LEVELS AND STABILITY OF EXPRESSION

Other than DNA and RNA-related factors, expression levels and stability of heterologous proteins in transgenic plants are also influenced by their correct assembly, post-translational modifications, quality control mechanisms, sub-cellular sorting and final deposition to the appropriate compartment(s) within cells and tissues. A combination of a broad range of biochemical methods (sub-cellular fractionation, pulse and chase experiments after radioactive labeling, glycan analysis by MALDI-TOF mass spectrometry, immunoprecipitation, western blot analysis) and immunolocalization techniques (confocal and electron microscopy) have been used to study the stability, molecular interactions, intracellular pathways, delivery and deposition of recombinant proteins expressed in plants.<sup>92</sup> Heterologous proteins also provide ideal tracers to investigate the complex networks of protein sorting and intracellular trafficking mechanisms. In this section, we focus on the cellular and post-translational factors that influence expression of recombinant proteins in plants.

**Cellular factors controlling the trafficking of recombinant proteins and effects on expression.** One of the most effective strategies to increase the stability of recombinant proteins is to combine an N-terminal leader peptide to direct the mRNAs to the rough endoplasmic reticulum (ER), and a C-terminal K/HDEL signal peptide to retrieve the proteins back into the protective environment of the ER. Thus,  $\beta$ -phaseolin-KDEL accumulates at 0.5% of total extractable protein in tobacco leaves as compared to  $\beta$ -phaseolin without KDEL that accumulates at 0.05%.<sup>93</sup> Efficient cisternal ER targeting, occasionally leading to the formation of protein bodies inside, and localization in ER domains, such as the nuclear envelope space and ribosome-coated vesicles, were observed in vegetative tissues of crops and model species.<sup>94-96</sup> In seed storage tissues of rice, wheat and tobacco, KDEL-tagged proteins accumulated in ER-derived protein bodies (PB) and in protein storage vacuoles<sup>96-98</sup> (PSV). The localization of ER resident proteins in rice endosperm PBs can be explained by passive trapping during the aggregation of prolamins within specific domains of the ER. Traveling beyond ER compartments to PSV was unexpected since rice glutelins (PSV-resident storage proteins) are known to pass through the Golgi complex en route to the vacuole. The co-localization of calreticulin, a native reticuloplasm, and a KDEL-tagged single chain antibody in PSVs together with the absence of labeling in the dense vesicles of the trans-Golgi network suggest a Golgi-independent route from the ER to PSV.<sup>95</sup> Such an alternative route has been proposed for the sorting of SH-EP, a vacuolar protease with a KDEL signal peptide, to the storage vacuole from the ER by a pathway that by-passes the Golgi cisternae<sup>99</sup> and for the accumulation of wheat prolamins in PSVs.<sup>100</sup> The same route directed a small proportion of phaseolin-KDEL to the lytic vacuole in transgenic tobacco leaves, where it was degraded,<sup>94</sup> following the transport pathway put forward for the turnover of reticuloplasm.<sup>101</sup> Therefore, presence of a KDEL signaling might not always guarantee ER sub-cellular targeting and consequently secretion or delivery to seed storage vacuoles might occur.

In other cases, secretion is preferred, but may not always be obtained. An efficient assimilation of organic phosphorus from phytate in the culture medium was achieved by transforming *Arabidopsis* plants with a phytase construct to target the protein to the ER. In the absence of any additional sorting signal,<sup>102</sup> this enzyme

would be expected to travel through the default secretory pathway<sup>103</sup> and directed to the extra-cellular space where it normally breaks down phytate releasing digestible phosphorus. Phytase is a glycoprotein exhibiting Golgi-modified complex glycans, which is normally secreted in different plant expression systems such as soybean cell cultures and *Medicago truncatula* leaves.<sup>104,105</sup> Phytase was not secreted rather it accumulated in PBs and PSVs in rice endosperm.<sup>106</sup> In situ hybridization at the electron microscopy level showed that accumulation of phytase in PBs did not require mRNA targeting to the specific ER domain where prolamin mRNA is normally directed.<sup>107</sup> Analysis of the glycan profiles showed two such types: [1] typical vacuolar glycans with  $\alpha(1,3)$ -fucose and  $\beta(1,2)$ -xylose residues that indicate Golgi modifications and [2] oligomannose type, a feature of early N-glycosylation in the ER. These two types of glycans could account for the phytase labeling in PSVs and PBs, respectively. As opposed to rice leaves, no terminal N-acetylglucosamine (GlcNAc) residues, as those found in secreted proteins, were detected in phytase accumulating in the endosperm. This suggests the existence of predetermined tissue-dependent intracellular routes of transport.<sup>97</sup> In fact, all recombinant proteins expressed in seed storage tissues to date were retained intra-cellularly in protein bodies of rice, soybean, wheat and tobacco<sup>108-112</sup> or surprisingly, in maize starch granules.<sup>113</sup> Seeds contain tissues which are developmentally programmed to synthesize and store proteins, starch and lipids to nurture the embryo upon germination. Consequently, the cell's architecture and function would prevail over protein targeting signals or host plant expression system to prevent secretion in accumulating tissues and avoid the escape of valuable reserves. This can be exploited for long-term storage of biologically functional recombinant proteins, for example in molecular pharming applications.

