EXPRESSION OF HPV-16 L2 IN PLANTS

RAMON PEREIRA

A thesis submitted to the Faculty of Science, University of Cape Town,
in fulfilment of the degree of Master of Science

February 2008

Dedication

This dissertation is dedicated to my grandfather, Clive DeWitt, who has always selflessly provided every opportunity to me, in the pursuit of my goals. Although at times unsure, he has always had my best interests at heart and I thank him for his unwavering belief in my ability to succeed.

TABLE OF CONTENTS

ABSTRACT

Cervical cancer is the second most prevalent cancer amongst women worldwide. Almost 99% of all cervical carcinomas have been linked to Human Papillomavirus (HPV). The various types of HPV can be distinguished between those that have a high-risk of causing cancer and low-risk. HPV-16 is high-risk and has been found to coincide with 60% of cervical carcinomas. The native virion contains two structural proteins; L1 and L2. L1 has been extensively characterized and has the ability to form virus-like particles (VLPs), a phenomenon currently available vaccines exploit. However, the immune response against L1 was found to be type-specific necessitating multivalent vaccines. In contrast, L2 has shown cross protection against different HPV types. In addition, it appears to enhance capsid formation making it an ideal candidate for a recombinantly expressed vaccine and/or vaccine enhancer.

Currently, the available vaccines are expensive, thus beyond the reach of developing countries where the vaccine is needed most. This is in part, due to the expensive expression system used in their production. Plant production provides a possibly cheaper platform. This study demonstrates high level expression of L2, 25 $mg.kg^{-1}$ of leaf material, is achievable. Interestingly, expression was best when coded for by a mammalian codon-optimized form of the L2 gene as opposed to the wildtype or plant codon-optimized (plantized) genes. Moreover, real time PCR revealed limited levels of transcript when coded for by the plantized gene in comparison to the other genes. A set of vectors which target the protein to the cytoplasm, the endoplasmic reticulum (ER), chloroplast or apoplastic space were

used and it was found that targeting of the protein had no effect on its expression levels. Furthermore, the results indicate a possible mitochondrial-targeting site in the N-terminus of the L2 protein, specifically recognized in plants, which may interfere with targeting to the ER and apoplastic space. However, targeting to the chloroplast appears to be unaffected.

Finally, co-expression of L1 and L2 was found to be possible in this system making formation of L1/L2 VLPs highly probable. This study represents the initial steps towards realizing the feasibility of a plant produced HPV L1/L2 vaccine that may be cheaper and more effective.

ACKNOWLEDGMENTS

I would like to formally thank the following:

My supervisor **Prof Ed Rybicki**, for his consistent help and steady guidance throughout.

My co-supervisor **Dr Inga Becker**, for her unwavering support and motivation when I needed it most.

My other co-supervisor, **Dr James Maclean**, for his valuable input and willingness to assist wherever possible.

The various **lab members** in the virology lab of the MCB department who were always willing to go beyond merely teaching a new technique.

Dr Suhail Rafudeen and **Dermot Cox** for their help in both formulating and interpreting the real time PCR work.

Keren Cooper and **Mohammed Jaffer** for their assistance in perfecting the electron microscopy technique, proving that with patience and tenacity it is possible to do your own EM work.

In addition, my grandfather, **Clive DeWitt**, to whom this thesis is dedicated, for his support, both financial and personal. Although, at times unsure, he has always had my best interests at heart and I thank him for the incredible opportunities he has constantly provided me with.

Finally, I would like to thank the **Poliomyelitis Research Fund** for providing the funding awarded to Ed Rybicki for this study.

1. CHAPTER 1

LITERATURE REVIEW

1.1 Introduction

The idea of a virus causing cancer is not a novel one; an estimated 15-20% of all cancers are directly linked to various viruses (41). Cervical cancer is the second most prevalent cancer amongst women worldwide with over 470,000 new cases reported annually and almost half that number dying in the same time (55). Cytological screens of cervical tumours have detected Human Papillomavirus (HPV) DNA in 99.7% of these cases (78). HPV has also been implicated in 85% of anal cancers; 50% of cancers of the vulva, vagina and penis; 20% of oropharyngeal cancers; and 10% of laryngeal and esophageal cancers (78). In addition, ~20 million people are infected with HPV in the US at any one time, making it the most commonly diagnosed, sexually transmitted disease (55,78).

HPVs belong to the genus Papillomavirus, which are known to infect a large number of mammals and birds (26). Papillomaviruses (PVs) have a natural tropism for epithelial cells, often entering through an abrasion that exposes the lower layers of the epithelium, more than likely targeting the stem cells in the basal layer (72,74). Humans are the most extensively studied host of PVs, with over 120 HPVs that have been characterized and many more thought to exist (26,78). HPVs are further divided into those infecting cutaneous or mucosal epithelium, with ~40 mucosal HPVs having been described (78). It is these HPVs infecting the mucosa of the anogenital tract that have raised concern.

Infection is characterized by unsightly genital warts, which are painful and costly to remove. Majority of infections are cleared naturally, with the wart spontaneously regressing; however, persistent infections could lead to carcinogenic development (55,74,78). HPVs that infect the mucosa are further subdivided into two groups, high-risk and low-risk, according to their propensity of infection to lead to a carcinoma developing. Low-risk HPVs such as HPV-6 and HPV-11 cause benign genital warts and are classified as possibly causing cancer (26). High-risk types include HPV-16, accounting for 60% of all cervical cancer cases, HPV-18, accounting for 10-20%, and HPV types 31, 33, 35, 39, 45, 51, 52, 56, 58, 59, 68, and 73, which account for the rest of cervical cancer cases (55,64,78). It is evident that a broad spectrum vaccine would be ideal to completely eradicate cervical cancer.

Two vaccines are presently available in certain countries which have proven highly effective at preventing HPV infection in human trials (74,78). A general vaccine model has been proposed: vaccination of females at age 12 and triennial cytological screening from age 25 (82). Using this model, a 94% decrease in cervical cancer deaths has been predicted if the vaccine is only 70% efficacious (82). Detection of abnormal cytology in papinocolau (pap) smears has lead to a marked decrease in cervical cancer cases in developed countries, coupled with the availability of vaccines; research in "first-world" countries will begin to wane. However, due to the socio-economic context within developing countries research into more effective, cheaper vaccines with possibly broader, cross-neutralizing activities will be emphasized.

Vaccine design has pivoted on the exquisite fusion of microbiology and immunology. As technology advances, "second generation" vaccines are inevitable, these will seek to decrease the cost of the vaccine and broaden the vaccine spectrum to include other HPV types (55). The minor capsid protein (L2)

of HPV has shown promise in both aspects. This chapter will attempt to explain where L2 fits into both the structure and the HPV life cycle and explain its relevant immunogenic properties.

1.2 The Viral Life Cycle: The role of L2

1.2.1 The Structure of the HPV Virion

The viral capsid structure is elegantly simple, yet at the same time complicated to conceptualize; understanding it, however, is necessary to get closer to designing a vaccine. The easiest way to comprehend the structure of the viral capsid is to use an analogy for the various terms used. The capsid proteins/molecules/monomers are the bricks; the capsomeres are the walls; the capsomers are the rooms; and the capsid is the house.

The HPV capsid is icosahedral in shape and arranged in a T=7 right-hand skew lattice (37). Only two proteins make up the capsid; the major capsid protein (L1) and the minor capsid protein (L2). Three hundred and sixty (360) molecules of L1 make up the capsid whilst the amount of L2 incorporated is unclear. The stoichiometric ratio of L1:L2 is speculated to be 30:1, meaning twelve L2 molecules per capsid (13,92). Five L1 molecules associate to form a pentamer/capsomere; each containing a hole in the centre in which the L2 may associate (13). L2 molecules appear to interact with each other in an intercapsomeric-dependent manner with the C-terminal of one molecule interacting with the N-terminal of another (13).

72 capsomeres make up the capsid; it is postulated that each L1 molecule in a capsomer contains a C-terminal tail that invades a neighbouring L1 molecule in an

adjacent capsomer forming a disulphide bond, which stabilizes the interaction upon maturation (14,56,69). In this way a three-fold and two-fold axes of symmetry are formed (Fig. 1) (19,56). The capsid is made up of both 60 hexavalent capsomers and 12 pentavalent capsomers which are stabilized by the inter-capsomeric disulphide bonds (Fig. 1) (18).

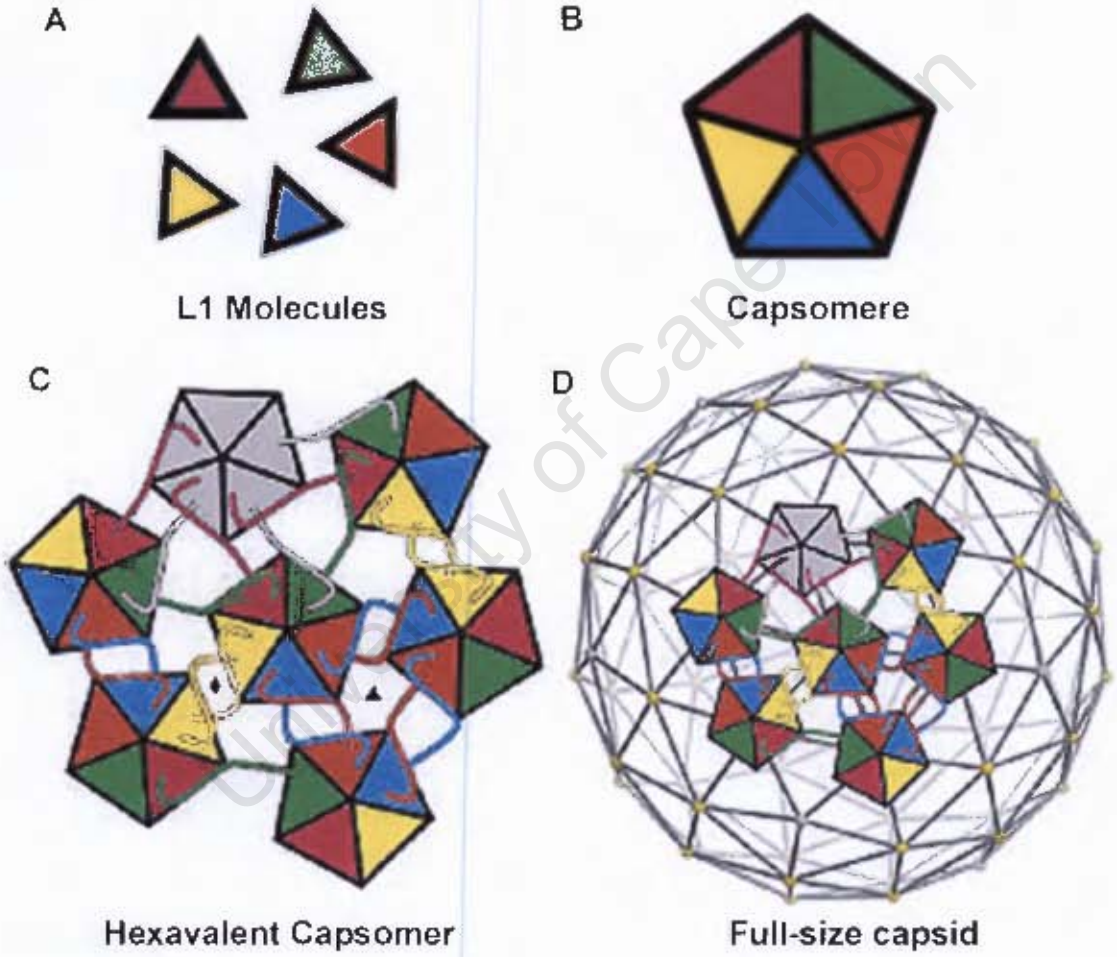

Figure 1. Modes of interaction between pentamers in subunit assemblies. (A) The L1 molecules are represented by triangles. (B) The L1 molecules associate to form pentameric capsomeres (pentamers). (C) Pentamer contacts in full-size (T = 7) papillomavirus particles form a hexavalent capsomer. (D) Diagram in (C) superimposed on the full T = 7 lattice. The six-fold symmetry is in the centre and the five-fold axis of symmetry is indicated by the grey capsomer. The two-fold and three-fold axes of symmetry are indicated in standard notation Taken and adapted from Modis et al., 2002 (57).

The L1 protein contains the intrinsic knowledge to form the capsid structure; allowing it to form virus-like particles (VLPs) that mimic the native virion structure (18). VLPs can form in the presence or absence of L2 and other cellular factors (18,37). Comparably, L1 alone VLPs have been noted to be more variable in size and shape than L1 and L2 expressed VLPs, leading to speculation that L2 enhances capsid formation (18,37). L2 may possibly act as a nucleating structure and allow formation at a more physiological pH (18,37). Cellular factors, such as hsp70 and several other chaperones have also been implicated and may aid capsid formation by facilitating the C-terminal invading arm (18,37).

The L2 protein acts to localize the various cellular components and viral components to nuclear domain 10 (ND10) sites within the nucleus which appears necessary for virion production *in vivo* (30,34). However, recent studies have shown that L2 in frozen wart and raft culture sections is localized rather to the nucleus than specifically to ND10 domains. L2 has also been implicated in aiding encapsidation of viral DNA. The HPV genome associates with cellular histones and may serve as a scaffold for capsid formation (18).

1.2.2 Viral Entry

The infectious virion is bound to the epithelial cell via the heparan sulphate receptor which recognizes a surface loop exposed on L1 (10). L1-alone VLPs are noted to be able to enter a number of cells; however, infectivity is increased with L2 co-expression (42). This may be due to a stabilizing effect of L2 or possibly indicate a secondary receptor binding motif within L2 (42). A surface binding region in L2 has been described but it is unclear to which receptor it binds (92).

The infectious virion is internalized by clathrin-dependent endocytosis (44). The initial cell-surface binding may confer a conformational change in the capsid,

exposing the N-terminal region of L2 – normally hidden by the capsomere (65). A proprotein convertase, furin, cleaves the L2 at a specific site, removing the first nine amino acids (65). This cleavage may facilitate a further conformational change allowing an as yet unidentified secondary-binding receptor recognition close to the N-terminal of the protein (92). This is evidenced by a monoclonal antibody against HPV-16 L2 that prevents internalization, most probably by preventing binding of the secondary receptor (24).

The cleavage and/or secondary-binding could allow the L2 and HPV genome complex to free itself of L1, and allow a membrane destabilizing peptide located in the C-terminal of L2 to allow egress of the complex from the endosome (44). A region in the N-terminal also interacts with tSNARE Syntaxin 18 which may facilitate movement to the nucleus (9). A C-terminal region has also been identified as binding to Dynein, a microtubule motor protein, which may also aid movement of the L2-HPV genome complex to the nucleus (Fig. 2) (29). This theory of viral entry has many unanswered questions; however, the importance of L2 is undeniable.

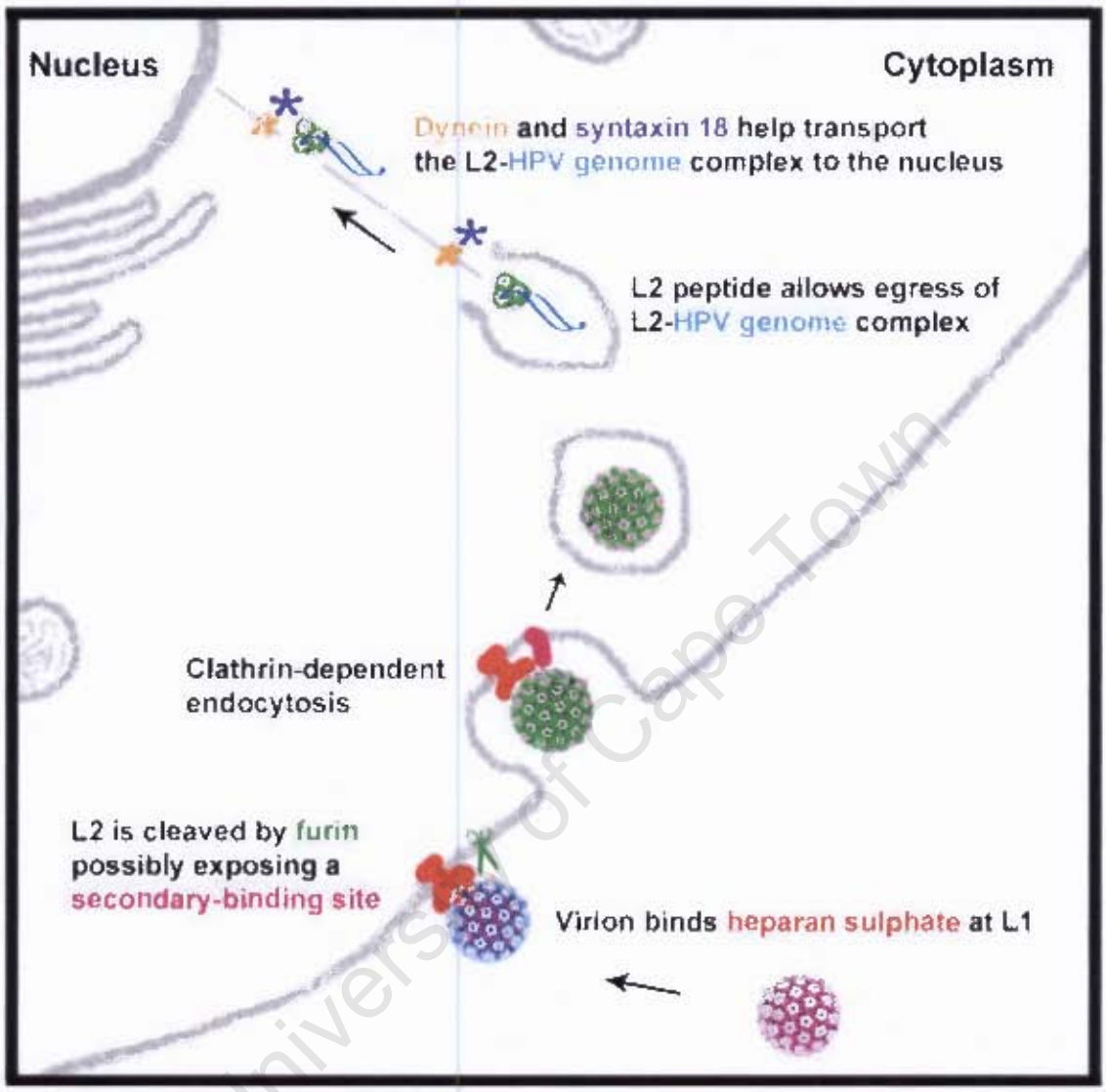

Figure 2. Theory of HPV cell entry. This is a simplified diagram to help explain the complex viral entry. Receptors on the cell membrane and various cellular proteins are indicated by symbols with corresponding colours. The change in colour of the virion indicates a possible conformational change initially with the binding of heparin sulphate (red) which may expose the region in L2 then cleaved by furin (green) further facilitating a conformational change exposing a region in L2 allowing binding of the secondary receptor (pink). Information from references (9,10,29,42,44,65,92) were used to make this diagram.

1.2.3 Propagation of Infection: Regulating the HPV Genome

HPVs are small, non-enveloped viruses that have a double-stranded DNA genome that is roughly 8 kb in size. In the capsid; cellular histones associate with the

genome to form a closed, circular mini-chromosome (72). Once inside the basal cell an initial burst of DNA replication serves to establish a copy number of the plasmid in the cell (8,72). The copy number varies between 50 and 100 copies, thereafter, they replicate once every cell cycle, essentially establishing a basal stem cell (8,72). A second round of replication occurs when the cell enters the suprabasal level and exits the cell cycle (72,74). Replication changes from theta to rolling-circle replication as a large number of copies are made (8). As the cell differentiates further the structural proteins are expressed and virions begin to form, finally resulting in the outer cell containing thousands of infectious virions (Fig. 3) (74).

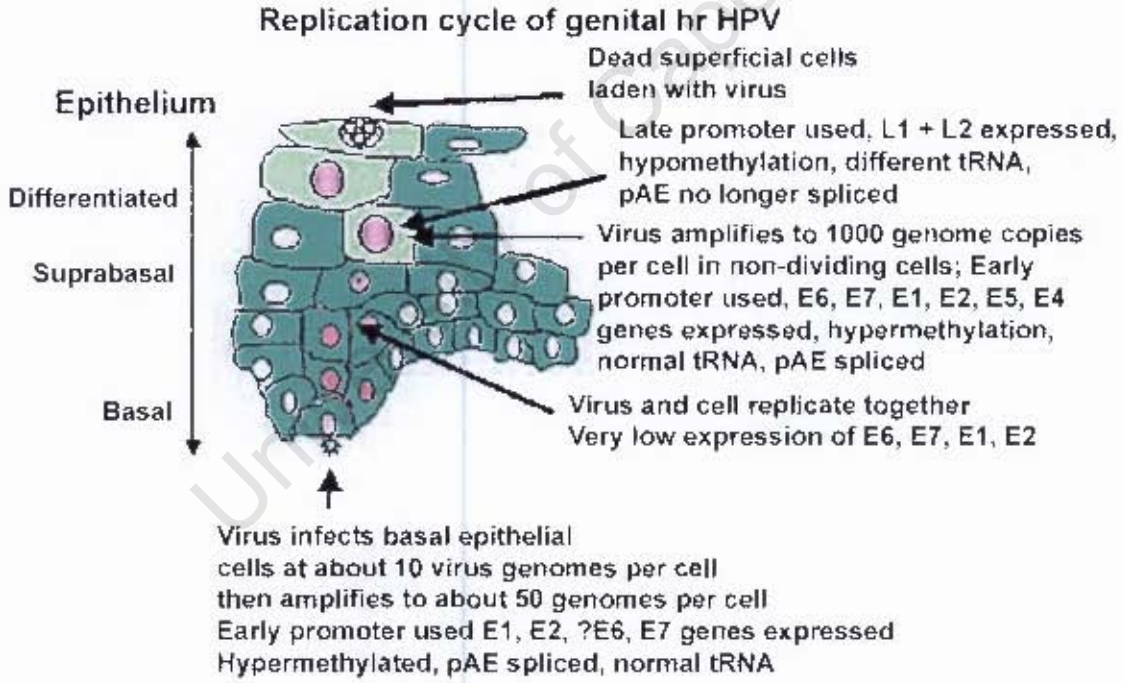

Figure 3. The HPV infectious cycle. The virus infects a primitive basal keratinocytes (probably a stem cell) via microabrasion of the mucosal epithelium. It is speculated that the immediate early events of virus growth involve an amplification of virus copy number from 1–10 to 50- 100 virus episomes/cell. The next phase of virus growth is one of plasmid maintenance in which the virus and cell replicate in tandem and there is no amplification of virus copy number. This occurs in the proliferative compartment of the epithelium. The infected keratinocyte then enters the differentiating strata of the epithelium, exiting the cell cycle. Virus gene expression is hugely upregulated with viral DNA amplification generating thousands of viral genomes. Late viral proteins L1, L2, and E4 are made, and virus assembly occurs in the superficial terminally differentiated squames. Taken and adapted from Stanley (2006) (74).

