

---

# UTILIZZAZIONE DI PIANTE TRANSGENICHE COME BIOFABBRICHE PER LA PRODUZIONE DI VACCINI ANTIVIRALI PER LA SALUTE UMANA

---

**Enrico Ferraiolo**

Dottorato in Scienze Biotecnologiche – XIX ciclo  
Indirizzo Biotecnologie Vegetali  
Università di Napoli Federico II





Dottorato in Scienze Biotechologiche – XIX ciclo  
Indirizzo Biotecnologie Vegetali  
Università di Napoli Federico II



---

# **UTILIZZAZIONE DI PIANTE TRANSGENICHE COME BIOFABBRICHE PER LA PRODUZIONE DI VACCINI ANTIVIRALI PER LA SALUTE UMANA**

---

Dottorando: Enrico Ferraiolo

Relatore: Prof. Luigi Monti

Co - Relatore: Dott. Teodoro Cardi

Coordinatore: Prof. Gennaro Marino



*Alla mia famiglia*



## INDICE

<b>RIASSUNTO</b>	<b>pag. 1</b>
<b>SUMMARY</b>	<b>pag. 2</b>
<b>1. INTRODUZIONE</b>	<b>pag. 7</b>
1.1 Piante come biofabbriche	pag. 7
1.2 Trasformazione genetica	pag. 12
1.3 HIV-1 ed HPV16	pag. 15
1.4 Scopo e contenuto della tesi	pag. 18
<b>2. MATERIALI E METODI</b>	<b>pag. 20</b>
2.1 Materiale utilizzato	pag. 20
2.2 Gene <i>gus-int</i>	pag. 20
2.2.1 Costruzione dei vettori pEF60, pEF61, pEF62, pEF63, pEF64	pag. 20
2.3 Gene <i>gag</i>	pag. 23
2.3.1 Costruzione dei vettori pEF40, pEF41, pEF42, pEF43, pEF44	pag. 23
2.3.2 Costruzione del vettore pEF45	pag. 23
2.3.3 Costruzione del vettore pEF46	pag. 23
2.3.4 Costruzione dei vettori pEF47, pEF48, pEF49	pag. 24
2.3.5 Costruzione dei vettori pEF470, pEF480, pEF490	pag. 24
2.4 Gene <i>L1<sup>v</sup></i>	pag. 30
2.4.1 Costruzione dei vettori pEF50, pEF51, pEF52, pEF53, pEF54	pag. 30
2.4.2 Costruzione dei vettori pEF55, pEF56	pag. 30
2.5 Gene <i>L1<sup>pt</sup></i>	pag. 30
2.5.1 Costruzione dei vettori pEF57, pEF58	pag. 30
2.5.2 Costruzione dei vettori pEF570 e pEF580	pag. 30
2.6 Gene <i>L1<sup>h</sup></i>	pag. 31
2.6.1 Costruzione dei vettori pEF500 e pEF501	pag. 31
2.7 Ottenimento dei plasmidi ricombinanti	pag. 36
2.8 Agroinfiltrazione di <i>N. tabacum</i> e <i>N. benthamiana</i>	pag. 39
2.9 Verifica dell'attività biologica del gene <i>gus</i>	pag. 41
2.10 Estrazione di RNA totale	pag. 41
2.11 Estrazione di proteine solubili e insolubili	pag. 42
2.12 Analisi molecolari	pag. 42

<b>3. RISULTATI</b>	<b>pag. 45</b>
3.1 Vettori con il gene reporter <i>gus-int</i> e ottimizzazione della tecnica di agroinfiltrazione	<b>pag. 45</b>
3.2 Gene <i>gag</i> (HIV-1)	<b>pag. 49</b>
3.2.1 pr55 ( <i>gag</i> )	<b>pag. 49</b>
3.2.2 Subunità del gene <i>gag</i> ( <i>p17</i> , <i>p24</i> e $\Delta p17$ )	<b>pag. 55</b>
3.2.2.1 Espressione nel citosol	<b>pag. 55</b>
3.2.2.2 Espressione nel cloroplasto	<b>pag. 55</b>
3.3 Gene <i>L1</i> (HPV16)	<b>pag. 60</b>
3.3.1 Espressione dei geni <i>L1<sup>v</sup></i> ed <i>L1<sup>pt</sup></i>	<b>pag. 60</b>
3.3.2 Espressione del gene <i>L1<sup>h</sup></i>	<b>pag. 61</b>
<b>4. DISCUSSIONE</b>	<b>pag. 68</b>
4.1 Vettori costruiti	<b>pag. 68</b>
4.2 Agroinfiltrazione dei vettori costruiti	<b>pag. 70</b>
4.3 Gene <i>gag</i>	<b>pag. 71</b>
4.4 Gene <i>L1</i>	<b>pag. 74</b>
<b>5. CONCLUSIONI</b>	<b>pag. 77</b>
<b>6. BIBLIOGRAFIA</b>	<b>pag. 78</b>
<b>APPENDICE</b>	<b>pag. 85</b>

## RIASSUNTO

L'uso delle piante come sistema di produzione di molecole di interesse farmaceutico sta diventando una delle maggiori attività attualmente studiate nel campo della genetica e delle biotecnologie vegetali (Ma *et al.*, 2005). L'obiettivo di questo progetto di dottorato è stato l'espressione in piante di tabacco di due geni di virus umani HIV-1/*gag* e HPV16/*L1*, responsabili il primo della Sindrome da Immunodeficienza Acquisita e il secondo del cancro della cervice uterina. In entrambi i casi, tra i numerosi vaccini a tutt'oggi in fase di studio, quello più promettente sembra essere basato sulla produzione di "Virus Like Particles" (VLPs), ossia particelle simil-virali altamente immunogeniche. L'espressione di questi geni in pianta (*N. tabacum* e *N. benthamiana*) è avvenuta tramite la trasformazione transiente mediata da *A. tumefaciens* ("agroinfiltration"). Per la trasformazione dei geni d'interesse sono stati utilizzati cinque vettori commerciali ([www.impactvector.com](http://www.impactvector.com)) che presentano il promotore, la 5'UTR ed il terminatore del gene *rbcS* di crisantemo, e due "tag" (c-myc e 6His) al COOH terminale. Questi vettori si differenziano per la presenza o meno di segnali di indirizzamento della proteina in cinque compartimenti subcellulari diversi (citoplasma, apoplasto, reticolo endoplasmico, cloroplasto e mitocondrio). Il gene *gag* è stato indirizzato in tutti i compartimenti, mentre le subunità di cui è composto (*p17*, *p24* e  $\Delta p17$ ) solo nel citoplasma e nel cloroplasto. Il gene  $L1^v$  (con la sequenza virale "wildtype"), è stato clonato negli stessi vettori ed indirizzato in tutti e cinque i compartimenti. Per esprimere una proteina più corta e potenzialmente più stabile, è stata fatta una delezione degli ultimi 205 aminoacidi dal COOH terminale, che è stata indirizzata solo nel citoplasma e nel cloroplasto. Essendo disponibili due sequenze sintetiche del gene *L1*, una ottimizzata per l'espressione plastidiale ( $L1^{pt}$ ) e l'altra per l'espressione in cellule umane ( $L1^h$ ), sono stati costruiti quattro vettori dove le due sequenze sono state indirizzate nel citoplasma e nel cloroplasto. Inoltre, anche per il gene  $L1^{pt}$  è stata fatta la delezione degli ultimi 205 aminoacidi come per il gene  $L1^v$ . Le analisi molecolari sui tessuti agroinfiltrati con tutti i vettori sono state effettuate 3 giorni dopo l'infiltrazione.

L'analisi RT-PCR effettuata per il gene *gag* e per le subunità hanno generalmente indicato la presenza di un messaggero stabile e di uguale intensità, indipendentemente dal vettore e dalla specie utilizzata. Le analisi western, invece, hanno mostrato la presenza di due bande di 40 e 24 KDa relative alla proteina Gag solo nei cloroplasti di *N. benthamiana*. La presenza di due bande ha indicato un processamento della proteina Gag (55 KDa) nei tessuti vegetali. Le analisi western condotte sui tessuti di *N. tabacum* e *N. benthamiana*, agroinfiltrate con le subunità P17, P24 e  $\Delta P17$  del gene *gag*, hanno evidenziato su *N. benthamiana*, la presenza di una banda di circa 24 KDa, sia con i vettori per l'espressione nel citoplasma che nel cloroplasto. Al contrario, la proteina P17 è stata osservata solo quando è stata indirizzata nei cloroplasti. Risultati simili si sono avuti anche in *N. tabacum*, anche se l'espressione è risultata essere molto più bassa.

Per il gene  $L1^v$  ed  $L1^{pt}$ , nonostante la presenza di un messaggero stabile, in nessun compartimento subcellulare è stato possibile evidenziare la proteina di interesse, sia in *N. tabacum* che in *N. benthamiana*. Solo utilizzando il gene  $L1^h$ , è stata evidenziata in *N. benthamiana* una debole banda del peso molecolare atteso per la proteina.

## SUMMARY

During the past decade, many changes have taken place in plant molecular biology. The improvement of techniques for plant genetic analysis has permitted the production of heterologous proteins in plants. Nowadays, different production systems are used to produce recombinant proteins such as bacteria, yeast and mammalian cells.

The use of microorganisms permits the manufacture on large scale. The bacteria system is inexpensive, easily to propagate, but prokaryotic cells are incapable of making most post-translational modifications necessary for many mammalian proteins, such as glycosylation and others. The eukaryotic proteins expressed in *E. coli* frequently acquire an unnatural conformation accompanied sometimes by intracellular precipitation. Using yeast as production system there are same advantages, such as high level of expression and correct protein folding, but also incorrect post-translational modifications. Mammalian cell cultures synthesize correctly eukaryotic proteins. They have a correct machinery for post-translational modifications and the protein product yield is high. However, the process to obtain the recombinant protein is expensive and the time to obtain them is long. The mammalian cultures are sensitive to temperature and pH; to have the final product, it is necessary to purify the protein (Fischer *et al.*, 2001).

Plants have many advantages to express recombinant proteins. First, plant systems are more economical than bioreactor systems used in industries. The technology is already available for processing and harvesting plant products on large scale. The cost to grow a plant is low and the storage is very cheap. Purification cost of heterologous proteins can be lowered, and it can be even eliminated when the plant tissue containing the recombinant protein is used as food, such as tomato, potato and rice. Plants have the correct machinery for post-translational modifications (glycosylation, disulfide bridge, correct folding, etc). Recombinant proteins can be easily targeted to subcellular compartments, such as apoplast, endoplasmic reticulum, chloroplast and mitochondria, where the proteins can be more stable. Also, the health risks from contaminating human pathogens or toxins are minimized, because plants are not hosts for human pathogens. Further, the storage and transports of proteins is cheaper, because the plants can be stored without the "cold chain", that is necessary for the proteins produced in other systems (Daniell *et al.*, 2001).

The first recombinant plant-derived pharmaceutical protein (PDP) was the human serum albumin, initially produced in 1990 in transgenic tobacco and potato plants (Sijmons *et al.*, 1990). Based on this result, the proof of concept was established for the production of many therapeutic proteins, including antibodies, blood products, growth factors, hormones, enzymes, and human and veterinary vaccines (Ma *et al.*, 2006).

Many papers reported the expression of antigenic proteins in a plant system to obtain a low cost vaccine (Haq *et al.*, 1995; Mason *et al.*, 1992; Obregon *et al.*, 2006; Richter *et al.*, 2000; Tackaberry *et al.*, 1999; Tuboly *et al.*, 2000). Several PDP products for the treatment of human diseases are approaching commercialization, including recombinant gastric lipase for the treatment of cystic fibrosis, and antibodies for the prevention of dental caries.

The objective of this project was the expression of two different antigens of human viruses, HIV-1/*gag* and HPV16/*L1* in tobacco plants (*N. tabacum* and *N.*

*benthamiana*). The HIV-1/gag causes the AIDS disease, and HPV16 is the virus responsible for the cervical cancer in women. The *gag* gene is composed of four distinct domains, the N-terminal matrix domain (p17), the central capsid domain (p24), the nucleocapsid domain (p7) and the C-terminal domain (p6). Two spacer peptides, p1 and p2, separate the p24/p7 and p7/p6 domains, respectively (Morikawa 2003). *L1* gene codes for the major capsid protein of HPV virus. Both genes are responsible for the auto-assembling and formation of Virus Like Particles (VLPs) that are a good candidate to produce an effective vaccine. VLPs mimic virus and the infection stimulating the production of antibodies for humoral and cellular immune response. At this moment, there are no papers that report the expression of Gag protein in a plant system, but some articles reported the expression of L1 protein in tobacco and potato plants, albeit at low levels.

To obtain the expression of these proteins in plant cells, it has been necessary to construct several vectors. The *gag* gene has been inserted into five different plasmids (*ImpactVector*, [www.impactvector.com](http://www.impactvector.com)) under the control of the promoter, 5'UTR and terminator of the *rbcS* gene, and two different tags (c-myc and 6His) at COOH terminus. These vectors have signal peptides for protein accumulation in different subcellular compartments such as cytosol, apoplast, endoplasmic reticulum, chloroplast and mitochondrion. In order to test a different promoter and 5'UTR, the *gag* gene has been cloned also in two different binary vectors with 35S<sup>2</sup> promoter and  $\Omega$  sequence of tobacco mosaic virus as 5'UTR, with or without the signal peptide to accumulate the protein in the chloroplasts.

To check whether the low expression of *gag* was due to a particular subunit of the same gene, in this work, sequences encoding the main subunits have been cloned in the same vectors. In particular, *p17*, *p24* and  $\Delta p17$  (*gag* gene without the p17 domain) subunits were cloned into plant expression vectors directing protein accumulation in the cytosol or in the chloroplasts.

The *L1<sup>v</sup>* (wildtype viral sequence) gene was cloned in the same vectors used for the *gag* gene, without the last 22 aminoacids, which code for the Nuclear Localization Signal (NLS). In order to facilitate the accumulation of the protein, it has been also used a truncated *L1* sequence, that lacks the last 205 aminoacids. This sequence has been targeted into the cytosol and in the chloroplasts. Furthermore, two optimized sequences of the *L1* gene, the first with a codon usage for plastid expression (*L1<sup>pt</sup>*) and the second with a sequence optimized for expression in human cells (*L1<sup>h</sup>*), both without the NLS, have been cloned in the vectors for protein expression in the cytosol and chloroplasts. The *L1<sup>pt</sup>* gene has been also cloned in the short version (devoid of the last 205 aminoacids) and targeted into cytosol and chloroplasts.

All vectors were tested in transient assay by agroinfiltration technique. To optimize the agroinfiltration procedure and check that the constructs worked correctly, a reporter *gus-int* gene was also cloned in the same five expression vectors.

Transient gene expression has several advantages over the generation of stably transgenic plants. It is rapid and results in protein expression in a few days (Kapila *et al.*, 1997). The transient expression of gene constructs may be desirable in some cases in which long-term expression is not required. This approach has been used to test the effectiveness of various gene constructs prior to stable transformation. Using the transient expression system it is also possible to express two or more genes contemporarily, and it is not required the regeneration of plant tissues.

In this thesis, the agroinfiltration was obtained using the protocol reported by Voinnet *et al.* (2003), and the agroinfiltrated tissue was harvested 3 days post infiltration.

The presence of the *gus* gene mRNA was obtained using RT-PCR. The results showed that the *gus* gene was correctly transcribed both in *N. tabacum* and *N. benthamiana*, at the same level. The presence of Gus protein was checked using anti - his or anti - c-myc antibodies. The presence of the protein was observed in all subcellular compartments albeit at different levels. Protein accumulation was higher in the cytosol and chloroplast compartments than in the apoplast, endoplasmic reticulum and mitochondria.

Istochemical assays, to check the functionality of Gus protein, were carried out directly on the agroinfiltrated leaves. The results showed the blue colour on the leaf indicating that the protein folding was correct and the enzyme worked correctly.

Agroinfiltration with all vectors carrying the *gag* gene or its subunits were carried out both in *N. tabacum* and in *N. benthamiana*. Using RT-PCR the presence of the transcript was observed in all agroinfiltrated tissues at the same level and with all constructs. In *N. tabacum* no Gag protein was observed in any subcellular compartment, but in *N. benthamiana* the protein was accumulated in chloroplasts. Using anti-p24 antibodies, the results showed two bands of about 40 and 24 KDa indicating the rapid processing of the Gag polypeptide by plastidial proteases, resembling what happens in human cells infected by HIV virus. The 40 KDa band was detected also using anti - p17 antibodies, indicating that the 40 KDa is composed by p17 and p24 domains.

The experiments with vectors containing the 35S promoter were carried out only in *N. tabacum*. The results did not show the protein in cytosol and in chloroplasts, although the mRNA was observed with both constructs. These results indicate a loose relationship between the presence of the mRNA and protein. Similar results have been reported by Wydro *et al.* (2004) and James *et al.* (2000).

The *p17*, *p24* and  $\Delta p17$  subunits were targeted to the cytosol and chloroplasts. The presence of the mRNA was detected by RT-PCR using specific primers and the results showed the transcripts at the same level in both species with the construct for cytosol expression or with the construct for chloroplast expression. The presence of the protein was checked using anti - p24 polyclonal antibodies to detect p24 and  $\Delta p17$  proteins, and anti - p17 antibodies to detect the p17 protein. Western blot analysis on agroinfiltrated tissue with p24 construct, cytoplasmic version, showed two bands of about 22 and 25 KDa. To understand the composition of these bands, another western blot was made on the same extract using anti - his and anti - c-myc antibodies. The results showed only the upper band of 25 KDa, that likely corresponds to the p24 protein with tag c-myc and 6His. When the p24 protein was targeted into chloroplasts, only the 24 KDa was observed in *N. benthamiana* plants. Agroinfiltration with the  $\Delta p17$  construct showed a 24 KDa band in the cytosol and in the chloroplasts, but in the chloroplasts the level of expression was higher than in the cytosol. This result indicate that the protein was cleaved after the p24 domain, as already observed with the Gag protein.

Results were similar in both species, but in agroinfiltrated *N. tabacum* plants signals were much weaker, suggesting a reduced protein synthesis and/or an increased protease activity in this species.

The presence of the p17 protein was detected only when it was accumulated in the *N. benthamiana* chloroplasts, showing two predominant bands of about 13 and 20 KDa. Using the anti - c-myc antibody only the 20 KDa band was observed, corresponding to the p17 protein plus the tag sequence (c-myc and 6His).

No papers have reported the expression of the Gag protein in plants so far, but an abstract of Meyers *et al.* (2005) reported the expression of Gag protein mediated by agroinfiltration or using recombinant TMV vector. The authors reported low levels of *Nicotiana* codon-optimised Gag protein in leaf extracts of *N. benthamiana* inoculated with recombinant TMV vector, but higher levels were measured after agroinfiltration. Meyer *et al.* (2005) conclude that agroinfiltration is better than viral vector to express the Gag protein in plant cells.

On the contrary, there are some articles reporting the expression of the p24 protein in plants, using a stable transformation by *A. tumefaciens* or transient expression mediated viral TMV vectors. In both cases, the presence of the p24 protein was observed in plant cells, at level of 0,35% or 0,1% of TSP, respectively (Zhang *et al.*, 2002; Perez-Filgueira *et al.*, 2004).

The agroinfiltrated tissue with L1 constructs were analyzed by RT-PCR and western blot analyses. Western blot analyses using all tested vectors with *L1<sup>v</sup>* gene did not show the protein in any subcellular compartment. To check if the gene was transcribed correctly, RT-PCR analyses were carried out on agroinfiltrated tissues with *L1<sup>v</sup>*, *L1<sup>v</sup>Δ205*, *L1<sup>pt</sup>* and *L1<sup>pt</sup>Δ205* sequences targeted into cytosol or chloroplasts. The results showed the presence of the mRNA, with all constructs in *N. tabacum* and *N. benthamiana* plants and their expression levels were comparable in the cytosol and chloroplasts. Again, the proteins could not be detected in Western blot experiments using anti - L1 or anti - c-myc antibodies. Despite the use of a synthetic optimized *L1* gene, it has not been possible to show the protein in plants cells. In literature, there are only five articles that reported the expression of L1 gene in a plant system (Biemelt *et al.*, 2003; Liu *et al.*, 2005; Varsani *et al.*, 2003; Varsani *et al.*, 2006; Warzecha *et al.*, 2003). All authors describe the difficulties to express this gene and reported the accumulation of the L1 protein only using a human optimized *L1* gene.

In this work, agroinfiltration of the *L1<sup>h</sup>* gene has been tested in *N. benthamiana* plants. The results of RT-PCR showed the presence of a strong transcript both in cytosol and chloroplasts, indicating that the transgene is transcribed at high level. Western Blot analysis, using anti - c-myc antibodies, reported, in the soluble protein fraction, only a weak band of 55 KDa both in cytosol and chloroplasts.

McLean *et al.* (2005) also used a human and plant optimized *L1* gene, and targeted the protein to the plastids or endoplasmic reticulum by a transit peptide. They demonstrated that a relatively high expression level was obtained only using the human *L1* gene. Furthermore, when the protein was targeted to chloroplasts, the accumulation level was better than expression in the cytosol and endoplasmic reticulum. So far, the best results in the literature have been achieved by Varsani *et al.* (2006) using a human *L1* gene and viral expression vector as transformation system. The expression level of L1 protein (about 20-37

ug/Kg of fresh infected tissue) has been reported by these authors. Santi *et al.* (2005) and Smidkova *et al.* (2005) reported the expression of the *L1* gene in *N. benthamiana* and tomato. Expression levels equal to 0,5 – 1% TSP were obtained. In the case of Smidkova *et al.* (2005), a human optimized sequence gave better results than a plant optimized sequence. Both authors conclude that viral vectors, as expression system, are somehow better than stable transformation to express this gene.

## 1 INTRODUZIONE

### 1.1 Piante come biofabbriche

Le piante ed i prodotti di origine vegetale sono da sempre utilizzati dall'uomo per la cura di malattie. Più di un quarto delle medicine utilizzate contengono, almeno in parte, prodotti di origine vegetale (Raskin *et al.*, 2002). Negli ultimi anni, lo sviluppo della genetica vegetale e la scoperta di tecniche sempre più efficienti per la trasformazione delle piante, hanno permesso di utilizzare le piante non solo per la produzione di molecole di origine vegetale, ma anche come sistema di produzione di molecole esogene di interesse industriale.

Fino ad oggi, tra i sistemi più utilizzati per la produzione di molecole di interesse farmaceutico, vi sono i microorganismi, come *Escherichia coli*, che è largamente utilizzato nelle industrie per la produzione di proteine su larga scala, in quanto è un sistema economico e flessibile. Il sistema batterico però ha lo svantaggio di non poter esprimere proteine che necessitano di modificazioni post-traduzionali come la glicosilazione o la formazione di ponti disolfuro, che in alcuni casi sono necessari per il corretto funzionamento delle proteine.

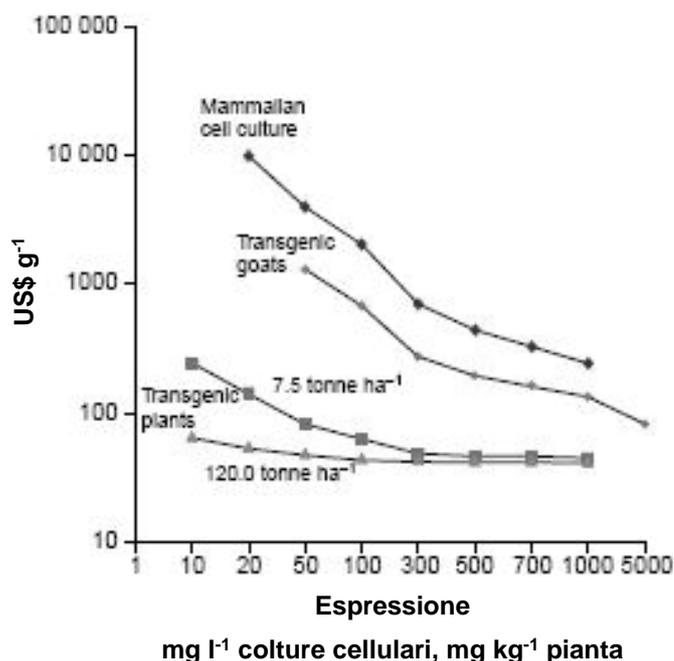
Nelle industrie è molto diffuso anche l'uso delle cellule di lievito per l'espressione di proteine ricombinanti. I vantaggi di questo sistema sono dati dalla facile propagazione delle colture cellulari, dalla grande quantità di proteina ricombinante che si riesce ad ottenere e dal corretto assemblamento delle proteine espresse. Il maggior problema risiede nell'espressione delle proteine che necessitano di glicosilazione, che in cellule di lievito non avviene correttamente.

Per superare questo problema esiste la possibilità di utilizzare cellule di coltura di mammifero. Infatti, queste colture cellulari hanno il vantaggio di assemblare correttamente le proteine ma lo svantaggio di avere sia un costo elevato e quindi l'impossibilità di avviare una produzione su larga scala, sia di possibili contaminazioni con patogeni che potrebbero risultare dannosi per l'uomo (Daniell *et al.*, 2001).

Il sistema vegetale rappresenta una valida alternativa per l'espressione di molecole di interesse industriale. Infatti, l'uso delle piante per l'espressione di proteine eterologhe garantisce la sintesi di molecole proteiche, con caratteristiche strutturali e funzionali assimilabili a quelle delle proteine native, ivi comprese le modificazioni post-traduzionali tipiche delle proteine eucariotiche. I vantaggi che il sistema vegetale offre sono numerosi, tra cui, i bassi costi di produzione, l'assenza o i bassi costi di purificazione, la possibilità di accumulo delle proteine in diversi compartimenti subcellulari e l'assenza di patogeni in comune con l'uomo. In particolare, i bassi costi di produzione sono possibili grazie all'assenza di fermentatori e di substrati per la crescita dei batteri. Autori come Kusnadi *et al.* (1997) hanno stimato che i costi di produzione di proteine ricombinanti in pianta potrebbero essere da 10 a 50 volte più bassi rispetto al costo delle stesse proteine espresse in *E. coli*. In Fig. 1 è riportato il costo per grammo di anticorpi IgA prodotti in diversi sistemi cellulari. Come si evince dalla figura, c'è un notevole abbattimento dei costi di produzione utilizzando il sistema vegetale come sistema di produzione.

Altro vantaggio è dato dalla possibilità di una produzione su larga scala in quanto è già disponibile la tecnica di raccolta e processamento delle proteine di origine vegetale, e inoltre l'aumento della produzione può essere ottenuto

soltanto aumentando il terreno di coltivazione e facendo crescere un numero di piante sufficiente alla richiesta di mercato (Ma *et al.*, 2003).



**Fig. 1: Costi per grammo di anticorpi IgA prodotti in vari sistemi di espressione**  
**Modificata da Daniell *et al.* (2001)**

La purificazione delle proteine prodotte in pianta può essere costosa, ma i costi possono essere notevolmente abbattuti se le proteine sono fatte accumulare nell'endosperma dei semi dal quale le proteine possono essere facilmente estratte (Wright *et al.*, 2001). Inoltre, il costo della purificazione può essere completamente abbattuto se la proteina d'interesse viene espressa in piante eduli; in questo caso la proteina non deve essere più estratta dai tessuti vegetali, ma può essere somministrata direttamente dopo semplici processi di concentrazione e dosaggio.

Altro vantaggio risiede nella possibilità di far accumulare le proteine in differenti compartimenti subcellulari, quali il cloroplasto o il reticolo endoplasmico, con la possibilità di ottenere un maggior accumulo e stabilità della proteina. Inoltre le piante possiedono il vantaggio della facilità di trasporto delle piante esprimenti proteine ricombinanti nei paesi in via di sviluppo, in quanto non sono necessarie le cosiddette "catene del freddo" indispensabili invece per il trasporto e la conservazione delle proteine prodotte con i sistemi tradizionali. Infine, l'assenza di patogeni in comune con l'uomo abbate i rischi di possibili contaminazioni da parte di patogeni umani (Daniell *et al.*, 2001).

In Tabella 1 è riportato un confronto tra i vari sistemi di espressione maggiormente utilizzati. Tutti questi vantaggi si traducono anche in un nuovo valore economico delle piante di interesse agrario e, di conseguenza, un aumento potenziale del consumo e della loro redditività (Richter, 1999).

Nel 1986, l'ormone della crescita umano è stato la prima proteina ad essere stata espressa in piante di tabacco (Barta *et al.*, 1986). Subito dopo, nel 1990, Sijmons *et al.*, (1990) hanno prodotto con successo l'albumina umana in piante transgeniche di tabacco e di patata. Nel corso degli anni sempre più numerosi sono stati i lavori pubblicati sull'espressione di proteine di interesse farmaceutico in pianta. Di grande interesse è stata la produzione in pianta dell'ormone della somatotropina, espresso nei plastidi di tabacco (Staub *et al.*, 2000) e dell'adenosina deamminasi (ADA) in mais (Petolino *et al.*, 2000). La produzione di collagene umano in tabacco è stata riportata dal gruppo di Ruggiero *et al.*, (2000) e grazie alla disponibilità di efficienti protocolli di estrazione di questo polimero dalla pianta si sono ridotti notevolmente i costi di produzione su larga scala di questa molecola utilizzata dalle industrie cosmetiche. In Tabella 2 sono riportate alcuni esempi di proteine di interesse farmaceutico espresse con un sistema vegetale.

**Tabella 1:** Confronto tra vari sistemi di espressione per proteine ricombinanti

	Piante transgeniche	Lievito	Batteri	Cellule mammifero
Costo / Conservazione	Economico / Temperatura ambiente	Economico / - 20°C	Economico / - 20°C	Costoso / N <sub>2</sub>
Distribuzione	Facile	Flessibile	Flessibile	Difficile
Dimensioni del gene	Non limitata	ND	ND	Limitata
Glicosilazione	Corretta?	Non corretta	Assente	Corretta
Assemblamento proteine multimeriche	Possibile	Impossibile	Impossibile	Impossibile
Costo di produzione	Basso	Medio	Basso	Alto
Produzione su larga scala	Mondiale	Limitata	Limitata	Limitata
Accuratezza "folding" delle proteine	Alto?	Medio	Basso	Alto
Livelli di espressione	Variabile	Alta	Media	Alta
Percezione pubblica del rischio <sup>¶</sup>	Alta	Media	Bassa	Media
Sicurezza	Alta	Alta	Bassa	Media
Rischi terapeutici	Sconosciuti	Sconosciuti	Presenti	Presenti
Tempo richiesto	Medio	Medio	Basso	Alto

Modificata da Fischer *et al.* (2001). ND= non determinata. ? = non confermato  
<sup>¶</sup> = residui di peptidi virali, endotossine.

**Tabella 2:** Produzione di molecole non vegetali in piante transgeniche

Proteina	Pianta usata	Applicazioni
Proteina C umana	Tabacco	Anticoagulante
Eritropoietina	Tabacco	Anemia
Interferone $\alpha$	Riso	Trattamento epatite B e C
Emoglobina	Tabacco	Sangue artificiale
Collagene I omotrimerico	Tabacco	Collagene
$\alpha$ -1 antitripsina	Riso	Peptidi inibitori
Apoproteina	Mais	Trapianti chirurgici
Glucocerebrosidasi	Tabacco	Malattia di Gaucher's
Fattore epidermide di crescita	Tabacco	Controllo e riparo della proliferazione cellulare
Lattoferrina	Patata	Antimicrobici
Eritropoietina	Tabacco	Anemia
Albumina umana	Tabacco	Cirrosi epatica
Angiotensina	Tabacco e pomodoro	Ipertensione

Modificata da Daniell *et al.* (2001)

Il primo approccio dell'espressione in pianta di un antigene utile per un vaccino umano risale al lavoro di Mason *et al.* (1992) sull'espressione dell'antigene di superficie del virus dell'epatite B in tabacco, ma con bassi livelli di espressione. Altri autori hanno riportato l'espressione in pianta dell'enterotossina di *E. coli* in piante di patata (Mason *et al.*, 1998; Tacket *et al.*, 1998) confermando l'immunogenicità e protezione quando somministrata oralmente. Anche l'antigene della proteina capsidica del virus Norwalk è stata espressa dal gruppo di Mason *et al.* (1996) in piante transgeniche di tabacco e dal gruppo di Tacket *et al.* (2000) in piante transgeniche di patata con livelli di espressione di circa lo 0,23% e lo 0,37% rispettivamente. In *Nicotiana benthamiana* Giddings *et al.* (2000) hanno espresso, con livelli di circa il 2% delle PTS la molecola  $\alpha$ -tricosantina, che sembra essere una buona candidata nella terapia contro l'HIV.

In Tabella 3 è riportato un elenco di alcune molecole antigeniche espresse in pianta. Alcune di queste molecole sono sul mercato (Hood *et al.*, 1997), mentre altre sono in fase di sperimentazione e pronte per la commercializzazione

(Ma *et al.*, 2003). E' il caso dell'antigene dello streptococco che causa le carie dentarie, che è stato espresso in tabacco fino a livelli di 500 µg per grammo di foglie fresche (Ma *et al.*, 1998) ed ora è in Fase II di sperimentazione. Anche altre proteine prodotte in un sistema vegetale come la lipasi gastrica per la cura della fibrosi cistica e anticorpi per il trattamento del linfoma non-Hodgkin sono in fase avanzata di sperimentazione. Molte industrie private hanno ottenuto le licenze ed i fondi necessari per avviare la produzione su larga scala delle proteine approvate per l'immissione in commercio. Esempio rilevante è dato dagli accordi stipulati tra la "Large Scale Biology Corp." ([www.lsbcc.com](http://www.lsbcc.com)) e la "Sigma-AldrichCo." ([www.sigmaaldrich.com](http://www.sigmaaldrich.com)) per la produzione dell'interferone  $\alpha$ -2a ed  $\alpha$ -2b in pianta.

**Tabella 3:** Proteine antigeniche prodotte in pianta.

Proteine o peptidi espressi	Origine della proteina	Sistema di espressione vegetale	Espressione PTS (%)
Subunità B	<i>Escherichia coli</i> enterotossico	Tabacco	0,01
Subunità B	<i>Escherichia coli</i> enterotossico	Patata	0,19
Subunità B	<i>Escherichia coli</i> enterotossico	Mais	ND
Subunità B	<i>Vibrio cholerae</i>	Patata	0,30
Proteina di rivestimento	Virus dell'epatite B	Tabacco	0,01
Proteina di rivestimento	Virus dell'epatite B	Patata	0,01
Proteina di rivestimento	Virus dell'epatite B	Lattuga	0,01
Proteina capsidica	Virus di Norwalk	Tabacco	0,23
Proteina capsidica	Virus di Norwalk	Patata	0,37
Glicoproteina	Virus della rabbia	Pomodoro	1,00
Anticorpo	<i>Streptococcus mutans</i>	Tabacco	500 ug / gr di foglie
Peptide terapeutico	Cancro dell'intestino	Mais	ND
Anticorpo	Non-Hodgking	Tabacco	ND

Modificata da Daniell *et al.* (2001).

PTS= Proteine totali solubili; ND = non determinata

Numerose specie vegetali sono state trasformate con successo fino ad oggi. Il tabacco rappresenta la specie su cui si stanno mettendo a punto molte delle tecnologie d'interesse, in quanto ben si presta a tutte le tecniche di coltura in *vitro*, ha una grande biomassa verde e produce una grande quantità di semi.

Inoltre, il tabacco non essendo una pianta edule, non presenta il rischio di contaminazione delle catene alimentari. Questo uso alternativo del tabacco ha permesso di rivalutare l'uso di questa specie, non più per la produzione dei prodotti da fumo, ma per la produzione di molecole utili per la salute umana. In questa specie sono state espresse numerose proteine a scopi terapeutici (Haq *et al.*, 1995; Mason *et al.*, 1992; Obregon *et al.*, 2006; Richter *et al.*, 2000; Tackaberry *et al.*, 1999; Tuboly *et al.*, 2000). Altre specie sono state utilizzate con successo per la produzione di sostanze utili, come la lattuga (Kapusta *et al.*, 1999), la soia (Zeitlin *et al.*, 1998), il riso (Giddings *et al.*, 2000), la patata (Chong e Langridge, 2000), arapidopsis (Kusnadi *et al.*, 1997), gli spinaci (Yusibov *et al.*, 2002), il pomodoro (McGarvey *et al.*, 1995) e mais (Streatfield *et al.*, 2001).

Utilizzando piante eduli si elimina la fase di estrazione della proteina dai tessuti vegetali, con il conseguente vantaggio che la proteina potrebbe essere somministrata oralmente dopo semplici processi di concentrazione e dosaggio (Daniell *et al.*, 2001). Il riso, oltre a presentare questi vantaggi, riveste grande interesse nei paesi in via di sviluppo nei quali è molto utilizzato a scopo alimentare. Attraverso l'uso di sequenze specifiche, si possono esprimere e far accumulare nei semi le proteine d'interesse (Cramer *et al.*, 1999). I vantaggi dell'espressione nei semi consiste nell'assenza di composti fenolici presenti invece nei tessuti fogliari, nella possibilità di avere una grande quantità di proteina in poco volume e nella estrema facilità di trasportare i semi senza bisogno di refrigerazione, abbattendo l'uso di costose "catene del freddo".