The structural information contained in a protein can also influence its sub-cellular fate. A decameric secretory antibody<sup>114</sup> (SIgA/G) comprising two immunoglobulins (two light chains and two heavy chains each), a joining chain (J) and a secretory component (SC) in tobacco leaves was only partially secreted, although all four moieties (light chain, heavy chain, J and SC) were expected to travel along the secretory route.<sup>115</sup> Most of the antibody was sequestered into the ER or delivered to the vacuole, with expression levels between 5–8 % of total soluble proteins.<sup>115</sup> However the parent tetrameric Guy's 13 IgG antibody<sup>116</sup> was efficiently secreted in tobacco leaves and expression levels of 1% of total soluble protein were achieved. This suggests that information for intra-cellular retention and vacuolar delivery resides in the  $\alpha$ -domains contained in the hybrid  $\alpha/\gamma$ -heavy chains of the SIgA/G antibody. A second example is zeolin. This chimeric storage protein contains phaseolin, from common bean, and the repeated Pro-rich regions of maize  $\gamma$ -zein.<sup>93</sup> Expression of recombinant phaseolin in tobacco leaves led to the formation of protein bodies in the vacuole<sup>117</sup> and expression levels of 0.05% of total soluble protein. However, when fused to domains of  $\gamma$ -zein alone, phaseolin accumulated in protein bodies in the cytoplasm and expression levels were elevated to 3.5%.<sup>90</sup> This suggests that the zein domains are dominant over the vacuolar signaling of phaseolin and provide sufficient information to induce the formation of such protein bodies in the cytoplasm. It is evident, therefore that some specific protein domains are involved in sub-cellular targeting. This is particularly important for proteins targeted to the vacuole as such targeting can lead to protein degradation and, therefore, influence expression levels. Although the mechanisms of vacuolar delivery are not fully understood, identification of such domains and development of an understanding of their interactions might predict where in the cell a recombinant protein is delivered.

Table 1 **Subcellular fate of recombinant proteins expressed in transgenic plants**

| Recombinant Protein                                                              | Sorting Signal                                            | Subcellular Localization                                                                                                                                                                                                                                                                                                                                       | References                   |
|----------------------------------------------------------------------------------|-----------------------------------------------------------|----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|------------------------------|
| Guy's 13 IgG to surface adhesin I/II of <i>Streptococcus mutants</i>             | None (default secretory route expected)                   | High efficient secretion in tobacco leaf protoplasts, independently of its glycosylation status                                                                                                                                                                                                                                                                | 115, 116, 122                |
| Guy's 13 secretory IgA/G to surface adhesin I/II of <i>Streptococcus mutants</i> | None (default secretory route expected)                   | Partial secretion in tobacco leaf<br>ER retention and vacuolar delivery of Ig fragments in tobacco leaf<br>Unassembled moieties in rice endosperm prolamins protein bodies (PB)                                                                                                                                                                                | 114, 115<br>120              |
| <i>Aspergillus niger</i> phytase                                                 | None (default secretory route expected)                   | Functional, assembled antibody in rice endosperm glutelin protein storage vacuoles (PSV)<br>Extracellular secretion in cultured cells of soybean and <i>Arabidopsis thaliana</i> plants<br>Prolamin bodies in wheat endosperm<br>Apoplast and vacuole in <i>Medicago truncatula</i> leaf<br>PBs and PSVs in rice endosperm<br>Secreted in rice callus and leaf | 102, 104<br>97<br>105<br>106 |
| Zeolin (chimeric protein composed of phaseolin and $\gamma$ -zein)               | None                                                      | ER-protein bodies in tobacco leaf                                                                                                                                                                                                                                                                                                                              | 93                           |
| <i>Pisum sativum</i> vicilin                                                     | ER (KDEL)                                                 | ER-protein bodies in tobacco leaf                                                                                                                                                                                                                                                                                                                              | 94                           |
| ScFvT84.66 antibody to the NA3 epitope of the carcinoembryonic antigen           | ER (KDEL)                                                 | PB and PSV in rice endosperm<br>ER-derived vesicles in rice callus and leaf                                                                                                                                                                                                                                                                                    | 96                           |
| Human serum albumin                                                              | ER (KDEL)                                                 | ER in wheat leaf<br>Prolamin bodies in wheat endosperm                                                                                                                                                                                                                                                                                                         | 97                           |
| <i>Phaseolus vulgaris</i> $\beta$ -phaseolin                                     | ER (K/HDEL)                                               | ER and PSVs in tobacco seeds<br>Intact protein in ER of tobacco leaf<br>Fragmented protein in the vacuole of tobacco leaf<br>PSVs of rice endosperm<br>Degradation products in protein bodies within the vacuoles of tobacco leaf<br>Secreted to the apoplast in tobacco leaf upon saturation of the vacuolar targeting mechanism when overexpressed           | 95,98<br>112<br>117          |
| <i>Pisum sativum</i> legumin                                                     | Vacuole                                                   | Inclusion bodies at the periphery of prolamins protein bodies in wheat endosperm                                                                                                                                                                                                                                                                               | 97,110                       |
| Human lysozyme                                                                   | Globulin signal peptide<br>Glutelin signal peptide        | PSVs in rice endosperm                                                                                                                                                                                                                                                                                                                                         | 108                          |
| Bovine $\beta$ -casein                                                           | Lectin signal sequence                                    | PSVs in soybean seeds                                                                                                                                                                                                                                                                                                                                          | 109                          |
| <i>E. coli</i> enterotoxin B (LT-B)                                              | Native LT signal peptide<br>$\gamma$ -zein signal peptide | Starch granules of amyloplasts in cells below the aleurone layer in maize endosperm                                                                                                                                                                                                                                                                            | 113                          |
| Glycoprotein B of Human cytomegalovirus (HCMV)                                   | Rice glutelin 3 signal sequence                           | PSVs in tobacco endosperm                                                                                                                                                                                                                                                                                                                                      | 111                          |