Although small, the HPV genome is tightly controlled and regulated, coding for only eight viral proteins that are read in all 3 frames. An untranslated long control region (LCR) rigorously controls gene expression (41). These proteins are further divided according to which stage of the viral life cycle they are expressed (72). The early proteins; E1, E2, E5, E6, and E7 are expressed in the basal and suprabasal cell layer and are involved in DNA replication, transcription and cell immortalization (Fig. 3) (55,72,74).

E1 and E2 are involved in DNA replication and transcription. HPV requires the host machinery to replicate its DNA; hence it needs to transform the differentiated cell, which has normally exited the cell cycle (8). E6 and E7 are involved in immortalization of the cell and interact with p53 and retinoblastoma (RB), respectively (55,72). E6 binds p53 and forces its degradation so disrupting its function as a tumour suppressor, it also acts to ensure telomeres are kept at their critical length (55,72). E7 competitively binds to RB, preventing it from negatively acting on the cell cycle (55,72). Both E6 and E7 have proven to be the potent oncogenes that lead to carcinoma development. Interestingly, the E6 and E7 proteins found in the high-risk HPVs differ by only one amino acid from those found in low-risk HPVs. (72) E5 is indirectly involved in tumour progression by sequestering the heavy chain of MHC I molecules in the Golgi, so aiding immune evasion (15).

Expression of all the HPV genes is exquisitely controlled, but the true success of the virus is determined by how well it can control the expression of its late genes, E4, L1, and L2. The late genes, in particular L1 and L2, are responsible for inducing an immune response to clear infection. E4 is believed to be involved in assisting escape of mature virions from the cell (55,72). As mentioned previously,

the capsid is made up of both L1 and L2; a number of methods are employed by the HPV genome to regulate the expression of these two genes.

Multiple promoters in the HPV genome allow the precise expression of the various genes. A family of promoters and regulatory regions are located in the LCR and function to control expression of early genes (72). Expression of late genes is controlled by regulatory regions in the LCR but the differentiation signal is thought to be closer to the late promoter located in the E7 gene (8). As a further control the late promoter is only functional in differentiated cells (8).

In addition, two separate polyadenylation signals are used by the early genes (pAE) and by the late genes (pAL). The pAE is located in the L2 gene; immediately downstream thereof are a number of GGG motifs that enhance the functioning of the pAE (61). If these GGG motifs are mutated read-through of the late genes occurs (61). The GGG motifs allow binding of hnRNP H which recruits various factors to allow splicing – another method of control (61). The HPV genome contains a number of splice sites that produce a varied number of mRNA products that can be selectively spliced according to the differentiation of the cell (68). For example, hnRNP H is abundantly found in the basal layer so the pAE is spliced into the early gene mRNAs leading to their expression (61). Expression of hnRNP H decreases as cells differentiate leading to the pAE not being used and allowing splicing of the late gene mRNAs with the pAL (61). This model supports the possibility of hnRNP H as a differentiation-dependent factor (61).

Recently methylation of the HPV genome has been implicated in controlling expression. Methylation may act to inhibit gene expression by directly preventing transcription factors from binding and/or indirectly allowing more complicated DNA structures to form that are not readily accessible by polymerases (4,79). The HPV genome has 40% GC content, therefore 400 sites for methylation (CpG sites)

are expected to occur randomly yet there are only 172 indicating their positions to be highly specific. Methylation was found predominantly in the LCR and the L1 gene. Interestingly, decrease in differentiated cells that are showing symptomatic infection; further enhancing the differentiation-dependent control. Furthermore, the HPV genome is hypermethylated with carcinoma development (80).

The HPV genome has the ability to integrate into the chromosomal DNA which more than likely happens due to the natural selection occurring due to the immunological response, such that integration is a means of escaping (74). Integration occurs at a particular site within the E2 gene disrupting its expression, this leads to an increase in E6 and E7 expression which is normally negatively-regulated by E2 (72). The L1 gene and rest of the genome have been noted to be hypermethylated upon integration but it is unsure if this is due to the integration into cellular DNA which is ordinarily tightly controlled by methylation.

Codon usage is another method many organisms use to control gene expression, including HPV (94). The eight HPV genes were analysed to determine how similar the codons were to those used in mammalian cells, it was found that L2 and L1 had the least similarity (94). Mammalian genes generally use a G or C at the 3rd position of the codon making them GC rich; HPV codons mainly employ a T at the 3rd position, making the HPV genome AT rich, the highest AT content was found in L1 and L2 genes (94). Changing the codon composition to those used in mammalian cells has greatly enhanced expression of L1 and L2 (20,58). Moreover, when cells differentiate codon usage has been shown to alter to those more similar to codons used in L1 and L2, further supporting differentiation-dependent control of expression (35,93).

Negative RNA elements found in L1 and L2 genes have also been implicated in regulating gene expression (60). Inhibitory RNA elements can form secondary

structures that inhibit polymerase-binding and act directly on the polymerase so inhibiting its functioning or possibly sequestering regulatory transcriptional factors (60). One RNA element has been described in L1 and two have been found in L2 (60).

All these regulatory factors coupled together have prevented production of the immunogenic L1 and L2 in various recombinant systems. Altering of the gene codons to those more readily used in mammalian cells has greatly enhanced production, possibly acting two-fold in negating negative RNA elements and in using more common codons. Overcoming regulation of these genes has allowed a greater understanding of their function and lead to the production of VLP-based vaccines.

1.3 Vaccine: Is L2 a Good Candidate?

Generation of a vaccine requires the fusion of two seemingly distinct disciplines, immunology and microbiology; thus a rudimentary knowledge of immunology is needed. A vaccine works by providing an antigen for the immune system so that an antibody "memory" is created that will prevent future infection. An immune response is developed by the T lymphocytes (T-cells) in the body that constantly sample for non-self antigens [for a review see reference (43)] (43,48). T cells do so by sampling the peptides bound to the major histocompatibility complex (MHC). There are two types of MHCs, MHC I and II, depending on the type of cell and the origin of the peptide will determine which MHC a peptide will bind. All nucleated cells have MHC I molecules and the peptides they bind are generated from endogenous proteins. MHC II molecules are found in professional antigen presenting cells (APCs) and are a representation of proteins that originate

outside of the cell. In these cells proteins are taken up and degraded in lysosomes and kept separate from cellular proteins, so preventing MHC I binding.

Peptide-MHC complexes are taken to the surface of the cell and presented for sampling by the T cells. (see reference (74) for a review with specific emphasis on HPV infection). MHC I molecules are sampled by $CD8^+$ cells otherwise known as cytotoxic T-lymphocytes (CTLs) (43). Stimulation of the CTLs leads to marking that cell for cell death (under the assumption that it is infected). MHC II molecules are sampled by $CD4^+$ cells. When activated $CD4^+$ cells release various cytokines to stimulate the immune system. $CD4^+$ cells are further subdivided into two sets of helper T cells (Ths), Th1 and Th2 (74). The type of infection will determine which pathway the non-activated $CD4^+$ cell will undertake which is determined by the APC interaction (74). A viral attack or vaccine will induce Th2 cells which will produce various cytokines stimulating B lymphocytes (B cells) (74).

B cells originate in the bone marrow and each contains a specific antigen recognition molecule similar to an MHC (74). Once a B cell is stimulated by an antigen and with the aid of cytokines from the Th2 cells it begins to divide and mature into a plasma cell. Plasma cells produce antibodies, which are a soluble form of the antigen recognition molecule. Antibodies will bind the antigen and hopefully prevent infection (74). There are two ways in which antibodies can prevent viral infection; (i) it can bind to a site in the capsid, inhibiting receptor binding; or (ii) bind and inhibit uncoating of the virus (19,25).

Two vaccines against HPV infection have recently passed rigorous human trials and are currently in production, namely, Gardasil and Cervarix which are produced by Merck and GSK, respectively. These vaccines exploit the fact that L1 expression alone leads to formation of VLPs which have proven highly effective at producing a strong immune response that can protect against infection in humans

(12,83). The immune response against L1 has been described as type-specific in a number of publications (16,21,62,86). The limiting type-specifity of L1 has been overcome by using multivalent vaccines incorporating more than one HPV type in the vaccine. Gardasil is a quadrivalent vaccine using VLPs, derived from recombinant yeast, with HPV types 16, 18, 6, and 11 and Cervarix a bivalent vaccine, derived from recombinant insect cells, incorporating VLPs from HPV-16 and -18.

Antibodies raised against VLPs that contain both L1 and L2 produce a predominantly type-specific response compared to the response against the L2 protein alone (66). This is due to immunodominance patterns that occur. Epitopes can be divided into those that produce a strong immune response (dominant epitopes) and those producing a weaker response (subdominant epitopes) (81). Three major factors that determine the dominance of a peptide are: (i) the antigen processing, (ii) the binding affinity to the MHC molecule, and (iii) the stability of the peptide-MHC complex (81). The quantity of the antigen has also been implicated in MHC I presentation but MHC II appears to be quantity-independent (71). It appears that L1 contains immunodominant epitopes that mask the antibody response to the subdominant epitopes found in L2. This could be due to the larger amount of L1 being present and/or may reflect a further viral evolution in protecting itself from the immune system.

A number of epitopes found on the L2 protein (~500 amino acids) have shown cross-reactivity. Analysis of the protein sequence of various HPVs has shown conserved regions at positions 1-12 and 56-81 (45). The surface regions have been characterized at positions 32-51, 69-81, 212-231, 279-291, and 362-381 (Fig. 4) (45). Surface regions are the most likely to be targets for neutralizing antibodies, and one that is conserved across HPVs would be ideal for cross-reactivity.

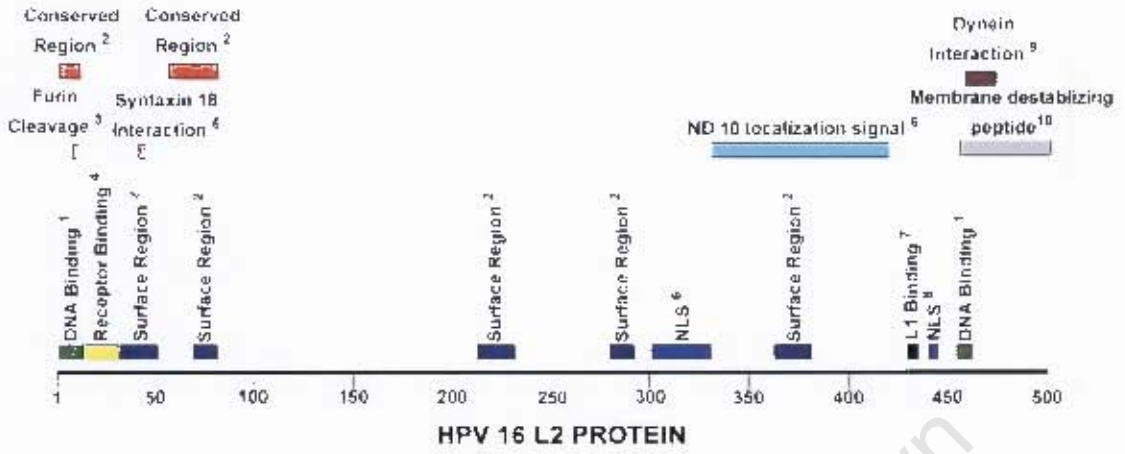

Figure 4. The Various Amino Acid Regions of the HPV 16 L2 Protein. 1 – The DNA binding domains (dark green); 1-13 and 454-462; were described by Boursaghin *et al.* (2003) (10). 2 – The surface expose regions of L2 (navy); 32-51, 69-81, 212-231, 279-291, and 362-381; and the conserved regions of L2 (red); 1-12 and 56-81; have been previously described by Kawana *et al.* (1998) (45). 3 – The furin cleavage site (light green); 8-11; has been determined by Richards *et al.* (2006) (65). 4 – The receptor binding site (yellow); 13-31; was defined by Roden *et al.* (2000) (66) and later confirmed to bind cells by Yang *et al.* (2003) (92). 5 – tSNARE Syntaxin 18 protein was found to interact with L2 at the residues 43-47 (pink) by Bossiss *et al.* (2005) (9). 6 – The NLS site (blue); 300-330; and ND10 localization signal (aqua); 330-420 were determined by Becker *et al.* (2004) (5). 7 – The L1 binding domain (black) was also initially described in HPV 11 L2 but from sequence homology of various PVs the corresponding site in HPV 16 was determined at residues 412-455. Work was done by Finnen *et al.* (2003) (28). 8 – Finnen *et al.* (2003) (28) also described an NLS (blue) present at sites 440-445. 9 – L2 was found to interact with dynein protein at the residues 457-473 (brown); by Florin *et al.* (2006) (29). 10 – A membrane destabilizing peptide (grey) was characterized to be in the C-terminal at residues 454-500 by Kämper *et al.* (2006) (44). The NLS present at 440-445 was determined in HPV 11, not HPV 16; therefore it may be the same as the NLS described in site 300-330.

The HPV-16 L2 region of 69-81 was inserted into the BPV L1 gene, to produce a chimeric VLP (cVLP) that displays the epitope on the VLP. Analysis of the immune response found neutralizing antibodies for both HPV-16 and HPV-11 infection (73). The region of 108-120 was previously described as a cross-neutralizing epitope (46), however, it appears inoculation with the cVLP with this peptide failed to produce neutralizing antibodies (73).

A receptor binding site is another potential region to elicit neutralizing antibodies. A possible site was determined in Bovine Papillomavirus (BPV) L2 in the region of 1-88, sequence homology was found in HPV-16 L2 to the region 13-31, and

cell-surface binding was later confirmed for this region (66,92). This potential site for a neutralizing epitope, however, may be hidden within the capsid and only become available upon furin digestion.

BPV L2 1-88 region was found to elicit an immune response that could neutralize a number of HPVs (63). More recently, epitopes HPV-16 L2 1-88, HPV 16 L2 11-200, Cotton-tail Rabbit Papillomavirus (CRPV) L2 1-88, and CRPV L2 11-200 were tested (32). The HPV-16 L2 11-200 epitope proved to elicit the best cross-protection against HPV 16, 18, 5, 31, 45, 52, 58, and BPV 1. (32) HPV 16 L2 1-88 epitope was found to elicit a similar response to that of BPV 1 1-88 (32).

RG-1 a monoclonal antibody raised against HPV-16 L2 amino acids 17-36 was found to cross-neutralize 16 and 18. This antibody was shown to be able to bind to the virion after the initial binding to the cell surface and thereby inhibiting internalization (70). Within the 17-36 amino acid sequence of HPV-16 L2 there are two cysteine residues at position 22 and 28 and these are highly conserved throughout different HPV types. After mutating these cysteines to serines it was still possible to obtain infectious virus in organotypic culture, but when both cysteines were mutated no infectious virus were obtained indicating the importance of these residues (22). In addition, the HPV-16 L2 17-36 peptide was fused to a toll-like receptor 2 ligand and a T helper epitope to obtain a lipopeptide. Sera from mice vaccinated with the lipopeptide was able to neutralize HPV16, 18, 45 and BPV1 (3).

Accumulatively this data indicates that the amino acid region between 17-36 of HPV-16 L2 has the potential to be used to raise antibodies that will be able to cross neutralize against various HPV types.

Peptides in the N-terminal of HPV 16 L2 were also analyzed to determine the antigenic regions (51). Using a computer program they determined which would be the most antigenic sites, i.e. which peptides are more readily processed by the immune system and elicit a cross-protective response (51). The peptide covering the region 65-71 was speculated to be the most immunogenic and produced the most cross-protective response, neutralizing HPV 16, 18, 31, and 58 infections (51).

1.4 Conclusion

Various elements in the L2 gene have been shown to be significant in the viral life cycle; controlling the expression of a number of genes. The tight control of its own expression allows the virus to remain undetected by the immune system. This control has also limited expression of L2 in a number of recombinant systems which has been overcome by gene codon optimization, increasing its production to levels that are easily detectable. Current recombinant technology allows the production of L2 in a number of different eukaryotic systems, including; mammalian cells, insect cells, and yeast cells.

The role of L2 in the viral entry is undeniably important, various residues within L2 and so present in the virion capsid are necessary for efficient cell entry. Theses residues remain attractive sites for neutralizing antibodies to bind; in particular the receptor-binding site found in the 17-36 aa region. L2 has also shown a number of other potential sites that have cross-neutralizing activity; in particular a site located in the surface region that is highly conserved amongst HPVs, the region 56-81.

Current information suggests that with expression in a good recombinant system the L2 protein, or parts thereof, are an excellent candidate for a cross-protective vaccine against HPV and cervical cancer.

2. CHAPTER 2

THE EXPRESSION OF L2 IN PLANTS

2.1 Introduction

The currently available HPV vaccines are projected to cost $120 per dose, making the total cost with three doses $360 (1,31). Considering the frequency of infection in developing countries and their socio-economic environments, this fee would be vastly out of their reach. The patents upon these vaccines will, however, eventually expire, leading to a race for alternatively cheaper generic vaccines. Plant production of an HPV vaccine may serve this purpose.

Vaccine production in plants could be cost effective for a number of reasons. First, plant systems are more cost effective than the use of fermentation bio-reactors, such as those used in bacterial production of heterologous protein (23). Secondly, the infrastructure and equipment for large scale plant production and harvesting are already in place (23). Thirdly, plants allow for the possibility of an edible vaccine, negating purification requirements and needle-associated complications (23). Last, the yield from plants has crossed the threshold to industrial levels, making their use economically viable (23). These factors have lead to a flurry of novel vaccines and pharmaceutical products being produced in plants.

The vast range of plant-based products is testament to the versatility the system endows. Edible vaccines currently being tested include protection against human infection by hepatitis B (77), animal infection by rotavirus (88), bacterial infection of humans by *Shigella* (87) and bioterrorism using the anthrax toxin (52). In addition, plant production has found a specialized niche in the production of

antibodies (11,23,33,49). Various mechanisms are employed to produce a protein in plants, including chloroplast transformation involving the transformation of the chloroplast genome as opposed to the nuclear plant genome. A number of advantages of this method have been argued, in particular the means of bio-containment of the foreign gene (23,47).

Transgenic plant lines, however, remain at present the predominant means of producing heterologous gene products in plants. Recently an alternative system, namely transient expression, has gathered interest. This involves the infiltration of plant material with acetosyringone-induced *Agrobacterium* containing the plasmid with the foreign gene. Acetosyringone is a plant phenolic that activates the *vir* genes on the tumour-inducing plasmid within *Agrobacterium* which facilitates the relocation of the transfer DNA (T-DNA) from the bacterium to the plant cell. A single strand copy of T-DNA (ssDNA) is transferred and localized to the nucleus of the plant cell (95). The ssDNA is made double-stranded at which point it can be used by the cell for transcription, a fact exploited by the transient expression system (95).

The T-DNA can stably integrate into the plant chromosome or, what is more often the case, remain in the nucleus as extrachromosomal DNA (95). Integration leads to the ability to make transgenic lines. One advantage of transient expression is that expression can be determined shortly after DNA delivery as compared to transgenic plant regeneration, which requires a number of months. In addition, gene silencing due to positional effect of integration, commonly seen in transgenics, is negated.

A number of vectors have been designed to facilitate the transfer of the T-DNA. Binary vectors are commonly used; however, various novel vectors have been designed to remove the cumbersome system of ensuring both plasmids are present.

Such an example is the pTRA vectors which have the necessary genes for selection of the recombinant bacteria and the foreign gene on one plasmid. The *vir* genes are contained within a stable helper plasmid found in the specific GV3101 *Agrobacterium*. These and other vectors have contributed to the increase in expression levels available in plant systems. Expression of foreign genes, however, can be limited by the natural defence mechanisms of the plant.