## 1.2 Trasformazione genetica

Nel corso degli ultimi venti anni, numerosi lavori ed applicazioni pratiche hanno dimostrato che la trasformazione delle piante è una tecnica riproducibile, che permette di poter esprimere in pianta numerose proteine di origine animale da utilizzare in campo biomedico.

Tra le specie trasformate con successo sono incluse quelle più importanti dal punto di vista economico, quali cereali, ortaggi, piante medicinali, pomodoro patata e tabacco (Birch *et al.*, 1997). Il trasferimento di uno o più geni in cellule vegetali può essere effettuato mediante metodi diretti (chimici o fisici) o indiretti (attraverso l'uso di vettori batterici o virali). Esso può essere stabile, se il gene d'interesse viene integrato stabilmente nel genoma vegetale, oppure transiente se il gene non viene stabilmente trasferito nel DNA dell'ospite e la sua espressione è limitata nel tempo. Numerosi sono i fattori da cui dipende il successo della trasformazione genetica delle piante, tra cui un efficiente protocollo di trasformazione e la capacità della specie di rigenerare in *vitro*. Anche se ormai sono disponibili numerosi protocolli di trasformazione per diverse specie vegetali, il tabacco rappresenta la specie modello che più è utilizzata per l'espressione di proteine ricombinanti per i motivi descritti nel § 1.1.

La metodologia più utilizzata si basa sull'impiego del batterio *Agrobacterium tumefaciens* (Hooykaas e Mozo, 1994). Questo batterio infetta normalmente piante dicotiledoni tra cui piante da frutto, pomodoro, soia, patata e molte altre (Lessard *et al.*, 2002), provocando la formazione di un tumore noto come "tumore del colletto". La formazione tumorale è dovuta al trasferimento da parte del batterio di una regione di DNA chiamata T-DNA, dal plamide Ti ("tumor inducing") al genoma nucleare della pianta. Il T-DNA contiene geni che regolano la sintesi di enzimi implicati nella divisione delle cellule vegetali e alcuni geni che producono derivati aminoacidici, chiamati opine, che diventano la fonte di azoto e

carbonio del batterio presente all'interno delle cellule infettate. Queste sequenze geniche nel T-DNA sono fiancheggiate da due regioni ad elevata omologia, il "left border" (LB) ed il "right border" (RB) lunghe circa 25 coppie di basi che sono responsabili dell'integrazione del T-DNA nel genoma nucleare. Esperimenti di ingegneria genetica hanno permesso di modificare il T-DNA, rimuovendo i geni che inducono il tumore e inserendo al loro posto un sito multiplo di clonaggio. Siccome il plasmide Ti è molto grande (200 Kb) ed è molto difficile modificarlo, sono stati sviluppati i cosiddetti vettori binari (Bevan *et al.*, 1984). Con questa strategia nell'agrobatterio sono presenti due vettori, uno con la regione T-DNA disarmata ed i geni *vir*, necessari per indurre il trasferimento, intatti; ed un altro con le regioni LB ed RB, la sequenza d'interesse e un'origine di replicazione riconosciuta sia in *E. coli* che in *A. tumefaciens*.

Per trasformare una pianta con *A. tumefaciens*, è quindi necessario creare un vettore binario contenente, tra le sequenze LB ed RB, una regione promotrice, la sequenza genica d'interesse e una sequenza terminatrice della trascrizione. Inoltre, è generalmente presente un gene marcatore completo di sequenze promotrici e di terminazione che consente di distinguere e far rigenerare solo le cellule trasformate. La trasformazione genetica consiste in una co-cultura degli espianti della pianta con l'agrobatterio contenente il plasmide d'interesse. Successivamente, i tessuti vegetali vengono fatti rigenerare su un substrato selettivo contenente sia l'antibiotico specifico sia gli ormoni che permettono lo sdifferenziamento e la successiva rigenerazione solo delle cellule trasformate.

Questa metodologia presenta dei notevoli vantaggi in termini di facilità d'impiego, economicità e di efficienza di trasformazione, tuttavia presenta dei limiti, in quanto possono essere trasferiti solo pochi geni contemporaneamente, possono verificarsi fenomeni di co-soppressione quando sono inserite più copie dello stesso gene, ed è necessario dover analizzare un elevato numero di piante, in quanto, essendo l'integrazione nel genoma casuale, l'espressione della proteina d'interesse varia tra tutte le piante ottenute.

Una tecnica di trasformazione stabile alternativa è rappresentata dal metodo biolistico. La tecnica si basa sull'uso di microproiettili di oro o di tungsteno rivestiti con il DNA plasmidico di interesse, che vengono accelerati contro una cellula. La velocità dei microproiettili è sufficientemente alta da superare, senza danneggiare, la parete e la membrana cellulare (Sanford, 1990). Il DNA all'interno della cellula si potrà integrare così nel genoma nucleare. Con questa tecnica è stato possibile trasferire più geni contemporaneamente nel nucleo della pianta. Questo è dimostrato in diversi lavori di Chen *et al.* (1998); Romano *et al.* (2003), dove è stato riportato l'inserimento di più geni contemporaneamente, rispettivamente nel genoma di riso e in quello di patata.

Un altro uso della tecnica biolistica è dato dalla possibilità di trasformare il genoma plastidiale. La trasformazione plastidiale è una tecnica descritta da molti autori in letteratura (Maliga, 2004; Daniell, 1993). Tuttavia, rispetto alla trasformazione nucleare, che è oggi un processo routinario, le caratteristiche strutturali e metaboliche dei plastidi comportano un aumento delle difficoltà in tutte le fasi di trasformazione (Heifetz, 2000). A tutt'oggi solo alcuni protocolli sono disponibili per un numero limitato di specie, tra le quali tabacco, patata, riso e pomodoro. I vantaggi che questo nuovo sistema offre sono un aumento del grado di espressione della proteina di interesse di circa 10 volte rispetto alla trasformazione nucleare, sequestro della proteina nell'organello, assenza di effetti di posizione in quanto l'integrazione nel genoma plastidiale è sito-specifica,

assenza di trasmissione indesiderata del transgene nel polline in quanto i plastidi sono ereditati per via materna. Quest'ultimo vantaggio è di notevole importanza in quanto riduce il problema dell'impatto ambientale delle piante transgeniche, favorendo il contenimento genico (Daniell *et al.*, 2002).

Tra i limiti della trasformazione plastidiale rientrano alcune modifiche post-traduzionali che non avvengono nei plastidi per la mancanza degli enzimi necessari, come la glicosilazione e la miristilazione. La mancanza di queste modificazioni comporta che proteine che necessitano di glicosilazione o di altre modifiche non potranno essere espresse correttamente nell'ambiente plastidiale.

Nell'ultimo decennio sono stati riportati in letteratura esempi di trasformazione plastidiale con geni di interesse biomedico e farmaceutico. Staub *et al.* (2000) hanno riportato la produzione della somatotropina umana in cloroplasti di tabacco con livelli di circa il 7% delle PTS, invece Tregoning *et al.*, (2003) hanno ottenuto, sempre nei cloroplasti di tabacco, l'espressione della tossina del tetano con livelli attorno al 25% delle PTS.

Sia la trasformazione mediata da *A. tumefaciens*, sia quella con il metodo biolistico, richiedono la rigenerazione dei tessuti vegetali *in vitro*, che comporta lunghi tempi di attesa che variano a seconda della specie da trasformare e considerevoli investimenti di tempo e costi per il mantenimento delle piante *in vitro*, prima che l'espressione delle proteine può essere analizzata (Fischer e Emans, 2000).

La tecnica di trasformazione transiente presenta diversi vantaggi rispetto alla trasformazione stabile. L'espressione è rapida e flessibile, non è richiesta la rigenerazione dei tessuti vegetali ed i risultati riguardo l'espressione proteica possono essere ottenuti in pochi giorni (Kapila *et al.*, 1997). Tre sono i principali sistemi di espressione transiente: bombardamento tramite il metodo biolistico; infezione dei tessuti vegetali tramite vettori virali modificati e infiltrazione dei tessuti vegetali intatti mediante agrobatterio (agroinfiltrazione). I livelli di espressione variano a seconda della tecnica utilizzata. Il bombardamento con il metodo biolistico raggiunge poche cellule e la trascrizione è limitata (Christou, 1993). I vettori virali infettano sistemicamente molte cellule e la trascrizione del gene d'interesse è amplificata anche dalla replicazione virale nel citoplasma, generando sempre più copie del trascritto, anche se è noto dalla letteratura che sono frequenti i fenomeni di ricombinazione che inducono ad un silenziamento genico. La tecnica dell'agroinfiltrazione, invece, infetta un elevato numero di cellule e la trascrizione del T-DNA avviene nel nucleo delle cellule infettate. Inoltre, l'agrobatterio contenente il plasmide di interesse può essere conservato ed essere utilizzato successivamente per generare piante transgeniche. Inoltre, la tecnica di agroinfiltrazione può essere utilizzata facilmente in specie recalcitranti alla trasformazione stabile (Kapila *et al.*, 1997). Due sono le tecniche disponibili per la trasformazione transiente mediata da agrobatterio. La prima descritta da Kapila e collaboratori (1997), dove la sospensione batterica opportunamente indotta con acetosiringone, messa in contatto con i tessuti vegetali, è mantenuta sottovuoto per alcuni minuti. I tessuti vengono mantenuti in ambiente umido per un tempo che varia dai 3 ai 5 giorni e successivamente analizzati per verificare l'espressione della proteina o la presenza del trascritto di interesse. L'altra metodica descritta da Voinnet *et al.* (2003), consiste nell'infiltrare la soluzione contenente l'agrobatterio indotto, direttamente, nell'epidermide delle foglie della pianta tramite l'uso di una siringa. Il vantaggio di quest'ultima risiede nel fatto che i tessuti da infettare rimangono attaccati alla

pianta in modo che l'analisi dei tessuti agroinfiltrati si possa fare anche a distanza di 10-20 giorni dall'inoculo senza che i tessuti vegetali muoiano. In entrambi i casi, comunque, è possibile esprimere più geni contemporaneamente, semplicemente unendo due ceppi di agrobatterio contenenti ognuno un plasmide con il gene d'interesse.

Un ulteriore vantaggio della trasformazione transiente è dato dalla possibilità di provare un gran numero di costrutti. E' possibile, infatti, indirizzare la proteina d'interesse in vari compartimenti cellulari in modo da verificare quale sia l'ambiente più adatto per un maggior accumulo, oppure confrontare i livelli di espressione di vari promotori e 5'UTR.

In letteratura sono riportati numerosi lavori dove le proteine sono espresse in vari compartimenti cellulari. Ad esempio, in alcuni lavori i livelli di espressione di diversi anticorpi in pianta variavano dall'1% al 13% a seconda se le proteine venivano fatte accumulare nel reticolo o nell'apoplasto.

Lavori che utilizzano l'espressione transiente di molecole antigeniche sono riportati da Li *et al.* (2006), che hanno espresso mediante agroinfiltrazione la proteina S1 del virus SARS fusa alla GFP. Anche Meyers *et al.* (2005) riportano l'espressione del gene *gag* e della sequenza corrispondente alla subunità p24 mediante infiltrazione con agrobatterio, mostrando buoni livelli di espressione. Numerosi sono i lavori di espressione transiente riguardanti il gene *gus* ed il gene della *gfp* (Wydro *et al.*, 2006; Lee e Yang, 2006; Joh *et al.*, 2005).

Una limitazione di questa tecnica è data dalla possibilità di silenziamento genico mediato da "RNA interfering" (Voinnet *et al.*, 2003). In questo caso, come riportano gli stessi autori, l'uso di geni soppressori del silenziamento (p19, HcPro) può prolungare l'espressione del gene d'interesse anche per 10 – 20 giorni dall'infiltrazione.

Chiaramente la trasformazione transiente potrebbe rappresentare una base di partenza per verificare il corretto funzionamento dei plasmidi costruiti, in modo da utilizzare per la trasformazione stabile solo i vettori che hanno dato risultati migliori nella trasformazione transiente, riducendo così tempi e costi.

### 1.3 HIV-1 ed HPV16

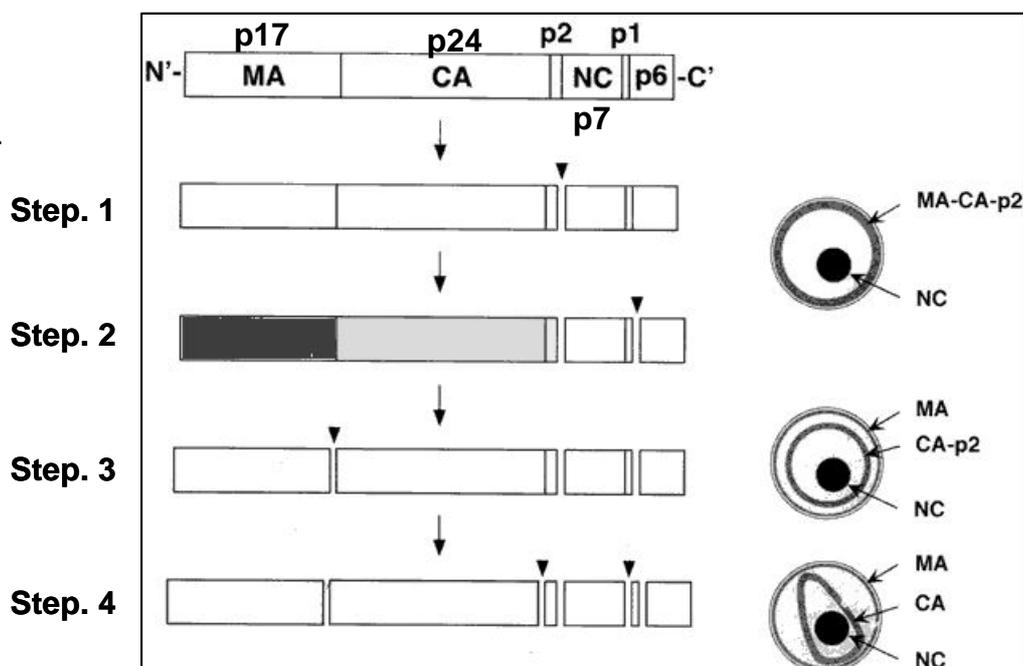
In questo lavoro di tesi, l'interesse è volto all'espressione nucleare in piante di tabacco di due geni di due virus umani: il gene *gag* del virus HIV-1 e il gene *L1* del Papillomavirus Umano (HPV16), responsabili rispettivamente della sindrome da immunodeficienza acquisita (AIDS) e del cancro della cervice uterina. Entrambi i geni sono responsabili delle formazione di strutture simil virali "Virus Like Particles" (VLPs), altamente immunogeniche in quanto si assemblano in modo simile al virus mimando l'infezione e capaci di presentare antigeni in superficie in forma nativa in grado di indurre la produzione di anticorpi neutralizzanti (Biemelt *et al.*, 2003). VLP del virus HIV-1 sono state già espresse in un sistema animale dall'INT Pascale di Napoli (Buonaguro *et al.*, 2001), mentre VLP del virus HPV16 sono riportate in lavori riguardanti l'espressione nucleare del gene L1 HPV16 in patata ed in tabacco, ma con bassi livelli di espressione (Biemelt *et al.*, 2003; Varsani *et al.*, 2003; Warzecha *et al.*, 2003).

Il virus dell'HIV-1 è un retrovirus della classe dei lentivirus. Ha un forte tropismo per i linfociti CD4, infettandoli ed integrando il suo genoma all'interno dei linfociti. Il materiale genetico del virione è costituito da due copie di RNA a polarità positiva, le quali sono legate a due proteine basiche del peso, rispettivamente, di 7 e 6 KDa (denominate P7 e P6). Come tutti i retrovirus, l'HIV

possiede i tre geni fondamentali per la sua replicazione: *gag*, *pol* ed *env*. Il gene *gag* codifica inizialmente per una proteina di 55 KDa che per tagli successivi ad opera della proteasi virale produce le proteine del virione, p17 (“matrix domain”), p24 (“capsid domain”), p7 (“nucleocapsid domain”) e p6 (“C-terminal domain”). Due sequenze (p2 e p1) separano rispettivamente il dominio p24/p7 e il dominio p7/p6 (Morikawa, 2003).

Dal gene *pol* derivano la trascrittasi inversa, la proteasi e l'integrasi, mentre il gene *env* codifica per le proteine dell'involucro esterno. Oltre a questi geni, l'HIV-1 contiene altri sette geni accessori che hanno funzioni regolatorie del ciclo virale e della sintesi proteica: *tat*, *rev*, *nef*, *vpr*, *tev*, *vif*, *vpu*.

Sia il gene *gag* che il gene *pol* sono trascritti in un mRNA unico che viene poi tradotto in una proteina di 180 KDa (P180) che viene poi tagliata tramite proteolisi. La sua scissione determina la formazione della proteasi (p10), della trascrittasi inversa (p51/p66), della integrasi e di una proteina di 55 KDa (pr55). Dalla pr55, sempre per proteolisi, derivano la p17, la p24 e la p15. La p15 è il progenitore della P6 e della P7, anch'esse ottenute tramite l'intervento della proteasi virale. Come riportato da Morikawa (2003), la proteasi virale effettua i tagli della pr55 in una successione precisa (Fig. 2). Il primo taglio è effettuato subito dopo il dominio p2, separando i domini p17+p24+p2 dai domini p7+p6. Successivamente, la proteasi virale separa il dominio p7 dal dominio p6. Nel terzo “step” la proteasi separa il dominio p17 dal dominio p24 ed infine il dominio p24 è separato dal dominio p2. La proteina p24 va a formare il “core” virale, una struttura conica all'interno della quale è presente tutto il genoma virale (9 Kb), insieme alla trascrittasi inversa (una DNA polimerasi RNA-dipendente), alla proteasi ed all'integrasi. Tra il “core” e l'involucro lipoproteico si trova la proteina virale P17 che nella sua forma nativa è miristilata all'N-terminale. La miristilazione è un fenomeno importante per la successiva interazione della p17 con la membrana cellulare e per dare avvio alla gemmazione del virus dalla cellula infettata.



**Fig. 2: Processamento della proteina pr55, codificata dal gene *gag*, nelle sue subunità, ad opera della proteasi virale. Modificata da Morikawa (2003)**

L'infezione inizia quando il virus penetra nei linfociti attraverso il legame della gp120 ai recettori CD4. Una volta che il virus è all'interno della cellula viene diretta la trascrizione inversa del RNA virale in DNA, che si integra nel genoma della cellula infettata. Questa fase di latenza può durare mesi o addirittura anni. La malattia è conclamata quando si ha l'espressione dei geni virali precoci e l'assemblaggio delle particelle virali in nuovi virus che usciranno dalla cellula infetta e potranno infettarne altre. Ripetuti cicli di infezione portano ad una distruzione completa dei linfociti T e successivamente dei linfociti B e macrofagi, rendendo così il sistema immunitario progressivamente meno attivo.

La malattia causata dal virus (AIDS) è causa di numerosi decessi che si verificano ogni anno ed in ogni parte del mondo. Circa 30,6 milioni di persone hanno contratto il virus soprattutto nelle regioni africane, dove la malattia è divenuta endemica. A tutt'oggi la malattia non è curabile ma sono presenti solo terapie farmacologiche HAART ("Highly Active Anti-Retroviral therapy") che prevedono la somministrazione contemporanea di farmaci inibitori della trascrittasi inversa e di inibitori della proteasi. L'assenza di un vaccino protettivo è dovuto principalmente alla grande variabilità genetica del virus soprattutto nel gene *env* che codifica per la proteina (gp120) responsabile del riconoscimento del recettore sui linfociti. La variabilità genetica determina, infatti, la continua generazione di varianti molecolari che sono state raggruppate in 10 sottotipi (Clades) principali (A-K) riuniti in un unico gruppo denominato M (Major), ed alcuni sottotipi altamente divergenti nel gruppo denominato O (Other). La ricerca di una strategia vaccinale è finalizzata alla ricerca dell'antigene più efficace, che induca una forte risposta immunitaria sia umorale che cellulo-mediata. Tra i modelli proposti fino ad oggi, quello più promettente sembra quello basato sulla produzione di VLPs. Per questo motivo il gene *gag* è un buon candidato per la produzione di un vaccino contro l'HIV-1. In letteratura non sono riportati lavori basati sull'espressione di VLPs del virus HIV-1 in pianta, ma sono presenti diverse pubblicazioni sull'espressione della subunità del "core" virale, la p24. Ultimamente il gruppo di Ma *et al.* (2006), ha ottenuto in piante di tabacco transgeniche un'espressione circa 13 volte più abbondante quando la proteina P24 è stata fusa con le regioni  $\alpha 2$  e  $\alpha 3$  delle immunoglobuline. Altri lavori (Zhang *et al.*, 2002; Perez-Filgueira *et al.*, 2004) riportano l'espressione della p24 con livelli circa dello 0,35% delle PTS sia in piante transgeniche, sia attraverso l'uso di un vettore virale TMV.

Il virus HPV (Human Papilloma Virus) è un virus appartenente al gruppo dei *papovavirus*. Le infezioni da HPV sono estremamente diffuse e possono causare malattie della pelle e delle mucose. Si conoscono oltre 100 tipi di HPV dei quali la maggior parte causa malattie lievi come verruche. Alcuni tipi, invece, sono responsabili di tumori benigni quali il condiloma genitale o di tumori maligni come il cancro al collo dell'utero o al pene.

I sottotipi del virus del papilloma umano che possono infettare le mucose genitali possono venir suddivisi in HPV a basso rischio (6, 11, 42, 43, 44) e HPV ad alto rischio (16, 18, 31, 33, 35, 39, 45, 51, 52, 56, 58, 59, 68). Si calcola che oltre il 70% delle donne contragga un'infezione genitale da HPV nel corso della propria vita, ma la grande maggioranza di queste infezioni è destinata a scomparire spontaneamente nel corso di pochi mesi. Solo in caso di persistenza nel tempo di infezioni di HPV ad alto rischio oncogenico è possibile, in una minoranza dei casi, lo sviluppo di un tumore maligno del collo uterino. Nel caso

delle infezioni benigne, il virus non integra il suo genoma all'interno della cellula infettata, mentre nei tumori maligni il DNA del virus si ritrova nel nucleo della cellula ospite.

Il virus dell'HPV possiede un genoma a DNA a doppio filamento di circa 8 Kb. Il genoma virale è suddiviso in una regione tardiva, una regione precoce e una regione distale. La prima comprende i geni L1 ed L2 che codificano per la maggiore o minore proteina capsidica, la seconda tutti i geni che controllano il ciclo cellulare, la terza gli elementi regolatori della trascrizione.

I geni E6 ed E7 del virus sono responsabili della produzione di due oncoproteine che interagiscono con le proteine cellulari p53 e Retinoblastoma che regolano il ciclo cellulare. L'interazione di queste proteine virali con quelle cellulari altera il ciclo cellulare dell'ospite promuovendo la crescita continua delle cellule infettate.

Sebbene l'infezione del virus è diminuita nel corso degli anni, è ancora molto frequente nelle donne (Parkin *et al.*, 2001). L'assenza di una terapia efficace, rende necessario lo sviluppo di un vaccino protettivo contro il virus. Negli ultimi anni, un vaccino che sembra dare ottimi risultati è stato accettato dalla "Food and Drug Administration" sulla base dei dati riportati nel lavoro di Ronco *et al.* (2006). Attualmente il vaccino è nella terza fase di sperimentazione clinica e dovrebbe essere attivo contro i sottotipi virali HPV16, 11, 8 e 6.

Altri vaccini sono in fase di ricerca. In particolare, anche in questo caso, come per il virus HIV-1, una strategia basata sulle VLPs sembra essere promettente per la produzione di un vaccino efficace contro il virus HPV. Per questo motivo, la proteina L1 è un'ottima candidata in quanto è responsabile dell'autoassemblamento e della formazione delle particelle simil virali. La sequenza che codifica per la proteina L1 è lunga circa 1500 coppie di basi e codifica per una proteina di 55 KDa. All'estremità C-terminale, presenta un segnale di localizzazione nucleare (NLS) che indirizza la proteina nel nucleo della cellula infettata (Florin *et al.*, 2002). Anche la proteina L2 viene indirizzata nel nucleo subito dopo la sua traduzione ed, insieme alla proteina L1, regola l'assemblamento del DNA (Bousarghin *et al.*, 2003).

Solo cinque lavori sono presenti in letteratura riguardo l'espressione della proteina L1 in pianta (Biemelt *et al.*, 2003; Liu *et al.*, 2005; Varsani *et al.*, 2003; Varsani *et al.*, 2006; Warzecha *et al.*, 2003). Da tutti i lavori si evince la difficoltà di esprimere questo gene in un sistema vegetale. Gli autori riportano che un'espressione dello 0,5% rispetto alle PTS si riesce ad ottenere utilizzando un gene con una sequenza nucleotidica ottimizzata per l'espressione in cellule umane. Solo Varsani *et al.* (2006) riportano l'espressione del gene L1 "wildtype" di HPV16 in tabacco utilizzando un vettore virale TMV, con livelli di espressione pari a 20-37 ug/Kg di tessuto fresco infettato.

#### **1.4 Scopo e contenuto della tesi**

Questo progetto di tesi è stato svolto presso l'Istituto di Genetica Vegetale di Portici, in collaborazione con l'I.N.T. Pascale di Napoli, ed ha previsto l'espressione in pianta di tabacco dei geni *gag* ed L1 di due virus umani HIV-1 e HPV16.

Per poter ottenere l'espressione in pianta di questi geni sono stati costruiti numerosi vettori di espressione. Per il gene intero *gag* sono stati prodotti in totale sette vettori. Cinque di questi plasmidi derivano dai vettori *ImpactVector* ([www.impactvector.com](http://www.impactvector.com)) e contengono lo stesso promotore 5'UTR e terminatore

del gene *rbcS* di crisantemo e due “tag” al 3’ (c-myc e 6His). Questi plasmidi si differenziano per la presenza o meno di peptidi segnale per l’indirizzamento delle proteine in differenti compartimenti subcellulari: citoplasma, apoplasto, reticolo endoplasmico, cloroplasto e mitocondrio. Altri due vettori contenenti il gene intero *gag* sono stati prodotti aventi però il promotore 35S<sup>2</sup> del virus del mosaico del cavolfiore, la 5’UTR  $\Omega$  del virus del mosaico del tabacco e le stesse “tag” al COOH terminale presenti nei vettori descritti precedentemente. Questi due ultimi vettori si differenziano per la presenza o meno di un peptide segnale per l’accumulo della proteina nei cloroplasti. In questo caso è stata utilizzata come sequenza segnale quella del gene *rbcS* di *Pisum sativum* (Anderson e Smith, 1986). Data la presenza dei pochi dati di letteratura riguardo l’espressione del gene intero *gag* in pianta ed al fine di capire se una subunità specifica del gene *gag* potesse essere responsabile del basso accumulo della proteina Gag nelle cellule vegetali, sono stati costruiti altri vettori contenenti le subunità *p17*, *p24* e  $\Delta p17$  (il gene *gag* intero senza la *p17*) del gene *gag*. Queste sequenze sono state clonate nei vettori ImpactVector ed indirizzate nel citoplasma e nel cloroplasto.

Per l’espressione del gene *L1* sono stati costruiti in totale 13 vettori. Il gene *L1<sup>v</sup>* (con la sequenza virale wild type), privo della sequenza di localizzazione nucleare (NLS), ed il gene *L1<sup>v</sup>* tronco degli ultimi 205 aminoacidi (*L1<sup>v</sup>* $\Delta$ 205) sono stati clonati in cinque plasmidi derivanti dai vettori ImpactVector in grado di indirizzare ognuno la proteina in cinque compartimenti differenti (citosol, apoplasto, reticolo endoplasmico, cloroplasto e mitocondrio). Per il gene *L1<sup>v</sup>* $\Delta$ 205, invece, la sequenza codificante è stata clonata solo nei vettori con il segnale di accumulo per il citoplasma e per il cloroplasto.

Inoltre, nel nostro laboratorio erano disponibili due sequenze del gene *L1*, ottimizzate per l’espressione plastidiale (*L1<sup>pt</sup>*) e per l’espressione in cellule umane (*L1<sup>h</sup>*). Sia il gene *L1<sup>pt</sup>* che il gene *L1<sup>h</sup>* sono stati clonati nei vettori ImpactVector che indirizzano le relative proteine nel citoplasma e nel cloroplasto. Solo per il gene *L1<sup>pt</sup>* è stata effettuata la stessa delezione del gene *L1<sup>v</sup>* degli ultimi 205 aminoacidi.

Per controllare il corretto funzionamento dei vettori ottenuti, è stato utilizzato il gene reporter *gus* contenente l’introne (int) PIV2 del gene *ST-LS1* (Vancanneyt *et al.*, 1990). Il gene reporter è stato clonato in tutti i vettori ImpactVector ed indirizzato negli stessi cinque compartimenti subcellulari utilizzati per i geni d’interesse.

Tutti i vettori sono stati provati in esperimenti di espressione transiente in *N. tabacum* e *N. benthamiana*. Gli espianti agroinfiltrati sono stati prelevati dopo 3 giorni dall’inoculo ed i tessuti sono stati analizzati a livello molecolare tramite RT-PCR per verificare la corretta trascrizione dei geni e con l’analisi Western per la verifica della produzione delle proteine nei tessuti vegetali.

## 2.0 MATERIALI E METODI

### 2.1 Materiale utilizzato

Per la trasformazione transiente di tabacco sono state utilizzate piante di *N. tabacum* e *N. benthamiana*. Le piante di *N. tabacum* sono state allevate in condizioni sterili in camera di crescita a 24 °C, propagate in barattoli di vetro su substrato contenente (Murashige, 1962) e saccarosio 3% (Svab e Maliga, 1993), e poi trasferite in vivo dopo circa 2 settimane. I semi di *N. benthamiana*, invece, sono stati fatti germinare *in vivo* in plateau. Le piantine germinate, dopo circa 1 settimana, sono state trasferite singolarmente in vaso e fatte crescere in camera di crescita in condizioni controllate.

I geni utilizzati in questa tesi sono il gene *gag* del virus HIV-1 e il gene *L1* “African Type (numero di accessione AF472508) del Papilloma Virus (HPV), forniti dal dott. F. Buonaguro dell’ I.N.T. G. Pascale di Napoli.

Come gene “reporter” è stato utilizzato il gene *gus-int*, dove la sequenza codificante l’enzima  $\beta$ -glucuronidasi è interrotta da un introne modificato “PIV2” del gene *ST-LS1* (Vancanneyt *et al.*, 1990)

I vettori utilizzati per gli esperimenti di trasformazione sono stati ottenuti mediante amplificazione PCR e subclonaggi delle sequenze di interesse in plasmidi commerciali ([www.impactvector.com](http://www.impactvector.com)) e in plasmidi già disponibili nel nostro laboratorio.

I vettori ImpactVector (I.V.), presentano il promotore, la 5’UTR ed il terminatore del gene della subunità piccola della RuBisCo di crisantemo (*rbcS*), un sito multiplo di clonaggio e due “tag” c-myc e 6His al COOH terminale. I vettori si differenziano in base ai peptidi segnale per l’indirizzamento delle proteine in diversi compartimenti subcellulari: il vettore 1.1tag non presenta alcun peptide segnale per l’accumulo nel citoplasma, il vettore 1.2tag presenta il peptide segnale del gene “*sea anemone equistatin*” per l’indirizzamento nell’apoplasto, il vettore 1.3tag presenta lo stesso peptide segnale del vettore 1.2, ma con il motivo KDEL al COOH terminale per la ritenzione della proteina nel reticolo endoplasmatico, il vettore 1.4tag presenta il peptide segnale del gene *rbcS* di *Chrysanthemum morifolium* per l’indirizzamento nel cloroplasto ed il plasmide 1.5tag presenta il peptide segnale del gene *CoxIV* di lievito per l’indirizzamento della proteina nella matrice mitocondriale. I vettori su descritti non essendo vettori binari, sono stati utilizzati solo per il clonaggio delle sequenze d’interesse. Le cassette di espressione, contenenti il promotore, il peptide segnale, il gene d’interesse, le “tag” ed il terminatore sono state infine clonate nel plasmide binario finale pBINplus, fornito insieme ai vettori ImpactVector su descritti.

### 2.2 Gene *gus-int*

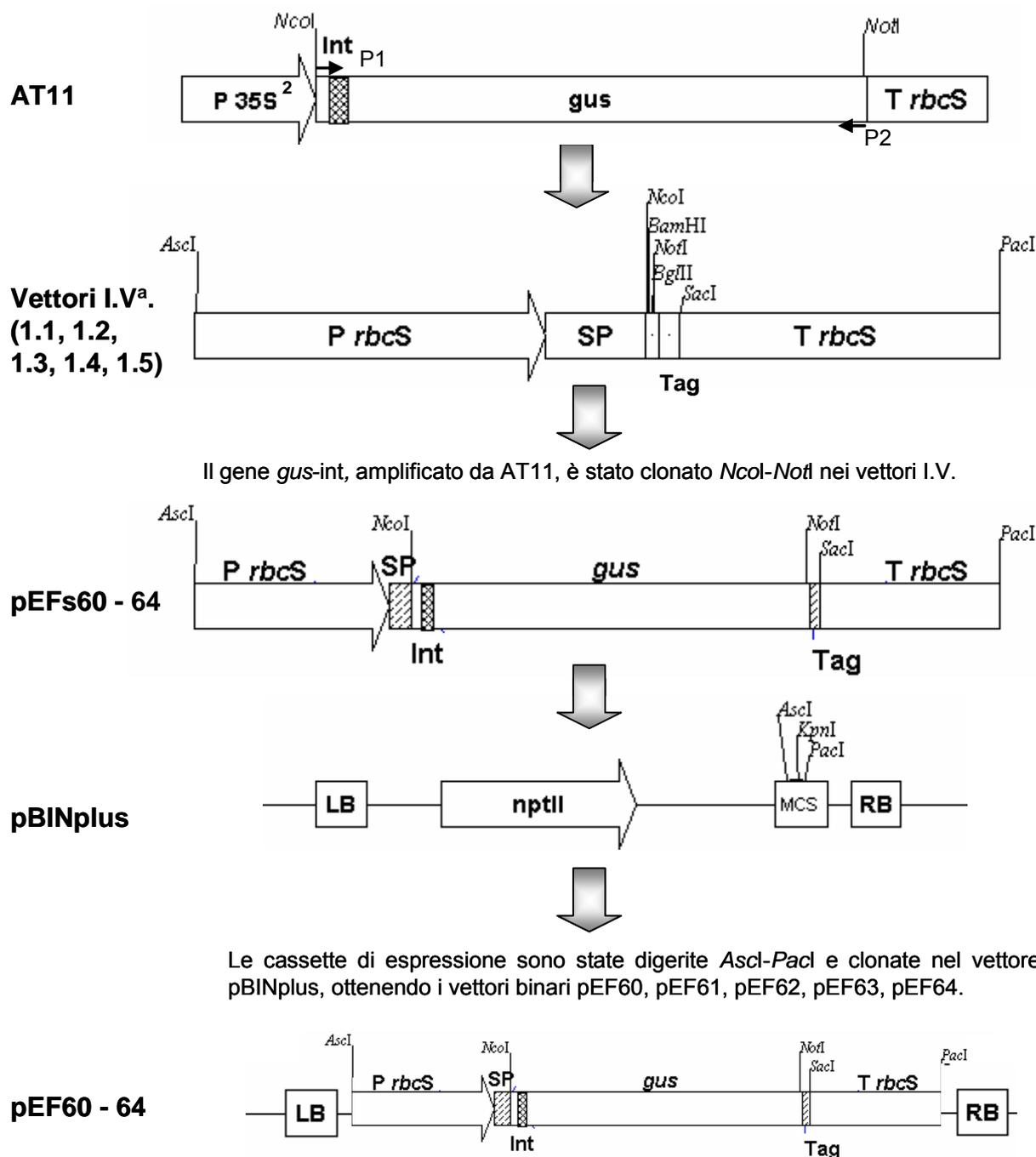
#### 2.2.1 Costruzione vettori pEF60, pEF61, pEF62, pEF63, pEF64

Il gene *gus-int*, è stato amplificato utilizzando i primer P1 e P2 mediante reazione PCR dal plasmide AT11 disponibile nel nostro laboratorio (Fig. 3), digerito con gli enzimi di restrizione *NcoI* - *NotI* e clonato nei cinque vettori I.V. come mostrato in Fig. 3. I cinque vettori intermedi, pEFs60, pEFs61, pEFs62, pEFs63, pEFs64, indirizzano la proteina ricombinante in compartimenti subcellulari diversi (citosol, apoplasto, reticolo endoplasmatico, cloroplasto e mitocondrio). Dai vettori

intermedi è stata rimossa l'intera cassetta di espressione utilizzando gli enzimi *Ascl* e *Pacl* ed essa è stata clonata nel vettore binario finale pBINplus, ottenendo i vettori binari pEF60, pEF61, pEF62, pEF63, pEF64. I "primer" utilizzati in tutto il lavoro di tesi sono elencati in Tabella 4.

