THE MOLECULAR CHARACTERISATION OF NARCISSUS LATENT VIRUS AND MACLURA MOSAIC VIRUS

Thesis submitted for the degree of Doctor of Philosophy at the University of Leicester

by

Joanne Louise Badge BA hons (Oxford)

Department of Botany

University of Leicester

September 1997

Page numbering as original

This thesis is dedicated to

Dr Richard Badge, my inspiration

Joanne Louise BADGE

The molecular characterisation of narcissus latent virus and Maclura mosaic virus

Abstract

Narcissus latent virus (NLV) and Maclura mosaic virus (MacMV) are serologically related. However, they have poor serological relationships with other plant viruses with which they have shared characteristics. Coat protein size, particle shape and structure, mode of vector transmission, cytology and serology proved insufficient to classify them. Molecular techniques were employed in order to create tools for the rapid and accurate classification of plant viruses. A carlavirus-specific PCR primer test failed to amplify NLV or MacMV but confirmed that several other viruses belonged to the carlavirus genus.

The nucleotide sequences of part of the nuclear inclusion body (NIb) gene, the complete coat protein gene and the 3' untranslated regions of narcissus latent virus (NLV) and Maclura mosaic virus (MacMV) were determined. Deduced amino acid sequences for the NIb and coat protein genes revealed that NLV and MacMV are closely related. Comparison of the NIb sequences with other viruses showed that NLV and MacMV have closer affinities with viruses of the *Potyviridae* than to those of the carlavirus genus with which they were initially classified. It is proposed that NLV and MacMV may form a new genus within the *Potyviridae*, the *Macluraviruses*. The viruses associated with narcissus yellow stripe disease were re-evaluated.

In order to identify further members for the new genus a second PCR primer was designed to amplify a region shared by the bymoviruses and macluraviruses. Sequence data obtained for the 3'-terminal region of rice necrosis mosaic virus (RNMV) using a fragment amplified by this primer confirmed that RNMV was a member of the bymovirus genus. MacMV and barley mild mosaic bymovirus replicase sequences were used to transform tobacco plants in an attempt to create transgenic resistance.

Acknowledgements

I am indebted Prof. Alan Brunt (HRI) and Dr David Robinson (SCRI) who kindly supplied infected material and antisera and I have greatly appreciated their criticism of my written work. My thanks go to the following for their kind gifts and hard work: Dr Sue Seal (NRI) and staff carried out similar RT-PCR tests on carlavirus-infected material. Dr Carol Brattey (SASA) supplied the red la soda virus and particle measurements. Dr Satoshi Kashiwazaki (National Agriculture Research Center, Japan) carried out RT-PCR tests on their stocks of bymovirus-infected material and supplied amplified RNMV fragment. I acknowledge the support of BBRSC in funding my studentship.

The fact that Gary Foster managed to continue to help me even after he moved to take up a post at Bristol University demonstrates his commitment to this project and to his student. I have greatly appreciated his sustained support and cheerful enthusiasm throughout the four years we have worked together.

Without the rest of the virus group I would have autoclaved many more bottles of acrylamide than is necessary for any scientific career, however short. They were great fun to work with and kept my spirits (alcoholic and otherwise) up through the endless sequencing. Pointy toe dancing and the Tom Jones door would never have brightened our lab without Sally, Emma, Roisin or Tim. My heartfelt thanks go to everyone in the department who ever got dragged to the pub 'just for one' because it was a bad sequencing day but especially Rachel Hodge (whose sympathetic alcohol consumption was usually beyond the call of duty) and Diane Hird (who is responsible for the inspiration of my veggie plot). I can only hope that someone will keep up the betting syndicates in my absence.

My parents have always encouraged me to follow my dreams and do my best. This is the best I have ever done and I hope they realise that it is as much their work as mine.

Finally, I am certain that without Richard's placid temperament, love and constant support I would have actually buried this thesis in the garden. Just saying thank you does not convey my gratitude to this incredible man. I can only hope that the rest of our lives is time enough for me to repay the debt I owe for his unerring encouragement and patience.

Now, I have to agree with Sally's final thesis comments - anyone for a nice gin and tonic?

Chapter 1

Introduction

This thesis is about the classification of two plant viruses, Maclura mosaic virus (MacMV) and narcissus latent virus (NLV). At present, neither can be classified satisfactorily since their biological and cytological characteristics seem to be anomalous when considered alongside their particle morphology. Traditionally, viruses have been classified by groupings based on structural features and biological characteristics. For example, coat protein molecular weight and virus particle size and type combined with mode of vector transmission and cytology have been used to delineate a large number of plant virus groups. However, in the case of NLV and MacMV, these characters when defined proved insufficient to classify them. The limit of the present knowledge about these two viruses is typical of many unassigned viruses. The aim of this thesis is to employ modern molecular level techniques as rapid tools to assign NLV and MacMV to a taxon.

NLV and MacMV appear to possess the characteristics normally attributed to two distinct genera. This raises the possibility that they represent an evolutionary link between the two genera. Even if they do not represent such a link, they are still worthy of investigation. Information about any virus can have repercussions on our understanding of other viruses. The problem of classifying viruses satisfactorily is one that has existed since they were first recognised and the challenge of classification can enlighten our views of other taxa.

Why classify viruses?

Classification provides a consistent naming system. This is a practical necessity, especially where disease is concerned. Viruses are pathogens and as such there is a practical need to identify them for diagnosis, in medicine, agriculture and horticulture. A nomenclature system aids identification by providing a concise summary of characteristics. A common universal nomenclature is essential for effective communication and discussion amongst the scientific community. Although there has been agreement for over half a century that a system of taxonomy and nomenclature for viruses is needed, there is still argument over how it should be used and defined.

The taxonomy of plant viruses

The theory

Taxonomy is used to provide a means of identification of organisms. A universal naming system, Linnean binomials, is used by the majority of biologists easily and without controversy (Ridley, 1993). Species can be practically defined in terms of easily identifiable morphologic characteristics. There are two methods of classification, phenotypic and phylogenetic. Phenotypic classification groups organisms together according to shared attributes. If organisms are grouped on one characteristic alone, then the use of a different characteristic may lead to a different grouping. Usually phenotypic classifiers use a large number of physical attributes to define a group, thereby diluting out those attributes which could lead to a radically different solution. Phylogenetic classification is an evolutionary method, which analyses how recently organisms have shared a common ancestor. Usually, for most biological organisms, the two methods give the same outcome, however there are important ideological arguments which distinguish the two methods that affects the classification of viruses.

There are three main schools of classification: numerical taxonomy (phenotypic), cladistic taxonomy (phylogenetic) and evolutionary taxonomy (both phenotypic and phylogenetic). Numerical taxonomy proposed by Sneat & Sokal (see Ridley, 1993 for review) is subjective in its application and can result in many different groupings of the same organisms each of which are as valid as the next with no means of distinguishing between them. Phenotypic classification is ambiguous because there is more than one way of measuring phenotypic similarity and the different measurements can disagree. Hennig was the main proponent of cladistic analysis (see Ridley, 1993 for review) which uses evolutionary theory to group organisms according to their most recent common ancestor. This method fits well with Linnean binomials. It is not always a facile method since the distinction between shared derived characteristics and ancestral characteristics requires the presence or knowledge of the ancestor. The phylogenetic relationship must often be inferred and is therefore not always consistently applied. The majority of biologists use evolutionary taxonomy (main proponents are Mayr, Simpson and Dobhanzsky: see Ridley, 1993 for review) since it agrees with most existing taxonomy and is easier to apply than cladistic analysis. However, it is not theoretically rigorous and suffers from the same problem with phenotypic relationships that numerical taxonomy does.

Many virologists have considered only one of these schools of thought, evolutionary taxonomy as a theory by which viruses should be analysed. In order to apply this

theory, they have looked to the definition of a species used by evolutionary taxonomists, the biological species concept:

> "species are groups of interbreeding natural populations that are reproductively isolated from other such groups" (Ridley, 1993)

The biological species concept is clearly problematic for the virologist. The notion of 'interbreeding' in an asexual entity, such as a virus is non-sensical. The use of the term 'natural population' is difficult to apply to an obligate parasite when individuals exist in separate hosts. Virus particles may exist for many years in a vector (e.g.: virus transmitted by fungal spores) in isolation. However, this form of isolation is difficult to compare to the reproductive isolation described above. An entire field of barley may be infected with the same viral strain, but the 'populations' within each individual plant are isolated from one another. If viruses are transmitted through asexually reproducing hosts (e.g.: bulb division) then the isolation may be limited to a single host for many years. Genetic drift or adaptation to the individual host may result in sufficient variations between the viruses present in individual bulbs to delineate strains.

Evolutionary taxonomists also use phenotypic characteristics to define a group. As viruses are simple organisms, the choice of characteristics is limited and therefore classification would appear to be simple. However, simplicity generates problems of its own. Some characteristics may be delineated by very small variations in the virus genome, for example, a single nucleotide change can affect virulence or vector transmission. Vector transmission, is therefore one objective phenotypic character that is difficult to use for analysis.

The most insurmountable problem for evolutionary taxonomy presented by viruses is their polyphyletic nature. Polyphyletic groups are formed when two lineages convergently evolve similar character states. Sequence data has demonstrated that viruses have acquired genes from host genomes or horizontally from other viruses through RNA recombination. A polyphyletic group does not contain a common ancestor and therefore evolutionary taxonomy cannot accommodate polyphyletic groups. Similarly, cladistics (phylogenetic analysis) cannot be applied to viruses either. If viruses are truly polyphyletic, as sequence data suggests they are, then the only option left open to virologists is the use of phenotypic classification.

The universal system of virus classification that has been internationally adopted does not involve any implication of evolutionary relationships. It is a purely phenotypic system. However, as higher taxa, above the level of species are being introduced, this

system is being challenged and evolutionary relationships are being inferred. Interestingly, the phenotypic characters that are used to imply these relationships are sequence data which have the potential to be truly objective. It appears that virologists could have moved to a form of phenotypic taxonomy that will be easy to define and apply, the ideal of many zoologists.

The history of plant virus taxonomy

As far back as 1930, a forum for the discussion of virus taxonomy was introduced. However, little progress was made since information was scarce and there was still argument over the intrinsic nature of viruses and even their existence. With improvements in electron microscope techniques and the discovery of DNA, new information on viruses was collected apace through the 1950s and early 1960s. In 1966, at the International Congress for Microbiology in Moscow, the International Committee for Nomenclature of Viruses (ICVN) was first established (Matthews, 1985). A system for the classification of viruses was suggested by Lwoff at the meeting (Lwoff *et al.*, 1962). It relied on two main features, the architecture of the virion capsid and the type of nucleic acid, to sort viruses into descending hierarchical divisions. There were several objections to the scheme. Although phylogenetic relationships were not implied by the system, it was thought that this would be easily forgotten and therefore eventually become misleading. There was concern that the characteristics used for the divisions were arbitrary and were unequally weighted, and the view was expressed that there was little reliable information available on the diversity of viruses. It was concluded that more data needed to be collected before a satisfactory system could be devised.

The 1966 ICVN meeting rejected the system proposed by Lwoff and adopted a code of nomenclature and a set of rules which included the rejection of the bacterial nomenclature system for viruses. It was agreed that all nomenclature would be international and applied to all viruses and that the law of priority would not be observed. No person's name would be used in naming a virus or group and that an effort would be made towards a latinized binomial nomenclature (Matthews, 1985).

The Committee continues to meet every three years and publishes a summary of the official approved viral taxonomy and nomenclature. It changed its name in 1973 to the International Committee on Taxonomy of Viruses (ICTV) to more closely reflect its objectives. The ICTV now operates through a series of committees and study groups which allows approval to be viewed by a wide cross section of the ICTV membership. This has been reflected in the unity of virus taxonomy, where for example, several plant

viruses have been classified with viruses infecting vertebrate hosts, which is unlikely to have happened in a more restrictive and directed environment (Matthews, 1985).

The current state of plant virus taxonomy

The latest edition of the ICTV publication, Murphy *et al.* (1995) states that the species taxon is still the hardest to apply to viruses. The ICTV has now adopted the following definition of a virus species:

> " A virus species is defined as a polythetic class of viruses that constitutes a replicating lineage and occupies a particular ecological niche." (van Regenmortel, 1990). Members of a polythetic class are defined by more than one property and no single property is essential or necessary (Murphy *et al.*, 1995).

The ICTV acknowledges that the species taxon is regarded as the most important hierarchical level in classification, but has proved the most difficult to define. The work of the ICTV study groups on the specific properties defining species suggests that the term species may eventually be defined similarly to the term virus (Murphy *et al.*, 1995). Van Regenmortel (1990) has suggested that the common use of the term virus and group could be accepted as species and genus respectively.

An approach towards a higher taxonomy for viruses is being made slowly by the ICTV. Progress has occurred at widely differing rates, reflecting the advances in different areas of research. Plant virus taxonomists have long preferred to use the non-committal term "group" and have been slow to apply the term genus. They have been reluctant to accept that a higher taxonomy above the level of family exists. Although there are several accepted families, at present only one order has been approved by ICTV, *Monoegavirales*, which are negative-sense single-stranded RNA viruses. Interestingly, this order comprises three families, two contain viruses which have vertebrate hosts (*Paramyxoviridae* and *Filoviridae*) and the third (*Rhabdoviridae*) includes those with vertebrate hosts and two genera of plant-infecting viruses (*Cytorhabdoviruses* and *Nucleorhabdoviruses*).

Several schemes using supergroups or classes above the level of order have been suggested for positive-sense single-stranded RNA viruses (Goldbach & Wellink, 1988; Habili & Symons, 1989; Ward, 1993; Koonin & Dolja, 1993). These relationships are based on comparative amino acid sequence data for several conserved viral genes encoding proteinases, helicases and RNA-dependent RNA polymerases. Plant virologists have been encouraged to explore the use of higher taxa (Mathews, 1985) yet

discouraged by others from using sequence data to imply evolutionary relationships in a hierarchical structure (Rybicki, 1990; Rybicki & Shukla, 1992).

The use of sequence data

The classification of viruses raises some fundamental questions about the procedures used to investigate new or previously unclassified viruses. Historically, morphological characteristics have been given much weight, but with the advent of sequencing technology, these properties are being re-examined.

Viral genome sequence data have become an accepted tool of the plant virus taxonomist. Shukla & Ward (1988) successfully employed amino acid sequence homology as a basis for identification and classification of the *Potyviridae*. The members of the potyvirus family are positive-sense single-stranded RNA viruses, which are encapsidated in flexuous rod-shaped particles. They form the largest of all the plant virus families, and include some of the most economically important viruses (Shukla *et al.*, 1994). Until the analysis of their coat protein sequences, it was generally recognised that the taxonomy of the potyviruses was in a very unsatisfactory state, due in part to its large size and the vast variation of its members. The analysis involved the coat protein sequences of 17 strains of 8 distinct potyviruses. When the 136 possible comparisons of the coat protein sequences were graphed as a frequency distribution, a clear bimodal distribution of sequence homologies showed that sequence homology between distinct members ranged from 38-71 % (average 54%) while that between strains of one virus from 90-99% (average 95%) (Shukla & Ward, 1988). This general rule has been validated as more sequence data emerged. The procedure was repeated with 56 strains of 25 distinct potyviruses (Ward *et al.*, 1995). The trypsin resistant core of the coat protein was used for this analysis. The resulting pattern of distribution was more complex than in the original report. Four levels of homology were observed, reflecting the taxa of genus, species, subspecies and strain, thus demonstrating that the coat protein amino acid sequence alone can reflect the taxonomic position of a particular potyvirus.

Other viral gene sequences have been used as taxonomic parameters. Kamer & Argos (1984) aligned the RNA-dependent polymerases of plant, animal and bacterial viruses and found a stretch of hydrophobic residues surrounding a Gly- Asp-Asp motif (GDD). This conserved region has been used to construct phylogenies for superfamilies or classes of positive-sense single-stranded RNA viruses (Ward, 1993).

Many virologists are naturally wary of classifying viruses by comparative sequence data on the basis of only one or two genes. Horizontal transfer of genes between viruses or

from host species and RNA recombination (for review see Simon & Bujarski, 1994) has been demonstrated and the polyphyletic nature of viruses precludes the inclusion of a single common ancestor for all viruses. However, at least at the level of superfamily or order, it does appear that there are certain valid 'ancestral' properties that were not apparent until the advent of sequence data. To avoid these problems, Ward (1993) proposes the use of a single gene, the RNA-dependent RNA polymerase, on which to base the higher taxa of RNA viruses.

Taxonomy of narcissus latent virus and Maclura mosaic virus

This thesis aims to investigate two plant viruses, narcissus latent virus (NLV) and Maclura mosaic virus (MacMV) by using molecular techniques to accumulate more information about them and to elucidate their taxonomic position.

Narcissus latent virus

The first report of a virus now known to fit the description of narcissus latent virus (NLV) was in 1966 (Brunt & Atkey, 1967). A previously unrecognised filamentous virus with particles of mean length 635 nm was observed in the sap of naturally-infected daffodils with conspicuous yellow stripe symptoms which also contained narcissus yellow stripe virus (NYSV). Further investigation by Brunt defined NLV as a virus with slightly flexuous filamentous particles *c*.657 nm long and *c*.13 nm wide, and with an estimated coat protein molecular weight of 32.6 kDa (Brunt, 1977).

NLV has been reported to induce very mild leaf chlorosis in the tips of narcissus leaves and affects many commercially important cultivars (Brunt, 1977) but is often described as symptomless. It also occurs naturally in bulbous iris, gladioli, nerine and *Acidanthera* (Brunt, 1977; Derks *et al.*, 1985) and there has been one report of its occurrence in garlic (Walkey, 1990). NLV commonly occurs as part of a mixed infection in narcissus, and it is possible that it exacerbates the symptoms of other viruses (Brunt, 1995).

Several properties of NLV led it to be classified as a carlavirus for almost 20 years (Francki *et al.*, 1991) . The size and morphology of its particles fell within the accepted range for carlaviruses (600-700 nm) and the inability to induce symptoms on infection in some cultivars mirrored the characteristically 'latent' carlaviruses. However, in preliminary tests NLV coat protein cross-reacted only with antiserum raised to one carlavirus, lily symptomless virus (LSV) (Brunt, 1977). NLV was later found to be serologically distinct from carnation latent virus and 13 other carlaviruses (Adams &

Barbara, 1982) as well as several carlaviruses found commonly in garlic (Van Duk, 1993).

In 1991, NLV was re-evaluated by Mowat and co-workers. They observed a coat protein of a much larger size (45 kDa) than previously reported and most importantly, they observed the presence of cylindrical cytoplasmic inclusions (CCIs) in NLV-infected *Nicotiana clevelandii* (Mowat *et al.*, 1991; Brunt *et al.*, 1994). A by-product of their translation strategy, CCIs are a distinctive feature of the potyviruses. NLV was subjected to new serological tests but NLV particles failed to react to antisera raised to 12 potyviruses and 9 carlaviruses (Mowat *et al.*, 1991). NLV RNA was used in time course *in vitro* translation experiments in an attempt to elucidate its translation strategy. If NLV used a translation strategy similar to the potyviruses, it would be expected to form a large polyprotein which was cleaved over time by viral proteinases into smaller mature products. If however, NLV utilised a strategy similar to the carlaviruses, the first product to accumulate would probably be the coat protein. The *in vitro* translation results proved inconclusive, as only a single product of 25 kDa was obtained which did not precipitate with antisera raised to NLV coat protein (Mowat *et al.*, 1991).

The first link between NLV and Maclura mosaic virus (MacMV) was discovered when NLV particles cross-reacted in enzyme linked immuno-sorbent assay (ELISA) and immuno-electron microscopy (IEM) with MacMV antiserum. Mowat *et al.* (1991) concluded that NLV was neither a potyvirus nor a carlavirus but could be a member of a new genus with MacMV. NLV therefore remained unclassified.

Maclura mosaic virus

A virus causing well marked mosaic symptoms on the leaves of the ornamental tree *Maclura pomifera* was first reported in 1973 (Plese & Milicic, 1973). The diseased leaves were often deformed and the tissues in the intercostal areas were sometimes yellow-green, while other parts were normal in colour, so that a symptom similar to vein banding occurred. The mosaic virus was mechanically transmitted to *Chenopodium amaranticolor* and a few flexuous virus particles of 700-800 nm were observed in the infected sap.

The virus was designated Maclura mosaic virus (MacMV) and when further characterised, was reported to induce CCIs (Plese & Wrischer, 1978). MacMV has flexuous filamentous particles 650-710nm long (Koenig & Plese, 1981), which fall within a similar range to those of NLV. Measurements of NLV and MacMV particles have always shown a single modal length (Brunt & Atkey, 1967; Brunt, 1977; Mowat *et al.*, 1991; Plese & Wrischer, 1978), although some anomalies in the appearance of

the isolated particles have been reported (granular bodies associated with particles). The coat protein size of MacMV is reported to be 45 kDa (Plese & Wrischer, 1978), the same size as that reported for NLV by Mowat *et al.* (1991). Due to the presence of CCIs in infected tissue and its weak cross reaction with antisera to bean yellow mosaic potyvirus (Plese *et al.*, 1979), MacMV was initially classified as a potyvirus but its position within the genus remained uncertain.

Current classification of NLV and MacMV

NLV and MacMV are currently classified as unassigned viruses within the *Potyviridae* (Murphy *et al.*, 1995). Mowat suggested they form a new a group (Mowat *et al.*, 1991) and Brunt has suggested that they form a distinct genus, *Macluraviruses* within the *Potyviridae* (Brunt, 1992; Brunt *et al.*, 1994; Shukla *et al.*, 1994).

Historically, NLV and MacMV have been associated with the two genera of viruses which their particle morphologies fall between, the carlaviruses (610-700 nm) and the potyviruses (680-900nm). Thus, both types of virus are briefly reviewed here in order that the criteria for inclusion in either group is clear when the position of NLV and MacMV is re-evaluated in this thesis.

Carlaviruses

Carlaviruses have a tendency to induce little or no symptoms. This may be a reason why they have long been ignored and are often omitted from major reviews of plant viruses (Foster, 1992). Recently however, a survey of the 1991 ICTV report shows that carlaviruses are the third most numerous genus of 35 recognised genera of plant viruses (Shukla *et al.*, 1994; Francki *et al.*, 1991). There is a comparatively large amount of sequence data available for carlaviruses. The complete genomes of two carlaviruses, potato virus M (PVM: Rupasov *et al.*, 1989; Zavriev *et al.*, 1991) and blueberry scorch virus (BBScV: Cavileer *et al.*, 1994) have been published and there are partial 3' terminal sequences available for seven others.

Carlaviruses have slightly flexuous filamentous particles of 610-700 nm (Wetter & Milne, 1981) and comprise a positive-sense single-stranded RNA genome of approximately 7kb when estimated by agarose gel analysis. The particles are encapsidated by a single protein species with a molecular weight of 32-34kDa. Encapsidated sub-genomic RNAs of approx. 2-3kb and 1.5 kb have been detected by northern hybridisation experiments (Foster & Mills, 1990 a, b). Carlaviruses utilise sub-genomic RNAs for the translation of internal open reading frames (ORFs). This has been confirmed by *in vitro* translation experiments (Foster & Mills, 1990 a, b, c; 1991 a, b; Meehan & Mills, 1991).

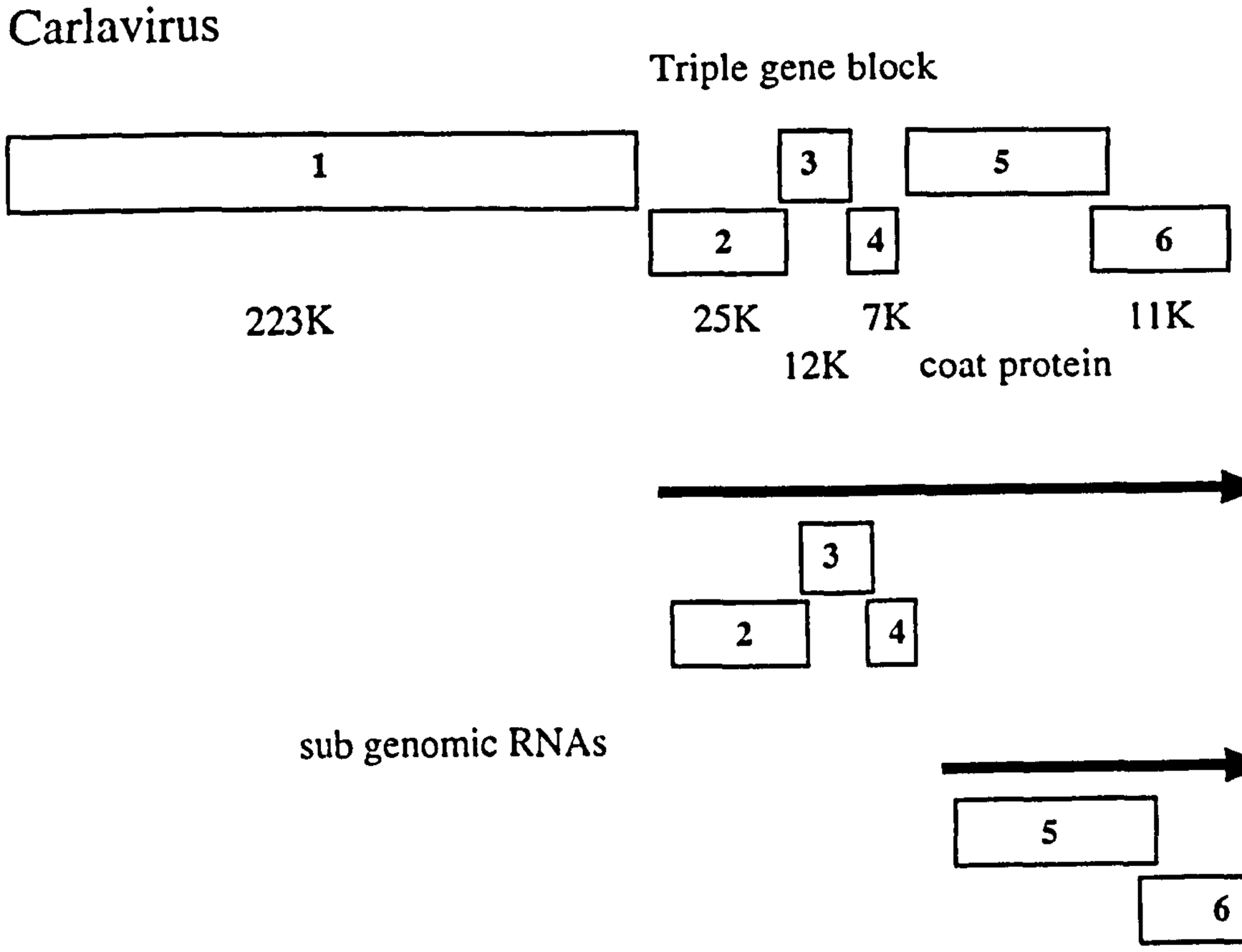


Figure 1.1: Diagrammatic representation of the genome of a typical carlavirus (after Cavileer *et al.*, 1994).

The final cistron encoded by the carlavirus genome, the 11K ORF is thought to be unique among plant viruses. It distinguishes the carlaviruses from the closely related potexviruses and contains a cysteine-rich region. This region is similar to the putative zinc-finger binding domain (Berg, 1986; Klug & Rhodes, 1987) which is proposed to bind nucleic acid. It has been suggested that the 11K protein is involved in the regulation of host-gene transcription or viral RNA replication (Foster, 1992). It is also interesting to note that carlaviruses are transmitted by aphids but their close relatives, the potexviruses are not and that the 11K ORF is the only major cistron absent from the latter. This could suggest a role in aphid transmission for the 11K ORF.

Although the carlaviruses are placed between the potexviruses and potyviruses because they have an intermediate particle size, they have a translation strategy which is different from the potyviruses and a different genome organisation from both genera.

Potyviridae

The *Potyviridae* is the largest family of plant viruses, containing 198 viruses. Its members are so numerous and the research into them has been so prolific that a substantial body of work has been recently produced on this family alone (Shukla *et al.*, 1994). The *Potyviridae* was created by the ICTV Potyvirus Study Group at its meeting in 1990 (Barnett, 1991; 1992) and comprises three accepted genera, potyviruses, bymoviruses and rymoviruses and one possible genus, ipomoviruses. Of these, the potyvirus genus is by far the largest, and although its exact size is always in doubt, due to the large numbers of possible strains, sub-species and unassigned viruses, it is accepted as the largest of the 47 plant virus genera recognised by the ICTV (Murphy *et al.*, 1995). The potyvirus family has been grouped together by the possession of a common genome organisation and high sequence identity within the RNA-dependent RNA polymerase protein. Each genus is transmitted by a different vector, although this was not a criterion for grouping which was based originally on coat protein sequence data (Shukla *et al.*, 1994). The potyviruses are transmitted by aphids, the bymoviruses by a soil-borne fungus, the rymoviruses by eriophyid mites and the ipomoviruses by whiteflies.

Potyviruses

The potyvirus group was first officially recognised by the ICTV in 1971 and the acronym from the type member, potato virus Y (PVY) was adopted (Harrison *et al.*, 1971). Potyviruses cause significant losses in a wide variety of crops, such as cereal, millet, fruit, vegetables, sugar cane, oilseed, ornamental, fodder and pasture world-wide (Shukla *et al.*, 1994). Potyviruses are transmitted by aphids in a non-persistent

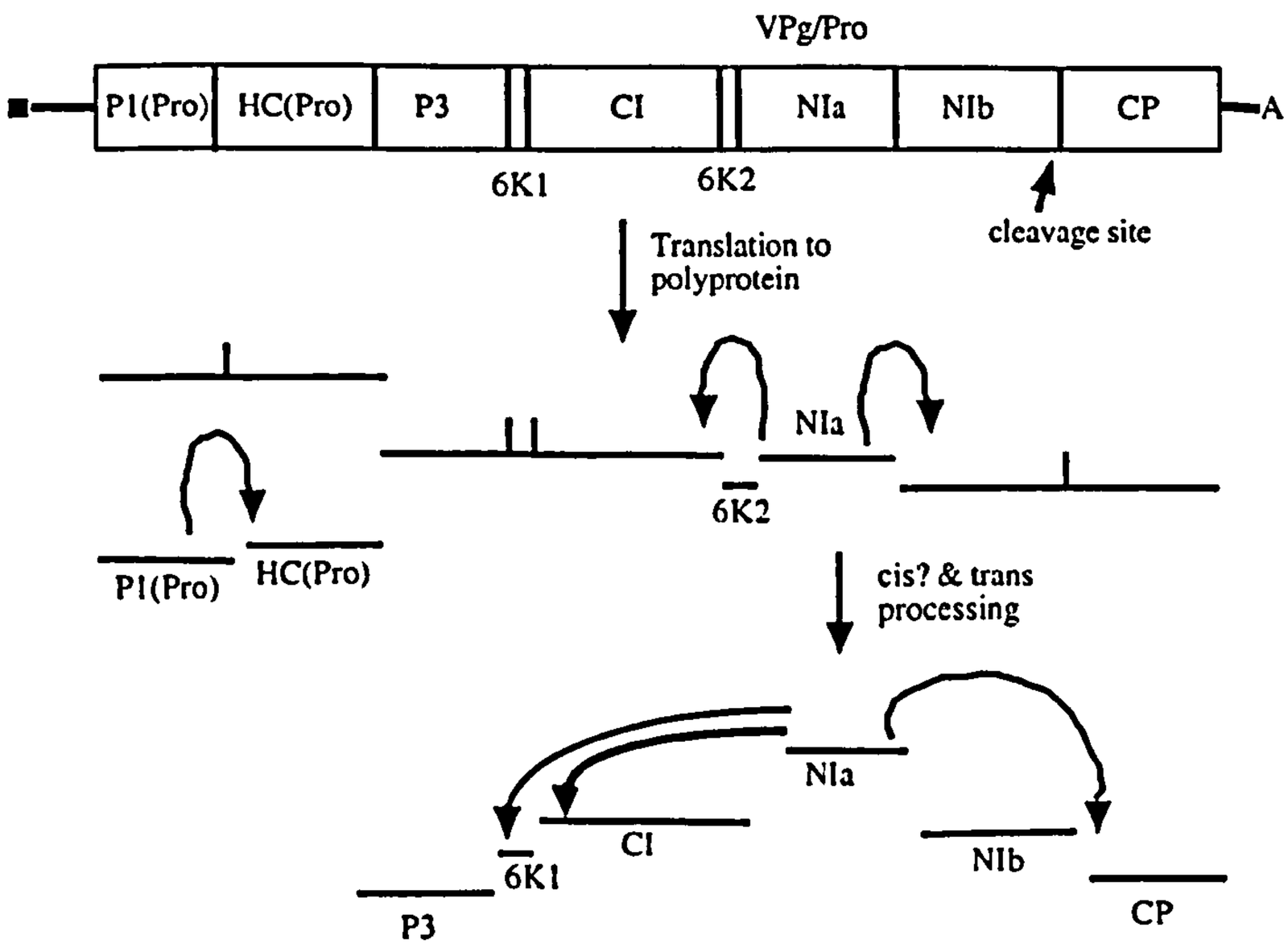
manner and have particles which are flexuous and fall within the range of length 720-770 nm, diameter 11-12 nm.

Many potyviruses have now been completely sequenced and as a consequence there is a great deal of information on their genetic organisation. A diagrammatic representation of a typical genome is shown in figure 1.2a. *In vitro* translation experiments have shown that a large polyprotein representing the entire coding region is initially translated (Dougherty & Carrington, 1988; Dougherty *et al.*, 1989). This product is then autocatalytically cleaved by three viral encoded proteinases (see figure 1.2a). The first proteinase to be characterised was the small nuclear inclusion proteinase (NIa) which is involved in processing the C-terminal two thirds of the genome (Dougherty & Carrington, 1988). It has a two domain structure, the N-terminal portion functions as the genome-linked VPg and the C-terminal domain functions as the proteinase. The NIa proteinase recognises a cleavage sequence that appears to be specific to each virus and is conserved at each processing junction. A group-specific consensus cleavage sequence has been defined as V-X-X-Q/[A,S, G or V]. Variations in the cleavage site can cause differential cleavage rates, and hence offer post-translational control of virus gene function and expression (Dougherty & Parks, 1989).

The second proteinase to be clearly defined was the Helper Component cysteine proteinase (HC-Pro) (Carrington *et al.*, 1989). Again this has a two domain structure with the C-terminal protein involved in proteolysis and the N-terminal domain having a role in vector transmission and possibly long distance movement. HC-Pro is responsible for the cleavage at one site only, HC-P3 (Carrington & Herndon, 1992). The final cleavage site between HC and P1 is processed by the proteinase domain of P1 which is a serine proteinase (Verchot *et al.*, 1991). This proteinase was difficult to identify due to variation of its function in different *in vitro* translation systems. It is possible that P1-Pro requires a host factor or a some other aspect of a subcellular environment for activation. The P1 protein has been implicated in viral cell-to-cell movement (Domier *et al.*, 1987).

The nuclear inclusion body b (NIb) protein is thought to be involved in viral genome replication. It contains the 'G-D-D' motif, which has been demonstrated to be essential to the function of RNA-dependent RNA polymerase (Inokuchi & Harishima, 1987; 1990). This domain is also found in many other viral RNA-dependent RNA polymerases (Domier *et al.*, 1987). The NIb is the most conserved of all the potyvirus proteins. The cytoplasmic inclusion protein (CI) is also thought to be involved in genome replication and has been shown to be a RNA helicase (Lain *et al.*, 1990). It contains the conserved nucleotide binding domain G-X-X-G-X-K-S (Domier *et al.*,

a) Potyvirus



b) Bymovirus

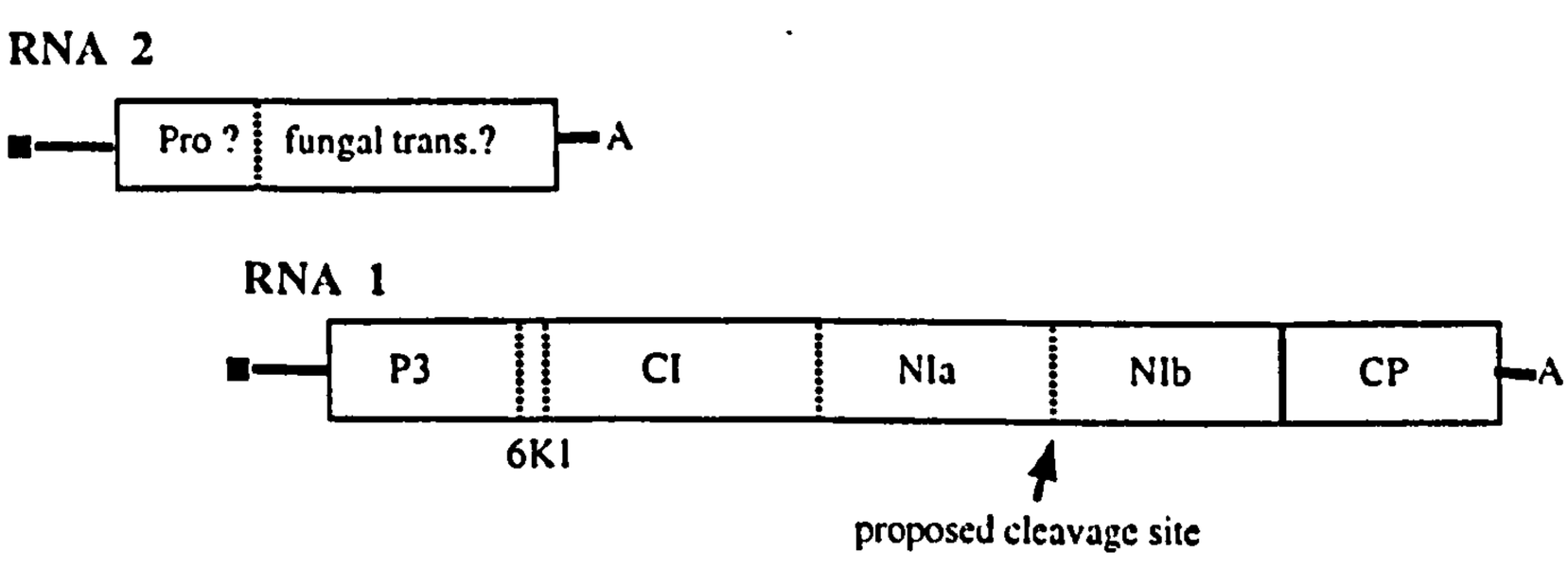


Figure 1.2: (a) Diagrammatic representation of the genome of a typical potyvirus and its translation strategy. A polyprotein is cleaved by viral encoded proteinases to produce mature proteins (after Riechmann *et al.*, 1992). (b) Diagrammatic representation of the genome of a typical bymovirus (after Kashiwazaki, 1996)

1987). The CI is the only component of the cytoplasmic cylindrical inclusions which are so characteristic of potyviral infections. It has been suggested that it is associated with plasmodesmata in order to facilitate cell-to-cell movement. VPg is the third protein thought to be involved in viral replication. VPg is covalently bound to the 5'-terminus of the genome and is thought to act as a primer for RNA replication.

The potyvirus coat protein is the only potyvirus product which shows little sequence identity to the corresponding protein of other virus groups (Domier *et al.*, 1987). Its amino acid composition is said to be characteristic of the genus and its use in the classification of potyviruses has already been discussed. It is known to be involved in the specificity of vector transmission and contains a conserved motif (D-A-G) which is essential for this function (Atreya *et al.*, 1990) and is normally placed in close proximity to the N-terminus.

Bymoviruses

The bymoviruses were first recognised as a cohesive group at the end of the 1980s (Usugi *et al.*, 1989) and were officially recognised as a distinct genus of the *Potyviridae* two years later (Barnett, 1991; 1992). To date, the genus includes five viruses, barley mild mosaic virus (BaMMV), rice necrosis mosaic virus (RNMV), oat mosaic virus (OMV), wheat streak spindle mosaic virus (WSSMV, synonym wheat yellow mosaic virus) and the type member, barley yellow mosaic virus (BaYMV). All bymoviruses are transmitted by the soil borne fungus, *Polymyxa graminis* by the *in vivo* method (Campbell, 1996). This is analogous to the persistent transmission of viruses by insect vectors. The virus particles are contained within the zoospores when they emerge from vegetative sporangia or resting spores.

Bymoviruses are unusual members of the *Potyviridae* because they have a bipartite particle and genome structure. The flexuous rod-shaped particles are 200-300 nm and 500-600 nm in length and encapsidate positive-sense single-stranded RNAs of approximately 4 and 8kb (Usugi *et al.*, 1989). Two bymoviruses have been completely sequenced on both RNAs, BaYMV (Kashiwazaki *et al.*, 1989; 1990 a; 1991; Davidson *et al.*, 1991; Peerenboom *et al.*, 1992) and BaMMV (Kashiwazaki *et al.*, 1992; Foulds *et al.*, 1993; Schlichter *et al.*, 1993; Timpe & Kuhne, 1994; Kashiwazaki, 1996; Meyer & Dessens, 1996; Peerenboom *et al.*, 1996). Analysis has shown that the cistrons encoded on RNA1 are similar to those encoded by the C-terminal two thirds of potyviruses, and that RNA2 corresponds in part to the N-terminal third (see figure 1.2b). Sequence analysis and *in vitro* translation studies (Prols *et al.*, 1990) have shown that BaYMV follows the same translation strategy as the potyviruses, producing a large polyprotein which is cleaved internally. RNA1 encodes a large polyprotein

(approx. 270kDa) which contains five internal cleavage sites to produce six putative functional proteins. These are equivalent to the potyviral proteins P3, 6K1, CI, NIa, NIb, and CP and contain all the major conserved domains found in the potyviruses (Kashiwazaki *et al.*, 1990 a).

RNA2 of BaYMV and BaMMV encodes a polyprotein of 98kDa which is autocatalytically cleaved into two functional proteins. The smaller of these two proteins, P1 (28kDa in BaYMV) contains proteinase domain similar to that found in the potyvirus HC-Pro (Kashiwazaki *et al.*, 1991; Davidson *et al.*, 1991). However the bymovirus P1 protein appears to have similarities to only the proteinase domain of HC-Pro and has no equivalent to the potyvirus Helper Component aphid transmission domain. The second RNA2 encoded protein, P2, appears to be associated with fungal transmission, sharing conserved motifs with a fungal transmission protein of furoviruses (Dessens & Meyer, 1996; Dessens *et al.*, 1995; Peerenboom *et al.*, 1996). Bymoviruses which are maintained in laboratory conditions by repeated mechanical transmission eventually become unable to be transmitted by fungal vectors (Dessens & Meyer, 1995; Jacobi *et al.*, 1995). Analysis of RNA2 of these mechanically transmitted isolates reveals a deletion in P2, supporting the suggestion that this is protein has a fungal transmission function (Dessens *et al.*, 1995; Jacobi *et al.*, 1995). Both P1 and P2 are thought to be involved in long-distance and cell-to-cell movement (Schenk *et al.*, 1995).

Rymoviruses

The rymoviruses were recognised as a separate genus of the *Potyviridae* at the same time as the bymoviruses (Barnett, 1992), however the accumulation of sequence data of these viruses has led to recent suggestions that this genus should be split (Salm *et al.*, 1996a). The Rymovirus genus currently contains five members, Agrypyron mosaic virus (AgMV), Hordeum mosaic virus (HoMV), oat necrosis mottle virus (ONMV), wheat streak mosaic virus (WSMV) and the type member ryegrass mosaic virus (RGMV). There are also several possible members, brome streak mosaic virus (BrSMV), onion mite-borne latent virus, shallot mite-borne latent virus and spartina mottle virus (Shukla *et al.*, 1994). The Rymoviruses are monopartite, like the potyviruses, but with distinctively shorter, wider particles (680-750 nm in length, 15 nm in diameter). They are transmitted by eriophyid mites and in general infect only graminaceous hosts.

Cloning and sequencing of the complete genome of BrSMV (Gotz & Maiss, 1995; Schubert & Rabenstein 1995) and the partial genomes of WSMV (Niblett *et al.*, 1991) AgMV, HoMV (Salm *et al.*, 1996b) and RGMV (Schubert *et al.*, 1995; Salm *et al.*,

1996c) has shown that they have a similar genome organisation and translation strategy to the potyviruses. Nucleotide and amino acid sequence comparisons of the 3' terminal regions of these viruses indicate that two distinct groups exist (Salm *et al.*, 1996b) and phylogenetic justification for splitting the Rymovirus genus has been presented (Salm *et al.*, 1996c). It is proposed that WSMV and BrSMV form a distinct cluster and are more similar to one another than they are to RGMV, AgMV and HoMV which form a second cluster. It is interesting to note that RGMV and AgMV are transmitted by the mite *Abacarus hystix*, whereas WSMV and BrSMV are transmitted by a different mite, *Aceria tulipae*. The authors suggested that a new genus *Whestreviruses* be created to include WSMV and BrSMV, while AgMV, RGMV and HoMV be retained in the *Rymovirus* genus (Salm *et al.*, 1996a).

Ipomoviruses

The ipomovirus genus is not yet recognised by ICTV and at present contains only one member, sweet potato mild mottle virus (SPMMV) and one possible member, sweet potato yellow dwarf virus (Shukla *et al.*, 1994). SPMMV is transmitted by whitefly and has particles which are longer and slightly wider than those of potyviruses (750-900nm, diameter 14 nm). Sequence analysis of the 3'-terminal region of SPMMV provided molecular data which suggested that it was a member of the *Potyviridae* (Colinet *et al.*, 1996).

It can be seen from this review of the carlaviruses and the *Potyviridae* that they can be easily distinguished by the presence of CCIs on infection. The limit of our knowledge about NLV and MacMV currently assigns them to the potyviridae family due to the presence of CCIs but not to a genus . Modern sequencing techniques should provide further information to compare NLV and MacMV to the potyviruses and carlaviruses to establish the validity of this classification, as a wealth of sequence data for these groups is now available.

Aims of the Thesis

The central aim of this thesis is to elucidate the taxonomic position of NLV and MacMV by the use of molecular techniques.

1. The assumption that NLV could no longer be assigned to the carlavirus group due to the observation of CCIs on infection (Mowat *et al.*, 1991; Brunt *et al.*, 1994) will be tested by the design of a PCR primer specific to carlaviruses which could be used to amplify viral first-strand cDNA. If successful, this rapid test on total RNA extracted from infected plant material would also provide an amplified fragment suitable for cloning. Sequence analysis could then confirm the RT-PCR test results. This

carlavirus-specific PCR primer would also be used to test several 'suspected' carlaviruses, red la soda virus, American hop latent virus, hop mosaic virus and cowpea mild mottle virus (Chapter 3) .

2. It was suspected that this strategy may prove unsuccessful on NLV and MacMV, therefore a second strategy was developed. An attempt would be made to obtain sequence data for both genomes by an alternative route. Viral genomic RNA can be used to create first-strand cDNA which when converted to double-stranded cDNA, can be cloned directly. A mixture of oligo-d(T) and random hexanucleotide primers can be used to create a wide range of clones. When transformed into competent cells a library of clones can be screened using radio-labelled first-strand viral cDNA. These clones should provide genome sequence data which when used in a database search would allow homologous sequences to be identified. If the coat protein gene was isolated then it could be used to determine the relationship of NLV and/or MacMV to other viruses according to the principles established by Shukla & Ward (1988) (Chapter 4).

3. A universal potyvirus primer could be used to further investigate the relationship between narcissus yellow stripe symptoms and NLV. Confusion exists over the link between disease symptoms described as "narcissus yellow stripe" and the infection of narcissus by NLV. NLV often occurs as part of a mixed infection, and as it was assumed to be a carlavirus for many years, it is possible that this confusion may be connected to its misdiagnosis. However, a specific test for NLV would enable it to be accurately identified in naturally infected narcissus plants (Chapter 5).

4. If sequence data for NLV and MacMV was obtained, another aim of this thesis is to produce a PCR primer specific to both viruses which could be used on other unassigned viruses. Other viruses similar to MacMV and NLV may be identified to aid the delineation of their shared characteristics further. If MacMV and NLV indeed represented a new genus, another member would facilitate the definition of the genus (Chapter 6).

5. Finally, if sequence data was obtained, it was intended that it be used to create constructs which could be transformed into tobacco plants to further study the effects of the viruses and in an attempt to create resistant plants (Chapter 7).

Chapter 2

Materials and Methods

This chapter details the methods used to carry out the work presented in this thesis. All commonly used solutions have been listed in Appendix I, and the names and addresses of suppliers in Appendix II.

1) Methods involving DNA

Unless otherwise stated, all glassware and disposable plastic equipment used in this section were autoclaved before use.

1.1. Purification of DNA

1.1.1 Mini-preparation of plasmid DNA

1. Inoculate 5ml of LB broth (see Appendix I) with the required clone from a plate or a previous overnight culture. Add the appropriate antibiotic selection (use ampicillin at 100µg/ml or kanamycin at 50-100µg/ml) and place in a shaking 37°C Gallenkamp orbital incubator overnight.
2. Take 1.5ml of the overnight culture and place in a 1.5ml sterile eppendorf. Spin at approx. 11,000 x g for 5 min (MSE microcentaur microfuge used throughout).
3. Pour off the supernatant and resuspend the pellet in 100µl solution I (50mM sucrose, 25 mM Tris HCl pH8.0, 10 mM EDTA).
4. Add 200µl solution II (0.2M NaOH, 1% SDS) and mix by inverting. Add 150µl of solution III (3M KoAc: 3M with respect to K, 5M with respect to Acetate) and leave on ice for 10 min.
5. Centrifuge for 10 min at approx. 11,000 x g .
6. Aspirate 450µl of the supernatant using a pipette into a fresh tube, leaving the white precipitate behind.
7. Add an equal volume of isopropanol .
8. Centrifuge for 10 min at approx. 11,000 x g .
9. Pour off the supernatant, absorbing the final few drops onto a paper towel to ensure that the pellet is dry. Resuspend the pellet in 400µl of sterile distilled water, transfer to a 1.5 ml sterile eppendorf tube. Add 1µl RNase A (50 mg/ml) and incubate at 37°C for 10 min
10. Add 400µl of phenol (equilibrated to pH8.0 using 0.5M Tris buffer), vortex briefly (Fisons whirlimixer) and centrifuge for 2 min at approx. 11,000 x g in a microfuge.

11. Carefully remove the upper aqueous layer, leaving behind any protein residue at the interface and place in a new eppendorf tube.
12. Add 200µl of phenol and 200µl of chloroform, vortex briefly and centrifuge for 2 min at approx. 11,000 x g .
13. Again remove the upper aqueous layer and place in a new eppendorf tube. Repeat the final extraction as before but this time with 400µl of chloroform.
14. Add 40µl of 3M NaOAc pH 4.8 and 1 ml of ice cold 100 % ethanol (EtOH). Precipitate at -80°C for 10 min. Centrifuge at approx. 11,000 x g for 10 min.
15. Pour off the supernatant and dry the pellet under vacuum, to ensure that it is free from ethanol.
16. Resuspend in 20µl Sterile Distilled Water (SDW).

1.1.2 Mini-preparation of plasmid DNA for automatic sequencing

Only small quantities of good quality DNA were required for automatic sequencing. This quick method was used which is similar to the mini-preparation method (section 1.1.1).

1. Use a single colony to inoculate 5ml LB broth (See Appendix I), add appropriate antibiotic selection and place overnight in a shaking 37°C incubator.
2. Take 1.5ml of the overnight culture and place in a 1.5ml sterile eppendorf. Centrifuge at approx. 11,000 x g for 5 min. Remove supernatent and repeat this process so that 3ml of cells are harvested.
3. Pour off the supernatant and resuspend the pellet in 200µl solution I (25mM Tris HCl pH8.0, 10mM EDTA, 50mM glucose).
4. Add 300µl solution II (0.2M NaOH, 1% SDS) and mix by inverting. Add 300µl of solution III (as mini-preparation section 1.1.1) and maxi-preparation, section 1.1.3) and leave on ice for 5 min.
5. Centrifuge for 10 min at approx. 11,000 x g.
6. Transfer the supernatant using a pipette into a fresh tube, leaving the white precipitate behind. Add 1µl of RNase A (50µg/ml) and incubate at 37°C for 20 min.
7. Perform 2 chloroform extractions and then precipitate nucleic acid by adding an equal volume of isopropanol.
8. Centrifuge immediately for 10 min at approx. 11,000 x g.
9. Pour off the supernatant and wash the pellet in 70 % EtOH. Dry the pellet under vacuum and resuspend in 32µl SDW.
10. Reprecipitate using an equal volume of PEG solution (13% PEG 8,000, 1.6M NaCl) and mix well. Place on ice for 20 min.
11. Centrifuge at approx. 11,000 x g for 15 min and remove supernatant. Wash in 80% EtOH, dry under vacuum and resuspend in 20µl SDW.

1.1.3 Maxi-preparation of plasmid DNA

1. Inoculate 100ml of LB broth (See Appendix I) with the required clone from a plate or a previous overnight culture. Add the appropriate antibiotic selection (use ampicillin at 100µg/ml or kanamycin at 50-100µg/ml) and place in a shaking 37°C incubator overnight.
2. Divide the culture equally between two 50ml tubes and centrifuge at 4°C for 10 min at 3,200 x g (Beckman benchtop centrifuge used throughout).
3. Pour off the supernatant and thoroughly resuspend the pellet in 2 ml solution I (as section 1.1.1).
4. Add 4 ml solution II (as section 1.1.1) and mix by inverting. Add 4 ml of solution III (as section 1.1.1) and leave on ice for 10-30 min.
5. Centrifuge at 4°C for 10 min at 3,200 x g.
6. Filter the supernatant through polymer wool into a fresh tube, leaving the white precipitate behind.
7. Add an equal volume of isopropanol and leave to precipitate for 1 hr at room temperature.
8. Centrifuge at 20°C for 15 min at 3,200 x g.
9. Pour off the supernatant, absorbing the final few drops onto a paper towel to ensure that the pellet is dry. Resuspend the pellet in 300µl of sterile distilled water (SDW) and transfer to a 1.5ml sterile eppendorf tube.
10. Add an equal volume of 10M LiCl. Centrifuge at approx. 11,000 x g for 10 min.
11. Take up the supernatant with a pipette and place in a fresh eppendorf tube. Add an equal volume of PEG (13% PEG 8,000, 1.6M NaCl) and leave at -80°C for 10 min or overnight.
12. Centrifuge at approx. 11,000 x g for 10 min.
13. Pour off the supernatant and ensure the pellet is free from any residual PEG solution. Resuspend the pellet in 400µl sterile distilled water and add 4µl of RNase A (50 µg/ml). Incubate at 37°C for 15 min.
14. Follow method as for mini-preparation (1.1.1) steps 10 -15.
15. Resuspend the pellet in 50µl of SDW.

1.1.4 Preparation of single-stranded plasmid DNA for sequencing

Some plasmid vectors contain an F1 origin of replication enabling the export of a single-stranded molecule including insert sequences. Single-stranded sequencing was used since it offers a speedier protocol, longer reads (up to 700 bp) and better clarity and resolution than double-stranded sequencing.

1. Add the following to a 50ml sterile screw top plastic tube:

   5ml LB Broth (See Appendix I)

   25µl M13KO7 helper phage

24µl Ampicillin (25mg/ml)
100µl fresh overnight culture of Bluescript clone (or other vector with F1 origin)

2. Incubate at 37°C overnight in a shaking incubator.
3. Centrifuge the whole 5ml culture at 3,200 x g for 5 min to pellet the cells.
4. Split the supernatant into 3 eppendorfs, placing 1.5ml in each and centrifuge for 5 min at 11,000 x g .
5. Remove 1ml of supernatant from each of the tubes to fresh eppendorfs containing 200µl of 20% PEG 6,000/2.5M NaCl and mix by inverting.
6. Leave at room temperature for 20 min.
7. Centrifuge for 10 min at approx. 11,000 x g and pour off most of the supernatant.
8. Centrifuge for a further 3 min and remove the remaining PEG mixture with a drawn out pasteur pipette. The white pellet of phage should be visible at this stage.
9. Add 100µl sterile distilled water to each pellet and add 50µl of phenol. Vortex for 10 s.
10. Centrifuge for 3 min at 11,000 x g and transfer upper aqueous layer to a fresh eppendorf, combining the contents of the three tubes.
11. Extract once more with an equal volume of chloroform.
12. Add 1/4 volume of 10M ammonium acetate pH4.8 and 2 volumes of absolute ethanol. Mix and store at -20°C until required.
13. Precipitate the ssDNA by centrifugation for 10 min at 11,000 x g. Wash the pellet in 1ml 70% ethanol, aspirate carefully and vacuum dry.
14. Resuspend the pellet in 40µl of water. Check the quality of the ssDNA by spectrophotometric analysis (section 1.2.2) or on a 1% agarose gel (section 1.2.1).