Transient expression has been characterized by limited levels of the gene product after a number of days due to post transcriptional gene silencing (PTGS) (75,84). Heterologous gene expression is often under the control of a strong, constitutive promoter leading to a large number of mRNAs being present in the plant cell. A certain amount of inter- and intra-molecular homology in the RNA will be present within these sequences which will associate and dissociate relative to the concentration of the mRNA in the cell. Any dsRNA present in the plant cell is recognized by an RNAse III-like protein that will cleave dsRNA into 21-25 bp short interfering RNAs (siRNAs) (84). An siRNA associates with a nuclease – giving the nuclease sequence specificity – which will in turn go on to degrade any RNA with sequence homology (Fig. 1) (84). A number of plant viruses have developed methods of avoiding the host defense by encoding proteins that interfere at various stages of the mechanism.

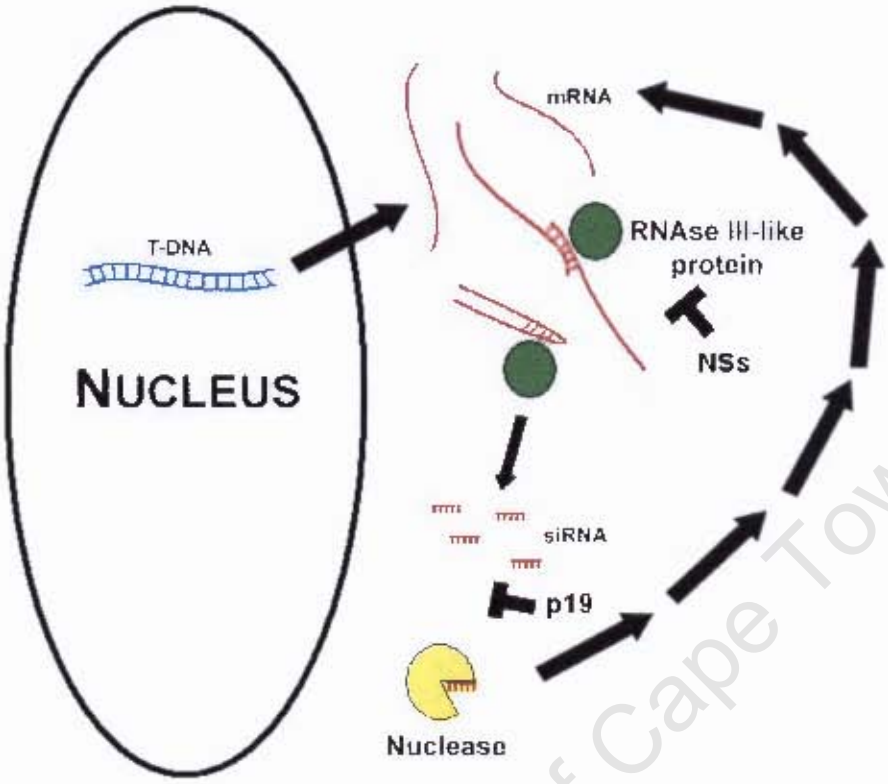

Figure 1. Diagrammatic view of PTGS. Exogenous mRNA is made from the T-DNA. These form dsRNA which are recognized by RNAse III-like proteins. NSs acts to inhibit these proteins and so prevent formation of siRNAs. The p19 suppressor acts to prevent these siRNAs from associating with the nucleases thus not giving them the necessary site specificity and preventing degradation.

In order for these other proteins to function in inhibiting PTGS they need to be present in the same cell as the foreign gene to be expressed. This tells us that in the proposed system two separate *Agrobacterium* will have to transfect the same cell with two separate pieces of T-DNA. Production of antibodies in plants have demonstrated that such a phenomenon can occur, with even three separate pieces of T-DNA encoding the heavy, light and constant chains transiently expressing antibodies (33,49). It then stands to reason that full HPV VLPs that incorporate both the L1 and L2 molecule should be possible in this system.

In addition to PTGS, protein production may also be limited by codon usage. As mentioned in Chapter 1 it is a well established principle that codon usage can limit gene expression in most systems, including plants. Thus, by converting the codons

of a gene to those more commonly used in plants, gene expression should, in theory, increase.

With these principles in mind this chapter reports on investigations aimed at establishing whether the L2 protein, like its counterpart L1, can be efficiently expressed in plants. If so, to determine which L2 gene would produce the highest yields, i.e. the wildtype, one whose codons have been optimized for plant expression or one which has been optimized for mammalian expression? And if one gene is found to produce greater amounts of L2 we sought to investigate why this would be so. In addition, once expression is established it would be interesting to determine where in the plant the L2 is targeted and whether specific targeting to different cell compartments would yield different amounts. Finally, this study further seeks to determine if full VLPs that incorporate L2 can be produced upon simultaneous transfection, with a view of making a cheaper vaccine for developing countries.

2.2 Methods

2.2.1 Plasmid Construction

Agrobacterium vectors pTRAc, pTRAkc-A, pTRAkc-ERH and pTRAkc-rbcs1-cTP were supplied by Rainer Fischer (Fraunhofer Institute for Molecular Biology and Applied Ecology IME, Aachen, Germany) (Fig. 2).

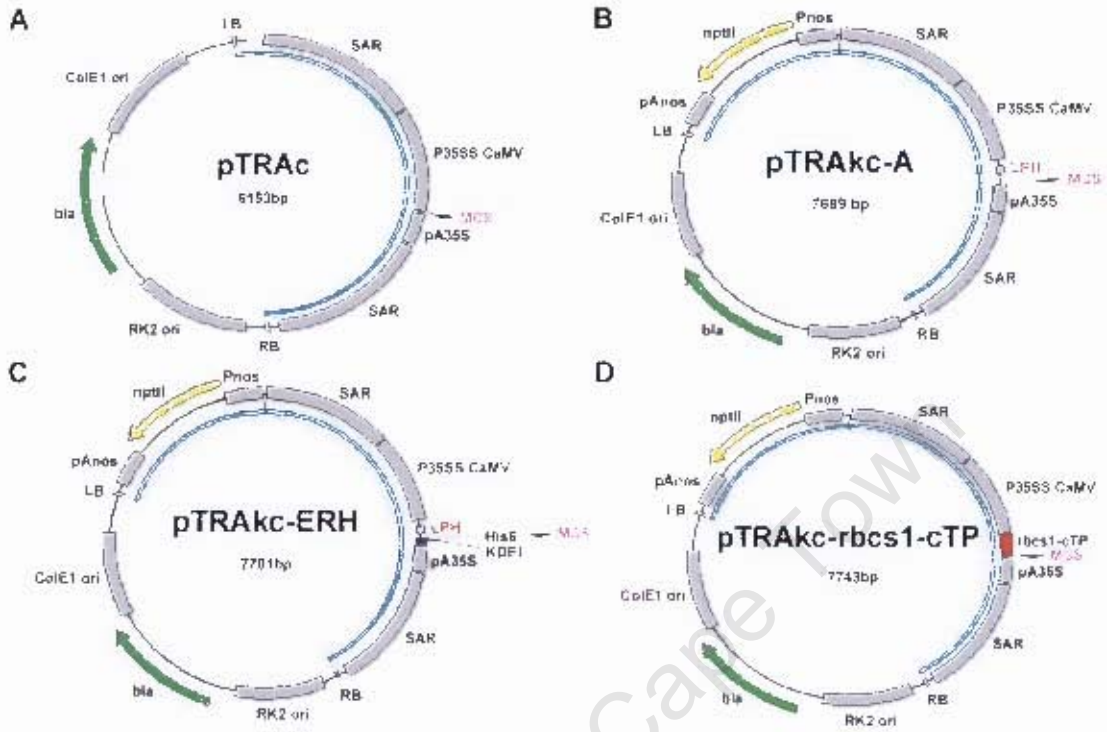

Figure 2. *Agrobacterium* vectors utilised in this study. They are pTRAc (A), pTRAkc-A (B), pTRAkc-ERH (C), and pTRAkc-rbcs1-cTP (D). The light blue bar indicates the T-DNA that is transferred to the plant cell at transfection. The heterologous gene is cloned into the vectors at the multiple cloning site (MCS). The vectors share a number of common features indicated by regions in grey. The T-DNA region is flanked by the left border (LB) and the right border (RB). At either end of the MCS are scaffold attachment regions (SAR). Expression of the transgene is controlled by a dual 35S Cauliflower Mosaic Virus promoter (P35SS CaMV) with attachment of the polyadenylation signal of the same CaMV gene (pA35S). Replication of the vector in *Agrobacterium* is initiated at the RK2 ori. A separate replication initiation site, ColE1, is used in *E. coli*. The pTRAc vector contains only one antibiotic resistance marker (*bla*) allowing selection with ampicillin/kanamycin. The pTRAkc-A, pTRAkc-ERH and pTRAkc-rbcs1-cTP contain a second antibiotic marker (*nptII*) permitting selection with kanamycin in plants. The *nptII* gene is controlled by the promoter of the *Agrobacterium* gene nopaline synthase (Pnos), the polyadenylation of the same gene is attached to nptII (pAnos). pTRAkc-A has a secretory signal (LPH) immediately upstream of the MCS. pTRAkc-ERH also has the LPH signal and additionally includes a His tag (His6) sequence and an endoplasmic reticulum retention signal (KDEL) sequence downstream of the MCS. pTRAkc-rbcs1-cTP includes a chloroplast signalling peptide (rbcs1-cTP) sequence upstream of the MCS.

The pTRA vectors hold a number of characteristics that optimize the expression of the foreign gene. pTRAc contains the skeleton features and is proficient at transgene expression in the cytoplasm of plant cells (Fig. 1). The other vectors pTRAkc-A, pTRAkc-ERH and pTRAkc-rbcs1-cTP include additional sequences for the targeting of the transgene expression to the apoplastic space, endoplasmic

reticulum (ER) and chloroplast, respectively. The left and right borders flank the region of DNA that will be transferred to the plant cell, the T-DNA. The scaffold attachment regions, also known as matrix attachment regions, are on either side of the transgene. These regions are thought to interact with nuclear matrix proteins forming loop domains in DNA which has been shown to increase expression with insertion in a plasmid (59).

Expression of the transgene is controlled by a dual P35S CaMV promoter, with a duplicated transcriptional enhancer. This promoter has been shown to increase expression over the traditional single P35S CaMV (67). The polyadenylation signal of the P35S gene is fused to the end of the foreign gene; this stabilizes RNA transcripts. Replication of the plasmid in *Agrobacterium* is initiated at the RK2 origin of replication. Due to the low copy number of the plasmid in *Agrobacterium*, the vector can alternatively be replicated in *Escherichia coli* using the ColE1 origin of replication (50). The development of resistance by a number of *Agrobacterium* to ampicillin has lead to the fusion of kanamycin and ampicillin resistance in the *bla* gene. Ampicillin is used as a marker in *E. coli* with kanamycin as a marker in *Agrobacterium*.

Three different L2 genes were used (i) the wildtype HPV-16 L2, (ii) the HPV-16 L2 codon optimized for expression in plants (plantized) and (iii) the HPV-16 L2 codon optimized for expression in mammalian cells (humanized). Codon optimization for the plantized gene was performed by GENEART® (Regensberg, Germany) to be optimized for *Nicotiana*. The humanised gene was generously supplied by Martin Müller (Deutsche Krebsforschungszentrum, Heidelberg, Germany). Restriction enzyme sites were included at either end of the L2 open reading frame (ORF) by polymerase chain reaction (PCR) to facilitate the directional cloning into the vectors.

Cloning into the pTRA-C vector requires the gene to be placed in frame on the plasmid and so BspHI is necessary as it contains an ATG site within its restriction site. The downstream restriction enzyme may vary. The wildtype L2 ORF (saL2) was amplified with the primer pair F saL2-pTRAc and R saL2-pTRAc (Table 1). The primer pair F pL2-pTRAc and R pL2-pTRAc were used to amplify the plantized L2 ORF (pL2). With the humanised L2 ORF (hL2) the primer pair of F hL2-pTRAc and R hL2-pTRAc was used. The 1.4 kb fragments were cloned into pGEM®-T easy (PROMEGA) and sequenced to confirm PCR fidelity. The saL2 was excised with restriction enzymes BspHI and XbaI and subcloned into the AflIII and XbaI sites of pTRAc making the plasmid pTRA-saL2-C (Table 2). Restriction enzymes BspHI and BamHI were used to excise pL2 which was subsequently subcloned into the AflIII and BamHI sites of pTRAc making the plasmid pTRA-pL2-C. Finally, restriction enzymes BspHI and XbaI were used to excise hL2 and subcloned into the AflIII and XbaI sites of pTRAc making the plasmid pTRA-hL2-C.

Similarly the three genes were subcloned into the pTRAkc-ERH vector. The in frame placement of the gene using BspHI in this vector becomes important due to the addition of the secretory signal, LPH. The same forward primers and so corresponding restriction sites, used to clone the genes into the pTRAc vector were utilised (Table 1). Due to the requirement of a translational fusion at the C-terminal the stop codons of the ORFs were mutated and replaced by a NotI site. The addition of a his6 tag allows for purification and the KDEL sequence facilitates retention of the protein in the ER. Primers R saL2-ERH, R pL2-ERH and R hL2-ERH (Table 1) were used to subclone saL2, pL2 and hL2 into the pTRAkc-ERH vector making the plasmids pTRA-saL2-E, pTRA-pL2-E and pTRA-hL2-E, respectively (Table 2). In addition, a fourth plasmid was made with the hL2 gene that incorporated the hL2 fragment from the pTRAc clone into the pTRA-kc-ERH vector which resulted in removal of the translational fusion

making the plasmid pTRA-hL2-A (Table 2). Thus, the protein would not be retarded in the ER due to the absence of the KDEL sequence and is allowed to be secreted.

Concurrently the three genes were subcloned into the pTRAkc-rbcs1-cTP vector which also requires a translational fusion of a chloroplast-targeting domain at the N-terminal. For this reason a longer forward primer is required that incorporates a MluI site into the fragment. Primers F saL2-CTP, F pL2-CTP and F hL2-CTP (Table 1) were used to subclone saL2, pL2 and hL2 into the pTRAkc-rbcs1-cTP vector to make plasmids pTRA-saL2-P, pTRA-pL2-P and pTRA-hL2-P, respectively (Table 2). The corresponding reverse primer used to clone the L2 genes into pTRAc were employed (Table 1).

Table 1: Primers used in cloning.

Primer Name	Sequence	Underlined Restriction Site
F-saL2-pTRAc	GCTCATGAGACAAACGTTCTGCAAAAG	BspHI
R-saL2-pTRAc	AATCTAGACTAGGCAGCCAAAGAGAC	XbaI
R-saL2-ERH	AAGCGGCCGCGGCAGCCAAAGAGACATC	NotI
F-saL2-CTP	AAACGCGTTAGGTGCATGAGACACAAACGTTCTGCAAAAC	MluI
F-pL2-pTRAc	AATCATGAGACATAAGAGATCTGCTAAG	BspHI
R-pL2-pTRAc	CGGGATCCCTAAGCAGCAAGAGAAACATC	BamHI
F-pL2-CTP	GGACGCGTTAGGTGCATGAGACATAAGAGATCTGC	MluI
F-hL2-pTRAc	ATTCATGAGGCACAAGAGGAGCGCC	BspHI
R-hL2-pTRAc	ATTCTAGATCAGGCGGCCAGGCTCAC	XbaI
R-hL2-ERH	ATGCGGCCGCGGCGGCCAGGCTCACGTC	NotI
F-hL2-CTP	ATACGCGTTAGGTGCATGAGGCACAAGAGGAGC	MluI

Table 2: *Agrobacterium* clones that were used.

Plasmid Name	Insert	*Agrobacterium* Strain	*Agrobacterium* Vector	Cell Compartment Targeted
pTRA-saL2-C	HPV-16 saL2	GV3101	pTRAc	cytoplasm
pTRA-saL2-E	HPV-16 saL2	GV3101	pTRAkc-ERH	endoplasmic reticulum
pTRA-saL2-P	HPV-16 saL2	GV3101	pTRAkc-rbcs1-cTP	chloroplast
pTRA-pL2-C	HPV-16 pL2	GV3101	pTRAc	cytoplasm
pTRA-pL2-P	HPV-16 pL2	GV3101	pTRAkc-rbcs1-cTP	chloroplast
pTRA-hL2-C	HPV-16 hL2	GV3101	pTRAc	cytoplasm
pTRA-hL2-E	HPV-16 hL2	GV3101	pTRAkc-ERH	endoplasmic reticulum
pTRA-hL2-P	HPV-16 hL2	GV3101	pTRAkc-rbcs1-cTP	chloroplast
pTRA-hL2-A	HPV-16 hL2	GV3101	pTRAkc-A	apoplastic space
*pTRA-hL1-C	HPV-16 hL1	GV3101	pTRAc	cytoplasm
*pTRA-hL1-P	HPV-16 hL1	GV3101	pTRAkc-rbcs1-cTP	chloroplast
*pTRA-GFP-P	GFP	GV3101	pTRAkc-rbcs1-cTP	chloroplast
*pTRA-p19-C	p19	GV3101	pTRAc	cytoplasm
#pBIN-NSs	NSs	LBA 4404	pBIN	cytoplasm

*Clones were supplied by James Maclean, this laboratory.
#Clone provided by Marcel Prins, Wageningen Agricultural University.

2.2.2 Generation of Recombinant Agrobacterium

Agrobacterium tumefaciens GV3101 was provided by Rainer Fischer (Fraunhofer Institute for Molecular Biology and Applied Ecology IME, Aachen, Germany). The GV3101 strain contains the helper plasmid, pMP90RK, which contains the crucial *vir* genes (50). The pTRA vectors can be replicated in *E. coli* as well as *Agrobacterium*. The vector constructs were initially cloned into DH5α cells, which are easier to culture and have a higher copy number of the plasmid in comparison with *Agrobacterium* (50). Plasmid DNA was isolated from DH5α cells and the concentration was determined by UV spectrometry at 260 nm.

GV3101 cells were grown to log phase at 0.8 OD_{600} at 26°C with shaking in Luria Broth (LB) containing antibiotics rifampicin (50 µg/ml) and kanamycin

(30μg/ml). Cells were subsequently made electrocompetent by washing three times with Milli-Q water and resuspended in 1/20 the culture volume in 10% glycerol to allow for storage at -20°C.

Plasmid DNA (400 ng) isolated from DH5α was mixed with 100 μl of electrocompetent GV3101 cells in a 0.1 cm cuvette (BIORAD®) and electroporated using the following parameters; 200Ω, 25 μF and 1.5 kV (Gene Pulse, BIORAD®). After incubation at 26°C in LB for one hour the electroporated cells were placed on Luria agar (LA) containing rifampicin (50 μg/ml), kanamycin (30 μg/ml) and carbenicillin (50 μg/ml) and allowed to grow at 26°C for 3-4 days. Successful transformation was determined by colony PCR or restriction enzyme analysis. Due to the low copy number in *Agrobacterium*, the plasmid was transformed into DH5α for restriction analysis.

Glycerol stocks of the suitable *Agrobacterium* clones were made. This resulted in a total of ten clones being made (Table 2). In addition, Marcel Prins (Laboratory of Virology, Wageningen University, Wageningen, Netherlands) provided the *Agrobacterium* recombinant with the pBin-NSs plasmid coding for the NSs protein (Table 2). James Maclean provided the *Agrobacterium* GV3101 clones with plasmids pTRA-hL1-C and pTRA-p19-C that code for the humanised L1 and p19 proteins, respectively (Table 2).

2.2.3 Plant Lines

Wild type *Nicotiana benthamiana* and *Nicotiana tabacum* cv Petite Havana were used and grown under conditions of 16h light and 8h dark at 22°C. Plants were utilised 14-28 days after being transplanted from seedling trays to pots.

2.2.4 Preparation of Agrobacterium for Infiltration

The infiltration protocol simulates the natural process of *Agrobacterium* induction by acetosyringone. The induction and infiltration medium both contain varying amounts of acetosyringone which activates the different *vir* genes according to concentration (95). Clones were grown from glycerol stocks in LB with rifampicin (50 µg/ml), kanamycin (30 µg/ml) and carbenicillin (50 µg/ml) to an OD_{600} between 1 and 2 at 26°C. *Agrobacterium* was centrifuged at 5,000xg for 10 min at room temperature (RT) and resuspended in induction medium [LB, 10 mM MES, 2 mM $MgSO_4$, 20 µM acetosyringone, rifampicin (50 µg/ml), kanamycin (30 µg/ml) and carbenicillin (50 µg/ml) at pH 5.6] and grown to an OD_{600} of 1-2 at 26°C. Cells were centrifuged at 5,000xg for 10 min at RT and resuspended in infiltration media (10 mM MES, 10 mM $MgCl_2$, 2.2105 g/l Murashige Skoog salts, 35 g/l sucrose, 150 µM acetosyringone at pH 5.6). The optical density of the cells were measured and diluted in infiltration media to an OD of 0.4-1. Cells were incubated for a minimum of 3h in the infiltration medium before infiltration.

2.2.5 Infiltration Procedure

Two different infiltration protocols were used; injection infiltration or infiltration under vacuum. The injection procedure requires a 1 ml syringe which is used to inject the *Agrobacterium* suspended in infiltration media directly into the abaxial air spaces of the plant leaves.

Vacuum infiltration required uprooting of the whole plant and suspending them in infiltration medium containing the *Agrobacterium* clone. This was placed under vacuum at 80 mbar for 5 min. Upon release of the vacuum the air spaces in the leaves were infiltrated and the plant was replanted. Plants were subsequently

covered in cellophane bags for 2 days, or longer, and incubated at 28°C in a humidity-controlled room. Vaquero *et al.* (1999) used an alternative method of using just the leaves not the whole plant; we modified that protocol for our own purposes.