**Post-translational factors affecting the sub-cellular deposition of recombinant proteins and effects on expression.** Correct folding and glycosylation of recombinant proteins also influence their sub-cellular fate and accumulation levels. Protein folding is assisted by the ER-resident chaperone BiP by binding to hydrophobic regions. Presence of permanently misfolded proteins, for example upon treatment with tunicamycin or by the expression of folding-defective mutant proteins, triggers a signaling pathway known as "unfolded protein response" (UPR), which upregulates the expression of BiP in yeast, mammals and plants.<sup>118</sup> Quality control in the ER is a two-step process: retention and degradation. The prolonged interactions of unfolded proteins with BiP bring about ER-associated protein degradation (ERAD). This involves the translocation of the unfolded proteins back into the cytoplasm where they are degraded by the proteasome. Quality control can also occur along the secretory pathway by targeting proteins to the vacuole for degradation after

passage through the Golgi.<sup>119</sup> A phaseolin variant encompassing a sulfur rich (high methionine) 15 amino acid sequence and a HDEL tag (HiMet phaseolin-HDEL) alters the structure of the protein by introducing an  $\alpha$ -helix domain.<sup>97</sup> HiMet phaseolin-HDEL travels along the secretory route and it is directed to the vacuole of tobacco embryo cells via the dense vesicles of the trans-Golgi, where it is degraded in the early developmental stages, whereas phaseolin-HDEL accumulates in PSVs. This route to the vacuole might coexist with a Golgi-independent pathway as discussed earlier.

Immuno-assays and co-immunoprecipitation showed that the assembly of all four components of a SIgA/G antibody occurred in rice endosperm.<sup>120</sup> Comparison of immunogold labeling in transgenic lines expressing either single or all four components showed an assembly-dependent distribution. Unassembled components were preferentially found in PBs whereas the assembled antibody with antigen-binding activity was localized to PSVs. Unassembled chains,

likely bound to BiP, would remain in the ER for a prolonged period of time and then become incorporated into prolamin PBs, whereas correct assembly would promote their efficient transport to PSVs.<sup>120</sup>

A number of proteins expressed in transgenic plants (for example, most pharmaceutically active proteins for molecular pharming applications) require enzymatic glycosylation to fold correctly and retain functionality.<sup>105,121</sup> This complex process is highly conserved among eukaryotes. It commences with N-glycosylation in the ER by the co-translational addition of a precursor oligosaccharide Glc3Man9GlcNAc2 from a specialized lipid in the ER membrane, called dolichol, to Asn residues of the sequence Asn-X-Ser/Thr. Then, further modifications occur in the Golgi apparatus, with some specific differences between plants and mammals. In plants,  $\beta(1,2)$ -xylose is linked to  $\beta$ -mannose whereas  $\alpha(1,3)$ -fucose is linked to proximal N-acetylglucosamine moieties. Furthermore, plants lack terminal  $\beta(1,3)$ -galactose moieties.<sup>105</sup> Inhibition of N-glycosylation by tunicamycin activates UPR, upregulates BiP and results in protein degradation by ERAD. However, when the heavy chain of the Guy's 13 IgG antibody was mutated to remove the two potential N-glycosylation sites this had no effect on the folding, assembly or secretion of the functional IgG molecule, with antigen-binding activity as demonstrated in tobacco mesophyll protoplasts.<sup>122</sup> Although tunicamycin treatment and transfection with plasmids encoding aglycosylated variants of the heavy chain activate UPR (co-immunoprecipitated BiP with aglycosylated heavy chain increased 1,3-fold as compared to the wild type) lack of N-glycosylation did not affect IgG folding and assembly. This could be an exception more than the general rule and might only occur in specific proteins (immunoglobulins) that are not affected by removal of N-glycans.

Table 1 summarizes the subcellular fate of a number of recombinant proteins expressed in model species and crops. Unfortunately, examples of heterologous proteins (containing the same targeting signals) that have been expressed in the same tissues in different host plant systems are rare. An example is phytase, which was secreted in rice<sup>106</sup> and *Medicago truncatula* leaves.<sup>105</sup> However, the scarce information in the literature does not rule out any influence of the host system. We reiterate that the information contained in the recombinant protein itself (targeting signals, specific domains), the tissue where the protein is expressed, its assembly status and other post-translational modifications clearly affect protein sorting and deposition in both crops and model species thus contributing to levels and stability of expression.

## CONCLUSIONS

Our understanding of factors that influence levels and stability of heterologous proteins introduced into crop plants, is improving continuously. This is a direct consequence of early intensive research focusing on DNA-related factors involving the introduced transgene(s) and the recipient host genome. Integration of transgenes into the nuclear genome of plants, with a very small number of exceptions is an entirely random process, based on illegitimate recombination. We do not have any control where in the genome a transgene will integrate. Furthermore, we cannot predict the number and integrity of integrated transgene copies. Consequently, the first step in assessing the long-term utility of a transgenic plant in terms of stable and predictable expression of a transgene(s) is detailed analysis of the transgene locus. Single copy integration events though generally useful are rare due to the host cell nucleic acid modifying enzyme

activities. Rearranged multi-copy integration events are more common. Such events do not lead to silencing if the expression unit in all copies is intact. Additionally, if higher amounts of the transgene product are required then multi-copy integration events are more useful. Analysis of structural features of transgenic loci reveals that the host-cell enzymatic machinery largely controls the integration event(s) independent of how the transgene is delivered.

Advances in our understanding of the mechanisms involved in the integration process, the influence of chromatin modifications and chromatin-mediated transgene regulation also appear to influence levels and stability of expression, post integration. At the protein level, the overall picture emerging from biochemical and cell biology studies in different systems suggests that a number of cellular and post-translational factors govern the fate of expressed recombinant proteins. An otherwise secreted recombinant protein can be retained, depending on default tissue-specific routes. This suggests that the cell's architecture and physiology play a crucial role in protein deposition. The state of assembly of complex macromolecules, such as immunoglobulins, influences their final destiny. ER quality control is upregulated when glycosylation is impaired, targeting the defective proteins for ERAD. Hence understanding the mechanisms and fundamental cellular processes that influence protein sorting, trafficking and deposition, also have a significant role to play in achieving stable and predictable levels of expression of recombinant proteins in plants.