1.2. Analysis

1.2.1 Agarose gel electrophoresis of DNA

For visual analysis of DNA, agarose gels of various sizes were used (1% gel for DNA, 1.5% or 2% for RNA or RT-PCR products). Agarose was melted in 1 x TAE (8mM Tris, 30mM sodium acetate trihydrate, 1mM EDTA, pH8.2) in a microwave oven on full power for 1 min and 30 s. The agarose was allowed to cool and then ethidium bromide (EtBr) was added to a concentration of 0.5µg/ml. The melted gel was poured into a casting tray of suitable size which had been taped at each end with masking tape and contained a comb in the required position. Once set, the tapes and comb were removed. The gel was placed in an electrophoresis tank containing 1 x TAE so that it covered the gel completely. Samples were prepared by adding 3x loading dye (1 x TAE, 10% glycerol, 0.5µg/ml bromophenol blue). Electrophoresis was carried out at a

constant voltage of 100V. When the dye front had run to the required point, the gel was viewed on a UV trans-illuminator (UVP dual intensity transilluminator) at 320nm and photographed using a video camera with an orange filter (UVP camera). Hard copy was obtained by printing the image on a thermal printer (Mitsubishi video copy processor).

1.2.2 Spectrophotometric analysis of DNA

2µl of DNA was diluted in 200µl of SDW. The spectrophotometer (Hewlett Packard 8452A Diode Array spectrophotometer) was blanked using sterile distilled water and readings taken of the DNA sample required at OD_{260}. The spectrophotometer was calibrated so that one OD_{260} unit is equal to 3.7 ng/µl of double-stranded DNA. Therefore, the OD_{260} value was multiplied by 3.7 to calculate the amount of DNA in µg/ml in the original sample. Oligonucleotides were sampled in the same way and OD_{260} value of 1 was taken to be equivalent to a concentration of 20µg/ml.

The purity of single-stranded DNA was ascertained by calculating the ratio of the values for OD_{260}/OD_{280}. For sequencing quality this value should be approximately 1.8.

1.2.3 Polymerase Chain Reaction (PCR)

A stock of 11x PCR buffer was prepared:

Amount	Stock solution	[Final] in 11x Buffer
165 µl	2M Tris HCl pH 8.8	50mM
85 µl	1M Ammonium sulphate	12mM
33.5 µl	1M Magnesium Chloride	4mM
3.6 µl	2-mercaptoethanol	0.5% (v/v)
3.4 µl	10mM EDTA pH 8.0	0.5mM
75 µl	each 100mM dNTP	11mM
85 µl	10 mg/ml BSA	1mg/ml
676.5 µl	(total volume)	

All stock solutions used in the buffer were filter sterilised. The buffer was stored at -20°C. All primers were used at the concentration of 10 µM (for a 25mer this was approximately 78 ng/µl).

1. In autoclaved 0.5 ml PCR tubes, place 1.8 µl of 11 x PCR buffer, 2 µl of each primer used, 4 µl of DNA (at 100ng/ml) and 1 unit of *Taq* polymerase (Advanced Biotechnologies Ltd). Make the reaction up to a volume of 20 µl.

2. Cover the reaction in 100µl of mineral oil (Sigma) to prevent evaporation and place tubes in Perkin Elmer 480 Thermal Cycler.
3. Select and run the required cycle. For routine re-amplificiation, colony screening (section 1.6.3) and amplification from plasmids use 25 cycles of 94°C for 1 min, 55°C for 1 min and 72°C for 2 min.

In general, 1 min per kilobase of target sequence to be amplified during the extension phase of the cycle (72°C) was allowed. The annealing temperature of the reaction was altered depending on the annealing temperature of the primer and the amount of degeneracy in the primer sequence.

When preparing large numbers of reactions, it was easier and quicker to prepare a master mix of 11 x PCR buffer, SDW and *Taq* polymerase. DNA was always added last after negative controls had been set up. A control where SDW replaces DNA was always included to ensure that the components of the reaction mix were not contaminated. Positive controls were the last reactions assembled to prevent aerosol contamination of the pipette coming into contact with other reactions.

1.2.4 Purification of DNA from agarose gels

The required DNA band was visualised by UV illumination and excised (using the 302nm wavelength of UV whilst cutting to minimise UV damage to the sample). The Bio-101 gene clean kit III (Bio 101 Inc.) was used in accordance with the manufacturers instructions. This protocol involved the binding of DNA to silica in the presence of high salt at temperature. The DNA was washed and then recovered in SDW by heat. A third to a half of the resuspended sample was analysed by agarose gel electrophoresis to estimate the concentration of DNA recovered.

1.2.5 Restriction Enzyme digestion of DNA

DNA was used in a volume of 20µl and an appropriate amount of 10x buffer (as recommended by manufacturers of the enzyme) was added, followed by 1 unit/µg DNA of the required restriction enzyme. Samples were incubated at 37°C for 1 hr and then analysed by agarose gel electrophoresis (section 1.2.1).

When digestion with two enzymes was required, the compatibility of the buffers used with each enzyme were compared. If they were compatible, then both enzymes were added as above. However, if they were incompatible, then the sample was digested first with one enzyme and then the sample was subjected to a phenol/chloroform extraction in a volume of 200µl and precipitated by salt and EtOH (see section 1.1 steps

10 - 16). The sample was resuspended in 20µl of SDW and the digest repeated with the second enzyme.

1.3. Cloning

Cloning was performed on fragments of DNA and vectors which had been digested with the same or compatible restriction enzymes. Once the digests had been completed and the results analysed by agarose gel electorphoresis, the required fragment and vector were then excised and recovered from the gel using the Gene Clean kit. Fragment and vector were then ligated together:

1. In a total volume of 20µl, combine a 1:8 molar ratio of vector to fragment .
2. Add 2.5µl of 10x buffer supplied with the enzyme and 1µl of T4 DNA ligase.
3. Incubate at 14°C overnight.

Half the ligation reaction was then transformed into competent cells and the remaining portion was stored at -20°C.

1.3.1 Cloning of PCR fragments

PCR fragments generated by conventional means or RT-PCR were cloned using a selection of specialised vectors available commercial. In each case the manufacturers instructions were followed. Vectors used were, pCRII and pCR2.i (Invitrogen) and pCRSCRIPT (Stratagene).

1.4 Transformations

1.4.1 Competent cells

1. Inoculate 10ml LB broth (See Appendix I) with a single colony and no antibiotic selection. Incubate overnight at 37°C.
2. Use 1ml of the overnight culture to inoculate 100ml LB broth. Incubate at 37°C until OD_{600} is between 0.45 and 0.60.
3. Transfer the cells to ice-cold 50ml tubes and centrifuge for 15 min at 3,200 x g at 4°C.
4. Resuspend the pellet in 10ml ice cold 50mM $CaCl_2$. Leave on ice for 1 hr.
5. Centrifuge at 4°C for 10 min and resuspend in 5 mls ice cold 50mM $CaCl_2$ if for immediate or short term use (store at 4°C for up to 4 days). For longer term storage, resuspend in 50mM $CaCl_2$, 10% sterile glycerol and aliquot into 300µl batches and immediately snap freeze in liquid nitrogen prior to storage at -80°C.

1.4.2 Transformation

1. Add 10µl of ligation mix to an aliquot of cells in a 1.5ml eppendorf tube and tap gently to mix. Incubate on ice for 30 min.

2. Heat shock the cells by incubating at 42°C for 90 s, and place back on ice for 2 min.
3. Add 1ml of LB broth and incubate at 37°C for 30 min in an orbital shaking incubator.
4. Centrifuge for 5 min at approx. 11,000 x g and resuspend pellet in 200µl of fresh LB broth. Plate out 180µl and 20µl on agar plates which contain the required antibiotic selection and X-Gal and IPTG if blue-white colour selection is necessary (See Appendix I).
5. Incubate overnight at 37°C.

1.5 Radio-labelling of DNA

The required amount of DNA was used in a volume of 7µl (at least 10 ng of DNA) and incubated at 100°C for 2 min to denature it. The sample was cooled briefly at 37°C and centrifuged to collect the contents of the tube. The sample was then added to a screw top 1.5ml sterile plastic tube containing 0.6µl Klenow (GibCo), 0.6µl bovine serum albumin (stock solution 10 mg/ml BSA), 3µl oligo labelling buffer (See Appendix I) and 1µl α–dCTP^{32} (3,000 Ci/mmole). The reaction was incubated at 37°C for 1 hr and then at 100°C for 2 min to denature before it was added to the required hybridisation solution.

1.6 Colony screening

1.6.1 Colony Screening by hybridisation

Colony screening was used as a rapid method of isolating transformants containing plasmids with a specific insert. This method eliminates the step of preparing plasmid DNA and digesting before the inserts can be identified, instead one or two positive colonies can be identified overnight and then large scale plasmid preparations can be made the following day.

1. Prepare two agar plates with appropriate media and selection. Draw a numbered grid onto the bottom of one petri dish, and cover the surface of the second plate with a piece of Hybond-N (Amersham) which has a duplicate grid drawn in ink or pencil.
2. Pick single transformed colonies and streak out on the duplicate plates. Always include a negative control colony (known to contain a plasmid with no insert).
3. Incubate overnight at 37°C.
4. Lift the filter off the agar using forceps and place onto Whatman paper (3M) soaked in 2xSSC (See Appendix I), 0.5% SDS for 2 min. Store the reference plate at 4°C.

5. Cross-link the lysed cells to the filter by either exposure to UV (UV Stratalinker 2400, Stratagene) or 1 min in a microwave on full power (Proline microwave, 950W)
6. Wash the filter in 2 x SSC (See Appendix I) and then pre-hybridise for 1 hr in a shaking water bath at 65°C in a solution of 6xSSC, 2 x Denhardts solution (See Appendix I), 1% SDS and 100µg/ml sheared single-stranded salmon sperm nucleic acid.
7. Add radio-labelled DNA fragment (as described above in section 1.5) and allow to hybridise overnight.
8. Wash the filter with 3x10 min (3 x SSC, 1% SDS). Monitor with a Geiger counter to ensure a low background signal is obtained.
9. Wrap the filter in Saran wrap and expose to X-ray film at -80°C for 4-6 hr using a high-intensity screen, or overnight as necessary.
10. Develop the X-ray film according to manufacturers instructions or by using an automatic developer. Positive colonies can be identified where a very strong signal is present in comparison to the negative control.

1.6.2 Colony Screening using viral cDNA

Colonies transformed using vectors with inserts of double-stranded viral cDNA were screened with radio-labelled first-strand cDNA of viral RNA in order to identify clones containing inserts of viral origin. The procedure described above (section 1.6.1 from steps 1 to 5) was followed and then:

1. Wash the filter in 2 x SSC and then pre-hybridise for 1 hr in a shaking water bath at 42°C in a solution of 50% formamide, 4 x SSC, 50mM NaH_2PO_4, 100µg/ml sheared single-stranded salmon sperm nucleic acid.
2. Prepare radio-labelled first-strand cDNA using viral genomic RNA and a cDNA synthesis kit (Stratagene) which supplies separate mixes of nucleotides. Follow the manufacturers instructions to anneal the RNA and oligo-dT primer or random primers together and then add all remaining components except dCTP which should be substituted with αP^{32}-dCTP. Incubate at 42°C for 1 hr.
3. Add the labelled first-strand cDNA to the pre-hybridisation solution and hybridise overnight.
4. Treat as described in the above method section 1.6.1 steps 8-10.

1.6.3 Colony screening by PCR

PCR reactions were prepared as described in section 1.2.3, except that no DNA was added. Instead, 4µl of SDW was used to make the volume up to 20µl. Using a sterile pipette tip or toothpick, a single colony was picked from the plate and was placed in the

reaction mixture. The colony was streaked out on a gridded reference plate. The PCR reaction was carried out in the usual way (section 1.2.3)

1.7 Southern analysis

To probe DNA which has been electrophoresed on an agarose gel it was treated to depurinate and denature the DNA prior to transfer to a membrane.

1. Take a photograph of the agarose gel, illuminated by UV, with a fluorescent ruler to allow direct comparison with the autoradiograph later.
2. Cut the top left corner of the agarose gel, so that it can be orientated when blotted.
3. Soak the gel for 10 min in depurinating solution (0.25M HCl), then 30 min in denaturing solution (0.5M NaOH, 1.5M NaCl) and finally for 30 min in neutralising solution (0.5M Tris HCl pH7.5, 3M NaCl).
4. Use a wick system to blot the gel. Almost fill a large tray containing a sponge with 20xSSC, cover the sponge with a piece of 3M Whatman paper. Place the treated agarose gel on top of the paper. Cover the rest of the area left around the gel with clingfilm. Cut a piece of Hybond-N (Amersham) to fit exactly over the gel. Pre-soak the filter in 2xSSC before placing carefully on top of the gel. Smooth the filter down with a clean glass rod and cover with 3 layers of 3M Whatman paper cut to the size of the gel which have been soaked in 2xSSC. Place 3 dry sheets of 3M Whatman paper on top and then add a stack of paper towels or kleenex tissues (depending on the gel size). Compress the whole stack with a weight (approximately 500g). Leave overnight to blot.
5. Remove the blotted filter and discard the agarose gel. Air dry the filter and cross-link using a UV Stratalinker (Stratagene).
6. Incubate the filter in a solution of 6 x SSC, 2 x Denhardts, 1% SDS and 100µg/ml sheared single-stranded salmon sperm nucleic acid at 65°C for 1 hr. Label the DNA probe as described in section 1.5 and add to the solution. Incubate overnight at 65°C in a shaking water bath or hybridisation oven.
7. Proceed to wash the filter and expose to film as described in section 1.6.1, steps 7-9.
8. Develop the autoradiograph and compare to the original photograph of the agarose gel.

2) Sequencing

2.1. Double-stranded sequencing

The T^7 (Pharmacia) or Sequenase (USB) sequencing kits were used according to the manufacturers instructions. For larger clones, a stock of DNA was made by the maxi-

preparation method and stored at -20°C until required. ^{35}S dATP was used to label the reactions. Compressions were resolved using either 7-deaza dGTP termination mixes or dITP depending on the kit used.

2.2. Single-stranded sequencing

The same method as described for double-stranded sequencing was followed, except that annealing of the primer to the template was performed at 60°C for 10 min, instead of the quick annealing and slow cooling of the dsDNA method. Quality of ssDNA was assessed by spectrophotometric analysis, where a value of 1.8-2.0 was expected for the A_{260}/A_{280} ratio.

2.3. Polyacrylamide gel electrophoresis

Reactions were electrophoresed at constant power (55W) using a polyacrylamide gel (6% acrylamide (19:1 acrylamide:bis-acrylamide), 50% Urea (w/v), 1 x TBE, 0.25% ammonium persulfate (AMPS), 0.1% TEMED) in 1 x TBE running buffer (89mM Tris, 89mM Boric acid, 2.5mM EDTA). BioRad sequencing gel kit was used. BioRad model 3000xi computer controlled electrophoresis power supply was used to maintain constant power.

2.3.1 Buffer gradient gels

A buffer gradient gel was used where long runs were required. This gel has a graded concentration of salt which increases towards the bottom of the gel which has the effect of creating a voltage gradient down the length of the gel. This slows the migration of low molecular weight fragments near the bottom of the gel. The bands are more evenly spaced on the final autoradiogram and thus more sequence can be read from one run (up to 600 nucleotides), as compression at the top of the gel is reduced, and resolution was improved.

1. Prepare and clean the gel kit (BioRad).
2. Mix two gel solutions:
   a) 75ml Light solution (0.5 x TBE, 6% acrylamide, 7.67M urea) and;
   b) 40ml Heavy solution (2.5 x TBE, 6% acrylamide, 7.67M urea, 10% sucrose, 1xbromophenol blue).
3. Pour the light and heavy solutions into separate beakers.
4. Add 10% AMPS and TEMED to achieve a final concentration of 0.25% AMPS and 0.1% TEMED to 20ml of the heavy solution to use for a casting plug. Once set, add AMPS and TEMED to both remaining solutions.
5. Take up 50 ml of the light solution into a 50 ml syringe and set aside.
6. Take up 12ml of light solution into a glass 25ml pipette. Then take up 12ml of the heavy solution into the same pipette. This should mix slightly on the interface to

form a gradient, if it does not, allow a couple of air bubbles to be drawn into the pipette.

7. Pour this gradient down the centre of the gel.
8. Pour the rest of the light solution into the gel, making sure that the flow is kept in the centre of the gel, so that the gradient forms evenly across the gel as well as vertically along its length. The bromophenol blue dye will show how well the gradient has formed.
9. The gel can be run in the usual way. The bromophenol blue dye will run out of the gel as electrophoresis proceeds, so the progress of the fragment dye front can be seen. The dark blue dye front should take at least 4 hr to run off the gel.

2.4 Automatic sequencing

Automatic sequencing was performed on an ABI 373 DNA sequencer using the PRISM kits provided according to the manufacturers instructions.

1. Combine 8µl of reaction mix with 0.5µg DNA, and 3.2pmole of primer. Make up to a volume of 20µl using SDW. Cover with mineral oil.
2. Place in a PCR machine perform 25 cycles of 96°C 30s, 50°C 15s, 60°C 4 min.
3. Transfer the reaction mixture into a clean tube containing 2µl 3M NaOAc pH8.4 and 50µl 95% EtOH. Place on ice for 10 min. Centrifuge at approx. 11,000 x g for 15 min to pellet the DNA.
4. Remove supernatant and rinse pellet in 70% EtOH. Vacuum dry and store at -20°C until loaded on the sequencer.
5. Sequence reactions were loaded and run by PNACL (University of Leicester) and chromatograms returned over the Macintosh computer network, and analysed using Gene Jockey II, or ABI PRISM Sequence Navigator, Version 1.01 (Perkin Elmer, Applied Biosystems Inc.).

3) Methods involving RNA

All work involving the preparation and handling of RNA was carried out with extreme care. In general, work was carried out on ice or at 4°C, and care was taken to wear clean gloves at all times. Glassware was not usually pre-treated with DEPC but was autoclaved or baked.

3.1 Extraction of RNA

3.1.1 Extraction of RNA from virus particles

1. Add 200µl of virus particle suspension to 200µl of isolation buffer (40mM Tris HCl pH 9.0 , 2mM EDTA, 2% SDS) in a sterile eppendorf.
2 Vortex briefly and incubate at 60°C for 5 min.

3. Allow to cool to room temperature and add 400µl of phenol (equilibrated using 0.5M Tris buffer to pH7.0). Vortex the sample and centrifuge at approx. 11,000 x g for 2 min.
4. Remove the upper aqueous layer carefully avoiding any precipitate at the interface and transfer to a clean eppendorf.
5. Repeat the extraction using 200µl of phenol and 200µl of chloroform. Vortex the sample and centrifuge at approx. 11,000 x g for 2 min.
6. Remove the upper aqueous layer and extract with 400µl of chloroform. Vortex briefly and spin for 2 min.
7. Transfer the upper aqueous layer to a clean eppendorf and add 40µl of 3M NaOAc (pH4.8) and 1ml of ice cold EtOH.
8. Place at -80°C for 10 min and then centrifuge the precipitated RNA at approx. 11,000 x g for 10 min.
9. Pour off the supernatant and dry the pellet under vacuum. Resuspend in 20 to 50µl of sterile distilled water containing 1µl of RNasin enzyme (RNase inhibitor, Promega).

3.1.2 Extraction of total RNA from frozen infected *N. clevelandii* and *N. tabacum* samples

1. Grind 1g of frozen leaf tissue in a pestle and motar with 10ml of buffer (10mM Tris HCl pH7.6, 5mM EDTA, 50mM NaCl, 0.5% SDS) and express through muslin into at 50ml sterile plastic tube.
2. Vortex for 30s and extract three times using an equal volume of phenol.
3. Precipitate the final aqueous layer with a tenth volume of 3M NaOAc pH 4.8 and two and a half volumes of ice cold ethanol.
4. Place at -80°C for 1 hr or overnight and centrifuge the precipitated RNA for 10 min at approx. 11,000 x g. Resuspend the pellet in up to a 1ml of sterile distilled water, depending on the size of the pellet.

3.1.3 Extraction of total RNA from fresh infected Narcissus samples

The method as described above in section 3.1.2 was used, but 0.1mg/ml of proteinase K was added to the buffer which did not contain SDS. Before grinding, the buffer and leaf tissue were left at room temperature for 10 minutes to allow the proteinase K to act. SDS was not included in the buffer during grinding because Narcissus leaves contain a high concentration of polysaccharides and other substances which cause the tissue to froth when ground. The SDS was added to the correct concentration after grinding to ensure that excessive frothing was avoided.

3.1.4 Extraction of total RNA from fresh infected *N. clevelandii* and *N. tabacum* samples

1. Grind 250-500 mg of sample to a fine powder in liquid nitrogen with a mortar and pestle.
2. Add 500µl of extraction buffer (25mM sodium citrate, pH 7.0, 4M Guanidinium isothiocyanate, 1.5% (w/v) sodium-lauryl sarcosine, 100mM β-mercaptoethanol) and grind in liquid nitrogen again to obtain an even mixture of frozen buffer and material.
3. Scrape into a 50ml sterile plastic tube and vortex until the mixture forms a solution. Transfer to a sterile eppendorf tube and add 500µl of phenol. Vortex for 15 s and centrifuge at approx. 11,000 x g for 5 min.
4. Transfer the upper aqueous phase and repeat the extraction twice more with 500µl phenol/chloroform and once with 500µl chloroform.
5. Transfer 400µl of aqueous phase to a fresh eppendorf and add 400µl of 6M LiCl, and incubate on ice for 1 hr. Centrifuge for 10 min at 4°C and remove the supernatant.
6. Disrupt the pellet with a sterile yellow pipette tip and add 1ml of 3M LiCl. Vortex to evenly resuspend the pellet and repeat the centrifugation. Remove the supernatant and repeat the wash.
7. Resuspend the pellet in 400µl 2% KoAc pH4.8 and heat to 55°C to dissolve the RNA. Centrifuge for 5 min to remove any remaining insoluble matter and transfer supernatant to a new tube.
8. Precipitate the RNA by adding 1ml 100% EtOH, incubating at -80°C for 15 min (or overnight at -20°C) and centrifuge for 15 min at approx. 11,000 x g. Dry the pellet under vacuum and resuspend in 50µl of sterile distilled water containing 1µl of RNasin.

3.2 Analysis of RNA

3.2.1 Formaldehyde gels

RNA was analysed on formaldehyde gels which are denaturing and therefore allow a more accurate representation of single-stranded RNA size.

1. Prepare electrophoresis tank by washing carefully in SDW. Soak gel tray and comb in 0.1M NaOH for 1 hr. Rinse well in SDW, handling all equipment with clean gloves at all times.
2. Make a 1% gel by adding 30 ml SDW to 0.5g agarose, melt using a microwave on full power. Allow to cool slightly.
3. In a fume hood, add 5ml 10 x MOPS buffer (See Appendix I) and 3.75ml filtered formaldehyde. Make up to 50 ml using SDW and pour into a small gel tray with comb which has already been taped at each end.

4. Place the set gel in an electrophoresis tank containing 1 x MOPS buffer. Prepare samples for loading.
5. Use 5-10µg of RNA in a volume not less than 8µl. Add an equal volume of formamide sample buffer (2.5xMOPS, 22% formamide, 10% formaldehyde) and incubate at 60°C for 5 min. Snap cool on ice. Add 3µl ficoll loading dye (100mM EDTA pH 7.5, 0.03% Ficoll type 400, 0.5% bromophenol blue, 0.5% xylene cyanol) and 1µl of 1mg/ml EtBr.
6. Load the gel and run at 75mA until the dye reaches the required position.
7. View the gel on a UV transilluminator.

3.2.2 Northern analysis

Formaldehyde gels were blotted and the filter probed. The procedure was similar to that in section 1.7 for Southern analysis.

1. Photograph the gel, including a fluorescent ruler for reference. Cut the top left corner from the gel for orientation. Wash the gel in SDW for 1 hr and then blot as described in section 1.7 (steps 4 & 5).
2. Incubate the filter at 42°C for 1 hr in a solution of 4 x SSC, 22% formamide, 50mM NaH_2PO_4, 5 x Denhardt's solution, 100ng/ml sheared single-stranded salmon sperm nucleic-acid.
3. Add the labelled DNA (section 1.5) and incubate at 42°C overnight. Wash as described in section 1.6.1 (steps 7-9). Expose to hyperfilm (Amersham) overnight at -80°C. Develop the autoradiograph and compare to the original photograph of the formaldehyde gel.

3.3 Reverse-Transcriptase PCR (RT-PCR)

RNA extracted from virus particles or total plant RNA was used in a RT-PCR reaction. First-strand cDNA was synthesised using the Promega cDNA riboclone kit according to manufacturers instructions. First-strand cDNA was used directly in a PCR reaction with the primer used to create the cDNA, and an upstream primer (or two internal primers). Usually 30 cycles of 94°C 1 min, 42°C 1 min, 72°C 2min were used. The annealing temperature was altered depending on the annealing temperature of the primer and its specificity to the target sequence.

4) Methods involving Protein

4.1 Protein extraction

4.1.1 Virus coat protein extraction

Isolation of viral coat protein was carried out according to the method described by Foster & Mills (1991e). Equal volumes of dissociation buffer (20mM NaH_2PO_4

pH7.0, 2% SDS (w/v), 0.4% dithiothreitol (v/v) and virus suspension were incubated together at 90°C for 2 min. Dissociated protein was either used immediately or stored at 4°C until required.

4.1.2 Protein extraction from plant tissue

Narcissus, *N. clevelandii* or *N. tabacum* leaf samples were collected by using a cork borer to cut two 1.5 cm diameter circles of tissue. These were ground in 3x protein loading buffer (0.125M Tris HCl, pH6.8, 5% glycerol, 1% SDS (w/v), 2% β-mercaptoethanol, 0.3µg/ml bromophenol blue) and incubated at 100°C for 2 min. Samples were centrifuged to pellet any tissue still present and the supernatant transferred to a clean tube. The preparation was either used immediately or stored at -20°C until required.

4.2 SDS-Polyacrylamide Gel Electrophoresis

SDS-Polyacrylamide Gel Electrophoresis (SDS-PAGE) was carried out as described by Laemmli (1970) using the BioRad Protean II gel kit. Gels were run in glycine buffer (0.1% SDS, 25mM TRIS, 200mM Glycine) at 100V through the stacking gel (375mM Tris HCl pH 8.8, 1% SDS, 12% acrylamide, 0.25% AMPS, 0.5% TEMED) and at 150V through the separating gel (125mM Tris HCl pH6.8, 1% SDS, 4% acrylamide, 0.05% AMPS, 0.5% TEMED). Wells were thoroughly rinsed with buffer before loading. The sample was mixed 1: 3 with 3x protein loading dye (0.185M TRIS pH 6.8, 6% SDS, 30% glycerol, 0.3mg/ml bromophenol blue and 2M β-mercaptoethanol) and boiled for 2 min prior to loading.

4.2.1 Coomassie staining

1. After electrophoresis (section 4.1) remove the gel from the kit and plates, and discard the stacking gel.
2. Place the gel in coomassie stain (5:5:1 water : methanol: acetic acid, 0.1% (w/v) coomassie blue) on a shaking table for 30 min.
3. Destain (50% methanol, 10% acetic acid) until the background colour of the gel is sufficient to visualise the stained protein bands.

4.2.2 Dry electro-blotting

1. After electrophoresis (section 4.1), remove the gel from the kit and plates, and discard the stacking gel
2. Soak the gel in cathode buffer (25mM Tris HCl pH 9.4, 40mM 6-amino-n-hexanoic acid, 20% methanol) for 15 min.
3. Place the gel on top of three sheets of Whatman paper cut to the size of the gel soaked in cathode buffer.

4. To prepare the membrane (PVDF, Amersham), soak first in methanol, then quench in SDW and finally equilibrate in cathode buffer. Place on top of the gel. Next layer one piece of Whatman paper soaked in anode 2 buffer (25mM Tris HCl pH 10.4, 10% methanol) and then finally two pieces of paper soaked in anode buffer 1 (300mM Tris HCl pH10.4, 10% methanol).
5. Place the whole stack in the electro-blotter, ensuring that the paper soaked in cathode buffer is in contact with the cathode electrode.
6. Blot at 150mA for 45 min.
7. Remove the filter and wash with Ponceau stain (0.2% Ponceau, 3% TCA) and rinse with SDW or 1xTBS (20mM Tris HCl pH7.6, 137mM NaCl) until the background colour is reduced sufficiently to visualise the stained protein bands.
8. If molecular weight protein markers were used, mark the position of each band of the markers on the filter with a pen or pencil.
9. Remove the rest of the stain by washing with water or 1 x TBS

4.2.3 Western Blotting

1. Block the filter by incubating for 1 hr (or overnight at 4°C) in 1xTBS, 5% (w/v) dried milk powder sealed in a plastic bag and placed on a rotating disc.
2. Remove the filter from the sealed bag and wash the filter for 10 min in 1xTBS-tween. Repeat the wash twice.
3. Incubate the filter in primary antibody appropriately diluted in 1xTBS, 3% (w/v) dried milk powder for 1 hr sealed in a plastic bag.
4. Wash as in step 2.
5. Incubate the filter in secondary antibody appropriately diluted in 1xTBS, 3% (w/v) dried milk powder for 1 hr sealed in a plastic bag.
6. Wash as in step 2.
7. Develop the filter by the addition of developing solution (3mg/ml nitro-blue tetrazolium (NBT: Sigma), 5mg/ml 5-Bromo-4-chloro-3-indoyl phosphate (BCIP: Sigma), 0.1M Tris HCl pH9.5, 1mM $MgCl_2$). Rinse in copious amounts of water once the result is visible. Air dry.

4.2.4 N-terminal micro-sequencing

N-terminal micro-sequencing was carried out on virus coat protein by PNACL (University of Leicester). Protein was electrophoresed and electro-blotted as described in section 4.2.1 and 4.2.2 except that rather than Ponceau staining of the blotted filter, coomassie staining was used instead.

Red la Soda virus coat protein was subjected to selective cleavage of the methionyl peptide in ribonuclease with cyanogen bromide (CNBr). This was carried out according

to method described by Gros & Witkop (1961) and analysed by the Cleveland gel system (Cleveland *et al.*, 1977), blotted and sequenced as before.

5) Plant Transformations

5.1 *Agrobacterium* transformations

5.1.1 Triparental matings

The relevant construct required for plant transformation was ligated into a Binary vector, such as pROK2.

1. Inoculate 10 ml LB (kanamycin selection, 50-100µg/ml) with a single transformed pROK2 colony and incubate overnight at 37°C. Inocluate 10 ml LB (kanamycin selection) with helper phage RK2013 and incubate overnight at 37°C. Inoculate 10 ml NB (rifamycin stock 50µg/ml for selection) with *Agrobacterium* LBA4044 and incubate for 48 hr at 30°C.
2. Mix 20µl of each culture in an eppendorf containing 100µl of LB with no selection.
3. Plate out onto NA plate without selection. Incubate for 2 days at 30°C.
4. Streak out onto an NA plate (See Appendix I) with rif/kan selection using an inoculating loop and incubate at 30°C for 2 days, or until single colonies are visible.

5.1.2 *Agrobacterium* total nucleic acid preparation

1. Inoculate 5 ml of kan/rif NB with a single *Agrobacterium* colony and incubate overnight at 30°C.
2. Transfer 1.5ml of culture to a sterile eppendorf tube and centrifuge at approx. 11,000 x g for 5 min to harvest cells.
3. Resuspend the pellet in 300µl of SDW. Add 100µl of 5% Sarcosyl and 150µl 5mg/ml Pronase. Incubate at 37°C for 1 hr.
4. Extract with an equal volume of phenol/chloroform three times and precipitate the nucleic acid with $^1/_{10}$ th vol 3M NaOAc and 2 $^1/_2$ vols 100% EtOH.
5. Centrifuge for 10 min, discard the supernatant and resuspend the pellet in 50µl of SDW.

5.1.3 Southern Analysis of *Agrobacterium*

In order to confirm that the transformed *Agrobacterium* colonies contained the correct insert, plasmid DNA was extracted (section 5.1.2) and digested (section 1.2.5) overnight at 37°C with enzymes which should excise the insert. Samples were analysed by electrophoresis as described in section 1.2.1 and the gel was photographed under UV illumination with a fluorescent ruler for reference. The gel was blotted and probed

as described in section 1.7 using the insert required for plant transformation as the probe.

5.2 Plant tissue culture

5.2.1 Plant transformation

Several types of tobacco were used for plant transformations, *N. tabacum* (SR1), *N. clevelandii*, and *N. benthamiana*. The same method was used for each with the exception that *N. benthamiana*. leaves were sterilised for less time (15 min) and a lower concentration of kanamycin (50µg/ml) was used for selection against *Agrobacterium*.

1. Use young tobacco leaves which are not fully expanded. Surface sterilise by placing in an autoclaved pyrex casserole dish containing 10% bleach for 30 min. Rinsé carefully three times in sterile distilled water.
2. Keep leaves wet while working in a flow hood. On a sterile tile, cut 1.5 cm squares of leaf and place on MSO plates (See Appendix I). Seal petri dishes and incubate at 25°C in a growth chamber for two days.
3. In a flow hood, remove the tissue samples from the plates and soak in fresh transformed *Agrobacterium* culture diluted 1/20 in MSO liquid for 10 min. Replace on the MSO plates and return to the growth chamber. Check the plates regularly for signs of *Agrobacterium* growth (usually 2-3 days).
4. Transfer tissue samples to MSD4x2 (See Appendix I) plates with kanamycin (100µg/ml) and augmentin (400µg/ml) selection . Return to the growth chamber and check regularly for shoots (usually 2-3 weeks).
5. Excise shoots and place in pots with MSO medium with kanamycin selection (100µg/ml). Seal pots and return to growth chamber, roots should appear within 10-12 days. Once shoots are large enough, pot up into soil.

5.2.2 Plant PCR

In order to confirm that plants have been transformed with the correct binary vector, PCR of total nucleic acid extracted from the transformants was used in PCR analysis using primers known to be present in the target sequence.

1. Collect two discs of plant material for each transformant tested using a cork borer. Snap freeze immediately in liquid nitrogen.
2. Grind using an eppendorf hand grinder in 250µl extraction buffer (100mM Tris HCl pH8.0, 50mM EDTA, 500mM NaCl, 10mM β-mercaptoethanol, 1.5% SDS). Vortex for 10 s and incubate at 65°C for 10 min.
3. Add 65µl 5M KoAc pH4.8 and incubate on ice for 10 min.
4. Centrifuge for 5 min and transfer supernatant to a clean tube. Add 320µl isopropanol and incubate at -20°C for 10 min.

5. Centrifuge for 5 min, discard supernatant and resuspend pellet in 100µl of SDW. Use 4µl in a 20µl PCR reaction (see section 1.2.3). For best results use AmpliTaq polymerase (Perkin Elmer).

6) Virus purification

6.1 Plant inoculation

Plants were manually inoculated by using infected sap or virus particles in suspension. Using a young plant, a leaf was selected for inoculation and marked with a piece of cotton tied loosely around the stem or by cutting a triangular section from the leaf. The leaf should be near the apex of the plant and not yet fully expanded. Carborundrum powder was sprinkled on the leaf. The virus preparation was diluted or infected tissue was ground in a little SDW and rubbed gently onto the leaf. After 15 min, the inoculated leaf was washed gently with water. Inoculated plants were placed in a growth chamber and inspected regularly for symptoms (usually 7-14 days). Infected leaves were harvested, frozen immediately in liquid nitrogen and stored at -80°C for future infections or analysis.

To confirm that plants were infected, western analysis was used with antiserum raised to the coat protein of the virus used for infection (section 4.2.3) or northern dot blots were performed (see below).

6.2 Northern dot blots

A small section of infected leaf tissue was ground in SDW and 5 to 10 µl was pipetted onto a gridded filter (Hybond-N, Amersham) which had been pre-soaked in 2xSSC. The filter was allowed to air dry and then it was cross-linked using the UV Stratalinker (Stratagene). The required probe was hybridised to the filter as described for Northern analysis (section 3.2.2).

6.3 NLV particle purification

Virus particles were extracted from Narcissus leaves which were known to be infected with NLV.

1. Grind in 5 vols of extraction buffer (See Appendix I), homogenise in a blender and filter the homogenate through muslin.
2. Add Triton X-100 to 2.5 % and stir at 4°C for 20 min.
3. Centrifuge at 800xg for 10 min. Transfer supernatant to a clean flask and add NaCl to 0.1M and PEG 6,000 to 4%. Stir for 45 min at 4°C.
4. Centrifuge at 800xg for 10 min and resuspend in borate buffer (See Appendix I). Stir for 45 min at 4°C.

5. Centrifuge at 800xg for 10 min, add 5 ml 30 % sucrose and re-centrifuge at 6,000xg for 2 hr and 30 min at 10°C to create a sucrose gradient.
6. Resuspend pellet in 1 ml borate buffer and add 3.5g caesium sulphate. Centrifuge at 12,000 x g for 16 hr at 10°C.
7. Remove the opalescent light scattering band from the caesium gradient, resuspend in borate buffer and check for virus particles under the electron microscope.

6.4 MacMV particle purification

Virus particles were extracted from *N. clevelandii* leaves which were known to be systemically infected with MacMV (20 days post-inoculation). This method described was the one used (Chapter 4) but is only part of the complete method as described by Plese, *et al.*, (1979).

1. Homogenise leaves in 1 volume of borate buffer (See Appendix I) and filter through muslin.
2. Add 0.4 vol chloroform and stir at 4°C for 1 hr.
3. Centrifuge at 6,000xg for 30 min at 10°C.
4. Remove aqueous layer and centrifuge at 10,000 x g for 15 min at 10°C.
5. Transfer the supernatant to a clean tube and re-centrifuge at 30,000xg for 90 min at 10°C.
6. Discard the supernatant and resuspend the pellet in 0.05M Borate buffer and check for virus particles under the electron microscope.

Chapter 3

The design and use of a carlavirus-specific PCR primer

Introduction

Carlaviruses

Often placed between the potexviruses and potyviruses due to a particle morphology and genome size intermediate between the two groups, carlaviruses are one of the largest yet least studied groups of plant viruses (Foster, 1992).

It is only within the last 6 -7 years that information on the organisation and expression of carlavirus genome has burgeoned. Sequence information now exists for 7 distinct carlaviruses and two strains of potato virus S. The complete genome of two carlaviruses, potato virus M (PVM: Rupasov *et al.*, 1989; Zavriev *et al.*, 1991) and blueberry scorch virus (BBScV: Cavileer, 1994) have been published, and there is partial sequence data available for several others: carnation latent virus (CLV: Meehan & Mills, 1991; Haylor *et al.*, 1990), *Helenium* virus S (*Hel*VS: Foster *et al.*, 1990a), lily symptomless virus (LSV: Memelink *et al.*, 1990), potato virus S, ordinary and Andean strains (PVS^O: Foster & Mills, 1992 and PVS^A: MacKenzie *et al.*, 1989; Foster *et al.*, 1990b), chrysanthemum virus B (CVB: Levay & Zavriev, 1991) and poplar mosaic virus (PopMV: Henderson *et al.*, 1992). This information has enabled areas of conservation to be identified within the carlavirus genus and between carlaviruses and other genera.

From sequence analysis it is clear that carlaviruses and potexviruses have a very similar genome organisation (Memelink *et al.*, 1990). The main difference between them being the additional 3' ORF present in all carlaviruses between the coat protein and the poly (A) tail. This cistron encodes a protein of approximately 11kDa. This protein contains a motif $C\text{-}X_2\text{-}C\text{-}X_{12}\text{-}C\text{-}X_4\text{-}C$ which is highly conserved among all carlaviruses. These four cysteine residues are arranged in a zinc-finger which conforms to a consensus sequence found in many nucleic acid binding proteins (Klug & Rhodes, 1987; Berg; 1986). It is this ORF that differentiates carlaviruses from potexviruses and any other group of plant viruses. The sequence of the 11K ORF was therefore an ideal candidate for the design of a PCR primer to rapidly detect carlaviruses and identify suspected members of the group.

Design of a carlavirus-specific PCR primer

When the nucleotide sequences representing the region from the AUG initiation codon of 11K ORF of the published carlavirus sequences are aligned, a striking region of homology at the 3' end can be seen (figure 3.1). A block of 20 nucleotides is conserved with only one variant base in 6 out of 8 carlaviruses sequenced. This unusual consensus has no known function and no set position in a coding or non-coding region. In some viruses it is present in the 11K ORF and includes its termination codon, in others it is in the 3' untranslated region between the 11K ORF and the poly (A) tail. However the distance from this block of homology to the poly (A) tail is remarkably constant, between 128 and 113 base pairs (bp). This block was therefore used to make a PCR primer, 'Carla-Uni', which contained the whole 20 nucleotides including the degeneracy at position 8 (G G A G T A A (C/T) G A G G T G A T A C C) (Badge *et al.*, 1996).

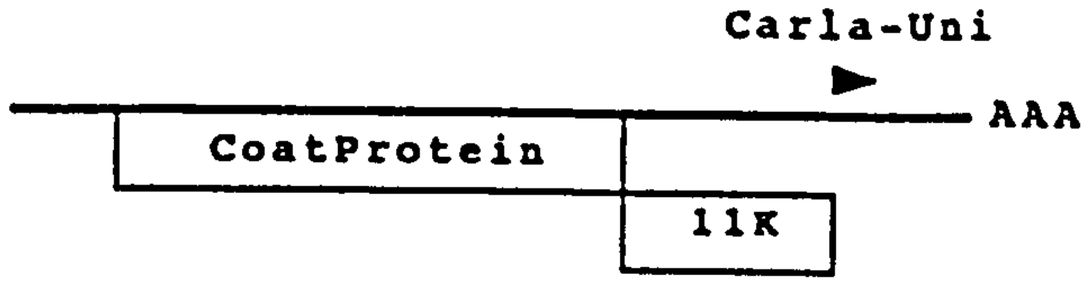
The Carla-Uni primer was used in conjunction with a variety of oligo-d(T) primers in an RT-PCR technique allowing for rapid testing of several known and suspected carlaviruses including NLV and MacMV.

Unusual carlaviruses to investigate

Conventional parameters used to classify viruses often prove to be inconclusive. Viruses are grouped into taxa by particle morphology, physiochemical characteristics, transmission vector and serology. By these criteria, a recognised member of the carlavirus genus has slightly flexuous filamentous particles in the size range 610-700nm by 13 nm, a coat protein sub-unit of 30-34 kDa (Wetter & Milne, 1981) and a plus-sense single-stranded RNA genome of *c.* 7-8 Kb. They are transmitted mechanically and in a non-persistent manner by aphid vectors and have a narrow host range with a tendency to induce little or no symptoms (Koenig, 1982). They generally have a good serological relationship with one another (Adams & Barabara, 1982). Carlaviruses utilise the sub-genomic method of translation (Foster, 1992) and do not usually produce inclusion bodies which are often a by-product of the polyprotein method of translation.

When new viruses are discovered they may often be tentatively classified on particle morphology until further data allows a more certain decision. However, this can lead to a biased investigation of further properties and many viruses have been misclassified in this way. Many viruses are only tentatively classified because there is only ambiguous information available. Those that are investigated here all have the type and size of particle structure within the range accepted for the carlavirus genus, but each has one or more features which do not fit into the norms expected of the genus. All have poor serological relationships with other recognised members of the genus and have

Figure 3.1: Position of the primer Carla-Uni on the carlavirus genome. The reported nucleotide sequences are aligned below in the enclosed box and show the homology in this region. The sequence distance from the start of the poly (A) tail is shown for chrysanthemum virus B (CVB), potato virus M (PVM), lily symptomless virus (LSV), *Helenium* virus S (*Hel*VS), potato virus S (PVS - this sequence is identical in both ordinary and Andean strains), blueberry scorch virus (BBScV - present in three strains of the virus) and cowpea mild mottle virus (CMMV, results presented here). Two published carlavirus sequences do not contain this sequence, poplar mosaic virus and carnation latent virus.



	5'																3'		PCR Fragment Size	
CVB	G	G	A	G	T	A	A	C	C	G	A	G	G	T	G	A	T	A	C	C	=	128	bp
PVM	G	G	A	G	T	A	A	C	C	G	A	G	G	T	G	A	T	A	C	C	=	116	bp
LSV	G	G	A	G	T	A	A	C	C	G	A	G	G	T	G	A	T	A	C	C	=	114	bp
HELVS	G	G	A	G	T	A	A	C	T	G	A	G	G	T	G	A	T	A	C	C	=	117	bp
PVS	G	G	A	G	T	A	A	C	T	G	A	G	G	T	G	A	T	A	C	C	=	113	bp
BScV	G	G	A	G	T	A	A	C	T	G	A	G	G	T	G	A	T	A	C	C	=	116	bp
CMMV	G	G	A	G	T	A	A	C	T	G	C	G	G	T	G	A	T	A	C	C	=	125	bp
Carla-Uni	G	G	A	G	T	A	A	C	C T	G	A	G	G	T	G	A	T	A	C	C	=	120	bp

therefore remained unclassified or under investigation. In cases such as these, when these conventional parameters do not lead to obvious conclusions, specific PCR primers and genome sequence data are proving to be decisive tools.

Cowpea mild mottle virus (CMMV)

Cowpea mild mottle virus (CMMV) causes either a mild systemic chlorotic mottle or, occasionally, a conspicuous chlorosis in its natural host, *Vigna unguiclata* (cowpea) (Brunt & Kenten, 1973). It appears to be a member of the carlavirus genus when analysed on the basis of its physiochemical properties. It has filamentous particles *c.* 650 nm in length consisting of a coat protein sub-unit of 31-33 kDa encapsidating a plus-sense single-stranded RNA genome (Jeyanandarajah & Brunt, 1993). However, it has two unusual features which cast doubt on its classification. Firstly, CMMV is transmitted in a non-persistent manner by whiteflies (*Bemisia tabaci*) (Jeyanandarajah & Brunt, 1993). Secondly, the presence of brush-like cytoplasmic inclusion bodies were reported in CMMV-infected tissue (Brunt *et al.*, 1983). This is usually a feature attributed to potyviruses, not carlaviruses. Cytoplasmic inclusion bodies are thought to occur as a by-product of the translation of potyviruses where a large polyprotein is cleaved by a viral protease to produce mature proteins thus forming a large excess of some proteins. Carlaviruses are thought to utilise the sub-genomic translation strategy, where shorter positive-sense RNAs are produced, presenting various internal ORFs at their 5' end, thus negating the need for excess viral protein production. The presence of cytoplasmic inclusion bodies in CMMV-infected tissue posed a serious problem for its classification as a carlavirus, and it was suggested that it remained unclassified.

Conventional methods to confirm that CMMV belonged to the carlavirus group failed when no serological relationship was found with 18 recognised members of the carlavirus group (Jeyanandarajah & Brunt, 1993). It was therefore an ideal candidate to be tested with our carlavirus-specific PCR primer.

Narcissus latent virus (NLV) and Maclura mosaic virus (MacMV)

Narcissus latent virus (NLV) was first described in 1966 (Brunt & Atkey, 1967) as a member of the carlavirus genus, again based on its physiochemical properties. However, it was serologically unrelated to 9 carlaviruses and 12 potyviruses (Mowat *et al.*, 1991). Like CMMV, NLV induces cylindrical cytoplasmic inclusions (CCIs) characteristic of potyviruses (Mowat *et al.*, 1991; Brunt *et al.*, 1994). Maclura mosaic virus (MacMV) was first reported to cause mosaic symptoms on the leaves of the ornamental tree *Maclura pomifera* (Plese & Milicic, 1973) and to induce CCIs (Plese & Wrischer, 1978). It also had the particle morphology of a carlavirus, but with no serological relationship to them.

Red la soda virus (RLSV)

A possible carlavirus was identified by the Scottish Agricultural Science Agency (SASA) in the potato cultivar Red La Soda imported from the USA. It had an unusually large host range for a carlavirus and a poor serological relationship with other members of the genus. The virus initially named red la soda virus (RLSV) was detected during routine testing in 1992 by the appearance of slight vein clearing in *Nicotiana debneyi* and local chlorotic spots in *Chenopodium murale* which had been inoculated with sap extract from the potato plant (Brattey *et al.*, 1996). The potato plant itself showed no symptoms but the presence of a virus was confirmed by immunosorbent electron microscopy using antisera raised against PVM and PVS. However, in further tests the purified RLSV failed to react to a wide range of antisera. Electron microscopy demonstrated that RLSV has filamentous particles of modal lengths 525nm and 650nm, within the range for a carlavirus or potexvirus. Further information in the form of sequence was essential to identify this new potato virus. The name 'potato latent virus' has now been officially proposed for this new virus, however for the purpose of this thesis it will be referred to as red la soda virus (RLSV).

Aims of the PCR test

In order to gain further information on all these viruses, and attempt to classify them where conventional methods had failed, the carlavirus-specific PCR primer, Carla-Uni was utilised. A positive result using the PCR primer would indicate that the virus was likely to be a carlavirus, and this could be quickly confirmed by cloning and sequencing the product.

Results

RT-PCR on RNA extracted from infected material

CMMV RNA was extracted from virus particles (Chapter 2: section 3.1.1). It was used to create first-strand cDNA (Chapter 2: section 3.3) by priming with either oligo-d(T)Not I, CN47, CN54 or CN55 (see below for sequences). This first-strand cDNA was included in an RT-PCR reaction with the primer Carla-Uni (figure 3.2 a,b).

Primer	Sequence 5' - 3'	Supplier/Reference
oligo-d(T)Not I	$CAATTCGCGGCCGC(T)_{15}$	Promega
CN47	$(T)_{21}A$	Pappu *et al.*, 1993
CN54	$(T)_{21}C$	Pappu *et al.*, 1993
CN55	$(T)_{21}G$	Pappu *et al.*, 1993

Carla-Uni was tested against a wide range of viruses (Chapter 2: section 3.3; PCR at 94°C, 30s; 50°C, 1min; 72°C, 2 mins; 35 cycles) in collaboration with Dr. S. Seal (NRI). The PCR primer Carla-Uni was tested to ensure that a genuine product was

created by reactions containing one or other of the two primers used, no primers, or both primers and no cDNA. No product was obtained in any of these reactions (figure 3.2b). The primer was also tested on a potexvirus (potato virus X, PVX) and a potyvirus (potato virus Y, PVY) to show that it was specific to carlaviruses (figure 3.2a). No products were obtained, but in parallel reactions, both viruses gave products of the expected sizes with primers designed from the sequence of each virus.

Several known and suspected carlaviruses were tested with Carla-Uni. Products of the expected size (*c.* 120bp) were obtained from the following known carlaviruses : potato virus S (PVS), potato virus M (PVM) and lily symptomless virus (LSV). Similarly products were obtained from the following suspected carlaviruses: American hop latent virus (AHLV), hop mosaic virus (HMV) and CMMV (figure 3.2a). No product was obtained with either MacMV (source: A. Brunt) or NLV (source: A.Brunt) RNA, even in the presence of an effective positive control (figure 3.2a).

Cowpea mild mottle virus

Northern and dot blot analysis

The PCR product obtained for CMMV was tested by RNA dot blots (Chapter 2: section 3.2.2). The CMMV fragment gave a positive result when radio-labelled and probed onto CMMV RNA, but not with several other virus RNAs (figure 3.3).

Sequencing

In order to confirm that Carla-Uni primer was working correctly, conventional ds-cDNA clones representing the 3'-terminal region of CMMV were generated and sequenced (Chapter 2: section 2). Double-stranded cDNA was synthesised from viral RNA and after the addition of *Bst X-I* linkers, cloned into *Bst X-I* cut pcDNAII (Invitrogen). Colonies were screened with radio-labelled first-strand cDNA to CMMV RNA, with positives being sized and sequenced. A clone with a 500bp insert was isolated that encoded the 11K ORF, and a primer was made to the 5' region of this clone in order to extend the sequence into the coat protein region. A second clone was thus isolated, extending the viral region sequenced to 958bp.

The nucleotide sequence obtained and its translation are shown in figure 3.4. The position of Carl-Uni is underlined (832-852 nucleotides), and is such that PCR would produce the expected and observed 125bp fragment. Only one base is different from the Carla-Uni consensus sequence; C has been substituted for A at position 843 nucleotides (nt). The primer sequence includes the termination codon for the 11K ORF. Analysis showed that a partial ORF encoding a 19.5 kDa protein corresponded to the coat protein and that a second ORF, with initiation codon at position 531 nt

Figure 3.2 (a): Agarose gel analysis of PCR products generated using Carla-Uni and anchored oligo-d(T)Not I primers on first-strand cDNA generated from RNA isolated from plant material infected with a range of viruses (known carlaviruses, possible carlaviruses and viruses of other genera with filamentous particles), as outlined below. Lane A, molecular weight markers (Superladder; Advanced Biotechnologies); lane B, molecular weight markers (kilobase ladder; Gibco BRL). Lanes 1-12 were PCR reactions performed with Carla-Uni and anchored oligo-d(T)Not I primers CN47, CN54 and CN55 (Pappu, *et al.*, 1993). Closed arrow indicates the predicted PCR product of approximately 120bp which should be obtained with Carla-Uni and oligo-d(T)Not I. Lane 1, potato virus S; lane 2, potato virus M; lane 3, lily symptomless virus; lane 4, American hop latent virus; lane 5, hop mosaic virus; lane 6, cowpea mild mottle virus; lane 7, narcissus latent virus; lane 8, Maclura mosaic virus; lane 9, potato virus X; lane 10, potato virus Y; lane 11, negative control (water). Lane 12, potato virus X positive control PCR with PVX specific primers. Lane 13, potato virus Y positive control PCR with CN48 potyvirus primer and CN47, CN54 and CN55.