2.2.6 GFP Imaging

As an easily traceable protein product which can be readily visualized, GFP was used to determine the progress of protein expression within the transient expression system. Two week old *N. benthamiana* plants were infiltrated with the *Agrobacterium* clone containing plasmid pTRA-GFP-C. Leaves were measured on a short wave UV light box (Gene genius Bio imaging systems, Syngene – a division of SYNOPTICS LTD.) after six days with the light source from below. Photos were adjusted using Adobe Photoshop Elements 2.0.

2.2.7 Protein Extraction and Western Blots

To determine whether expression was present within the plant standard techniques of SDS-PAGE and western blot analysis were used. The size of the sample determined the method of homogenization: if it was below 500 mg fresh weight eppendorf pestles were used and if above 500 mg the leaf tissue was frozen in liquid nitrogen and ground in a mortar and pestle. Homogenized sample was suspended in 2 $\mu l.mg^{-1}$ of buffer. Buffers used were either PBS at pH 6.7, high salt PBS (0.5M NaCl) or 8M urea. Cell debris and other larger molecules were separated either by a mesh filter, if the samples were large enough, and/or underwent 2 centrifugation rounds (10,000xg, 10 min at RT).

SDS-PAGE loading buffer was added to the samples and boiled for 10 min. Sample was then loaded onto a 10% SDS-PAGE and run at 20mA/gel for ~2.5h. Protein was transferred to nylon membrane (Nitrobind, Cast, Pure nitrocellulose, 0.45 micron, OSMONICS INC.) by semi-dry blotting at 15V and 400 mA for 2h (Trans Blot® semi-dry, BIORAD® with the power supply Electrophoresis Power Supply, AMERSHAM®). The success of the transfer was measured by coomassie blue staining of the gel after transfer.

The membrane was blocked O/N in PBS (5% skimmed milk, 0.05% Tween-20). The membrane was then incubated for 2h at RT with the primary antibody followed by 4 repeated washes in PBS (5% skimmed milk, 0.05% Tween-20). For detection of L2 rabbit polyclonal serum was used at 1:1,000 and for L1 mouse monoclonal antibody (J4) was used at 1:3,000; diluted in PBS (5% skimmed milk, 0.05% Tween-20). Subsequently, the secondary antibody was diluted in PBS (5% skimmed milk, 0.05% Tween-20) and added to the membrane and incubated at RT for 1h followed by 4 repeated washes over 1h in PBS (0.05% Tween-20). Antibodies goat-α-rabbit and goat-α-mouse conjugated with alkaline phosphatase (SIGMA®-Aldrich) were used as secondary antibodies for L2 and L1, respectively; both were diluted to 1:10,000. Immunodetection was done using 5-Bromo-4-Chloro-3-indolyl phosphate/Nitroblue tetrazolium (BCIP/NBT) made according to manufacturer's instructions (ROCHE® Diagnostics).

2.2.8 Enzyme Linked Immunosorbent Assay (ELISA)

In an attempt to quantitate the expression of L2 and hopefully detect expression unable to be seen with a western blot the more sensitive ELISA was employed. Detection of HPV-16 L2 was done by coating a Polysorb® 96-well plate (NUNC®) with L2 samples either in PBS or 8M urea O/N at 4°C. Plates were blocked in PBS with 5% skimmed milk for 1h. Non-specific binding was determined by coating

wells with only PBS or 8M urea. The plate was incubated with rabbit polyclonal antiserum raised against HPV-16 L2 (1:1,000) for 3h followed by 4 repeated washes in PBS (5% skimmed milk, 0.05% Tween-20). Plates were then incubated with goat-α-rabbit antibody conjugated with alkaline phosphatase (SIGMA®-Aldrich) diluted (1:5,000) in PBS (5% skimmed milk) for 1.5h. The extinction was measured at 405 nm at 30 min after adding the substrate Sigma FAST™ p-Nitrophenyl Phosphate (pNPP) made as per manufacturer's instructions (SIGMA®-Aldrich). The extinction coefficient was measured on the Multiskan® PLUS, Titrek®.

2.2.9 Concentrating the L2 Protein

2.2.9.1 Tri-Chloro Acetic Acid (TCA) Precipitation

TCA precipitation is a commonly used technique to concentrate proteins. It acts to precipitate all soluble proteins which may then be pelleted and thus allow resuspension in a smaller volume. Sample was extracted from leaf material as per the extraction method for material >500mg. 112.5µl of 50% TCA was added per 1ml of sample. The protein was pelleted by centrifugation at 13,000xg for 8 min. The pellet was washed twice, first with acetone containing 20mM HCl and then with acetone alone after each wash the sample was centrifuged at 13,000xg for 8 min. The pellet was then dried for 10 min subsequently broken apart and resuspended in 25 mM Tris (2% SDS) at 4°C O/N with shaking. The pellet was resuspended in a tenth of the original starting volume resulting in a 10-fold concentration.

2.2.9.2 Chloroplast Isolation

Due to the targeting of the protein facilitated by the pTRAkc-rbcs1-cTP vector it was thought that by specifically isolating the chloroplasts and then resuspending them in a lower volume that an increase in concentration may make it possible to visualize the protein on a western. The clones with plasmids pTRA-saL2-P, pTRA-pL2-P and pTRA-hL2-P were used to infiltrate plants. Chloroplasts were isolated from infiltrated plant material using the Chloroplast Isolation Kit from SIGMA®-Aldrich as per manufacturer's instructions. Chloroplasts were either lyzed in SDS-PAGE loading buffer or lyzed and butanol extracted to remove chlorophyll. Subsequently samples were run on a western.

2.2.9.3 His-tag purification

The clones with plasmids pTRA-saL2-E and pTRA-pL2-E express the L2 protein fused to a His tag. It was thought that by isolating the His tagged proteins and resuspending them in a lower volume that an increase in concentration may allow detection. Extracted leaf sample was used in conjunction with the MagneHis™ protein purification kit (PROMEGA®) which uses the principle that His-tagged proteins bind to Nickel (Ni) particles and the Ni-particles can be magnetically separated out of solution. Once separated the protein is removed from the Ni particles by competitive binding of imidazole. The manufacturer's protocol is specific for microbial isolation thus it was modified to be used in plants. Plant material was extracted in Wash Buffer solution under denaturing conditions and the manufacturer's protocol followed.

2.2.10 RNA Extraction

RNA is notoriously difficult to extract mainly due to the abundance of RNAses in the environment, for this reason immense care needs to be taken. Two-week old *N. benthamiana* plants were injection infiltrated with *Agrobacterium* clones containing the plasmids pTRA-hL2-P, pTRA-saL2-P and pTRA-pL2-P. Total RNA was extracted from the plants at six days post infiltration using the RNeasy® Plant Mini Kit (Qiagen). The protocol was followed according to the manufacturer using the RLT buffer and included a DNase step. RNA was quantified using UV spectrometry at 260nm. RNA was subsequently either stored at -70°C or immediately transcribed into cDNA.

2.2.11 cDNA Synthesis

As mentioned RNA is considered highly unstable making long term storage difficult and risky, for this reason the mRNA was converted to cDNA which is easier to store and less susceptible to degradation. The mRNA was transcribed into cDNA using the Transcriptor First Strand cDNA Synthesis Kit (Roche® Applied Science). The protocol was followed per manufacturer specifications utilizing the supplied anchored-oligo$(dT)_{18}$ primer. Due to the addition of the PolyA tail on all mRNAs to enhance stability this primer ensures only the mRNA is transcribed. The cDNA was subsequently stored at -70°C or immediately used for real time PCR.

2.2.12 Real Time PCR

The cDNA was used for real time PCR utilizing the LightCycler® FastStart DNA Master^{PLUS} SYBR Green I kit (Roche® Applied Science). The protocol was

followed in accordance with those specified by the manufacturer and run using the Rotor Gene RG-3000A (Corbett Research).

All reactions under went an initial pre-incubation step of 95°C for 10 min to ensure the cDNA was denatured and activate the FastStart DNA polymerase. Thirty run cycles were used for all reactions which comprised of a denaturation step followed by an annealing step and finally an extension step. Each reaction used the same denaturation step of 95°C for 10s. The hL2 (F-hL2 and R-hL2), saL2 (F-saL2 and R-saL2) and pL2 (F-pL2 and R-pL2) real time PCR primers had an annealing temperature of 62°C, 54°C and 54°C, respectively, and were allowed to anneal for 10s (Table 3). Two housekeeping genes that did not change with experimental conditions were chosen to normalize the data. They were 18S and Actin-2. 18S (F-18S and R-18S) and Actin-2 (F-Act2 and R-Act2) real time PCR primers with annealing temperatures of 54°C and 55°C, respectively, were also allowed to anneal for 10s within their respective reactions. All reactions had the same extension step of 72°C for 10s. At the end of each run a melt curve was assembled between 72°C and 95°C.

Samples were run in triplicate where possible and a standard curve was generated for each batch of reactions using quantified dilutions of DNA from the corresponding gene of interest, i.e. dilutions of a known concentration of pTRA-hL2-P plasmid were run to generate a standard curve when detecting hL2 mRNA. Data was analyzed using the Rotor-Gene 6.0.41 software.

Table 3: Primers used for RT-PCR

Primer Name	Sequence	Annealing Temperature	Size of Amplified Region
F-hL2	GCAACGACAACAGCATCAAC	62°C	133 bp
R-hL2	GTCCTCAGGGTCTGCTTGTT		
F-saL2	TCCCACTTTCACTGACCC	54°C	221 bp
R-saL2	TGTTGTGTTGTGCGACTA		
F-pL2	CCAGTTAGACCACCACTTAC	54°C	252 bp
R-pL2	GTGGTTGAAGAACAGATGG		
F-18S	GTAAGGATTGACAGACTGAG	54°C	212 bp
R-18S	CAGACCTGTTATTGCCTCA		
F-Act2	ATGCCATCTTGCGGTTAG	55°C	108 bp
R-Act2	CGATTTCCCGTTCAGCAG		

2.2.13 Arabidopsis Small RNA Project (ASRP) Data Analysis

The sequences of the various L2 genes were in essence compared by BLAST analysis (36) against the database created by the ASRP. By entering the user-defined sequence the TargetFinder program searches the sequence for potential small RNA target sites. Using the parameters: mispair, bulge or gap – 1.0; G:U pair – 0.5; with a double penalty if the mispair, bulge, gap or G:U pair occurs within the 2 to 13 region relative to the 5' end of the small RNA; and a maximum of one single nucleotide bulge or gap allowed. The results are given as predicted scores and if this value is equal to or below four it is a plausible small RNA target site. The database comprises of 1920 unique small RNA sequences from various sources that have been discovered in *Arabidopsis* (36).

2.2.14 Preparation of Plant Material for Electron Microscopy

One of the primary concerns with electron microscopy is preservation of the sample material such that what is seen is a true reflection of what is happening *in vivo*. Two week old *N. benthamiana* plants were infiltrated with *Agrobacterium* clones with plasmids pTRA-hL2-A, pTRA-hL2-C, pTRA-hL2-E and pTRA-hL2-P. As a negative control plant material that had not been infiltrated was used. Leaves were cut into strips after six days and immediately fixed overnight in 2.5% gluteraldehyde made in 0.1 M PBS and containing 0.5% caffeine. The samples were then washed three times in 0.1 M PBS followed by a dehydration sequence. The sequence consisted of 2 x 5 min in 30% ethanol, 2 x 5 min in 50% ethanol, 2 x 5 min in 70% ethanol, 2 x 5 min in 90% ethanol, 2 x 5 min in 95% ethanol and finally 2 x 10 min in 100% ethanol. The samples were then gradually impregnated with LR white resin before placing them in pure resin in beem capsules and allowing it to set at 60°C for 16 hours.

Knives were made from glass and used in the microtome to section the samples. Two thicknesses of sections were made, 120 nm and 90 nm, with the finding that 90 nm sections were ideal for our purposes. Ribbons of section were placed on nickel grids that had been coated with butivar. Butivar serves to create a scaffold in between the spaces of the grid to reinforce the thin, fragile sections.

Grids were then blocked with a PBST solution with 5% skim milk for 30 min. The primary antibody (polyclonal rabbit αL2 antibody) was then allowed bind at a dilution of 1:300 in PBST with 5% skim milk at 4°C overnight. Grids were then washed 20 times in PBST with 5% skim milk followed by binding of the secondary antibody (monoclonal goat αRabbit antibody conjugated with 10 nm gold particles) at a dilution of 1:50 in PBST with 5% skim milk at room temperature for 60 min. Grids were later washed five times in PBST and then

fixed in 1% gluteraldehyde in PBS for 5 min. This second fixation step ensures the target – 1° Ab – 2° Ab complex is preserved throughout the subsequent staining.

Uranyl acetate and lead citrate staining was used. First, the grids were stained with uranyl acetate for 10 min; then washed five times with water. Then stained with lead citrate for 10 min and finally washed 40 times with water. Grids were allowed to dry and then used for electron microscopy.

2.2.15 Electron Microscopy

Grids with plant sample were visualized using the JEOL 200CX transmission electron microscope available at the Electron Microscope Unit of UCT.

2.2.16 Butanol Extraction

Butanol extraction serves to remove various plant compounds that may interfere with later applications such as polyphenols and chlorophyll. For this reason it was used in the purification of VLPs and in chloroplast isolation. Leaves were macerated in high salt PBS buffer in a Woring blender and homogenized using a Silverston homogenizer. Tween-80 (final 0.05%) was added to solubilize L2 and to remove proteins that are loosely attached to membranes. The solution was subsequently put through a vinyl mesh to remove large debris. Butanol was added to the filtrate at a ratio of 1:1 and left to separate into the aqueous and non-polar phase. The aqueous phase should contain all soluble protein and was subsequently removed for analysis.

2.2.17 VLP Purification

2.2.17.1 High Salt PBS Buffer

A specific buffer was used in the extraction of VLPs. This buffer contained 0.2 M – 0.5 M PBS with the concentration varying in accordance with the amount of plant material i.e. a higher concentration was needed with more plant material in a lower volume. In addition, 0.1 M NaCl was added which increases the ionic strength and was found to maintain the integrity of the VLPs. 2% PVPP (mass to volume) was also added which acts on polyphenols to prevent them from being oxidized to quinones which releases free radicals and so damages other proteins. 0.1 M sodium metabisulphate was added to inhibit proteases by chelating metal ions and preventing oxidases from releasing free radicals. Finally 0.01% sodium azide was added to remove any bacteria and fungi.

2.2.17.2 Differential Centrifugation

Differential centrifugation was used to isolate VLPs capitalizing on the fact that they are denser than single molecule proteins. First, leaf material either derived from physical extraction or butanol extraction was subjected to a low speed spin at 3,800xg. The supernatant was removed and spun for 4h at 28,000 rpm. The pellet was resuspended in buffer and the supernatant kept for analysis. At this stage the VLPs are presumed to be present in the pellet whilst single molecule proteins remain in the supernatant.

2.2.17.3 Sucrose Gradient Fractionation

Sucrose gradient fractionation was further used to purify VLPs. Either the pellet from the differential centrifugation or sample directly extracted was used. The method behind this technique is to create a density gradient using sucrose and then load the sample on top of the gradient and allow it to "run" through the gradient. In principle, the higher density structures should migrate quicker through the gradient whilst the less dense molecules remain closer to the top of the gradient. In this way it allows separation of the denser VLP structure from the less dense single protein. The pellet from the VLP isolation was thus loaded on a 5-50% sucrose gradient and spun at 37,000 rpm for 4h. Fractions were taken and analyzed.

2.2.17.4 Immunoprecipitation

Immunoprecipitation was another method used to try and isolate VLPs. Guinea pig-derived α-L1 antibody (1:50) was mixed with extracted leaf material. Protein A conjugated to agarose beads (SIGMA®-Aldrich) was subsequently added and centrifuged to pellet the beads. The supernatant was removed and the pellet, directly run on a western.

2.3 Results

2.3.1 A silencing suppressor was used in infiltration.

In this system, GFP-imaging was used to determine if a silencing pattern was present. GFP expression was used as it is an easily observable positive control for the infiltration procedure. Two week old *N. benthamiana* plants were infiltrated with the *Agrobacterium* GV3101 strain containing GFP (pTRA-GFP-P) alone or co-infiltrated with *Agrobacterium* LBA4404 coding for NSs (pBIN-NSs). GFP was visualized in infiltrated leaves at 2, 6, and 10 days post infiltration (dpi) under UV light.

At 2 dpi the GFP expression was similar whether NSs was present or not (Fig. 3). At 6 dpi there was an apparent difference in fluorescence with the GFP co-infiltrated with NSs having noticeably more expression. A marked difference in expression can be seen at 10 dpi between GFP infiltrated alone and GFP co-infiltrated with NSs, with significantly more fluorescence when co-infiltrated.

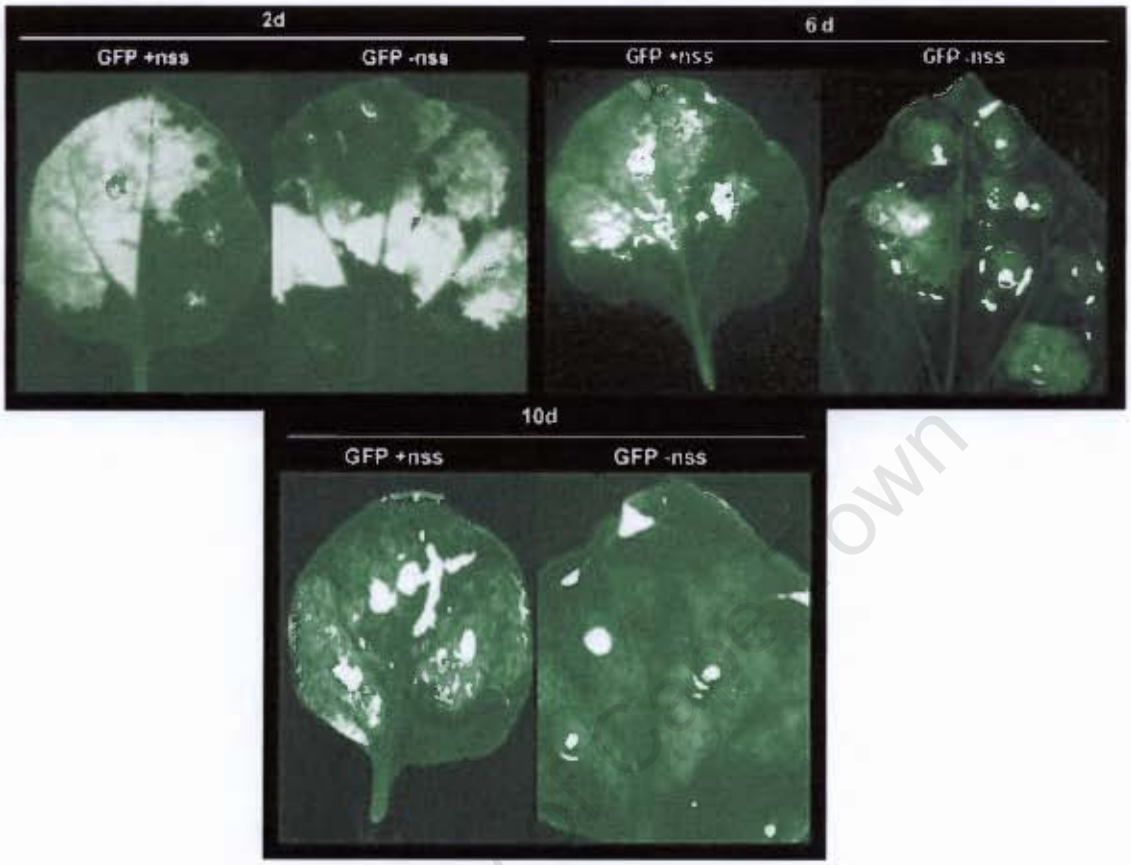

Figure 3. Expression of GFP at 2 days, 6 days, and 10 days post infiltration. The *Agrobacterium* GV3101 clone containing the pTRA-GFP-P (GFP) plasmid was either co-infiltrated with the *Agrobacterium* LBA4404 clone with plasmid pBIN-NSs (−nss) or without (-nss). All *Agrobacterium* clones were infiltrated at a final concentration of 0.5 OD_{600}. UV-imaging was done on infiltrated leaf samples at 2 days (2d), 6 days (6d), and 10 days (10d) post infiltration. Images were captured with a short wave UV light box.

The NSs protein came from a gene from an isolate of the *Tomato spotted wilt virus* (76). A similar silencing suppressor, p19, was isolated from *Tomato bushy stunt virus* (85). The p19 gene was subsequently cloned into the pTRAc vector and was found to have marginally better silencing suppression than NSs (unpublished in lab). Use of the same *Agrobacterium* GV3101 also proved easier to work with. For these reasons, all future infiltrations were done with the *Agrobacterium* clone containing the plasmid with the p19 gene (pTRA-p19-C).

2.3.2 Comparison of L2 protein expression in plants when coded for by the wildtype and plant-optimized gene.

Expression of L1 in plants was very low when coded for by the wildtype gene and a plant-optimized (plantized) version of the gene (6). We sought to determine if a similar phenomenon occurs with respect to L2 protein expression when coded by the wildtype gene (saL2) and a plantized version (pL2) thereof.