## References

- Altpeter F, Baisakh N, Beachy R, Bock R, Capell T, Christou P, Daniell H, Datta S, Dix, PJ, Fauquet C, Huang N, Kohli A, Mooibroek H, Nicholson L, Nguyen TT, Nugent G, Raemakers K, Romano A, Somers AD, Stoger E, Taylor N, Visser R. Particle bombardment and the genetic enhancement of crops: Myths and realities. *Mol Breed* 2005; 15:305-27.
- Hensgens LAM, Debakker EPHM, Vanosuygrok EP, Rueb S, Vandemark F, Vandermaas HM, Vanderveen S, Koomangersmann M, Thart L, Schilperoort RA. Transient and stable expression of *gusA* fusions with rice genes in barley and perennial ryegrass. *Plant Mol Biol* 1993; 23:643-69.
- Salomon S, Puchta H. Capture of genomic and T-DNA sequences during double-strand break repair in somatic plant cells. *EMBO J* 1998; 17:6086-95.
- Barakat A, Gallois P, Raynal M, Mestre-Ortega D, Sallaud C, Guiderdoni E, Delseny M, Bernardi G. The distribution of T-DNA in the genomes of transgenic *Arabidopsis* and rice. *FEBS Lett* 2000; 47:161-4.
- Kohli A, Twyman RM, Abranches R, Wegel E, Stoger E, Christou P. Transgene integration, organization and interaction in plants. *Plant Mol Biol* 2003; 52:247-58.
- Kohli A, Gahakwa D, Vain P, Laurie DA, Christou P. Transgene expression in rice engineered through particle bombardment: molecular factors controlling stable expression and transgene silencing. *Planta* 1999; 208:88-97.
- Kohli A, Leech M, Vain P, Laurie DA, Christou P. Transgene organization in rice engineered through direct DNA transfer supports a two-phase integration mechanism mediated by the establishment of integration hot-spots. *Proc Natl Acad Sci USA* 1998; 95:7203-8.
- Kohli A, Griffiths S, Palacios N, Twyman RM, Vain P, Laurie DA, Christou P. Molecular characterization of transforming plasmid rearrangements in transgenic rice reveals a recombination hotspot in the CaMV 35S promoter and confirms the predominance of microhomology-mediated recombination. *Plant J* 1999; 17:591-601.
- Arencibia A, Gentinetta E, Cuzzoni E, Castiglione S, Kohli A, Vain P, Leech M, Christou P, Sala F. Molecular analysis of the genome of transgenic rice (*Oryza sativa* L.) plants produced via particle bombardment or intact cell electroporation. *Mol Breed* 1998; 4:99-109.
- McCabe DE, Swain WF, Martinell BJ, Christou P. Stable transformation of soybean (*Glycine max*) by particle acceleration. *Bio/Technology* 1988; 6:923-6.
- Christou P, Swain WF, Yang NS, McCabe DE. Inheritance and expression of foreign genes in transgenic soybean plants. *Proc Natl Acad Sci USA* 1989; 86:7500-4.
- Christou P, Ford TL, Kofron M. Genotype-independent stable transformation of rice (*Oryza sativa*) plants. *Bio/Technology* 1991; 9:957-62.
- Oard JH, Linscombe SD, Braverman MP, Jodari F, Blouin DC, Leech M, Kohli A, Christou P. Development, field evaluation and agronomic performance of transgenic herbicide resistant rice. *Mol Breed* 1996; 2:359-68.
- Christou P, Ford TL, Kofron M. Rice Genetic Engineering: a Review. *Trends Biotech* 1992; 10:239-46.
- Gahakwa D, Maqbool SB, Fu X, Sudhakar D, Christou P, Kohli A. Transgenic rice as a system to study the stability of transgene expression: multiple heterologous transgenes show similar behaviour in diverse genetic backgrounds. *Theor Appl Genet* 2000; 101:388-99.
- Stöger E, Williams S, Keen D, Christou P. Constitutive versus seed specific expression in transgenic wheat: temporal and spatial control. *Trans Res* 1999; 8:73-82.
- Stöger E, Williams S, Keen D and Christou P. Molecular characteristics of transgenic wheat and the effect on transgene expression. *Trans Res* 1999; 7:463-71.