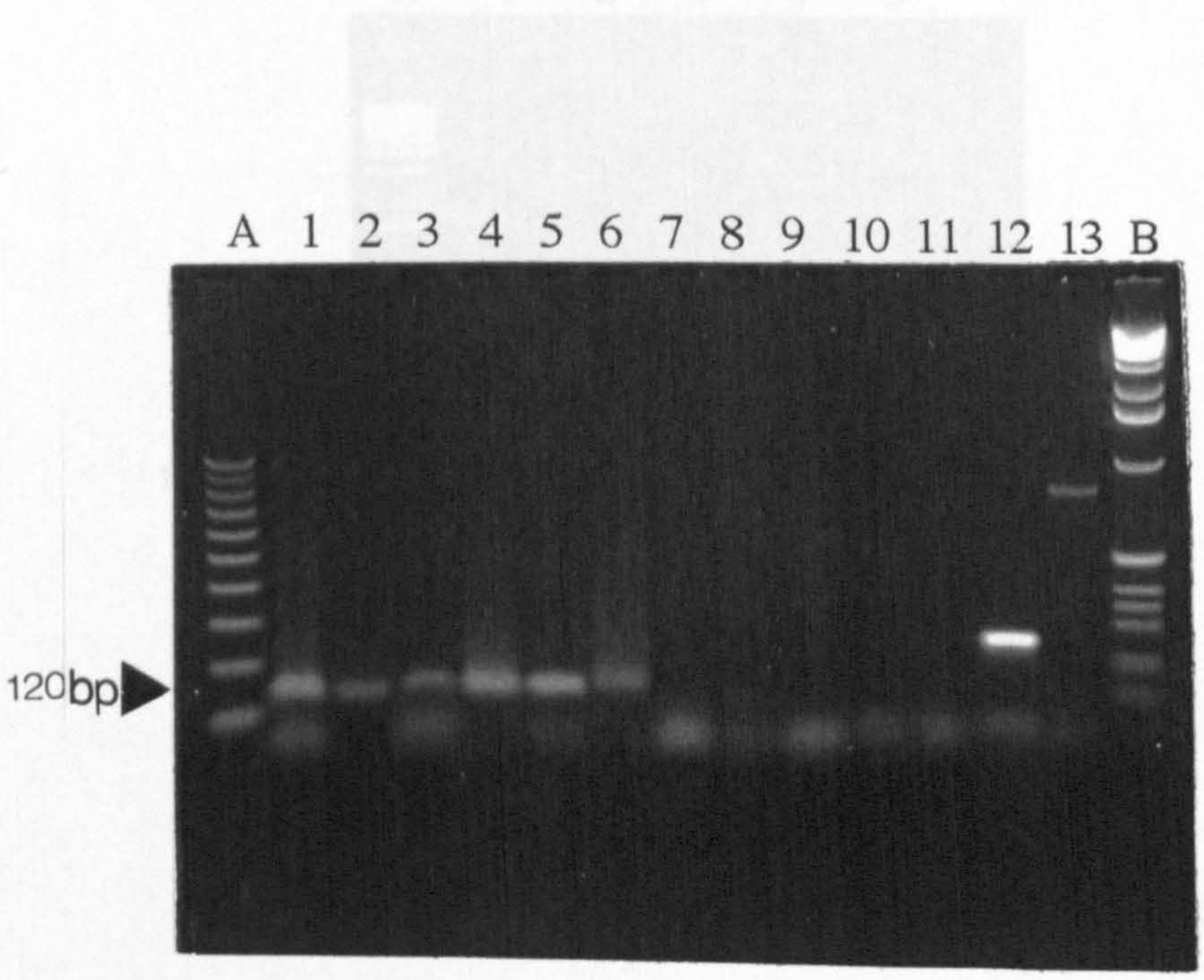


Figure 3.2 (b): Agarose gel analysis of PCR products and control PCR reactions generated from first strand cDNA primed with oligo-d(T)Not I generated from RNA extracted from purified particles of CMMV. Lane A - molecular weight markers (kilobase ladder; Gibco BRL). Lane 1, PCR reaction containing cDNA and both primers Carla-Uni and oligo-d(T)-Not I; lane 2, PCR reaction containing cDNA and oligo-d(T)-Not I alone; lane 3, PCR reaction containing cDNA and Carla-Uni alone; lane 4, PCR reaction containing both primers both no cDNA; lane 5, PCR reaction containing cDNA but no primers. Closed arrow indicates the PCR product of approximately 120bp which should be obtained with Carla-Uni and oligo-d(T)-Not I.

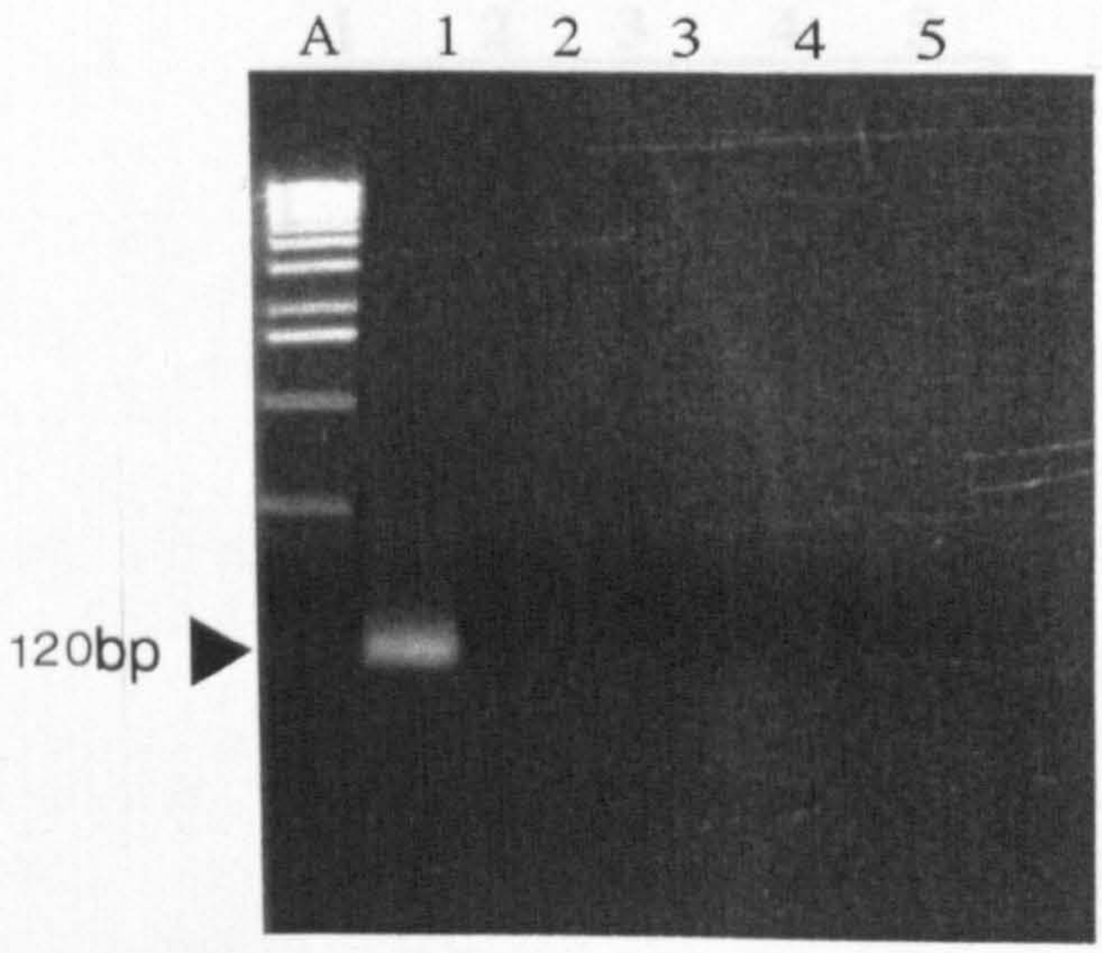


Figure 3.3: RNA dot-blot analysis using radiolabelled 120bp PCR fragment generated from CMMV as probe. Sample 1, Maclura mosaic virus RNA; sample 2, cowpea mild mottle virus RNA; sample 3, narcissus latent virus RNA; sample 4, brome mosaic virus RNA (BMV); sample 5, tobacco mosaic virus RNA (TMV).

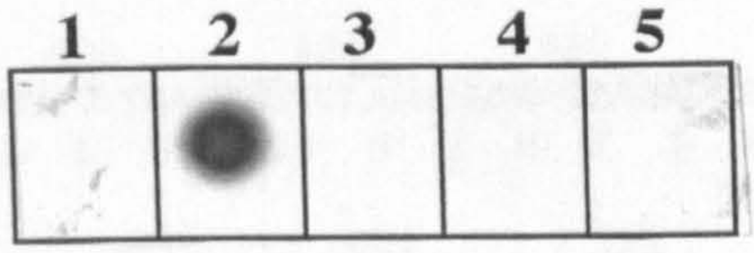


Figure 3.4: The DNA sequence corresponding to the 3'-terminal region of CMMV positive-sense RNA. Derived amino acid sequence indicated below with stop codons underlined. A partial ORF of approx. 19.5 kDa is evident followed by an ORF encoding a cysteine-rich protein of 11.7kDa. The position of Carla-Uni is underlined (832-852 nt)

```
          10         20         30         40         50         60         70
GTGCAAATCACTTTAGAAGGTCTTGGAGTCCCAACAGAACATGTGGCTGAAGTGCTGCTACAAGTTGCCATTTACTGC
V  Q  I  T  L  E  G  L  G  V  P  T  E  H  V  A  E  V  L  L  Q  V  A  I  Y  C

 80         90         100        110        120        130        140        150
AAGGATCTAAGCTCTTCGAGCTTTATGGATTCAAGTGGCACCTTCGATTGGAAGGGCGGATCTATACTATCTGATTCA
K  D  L  S  S  S  S  F  M  D  S  S  G  T  F  D  W  K  G  G  S  I  L  S  D  S

   160        170        180        190        200        210        220        230
GTCTTAGCAGCCTTACGCAAGGATGACAACACCCTGAGGAGGGTGTGCCGCCTTTATGCTCCCATTACTTGGAACTTC
V  L  A  A  L  R  K  D  D  N  T  L  R  R  V  C  R  L  Y  A  P  I  T  W  N  F

     240        250        260        270        280        290        300        310
ATGTTAACTCATAAAGCCCCACCTTCTGATTGGGCCGCAATGGGCTTCAAATGGGACGATCGATATGCTGCTTTCGAT
M  L  T  H  K  A  P  P  S  D  W  A  A  M  G  F  K  W  D  D  R  Y  A  A  F  D

          320        330        340        350        360        370        380        390
TGTTTCGAATATGTTGAGAATCCAGCTGCCATTCAACCAGCTGAAGGATTAATAAGGAAGCCGACCTCAAGTGAGAAG
C  F  E  Y  V  E  N  P  A  A  I  Q  P  A  E  G  L  I  R  K  P  T  S  S  E  K

          400        410        420        430        440        450        460
ATCGCCCACAATACATACAAGGGTCTAGCTCTTGATCGATCAAACAGGAATAAGATCTACTCCAATCTGAACACTGAA
I  A  H  N  T  Y  K  G  L  A  L  D  R  S  N  R  N  K  I  Y  S  N  L  N  T  E

 470        480        490        500        510        520        530        540
GTCACAGGCGGCACTTTAGGCCCAGAGATTTCTAGAAATTTCAATCACACCAAGAAGTGACCATGATTGGGTACAAAC
V  T  G  G  T  L  G  P  E  I  S  R  N  F  N  H  T  K  K  •
                                                        M  I  G  Y  K  R

   550        560        570        580        590        600        610        620
GCATTGCTATCTTGCTTCACTTATGTTGTTTAAAGATGAATAAAATGTTGCCTTTAGATTTGTGTATTTTAATTGCCC
  I  A  I  L  L  H  L  C  C  L  K  M  N  K  M  L  P  L  D  L  C  I  L  I  A  L

     630        640        650        660        670        680        690        700
TTAAGGCAGGGCCAACTAGTCTTTATTCTGGATCTTCAAGTTACGCACGACGTAGACGTGCTAAGTTGATAGGTCGAT
  K  A  G  P  T  S  L  Y  S  G  S  S  S  Y  A  R  R  R  R  A  K  L  I  G  R  C

          710        720        730        740        750        760        770        780
GTCATCGTTGTTATAGAGTCAGTCCTGGCTTTTACTTTACTACTAGGTGTGATGGTTTGTCTTGTGTGCCTGGTATAA
  H  R  C  Y  R  V  S  P  G  F  Y  F  T  T  R  C  D  G  L  S  C  V  P  G  I  S

          790        800        810        820        830        840        850
GCTACAAACAATGGGTTGAGTCTTTCATTAGATTTGGTCACTTTACAAGGGACGGAGTAACTGCGGTGATACCCCCTTT
  Y  K  Q  W  V  E  S  F  I  R  F  G  H  F  T  R  E  E  •

   860        870        880        890        900        910        920        930
AAACTCTTGAACCAAGGAGAGAGTATAAAGAGTCCTTGTGTAGCTTCAGACCTAACTGTTGAAGTTTGTACCGGTTTT

     940        950
AAAGTTATTTTCCTGGTTTTTA(n)
```

encoded a cysteine rich protein of 11.7 kDa. A 3' untranslated region of 119 nt is present between the 11K ORF and the poly (A) tail.

Amino acid sequence analysis

Figure 3.5 shows an alignment of the partial coat protein (figure 3.5a) and 11K ORFs (figure 3.5b) of CMMV with several recognised carlaviruses.

Coat protein sizes

Virus coat protein was obtained using the method as described by Foster & Mills, (1991e) (Chapter 2: section 4.1.1) and analysed by SDS-PAGE (Chapter 2: section 4.2), and stained with Coomassie blue (figure 3.6). CMMV coat protein appeared as a doublet. The larger of the two proteins was sized at *c.* 24 kDa (figure 3.6) when compared to SDS-7 molecular weight markers (Sigma).

Virus particle measurements

Measurements of particles from a purified virus preparation were obtained from electron micrographs of CMMV. A total of 300 particles were measured and a running mean used to construct the graph shown. The modal length of particles was 650 nm (figure 3.7). Shorter particles (130 nm) which were not an exact fraction of full length particles, and therefore unlikely to be broken or damaged particles were also detected at a high frequency.

<u>red la soda virus, RLSV</u>

In January 1996, a sample of red la soda virus (RLSV) was received from Dr C. Brattey (SASA). Suspected of belonging to the carlavirus genus, it was tested using the Carla-Uni primer. RT-PCR using oligo d(T)Not I and Carla-Uni on oligo-d(T)Not I primed first-strand RLSV cDNA resulted in a 800bp product (figure 3.8), much larger than the expected size. The product was cloned (Chapter 2: section 1.3.1.) after the reaction was shown to be reproducible and no bands were observed in control reactions. The RLSV fragment gave a positive result when probed onto viral RNA and infected, but not healthy sap from potato and *N. debneyi.* (figure 3.9).

Figure 3.10 shows the sequence obtained from the two independent clones, containing the insert of the RT-PCR product, pRLSV1 and pRLSV3. Both clones were presented in opposite orientations, and thus subcloning allowed the sequencing of opposite strands of different clones. Hence, 86% of sequence was obtained from complementary strands of two clones, and 14% on both strands of one clone. A total of 877 nt was sequenced which represented the 3' terminal portion of RLSV. Analysis showed that two consecutive open reading frames (ORF) were present. The first was a

Figure 3.5 (a): Alignment of the partial ORF of CMMV and RLSV with the C-terminal regions of the coat protein genes of BBScV, CLV, CVB, *Hel*VS, LSV, PVS and PVM. Gaps (-) have been introduced for maximum alignment, with identical amino acids in all virus sequences indicated by a (*), chemically similar residues in all sequences indicated by (|).

```
                    •••••     |        •     |• | •| •  |   •||   • | | |
BBScV       ISSAISGLGVPTEEVAGIIVRMAFMCASVSSSVYLDPDGSIEFEGGAVS-VDSIAAIMKRE
CLV         ITAKLEGLGVPTERSAPICWAVARYCANTSTSPYTDPKGVFEFPGGAITR-DAVFAVIREV
CVB         IQVALEGLGVPTESVKEVIIRLVLNCANTSSSVYQDPKGVIEWDGGAII-ADDVVGVITKH
HelVS       IKVALEGLGVPTEHITGIIYQMCFYCASTSSSSYQDPKGVFEWPGGAIM-VDDVMGKVQEI
LSV         IASDIAGLGVPTEHVASVILQMVIMCACVSSSAFLDPEGSIEFENGAVP-VDSIAAIMKKH
PVM         IYVNLEGLGVPTEHVQQVVIQAVLFCKDASSSVFLDPRGSFEWPRGAITA-DAVLAVLKKD
PVS         ITADIAGLGVPTEHVAGVILKVVIMCASVDSSVYLDPAGTVEFPTGAVPTLDSIIAIMKNR
CMMV        VQITLEGLGVPTEHVAEVLLQVAIYCKDLSSS-FMDSSGTFDWKGGSILS-DSVLAALRDD
RLSV cp     SNR-----GVPTEQVANVVLQAVIYCASASSSVYLDPHGTIEYTGGAVVR-DSVVAIIKRD

                  |•  ••   |•|| •• •   | •• |• | •|  | | •••• ||• • ••||•  •
BEScV       ASLRKVCR--LYAPVVWNFMLVKNQPPSDWQAMGFSFETRFAAFDTSDYVTNEAAVQPVEG
CLV         TTLRAFCRRDAFAPITWNQMLFAKSPPENWQAKGYTYETRYAAFDVFDFVQNPAAIQPLEG
CVB         STLRKVCR--LYAAVAWNYMHLQQTPPSDWSAMGFHPNVKYAAFDFFDYVENGAAIRPSGG
HelVS       AGIRRVCR--LYAPVTWNYMHIHDSPPSDWASMGFAPNVKYAAFDCFDYVENPAAVQPLGG
LSV         AGLRKVCR--LYAPIVWNSMLVRNQPPADWQAMGFQYNTRFAAFDTFDYVTNQAAIQPVEG
PVM         A-LRRVCR--LYAPVTWNHMLTHNAPPADWAAMGFQYEDRFAAFDCFDYVENTAAVQPLEG
PVS         AGLRKVCR--LYAPVVWNYMLVQNRPPSDWQAMGFQWNARCAAFDTFDYVTNGAAVQPVEG
CMMV        NTLRRVCR--LYAPITWNFMLTHKAPPSDWAAMGFK-DDRYAAFDCFEYVENPAAIQPAEG
RLSV cp     AGLRRVCR--LFAPLVWNHMLVHNSPPSDWAAMGFQWNDRFAAFDFFDYVENEAAIQPLDG

                  ||  ••  • |•  |         |  |  |  | |     •••  •  |  |
BBScV       LIRRPTPAECIAHNAH---KRLALDRSNRNEKYTNLETEYTGGLQGAEIVRNHRNAGNGSA
CLV         LLRIPTAEEKIAHAT---NKRLALDRNRRNARFSSTDSLVTGGMYGKDIKTNF-NGSNNSD
CVB         IVPKPTRAEYVAYNTY---KMLALNKANNNDTFGNFDSAITGGRQGPAIHNNLNNANNKTL
HelVS       VIPRPTRDEYVAYNAY---KLIVLNKANNNDTYGNFSAQITGGRMGPTIEHNFNNANNKKQ
LSV         IIRRPTSAEVIAHNAH---KQLALDRSNRNERLGSLETEYTGGVQGAEIVRNHRYANNG
PVM         LIRRPTPREKVAHNTH---KDIAVRGANRNQVFSSLNAEVTGGMNGPELTRDY--VKSNRK
PVS         LIRRPTPEETIAHNAHNAHKSMAIDKSNRNERLANTNVEYTGGMLGAEIVRNHRNAIN--Q
CMMV        LIRKPTSSEKIAHNTYKG----ALDRSNRNKIYSNLNTEVTGGTLGPEISRNFNHT---KK
RLSV cp     LIRRPTRSEKIAHNTH---KRLALDKSNRDEVFASLETEITGGKRGPEISRNFRNAAN
```

Figure 3.5 (b): Alignment of the predicted amino acid sequences encoded by the 3' 11K ORF of CMMV and RLSV with the equivalent ORFs of BBScV, CVB, *Hel*VS, PVS and PVM. Gaps (-) have been introduced for maximum alignment, with identical amino acids in all virus sequences indicated by a (*), chemically similar residues in all sequences indicated by (|).

```
BBScV   MLTKGMRWIDPIAMRSTGISKRSTLVGFKVQKSSAIIEMRAMDQHDQRRLLALIKVCRSF----NLYDV
CVB     MD-------------------------VIVK---MLIL-----------------RKFVEQGNVCPI
HelVS   MD-------------------------KRNKANVVLSL----------------CSMFASRGNCIPI
PVM     M--------------------------KDVTKVALLIARAM-------------CAS--SGTFVFEL
PVS     M--KADR-------------------------------------------------LAMLLLCV--HRLGYVLPV
RLSV    M--------QPIS-------------------------------HYEAKLLAVSLAMYKFTGR---CEP
CMMV    MI-GYKR-------------------------------------IAILLHLCCLKM---NKMLPL

BBScV   GVCIHILNKVPPREVGNGCSSYAAKRRAKSIGRCERCYRVFPI----GASKCYNRTCVPGISYNEKVAN
CVB     HLCVDIYKRAFPRSVNKGRSSYARRRRALELGRCHRCYRVYPP-LFPEISRCDNRTCVPGISYNSKVRD
HelVS   PIVFNIYMRAFPKLVGRGTSTYARRRRARSILRCERCYRVYPP-L-PFSKKCDNRTCVPGISYNIKVAD
PVM     AFSITEYT---GRPLGGGRSKYARRRRAISIARCHRCYRLWPPTVF--TTRCDNKHCVPGISYNVRVAQ
PVS     EVCVNIIS-LSAGPVSGGRSTYARKRRARSIGRCWRCYRVYPP-I--CNSKCDNRTCRPGISQNYKVVT
RLSV    AVALNIVNKAC--NVGMGKSSFARRARAALLGRCHRCFRT------SMATRCNGVTCYPGIGAKPKIEM
CMMV    DLCILIALKAGPTSLYSGSSSYARRRRAKLIGRCHRCYRVSP--GFYFTTRCDGLSCVPGISYKQWVES

BBScV   FIRCGVTEVIPHP-GFNF
CVB     YILWGVTEVIPHP-GYNF
HelVS   FIKWGVTEVIPHP-GFNF
PVM     FIDEGVTEVIPSVINKRE
PVS     FIR----------GWSN
RLSV    FIKYGVSELKP
CMMV    FIRFGHFTREE
```

Figure 3.6: Coomassie stained SDS-PAGE of CMMV coat protein isolated form virus particles. Lane 1, SDS-7 protein molecular weight markers (Sigma); Lanes 2 & 3, CMMV coat protein. Sizes of markers are indicated.

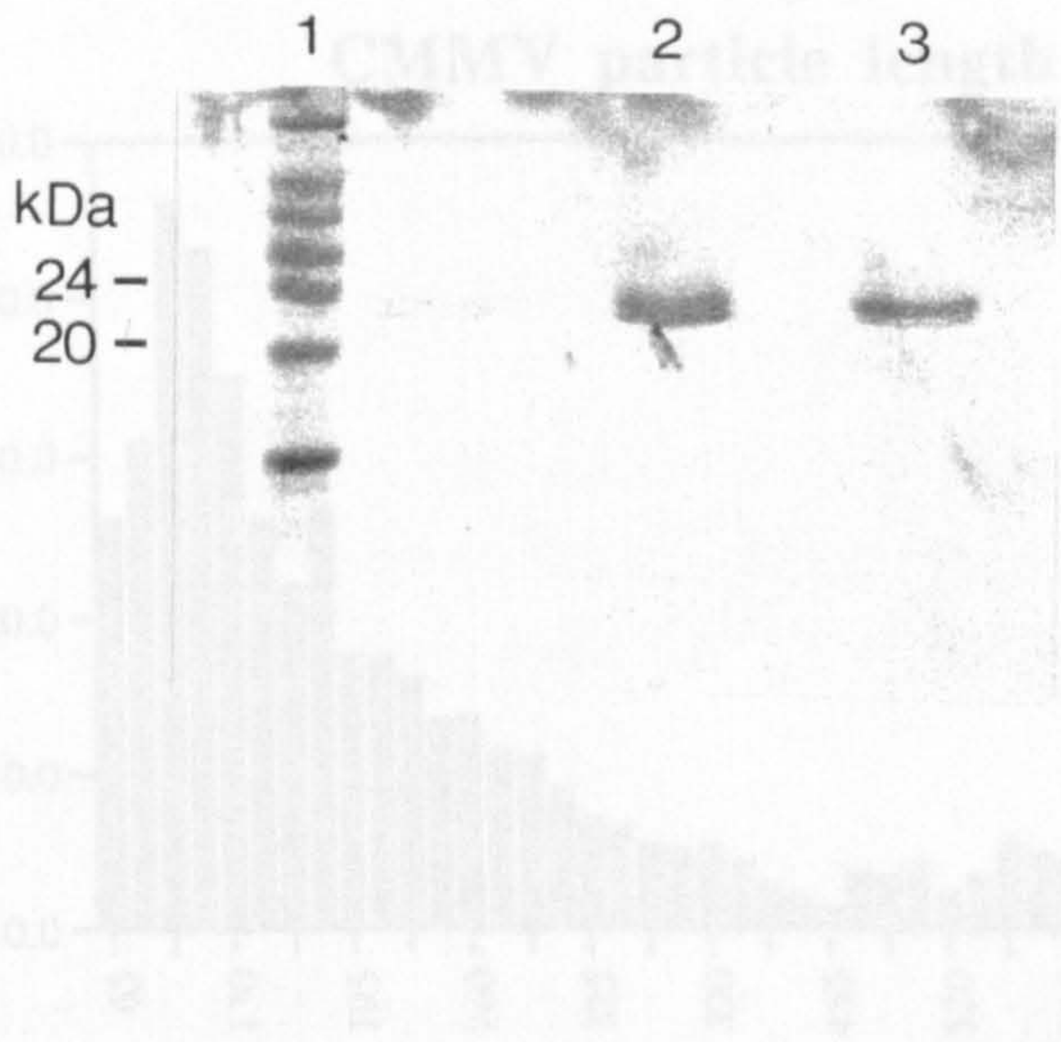
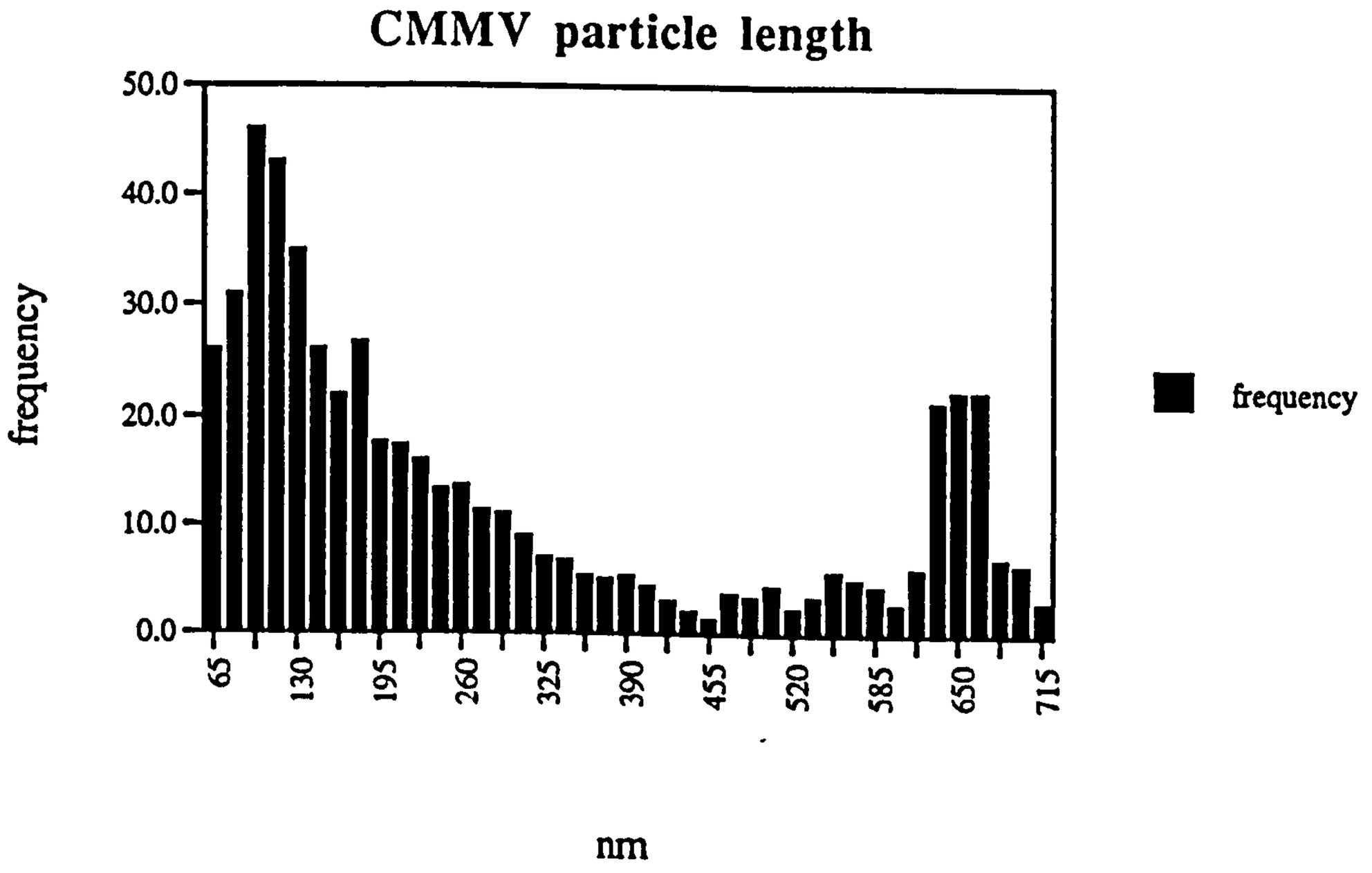


Figure 3.7: Graph to show the frequency distribution of particle size of CMMV. Two peaks, one at 650 nm is the complete particle, the second at 130 nm is the proposed encapsidated coat protein sub-genomic RNA.



partial ORF of 169 amino acids (aa) with a termination codon at 510 nt and the second was 99aa from a methionine start codon (499nt). A short untranslated region of 81nt was present before the poly (A) tail.

The primer Carla-Uni was present, as expected, in the first 19 nucleotides of the insert, with one nucleotide missing from the primer sequence (A missing at position 16). However, a second more degenerate sequence of Carla-Uni is present at position 775-794 nt (figure 3.10). This is just within the 11K ORF and is degenerate at 4 positions.

Amino acid sequences obtained were aligned with other known carlaviruses sequences for the partial coat protein and 11K ORF (figure 3.5 a,b).

RLSV coat protein was sized at 32 kDa, the largest of the several proteins observed on the gel (figure 3.11). The complete protein was found to be N-terminally blocked, so a CNBr digest was performed, and a 7.5kDa product was sequenced (Chapter 2: section 4.2.4). This successfully produced 30 residues (figure 3.10) which were found to be present in the deduced amino acid sequence.

Particle measurements of RLSV were carried out at SASA from crude sap extractions of infected material, and results are displayed for completeness (figure 3.12). The modal lengths of particles were 525 nm and 650 nm.

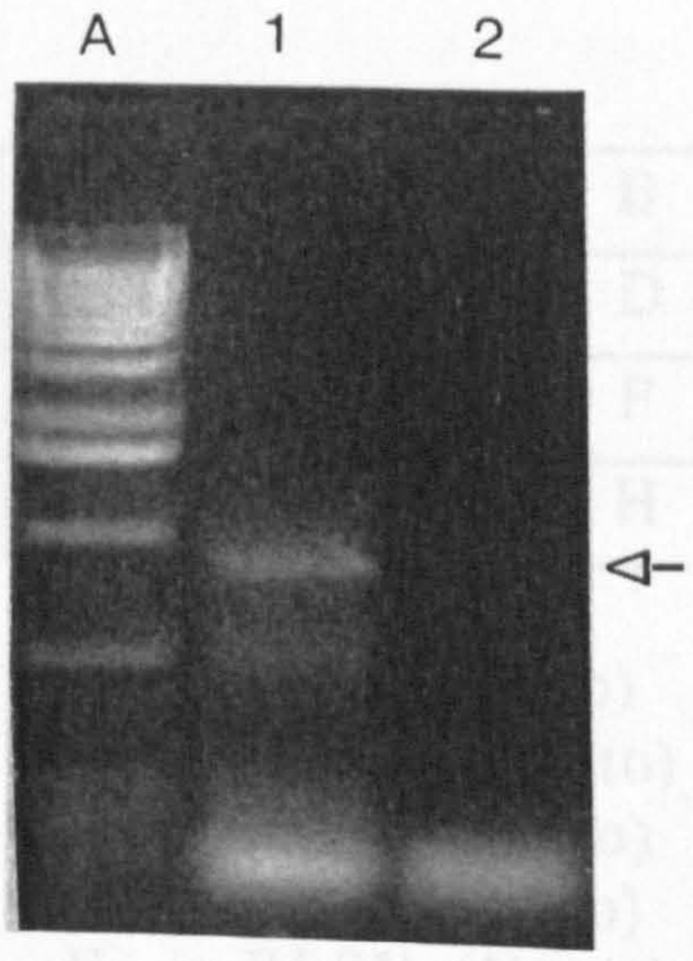
Discussion

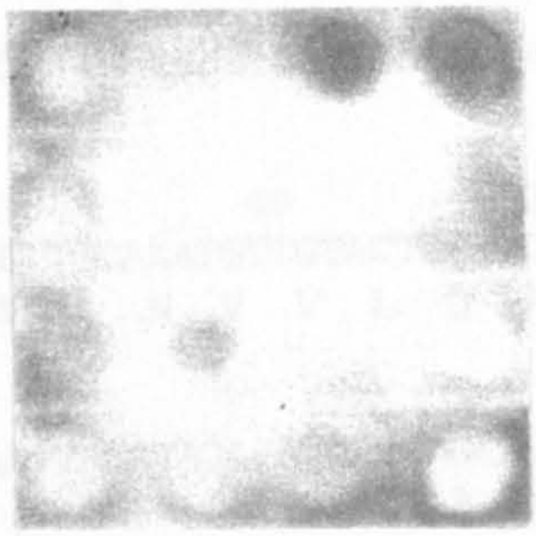
Primer design

A carlavirus-specific PCR primer, Carla-Uni, was designed in order to rapidly test viruses suspected to belong to the genus. The primer was specific to carlaviruses, not binding to a closely related potexvirus or a typical potyvirus. It was effective on recognised carlaviruses, PVS, PVM and LSV and demonstrated that two viruses, AHLV and HMV belong to the carlavirus genus, confirming the evidence gained from particle morphology and serology.

Negative results were obtained with NLV and MacMV, both of uncertain classification, but with particle morphology within the accepted range for carlaviruses. Although this result was carried out at the same time as a positive control, the lack of product cannot conclusively determine that MacMV and NLV are not carlaviruses. It is possible that the PCR reaction was disrupted by the presence of *Taq* polymerase inhibitors. The target sequence may not be present, be degenerate, or the viruses may be deleted mutants of a wider population. Further investigation of these two viruses by

Figure 3.8: Agarose gel analysis of PCR products and control PCR reactions generated from first strand cDNA primed with oligo-d(T) generated from RNA extracted from purified particles of RLSV. Lane A - molecular weight markers (kilobase ladder; Gibco BRL). Lane 1, RLSV; Lane 2, negative control - no cDNA.





A	A	B	B
C	C	D	D
E	E	F	F
G	G	H	H

A PVS (potato)
B RLSV (potato)
C PVV (potato)
D PVY (potato)
E RLSV (*N. debneyei*)
F PVM (potato)
G Healthy *N. debneyei*
H. Healthy potato

Figure 3.9: Sap dot blots using radio-labelled 800bp PCR fragment generated from RLSV as a probe. The filter was prepared by SASA for probing in Leicester. The grid shows the position of the various samples used, detailed underneath.

Figure 3.10: The DNA sequence corresponding to the 3' terminal region of RLSV positive-sense RNA. Derived amino acid sequence is shown below with stop codons underlined. A partial ORF of *c*. 18.8 kDa is evident followed by an ORF encoding a cysteine-rich protein of 10.8 kDa. The amino acid sequence underlined was gained from N-terminal micro-sequencing of CNBr digested RLSV coat protein. The primer sequence for Carla-Uni is present at the beginning of the clone, underlined (1-19 nt), one nucleotide (A) is missing from position 16. A second site where a degenerate form of Carla-Uni is present is at 775-794 nt, also underlined.

```
                10                  20                  30                  40                  50                  60                  70
GGAGTAACCGAGGTGTACCCACTGAACAAGTTGCAAATGTTGTACTGCAGGCGGTGATTTACTGTGCTAGTGCAAGTA
  S   N   R   G   V   P   T   E   Q   V   A   N   V   V   L   Q   A   V   I   Y   C   A   S   A   S   S

 80                  90                 100                 110                 120                 130                 140                 150
GCTCCGTGTATCTGGATCCACATGGTACAATTGAGTACACAGGGGGAGCAGTCGTTCGTGATTCTGTTGTAGCCATCA
  S   V   Y   L   D   P   H   G   T   I   E   Y   T   G   G   A   V   V   R   D   S   V   V   A   I   I

  160                 170                 180                 190                 200                 210                 220                 230
TCAAAAGGGATGCCGGGTTGAGGAGGGTGTGTCGGCTTTTTGCGCCGTTGGTCTGGAATCACATGTTGGTGCACAACT
  K   R   D   A   G   L   R   R   V   C   R   L   F   A   P   L   V   W   N   H   M   L   V   H   N   S

  240                 250                 260                 270                 280                 290                 300                 310
CACCTCCCTCTGATTGGGCTGCTATGGGCTTCCAATGGAATGACAGATTTGCAGCCTTCGACTTTTTCGACTATGTTG
  P   P   S   D   W   A   A   M   G   F   Q   W   N   D   R   F   A   A   F   D   F   F   D   Y   V   E

  320                 330                 340                 350                 360                 370                 380                 390
AAAATGAAGCTGCGATACAGCCGTTGGATGGGTTGATTCGGAGGCCCACAAGGTCTGAAAAGATTGCGCACAACACAC
  N   E   A   A   I   Q   P   L   D   G   I   I   R   R   P   T   R   S   E   K   I   A   H   N   T   H

  400                 410                 420                 430                 440                 450                 460
ACAAGCGACTTGCACTAGACAAATCGAATAGGGATGAAGTGTTCGCAAGTTTGGAAACCGAAATCACAGGTGGGAAAA
  K   R   L   A   L   D   K   S   N   R   D   E   V   F   A   S   L   E   T   E   I   T   G   G   K   R

  470                 480                 490                 500                 510                 520                 530                 540
GAGGACCAGAAATCAGCAGGAATTTCCGAAATGCAGCCAATTAGCCACTATGAAGCCAAGCTCCTAGCCGTGTCTTTA
  G   P   E   I   S   R   N   F   R   N   A   A   N   •
                                                  M   Q   P   I   S   H   Y   E   A   K   L   L   A   V   S   L

  550                 560                 570                 580                 590                 600                 610                 620
GCTATGTATAAGTTTACGGGACGTTGTGAGCCCGCTGTTGCTTTAAATATTGTTAATAAAGCTTGTAATGTTGGTATG
  A   M   Y   K   F   T   G   R   C   E   P   A   V   A   L   N   I   V   N   K   A   C   N   V   G   M

  630                 640                 650                 660                 670                 680                 690                 700
GGCAAGAGCAGCTTCGCGCGCAGGGCTAGGGCGGCTCTCCTGGGGCGTTGTCATAGATGCTTTAGGACTTCCATGGCC
  G   K   S   S   F   A   R   R   A   R   A   A   L   L   G   R   C   H   R   C   F   R   T   S   M   A

  710                 720                 730                 740                 750                 760                 770                 780
ACACGTTGCAATGGTGTGACTTGTTATCCGGGCATTGGAGCTAAGCCAAAAATCGAGATGTTCATCAAATACGGTGTA
  T   R   C   N   G   V   T   C   Y   P   G   I   G   A   K   P   K   I   E   M   F   I   K   Y   G   V

  790                 800                 810                 820                 830                 840                 850
AGTGAGCTGAAACCGTAATTGAGTTGTGCATAAAACCTAATTAATTTTAAGTGTGCAACTATAAAATAAATAAAGTTT
  S   E   L   K   P   •

  860                 870
TAAATATTTTTCCTTTAGA(n)
```

Figure 3.11: Coomassie stained SDS-PAGE of RLSV coat protein isolated form virus particles. Lanes 1 & 2, RLSV coat protein; Lane 3, SDS-7 protein molecular weight markers (Sigma). Sizes of markers are indicated.

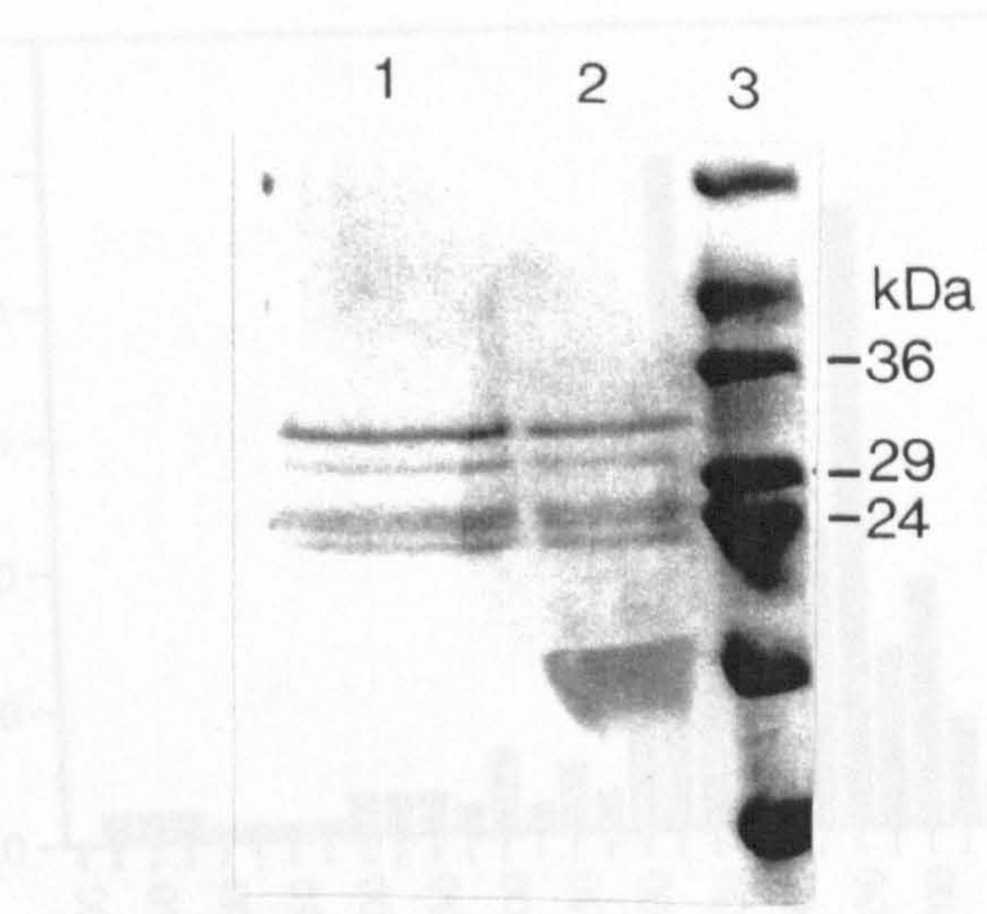
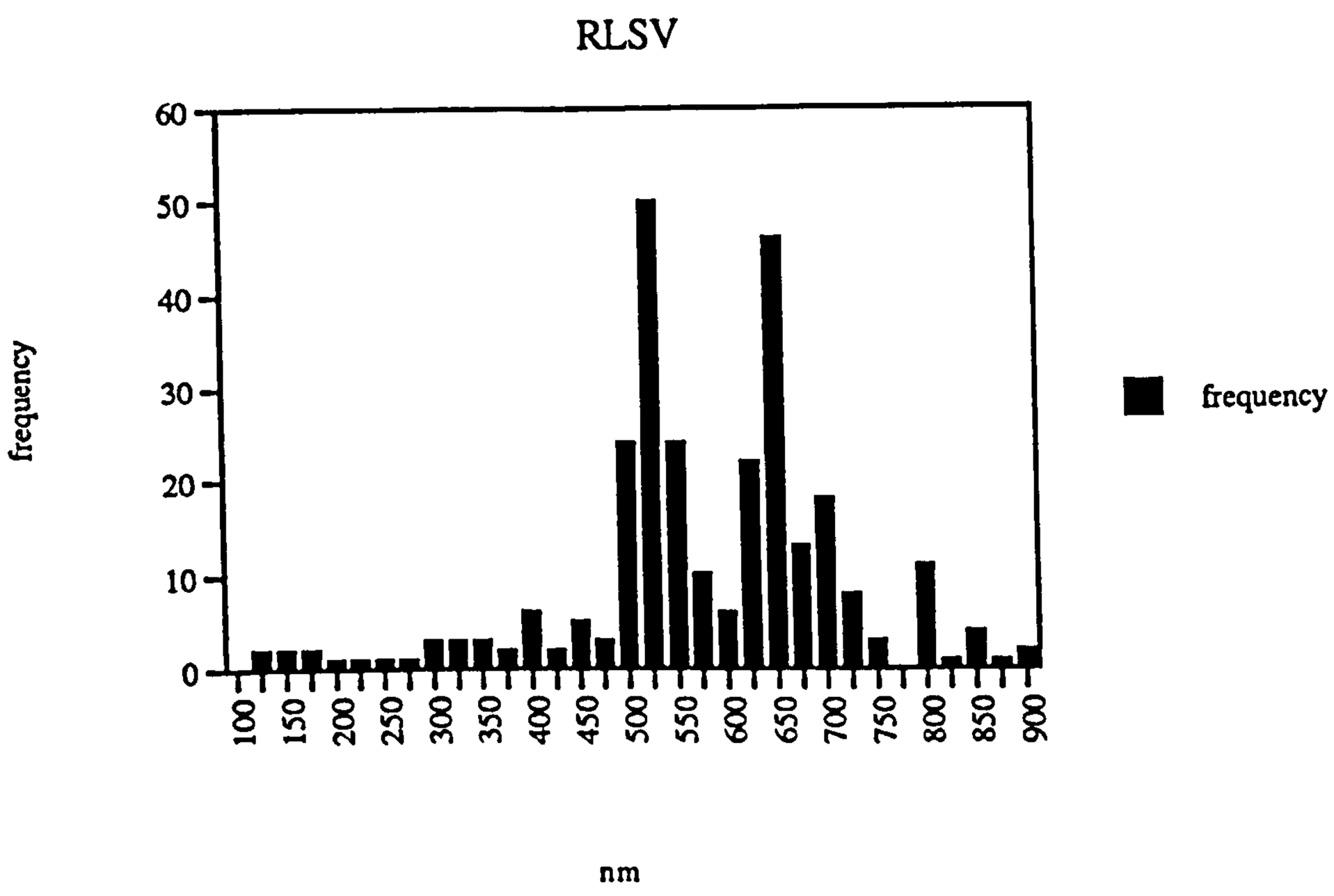


Figure 3.12: Graph to show the frequency distribution of particle size of RLSV. Two peaks, one at 650 nm, the second at 525 nm.



conventional cloning and sequencing was needed to evaluate their taxonomic position (see Chapter 4).

A positive result was obtained for CMMV, with a product of the correct size from RT-PCR with Carla-Uni. The product hybridised with CMMV RNA on dot blots. Conventional cloning and sequencing methods confirmed the presence of the Carla-Uni sequence in the 3' terminal region of the viral genome. It is positioned 125 bp upstream from the poly (A) tail which is in agreement with the size of the RT-PCR product (*c.* 120 bp). The only difference between the primer sequence and that encoded by CMMV is a base substitution at position 843 nt (C to A). Since the significance of this block of homology is not known (Foster, 1992), then the importance of this substitution cannot be assessed. However, the presence of the product confirms that the sequence has not been altered sufficiently to disrupt the binding of the primer.

Although a product of 800 bp was obtained with RLSV, it was not within the expected size range (100-150 bp). However, the result was repeatable, and, after the possibility of contamination of primers or stocks was eliminated, and the product was shown to hybridise to infected sap and viral RNA, it was cloned and sequenced.

The Carla-Uni primer sequence was present at the 5' end of the clone, as would be expected. However it was incomplete, missing a base, A, at position 16 nt. The position of the primer within the genome is not as predicted, hence the larger PCR product size. It is placed within the coat protein, and not near the 3' terminus of the 11K ORF. It is impossible to tell whether this sequence is exactly the same as that present in the genome of RLSV. Once bound, to whatever sequence, it is the sequence of the primer that will be copied in the process of PCR.

A degenerate version of the primer sequence is present in this expected position which could produce a fragment of 102 bp, within the accepted range for Carla-Uni. However, the primer does not bind here, as no fragment of this size was ever observed. It may be possible to encourage binding at this position by altering PCR conditions. RLSV may be a carlavirus that does not conform to this consensus block like poplar mosaic virus and carnation latent virus.

<u>CMMV and RLSV share sequence homology with the carlaviruses</u>

Comparison of the partial coat protein and 11K ORF with those of recognised carlaviruses demonstrates that RLSV and CMMV share a good degree of sequence identity with the genus (figure 3.5 a,b).

Both CMMV and RLSV show levels of sequence identity in the coat protein region to the carlaviruses that supports the view that they are new and distinct members of the genus (Shukla & Ward, 1988). In general, carlaviruses have high homology in the coat protein (*c*.65-70%) (Foster *et al.*, 1990a; Foster & Mills, 1992) and very good serological relationships. CMMV has a lower than average amino acid sequence identity to the other carlaviruses (e.g.: 56% to PVS and 54% to *Hel*VS). Although this does not represent the whole coding region for the coat protein, this lower homology may offer some explanation for the lack of serological relationship between CMMV and the other carlaviruses.

The sequence presented here accounts for 19.5 kDa of the CMMV coat protein. Analysis by SDS-PAGE suggested that CMMV coat protein was only 24 kDa, smaller than that expected of a carlavirus and smaller than that previously reported (32-33 kDa, Brunt & Kenten, 1973). It is possible that the coat protein had been degraded during preparation or analysis.

RLSV, is most closely related to the Andean strain of PVS (68% amino acid identity in the partial coat protein region). In studies on the amino acid sequences of potyvirus coat proteins, Shukla and Ward (1988) proposed that these data could be used as a basis for classification. A comparison of a wide range of complete potyviral coat protein sequences were graphed as a frequency distribution and revealed a clear bimodal distribution. This demonstrated that sequence homology between distinct members of the genus ranged from 38 to 71 % (average 54%) while that between strains of one virus from 90 to 99% (average 95%). If this theory is applied to the carlaviruses, then it is possible to propose that RLSV is a distinct species from the other potato carlaviruses and not another strain of PVS. This is supported by serological evidence. Although it should be remembered that these sequence identities are only based on a partial coat protein sequence (18.8kDa sequenced, 32kDa predicted from SDS-PAGE).

There is some evidence to suggest that post-translational proteolytic processing may occur *in vitro* in carlaviruses (Foster & Mills, 1991d). Homology between PVM C-terminal region and a number a serine proteases was discovered and it was suggested that the source of the viral protease required for this processing could be the coat protein (Foster & Mills, 1991d). The protease reactive catalytic triad (His_{57}-Asp_{102}-Ser_{195}) was found to be equivalent to the PVM sequence His_{261}-Asp_{266}-Thr_{286}. This homology was also found in several other carlaviruses and is also present in RLSV and CMMV (figure 3.13). It was proposed that the essential serine residue had been substituted for a threonine in the carlaviruses as this was mirrored by the presence of the sequence Thr-Gly-Gly found in all carlaviruses which is matched by Ser-Gly-Gly in

```
                    ••          •  |        |              |••
PVM                 AHNTHK----DIAVRGANRNQVFSSLNAEVTGG
RLSV                AHNTHKRLALDKSNRDEVFASLETEI----TGG
CMMV                AHNTYKG-ALDRSNRNKIYSNLNTEV----TGG
ELA                 AHC-------DIALLRLAQS---------GDSGG
```

Figure 3.13 : Alignment of the C-terminal portions of PVM, RLSV, CMMV and elastase (ELA), a mammalian serine proteinase. The catalytic triad is indicated with closed arrows.

five cellular proteases (Foster, 1992). The significance of this unprecedented substitution is not known.

Translation and expression of 11K ORF

The 11K ORF is unique to the carlaviruses and distinguishes them from the otherwise closely related potexviruses. The potexviruses are transmitted mechanically, not by insect vectors, and this had led to the suggestion that the 11K ORF could be responsible for vector transmission or recognition in the carlaviruses (Foster & Mills, 1992). Other viruses which are transmitted by aphid vectors, such as the potyviruses and caulimoviruses, have been shown to encode helper component proteins. Transmissibility has been shown to be virus-encoded in these proteins (Atreya *et al.*, 1990). Although there is no sequence homology between the carlavirus 11K ORF and those of the potyviruses or caulimoviruses, all contain a cysteine motif ($C\text{-}X_2\text{-}C\text{-}X_{12}\text{-}C\text{-}X_4\text{-}C$), a zinc-finger, which is associated with nucleic acid binding (Klug & Rhodes, 1987; Berg, 1986).

CMMV is unusual amongst the carlaviruses as it is transmitted by whiteflies. If function was conserved in overall homology, we would expect to see a higher than normal variation between the 11K ORF of CMMV and those of other carlaviruses. This is in fact not the case, as CMMV shares 48% amino acid sequence identity with PVS and 46% identity with *Hel*VS in the 11K ORF. This is higher than the average level of identity seen in this region (*c.* 40%). This may suggest that either the 11K ORF is not involved in vector specificity or that this specificity is not based on sequence identity.

An unusual feature of the genome organisation of CMMV is that there is a small intergenic gap between the 11K and coat protein ORFs. In the majority of sequenced carlaviruses, there is an overlap between these two reading frames, with the start methionine of the 11K ORF one frameshift back from the termination codon of the coat protein (Foster, 1992). CMMV has a gap of 2 nucleotides between the termination codon of the coat protein and the AUG of the 11K ORF, shifting the frame forward, rather than back. This may indicate that the translation levels of the 11K ORF are different from those normally seen in carlaviruses. If the levels are raised, then excess protein could be produced and perhaps account for the brush-like cytoplasmic inclusion bodies observed in CMMV infections. If this is the case, it is possible to speculate that it is this excess protein which may have an influence vector specificity rather than sequence conservation. This speculation would require further investigation. Purification of the CMMV inclusion bodies would enable them to be tested serologically against antisera raised to 11K protein.

The level of sequence identity in the 11 kDa protein of RLSV is very low (*c*.25%) compared to that normally seen between carlaviruses (*c*. 40%). It contains the cysteine zinc-finger motif. Its mode of transmission has not been established, and it has been detected in a wide range of potato tissue culture samples (per. com. Dr.Brattey). If this virus has been maintained in tissue culture for many years, previously undetected, it is possible that low vector selection pressure may have lead to drift in sequence homology between RLSV 11K ORF and that of other carlaviruses.

Particle sizes

CMMV

The length of CMMV particles was confirmed to be in agreement with published measurements (650 nm). However, a high frequency of 130 nm particles was observed (figure 3.7). It is possible that this could represent the encapsidated sub-genomic RNA which encodes the coat protein and 11K genes (Foster & Mills, 1991b). The length of the particle (1/5th of the full length particle) is in proportion to that expected of a 1.4 kb RNA, which is in agreement with an estimated size of the sub-genomic RNA representing the coat protein and 11K ORFs. No evidence to support the presence of encapsidated sub-genomics was found when a northern blot of CMMV RNA isolated from the same particles used for the measurements was probed with a fragment representing the 11K ORF.

RLSV

Two modal lengths of particle were observed from RLSV, 525 nm and 650 nm (figure 3.12). This is unusual in that the first is within the size range for potexviruses, and the second within that for carlaviruses. RLSV has been confirmed as a member of the carlavirus group which could suggest that the shorter particles may not be RLSV. However, both type of particles reacted similarly to the antisera used for ISEM and only one species of RNA was isolated from the virus preparation. A second *c*.500 bp fragment was consistently observed on RT-PCR, and although this may be single-stranded DNA from an incomplete PCR reaction, it could also be a genuine product. Attempts were made to clone this fragment, though none were successful. A 594 bp fragment could be produced if Carla-Uni bound in the same position on a virus deleted for the 11K ORF. Such a population of deleted viruses would appear shorter than the full length particles. Since RLSV has been observed in many potato tissue culture stocks, it is possible that a sub-population of virus particles without an 11K ORF exist in this environment. There would be no selection pressure for transmission by vectors in this environment, and if the 11K is involved in vector transmission and/or specificity, then it is plausible that such a deleted population could exist. However, as

stated earlier, experimental data have not yet confirmed that RLSV can be transmitted by aphids, and RNA isolated solely from these smaller particles has not been obtained.

In conclusion, the PCR primer Carla-Uni has been shown to be specific to the carlaviruses, not reacting to the potexviruses or potyviruses. It can be used rapidly in one RT-PCR reaction to confirm supporting evidence that a virus is a member of the carlavirus genus. This can be validated by sequence analysis of the 3' terminal region of the genome. The Carla-Uni primer has confirmed that AHLV, HMV, CMMV and RLSV are carlaviruses. However, the negative result obtained with MacMV and NLV required further investigation and thus a different approach was adopted. Conventional double-stranded cDNA cloning of MacMV and NLV was attempted after the failure of RT-PCR with Carla-Uni (Chapter 4).