Agrobacterium clones were made that contain the L2 wildtype or L2 plantized gene in each of the various pTRA vectors. These vectors differ to each other in that each attaches a targeting signal to the protein which directs it to various compartments in the cell. The pTRAc vector (clones represented as pTRA-"gene"-C) attaches no signal to the protein and thus as a default the protein is presumed to be targeted to the cytoplasm. The pTRAkc-ERH vector (clones represented as pTRA-"gene"-E) attaches a secretory signal to the N-terminus of the protein and a KDEL sequence to the C-terminus which are presumed to target the protein along the secretory pathway and retain the protein within the endoplasmic reticulum (ER), respectively. The pTRAkc-rbcs1-cTP vector (clones represented as pTRA-"gene"-P) attaches a chloroplast-targeting signal to the N-terminus of the gene that facilitates targeting to the chloroplast.

Clones used included pTRA-saL2-C, pTRA-saL2-E, pTRA-saL2-P, pTRA-pL2-C, pTRA-pL2-E, and pTRA-pL2-P. All six of these clones were agro-infiltrated into *N. benthamiana* plants and transient expression was measured at various time intervals from infiltration. Leaf material was extracted in 8M urea and samples were detected using western blot analysis. No L2 signal was detected in any of the samples; as a positive control *E. coli*-expressed L2 (200 ng) was loaded and an L2-specific ~72 kDa signal was seen in each western (data not shown).

We thought L2 expression may be too low to be detected in a western, for this reason we used various methods to determine if any L2 protein was present. The ELISA detection method is known to be more sensitive than a western; however, an L2 ELISA of the samples produced no meaningful results. The 8M urea could play a role in inhibiting detection through ELISA and so different buffers were used yet this had no significant effect on detection either.

A number of concentrating methods were also employed to increase the amount of L2 to be detected. These included chloroplast isolation, his-tag purification, and TCA-precipitation. No L2 protein was able to be detected in any of the leaf samples using either of these methods.

Thus, it can be confidently concluded that no L2 protein expression in plants is present or is able to be detected when coded for by the wildtype gene or a plantized version thereof.

2.3.3 Expression from mammalian codon-optimized L2 in plants.

Expression of L1 in plants was found to be best when coded for by the human codon-optimized version of the gene (6). Again, we sought to determine if a similar phenomenon can be seen in L2 expression in plants.

Agrobacterium clones were made that contain each of the pTRA vectors each containing the mammalian codon-optimized (humanized) L2 gene (hL2). Initial determination of expression was done using the pTRA-hL2-C *Agrobacterium* clone which was infiltrated into two-week old *N. benthamiana* plants. At 5 dpi protein was extracted from plant material in 8M urea and expression was determined by western blot analysis (Fig 4). A positive control, 200 ng of *E. coli*-

expressed L2 protein, was also loaded. A clearly visible band representing the L2 protein can be seen that was present in the positive control and the plant material infiltrated with the *Agrobacterium* clone containing pTRA-hL2-C but absent in the negative control.

Two methods to quantitate the amount of L2 protein present were employed, namely ELISA and western blot analysis. Protein was extracted in 8M urea from leaf material infiltrated with the *Agrobacterium* clone containing plasmid pTRA-hL2-C at 5dpi. The leaf material was initially weighed prior to extraction with a 10-fold and then 2-fold serial dilution of the sample being made ranging from 8mg (undiluted) – 0.2 mg (40-fold dilution). An L2 ELISA was attempted on these dilutions but, as previously found, the amount of L2 was unable to be determined. Again this may be due to the 8M urea buffer being used or possibly interference created by other plant proteins. A crude quantitation was completed utilizing western blot analysis (Fig. 4).

The dilutions were electrophoresed and transferred to nylon membrane for detection through western blot. The undiluted sample (8 mg), 10-fold dilution (0.8 mg) and 20-fold dilution (0.4 mg) all produced a detectable protein band for L2 confirmed by the positive control of 200 ng of *E. coli*-expressed L2 (Fig. 4). This L2 band was absent in the negative control (non-infiltrated plant material) (Fig. 4). Upon analysis of the intensity of the band in the positive control and that of the undiluted sample it can be speculated that they are roughly equivalent. Thus, it is possible to deduce that the amounts of L2 in both are roughly the same, i.e. both appear to contain 200 ng of L2 protein. With this in mind it is possible to estimate that the amount of L2 is 200 ng per 8mg of plant material or 25mg.kg^{-1} of plant material. This is lower than expression levels of L1 which had the highest value of 533 mg.kg^{-1} (54).

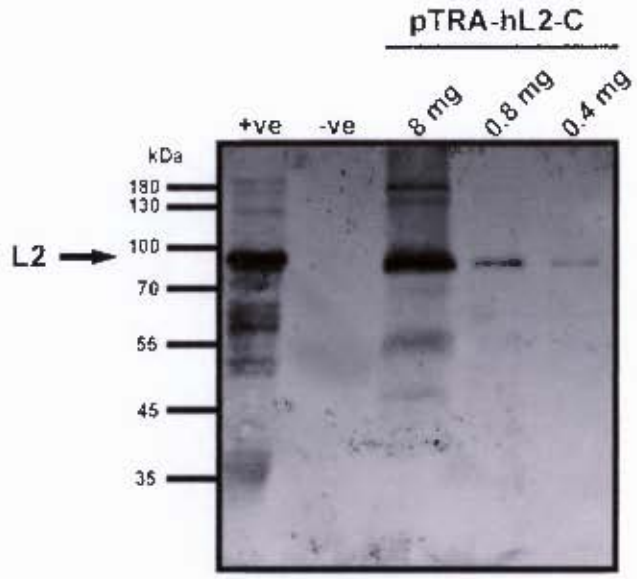

Figure 4. Determination and quantitation of L2 protein expression in plants. A two week old *N. benthamiana* was infiltrated with the *Agrobacterium* clone containing plasmid pTRA-hL2-C. At 5 dpi, plant material was weighed and extracted in 8M urea. Serial dilutions were made of the extracted leaf material and electrophoresed. The amount of leaf material loaded in each well (8 mg, 0.8 mg, and 0.4 mg) was extrapolated from the original weight. L2 expression was determined by western blot analysis. An L2-specific signal was detected using rabbit polyclonal α-L2 and mouse monoclonal α-rabbit as the primary and secondary antibodies, respectively. Plant material that had not been infiltrated was also extracted in the same method and loaded on the gel as a negative control (-ve). As a positive control 200 ng of *E. coli*-expressed L2 was loaded on the gel (+ve).

The data so far indicates high L2 protein expression in plants when it is coded for by the humanized gene, and undetectable expression when it is coded for by either the wildtype or plantized sequences of the gene.

2.3.4 Comparison of RNA levels of the various L2 genes.

Given the differential protein expression between the different sequences of L2 we sought to determine whether expression is limited at the RNA level due to limited amounts of transcript being present for translation.

Real time PCR employs certain key principles in the technique, namely, the immense power of amplifying small quantities of template and the high specificity

that accompanies PCR in addition to spectroscopy of fluorophores that fluoresce when they interact with DNA. SYBR green is such a fluorophore that intercalates with double stranded DNA (dsDNA) and fluoresces when it does so. Thus, it becomes possible to determine at what time the dsDNA within a PCR reaction is formed. This time point is important in determining the initial concentration of the template at the start of the reaction. Therefore by running various standards of known concentration simultaneously with the sample of unknown concentration and measuring at what time these dsDNA products of the standards are formed it becomes possible to construct a standard curve to determine the concentration of the unknown.

Agrobacterium clones containing the pTRA-hL2-P, pTRA-pL2-P and pTRA-saL2-P plasmids were infiltrated into two week old *N. benthamiana* plants. RNA was extracted at various time points to determine the optimal time at which to harvest the RNA. An incubation period of six days was found to be ideal for our purposes.

RNA samples were first transcribed into cDNA and then run through Real Time PCR amplifying mRNA for pL2, saL2 and hL2. All reactions were run simultaneously with respective standards of the genes of known concentrations. In addition, mRNA from two reference genes, namely 18S and Actin-2, were also amplified from the same RNA samples along with their own standards. These two reference genes were chosen based on the fact that their levels did not change with a change in experimental conditions and may also be termed housekeeping genes.

The concentrations of pL2, saL2 and hL2 mRNA were calculated using the Rotor-Gene 6.0.41 software program. Concentrations of the two reference genes in the RNA samples were also calculated in the same way. The program measures the time points at which the DNA products form both within the unknowns and the

standards and constructs a standard curve that best fits both the unknown time points and the standard time points. In this way the time points are translated into concentrations. This highlights the importance of the standard curve and in this study where the unknown time points did not fall within two standard time points these results were not considered.

In addition, a melt curve was constructed for each reaction between 72°C and 95°C. Fluorescence is detected at incremental increases in temperature, i.e. it is measured at 72°C then the temperature is increased to 73°C and measured again. As SYBR green only fluoresces when intercalated with dsDNA with an increase in temperature so the dsDNA is dissociated into ssDNA at which point the SYBR green no longer fluoresces. This allows you to determine what products were formed within the PCR reaction by plotting the derivative of fluorescence divided by the derivative of temperature (Y-axis) against temperature (X-axis). Thus, peaks are formed at the temperature that products dissociated and so served as a further stringency parameter to ensure the results were reliable, i.e. if additional peaks formed at lower temperatures indicating primer dimers or non-template specific product these results were discarded.

The concentrations of the mRNA of the various L2 genes was subsequently normalized using the concentrations of each of the reference genes. The amount of pL2 mRNA was found to be the lowest using both reference genes and thus was used to calculate the relative amounts in comparative analysis. The mean was calculated between the normalized data generated from both reference genes and found to be in the ratio of 1 : 11.59 : 21.84 for pL2 : saL2 : hL2 (Fig. 5).

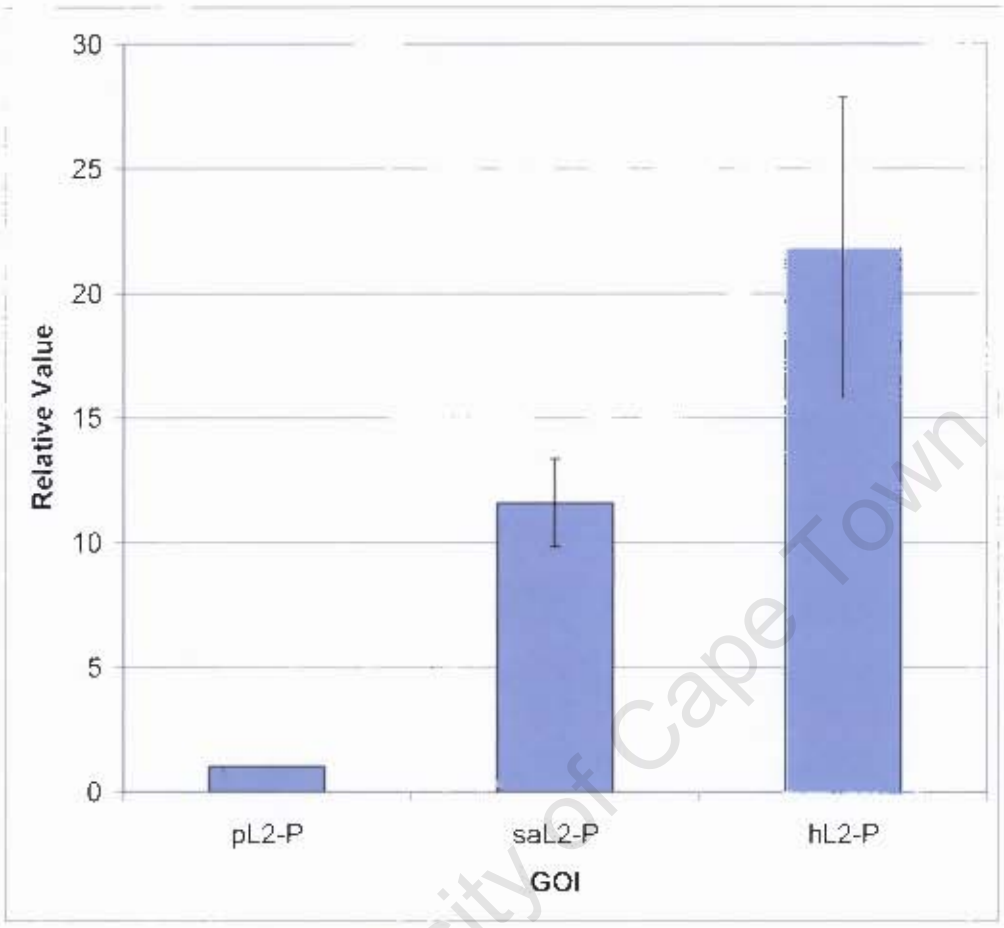

Figure 5. Relative levels of mRNA coded for by the various sequences of the L2 gene. Two week old *N. benthamiana* plants were injection infiltrated with *Agrobacterium* clones containing plasmids pTRA-pL2-P (pL2-P), pTRA-saL2-P (saL2-P) and pTRA-hL2-P (hL2-P). Total RNA was extracted at 6 dpi and the mRNA was converted to cDNA. The levels of transcript of the various genes (GOI) were determined by RT PCR. Using both 18S and Actin-2 as reference genes the relative amounts of transcripts were compared. The mean between normalization using 18S and Actin-2 was calculated with the error bars indicating the variation between the two.

2.3.5 *In silico* comparison of siRNA binding to the various L2 genes.

It is well established that siRNAs act to limit gene expression and regulate RNA amounts through PTGS in plants. This includes regulation of over-expressed viral genes and endogenous genes alike. Thus a pool of siRNAs are present within a plant cell and are indeed able to move to other cells giving the system immune response-type capabilities (89). With this in mind we sought to determine if this pool of endogenous siRNAs may help to explain why L2 expression was markedly less when coded by the plantized L2 in comparison to the humanized and wildtype genes.

At present there are no databases for small RNAs isolated from Tobacco. However, there is a database that has been recently constructed that contains some of the characterized small RNAs in *Arabidopsis*. The sequences of the hL2, saL2 and pL2 genes were entered as user-defined sequences that were then blasted by the TargetFinder program to determine possible small RNA target sites. Results are given in predicted scores whereby a value equal to or lower than 4 are considered plausible small RNA target sites. Interestingly, pL2 had the highest number of hits with eleven small RNAs with a score of 4, four with a score of 3.5 and one with a score of 3 (Fig. 6). The saL2 and hL2 number of hits were comparably similar with saL2 having two hits at a value of 4 and one at a value of 3.5 (Fig. 6). Whilst hL2 had three hits with a score of 4 and one hit with a score of 3 (Fig. 6).

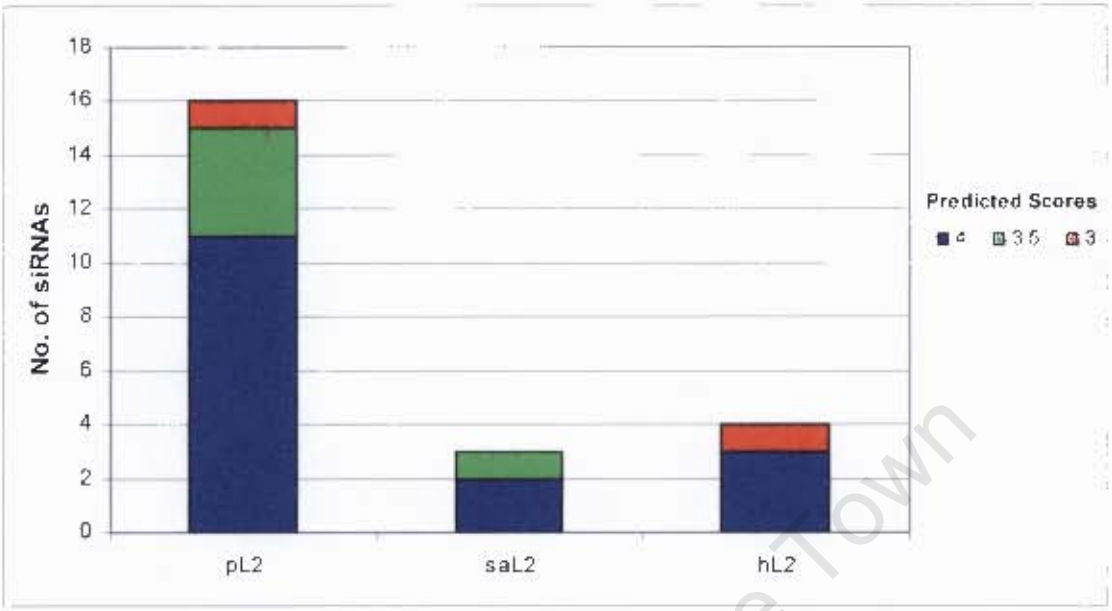

Figure 6. ASRP analysis of the various RNA sequences of the different L2 genes. The RNA sequences of the plantized L2 gene (pL2), wildtype L2 gene (saL2) and the humanized L2 gene (hL2) were determined and entered as user-defined sequence in the ASRP web interface. The TargetFinder program, which the web page uses, searches their small RNA database for matches with the sequence and produces likely binding sites with prediction scores below or equal to 4 indicating a probable site for small RNA binding. A score of 3.5 indicates a more likely site of binding than a score of 4 and likewise a score of 3 indicates an even higher likelihood than 3.5 or 4.

This data suggests that the L2 mRNA from the plantized gene may preferentially be targeted for degradation by the endogenous siRNAs leading to limited levels of protein expression.

2.3.6 Expression of L2 protein using the pTRA vectors.

After determining a plausible reason for limited expression when coded by different genes we turned our attention to protein expression that could be seen i.e. expression of L2 from the humanized gene. Interestingly, L1 expression when coded for by the humanized version of the gene produced different levels of expression across the different pTRA vectors (54). More specifically, expression of L1 was greatest when the pTRAkc-rbcs1-cTP vector was used; expression when

using the pTRAc vector was less and expression utilizing the pTRAkc-ERH vector was barely detectable (54). These vectors are identical in all respects except for the targeting peptides attached. For this reason, the variation can only be explained by the different signaling pathways the proteins would be targeted to and indicates a possible protection from degradation in different compartments of the cell.

We thus sought to determine if a similar differential pattern in L2 protein expression was seen when using the different pTRA vectors.

As mentioned previously, the different pTRA vectors fuse signaling peptides that target the protein to various compartments within the plant cell. The pTRAc does not attach a signal to the protein and thus as a default the protein is thought to be targeted to the cytoplasm. The pTRAkc-A vector attaches a secretory signal (LPH) to the N-terminus which targets the protein along the secretory pathway and is assumed to be targeted to the apoplastic space. The pTRAkc-ERH vector also attaches the same secretory signal to the N-terminus but in addition attaches a KDEL sequence to the C-terminus which serves to retard the protein in the ER. The pTRAkc-rbcs1-cTP attaches a chloroplast transit peptide to the N-terminus and is thus assumed to target the protein to the chloroplast.

In addition to the *Agrobacterium* clone containing pTRA-hL2-C, clones were made that contain the humanized L2 (hL2) gene within the pTRAkc-A, pTRAkc-ERH and pTRAkc-rbcs1-cTP vectors; creating the plasmids pTRA-hL2-A, pTRA-hL2-E and pTRA-hL2-P, respectively. These clones were infiltrated into two week old *N. benthamiana* and protein was extracted in 8M urea at various time points after initially weighing them to ensure all leaf material was standardized. Presence of L2 in the samples was detected using western blot analysis (Fig. 7). 200 ng of *E. coli*-expressed L2 was also loaded to confirm the L2 band which was not present in the negative control (non-infiltrated plant material) (Fig. 7). All SDS-

PAGE gels were stained with coomassie blue post-transfer where similar amounts of rubisco in each of the samples, including the non-infiltrated plant material, could be confirmed. This allows us to reliably infer an expression pattern from these western results.

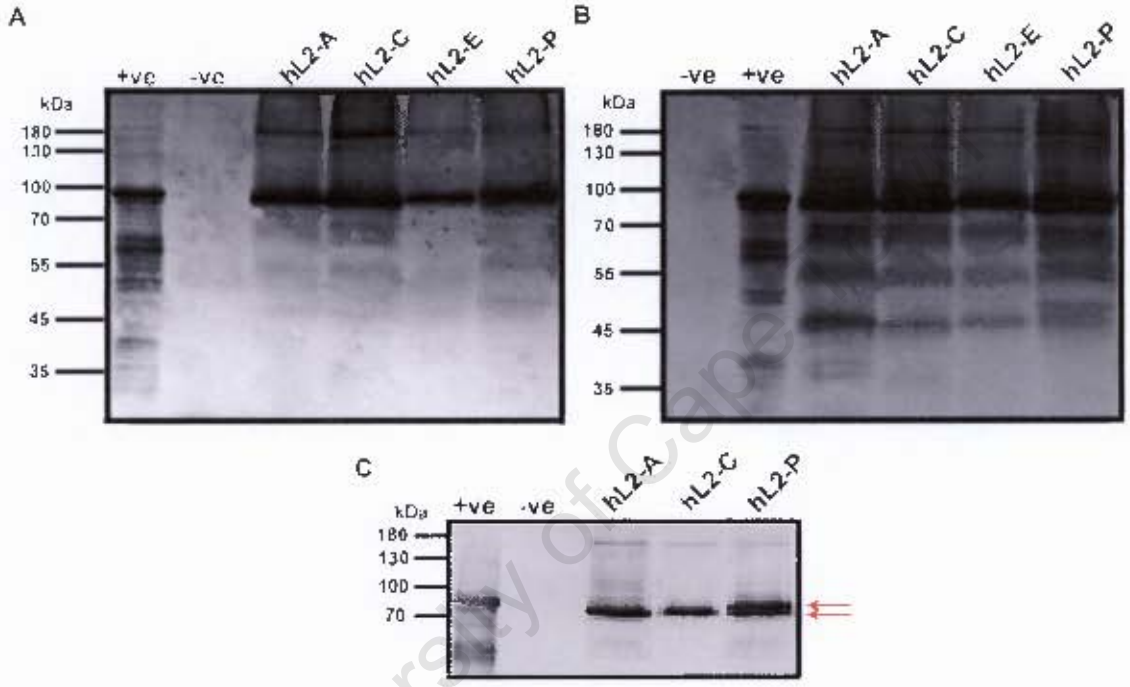

Figure 7. Expression of L2 protein detected by western blot. Two week old *N. benthamiana* plants were infiltrated with *Agrobacterium* clones containing plasmids pTRA-hL2-A (hL2-A), pTRA-hL2-C (hL2-C), pTRA-hL2-E (hL2-E), or pTRA-hL2-P (hL2-P). Leaf material was extracted in 8M urea at 2 dpi (A and C) and at 5 dpi (B). The doublet present in hL2-P samples has been clearly indicated by the arrows (C). Samples (32 µl) were loaded on an SDS-PAGE and subsequently detected using the western blot method. As a negative control, non-infiltrated plant material was extracted in the same method and electrophoresed concurrently (-ve). As a positive control 200 ng of *E. coli*-expressed L2 were loaded (+ve). An L2-specific signal was detected using rabbit polyclonal α-L2 and mouse monoclonal α-rabbit as the primary and secondary antibodies, respectively.