18. Mehlo L, Mazithulela G, Twyman RM, Boulton MI, Davies J and Christou P. Structural analysis of transgene rearrangements and effects on expression in transgenic maize plants generated by particle bombardment. *Maydica* 2000; 45:277-87.
19. Mehlo L, Gahakwa D, Nghia PT, Loc NT, Capell T, Gatehouse J, Gatehouse A, Christou P. An alternative strategy for sustainable pest resistance in genetically enhanced crops. *Proc Nat Acad Sci USA* 2005; 102:7812-6.
20. Drakakaki G, Marcel S, Glahn RP, Lund L, Periaigh S, Fischer R, Christou P, Stoger E. Endosperm specific co-expression of recombinant soybean ferritin and Aspergillus phytase in maize results in significant increases in the levels of bioavailable iron. *Plant Mol Biol* 2005; 59:869-80.
21. Meza TJ, Kamford D, Hakelien AM, Evans I, Godager LH, Mandal A, Jakobsen KS, Aalen RB. The frequency of silencing in *Arabidopsis thaliana* varies highly between progeny of siblings and can be influenced by environmental factors. *Trans Res* 2001; 10:53-67.
22. Alberts B, Bray D, Johnson A, Lewis J, Raff M, Roberts K, Walter P. Intracellular compartments and transport. In: Robertson M, ed. *Essential Cell Biology. An introduction to the molecular biology of the cell*. New York: Garland Publishing Inc. 1998:447-79.
23. Lippincott BB, Lippincott JA. Genus *Agrobacterium* and plant tumorigenesis. *Ann Rev Microbiol* 1975; 29:377-405.
24. Hunold R, Bronner R, Hahne G. Early events in microprojectile bombardment-cell viability and particle location. *Plant J* 1994; 5:593-604.
25. Wessler SR. Plant retrotransposons: Turned on by stress. *Current Biol*. 1996; 6:959-961.
26. Jeung AU, Cho SK, Lee SJ, Shin JS. Characterization of Ty3-gypsy-like elements in *Hibiscus syriacus*. *Molecules Cells* 2005; 19:318-27.
27. Puchta H, Hohn B. From centiMorgans to base pairs: Homologous recombination in plants. *Trends Plant Sci* 1996; 1:340-8.
28. Gheysen G, Villarroel R, Vanmontagu M. Illegitimate recombination in plants—A model for T-DNA integration. *Genes & Dev* 1991; 5:287-97.
29. Iida A, Yamashita T, Yamada Y, Morikawa H. Efficiency of particle bombardment-mediated transformation is influenced by cell-cycle stage in synchronized cultured cells of tobacco. *Plant Physiol* 1991; 97:1585-7.
30. Arias RS, Filichkin SA, Strauss SH. Divide and conquer: development and cell cycle genes in plant transformation. *Trends Biotech* 2006; 24:267-73.
31. Villemont E, Dubois F, Sangwan RS, Vasseur G, Bourgeois Y, SnagwanNorreel BS. Role of the host cell cycle in the *Agrobacterium*-mediated genetic transformation of *Petunia*: Evidence of an S-phase control mechanism for T-DNA transfer. *Planta* 1997; 201:160-72.
32. Fu X, Kohli A, Twyman RM, Christou P. Alternative silencing effects involve distinct types of non-spreading cytosine methylation at a three-gene, single-copy transgenic locus in rice. *Mol Gen Genet* 2000; 263:106-18.
33. Ehrlich M. Expression of various genes is controlled by DNA methylation during mammalian development. *Lect Cell Biochem* 2003; 88:899-910.
34. Nagaya S, Kato K, Ninomiya Y, Horie R, Sekine M, Yoshida K, Shinmyo A. Expression of randomly integrated single complete copy transgenes does not vary in *Arabidopsis thaliana*. *Plant Cell Physiol* 2005; 46:438-44.
35. Howarth JR, Jacquet JN, Doherty A, Jones HD, Cannell ME. Molecular genetic analysis of silencing in two lines of *Triticum aestivum* transformed with the reporter gene construct pAHC25. *Ann Appl Biol* 2005; 146:311-20.
36. Vaughn, T Cavato T, Brar G, Coombe T, DeGooyer T, Ford S, Groth M, Howe A, Johnson S, Kolacz K, Pilcher C, Purcell J, Romano C, English L, Pershing J. A method of controlling corn rootworm feeding using a *Bacillus thuringiensis* protein expressed in transgenic maize. *Crop Sci* 2005; 45:931-8.
37. U.S. Environmental Protection Agency. 2002. [www.epa.gov/pesticides/biopesticides/ingredients/factsheets/factsheet\\_006487.htm](http://www.epa.gov/pesticides/biopesticides/ingredients/factsheets/factsheet_006487.htm)
38. Tang J, Searth R, Fristensky B. Effects of genomic position and copy number of Acyl-ACP thioesterase transgenes on the level of the target fatty acids in *Brassica napus* L. *Mol Breed* 2003; 12:71-81.
39. Huang JM, Nandi S, Wu Ly, Yalda D, Bartley G, Rodriguez R, Lonnerdal B, Huang N. Expression of natural antimicrobial human lysozyme in rice grains. *Mol Breed* 2002; 10:83-94.
40. Tu JM, Zhang GA, Datta K, Xu CG, He YQ, Zhang QF, Khush GS, Datta SK. Field performance of transgenic elite commercial hybrid rice expressing *Bacillus thuringiensis* delta-endotoxin. *Nature Biotech* 2000; 18:1101-4.
41. Ye GY, Shu QY, Yao HW, Cui HR, Cheng XY, Hu C, Xia YW, Gao MW, Altosaar I. Field evaluation of resistance of transgenic rice containing a synthetic cry1Ab gene from *Bacillus thuringiensis* Berliner to two stem borers. *J Econ Entomol* 2001; 94:271-6.
42. Nandi S, SuzukiYA, Huang JM, Yalda D, Pham P, Wu LY, Bartley G, Huang N, Lonnerdal B. Expression of human lactoferrin in transgenic rice grains for the application in infant formula. *Plant Sci* 2002; 163:713-22.
43. The TIGR *Arabidopsis thaliana* database <http://www.tigr.org/tdb/e2k1/ath1/>; The TIGR rice database <http://www.tigr.org/tdb/e2k1/osa1/>
44. Dietz KJ. Plant peroxiredoxins. *Ann Rev Plant Biol* 2003; 54:93-107.
45. Richter TE, Ronald P. The evolution of disease resistance genes. *Plant Mol Biol* 2000; 42:195-204.
46. Gantt JS, Lenvik TR. *Arabidopsis thaliana* H1 Histone - Analysis of 2 members of a small gene family. *Eur J Biochem* 1991; 202:1029-39.
47. Jamet E, Paramentier Y, Durr A, Fleck J. Gene encoding the small subunit of Rubisco belong to 2 highly conserved subfamilies in Nicotianaceae. *J Mol Evol* 1991; 33:226-36.
48. Cooke R, Raynal M, Laudie M, Delseny M. Identification of members of gene families in *Arabidopsis thaliana* by contig construction from partial cDNA sequences: 106 genes encoding 50 cytoplasmic ribosomal proteins. *Plant J* 1997; 11:127-40.