Chapter 4

Molecular analysis of narcissus latent virus (NLV) and Maclura mosaic virus (MacMV)

Introduction

Narcissus latent virus (NLV)

Narcissus latent virus (NLV) was first described in 1966 (Brunt & Atkey, 1967) as a member of the carlavirus genus. It has been reported to induce very mild leaf chlorosis in the tips of narcissus leaves and affects many commercially important cultivars (Brunt, 1977). It also occurs naturally in bulbous iris, gladioli, nerine and *Acidanthera* (Brunt, 1976; 1977; Derks *et al.*, 1985). NLV has flexuous filamentous particles *c.*657 nm long and *c.*13 nm wide, and a coat protein estimated to be 32.6 kDa (Brunt, 1977), properties suggesting that NLV is a carlavirus. However, in preliminary tests its coat protein cross-reacted only with antiserum raised to one carlavirus, lily symptomless virus (LSV) (Brunt, 1977). However it was later reported that NLV had a coat protein of 45 kDa and that in *Nicotiana clevelandii* it induces cylindrical cytoplasmic inclusions (CCIs) characteristic of potyviruses (Mowat *et al.*, 1991; Brunt *et al.*, 1994). In further serological tests NLV particles failed to react to antisera raised to 12 potyviruses and 9 carlaviruses (Mowat *et al.*, 1991). However, NLV particles cross-reacted in ELISA and IEM with Maclura mosaic virus (MacMV) antiserum, results establishing a link between them. It was concluded that NLV was neither a potyvirus nor a carlavirus but could be a member of a new genus with MacMV (Mowat *et al.*, 1991). Moreover, a PCR primer designed to a sequence present in 80% of all carlaviruses sequenced to date failed to amplify a product with NLV or MacMV cDNA (Badge *et al.*, 1996; Chapter 3).

Maclura mosaic virus (MacMV)

MacMV was first reported to cause mosaic symptoms on the leaves of the ornamental tree *Maclura pomifera* (Plese & Milicic, 1973) and to induce CCIs (Plese & Wrischer, 1978). Like NLV, MacMV has flexuous filamentous particles 650-710nm long. Although some anomalies in the appearance of the isolated particles have been reported, measurements of NLV and MacMV particles have always shown a single modal length, suggesting that the viruses have a monopartite genome (Brunt & Atkey, 1967; Brunt, 1977; Mowat *et al.*, 1991; Plese & Wrischer, 1978). The coat protein size of MacMV is reported to be 45 kDa (Plese & Wrischer, 1978). MacMV was initially classified as a

member of the potyvirus genus due to the presence of CCIs in infected tissue and its weak cross-reaction with antisera to bean yellow mosaic potyvirus (Plese *et al.*, 1979).

Aims

Due to the uncertain taxonomic status of NLV and MacMV, it was decided to investigate further using molecular techniques, since more traditional methods had failed to provide sufficient information.

It is important to establish the size of the coat proteins, as this has been a contentious issue, hence, SDS-PAGE and western analyses were carried out. In order to finally quantify the size of the coat proteins, it was necessary to sequence them. Since RT-PCR techniques using universal primers had failed in the past (Chapter 3, Carla-Uni), a more conventional approach was adopted. Double-stranded cDNA was generated from viral RNA and blunt cloned into dephosphorylated vector. Once clones were isolated and confirmed as of viral origin, primers were designed to complete the 3' terminal region of each genome. In this way, sequence data was obtained for the 3' terminal region of NLV and MacMV, enabling objective analysis of the taxonomic status of these two viruses.

The relationship between NLV found in bulbous Iris and Gladioli was investigated by RT-PCR and sequence analysis. Once considered three distinct viruses it was necessary to confirm the suggestion that they were synonymous (Brunt, 1977).

Results

RNA

Virus particle suspension (supplied by Prof. A Brunt, according to method outlined Brunt, 1977) was used to isolate RNA from NLV and MacMV. These virus particles had been stored at 4°C for several years. NLV and MacMV RNA analysed on both denaturing and non-denaturing gels (Chapter 2: section 3.2) was shown to be a single species estimated to be *c.* 8 kb in length from its position between the 9.5 kb and 7 kb RNA markers (figure 4.1).

Coat protein sizes

Coat protein isolated from MacMV particles when analysed by SDS-PAGE and stained with Coomassie Blue, produced two or more bands ranging from 32 to 20 kDa (figure 4.2). When this 32 kDa protein was subjected to N-terminal micro-sequencing the sequence 'SDPEE' was obtained. This sequence was submitted to a database search but without success. Further analysis using western blotting (Chapter 2: section 4.2.3) of MacMV infected *N. clevelandii* tissue showed that coat protein of the purified virus

Figure 4.1: Genomic viral RNA extracted from virus particles. Lane 1: narcissus latent virus (NLV); Lane 2: Gibco RNA molecular weight markers; Lane 3: Maclura mosaic virus (MacMV). NLV and MacMV single-stranded RNA genomes were estimated to be *c*. 8 kb, closed arrows show the position of the two molecular weight markers closest in size.

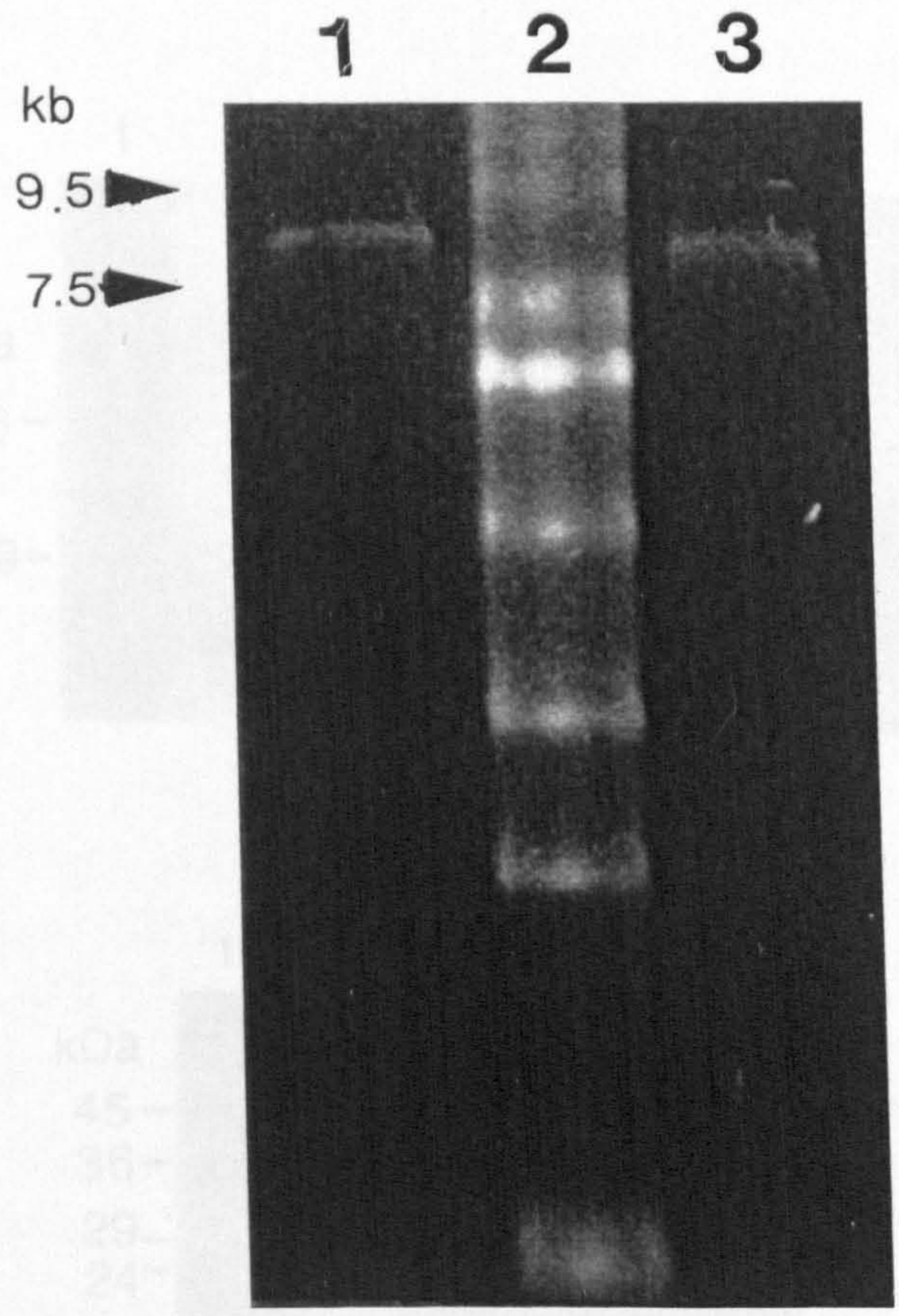
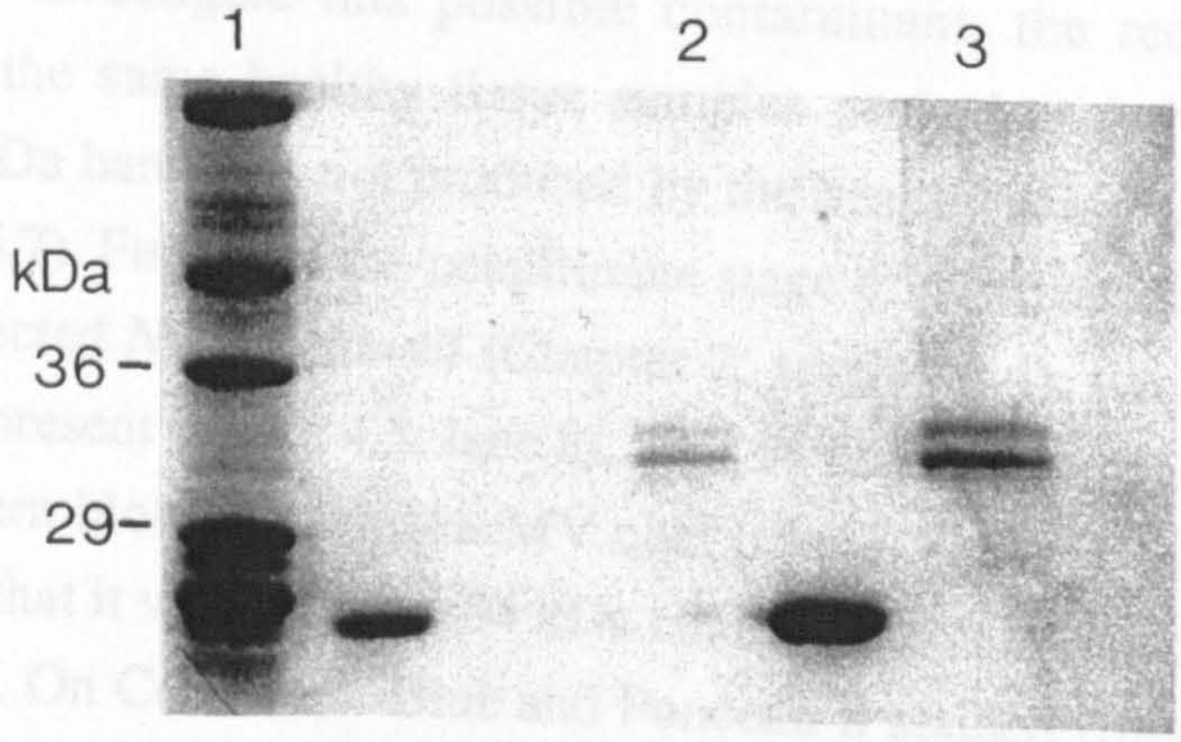
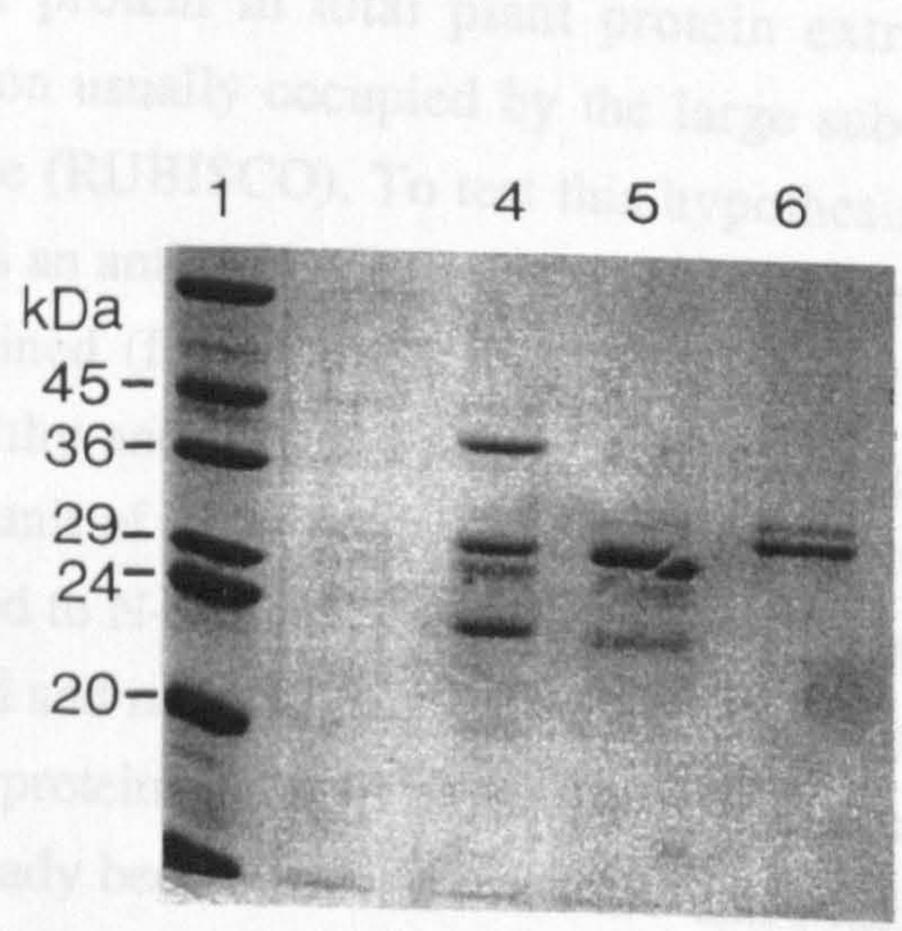


Figure 4.2: Coomassie-Blue stained SDS-PAGE analysis of coat protein extracted from virus particles. Lane 1: SDS-7 molecular weight markers, lanes 2 & 3: MacMV, lane 4: NLV extracted form narcissus, lane 5: NLV extracted from gladioli, lane 6: NLV extracted from iris.





particles was probably degraded, since a single protein of 40 kDa (figure 4.3; lane 1) was observed. N-terminal micro-sequencing of this 40 kDa was unsuccessful as the product was N-terminally blocked. Similarly, NLV coat protein isolated from virus particles was estimated to be 32 kDa, whereas western analysis of coat protein from infected narcissus tissue indicated that the coat protein was 39.5 kDa (fig 4.3; lane 2).

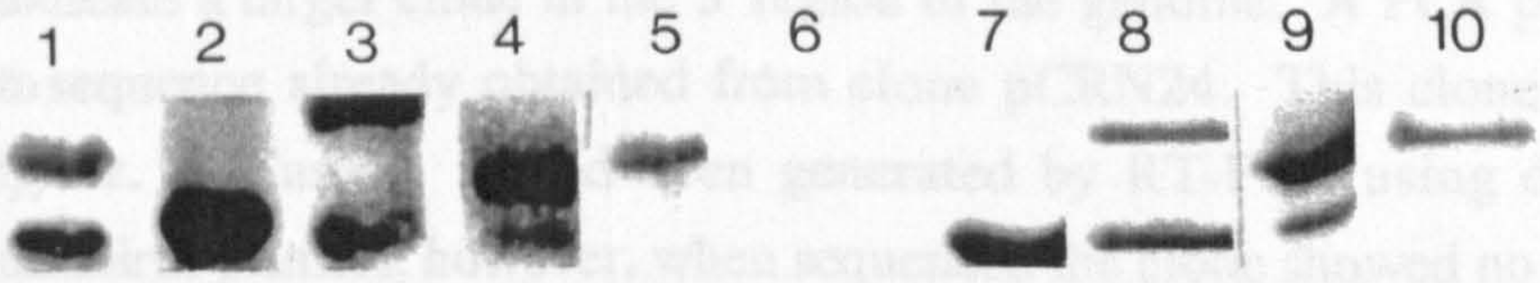
A reciprocal reaction between NLV and MacMV from *N. clevelandii* was observed when using antisera raised to either coat protein. Antiserum raised to MacMV coat protein also reacted strongly with NLV infected material from narcissus (figure 4.3; lane 3). However, we consistently observed some contamination by the reaction of a band present in *N. clevelandii* and narcissus tissue of *c.* 54 kDa (figure 4.3; lanes 3,4 & 5). To further investigate this possible contaminant, the reciprocal blot was performed using the same healthy tissue samples probed with NLV coat protein antisera. The 54 kDa band was not produced by the healthy samples, nor by MacMV (figure 4.3; lanes 6,7). Further at the penultimate stage of purification of MacMV virus particles from infected *N. clevelandii* (Chapter 2: section 6.4), two proteins, 40 kDa and 54 kDa, were present (figure 4.3; lane 8). This 54 kDa co-purifying protein reacted positively in western blotting with MacMV coat protein antisera (figure 4.3: lane 9), results suggesting that it was the same 54 kDa contamination band observed in healthy and infected tissue. On Coomassie-Blue and Ponceau-S stained filters, a 54 kDa band was the most abundant protein in total plant protein extracts, and appeared to correspond to the position usually occupied by the large sub-unit of Ribulose-1, 5-bisphosphate carboxylase (RUBISCO). To test this hypothesis, the partially purified MacMV was probed with an antibody raised to the large sub-unit of RUBISCO, and a positive result was obtained (figure 4.3: lane 10). We concluded that this 54 kDa protein observed in healthy and infected western blots using MacMV coat protein antisera is the large sub-unit of RUBISCO. in order to verify this conclusion, the 54 kDa protein was submitted to N-terminal micro-sequencing. Unfortunately the protein was N-terminally blocked and no sequence was obtained. If antiserum was raised to a complex of the virus coat protein and RUBISCO, then it would react with RUBISCO in healthy tissue. It has already been observed that NLV and MacMV particles are often coated with granules in infected sap (Mowat *et al.*, 1991; Plese *et al.*, 1979) which could provide an explanation not only for the presence of some host protein/viral coat protein complex but the previous sizing of NLV and MacMV at 54 kDa.

<u>Clones isolated for sequencing</u>

To determine the size of the coat protein genes, clones representing the 3' terminal regions of both viruses were obtained with the assumption that the coat protein gene would be encoded in this region.

Figure 4.3: Analysis of the serological relationship between NLV and MacMV coat proteins by western blotting. Samples of total protein extracted from healthy and infected plant material were separated by SDS-PAGE, blotted and probed with antiserum raised to either NLV or MacMV coat protein. The source of plant material is shown below each sample and the antibody used above them. Lane 1: MacMV-infected *N. clevelandii* blotted with MacMV antisera (two bands 40 and 54 kDa); Lane 2: NLV-infected *N. clevelandii* blotted with NLV antisera (one band 40 kDa); Lane 3: NLV-infected *N. clevelandii* blotted with MacMV antisera (two bands 40 and 54 kDa); Lane 4: NLV-infected Narcissus blotted with MacMV antisera (two bands 40 and 54 kDa); Lane 5: Healthy *N. clevelandii* blotted with MacMV antisera (one band 54 kDa); Lane 6: Healthy *N. clevelandii* blotted with NLV antisera (no bands present); Lane 7: MacMV-infected *N. clevelandii* blotted NLV antisera (one band 40 kDa).

Lanes 8-10 show total protein extracts from a bulk scale MacMV virus particle preparation: Lane 8: stained with Coomassie Blue (two bands at 40 and 54kDa); Lane 9: analysed by western blotting with MacMV coat protein antibody (40 and 54kDa bands); Lane 10: western analysis using antisera raised to the large sub-unit of RUBISCO (single positive 54 kDa band).



MacMV clones

First-strand cDNA was generated using a mixture of oligo-d(T) primer which incorporated a *Not I* restriction site at its 3' end (denoted oligo-d(T)Not I: see Chapter 3 for sequence) and random hexamers from MacMV RNA obtained from purified virus particles. Double-stranded cDNA was created and blunt cloned into *Sma I* cut, dephosphorylated pUC18. The resultant transformants were screened using radio-labelled first-strand cDNA generated from viral RNA (Chapter 2: section 1.6.2). Initially, four clones were analysed, pM3, pM7, pM91 and pM99, though none contained the oligo-d(T)Not I primer. Internal primers were designed from sequence obtained from these four clones, in order to completely sequence both strands of each clone. Two of these PCR primers (99.1 and 99.3) were used as the upstream primers for RT-PCR using oligo-d(T)Not I-primed first-strand cDNA to create further clones, pM99T1 and pM99T2. Figure 4.4a shows the relative positions of the cDNA clones.

A clone isolated from the original screen of the double-stranded cDNA clones, pM4 with an insert of *c.* 1.4 kb, did not show any homology with the other clones obtained. Re-screening the original double-stranded cDNA clones with the insert from pM4 led to the identification and isolation of pM29. Sequence analysis showed that pM29 overlapped with pM4. Figure 4.4b shows the relative positions of these cDNA clones.

NLV clones

Double-stranded cDNA clones were generated in the same way for NLV, however only one clone was successfully isolated and sequenced, pNLV39. An RT-PCR approach was used to create a larger clone in the 3' region of the genome. A PCR primer was designed to sequence already obtained from clone pCRN24. This clone had been obtained by Dr. R. Turner. It had been generated by RT-PCR using degenerate universal carlavirus primers, however, when sequenced the clone showed no homology to the carlavirus region usually amplified by these primers. Both primers used in the amplification were not present in the sequenced clone but degenerate copies of only one of the primers were detected at both ends of the insert (primer sequences found were degenerate at 2 or 3 positions of the 14-mer primer). However, the insert reacted positively when radio-labelled and probed onto NLV-infected sap blots (Chapter 2: section 6.2) and therefore it was considered a genuine NLV clone and used to create primers NLV 1 and NLV 2 (figure 4.4a).

NLV1 was used in an RT-PCR reaction with oligo-d(T)Not I-primed first-strand cDNA. A fragment of *c.*1.5 kb was obtained and cloned into pCRII (Invitrogen). Resultant colonies were screened with the radio-labelled insert from pCRN24 and one

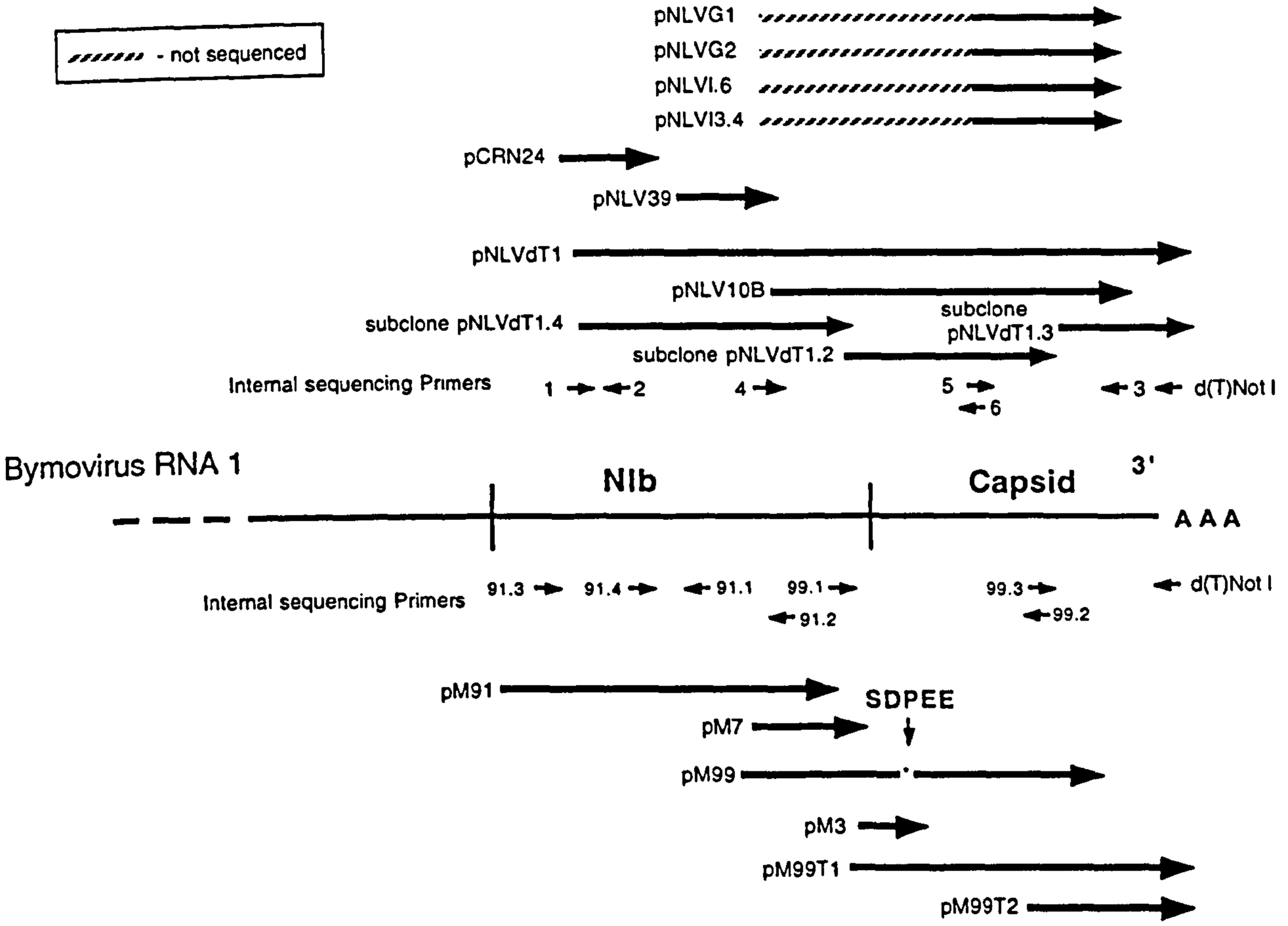


Figure 4.4 a: Positions of cDNA clones relative to one another and in reference to the genomic organisation of a typical bymovirus. The clones obtained from NLV are shown above the bymovirus genome, those from MacMV below. These clones were used to obtain the 3' terminal regions of NLV and MacMV genomes. Positions and names of primers used for internal sequencing are indicated.

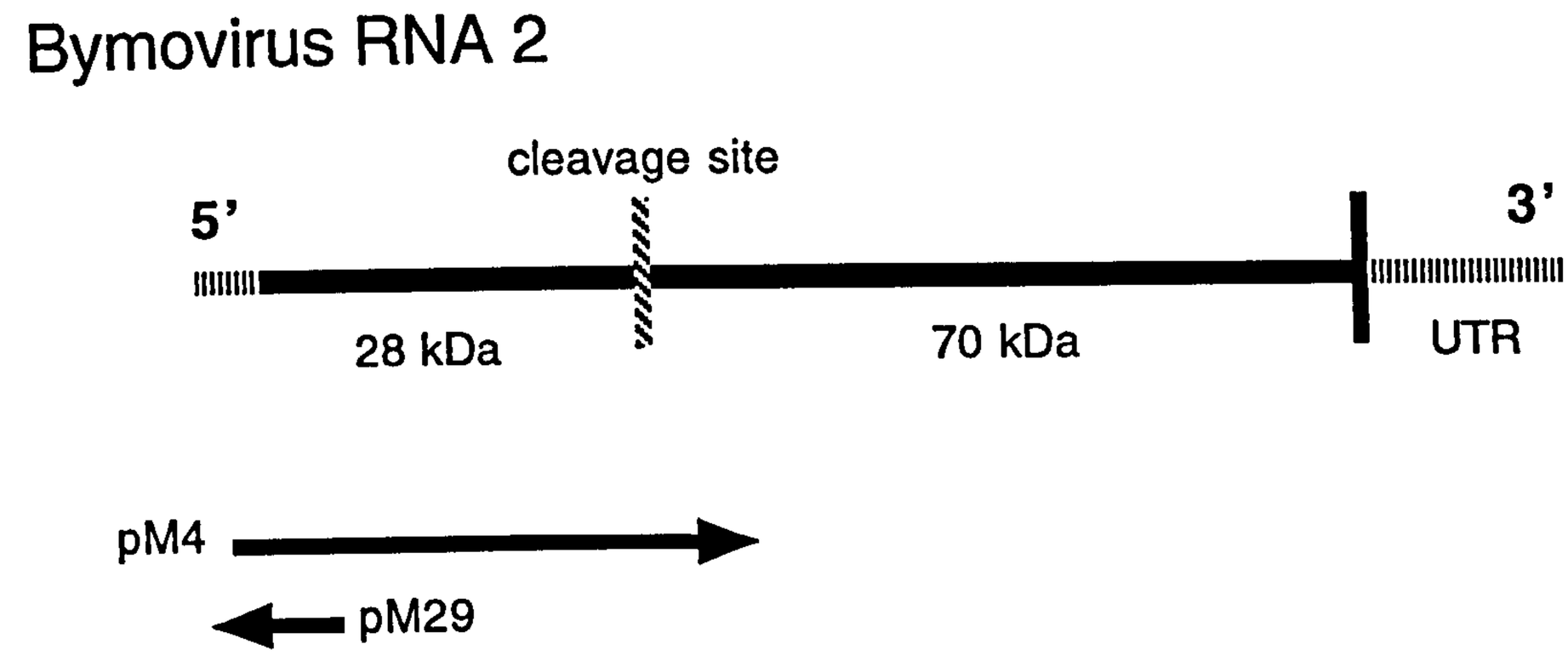


Figure 4.4 b: Positions of cDNA clones relative to one another obtained from the 5' terminus of MacMV

clone was isolated, pNLVdT1. In order to sequence this large clone, three subclones (pNLVdT1.2, pNLVdT1.3, pNLVdT1.4) were created using two *EcoRI* sites within pNLVdT1. The junctions between these subclones were sequenced using internal primers on the full length clone in order to ensure that only one restriction site was present in each position.

A fourth clone in this region of the genome (pNLV10B) was isolated by screening colonies obtained from cloning another RT-PCR fragment (Chapter 2: section 1.6.1). This fragment was amplified from first-strand cDNA generated from the total RNA of NLV-infected narcissus leaves. The first-strand was primed using a primer (NLV3) designed to sequence obtained from pNLVdT1 and a second upstream primer (NLV4), also from NLVdT1 was used to amplify the fragment. A fragment of the predicted size (1.1 kb) was obtained and cloned into pCRII (Invitrogen).

Clones from Iris and Gladioli

NLV has been reported to infect gladioli and bulbous iris (Brunt, 1977) but was only diagnosed by antisera tests and electronmicrographs of particles. In order to confirm that these particles found in iris and gladioli were NLV, sequence data was obtained from them.

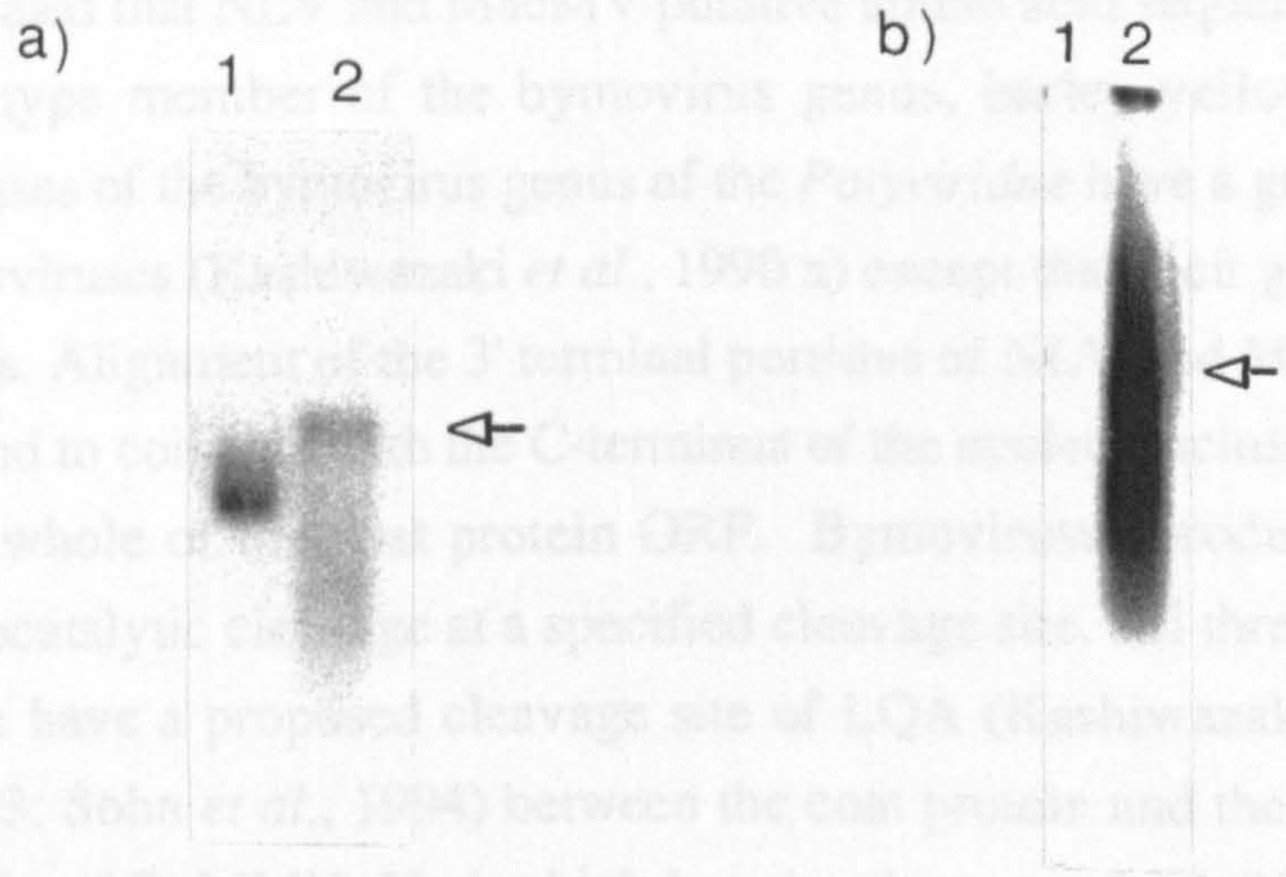
Viral genomic RNA was extracted from virus particles suspected to be NLV purified from infected Iris and Gladioli samples (kind gift of Prof. A. Brunt). First-strand cDNA was generated using NLV3 as the downstream primer. This cDNA was used in an RT-PCR reaction with an upstream primer, NLV4 to create a fragment of the size predicted by the sequence obtained from pNLVdT1 (*c.* 1.1 kb). The fragments were cloned into pCRII. Two independent clones containing the fragment obtained from PCR of the Gladioli sample were isolated (pNLVG1, pNLVG2). Similarly, two independent clones from the Iris sample were obtained (pNLVI3.4, pNLVI6).

Northern analysis

A northern analysis (figure 4.5) (Chapter 2: section 3.2.1) was performed on NLV-infected *N. clevelandii* tissue (Chapter 2: section 3.1.4) and MacMV RNA purified from virus particles using two cDNA clones, which were later shown to be in the coat protein region of the viral genome, pNLVdT1 and pM91 (respectively). A single band for each virus was observed with a slight smear of smaller products decreasing in size, however, there was no evidence of any other distinct bands. The largest band for each virus was estimated to be *c.* 8 kb by comparison with the ethidium bromide stained RNA molecular weight markers used on the original gels.

Figure 4.5 (a): Northern blot of viral RNA extracted from MacMV particles and probed using the radiolabelled insert from the cDNA clone, pM91. Lane 1: RNA molecular weight markers (GibCo), Lane 2: MacMV RNA.

Figure 4.5 (b): Northern blot of NLV-infected total nucleic acid extracted from *Nicotiana clevelandii* and probed using the radiolabelled insert from the cDNA clone, pNLVdT1. Lane 1: uninfected narcissus RNA, Lane 2: NLV RNA.



Nucleotide and deduced amino acid sequence analysis

3' terminal region of NLV and MacMV

The deduced sequence for the 3' terminal portions of MacMV and NLV RNAs are shown in Figures 4.6 and 4.7 (respectively). A total of 1797 nucleotides (nt) were sequenced for NLV and 2413 nt for MacMV. Analysis showed that each sequence contained a single continuous open reading frame (ORF). This was 727 amino acids (aa) long in the case of MacMV and 513 aa long in the case of NLV. Stop codons were found at position 2182nt (MacMV) and 1541nt (NLV), with a 3' untranslated region (UTR) of 231nt (MacMV) and 256nt (NLV). Alignment of the coding regions indicated that NLV and MacMV RNAs are 52.4% identical at the nucleotide level.

From the nucleotide sequences amino acids sequences were deduced. Database searches revealed that NLV and MacMV putative amino acid sequences had highest similarity to the type member of the bymovirus genus, barley yellow mosaic virus (BaYMV). Viruses of the bymovirus genus of the *Potyviridae* have a genetic organisation similar to potyviruses (Kashiwazaki *et al.*, 1990 a) except that their genomes are divided into two parts. Alignment of the 3' terminal portions of NLV and MacMV putative proteins was found to coincide with the C-terminus of the nuclear inclusion body b (NIb) protein and the whole of the coat protein ORF. Bymoviruses produce mature viral proteins by autocatalytic cleavage at a specified cleavage site. All three bymoviruses sequenced to date have a proposed cleavage site of LQA (Kashiwazaki *et al.*, 1989; Foulds *et al.*, 1993; Sohn *et al.*, 1994) between the coat protein and the NIb proteins except for one strain of BaMMV, Na I which has the cleavage site LQS (Kashiwazaki *et al.*, 1992; Kashiwazaki, 1996).

Figure 4.8 shows an alignment using CLUSTAL of the deduced NIb and coat protein amino acid sequences of NLV, MacMV, two members of the bymovirus genus, (BaYMV: Kashiwazaki *et al.*, 1990a) and barley mild mosaic virus (BaMMV: Kashiwazaki *et al.*, 1992), a typical potyvirus, potato virus Y (PVY : Robaglia *et al.*, 1989), the type members of the two other genera recognised in the *Potyviridae* family, ryegrass mosaic virus, RGMV (rymoviruses: Salm *et al.*, 1996) and sweet potato mild mottle virus, SPMMV (ipomoviruses: Colinet *et al.*, 1996). Comparison of the sequence of NLV and MacMV with that of the members of the bymoviruses suggests a possible cleavage site between the coat protein and the NIb protein at position 215aa (NLV) and 423aa (MacMV) around the tripeptide sequence 'LQM' (Figures 4.6 & 4.7). It is estimated from the deduced amino acid sequence, using the LQM site as its N-terminus, that the coat protein size for NLV is 32.8 kDa and for MacMV is 34.1 kDa. N-terminal sequencing of MacMV coat protein revealed the sequence SDPEE which was present within the deduced amino acid sequence (figure 4.7).

Figure 4.6: The DNA sequence corresponding to the 3' terminal region of NLV positive-sense RNA, with derived amino acid sequence indicated below. A possible cleavage site, LQM (position: 645nt) between NIb protein and coat protein has been underlined. Termination codon is marked *.

```
 10        20        30        40        50        60        70        80
TAGTGGCCAGCCCAGCACAGTTGTGGACAACACTTTGGTGCTAATGCTGTCCTTCTATTACGCATATGCCAAGAAAACCA
  S  G  Q  P  S  T  V  V  D  N  T  L  V  L  M  L  S  F  Y  Y  A  Y  A  K  K  T

         90        100       110       120       130       140       150       160
ATGACCTGGCCTTCGAGCACATAGACGAGCGTTTTCGCTTTGTTTGCAATGGTGACGACAACAAATTCTCAGTGTCGCCT
N  D  L  A  F  E  H  I  D  E  R  F  R  F  V  C  N  G  D  D  N  K  F  S  V  S  P

        170       180       190       200       210       220       230       240
GAATTTGTCGCCGAATTTGGGGGTTCGTTTGGGGATGAGATAAGTGAACTAGGCTCACTTACGAGTTTGATGATCCTAAC
 E  F  V  A  E  F  G  G  S  F  G  D  E  I  S  E  L  G  S  L  T  S  L  M  I  L  T

        250       260       270       280       290       300       310       320
GCCGGACATAATGAGAAACCCATACATGAGCCTGACAATAGTTCAAGTGGGCGATCGCATTGGCTTTCAGCTCAACCCTG
  P  D  I  M  R  N  P  Y  M  S  L  T  I  V  Q  V  G  D  R  I  G  F  Q  L  N  P

        330       340       350       360       370       380       390       400
AACGAATCGTTGGGATTGTTCAATGGATCAAGAAAGGGGGCGTACTTCATGCAGCCCAGGCAGCATTTGCAGCAATGGTT
E  R  I  V  G  I  V  Q  W  I  K  K  G  G  V  L  H  A  A  Q  A  A  F  A  A  M  V

        410       420       430       440       450       460       470       480
GAAGCATTCAACGACCCAAGTCTATTCACTGTCATGCACACATATTTGGTGTGGCTGTTGGTGACTCACAAGGATGTTCT
 E  A  F  N  D  P  S  L  F  T  V  M  H  T  Y  L  V  W  L  L  V  T  H  K  D  V  L

        490       500       510       520       530       540       550       560
GCTGTACGCACAAGAGAACGGTCTGGGTTCGGTTTGTTACATGGACCCGTGTCAGGTGTTTGCGTTGCACAATGGCTCTT
  L  Y  A  Q  E  N  G  L  G  S  V  C  Y  M  D  P  C  Q  V  F  A  L  H  N  G  S

        570       580       590       600       610       620       630       640
CTAAAGGCTTAGAAGATGTCAAGTTCGACAATGAAGATGAGAGTGCGGATGAGGATGATGGTAACATTACACCCGACCTG
S  K  G  L  E  D  V  K  F  D  N  E  D  E  S  A  D  E  D  D  G  N  I  T  P  D  L

        650       660       670       680       690       700       710       720
GAACTCCAAATGGACGTAGGCAACCTCATTCCTGAGAAGGAGAAGAATTCCCAAAATGTGAATACAAGTGACGGAGGCAA
 E  L  Q  M  D  V  G  N  L  I  P  E  K  E  K  N  S  Q  N  V  N  T  S  D  G  G  K

        730       740       750       760       770       780       790       800
GAACGCCAGCAACTCTGCCACAGGCGAGTCAAGTAAACCTCCTGAGAATAAAGCAGGGAAGGGCGCTGATCAAGGTGATG
  N  A  S  N  S  A  T  G  E  S  S  K  P  P  E  N  K  A  G  K  G  A  D  Q  G  D

        810       820       830       840       850       860       870       880
TAGATCCACCACAGGGTGACCCCTTAGTTGATGATGAGGTGGTTGAGTGGGTCATTCCAAAAATGAGCCCAAACATTGGA
V  D  P  P  Q  G  D  P  L  V  D  D  E  V  V  E  W  V  I  P  K  M  S  P  N  I  G

        890       900       910       920       930       940       950       960
ACTTCACCGATACCGGTCATCAACGGGAAGAGGCTGTGGAAGCGAGGAGTATTGAAGAGCATACCCAAAATGATGTTCAA
 T  S  P  I  P  V  I  N  G  K  R  L  W  K  R  G  V  L  K  S  I  P  K  M  M  F  N

        970       980       990      1000      1010      1020      1030      1040
CACGACTTCAACCATGGCAACTCAAGCCCAACTCACCTCATGGGTGGAGGAGGTGAAGCAAGCGCTTGCGCTTAAAACAG
  T  T  S  T  M  A  T  Q  A  Q  L  T  S  W  V  E  E  V  K  Q  A  L  A  L  K  T

       1050      1060      1070      1080      1090      1100      1110      1120
ATGATGCATGGACGGTTGTCATCACGAATTGGTGCATTTGGTGCGCTAACAACGGCACATCGTCTGAAGTTGATACGTCG
D  D  A  W  T  V  V  I  T  N  W  C  I  W  C  A  N  N  G  T  S  S  E  V  D  T  S

       1130      1140      1150      1160      1170      1180      1190      1200
CAGACTATGGAGATTCGCGATGGTTTTGGGAAGGTTCAGGCTATCCCCATTGAGGTTTTTGTGAACCCAGCTGTTGAAAA
 Q  T  M  E  I  R  D  G  F  G  K  V  Q  A  I  P  I  E  V  F  V  N  P  A  V  E  N

       1210      1220      1230      1240      1250      1260      1270      1280
TGGCGGTTTACGAAAGATAATGAGGCACTTCTCTGGCATAACTCACGAGATTTTAAAGGCTGGTAAGCGAATGACAGCCT
  G  G  L  R  K  I  M  R  H  F  S  G  I  T  H  E  I  L  K  A  G  K  R  M  T  A
```

```
      1290       1300       1310       1320       1330       1340       1350       1360
GGGGAAATAAGAGGGGCTTTACTGAGAAATCAATGATCCCCTACGCTTTTGATTACTATGTTGTCACAAACACAACGCCC
W  G  N  K  R  G  F  T  E  K  S  M  I  P  Y  A  F  D  Y  Y  V  V  T  N  T  T  P

      1370       1380       1390       1400       1410       1420       1430       1440
AAGACTGTGCGCGAGCAATTGGCACAGTCGAAGGCAGCAGCAATTGGTTCAGGTGTGACGAGGAAGATGGTTCTTGATGG
 K  T  V  R  E  Q  L  A  Q  S  K  A  A  A  I  G  S  G  V  T  R  K  M  V  L  D  G

      1450       1460       1470       1480       1490       1500       1510       1520
GAACATACAAGGGTCTCATGCCAGTTACGAACGCCATGTCGATACGGACAACAGCGAGTACGAGCATGGCAATGATGTTG
  N  I  Q  G  S  H  A  S  Y  E  R  H  V  D  T  D  N  S  E  Y  E  H  G  N  D  V

      1530       1540       1550       1560       1570       1580       1590       1600
ATCAAAGGCCCTATCTTACTTGAATTCTGCGTTCTTGAACGCTCGAATTCAACTATTTGCATTCGCTCTTTATAAGTTTT
D  Q  R  P  Y  L  T  *

      1610       1620       1630       1640       1650       1660       1670       1680
CTCCTAAGCTTATTTCTTTTCTCTCTCTCTTGAGCGTTAAGGGTTAAATGGGGGTTTATTTGCCATTGTTCCTCCAATGC

      1690       1700       1710       1720       1730       1740       1750       1760
CCTGAAATGGCAAGCTCATTCTAGCGATGTCTATCTACAAAGAGTCTTAATCGCTTTCACTTAAGTCTGATCTGGATGAT

      1770       1780       1790
CGACTGCGTTTAAACGTATCCTACGTCGTAGAAGACA(n)
```

Figure 4.7: The DNA sequence corresponding to the 3' terminal region of MacMV positive-sense RNA, with derived amino acid sequence indicated below. The motif GDD found in all positive-sense RNA viruses is underlined. The amino acids identified by N-terminal sequencing of a partially degraded MacMV coat protein sample, SDPEE is underlined.

```
        10         20         30         40         50         60         70         80
AGGGCTGGGATGACCCAAACGAGGATTAGAACAACAATGGAGGTTCTGGAGGACATACAGTGGGGTAAGGCCGCTGGACC
 R  A  G  M  T  Q  T  R  I  R  T  T  M  E  V  L  E  D  I  Q  W  G  K  A  A  G  P

        90        100        110        120        130        140        150        160
CCTGTACGCAATGAAGAAGCGGGACCTCTGCAAGAATCTCACAGAGGAGGAGTTAGTTGCCCTTGGCGCACATTGTAGAA
 L  Y  A  M  K  K  R  D  L  C  K  N  L  T  E  E  E  L  V  A  L  G  A  H  C  R

       170        180        190        200        210        220        230        240
GCGAGTTGAACAAGGGTAAAAACGCTGGGTTGTGGAATGGATCTCTCAAAGCTGAGCTGCGGCCTAAAGAGAAGGTTGAC
S  E  L  N  K  G  K  N  A  G  L  W  N  G  S  L  K  A  E  L  R  P  K  E  K  V  D

       250        260        270        280        290        300        310        320
TTAAACAAAACGCGTGTCTTCACACCAGCCCCCATCACGACGCCTCATTGCGCAAAGTATTTTGTGGACGACTTCAACAA
 L  N  K  T  R  V  F  T  P  A  P  I  T  T  P  H  C  A  K  Y  F  V  D  D  F  N  K

       330        340        350        360        370        380        390        400
GCAGTTCTACAAATCCCACCTCAAAGCCACCTCACCACTGTTGGCATTAATAAGTTCCCAGAATGGGTGGTCAAAGGTAT
  Q  F  Y  K  S  H  L  K  A  T  S  P  L  L  A  L  I  S  S  Q  N  G  W  S  K  V

       410        420        430        440        450        460        470        480
ACAACAAGCTCAACAAAGACCGGCTGCTACATGGAAGTGGAGATGGATCACGCTTTGATTCATCCATCGATCCATACCTA
Y  N  K  L  N  K  D  R  L  L  H  G  S  G  D  G  S  R  F  D  S  S  I  D  P  Y  L

       490        500        510        520        530        540        550        560
TTTGACATGATTTACACCATCCGGTGCCATTTCATGCATGAGGATGATAAACGGGAAAGTGAAAGAGCCATGAGCAACAT
 F  D  M  I  Y  T  I  R  C  H  F  M  H  E  D  D  K  R  E  S  E  R  A  M  S  N  M

       570        580        590        600        610        620        630        640
GTTTCGCGAGTTTGTTTTTTACTCCAATTCACACGATCAGTGGAACATCCTGGTCAAGCAAGTTGGCAATAACAGCGGAC
  F  R  E  F  V  F  Y  S  N  S  H  D  Q  W  N  I  L  V  K  Q  V  G  N  N  S  G

       650        660        670        680        690        700        710        720
AACCAAGTACAGTCGTGGATAATACGCTGGTACTCATGATCTCATTTTATTACGCATACGCTGTAAAGACGAGGGATTAT
Q  P  S  T  V  V  D  N  T  L  V  L  M  I  S  F  Y  Y  A  Y  A  V  K  T  R  D  Y

       730        740        750        760        770        780        790        800
ACGTTTGATAAAATCGATGAAAGATTCGTCTTTGTGTGTAATGGTGATGACAACAAATTCGCGGTGTCACCGGAATTCGT
 T  F  D  K  I  D  E  R  F  V  F  V  C  N  G  D  D  N  K  F  A  V  S  P  E  F  V

       810        820        830        840        850        860        870        880
TAAGGAGTTTGGTGGGTCCTTCACTGACGAGATTGCACAGCTCGGTCTTCATTACGAATTCGATGAGCTCACAACCGACA
  K  E  F  G  G  S  F  T  D  E  I  A  Q  L  G  L  H  Y  E  F  D  E  L  T  T  D

       890        900        910        920        930        940        950        960
TAACAGCGAATCCTTACATGAGCTTAACAATGATAGACATAGGAGGGCGCATAGGGTTCCAACTCAATCCCGAACGTATT
I  T  A  N  P  Y  M  S  L  T  M  I  D  I  G  G  R  I  G  F  Q  L  N  P  E  R  I

       970        980        990       1000       1010       1020       1030       1040
CTAGGAATCGTTCAGTGGATCAAGAAAGGTGGAATTGTGCACGCCGCACAAGCCGCCTTTGCAGCAATGATCGAATCATT
 L  G  I  V  Q  W  I  K  K  G  G  I  V  H  A  A  Q  A  A  F  A  A  M  I  E  S  F

      1050       1060       1070       1080       1090       1100       1110       1120
CAATGATCCAGACCTGTTTTGTGTGATGCACTCGTATTTGGTGTGGCTGCTTGTAACATACCGAAGTGAACTCGTTTATG
  N  D  P  D  L  F  C  V  M  H  S  Y  L  V  W  L  L  V  T  Y  R  S  E  L  V  Y

      1130       1140       1150       1160       1170       1180       1190       1200
CCATGCACAATGATCTGGTGTCAGTGGTTTACATGGACCCATGCCAAGTCTTTGCCTTGCACTACAACGACAGCGAGGAC
A  M  H  N  D  L  V  S  V  V  Y  M  D  P  C  Q  V  F  A  L  H  Y  N  D  S  E  D
```

```
      1210       1220       1230       1240       1250       1260       1270       1280
GTAAGGGAGTGGTTTGATGAGGACGATGAGTCAAGTGATGACGAAGACGAAGAGCCAACTCAAGTGCTACAGATGGACGC
  V  R  E  W  F  D  E  D  D  E  S  S  D  D  E  D  E  E  P  T  Q  V  L  Q  M  D  A

      1290       1300       1310       1320       1330       1340       1350       1360
GGAAACCCTTGCGAAGGATGGAGAGGCGAAGAAAGAGAAGGACGAGAAAGAAAGGGAGAAGGCTGAACAAAGGCGAGTAG
  E  T  L  A  K  D  G  E  A  K  K  E  K  D  E  K  E  R  E  K  A  E  Q  R  R  V

      1370       1380       1390       1400       1410       1420       1430       1440
AAGTCGAGAAGGCAAGAGCAGAAAAAGCTCAAGTGTCAGATGGCGCCAAGGAACCACAGCCGGAGATCAAAGGGAACGAG
E  V  E  K  A  R  A  E  K  A  Q  V  S  D  G  A  K  E  P  Q  P  E  I  K  G  N  E

      1450       1460       1470       1480       1490       1500       1510       1520
GACGTTGAGCAGCCAGCTAGTGACCCTGAAGAAAAGGAAGAAGAGGTCAAATGGGTCATGCCTTCAATAAATCCCAACAG
 D  V  E  Q  P  A  S  D  P  E  E  K  E  E  E  V  K  W  V  M  P  S  I  N  P  N  R

      1530       1540       1550       1560       1570       1580       1590       1600
AGGTAGTAATGCCATTCCAACGGTAAATGGCAAGAAATTGTGGAAGAGAGGAATACTGAAGCACATTCCCAAACAGCAGT
  G  S  N  A  I  P  T  V  N  G  K  K  L  W  K  R  G  I  L  K  H  I  P  K  Q  Q

      1610       1620       1630       1640       1650       1660       1670       1680
ATGATGCATCCACAACAAAAGCAACAAGTGCGCAACTAGCTGCATGGGTTGAGGCCGTTAAGAAGGACCTCAAGATTCGG
Y  D  A  S  T  T  K  A  T  S  A  Q  L  A  A  W  V  E  A  V  K  K  D  L  K  I  R

      1690       1700       1710       1720       1730       1740       1750       1760
AACGACGATGCATGGTCAATTGTCCTGACGGCATGGTGTATCTGGTGTGCAAACAACGGAACCTCTTCTGAGGTGGACAC
 N  D  D  A  W  S  I  V  L  T  A  W  C  I  W  C  A  N  N  G  T  S  S  E  V  D  T

      1770       1780       1790       1800       1810       1820       1830       1840
CAACCAGGACATGGAATCTGACAGCTTGGGGAAAGTGCAGACTGTTCGCATCGATTCATTCGTTGAGCCAGCCATTGAAA
  N  Q  D  M  E  S  D  S  L  G  K  V  Q  T  V  R  I  D  S  F  V  E  P  A  I  E

      1850       1860       1870       1880       1890       1900       1910       1920
ATGGTGGTTTGAGGAAAATCATGAGGCTACTTTTCCGGTATCACTCAGGAAATCTTGGCCAAAGGGGGAAAAATGACAGC
N  G  G  L  R  K  I  M  R  L  L  F  R  Y  H  S  G  N  L  G  Q  R  G  K  N  D  S

      1930       1940       1950       1960       1970       1980       1990       2000
TTATGGAATCAAGCAGGTTTCACAGAGAAAGCGATGACTACCCTACCCTTTGATTTTGTTGAAGTCACCAAAACAACCCC
 L  W  N  Q  A  G  F  T  E  K  A  M  T  T  L  P  F  D  F  V  E  V  T  K  T  T  P

      2010       2020       2030       2040       2050       2060       2070       2080
CAAAACAGTCAAGGAACAACTTGCTCAGGCTAAAATTGCAGCTATAGGGCATGGAACTCGCAGAGCTATGGTGACGGATG
  K  T  V  K  E  Q  L  A  Q  A  K  I  A  A  I  G  H  G  T  R  R  A  M  V  T  D

      2090       2100       2110       2120       2130       2140       2150       2160
GCAGCGTTCATGGAAACAAGACAAGTTACGAACGCCATGTTGACACAGACAACGATGAATCTGAGCATGGTAAGGACATC
G  S  V  H  G  N  K  T  S  Y  E  R  H  V  D  T  D  N  D  E  S  E  H  G  K  D  I

      2170       2180       2190       2200       2210       2220       2230       2240
GATTACAGACCACATCTATCGTAAGTTATTTAAATGAATGAATTTGAATTTATTTATTTAGAGTCCTTTAATAGTTTAAA
 D  Y  R  P  H  L  S  *

      2250       2260       2270       2280       2290       2300       2310       2320
TTAAATAAGCACAAGGTTTGAATGGGGGTTTGTGCCATTGTTCCTCCAATGACCTGAAATGGGTGTTGTTTTATTAACTC

      2330       2340       2350       2360       2370       2380       2390       2400
TGTTTTAATCTTTTCGATGCTTTAAGAGTCGTTTTCTTTAGGCCTGATCTGGATGATCGGCTGCACTTTGGTGAATCCTA

      2410
TTCCGTAGAAGGA(n)
```