It is notable that the expression pattern at two dpi indicates the levels are relatively similar. In addition, at two dpi there are relatively few broken down products below the L2 band. At five dpi expression appears to follow the same pattern of relatively similar levels. At this time point more broken down products appear to be present and following a similar pattern indicating the same pathway of degradation.

Interestingly, a doublet of the L2 band can be seen in the chloroplast targeted expression (pTRA-hL2-P) (Fig. 5C). This phenomenon was seen in all westerns containing this sample and can be explained by the presence of the chloroplast-targeting signal. The signal is 58 amino acids in size and so is big enough to alter the migration pattern. In addition once the protein enters the chloroplast this signal is cleaved off. Therefore two bands corresponding to L2 would be expected if the protein was indeed targeted to the chloroplast.

These results indicate no difference between expression of L2 when incorporated into the different pTRA vectors. In addition, the migration pattern of L2 in pTRA-hL2-P samples indicates the protein is being targeted to the chloroplast.

2.3.7 *In silico* predictions of L2 targeting with attached signal peptides.

Having confirmed no difference in expression of L2 when using the different pTRA vectors we sought to determine possible reasons for this. More specifically, we hypothesized that the reason expression was not variable between the different vectors was because signaling peptides, such as a nuclear localization signal (NLS) within the L2 protein were interfering with targeting. Consequently, we thought, the L2 protein was in fact accumulating at the same site independent of the attached signal peptides fused to the protein by the pTRA vectors.

A number of targeting prediction programs exist each using different methods and parameters of predicting localization. They essentially rely on datasets of signaling peptides from proteins of known localizations and search for sequence homology between these and the user-defined sequence. These programs also differ in the

compartments they are able to predict localization to (39). Seven programs were used in this study that are freely available to the public through the internet. They include Protein Prowler (http://pprowler.imb.uq.edu.au/index.jsp), TargetP (http://www.cbs.dtu.dk/services/TargetP), SignalP (http://www.cbs.dtu.dk/services/ SignalP), MitoProt II (http://ihg.gsf.de/ihg/mitoprot.html), Predotar (http://urgi.Versailles.inra.fr/predotar/predotar.html) and SubLoc (http://www.bioinfo.tsinghua. edu.cn/SubLoc).

The amino acid sequences of the signaling peptides were attached to the sequence of the HPV-16 L2 protein *in silica* using DNAMAN (Fig. 8). These sequences were then loaded as user-defined sequences in all of the programs and subsequently "blasted" to predict the localization with the organism specified as plant in all programs. The output formats differ across the programs; however, I have taken the qualitative information given by each of them (Table 4).

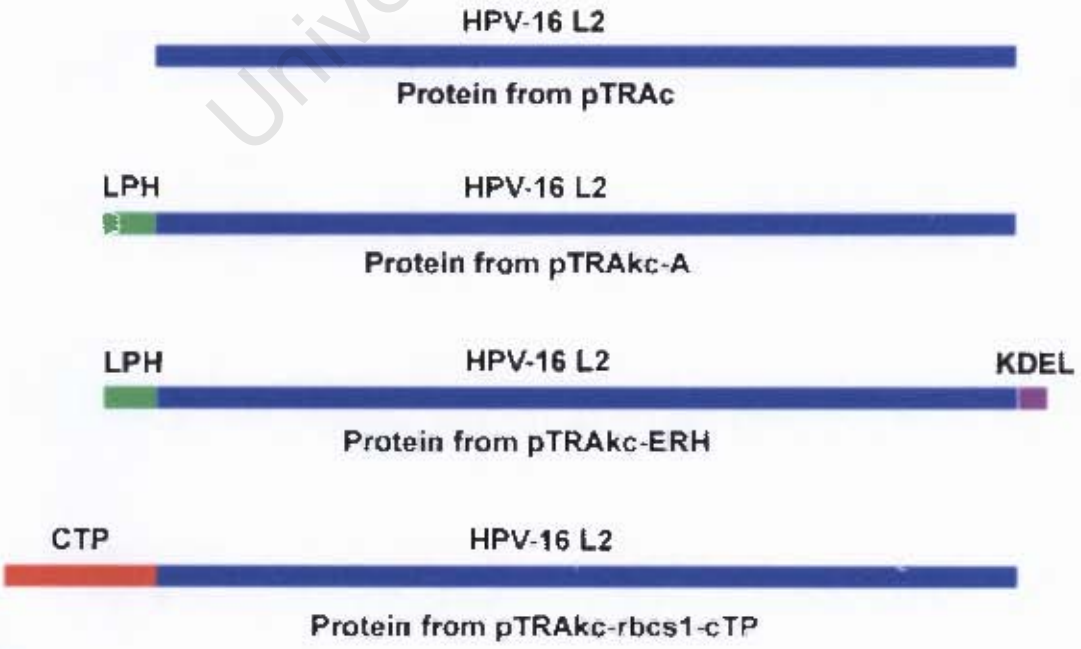

Figure 8. A diagrammatic view of the L2 protein with attached signaling peptides. The secretory signal (LPH) from the pTRAkc-A and pTRAkc-ERH vectors is attached to the N-terminus. The pTRAkc-ERH vector further attaches a KDEL sequence at the C-terminus to retain the protein in the ER. The pTRAkc-rbcs1-cTP vector attaches a chloroplast targeting signal (CTP) to the N-terminus of the protein.

Table 4: The qualitative information obtained from the various targeting prediction programs.

Prediction Program	L2 protein alone	LPH + L2	LPH + L2 + KDEL	cTP + L2
Protein Prowler	Mito	Sec/Mito	Sec/Mito	Chlo
TargetP	XXX	XXX	XXX	Chlo
SignalP	none	Sec	Sec	none
SubLoc	XXX	ExCell	ExCell	XXX
Predotar	XXX	ER	ER	Chlo
MitoProt II	Mito	Mito	Mito	Mito

I have removed predictions that I did not consider reliable. This was facilitated by creating parameters I considered relevant to determine accuracy. These included: the setting of specificity at 0.9 for TargetP which excluded the results of 3 of the amino acid sequences from being included (indicated by XXX) (Table 4). For SubLoc if the expected accuracy was below 70% it was excluded (marked by XXX in table 4). If Predotar predicted the localization to be "elsewhere" it too was not considered (also marked by XXX in table 4). If the percentage probability of the protein being targeted to the mitochondria in MitoProt II was below 0.7 it too was removed.

Upon analysis of the results two of the programs predicted the L2 protein alone may be targeted to the mitochondria (Mito), but one (MitoProt II) seems to indiscriminately assign this same prediction to all the sequences. There is a plausible reason for this, however; MitoProt II can only predict the probability of

targeting to the mitochondria. Consequently, it predicts the same mitochondrial targeting peptide in all the sequences as the signal is within the N-terminus of the L2 protein which it consistently recognizes. It fails however, to recognize other signals that may influence targeting such as a chloroplast transit peptide or secretory signal.

The sequences from the pTRA-hL2-A and pTRA-hL2-E vectors indicate targeting along the secretory pathway (ExCell, Sec, ER). A signal peptide, responsible for secretion, was found by SignalP in both these sequences at the position of the LPH. Interestingly, TargetP, which is closely linked to SignalP, failed to recognize this signal peptide and indicated targeting to the mitochondria instead (although the value was below the accepted parameter). This same phenomenon was found in Protein Prowler.

This indicates that the two targeting peptides may be competing. Moreover, where the target group was changed from plant to animal; recognition of the mitochondrial-targeting peptide was removed, indicating specific targeting in plant cells. Most of the programs analyze the N-terminal region only, making a distinction of the sequence containing KDEL difficult from those which do not. However, it is well established in the literature that if a KDEL sequence is present and the protein is in the ER it will be retarded.

Finally, most of the programs specifically recognized the chloroplast transit peptide attached to the N-terminus of the sequence derived from pTRA-hL2-P, with all confidently predicting targeting to the chloroplast.

In summary, these results suggest that the L2 protein alone when in the plant cell is targeted to the mitochondria by a strong targeting peptide within the N-terminus of the protein. Furthermore, they indicate that when the LPH signal is attached it

may conflict with the mitochondrion targeting and thus it is questionable which targeting succeeds over the other. From the prediction programs it is difficult to say whether the KDEL sequence has influenced the results but it can be assumed that if the protein finds itself in the ER it will be retarded by this sequence. Lastly, with attachment of the chloroplast transit peptide it appears that L2 will be targeted to the chloroplast.

2.3.8 Electron Microscopy to elucidate targeting of L2 in the plant cell.

Having established potential sites for targeting through *in silico* prediction we sought to determine if these can be verified through empirical data.

Two-week old *N. benthamiana* plants were infiltrated with the various *Agrobacterium* clones containing plasmids pTRA-hL2-A, pTRA-hL2-E, pTRA-hL2-C and pTRA-hL2-P. plant material was fixed with gluteraldehyde and dehydrated at 5 dpi. Samples were then embedded in LR white resin and 90 nm sections were made using the microtome. Sections were loaded onto Nickel grids that had been coated with butivar. These were subsequently immunolabelled for L2 with the polyclonal rabbit α-L2 antibody and monoclonal goat α-rabbit conjugated with 10 nm gold particles as the primary and secondary antibodies respectively. Electron micrographs were visualized using the transmission electron microscope. Sections were analyzed for clearly defined gold label indicating presence of L2.

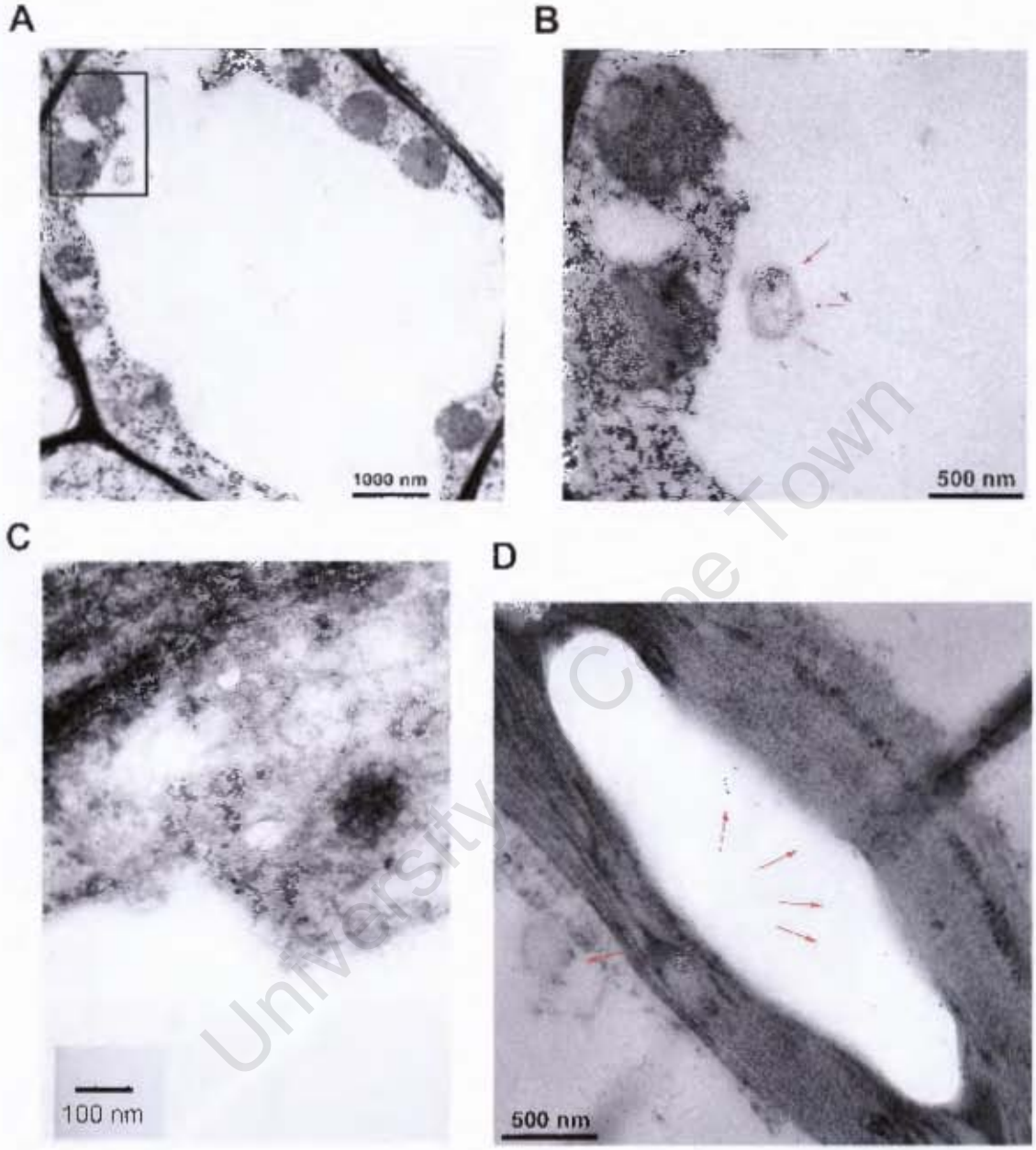

Figure 9. Electron Micrograph of L2 expression. Two week old *N. benthamiana* plants were infiltrated with *Agrobacterium* clones containing plasmids pTRA-hL2-C and pTRA-hL2-P. Plant material was fixed with gluteraldehyde and dehydrated using ethanol after 5 dpi. Samples were then embedded in LR white resin and 90 nm sections were made using a microtome. Sections were loaded onto Nickel grids coated with butivar. Immunolabelling was done with rabbit polyclonal α-L2 antibody and goat monoclonal α-rabbit conjugated with 10 nm gold particles as the primary and secondary antibodies, respectively. Imaging was done with the transmission microscope at various magnifications. (A) Plant material infiltrated with pTRA-hL2-C clone. (B) A higher magnification of the boxed area in (A) with arrows indicating specific immunolabel of L2. (C) Plant material also infiltrated with the pTRA-hL2-C clone with two specific immunolabels present in the centre of the micrograph. (D) Plant material infiltrated with the pTRAhL2-P clone. Thylakoid membranes indicate a chloroplast and specific immunolabel both within and outside of the chloroplast are indicated by the arrows.

Label was found to be present in sections obtained from leaf material infiltrated with pTRA-hL2-C (Fig. 9). The integrity of cells was found to be sufficiently preserved evidenced by the intact cells with large central vacuoles and surrounding plant structures presumed to be various plastids (Fig. 9A). Upon closer magnification of the area enclosed in the black box of figure 9A specific label can be seen to accumulate within a membrane-bound organelle (Fig. 9B). This suggests the possibility of the L2 protein being targeted to a specific organelle. It may be speculated to be a mitochondrion as it is smaller than surrounding organelles which could represent chloroplasts of which, in addition, there are more present in the cell. Specific label was, however, also found in the cytoplasm of various samples (Fig. 9C).

In the leaf material infiltrated with pTRA-hL2-P specific label was consistently found to be located to an organelle structure (Fig. 9D). The presence of thylakoid membranes within the structure confirms the organelle to be a chloroplast. Specific label was found both within and outside of the chloroplast confirming earlier results of western analysis (Fig. 9D). Interestingly, the large, central, clear area within the chloroplast was found to be present in samples infiltrated with pTRA-hL2-P and absent in leaf material from non-infiltrated plants and plants infiltrated with the other clones. In addition, specific label was commonly found within this clear area. At first it was presumed to be an artifact and thought to occur as a result of the resin splitting in certain areas. However, this phenomenon was consistently found in subsequent results with the clear area always being perfectly enclosed within the chloroplast and only when infiltrated with the pTRA-hL2-P clone. We thus suggest that the area is in fact a result of accumulation of the L2 protein within the chloroplast.

Leaf material obtained from infiltration with pTRA-hL2-A and pTRA-hL2-E were also analyzed by electron microscopy. However, no conclusive information as to

their specific targeting could be obtained. In the case of pTRA-hL2-A this may be as a result of where it is presumed to be targeted, namely the apoplastic space. Due to the dense nature of cell walls they appear as thick black lines in an electron micrograph. This makes it impossible to determine a clearly defined label in between the cell walls. In the case of material infiltrated with pTRA-hL2-E samples indicated targeting to the cytoplasm, however, the amount of label was found to be dramatically low and close to the number of non-specific label found in the non-infiltrated material. For this reason the results were excluded.

In summary, these results suggest the localization of the L2 protein with no attached targeting peptides is either the cytoplasm or possibly a membrane enclosed vacuole or organelle which may be mitochondria. L2 protein with an attached chloroplast transit peptide is localized to the chloroplast where it accumulates. Localization of the L2 protein with either the LPH signal alone or in addition to the KDEL sequence needs to still be determined.

2.3.9 Co-expression of L1 and L2 in plants.

As previously mentioned L1 contains the intrinsic knowledge to form virus-like particles that mimic the HPV virion structure and so provide a perfect candidate for a vaccine (17). L2 has been found to enhance VLP formation when co-expressed with L1 in mammalian cells (17,38). In addition, L1 expressed alone in plants has been found to form VLPs (6,54). Having sufficiently characterized the expression of L2 in plants we turned our attention in assessing whether it is possible to form full VLPs that incorporate L2 when co-expressed with L1 in plants.

N. benthamiana plants were co-infiltrated with *Agrobacterium* clones containing plasmids pTRA-hL2-C and pTRA-hL1-C. Plant material was subsequently extracted at 5 dpi in high-salt PBS which we found to be best for VLP extraction. The presence of L1 and L2 was confirmed by western blot analysis (fig. 10). An L1 and L2 signal was able to be detected in the leaf material that was co-infiltrated.

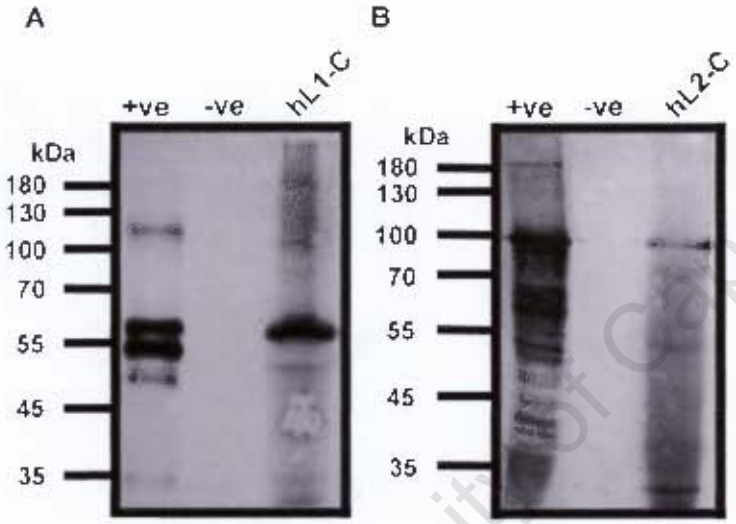

Figure 10. Co-expression of L1 and L2. Two week old *N. benthamiana* plants were co-infiltrated with *Agrobacterium* clones with plasmids pTRA-hL2-C (hL2-C) and pTRA-hL1-C (hL1-C). At 5 dpi leaf material was analyzed for both L1 (A) and L2 (B) expression using western blot analysis. L1 was detected using α-L1 mouse monoclonal antibody (J4) and α-mouse monoclonal antibody as the primary and secondary antibodies, respectively in (A). L2 was detected using α-L2 rabbit polyclonal antibody and α-rabbit monoclonal antibody as the primary and secondary antibodies, respectively in (B).Non-infiltrated plant material was extracted and used as a negative control (-ve). As positive controls (+ve) 100 ng of baculovirus-expressed L1 was loaded in gel (A) and 200 ng of *E. coli*-expressed L2 was loaded in (B).

A number of techniques were looked at to separate the more dense VLPs from the lighter L2 and L1 monomers. We theorized that if L2 was present in fractions that are found to contain the VLPs then it can be assumed that L2 is incorporated into the VLPs. Plant material was co-infiltrated with L1 and L2 expressing plasmids and as controls L1 and L2 expressing clones were also individually infiltrated separately. Plant material with L1 alone was used as a positive control to

determine the location of VLPs in fractions and plant material with L2 alone served as a negative control to determine if the presence of L2 in a VLP fraction is due to the VLP or to another factor.