49. Matzke AJM, Neuhuber F, Park YD, Ambros PF, Matzke MA. Homology-dependent genes silencing in transgenic plants-epistatic silencing loci contain multiple copies of methylated transgenes. *Mol Gen Genet* 1994; 244:219-29.
50. Jorgensen RA. Cosuppression, flower colour patterns and metatable gene expression states. *Science* 1995; 268:686-91.
51. Lu SF, Shi R, Tsao CC, Yi XP, Li LG, Chaing VL. RNA silencing in plants by the expression of siRNA duplexes. *Nuc Acids Res* 2004; 32:e171.
52. Vaucheret H, Beclin C, Elmayan T, Feuerbach F, Godon C, Morel JB, Mourrain P, Palauqui JC, Vernhettes S. Transgene-induced gene silencing in plants. *Plant J* 1998; 16:651-9.
53. Fu X, Duc LT, Fontana S, Bong BB, Tinjuanglin P, Sudhakar D, Twyman RM, Christou P, Kohli A. Linear transgene constructs lacking vector backbone sequences generate low-copy-number transgenic plants with simple integration patterns. *Trans Res* 2000; 9:11-9.
54. Loc NT, Tinjuanglin P, Cohen M, Gatehouse AMR, Christou P, Gatehouse JA. Linear transgene constructs lacking vector backbone sequences generate transgenic rice plants which accumulate higher levels of proteins conferring insect resistance. *Mol Breed* 2002; 9:231-44.
55. Agrawal PK, Kohli A, Twyman RM, Christou P. Transformation of plants with multiple cassettes generates simple transgene integration patterns and high expression levels. *Mol Breed* 2005; 16:247-60.
56. Pawlowski WP, Somers DA. Transgenic DNA integrated into the oat genome is frequently interspersed by host DNA. *Proc Natl Acad Sci USA* 1998; 95:12106-10.
57. Jackson SA, Zhang P, Chen WP, Phillips RL, Friebe B, Muthukrishnan S, Gill BS. High-resolution structural analysis of biolistic transgene integration into the genome of wheat. *Theor Appl Genet* 2001; 103:56-62.
58. Christou P and Swain WF. Co-transformation frequencies of foreign genes in soybean tissues. *Theor Appl Genet* 1990; 79:337-41.
59. Christou P, Murphy JE and Swain WF. Stable transformation of soybean callus by electroporation and root formation from transformed callus. *Proc Natl Acad Sci USA* 1987; 84:3962-6.
60. Sawasaki T, Takahashi M, Goshima N, Morikawa H. Structures of transgene loci in transgenic *Arabidopsis* plants obtained by particle bombardment: Junction regions can bind to nuclear matrices *Gene* 1998; 218:27-35.
61. Craig W, Gargano D, Scotti N, Nguyen TT, Lao NT, Kavanagh TA, Dix PJ, Card T. Direct gene transfer in potato: A comparison of particle bombardment of leaf explants and PEG-mediated transformation of protoplasts. *Plant Cell Rep* 2005; 24:603-11.
62. Radchuk VV, Van DT, Klocke E. Multiple gene co-integration in *Arabidopsis thaliana* predominantly occurs in the same genetic locus after simultaneous in planta transformation with distinct *Agrobacterium tumefaciens* strains. *Plant Sci* 2005; 168:1515-23.
63. Huang CY, Ayliffe MA, Timmis JN. Simple and complex nuclear loci created by newly transferred chloroplast DNA in tobacco. *Proc Natl Acad Sci USA*. 2004; 101:9710-5.
64. Jones JDG, Gilbert DE, Grady KL, Jorgensen RA. T-DNA Structure and gene-expression in *Petunia* plants transformed by *Agrobacterium-tumefaciens* C58 derivatives. *Mol Gen Genet* 1987; 207 (2-3):478-85.
65. Ye XD, Al-Babili S, Kloti A, Zhang J, Lucca P, Beyer P, Potrykus I. Engineering the provitamin A (beta-carotene) biosynthetic pathway into (carotenoid-free) rice endosperm. *Science* 2000; 287:303-5.
66. Fu X, Sudhakar X, Peng J, Richards DE, Christou P, Harberd NP. Expression of *Arabidopsis GAI* in transgenic rice represses multiple rice gibberellin responses. *Plant Cell* 2001; 13:1791-802.
67. Peng J, Richards DE, Hartley NM, Murphy GP, Devos KM, Flintham JE, Beales J, Fish LJ, Worland AJ, Pelica F, Sudhakar D, Christou P, Snape JW, Gale MD, Harberd NP. Green revolution' genes encode mutant gibberellin response modulators. *Nature* 1999; 400:256-61.
68. James VA, Avar C, Worland B, Snape JW, Vain P. The relationship between homozygous and hemizygous transgene expression levels over generations in populations of transgenic rice plants. *Theor Appl Genet* 2002; 104:553-61.
69. Capell T, Escobar C, Liu H, Burtin D, Lepri O, Christou P. Over-expression of the oat arginine decarboxylase cDNA in transgenic rice (*Oryza sativa* L.) affects normal development patterns in vitro and results in putrescine accumulation in transgenic plants. *Theor Appl Genet* 1998; 97:246-54.
70. Noury M, Bassie L, Lepri O, Kurek I, Christou P, Capell T. A transgenic rice cell lineage expressing the oat arginine decarboxylase (adc) cDNA constitutively accumulates putrescine in callus and seeds but not in vegetative tissues. *Plant Mol Biol* 2000; 43:537-44.
71. Bassie L, Noury M, Lepri O, Lahaye T, Christou P, Capell T. Promoter strength influences polyamine metabolism and morphogenic capacity in transgenic rice tissues expressing the oat adc cDNA constitutively. *Trans Res* 2000; 9:33-42.
72. Lepri O, Bassie L, Safwat G, Pham TH, Pham TN, Hölttä E, Christou P, Capell T. Over-expression of the human ornithine decarboxylase cDNA in transgenic rice plants alters the polyamine pool in a tissue-specific manner. *Mol Gen Genet* 2001; 266:303-12.
73. Pham TN, Bassie L, Safwat G, Pham TH, Lepri O, Rocha P, Christou P, Capell T. Reduction in the endogenous arginine decarboxylase transcript levels in rice leads to depletion of the putrescine and spermidine pools with no concomitant changes in the expression of downstream genes in the polyamine biosynthetic pathway. *Planta* 2003; 218:125-34.
74. Pham TH, Bassie L, Safwat G, Pham TN, Christou P, Capell T. Expression of a heterologous S-adenosylmethionine decarboxylase cDNA in plants demonstrates that changes in S-Adenosyl-L-Methionine Decarboxylase activity determine levels of the higher polyamines spermidine and spermine. *Plant Physiol* 2002; 129:1744-54.