Figure 4.8 : CLUSTAL W alignment of the partial ORFs of NLV (GenBank U5870) and MacMV (GenBank U58771) with the C-terminal regions of two bymoviruses, barley mild mosaic virus (BaMMV, GenBank D10947), barley yellow mosaic virus (BaYMV, GenBank D01091), one rymovirus, ryegrass mosaic virus (RGMV, GenBank U27383), one ipomovirus, sweet potato mild mottle virus (SPMMV, GenBank Z48058) and one typical potyvirus, potato virus Y (PVY, GenBank A08776). Gaps (-) have been introduced for maximum alignment and the program BOXSHADE was used to create boxing. Residues identical to MacMV are boxed with a black background, chemically similar residues are boxed in grey.

```
PVY      2580  GNNSGQPSTVVDNSLMVVLAMHYALIKEC--VEFEEIDSTCVFFVNGDDLLIAVNPEKESIL-DRMSQHFSDLGLNYDFSSRTRRKEELWFMSHRGLLIE
RGMV      100  GNNSGQPSTVVDNTIMVIIAMQYAISKAE--FPAGRLRDQIRYFANGDDLVVAVEPSLSDKI-SSFSASFAELGLSYDFSNKVNVRSELQFMSHTGKLID
SPMMV      99  GNNSGQPSTVVDNTLILMIAMEYAIAKVF--VT----RPDIKYVCNGDDLLINCPRSTANAISEHFKDVFADLSLNYDFDHVCDKITDVDFMSHSFMWLD
BaMMV    1808  GNNSGQPSTVVDNTLALMTAFLYAYARLTGDHAFELMDENFVFVCNGDDNKFSMSPSFMAKFGCDFSPFLSELGLTYEFDEATEDICENPYMSLTMVRTS
BaYMV    1972  GNNSGQPSTVVDNTLVLMTAFLYAYIHKTGDRELALLNERFIFVCNGDDNKFALSPQFDEEFGHDFSPELVELGLTYEFDDITSDICENPYMSLTMVKTP
MacMV       1  GNNSGQPSTVVDNTLVLMISFYYAYAVKTRDYTFDKIDERFVFVCNGDDNKFAVSPEFVKEFGGSFTTDEIAQLGLHYEFDELTTDITANPYMSLTMIDIG
NLV         1  ---SGQPSTVVDNTLVLMLSFYYAYAKKTNDLAFEHIDERFRFVCNGDDNKFSVSPEFVAEFGGSFGDEISELGSLTSLMILTPDIMRNPYMSLTIVQVG

PVY      2678  G--MYVPKLEEERIVSILQWDRADLPEHRLEAICAAMIESWGYSELTHQIRRFYSWLLQQQP----FATIAQEGKAPYIASMALRKLYMDRAVDEEELRA
RGMV      198  G--MYIPMLERERICAILEWSRSDEPQFQLDAISAAMIEAWDDDYLLYQIRRYYSWLLDEEP----IKSIAELGHAPYLAEAALKALYTGKNPDAELIAI
SPMMV     194  TEQMYIPKLDKERIVAILEWERSDEQFRTRSALNAAYIESFGYEDLMTEIEKFAHFWAKKHG----LNDV--------LMEREKVRSLYVDENFDASRFEK
BaMMV    1909  FGIGF--SLSIERIVAILQWSRAGGVLHAYLSGIAALFESFNTPKLFNLVHTYLLWLVTEHEEELFSMMELKDMFMPLPTREQIALLHYVGT----E---
BaYMV    2073  FGVGF--SLPVERIIAIMQWSKKGGVLHSYLAGISAIYESFNTPKLFKSIYAYLLWLTEEHEAEILAAMTQSSTALPIPSMLDVYRLHYGDD----EIWL
MacMV     101  GRIGF--QLNPERILGIVQWIKKGGIVHAAQAAFAAMIESFNDPDLFCVMHSYLVWLLVTYRSELVYAMHNDLVSVVYMDPCQVFALHYNDSEDVRE-WF
NLV        98  DRIGF--QLNPERIVGIVQWIKKGGVLHAAQAAFAAMVEAFNDPSLFTVMHTYLVWLLVTHKDVLLIAQENGLGSVCYMDPCQVFALHNGSSKGLEDVKF

PVY      2772  FTEMMVALDDEFELDSYEVHHQANDTIDAGGSNKKDAKPE----------------------QGSIQPNPN----------------------------
RGMV      292  YERAMLNTPPTEDRPTKVVHEANVTAASSAATQTSATSPTVTSTSGASTLTSSGTTSAPIASTTPPVPTATTPPTGTTAPTTPAVRAANLDLAGHRKAKA
SPMMV     283  FY---------PESFSPFDVYVEPHASTSKTIEELQQEMEDLDADT-TITVVQRETQKAGIRDQIEALRAQQI----------VRPPEAQL---------
BaMMV    2000  ----PI----------------------------------VEE---------TFLQAGH--EEPDPIVPPVPDTDLTNMAAAPPD----------------NR
BaYMV    2167  QAADPLTDAQKE---AAHTAAADRARLDLADADRRRKVEA---------DRVEAARVKKAADAVLKPVTLTA-TRMPTEDDG----------------KL
MacMV     198  DEDDESSDDEDEE--PTQVLQMDAETLAKDGEAKKEKDEK---------EREKAEQRRVEVEK--ARAEKAQVSDGAKEPQP----------------EI
NLV       196  DNEDESADEDDGNITPDLELQMDVGNLIPE----KEKNSQ---------NVNTSDGGKNASNS--ATGESSKPPENKAGKGA----------------D-

PVY      2821  ---------KGKDKDVNAGTSGTHTVPRIKAITSKMRMPTSKG--ATVLNLEHLLEYAPQQIDISNTRATQSQFDTWYEAVRMAYDIGETE-MPTVMNGLM
RGMV      392  NGESQLNVRGENDDEDVPAASEFALPRLPTLGAKIRVPKFKG--VIVLNKDHLIKYTPDQRDLSNTRATQEQFEKWYSGVRNEVEKTDEE-MALLLNGFM
SPMMV     354  ----------QPDVTPAQIVTFEPPRVTGFGA-LWIPRQQRNYMTPSYIEKIKAYVPHSNLIESGLASEAQLTSWFENTCRDYQVSMDVFMSTILPAWI
BaMMV    2038  KSRAV--IPRGTSD---------WSLPEPKMRTLGFKSKINIET----------LANVPDGYMNTFASVATESQRRKWEEAARGDFGITDDEKWEKLLIAAC
BaYMV    2238  KTPSGARIPSSAADGN-------WSVPATKQVNAGLTLKIPLNK---------LKSVPKSVMEHNNSVALESELKAWTDAVRTSLGITTDEAWIDALIPFI
MacMV     269  KGNEDVEQPASDPEEK-EEEVKWVMPSINPNRGSNAIPTVNGK--KLWKRGILKHIPKQQYDASTTKATSAQLAAWVEAVKKDLKIRNDDAWSIVLTAWC
NLV       264  --QGDVDPPQGDPLVD-DEVVEWVIPKMSPNIGTSPIPVINGK--RLWKRGVLKSIPKMMFNTTSTMATQAQLTSWVEEVKQALALKTDDAWTVVITNWC

PVY      2910  VWCIENGTSPNVNG--VW-----VMMDGNEQVE-YPLKPIVENAK--PTLRQIMAHFSDVAEA-YIEMRNKKEPYMPRYGLIRNLRDMGLARYAFDFYEV
RGMV      489  VWCMENGTSPDLSG--SW-----TMMEGEEQIS-YPLGPFCRHAQ--PTLRSIMAHFSDAATA-YVVLRNQKSRYMPRYGLKRGLNDYSLAPYAFDFYEI
SPMMV     442  VNCIINGTSQERTNEHTWRAVIMANMEDQEVLY-YPIKPIIINAQ--PTLRQVMRHFGEQAVAQYMNSLQVGKPFTVKGAVTAGYANVQDAWLGIDFLRD
BaMMV    2119  IYFADNGTSPNFDEELT-----MEVNGGLNSIKEYPVRPFVVRAKKISTLRRIFRC-YSIET----KLMFVKLRRVPHWAIKHGCLDE--I--VFDFMIP
BaYMV    2323  GWCCNNGTSDKHAENQV-----MQIDSGKGAVTEMSLSPFIVHARMNGGLRRIMRN-YSDET----VLLITNNKLVAHWSMKHGASAN--AKYAFDFFVP
MacMV     366  IWCANNGTSSEVDTNQD-----ME-SDSLGKVQTVRIDSFVEPAIENGGLRKIMRLLFRYHS----GNLGQRGKNDSLW-NQAGFTEKAMTTLPFDFVEV
NLV       359  IWCANNGTSSEVDTSQT-----MEIRDGFGKVQAIPIEVFVNPAVENGGLRKIMRH-FSGIT----HEILKAGKRMTAWGNKRGFTEKSMIPYAFDYYVV

PVY      2999  TSRTTPVRAREAHIQMKAAALKSAQPRLFGLDGGISTQEENTERHTTEDVSPSMHTLLGVK---NM
RGMV      578  TSTSPLRARERHAHMKAAAAIRGKASRMFGLDGNVSAQSENTERYTVEDVNTRVHSLSGAN---ML
SPMMV     539  TMKLTTKQMEVKHQIIAANVTRRKIRVFALAAPGDGDELDTERHVVDDVARGRHSLRGAQ---LD
BaMMV    2205  DQFTSRTALETLKQTKLAAIGVGTSNSLLTSEQTNMRTTETRRRNDYD----------GHEALLR
BaYMV    2411  RSWMNPQDIEVSKQARLAALGTGTYNTMLTSDTTNLRKTTNHRVLDSD----------GHPEL-T
MacMV     455  TKTTPKTVKEQLAQAKIAAIGHGTRRAMVTDGSVHGNKTSYERHVDTDNDESEHGKDIDYRPHLS
NLV       449  TNTTPKTVREQLAQSKAAAIGSGVTRKMVLDGNIQGSHASYERHVDTDNSEYEHGNDVDQRPYLT
```

Figure 4.9 shows a neighbour-joining dendogram constructed using the percentage identities obtained for the core coat protein showed in table 4.1. It demonstrates that NLV and MacMV are located within a separate grouping from the bymoviruses, rymoviruses, ipomoviruses and potyviruses.

NLV isolated from Iris and Gladioli

Sequence was obtained from the Iris and Gladioli clones using two internal primers designed to pNLVdT1 (NLV3, NLV5). Of the 1.1kb inserts of these clones, 240bp was sequenced on both strands of two clones. Figure 4.10 shows the sequence obtained in this region. As there were no nucleotide differences between the two Iris and the two Gladioli clones, this sequence represents all four clones. A partial ORF of 37 amino acids is present. There were no nucleotide differences between these clones and the same region in NLV, and the position of the Gladioli and Iris clones are shown on figure 4.4a. The sequence obtained from the limited section of Iris and Gladioli clones was identical to that obtained for NLV isolated from *N. clevelandii* and narcissus material.

pM4 and pM29

Figure 4.11 shows the nucleotide sequence obtained for pM4 and pM29. The clone pM29 was found to be equivalent to position 4 - 239 nt within pM4. There were no nucleotide differences between pM29 and pM4. A total of 1149 bp were sequenced from pM4, and analysis showed a partial ORF from an initiating methionine. Two AUG codons could be the initiating codon, producing a protein of either 359 or 356 amino acids (40.2 kDa or 39.8 kDa). There was therefore a possible 5' untranslated region (UTR) of either 70 or 79 nt. The ORF did not contain a termination codon, and therefore there was no 3' UTR.

When the partial ORF represented by these two clones was submitted to a database search, it was shown to have homology with the N-termini of the polyproteins generated by RNA2 of BaMMV and BaYMV. It also had homology in a limited area with several potyvirus helper component proteinases (HC-Pro) around the motif recognised as the active site of the proteinase encoded by these ORFs.

Figure 4.12 shows an alignment of the partial ORF encoded by pM4 with the equivalent region encoded by RNA2 of BaMMV and BaYMV. The second methionine residue encoded by the pM4 sequence has been used as the initiating methionine, as this gives maximum alignment with the bymoviruses. It has been suggested that BaYMV and BaMMV could produce two mature proteins from the *c.* 98 kDa polyprotein translated from RNA2 (Davidson *et al.*, 1991; Kashiwazaki *et al.*, 1991; Timpe & Kuhne, 1994).

Figure 4.9: Neighbour-joining tree produced using CLUSTAL W to demonstrate the taxonomic relationship between the core of the coat proteins of MacMV and NLV and those of selected members of the *Potyviridae* (as used in Table 4.1). Horizontal distances are proportional to percent divergence of tip species from nodes (scale given), vertical distances are arbitrary. Numbers shown on the branches are the results from bootstrapping using CLUSTAL W for 1000 replicates.

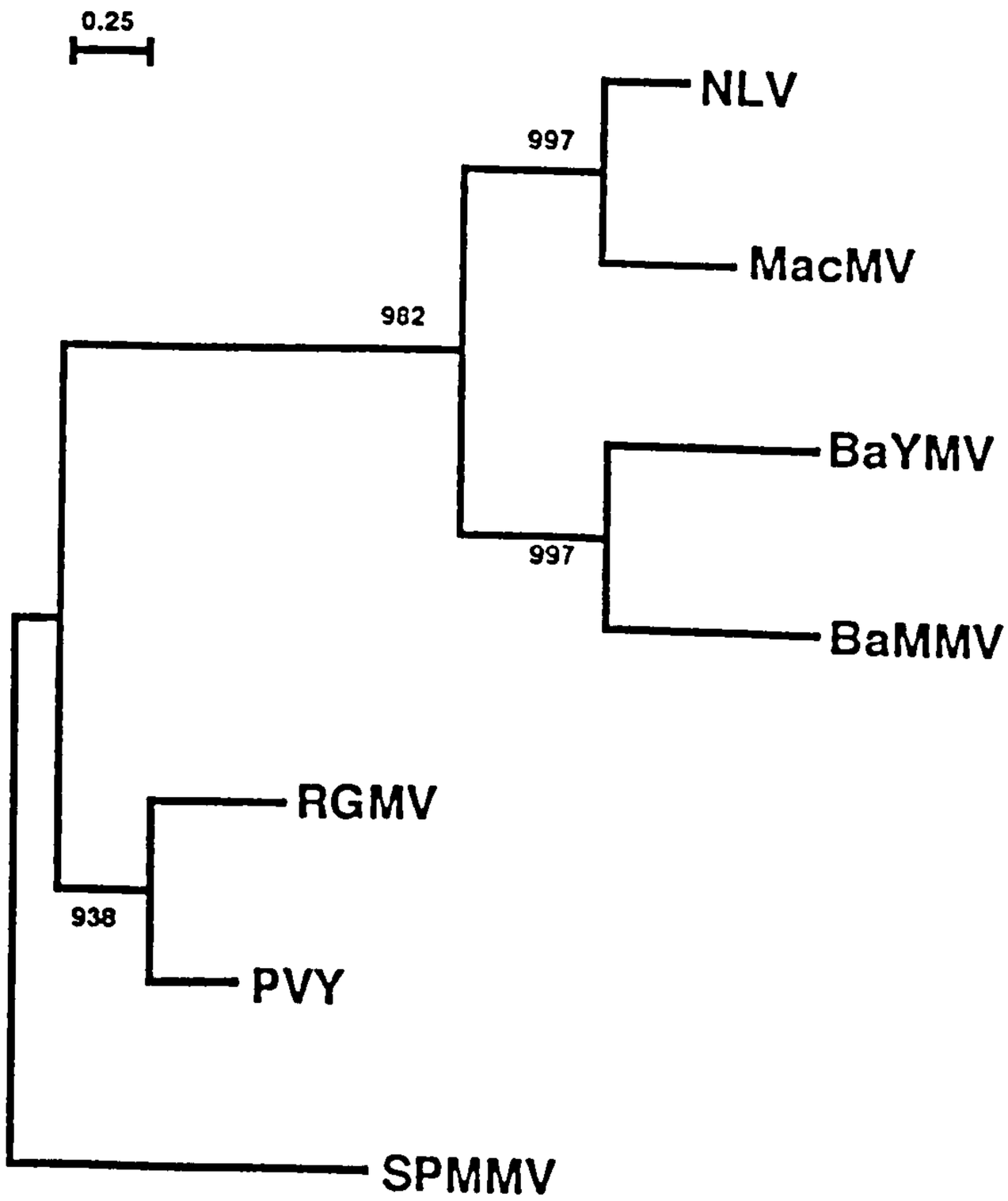


Figure 4.10 : The DNA sequence corresponding to the 3' terminal region of NLV positive-sense RNA from infected Gladioli and Iris samples, with derived amino acid sequence indicated below. The positions and sequence of NLV 3 and NLV 5, primers used to sequence this section are underlined.

```
          10         20         30         40         50         60         70
GGTTCTTGATGGGAACATACAAGGGTCTCATGCCAGTTACGAACGCCATGTCGATACGGACAACAGCGAGTACGAGCA
 V  L  D  G  N  I  Q  G  S  H  A  S  Y  E  R  H  V  D  T  D  N  S  E  Y  E  H

 80         90         100        110        120        130        140        150
TGGCAATGATGTTGATCAAAGGCCCTATCTTACTTGAATTCTGCGTTCTTGAACGCTCGAATTCAACTATTTGCATTC
 G  N  D  V  D  Q  R  P  Y  L  T  *

   160        170        180        190        200        210        220        230
GCTCTTTATAAGTTTTCTCCTAAGCTTATTTCTTTTCTCTCTCTCTTGAGCGTTAAGGGTTAAATGGGGGTTTATTTG

      240
CCATT
```

Figure 4.11: The DNA sequence corresponding to part of the 5' terminal region of MacMV positive-sense RNA, with derived amino acid sequence indicated below. Two possible methionine initiation codons are indicated in the translation, however, the second (underlined) appears to conform to a consensus shared with BaMMV and BaYMV RNA2 initiation codons. A possible cleavage site (VG/G) is underlined.

```
          10         20         30         40         50         60         70
GCAATGGCAACTGTTAATAGCAAACGTCAATTTTAAAGTTGAAAACGTCGACACCCTAAACACTCACCTCATGCAACT
                                                                  M  Q  L

 80         90         100        110        120        130        140        150
CATGTCGGGGTCTGCTCTAACTGTGCAGCCAATAAGTGCTCAAGCATTAATGCGCATGGTTAAAGCTCACTTTGAGCA
 M  S  G  S  A  L  T  V  Q  P  I  S  A  Q  A  L  M  R  M  V  K  A  H  F  E  Q

   160        170        180        190        200        210        220        230
AAACAAGATGATTAAGGGGAGTGAATTTGTCATGGTGGATGGCAATGGGAAGACCCTCGCAACATTCAACACATCAAT
 N  K  M  I  K  G  S  E  F  V  M  V  D  G  N  G  K  T  L  A  T  F  N  T  S  M

     240        250        260        270        280        290        300        310
GAGGAGCAACGCAGATTTGCAAACCTTTATGAAAGGAGAAATTAAATTCCCAGCACAGCAGAGTTTTGCAGACTGTGA
 R  S  N  A  D  L  Q  T  F  M  K  G  E  I  K  F  P  A  Q  Q  S  F  A  D  C  D

       320        330        340        350        360        370        380        390
CAAGATTGTTAAGTGCACCAAAATGCACCAACAGCATGTTGCTCACACACGACTGCCTGCAGCTAACTATGTGTTCGG
 K  I  V  K  C  T  K  M  H  Q  Q  H  V  A  H  T  R  L  P  A  A  N  Y  V  F  G

         400        410        420        430        440        450        460
GGGTTACTCAAAATCATGCCCCATCCCACTCTTCAAGGAGCAGGTAGCTAAACAGGGGAGATGTGTGAGCACATTACT
 G  Y  S  K  S  C  P  I  P  L  F  K  E  Q  V  A  K  Q  G  R  C  V  S  T  L  L

 470        480        490        500        510        520        530        540
GCTGGCTCTAAGTTACTTTGTTACAAAAGAGTTTGATGAGGTTTACGCAAGAATCATAGATGAGGTTATCACAACAAT
 L  A  L  S  Y  F  V  T  K  E  F  D  E  V  Y  A  R  I  I  D  E  V  I  T  T  I

   550        560        570        580        590        600        610        620
CAAGAGTTGGCCCACACTCCATGATGTTTCTAAACTCTGCCAGTACATCATCTCGAAAGTCCCAATTTTAGGGCCGGT
 K  S  W  P  T  L  H  D  V  S  K  L  C  Q  Y  I  I  S  K  V  P  I  L  G  P  V

     630        640        650        660        670        680        690        700
GCCTATTCCCGTGGTAGCTGTTGATCATAGCAAAAAACTAATTCACATCTGTGATCAGCGAGGCATACCTGATGGGTG
 P  I  P  V  V  A  V  D  H  S  K  K  L  I  H  I  C  D  Q  R  G  I  P  D  G  W

       710        720        730        740        750        760        770        780
GCATCAGCTGAAGATGGGGACGCTCGTTGAAATCGCAAACGCTGGGCTCCTCAAACAAAGCGTGCTCCATTCATATTT
 H  Q  L  K  M  G  T  L  V  E  I  A  N  A  G  L  L  K  Q  S  V  L  H  S  Y  F

         790        800        810        820        830        840        850
TGTTGGTGGGACTAACGAGCAACCGGAGGACCTCCTATACGTGCGTGGTTTAAGTAAGATCAAACGCAAGTTGCGTGG
 V  G/ G  T  N  E  Q  P  E  D  L  L  Y  V  R  G  L  S  K  I  K  R  K  L  R  G

 860        870        880        890        900        910        920        930
CTGGGTCACGCAGAGGAACATGATGGATGATTTTCGTGGAAGATTTAGACTTAGCAGCATTTTCATGATGCTATCACC
 W  V  T  Q  R  N  M  M  D  D  F  R  G  R  F  R  L  S  S  I  F  M  M  L  S  P

   940        950        960        970        980        990        1000       1010
ATCAATGTTGTCGCGACTCAAAGCGCTCATTGAGGCGGAGGACGCATCGTCGCTCAAGCTCATTATGATGGATGATGC
 S  M  L  S  R  L  K  A  L  I  E  A  E  D  A  S  S  L  K  L  I  M  M  D  D  A

     1020       1030       1040       1050       1060       1070       1080       1090
ATGCTCAAATAAAATTGTCGTTGCTGCTGCCCTGCGAACAGCAATACAAGGCGTGATACTCCAAACTGGTGAGAGCCG
 C  S  N  K  I  V  V  A  A  A  L  R  T  A  I  Q  G  V  I  L  Q  T  G  E  S  R

         1100       1110       1120       1130       1140
AATAGAGAGGATCTGGCTAAATGTGAACAACATTTTGCAGGATCAATTGAATGGAGA
 I  E  R  I  W  L  N  V  N  N  I  L  Q  D  Q  L  N  G
```

Figure 4.12: CLUSTAL W alignment of the partial ORFs of MacMV with the N-terminal regions of RNA 2 of two bymoviruses, barley mild mosaic virus (BaMMV, GenBank X75933; German ASL strain; Timpe & Khune, 1994), barley yellow mosaic virus (BaYMV, GenBank Q01207: Japanese II-1 strain; Kashiwazaki *et al.*, 1991). Gaps (-) have been introduced for maximum alignment. Residues identical to MacMV are marked (*), chemically similar residues are are marked (|).

```
          •                              |   |    |           |    | |  |
BaYMV   1 MSASSSRQLFDCGSLDWPNKSLFGDPTTRDVMHEHISSTWNAVIRRHMLAPNADAETILGRDGL
BaMMV   1 M-----------------------------MMNSMIRQGWQQVLRRFSIPTSGDRLIVSNSTDQ
MacMV   1 MSGSA----------------LTVQPISAQALMRMVKAHFEQ-----NKMIKGSEFVMVDGNGK

            | •||       | •        |      |                |  ||     | |
BaYMV  65 PSAQFDAYGAMLPSFIQALNAPTTRLRITHRCPT---AESILCADASHAPWLYMANNVCAYEAT
BaMMV  36 PIGLFGAFDTSLQTLSQVTNDPEV-LKQKSNIPTHLDVASVLEASPRSFPWVFLTNSFCTFGGS
MacMV  44 TLATFNTSMRSNADL-QTFMKGEIKFPAQQSFADCDKIVKCTKMHQQHVAHTRLPAANYVFGGY

               |  •       • •   •|||••| |       |  |||||  |  | •|•  | |
BaYMV 126 -HLKPVQTFIAFDFAHGYCYLSLFIPLSFRITFENARSFSRFLEQLPDILGAYPTLAAIYKTML
BaMMV  99 IHAQNLQAFATAEFKSGFCYMNLLIPLSFDIIDAHADSFRGFVEQLPDTLGAYPSLSMVLNVML
MacMV 107 SKSCPIPLFKEQVAKQGRCVSTLLLALSYFVTKEFDEVYARIIDEVITTIKSWPTLHDVSKLCQ

                • |   ••• |•       |•| • ••|•  •  |•     || |     |
BaYMV 189 FAIRLFPEVLQAPIPIIAKRPGVLQFHVSDARGLPPSWFPMKCGSVRSFVALITNNLNSDLLDG
BaMMV 163 HAATRFPEIVASPIPTIAFDAESLQFHVTDKRGVPGMWNILKACRVYELLSLAADGIGCEYM--
MacMV 171 YIISKVPILGPVPIPVVAVDHSKKLIHICDQRGIPDGWHQLKMG---TLVEIANAGL---LKQS

           |         |    | ||   | ||    |     • |       ||   ||
BaYMV 253 IVGSNGDGEHYTNWNSGHDHWIVNRFITVRDLHSSLKSALDVDLDTEGGRNAVLDLLLDLGVTN
BaMMV 225 -LYPVGAAPQYSFWKKSMDHFTSDRFVEFLAMQDLLASALEQDYATHDARDALLSALQNAGYTN
MacMV 229 VL--------HSYFVGGTNE-QPEDLLYVRGL-----SKIKRKLRGWVTQRNMMD---------

                 •    |    | | ••   ||         •      | |  |   | |
BaYMV 317 LVRREKRFPAHFQGAESVYLLLSCERVGNELVAVQDALQEPLANHSGLDLRALIIN--LGGLPS
BaMMV 288 VVARERRFPNG-HDPSIVWLNLSEAPISEKLTELKRYLLVGHRSDDTADITHNVHQHVFEVLKT
MacMV 270 ------DFRGRFR-LSSIFMMLSPSMLSR--------LKALIEAEDASSLKLIMMDDACSNKIV

              |            •     |•                       ||
BaYMV 379 RHSDIC-----YTRNIFENDNHLVWN------------------F
BaMMV 351 MSVQFSKTTNAYNRARFEVNHKVIWNAEYGRGPQQNAELEALVLFLNRQS
MacMV 319 VAAALRTAIQGVILQTGESRIERIWL-------NVNNILQDQL-----NG
```

Figure 4.13: CLUSTAL W alignment of the partial ORFs of and MacMV with the N-terminal regions of RNA 2 of two bymoviruses, barley mild mosaic virus (BaMMV, GenBank X75933; German ASL strain), barley yellow mosaic virus (BaYMV, GenBank Q01207: Japanese II-1 strain), and the Helper Component-Proteinase (HC-Pro) regions from the polyprotein of several potyviruses, potato virus Y (PVY, GenBank A08776; Robaglia *et al.*, 1989), plum pox virus (PPV, PIR S47508; Maiss *et al.*, 1989) and tobacco vein mottling virus (TVMV, PIR g75614; Domier *et al.*, 1986). Gaps (-) have been introduced for maximum alignment. Residues identical to MacMV are marked (*), chemically similar residues are are marked (|).

```
                          • •   |||||   |         |   | |      |   •||   |
PVY          613 GDSEMLYIAKQGYCYINVFLAMLINISEEDAKDFTKKVRDNCVPKLGTWPTMMDLATTCAQMRIF
PPV          639 AKGGAMFIAKAGYCYINIFLAMLININEDEAKSFTKTVRDTLVPKLGTWPSMMDLATACHFLAVL
TVMV         586 EISNLMYIAKEGYCYINIFLAMLVNVDEANAKDFTKRVRDESVQKLGKWPSLIDVATECALLSTY
BaYMV        130 VQTFIAFDFAHGYCYLSLFIPLSFRITFENARSFSRFL-EQLPDILGAYPTLAAIYKTMLFAIRL
BaMMV        104 LGAFATAEFKSGFCYMNLLIPLSFDIIDAHADSFRGFV-EQLPDTLGAYPSML-VLNVMLHAATR
MacMV        112 IPLFKEQVAKQGRCVSTLLLALSYFVTKEFDEVYARII-DEVITTIKSWPTLHDVSKLCQYIISK

                  •       |• |          •| •  •          |•       |  |    |          ••
PVY          678 YPDVHDAELPRILVDHDTQTCHVVDSFGSQTTGYHILKASSVSQLILFANDEL--ESDIKHYRVGG
PPV          704 YPETRNAELPRILVDHEAKIFHVVDSFGSLSTGMHVLKANTINQLISFASDTL--DSNMKTYLVGG
TVMV         652 YPAAASAELPRLLVDHAQKTIHVVDSYGSLNTGYHILKANTVSQLEKFASNTL--ESPMAQYKVGG
BaYMV        194 FPEVLQAPIPIIAKRPGVLQFHVSDARG-LPPSWFPMKCGSVRSFVALITNNL--NSDLLDGIVGA
BaMMV        168 FPEIVASPVPTIAFDAESLGFHVTDKRG-VPGMWNILKACRVYELLSLAADGI--GCEYMLYPVGS
MacMV        176 VPILGPVPIPVVAVDHSKKLIHICDQRG-IPDGWHQLKMG---TLVEIANAGLLKQSVLHSYFVGG
```

Table 4.1 (a) Pairwise percent amino acid sequence identities between the core coat protein* of MacMV and NLV and other selected members of the *Potyviridae*.

	MacMV	BaYMV	BaMMV	SPMMV	RGMV	PVY
NLV	53.4	23.2	22.2	14.6	22.2	21.2
MacMV		22.2	21.2	15.1	18.0	19.4
BaYMV			33.7	11.0	11.5	13.4
BaMMV				11.1	12.9	14.8
SPMMV					18.0	18.9
RGMV						54.6

* The core coat protein is equivalent to D_{2748} to R_{3043} in PVY.

Table 4.1 (b) Pairwise percent amino acid sequence identities between the †partial nuclear inclusion body b (NIb) protein of MacMV and NLV and other selected members of the *Potyviridae*.

	MacMV	BaYMV	BaMMV	SPMMV	RGMV	PVY
NLV	76.3	57.2	60.3	23.0	21.3	26.7
MacMV		60.4	59.7	30.7	28.2	31.2
BaYMV			73.1	27.6	28.2	28.2
BaMMV				30.7	29.0	32.0
SPMMV					38.4	35.3
RGMV						51.1

† This region is equivalent to G_{2580} to A_{2711} in PVY. This region was used since cleavage sites between NIb and coat proteins have not been proposed for RGMV or SPMMV.

A 28 kDa protein has been observed in *in vitro* translation studies, and deletion analysis has confirmed that this protein is produced from the N-terminus of the polyprotein (Davidson *et al.*, 1991). This implies that a cleavage site (VG/S) which is similar to that found in HC-Pro of potyviruses is responsible for the production of this 28 kDa product. A similar motif exists in the MacMV sequence, VG/G (figure 4.11: position 782 nt), which would create a protein of 26.2 kDa that included the second methionine initiation residue. This cleavage site is identical to that found in potyviruses (VG/G) (Carrington & Herndon, 1992), and distinct from the bymovirus motif (VG/S).

Within this mature protein of 26 kDa, there are two motifs that are conserved between MacMV, the bymoviruses and the potyviruses. These centre around the two active residues of the Helper Component proteinase (HC-Pro), cysteine (C) and histidine (H) (Oh & Carrington, 1989). Figure 4.13 shows an alignment of this conserved region, and demonstrates that although the bymoviruses have bipartite genomes, they share homology with the potyvirus HC-Pro. However, neither the bymoviruses nor MacMV show homology to the N-terminus of HC-Pro which is involved with aphid transmission.

Overall in this 5' region, there is only limited identity between MacMV and the bymoviruses (13%) although this is not a region of high homology between BaYMV and BaMMV (25%). In the HC-Pro region, around the two active residues (C, H), there is a greater degree of identity between the bymoviruses (41%). In this region, MacMV is more similar to the bymoviruses (21-22%) than it is to the potyviruses (15-19%). However, this homology is low in comparison to that displayed amongst the potyviruses in the HC-Pro active site region shown in figure 4.13 which is 60-62%.

Serological relationship with BaMMV

NLV and MacMV coat proteins did not cross-react with antisera raised to BaMMV coat protein. This antisera was supplied by John Antoniw (Rothamsted). The reciprocal western analysis where BaMMV coat protein was subjected to NLV antisera also proved to be negative. Since both these results were negative, they are not presented here.

Discussion

The conflicting morphological and physio-chemical properties of NLV and MacMV in combination with a lack of serological relationships had led to the suggestion that they remain unclassified (Mowat *et al.*, 1991). Further investigation of their coat protein size and genome sequences has enabled a more definite taxonomic proposal.

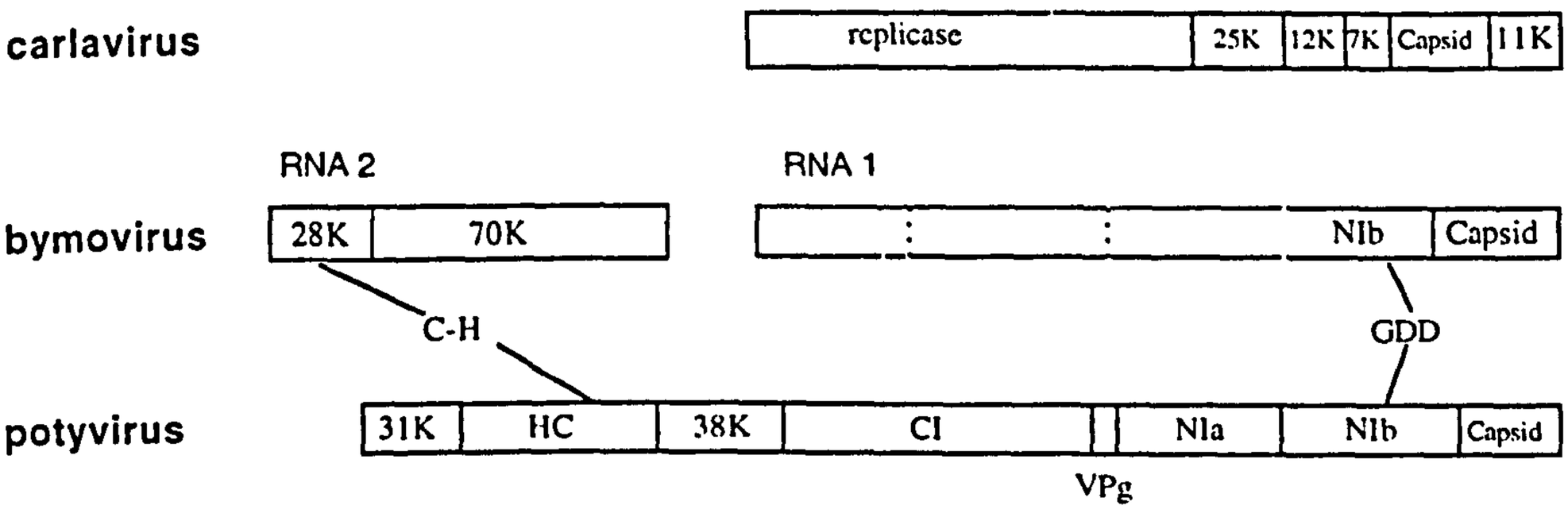
Previous reports have suggested that NLV and MacMV have properties which appear to be intermediate between the carlaviruses and the potyviruses. With the size of their coat proteins larger than those expected of either group, more information was needed to establish if the proposed size of 45 kDa (Mowat *et al.*, 1991; Plese & Wrischer, 1978) was correct.

Coat protein sizes

By generating cDNA clones of the coat protein coding region of both viruses, it was possible to estimate more accurately the encoded molecular mass of the virus sub-units. The deduced amino acid sequences from this region contained a single ORF. This is analogous to a potyviral genome organisation (figure 4.14), rather than a carlavirus one which has an 11K protein with a methionine start and usually a stop codon at the end of the coat protein which proceeds it (Foster, 1992). Alignments using all carlavirus coat proteins sequenced to date show only 9-12% sequence identity to NLV and MacMV coat proteins (alignment not presented). Database searches over the whole region sequenced for NLV and MacMV revealed a much higher level of identity to the bymoviruses, which also utilise a polyprotein processing system. We therefore concluded that the ORFs found in the deduced amino acid sequence of the 3' terminal region of NLV and MacMV were the C-termini of large polyproteins. To estimate the size of NLV and MacMV coat proteins, it is therefore necessary to know the position of the cleavage site between the nuclear inclusion body b (NIb) protein and the coat protein.

Alignment of the deduced amino acid sequence for NLV and MacMV partial polyprotein shows that in both sequences there is a possible cleavage site LQM (figure 4.8). Although the consensus potyvirus cleavage site is a heptapeptide cleaving between glutamine (Q) and serine (S) or glycine (G), the bymoviruses are thought to have a consensus cleavage site between the coat protein and NIb of glutamine and alanine (A) (Kashiwazaki *et al.*, 1990 a; Foulds *et al.*, 1993). Only one exception to this consensus has so far been found, a Japanese strain of BaMMV, Na I, has a cleavage site of Q/S (Kashiwazaki *et al.*, 1992; Kashiwazaki, 1996). Particular N-termini may be more resistant to degradation than others, including alanine (A) and methionine (M) (Dinant *et al.*, 1991), observations thus supporting the hypothesis that LQM is a potential cleavage site. The N-termini of potyvirus coat proteins are highly diverse and the C-terminus of the NIb protein is known to be well conserved (Shukla & Ward, 1988). With a cleavage site of LQM, NLV and MacMV deduced amino sequences follow this principle, with a good level of similarity in the region up to the cleavage site, and much lower identity downstream (see figure 4.8).

Figure 4.14: Diagrammatic representation of the genetic organisation of carlaviruses, bymoviruses and potyviruses. Common motifs which have been sequenced in MacMV are shown in the centre with their relative positions in the genome indicated.



This putative cleavage site would produce coat proteins of a predicted size (32.8 kDa NLV, 34.1 kDa MacMV) smaller than those we have observed by Coomassie-Blue stained SDS-PAGE or western analysis. The difference in size of coat proteins observed by western analysis and that expected by deduced amino acid sequence has been noted previously (Shukla *et al.*, 1994), and an artefact of gel analysis could account for this discrepancy. The N-terminal micro-sequencing of MacMV coat protein did not yield the sequence expected for the LQM cleavage site, but instead a sequence of 'SDPEE' which is situated elsewhere the coat protein. However it was later found that the 32 kDa protein that was used for this analysis was probably degraded when it was compared to the size of the coat protein estimated by western analysis (40 kDa).

Western analysis using the antisera raised to MacMV coat protein consistently produced a contaminating band of 54 kDa. This contamination did not occur when antisera raised to NLV coat protein was used on MacMV-infected samples, where a single band of 40 kDa was obtained due to the serological relationship between NLV and MacMV. This contaminating band present in healthy tissue was assumed to be RUBISCO after it reacted with antisera raised to the large sub-unit of RUBISCO. Pre-absorption of the MacMV antisera with healthy sap may have eliminated this problem. It is possible that RUBISCO could have been co-purified with a virus preparation of MacMV particles, and thus contaminated the original antiserum. It has been previously noted that both NLV and MacMV virus particles have shown unusual granules attached to the particle surface on electron micrographs (Mowat *et al.*, 1991; Plese *et al.*, 1979). These particles may be the result of the association of a host protein, such as RUBISCO and therefore may be expected to co-purify and offer an alternative explanation for the problems associated with the analysis of these coat proteins.

Taxonomy of MacMV and NLV by sequence analysis

Coat protein sequence

NLV and MacMV have few serological relationships with any plant viruses other than themselves. It was previously observed that MacMV antiserum decorated NLV virus particles in infected sap (Mowat *et al.*, 1991). We have also demonstrated that NLV and MacMV are serologically related. NLV or MacMV coat proteins do not react with BaMMV antisera, and BaMMV coat protein does not react with NLV antisera. This lack of serology has meant that NLV and MacMV coat proteins have not been useful tools for their classification (Mowat *et al.*, 1991; Plese *et al.*, 1979).

The elucidation of the amino acid sequence of their coat proteins and their comparison to established databases of other viral coat proteins has shown that NLV and MacMV are not related to the carlavirus genus but to a genus of the *Potyviridae*.

In general, the potyviruses have a very good degree of identity in the coat protein ORFs, and this is particularly striking in a defined core region (Shukla *et al.*, 1994). This homology has allowed the construction of a set of parameters by which the relationships of potyviruses can be judged. Shukla and Ward (1988) have shown that distinct species of the potyvirus genus have coat protein sequence similarities within the range of 38% to 71% (average 54%). Later studies showed that if a trypsin resistant core was used, the distribution of frequencies was refined to allow the delineation of four separate taxa (Shukla *et al.*, 1994). Using an equivalent to this core coat protein region for percentage amino acid identity analysis, NLV and MacMV are highly related to one another (53.4%) and have very low identity to all the other accepted genera within the *Potyviridae* family (14-23%) (see table 4.1a). This result would support the hypothesis that NLV and MacMV should form a new genus within the family *Potyviridae*. This is further supported by the dendogram of sequence relationships generated from the alignment of the coat protein core amino acid sequences which demonstrates that NLV and MacMV do not belong to any recognised genus in the *Potyviridae*.

There is further evidence to support their inclusion in the potyviridae family. An important motif, 'NGTS' (NLV, 363 aa; MacMV, 578 aa) is found in NLV, MacMV, bymoviruses and potyviruses and has been used as a part of a universal potyvirus primer (Pappu *et al.*, 1993). Moreover, Gibbs and MacKenzie (1997) recently subjected NLV and MacMV to analysis using a pair of RT-PCR primers designed to amplify part of the genome of all potyvirids. The primers amplified a fragment of the expected size for NLV and MacMV, agreeing with the suggestion that they are probably unassigned members of the *Potyviridae*.

Nuclear inclusion body b (NIb) protein

Since the core coat protein sequence gives information only that NLV and MacMV should form a new genus, it is necessary to look at other proteins for further information. The NIb proteins (figure 4.8) support the new genus theory. Identity between NLV and MacMV of 76.3% is of a similar order to that observed between distinct members of the same genus (Shukla & Ward, 1988). It is here that there is a marked difference between the degree of relatedness of NLV and MacMV to potyviruses, bymoviruses, ipomoviruses and rymoviruses. On average, NLV and MacMV show 26-31% similarity to typical potyviruses, such as PVY, and similar

levels of identity to ipomoviruses and rymoviruses. They not only show a much greater degree of identity to the bymoviruses (57-60%), but are also similarly conserved around important motifs. The motifs ([T/S]GXXXTXXXN[T/S]) and GDD are highly conserved in many RNA viruses, including all potyviruses (Kamer & Argos, 1984; Domier *et al.*, 1987; Poch *et al.*, 1989) and GDD is thought to be part of the active site for the RNA-dependent RNA polymerase activity of the NIb protein (Inokuchi & Harishima, 1987;1990). NLV and MacMV also contain these motifs and share identity around the central GDD domain with the bymoviruses, but not with the potyviruses, rymoviruses or ipomoviruses. This area unique to the bymo- and macluraviruses is a region suitable for the design of a PCR primer which would allow the identification of new macluraviruses. The identification and sequencing of more viruses similar to MacMV and NLV would aid the definition of the parameters of this new genus. Such a primer was constructed and tested using rice necrosis mosaic virus (RNMV), a member of the bymovirus genus which had not been sequenced (see Chapter 6).

Sequence evidence for aphid transmission

This degree of homology between MacMV, NLV and the bymoviruses leads us to question the biological and morphological characteristics of NLV and MacMV. The bymoviruses are bi-partite, fungus transmitted viruses, and NLV and MacMV have always been defined as monopartite aphid transmitted viruses. There are several amino acid motifs associated with aphid transmission which the bymoviruses do not share with the potyviruses.

Coat protein

A tripeptide motif for aphid transmission, 'DAG', is always found within a heptapeptide block at or near the N-terminal of the coat protein of aphid transmitted potyviruses (Atreya *et al.*, 1990; 1995). It was demonstrated by mutagenesis that this motif was essential for vector transmission (Atreya *et al.*, 1995). Two motifs similar to 'DAG' can be found after the putative cleavage site in NLV and MacMV. At position 653nt, NLV has the sequence DVG, and MacMV has the sequence DAE at position 1276nt (figures 4.6 & 4.7). However, both tripeptides are positioned within the first 2 residues of the putative coat protein and therefore may not form part of the suggested heptapeptide motif. It was also shown that in tobacco vein mottling potyvirus (TVMV) a mutation to DAE from DAG as observed for MacMV, gave a non-transmissible product, and that a mutation to DVG, which is found in NLV left only minimal transmission function (Atreya *et al.*, 1990; 1995). This evidence suggests that these isolates of NLV and MacMV may have once been aphid transmissible, but now exist as mutated non-transmissible forms. The isolates sequenced have not been tested for aphid transmission, so this theory cannot be confirmed.

Although two major motifs recognised to be necessary for aphid transmission are missing or mutated from NLV and MacMV, and there is no evidence to suggest that they have bi-partite genomes, there is equally no evidence to suggest that they are transmitted by the same vector as the bymoviruses. Clearly more work is necessary in this area, including the testing of the sequenced isolates of NLV and MacMV reported in this work for aphid transmission. Original work showing that NLV was aphid transmitted used the NLV-infected *N. clevelandii* , not narcissus and transmission by *Myzus persicae* occurred in only 1/10, 2/4 and 7/9 plants tested (Mowat *et al.*, 1991; Brunt, 1977). Since, at the time of testing, it was assumed that NLV belonged to an aphid transmitted virus genus, then a low transmission rate was taken as evidence for vector transmission. Until these isolates are examined in a new light, it is impossible to tell whether these results were misleading.

Helper Component (HC)

In potyviruses, a second protein, the Helper Component (HC) is usually associated with vector transmission and specificity. This protein is situated at the 5' end of the genome (figure 4.14). The N-terminus of the mature HC protein is thought to be responsible for vector specificity, though the mechanism used is not known. Naturally occurring virus isolates which are non-aphid transmissible (NAT) have been sequenced in order to establish whether a mutation is responsible for this loss of function. Potato virus C, a NAT strain of PVY, was sequenced and although there were several amino acid differences from the aphid transmissible strain, only two were unique to the mutant (Thornbury *et al.*, 1990). The first of these substitutions occurred in the N-terminus of HC protein, where K_{307} was changed to E. The second substitution was I_{482} to V, a conservative change. The effect of these mutations was tested mutagenesis of TVMV (Atreya *et al.*, 1992). It was shown that the substitution at 482 had no effect on vector transmission, whereas the 307 residue was essential for function. Many potyviruses have now been sequenced in this region and all have K at this position, with the exception of pea seed borne mosaic virus, which has a conservative substitution, R.

MacMV has homology to potyvirus HC

The studies of natural and induced mutations in the HC protein suggests that it is important for aphid vector transmission and it would therefore be informative to analyse a comparable region in NLV and MacMV. A MacMV cDNA clone (pM4) was isolated from the original screen of ds cDNA clones. The sequence obtained from clone pM4 has homology with the N-termini of the polyproteins translated from RNA2 of BaMMV and BaYMV, and to limited regions of the HC proteins of potyviruses (figure 4.12) and it can therefore be tentatively placed at the 5' terminus of the MacMV genome.

The bymovirus genome is bipartite and the majority of RNA1 has a similar genetic organisation to that found in the C-terminal two thirds of the potyvirus genome (figure 4.14). However, there is homology between part of the polyprotein encoded by BaYMV RNA2, BaMMV RNA2 and the helper component proteinase (HC-Pro) of potyviruses (Davidson *et al.*, 1991; Kashiwazaki *et al.*, 1991; Timpe & Kuhne, 1994). This potyvirus protein contains a papain-like cysteine proteinase motif (Kamphuis *et al.*, 1985) which is conserved in the bymoviruses and MacMV (figure 4.13). There is consensus around the active cysteine residue (Oh & Carrington, 1989) of the proteinase between the bymoviruses, potyviruses and MacMV. In potyviruses, this proteinase is responsible for the autocatalytic cleavage of a protein, P3 which adjoins HC (figure 4.14). The cleavage site recognised by this proteinase between P3 and HC conforms to the consensus Y-X-V-G-G, with cleavage occurring between the two glycine residues (Carrington & Herndon, 1992). This sequence is present in MacMV (figure 4.13), conforming exactly with the potyvirus consensus, where the bymoviruses do not. In BaMMV the second G is changed to A, and in BaYMV, this residue is S. Cleavage at VG/G in MacMV would create a mature protein from the N-terminus of 26.2 kDa, which is a similar size to that predicted for BaYMV (28 kDa) and BaMMV (25kDa). *In vitro* translation experiments using full length and N-terminally deleted BaYMV RNA2 have shown that this 28 kDa protein is produced by autocatalytic cleavage from the N-terminus of the 98 kDa polyprotein (Davidson *et al.*, 1991). When full length NLV RNA was translated *in vitro* a distinct product of 25kDa was obtained (Mowat *et al.*, 1991). This protein did not precipitate with antisera raised to NLV coat protein or cytoplasmic inclusion protein. It is possible that this 25kDa found in *in vitro* translations could be synonymous with the translation and autocatalytic cleavage of the 5' terminus of MacMV.

The homology between MacMV and potyviral HC protein only extends from the proteinase cysteine residue to the cleavage site. The K_{307} residue important for aphid transmission is situated approximately 300 residues towards the N-terminus of the HC protein. No comparable region exists in MacMV, BaMMV or BaYMV since the proteinase cysteine residue is found around position 130 after the initiation methionine. This conforms to the genome organisation shown by the fungal transmitted bymoviruses in this region and therefore supports the suggestion that MacMV may not be transmitted by aphids.

<u>NLV and MacMV form a new genus, the *Macluraviruses*</u>

Although NLV and MacMV show good identity with bymoviruses based on amino acid sequence comparisons in the NIb gene, they show little similarity in the coat protein

and hence are serologically unrelated. They differ morphologically from the bymoviruses, with different coat protein sub-unit sizes and probable monopartite genome structure. This evidence, and the observation the NLV and MacMV have a similar level of amino acid sequence identity to the potyviruses that is shown by the bymoviruses strongly support the suggestion (Mowat *et al.*, 1991) that they should be classified as the fifth genus of the *Potyviridae*, tentatively designated the *Macluraviruses* (Shukla *et al.*, 1994).

NLV in Iris and Gladioli

It was suggested after serological testing that NLV was synonymous with Irisbontvirus and Iris mild mosaic virus which is found in bulbous iris (Brunt, 1977). Sequence data in the coat protein and 3'UTR from NLV particles obtained from infected Iris has confirmed this suggestion. NLV from Iris was indistinguishable from that analysed from narcissus or *N. clevelandii*. Similarly, NLV particles extracted from Gladioli samples were also found to be identical to NLV from narcissus and *N. clevelandii*.

This result was achieved using NLV- specific PCR primers (designed originally to aid sequencing). The downstream primer was used to prime first-strand cDNA from RNA extracted from virus particles. This cDNA was then used in an RT-PCR reaction with the first primer and an NLV-specific upstream primer. This technique could be repeated to examine the prevalence of NLV in natural narcissus populations. The difficulties presented by the use of synonyms in different plant species infected with NLV is only the beginning of the complex problems associated with the identification of NLV in narcissus, where symptoms can vary between cultivars. The availability of an RT-PCR test and DNA probe for NLV allowed the investigation of the nature of narcissus yellow stripe disease (Chapter 5).

Future work

The suggestion that clone pM4 represents sequences which are encoded by the 5' terminus of a monopartite genome, and not the terminus of a separate RNA from MacMV needs to be confirmed by northern analysis. This experiment was attempted once towards the end of the project but without success due to poor RNA quality. RNA from infected plant material should be used. A positive result would indicate the size of RNA to which pM4 hybridised and hence confirm whether MacMV has monopartite genome of 8 kb.

More work is required to elucidate the mode of transmission of NLV and MacMV. Sequence data suggest that they are not transmitted by aphids, however, there is no evidence that identifies a region of genome responsible for fungal transmission. The

aphid transmissibility of the isolates of NLV and MacMV should be accurately obtained. The possibility that MacMV and NLV are transmitted by *Polymyxa graminis* could also be investigated.

Chapter 5

A re-evaluation of the viruses associated with narcissus yellow stripe disease

Introduction

There has been a long history of conflicting and confusing data on the viral diseases of narcissus. This confusion is due to several factors: the use of many synonyms in different countries (Brunt, 1966), description of symptoms which vary between cultivars and problems in securing singly infected plants. This is compounded by a lack of herbaceous alternative hosts and the misleading classification of some viruses.

Twenty-one viruses have been reported to occur naturally in narcissus (Brunt, 1995). Of these, seven viruses are generally considered to be specific pathogens of narcissi; they frequently occur in complexes, with up to five viruses present in any one infection. The most confusion surrounds those viruses associated with yellow stripe symptoms, namely, narcissus latent virus (NLV), narcissus yellow stripe virus (NYSV), narcissus late seasons yellows virus (NLSYV) and narcissus mosaic virus (NMV). All these viruses have been reported to cause chlorotic leaf striping of varying colours from grey-green to bright yellow. This is dependent on the cultivar infected, hence the scope for confusion.

A pivotal factor appears to be the assignment of symptoms to NYSV. Although yellow stripe disease is the oldest recorded viral disease of narcissus (Dod, 1894), it is the one about which least is known. NYSV has no known herbaceous indicator plant (Brunt & Atkey, 1967). It is, therefore, the most elusive narcissus virus in this group and little further has been discovered since it was described in 1964 (Cremer & van der Veken, 1964).

<u>Narcissus mosaic virus (NMV)</u>

The naming of NMV and NYSV has long been confused. A mosaic disease was originally distinguished from the yellow stripe or 'greys' disease by van Slogteren & Ouboter (1946); however, 'narcissus mosaic virus' had previously been used as a synonym for NYSV (Haasis, 1939; McWhorter, 1932). This distinction was made clear by the isolation of a virus by single lesion transfers from *Gomphrena globosa* and *Chenopodium amaranticolor* which induced distinct mosaic symptoms that differed from those caused by NYSV in narcissus cv. Minister Talma (Brunt, 1966). This isolation of NMV in a herbaceous host allowed antiserum to be raised and tested against

NYSV particles. No reaction was observed (Brunt, 1966), confirming that NYSV and NMV were two separate viruses.