VLPs were extracted using two different methods; butanol and high salt PBS. These were later pelleted at high speed (55K rpm). An L2 signal was detected, using western blot analysis, in the pellet of plant material that was co-infiltrated with L1. However, an L2-specific signal was also seen in the pellet derived from plant material infiltrated with L2 alone. This suggests that the presence of L2 in the VLP pellet may be due to another reason other than incorporation into the VLP.

An alternative method was thus attempted, namely immunoprecipitation. An antibody that recognizes L1 molecules and more importantly VLPs was provided and used for immunoprecipitation. This antibody serves as an immunoprecipitating antibody binding specifically to the VLPs. Protein A specifically binds to the F_{AB} of any antibodies thus when conjugated to agarose can be used to specifically precipitate a molecule that an antibody is bound to. The idea was to immnoprecipitate VLPs from solution and test for L2 using western blot detection. The immunoprecipitation antibody was derived from guinea pig so to not interfere with the western blot later used for detection. An L2 signal was seen in the immunoprecipitated sample of L2 and L1 expressed plant material, however, this signal was also seen in the L2 expressed alone plant material. Again, suggesting an alternative reason for the L2 being present in the VLP pellet.

Finally, sucrose gradient fractionation was also used to separate VLPs in solution from protein molecules. This form of separation is more sensitive than the others used thus by creating a density gradient so molecules of various sizes can be distinguished. In this way the less dense single molecules of L2 and L1 can be

separated from the denser VLPs. Plant material was again extracted in high salt PBS and run in a 5-50% sucrose gradient. Fractions of L2 and L1 expressing plant material were analyzed by L2 dot blot – L2 was again found to be present in every fraction. In addition, the L2 alone fractions produced a similar pattern. An L2 ELISA was also attempted using these fractions; however the same problems as previously encountered of no detection or interference was found.

Therefore these results indicate that L1 and L2 can be successfully co-expressed in the same cell in the plant. However, whether or not they form VLPs that incorporate L2 is unclear. In addition, the presence of L2 at various densities is a further problem that needs to be addressed.

2.4 Discussion

Due to the relatively new technology of transient expression an investigation into its limiting factors was required before the technique could be objectively used. Post-transcriptional gene silencing (PTGS) is a major limiting factor of transient expression in plants, with various silencing suppressors having been characterized in a number of plant viruses (90). These act in different ways and at different points within the system. For example, the NSs protein inhibits the functioning of the RNAse III-like protein from making siRNAs (75). In addition, p19 prevents association of siRNAs with the dicer nucleases (84). The results in this study show the infiltration procedure to be effective, silencing to be suppressed with a silencing suppressor, and the transient expression system to be viable. For these reasons a silencing suppressor was subsequently used in every infiltration.

The L1 gene expression was limited in mammalian cells when coded by the wildtype gene which was subsequently overcome by codon optimization (58).

Expression of L1 was similarly unable to be detected in plants at the protein level when coded by the L1 wildtype or plantized gene (7). In both this lab and others expression of L1 in plants was confirmed to dramatically increase when coded for by a mammalian-codon optimized form of the L1 gene (7).We sought to determine if the L2 gene follows a similar pattern of limited protein expression when coded by the wildtype and plantized gene and proposed that a similar increase in L2 protein expression would be seen when coded for by the equivalent mammalian-codon optimized L2 gene.

The results show a comparatively similar pattern of protein expression between L1 and L2; however, some differences were apparent. Our lab has shown marginal expression of L1 utilizing the wildtype gene (54), contrary to a previous publication (6) – such a discrepancy could be due to different methods used and different expression vectors employed. In addition, no expression of L1 protein was seen when coded by the plantized version of the gene. In the case of L2, no protein expression was detectable when coded for by the wildtype or plantized sequence of the gene. Expression was seen, however, when coded for by the humanized sequence of the gene. This prompted an enquiry at the RNA level to elucidate at what stage expression is limited.

Real time PCR (RT-PCR) is a relatively new technique and as such is prone to a number of difficulties. Due to the sensitive nature of RT-PCR, stability of the template is paramount, particularly when the same template is to be used in two or even three separate batches of reactions. For this reason we chose to convert the mRNA to cDNA which is notably more stable and easier to work with. In addition, the technique utilizes complicated empirical formulae that would confuse even an ardent biological statistician. This requires, in turn, disciplined optimization and the acceptance of a large portion of troubleshooting (53).

This, in turn, necessitates stringent parameters set by the researcher to evaluate which results they feel confident about presenting. In this study, I chose to limit the results to those that fell into the parameters I set for the standard curve and the melt curve. This consequently led to a number of experiments being discarded which may otherwise exaggerate the results presented in this study. Such an example would be where the concentration found in a separate set of results would dramatically increase the error bars either side of the mean but the calculated concentration did not lie within two points of the standards used or an extra peak was found within the melt curve. For this reason these results can only be termed preliminary as they have yet to be verified by further reactions and so be statistically validated by reproducibility.

An alternative to addressing this problem would be to use the $\Delta\Delta C_T$ method of quantification. This negates the need for a standard curve to be generated in each experiment and is relatively simpler to develop and optimize (53). It requires an initial construction of standard curves of all the genes of interest and the reference gene (2). Then a determination of the efficiencies of the reactions which takes into account that different primers and templates will have different kinetics with changes in template amounts (2). If, when plotted the efficiencies are determined to be comparable this method can be used with confidence (2). However, if they are not comparable then a normalization factor is determined to take into account the variation in efficiencies (2). However, it needs to be accepted that a certain percentage of error will still be present, but this may also be addressed by certain programs such as REST that take this error into account (40). This method may solve or negate some of the problems found in this study but due to the presence of a percentage of error it may also create some.

Ultimately, despite the fact that these results can only be termed preliminary they do nevertheless represent a trend that was present throughout the analysis.

Specifically, that the pL2 gene consistently had the lowest levels of mRNA in comparison to the other L2 genes. A possible reason for this may be linked to the plants own immune response system.

It is well known that siRNAs are used in both plants and animals as part of an immune response to limit viral infection through PTGS (89). In addition, the plant utilizes endogenous siRNAs to regulate gene expression (91). It is uncontestable that the various L2 genes are exogenous genes within the plant cell and so are susceptible to degradation through the PTGS system in much the same way as plant viral genes. However, with the addition of p19 or NSs this pathway is disrupted. But the endogenous siRNA pool may not be disrupted as the siRNAs were already present and so able to associate with the nucleases to give them site specificity prior to addition of either of these suppressors. Thus, I hypothesize that the exogenous genes are still susceptible to endogenous siRNAs if they share any sequence homology.

The process of optimizing a gene for expression in a particular system involves the changing of the codons of the gene to those that are more commonly recognized in the cells of that system. Thus, it is possible that in optimizing the L2 gene for expression in plants, in this case specifically for *Nicotiana benthamiana* has lead to an increased chance of the endogenous siRNAs having sequence homology with the gene. This theory is also supported by the findings that L1 expression is markedly decreased when coded by a plantized version of the gene in comparison to the wildtype and humanized versions (54).

An obvious flaw in the ASRP analysis is the fact that these siRNAs are from *Arabidopsis* and not tobacco. It is important to remember that the tobacco genome has yet to be sequenced and so can be thought to be evolutionarily impaired with respect to scientific knowledge in comparison to *Arabidopsis*. Thus, a comparable

type of application that contains a database of tobacco siRNAs is still a considerable time away. However, this should not prevent a certain amount of inference from the data. In addition, statistically-speaking if you consider the three different L2 genes to be random sequences they would each have as much chance of having the same number of hits. This, however, was not the case and the plantized L2 sequence was found to have a higher number of probable binding sites than the wildtype or humanized sequences.

Consequently, I believe I have managed to construct a plausible theory of why the wildtype and plantized L2 genes are not very well expressed in plants in comparison to the humanized version of the gene. I postulate that the negative RNA elements present within the wildtype L2 gene prevent expression in plants. This relies on an arguable premise that similar negative RNA element systems exist between plants and animals, of which I have found no supporting evidence in the literature. It is plausible, however, if the negative elements form structural inhibitions preventing ribosomal binding or read-through thus limiting expression. As opposed to a system where specific translational regulating factors that recognize the region and aid inhibition, in which case similar proteins that recognize the same sequence in both animals and plants is unlikely.

Thus, assuming the negative elements are recognized within the plant cell; a change in sequence should remove these elements. This is the case for both the plantized L2 (pL2) and the humanized L2. However, when coded for by the pL2 gene detectable levels of the protein are not realized. A reason, I believe, is due to the increased probability of small endogenous RNAs binding to the gene as a result of optimizing the gene to be more similar to the endogenous plant genes.

Finally, I propose that hL2 is successfully expressed as it changes the sequence of the wildtype gene to remove the negative RNA elements. In addition, it is

sufficiently different to endogenous plant genes to not be recognized by the small endogenous RNAs that would otherwise preferentially target it for degradation.

Turning then to the protein expression of L2 when coded for by the humanized gene cloned into the various pTRA vectors we noticed no differential pattern in expression. This contradicts with the parallel expression pattern of L1 when using the different pTRA vectors (54). Expression levels of L1 when coded for by the humanized version of the gene were 533 $mg.kg^{-1}$ of plant material and 379.9 $mg.kg^{-1}$ when cloned into the pTRAkc-rbcs1-cTP and pTRAc vectors, respectively (54). In addition, the expression levels of L1 when targeted to the ER were found to be dramatically lower (54).

In comparison we estimate the expression levels of L2 to be roughly 25 $mg.kg^{-1}$ irrespective of which vector was used. However, it must be acknowledged that this value is purely an estimate and has as yet to confirmed by additional, more reliable and sensitive quantifying techniques such as ELISA. These results do, nevertheless, represent a general trend that was consistently observed, namely expression of L2 does not vary across the different pTRA vectors.

A possible reason to explain the different expression patterns seen with L1 when utilizing the various pTRA vectors is that the protein is targeted to different sites within the plant cell that confer either more or less protection. Concurrently, the fact that expression of L2 does not vary with the different pTRA vectors may indicate targeting to the same site or to different sites that confer equal protection. In an attempt to answer this question we used two methods to devise a theory of where the protein is being targeted *in vivo*. These were to use *in silico* predictions which could either be confirmed or refuted through electron microscopy.

Protein targeting signals have only recently been characterized with the Nobel prize being awarded to their discoverer in 1999 (27). With the advent of high-throughput techniques localizations of many proteins have been subsequently discovered (27). By combining this data and computational power various prediction programs have become available. But, it must be noted that these programs can only serve as preliminary data with a certain expected degree of accuracy. This is evidenced in Heazlewood *et al.* which compared experimental and predicted datasets of various *Arabidopsis* genes (39). Interestingly, they found, on average, prediction programs conflicted with experimental data 50% of the time, with some achieving 70% accuracy upon prediction (39). It must be noted, however, that even amongst different experimental data there were also many contradictions as to localization (39). This highlights a degree of discretion to be used when looking at both predicted data and experimental data.

The prediction programs used indicated the presence of a mitochondrial signal within the N-terminus of the L2 protein. Interestingly, when the programs which allowed for a change in defining the organism of origin, i.e. the user can set whether they wish to use plant sequences or animal sequences, this mitochondrial signal was not recognized when animal sequences were used. This indicates a specific plant mitochondrial signal. Thus the prediction data does not conflict with previous findings in animal expression systems where L2 was not seen to be targeted to the mitochondrion. It is possible that this mitochondrial signal may conflict with other signals present in the protein attached by the other pTRA vectors, such as the LPH, KDEL and cTP. In an attempt to confirm or refute this hypothesis we turned to electron microscopy.

Electron microscopy is an inherently difficult technique to master as any technician sitting before a microtome "scooping up" microscopic ribbons can confirm. A number of trade-offs need to be assessed, such as, what degree of

distortion is permissible to obtain a general snap shot of the cell versus a cumbersome method to attain a perfect snap shot. We chose to use gluteraldehyde fixation which cross-links all proteins in the cell and mounted the samples in LR white resin necessary for our purposes of immunolabelling. Our polyclonal rabbit α-L2 antibody was used to initially label the L2 protein and monoclonal goat α-rabbit conjugated with 10 nm gold particles was used as the secondary antibody. This gold particle is important as it does not allow electrons to pass through it and is bigger than the proteins thus allowing specific visualization of the target protein. We chose to use Nickel grids coated with butivar as initial experiments found our sample unable to be preserved sufficiently with uncoated grids.

In addition, it is important to remember that an electron micrograph (EM) is a snap shot of the cell and as such any results could indicate a step within the targeting pathway as the protein is on its way to a different compartment. The snap shot analogy becomes particularly relevant when considering the number of variables implicit in immunolabelling. First, the leaf sample needs to be sufficiently preserved and embedded. Secondly, when sectioning the plane of the sample becomes relevant and affects what structures will be visible. Lastly, the section needs to be taken in a plane that allows exposure of an epitope that allows for recognition and binding by the antibody.

With this in mind EMs of plant material suggest the L2 protein alone (without targeting signals) is present in the cytoplasm. They further indicate the possibility of targeting to a membrane-sac that may be a mitochondrion. With the attachment of a chloroplast transit peptide to L2 it appears the protein is localized to the chloroplast. Unfortunately EMs to elucidate targeting of L2 when the LPH is attached with or without the KDEL sequence proved inconclusive.

Accumulatively these results allow the surmise of a hypothesis of targeting which may explain the pattern of expression. I suggest that the mitochondrial signal present in L2 is recognized within the plant cell and thus targets it to the mitochondrion. In addition, this signal interferes with the secretory signal (LPH) and thus the protein does not follow the secretory pathway instead also localizing to the mitochondrion. Thus, the protein may not be targeted to the ER and so the KDEL sequence cannot act to retain it therein. The cTP signal, however, is strong enough to overcome the mitochondrial targeting and so L2 accumulates in the chloroplast when attached. I surmise that similar expression levels are seen as both the mitochondrion and chloroplast confer similar protection to allow accumulation of the protein.

Finally, turning to co-expression of L1 and L2 in this system necessitates an initial exploration of co-expression in mammalian systems which have shown the ability of producing VLPs that incorporate L2 (37). Interestingly, VLPs produced from co-expression of L2 were found to be less variable in size, leading to speculation that L2 enhances capsid formation (37) (as mentioned in Chapter 1). In this lab and others, VLPs have been shown to form when L1 is expressed alone in plants (7,54). We hypothesized that L1 co-expressed with L2 in plants would lead to VLPs incorporating L2 and a similar enhancement of VLP formation as seen in mammalian cells.

Our results indicate that L2 and L1 can successfully be co-expressed in the same plant. It is probable that both proteins are able to be present in the same plant cell based on corollary findings that antibodies with separate constituent parts can be formed within the same cell (33,49). However, whether L1 and L2 are able to form full VLPs that incorporate L2 is unclear. We consistently found L2 alone to sediment with the higher density VLPs and present at various densities within the sucrose gradient fractions. A possible explanation for this may be that L2 forms

aggregates in the high salt buffer used; a problem not previously encountered due to the use of 8M urea in prior isolation of L2. This same problem was encountered within this lab when L2 was expressed in *E. coli*; hence the use of 8M urea in all experiments.

Butanol extraction relies on the separation of all aqueous molecules from non-polar substances such as polyphenols. In theory, the VLPs and L2 protein should be found in the aqueous phase and the commonly interfering substances of plant material to be separated into the non-polar phase. Unfortunately, this extraction method did not seem to improve our results.

These results indicate a good starting point from which to begin to investigate the feasibility of a plant-produced HPV L1/L2 vaccine. It is clear that alternative buffers should be assessed that solubilize and preserve the VLPs and in addition that solubilize L2. Whether butanol extraction should be subsequently used remains unsure, however, if no dramatic loss in protein is detected it could be an invaluable method of removing plant material that interferes with downstream applications.

2.5 Conclusion

I found it is possible to express L2 in plants, with expression being best when coded for by the mammalian codon optimized sequence of the gene. In addition, they indicate that high levels of co-expression of L1 and L2 are possible through transient expression making the likelihood of VLPs that incorporate L2 highly probable.

Further investigations should target, first, the quantitation of L2 protein expression within this system which necessitates improved extraction and solubilization of undenatured L2. This will have a knock on effect on the L2 ELISA in possibly allowing detection of small quantities without interference of plant material. In addition, another monoclonal or polyclonal α-L2 antibody should be sourced which would allow the use of the capture ELISA method known to be more sensitive and remove the problem of interfering plant products.

Secondly, additional RT PCR experiments should be completed to confirm the results represented in this study. This may include a different or even concurrent approach to RNA quantitation using a different calculation method, such as the $\Delta\Delta C_T$ method.

Thirdly, additional electron micrographs should be completed to determine localization of the ER-targeted and apoplastic-targeted L2. Alternative methods may also be utilized such as immunofluorescence which uses a secondary antibody attached to a fluorophore. However, due to the use of spectrometry as opposed to electron microscopy less detail will be visible, but it may allow determination of localization of the apoplastic-targeted protein between the cell walls.

Finally, extraction and separation methods of VLPs needs to be refined to remove the false positive seen in the L2 alone expressed material. This would allow confirmation of L1/L2 VLPs. This may include a series of different buffers that have various pHs, different salt concentrations and/or possible additional agents that may aid solubility, such as various ionic and non-ionic detergents.

The immunogenic properties of L2 and stabilizing properties with respect to VLPs make L2 an attractive recombinantly expressed vaccine or vaccine enhancer, respectively or in addition. Thus, these results in combination with the cheaper

expression system of plant production represent the initial steps needed in realizing a cheaper HPV vaccine for developing countries.

3. REFERENCES

1. National Immunization Program: HPV Vaccines, FAQs. The CDC Website . 2007.
Ref Type: Electronic Citation

2. Critical Factors for Successful Real-Time PCR. 2004. Qiagen.
Ref Type: Catalog

3. Alphs, H. International Papillomavirus Conference. 2007.
Ref Type: Conference Proceeding

4. **Badal, S., V. Badal, I. E. Calleja-Macias, M. Kalantari, L. S. H. Chuang, B. F. L. Li, and H.-M. Bernard**. 2004. The Human Papillomavirus-18 genome is efficiently targeted by the cellular DNA Methylation. Virology **324**:483-492.

5. **Becker, K. A., L. Florin, C. Sapp, G. G. Maul, and M. Sapp**. 2004. Nuclear Localization but not PML Protein is Required for Incorporation of the Papillomavirus Minor Capsid Protein L2 into the Virus-like Particles. Virology **78**:1121-1128.

6. **Biemelt, S., P. Sonnemald, P. Galmbacher, L. Willmitzer, and M. Müller**. 2003. Production of Human Papillomavirus Type 16 Virus-like particles in transgenic plants. Journal of Virology **77**:9211-9220.

7. **Biemelt, S., U. Sonnemald, P. Galmbacher, L. Willmitzer, and M. Müller**. 2003. Production of Human Papillomavirus type 16 Virus-like Particles in Transgenic Plants. Journal of Virology **77**:9211-9220.

8. **Bodily, J. and C. Meyers**. 2005. Genetic Analysis of the Human Papillomavirus Type 31 Diffrentiation-Dependent Late Promoter. Journal of Virology **79**:3309-3321.

9. **Bossis, I., R. B. S. Roden, R. Gambhira, R. Yang, M. Tagaya, P. M. Howley, and P. I. Meneses**. 2005. Interaction of tSNARE Syntaxin 18 with the Papillomavirus Minor Capsid Protein Mediates Infection. Journal of Virology **79**:6723-6731.

10. **Bousarghin, L., A. Touzé, A. Combita-Rojas, and P. Coursaget**. 2003. Positively charged sequence of human papillomavirus type 16 capsid proteins are sufficient to mediate gene transfer into target cells via the heparan sulphate receptor. Journal of General Virology **84**:157-164.

11. **Brodzik, R., M. Glogowska, K. Bandurska, M. Okulicz, D. Deka, K. Ko, J. van der Linden, J. H. Leusen, N. Pogrebnyak, M. Golovkin, Z. Steplewski, and H. Koprowski**. 2006. Plant-derived anti-Lewis Y mAb Exhibits Biological Activities for Efficient Immunotherapy against Human Cancer Cells. Procedures in Natural and Applied Sciences **103**:8804-8809.

12. **Brown, D. R., K. H. Fife, C. M. Wheeler, L. A. Koutsky, L. Lupinacci, R. Railkar, G. Suhr, E. Barr, A. Dicello, W. Li, J. F. Smith, A. Tadesse, and K. U. Jansen**. 2004. Early Assessment of the Efficacy of a Human Papillomavirus Type 16 L1 Virus-like Particle Vaccine. Vaccine **22**:2936-2942.

13. Buck, C. B., N. Cheng, C. D. Thompson, D. Lowy, A. C. Steven, J. T. Schiller, and B. L. Trus. Arrangement of L2 within the Papillomavirus Capsid. International Papillomavirus Conference. 2006.
Ref Type: Conference Proceeding

14. **Buck, C. B., C. D. Thompson, Y.-Y. S. Pang, D. Lowy, and J. T. Schiller**. 2005. Maturation of Papillomavirus Capsids. Journal of Virology **79**:2839-2846.