**Tabella 4:** Oligonucleotidi utilizzati in questo lavoro di tesi per le analisi PCR e per i clonaggi.

Codice	Nome	Sequenza (5'-3')	Tm
P1	GUSNcoF	ACGCGTCGACCATGGTACGTCCTGTAGAAACCCCA	72
P2	GUSNotR	CGGAAGATCTGCGGCCGCTTGTGGCTCCCTGGTGCGG	71
P3	GagNcoF	ACGCGTCGACCATGGGCGCGAGAGCGTCAGT	76,5
P4	GagNotRev	AAGGAAGAAAGCGGCCGCTTGTGACGAGGGGTCGTTGCC	77,3
P5	GAGSacRev	CGAGCTCTTATTGTGACGAGGGGTCGTTGCC	72,6
P6	OmegaFor	CCGCTCGAGAAGCTTTATTTTTACAACAAT	73
P7	GAGSalF	ACGCGTCGACAAATGGGCGCGAGAGCGTCA	74,2
P8	GAGEcoRI	CGGAATTCTTATTGTGACGAGGGGTCG	67,6
P9	NcoI-TP	CATGCCATGGCTTCTATGATATCCTCTTCA	66
P10	UP	ATGGGCGCGAGAGCGTCAGT	66
P11	C10	GGTGATATGGGCTGATGTACC	62,6
P13	P17 Rev	CGGAAGATCTGCGGCCGCGTAATTTTGGCTGGCCTGGC	77,6
P14	P24 Forw	CGTCGACCATGGCTATAGTGCAAAACCTCCAGG	72,1
P15	P24 Rev	CGGAAGATCTGCGGCCGCCAAAATTCTTGCTTTATGGC	73,3
P16	L1NcoF	ACGCGTCGACCATGGCTCTTTGGCTGCCTAGCGGAG	76
P17	L1NotRev	CGGAAGATCTGCGGCCGCTCCTAATGTAAATTTGGTT	75
P18	L1205NotR	CGGAAGATCTGCGGCCGCAACCATAGAACCACTAGGTGT	75,2
P19	L1ptUP	ATCATGCCATGGCTTTGTGGTTACCGAATTCAGACCTAGGAG ATCTGAGC	74,6
P20	L1ptDOWN	GCTCAGATCTCCTAGGTCTGAATTCGGTAACCACAAAGCCAT GGCATGAT	74,6
P21	L1ptNcoF	CGTCGACCATGGCTATAGTGCAAAACCTCCAGG	72,1
P22	L1pt205R	GCATCAGATCTGCAACCATCGATCCAGATGGAG	70,9
P23	RbcSTemR	AAATAGTAAACTTAGTAGCCATC	56
P24	RbcSPromF	CCGTTAGATAGCAAACAACA	56
P25	Gus3'	CCCGGCAATAACATACGGCGTG	61
P26	Gus5'	CCTGTAGAAACCCCAACCCGTG	62
P27	Gag seq rev	TCTGATCCTGTCTGAAGGGCTGCT	66,3
P28	Gag1	CAAATGGTACATCAGCCCATA	55,9
P29	Gag2 anti	CTCTAAAGGGTTCCTTTGGTCC	60,3
P30	L1h F	AGTACGTGGCCAGAACCAAC	60
P31	L1h R	GTCGTCCAGCTTGTTCAGC	61
P32	L1wtF5'	TTGCCTCCTGTCCCAGTATC	62
P33	L1 1.2 Rev	GCCGCTGTGTATCTGGATTGTAAAA	62,9
P34	L1ptF5'	CCTGTTTCAAAGTCGTTTCG	58,7
P35	L1ptR5'	TGTTTAAACAACGGGTGTCCA	58,4



**Fig. 3: Costruzione dei vettori pEF60, pEF61, pEF62, pEF63, pEF64.**

a) Sono stati utilizzati cinque vettori ImpactVector ([www.impactvector.com](http://www.impactvector.com)) contenenti o meno il segnale (SP) per i compartimenti subcellulari:

1.1 = citoplasma, 1.2 = apoplasto, 1.3 = reticolo endoplasmatico, 1.4 = cloroplasto, 1.5 = mitocondrio.

Tutti i vettori possiedono il promotore, la 5'UTR ed il terminatore del gene *rbcS*, e le "tag" c-myc e 6His

AT11 = vettore binario sorgente contenente il gene *gus-int*

Int = introne modificato PIV2 del gene *ST-LS1* (Vancanneyt *et al.*, 1990)

## 2.3 Gene *gag*

### 2.3.1 Costruzione dei vettori pEF40, pEF41, pEF42, pEF43, pEF44

Il gene di interesse *gag* è stato amplificato con gli oligonucleotidi P3 e P4 dal vettore binario pEF1 mostrato in Fig. 4 digerito con le endonucleasi di restrizione *NcoI* - *NotI* e clonato tra il promotore e terminatore *rbcS* in tutti i vettori della serie I.V., ottenendo così i vettori pEFs40 (citosol), pEFs41 (apoplasto), pEFs42 (reticolo endoplasmatico), pEFs43 (cloroplasto), pEFs44 (mitocondrio). Successivamente questi vettori sono stati digeriti con le endonucleasi di restrizione *Ascl* - *Pacl*, rimuovendo così l'intera cassetta di espressione che è stata clonata nel vettore pBINplus, ottenendo i plasmidi binari pEF40, pEF41, pEF42, pEF43, pEF44. In Fig. 4 è schematizzato il clonaggio su descritto.

### 2.3.2 Costruzione del vettore pEF45

#### pEF100

La sequenza  $\Omega$ -luciferasi è stata digerita dal vettore pCAMBIA 1381Z usando gli enzimi *HindIII* - *SacI* e clonata nel plasmide commerciale pKS (pBlueScript; [www.fermentas.it](http://www.fermentas.it)), ottenendo così il plasmide pEF100 (Fig. 5)

#### pEF101

La sequenza di interesse *gag* è stata amplificata dal vettore pEF1 con gli oligonucleotidi P3 e P5, digerita con gli enzimi di restrizione *NcoI* - *SacI* e clonata nel plasmide pEF100, sostituendo il gene *luciferasi*, ottenendo così pEF101 (Fig 5).

#### pEF102

Per aggiungere le stesse "tag" (c-myc e 6His) presenti nei vettori pEF40, pEF41, pEF42, pEF43, pEF44 al vettore finale pEF45, parte della sequenza *gag* è stata digerita dal vettore pEF40 come frammento *AvrII* - *SacI* e clonata nel vettore intermedio pEF101, ottenendo così pEF102 (Fig. 5).

#### pEF45

Per ottenere il vettore binario finale, la sequenza " $\Omega$ -*gag*-tag" è stata digerita completamente dal vettore pEF102 con l'enzima *SacI* e mediante digestioni parziali con endonucleasi di restrizione *HindIII* e clonata nel MCS del vettore pKYLX71:35S<sup>2</sup> (Fig. 5).

### 2.3.3 Costruzione del vettore pEF46

#### pEF200

La sequenza *gag* è stata amplificata dal plasmide pEFs40 con i "primer" P7 e P8, digerita con le gli enzimi di restrizione *SalI* - *EcoRI* e clonata nel vettore pTP, ottenendo così pEF200 (Fig. 6).

#### pEF201

La sequenza TP-*gag* è stata riamplicata dal vettore pEF200 mediante gli oligonucleotidi P9 e P5, digerita *NcoI* - *SacI*, e clonata nel vettore pEF100 sostituendo il gene della *luciferasi*, ottenendo così il vettore pEF201 (Fig. 6)

### **pEF202**

Per aggiungere le due “tag” (c-myc e 6His) presenti nel vettore pEF40 alla sequenza d’interesse presente nel vettore pEF201, è stata utilizzata la stessa strategia descritta per la costruzione del vettore pEF102. (Fig. 6)

### **pEF46**

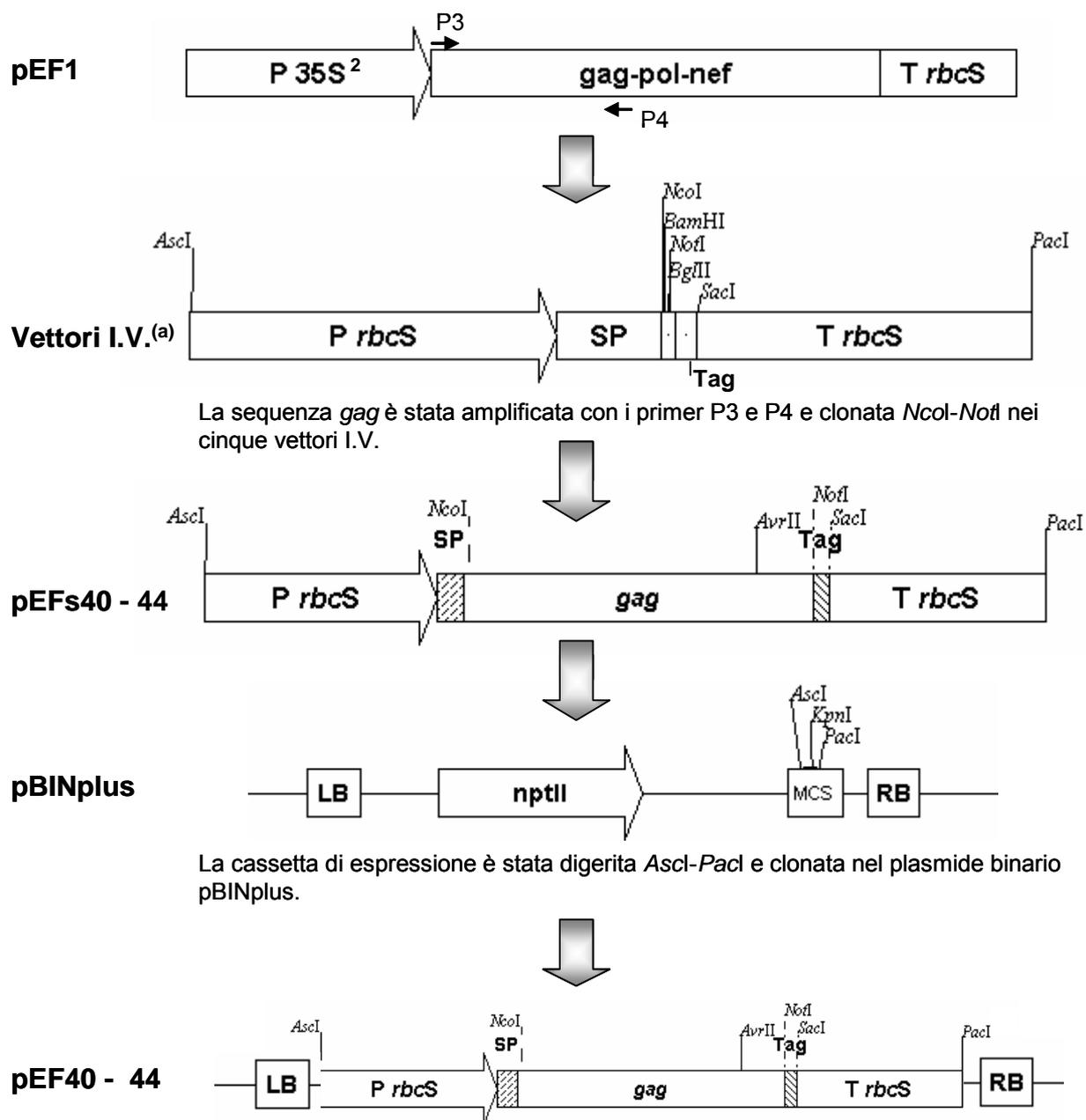
La sequenza  $\Omega$ -TP-*gag*-tag è stata digerita dal vettore pEF202 mediante digestione completa con l’enzima *SacI* e successivamente mediante digestione parziale con l’enzima *HindIII*. Il frammento  $\Omega$ -TP-*gag*-tag è stato clonato nel vettore binario pKYLX71:35S<sup>2</sup>, ottenendo il vettore binario finale pEF46. (Fig. 6).

#### **2.3.4 Costruzione dei vettori pEF47, pEF48, pEF49**

Le sequenze codificanti le subunità *p17*, *p24* e  $\Delta$ *p17* del gene *gag* sono state amplificate dal vettore pEF40, utilizzando la reazione della PCR mediante l’uso di diverse combinazioni di oligonucleotidi. I primer P12 e P13 sono stati utilizzati per amplificare il gene che codifica la *p17*; il primer P14 e P15 per la *p24* e i primer P14 e P4 per amplificare la sequenza  $\Delta$ *p17*. Alle sequenze codificanti le subunità *p24* e  $\Delta$ *p17*, è stato necessario aggiungere l’ATG mediante l’uso del primer P14. Le sequenze così amplificate sono state digerite con gli enzimi di restrizione *NcoI* – *NotI* e clonate nel vettore 1.1tag, ottenendo i plasmidi intermedi pEFs47, pEFs48, pEFs49 (Fig. 7). Le cassette di espressione contenenti i tre geni, sono state digerite *Ascl* – *PacI* e clonate nel vettore binario pBINplus, ottenendo i vettori finali pEF47, pEF48, pEF49 (Fig. 7).

#### **2.3.5 Costruzione dei vettori pEF470, pEF480, pEF490**

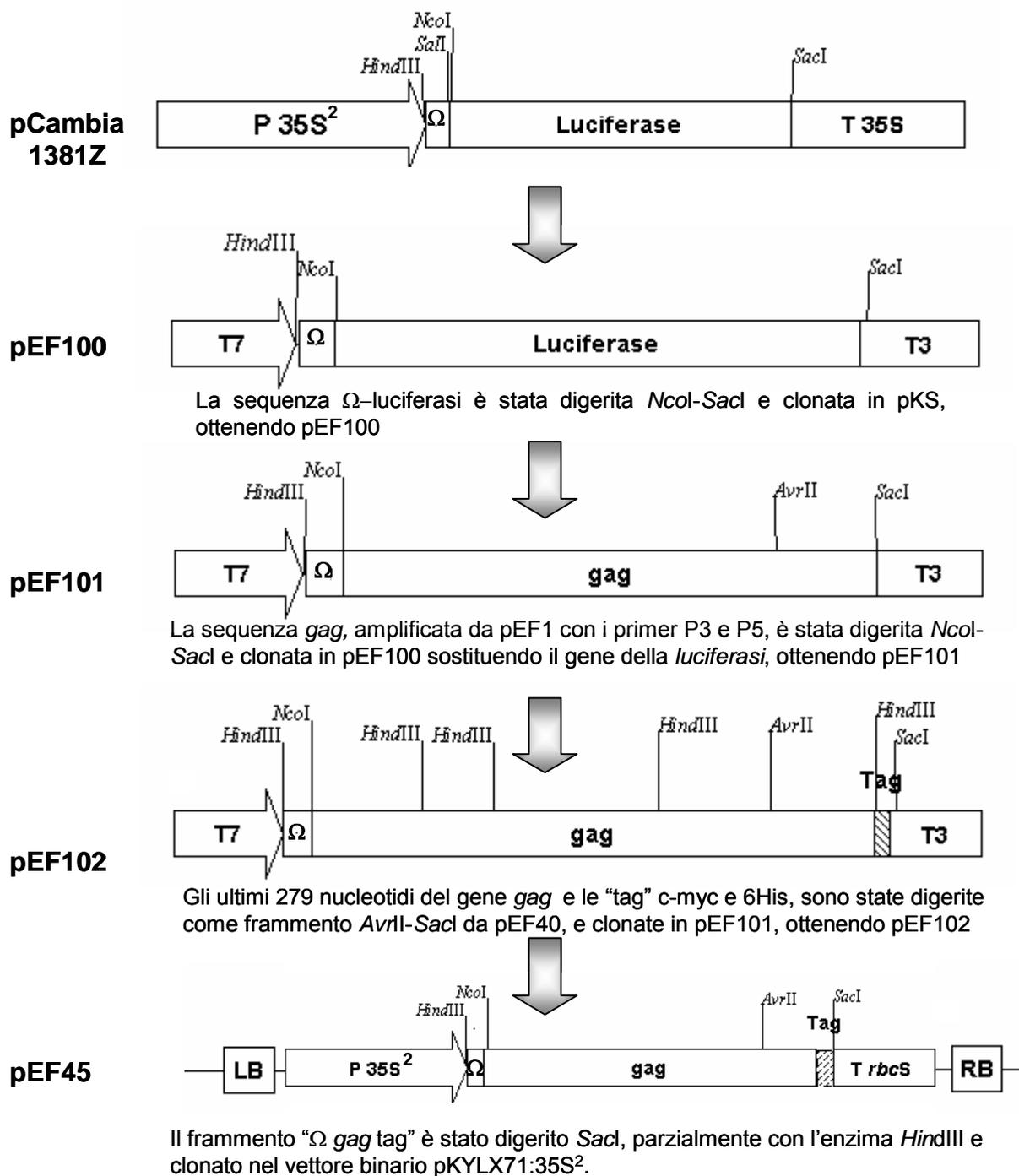
Successivamente, come mostrato in Fig. 8, dai vettori pEFs47, pEFs48, pEFs49, le sequenze virali sono state digerite con gli enzimi *NcoI* - *NotI* e clonate nel vettore 1.4tag, indirizzando le proteine codificate da questi geni nei cloroplasti, ottenendo i vettori pEFs470, pEFs480, pEFs490. Per ottenere i vettori finali, anche in questo caso le cassette di espressione sono state rimosse dai vettori intermedi pEFs470, pEFs480, pEFs490 come frammenti *Ascl* - *PacI* e clonate nel vettore binario pBINplus (Fig. 8).



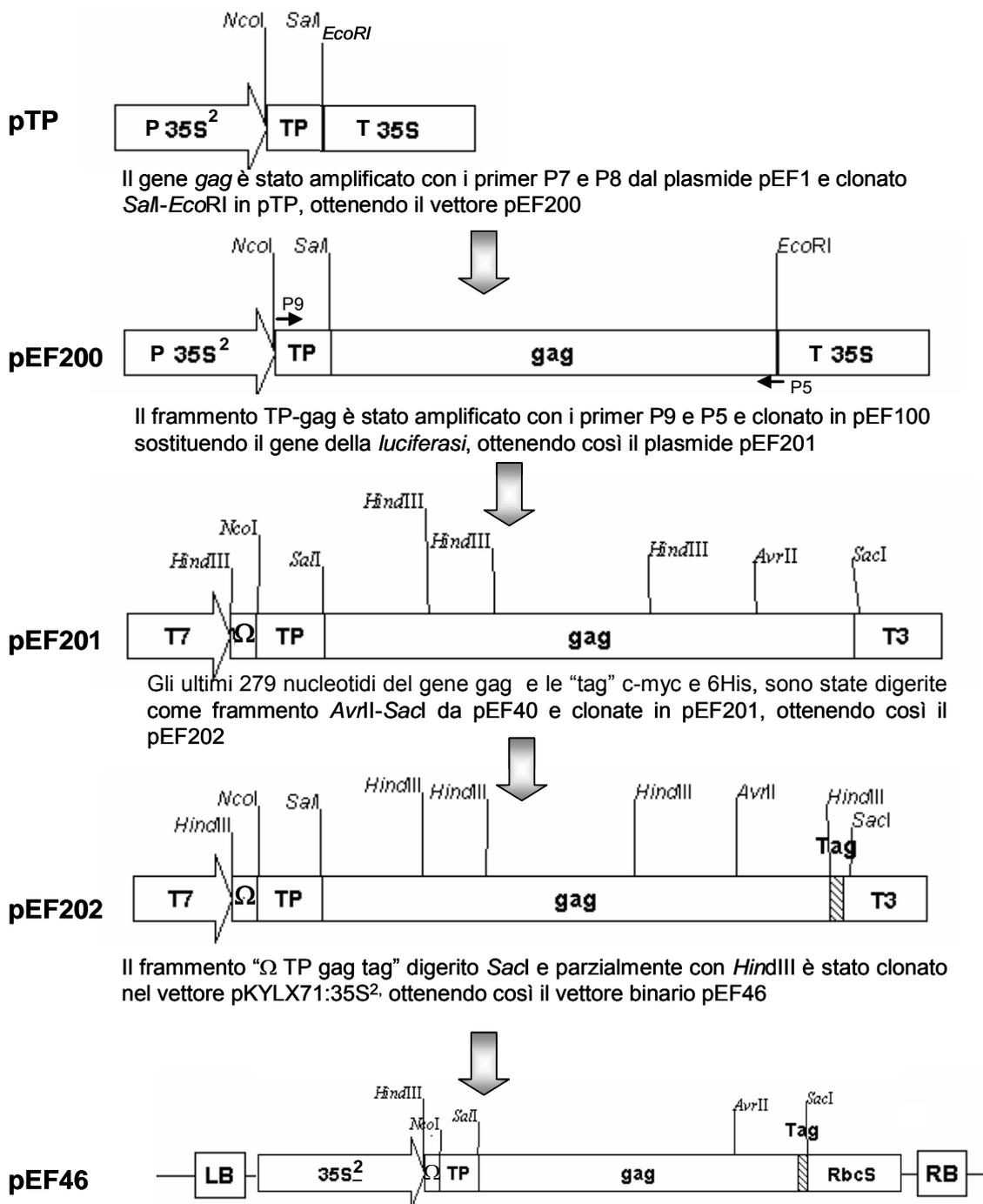
**Fig. 4: Costruzione vettori pEF40, pEF41, pEF42, pEF43, pEF44.**

<sup>a)</sup> I dettagli dei vettori ImpactVector utilizzati sono riportati in Fig. 3

pEF1 = vettore binario contenente i geni *gag-pol-nef* dell'HIV-1.



**Fig. 5: Rappresentazione schematica della costruzione del vettore pEF45.**  
 pCAMBIA 1381Z = vettore contenente la sequenza  $\Omega$  del TMV.  
 pKYLX71:35S<sup>2</sup> = Vettore binario vuoto contenente il promotore 35S<sup>2</sup>  
 Tag = sequenza c-myc e 6His derivata dai vettori I.V.  
 I plasmidi pEF100, 101, 102, sono derivati del plasmide pKS (pblueScript)



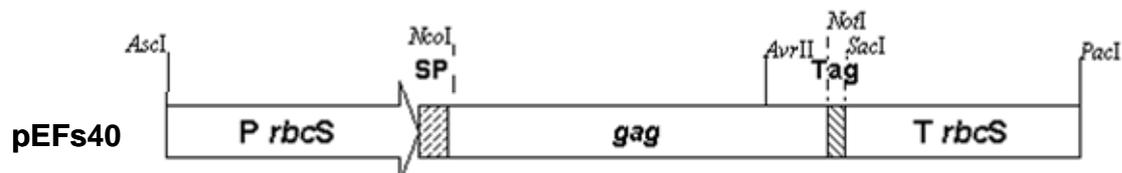
**Fig. 6: Rappresentazione schematica della costruzione del vettore pEF46.**

pTP = Plasmide con il peptide segnale (TP) per il cloroplasto del gene *rbcS* di *Pisum sativum*.

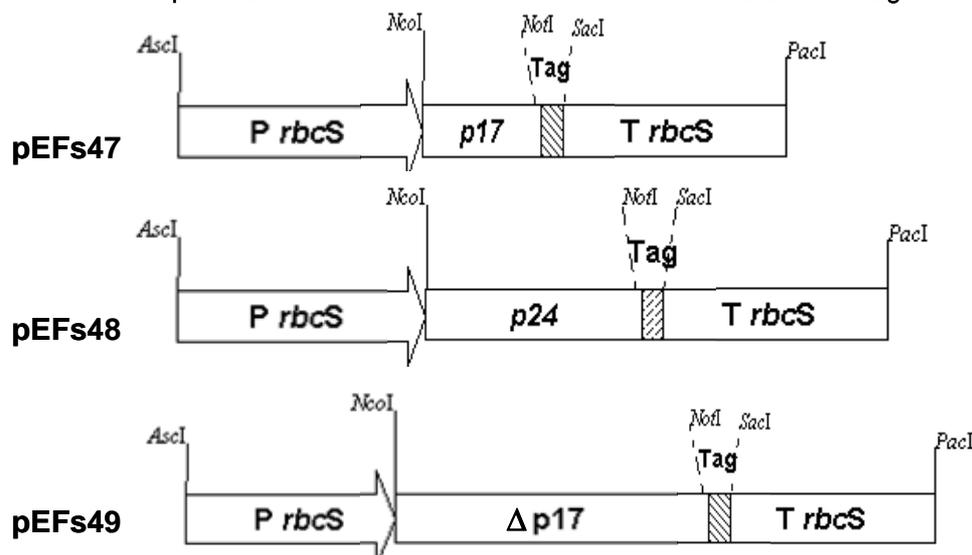
Sequenza Ω = regione 5' non tradotta del TMV

Tag = sequenza c-myc e 6His derivata dai vettori I.V

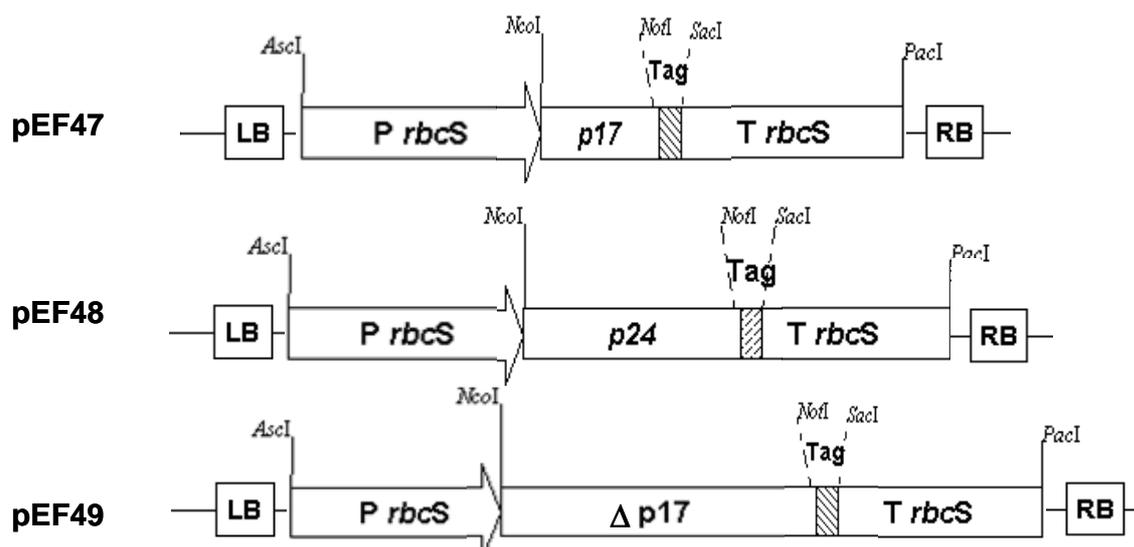
I vettori pEF200, 201 e 202 sono derivati del pKS (pBluescript)



Le subunità del gene *gag* (*p17*, *p24* e  $\Delta p17$ ), sono state amplificate dal vettore pEFs40 tramite PCR utilizzando e clonate nel vettore I.V. 1.1tag<sup>a</sup>



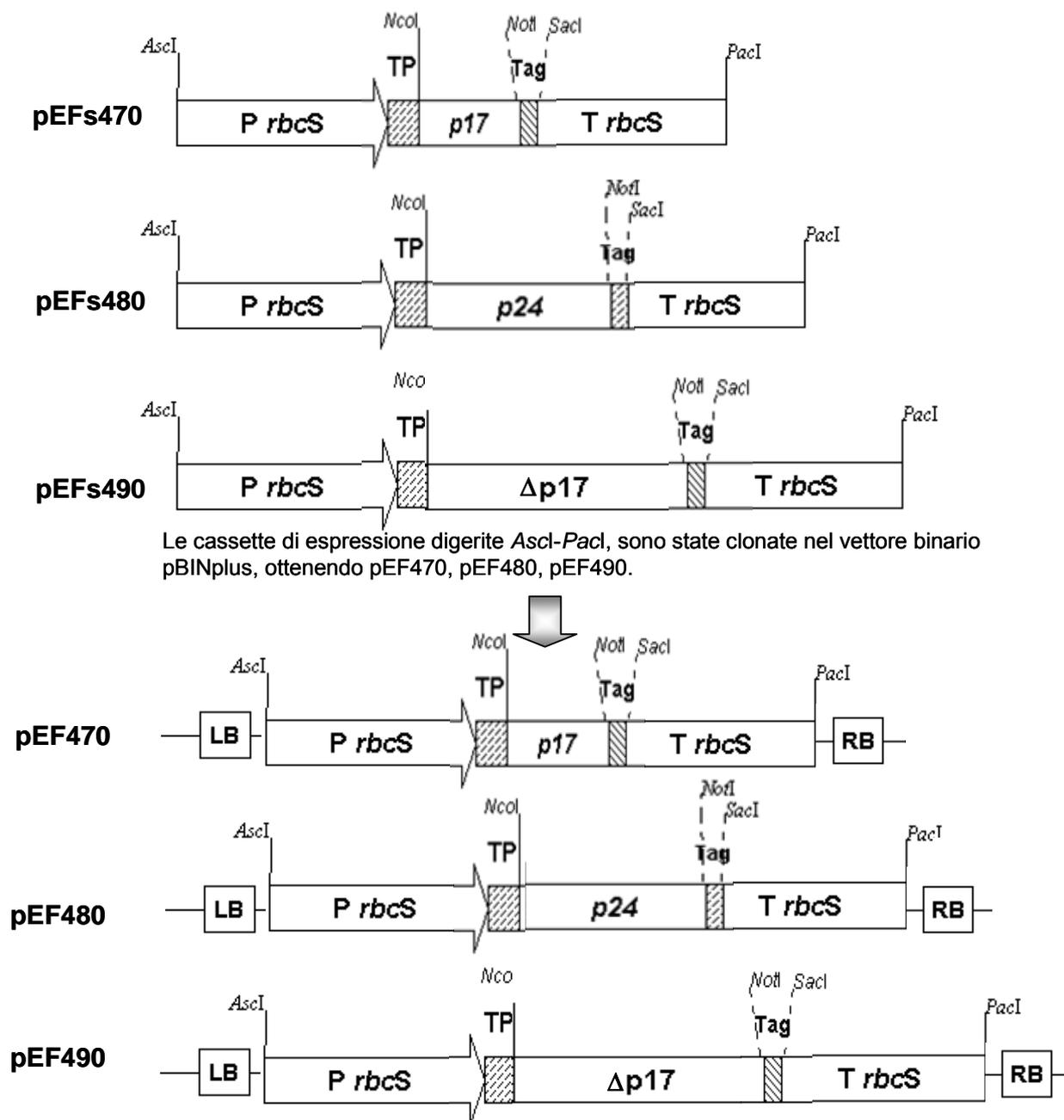
Le diverse cassette di espressione sono state digerite *Ascl*-*PacI* e clonate nel plasmide binario pBINplus, ottenendo i vettori pEF47, pEF48, pEF49



**Fig. 7: Rappresentazione schematica della costruzione dei vettori pEF47, pEF48, pEF49.**

- a) Le caratteristiche del vettore I.V. 1.1 tag sono descritte nella Fig. 3  
*p17* = gene codificante l' "N-terminal matrix domain" del gene *gag*  
*p24* = gene codificante il dominio "capsid nuclear domain" del gene *gag*  
 $\Delta p17$  = gene *gag* intero mancante del dominio *p17*

I geni codificanti le subunità *p17*, *p24* e  $\Delta p17$ , digeriti come frammenti *NcoI*-*NotI* dai vettori pEFs47, pEFs48, pEFs49, sono stati clonati nei vettori I.V.1.4tag<sup>a</sup>



Le cassette di espressione digerite *AsclI*-*PacI*, sono state clonate nel vettore binario pBINplus, ottenendo pEF470, pEF480, pEF490.

**Fig. 8: Rappresentazione schematica della costruzione dei vettori pEF470, pEF480, pEF490**

a) I dettagli del vettore I.V. 1.4tag sono descritti nella Fig. 3  
 TP = Peptide segnale per il cloroplasto del gene *rbcS* di *Chrysanthemum morifolium*

## 2.4 Gene $L1^V$

### 2.4.1 Costruzione dei vettori pEF50, pEF51, pEF52, pEF53, pEF54

La sequenza di interesse  $L1^V \Delta 22$  ("wild type") è stata amplificata mediante PCR con gli oligonucleotidi P16 e P17 dal vettore pGEX-L1, digerita con gli enzimi di restrizione *NcoI* - *NotI* e clonata in tutta la serie dei vettori I.V., ottenendo, come mostrato in Fig. 9, i vettori pEFs50, pEFs51, pEFs52, pEFs53, pEFs54. Questi plasmidi sono stati digeriti con gli enzimi di restrizione *Ascl* - *PacI* per rimuovere le cinque cassette di espressione, che sono state successivamente clonate nel vettore binario finale pBINplus (Fig. 9).

### 2.4.2 Costruzione dei vettori pEF55, pEF56

La costruzione di questi due vettori è schematizzata in Fig. 10.

La sequenza  $L1^V \Delta 205$  ( $L1$  mancante degli ultimi 205 aminoacidi a partire dal COOH terminale) è stata amplificata dal vettore pEFs50 usando i "primer" P16 e P18, digerita con le endonucleasi di restrizione *NcoI* - *NotI* e clonata nei vettori 1.1tag e 1.4tag della serie I.V. per l'espressione nel citosol e cloroplasto, ottenendo i plasmidi intermedi pEFs55 e pEFs56. Successivamente, le due cassette di espressione sono state digerite come frammenti *Ascl* - *PacI* e clonate nel vettore binario finale pBINplus, ottenendo pEF55 e pEF56.

## 2.5 Gene $L1^{pt}$

### 2.5.1 Costruzione dei vettori pEF57, pEF58

Per la costruzione del vettore intermedio pEF/oligo, sono stati sintetizzati due oligo complementari P19 e P20 contenenti parte della sequenza di  $L1^{pt}$ . Gli oligo, fatti reagire come descritto in § 2.7, sono stati successivamente digeriti con gli enzimi *NcoI* - *BglII* e clonati nel vettore 1.1tag, come indicato in Fig. 11.

La sequenza del gene  $L1^{pt} \Delta 22$  ("codon usage" ottimizzato per l'espressione plasmidiale), è stata digerita usando le endonucleasi di restrizione *BstEII* - *AvrII* dal plasmide pPL50 disponibile nel nostro laboratorio e clonata nel vettore pEF/oligo, ottenendo così il vettore pEFs57 (Fig. 11). Dal vettore pEFs57 il gene  $L1^{pt}$  è stato digerito come frammento *NcoI* - *BglII* e clonato nel vettore intermedio 1.4tag con il "transit peptide" per il cloroplasto, ottenendo così il vettore intermedio pEFs58. Successivamente, come indicato in Fig. 11, le cassette di espressione sono state digerite *Ascl* - *PacI* e clonate nel vettore binario pBINplus, ottenendo i plasmidi finali pEF57 e pEF58.

### 2.5.2 Costruzione dei vettori pEF570 e pEF580

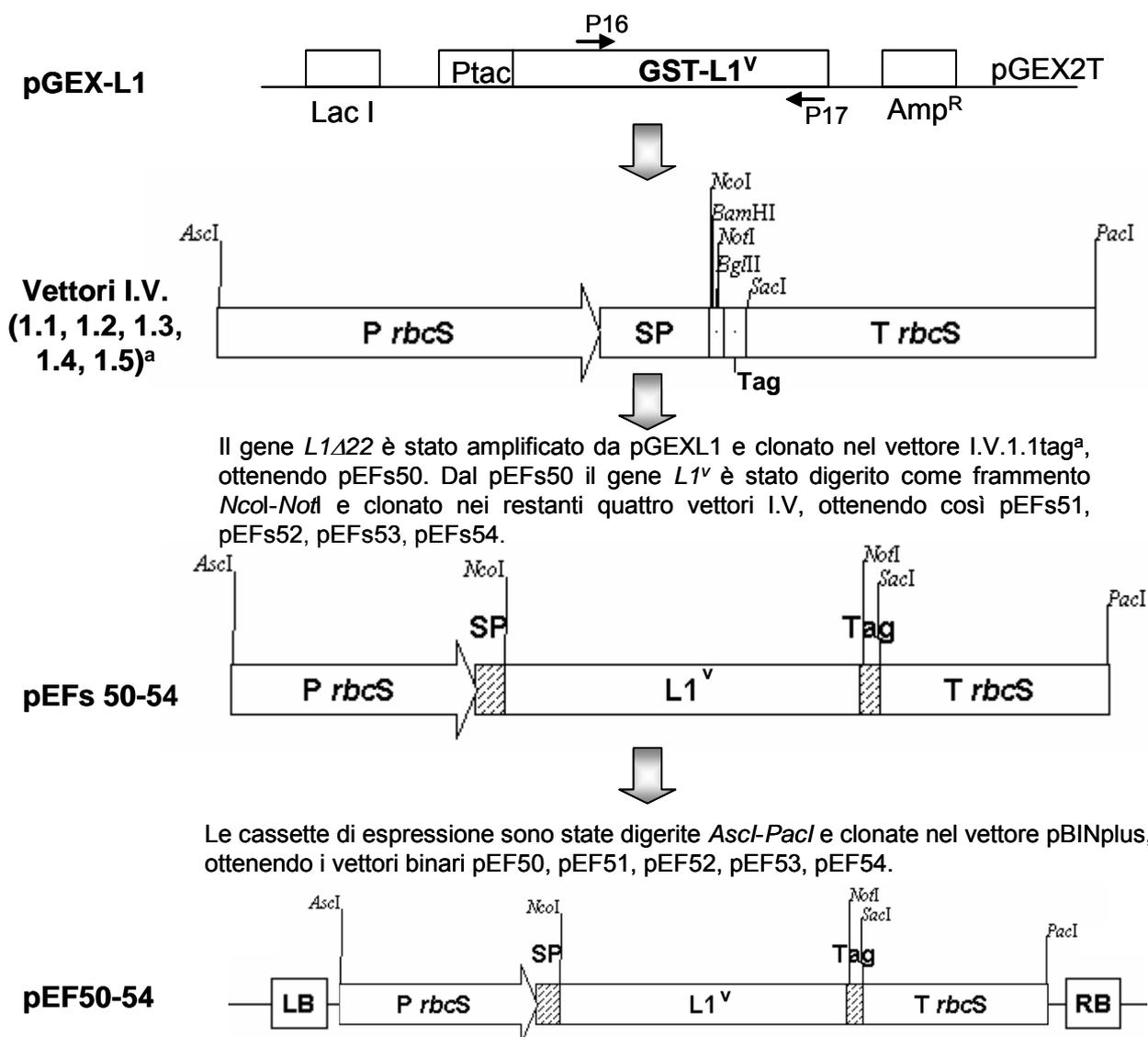
La sequenza  $L1^{pt} \Delta 205$  ( $L1^{pt}$  mancante degli ultimi 205 aminoacidi) ottenuta mediante PCR facendo legare gli oligonucleotidi P21 e P22 sul plasmide pEF57 è stata clonata come frammento *NcoI* - *BglII* nei due vettori della serie I.V. 1.1tag e 1.4tag già illustrati in Fig. 3, ottenendo i vettori pEFs570 e pEFs580 (Fig.12). Successivamente, la cassetta di espressione di questi due vettori è stata digerita come frammento *Ascl* - *PacI* e clonata nel vettore pBINplus, ottenendo i vettori finali pEF570 e pEF580 (Fig. 12).