NYSV is reported to have filamentous flexuous particles measuring 12x750 nm (Brunt & Atkey, 1967), typical of a potyvirus, whereas, NMV was observed to have particles 548 nm long. The particle morphology of NMV suggested that it belonged to the potexvirus genus yet it failed to react to antisera to seven different members of that genus (Brunt, 1966). Its classification was finally settled when its genome was cloned and sequenced (Zuidema *et al.*, 1989). When this sequence was aligned with those of other known potexviruses it is clear that NMV is a member of this genus.

NMV frequently occurs in complexes with other viruses in narcissus. Alone, however, it induces inconspicuous mild mosaic leaf symptoms or dark brown oval spots (Brunt, 1966; 1995).

Narcissus latent virus (NLV)

The suggestion that NYSV is synonymous with NLV (Mowat *et al.*, 1991) was based on the evidence presented by the serological reactions of all coat protein antisera available for NYSV and NLV. This showed that two NYSV antisera reacted to NLV particles isolated from narcissus or *Nicotiana clevelandii*, but that a third reacted to NLSYV and did not react to NLV. A single infection of NYSV is difficult to obtain as it has no herbaceous indicator host species, and therefore mixed infections are likely to produce mixed antisera. If it cannot be assumed that any NYSV antisera has been raised from a pure infection source, then logically, conclusions cannot be drawn about the serological relationship between NYSV and NLV.

Further problems in identifying NYSV occurred when an antiserum raised to NYSV cytoplasmic inclusion body protein (CIP) (Mowat *et al.*, 1988a; 1989) reacted to NLV-infected samples. When isolating the NYSV CIP, Mowat *et al.* (1989) used virus-tested narcissus plants which were inoculated by aphid transmission using a natural disease source. They detected only the shorter NLV particles but observed the presence of cytoplasmic inclusion bodies and attributed those to the presence of a potyvirus, NYSV. At that time, NLV was characterised as a carlavirus, and it was assumed that the NLV particles would therefore not induce CIPs. This classification is no longer correct (Badge *et al.*, 1997a; Chapter 4), and we can assume that the NYSV CIP antisera was actually raised to the CIP of NLV.

NYSV is assumed to be a potyvirus because of its particle morphology (12x750nm), the main feature by which it is distinguished from NLV which has shorter particles

(13x650nm). NLV can be isolated in herbaceous indicator host species and can be purified for antiserum production, and has consequently been further investigated by sequence analysis of its genome. NLV is therefore established as a distinct virus from NYSV. NLV often occurs in complexes in narcissus, but alone induces light and dark green chlorosis especially at the leaf tips in some cultivars (Brunt, 1977).

Narcissus yellow stripe virus (NYSV)

Although narcissus yellow stripe potyvirus has been associated with the well characterised yellow stripe disease, its occurrence in complexes with other viruses in narcissus and problems with its isolation in indicator species have led to confusion. Haasis (1939) describes a mosaic pattern of light green to greyish green streaks on narcissus leaves, which may be dull to bright yellow on some cultivars, and was designated narcissus mosaic virus. However, common symptoms reportedly associated with NYSV are leaf distortion and epidermal roughening (Haasis, 1939; Cremer & van der Veken, 1964) which are not associated with NMV. It is generally accepted by the range of symptoms described that the mosaic virus reported by Haasis is actually NYSV (Brunt, 1977). The symptoms induced by NYSV intensify throughout the season and are accompanied by severe stunting and flower breaking (Brunt, 1977; Haasis, 1939).

Narcissus late seasons yellows virus (NLSYV)

NLSYV was characterised as a potyvirus with filamentous flexuous particles of 12x750nm (Mowat *et al.*, 1988b) but could not be transmitted by aphid or manual inoculation to 12 herbaceous species. Antiserum raised to the coat protein of NLSYV was found to react to particles of three typical potyviruses, potato virus Y (PVY), turnip mosaic virus (TuMV) and alstroemeria mosaic virus (AlMV), though it did not react with NLV particles or some plants that displayed yellow stripe symptoms. There are no sequence data available for NLSYV, which like NYSV, has no known alternative experimental herbaceous hosts. The division between NYSV and NLSYV is still one of aetiology, as their serological relationship remains unclear.

Aims of the investigation of narcissus yellows disease

The problems associated with the isolation of these narcissus-infecting viruses into herbaceous indicator hosts is at the centre of the confusion surrounding them. When symptoms are latent, unreliable, masked by other infections or vary from variety to variety, as is the case here, then identification can be extremely difficult. Isolation of single infections would aid this process greatly, allowing direct comparison of symptoms on a chosen variety of narcissus after inoculation from a single source.

Such isolation would also lead to the production of specific antisera. It is clear from a review of the literature that mixed antisera can generate misleading results, and it is difficult to prove that an antiserum is mixed when single infections cannot be confirmed. However, no such herbaceous hosts have been found, so an alternative method was required.

Several potyvirus-like viruses infecting narcissus need further investigation and characterisation. NLSYV is distinct from NYSV only in its aetiology and NYSV is reported to be synonymous with NLV. A more objective and reliable method of testing was required to distinguish these viruses from one another, or to establish if they were synonymous. Since sequence data are now available for NLV (Chapter 4) and all three viruses were considered to have particle morphology similar to virus of the potyvirus genus, it was decided that a molecular approach would be appropriate.

Following the success of the Carla-Uni PCR primer (Chapter 3) to identify known and suspected carlaviruses, it was decided to use a similar approach. A universal potyvirus PCR primer, CN48, (Pappu *et al.*, 1993) had also been successful in similar experiments and this was tested on infected narcissus samples. The aim was to generate fragments which could be sequenced in order to provide more information about the viruses associated with yellow stripe symptoms.

Results

Samples used for analysis

A wide variety of infected material was available from the stocks kept in gauze houses and in the field at the Scottish Crop Research Institute, Invergowrie (Dundee). Samples were chosen and used for analysis in two consecutive seasons (Spring 1995, Spring 1996). A summary of the type of material used is shown in table 5.1 and colour photographs of the samples are presented in figure 5.1. Two samples (1&2) were natural field infections of yellow stripe disease which had been maintained for a number of years in a gauze house. Previous experiments had determined that they were infected with NLV and showed yellow stripe symptoms. Two samples were the product of transmission of yellow stripe disease from sample 2: sample 3 was a manual transmission and sample 4 was infected by aphid transmission. Sample 5 was inoculated by mechanical transmission of NLV from infected *N. clevelandii* and showed no yellow stripe symptoms. Three samples were obtained from field narcissi. One, sample 8, displayed NLYSV symptoms and had become infected as part of a field trial on the transmission of NLYSV. Samples 6 and 7 had yellow stripe symptoms.

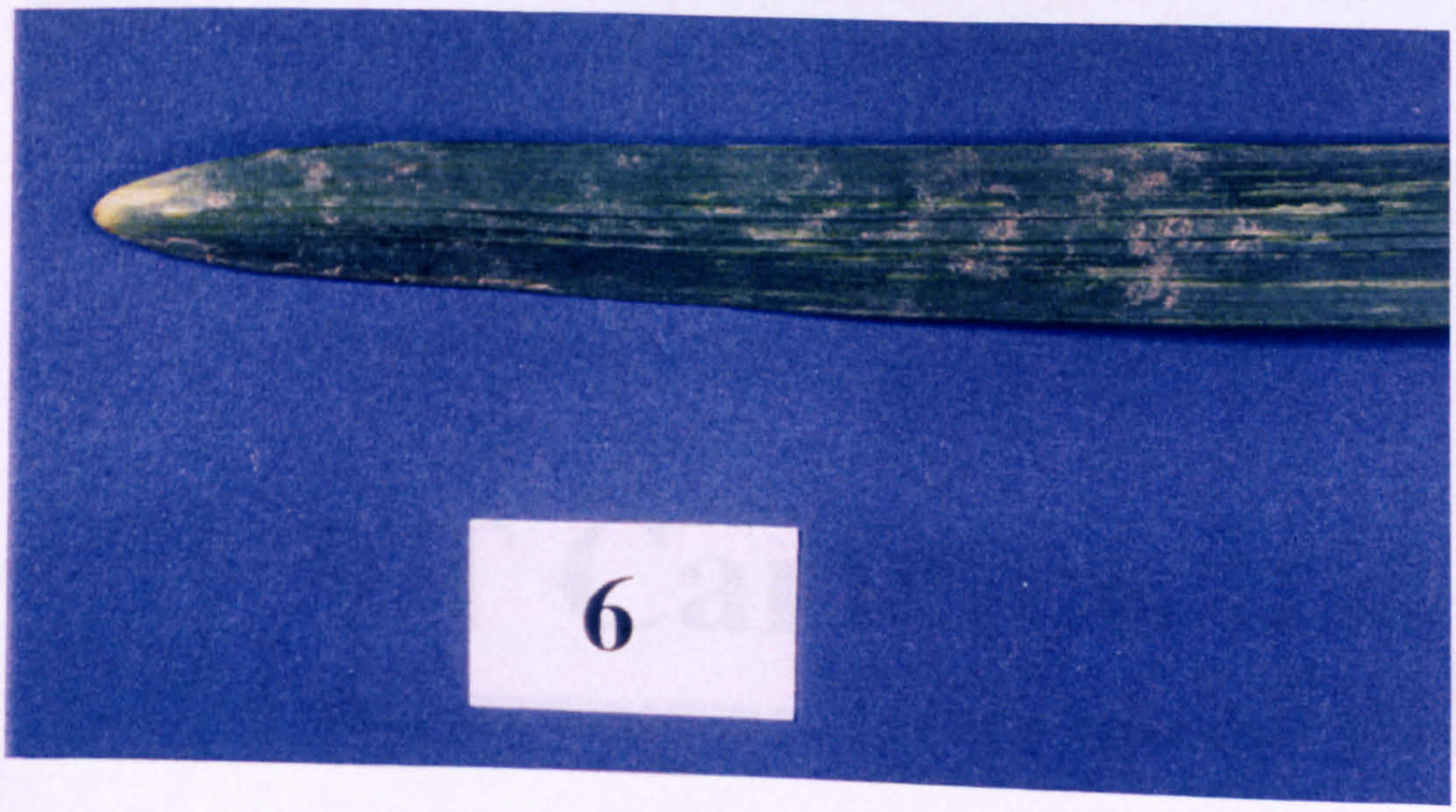
Sample Number	SCRI code	sample type	Virus originally diagnosed as present	Infection source
1	10/87	narcissus	NYSV/NLV	Natural field infection
2	12/87	narcissus	NYSV/NLV	Natural field infection
3	15/87	narcissus	NLV	Manual transmission from 2 (12/87)
4	16/87	narcissus	NLV	Aphid transmission from 2 (12/87)
5	GM1.8	narcissus	NLV	Manual transmission from NLV-infected *N. clevelandii*
6	R7B1	narcissus	NYSV	Natural field infection
7	R11B6	narcissus	NYSV	Natural field infection
8	R5B6	narcissus	NLSYV	Natural field infection
9	NLV	*N. clevelandii*	NLV	Manual transmission
10	PVY	*N. tabacum*	PVY	Virus stocks (Leics)
11	Healthy	*N. clevelandii*	none	Healthy *N. clevelandii*
12	VT	narcissus	none	virus tested healthy narcissus
13	790.1	narcissus	NYSV	Natural field infection HRI
14	790.2	narcissus	NYSV	Natural field infection HRI
15	790.3	narcissus	NYSV	Natural field infection HRI
16	790.4	narcissus	NYSV	Natural field infection HRI

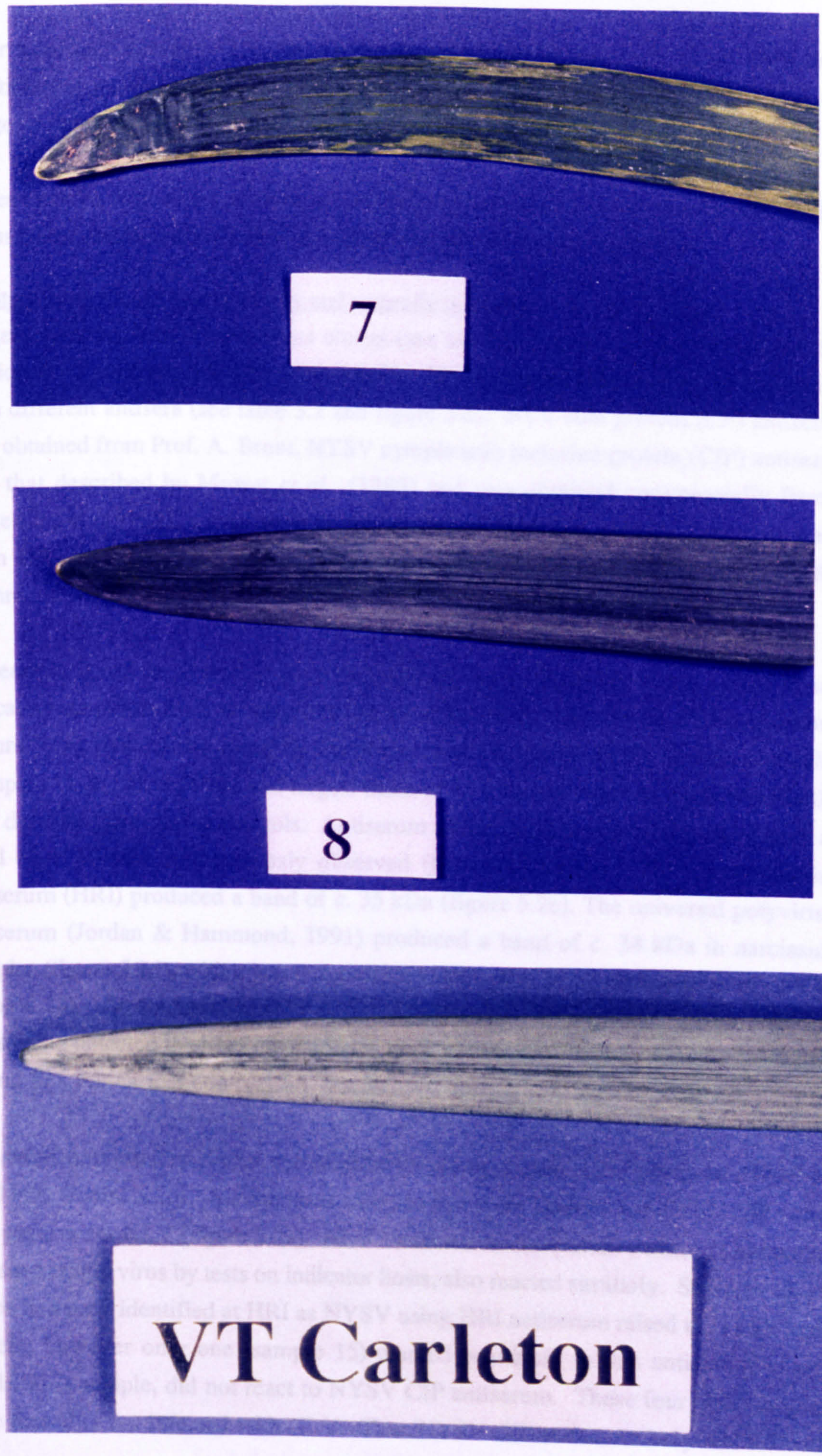
Table 5.1: Infected samples used for western, northern and dot-blot analysis. Samples 1 to 5, 9 and 11 were taken from the stocks held in gauze houses at SCRI, Invergowrie, samples 6 to 8 were from the field (SCRI), sample 12 was the available frozen lab stock and samples 13 to 16 were kindly supplied by Dr. N. Spence, HRI, Wellesbourne.

Figure 5.1 : Colour photographs of narcissus samples used for western and RT-PCR analysis. Each sample is labelled according to the SCRI reference code listed in table 5.1.


16/87


GM1-8


6


7
8
VT Carleton

Four narcissus samples (13, 14, 15 & 16) were kindly donated by Dr. Spence, Horticulture Research Institute (Wellesbourne), which had been ELISA tested for a range of narcissus viruses, and were reported to contain NYSV, but not NLV.

Other samples used as controls were obtained from virus stocks at Leicester or SCRI. Virus tested narcissus cv. Carlton was used a healthy control (sample 12).

Western analysis of experimentally and naturally infected narcissus samples

Protein samples from 13 narcissus clones (see table 5.1) were prepared (Chapter 2: section 4.1.2) and analysed by SDS-PAGE. Identical gels were blotted and probed with different antisera (see table 5.2 and figure 5.2). NLV coat protein (CP) antisera was obtained from Prof. A. Brunt, NYSV cytoplasmic inclusion protein (CIP) antisera was that described by Mowat *et al.*, (1989) and was obtained commercially from Adgen, as was NLSYV coat protein antiserum. NYSV coat protein antiserum was from HRI and the universal potyvirus antiserum was that selected by Jordan & Hammond (1991) and obtained commercially from Adgia Inc (USA).

In general, it can be seen that specific reactions were obtained, with a single band indicating the presence of an antigen to the antiserum used. Only the NLSYV antiserum (figure 5.2c) was contaminated by the presence of two bands in the healthy controls (samples 11 & 12), the third and largest band, at 35 kDa was taken as a positive result as it did not appear in the controls. Antiserum raised to NLV coat protein produced a band of *c*. 40 kDa as previously observed (figure 5.2a), the NYSV coat protein antiserum (HRI) produced a band of *c*. 35 kDa (figure 5.2e). The universal potyvirus antiserum (Jordan & Hammond, 1991) produced a band of *c*. 34 kDa in narcissus samples (figure 5.2d), which was comparable to that seen in the positive control, PVY (sample 10). The controls used for each antiserum gave the expected results, with no virus being detected in either the narcissus or *N. clevelandii* healthy samples (samples 11 and 12). For ease of comparison, results have been summarised in table 5.3.

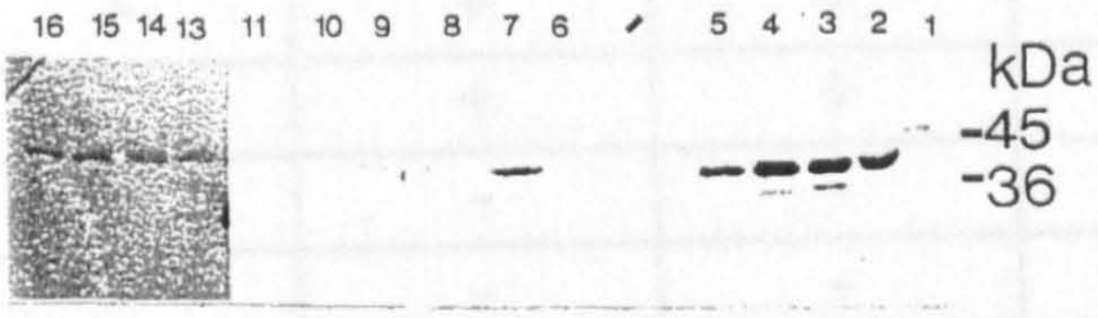
The antiserum raised to NYSV cytoplasmic inclusion protein (CIP) produces a band of *c*. 70 kDa (figure 5.2b). All narcissus samples that were positive for NYSV CIP were also positive for NLV (figure 5.2a). NLV - infected *N. clevelandii* , which is known to contain a single virus by tests on indicator hosts, also reacted similarly. Samples 13-16 which had been identified at HRI as NYSV using HRI antiserum raised to NYSV coat protein, however only one (sample 15) reacted positively to this antiserum in our hands. This sample, did not react to NYSV CIP antiserum. These four HRI samples were therefore not infected with NLV. Thus NYSV CIP antiserum only reacted with

Antisera	Protein source	Source	Secondary Ab	Titre
NLV CP	coat protein	A. Brunt, HRI	rabbit	1/10,000
NLSYV CP	coat protein	Adgen, SAC diagnostics (SCRI)	rabbit	1/5,000
NYSV CIP	cytoplasmic inclusion protein	Adgen, SAC diagnostics (SCRI)	rabbit	1/5,000
NYSV CP	coat protein	A. Brunt, HRI	rabbit	1/5,000
Uni-Poty		Adgia, Inc. (Jordan & Hammond, 1991)	mouse	1/250

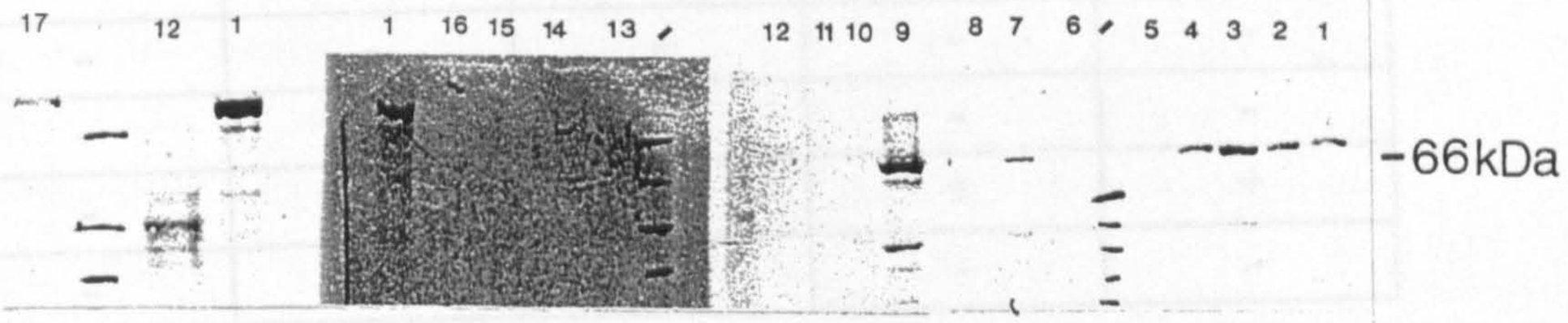
Table 5.2: Antisera used for western blotting in figure 5.2, results shown in table 5.3

Figure 5.2 : Western blots showing reaction between narcissus and tobacco protein samples and different antisera. Lane numbers 1-16 on each gel correspond to the sample numbers identified in table 5.1. Lane numbered 17 is MacMV coat protein. Antisera and dilution used: (a) NLV coat protein, HRI 1/10,000; (b) NYSV cytoplamsic inclusion body protein, SCRI, 1/5,000; (c) NLSYV coat protein, SCRI, 1/5,000; (d) universal potyvirus coat protein polyclonal antiserum, Adgia, 1/250; (e) NYSV coat protein, Brunt, 1/5,000.

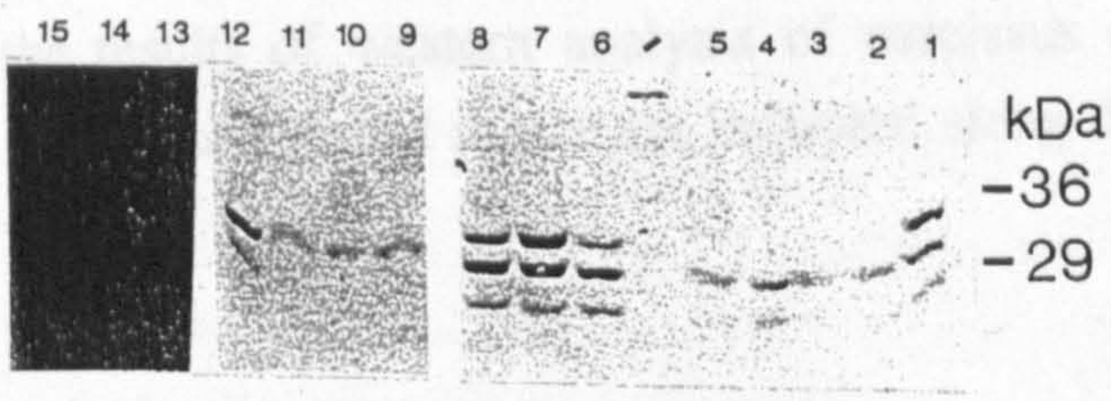
a NLV



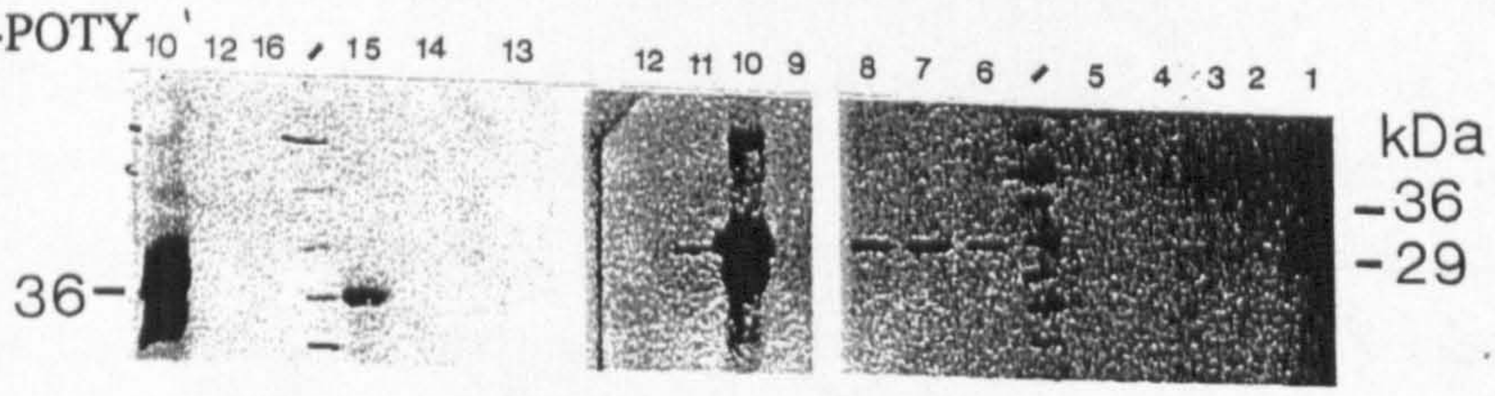
b NYSV CI



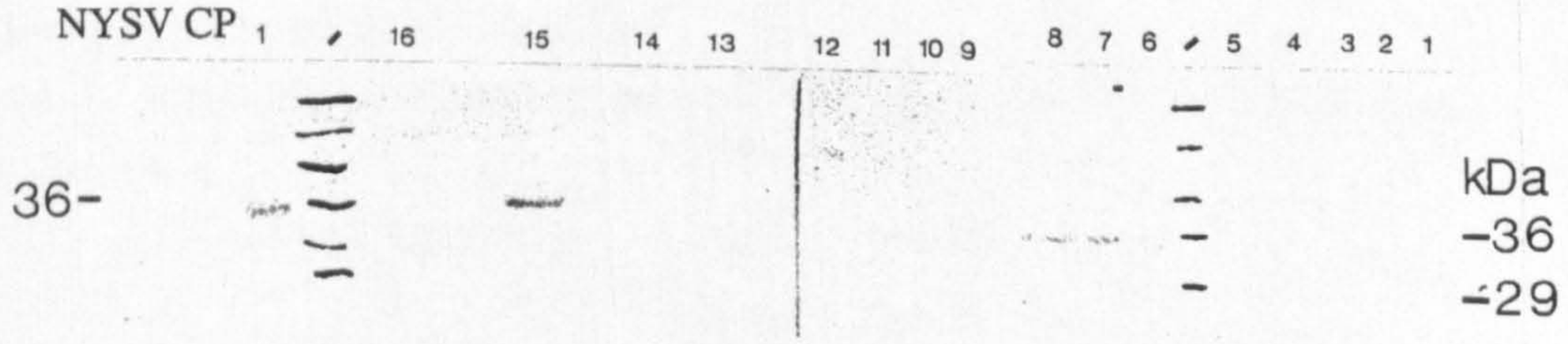
c NLSYV



d UNI-POTY



e NYSV CP



Sample Number	NLV CP Brunt	NYSV CIP SCRI	NLSYV CP SCRI	NYSV CP Brunt	Uni-Poty Adgia
1	+	+	+	+	+
2	+	+	-	-	-
3	+	+	-	-	-
4	+	+	-	-	-
5	+	+	-	-	-
6	-	-	+	+	+
7	+	+	+	+	+
8	-	-	+	+	+
9	+	+	-	-	-
10	-	-	-	-	+
11	-	-	-	-	-
12	-	-	-	-	-
13	-	-	-	-	-
14	-	-	-	-	-
15	-	-	+	+	+
16	-	-	-	-	-

Table 5.3: A summary of the results of western analysis of narcissus and tobacco samples shown in table 5.1. Antiserum and its source are indicated along the top of the table.

NLV positive samples. It also cross-reacted to a MacMV-infected *N. clevelandii* sample (figure 5.2b).

Northern and dot blot analysis of infected narcissus samples

Total plant RNA was extracted from all infected samples (Chapter 2: section 3.1.3) and used in northern dot blots. The same RNA samples were analysed by agarose gel electrophoresis and then blotted for northern analysis. Both filters were probed using the radio-labelled insert from clone pCRN24 which represented part of the nuclear inclusion body b (NIb) ORF of NLV. Positive results were obtained for samples in both dot blots and by northern analysis (figure 5.3 a,b). These results agreed with those obtained by western analysis using the NLV CP antiserum, except those for samples 8 and 2. RNA samples were prepared and analysed from the samples harvested in 1995, and the protein analysis presented here was carried out on new samples harvested in 1996. Since sample 8 was a field isolate from a naturally infected plant, it is likely that the 1995 sample used for RNA is not the same individual plant used for the protein (and consequently RT-PCR) analysis in 1996. Therefore it may be suggested that, the 1995 sample 8 was infected with NLV, but the 1996 was not. Sample 2 however, gave a positive dot blot result when a fresh sample was used (results not shown). Any following analysis of sample 8 described below was carried out on the sample taken in 1996 which was not infected with NLV.

Further investigation of key samples

Three samples, 6, 8 and 15 were negative for NLV by dot blot, northern and western analysis. None of these samples appeared to be a single infection on first analysis, since all reacted with three other antisera, NYSV (HRI), NLSYV (SCRI) and the universal potyvirus antiserum (figure 5.2c, e and table 5.3). It was possible that these samples contained only one potyvirus and that the two antisera were to closely related viruses and therefore cross-reacted, or one or both were raised to mixed infections. Without a different approach, no conclusions could be drawn about the number or nature of the potyvirus infection. However, it is clear from antisera results that at least one potyvirus, which was not NLV was present in these samples. Therefore it was decided to use the universal potyvirus PCR primer to obtain further information about these infections.

RT-PCR using a universal potyvirus primer

The universal potyvirus primer designated CN48 by Pappu *et al.* (1993) has the following sequence:

5'-TGG TGY ATH GAN AAT GG-3' where Y = C or T
H = A, C or T
N = A,C,G or T

Figure 5.3 (a): Dot blots of narcissus RNA samples probed with pCRN24, a cDNA clone representing part of the NIb ORF of NLV.

(a)

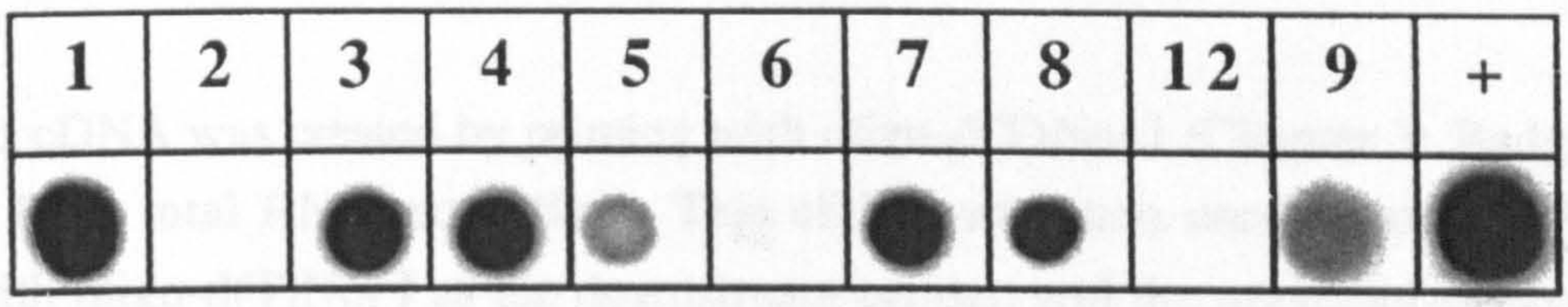
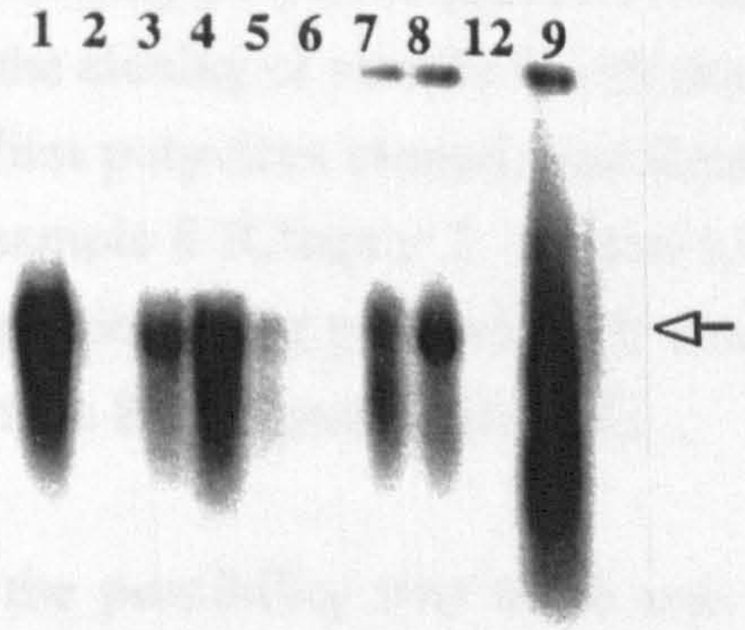


Figure 5.3 (b): Northern analysis of narcissus RNA samples electrophoresed on a non-denaturing agarose gel and probed with pCRN24, a cDNA clone representing part of the NIb ORF of NLV.

(b)



This potyvirus primer was selected since no sequence data are available for NYSV or NLYSV and all samples reacted positively to the universal potyvirus antiserum. Sample 8 which reacted positively to NYSV CP, NLSYV CP and the universal potyvirus antisera was selected. It was the only plant showing distinct narcissus late seasons yellows symptoms, and thus offered the greater possibility of containing only one potyvirus.

First-strand cDNA was created by priming with oligo-d(T)Not I (Chapter 3; Badge *et al.*, 1996) from total RNA extractions. This cDNA was then used in an RT-PCR reaction with oligo-d(T)Not I as the downstream primer, and the universal potyvirus primer upstream. Negative and positive controls were included to ensure a genuine product was obtained. Sample 8 produced a band of 600bp, slightly smaller than the expected size (see figure 5.4). No product was detected in the negative control (no cDNA) or healthy control (cDNA made to healthy narcissus RNA) and an 800bp band was produced by PVY cDNA. The PCR product from sample 8 was recovered from the gel and cloned into pCR-SCRIPT (Stratagene). Three independent clones, pNP8.4, pNP8.19 and pNP8.36 were sequenced in both directions, and were found to be identical. When subjected to a BLAST search, the nucleotide sequence was confirmed to be a potyviral coat protein sequence due to its high level of sequence identity with other known potyviruses.

In order to confirm that this was the only potyvirus present in the RT-PCR product, all 61 white colonies resulting from the cloning of sample 8 with possible inserts were first screened with the insert of the first potyvirus cloned, and then re-screened with the purified RT-PCR product from sample 8 (Chapter 2: section 1.6.1). Both screenings gave the same result, with 10 colonies being positive each time. Therefore only one potyvirus was amplified from sample 8 (designated Black.8).

In order to investigate further the possibility that there was a mixed infection of potyviruses present in some narcissi, the same RT-PCR technique was used on sample 1. This sample was selected because it was naturally infected, had yellow stripe symptoms and was positive to all antisera tests. A 600bp product was obtained using the universal potyvirus primer (see figure 5.4). This was recovered from the gel and cloned into pCR-SCRIPT (Stratagene). The resultant 60 colonies were screened with the insert from the first potyvirus cloned from sample 8, pNP8.4. Thirteen positive colonies were obtained (see figure 5.5); two, pNP1.4 and pNP1.7 (numbered 4 and 7 on figure 5.5 a) were picked for sequencing. They were found to be identical and the sequence was of potyviral origin, the virus from sample 1 was therefore designated Black.1.

However, it can be seen on figure 5.5a that an additional 6 colonies reacted only weakly to the probe, appearing grey on the autoradiograph (numbered 9, 29, 34, 35, 47 and 48 on figure 5.5a). If sample 1 contained a mixed infection of two potyviruses, then it is possible that the universal PCR primer bound to both viruses and generated a mixed population of PCR fragments from the two putative viruses present. If this were the case then the original PCR product would hybridise to all transformed clones which would represent both viruses, whereas pNP8.4 representing only one virus, would only hybridise strongly to some of the clones. If the two viruses were closely related, then there may be some weak cross-hybridisation between them, as seen when the colonies from sample 1 were probed with pNP8.4. All colonies were therefore re-screened using the RT-PCR product from sample 1, and 20 positives were obtained and those 6 that had previously appeared faint, were now strong positives (figure 5.5b). Three of these 'new' positive clones, pNP1.34, pNP1.35, pNP1.47 (numbered 34, 35 and 47 on figure 5.5b), were sequenced and shown to be identical to each other and of potyviral origin. This virus was designated Grey.1.

Nucleotide and amino acid sequence analysis of three narcissus clones

The nucleotide sequence and deduced amino acid translation of the clones obtained from narcissus samples 1 and 8 are shown in figure 5.6. Each clone is 665 nucleotides (nt) and has a ORF of 150 amino acids (aa), with a termination codon at position 451 nt and an untranslated region of 211 nt. At the nucleotide level, Black.1 and Black.8 are 95 % identical, whilst both Black viruses are only 68% identical to Grey.1.

A BLAST database search revealed that all three viral sequences share identity with the potyvirus, turnip mosaic virus (TuMV) and the deduced amino acid sequence of the large ORF corresponds to its coat protein C-terminus. Figure 5.7a shows an alignment of all the available sequence for the three clones with a corresponding region for tobacco etch virus (TEV), a typical potyvirus and PVY, the type member of the potyvirus genus. Figure 5.7b shows that although sample 1 was infected with NLV (as demonstrated by dot blot, western and northern analysis), Black.1 is a virus distinct from NLV and it is more closely related to the potyviruses (77.3%-79.1%) than to NLV(22.6%). The Black viral sequences obtained from samples 1 and 8 are 96.6% identical to one another in this region, but they are only 88% identical to the viral sequence Grey.1.

An alignment of the sequenced narcissus viral cDNA and its closest relatives, TuMV, tulip top-breaking virus (TTBV) and tulip breaking virus (TBV) is shown in figure 5.8. The sequence data available for TTBV and TBV were obtained by a similar method of

Figure 5.4 : Agarose gel electrophoresis of the products of RT-PCR using oligo-d(T)Not I primed first-strand cDNA and universal potyvirus primer. Lane 1: kilobase marker; Lane 2: narcissus sample 1; Lane 3: narcissus sample 8; Lane 4: healthy sample; Lane 5: negative water control; Lane 6: PVY RNA as positive control.

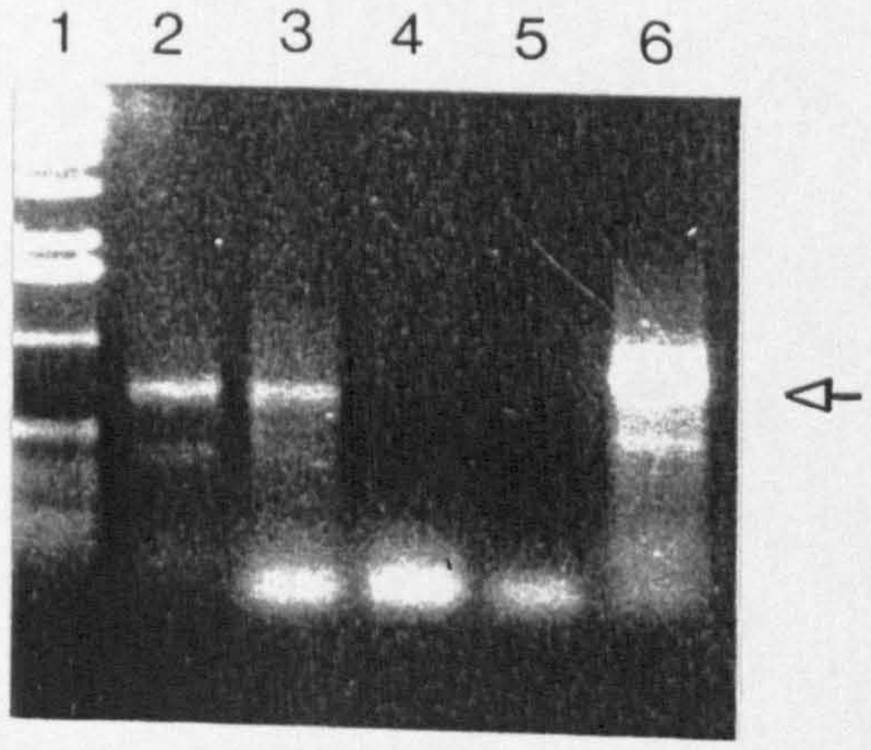
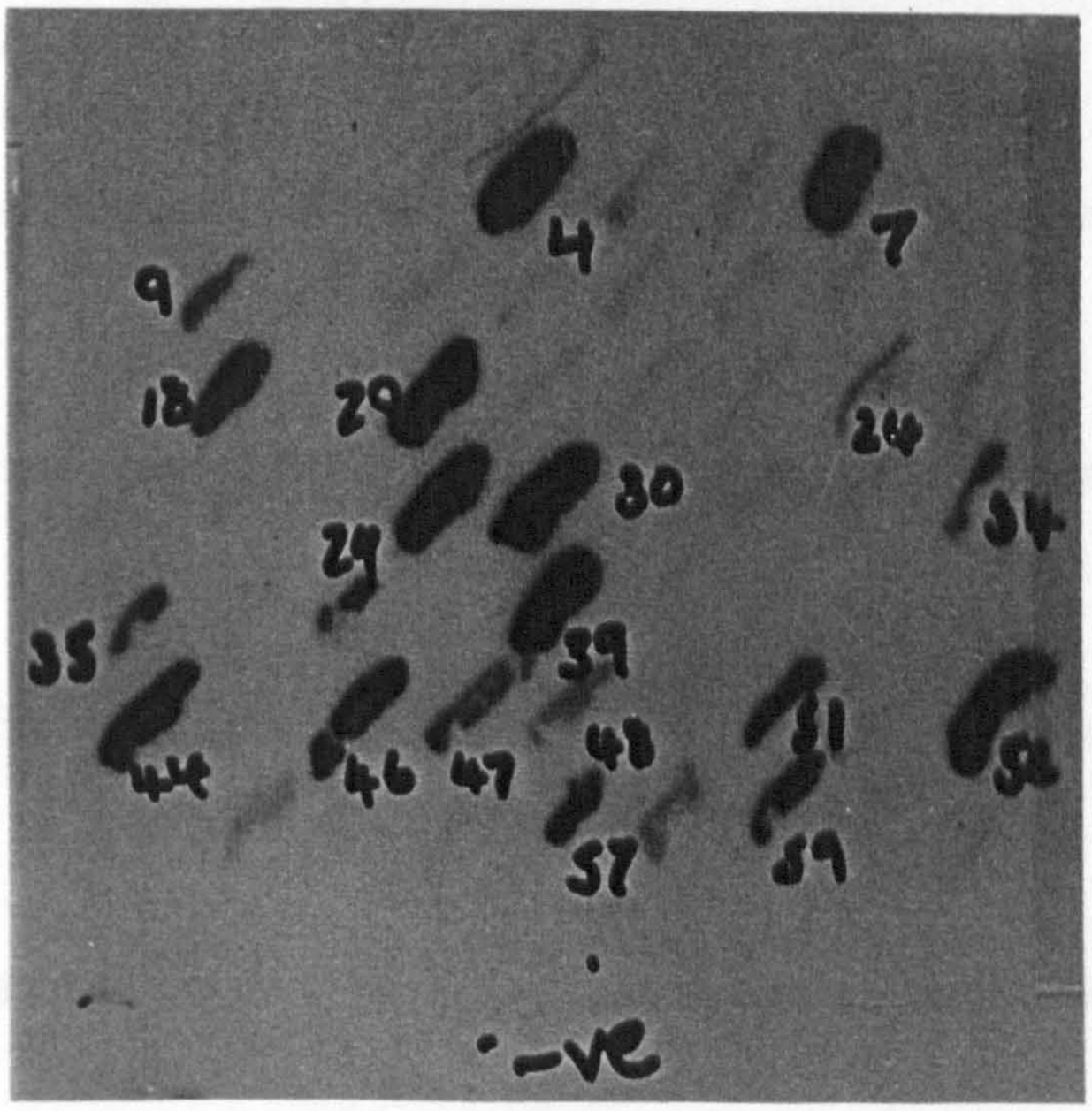


Figure 5.5 (a) : Colonies resultant from cloning of PCR fragment generated by RT-PCR using uni-poty and oligo-d(T)Not I primers on narcissus sample 1. Colonies were screened with radio-labelled insert from pNP8.4.

Figure 5.5 (b) : Colonies resultant from cloning of PCR fragment generated by RT-PCR using uni-poty and oligo-d(T)Not I primers on narcissus sample 1. Colonies were screened with the original PCR product from sample 1.

a)



b)

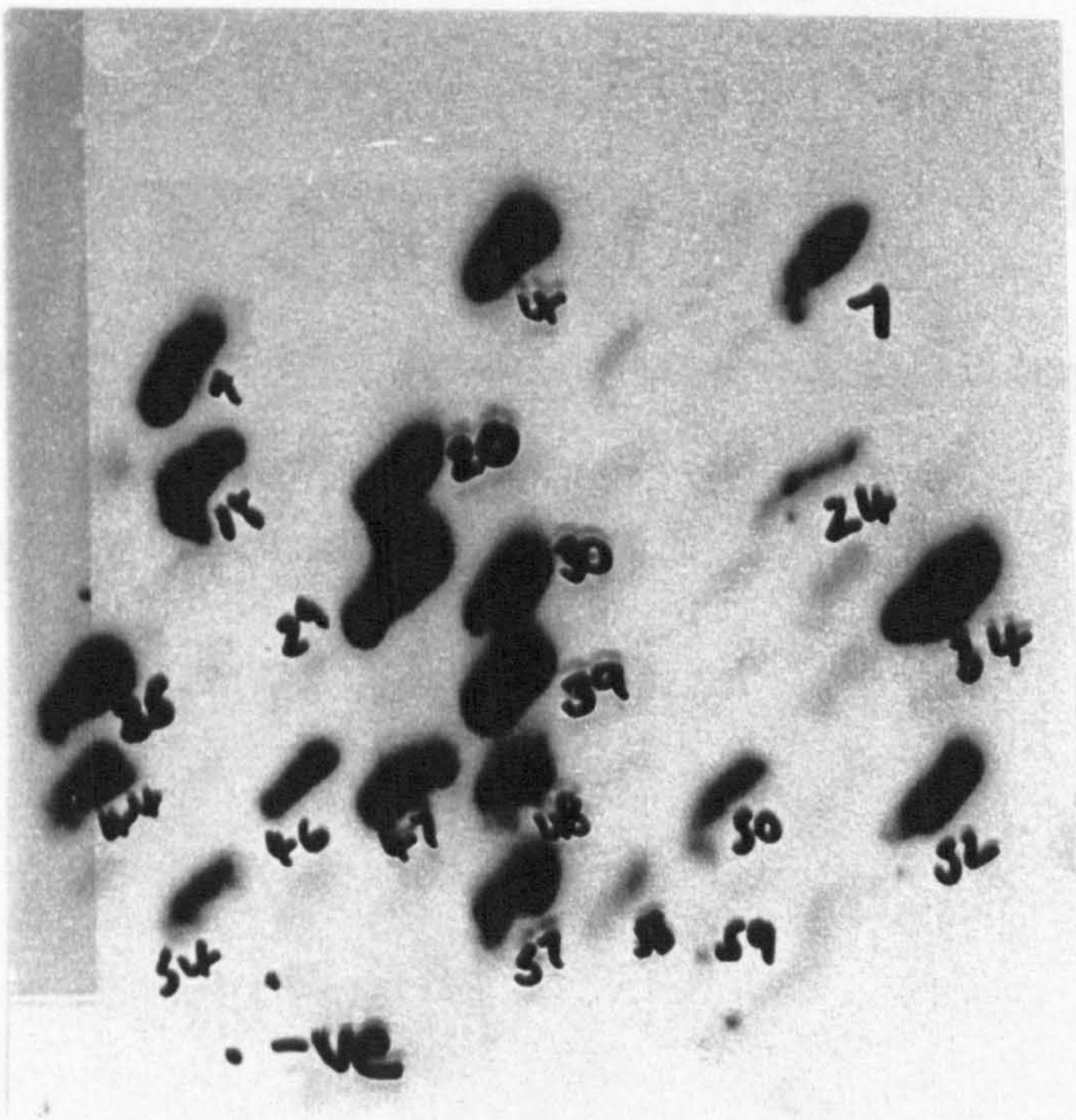


Figure 5.6: Nucleotide sequence of cloned RT-PCR products obtained from infected narcissus RNA translation is shown underneath. (a) pNP1.4, pNP1.7 (viral sequence Black.1); (b) pNP8.4, pNP8.19, pNP8.36 (viral sequence Black.8); (c) pNP1.34, pNP1.35, pNP1.47 (viral sequence Grey.1).

(a)

```
          10         20         30         40         50         60         70
          |          |          |          |          |          |          |
TGGTGTATCGAGAATGGAACATCACCAAACATCAATGGAGTGTGGGTTATGATGGACGGAGAGGAACAAGTTGAATTT
W  C  I  E  N  G  T  S  P  N  I  N  G  V  W  V  M  M  D  G  E  E  Q  V  E  F

 80         90         100        110        120        130        140        150
 |          |          |          |          |          |          |          |
CCTCTAAAACCAGTCGTTGAACACGCCAAACCAACACTCAGGCAAATAATGGCCCATTTCAGTAACGGTGCTGAAGCG
P  L  K  P  V  V  E  H  A  K  P  T  L  R  Q  I  M  A  H  F  S  N  G  A  E  A

   160        170        180        190        200        210        220        230
   |          |          |          |          |          |          |          |
TATATTGAGATGAGAAACTACCAACAACCGTACATGCCACGGTACGGCCTGCAGCGAAATTTAACCGACATGAGCCTG
Y  I  E  M  R  N  Y  Q  Q  P  Y  M  P  R  Y  G  L  Q  R  N  L  T  D  M  S  L

      240        250        260        270        280        290        300        310
      |          |          |          |          |          |          |          |
GCACGCTTTGCTTTTGATTTCTATGAGATGACTTCAAAAACACCAATCCGAGCACGAGAAGCACACCTGCAGATGAAA
A  R  F  A  F  D  F  Y  E  M  T  S  K  T  P  I  R  A  R  E  A  H  L  Q  M  K

         320        330        340        350        360        370        380        390
         |          |          |          |          |          |          |          |
GCAGCAGCTCTTCGGGGCACGCAGAATCATTTATTTGGCTTGGATGGAAATGTTGGTACAACAGAAGAGAACACGGAG
A  A  A  L  R  G  T  Q  N  H  L  F  G  L  D  G  N  V  G  T  T  E  E  N  T  E

            400        410        420        430        440        450        460
            |          |          |          |          |          |          |
AGACACACCACCGACGATGTAAATAGGAATATGCATACCCTTTTAGGCGTAAGGGGCGTGTAATATTGTGTGTTGGTA
R  H  T  T  D  D  V  N  R  N  M  H  T  L  L  G  V  R  G  V  •

 470        480        490        500        510        520        530        540
 |          |          |          |          |          |          |          |
TTTATAGCCAAATATATAAATATTCGTTAGTATTTTCAACTATAGAGAAATATTATCACACAAGGATGCTAGTTTGGC

   550        560        570        580        590        600        610        620
   |          |          |          |          |          |          |          |
ACAGCCAGGAATGTAGTCTTGTGTTTATATATATTTCGATCGAAAACGGCGGATTTGCGGAAGAACCAGTGAGAGGTT

      630        640        650        660
      |          |          |          |
TTTTCCAACATTTTTCACCAAGTTCGGTGTTTTGGGGACCA(n)
```

(b)

```
          10         20         30         40         50         60         70
          |          |          |          |          |          |          |
TGGTGTATCGAGAATGGAACATCACCAAACATCAATGGAGTGTGGGTTATGATGGACGGAGAGGAACAAGTTGAATTT
W  C  I  E  N  G  T  S  P  N  I  N  G  V  W  V  M  M  D  G  E  E  Q  V  E  F

 80         90         100        110        120        130        140        150
 |          |          |          |          |          |          |          |
CCTCTAAAACCAGTCATTGAACACGCCAAACCAACACTCAGGCAAATAATGGCCCATTTCAGTAACGGCGCTGAAGCG
P  L  K  P  V  I  E  H  A  K  P  T  L  R  Q  I  M  A  H  F  S  N  G  A  E  A

    160        170        180        190        200        210        220        230
    |          |          |          |          |          |          |          |
TATATTGAGATGAGAAACTACCAACAACCGTACATGCCACGGTACGGCCTGCAGCGAAATTTAACCGACATGAGCCTG
Y  I  E  M  R  N  Y  Q  Q  P  Y  M  P  R  Y  G  L  Q  R  N  L  T  D  M  S  L

      240        250        260        270        280        290        300        310
      |          |          |          |          |          |          |          |
GCACGCTTTGCTTTTGATTTCTATGAAATGACTTCAAAAACACCAATCCGAGCACGAGAAGCACACCTGCAGATGAAA
A  R  F  A  F  D  F  Y  E  M  T  S  K  T  P  I  R  A  R  E  A  H  L  Q  M  K

       320        330        340        350        360        370        380        390
       |          |          |          |          |          |          |          |
GCGGCAGCTCTTAGGGGCACGCACAATCATCTATTTGGCTTGGATGGAAATGTTGGTACAACCGAAGAGAACCCGGAG
A  A  A  L  R  G  T  H  N  H  L  F  G  L  D  G  N  V  G  T  T  E  E  N  P  E

          400        410        420        430        440        450        460
          |          |          |          |          |          |          |
AGACACACCACCGACGATGTAAATGGGAATATGCATAGCCTTTTAGGCGTAAGGGGGGGTGTAATATTGTGTGTTGGTA
R  H  T  T  D  D  V  N  G  N  M  H  S  L  L  G  V  R  G  V  •

 470        480        490        500        510        520        530        540
 |          |          |          |          |          |          |          |
TTTATAGTAAAATATATAAATACCCGTTAGTATTTTCAACTATAGAGAAATATTATCACACAAGGATGTTAGTTTGGC

   550        560        570        580        590        600        610        620
   |          |          |          |          |          |          |          |
ACACCCAAGAACGTAGTCTTGTGTTTATATATATTTCGATCGAAAACGGCGGATTTGCTGAAGAACCAGTGAGAGTTT

      630        640        650        660
      |          |          |          |
TTTTCCAACAGTTTTCACCAAGTTCGGTGTTTTAGGAGACA(n)
```

(c)

```
          10         20         30         40         50         60         70
          |          |          |          |          |          |          |
TGGTGCATAGAGAATGGTACCTCACCGAACATAAACGGAGTGTGGGTGATGATGGACGGGGACAAACAATTTGAATTC
W  C  I  E  N  G  T  S  P  N  I  N  G  V  W  V  M  M  D  G  D  K  Q  F  E  F

 80         90         100        110        120        130        140        150
 |          |          |          |          |          |          |          |
CCACTCAAACCCATCATTGACCACGCCAAACCCACACTCAGGCAAATAATGGCCCATTTCAGTAACGTCGCTGAAGCG
P  L  K  P  I  I  D  H  A  K  P  T  L  R  Q  I  M  A  H  F  S  N  V  A  E  A

    160        170        180        190        200        210        220        230
    |          |          |          |          |          |          |          |
TATATTGAAAGAAGGAACTATGACAGACCGTATATGCCACGTTATGGTCTTCAGCGTAATTTAACCGACATGAGCTTG
Y  I  E  R  R  N  Y  D  R  P  Y  M  P  R  Y  G  L  Q  R  N  L  T  D  M  S  L

      240        250        260        270        280        290        300        310
      |          |          |          |          |          |          |          |
GCACGCTTTGCCTTTGACTTCTATGAAATGACGTCAAAGACACCAACCCGAGCGCGAGAGGCGCACATCCAGATGAAA
A  R  F  A  F  D  F  Y  E  M  T  S  K  T  P  T  R  A  R  E  A  H  I  Q  M  K

       320        330        340        350        360        370        380        390
       |          |          |          |          |          |          |          |
GCGGCGGCACTACGTGGTGCTAAGAATCACTTATTTGGGTTGGATGGAAATGTTGGATCGGCAGATGAGAACACGGAG
A  A  A  L  R  G  A  K  N  H  L  F  G  L  D  G  N  V  G  S  A  D  E  N  T  E

           400        410        420        430        440        450        460
           |          |          |          |          |          |          |
AGACACACCACCGATGATGTAAATAGGAACATGCACAACCTTCTAGGCGTGAGGGGTGTGTAATTGTTGTGTCGATAT
R  H  T  T  D  D  V  N  R  N  M  H  N  L  L  G  V  R  G  V  •

 470        480        490        500        510        520        530        540
 |          |          |          |          |          |          |          |
TTATAGTGAGATATATCCATATAGATATCCGTTAGTATTTTCATCGTAGAGAAGTATTTTATATAGTGTTTTAGCCAG

   550        560        570        580        590        600        610        620
   |          |          |          |          |          |          |          |
TTGGGTTGGACCCAGAGTGCACCACTCTATCTATATACTTTGTTCGAATACATTGTACTTACTAAGGGAACAGTGAGA

      630        640        650        660
      |          |          |          |
GCCTTTTTCGAACTCCTTTACTAAAGACAGTGTGCTAGAGA(n)
```

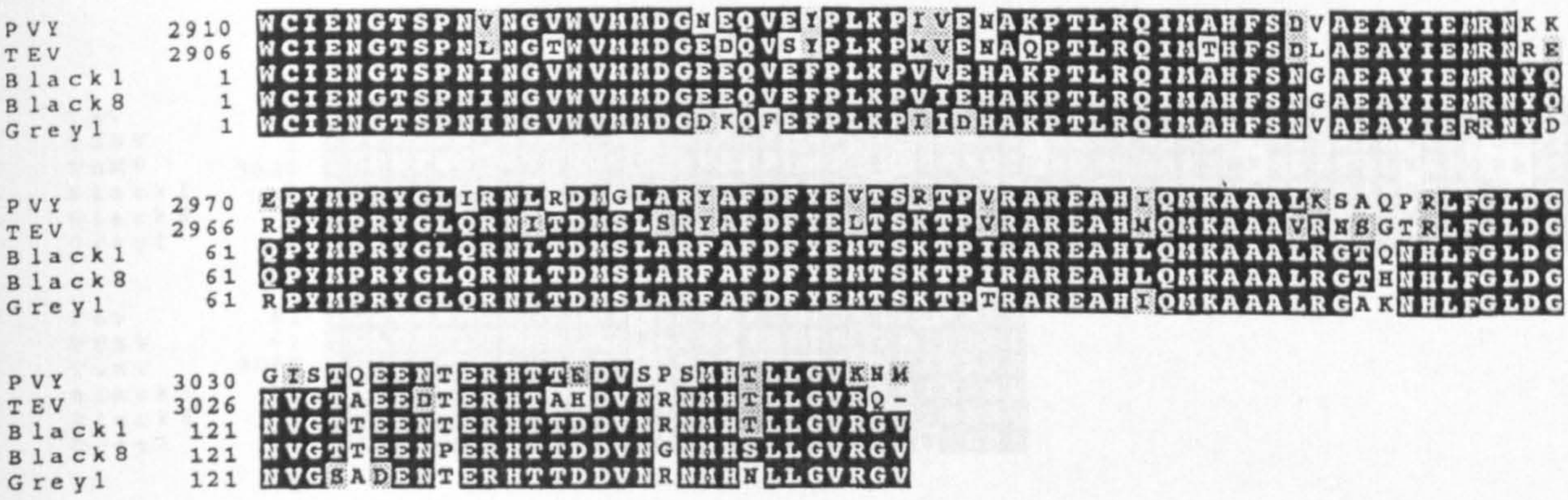


Figure 5.7 (a) : CLUSTAL W alignment of the partial coat protein ORFs of the three narcissus virus cDNA sequences, Black.1, Black.8 and Grey.1 with the C-terminal regions of the coat proteins of two typical potyviruses, potato virus Y (PVY, GenBank A08776) and tobacco etch virus (TEV, GenBank M15239). Gaps (-) have been introduced for maximum alignment and the program BOXSHADE was used to create boxing. Residues identical to Black.1 are boxed with a black background, chemically similar residues are boxed in grey.