15. Campo, M. S., G. H. Ashrafi, B. Marchetti, M. Haghshenas, and E. Arabi. Papillomavirus E5 Proteins and MHC Class I: A Common way to Evade the Immune Response. International Papillomavirus Conference. 2006.
Ref Type: Conference Proceeding

16. **Carter, J., G. C. Wipf, M. M. Madeleine, S. Schwartz, L. A. Koutsky, and D. A. Galloway**. 2006. Identification of Human Papillomavirus Type 16 L1 Surface Loops Required for Neutralization by Human Sera. Journal of Virology **80**:4664-4672.

17. **Casini, G., D. Graham, D. Heine, R. L. Garcea, and D. T. Wu**. 2004. In vitro Papillomavirus Capsid Assembly Analyzed by Light Scattering. Virology **325**:320-327.

18. **Casini, G., D. Graham, D. Heine, R. L. Garcea, and D. T. Wu**. 2004. In vitro Papillomavirus Capsid Assembly Analyzed by Light Scattering. Virology **325**:320-327.

19. **Chen, X. S., R. L. Garcea, I. Goldberg, G. Casini, and S. C. Harrison**. 2000. Structure of Small Virus-like Particles Assembled from the L1 Protein of Human Papillomavirus 16. Molecular Cell **5**:557-567.

20. **Cheung, Y.-K., S. C.-S. Cheng, F. W.-Y. Sin, and Y. Xie**. 2004. Plasmid Encoding Papillomavirus Type 16 (HPV 16) DNA Constructed with Codon Optimization Improved the Immunogenicity against HPV Infection. Vaccine **23**:629-638.

21. **Christensen, N. D., J. Dillner, C. Eklund, J. Carter, G. C. Wipf, C. A. Reed, N. M. Cladel, and D. A. Galloway**. 1996. Surface Conformational and Linear Epitopes on the HPV-16 and HPV-18 L1 Virus-like Particles as Defined by Monclonal Antibodies. Virology **223**:174-184.

22. Conway and C. Meyers. International Papillomavirus Conference. 2007.
Ref Type: Conference Proceeding

23. **Daniell, H., S. J. Streatfield, and K. Wycoff**. 2001. Medical Molecular Farming: Production of Antibodies, Biopharmaceuticals and Edible Vaccines in Plants. Trends in Plant Sciences **6**:219-226.

24. Day, P. M. International Papillomavirus Conference. 2007.
Ref Type: Conference Proceeding

25. Day, P. M., C. D. Thompson, R. Gambhira, R. B. S. Roden, D. Lowy, and J. T. Schiller. Insights into the Process of HPV-16 Infection Revealed by the Discovery of Multiple Distinct Mechanisms of Antibody-mediated Virus Neutralization. International Papillomavirus Conference. 2006.
Ref Type: Conference Proceeding

26. **de Villiers, E.-M., C. Fauquet, T. R. Broker, H.-M. Bernard, and H. zur Hausen**. 2004. Classification of Papillomaviruses. Virology **324**:17-27.

27. **Emanuelsson, O., S. Brunak, G. von Heijne, and H. Nielsen**. 2007. Locating proteins in the cell using TargetP, SignalP and related tools. Nature Protocols **2**:953-971.

28. **Finnen, R. L., K. D. Erikson, X. S. Chen, and R. L. Garcea**. 2003. Interactions between Papillomavirus L1 and L2 Capsid Proteins. Journal Virology **77**:4818-4826.

29. **Florin, L., K. A. Becker, C. Lambert, T. Nowak, C. Sapp, D. Strand, R. E. Streek, and M. Sapp**. 2006. Identification of a Dynein Interacting Domain in the Papillomavirus Minor Capsid Protein L2. Journal of Virology **80**:6691-6696.

30. **Florin, L., K. Schafer, K. Sotlar, R. E. Streek, and M. Sapp**. 2002. Reorganization of Nuclear Domain 10 Induced by the Papillomavirus Capsid Protein L2. Virology **295**:97-107.

31. Friedman, A. L. and H. Shepeard. Applying Lessons Learned from HPV Audience Research to HPV Vaccine Education Efforts. International Papillomavirus Conference. Vaccines II: Preventative Vaccines - Models . 2006. 2006.
Ref Type: Conference Proceeding

32. Gambhira, R., C. B. Buck, I. Bossis, H. Alphs, T. D. Culp, N. D. Christensen, and R. B. S. Roden. Cross Protection of Rabbits Against Challenge with Rabbit Papillomaviruses by Immunization with the N-terminus of HPV 16 Minor Capsid Antigen L2. International Papillomavirus Conference. 2006.
Ref Type: Conference Proceeding

33. **Giritch, A., S. Marillonnet, C. Engler, G. van Eldik, J. Botterman, V. Klimyuk, and Y. Gleba**. 2006. Rapid High-yield Expression of Full Size IgG Antibodies in Plants Coinfected with Noncompeting Viral Vectors. Procedures in Natural and Applied Sciences1-6.

34. **Görnemann, J., T. G. Hofmann, T. Will, and M. Müller**. 2002. Interaction of Human Papillomavirus Type 16 L2 with Cellular Proteins: Identification of Novel Nuclear Body-associated Proteins. Virology **303**:69-78.

35. **Gu, W. Y., M. Li, W. M. Zhao, N. X. Fang, S. Bu, I. H. Frazer, and K.-N. Zhao**. 2004. $tRNA^{Ser}$(CGA) differentially regulates expression of wild-type and codon-modified papillomavirus L1 genes. Nucleic Acids Research **32**:4448-4461.

36. **Gustafson, A. M., E. Allen, S. Givan, D. Smith, J. C. Carrington, and K. D. Kasschau**. 2005. ASRP: the *Arabidopsis* Small RNA Project Database. Nucleic Acids Research **33**:D637-D640.

37. **Hagensee, M. E., N. Yaegashi, and D. A. Galloway**. 1993. Self-Assembly of Human Papillomavirus type 1 Capsids by Expression of the L1 Protein Alone or by Coexpression of L1 and L2 Capsid Proteins. Journal of Virology **67**:315-322.

38. **Hagensee, M. E., N. Yaegashi, and D. A. Galloway**. 1993. Self-Assembly of Human Papillomavirus type 1 Capsids by Expression of the L1 Protein Alone or by Coexpression of L1 and L2 Capsid Proteins. Journal of Virology **67**:315-322.

39. **Heazlewood, J. L., J. Tonti-Filippini, R. E. Verboom, and A. H. Millar**. 2005. Combining experimental and predicted datasets for determination of the subcellular location of proteins in *Arabidopsis*. Plant Physiology **139**:598-609.

40. Herrmann, M. and M. Pfaffl. REST 2005: Greater certaint in expression studies. 2005. Corbett Research Pty Ltd.
Ref Type: Report

41. **Hilleman, M. R.** 2000. Overview of Vaccinology with Special Reference to Papillomavirus Vaccines. Journal of Clinical Virology **19**:79-90.

42. **Holmgren, S. C., N. A. Patterson, M. A. Ozbun, and P. F. Lambert**. 2005. The Minor Capsid Protein L2 Contributes to two Steps in the Human Papillomavirus Type 31 Life Cycle. Journal of Virology **79**:3938-3948.

43. **Kalish, R.** 1995. Antigen Processing: the Gateway to the Immune Response. Journal of American Academy of Dermatology **42**:640-652.

44. **Kämper, N., P. M. Day, T. Nowak, H.-C. Selinka, L. Florin, J. Bolscher, L. Hilbig, J. T. Schiller, and M. Sapp**. 2006. A Membrane-Destabilizing Peptide in Capsid Protein L2 is Required for Egress of Papillomavirus Genomes from Endosomes. Journal of Virology **80**:759-768.

45. **Kawana, K., K. Kawana, H. Yoshikawa, Y. Taketani, T. Kawana, K. Yoshiike, and T. Kanda**. 1998. A surface Immunodeterminant of Human Papillomavirus Type 16 Minor Capsid Protein L2. Virology **245**:359.

46. **Kawana, K., H. Yoshikawa, Y. Taketani, K. Yoshiike, and T. Kanda**. 1999. Common Neutralization Epitope in Minor Capsid Protein L2 of Human Papillomavirus Types 16 and 6. Journal of Virology **73**:6188-6190.

47. **Khan, M. S.** 2006. Hybrid Transcription-mediated Transgene Regulation in Plastids. Trends in Biotechnology **24**:479-482.

48. **Kloetzel, P.-M. and F. Ossendorp**. 2004. Proteasome and Peptidase Function in MHC-Class-I-Mediated Antigen Presentation. Current Oppinion in Immunology **16**:76-81.

49. **Ko, K., Z. Steplewski, M. Glogowska, and H. Koprowski**. 2005. Inhibition of Tumour Growth by Plant-derived mAb. Procedures in Natural and Applied Sciences **102**:7026-7030.

50. **Koncz, C. and J. Schell**. 1986. The Promoter of T_L-DNA gene *5* controls the tissue-specific expression of chimaeric genes carried by a novel type of *Agrobacterium* binary vector. Mol Gen Genet **204**:383-396.

51. **Kondo, K., Y. Ishii, H. Ochi, T. Matsumoto, H. Yoshikawa, and T. Kanda**. 2007. Neutralization of HPV 16, 18, 31, and 58 Pseudovirions with Antisera induced by immunizing Rabbits with Synthetic Peptides Representing Segments of the HPV 16 Minor Capsid Protein L2 Surface Region. Virology **358**:266-272.

52. **Koya, V., M. Moayeri, S. H. Leppla, and H. Daniell**. 2005. Plant-based Vaccine: Mice Immunized with Chloroplast-derived Anthrax Protective Antigen Survive Anthrax Lethal Toxin Challenge. Infection and Immunity **73**:8266-8274.

53. **Mackay, I. M., K. E. Arden, and A. Nitsche**. 2002. Survey and Summary: Real-time PCR in Virology. Nucleic Acids Research **30**:1292-1305.

54. **Maclean, J., M. Koekemoer, A. J. Olivier, D. Stewart, I. I. Hitzeroth, T. Rademacher, R. Fischer, A.-L. Williamson, and E. P. Rybicki**. 2007. Optimization of human papillomavirus type 16 (HPV-16) L1 expression in plants: comparison of the suitability of different HPV-16 L1 gene variants and different cell-compartment localization. Journal of General Virology **88**:1460-1469.

55. **Mahdavi, A. and B. J. Monk**. 2005. Vaccines Against Human Papillomavirus and Cervical Cancer: Promises and Challenges. The Oncologist **10**:528-538.

56. **Modis, Y., B. L. Trus, and S. C. Harrison**. 2002. Atomic Model of the Papillomavirus Capsid. The European Molecular Biology Organisation Journal **21**:4754-4762.

57. **Modis, Y., B. L. Trus, and S. C. Harrison**. 2002. Atomic Model of the Papillomavirus Capsid. The European Molecular Biology Organisation Journal **21**:4754-4762.

58. **Mossadegh, N., L. Gissman, M. Müller, H. Zentgraf, A. Alonso, and P. Tomakidi**. 2004. Codon Optimization of the Human Papillomavirus 11 (HPV 11) L1 Gene Leads to Increased Gene Expression and Formation of Virus-like Particles in Mammalian Epithelial Cells. Virology **326**:57-66.

59. **Nowak, W., M. Gawlowska, A. Jarmolowski, and J. Augustyniak**. 2001. Effect of nuclear attachment regions on transgene expression in tobacco plants. Acta Biochim Pol **48**:637-646.

60. **Õberg, D., B. Collier, X. Zhao, and S. Schwartz**. 2003. Mutational Inactivation of Two Distinct Negative RNA elements in the Human Papillomavirus Type 16 L2 Coding Region Induces Production of High Levels of L2 in Human Cells. Journal of Virology **77**:11674-11684.

61. **Õberg, D., J. Fay, H. Lambkin, and S. Schwartz**. 2005. A Downstream Polyadenylation Element in Human Papillomavirus Type 16 L2 encodes Multiple GGG Motifs and Interacts with hnRNP H. Journal of Virology **79**:9254-9269.

62. Olcese, V. A., Y. Chen, R. Schlegel, and H. Yuan. Characterization of HPV 16 L1 Loop Domains in the Formation of a Type-specific, Conformational Epitope. BioMed Central Microbiology . 2004.
Ref Type: Internet Communication

63. **Pastrana, D. V., R. Gambhira, C. B. Buck, Y.-Y. S. Pang, C. D. Thompson, T. D. Culp, N. D. Christensen, D. Lowy, J. T. Schiller, and R. B. S. Roden**. 2005. Cross-Neutralization of Cutaneous and Mucosal Papillomavirus types with Anti-Sera to the Amino Terminus of L2. Virology **337**:365-372.

64. **Quint, W. G. V., S. R. Pagliusi, N. Lelie, E.-M. de Villiers, C. M. Wheeler, and World Health Organization Human Papillomavirus DNA International**

Collaborative Study Group. 2006. Results of the First World Health Organization International Collaborative Study of Detection of Human Ppaillomavirus DNA. Journal of Clinical Microbiology **44**:571-579.

65. **Richards, R. M., D. Lowy, J. T. Schiller, and P. M. Day**. 2006. Cleavage of the Papillomavirus Minor Capsid Protein, L2, at a Furin Consensus Site is Necessary for Infection. Procedures in National and Academic Sciences **103**:1522-1527.

66. **Roden, R. B. S., W. Yuty, R. Fallon, S. Inglis, D. Lowy, and J. T. Schiller**. 2000. Minor Capsid Protein of Human Genital Papillomavirus Contains Subdominant, Cross-Neutralizing Epitopes. Virology **270**:254-257.

67. **Rukatsova, E., O. Zolova, N. Buryanova, V. Borisova, V. Bykova, and Y. Burynova**. 2003. Russ J Genet **39**:41-45.

68. **Rush, M., X. Zhao, and S. Schwartz**. 2005. A Splicing Enhancer in the E4 Coding Region of Human Papillomavirus Type 16 is Required for Early mRNA Splicing and Polyadenylation as well as Inhibition of Premature Late Gene Expression. Journal of Virology **79**:12002-12015.

69. **Sapp, M., C. Fligge, J. Petzak, J. Harris, and R. E. Streek**. 1998. Papillomavirus Assembly Requires Trimerization of the Major Capsid Protein by Disulfides between Two Highly Conserved Cysteines. Journal of Virology **72**:6186-6189.

70. Schiller, J. T. International Papillomavirus Conference. 2007.
Ref Type: Conference Proceeding

71. **Sette, A. and J. Fikes**. 2003. Epitope-based Vaccines: an Update on Epitope Identification, Vaccine Design and Delivery. Current Oppinion in Immunology **15**:470.

72. **Shah, K. V. and P. M. Howley**. 1996. Papillomaviruses, p. 2077-2102. *In* B. N. Fields, D. M. Knipe, and P. M. Howley (ed.), Fields Virology. Lippincott - Raven Publishers, Philadelphia.

73. Slupetzky, K. Chimeric VLPs that Display HPV 16 L2 Peptides Induce Cross-neutralizing Ab to HPV 11. International Papillomavirus Conference. 2005.
Ref Type: Conference Proceeding

74. **Stanley, M.** 2006. Immune Responses to Human Papillomavirus. Vaccine **24S1**:S1-16-S1/22.

75. **Takeda, A., K. Sugiyama, H. Nagano, M. Mori, M. Kaido, K. Mise, S. Tsuda, and T. Okuno**. 2002. Identification of a Novel RNA Silencing Suppressor, NSs Protein of *Tomato spotted wilt virus*. FEBS **532**:75-79.

76. **Takeda, A., K. Sugiyama, H. Nagano, M. Mori, M. Kaido, K. Mise, S. Tsuda, and T. Okuno**. 2002. Identification of a Novel RNA Silencing Suppressor, NSs Protein of *Tomato spotted wilt virus*. Federation of European Biochemical Societies **532**:75-79.

77. **Thanavala, Y., M. Mahoney, S. Pal, A. Scott, L. Richter, N. Natarajan, P. Goodwin, C. J. Arntzen, and H. S. Mason**. 2005. Immunogenicity in Humans of an Edible Vaccine for Hepatitis B. Procedures in Natural and Applied Sciences **102**:3378-3382.

78. **Trottier, H. and E. L. Franco**. 2006. The Epidemiology of Genital Human Papillomavirus Infection. Vaccine **24S1**:S1-4-S1/15.

79. **Turan, T., M. Kalantari, I. E. Calleja-Macias, H. A. Cubie, K. Cuschieri, L. L. Villa, H. Skomedal, H. A. Barrera-Saldaña, and H.-M. Bernard**. 2006. Methylation of the Human Papillomavirus-18 L1 Gene: A Biomarker of Neoplastic Progression. Virology **349**:175-183.

80. **Turan, T., M. Kalantari, I. E. Calleja-Macias, H. A. Cubie, K. Cuschieri, L. L. Villa, H. Skomedal, H. A. Barrera-Saldaña, and H.-M. Bernard**. 2006. Methylation of the Human Papillomavirus-18 L1 Gene: A Biomarker of Neoplastic Progression. Virology **349**:175-183.

81. **van der Most, R. B., K. Murali-Krishna, J. G. Lanier, E. J. Wherry, M. T. Puglielli, J. N. Blattman, A. Sette, and R. Ahmed**. 2003. Changing Immunodominance Patterns in Antiviral CD8 T-Cell Responses after loss of Epitope Presentation or Chronic Antigenic Stimulation. Virology **315**:93-102.

82. **Villa, L. L.** 2006. Prophylactic HPV Vaccines: Reducing the Burden of HPV-related Diseases. Vaccine **24S1**:S1-23-S1/28.

83. **Villa, L. L., K. A. Ault, A. R. Giuliano, R. L. R. Costa, D. R. Brown, A. Ferenczy, D. M. Harper, L. A. Koutsky, R. J. Kurman, M. Lehtinen, C. Malm, S.-E. Olsson, B. M. Ronnett, F. E. Skjeldestad, M. Steinwall, M. H. Stoler, C. M. Wheeler, F. J. Taddeo, J. Yu, L. Lupinacci, R. Railkar, R. Marchese, M. T. Esser, J. Bryan, K. U. Jansen, H. L. Sings, G. M. Tamus, A. J. Saah, and E. Barr**. 2006. Immunologic Responses Following Administration of a Vaccine targeting Human Papillomavirus Types 6, 11, 16, and 18. Vaccine **24**:5571-5583.

84. **Voinnet, O., S. Rivas, P. Mestre, and D. Baulcombe**. 2003. An Enhanced Expression System in Plants Based on Suppression of Gene Silencing by the p19 Protein of *Tomato bushy stunt virus*. Plant Journal **33**:949-956.

85. **Voinnet, O., S. Rivas, P. Mestre, and D. Baulcombe**. 2003. An Enhanced Transient Expression System in Plants Based on Suppression of Gene Silencing by the p19 Protein of *Tomato bushy stunt virus*. The Plant Journal **33**:949-956.

86. **Volpers, C., M. Sapp, P. J. F. Snijders, J. M. M. Walboomers, and R. E. Streek**. 1995. Conformational and Linear Epitopes on Virus-like Particles of Human Papillomavirus type 33 identified by Monoclonal Antibodies to the Minor Capsid Protein L2. Journal of General Virology **76**:2661-2667.

87. **Wen, S. X., L. D. Teel, N. A. Judge, and A. D. O'Brien**. 2006. A Plant-based Oral Vaccine to Protect against Systemic Intoxication by Shiga Toxin Type 2. Procedures in Natural and Applied Sciences **103**:7082-7087.

88. **Wigdorovitz, A., M. Mozgovoj, M. J. Dus Santos, V. Parreño, C. Gómez, D. M. Pérez-Filgueira, K. G. Trono, R. D. Ríos, P. M. Franzone, F. Fernández, C. Carrilo, L. A. Babiuk, J. M. Escribano, and M. V. Borca**. 2004. Protective Lactogenic Immunity Conferred by an Edible Peptide Vaccine to Bovine Rotavirus Produced in Transgenic Plants. Journal of General Virology **85**:1825-1832.

89. **Xie, Q. and H.-S. Guo**. 2006. Systemic antiviral silencing in plants. Virus Research **118**:1-6.

90. **Xie, Q. and H.-S. Guo**. 2006. Systemic Antiviral Silencing in Plants. Virus Reserch **118**:1-6.

91. **Xie, Z., L. K. Johansen, A. M. Gustafson, K. D. Kasschau, A. D. Lellia, D. Zilberman, S. E. Jacobsen, and J. C. Carrington**. 2004. Genetic and functional diversification of small RNA pathways in plants. PLoS Biology **2**:e104.

92. **Yang, R., P. M. Day, W. H. Yutzy IV, K. Lin, C.-F. Hung, and B. S. Roden**. 2003. Cell Surface-binding Motifs of L2 that Facilitate Papillomavirus Infection. Journal of Virology **77**:3531-3541.

93. **Zhao, K.-N., W. Y. Gu, N. X. Fang, N. A. Saunders, and I. H. Frazer**. 2005. Gene Codon Composition Determines Differentiation-Dependent Expression of a Viral Capsid Gene in Keratinocytes in vitro and in vivo. Molecular and Cellular Biology **25**:8643-8655.

94. **Zhao, K.-N., W. J. Liu, and I. H. Frazer**. 2003. Codon Usage Bias and A + T Content Variation in Human Papillomavirus Genomes. Virus Reserch **98**:95-104.

95. **Zupan, J., T. Muth, O. Draper, and P. Zambryski**. 2000. The Transfer of DNA from *Agrobacterium tumefaciens* into Plants: a Feast of Fundamental Insights. The Plant Journal **23**:11-28.