## 2.6 Gene $L1^h$

### 2.6.1 Costruzione dei vettori pEF500 e pEF501

La sequenza di interesse  $L1^h\Delta22$  ( $L1$  con “codon usage” umanizzato), è stata digerita dal vettore pUC- $L1^h$  con gli enzimi di restrizione *Bst*II - *Avr*II e clonata nel vettore pEF/oligo, ottenendo così il vettore pEFs510 mostrato in Fig. 13.

La sequenza completa di  $L1^h$  è stata digerita dal vettore pEFs510 come frammento *Nco*I - *Bgl*II e clonata nel vettore I.V. 1.4tag, ottenendo il vettore pEFs540. La cassetta di espressione di questi due vettori è stata digerita *Asc*I - *Pac*I e clonata nel vettore binario pBINplus, ottenendo i vettori finali pEF510 e pEF540 illustrati in Fig. 13.

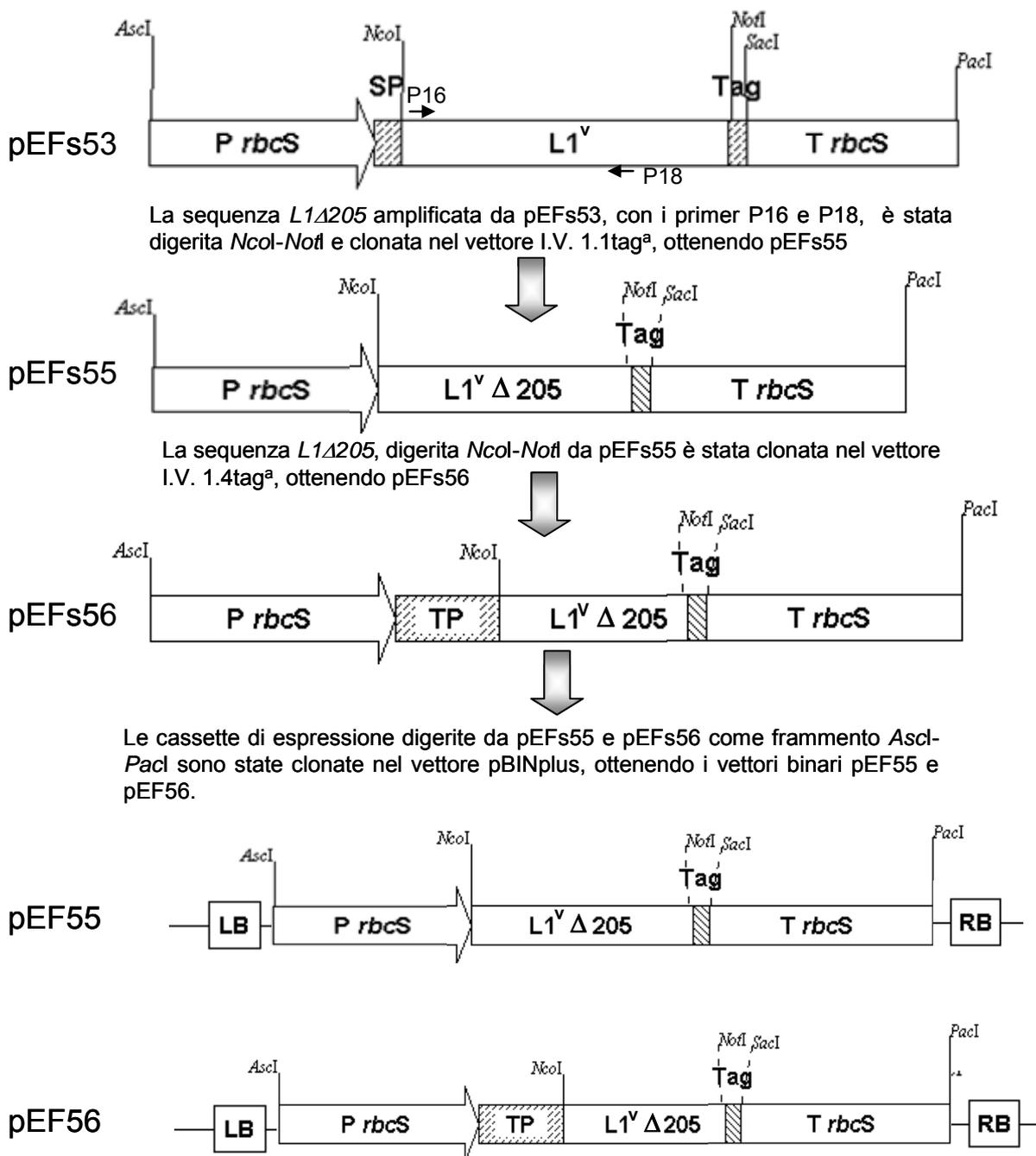


**Fig. 9: Rappresentazione schematica della costruzione dei vettori pEF50, pEF51, pEF52, pEF53, pEF54.**

<sup>a</sup>) I dettagli dei vettori I.V. sono descritti nella Fig. 3

pGEXL1 = plasmide contenente la sequenza del gene  $L1^V$  “wild type”

$L1^V$  = gene  $L1\Delta22$  “wild type” mancante del Segnale di Localizzazione Nucleare (NLS)



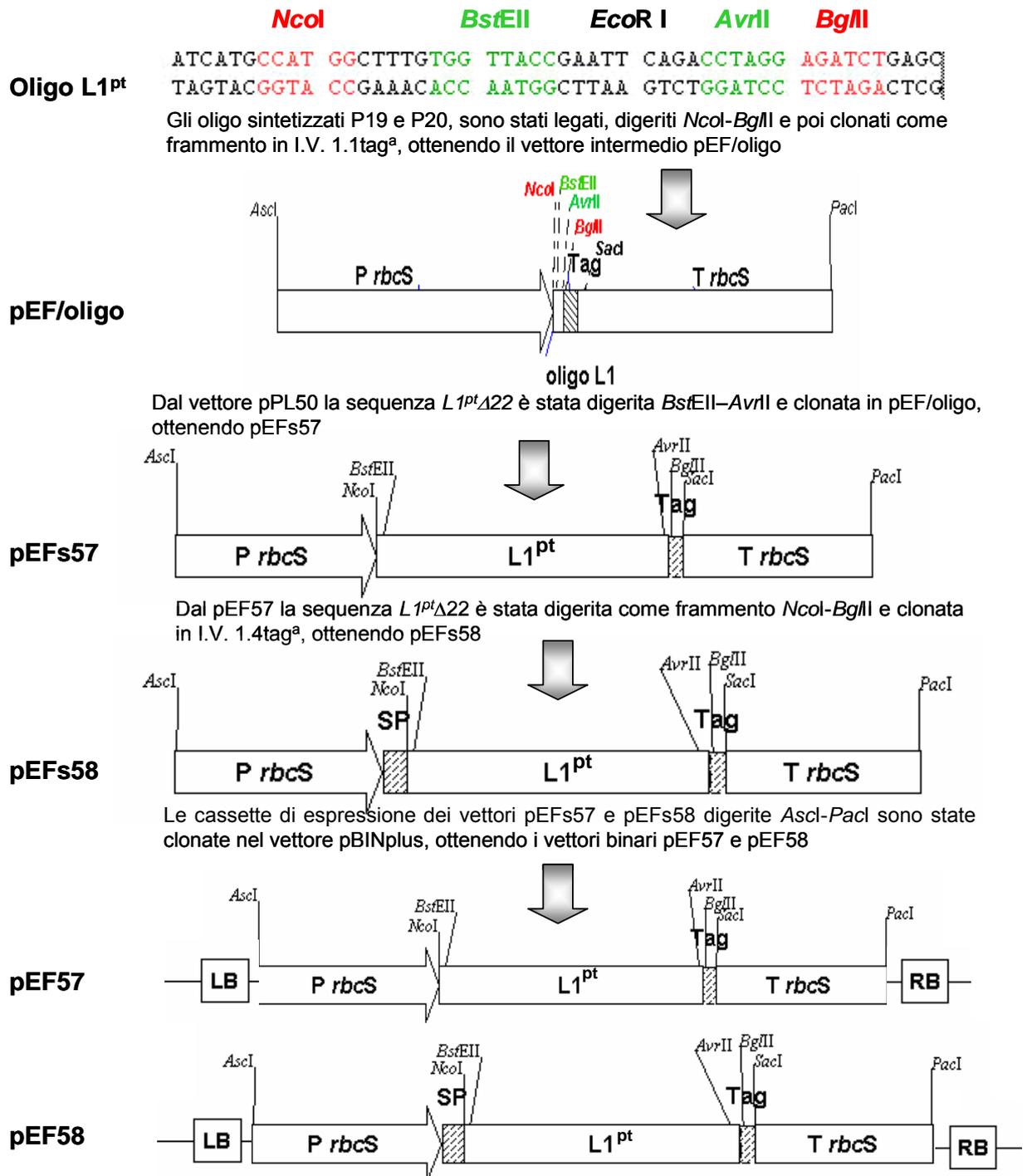
**Fig. 10: Rappresentazione schematica della costruzione dei vettori pEF55 e pEF56.**

a) I dettagli del vettore I.V. 1.4 tag sono descritti nella Fig. 3

*L1<sup>v</sup>* = gene *L1Δ22* "wild type" mancante del Segnale di Localizzazione Nucleare (NLS)

*L1<sup>v</sup>Δ205* = gene *L1* "wild type" mancante degli ultimi 205 aminoacidi

TP = Peptide segnale per il cloroplasto del gene *rbcS* del *Chrysanthemum morifolium*

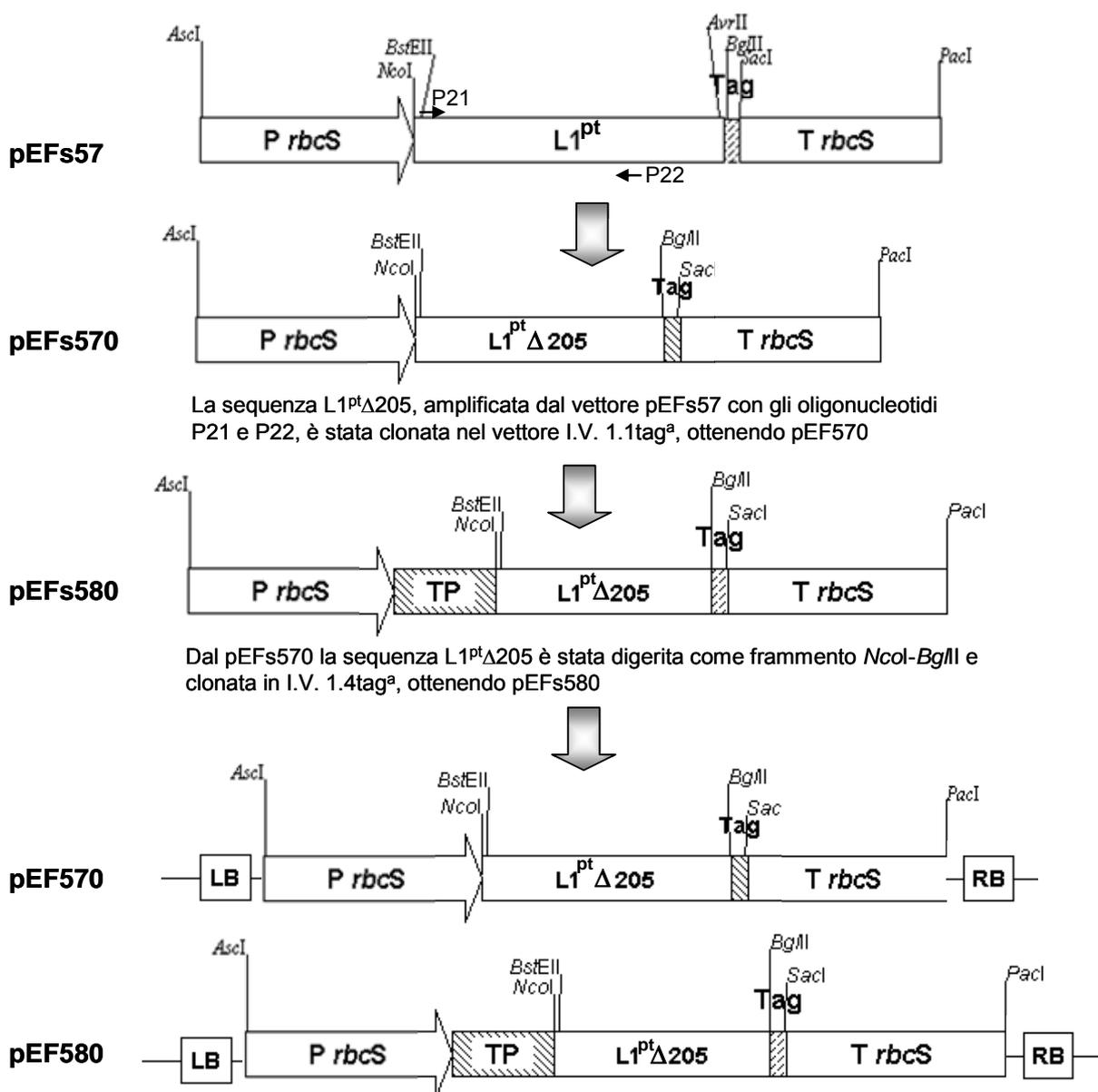


**Fig. 11: Rappresentazione schematica della costruzione dei vettori pEF57 e pEF58.**

a) I dettagli dei vettori I.V. sono descritti nella Fig.3

Oligo L1<sup>pt</sup> = oligo sintetizzato con la sequenza parziale del gene sintetico tra *NcoI*-*BstEII* e *AvrII*-*BgIII*

L1<sup>pt</sup> = gene *L1Δ22* con un "codon usage" ottimizzato per l'espressione plasmidiale.



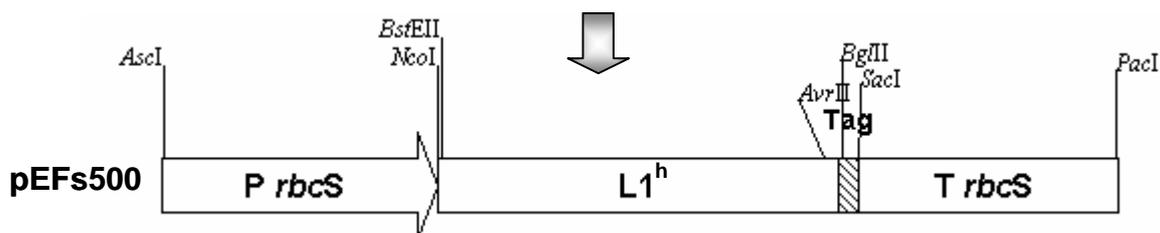
**Fig. 12: Rappresentazione schematica della costruzione dei vettori pEF570 pEF580.**

<sup>a</sup>) I dettagli dei vettori I.V. sono descritti in Fig. 3

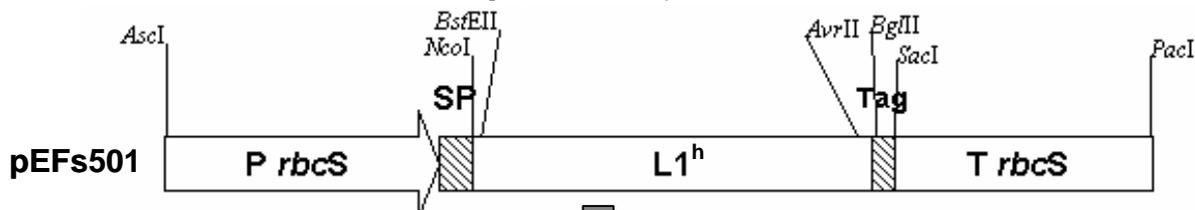
$L1^{pt}\Delta205$  = gene *L1* con un "codon usage" ottimizzato per l'espressione in cloroplasti, mancante degli ultimi 205 aminoacidi.

TP = peptide segnale per il cloroplasto del gene *rbcS* del *Chrysanthemum morifolium*

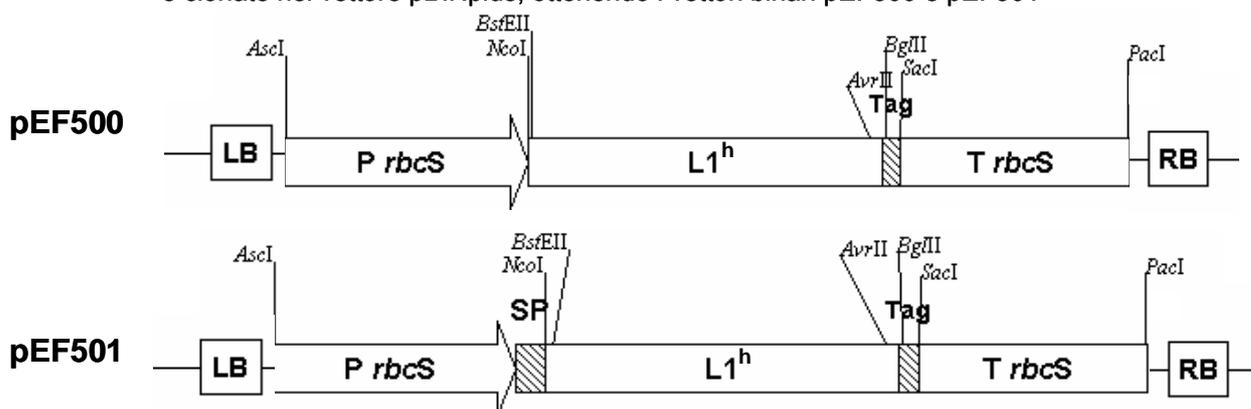
La sequenza  $L1^{\Delta 22}$  è stata digerita dal vettore  $pL1^{h(a)}$  come frammento *BstEII*-*AvrII* e clonata nel vettore *pEF/oligo*, ottenendo *pEF500*



La sequenza  $L1^h$  digerita dal vettore *pEF500* come frammento *NcoI*-*BglII* è stata clonata nel vettore *I.V.1.4tag<sup>(b)</sup>*, ottenendo *pEF501*



Le cassette di espressione dei vettori *pEF500* e *pEF501* sono state digerite *AscI*-*PacI* e clonate nel vettore *pBINplus*, ottenendo i vettori binari *pEF500* e *pEF501*



**Fig.13: Rappresentazione schematica della costruzione dei vettori pEF500 e pEF501**

a) vettore sorgente contenente la sequenza  $L1\Delta 22$  con un “codon usage” umanizzato

b) i dettagli dei plasmidi I.V. sono descritti in Fig. 3

$L1^h$  = gene  $L1\Delta 22$  con “codon usage” umanizzato

SP = peptide segnale per il cloroplasto del gene *rbcS* di *Chrysanthemum morifolium*

## 2.7 Ottenimento dei plasmidi ricombinanti

### Crescita dei batteri

#### Substrato LB (Lauria-Bertani)

bacto-triptone	10 g
bacto-estratto di lievito	5 g
NaCl	10 g
H <sub>2</sub> O milliQ (Millipore)	1 L

I batteri del ceppo di *E. coli* DH5 $\alpha$  contenenti i vettori plasmidici sono stati fatti crescere a 37 °C per tutta la notte in agitazione. Il substrato di crescita è stato addizionato, a seconda dei casi, con l'antibiotico ampicillina 100 mg/ml o kanamicina 50 mg/ml.

#### Substrato YEB

bacto-triptone	5 g
bacto-peptone	5 g
estratto di lievito	1 g
Saccarosio	5 g
MgSO <sub>4</sub>	2mM
pH	7.5
H <sub>2</sub> O milliQ (Millipore)	1 L

I batteri del ceppo *A. tumefaciens* LBA4404 contenenti i vettori plasmidici sono stati fatti crescere a 28 °C in agitazione fino all'O.D. desiderato (0,5-0,6). Il substrato di crescita è stato addizionato con l'antibiotico kanamicina 50 mg/ml.

### Digestione del DNA con enzimi di restrizione

#### Tampone di digestione universale

1M Tris, pH 7.5	1 ml
1M MgCl <sub>2</sub>	0.35 ml
1M KCl	5 ml
14.6M $\beta$ -mercaptoetanololo	10 $\mu$ l
BSA (10mg/ml)	0.5 ml
H <sub>2</sub> O	10 ml

Il DNA plasmidico e i frammenti da clonare sono stati sottoposti a digestione con differenti endonucleasi di restrizione. Le reazioni di digestione sono avvenute alla temperatura di 37 °C per 3 ore in un volume finale di 20  $\mu$ l, addizionando a ~700 ng di DNA 4  $\mu$ l del tampone di digestione universale (5X), 2 U/ $\mu$ g DNA dell'enzima di restrizione e acqua MilliQ (Millipore) fino a 20  $\mu$ l.

Per il clonaggio dei frammenti " $\Omega$ -gag-tag" e " $\Omega$ -TP-gag-tag" nei vettori finali pEF45 e pEF46, sono state effettuate delle digestioni parziali con l'enzima *Hind*III. L'enzima *Hind*III è stato diluito ad 1U/ $\mu$ l nel tampone universale di digestione. Le digestioni parziali sono avvenute per 30 min a 37°C con circa 1 $\mu$ g di vettore pEF102 e pEF202 linearizzati precedentemente con l'enzima *Sac*I.

### Isolamento e purificazione degli inserti

#### Tampone TAE 1X

40 mM Tris-acetato
10 mM EDTA

Per i vettori pEF40, pEF41, pEF42, pEF43, pEF44, pEF50, pEF55, pEF56, pEF510, pEF570, pEF580 il gene d'interesse è stato amplificato mediante Reazione a Catena della Polimerasi (PCR).

L'amplificazione degli inserti da clonare è stata condotta utilizzando l'enzima Taq polimerasi "AccuPrime" (Invitrogen) con attività "proof reading" di correzione delle bozze.

Per ogni reazione di amplificazione sono stati impiegati 50 µl di una miscela di reazione costituita da:

DNA	50 ng
Buffer 10x	5 µl
Oligo "Forward" 10 µM	1,5 µl
Oligo "Reverse" 10 µM	1.5 µl
Taq polimerasi <i>Pfu</i> (2,5 U/ µl)	0,4 µl
H <sub>2</sub> O sterile fino a 50 µl	

Per la reazione di amplificazione è stato utilizzato il seguente ciclo:

denaturazione	a 92°C per 3 minuti;	1 ciclo
denaturazione	a 92°C per 15 secondi;	30 cicli
"annealing" (variabile in funzione della T <sub>m</sub> degli oligo) per 30 sec;		30 cicli
estensione	a 68°C per 1 minuto per ogni Kb;	30 cicli

Al termine dell'amplificazione, i prodotti sono stati analizzati su di un gel di agarosio all'1% in tampone 1X TAE, in presenza di marcatori di peso molecolare noto.

I prodotti di amplificazione sono stati purificati mediante l'uso del kit "PCR purification" (Qiagen). Infine, la qualità e la quantità dei frammenti così ottenuti è stata controllata come descritto precedentemente.

Per la costruzione dei vettori pEF51, pEF52, pEF53, pEF54, pEF57, pEF58, pEF61, pEF63, pEF63, pEF64, pEF540 gli inserti sono stati ottenuti per digestione a partire da plasmidi intermedi progenitori.

I prodotti di restrizione così ottenuti sono stati separati su gel di agarosio all'1%, la banda relativa all'inserto è stata tagliata dal gel con uno scalpello e il DNA è stato eluito servendosi del "GeneClean kit" (Q-BIOgene).

Nel caso del plasmide pEF/oligo, l'inserto era costituito da due oligo complementari fatti reagire secondo la seguente procedura: gli oligo sono stati risospesi alla concentrazione di 1µg/µl, messi a reagire in tubi eppendorf nella seguente miscela:

Oligo "Forward"	10 µl
Oligo "Reverse"	10 µl
Buffer	1 µl (100 mM TrisHCl pH 7.5, 1M NaCl, 10 mM EDTA)
H <sub>2</sub> O	9 µl

e incubati a 85 °C per 10 min. Il bagnetto termico è stato poi spento, lasciando la reazione per tutta la notte in modo da far scendere la temperatura lentamente e permettere l'annealing alla temperatura ottimale per gli oligo. Il prodotto della reazione è stato digerito in 50 µl finali e gli oligo sono stati purificati servendosi del "Nucleotide Removal" kit (Qiagen).

### **Subclonaggio dell'inserto in plasmidi**

Gli inserti ottenuti dopo digestione e purificazione, sono stati clonati in appropriati vettori di espressione. La procedura seguita per il clonaggio ha previsto:  
l'inserzione stabile del frammento di DNA in un vettore plasmidico;  
la trasformazione di cellule batteriche opportunamente preparate;  
la crescita dei batteri trasformati su terreno selettivo;  
il controllo del subclonaggio.

### **Ligazione dell'inserto in vettori plasmidici**

10X ligasi buffer:

300 mM Tris-HCl pH 8

100 mM MgCl<sub>2</sub>

100 mM DTT

10 mM ATP

Gli inserti sono stati incubati con i vettori opportunamente predigeriti e la DNA ligasi del batteriofago T4 (Roche). Le reazioni di ligazione sono state effettuate utilizzando una concentrazione di vettore pari a circa 25 – 50 ng e con diversi rapporti tra il vettore e l'inserto, così come raccomandato dalla casa.

Le reazioni di ligazione sono state incubate generalmente a temperatura ambiente per 15 minuti secondo il protocollo suggerito dalla casa (Roche).

### **Trasformazione delle cellule competenti mediante "shock" termico**

Le reazioni di ligazione sono state aggiunte a 100 µl di cellule competenti precedentemente preparate seguendo il protocollo riportato da (Sambrook, 2001). Il DNA e le cellule batteriche, dopo 30 minuti di incubazione in ghiaccio, sono stati sottoposti a "shock" termico per due minuti a 42 °C. Successivamente, i batteri sono stati incubati a 37 °C per un'ora e 30 minuti in agitazione. I batteri sono stati piastrati su terreno LB, contenente 100 mg/l di ampicillina, o 50 mg/l di kanamicina. Le piastre sono state incubate a 37 °C per tutta la notte. Le colonie contenenti i frammenti d'interesse (ricombinanti) sono state analizzate mediante analisi PCR con gli oligonucleotidi P23 e P24. Le colonie positive all'analisi PCR sono state propagate mediante coltura "overnight (o/n)" in substrato LB addizionato con l'antibiotico appropriato, per l'estrazione e l'analisi del DNA plasmidico in esse contenuto; i prodotti di restrizione sono stati visualizzati su gel d'agarosio allo 0,8% in tampone TAE 1X. Le dimensioni dei frammenti ottenuti sono state stimate confrontando la migrazione delle bande con un marcatore di peso molecolare noto. Quando il frammento clonato è stato amplificato per PCR, si è proceduto con la determinazione della sequenza nucleotidica del DNA plasmidico. Quest'ultima è stata eseguita presso i laboratori dell'ENEA di Roma, con metodo Sanger *et al.* (1977) automatizzato.

## Trasformazione delle cellule competenti di *A. tumefaciens* mediante “Freeze and Thaw”.

### KL

lattosio	10g/L
estratto di lievito	1g/L
agar	12g/L

### Reagente di Benedict

citrato sodico	17.3 g/L
carbonato bisodico	10.0g/L
solfo rameico	1.68g
H <sub>2</sub> O milliQ (Millipore) fino a	100 mL

Tutti i plasmidi binari ricombinanti ottenuti e controllati nel ceppo di *E. coli*, sono stati successivamente trasferiti nel ceppo di *A. tumefaciens* LBA4404. cento µl di cellule competenti di *A. tumefaciens* sono state aggiunte a circa un µg di DNA plasmidico. Le cellule sono state congelate immediatamente in N<sub>2</sub> liquido e scongelate a 37 °C per 5 min. Successivamente ai batteri è stato aggiunto un ml di YEB ed essi sono stati incubati in agitazione per 2-3 ore a 28 °C. I batteri sono stati piastrati su terreno selettivo LB contenente gli antibiotici 50 mg/L di rifampicina, 50 mg/L di kanamicina, e 300 mg/L di streptomycina.

Le piastre sono state incubate per 2 giorni a 28°C e le colonie ottenute sono state analizzate mediante analisi PCR con gli oligonucleotidi P23 e P24 e quelle risultate positive sono state fatte crescere su substrato KL e sottoposte a test colorimetrico mediante l'uso del reattivo di Benedict per confermare la presenza di *A. tumefaciens* (Hooykaas e Mozo, 1994). Le colonie che in presenza del reattivo sono virate da incolore a giallo intenso, indicando la presenza di *A. tumefaciens*, sono state utilizzate per l'agroinfiltrazione seguendo il protocollo descritto nel § 2.8

## 2.8 Agroinfiltrazione di *N. tabacum* e *N. benthamiana*

### **Substrati per l'induzione di *A. tumefaciens*.**

MES	10 mM
Acetosiringone	200 µM
pH 5.5	

### **Agroinfiltrazione delle piante di tabacco**

Per l'infiltrazione degli espianti fogliari di tabacco di *N. tabacum* e *N. benthamiana* è stato utilizzato il protocollo di agroinfiltrazione descritto da Voinnet *et al.*, 2003). *A. tumefaciens*, contenente il vettore ricombinante, è stato fatto crescere in terreno liquido YEB addizionato con 50 mg/L di kanamicina fino ad un'O.D.<sub>600</sub> di 0,5-0,6 e risospeso nel mezzo di induzione MES addizionato con l'acetosiringone fino ad avere un'O.D.<sub>600</sub> pari a 0,3. La coltura, lasciata per almeno un'ora a temperatura ambiente, è stata utilizzata per l'infiltrazione delle aree fogliari. Per entrambe le specie sono state infiltrate foglie espanse di piante cresciute in camera di crescita in condizioni controllate per circa 3 settimane. Il punto di infiltrazione, localizzato lontano dalle nervature centrali, è stato “graffiato” leggermente con l'aiuto di un ago. L'infiltrazione è avvenuta in questo punto,

mediante l'uso di una siringa da un ml (Fig. 14). I bordi delle aree infiltrate sono stati marcati dopo l'infiltrazione. Per ogni costrutto sono state agroinfiltrate in media due foglie. I tessuti fogliari agroinfiltrati prelevati dopo 3 giorni dall'infiltrazione sono stati congelati in N<sub>2</sub> liquido e conservati a -80 °C per le analisi molecolari, oppure utilizzati immediatamente per il saggio istochimico *gus*.



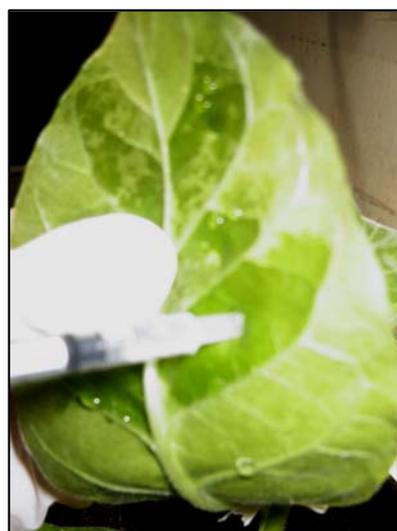
**Pianta di *N. benthamiana***



**Pianta di *N. tabacum***



**Foglia di *N. benthamiana* agroinfiltrata**



**Foglia di *N. tabacum* agroinfiltrata**

**Fig. 14: Piante di *N. benthamiana* e *N. tabacum* e relative foglie agroinfiltrate**

## 2.9 Verifica dell'attività biologica del gene *gus*

La verifica della funzionalità del gene *gus* è stata verificata mediante saggio istochimico (Jefferson, 1987) dopo tre giorni dall'infiltrazione dei tessuti. Il saggio ha previsto la preparazione di una soluzione per il saggio mescolando il tampone fosfato 50 mM a pH 7,0, 10% di TritonX100 (Sigma) e la soluzione X-GLUC preparata sciogliendo in un ml di dimetilformamide 25 mg/ml di X-GLUC. Le aree agroinfiltrate sono state prelevate e lesionate con i bisturi, in piastre Petri, contenente ciascuna 5-7 ml della soluzione X-GLUC. Le piastre contenenti il substrato XGLUC e le aree agroinfiltrate sono state poste con il coperchio aperto in una camera da vuoto. Il vuoto è stato creato tramite una pompa e le piastre sono state lasciate sottovuoto per circa 2-3 minuti. Infine i tessuti sono stati incubati a 37 °C per tutta la notte. Il giorno successivo la soluzione di XGLUC è stata allontanata e i tessuti sono stati decolorati aggiungendo etanolo assoluto in modo da rimuovere la clorofilla e rendere più evidente il precipitato blu. Come controllo positivo sono state saggiate sezioni di foglie di *Nicotiana tabacum* transgenica che esprimono costitutivamente il gene *gus-int*.

## 2.10 Estrazione di RNA totale

Prima dell'estrazione sia le soluzioni che la vetreria utilizzate sono state trattate con dietil pirocarbonato (DEPC) allo 0,1%, per inattivare eventuali RNAsi cellulari contaminanti, e poi sterilizzate in autoclave, a 120 °C per 30 min, per disattivare il DEPC.

L'RNA è stato estratto usando il KIT "Nucleo-spin RNA plant" (Macherey-Nagel), seguendo il protocollo descritto dalla casa. La concentrazione dei campioni di RNA estratti è stata determinata mediante lettura allo spettrofotometro.

La purezza dell'estrazione (assenza di eventuali sostanze contaminanti, quali proteine, polisaccaridi e fenolo), è stata stimata dalla misura del rapporto Assorbanza O.D.<sub>260</sub>/Assorbanza O.D.<sub>280</sub>. Tale rapporto deve avere un valore compreso tra 1,7 e 2,0 per una buona preparazione di RNA.

### Tampone MOPS 10X

Acido N-morfolino-propansulfonilico pH 7	0,2 M
Sodio acetato	80 mM
EDTA	10 mM

### Tampone denaturante

Formammide	750 µl
10X MOPS	120 µl
37% formaldeide	250 µl
10 µg/µl bromuro di etidio	1 µl

### Gel di agarosio/formaldeide (60 ml)

Agarosio 1,2%	0,75 g
1X MOPS	6 ml
Formaldeide	3,5 ml
H <sub>2</sub> O sterile	50,5 ml

## 2.11 Estrazione di proteine solubili e insolubili

### Soluzione High Salt.

PBS	1X
NaCl	0.5 M
DTT	5 mM
EDTA	2 mM
PMSF	1 mM

### Soluzione di lavaggio:

Hepes/KOH pH 7,5	50 mM
Potassio Acetato	10 mM
Magnesio Acetato	5 mM
EDTA	1 mM
PMSF	2 mM

### Soluzione di solubilizzazione per proteine insolubili:

Hepes/KOH pH 7,5	50 mM
LDS	2%

L'estrazione delle proteine totali dalle foglie è avvenuta secondo il seguente schema:

sono state unite circa due, tre aree agroinfiltrate e ad esse sono state aggiunte 300-400 µl di soluzione High Salt;

la miscela è stata macerata in N<sub>2</sub> liquido servendosi di mortai e pestelli, centrifugata a 10000xg per 5 minuti a 4°C;

il surnatante, contenente le proteine solubili, è stato trasferito in un nuovo tubo e conservato a -80 °C.

Per le proteine insolubili il pellet ottenuto è stato lavato quattro-cinque volte con tampone di lavaggio, centrifugato più volte per 5 minuti alla massima velocità, e solubilizzato con circa 100-200 µl di tampone di solubilizzazione a 100 °C per 10 minuti. Le proteine sono state centrifugate per 5 min alla massima velocità, il sovrnatante è stato trasferito in un nuovo tubo e conservato a -80 °C. Per la determinazione quantitativa delle proteine è stato seguito il metodo di Bradford (1976).

## 2.12 Analisi Molecolari

### **Analisi RT-PCR**

L'RNA è stato sottoposto ad analisi RT-PCR mediante l'uso del KIT "Superscript II one step" (Invitrogen).