	Black 8	Grey 1	NLV	TEV	PVY
Black 1	96.6	87.3	22.6	79.1	77.3
Black 8		88.0	22.6	76.5	74.6
Grey 1			24.0	75.1	74.1
NLV				25.5	26.6
TEV					74.5

Figure 5.7 (b) : Pairwise percent amino acid sequence identities between the partial coat protein sequences of the three narcissus virus cDNA sequences Black.1, Black.8 and Grey.1, two typical potyviruses, potato virus Y (PVY; Robaglia *et al.*, 1989) and tobacco etch virus (TEV; Allison *et al.*, 1986) and NLV.

Figure 5.8 (a) : CLUSTAL W alignment of the partial coat protein ORFs of the three narcissus virus cDNA sequences, Black.1, Black.8 and Grey.1 with the C-terminal regions of the coat proteins of tulip top-breaking virus (TTBV, GenBank S60806), tulip breaking virus (TBV, GenBank S44147) and turnip mosaic virus (TuMV, GenBank X83968). Gaps (-) have been introduced for maximum alignment and the program BOXSHADE was used to create boxing. Residues identical to Black.1 are boxed with a black background, chemically similar residues are boxed in grey.

	Black 1	Grey 1	TTBV	TBV	TuMV
Black 8	98.8	89.1	82.6	85.8	83.6
Black 1		88.0	81.5	84.7	82.6
Grey 1			85.8	84.7	86.9
TTBV				84.7	98.9
TBV					85.8

Figure 5.8 (b): Pairwise percent amino acid sequence identities between the partial coat protein sequence of the three narcissus virus cDNA sequences, Black.1, Black.8 and Grey.1, and tulip top-breaking virus (TTBV: Dekker, *et al*., 1993), tulip breaking virus (TBV; Dekker, *et al*., 1993) and turnip mosaic virus (TuMV; Tremblay, *et al*., 1990).

RT-PCR using a universal primer by Dekker *et al.* (1993), although the primers used fall within the same C-terminal region of the coat protein ORF, they generate a smaller region for sequencing. Thus, the alignment shows only the available data for TBV and TTBV and the corresponding region for the narcissus clones and TuMV. The percentage amino acid identity in this C-terminal region between the Black and Grey virus cDNA sequences in this region is higher (88%-89.1%) than over the whole region that has been sequenced (86%-86.5%) (figure 5.8b).

Discussion

Further investigation of NYSV and NLSYV was needed in order to establish whether they were distinct viruses, and to determine the nature of their relationship to NLV. Historically, they have always been considered distinct potyviruses due to their very different aetiology, but lack of a reliable experimental host for either virus has led to the production of unreliable antisera and therefore problems of identification. Careful analysis was needed in order to discover which antisera were specific, and whether these results could be confirmed by RT-PCR.

NYSV cytoplasmic inclusion protein (CIP) antiserum

The NYSV CIP antiserum raised by Mowat *et al.* (1989) reacted only when NLV was present. It did not react with the samples 6 or 8 (isolated in 1996) or those from HRI (13-16) which were not infected with NLV. NYSV CIP antiserum also reacted to MacMV-infected *N. clevelandii* (figure 5.2b) These results suggest that NYSV CIP antiserum was raised to NLV cytoplasmic inclusion protein, though the western results do not alone conclusively prove this.

NYSV coat protein antiserum

The NYSV coat protein antiserum from HRI gave a positive result with only one of the four NYSV samples from HRI (sample 15). This could be due to differences in protein preparation methods, or antibody concentration used. HRI samples tested by ELISA and the same antiserum was used at a very low titre to ensure a clean western result. Positive results using this antiserum were also obtained with some of the SCRI narcissus samples. However, it was not possible to obtain a narcissus sample that would react only to NYSV and not to NLSYV antiserum. Samples that were negative for NLV were positive for NYSV, NLSYV and the universal potyvirus antiserum. Although different in their aetiology, western blotting could not distinguish between these two viruses. This could be because either NYSV and NLSYV are closely related to one another and therefore their antisera would cross-react, or one or both of the antisera were raised from mixed infections. Narcissus late seasons yellows symptoms appear very late in the season and could be masked by an early NYSV infection, thus

making it more likely that the NYSV antiserum was raised from a mixed infection. However, it is impossible to investigate the NYSV antiserum further since the virus cannot be isolated in an experimental host. These problems with antisera lead us to use an RT-PCR technique where single or mixed infections could be qualified.

RT-PCR of NYSV AND NLSYV

Since all three samples (6, 8 (1996) and 15) which were positive for NYSV and NLSYV reacted positively with the universal potyvirus antiserum, and both viruses were described as potyviruses in the literature, the use of a universal potyvirus PCR primer was appropriate. This primer has been used on known potyviruses (Pappu *et al.*, 1993) as previously described in Chapter 3.

RT-PCR on sample 8

The universal potyvirus primer proved successful when used on RNA extracted from sample 8. The resulting product was cloned, sequenced and designated Black.8. Since sample 8 reacted with both NYSV and NLSYV antisera, it was necessary to establish whether the potyvirus primer produced a mixed population of products from two or more potyviruses. The transformants resulting from the ligation of the RT-PCR product of sample 8 were screened first with the insert from the sequenced clone (pNP8.4) and secondly with the original RT-PCR product. Both screenings gave the same results, establishing that only one potyvirus was detected in sample 8. The narcissus plant from which the material was taken displayed symptoms only late in the season and had been identified as NLSYV from pathology alone. It is therefore highly unlikely that the sample contained NYSV in a mixed infection since yellow stripe symptoms would have been induced early in the season. Therefore, it is probable that the sequence Black.8 represents the C-terminus of the coat protein of NLSYV. A BLAST search demonstrated that Black.8 had good homology to other potyvirus coat proteins.

RT-PCR on Sample 1

The same RT-PCR experiment was repeated using sample 1, which reacted positively with all antisera used. This time the colony screening showed that there were two potyviruses present in the RT-PCR product. It also showed that the two clones were related at the nucleotide level since some faint hybridisation had occurred between them. The first clone (Black.1) isolated from the screen using the insert from pNP8.4 as a probe, was 95 % identical at the nucleotide level to the viral sequence Black.8. The second group of clones, pNP1.34, pNP1.35 and pNP1.47, were isolated from those colonies which hybridised only faintly to pNP8.4 but which hybridised well to the total RT-PCR product. These clones represented the viral sequence Grey.1 and this was only 68% identical to the viral cDNA sequence Black.8.

This new viral sequence, Grey.1, was isolated from a naturally infected narcissus (sample 1). Sequence data showed that this was not a clone of NLV, therefore this newly cloned potyvirus was probably not responsible for the positive reaction to NLV antisera. None of the three virus cDNA clones isolated showed any homology with NLV. The four samples obtained from HRI (13-16) were not positive for NYSV CIP antisera, nor NLV, and likewise, NLV samples did not react to NYSV CP antisera. This supports the suggestion, that NLV and NYSV are distinct viruses and that NYSV CIP antiserum was raised from NLV cytoplasmic inclusion protein. Mowat *et al.* came to a different conclusion, suggesting that NLV and NYSV were synonymous (Mowat, *et al.*, 1991). They concluded that the reported 750 nm particles of NYSV could be attributed to NLSYV, since one antiserum to NLSYV contained antibodies to NYSV. Sequence data has demonstrated that this is incorrect and there is no such relationship between NYSV and NLV.

The viral sequence Grey.1 showed a lower level of identity (68%) to the viral sequence Black.8 (NLSYV) than would be expected of a strain of NLSYV. The results from sample 8, suggest that the NYSV antiserum could be contaminated with antibodies to NLSYV, since only one potyvirus was detected by RT-PCR, yet both NLSYV and NYSV antisera gave positive results. However, the NYSV antiserum, if mixed, should still contain antibodies to NYSV, and should not be discounted. It is also a possibility that NLSYV and NYSV are closely related viruses and hence would have a close serological relationship, causing a cross-reaction.

Sample 1 is known to contain two potyviruses by RT-PCR and colony screening. The sequence Black.1 was demonstrated to represent a strain of Black.8, which is probably NLYSV. The sequence Grey.1 could therefore represent the C-terminus of the coat protein region of a second potyvirus, NYSV. This hypothesis is not supported by aetiology, unlike the designation of NLSYV, since it cannot be shown that NYSV is the cause of yellow stripe symptoms in this naturally infected plant. At the amino acid level, there is a higher degree of identity between NYSV and NLSYV (88%) which could offer an alternative explanation to the theory that the serological cross-reaction between them occurs as a result of mixed antisera.

<u>Amino acid sequence alignments</u>

A BLAST search revealed that all three clones had homology to the potyvirus, TuMV, which is closely related to several tulip-infecting viruses (Dekker *et al.*, 1993). Tulip top-breaking virus (TTBV) is strain-related to TuMV with only one amino acid difference between them, and tulip breaking virus (TBV) was identified as a related but

distinct virus. When the available sequence for TuMV, TTBV and TBV is aligned with that for the three narcissus viral cDNA sequences (see figure 5.8), the homologies between them become apparent. According to Ward *et al.* (1995), four levels of classification can be observed when amino acid sequences are aligned. These are, genus, species, sub-species and strain. The narcissus viral sequences Black (NLSYV) and Grey (NYSV) have a level of percentage sequence identity to TBV, TTBV and TuMV in the order of sub-species by this method of classification. However, it should be noted that the region of the coat protein available for analysis is smaller than the core coat protein region usually used for this analysis and may be unintentionally weighted towards an area of high consensus at the coat protein C-terminus.

It is interesting that two narcissus viruses should be related to a group of tulip-infecting viruses. It has been generally stated that NYSV and NLSYV are restricted to narcissus in their host range, and it was surprising to find such closely related viruses in other bulb crops. Although both viruses are transmitted by aphids (Brunt, 1971; Mowat *et al.*, 1988b), NYSV is reported to spread slowly in Northern Europe (Hawker, 1943) because apterous aphids rarely colonise daffodils and migratory alatae are usually present for only 4-8 weeks before leaves senesce (Brunt, 1971). Selection pressure on these bulbous viruses could therefore be highly restricted to the infected bulb and its clones. We might therefore expect to see a degree of variation between the same viruses isolated from different clones, as is seen between Black.1 and Black.8. It is possible that NYSV and NLSYV are strains of TBV and TTBV and that they share a common evolutionary origin. Perhaps the taxonomic level of strain and sub-species is not applicable to these bulbous-infecting viruses where their gene pools have been isolated for many years due to lack of movement between hosts. These viruses, due to the use of vegetative propagation of many bulb crops could be sustained by clonal transmission at a higher rate than by aphids. Many may have become non-transmissible due to the lack of vector selection pressure. This may be indicated by the mutation of an aphid transmission motif found near the N-terminus of all aphid transmitted viruses (Atreya *et al.*, 1990; 1995), however, sequence data do not extend this far as yet. Interestingly, the sequenced isolate of NLV, which has not been tested for aphid transmission, has a mutated form of the DAG motif, and is missing another important domain associated with transmission (Chapter 4).

The usefulness of the universal potyvirus primer in this situation where difficulties presented by mixed antisera, mixed infections and closely related viruses was considerable. It was therefore decided to attempt to design a macluravirus specific primer which could be used to isolate new members of the genus (Chapter 6).

Future work

In order to conclusively establish that the viral sequences Black.1 and Black.8 are NLSYV and that Grey.1 is NYSV further analysis must be carried out. Now that clones are available, they could be used to probe RNA samples from the range of infected narcissi available to confirm the serological results and establish if mixed infections are present. RNA dot blots would be more accurate than serological screening and could therefore be used to survey naturally infected populations of narcissus for a NYSV infection that did not contain NLSYV. Further work with infected narcissi was outside the scope of this project as several years would be required to obtain manual infections. Work to establish disease aetiology by correlating the presence of a virus with symptoms should ultimately fulfil Koch's postulates.

Populations could be rapidly screened by RT-PCR using primers designed specifically to differentiate between NYSV and NLYSV in order to gain further information of the level of homology between viruses isolated from different clones. More sequence information could be used to search for aphid transmission motifs.

The narcissus samples used to obtain the Grey and Black cDNA clones could be re-tested with other universal primers, or conventional double-stranded cDNA clones could be made in order to look for further narcissus viruses. There are still other narcissus 'poty'-viruses that require investigation, such as narcissus white streak virus (NWSV) and narcissus degeneration virus (NDV). This task may be easier now that probes and sequence data exist for four narcissus viruses (NMV, NLV, NYSV and NLYSV), thus allowing them to be eliminated from mixed infections.

Chapter 6

The design and use of a maclura- and bymovirus specific primer

Introduction

In Chapter 3, a universal carlavirus primer failed to amplify MacMV or NLV cDNA leading to the conventional cloning of both viruses (Chapter 4; Badge *et al.*, 1997a). In order to confirm the proposal that the macluraviruses should form a distinct genus within the *Potyviridae*, more information is required. More macluraviruses need to be identified, characterised and sequenced to fully establish the parameters required to define the genus. The identification of plant viruses, especially those that have been previously unclassified or properly characterised is time consuming. RT-PCR now offers a rapid solution and allows sequencing to quickly confirm a positive result with more certainty.

A region specific to the macluraviruses suitable for such a primer has already been identified (Chapter 4). As there were no potential macluraviruses available for testing, an alternative test was sought. The bymoviruses share this specific region of homology with the macluraviruses and therefore a primer could be tested on rice necrosis mosaic virus (RNMV) which had not been sequenced and was already classified as a bymovirus.

The bymoviruses

The bymoviruses are considered a distinct genus of the *Potyviridae* (Usugi *et al.*, 1989; Barnett, 1992; Murphy *et al.*, 1995) due to their unusual particle morphology and mode of transmission. The viruses considered members of the group are: barley yellow mosaic virus (BaYMV, the type member), barley mild mosaic virus (BaMMV), wheat spindle streak mosaic virus (WSSMV), rice necrosis mosaic virus (RNMV) and oat mosaic virus (OMV). They have a narrow host range, restricted to cereals, and are transmitted by the soil-borne fungus *Polymxa graminis* (Inouye & Saito, 1975; Herbert & Panizo, 1975; Inouye & Fujii, 1977). All have flexuous filamentous particles with two modal lengths of 200-300 nm and 500-600 nm, except OMV for which only one particle size has been recorded, 600-750nm (Herbert & Panizo, 1975). All bymoviruses have two species of single-stranded positive-sense RNA of 8 kb and 4kb (RNA 1 and 2 respectively), although those of OMV are larger (Usugi *et al.*, 1989). The bymoviruses have a poor serological relationship to the potyviruses. BaYMV, WSSMV, RNMV and BaMMV are more closely related to one another than to OMV.

The nucleotide sequence of the complete genomes of two bymoviruses have been published, BaYMV (Kashiwazaki *et al.*, 1989; 1990a; 1991; Davidson *et al.*, 1991; Peerenboom *et al.*, 1992) and BaMMV (Kashiwazaki *et al.*, 1992; Foulds *et al.*, 1993; Schlichter *et al.*, 1993; Timpe & Kuhne, 1994; Kashiwazaki, 1996; Meyer & Dessens, 1996; Peerenboom *et al.*, 1996) and partial sequence data is available for WSSMV (Sohn *et al.*, 1994). The polyprotein encoded by BaYMV RNA 1 is similar to the C-terminal three-quarters of the polyprotein of potato virus Y (PVY) (Kashiwazaki *et al.*, 1990a). Homology is low in the coat protein region, but higher in the regions representing the cytoplasmic and nuclear inclusion proteins. The polyprotein encoded by RNA2 of BaYMV is very different from the N-terminal third of that expressed by PVY. This region in the potyviruses contains the helper component protein (HC) which has been shown to encode aphid transmissibility. It is therefore not surprising that the fungus-transmitted bymoviruses should have such a diverse polyprotein expressed by a separate RNA. However, the HC protein also encodes a protease (Oh & Carrington, 1989) which centres around two active residues, cysteine and histidine. This motif is conserved in the bymovirus RNA2 ORF, suggesting a common origin or function.

Further information about this unusual group of viruses may elucidate more clearly the function of RNA 2 and the nature of their narrow host range. It has been suggested that the bymoviruses could consitute one of the most ancient forms of plant viruses, transmitted by relatively immobile vectors, and that the vast diversity of host range shown by potyviruses represents a move towards adaptive radiation in the face of greater selection pressure, namely aphid transmission, a highly mobile and effective vector (Shukla *et al.*, 1994).

Design of a maclura- and bymovirus PCR primer

In order to provide a rapid method to detect new macluraviruses which could also be utilised to provide sequence information, it was decided to design a universal macluravirus PCR primer. Such a technique has been proved to be a powerful tool in the identification of plant viruses (Rybicki & Hughes, 1990; Langeveld *et al.*, 1991; Badge *et al.*, 1996). After close examination of the macluravirus and bymovirus coat protein genes, it was decided that this was not a suitable region for the design of a universal PCR primer since there is little homology between the bymoviruses in this area. Instead, the nuclear inclusion body b (NIb) ORF was utilised to design a primer specific to the bymoviruses and that would not cross react to the potyviruses. As sequence data in this region was available for only two macluraviruses (Badge *et al.*, 1997a; Chapter 4), it was decided to compare this region to three members of the

bymoviruses, as there is high homology between the macluraviruses and the bymoviruses within the NIb. The macluraviruses share a consensus sequence with the bymoviruses in the NIb ORF around the NGDD motif (Badge *et al.*, 1997a; Chapter 4), the RNA-dependent RNA polymerase binding site (Domier *et al.*, 1987; Dougherty & Carrington, 1988), and it was to this region that the primer was designed and named after (NGDD primer).

RNMV

RNMV has not received the same attention as the barley-infecting bymoviruses. It occurs in Japan and India in rice, the only known host, and causes mosaic symptoms characterised by spindle shaped yellow flecks and streaks on the lower leaves (Inouye & Fujii, 1977). RNMV is closely serologically related to barley yellow mosaic virus (BaYMV) and wheat spindle streak mosaic virus (WSSMV) (Usugi & Saito, 1976). It was decided to test the NGDD primer on this bymovirus which had not yet been sequenced.

OMV

OMV is an unusual member of the bymovirus genus, with a larger than normal particle and genome size, yet a slightly smaller coat protein (30 kDa). It is generally reported as having only one modal length of particle of 600-750 nm, although unpublished reports cited by Usugi *et al.* (1989) suggest that a second modal length of 250-300 nm is observed in some cases. Despite the problems associated with the identification of one or two particle lengths, it has been established that OMV is encoded by a bi-partite genome (Usugi *et al.*, 1989), though each RNA is larger than that normally expected of the bymoviruses. OMV is the most distantly related virus of the genus by serology. A PCR test would be ideal to use on this bymovirus, however, it was uncertain whether infected material was available.

Aims

In order to gain sequence information for a fourth bymovirus, RNMV, and to establish a rapid PCR test for the identification of macluraviruses and bymoviruses, the NGDD PCR primer was utilised. Such a PCR test could be used to identify new macluraviruses and bymoviruses.

Results

Primer design

When the amino acid sequences of the NIb protein available for bymoviruses and macluraviruses are aligned, a clear consensus around the motif NGDD which is unique to these viruses and distinct from the potyviruses (figure 6.1) is apparent. The

Figure 6.1: Diagram to show the design of the NGDD PCR primer. Alignment of the amino acid sequence around the NGDD motif in the nuclear inclusion body b (NIb) ORF shows that there is consensus among the bymoviruses and macluraviruses, but variation among the potyviruses, with nucleotide sequences also being conserved in this region. The NGDD primer based on its consensus sequence is indicated (Y=C/T, B=C/T/G). Three strains of barley mild mosaic virus, BaMMV are presented: Japanese isolates, Na I (GenBank: D10949) and Ka I (GenBank: D10947; Kashiwazaki *et al.*, 1992) and one German (G) isolate (GenBank: X69203; Schlichter *et al.*, 1993), two strains of barely yellow mosaic virus, BaYMV: Japanese (J) (GenBank: D01091; Kashiwazaki *et al.*, 1990) and German (G) (GenBank: X6975; Peerenboom *et al.*, 1992) and only one isolate of wheat spindle streak mosaic virus (WSSMV, GenBank X73883; Sohn *et al.*, 1994).

BYMO- & MACLURAVIRUSES		F	V	C	N	G	D	D	N	K
		|	|	|	*	*	*	*	|	|
POTYVIRUSES		F Y M	F Y	V A G I	N	G	D	D	L I	L I C V
		F	V	C	N	G	D	D	N	
BaMMV	(Ka1)	T T C	G T C	T G C	A A T	G G T	G A T	G A C	A A C	
	(Na1)	T T C	G T C	T G C	A A C	G G C	G A C	G A C	A A C	
	(G)	T T C	G T C	T G C	A A T	G G T	G A T	G A C	A A C	
BaYMV	(J)	T T T	G T T	T G C	A A T	G G C	G A T	G A C	A A C	
	(G)	T T T	G T T	T G C	A A C	G G T	G A T	G A C	A A T	
WSSMV		T T T	G T C	T G C	A A T	G G T	G A C	G A C	A A C	
MacMV		T T T	G T G	T G T	A A T	G G T	G A T	G A C	A A C	
NLV		T T T	G T T	T G C	A A T	G G T	G A C	G A C	A A C	
consensus		T T T C	G T C T G	T G C T	A A T C	G G T C	G A C T	G A C	A A C	
NGDD primer		T T Y	G T B	T G Y	A A Y	G G T	G A Y	G A C	A A	

Figure 6.2: Analysis of RT-PCR reaction using narcissus latent virus (NLV) (known to contain the NGDD primer sequence), NGDD and d(T)Not I primers. Lane A, kilobase DNA molecular weight marker, (GibCo); lanes 1 & 2, total plant RNA extracted from NLV-infected *N. clevelandii* (twice as much cDNA used in PCR reaction in lane 2 than 1); lane 3, healthy *N. clevelandii* RNA; lane 4, positive control (pNLVdT1 plasmid known to contain both NGDD and oligo-d(T)Not I sequences); lane 5, negative control (no cDNA in PCR reaction).

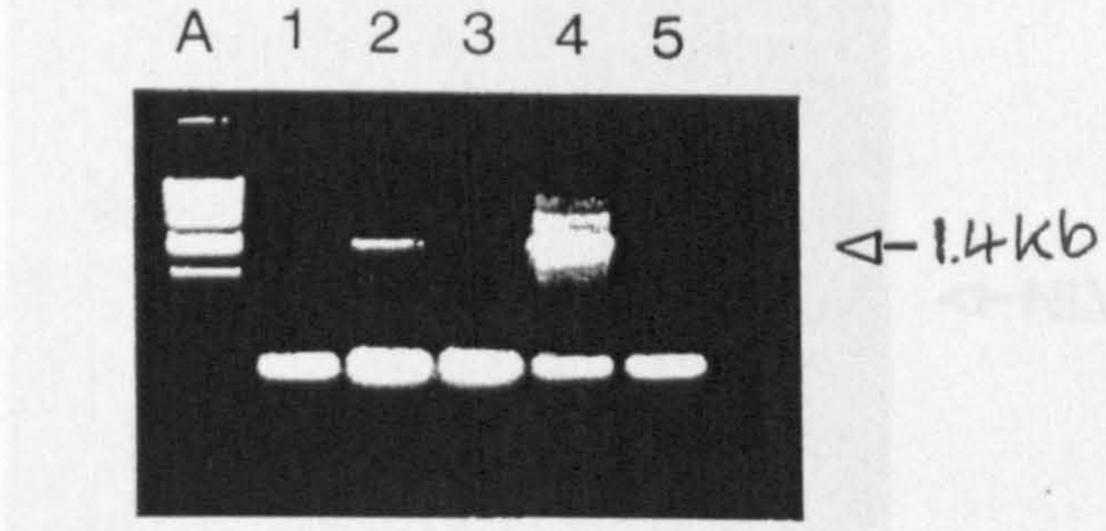


Figure 6.3: Analysis of RT-PCR reaction using NGDD and oligo-d(T)Not I and; lane 1: total plant RNA extracted from RNMV infected rice, lane 2 : total plant RNA extracted from healthy rice, lane 3: total RNA extracted from BaMMV-Ka I-infected barley, lane 4: total RNA extracted from BaMMV-Na I-infected barley, lane 5: total RNA extracted from BaYMV-II-i -infected barley, lane 6: total RNA extracted from healthy barley, lane 7: pNLVdT1 positive control, lane 8: negative water control containing all RT-PCR components except cDNA. (reactions carried out by Dr Kashiwazaki)

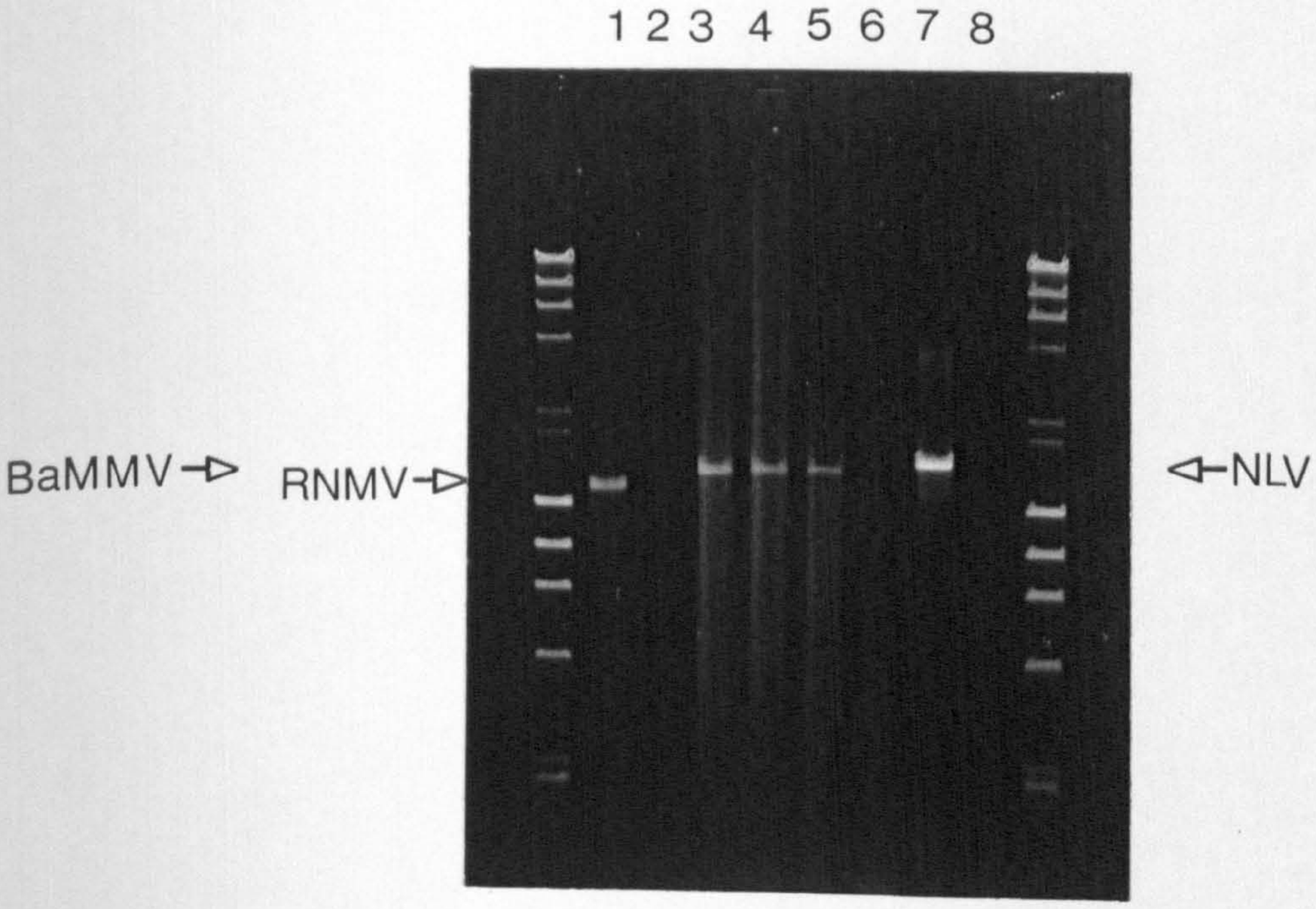


Figure 6.4 : Denaturing gel of total plant RNA extracted from: Lanes 1 & 2, Healthy oat sample (twice the amount loaded in lane 2 to lane 1) ; Lanes 3 & 4, OMV-infected oats

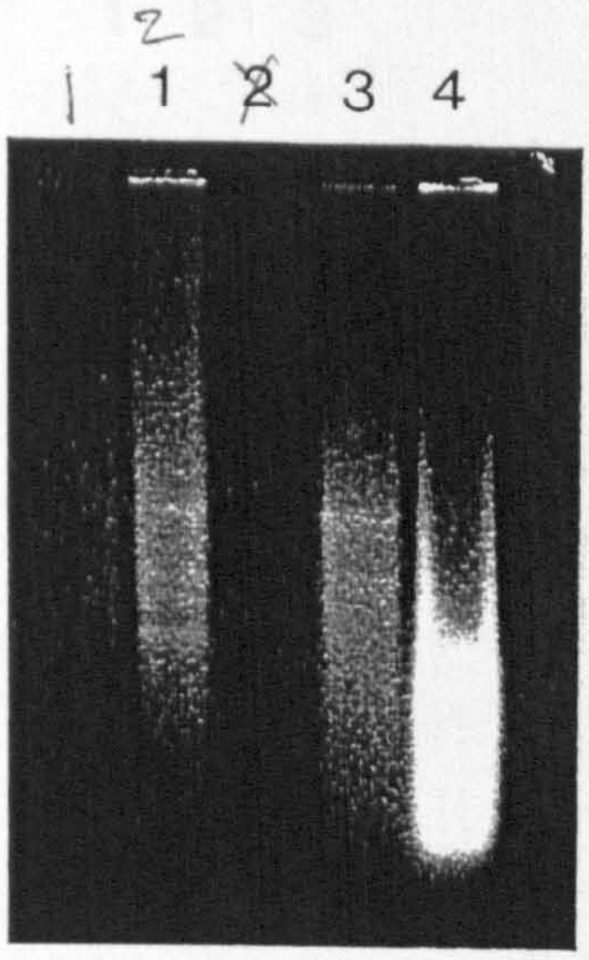
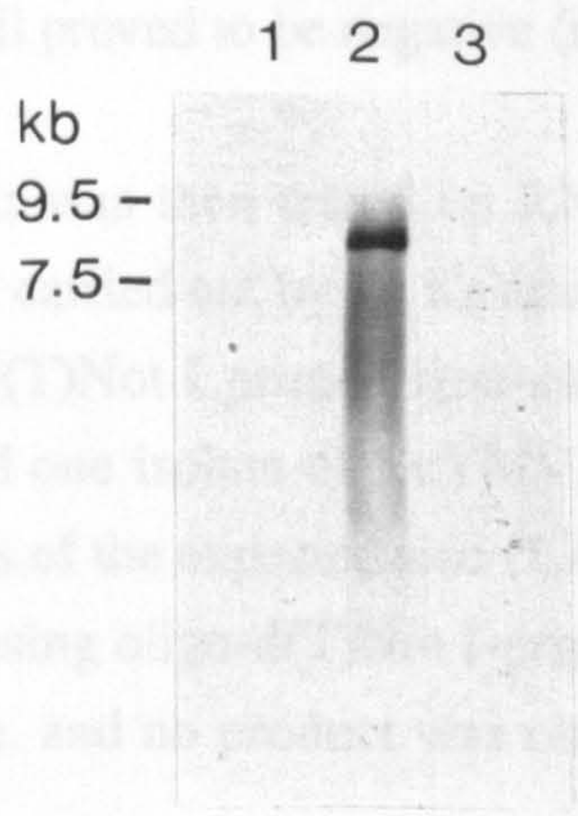


Figure 6.5: Northern blot of RNA extracted from virus particles and probed with *EcoRI* fragment of pRNMV4. Lane 1: RNA marker (GibCo BRL) ; Lane 2: RNMV; Lane 3: BaMMV. Size of the fragment is shown relative to the two closest markers.



nucleotide sequences are also well conserved and is degenerate at only 5 positions out of 24, and was therefore used as a primer for the rapid isolation of the 3' terminus of the genome of RNMV.

RNA and RT-PCR

Initially the NGDD PCR primer was tested on a virus known to contain the sequence from which the primer was derived. Figure 6.2 shows the results of using the NGDD primer on first-strand cDNA primed with oligo-d(T)Not I generated from NLV-infected total plant RNA (PCR conditions as for Carla-Uni, Chapter 3). A fragment of approximately 1.6 kb was generated which agreed with the size estimated by genome analysis. A corresponding reaction using healthy *N. clevelandii* RNA produced no fragment. Control reactions using only one or other of the primers, or no cDNA were also carried out, which all proved to be negative (results not presented).

The NGDD PCR primer was then tested on RNA from tissue infected with known bymoviruses (figure 6.3; carried out by Dr Kashiwazaki, National Agriculture Research Center, Japan). Oligo-d(T)Not I primed first-strand cDNA was generated from two isolates of BaMMV and one isolate of BaYMV. RT-PCR using oligo-d(T)Not I and NGDD resulted products of the expected size (1.4kb) (figure 6.3). Similarly, a product of 1.4kb was obtained using oligo-d(T)Not I-primed first-strand cDNA obtained from RNMV-infected sample, and no product was obtained from RT-PCR on healthy rice samples.

A second bymovirus, oat mosaic virus (OMV), which has not yet been sequenced was also tested using the NGDD PCR primer. Figure 6.4 shows the total plant RNA extracted from healthy and OMV-infected RNA, however, no PCR product was obtained despite repeated attempts.

The RNMV product was cloned. The insert was subsequently confirmed to be of viral origin by northern blotting, producing a 8 kb band on RNMV RNA only, and not reacting with BaMMV RNA or the RNA marker (figure 6.5).

Sequencing

Figure 4 represents the sequence obtained from the three independent clones, containing the insert of the RT-PCR product, pRNMV4, pRNMV1A29 and pRNMV2A18. These clones yielded sequence for the 3' terminal portion of RNMV, a total of 1431 nt were sequenced (figure 6.6). Two nucleotide differences were found between the three clones. In both cases a consensus of two nucleotides is shown in the sequence presented (figure 6.6), and the third alternative base is shown above. The first of these

Figure 6.6: Nucleotide sequences of pRNMV4 and pRNMV1A29 and pRNMV2A18 representing the 3' terminal region of RNMV. Where there were nucleotide differences between clones, the alternative base is shown above and its translation is shown below the line of data. Important motifs are also underlined, namely 'NGDD' which forms the core of the primer, 'LQA' as a putative cleavage site and NGTS which is found in other members of the *Potyviridae* family.

```
                10                  20                  30                  40                  50                  60                  70
                |                   |                   |                   |                   |                   |                   |
TTCGTTTGCAATGGTGACGACAACAAGTTTGCTATATCTCCAGAATTCTACTGGAAATATGGCTGTGACTTCTCACCG
F  V  C  N  G  D  D  N  K  F  A  I  S  P  E  F  Y  W  K  Y  G  C  D  F  S  P

                80                  90                 100                 110                 120                 130                 140                 150
                |                   |                   |                   |                   |                   |                   |                   |
TATCTCAGTGAGCTTGGTTTGAAGTACGAGTTTGACGAGATCACAGATGACATCTGTCTCAACCCGTACATGTCACTG
Y  L  S  E  L  G  L  K  Y  E  F  D  E  I  T  D  D  I  C  L  N  P  Y  M  S  L

                160                 170                 180                 190                 200                 210                 220                 230
                |                   |                   |                   |                   |                   |                   |                   |
ACCATGACACCAACTGAGCTTGGAATTGGGTTTAGCCTTGCCCCACAACGCATTGTTGCAATAGTTCAATGGAGTAGG
T  M  T  P  T  E  L  G  I  G  F  S  L  A  P  Q  R  I  V  A  I  V  Q  W  S  R

                240                 250                 260                 270                 280                 290                 300                 310
                |                   |                   |                   |                   |                   |                   |                   |
GCTGGGGGAGTTTTGCATGCATACCTTGCAGCAATTGCAGCTTGCATTGAATCCTTCAACACGCAACGACTCTACATG
A  G  G  V  L  H  A  Y  L  A  A  I  A  A  C  I  E  S  F  N  T  Q  R  L  Y  M

                320                 330                 340                 350                 360                 370                 380                 390
                |                   |                   |                   |                   |                   |                   |                   |
TTGATGAGAACTTATGCACTATATCTGATGGACAGGCATAAGCACGAAATAGCGGCGTTGGCGGAGCTGAAGGGCATG
L  M  R  T  Y  A  L  Y  L  M  D  R  H  K  H  E  I  A  A  L  A  E  L  K  G  M

                400                 410                 420                 430                 440                 450                 460
                |                   |                   |                   |                   |                   |                   |
GAGGCCATGAAACTCCCAACTTCCTGGGATATCGCCATCATGCATTATGTGGACATGCCAGACTACTATCAGTACGGT
E  A  M  K  L  P  T  S  W  D  I  A  I  M  H  Y  V  D  M  P  D  Y  Y  Q  Y  G

                470                 480                 490                 500                 510                 520                 530                 540
                |                   |                   |                   |                   |                   |                   |                   |
GATATGACAACAGCAAACATTGTGATTCTTGATGACGGGGTTCCCGACTCTGAAGACTCAATGACTTATCTCCAAGCT
D  M  T  T  A  N  I  V  I  L  D  D  G  V  P  D  S  E  D  S  M  T  Y  L  Q  A

                550                 560                 570                 580                 590                 600                 610                 620
                |                   |                   |                   |                   |                   |                   |                   |
GGTCCCGGGGATGGAGAAAATCTTGCTGAAACTGAGAACCTTCCCAGAACCGGAGATCTGCCCATTATACCCAAAACT
G  P  G  D  G  E  N  L  A  E  T  E  N  P  P  R  T  G  D  L  P  I  I  P  K  T

                630                 640                 650                 660                 670                 680                 690                 700
                |                   |                   |                   |                   |                   |                   |                   |
CAAGCAACATGGAGTCTCCCAACTATCAAACCAAAAGCTATCGGGTTCACCACAAAGATCCCCATTGATCTCATTGCC
Q  A  T  W  S  L  P  T  I  K  P  K  A  I  G  F  T  T  K  I  P  I  D  L  I  A

                710                 720                 730                 740                 750                 760                 770                 780
                |                   |                   |                   |                   |                   |                   |                   |
CAAGTTCCAGAGGGATACGTTGAAACACTCTCAATCAAGGCGACAACAAGCCAGCATCAGAAATGGGTTCAGGATGTG
Q  V  P  E  G  Y  V  E  T  L  S  I  K  A  T  T  S  Q  H  Q  K  W  V  Q  D  V
```

```
          790        800        810        820        830        840        850
          T          |          |          |          |          |          |
AGGTCAGACCTCGGGATAACAGACGATGACACGTGGCACAAGGTTATTGCACCCGCATGTATTTTCTTTGGAGACAAT
R  S  D  L  G  I  T  D  D  D  T  W  H  K  V  I  A  P  A  C  I  F  F  G  D  N
          P

 860        870        880        890        900        910        920        930
 |          |          |          |          |          |          |          |
GGAACATCTGAAGACATGGACGAAGATCAGGTCATTGAGGTGAAAAGCGGATTAAATTCAACTCAGGAACTGCCAGCC
G  T  S  E  D  M  D  E  D  Q  V  I  E  V  K  S  G  L  N  S  T  Q  E  L  P  A

  940        950        960        970        980        990       1000       1010
  |          |          |          |          |          |          |          |
AAACCATTCATACGGCACGCCCGTCGCAATGCAACACTTAGAAAGATCATGAGGCATTATTCATTTGAAACTAAGCTG
K  P  F  I  R  H  A  R  R  N  A  T  L  R  K  I  M  R  H  Y  S  F  E  T  K  L

      1020       1030       1040       1050       1060       1070       1080       1090
      |          |          |          |          |          |          |          |
CTCTTCATGAAAATGCGACGCATTCCTCATTGGGCTGTTAAGCACGGCTGTGTGGATGAAATTGTGTTTGACTTTATG
L  F  M  K  M  R  R  I  P  H  W  A  V  K  H  G  C  V  D  E  I  V  F  D  F  M

       1100       1110       1120       1130       1140       1150       1160       1170
       |          |          |          |          |          |          |          |
ATTCCCGACGCTTTCACGTCAAGGCAGGCAGTTGAGAAGCTCAAACAAACCAAAATCAGCGCCATTGGTGTTGGCACA
I  P  D  A  F  T  S  R  Q  A  V  E  K  L  K  Q  T  K  I  S  A  I  G  V  G  T

            1180       1190       1200       1210       1220       1230       1240
            |          |          |          |          |          |          |
TCCAACTATATGCTCACATCTGAAACAACCAACTTAAGGCGCACCGAAACACGTAGAAGAAACGACTATGATGGGCAT
S  N  Y  M  L  T  S  E  T  T  N  L  R  R  T  E  T  R  R  R  N  D  Y  D  G  H

 1250       1260       1270       1280       1290       1300       1310       1320
 |          |          |          |          |          |          |          |
GAAGGGCTCATCCATTAGAACTTGCTTCGAACTTTAAATAAAGTTGTTCTATCGAGATATACTCGAGCACCCAGCTTA
E  G  L  I  H  •

   1330       1340       1350       1360       1370       1380       1390       1400
   |          |          |          |          |          |          |          |
AGACTAACCCGAAAAGGGTGAAGGTGTCTATCACCAGTCTGATTTCCTAATCAACCATCGCTGGTTGGTGGCATGGTT

      1410       1420       1430
      |          |          |
CGCCATGTTTGCGTACACCGGAAGGTG
```

two differences at position 587 nt represents no change in amino acid encoded, but the second at position 790 nt, the consensus amino acid is leucine (L), and the alternative represents no overall change in charge (proline, P). Analysis showed that a large single open reading frame (ORF) of 421 aas was present, with a stop codon at position 1264 nt and hence an untranslated region (UTR) of 167 nt.

Amino acid sequence analysis

From the nucleotide sequence the putative amino acid sequence was deduced (figure 6.6). Figure 6.7 shows an alignment of RNMV sequence obtained and the corresponding sequence for the other bymoviruses, barely mild mosaic virus (BaMMV : Kashiwazaki *et al.*, 1992), barley yellow mosaic virus (BaYMV : Kashiwazaki *et al.*, 1990a), wheat spindle streak mosaic virus (WSSMV : Sohn *et al.*, 1994) and a typical potyvirus, potato virus Y (PVY : Robaglia *et al.*, 1989). All potyviruses are translated into a large polyprotein which is then cleaved by a viral encoded protease to form mature products. A heptapeptide consensus amino acid sequence required for a potyviral cleavage site has been determined by mutational analysis and *in vitro* translation studies (Dougherty & Carrington, 1988). The point of cleavage is between glutamine (Q) and serine (S) or glycine (G) in potyviruses but bymoviruses are thought to have a consensus cleavage site LQA (Kashiwazaki *et al.*, 1991; Foulds *et al.*, 1993; Sohn *et al.*, 1994) between NIb protein and the coat protein (except for one strain of BaMMV, Na I which has the cleavage site LQS (Kashiwazaki *et al.*, 1992; Kashiwazaki, 1996)). RNMV has the motif, LQA at position 240 aas which would generate a coat protein of 27.1 kDa.

Discussion

The use of a specific PCR primer in a an RT-PCR reaction is a rapid method to obtain information about new viruses. The primer NGDD was designed to a sequence unique to the macluraviruses and their close relatives, the bymoviruses. This enabled an RT-PCR product to be obtained and the 3' terminus of the genome of RNMV to be analysed.

NGDD PCR primer is effective on RNMV

The NGDD PCR primer was shown to be effective on infected total plant RNA. It detected the presence of NLV in *N. clevelandii* and RNMV in infected rice material. The primer failed to be effective on what is now the only member of the group for which there is no sequence data, OMV. There are many reasons why this reaction may have failed. Firstly, there could be problems with the PCR reaction itself, the annealing temperature or concentrations of magnesium chloride could be sub-optimal. *Taq* or reverse-transcriptase inhibitors could be present in the RNA extracted (see future

Figure 6.7: CLUSTAL W alignment of the partial ORF of RNMV (GenBank U95205) with the bymoviruses sequenced to date, barley mild mosaic virus (BaMMV: GenBank D10947), barley yellow mosaic virus strain II-1 (BaYMV: GenBank D01091), wheat spindle steak mosaic virus (WSSMV: GenBank X73883) and a typical potyvirus, potato virus Y (PVY : GenBank A08776). Gaps (-) have been introduced for maximum alignment and the program BOXSHADE was used to create boxing. Residues identical to RNMV are boxed with a black background, chemically similar residues are boxed in grey.

```
PVY       2625  FV-NGDDLLIAVNPEKRSILDR-MSQHFSDLGLNYDFSSRTRRKEELWEMSHRGLLIEGMYVP--KLEEERIVSILQWDRADLPEHRLEAICAAM
WSSMV     1024  FVCNGDDNKFSMSPSSFMAKDGCDFSPFLSELGLTYEFDEATHDICENPYMSLT--MVRTSFGIGFSLSIERIVAILQWSRAGGVLHAYLSGIAAL
BaYMV     1960  FVCNGDDNKFAISPQFDEEDGHDFSPELVELGLTYEFDDITSDICENPYMSLT--MVKTPFGVGFSLPVERIHAIMQWSKKGGVLHSYLAGISAI
BaMMV     1850  FVCNGDDNKFAISPEFNAQDGHDFSPELIELGLTYEFDDITDDICENPYMSLT--MVRTPFGIGFSLPVERIVAIMQWARRGGVLHSYLAGISAI
RNMV         1  FVCNGDDNKFAISPEFYWKYGCDFSPYLSELGLKYEFDEITDDICLNPYMSLT--MTPTELGIGFSLAPQRIVAIVQWSRAGGVLHAYLAAIAAC

PVY       2719  IESWGYSELTHQIRREYSWLLQQQP--FATIAQEGKAPYIASMALRKLYMDRAVDEEELRAFTHMMVALDDE-----------FKLDSYHVHHQA
WSSMV     1120  FESFNTPKLFNLVHTYLLWLVTEHEEELFSMMELKDM-FMPLPTREQIALLHYV--------GTEPI-----------VEETFLQAGHEEPDPIY
BaYMV     2056  YESFNTPKLFKSIYAYLLWLTEEHEAEILAAMT-QSSTALPIPSMLDVYRLHYGDDEIWLQAADPLTDAQKEAAHTAAADRARLDLADADRRRKY
BaMMV     1946  YESFNTPKLFKSIYAYLLWLTEEYEAQILAAMK-DTATALPIPSMLDVYRLHYGGCDIELQAADTQTDAQKEAARVAAADKARADAADAARKQKY
RNMV        96  IESFNTQRLYMLMRTYALYLMDRHKHEIAALAELKGMEAMKLPTSWDIAIMHYVDMPDYYQYGDMTT-----------ANIVILDDGVPDSEDSM

PVY       2813  NDTIDAGGSNKKDAKPEQGSIQPNPNKGKDKDVNAGTHGTHTVPRIKAITSKMRMPTSKGATV-----LNLEHLLEYAPQQIDISNTRATQSQFD
WSSMV     1214  ----------------P-----PVPDTDLTNMAAAP-PDNRKSRAVIPRGTSDWSLPEPKMRTLGFKSKINIETLANVPDGYMNTFASVATESQRR
BaYMV     2150  EADRVEAARVKKAADAVLKPVTLTATRNPTEDDGKLKTPSGARIPSSAADGNWSVPATKQVNAGLTLKIPLNKLKSVPKSVNEHNNSVALESELK
BaMMV     2027  EADRVEAARVKKAAAD---TANLTATKVTATEDGKVTTDSGTKRTSAAAEVTWTLPTMKQANAGLKLRIPIAKLKSVPKSVMQHDNSIALDSELT
RNMV       178  ----------------TYLQAGPGDGENLAETENPP-RTG--DLPIIPKTQATWSLPTIKPKAIGFTTKIPIDLIAQVPRGYVETLSIKATTSQHQ

PVY       2891  TWYEAVRMAYDIGETEM--PTVMNG--LMVWCIKNGTSPNVNGVWVMM--DGNEQV-EYPLKPIVENAK--PTLRQIMAHES-DVAEAYIEMRNK
WSSMV     1306  KWEEAARGDFGITDDEKWEKLLIAAC IYFA---DNGTSPNFDEELTMEVNGGLNSIKEYPVRPFVVRAKKISTLRHIFRCYSIETKLMFVKLRR-
BaYMV     2245  AWTDAVRTSLGITTDEAWIDALIP---FIGWCCNNGTSDKHAENQVMQIDSGKGAVTEMSLSPFIVHARMNGGLRRIMRNYSDETVLLITNNKL-
BaMMV     2095  AWADAVRTSLGITTDEAWQNTLIP---FLGWCCNNGASDKHSENQKMQVDAGKATLSEVSLSPFIVHARLHGGLRRIMRAYSDETVLLISEGKL-
RNMV       257  KWVQDVRSDLGITDDDTWHKVIAPACIFFG---DNGTSEDMDEDQVIEVKSGLNSTQELPAKPFIRHARRNATLRKIMRHYSFETKLLFMKMRR-

PVY       2980  KEPYMPRYGLIRNLRDNGLARYAFDFYEVTSRTPVRAREAHIQMKAAALKSAQPRLFGLDGGISTQEENTERHTTEDVSPSNHTLLGVKNM
WSSMV     1397  ----VPHWAIKHG--CLDEI--VFDFMIPDQFTSRTALETLKQTKLAAIGVGTSNSLLTSEQTNMRTTETRRRNDYDGHEALLR--------
BaYMV     2336  ----VAHWSMKHG--ASANAKYAFDFFVPRSWMNPQDIEVSKQARLAALGTGTYNTMLTSDTTNLRKTTNHRVLDSDGHPELT--------
BaMMV     2186  ----VPRWAMRHG--ASANAAYAFDFFVPRPWMNPQDIEISKQARLAALGTGTNNTMLTSDTTNLRKTTNHRVLDTDGHPELT--------
RNMV       348  ----IPHWAVKHG--CVDEI--VFDFMIPDAFTSRQAVEKLKQTKISAIGVGTSNYMLTSETTNLRRTTETRRRNDYDGHEGLIH--------
```

Table 6.1: Pairwise percent amino acid sequence identities between the core coat protein* of RNMV, a typical potyvirus, PVY and three bymoviruses, BaMMV, BaYMV and WSSMV.

	BaYMV	BaMMV	WSSMV	PVY
RNMV	52.5	57.8	53.2	20.8
BaYMV		60.3	85.7	21.4
BaMMV			61.7	21.3
WSSMV				21.4

* The core coat protein is equivalent to D_{2748} to R_{3043} in PVY.

work). The sequence may not be present in the viral genome or be in such a degenerate form that binding does not occur, yet OMV could still be a bymovirus. Finally, OMV may not belong to the group at all, and therefore may not contain the consensus sequence.

There are some doubts over the characteristics of OMV. It is often reported as having only one modal length of particle, *c.* 650-750 nm, larger than those normally seen within the genus. Unpublished reports (cited in Usugi *et al.*, 1989) state that two particles are occasionally observed in some experiments, supporting the observation that it has a bi-partite RNA genome. Further work is necessary to establish the relationship of OMV to the bymoviruses. Although it is transmitted by the same vector, has the same genome structure and similar limited cereal host range to the bymoviruses, it is perhaps the least characterised member of the genus.