75. Ingelbrecht I, Vanhoudt H, Van Montagu M, Depicker A. Post-transcriptional silencing of reporter transgenes in tobacco correlates with DNA methylation. *Proc Natl Acad Sci USA* 1994; 91:10502-6.
76. Tyler JK, Kadonaga JT. The "dark side" of chromatin remodeling: repressive effects on transcription. *Cell* 1999; 99:443-6.
77. Lusser A. Acetylated, methylated, remodeled: chromatin states for gene regulation. *Curr Opin in Plant Biol* 2002; 5:437-43.
78. Loidl P. A plant dialect of the histone language. *Trends Plant Sci* 2004; 9:84-90.
79. Dobosy JR, Selker EU. Emerging connections between DNA methylation and histone acetylation. *Cellular Mol Life Sci* 2001; 58:721-7.
80. Agalioiti T, Lomvardas S, Parekh B, Yie J, Maniatis T, Thanos D. Ordered recruitment of chromatin modifying and general transcription factors to the IFN-beta promoter. *Cell* 2000; 103:667-78.
81. Fernandez LA, Winkler M, Grosschedl R. Matrix attachment region-dependent function of the immunoglobulin mu enhancer involves histone acetylation at a distance without changes in enhancer occupancy. *Mol Cell Biol* 2001; 21:196-208.
82. Chua YL, Watson LA, Gray JC. The transcriptional enhancer of the pea plastocyanin gene associates with the nuclear matrix and regulates gene expression through histone acetylation. *Plant Cell* 2003; 15:1468-79.
83. Richards EJ, Elgin SCR. Epigenetic codes for heterochromatin formation and silencing: rounding up the usual suspects. *Cell* 2002; 108:489-500.
84. Fahrner JA, Baylin SB. Heterochromatin: stable and unstable invasions at home and abroad. *Gene Dev* 2003; 17:1805-12.
85. Jasencakova Z, Soppe WJ, Meister A, Gernand D, Turner BM, Schubert I. Histone modifications in Arabidopsis-high methylation of H3 lysine 9 is dispensable for constitutive heterochromatin. *Plant J* 2003; 33:471-80.
86. Houben A, Demidov D, Gernand D, Meister A, Leach CR, Schubert I. Methylation of histone H3 in euchromatin of plant chromosomes depends on basic nuclear DNA content. *Plant J* 2003; 33:967-37.
87. Riddihough G, Pennisi E. The evolution of epigenetics. *Science* 2001; 293:1063.
88. Sproul D, Gilbert N, Bickmore WA. The role of chromatin structure in regulating the expression of clustered genes. *Nat Rev Genet* 2005; 6:775-81.
89. Abranches R, Santos AP, Wegel E, Williams S, Castilho A, Christou P, Shaw P, Stoger E. Widely separated multiple transgene integration sites in wheat chromosomes are brought together at interphase. *Plant J* 2000; 24:713-23.
90. Pecinka A, Kato N, Meister A, Probst AV, Shubert I, Lam, E. Tandem repetitive transgenes and fluorescent chromatin tags alter local interphase chromosome arrangement in Arabidopsis thaliana. *J Cell Sci* 2005; 118:3751-8.
91. Yang G, Lee Y-H, Jiang Y, Kumpatla SR, Hall TC. Organization, not duplication, triggers silencing in a complex transgene locus in rice. *Plant Mol Biol* 2005; 58:351-66.
92. Hood E. Where, oh where has my protein gone? *Trends in Biotechnology* 2004; 22:53-5.
93. Mainieri D, Rossi M, Archinti M, Beluci M, De Marchis F, Vavassori S, Pompa A, Arcioni S, Vitale A, Zeolin. A new recombinant storage protein constructed using maize  $\gamma$ -zein and bean phaseolin. *Plant Physiol* 2004; 136:3447-56.
94. Wandelt CI, Khan MRI, Craig S, Schroeder HE, Spencer D, Higgins TJV. Vicilin with carboxy-terminal KDEL is retained in the endoplasmic reticulum and accumulates to high levels in the leaves of transgenic plants. *Plant J* 1992; 2:181-92.
95. Frigerio L, Pastres A, Prada A, Vitale A. Influence of KDEL on the fate of trimeric or assembly-defective phaseolin: selective use of and alternative route to vacuoles. *Plant Cell* 2001; 13:1109-26.
96. Torres E, González-Melendi P, Stöger E, Shaw P, Twyman R, Nicholson L, Vaquero C, Fischer R, Christou P, Perrin Y. Native and artificial reticuloplasmins co-accumulate in distinct domains of the endoplasmic reticulum and in post-endoplasmic reticulum compartments. *Plant Physiol* 2001; 127:1212-23.
97. Arcalis E, Marcel S, Altmann F, Kolarich D, Drakakaki G, Fischer R, Christou P, Stöger E. Unexpected deposition patterns of recombinant proteins in post-endoplasmic reticulum compartments of wheat endosperm. *Plant Physiol* 2004; 136:3457-66.
98. Pueyo JJ, Chrispeels MJ, Herman EM. Degradation of transport-competent destabilized phaseolin with a signal for retention in the endoplasmic reticulum occurs in the vacuole. *Planta* 1995; 196:586-96.
99. Toyooka K, Okamoto T, Minamikawa T. Mass transport of proform of a KDEL-tailed cysteine proteinase (SH-EP) to protein storage vacuoles by endoplasmic reticulum-derived vesicle is involved in protein mobilization in germinating seeds. *J Cell Biol* 2000; 148:453-64.