Per ogni reazione di amplificazione sono stati impiegati 25 µl di una miscela di reazione costituita da:

RNA	100 ng
Buffer MIX 2x	12,5 µl
Oligo "Forward" 10 µM	0,5 µl
Oligo "Reverse" 10 µM	0,5 µl
Taq polimerasi <i>Superscript II</i>	0,5 µl
H <sub>2</sub> O sterile fino a 25 µl	

Per la reazione di amplificazione è stato utilizzato il seguente ciclo:

reverse	a 48°C per 20 minuti;	1 ciclo
denaturazione	a 94°C per 3 minuti	1 ciclo
denaturazione	a 94°C per 15 sec	35 cicli
“annealing”	(variabile in funzione della T <sub>m</sub> degli oligo) per 30 sec;	35 cicli
estensione	a 72°C per 1 minuto per ogni Kb;	35 cicli

Al termine dell'amplificazione, i prodotti sono stati analizzati su di un gel di agarosio all'1% in tampone 1X TAE, in presenza di marcatori di peso molecolare noto.

## Analisi Western

### Separazione elettroforetica delle proteine

Soluzione stock acrilammide/bis

Acrilammide:Bisacrilammide (37:1)

Tampone dello “stacking” gel

Tris 1,5 M

SDS 0,4%

pH 8,8

Tampone del gel di separazione

Tris 0,5 M

SDS 0.4%

pH 6,8

Ammonio persolfato 10%

Tampone corsa elettroforetica 10X

SDS 1%

Tris 0,25 M

Glicina 1.92 M

Tampone di trasferimento

Tris-HCl 50 mM

Glicina 40 mM

Metanolo 10%

Le proteine estratte e quantificate sono state sottoposte a separazione elettroforetica in presenza di SDS, su gel di acrilammide/bis (12 %), come descritto da Laemmli (Laemmli, 1970), utilizzando il sistema elettroforetico “mini protean” della Biorad. Prima della corsa, i campioni sono stati bolliti per 5 min e in ogni pozzetto del gel sono state caricate uguali quantità di proteine (40 µg). La separazione elettroforetica è avvenuta a 80 V fino allo “stacking gel” e 150 V per circa 1,5 ore circa nel gel di separazione. Per ogni elettroforesi sono stati caricati contemporaneamente due gel. Uno dei due è stato sottoposto a colorazione con blue Coomassie per controllare la qualità dell'estrazione e della separazione

elettroforetica, l'altro è stato sottoposto a trasferimento delle proteine su filtro di nitrocellulosa.

### **Trasferimento su filtro di nitrocellulosa**

Le proteine totali sono state trasferite su filtro di nitrocellulosa (ECL-Amersham), utilizzando il sistema "mini trans blotter" della Biorad. Il trasferimento è avvenuto in presenza del tampone di trasferimento per 1 ora a 100 V costanti.

### **Immunolocalizzazione delle proteine con anticorpi**

#### PBS 10X

KCl	2g/l;
NaCl	80g/l;
Na <sub>2</sub> HPO <sub>4</sub>	14,4 g/l;
KH <sub>2</sub> PO <sub>4</sub>	2,4 g/l
pH 7,5	

#### TPBS:

PBS	1X
Tween 20	0,1%

I filtri sono stati trattati in una soluzione bloccante costituita da PBS 1X + BSA 3% per tutta la notte. Successivamente, essi sono stati lavati per una volta in una soluzione TPBS. I filtri sono stati poi incubati per due ore a temperatura ambiente, in agitazione, nella soluzione bloccante contenente l'anticorpo primario. Per le diverse analisi sono stati utilizzati un anticorpo anti - istidine (Tetra anti-His antibody – Qiagen) diluito 1:2000, un anticorpo policlonale anti-p24 (ARP432) o anti-p17 (ARP431) forniti da NIBSC (Centralised Facility for AIDS Reagent – EU programme EVA) diluiti 1:10000 e un anticorpo policlonale anti - c-myc (polyclonal c-myc antibody – SantaCruz) diluito 1:500. Dopo 3 lavaggi di 10 minuti in TPBS, il filtro è stato incubato per un'ora a temperatura ambiente in PBS1X contenente gli anticorpi secondari anti-IgG coniugati con la perossidasi: per i Western anti-istidine è stato usato un anticorpo "anti-mouse" (Sigma-Aldrich), diluito 1:20000 e per quelli anti-p24, anti-p17 e anti - c-myc un secondario "anti-rabbit" (Sigma-Aldrich) diluito 1:60000 quando abbinato all'anticorpo primario anti-p24 e 1:30000 quando abbinato agli anticorpi primari anti-p17 e anti-cmyc. Prima della rivelazione i filtri sono stati lavati 3 volte ogni 10 minuti in TPBS. Per la rivelazione è stato seguito il metodo chemiluminescente, utilizzando il kit "ECLplus Detection Reagent" (Amersham). I filtri lasciati a contatto con la soluzione di sviluppo per 5 min a temperatura ambiente sono stati esposti con un film autoradiografico (Amersham) per 15 minuti a temperatura ambiente.

### 3 RISULTATI

#### 3.1 Vettori con il gene reporter *gus-int* e ottimizzazione della tecnica di agroinfiltrazione

Al fine di verificare il corretto funzionamento dei vettori costruiti in questo lavoro, e per ottimizzare la tecnica dell'agroinfiltrazione, sono stati costruiti cinque plasmidi ricombinanti con il gene reporter *gus* (pEF60, 61, 62, 63, 64), derivati dai vettori commerciali ImpactVector. La sequenza codificante di questo gene, interrotta da un introne (*PIV2* del gene *ST-LS1*), consente di discriminare l'espressione della proteina in batteri da quella in pianta. Questi vettori, descritti in Fig. 3, presentano il promotore luce-inducibile e la 5'UTR del gene *rbcS* di crisantemo e si differenziano per la presenza di un diverso peptide segnale per l'indirizzamento della proteina in vari compartimenti cellulari, come citoplasma, apoplasto, reticolo endoplasmatico, cloroplasto e mitocondrio. In questo lavoro di tesi, in tutti i casi, il gene d'interesse è stato sempre fuso con due tag (c-myc e 6His) al COOH terminale, per facilitare la rilevazione e la purificazione della proteina esogena. In Tabella 5 sono riportati i vettori costruiti con il gene *gus-int*.

Tutti i vettori controllo ottenuti sono stati inoculati in foglie espanse di *N. tabacum* e *N. benthamiana* cresciute in camera di crescita in condizioni controllate per circa tre settimane, seguendo il protocollo descritto nel § 2.8. Per ottimizzare la tecnica dell'agroinfiltrazione, inizialmente i prelievi delle aree agroinfiltrate sono stati effettuati a distanza di 3, 5 e 7 giorni dall'inoculo. I risultati migliori, sulla base del saggio istochimico Gus, si sono osservati su espianti analizzati dopo 3 giorni dall'infiltrazione.

Per confermare la corretta trascrizione del gene è stata usata la tecnica della RT-PCR semiquantitativa. L'RNA estratto dai diversi tessuti agroinfiltrati con i vettori contenenti il gene reporter *gus-int* è stato analizzato utilizzando la coppia di primers P25 e P26 (Fig. 15a). Con tutti i vettori è stata ottenuta una banda di amplificazione delle dimensioni attese di circa 350 bp, sia in *N. tabacum* che in *N. benthamiana*, assente nei controlli non agroinfiltrati. I livelli di espressione del gene *gus*, standardizzati con l'espressione del gene costitutivo *18S*, sono risultati uguali in tutti i vettori.

Per verificare il livello di espressione della proteina Gus dopo l'infiltrazione, sono state effettuate analisi western sulle proteine estratte dai tessuti di *N. tabacum* e *N. benthamiana* agroinfiltrate con i cinque vettori contenenti il gene reporter *gus*. Le proteine totali, separate su un gel di acrilammide e trasferite su membrana di nitrocellulosa, sono state incubate con l'anticorpo primario anti - His oppure anti - c-myc. Con tutti i vettori si è evidenziato la presenza della proteina ricombinante, ma con differenti livelli di espressione nei vari compartimenti (Fig. 16). L'analisi western ha mostrato un maggior accumulo della proteina quando è stata espressa nel citoplasma e nel cloroplasto, mentre, come mostrato in Fig. 16a, l'apoplasto, il mitocondrio ed il reticolo sono risultati compartimenti meno adatti per l'espressione della proteina di controllo Gus.

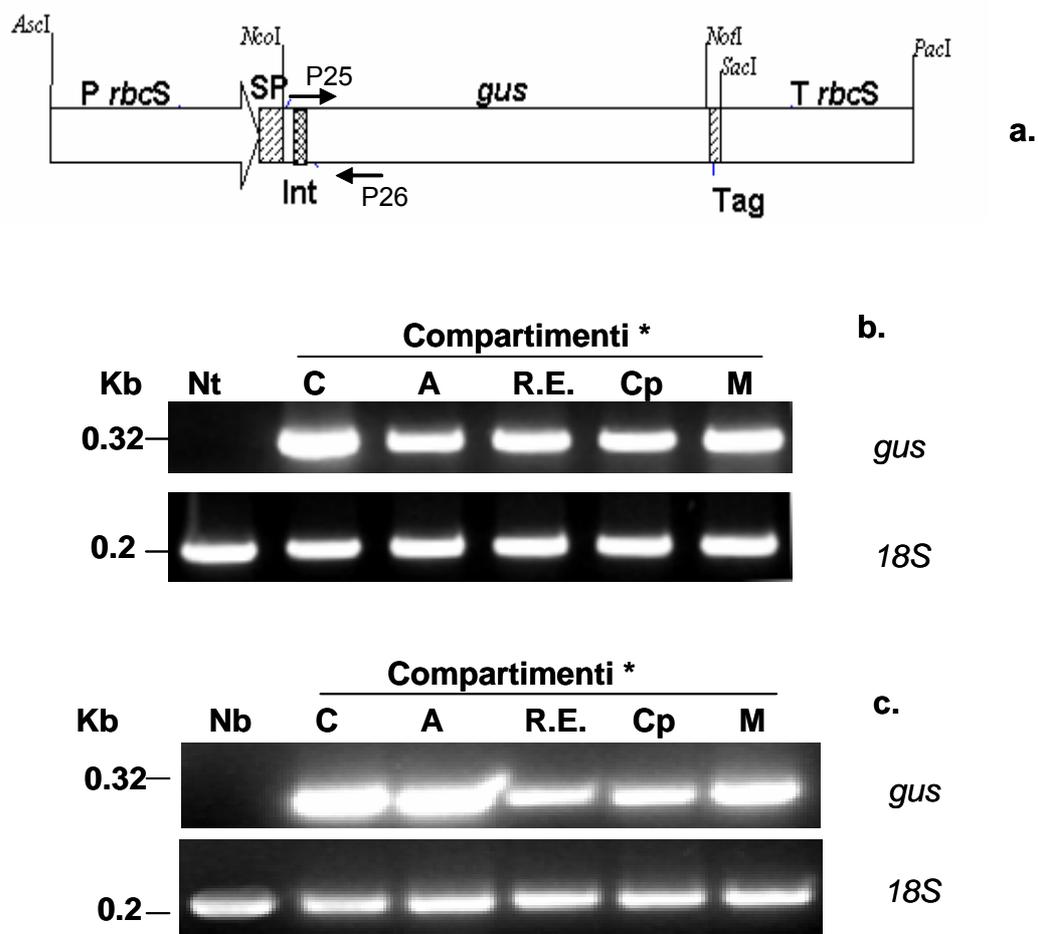
Per verificare la funzionalità del gene *gus*, i tessuti agroinfiltrati con i cinque vettori, sono stati sottoposti al saggio istochimico Gus, seguendo il protocollo descritto nel § 2.9. Come mostrato in Fig. 16c i tessuti analizzati hanno mostrato, come atteso, una diffusa colorazione blu, uniforme e indipendente dal vettore utilizzato, confermando l'espressione del gene *gus*.

Infine, in Tabella 6, è riportato un confronto tra i livelli di espressione del gene *gus* sia in *N. tabacum* che in *N. benthamiana*, sulla base dei risultati ottenuti dalle analisi western.

**Tabella 5:** Vettori con il gene reporter *gus-int*

Vettori	Compartimento subcellulare	Transgene
pEF60	Citoplasma	<i>gus-int</i> <sup>a</sup>
pEF61	Apoplasto	<i>gus-int</i>
pEF62	Reticolo endoplasmico	<i>gus-int</i>
pEF63	Cloroplasto	<i>gus-int</i>
pEF64	Mitocondrio	<i>gus-int</i>

a) gene reporter *gus* con l'introne modificato PIV2 del gene *ST-LS1* Vancanneyt *et al.*, 1990). Tutti i vettori presentano il promotore, 5'UTR e terminatore del gene *rbcS* di crisantemo (Outchkourov *et al.*, 2003) e due "tag" al COOH terminale (c-myc e 6His).



**Fig. 15: Analisi RT-PCR dei tessuti agroinfiltrati con i vettori pEF60, pEF61, pEF62, pEF63 e pEF64 (*gus-int*) per i compartimenti subcellulari indicati.**

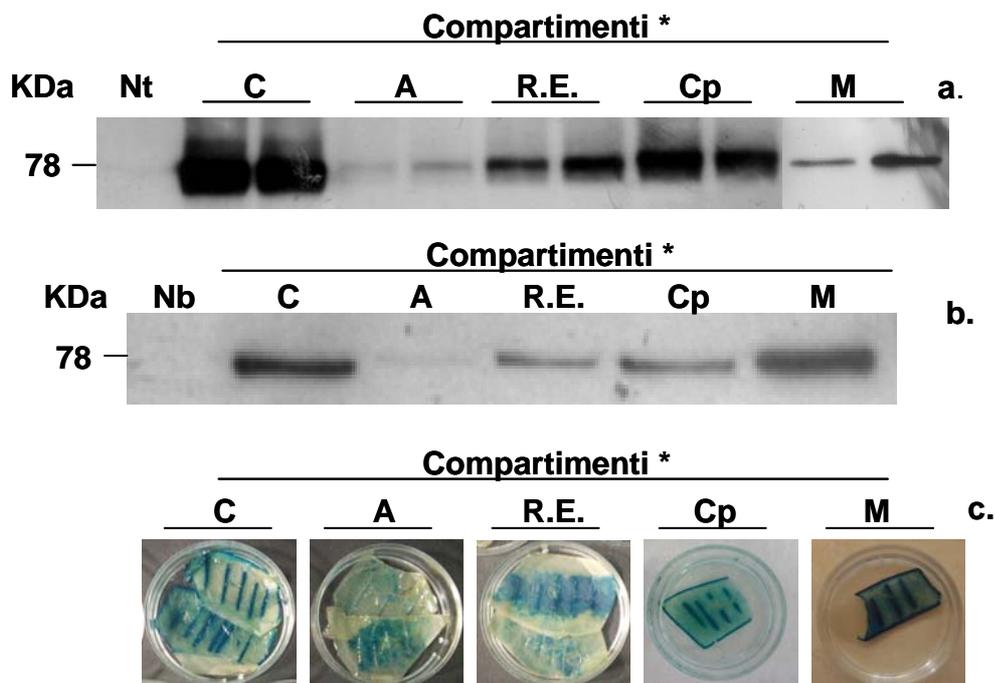
a. Rappresentazione schematica dei vettori con il gene *gus-int*

b. Analisi RT-PCR su tessuti di *N. tabacum* con la coppia di primer P25 e P26.

c. Analisi RT-PCR su tessuti di *N. benthamiana* con la coppia di primer P25 e P26.

Nt = *N. tabacum* non agroinfiltrata; Nb = *N. benthamiana* non agroinfiltrata  
18S = gene ribosomiale utilizzato come controllo

\* C = citosol, A = apoplasto, R.E. = reticolo endoplasmico,  
Cp = cloroplasto, M = mitocondrio



**Fig. 16: Analisi Western e saggio istochimico sui tessuti agroinfiltrati con i vettori pEF60, pEF61, pEF62, pEF63 e pEF64 per i compartimenti subcellulari indicati.**

- a. Analisi Western con anticorpi anti - his su tessuti di *N. tabacum*  
 b. Analisi Western con l'anticorpi anti - his su tessuti di *N. benthamiana*  
 Nt = tessuti di *N. tabacum* non agroinfiltrati  
 Nb = tessuti di *N. benthamiana* non agroinfiltrati  
 c. Saggio istochimico Gus sui tessuti di *N. tabacum* agroinfiltrati  
 \* C = citosol, A = apoplasto, R.E. = reticolo endoplasmatico,  
 Cp = cloroplasto, M = mitocondrio

**Tabella 6:** Risultati dell'espressione transiente del gene *gus* in *N. tabacum* (a) e in *N. benthamiana* (b) in base ai risultati mostrati dall'analisi Western.

Transgene	Compartimento	Anticorpi	
		c-myc	his
<i>gus-int</i>	Citoplasma	++	++
	Apoplasto	+	+
	Reticolo	+	+
	Cloroplasto	++	++
	Mitocondrio	+	+
<i>gus-int</i>	Citoplasma	++	++
	Apoplasto	-	+
	Reticolo	+	+
	Cloroplasto	+	+
	Mitocondrio	+	++

++ = alti livelli di espressione;  
 + = bassi livelli di espressione  
 - = nessuna espressione

### 3.2 Gene gag (HIV-1)

I vettori pEF40, 41, 42, 43, 44, costruiti con il gene d'interesse *gag*, derivano dai plasmidi *ImpactVector*, così come illustrato nel § 2.3. In questi vettori il gene d'interesse è sotto la guida del promotore e della 5'UTR del gene *rbcS* di crisantemo, e di quattro diversi peptidi segnale, per indirizzare la proteina nei differenti compartimenti cellulari quali apoplasto, reticolo endoplasmatico, cloroplasto e mitocondrio. Inoltre il gene *gag* è stato clonato anche nel vettore senza nessun peptide segnale per l'accumulo della proteina nel citoplasma. In tutti i vettori, il gene *gag*, è stato fuso al COOH terminale con due "tag", (c-myc e 6His) per facilitare il rilevamento e la purificazione della proteina esogena.

Per confrontare i livelli di espressione tra vari promotori e 5'UTR sono stati costruiti altri due vettori contenenti il gene *gag*, pEF45 e pEF46. Questi due vettori derivano dal plasmide binario pKYLX71 che presenta il forte promotore costitutivo 35S<sup>2</sup> del virus del mosaico del cavolfiore e 5'UTR  $\Omega$  del TMV. Nel vettore pEF46 inoltre è stato clonato anche il peptide segnale del gene *rbcS* di *Pisum sativum* (Anderson e Smith, 1986) per l'accumulo della proteina nel cloroplasto. Anche in questi due ultimi vettori il gene *gag* è stato fuso alle stesse "tag" presenti nei vettori *ImpactVector*.

Inoltre, sono stati costruiti altri vettori per l'espressione transiente nucleare, contenenti le sequenze delle subunità del gene *gag*: *p17*, *p24*, e  $\Delta p17$  (tutti i domini eccetto il p17).

Come già descritto nel § 1.3, il gene *gag* è composto da quattro diversi domini, p17 (MA), p24 (CA), p7 (NC) e p6 (dominio terminale). I geni che codificano per l'espressione della *p17*, della *p24* e del dominio  $\Delta p17$ , sono stati clonati nei vettori nucleari *ImpactVector* e i plasmidi ottenuti sono stati denominati pEF47, pEF48, pEF49, pEF470, pEF480, pEF490. In particolare, per utilizzare il sito *NcoI* presente nei vettori di clonaggio, sia per il gene *p24* che per  $\Delta p17$ , è stato necessario cambiare la quarta base della sequenza nucleotidica in modo da ricreare il sito *NcoI* (CCATGG) con l'ATG del gene d'interesse. Questa modifica ha prodotto un cambio di aminoacido da prolina ad alanina.

Per le sequenze relative alle subunità del gene *gag*, sono stati utilizzati solo due tipologie di vettori, quelli senza peptide segnale per l'indirizzamento della proteina nel citoplasma, e quelli con il "transit peptide" per l'accumulo della proteina nei plastidi.

In tutti i casi, i geni sono sotto la guida dello stesso promotore e 5'UTR *rbcS* e presentano al COOH terminale due "tag" (c-myc e 6His).

Tutti i vettori costruiti per il gene *gag* sono stati inseriti in *A. tumefaciens*, ceppo LBA4404. Le colonie risultate positive alla reazione di PCR e al test colorimetrico sono state utilizzate per l'infiltrazione degli espianti fogliari di tabacco sia di *N. tabacum* che *N. benthamiana*.

In Tabella 7 sono schematizzati tutti i plasmidi contenenti il gene *gag* costruiti in questo lavoro.

**Tabella 7:** Vettori con il gene *gag* e sue differenti subunità costruiti in questo lavoro di tesi

Vettore	Transgene *	Compartimento subcellulare	Promotore	5'UTR	Terminatore
pEF40	pr55	Citoplasma	<i>rbcS</i>	<i>rbcS</i>	<i>rbcS</i>
pEF41	pr55	Apoplasto <sup>¶</sup>	<i>rbcS</i>	<i>rbcS</i>	<i>rbcS</i>
pEF42	pr55	R.E. <sup>¥</sup>	<i>rbcS</i>	<i>rbcS</i>	<i>rbcS</i>
pEF43	pr55	Cloroplasto <sup>§</sup>	<i>rbcS</i>	<i>rbcS</i>	<i>rbcS</i>
pEF44	pr55	Mitocondrio <sup>†</sup>	<i>rbcS</i>	<i>rbcS</i>	<i>rbcS</i>
pEF45	pr55	Citoplasma	35S <sup>2</sup>	Ω	<i>rbcS</i>
pEF46	pr55	Cloroplasto <sup>‡</sup>	35S <sup>2</sup>	Ω	<i>rbcS</i>
pEF47	p17	Citoplasma	<i>rbcS</i>	<i>rbcS</i>	<i>rbcS</i>
pEF48	p24	Citoplasma	<i>rbcS</i>	<i>rbcS</i>	<i>rbcS</i>
pEF49	Δp17	Citoplasma	<i>rbcS</i>	<i>rbcS</i>	<i>rbcS</i>
pEF470	p17	Cloroplasto	<i>rbcS</i>	<i>rbcS</i>	<i>rbcS</i>
pEF480	p24	Cloroplasto	<i>rbcS</i>	<i>rbcS</i>	<i>rbcS</i>
pEF490	Δp17	Cloroplasto	<i>rbcS</i>	<i>rbcS</i>	<i>rbcS</i>

\* Pr55 = gene *gag*; p17 = "N-terminal matrix domain" del gene *gag*; p24 = "central capsid domain" del gene *gag*; Δp17 = gene *gag* mancante del dominio p17.

Per l'accumulo delle proteine nei vari compartimenti subcellulari sono stati utilizzati i seguenti peptidi segnale:

¶ = peptide segnale per l'indirizzamento nell'apoplasto del gene "sea anemone equistatin" di *Actinia equina*

¥ = stesso peptide segnale di (¶), con sequenza KDEL al COOH terminale

§ = peptide segnale per l'indirizzamento nel cloroplasto del gene *rbcS* di *Chrysanthemum morifolium*

† = Peptide segnale per l'indirizzamento nella matrice mitocondriale del gene *CoxIV* di *Saccharomyces cerevisiae*

‡ = peptide segnale per l'indirizzamento nel cloroplasto del gene *rbcS* di *Pisum sativum*

Tutti i vettori presentano due "tag" (c-myc e 6His) al COOH terminale.

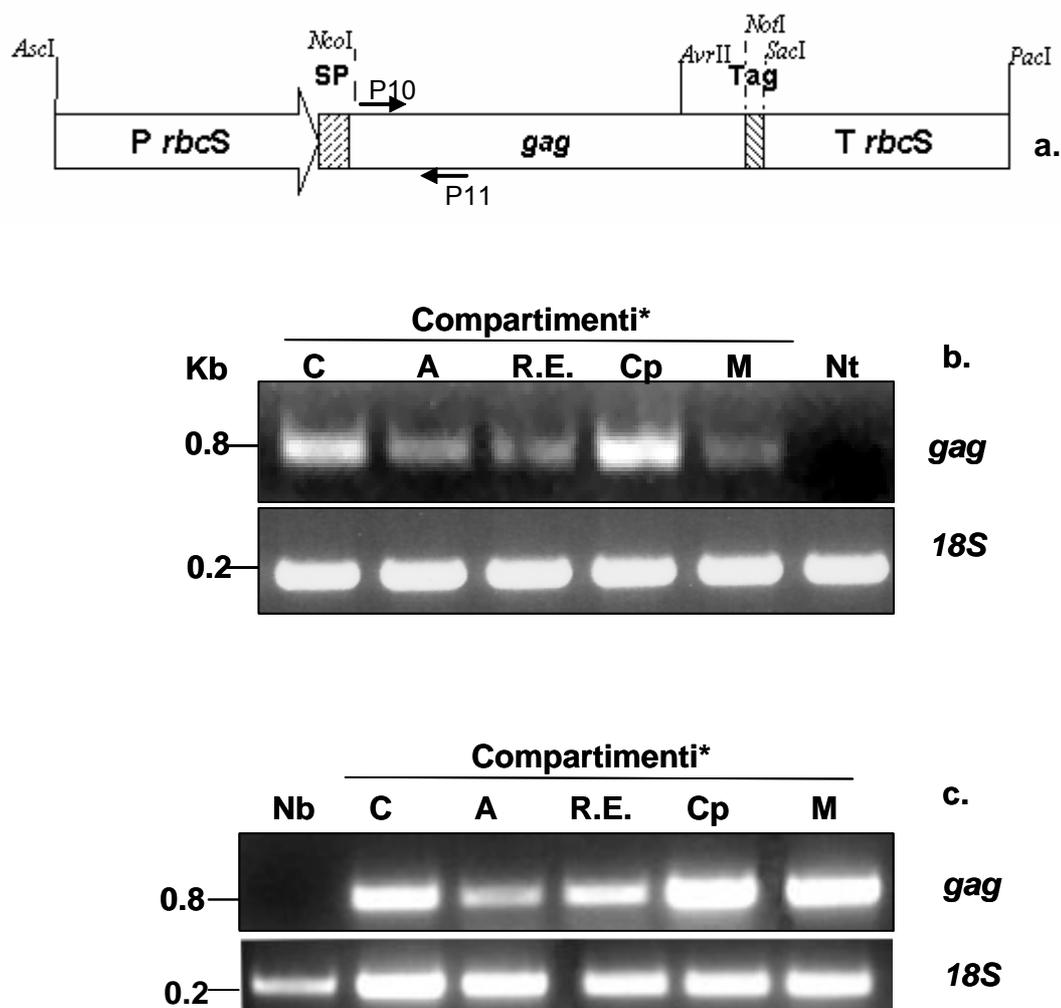
### 3.2.1 Pr55 (*gag*)

L'inoculo nei tessuti vegetali di tutti i vettori contenenti il gene *gag* è stato eseguito con la tecnica dell'agroinfiltrazione descritta nel § 2.8. Per controllare la presenza del trascritto relativo al gene di interesse in entrambe le specie vegetali, è stata utilizzata la tecnica della RT-PCR semiquantitativa sull'RNA estratto dai tessuti agroinfiltrati con i diversi vettori. L'analisi è stata condotta usando la coppia di oligonucleotidi P10 e P11, che amplificano una regione di circa 800 bp dell'RNA messaggero del gene *gag* a partire dall'inizio della regione codificante. I risultati hanno mostrato la presenza di una banda delle dimensioni attese in tutti i costrutti come mostrato in Fig. 17. I livelli di espressione del gene *gag* in *N. tabacum*, confrontati con l'espressione del gene costitutivo *18S*, sono risultati differenti tra i vari costrutti utilizzati, evidenziando un'espressione più abbondante con i vettori per l'espressione nel citosol e nel cloroplasto, mentre in *N. benthamiana* i livelli di espressione del gene *gag* sono risultati costanti, indipendentemente dal vettore utilizzato.

La presenza della proteina esogena nei tessuti di *N. tabacum* è stata condotta utilizzando l'analisi Western blot. Le proteine solubili sono state estratte dai tessuti agroinfiltrati sia di *N. tabacum* che di *N. benthamiana*, trasferite sul filtro di nitrocellulosa ed incubate con l'anticorpo primario policlonale anti - p24 e secondario anti-rabbit coniugato con la perossidasi. In *N. tabacum*, l'analisi western, come mostrato in Fig. 18a, non ha evidenziato la presenza della proteina in nessun compartimento cellulare, ma solo una banda attesa di 24 KDa nel controllo positivo, mentre nei tessuti di *N. benthamiana* agroinfiltrati con il vettore pEF43 (Cp), che indirizza la proteina nel cloroplasto, si è evidenziata la presenza della proteina d'interesse. In particolare, si sono evidenziate due bande di circa 40 e 24 KDa.

Per verificare da quali domini era composta la banda da 40 KDa, è stata effettuata un'altra analisi western sullo stesso estratto proteico usando l'anticorpo policlonale anti - p17. Come mostrato in Fig. 18c, i risultati hanno evidenziato solo la banda di 40 KDa, suggerendo che la banda da 40 KDa è composta dal dominio p17 e dal dominio p24.

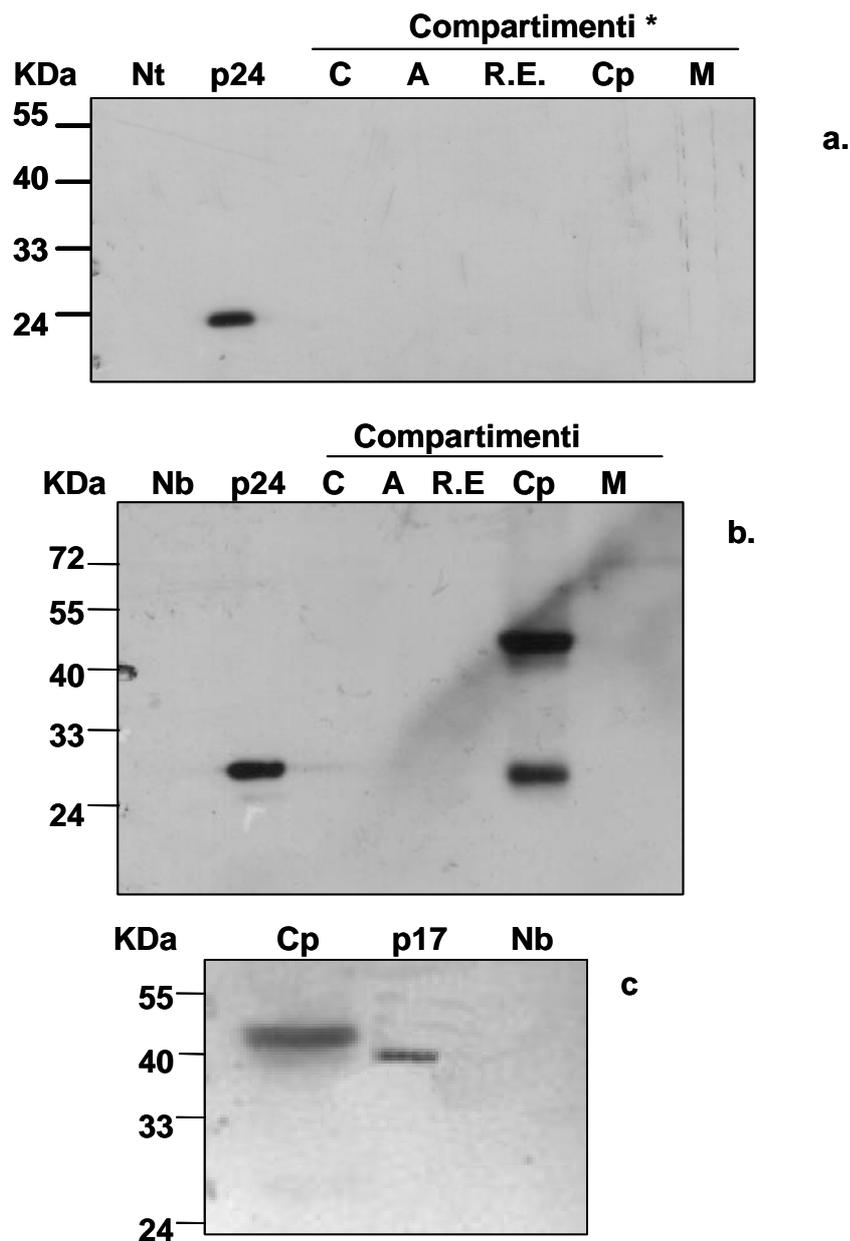
Inoltre, in *N. tabacum*, sono stati agroinfiltrati altri due vettori contenenti il gene *gag*, pEF45 e pEF46. Anche in questo caso per verificare l'espressione del gene *gag* è stata effettuata un'analisi RT-PCR semiquantitativa, con la stessa coppia di primer utilizzata precedentemente. Come evidenziato in Fig. 19b, i risultati hanno mostrato la presenza del trascritto con entrambi i vettori. La presenza della proteina esogena nei tessuti vegetali è stata effettuata utilizzando anticorpi anti - p24, ma in nessun caso è stato possibile osservarne l'accumulo così come mostrato in Fig. 19c.



**Fig. 17: Analisi RT-PCR dei tessuti agroinfiltrati con i vettori pEF40, pEF41, pEF42, pEF43 e pEF44 per i compartimenti subcellulari indicati.**

- a. Rappresentazione schematica dei vettori con il gene *gag*
- b. Analisi RT-PCR con la coppia di primer P10 e P11 sui tessuti di *N. tabacum*. Nt = tessuti di *N. tabacum* non agroinfiltrati
- c. Analisi RT-PCR con la coppia di primer P10 e P11 sui tessuti di *N. benthamiana*. Nb = tessuti di *N. benthamiana* non agroinfiltrati
- \* C = citosol, A = apoplasto, R.E= reticolo endoplasmatico, Cp = cloroplasto, M = mitocondrio

Il gene ribosomiale 18S è stato utilizzato come controllo.



**Fig. 18: Analisi Western sui tessuti agroinfiltrati con i vettori pEF40, pEF41, pEF42, pEF43 e pEF44 (*gag*) per i compartimenti subcellulari indicati.**

a. Analisi Western con anticorpi anti-p24 sui tessuti di *N. tabacum*

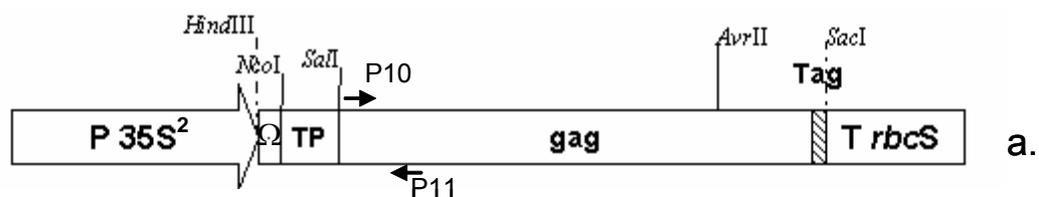
b. Analisi Western con anticorpi anti-p24 sui tessuti di *N. benthamiana*

c. Analisi Western con anticorpi anti-p17 sullo stesso estratto "Cp" usato in b.

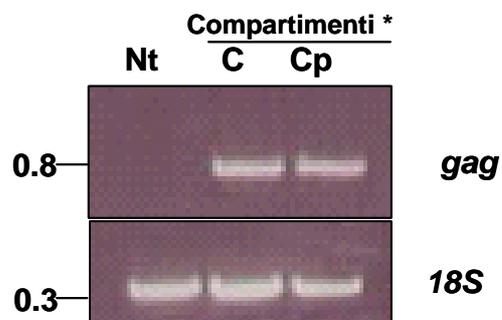
p17 = proteina GST-p17 ricombinante; p24 = proteina p24 ricombinante

Nt = tessuti di *N. tabacum* non agroinfiltrati; Nb = tessuti di *N. benthamiana* non agroinfiltrati;

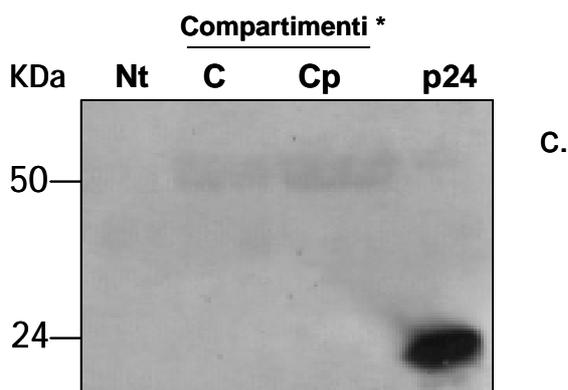
\* C = citosol, A = apoplasto, R.E = Reticolo Endoplasmatico, Cp = cloroplasto, M = mitocondrio.



## RT-PCR



## Western



**Fig. 19: Analisi molecolari dei tessuti di *N. tabacum* agroinfiltrati con i vettori pEF45 e pEF46 (*gag*) per i compartimenti subcellulari indicati.**

a. Rappresentazione schematica del vettore con il gene *gag* e il promotore 35S<sup>2</sup>

b. Analisi RT-PCR con la coppia di primer P10 e P11. Il gene ribosomiale *18S* è stato utilizzato come controllo

c. Analisi Western con l'anticorpo anti - p24. p24 = proteina ricombinante

\* C = citosol, Cp = Cloroplasto

Nt = tessuti di *N. tabacum* non agroinfiltrati

### 3.2.2 Subunità del gene *gag* (*p17*, *p24* e $\Delta p17$ ).

#### 3.2.2.1 Espressione nel citosol

I vettori contenenti le subunità del gene *gag* (*p17*, *p24* e  $\Delta p17$ ) sono stati inoculati sia in *N. tabacum* che in *N. benthamiana*. L'espressione di queste subunità è stata verificata sia nel citosol che nel cloroplasto.

L'espressione dei geni nel citoplasma è stata controllata mediante analisi RT-PCR semiquantitativa usando come controllo il gene costitutivo *18S*. Le analisi sono state condotte utilizzando gli oligonucleotidi specifici mostrati in Fig. 20a che amplificano una regione di circa 200 bp per il gene *p17* e una regione di circa 400 bp per i geni *p24* e  $\Delta p17$ . I risultati, mostrati in Fig. 20b, hanno evidenziato in tutti i casi la presenza di una banda corrispondente al messaggero relativo ai geni d'interesse. I livelli di espressione sono risultati all'incirca uguali se confrontati con il gene costitutivo *18S*.

La ricerca delle proteine d'interesse è stata effettuata mediante analisi western su estratti proteici ottenuti dai tessuti prelevati dopo 3 giorni dall'inoculo con i diversi vettori. Le proteine estratte dai tessuti agroinfiltrati con il vettore contenente la *p17* e trasferite sul filtro di nitrocellulosa, sono state ibridate con l'anticorpi primario anti - *p17* e secondario anti - "rabbit" coniugato con la perossidasi. Dopo la reazione di chemiluminescenza, non è stato possibile osservare nessuna banda relativa alla proteina P17 sia nel citoplasma di *N. tabacum* che di *N. benthamiana* come mostrato in Fig. 20f. Quando, le proteine, estratte dai tessuti vegetali inoculati con i vettori contenenti il gene *p24*, sono state ibridate con l'anticorpo primario anti - *p24* e secondario anti - "rabbit", i risultati hanno evidenziato, sia in *N. tabacum* che in *N. benthamiana*, la presenza di due bande di circa 22 e 25 KDa. Sfruttando le "tag" presenti al COOH terminale, sono state eseguite altre analisi western sugli stessi estratti proteici incubando il filtro con l'anticorpo primario anti - his oppure anti - c-myc. Come evidenziato in Fig. 20d-e, i due anticorpi hanno evidenziato solo la banda più alta di 25 KDa in *N. benthamiana*.

Una sola banda di circa 22 KDa, uguale a quella osservata nei campioni con la *p24*, è stata osservata anche in *N. benthamiana* inoculata con i vettori contenenti la  $\Delta p17$ . Come si evince dalla Fig. 20c, l'espressione in *N. tabacum* della  $\Delta p17$  è risultata essere molto più bassa rispetto a quella ottenuta in *N. benthamiana*.

#### 3.2.2.2 Espressione nel cloroplasto

L'espressione e l'accumulo delle subunità *gag* (*p17*, *p24* e  $\Delta p17$ ) sono stati anche analizzati indirizzando le subunità nei cloroplasti. I vettori costruiti sono stati inoculati in tessuti fogliari di *N. tabacum* e *N. benthamiana* e le analisi sono state eseguite dopo tre giorni dall'infiltrazione.

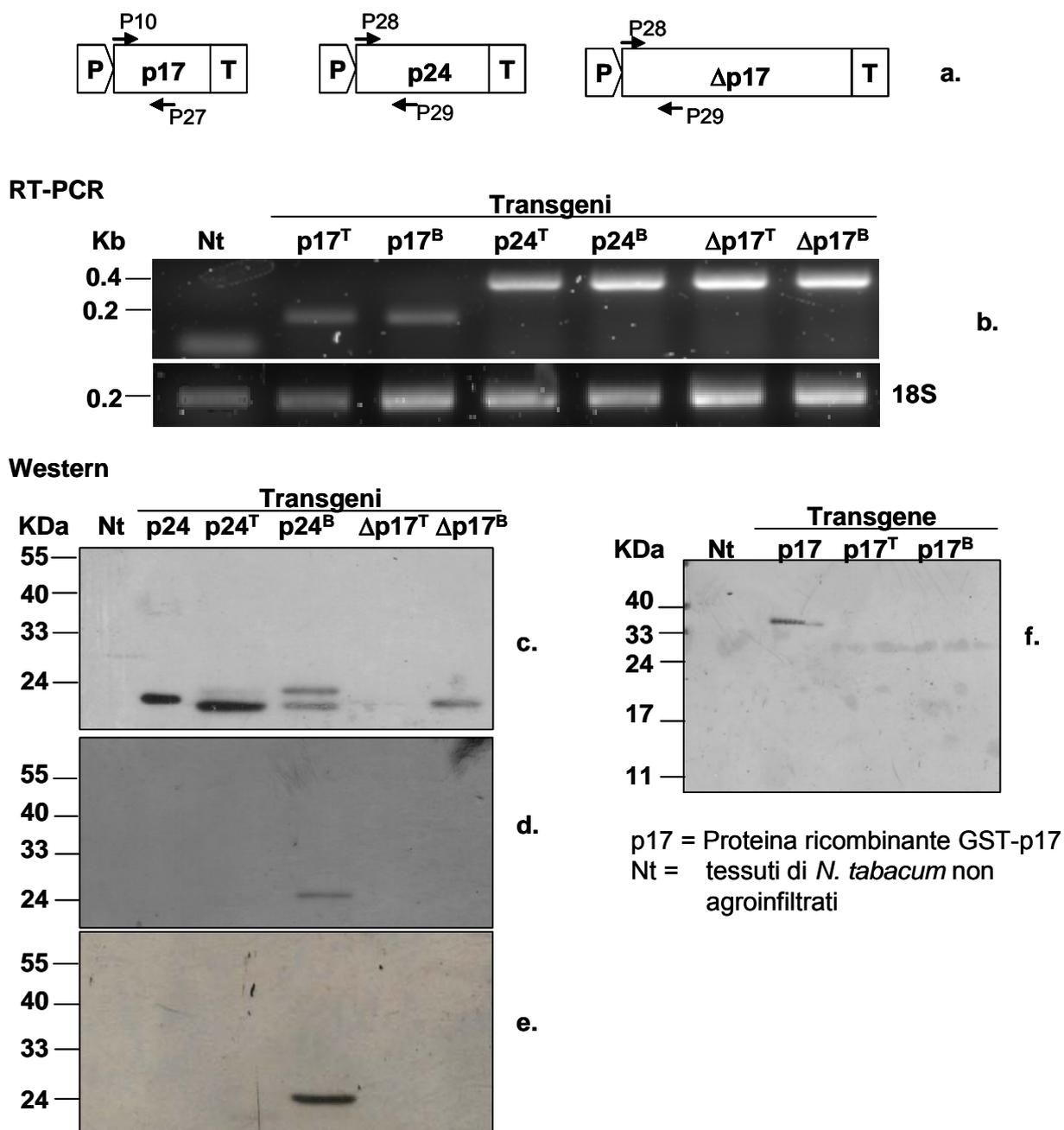
La verifica dell'espressione dei geni codificanti le subunità del gene *gag*, è stata eseguita mediante analisi RT-PCR semiquantitativa, sfruttando gli stessi oligonucleotidi utilizzati per l'espressione delle subunità nel citosol. Come mostrato in Fig. 21b, l'espressione dei diversi geni è risultata costante in entrambe le specie vegetali e uguale tra le subunità *p24* e  $\Delta p17$ . Al contrario, l'espressione della *p17* è risultata essere più bassa rispetto alle altre subunità, anche se costante tra le due specie.

L'accumulo delle subunità proteiche nei cloroplasti di *N. tabacum* e *N. benthamiana* è stata verificata mediante analisi western con anticorpi anti - *p24* ed anti - *p17*. Le analisi western mostrate in Fig. 21c evidenziano la presenza di

una forte banda di 24 KDa in tessuti di *N. benthamiana* agroinfiltrati sia con le subunità *p24* che  $\Delta p17$ . Gli stessi risultati si sono ottenuti anche in *N. tabacum* ma con un livello di espressione molto inferiore rispetto a quelli ottenuti in *N. benthamiana*.

La presenza della subunità *p17* nei cloroplasti di *N. tabacum* e di *N. benthamiana* è stata verificata mediante analisi western utilizzando anticorpi anti - *p17*. I risultati hanno mostrato due bande di circa 13 e 20 KDa solo quando la subunità *p17* è accumulata nei cloroplasti di *N. benthamiana* (Fig. 21d). Lo stesso filtro, è stato poi incubato con l'anticorpo primario anti - c-myc. I risultati, illustrati in Fig.21e, hanno mostrato solo la banda di circa 20 KDa, indicando che quella di 13 KDa è dovuta probabilmente ad un parziale processamento della subunità *p17*.

In Tabella 8 sono stati schematizzati tutti i risultati delle analisi western effettuate sul gene *gag* e sulle subunità *p17*, *p24*,  $\Delta p17$ , sia in *N. tabacum* che in *N. benthamiana*.



**Fig. 20: Analisi molecolari dei tessuti di *N. tabacum* (<sup>T</sup>) e di *N. benthamiana* (<sup>B</sup>) agroinfiltrati con i vettori pEF47 (*p17*), pEF48 (*p24*), pEF49 ( $\Delta p17$ ) contenenti le subunità indicate espresse nel citoplasma.**

a. Rappresentazione schematica delle subunità del gene *gag*

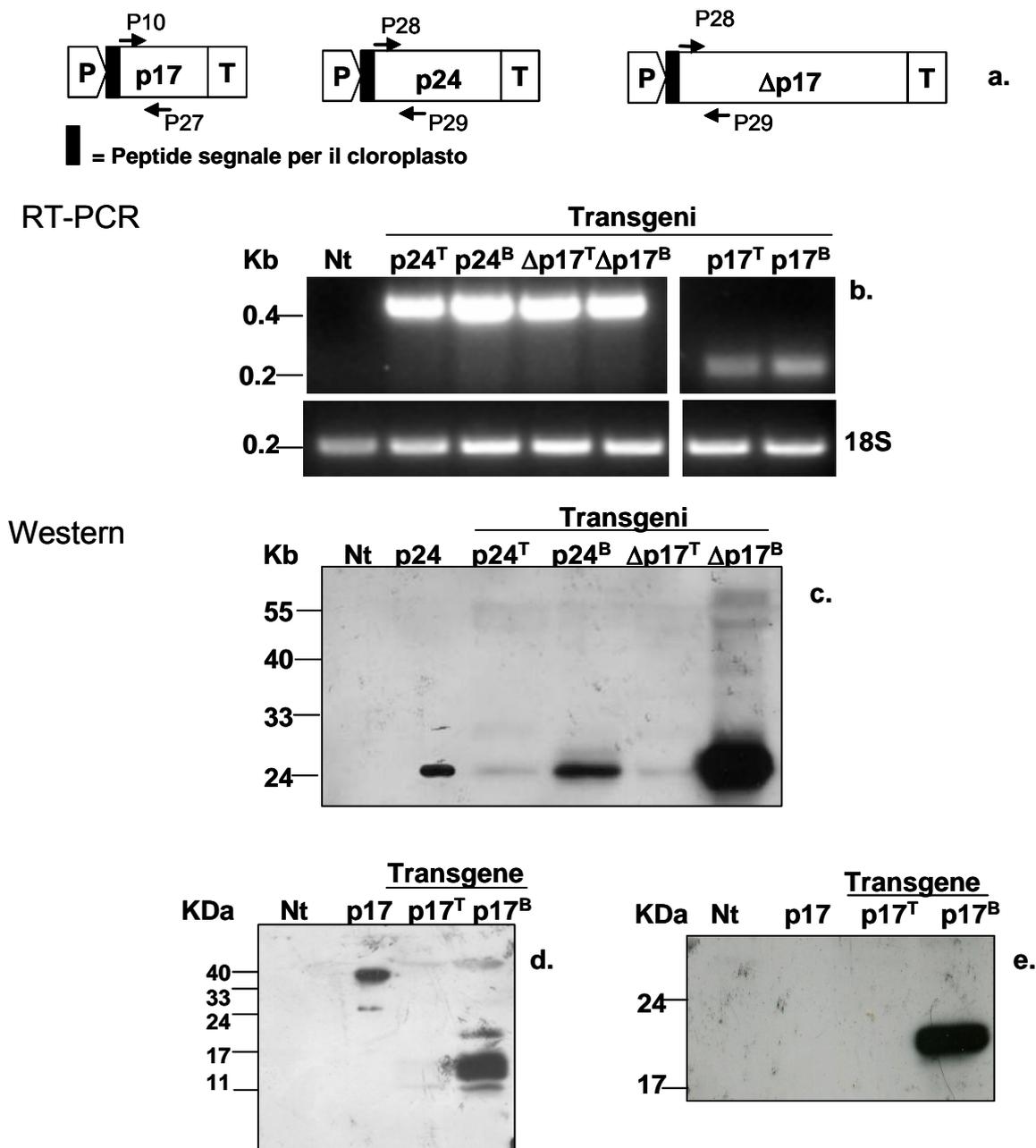
b. Analisi RT-PCR con la coppia di primer indicati in Figura 17a. Il gene ribosomiale *18S* è stato utilizzato come controllo

c. Analisi Western Blot con l'anticorpo anti - p24. p24 = proteina ricombinante p24

d. Analisi Western Blot con l'anticorpo anti - his

e. Analisi Western Blot con l'anticorpo anti - c-myc

f. Analisi Western Blot con l'anticorpo anti - p17



**Fig. 21:** Analisi molecolari sui tessuti agroinfiltrati in *N. tabacum*<sup>(T)</sup> e in *N. benthamiana*<sup>(B)</sup> con i vettori pEF470 (p17), pEF480 (p24), pEF490 (Δp17) contenenti le subunità indicate espresse nel cloroplasto.

- a. Rappresentazione schematica delle subunità del gene *gag*.
- b. Analisi RT-PCR con la coppia di primer indicati in figura a. Il gene ribosomiale 18S è stato utilizzato come controllo
- c. Analisi Western con l'anticorpo anti - p24. p24 = proteina p24 ricombinante
- d. Analisi Western con l'anticorpo anti - p17. p17 = proteina GST-p17 ricombinante
- e. Analisi Western con l'anticorpo anti - c-myc
- Nt = tessuti di *N. tabacum* non agroinfiltrati

**Tabella 8:** Riepilogo dei risultati delle analisi western del gene *gag* espresso in *N. tabacum* (a) e *N. benthamiana* (b)

Transgene	Compartimento subcellulare	Anticorpi	Bande osservate (KDa)	
<i>pr55</i>	Citoplasma	p24	-	
	Apoplasto	p24	-	
	Reticolo	p24	-	
	Cloroplasto	p24	-	
	Mitocondrio	p24	-	
<i>p17</i>	Citoplasma	p17	-	
	Cloroplasto	p17	-	
<i>p24</i>	Citoplasma	p24	22; 25	
	"	His	-	
	"	c-myc	-	
	Cloroplasto	p24	24	
$\Delta p17$	Citoplasma	p24	-	
	"	His	-	
	"	c-myc	-	
	Cloroplasto	p24	24	
<b>b.</b>				
<i>pr55</i>	Citoplasma	p24	-	
	Apoplasto	p24	-	
	Reticolo	p24	-	
	Cloroplasto	p24	40, 24	
	"	p17	40	
Mitocondrio		p24	-	
	<i>p17</i>	Citoplasma	p17	-
		Cloroplasto	p17	13; 20
"	c-myc	20		
<i>p24</i>	Citoplasma	p24	22; 25	
	"	His	25	
	"	c-myc	25	
	Cloroplasto	p24	24	
$\Delta p17$	Citoplasma	p24	22	
	"	His	-	
	"	c-myc	-	
	Cloroplasto	p24	24	

### 3.3 Gene L1 (HPV16)

I vettori costruiti per il gene *L1*, derivando tutti dai plasmidi *ImpactVector*, possiedono lo stesso promotore e 5'UTR del gene *rbcS* di crisantemo, due "tag" al 3', c-myc e 6His e quattro segnali di indirizzamento della proteina in differenti compartimenti subcellulari quali, apoplasto, reticolo endoplasmatico, cloroplasto e mitocondrio. Inoltre, il gene *L1* è stato clonato anche nel vettore senza il segnale di indirizzamento per l'espressione nel citoplasma.

In tutto il lavoro di tesi, è stato sempre utilizzato il gene *L1* privo degli ultimi 22 aminoacidi ( $L1\Delta22$ ), codificanti per il Segnale di Localizzazione Nucleare (NLS). Anche in questo caso, come per i clonaggi delle subunità *p24* e  $\Delta p17$  del gene *gag*, per sfruttare il sito *NcoI* presente nei vettori di clonaggio, è stato necessario cambiare la quarta base nella sequenza nucleotidica di *L1*, da  $ATGT$ , ad  $ATGG$ , ricreando il sito di clonaggio *NcoI* con l' $ATG$  del gene *L1*. Questa modifica ha prodotto un cambio di aminoacido da Serina ad Alanina, come peraltro riportato anche da Biemelt *et al.* (2003).

I primi vettori costruiti, pEF50, 51, 52, 53, 54, presentano il gene  $L1^V\Delta22$  (gene *L1* con la sequenza virale "wild type"), indirizzato rispettivamente nel citosol, nell'apoplasto, nel reticolo endoplasmatico, nel cloroplasto e nel mitocondrio. Inoltre, al fine di esprimere una proteina più corta e potenzialmente più stabile, è stata clonata nei vettori per l'espressione nel citosol e nel cloroplasto, la sequenza  $L1^V$  priva degli ultimi 205 aminoacidi ( $L1^V\Delta205$ ). Questi due vettori, pEF55 e pEF56, possiedono le stesse caratteristiche dei vettori *ImpactVector* su descritti.

Nel nostro laboratorio erano disponibili altre due sequenze del gene *L1*: la sequenza  $L1^{pt}$ , ottimizzata per l'espressione nei plastidi, e la sequenza  $L1^h$ , ottimizzata per l'espressione in cellule umane. Il gene  $L1^{pt}$  anch'esso privo del segnale NLS, è stato clonato in due vettori (pEF57, pEF58), con lo stesso promotore e 5'UTR del gene *rbcS*, due "tag" c-myc e 6His al 3' e due sequenze segnale per l'accumulo della proteina nel citoplasma o nel cloroplasto. Inoltre, negli stessi due vettori, il gene  $L1^{pt}$  ottimizzato per i plastidi è stato clonato anche nella forma tronca ( $L1^{pt}\Delta205$ ), ottenendo così in totale quattro vettori per il gene  $L1^{pt}$ : pEF57, pEF58, pEF570 e pEF580.

Infine, la sequenza genica  $L1^h$ , ottimizzata per l'espressione nelle cellule umane e mancante del segnale NLS, è stata clonata nei due vettori nucleari *ImpactVector* per l'espressione nel citosol o nel cloroplasto. Anche in questi plasmidi sono presenti due "tag" al COOH terminale, c-myc e 6His.

In Tabella 9 sono riportati i vettori contenenti il gene *L1* costruiti in questo lavoro di tesi.

#### 3.3.1 Espressione dei geni $L1^V$ ed $L1^{pt}$

I vettori contenenti i geni  $L1^V\Delta22$  e  $L1^V\Delta205$ , sono stati testati mediante la tecnica dell'agroinfiltrazione sia in *N. tabacum* che in *N. benthamiana*. Per verificare velocemente i primi vettori costruiti, è stata effettuata un'analisi western, con i cinque vettori contenenti il gene  $L1^V\Delta22$  (pEF50, pEF51, pEF52, pEF53, pEF54), inoculati in *N. tabacum*. Come mostrato in Fig. 22, i risultati non hanno evidenziato la presenza della proteina in nessun compartimento cellulare, sia nella frazione solubile che insolubile degli estratti proteici.

Inizialmente, per verificare la corretta trascrizione dei geni d'interesse, dato l'elevato numero di vettori costruiti, è stato necessario cominciare ad analizzare l'espressione dei geni virali  $L1^V\Delta22$ ,  $L1^V\Delta205$ ,  $L1^{pt}\Delta22$  e  $L1^{pt}\Delta205$ , clonati sia nei vettori per l'indirizzamento delle relative proteine nel citosol e nei cloroplasti. Per

confermare la trascrizione dei geni d'interesse, è stata utilizzata la tecnica di RT-PCR semiquantitativa, utilizzando oligonucleotidi complementari ai geni d'interesse che amplificano una regione di circa 300 bp per il gene  $L1^V$  e circa 400 bp per il gene  $L1^{Pt}$ . Come controllo è stata amplificata una regione di circa 200 bp del gene costitutivo 18S. Come mostrato in Fig. 23b, in *N. tabacum* si evidenzia un trascritto abbondante e uguale con tutti i geni e vettori utilizzati. I risultati in *N. benthamiana*, relativamente alla trascrizione dei geni  $L1$  mostrata in Fig. 23c, evidenziano per il gene  $L1^V$  un trascritto abbondante e di uguale intensità in tutti i compartimenti utilizzati, mentre l'espressione dei geni  $L1^{Pt}\Delta22$  e  $L1^{Pt}\Delta205$ , indirizzati nei cloroplasti, è risultata meno abbondante, se rapportata all'espressione del gene costitutivo 18S.

La presenza della proteina L1 nei tessuti agroinfiltrati, è stata verificata mediante analisi western sulle proteine solubili ed insolubili estratte dai tessuti agroinfiltrati. Le piante di *N. tabacum* sono state agroinfiltrate con le sequenze virali  $L1^V\Delta22$ ,  $L1^V\Delta205$ ,  $L1^{Pt}\Delta22$  indirizzate nel citoplasma e nel cloroplasto. Le proteine solubili ed insolubili estratte dai tessuti di *N. tabacum* sono state incubate successivamente con l'anticorpo primario anti - L1 e secondario anti-"rabbit". Come mostrato in Fig. 24, non si è evidenziata la presenza della proteina sia nel citoplasma che nel cloroplasto, ma solo una banda relativa alla proteina L1 espressa in batteri.

I vettori contenenti le sequenze  $L1^V\Delta22$ ,  $L1^V\Delta205$ ,  $L1^{Pt}\Delta22$  e  $L1^{Pt}\Delta205$ , sono stati agroinfiltrati anche in *N. benthamiana* e le analisi western per il controllo della presenza della proteina sono state effettuate utilizzando l'anticorpo primario anti - c-myc. Le proteine solubili ed insolubili, sono state separate su gel di acrilammide e trasferite sul filtro di nitrocellulosa. In questo caso come controllo positivo è stata utilizzata la proteina Gus espressa in *N. benthamiana* con il vettore pEF60. I risultati, anche in questo caso, non hanno evidenziato la presenza di una banda del peso molecolare atteso, sia con le proteine indirizzate nel citosol, che nei cloroplasti.

### 3.3.2 Espressione dei geni $L1^h$

Infine, i vettori pEF500 e pEF501, contenenti la sequenza  $L1^h$  ottimizzata per l'espressione nelle cellule umane, con o senza segnale di indirizzamento nel cloroplasti, sono stati agroinfiltrati in *N. benthamiana*. I risultati sono mostrati in Fig. 25. In particolare, l'analisi RT-PCR è stata effettuata per verificare la corretta trascrizione del gene  $L1^h$ , utilizzando oligonucleotidi specifici che amplificano una regione di circa 350 bp. Anche in questo caso, l'espressione del gene è stata confrontata con il controllo del gene ribosomiale 18S. I dati hanno mostrato un'espressione del messaggero abbondante e comparabile con quella ottenuta con i geni  $L1^V$  ed  $L1^{Pt}$ .

L'accumulo della proteina nel citoplasma e nei cloroplasti dei tessuti fogliari di *N. benthamiana* è stata verificata mediante analisi western utilizzando anticorpi anti - c-myc sia sulla frazione proteica solubile che su quella insolubile. I risultati mostrati in Fig. 25c, hanno evidenziato solo nella frazione solubile una banda molto debole del peso molecolare atteso assente nel controllo non agroinfiltrato, sia quando la proteina è accumulata nel citoplasma che quando accumulata nel cloroplasto.

In Tabella 10 sono riportati tutti i risultati delle analisi western effettuate sul gene  $L1^V$ ,  $L1^{Pt}$ , ed  $L1^h$ , riassumendo gli esperimenti fatti, gli anticorpi utilizzati e i compartimenti subcellulari analizzati.

**Tabella 9:** Vettori con il gene  $L1^v$ ,  $L1^{pt}$ ,  $L1^h$ 

Vettore	Transgene	Compartimento subcellulare
pEF50	$L1^v \Delta 22$	Citoplasma
pEF51	$L1^v \Delta 22$	Apoplasto <sup>¶</sup>
pEF52	$L1^v \Delta 22$	R.E <sup>¥</sup>
pEF53	$L1^v \Delta 22$	Cloroplasto <sup>§</sup>
pEF54	$L1^v \Delta 22$	Mitocondrio <sup>†</sup>
pEF55	$L1^v \Delta 205$	Citoplasma
pEF56	$L1^v \Delta 205$	Cloroplasto
pEF57	$L1^{pt} 22$	Citoplasma
pEF58	$L1^{pt} 22$	Cloroplasto
pEF570	$L1^{pt} \Delta 205$	Citoplasma
pEF580	$L1^{pt} \Delta 205$	Cloroplasto
pEF500	$L1^h \Delta 22$	Citoplasma
pEF501	$L1^h \Delta 22$	Cloroplasto

$L1^v \Delta 22$  = gene  $L1$  "wild type" mancante degli ultimi 22 aminoacidi

$L1^v \Delta 205$  = gene  $L1$  "wild type" mancante degli ultimi 205 aminoacidi

$L1^{pt} \Delta 22$  = gene  $L1$  "codon usage" plastidiale mancante degli ultimi 22 aminoacidi

$L1^{pt} \Delta 205$  = gene  $L1$  "codon usage" plastidiale mancante degli ultimi 205 aminoacidi

$L1^h \Delta 22$  = gene  $L1$  "codon usage" umanizzato mancante degli ultimi 22 aminoacidi

Per l'accumulo delle proteine nei vari compartimenti subcellulari sono stati utilizzati i seguenti peptidi segnale:

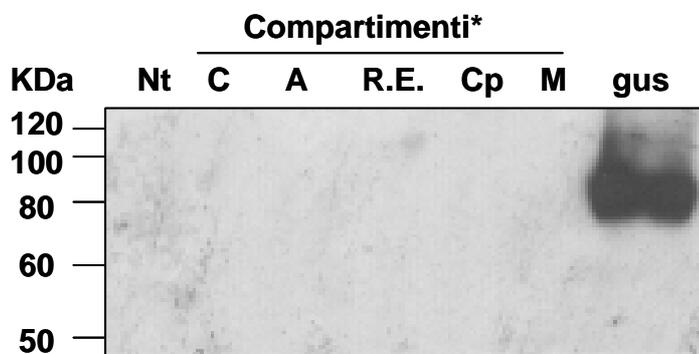
<sup>¶</sup> = Peptide segnale per l'indirizzamento nell'apoplasto del gene "sea anemone equistatin" di *Actinia equina*

<sup>¥</sup> = Stesso peptide segnale di (<sup>¶</sup>), con sequenza KDEL al COOH terminale

<sup>§</sup> = Peptide segnale per l'indirizzamento nel Cloroplasto del gene *rbcS* di *Chrysanthemum morifolium*

<sup>†</sup> = Peptide segnale per l'indirizzamento nella matrice mitocondriale del gene *CoxIV* *Saccharomyces cerevisiae*

Tutti i vettori presentano lo stesso promotore e 5'UTR del gene *rbcS* e due "tag" (c-myc e 6His) al COOH terminale.



**Fig. 22:** Analisi western sui tessuti di *N. tabacum* agroinfiltrati con i vettori pEF50, pEF51, pEF52, pEF53 e pEF54 (L1') per i compartimenti subcellulari indicati:

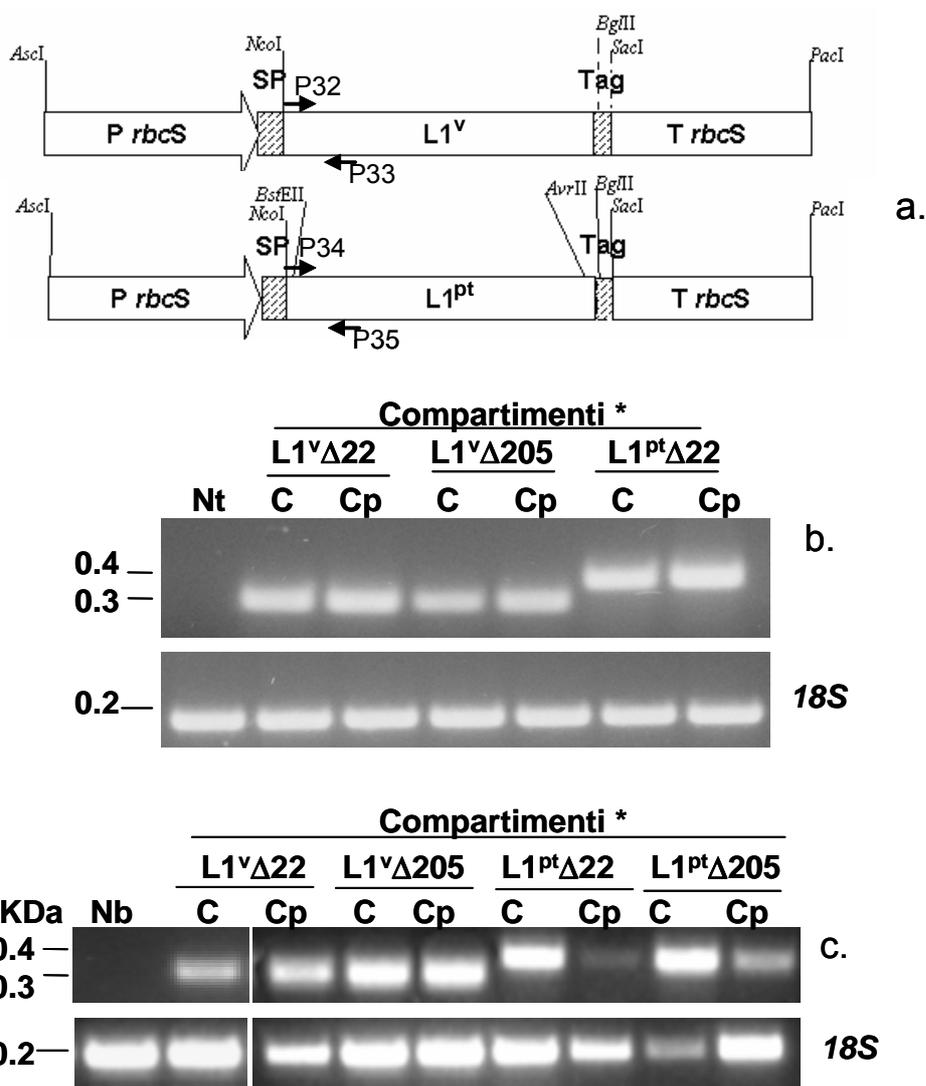
Analisi Western con l'anticorpo anti - his

GUS = Proteina Gus espressa con il vettore pEF60 (citosol)

\* C = citosol, A = apoplasto, R.E = reticolo endoplasmico,

Cp = cloroplasto, M = mitocondrio.

Nt = tessuti di *N. tabacum* non agroinfiltrati



**Fig. 23: Analisi RT-PCR sui tessuti agroinfiltrati con i vettori pEF50, pEF53, pEF55, pEF56, pEF57, pEF58, pEF570 e pEF580 ( $L1^v$  e  $L1^{pt}$ ) per i compartimenti subcellulari indicati**

a. Rappresentazione schematica dei vettori con il gene  $L1^v\Delta22$ ,  $L1^{pt}\Delta22$

b. Analisi RT-PCR con la coppia di primer P32 / P33 e P34 / P35, sui tessuti di *N. tabacum*

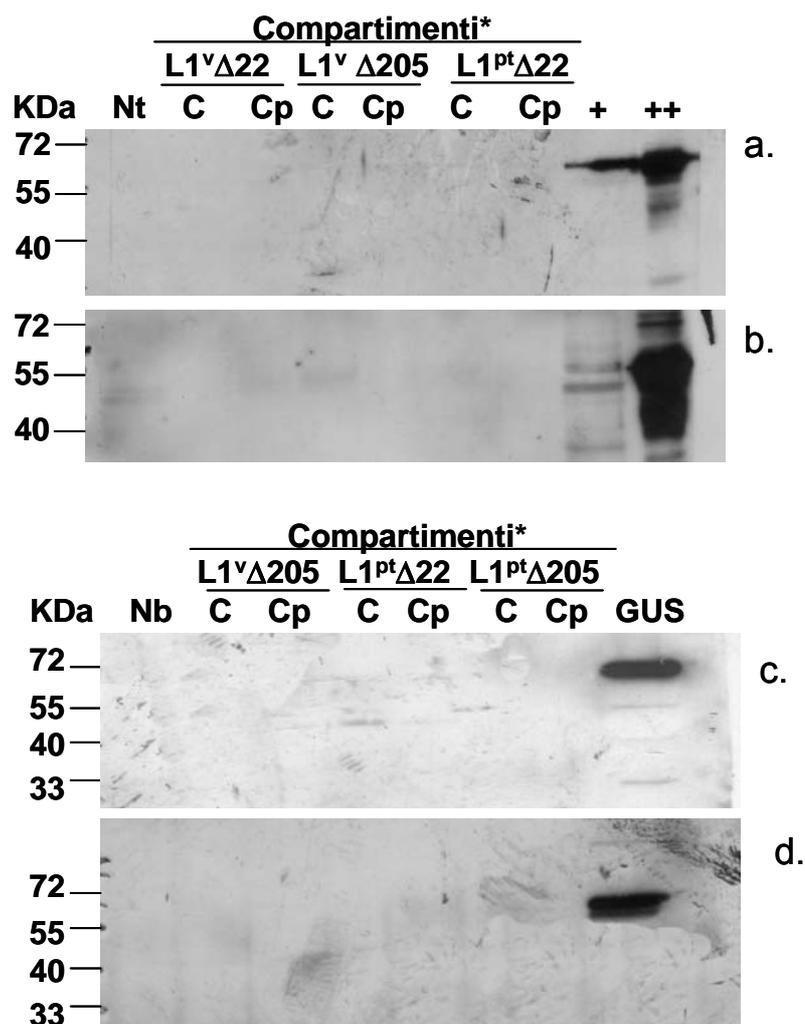
c. Analisi RT-PCR con la coppia di primer P32 / P33 e P34 / P35 sui tessuti di *N. benthamiana*

Nt = tessuti di *N. tabacum* non agroinfiltrata

Nb = tessuti di *N. benthamiana* non agroinfiltrata

\* C = Citosol ; Cp = Cloroplasto

Il gene ribosomiale 18S è stato utilizzato come controllo



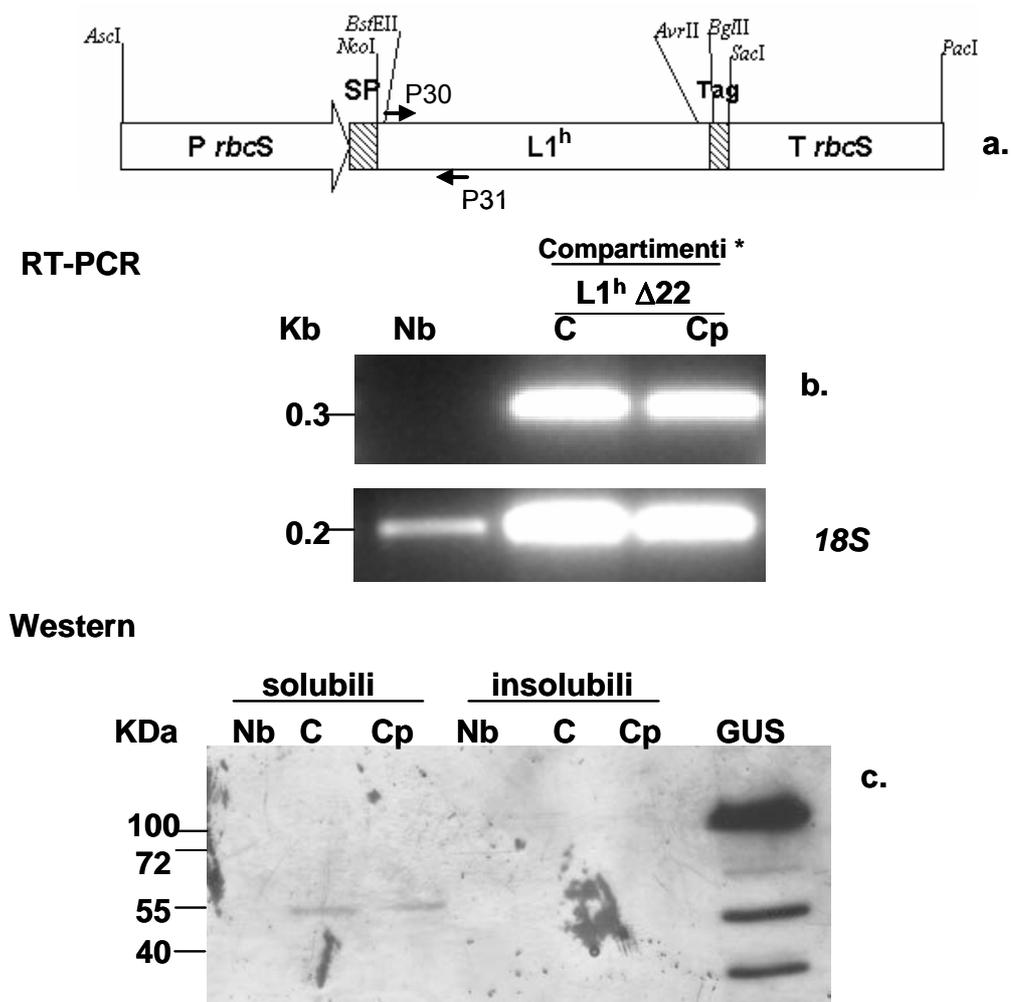
**Fig. 24:** Analisi western sui tessuti agroinfiltrati con i vettori pEF50, pEF53, pEF55, pEF56, pEF57, pEF58, pEF570 e pEF580 (*L1<sup>v</sup>* e *L1<sup>pt</sup>*) per i compartimenti subcellulari indicati.