100. Levanony H, Rubin R, Altschuler Y, Galili G. Evidence for a novel route of wheat storage proteins to vacuoles. *J Cell Biol* 1992; 119:1117-28.
101. Tamura K, Yamada K, Shimada T, Hara-Nishimura I. Endoplasmic reticulum-resident proteins are constitutively transported to vacuoles for degradation. *Plant J* 2004; 39:393-402.
102. Richardson AE, Hadobas PA, Hayes JE. Extracellular secretion of Aspergillus phytase from Arabidopsis roots enables plants to obtain phosphorus from phytate. *Plant J* 2001; 25:641-9.
103. Vitale A, Denecke J. The endoplasmic reticulum-gateway of the secretory pathway. *Plant Cell* 1999; 11:615-28.
104. Li J, Hegeman CE, Hanglon RW, Lacy GH, Denbow DM, Grabau EA. Secretion of active recombinant phytase from soybean cell-suspension cultures. *Plant Physiol* 1997; 114:1103-11.
105. Abranches R, Marcel S, Arcalis E, Altmann F, Feveiro P, Stöger E. Plants as bioreactors: a comparative study suggests that *Medicago truncatula* is a promising production system. *J Biotech* 2005; 120:121-34.
106. Drakakaki G, Marcel S, Arcalis E, Altmann F, González-Melendi P, Fischer R, Christou P, Stöger E. The intracellular fate of recombinant phytase is tissue-dependent. *Plant Physiol* 2006; 141:578-86.
107. Choi SB, Wang C, Muench DG, Ozawa K, Franceschi VR, Wu Y, Okita TW. Messenger RNA targeting of rice seed storage proteins to specific ER subdomains. *Nature* 2000; 407:765-7.
108. Yang D, Gou F, Liu B, Huang N, Watkins SC. Expression and localization of human lysozyme in the endosperm of transgenic rice. *Planta* 2003; 216:597-603.
109. Philip R, Darnowski DW, Maughan PJ, Vodkin LO. Processing and localization on bovine  $\beta$ -casein expressed in transgenic soybean seeds under control of a soybean lectin expression cassette. *Plant Sci* 2001; 161:323-35.
110. Stöger E, Parker M, Christou P, Casey R. Pea legumin overexpressed in wheat endosperm assembles into an ordered paracrystalline matrix. *Plant Physiol* 2001; 125:1732-42.
111. Wright KE, Prior F, Sardana R, Altosaar I, Dudani AK, Ganz PR, Tackaberry ES. Sorting of glycoprotein B from the human cytomegalovirus to protein storage vesicles of transgenic tobacco. *Trans Res* 2001; 10:177-81.
112. Zheng Z, Sumi K, Tanaka K, Murai N. The bean seed storage protein  $\beta$ -phaseolin is synthesized, processed and accumulated in the vacuolar type-II protein bodies of transgenic rice endosperm. *Plant Physiol* 1995; 109:777-86.
113. Chikwamba RK, Scott MP, Mejia LB, Mason HS, Wang K. Localization of a bacterial protein in starch granules of transgenic maize kernels. *Proc Natl Acad Sci USA* 2003; 100:11127-32.
114. Ma JK-C, Hikmat BY, Wycoff K, Vine ND, Chargelegue D, Yu L, Hein MB, Lehner T. Characterization of a recombinant plant monoclonal secretory antibody and preventive immunotherapy in humans. *Nature Med* 1998; 4:601-6.
115. Frigerio L, Vine ND, Pedrazzini E, Hein MB, Wang F, Ma JK-C, Vitale A. Assembly, secretion and vacuolar delivery of a hybrid immunoglobulin in plants. *Plant Physiol* 2000; 123:1483-93.
116. Ma JK-C, Hiatt A, Hein MB, Vine ND, Wang F, Stabila P, van Dolleweerd C, Mostov K, Lehner T. Generation and assembly of secretory antibodies in plants. *Science* 1995; 268:716-9.
117. Frigerio L, de Virgilio M, Prada A, Faoro F, Vitale A. Sorting of phaseolin to the vacuole is saturable and requires a short C-terminal peptide. *Plant Cell* 1998; 10:1031-42.
118. Carolino SMB, Vaez JR, Irsigler AST, Valente MAS, Rodrigues LAZ, Fontes EPB. Plant BiP gene family: differential expression, stress induction and protective role against physiological stresses. *Braz J Plant Physiol* 2003; 15:59-66.
119. Pimpl P, Taylor JP, Snowden C, Hillmer S, Robinson DG, Denecke J. Golgi-mediated vacuolar sorting of the endoplasmic reticulum chaperone BiP may play an active role in quality control within the secretory pathway. *Plant Cell* 2006; 18:198-211.
120. Nicholson L, González-Melendi P, van Dolleweerd C, Tuck H, Perrin Y, Ma JK-C, Fischer R, Christou P, Stöger E. A recombinant multimeric immunoglobulin expressed in rice shows assembly-dependent subcellular localization in endosperm cells. *Plant Biotechnol J* 2005; 3:115-28.
121. Faye L, Boulaflos A, Benchabane M, Gomord V, Michaud D. Protein modifications in the plant secretory pathway: current status and practical implications in molecular pharming. *Vaccine* 2005; 23:1770-8.
122. Nuttall J, Ma JKC, Frigerio L. A functional antibody lacking N-linked glycans is efficiently folded, assembled and secreted by tobacco mesophyll protoplasts. *Plant Biotechnol J* 2005; 3:497-504.