- a. Analisi western con l'anticorpo anti - L1 su proteine solubili estratte da tessuti di *N. tabacum*  
 b. Analisi western con l'anticorpo anti - L1 su proteine insolubili estratte da tessuti di *N. tabacum*  
 c. Analisi western con l'anticorpo anti - c-myc su proteine solubili estratte da tessuti di *N. benthamiana*  
 d. Analisi western con l'anticorpo anti - c-myc su proteine insolubili estratte da tessuti di *N. benthamiana*

Nt = tessuti di *N. tabacum* non agroinfiltrati; Nb = tessuti di *N. benthamiana* non agroinfiltrati

\* C = citosol; Cp = cloroplasto; GUS = proteina GUS espressa con il vettore pEF60; + = Proteina GST-L1 espressa in batteri non indotti

++ = Proteina GST-L1 espressa in batteri indotti con IPTG.



**Fig. 25: Analisi molecolari sui tessuti di *N. benthamiana* agroinfiltrati con i vettori pEF500 e pEF501 ( $L1^h$ ) per i compartimenti subcellulari indicati.**

a. Rappresentazione schematica del vettore con il gene  $L1^h$

b. Analisi RT-PCR con la coppia di primer P30 e P31.

c. Analisi Western con l'anticorpo anti - c-myc su proteine solubili ed insolubili estratte da tessuti di *N. benthamiana*

\* C = citosol, Cp = Cloroplasto

GUS = proteina Gus espressa con il vettore pEF60

Nb = tessuti di *N. benthamiana* non agroinfiltrati

**Tabella 10:** Riepilogo dei risultati delle analisi western del gene *L1* espresso in *N. tabacum* (a) e *N. benthamiana* (b)

Transgene	Compartimento subcellulare	Anticorpi	Bande osservate (KDa)
<i>L1<sup>v</sup>Δ22</i>	Citoplasma	His	-
	Apoplasto	His	-
	Reticolo	His	-
	Cloroplasto	His	-
	Mitocondrio	His	-
<i>L1<sup>v</sup>Δ205</i>	Citoplasma	L1	-
	"	His	-
	Cloroplasto	L1	-
	"	His	-
<i>L1<sup>pt</sup>Δ22</i>	Citoplasma	L1	-
	"	c-myc	-
	Cloroplasto	L1	-
	"	c-myc	-
<i>L1<sup>pt</sup>Δ205</i>	Citoplasma	L1	-
	"	c-myc	-
	Cloroplasto	L1	-
	"	c-myc	-
<b>b.</b>			
<i>L1<sup>v</sup>Δ22</i>	Citoplasma	L1	-
	"	c-myc	-
	Cloroplasto	L1	-
	"	c-myc	-
<i>L1<sup>v</sup>Δ205</i>	Citoplasma	L1	-
	"	c-myc	-
	Cloroplasto	L1	-
	"	c-myc	-
<i>L1<sup>pt</sup>Δ22</i>	Citoplasma	L1	-
	"	c-myc	-
	Cloroplasto	L1	-
	"	c-myc	-
<i>L1<sup>pt</sup>Δ205</i>	Citoplasma	L1	-
	"	c-myc	-
	Cloroplasto	L1	-
	"	c-myc	-
<i>L1<sup>h</sup>Δ22</i>	Citoplasma	c-myc	55 (fraz Solub)
	Cloroplasto	c-myc	55 (fraz Solub)

## 4. DISCUSSIONE

### 4.1 Vettori costruiti

In questo lavoro di tesi sono stati costruiti diversi vettori per l'espressione nucleare transiente in piante di tabacco di due geni di due virus umani, HIV-1/*gag* e HPV16/L1. I plasmidi utilizzati derivano dai vettori commerciali *ImpactVector* (I.V.) e dal vettore binario pKYLX71:35S<sup>2</sup> ottenuto dal Prof. A.G. Hunt (Università del Kentucky).

Tutti i vettori I.V. utilizzati presentano il promotore del gene *rbcS* di crisantemo. Il lavoro pubblicato da Outchkourov *et al.* (2003) mette a confronto i livelli di espressione del gene *gus* utilizzando il promotore *rbcS* di crisantemo ed il promotore 35S del virus del mosaico del cavolfiore. I risultati riportati hanno evidenziato che il promotore *rbcS* di crisantemo è risultato essere molto più efficiente del promotore 35S, mostrando, in tabacco, un'espressione del gene *gus* circa 8 volte maggiore. Inoltre, gli stessi autori dimostrano che in crisantemo l'espressione del gene *gus* è stata possibile solo utilizzando il promotore *rbcS*.

I due vettori pEF45 e pEF46, contenenti il gene *gag*, derivano invece dal vettore binario pKYLX71:35S<sup>2</sup>, che presenta il promotore 35S<sup>2</sup> duplicato del virus del mosaico del cavolfiore. I livelli di espressione in pianta ottenuti con questo promotore sono ampiamente descritti in letteratura (Daniell *et al.*, 2001; Streatfield *et al.*, 2001). Anche in questo caso la proteina è stata indirizzata nel citoplasma e nel cloroplasto utilizzando il peptide segnale del gene *rbcS* di *Pisum sativum* (Anderson e Smith, 1986).

In tutti i vettori il terminatore usato è sempre stato quello del gene *rbcS*. Il lavoro di Dean *et al.* (1989) mette in risalto l'importanza della combinazione tra promotore e terminatore. Dean ed altri (1989) hanno dimostrato che la regolazione dell'espressione in petunia di due geni endogeni è dovuta a sistemi di regolazione al 3' in combinazione con le regioni promotrici al 5', indicando che una giusta combinazione tra sequenze a valle ed a monte del gene d'interesse possono aumentare i livelli di espressione.

Per aumentare la stabilità del mRNA è stata utilizzata la 5'UTR del gene *rbcS* di crisantemo, oppure, nel caso dei vettori pEF45 e pEF46, la sequenza  $\Omega$  del virus del mosaico del tabacco (TMV). Numerose e varie 5'UTR sono state utilizzate in altri studi. In particolare, Zhang *et al.*, (2002) hanno ottenuto l'espressione della proteina P24 in piante transgeniche di tabacco utilizzando come 5'UTR la sequenza TL del TEV, mentre Biemelt *et al.* (2003), hanno confermato che la fusione della sequenza  $\Omega$  del TMV con il gene *L1* dell'HPV16 ha portato ad un maggiore stabilità del messaggero e ad un accumulo della proteina sia in tabacco che in patata. Anche James *et al.* (2000), hanno dimostrato che l'uso al 5' della sequenza non tradotta TL del TEV, ha incrementato, in cellule di tabacco, la produzione del fattore di crescita umano (GM-CSF) più di due volte.

In tutti i vettori costruiti, i geni sono sempre stati fusi con due "tag", c-myc e 6His, al COOH terminale per facilitare l'analisi western e la purificazione della proteina prodotta. L'uso della "tag" His in pianta è riportata in molti lavori presenti in letteratura. In particolare, nel lavoro di Perez-Filgueira *et al.* (2004), è riportata la fusione del gene *p24* del virus HIV-1 con una coda di sette His. In questo lavoro, gli autori, sfruttando la "tag" di istidine, hanno purificato con successo la proteina P24 espressa in piante di tabacco. Altri lavori riguardanti l'espressione di geni esogeni in pianta fusi con la tag His sono presenti in letteratura, tra cui quello di (Reed *et al.*, 2005; Rumeau *et al.*, 2005 e di Valdez-Ortiz *et al.*, 2005). In

quest'ultimo la "tag" His è stata utilizzata per purificare una globulina espressa in piante di tabacco transgeniche. Anche l'uso della "tag" c-myc è stata utilizzata con successo in molti lavori presenti in letteratura, tra cui il lavoro di Ferrando *et al.* (2001).

Tutti i vettori costruiti in questo lavoro sono strutturalmente simili, differenziandosi per la presenza o meno di sequenze segnale di indirizzamento delle proteine in differenti compartimenti cellulari. I vettori utilizzati presentavano 4 diverse sequenze segnale per l'indirizzamento della proteina nell'apoplasto, nel reticolo endoplasmico, nel cloroplasto o nel mitocondrio e nessuna sequenza segnale per indirizzare la proteina nel citoplasma. La scelta di esprimere una proteina in un determinato compartimento cellulare può portare in alcuni casi ad un aumento di espressione della stessa per una maggiore stabilità e/o condizioni cellulari più adatte. Per esempio in letteratura ciò è riportato da Ziegelhoffer *et al.* (2001), dove l'espressione del gene *E1* codificante per una  $\beta$ -1,4 glucanase, è risultata essere 100 volte maggiore quando indirizzata nell'apoplasto che quando espressa nel citosol, e 10 volte maggiore quando accumulata nei cloroplasti. Nello stesso lavoro, per l'accumulo della proteina nell'apoplasto e nel cloroplasto gli autori hanno usato il peptide segnale del gene *VSP $\beta$*  e del gene *rbcS*, rispettivamente. Nei vettori utilizzati in questa tesi, invece, per l'accumulo delle proteine nell'apoplasto o nel reticolo endoplasmico sono state utilizzate le sequenze del gene "sea anemone equistatin" di *Actinia equina* senza o con il segnale KDEL; la sequenza del gene *rbcS* di *Chrysanthemum morifolium* è stata invece utilizzata per l'indirizzamento nel cloroplasto e la sequenza del gene *CoxIV* di *Saccharomyces cerevisiae* per l'accumulo della proteina nei mitocondri.

Alcuni lavori in letteratura hanno dimostrato che l'uso della sequenza KDEL per la ritenzione nel reticolo endoplasmico ha portato in alcuni casi ad un aumento dei livelli di espressione (Schouten *et al.*, 1996; Conrad e Fiedler, 1998). In entrambi i lavori citati gli autori hanno espresso la singola catena Fv degli anticopi umani (scFv) in piante transgeniche di tabacco, indirizzandola nel citoplasma, nell'apoplasto oppure nel reticolo endoplasmico con l'aggiunta della sequenza KDEL. Schouten *et al.* (1996) indirizzando la proteina scFv nel citoplasma, non hanno rilevato nessun accumulo della proteina. Aggiungendo allo stesso costrutto solo il segnale KDEL, l'accumulo della proteina scFv è stato stimato sull'0,2% delle PTS. Inoltre, quando la stessa sequenza è stata indirizzata nell'apoplasto o nel reticolo endoplasmico, i livelli di espressione sono variati dallo 0,01% all'1% delle PTS rispettivamente.

L'uso del peptide segnale del gene *rbcS* di *Chrysanthemum morifolium* è stata descritto in letteratura da Wong *et al.* (1992). Gli autori hanno mostrato come i livelli della proteina *cryIA* nel cloroplasto aumentano di circa il doppio rispetto a quando espressa nel citoplasma. Altri dati di letteratura mediante l'uso di diverse sequenze segnale per il cloroplasto riportano che l'espressione della GFP in riso raggiunge circa il 10% delle proteine totali solubili quando è accumulata nei cloroplasti, rispetto allo 0,5% delle proteine totali solubili quando è espressa nel citoplasma (Jang *et al.*, 1999). L'uso del peptide segnale per il mitocondrio usato in questi vettori è stato descritto da Köhler *et al.* (1997), dove gli autori hanno espresso e localizzato con successo nella matrice mitocondriale la proteina GFP fusa con il segnale di indirizzamento del gene *CoxIV*.

## 4.2 Agroinfiltrazione dei vettori costruiti

I vettori costruiti in questo lavoro di tesi, contenenti il gene reporter *gus* e i geni d'interesse, sono stati inseriti nel ceppo di *A. tumefaciens* LBA4404 (Hoekema *et al.*, 1983) e provati in transiente mediante la tecnica dell'agroinfiltrazione. L'uso del ceppo LBA4404 di *A. tumefaciens* per l'espressione transiente è riportato in letteratura in numerosi lavori tra cui quelli di Wroblewski *et al.* (2005) Il ceppo LBA4404 da noi utilizzato ha mostrato ottimi risultati con i vettori contenenti il gene reporter *gus*. L'uso di altri ceppi di agrobatterio per l'agroinfiltrazione è riportato da Wydro *et al.* (2006), i quali, hanno evidenziato dopo quattro giorni dall'infiltrazione della GFP in *N. benthamiana* un buon accumulo della proteina anche con il ceppo di *A. tumefaciens* AGL0; invece, il ceppo EHA105 ha mostrato livelli di espressione della GFP più bassi ma comparabili con quelli del ceppo LBA4404. Anche la concentrazione di acetosiringone e la densità dei batteri sembra essere importante per ottenere una buona espressione. In tutti i nostri esperimenti abbiamo utilizzato un'O.D.<sub>600</sub> della sospensione batterica di circa 0.3 ed una concentrazione di acetosiringone di 200  $\mu$ M. Wydro, Wroblewski e rispettivi collaboratori, hanno dimostrato che in *N. benthamiana* differenti densità della sospensione batterica (0.1, 0.5, 1, 1.5) non hanno un effetto significativo sull'espressione transiente del gene della *gfp*. Sempre nel lavoro di Wydro e collaboratori si sono provate differenti concentrazioni di acetosiringone per l'induzione dell'agrobatterio ricombinante. Le concentrazioni da loro provate variavano da 50  $\mu$ M a 600  $\mu$ M ed i risultati hanno mostrato solo un lieve aumento del gene *gfp* variando la concentrazione da 150 a 300  $\mu$ M.

Come già ampiamente dimostrato dalla letteratura, è noto che in molti casi l'espressione transiente mediante agroinfiltrazione è soggetta a fenomeni di "RNA silencing" che riducono fortemente l'espressione già dopo 2 giorni dall'inoculo nei tessuti vegetali (Voinnet *et al.*, 2003; Yang *et al.*, 2000). In questi caso, l'uso di geni soppressori del silenziamento come il gene *p19* del TBSV, oppure il gene *HcPro* del PVY, sono fondamentali per ottenere un'elevata espressione del gene d'interesse anche fino a 20 giorni dall'inoculo (Voinnet *et al.*, 2003). In questo lavoro di tesi non è stato effettuato un "time course" dell'espressione dei geni d'interesse a vari giorni dall'inoculo. I risultati possono confermare che entro i tre giorni dall'inoculo non c'è stato un forte silenziamento genico in quanto sia il trascritto che la proteina relativa al gene *gus-int* sono stati rilevati mediante le relative analisi molecolari.

I risultati hanno invece evidenziato una marcata differenza dei livelli di espressione della proteina Gag tra *N. tabacum* e *N. benthamiana*. La spiegazione di questo risultato non trova riscontro con molti dati di letteratura. Il lavoro pubblicato da Koscianska *et al.* (2005) però, mette a confronto la comparsa di siRNA specifici dopo infiltrazione del gene *gfp* sia in *N. tabacum* che in *N. benthamiana*. I risultati riportati nel lavoro citato hanno mostrato in *N. tabacum* una più rapida comparsa e scomparsa di siRNA già dopo 2 giorni dall'infiltrazione rispetto a quanto accade in *N. benthamiana*, evidenziando una più veloce degradazione del messaggero specifico nella prima specie vegetale.

I risultati ottenuti in questo lavoro con il gene *gus* in *N. tabacum* e in *N. benthamiana* hanno mostrato un trascritto relativo al gene *gus* costante e paragonabile sia nelle due specie vegetali che tra i vettori con i differenti peptidi segnale. L'intensità dell'espressione della proteina Gus è stata invece differente a seconda del compartimento di accumulo. I risultati hanno evidenziato che l'accumulo della proteina Gus è ottimale nel citoplasma e nel cloroplasto, mentre il

reticolo endoplasmico, l'apoplasto e il mitocondrio non sembrano essere buoni siti di accumulo per la proteina Gus.

### 4.3 Gene *gag*

In seguito ad infiltrazione in *N. tabacum* e in *N. benthamiana* di tutti i vettori contenenti il gene *gag* indirizzato nei cinque compartimenti subcellulari, sono stati analizzati i tessuti agroinfiltrati per la ricerca dell'mRNA e della proteina. L'espressione del transgene *gag* è stata studiata mediante analisi RT-PCR, utilizzando due oligonucleotidi che amplificano una regione di circa 800 basi a partire dal sito di inizio della traduzione. Questo approccio ha permesso di verificare la presenza di parte del trascritto relativo al gene *gag*. I risultati in *N. tabacum* hanno mostrato un messaggero più abbondante nel citosol e nel cloroplasto, mentre in *N. benthamiana* non ci sono state differenze nei livelli dell'mRNA tra i vari costrutti analizzati.

Le analisi western condotte sui tessuti di *N. tabacum* non hanno mostrato la presenza della proteina Gag in nessun compartimento cellulare, evidenziando che non esiste una diretta correlazione tra trascrizione e traduzione. Risultati simili sono riportati in diversi lavori presenti in letteratura. Nel lavoro di Wirth *et al.* (2004), ad esempio, l'espressione in piante transgeniche di tabacco del fattore di crescita hEGF ha mostrato stessi livelli di trascritto con tutti i vettori costruiti, ma l'accumulo della proteina è stato evidenziato soltanto quanto il gene hEGF è stato indirizzato nell'apoplasto. Anche nel lavoro di James e altri (2000), gli autori riportano l'espressione del gene GM-CSF fuso o meno con la 5'UTR. I livelli del messaggero del gene sono risultati essere costanti in tutte le cellule di tabacco trasformate con il vettore senza la 5'UTR, ma più bassi nelle cellule trasformate con il vettore contenente la 5'UTR. Anche in questi casi i livelli di proteina non erano correlati ai livelli di espressione dell'RNA, che era rilevabile solo utilizzando il costrutto dove il gene GM-CSF è stato fuso con la 5'UTR. Gli autori giustificano questi risultati commentando che la presenza di 5'UTR nel messaggero porta ad una bassa ma più stabile espressione dei livelli di RNA, che si traduce in un maggiore accumulo della proteina. Apparentemente, la 5'UTR ha due effetti, bassa trascrizione e significativa stabilizzazione della traduzione (Gallie e Walbot, 1992).

In questa tesi, l'espressione della proteina Gag in *N. tabacum* è stata analizzata anche utilizzando i vettori pEF45 e pEF46 dove l'espressione del gene *gag* è guidata dal promotore 35S<sup>2</sup>. Di nuovo, pur in presenza di un messaggero stabile e abbondante, non è stato possibile rilevare la presenza della proteina.

I risultati delle analisi western in *N. benthamiana* hanno mostrato invece, la presenza della proteina solo quando il gene *gag* è stato indirizzato nei cloroplasti. Nessuna banda relativa alla proteina Gag è stata osservata negli altri compartimenti cellulari, suggerendo che l'ambiente plastidiale sembra essere in grado di permettere l'accumulo della proteina virale. In particolare, le due bande osservate di 40 e 24 KDa, evidenziano un rapido processamento della proteina intera (55 KDa) nei cloroplasti di *N. benthamiana*. La banda da 24 KDa, corrispondente al dominio p24 del gene *gag*, deriva dai tagli della proteina precursore Gag tra i domini p24 e p7 e tra i domini p17 e p24. La banda da 40 KDa, invece, è composta solo dal dominio p17 più il dominio p24 e manca dei domini p7 e C-terminale. Questo processamento parziale della proteina immatura Gag rispecchia quanto avviene nelle cellule umane infettate dal virus (Morikawa, 2003). In pianta queste modifiche potrebbero essere effettuate da proteasi vegetali, in particolare plastidiali, che riconoscono le stesse sequenze di taglio

della proteasi virale. Per supportare questa ipotesi è stato utilizzato il programma "NetPicoRNA" ([www.expasy.org](http://www.expasy.org)) che analizza potenziali siti di taglio riconosciuti da proteasi. Inserendo la sequenza proteica del gene *gag*, i risultati hanno mostrato che la sequenza aminoacidica YPIVQNLQG situata in posizione +136 a partire dall'inizio della traduzione, è riconosciuta come un probabile sito di taglio, con un indice di affidabilità dello 0,86, ben al di sopra del limite minimo di affidabilità (0,5) suggerito dal programma utilizzato. Questa sequenza, come atteso, corrisponde alla giunzione tra il dominio p17 e p24.

In letteratura, fino ad oggi, non sono ancora presenti lavori pubblicati riguardanti l'espressione in pianta del gene intero *gag* del virus HIV-1. Un solo riassunto, riporta l'espressione del gene *gag* in *N. tabacum* mediante l'uso di un vettore virale TMV, ed in *N. benthamiana* mediante la tecnica dell'agroinfiltrazione (Meyers *et al.*, 2005). Gli autori hanno anche espresso, oltre al gene *gag* "wild type", anche un gene *gag* modificato, sia con un "codon usage" ottimizzato per pianta sia con una sequenza ottimizzata per l'espressione in cellule umane. I dati riportati, ancora incompleti, hanno evidenziato una bassa espressione sia del gene *gag* "wild type" che del gene ottimizzato per pianta, usando come sistema di espressione il vettore virale TMV. La tecnica dell'agroinfiltrazione in *N. benthamiana*, ha dato livelli di espressione del gene *gag* ottimizzato per pianta, circa 10 volte maggiore rispetto a quanto ottenuto con il vettore virale TMV.

Sono presenti, invece, lavori in letteratura riguardanti l'espressione in pianta di altre proteine ricombinanti utili per la salute umana (Tacket *et al.*, 2000; Kapusta *et al.*, 1999; Smith *et al.*, 2002; Huang *et al.*, 2006). Nel lavoro di Huang e altri (2006), gli autori riportano l'espressione del gene HBs del virus HBV attraverso l'uso di un vettore virale. In questo lavoro sono riusciti ad ottenere un livello di espressione in *N. benthamiana* di circa il 7% delle proteine totali solubili. Questo risultato è molto promettente in quanto altri autori come Richter *et al.* (2000) riportano nei loro lavori livelli di espressione in pianta molto bassi, intorno allo 0.1 mg di proteina ricombinante per grammo di foglie. Una nuova strategia, descritta da molti autori come (Koya *et al.* 2005; Tregoning *et al.*, 2003; Maliga, 2004), è la trasformazione plastidiale, che potrebbe permettere un maggior accumulo delle proteine esogene. Questo è dimostrato dai lavori pubblicati da Koya *et al.* (2005), i quali hanno espresso l'antigene protettivo dell'antrace nei cloroplasti di tabacco fino ad ottenere il 14,2% delle proteine totali solubili. Tregoning e altri (2003), invece, hanno espresso, sempre nei cloroplasti di *N. tabacum*, la tossina del tetano, fino ad un livello pari al 25% delle proteine totali solubili.

In questo lavoro di tesi sono state espresse sia in *N. tabacum* che in *N. benthamiana* anche le subunità *p17*, *p24* e  $\Delta p17$  del gene *gag* separatamente per verificare se la mancata espressione del gene intero *gag* in *N. tabacum* fosse dovuta ad una particolare subunità. L'accumulo dei trascritti, a livelli paragonabili tra di loro, è stato confermato in *N. tabacum* e in *N. benthamiana*, sia quando le subunità sono state espresse nel citoplasma che quando indirizzate nei cloroplasti.

La presenza della proteina P24 è stata verificata mediante analisi western. Sono state osservate due bande di 22 e 25 KDa sia in *N. tabacum* che in *N. benthamiana*, sebbene la presenza di queste due bande è molto più marcata nella seconda specie. Questo risultato fa supporre un processamento della proteina nei tessuti vegetali. Questa nostra ipotesi è stata confermata sfruttando le due "tag" c-myc e 6His. Infatti, utilizzando anticorpi diretti contro le due "tag" sui tessuti di *N. benthamiana* agroinfiltrati con la *p24*, è stata osservata solo la banda di 25 KDa, che corrisponde al peso molecolare della proteina P24 fusa alle due "tag" e 6His.

Quando la *p24* è stata indirizzata nei cloroplasti di *N. tabacum* e di *N. benthamiana*, è stata evidenziata solo una banda di 24 KDa in entrambe le specie, ma di intensità molto più forte rispetto a quanto osservato nel citoplasma.

In letteratura, al contrario di quanto presente per il gene *gag*, sono presenti alcuni lavori riguardanti l'espressione della *p24* in pianta attraverso diversi sistemi di espressione (Zhang *et al.*, 2002; Badillo *et al.*, 2005; Perez-Filgueira *et al.*, 2004 e Meyers *et al.*, 2005). Il primo lavoro di Zhang e collaboratori nel 2002 riporta l'espressione della *p24* in piante transgeniche di tabacco con livelli di espressione dello 0,35% sulle proteine solubili totali. Utilizzando invece il vettore virale TMV come descritto da Perez-Filgueira *et al.* (2004), la quantità di *p24* che è possibile ottenere scende allo 0,1% delle proteine *totali* solubili. Questo dato è confermato anche da Meyers *et al.* (2005) che hanno espresso la *p24* "wildtype", la *p24* con un "codon usage" per pianta e la *p24* con un "codon usage" umanizzato, tramite l'uso sia di un vettore virale che dell'agroinfiltrazione. I risultati riportati dall'autore, indicano che utilizzando l'agroinfiltrazione è possibile ottenere un'espressione della *p24* ottimizzata per pianta di circa 100 volte maggiore rispetto a quanto ottenuto con il vettore virale.

La presenza delle sequenze peptidiche corrispondenti alle subunità  $\Delta p17$  sia nel citoplasma che nel cloroplasto è stata verificata utilizzando gli stessi anticorpi policlonali anti - *p24*. In questo caso, in *N. tabacum* ed in *N. benthamiana*, è stata osservata solo una banda di 24 KDa. Questi risultati confermano che la proteina precursore  $\Delta p17$ , come per la proteina intera Gag, perde i domini p7 e p6 per il taglio proteolitico che subisce dopo il dominio *p24*. Nessun lavoro in letteratura riporta l'espressione del gene Gag senza il dominio p17, ma questi dati riconfermano parte del processamento osservato precedentemente con il gene Gag intero.

I livelli di espressione della subunità  $\Delta p17$  sono risultati molto più alti quando la proteina è stata accumulata nei cloroplasti.

Infine, abbiamo verificato la presenza della *p17* nei tessuti agroinfiltrati. È stato osservato l'accumulo della proteina solo quando la proteina è stata indirizzata nei cloroplasti di *N. benthamiana*, mentre nel citoplasma non è stata rilevata la presenza della proteina. In particolare sono state osservate due bande del peso molecolare di 13 e 20 KDa, indice di un processamento della stessa nei tessuti vegetali ed una banda di peso molecolare ancora inferiore, che potrebbe essere un prodotto di degradazione. Utilizzando anticorpi anti - c-myc si è evidenziata solo la banda di 20 KDa che corrisponde alla proteina P17 fusa alle due "tag" c-myc e 6His presenti al COOH terminale. La mancata espressione della proteina P17 nel citoplasma conferma ancora che l'ambiente plastidiale sembra il miglior sito di accumulo della proteina Gag e delle sue subunità. Dati di letteratura indicano che la subunità p17 è miristilata all'N-terminale (Kaushik e Ratner, 2004; Morikawa, 2003). La miristilazione è una modificazione post-traduzionale, fondamentale per i primi stadi della replicazione virale, per permettere l'inserimento della proteina gp120 nella matrice virale (Casson, 1996) e per portare il virione a contatto con la membrana plasmatica (Morikawa, 2003). Questa modifica post-traduzionale potrebbe essere non tollerata dalla cellula vegetale. Infatti, i nostri risultati suggeriscono che indirizzando la proteina P17 nei cloroplasti, la sequenza di miristilazione presente all'N-terminale (MGARASV), è preceduta dalla sequenza del peptide segnale per il cloroplasto. Questo comporta che la proteina non è più miristilata in quanto la prima glicina non è più inserita nel contesto di miristilazione. Inoltre, la miristilazione è una modifica post-traduzionale

che non avviene nei cloroplasti, anche se fossero presenti le sequenze corrette. Questo potrebbe essere un motivo per cui la proteina P17 riesce ad accumularsi quando indirizzata nei plastidi. Non esistono dati di letteratura che riportino l'espressione della *p17* in pianta, anche se molte proteine di origine vegetale sono miristilate e vengono espresse normalmente nelle cellule vegetali. Un lavoro pubblicato da Loh *et al.* (1998) dimostra come una proteina di resistenza a patogeni *Pto* è espressa in pomodoro indipendentemente se sia miristilata o meno.

#### 4.4 Gene L1

L'espressione del gene *L1* del virus HPV16 è stata testata mediante agroinfiltrazione sia in *N. tabacum* che in *N. benthamiana*. Sono state clonate nei vettori I.V. tre sequenze del gene *L1*, la sequenza  $L1^V$ , di origine virale, la sequenza  $L1^{pt}$ , ottimizzata per l'espressione plastidiale, e la sequenza  $L1^h$  con un "codon usage" ottimizzato per l'espressione in cellule umane. La sequenza  $L1^V$ , è stata clonata nei cinque vettori I.V. in modo da indirizzarla nel citoplasma, nell'apoplasto, nel reticolo endoplasmatico, nel cloroplasto e nel mitocondrio. Dovendo indirizzare la proteina in differenti compartimenti subcellulari, da tutte le sequenze del gene *L1* è stato rimosso il segnale di localizzazione nucleare (NLS). Inoltre, tale segnale è stato considerato tossico per la vitalità delle cellule vegetali (Warzecha *et al.*, 2003), e non importante per la formazione delle VLPs (Paintsil *et al.*, 1996). L'NLS è stato rimosso anche dalla proteina L1 di HPV16 da Warzecha *et al.* (2003) e anche dalla proteina L1 di HPV11 da Varsani *et al.* (2003).

Inoltre, per esprimere una proteina più corta e potenzialmente più stabile, è stata clonata nei vettori per l'espressione nel citoplasma e nei plastidi anche una sequenza di  $L1^V$  mancante degli ultimi 205 aminoacidi. Questa proteina non possiede più la sua capacità autoassemblante ma conserva ancora le proprietà immunogeniche

I pochi dati di letteratura e risultati non promettenti ottenuti nel nostro laboratorio con il gene  $L1^V$  hanno portato alla costruzione di due sequenze sintetiche del gene *L1*: la prima dove la sequenza genica è stata ottimizzata per l'espressione nei cloroplasti ( $L1^{pt}$ ) e la seconda dove il "codon usage" del gene *L1* è stato ottimizzato per l'espressione in cellule umane ( $L1^h$ ).

La sequenza del gene  $L1^{pt}$ , già disponibile nel nostro laboratorio, è stata disegnata in modo da ottimizzare il "codon usage" per l'espressione plastidiale, eliminando anche dalla sequenza segnali di terminazione precoce della trascrizione (TTTTNT), segnali di poliadenilazione eucariotici (AAUAAA), che normalmente si trovano nella sequenza del gene *L1* (Earl *et al.*, 1990; Zhang *et al.*, 1998). Inoltre, per sfruttare il sito *NcoI* presente nei vettori di clonaggio utilizzati, è stata modificata la 4° base della sequenza di *L1* cambiando l'aminoacido da serina ad alanina. Questa sostituzione aminoacidica è stata già descritta nel lavoro di Biemelt *et al.* (2003) e non ha prodotto nessun cambiamento negativo nei livelli di espressione. Inoltre, come descritto da Kozak (1986), la presenza di una G in posizione +4 può migliorare l'efficienza di traduzione. Il gene  $L1^{pt}$ , anch'esso mancante del segnale NLS, è stato clonato sia nel vettore per l'espressione nel citoplasma, che nel vettore con il segnale di indirizzamento per il cloroplasto. Inoltre, per evitare che sequenze particolari potessero formare delle strutture che bloccano la traduzione, anche in questo caso, è stata effettuata la delezione degli ultimi 205 aminoacidi per esprimere una proteina più piccola.

La sequenza  $L1^h$  è stata disegnata considerando un "codon usage" umanizzato. L'espressione in pianta del gene sintetico *L1* di HPV16 ottimizzato per

l'espressione in cellule umane, è descritta nel lavoro di Biemelt *et al.*, (2003). Questo gene, è stato disponibile nel nostro laboratorio negli ultimi mesi di dottorato, ed è stato clonato nei vettori I.V. per l'espressione nel citoplasma e nel cloroplasto.

In totale sono stati costruiti 13 vettori contenenti il gene *L1*.

In seguito ad agroinfiltrazione in *N. tabacum* di tutti i vettori contenenti il gene *L1<sup>v</sup>*, è stata ricercata la presenza della proteina ma in nessun compartimento cellulare è stato possibile osservare la presenza della proteina.

Essendo stati costruiti numerosi vettori per il gene *L1*, si è deciso di cominciare ad analizzare più in dettaglio l'espressione dei geni virali *L1<sup>v</sup>Δ22*, *L1<sup>v</sup>Δ205*, *L1<sup>pt</sup>Δ22* e *L1<sup>pt</sup>Δ205* clonate indirizzate sia nel citosol che nei cloroplasti. I vettori sono stati agroinfiltrati sia in *N. tabacum* che in *N. benthamiana*. L'espressione dei trascritti relativi ai diversi geni è stata verificata mediante RT-PCR ed i risultati hanno mostrato un trascritto abbondante con tutti i vettori analizzati. Tuttavia, in entrambe le specie vegetali, non è stato possibile rilevare la presenza della proteina, utilizzando sia anticorpi policlonali contro il gene *L1*, sia anticorpi diretti contro le "tag" c-myc e 6His.

Un altro caso simile, oltre a quelli descritti in precedenza per il gene *gag*, è riportato nel lavoro di Ziegelhoffer *et al.* (2001) dove il gene *E1* è stato indirizzato in diversi compartimenti subcellulari. Anche in questo caso, i livelli di mRNA relativi al gene d'interesse erano costanti per tutti i costrutti analizzati, ma l'accumulo della proteina si è evidenziato solo quando il gene *E1* è stato indirizzato nella via di secrezione.

Gli unici lavori di *L1* espresso in un sistema vegetale riportati fino ad oggi in letteratura sono cinque: (Biemelt *et al.*, 2003; Liu *et al.*, 2005; Varsani *et al.*, 2003; Varsani *et al.*, 2006; Warzecha *et al.*, 2003). Biemelt e altri (2003) hanno evidenziato nel loro lavoro l'importanza della 5'UTR per ottenere una proteina stabile. Infatti senza la 5'UTR, l'RNA si degradava facilmente e non era possibile evidenziare l'accumulo della proteina *L1* nei tessuti vegetali. Nei nostri costrutti è stata utilizzata la 5'UTR del gene *rbcS*, mentre Biemelt e collaboratori (2003) hanno utilizzato la 5'UTR  $\Omega$  del virus del mosaico del tabacco (TMV). Gli stessi autori hanno utilizzato nel loro lavoro un gene *L1* ottimizzato per l'espressione in pianta, ma con scarsi risultati. Liu ed altri (2005) invece, hanno espresso in piante transgeniche di tabacco il gene *L1* "wildtype" di HPV16 anche se a livelli molto bassi (0.034%–0.076% delle PTS).

Il lavoro di Varsani e collaboratori del 2006 è il primo a riportare l'espressione del gene *L1* "wildtype" in *N. benthamiana* attraverso l'uso di un vettore virale TMV. Gli autori sono riusciti ad accumulare la proteina *L1* fino a livelli di 20–37 ug/kg di tessuto fresco infettato, circa 10 volte in più di quanto ottenuto dagli stessi autori nel lavoro pubblicato nel 2003 utilizzando la trasformazione stabile. In questo lavoro, inoltre, hanno dimostrato che la maggior parte di *L1* prodotta nei tessuti vegetali non è assemblata a formare VLPs, ma che sono presenti per la maggior parte solo capsomeri. Ciò può essere dovuto alla bassa concentrazione di monomeri presenti nelle cellule vegetali (Varsani *et al.*, 2003).

Sulla base della differenza di espressione tra le due specie vegetali riguardo al gene *gag* descritta precedentemente, l'espressione del gene *L1<sup>h</sup>* è stata provata solo in *N. benthamiana*. I tessuti sono stati inoculati con i vettori contenenti la sequenza *L1<sup>h</sup>* indirizzata sia nel citoplasma che nel cloroplasto.

I risultati hanno mostrato la presenza di un trascritto delle dimensioni attese con entrambi i vettori. Al contrario di quanto osservato con le altre due sequenze di

*L1*, in questo caso si è riusciti a visualizzare una banda molto debole delle dimensioni attese corrispondente alla proteina *L1*. L'accumulo della proteina, seppure molto modesto, è paragonabile sia se *L1* è espressa nel citoplasma che nel cloroplasto.

I nostri risultati sono in linea con quanto riportato da Biemelt *et al.* (2003). Infatti, gli autori di questo lavoro, oltre ad utilizzare un gene ottimizzato per l'espressione in pianta, hanno espresso in *N. tabacum* anche la sequenza *L1* ottimizzata per l'espressione in cellule umane, riuscendo ad ottenere, fino allo 0,4% delle PTS. Le maggiori differenze tra questo lavoro e quello di Biemelt e altri (2003) riguardano il sistema di espressione, il promotore utilizzato, e la 5'UTR. Nel loro lavoro gli autori citati hanno ottenuto piante transgeniche di tabacco, dove il gene *L1* umanizzato è guidato dal promotore 35S e dalla 5'UTR  $\Omega$  del TMV. Probabilmente, l'uso di diversi promotori e 5'UTR nel nostro caso potrebbe portare ad un aumento dell'accumulo della proteina nei tessuti vegetali, almeno ai livelli ottenuti dagli autori citati.

Nel riassunto presentato da MacLean *et al.* (2005) si riporta l'espressione transiente mediante agroinfiltrazione del gene *L1* di HPV16, sia "wildtype", sia ottimizzato per pianta, o ottimizzato per cellule umane, mancante del segnale di localizzazione nucleare. Inoltre, tutte le sequenze geniche sono state indirizzate nel citoplasma, nel reticolo endoplasmico e nei plastidi e testate in *N. tabacum* e *N. benthamiana*. Gli autori riportano che l'accumulo della proteina si è ottenuta in entrambe le specie vegetali, ma la sequenza *L1* ottimizzata per l'espressione in cellule umane, ha dato livelli di espressione circa 300 volte più alta rispetto al gene *L1* non ottimizzato o con il "codon usage" per pianta. Inoltre, tra i differenti compartimenti subcellulari dove la proteina è stata indirizzata, i cloroplasti hanno dato i migliori risultati con livelli di 50 mg di proteina ricombinante per Kg di *N. benthamiana* infettata.

Tuttavia, fino ad oggi il migliore sistema di espressione in pianta per la proteina *L1* sembra essere quello dei vettori virali, come dimostrato dai riassunti presentati da Santi *et al.* (2005) e Smidkova *et al.* (2005). Smidkova e collaboratori (2005) riportano l'espressione in *N. benthamiana* e pomodoro della proteina *L1*. Anche qui hanno espresso due sequenze del gene *L1*, una ottimizzata per l'espressione in cellule umane ed un'altra ottimizzata per l'espressione in pianta attraverso l'uso di due vettori virali, PVX e tobavirus (PTV100). Gli autori indicano che con entrambi i geni virali sono stati ottenuti livelli di trascritti comparabili, ma l'accumulo della proteina è stato notevolmente più alto (> 0,5% TPS) utilizzando il gene *L1* umanizzato. Santi ed altri (2005), invece, utilizzando due tipi di vettori virali per esprimere la sequenza *L1* di HPV16 in *N. benthamiana*, hanno ottenuto livelli di espressione di circa l'1% delle PTS, confermando anche in questo caso, che l'uso di vettori virali sembra essere una ottima strategia per l'espressione del gene *L1* in pianta.

## 5. CONCLUSIONI

Obiettivo di questo progetto di Dottorato era l'espressione in pianta della proteina Gag ed L1 del capsido del virus HIV-1 e HPV16 rispettivamente. La produzione in pianta di tali proteine potrebbe portare allo sviluppo di un vaccino di derivazione vegetale.

A tal fine sono stati costruiti 13 vettori contenenti il gene *gag* e le relative subunità p17, p24,  $\Delta$ p17 (tutto il gene *gag* mancante del dominio p17), 13 vettori per il gene *L1* e 5 vettori contenenti il gene reporter *gus-int*. Le proteine d'interesse sono state indirizzate in diversi compartimenti cellulari come: citoplasma, apoplasto, reticolo endoplasmico, cloroplasto e mitocondrio.

Tutti i vettori sono stati testati in esperimenti di espressione transiente mediante agroinfiltrazione in piante di *Nicotiana tabacum* e *N. benthamiana*. I risultati ottenuti con i vettori contenenti il gene reporter *gus-int* hanno mostrato il corretto funzionamento dei vettori costruiti ed hanno permesso l'ottimizzazione della tecnica di agroinfiltrazione.

I risultati ottenuti con i vettori contenenti le sequenze d'interesse hanno evidenziato, indipendentemente dalla specie vegetale utilizzata, un trascritto abbondante e stabile relativo ai geni d'interesse.

L'analisi western ha mostrato la presenza della proteina Gag solo quando la proteina è stata fatta accumulare nei cloroplasti di *N. benthamiana*. In particolare si è evidenziato un parziale processamento di tale proteina similmente a quanto osservato in cellule umane infettate dal virus.

I risultati delle analisi western relativi alla subunità p24 hanno mostrato la presenza della proteina p24, mentre le analisi molecolari condotte sui tessuti agroinfiltrati con la subunità  $\Delta$ p17 hanno evidenziato sempre una banda di 24 KDa, indicando che la proteina era processata dopo il dominio p24, come il gene intero *gag*. Nonostante l'espressione dei geni sia stata ottenuta in entrambe le specie vegetali, i livelli d'espressione in *N. tabacum* sono risultati essere molto più bassi rispetto a quanto osservato in *N. benthamiana*. Infine, l'accumulo della proteina P17 è stato ottenuto solo quando la subunità è stata indirizzata nei cloroplasti di *N. benthamiana*.

L'accumulo della proteina L1, invece, non è stato osservato in nessun compartimento cellulare, utilizzando sia la sequenza  $L1^V$ , sia la sequenza  $L1^{pt}$ . Gli stessi risultati negativi sono stati ottenuti utilizzando due sequenze nucleotidiche codificanti per la stessa proteina senza gli ultimi 205 aminoacidi. Solo utilizzando la sequenza *L1* ottimizzata per l'espressione in cellule umane, è stato possibile osservare in *N. benthamiana* una debole banda del peso molecolare atteso.

In conclusione, i risultati ottenuti indicano la possibilità di esprimere in pianta la proteina Gag, sebbene ulteriori studi sono necessari per quanto riguarda il suo processamento e la possibilità di ottenere VLPs (Virus Like Particles) per la produzione di un vaccino profilattico. Il cloroplasto è risultato il miglior compartimento subcellulare per l'accumulo della proteina.

Al contrario, notevoli difficoltà sembrano legate all'espressione stabile di L1 in cellule vegetali.

## 6. BIBLIOGRAFIA

- Anderson, S. and Smith, S.M. (1986) Synthesis of the small subunit of ribulose-bisphosphate carboxylase from genes cloned into plasmids containing the SP6 promoter. *Biochem J*, **240**, 709-715.
- Badillo, J., Newell, C., and Gray J. (2005) Expression of HIV-1 p24 in chloroplasts of transplastomic tobacco plants. *Plant-Based Vaccines & Antibodies. PBVA 2005 – 8-10 June 2005, Prague, Czech Republic*.
- Barta, A., Sommerbruger, K., Thompson, D., Hartmuth, K., Matzke, M.A., Matzke, A.J.M. (1986) The expression of a nopaline synthase — human growth hormone chimaeric gene in transformed tobacco and sunflower callus tissue. *Plant Mol Biol* **6**, 347-357.
- Bevan, M. (1984) Binary *Agrobacterium* vectors for plant transformation. *Nucleic Acids Res*, **12**, 8711-8721.
- Biemelt, S., Sonnewald, U., Galmbacher, P., Willmitzer, L., Muller, M. (2003) Production of human papillomavirus type 16 virus-like particles in transgenic plants. *J Virol*, **77**, 9211-9220.
- Birch, R.G. (1997) Plant transformation: Problems and Strategies for Practical Application. *Annu Rev Plant Physiol Plant Mol Biol*, **48**, 297-326.
- Bousarghin, L., Touze, A., Combata-Rojas, A.L. and Coursaget, P. (2003) Positively charged sequences of human papillomavirus type 16 capsid proteins are sufficient to mediate gene transfer into target cells via the heparan sulfate receptor. *J Gen Virol*, **84**, 157-164.
- Buonaguro, L., Buonaguro, F.M., Tornesello, M.L., Mantas, D., Beth-Giraldo, E., Wagner, R., Michelson, S., Prevost, M.C., Wolf, H. and Giraldo, G. (2001) High efficient production of Pr55 (*gag*) virus-like particles expressing multiple HIV-1 epitopes, including a gp120 protein derived from an Ugandan HIV-1 isolate of subtype A. *Antiviral Res*, **49**, 35-47.
- Casson, P. (1996) Direct interaction between the envelope and matrix protein of HIV-1. *EMBO J.*, **15**, 5783-5788.
- Chen, L., Marmey, P., Taylor, N.J., Brizard, J.P., Espinoza, C., D'Cruz, P., Huet, H., Zhang, S., de Kochko, A., Beachy, R.N. and Fauquet, C.M. (1998) Expression and inheritance of multiple transgenes in rice plants. *Nat Biotechnol*, **16**, 1060-1064.
- Chong, D.K. and Langridge, W.H. (2000) Expression of full-length bioactive antimicrobial human lactoferrin in potato plants. *Transgenic Res*, **9**, 71-78.
- Christou, P. (1993) Particle gun-mediated transformation. *Curr Opin Biotech*, **4**, 135-141.
- Conrad, U. and Fiedler, U. (1998) Compartment-specific accumulation of recombinant immunoglobulins in plant cells: an essential tool for antibody production and immunomodulation of physiological functions and pathogen activity. *Plant Mol Biol*, **38**, 101-109.
- Cramer, C.L., Boothe, J.G. and Oishi, K.K. (1999) Transgenic plants for therapeutic proteins: linking upstream and downstream strategies. *Curr Top Microbiol Immunol*, **240**, 95-118.
- Daniell, H. (1993) Foreign gene expression in chloroplasts of higher plants mediated by tungsten particle bombardment. *Methods Enzymol*, **217**, 536-556.
- Daniell, H., Khan, M.S. and Allison, L. (2002) Milestones in chloroplast genetic engineering: an environmentally friendly era in biotechnology. *Trends Plant Sci*, **7**, 84-91.

- Daniell, H., Streatfield, S.J. and Wycoff, K. (2001) Medical molecular farming: production of antibodies, biopharmaceuticals and edible vaccines in plants. *Trends Plant Sci*, **6**, 219-226.
- Dean, C., Favreau, M., Bond-Nutter, D., Bedbrook, J. and Dunsmuir, P. (1989) Sequences downstream of translation start regulate quantitative expression of two petunia *rbcS* genes. *Plant Cell*, **1**, 201-208.
- Earl, P.L., Hugin, A.W. and Moss, B. (1990) Removal of cryptic poxvirus transcription termination signals from the human immunodeficiency virus type 1 envelope gene enhances expression and immunogenicity of a recombinant vaccinia virus. *J Virol*, **64**, 2448-2451.
- Ferrando, A., Koncz-Kalman, Z., Farras, R., Tiburcio, A., Schell, J. and Koncz, C. (2001) Detection of in vivo protein interactions between Snf1-related kinase subunits with intron-tagged epitope-labelling in plants cells. *Nucleic Acids Res*, **29**, 3685-3693.
- Fischer, R. and Emans, N. (2000) Molecular farming of pharmaceutical proteins. *Transgenic Res*, **9**, 279-299; discussion 277.
- Florin, L., Sapp, C., Streeck, R.E. and Sapp, M. (2002) Assembly and translocation of papillomavirus capsid proteins. *J Virol*, **76**, 10009-10014.
- Gallie, D.R. and Walbot, V. (1992) Identification of the motifs within the tobacco mosaic virus 5'-leader responsible for enhancing translation. *Nucleic Acids Res*, **20**, 4631-4638.
- Giddings, G., Allison, G., Brooks, D., Carter, A. (2000) Transgenic plants as factories for biopharmaceuticals. *Nat Biotechnol* **18**, 1151-1155.
- Haq, T.A., Mason, H.S., Clements, J.D. and Arntzen, C.J. (1995) Oral immunization with a recombinant bacterial antigen produced in transgenic plants. *Science*, **268**, 714-716.
- Heifetz, P.B. (2000) Genetic engineering of the chloroplast. *Biochimie*, **82**, 655-666.
- Hoekema, A., Hirsch, P.R., Hooykaas, P.J.J., Schilperoort, R.A. (1983) A binary plant vector strategy based on separation of vir and T-region of the *Agrobacterium tumefaciens* Ti-plasmid. *Nature*, **303**, 179-180.
- Hood, E.E., Witcher, D.R., Maddock, S., Meyer, T., Baszczyński, C., Bailey, M., Flynn, P., Register, J., Marshall, L., Bond, D., Kulisek, E., Kusnadi, A., Evangelista, R., Nikolov, Z., Wooge, C., Mehig, R.J., Hernan, R., Kappel, W.K., Ritland, D., Ping Li, C., Howard, J.A. (1997) Commercial production of avidin from transgenic maize: characterization of transformant, production, processing, extraction and purification. *Mol Breeding* **3**, 291-306.
- Hooykaas, P.J.S., Mozo T. (1994). *Agrobacterium* molecular genetics. *Plant Molecular Biology Manual* **B3**, 1-9.
- Huang, Z., Santi, L., LePore, K., Kilbourne, J., Arntzen, C.J. and Mason, H.S. (2006) Rapid, high-level production of hepatitis B core antigen in plant leaf and its immunogenicity in mice. *Vaccine*, **24**, 2506-2513.
- James, E.A., Wang, C., Wang, Z., Reeves, R., Shin, J. H., Magnuson, N. S., Lee, J. M. (2000) Production and characterization of biologically active human GM-CSF secreted by genetically modified plant cells. *Protein Expr Purif*, **19**, 131-138.
- Jang, C., Nahm, B.H., Kim J.K. (1999) Subcellular targeting of green fluorescent protein to plastids in transgenic rice plants provides a high-level expression system. *Molecular Breeding*, **5**, 453-461.
- Jefferson, R.A. (1987) Assaying chimeric genes in plants: the *gus* gene fusion system. *Plant Mol Biol Rep* **5**, 387-405.

- Joh, L.D., Wroblewski, T., Ewing, N.N. and Vander Gheynst, J.S. (2005) High-level transient expression of recombinant protein in lettuce. *Biotechnol Bioeng*, **91**, 861-871.
- Kapila, J., De Rycke, R., Van Montagy, M., and Angenon, G. (1997) An *Agrobacterium*-mediated transient gene expression system for intact leaves. *Plant Science*, **122**, 101-108.
- Kapusta, J., Modelska, A., Figlerowicz, M., Pniewski, T., Letellier, M., Lisowa, O., Yusibov, V., Koprowski, H., Plucienniczak, A., Legocki, A. B. (1999) A plant-derived edible vaccine against hepatitis B virus. *Faseb J*, **13**, 1796-1799.
- Kaushik, R. and Ratner, L. (2004) Role of human immunodeficiency virus type 1 matrix phosphorylation in an early postentry step of virus replication. *J Virol*, **78**, 2319-2326.
- Köhler, R., Zipfel, W., Webb, W.W., Warren, R., Hanson, M.R. (1997) The green fluorescent protein as a marker to visualize plant mitochondria in vivo. *PBJ* **11**, 613-619.
- Koscianska, E., Kalantidis, K., Wypijewski, K., Sadowski, J. and Tabler, M. (2005) Analysis of RNA silencing in agroinfiltrated leaves of *Nicotiana benthamiana* and *Nicotiana tabacum*. *Plant Mol Biol*, **59**, 647-661.
- Koya, V., Moayeri, M., Leppla, S.H. and Daniell, H. (2005) Plant-based vaccine: mice immunized with chloroplast-derived anthrax protective antigen survive anthrax lethal toxin challenge. *Infect Immun*, **73**, 8266-8274.
- Kusnadi, A., Nikolov, Z. and Howard, J. (1997) Production of recombinant proteins in transgenic plants: practical considerations. *Biotechnol Bioeng*, **56**, 473-484.
- Laemmli, U.K. (1970) Cleavage of structural proteins during the assembly of the head of bacteriophage T4. *Nature*, **227**, 680-685.
- Lee, M.W. and Yang, Y. (2006) Transient expression assay by agroinfiltration of leaves. *Methods Mol Biol*, **323**, 225-229.
- Lessard, P.A., Kulaveerasingam, H., York, G.M., Strong, A. and Sinskey, A.J. (2002) Manipulating gene expression for the metabolic engineering of plants. *Metab Eng*, **4**, 67-79.
- Li, H.Y., Ramalingam, S. and Chye, M.L. (2006) Accumulation of recombinant SARS-CoV spike protein in plant cytosol and chloroplasts indicate potential for development of plant-derived oral vaccines. *Exp Biol Med* **231**, 1346-1352.
- Liu, H.L., Li, W. S., Lei, T., Zheng, J., Zhang, Z., Yan, X. F., Wang, Z. Z., Wang, Y. L., and Si, L. S. (2005) Expression of human papillomavirus type 16 L1 protein in transgenic tobacco plants. *Acta Biochim Biophys Sin* **37**, 153-158.
- Loh, Y.T., Zhou, J. and Martin, G.B. (1998) The myristylation motif of Pto is not required for disease resistance. *Mol Plant Microbe Interact*, **11**, 572-576.
- Ma, J.K., Chikwamba, R., Sparrow, P., Fischer, R., Mahoney, R. and Twyman, R.M. (2005) Plant-derived pharmaceuticals--the road forward. *Trends Plant Sci*, **10**, 580-585.
- Ma, J.K., Drake, P.M. and Christou, P. (2003) The production of recombinant pharmaceutical proteins in plants. *Nat Rev Genet*, **4**, 794-805.
- Ma, J.K., Hikmat, B.Y., Wycoff, K., Vine, N.D., Chargelegue, D., Yu, L., Hein, M.B. and Lehner, T. (1998) Characterization of a recombinant plant monoclonal secretory antibody and preventive immunotherapy in humans. *Nat Med*, **4**, 601-606.
- Ma, J.K., Barros, E., Bock, R., Christou, P., Dale, P.J., Dix P.J., Fischer, R., Irwin, J., Mahoney, R., Pezzotti, M., Schillberg, S., Sparrow, P., Stoger, E. and Twyman, R.M. (2006) Molecular farming for new drugs and vaccines. *Embo Rep* **6**, 592-599.

- MacLean, J., Becker, I., Williamson, A.L., Rybicki, E. (2005) Towards optimal HPV-16 L1 expression in plants: comparison of the suitability of different HPV-16 L1 variants and cell compartment localization by transient expression in plants. *Proceedings of the 22nd International Papillomavirus Conference and Clinical Workshop, Vancouver, British Columbia, Canada. April 30 - May 6.*
- Maliga, P. (2004) Plastid transformation in higher plants. *Annu Rev Plant Biol*, **55**, 289-313.
- Mason, H.S., Ball, J.M., Shi, J.J., Jiang, X., Estes, M.K. and Arntzen, C.J. (1996) Expression of Norwalk virus capsid protein in transgenic tobacco and potato and its oral immunogenicity in mice. *Proc Natl Acad Sci U S A*, **93**, 5335-5340.
- Mason, H.S., Haq, T.A., Clements, J.D. and Arntzen, C.J. (1998) Edible vaccine protects mice against *Escherichia coli* heat-labile enterotoxin (LT): potatoes expressing a synthetic LT-B gene. *Vaccine*, **16**, 1336-1343.
- Mason, H.S., Lam, D.M. and Arntzen, C.J. (1992) Expression of hepatitis B surface antigen in transgenic plants. *Proc Natl Acad Sci U S A*, **89**, 11745-11749.
- McGarvey, P.B., Hammond, J., Dienelt, M.M., Hooper, D.C., Fu, Z.F., Dietzschold, B., Koprowski, H. and Michaels, F.H. (1995) Expression of the rabies virus glycoprotein in transgenic tomatoes. *Biotechnology*, **13**, 1484-1487.
- Meyers, A., Lynch, A., Tanzer, F., Williamson, A.L., Rybicki EP (2005) Comparisons of human immunodeficiency virus type 1 subtype C Pr55Gag production in plants using different expression techniques. *Plant Expression Systems for Recombinant Pharmacologics PBVA 2005 – 8-10 June 2005, Prague, Czech Republic, Poster 3.*
- Morikawa, Y. (2003) HIV capsid assembly. *Curr HIV Res*, **1**, 1-14.
- Murashige, T., and Skoog, F. (1962) A revised medium from rapid growth and bioassay with tobacco tissue cultures. *Physiol Plant*, **15**, 473-497.
- Obregon, P., Chargelegue, D., Drake, M. W., Prada, A., Nuttall, J., Frigerio, L., Ma J.K.C. (2006) HIV-1 p24-immunoglobulin fusion molecule: a new strategy for plant-based protein production. *Plant Biotechnology Journal* **4**, 195-207.
- Outchkourov, N.S., Peters, J., de Jong, J., Rademakers, W., Jongsma, M. A. (2003) The promoter-terminator of chrysanthemum *rbcS1* directs very high expression levels in plants. *Planta*, **216**, 1003-1012.
- Paintsil, J., Muller, M., Picken, M., Gissmann, L. and Zhou, J. (1996) Carboxyl terminus of bovine papillomavirus type-1 L1 protein is not required for capsid formation. *Virology*, **223**, 238-244.
- Parkin, D.M., Bray, F.I. and Devesa, S.S. (2001) Cancer burden in the year 2000. The global picture. *Eur J Cancer*, **37 Suppl 8**, S4-66.
- Perez-Filgueira, D.M., Brayfield, B. P., Phiri, S., Borca, M. V., Wood, C., Morris, T. J. (2004) Preserved antigenicity of HIV-1 p24 produced and purified in high yields from plants inoculated with a tobacco mosaic virus (TMV)-derived vector. *J Virol Methods*, **121**, 201-208.
- Petolino, J.F., Young, S., Hopkins, N., Sukhapinda, K., Woosley, A., Hayes, C. and Pelcher, L. (2000) Expression of murine adenosine deaminase (ADA) in transgenic maize. *Transgenic Res*, **9**, 1-9.
- Raskin, I., Ribnicky, D.M., Komarnytsky, S., Ilic, N., Poulev, A., Borisjuk, N., Brinker, A., Moreno, D.A., Ripoll, C., Yakoby, N., O'Neal, J.M., Cornwell, T., Pastor, I. and Fridlender, B. (2002) Plants and human health in the twenty-first century. *Trends Biotechnol*, **20**, 522-531.

- Reed, D.G., Nopo-Olazabal, L.H., Funk, V., Woffenden, B.J., Reidy, M.J., Dolan, M.C., Cramer, C.L. and Medina-Bolivar, F. (2005) Expression of functional hexahistidine-tagged ricin B in tobacco. *Plant Cell Rep*, **24**, 15-24.
- Richter, L.J., Thanavala, Y., Arntzen, C. J., Mason, H. S. (2000) Production of hepatitis B surface antigen in transgenic plants for oral immunization. *Nat Biotechnol*, **18**, 1167-1171.
- Richter, L.K. (1999) Transgenic plants as edible vaccines. *Curr Top Microbiol Immunol.*, **240**, 159-176.
- Romano, A., Raemakers, K., Bernardi, J., Visser, R. and Mooibroek, H. (2003) Transgene organisation in potato after particle bombardment-mediated co-transformation using plasmids and gene cassettes. *Transgenic Res*, **12**, 461-473.
- Ronco, G., Segman, N., Giorgi, P., Zappa, M., Casadei, P. and Carrozzi, F. (2006) Human Papillomavirus testing and liquid-based cytology: results at recruitment from the New Technologies for Cervical Cancer randomized controlled trial. *J Natl Cancer Inst.* **98**: 765-774.
- Ruggiero, F., Exposito, J.Y., Bournat, P., Gruber, V., Perret, S., Comte, J., Olagnier, B., Garrone, R. and Theisen, M. (2000) Triple helix assembly and processing of human collagen produced in transgenic tobacco plants. *FEBS Lett*, **469**, 132-136.
- Rumeau, D., Becuwe-Linka, N., Beyly, A., Louwagie, M., Garin, J. and Peltier, G. (2005) New subunits NDH-M, -N, and -O, encoded by nuclear genes, are essential for plastid Ndh complex functioning in higher plants. *Plant Cell*, **17**, 219-232.
- Sambrook, J., Rusell, D. W. (2001) Molecular cloning: a laboratory manual, 3rd ed. *Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.*
- Sanford, J. (1990) Biolistic plant transformation. *Physiol. Plant.* **79**, 206-209.
- Sanger, F., Nicklen, S. and Coulson, A.R. (1977) DNA sequencing with chain-terminating inhibitors. *Proc Natl Acad Sci U S A*, **74**, 5463-5467.
- Santi, L., Kim, H., Gonzales, B., Yuon, S., Rose, R., Mason, H. (2005) Viral vector expression of HPV 16 L1 Plants. *Proceedings of the 22nd International Papillomavirus Conference and Clinical Workshop, Vancouver, British Columbia, Canada. April 30 - May 6.*
- Schouten, A., Roosien, J., van Engelen, F.A., de Jong, G.A., Borst-Vrenssen, A.W., Zilverentant, J.F., Bosch, D., Stiekema, W.J., Gommers, F.J., Schots, A. and Bakker, J. (1996) The C-terminal KDEL sequence increases the expression level of a single-chain antibody designed to be targeted to both the cytosol and the secretory pathway in transgenic tobacco. *Plant Mol Biol*, **30**, 781-793.
- Sijmons, P.C., Dekker, B.M., Schrammeijer, B., Verwoerd, T.C., van den Elzen, P.J. and Hoekema, A. (1990) Production of correctly processed human serum albumin in transgenic plants. *Biotechnology* **8**, 217-221.
- Smidkova, M., Mueller, M., Thoenes, N., Veleminsky J., Angelis K.J. (2005) Production of Human Papilloma Virus 16 protein L1h 16 in plants. *Plant-Based Vaccines & Antibodies PBVA 2005 – 8-10 June 2005, Prague, Czech Republic.*
- Smith, M.L., Mason, H.S. and Shuler, M.L. (2002) Hepatitis B surface antigen (HBsAg) expression in plant cell culture: Kinetics of antigen accumulation in batch culture and its intracellular form. *Biotechnol Bioeng*, **80**, 812-822.
- Staub, J.M., Garcia, B., Graves, J., Hajdukiewicz, P.T., Hunter, P., Nehra, N., Paradkar, V., Schlittler, M., Carroll, J.A., Spatola, L., Ward, D., Ye, G., Russell, D.A. (2000) High-yield production of a human therapeutic protein in tobacco chloroplasts. *Nat Biotechnol*, **18**, 333-338.

- Streatfield, S.J., Jilka, J. M., Hood, E. E., Turner, D. D., Bailey, M. R., Mayor, J. M., Woodard, S. L., Beifuss, K. K., Horn, M. E., Delaney, D. E., Tizard, I. R., Howard, J. A. (2001) Plant-based vaccines: unique advantages. *Vaccine*, **19**, 2742-2748.
- Svab, Z. and Maliga, P. (1993) High-frequency plastid transformation in tobacco by selection for a chimeric aadA gene. *Proc Natl Acad Sci U S A*, **90**, 913-917.
- Tackaberry, E.S., Dudani, A. K., Prior, F., Tocchi, M., Sardana, R., Altosaar, I., Ganz, P. R. (1999) Development of biopharmaceuticals in plant expression systems: cloning, expression and immunological reactivity of human cytomegalovirus glycoprotein B (UL55) in seeds of transgenic tobacco. *Vaccine*, **17**, 3020-3029.
- Tacket, C.O., Mason, H. S., Losonsky, G., Clements, J. D., Levine, M. M., Arntzen, C. J. (1998) Immunogenicity in humans of a recombinant bacterial antigen delivered in a transgenic potato. *Nat Med*, **4**, 607-609.
- Tacket, C.O., Mason, H. S., Losonsky, G., Estes, M. K., Levine, M. M., Arntzen, C. J. (2000) Human immune responses to a novel norwalk virus vaccine delivered in transgenic potatoes. *J Infect Dis*, **182**, 302-305.
- Tregoning, J.S., Nixon, P., Kuroda, H., Svab, Z., Clare, S., Bowe, F., Fairweather, N., Ytterberg, J., van Wijk, K. J., Dougan, G., Maliga, P. (2003) Expression of tetanus toxin Fragment C in tobacco chloroplasts. *Nucleic Acids Res*, **31**, 1174-1179.
- Tuboly, T., Yu, W., Bailey, A., Degrandis, S., Du, S., Erickson, L. and Nagy, E. (2000) Immunogenicity of porcine transmissible gastroenteritis virus spike protein expressed in plants. *Vaccine*, **18**, 2023-2028.
- Valdez-Ortiz, A., Rascon-Cruz, Q., Medina-Godoy, S., Sinagawa-Garcia, S.R., Valverde-Gonzalez, M.E. and Paredes-Lopez, O. (2005) One-step purification and structural characterization of a recombinant His-tag 11S globulin expressed in transgenic tobacco. *J Biotechnol*, **115**, 413-423.
- Vancanneyt, G., Schmidt, R., O'Connor-Sanchez, A., Willmitzer, L., Rocha-Sosa, M. (1990) Construction of an intron-containing marker gene: splicing of the intron in transgenic plants and its use in monitoring early events in *Agrobacterium*-mediated plant transformation. *Mol Gen Genet* **220**, 245-250.
- Varsani, A., Williamson, A. L., Rose, R. C., Jaffer, M., Rybicki, E. P. (2003) Expression of Human papillomavirus type 16 major capsid protein in transgenic *Nicotiana tabacum* cv. Xanthi. *Arch Virol*, **148**, 1771-1786.
- Varsani, A., Williamson, A. L., Stewart, D., Rybicki, E. P. (2006) Transient expression of Human papillomavirus type 16 L1 protein in *Nicotiana benthamiana* using an infectious tobamovirus vector. *Virus Res*, **120**, 91-96.
- Voinnet, O., Rivas, S., Mestre, P., Baulcombe, D. (2003) An enhanced transient expression system in plants based on suppression of gene silencing by the p19 protein of tomato bushy stunt virus. *Plant J*, **33**, 949-956.
- Warzecha, H., Mason, H. S., Lane, C., Tryggvesson, A., Rybicki, E., Williamson, A. L., Clements, J. D., Rose, R. C. (2003) Oral immunogenicity of human papillomavirus-like particles expressed in potato. *J Virol*, **77**, 8702-8711.
- Wirth, S., Calamante, G., Mentaberry, A., Bussmann, L., Lattanzi, M., Barañao, L., Bravo-Almonacid, F. (2004) Expression of active human epidermal growth factor (hEGF) in tobacco plants by integrative and non-integrative systems. *Mol Breed* **13**, 23-35.
- Wong, E.Y., Hironaka, C.M. and Fischhoff, D.A. (1992) *Arabidopsis thaliana* small subunit leader and transit peptide enhance the expression of *Bacillus thuringiensis* proteins in transgenic plants. *Plant Mol Biol*, **20**, 81-93.

- Wright, K.E., Prior, F., Sardana, R., Altosaar, I., Dudani, A.K., Ganz, P.R. and Tackaberry, E.S. (2001) Sorting of glycoprotein B from human cytomegalovirus to protein storage vesicles in seeds of transgenic tobacco. *Transgenic Res*, **10**, 177-181.
- Wroblewski, T., Tomczak, A., Michelmore, R. (2005) Optimization of *Agrobacterium* - mediated transient assays of gene expression in lettuce, tomato and *Arapidopsis*. *Plant Biotech J* **3**, 259-273.
- Wydro, M., Kozubek, E., Lehmann, P. (2006) Optimization of transient *Agrobacterium*-mediated gene expression system in leaves of *Nicotiana benthamiana*. *Acta Biochim Pol*, **53**, 289-298.
- Yang, Y., Li, R. and Qi, M. (2000) *In vivo* analysis of plant promoters and transcription factors by agroinfiltration of tobacco leaves. *Plant J*, **22**, 543-551.
- Yusibov, V., Hooper, D.C., Spitsin, S.V., Fleysh, N., Kean, R.B., Mikheeva, T., Deka, D., Karasev, A., Cox, S., Randall, J. and Koprowski, H. (2002) Expression in plants and immunogenicity of plant virus-based experimental rabies vaccine. *Vaccine*, **20**, 3155-3164.
- Zeitlin, L., Olmsted, S.S., Moench, T.R., Co, M.S., Martinell, B.J., Paradkar, V.M., Russell, D.R., Queen, C., Cone, R.A. and Whaley, K.J. (1998) A humanized monoclonal antibody produced in transgenic plants for immunoprotection of the vagina against genital herpes. *Nat Biotechnol*, **16**, 1361-1364.
- Zhang, G.G., Rodrigues, L., Rovinski, B., White, K. A. (2002) Production of HIV-1 p24 protein in transgenic tobacco plants. *Mol Biotechnol*, **20**, 131-136.
- Zhang, W., Carmichael, J., Ferguson, J., Inglis, S., Ashrafian, H. and Stanley, M. (1998) Expression of human papillomavirus type 16 L1 protein in *Escherichia coli*: denaturation, renaturation, and self-assembly of virus-like particles in vitro. *Virology*, **243**, 423-431.
- Ziegelhoffer, T., Raasch, J.A., Phillips-Austin, S. (2001) Dramatic effects of truncation and sub-cellular targeting on the accumulation of recombinant microbial cellulase in tobacco. *Mol Breed* **8**, 147-158.

**APPENDICE**

Le attività svolte durante questo progetto di dottorato hanno permesso la presentazione dei risultati ottenuti a convegni nazionali ed internazionali sotto forma di poster o presentazioni orali quali:

**Ferraiolo E.**, Scotti N., Alagna F., Formisano G., Buonaguro L., Buonaguro F.M., Grillo S., Cardi T. (2006). *HIV-1 Gag protein processing in tobacco cells*. In 50° Annual Congress SIGA, Società Italiana di Genetica Agraria. Ischia, 10-14 settembre 2006.

**Ferraiolo E.** (2006 - **Presentazione Orale**). *Espressione di geni del Virus HIV-1 in piante di tabacco per la produzione di un vaccino utile per la salute umana*. In V° corso estivo di approfondimento sulla genetica vegetale. Isola Polvese (PG), 5-7 Giugno 2006.

Scotti N., Alagna F., **Ferraiolo E.**, Formisano G., Tornesello M.L., Buonaguro L., Buonaguro F.M., Monti L., Grillo S., Cardi T. (2006). *Tabacco plants as biofactories for the production of HIV-1 gag antigen*. In International Cancer Vaccine Conference in association with 8<sup>th</sup> ECEAR Conference. Napoli, 26-28 maggio 2006.

Lenzi P., Alagna F., **Ferraiolo E.**, Formisano G., Gargano D., Pompa A., Vitale A., Tornesello M.L., Maliga P., Buonaguro L., Buonaguro F.M., Grillo S., Scotti N., Cardi T. (2006). *Plastid and nuclear transformation for production of tobacco and potato plants expressing the HPV L1 gene*. In International Cancer Vaccine Conference in association with 8<sup>th</sup> ECEAR Conference. Napoli, 26-28 maggio 2006.

Formisano G., **Ferraiolo E.**, Scotti N., Tornesello M.L., Buonaguro L., Buonaguro F.M., Grillo S., Cardi T. (2005). *Production of transgenic tomato plants expressing the HIV-1 gag gene*. In XLIX Congresso annuale – SIGA, Società Italiana di Genetica Agraria. Potenza, 12-15 settembre 2005.

**Ferraiolo E.**, Formisano G., Scotti N., Tornesello M.L., Buonaguro L., Buonaguro F.M., Grillo S., Cardi T. (2005). *Production of HIV antigens in tobacco and tomato transgenic plants*. In PBVA, Plant-Based Vaccines & Antibodies. Praga, 8-10 giugno 2005.

Desidero ringraziare il *Dott. Teodoro Cardi* per avermi permesso di lavorare su questo progetto e per la preziosa guida scientifica che mi ha dato in questi anni.

Un ringraziamento alla *Dott.ssa Nunzia Scotti*, per i consigli scientifici e per la sua grande disponibilità.

Desidero ringraziare anche la *Dott.ssa Stefania Grillo*, per la fiducia che ha mostrato di avere nei miei confronti e per consigli che mi ha dato da quando frequento il laboratorio di Portici.

Ringrazio i *Dott. Franco Buonaguro, Luigi Buonaguro* e la *Dott.ssa M.L. Tornesello*, dell' I.N.T Pascale di Napoli per la collaborazione con questo progetto.

Un ringraziamento particolare e parte di questo lavoro va alla *Dott.ssa Gelsomina Formisano*, amica e collega di lavoro indispensabile. Il suo aiuto e i suoi consigli sono stati, per me, fondamentali in questi anni.

Un ringraziamento anche a Paolo, Daniela, Anna, Manuela, Vladimir, Carmela, Fiammetta, Alfonso, Roberta (grazie anche per i "frozen"), Rossella, Alfredo, Giorgia, Imma, Floriana, Paola, Antonello, Teresa e Mena. Grazie per l'aiuto, i consigli e per l'affiatamento che è nato tra di noi.

Un pensiero a mio padre, mia madre e mio fratello, per avermi sostenuto ed aiutato in questi anni.

Infine ringrazio Rosa, per i numerosi consigli che mi ha dato, per avermi ascoltato ogni volta che ne avevo bisogno e per essermi accanto da sempre.