RNMV shares homology with the bymoviruses

The sequence data from the C-termini of the predicted large polyprotein encoded by RNMV allows confirmation of its classification. Figure 6.7 shows an alignment using CLUSTAL W of the deduced amino acid sequences of RNMV, three members of the bymovirus genus, BaYMV, BaMMV, WSSMV and a typical potyvirus, potato virus Y (PVY). This region represents the C-terminus of the nuclear inclusion body b (NIb) protein and the complete coat protein. RNMV contains several important motifs conserved within the bymoviruses and potyviruses. The GDD motif conforms, as expected by the binding of the primer sequence, to the consensus of the bymoviruses. This motif is thought to be the part of the binding site of the viral RNA-dependent RNA polymerase encoded by the NIb ORF (Domier *et al.*, 1987). However, it should be noted that as this motif is part of the primer, the amplified version will usually be true to the primer and not necessarily to the target sequence in the genome A second motif, NGTS within the coat protein is also well conserved, though its significance is not known.

Potyviruses use a polyprotein translation method which involves a viral encoded protease to generate mature products by proteolytic cleavage. A cleavage site between the NIb and coat protein, with residues conserved around glutamine (Q) at the point at which cleavage occurs, can be observed in all potyviruses. A group-specific consensus cleavage sequence has been suggested V-X-X-Q/[A,S, G or V] based on sequence homology (Shukla *et al.*, 1994). However, of the six published bymovirus sequences, five conform to the consensus L-Q/A and one has the sequence L-Q/S at the NIb/coat protein cleavage site (Kashiwazaki *et al.*, 1990a; 1992; Kashiwazaki, 1996; Schlichter *et al.*, 1993; Peerenboom *et al.*, 1992; Sohn *et al.*, 1994). RNMV conforms to the

bymovirus consensus with a putative cleavage site of LQA (538 nt). This would generate a coat protein of 27.1 kDa, slightly smaller than the reported 33 kDa from SDS-PAGE analysis (Usugi *et al.*, 1989).

In general, the potyviruses have a very good degree of identity in the coat protein ORFs, and this is particularly striking in a defined core region (Shukla *et al.*, 1994). Shukla and Ward (1988) have shown that distinct species of the potyvirus genus have coat protein core sequence similarities within the range of 38% to 71% (average 54%). Table 6.1 shows that RNMV has a percentage similarity to the other members of the bymovirus genus of 52% to 57% in an equivalent to the core coat protein, and should therefore be classified as a distinct species of the bymovirus genus.

To conclude, our results using the NGDD primer have provided useful information about RNMV and could be further used to investigate potential macluraviruses or bymoviruses, although negative results would clearly require further investigation.

Future work

In order to establish whether it is the failure of the NGDD primer to detect OMV in infected oats, it will be necessary to test the PCR process. Simultaneous positive reactions could be carried out, using primers designed to bind to a well characterised oat gene, which would work in healthy and infected tissue. If this reaction also fails, then it may be due to *Taq* or reverse-transcriptase inhibitors present in the extracted RNA. Different extraction methods could be tried or the cDNA could be more dilute before use in the PCR reaction. If this 'healthy' PCR reaction is successful, then it can be assumed that either OMV is not a bymovirus and does not contain the sequence for the primer to bind, or that does belong in the group, being an exception to the rule and contains a degenerate form of the sequence.

Chapter 7

Can the replicase genes of MacMV and BaMMV confer resistance to macluraviruses in tobacco?

Introduction

A survey of older commercial narcissus stocks in 1980 revealed that many were infected with one or more viruses (Brunt, 1980) and observations have shown that virus free stocks are more productive and vigorous than infected ones (Brunt, 1980; Rees, 1966). NLV is thought to be one of the most widespread diseases of commercial narcissus crops, occurring alone and in complex with other viruses. The vegetative method of propagating bulbs for retail market scales up the problem of testing for viruses. Current methods for the production of virus free bulb stocks involve repeated testing of an elite foundation stock which are maintained in a protected environment (e.g.: gauze houses to protect from aphids). However, the production of resistant bulbs from which others could be propagated would be useful to the grower, and may possibly reduce testing costs and improve bulb and flower quality and vigour.

The amino acid identity shared by the macluraviruses and the bymoviruses in the nuclear inclusion body b (NIb) ORF is the highest that the bymoviruses share with any member of the *Potyviridae*. This unusual relationship led us to speculate that the macluraviruses may be suitable for use as an experimental model for the engineering of resistance to barley yellow mosaic disease, as well as investigating constructs for NLV engineered resistance. The bymoviruses have a very limited host range, infecting only graminaecious hosts in the field, making them difficult to work with in the laboratory and slow to propagate. The macluraviruses however, readily infect some tobacco species, which are widely available and have a year round growing season under glass.

Barley yellow mosaic disease

This disease causes serious losses in winter barley in Western Europe and Eastern Asia (Huth & Adams, 1990; Kashiwazaki *et al.*, 1990a). The agents of the disease, BaMMV and BaYMV are often found as a mixed infection. At one time, BaMMV was thought to be a strain of BaYMV but they are now recognised as distinct viruses (Huth & Adams, 1990). They are both transmitted by the soil-borne *Polymyxa graminis* by the *in vivo* method (for review see Campbell, 1996). This means that the virions are retained within the zoospores of the vector and hence can survive in infected soil for many years. Total crop protection is therefore extremely difficult to achieve, and the only practical method available is the use of resistant barley cultivars. However, all

German resistant winter barley cultivars possess the same resistance gene to BaYMV and resistant breaking strains have been reported in several European countries in 1988 (Adams, 1989; Hariri *et al.*, 1990; Bendiek *et al.*, 1993). Cultivars that are resistant to BaYMV can still be susceptible to infection by BaMMV (Kashiwazaki *et al.*, 1990b). New resistance genes have only been identified in "exotic" barley lines and these will require considerable time to be transferred to European high yielding cultivars (Bendiek *et al.*, 1993). Hence the need to use a genetic engineering strategy to create plants resistant to BaYMV and BaMMV.

Using viral sequences to confer resistance

In the past few years, virus-derived sequences have been successfully employed to engineer transgenic plants resistant to viral infections (Lomonossoff, 1995). This method of "pathogen-derived resistance" was first suggested by Sanford and Johnson (1985) and to date the main focus of the work has been on coat-protein mediated resistance (for review see: Hull & Davies, 1992). This involves the expression in transgenic plants of the coat protein genes of plant RNA viruses. However there are problems with this form of resistance. In some cases, protection can be overcome at high levels of inoculum, or by viral RNA. The possibility of heteroencapsidation, where the RNA genome of one virus is encapsidated by the coat protein of another virus expressed by the transgenic plant, could lead to the creation of new chimaeric viruses. Several strategies have been investigated to overcome these difficulties, such as the use of satellite RNA sequences or transformation of plants with defective coat protein genes which may still confer resistance. Attention shifted to a different mechanism to engender resistance using non-structural pathogen-derived genes, which overcomes some of these problems.

Replicase-mediated resistance

The use of viral replicase genes has enormous potential because it offers the possibility of resistance at an early phase of the viral infection cycle. If a virus cannot replicate, it cannot increase in number and colonise the plant. All RNA positive-sense single-stranded viruses share homology in their viral replicases (Kamer & Argos, 1984; Poch *et al.*, 1989). This conservation of the RNA-dependent RNA polymerases is striking around two motifs, the nucleotide triphosphate binding domain, and the 'GDD' motif which appears to be essential for polymerase activity (Inokuchi & Harishima, 1987; 1990; Li & Carrington, 1995). Complete viral replicase genes were found to confer resistance which was not overcome by high levels of inoculum or viral RNA (Lomonossoff & Davies, 1992) and later, defective replicase genes were also successfully used (Anderson *et al.*, 1992).

Resistance to potyviruses

The first potyvirus replicase gene successfully used to confer resistance was from potato virus Y (PVY). Audy *et al.* (1994) transformed tobacco plants with the nuclear inclusion b (NIb) gene which contains the GDD sequence. Several lines were found to be resistant to infection, but transformation with a NIb sequence deleted for the GDD motif did not result in resistance in any of the 30 lines which were tested. Whereas resistance was obtained using 3'- or 5'- deleted versions of the NIb gene which still contained the GDD motif.

Attempts have been made to further investigate the necessity of the GDD motif in potyvirus replicase-mediated resistance. Tobacco plants transformed with an intact but mutated form of the NIb gene from plum pox potyvirus (PPV-R, Rankovic isolate) demonstrated some form of protection against PPV infection (Guo & Garcia, 1997). An unusual phenotype was observed in a high percentage of plants of two lines transformed with a version of the NIb gene that contained the sequence VDD instead of GDD. Well-defined patches of dark green tissue emerged on the chlorotic infected leaves. These dark green areas, which grew vigorously and gave the appearance of bubbling on the leaf surface, were shown to contain lower levels of virus than the chlorotic areas. The recovery phenotype is a delayed, very specific, and highly resistant phenotype which was induced by the initial infection. The mutation of the GDD motif to ADD in a full length replicase construct of potato virus X (PVX) transformed into tobacco plants resulted in resistance to several strains of PVX (Longstaff *et al.*, 1993).

<u>Replicase-mediated resistance to macluraviruses and bymoviruses</u>

The production of transgenic barley is still at an early stage, and to date there are no reports of narcissus transformation, so it was prudent to test viral-derived sequences for resistance in tobacco first. BaMMV and BaYMV were so often found as a mixed infection that initially it was assumed that BaMMV was only a strain of BaYMV (Huth & Adams, 1990). Ideally plants that were resistant to both viruses should be generated. Since neither bymovirus infects tobacco, the related macluraviruses were used as a model system for the transformation of tobacco plants with full length and truncated replicase genes. Bymovirus (BaMMV) and macluravirus (MacMV) sequences were used for the transformation and they were challenged with two macluraviruses to test for dual resistance.

<u>Aims</u>

The aim of this experiment was to transform tobacco plants with replicase genes from MacMV and BaMMV. The BaMMV construct contained the complete NIb replicase

gene and the MacMV construct represented a 3'-deleted version of the NIb gene which retained the GDD motif. MacMV was used for the deleted version of the NIb gene, since sequence data for the complete gene were not available. Three different species of tobacco were transformed, *Nicotiana tabacum, N. clevelandii* and *N. benthamiana*. The transformants were challenged with three different viruses to look for resistance to MacMV, NLV or PVY. The two macluraviruses were used to identify dual resistance to two related viruses, a necessity for resistance to the barley-infecting bymoviruses which usually occur together.

Results

Cloning of the replicase gene

MacMV

A cDNA clone of MacMV RNA, which encodes a transcript of part of the NIb gene was generated by RT-PCR (conditions as for Carla-Uni, Chapter 3). Primers (M1 and M2.2) were designed to amplify two-thirds of the NIb gene, including the GDD region. Since the macluraviruses contain potential cleavage sites between the NIb protein and the coat protein, and are thought to be translated as a polyprotein, there is no termination or initiation codon for the NIb protein. However, the NIb sequence does contain a methionine residue upstream of the GDD motif which was suitable for incorporation into the primer sequence as an initiation codon. A termination codon was engineered into the second primer (M2.2). Each primer was constructed so that a restriction site (*BamH* I or *Kpn* I) was created at each end of the amplified fragment (figure 7.1).

RT-PCR on oligo-d(T)Not I-primed first-strand cDNA using the M1 and M2.2 primers generated a product of approx. 600bp, as predicted by the nucleotide sequence (figure 7.2). Control reactions, using each primer alone or both primers and no cDNA produced no fragments as expected. The fragment was recovered from agarose gel, digested with *BamH* I and *Kpn* I and cloned into a Bluescript vector. The insert was completely sequenced in both directions using internal primers and was found to be identical to the original viral cDNA sequence.

A binary plasmid (pNIb) was derived by inserting the partial NIb MacMV gene into a *BamH* I /*Kpn* I digested pROK2 binary plant transformation vector (Baulcome *et al.*, 1986). Before plant transformation, this construct was transferred to *Agrobacterium tumefaciens* strain LBA4404 by tri-parental mating mediated by an *Escherichia coli* strain containing the RK2013 helper plasmid (Chapter 2:section 5). The transconjugants were selected by resistance to kanamycin and rifamycin. They were shown to contain the partial NIb MacMV gene by PCR analysis (figure 7.3 a) (Chapter 2: section 1.6.3).

Figure 7.1: Diagrammatic representation of the position of the two primers used to amplify the partial NIb gene from MacMV by RT-PCR. Primer sequences are shown in bold. Primer M1 was designed to include a *BamH* I restriction enzyme site (underlined) and an initiation codon in frame with the NIb sequence and primer M2.2 contained a termination codon and *Kpn* I restriction enzyme site (underlined).

BamH I

CCGGATCCTACGCTGGTACTCATGATCTCATTTTATTACGCATACGCTGTAAAGACGAGGGATTATACGTTTGATAAA
M I S F Y Y A Y A V K T R D Y T F D K

ATCGATGAAAGATTCGTCTTTGTGTGTAATGGTGATGACAACAAATTCGCGGTGTCACCGGAATTCGTTAAGGAGTTT
I D E R F V F V C N G D D N K F A V S P E F V K E F

GGTGGGTCCTTCACTGACGAGATTGCACAGCTCGGTCTTCATTACGAATTCGATGAGCTCACAACCGACATAACAGCG
G G S F T D E I A Q L G L H Y E F D E L T T D I T A

AATCCTTACATGAGCTTAACAATGATAGACATAGGAGGGCGCATAGGGTTCCAACTCAATCCCGAACGTATTCTAGGA
N P Y M S L T M I D I G G R I G F Q L N P E R I L G

ATCGTTCAGTGGATCAAGAAAGGTGGAATTGTGCACGCCGCACAAGCCGCCTTTGCAGCAATGATCGAATCATTCAAT
I V Q W I K K G G I V H A A Q A A F A A M I E S F N

GATCCAGACCTGTTTTGTGTGATGCACTCGTATTTGGTGTGGCTGCTTGTAACATACCGAAGTGAACTCGTTTATGCC
D P D L F C V M H S Y L V W L L V T Y R S E L V Y A

ATGCACAATGATCTGGTGTCAGTGGTTTACATGGACCCATGCCAAGTCTTTGCCTTGCACTACAACGACAGCGAGGAC
M H N D L V S V V Y M D P C Q V F A L H Y N D S E D

GTAAGGGAGTGGTTTGATGAGGACGATGAGTCAAGTGATGACGAAGACGAAGAGCCAACTCAAGTGCTACAGTAGGTACCGG
V R E W F D E D D E S S D D E D E E P T Q V L Q • *Kpn I*

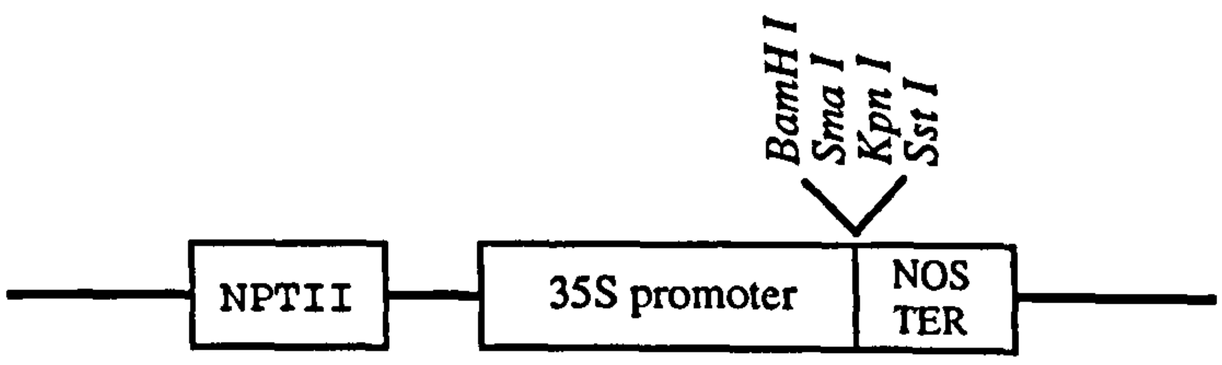


Figure 7.2: Electrophoresis of RT-PCR products from reaction using primers M1 and M2.2 with oligo-d(T)Not I MacMV cDNA. Lane 1: kilobase molecular weight marker ; Lane 2: cDNA and primer M1 alone; Lane 3: cDNA and primer M2.2 alone; Lane 4: cDNA and primer M1 and M2.2.

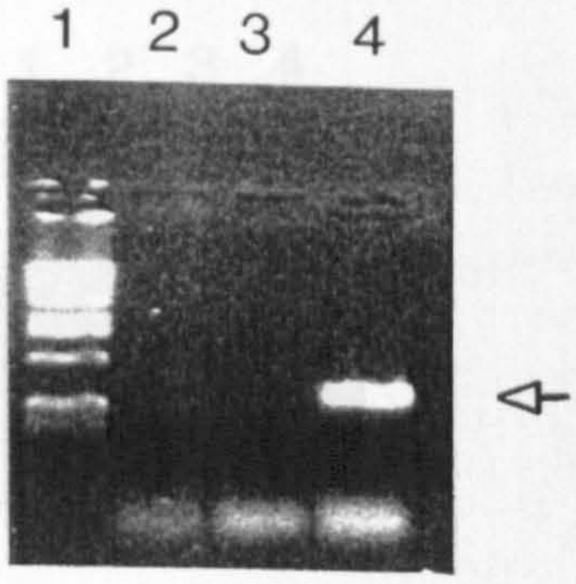
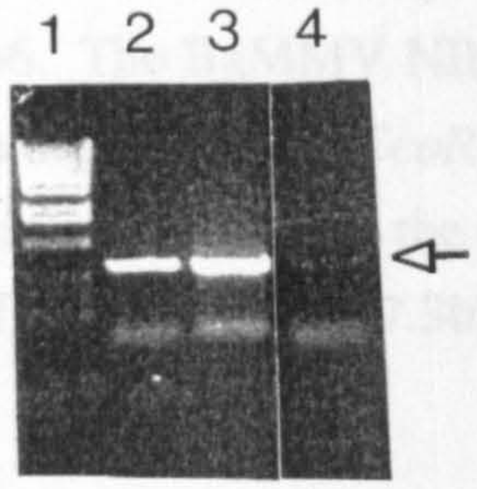
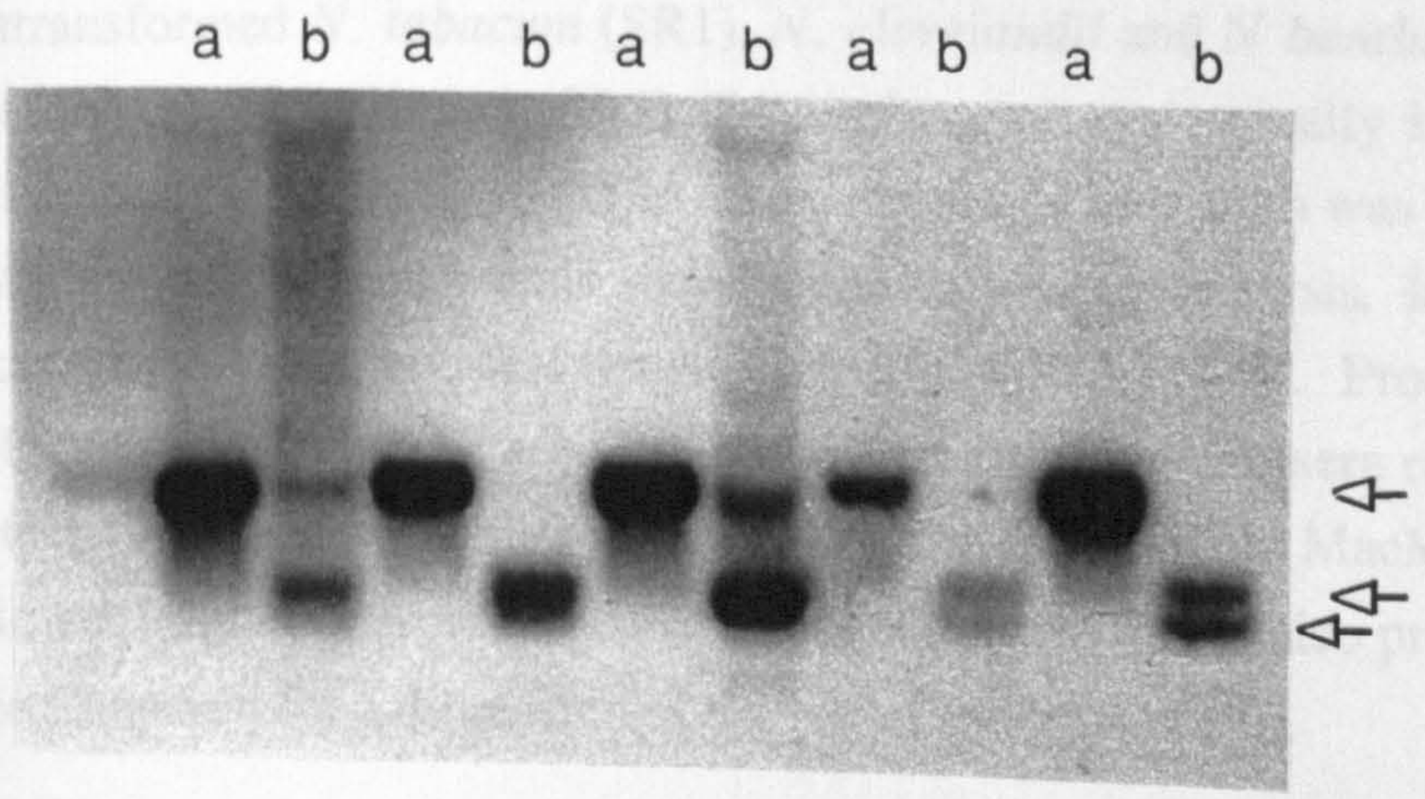


Figure 7.3 a: Electrophoresis of PCR products from reaction using primers M1 and M2.2 with total nucleic acid from *Agrobacterium* transconjugants with pROK2-MacMV. Lane 1: kilobase molecular weight marker ; Lane 2: *Agrobacterium* nucleic acid with primers M1 and M2.2; Lane 3: positive control (pNIb and primers M1 and M2.2); Lane 4: negative control (no DNA, both primers).

Figure 7.3 b: Southern analysis of *Agrobacterium* total nucleic acid digested in alternate lanes with *BamH* I and *Sst* I (a) or *Hind* III and *EcoR* I (b), probed with whole insert from pBMRep6.





BaMMV

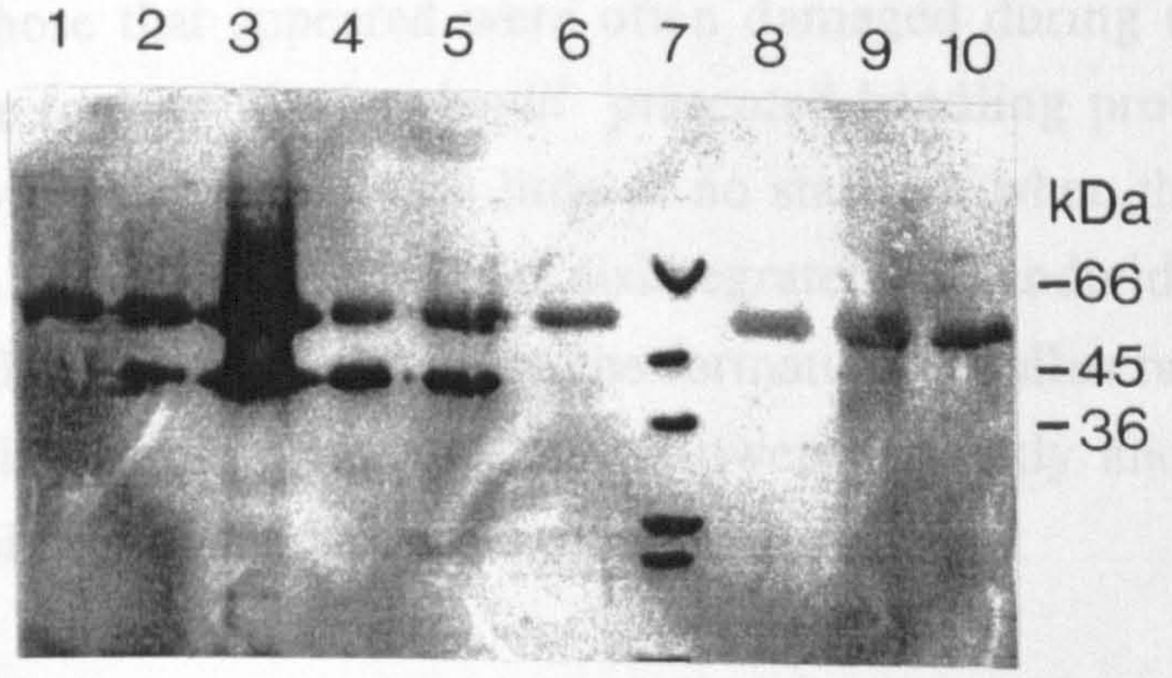
The plasmid pBMRep6 was a kind gift from Dr R. Stratford (PBI). This Bluescript-derived vector contains a 1.6kb insert, comprising the complete cDNA of the NIb gene of BaMMV. This replicase cDNA was produced by RT-PCR using primers designed to place an initiation codon at the 5' end of the NIb gene and a termination codon at the 3' end. The fragment was excised by the enzymes *BamH* I and *Sst* I and cloned into pROK2. This construct (pBMRep6) was transferred to *A. tumefaciens* strain LBA4404 and transconjugants were selected by the same method as for the MacMV construct. They were shown to contain the BaMMV NIb gene by Southern analysis (figure 7.3 b). Total nucleic acid from several *Agrobacterium* transconjugants was digested with *BamH* I and *Sst* I or *Hind* III and *EcoR* I, and probed using a the original insert from Bluescript plasmid pBMRep6. The BaMMV NIb gene contained two internal *EcoR* I restriction sites and therefore digestion with *EcoR* I produced three fragments (two 1kb and one 500bp fragment) which hybridised to the probe. The *BamH* I and *Sst* I excised the complete insert, a 1.5kb fragment (figure 7.3b).

Virus inoculations

Prior to plant transformations, the three tobacco species were assessed for their susceptibility to MacMV infection (Chapter 2: section 6.1.). Figure 7.4 shows a western analysis at 14 days post inoculation (dpi) of systemic and inoculated leaf material from untransformed *N. tabacum* (SR1), *N. clevelandii* and *N benthamiana* inoculated with MacMV. *N. tabacum* (SR1) did not become systemically infected unlike *N. clevelandii* and *N. benthamiana*. The same amount of inoculum was used in each case and equal amounts of total protein were loaded for western analysis. It can be seen that the *N. clevelandii* was the most heavily infected with MacMV. Proteins of two different molecular weight can be seen on the western using the antisera raised to MacMV coat protein. The faster migrating protein is that corresponding to MacMV coat protein and the slower migrating protein species is a contaminant that is also present in healthy tissue (see Chapter 4 for a discussion of the use of this antiserum).

The inoculum used was from the same source as that used throughout the following experiments. Frozen MacMV-infected *N. clevelandii* leaf material which had been confirmed to be systemically infected by western analysis prior to snap freezing in liquid nitrogen for long term storage at -80°C. When used as inoculum, a single leaf of the frozen material was ground in a pestle and mortar with 1ml of sterile distilled water. A drop of 50µl was used for inoculation of a single leaf. NLV-infected *N. clevelandii* and PVY^N-infected *N. tabacum* material was used in the same manner.

Figure 7.4: Western blot of systemic and inoculated leaf tissue from three different species of tobacco infected with MacMV, 14 days post inoculation, probed with antisera raised to MacMV coat protein. Lane 1: *N. tabacum* (SR1) systemic leaf; Lane 2: *N. tabacum* (SR1) inoculated leaf; Lane 3: *N. clevelandii* systemic leaf; Lane 4: *N. clevelandii* inoculated leaf; Lane 5: *N. benthamiana* systemic leaf; Lane 6: *N. benthamiana* inoculated leaf; Lane 7: SDS-7 protein molecular weight marker; Lane 8: *N. tabacum* (SR1) healthy control; Lane 9: *N. clevelandii* healthy control; Lane 10: *N. benthamiana* healthy control. The larger of the two proteins, present in all samples is a contaminant (see Chapter 4).



Plant transformations

Three different tobacco species (*N. clevelandii* , *N. tabacum* and *N. benthamiana*) were transformed using the MacMV partial and the BaMMV complete NIb constructs.

Nicotiana clevelandii

Nicotiana clevelandii was used for transformation since it was known that both NLV and MacMV readily infected this species by mechanical inoculation. However, no transformants containing the BaMMV construct were obtained and only two lines of transformants containing MacMV constructs were recovered.

So few lines were obtained from *N. clevelandii* due to experimental problems with the transformation protocol. A very low percentage of shoots were observed on leaf excisions. Those that appeared were often damaged during transfer to pots. The stunted rosette form of *N. clevelandii* presented handling problems. Transformed shoots were often very small with little or no stalk, so when they were excised and transferred to pots, the rosette often disintegrated and individual leaves had to be inserted into the medium. This led to the formation of callus rather than roots. Any tiny plants which were recovered often flowered quickly and formed poor roots, decreasing their rate of survival when transferred to soil.

Only two lines were obtained from the transformation of the partial NIb MacMV construct. Seeds were collected from line 3 and grown on kanamycin selection. These selfed progeny were grown on kanamycin selective medium and tested for the presence of the MacMV construct by plant PCR. Total DNA extracted from each plant was tested in an RT-PCR reaction using the original M1 and M2.2 primers used to create the MacMV NIb construct (figure 7.5). An amplified fragment of 600bp was obtained for each of the seven seedlings from line 3, confirming the presence of the MacMV NIb partial construct in these transformants (only five of the seven plants are shown on figure 7.5).

MacMV *N. clevelandii* challenged with MacMV

These seven selfed progeny from line 3 were challenged with MacMV inoculum. Systemic and inoculated leaf samples were taken at 0, 7 and 14 dpi. Six of the seven plants became systemically infected by 14 dpi (figure 7.6) as is shown by western analysis of the systemic leaf samples taken at this time point. Only plant 3.3 showed a delay in symptoms and the systemic leaf was free from virus at 14 dpi. This result was disregarded, since the delay in virus accumulation in only one plant of seven from a single line was not thought to be significant. The untransformed inoculated control was

Figure 7.5 : PCR analysis of total nucleic acid extracted from transformed plants to confirm the presence of an NIb construct. One line (3) of *N. clevelandii*, second generation transformed with partial NIb MacMV construct. PCR using primers M1 and M2.2. This was run as part of the gel shown in figure 7.7 and shares the same controls which are included on that figure.

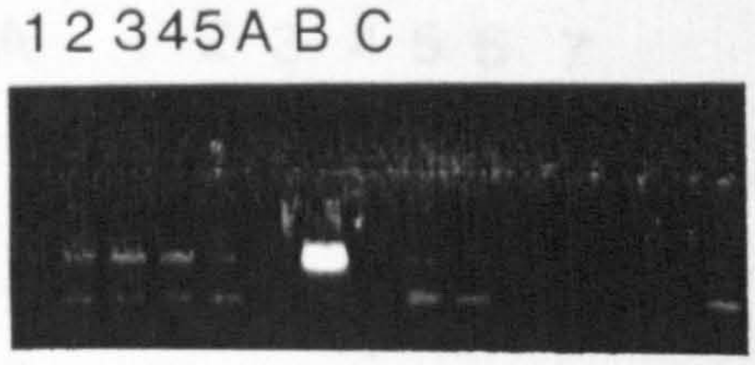


Figure 7.6: Second generation *N. clevelandii* transformed line 3 plants (MacMV partial NIb construct) 14 days post inoculation with MacMV. Western blot of systemic leaf protein samples probed with antisera raised to MacMV coat protein. Lane A: untransformed inoculated control; Lane B healthy uninoculated control. Lanes 1 to 7 individual plants from line 3.

systemically infected by 14 dpi, and the healthy untransformed, uninoculated control remained virus free.

Nicotiana tabacum (SR1)

Given the technical problems encountered with *N. clevelandii*, a tobacco species known to be easily transformed and easily handled was used to improve technique. *Nicotiana tabacum* (SR1) was used as this test species. MacMV and NLV do not systemically infect *N. tabacum* and as an alternative a potyvirus was used for challenging the transformants obtained. Using the MacMV construct, 6 lines of primary transformed *N. tabacum* were obtained. Each line was selfed and the seed collected. Progeny from the 6 lines containing the MacMV partial NIb construct were grown from seed on kanamycin selection. All were tested by PCR for the presence of the construct. Figure 7.7 shows the results of this plant PCR where all transformants tested produced a fragment of the expected size for the NIb construct (600bp) when amplified with the primers M1 and M2.2.

MacMV *N. tabacum* challenged with PVY^N

Table 7.1 shows the numbers of selfed progeny from each line which were inoculated with potato virus Y strain N (PVY^N). This virus was chosen since neither BaMMV, MacMV or NLV infects *N. tabacum* systemically and PVY^N has some degree of sequence homology around the GDD motif of the macluravirus replicase. It is also possible that the replicase constructs may interfere with replication of other viruses by binding to host factors involved in replication. The construct was designed with the intention of producing dual resistance to two related viruses, MacMV and NLV, and PVY^N was used to look for a more broad spectrum resistance.

Samples were collected from systemic and inoculated leaves at 0, 8, 12 and 19 dpi. All plants tested became systemically infected by 19dpi. However, there was a slight delay in the accumulation of virus in the systemic leaf of several plants. Western analysis of systemic leaves at 12 dpi (figure 7.8) demonstrates that in only line 6 were all plants systemically infected (3/3). Three lines (2, 5 and 8) had only one plant systemically infected (1/3, 1/4, 1/3, respectively) at 12 dpi. The remaining two lines showed only one plant with delayed systemic accumulation. It is possible that the leaves were infected at a level too low to be detected by western analysis at 12 dpi, since by the next sampling day (19 dpi) all plants from every line were systemically infected. Although this delay in systemic virus accumulations was noted in all 6 lines to varying degrees, no single plant maintained this delay beyond 19 dpi and therefore none were considered resistant.

Table 7.1: *N. tabacum* (SR1) transformed lines containing the partial NIb MacMV construct inoculated with PVY^N

	MacMV partial NIb					
Line	2	3	5	6	7	8
PVY_N	4	4	4	3	4	3
control	1	1	1	1	1	1

Figure 7.7: PCR analysis of total nucleic acid extracted from transformed plants to confirm the presence of an NIb construct. Six lines of *N. tabacum* (SR1), selfed progeny from primary transformants containing the partial NIb MacMV construct. PCR using primers M1 and M2.2. Each plant line is labelled at top of gel and four plants from each line are labelled below.

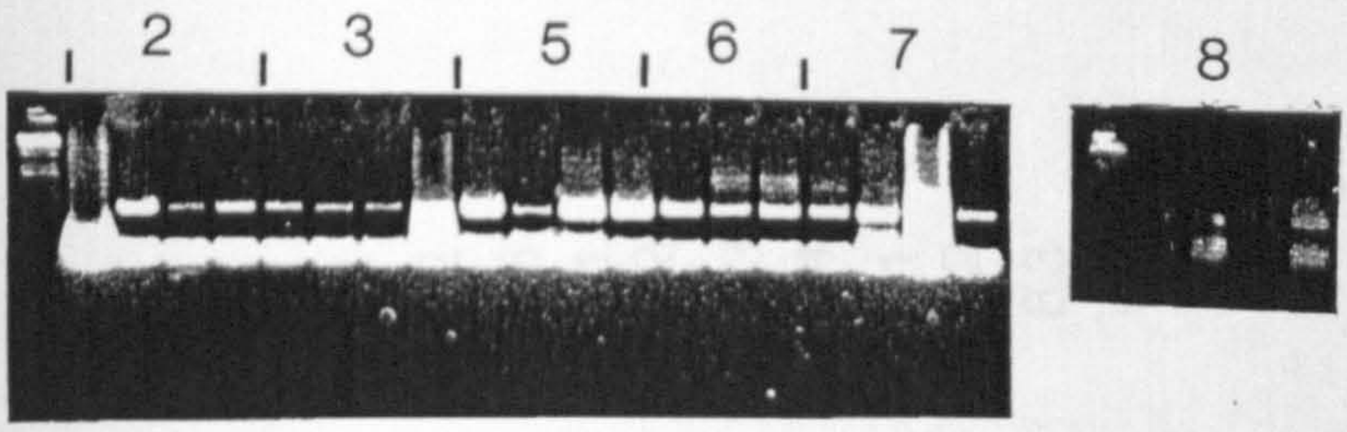
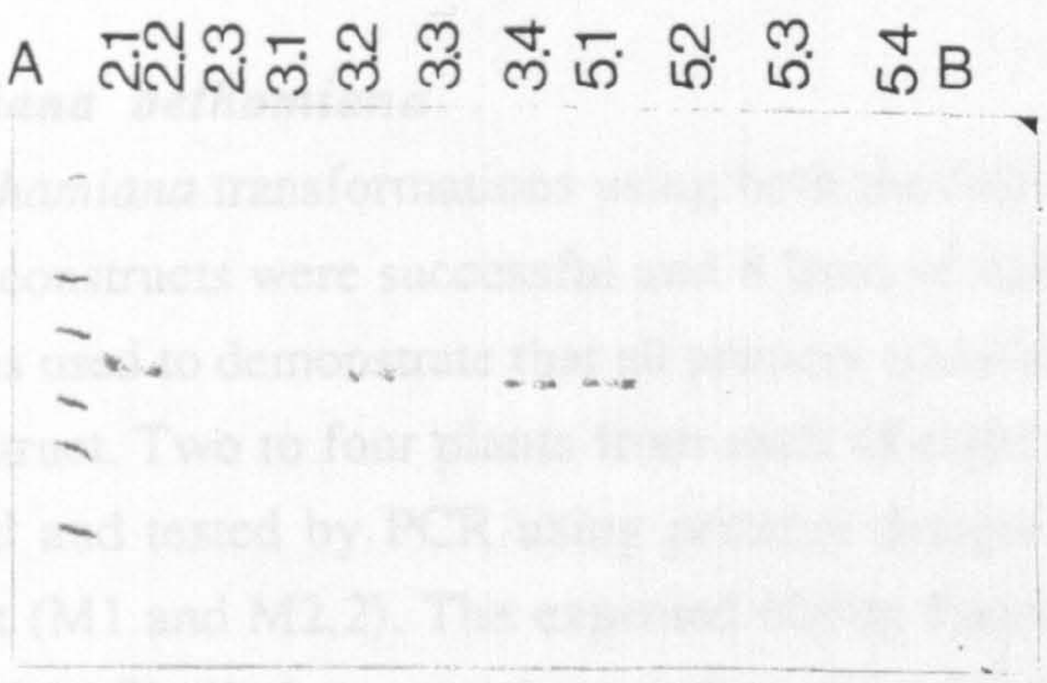
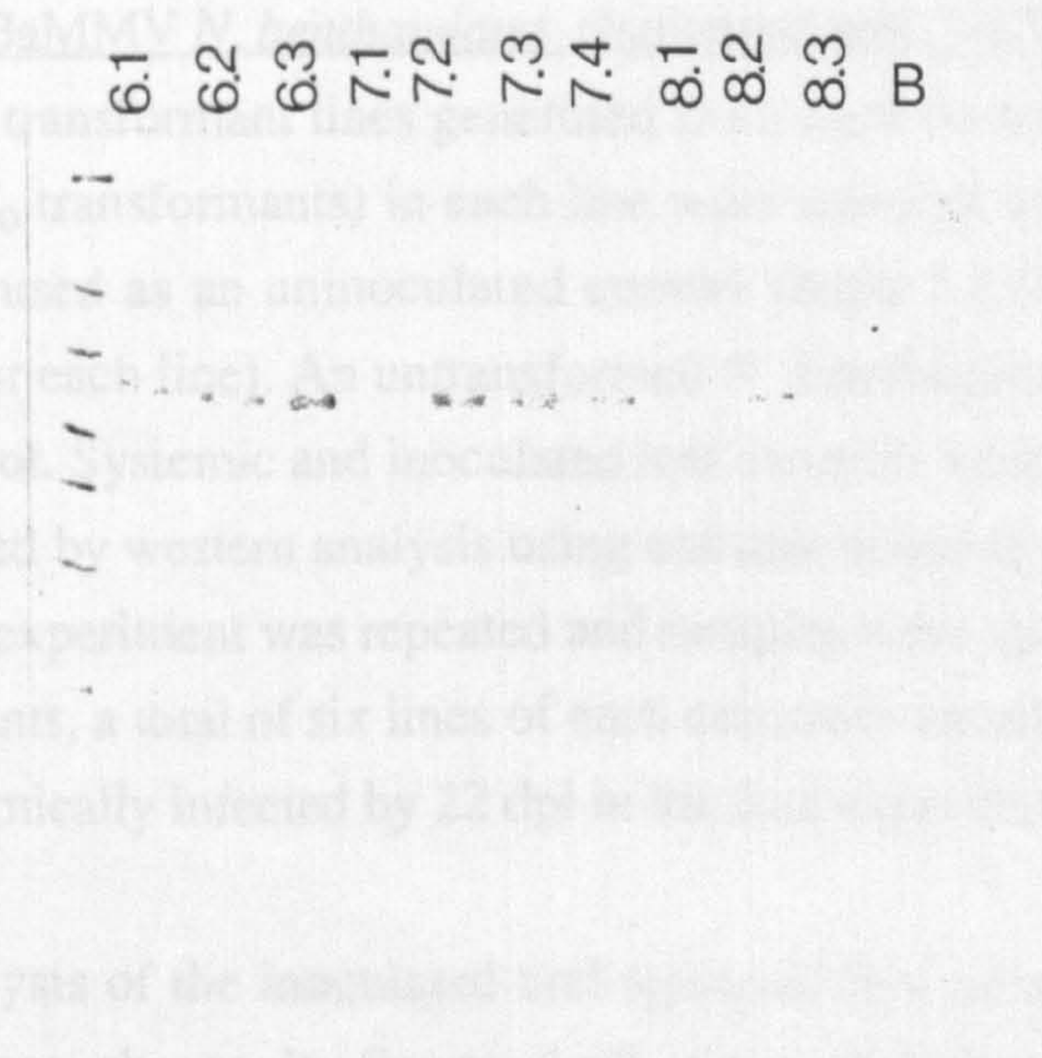


Figure 7.8: *N. tabacum* selfed progeny of 6 different lines of transformed plants (MacMV partial NIb construct) 12 dpi with PVY^N. Western blot of systemic leaf protein samples probed with antisera raised to PVY^N coat protein. Lane A: untransformed inoculated control; Lane B healthy uninoculated control. The line number is indicated above the lanes which contain samples from individual plants.





Attempts were made to transform *N. tabacum* plants with the BaMMV full length NIb construct. However these were unsuccessful due the presence of fungal infection in the *Agrobacterium* transconjugant. When clean *Agrobacterium* containing the BaMMV construct was finally obtained the time limits of the project dictated that this experiment could not be repeated.

Nicotiana bethamiana

Nicotiana bethamiana transformations using both the full length BaMMV and partial MacMV NIb constructs were successful and 8 lines of each construct were obtained. Plant PCR was used to demonstrate that all primary transformants tested contained the MacMV construct. Two to four plants from each of eight primary transformant lines were obtained and tested by PCR using primers designed to sequences within the original insert (M1 and M2.2). The expected 600bp fragment was obtained in every case (figure 7.9a). Similarly, to test the transformants containing the BaMMV construct the primers BAM-ATG and BAM-15 (supplied by Dr R. Stratford, PBI) were used. These amplified a 1.2kb fragment (figure 7.9b) from primary transformants, as predicted by the nucleotide sequence of pBMRep6 .

MacMV and BaMMV *N. benthamiana* challenged with NLV and MacMV

Four primary transformant lines generated from each construct were tested. In general, two plants (R_0 transformants) in each line were infected with MacMV, two with NLV and one was used as an uninoculated control (table 7.2 shows the exact numbers of plants used for each line). An untransformed *N. benthamiana* plant was inoculated as a positive control. Systemic and inoculated leaf samples were taken at 0, 7, 10 and 22 dpi and were tested by western analysis using antisera raised to either NLV or MacMV coat protein. The experiment was repeated and samples were taken at 7 and 13 dpi. Over the two experiments, a total of six lines of each construct were tested. All inoculated plants became systemically infected by 22 dpi in the first experiment and 13 dpi in the second.

Western analysis of the inoculated and systemic leaf samples at 7 dpi for the second experiment are shown in figure 7.10. Several individual plants showed virus accumulation in the systemic leaf, but not in the inoculated leaf (B1.1, B34.1, B61.1) when challenged with NLV or when challenged with MacMV (B34.5). It is more likely that this was due poor quality of plant protein samples available for western analysis rather than any effect of the construct, since these inoculated leaves suffered heavy necrosis. Only two individual plants (M2.1 and B2.3) from different lines show no virus accumulation in the systemic or inoculated leaves at 7 dpi (figure 7.10). However, it should be noted that samples from both of these plants were analysed in the first two wells of the gel, and hence may have not transferred as efficiently as other samples.

Table 7.2: *N benthamiana* transformed lines containing an NIb construct inoculated with MacMV and NLV, showing the number of plants of each line used. Each table represents a different experiment carried out at different times.

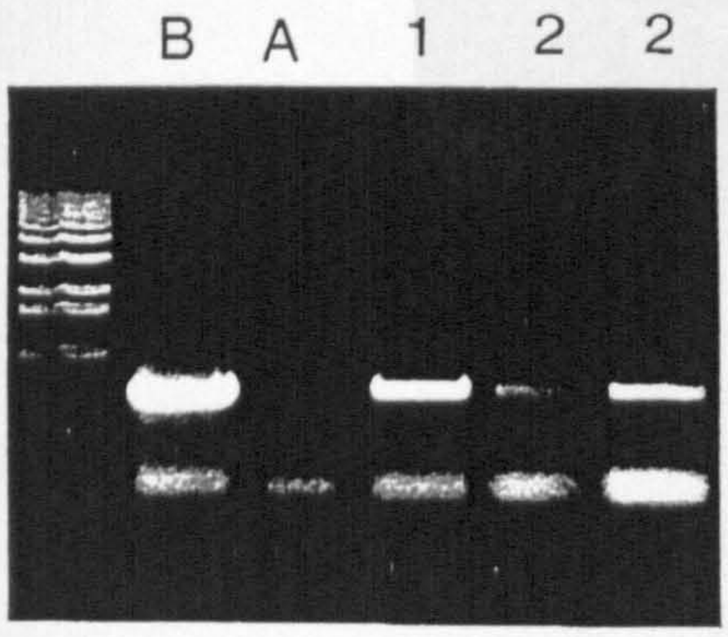
	MacMV partial NIb				BaMMV complete NIb				
Line	1	2	3	9	1	3	7	34	61
MacMV inoculum	2	2	2	2	2	2	2	2	2
NLV inoculum	2	2	2	2	2	2	1	2	2
transformed control	1	1	1	1	1	1	1	1	1

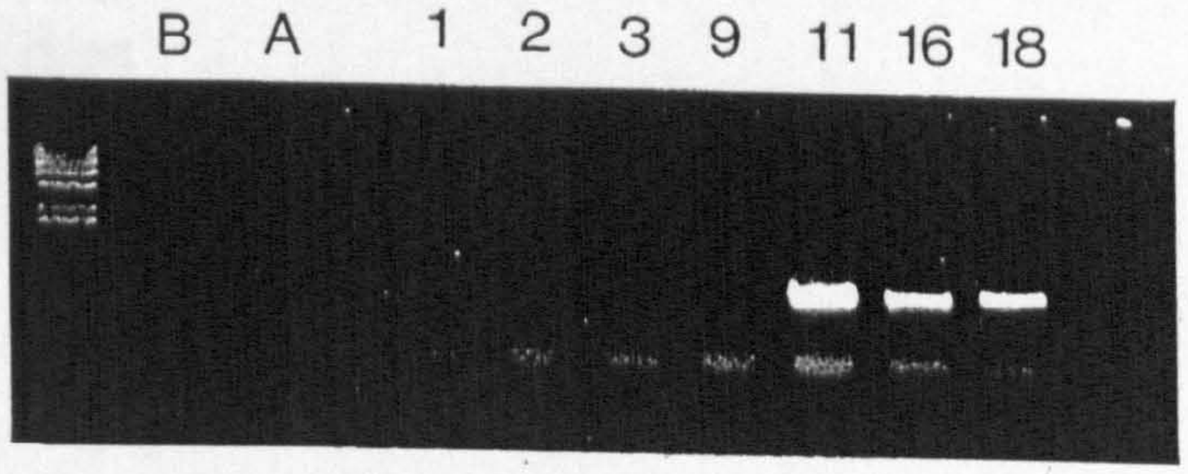
	MacMV partial NIb				BaMMV complete NIb			
Line	2	9	10	11	1	2	34	61
MacMV inoculum	1	1	1	1	3	4	3	1
NLV inoculum	1	1	2	2	2	3	3	1
control	1	1	1	1	1	1	1	1

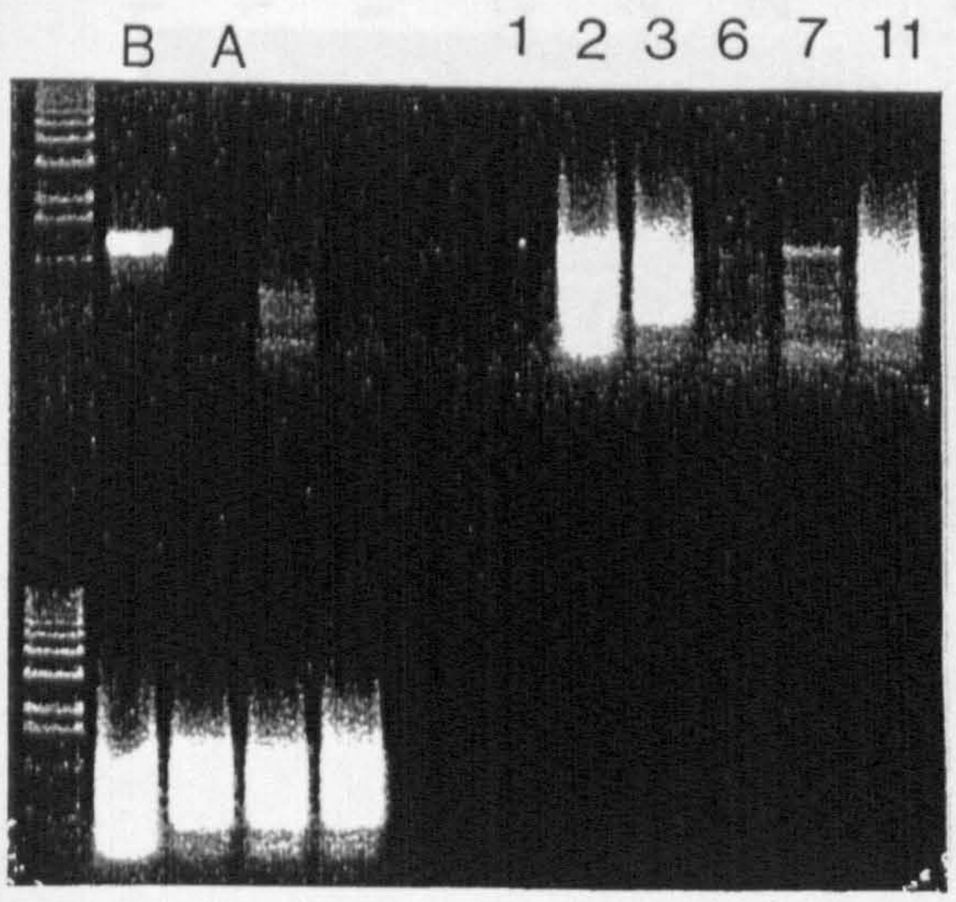
Figure 7.9: PCR analysis of total nucleic acid extracted from transformed plants to confirm the presence of an NIb construct.

a) Individual primary transformants of *N. benthamiana* containing the partial NIb MacMV construct. PCR using primers M1 and M2.2. Lanes are labelled according to transformants line number. Lane A: negative control (no DNA, both primers); Lane B: positive control (pNIbM, both primers); Lane C: non-transformed control.

b) Individual primary transformants of *N. benthamiana* containing the complete NIb BaMMV construct. PCR using primers BAM-ATG and BAM-15. Lanes are labelled according to transformants line number. Lane A: negative control (no DNA, both primers); Lane B: positive control (pBMRep6, both primers); Lane C: non-transformed control, both primers.







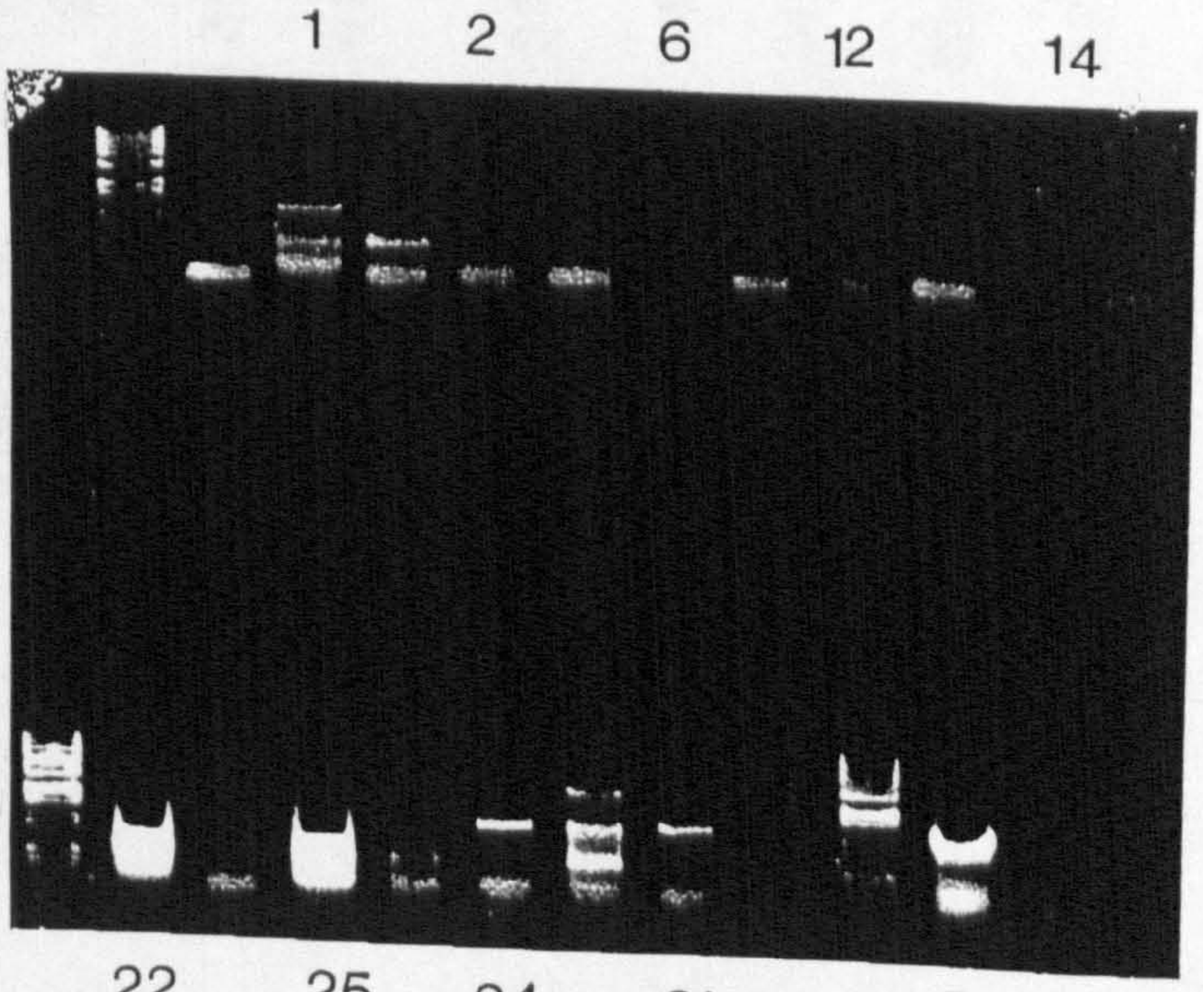
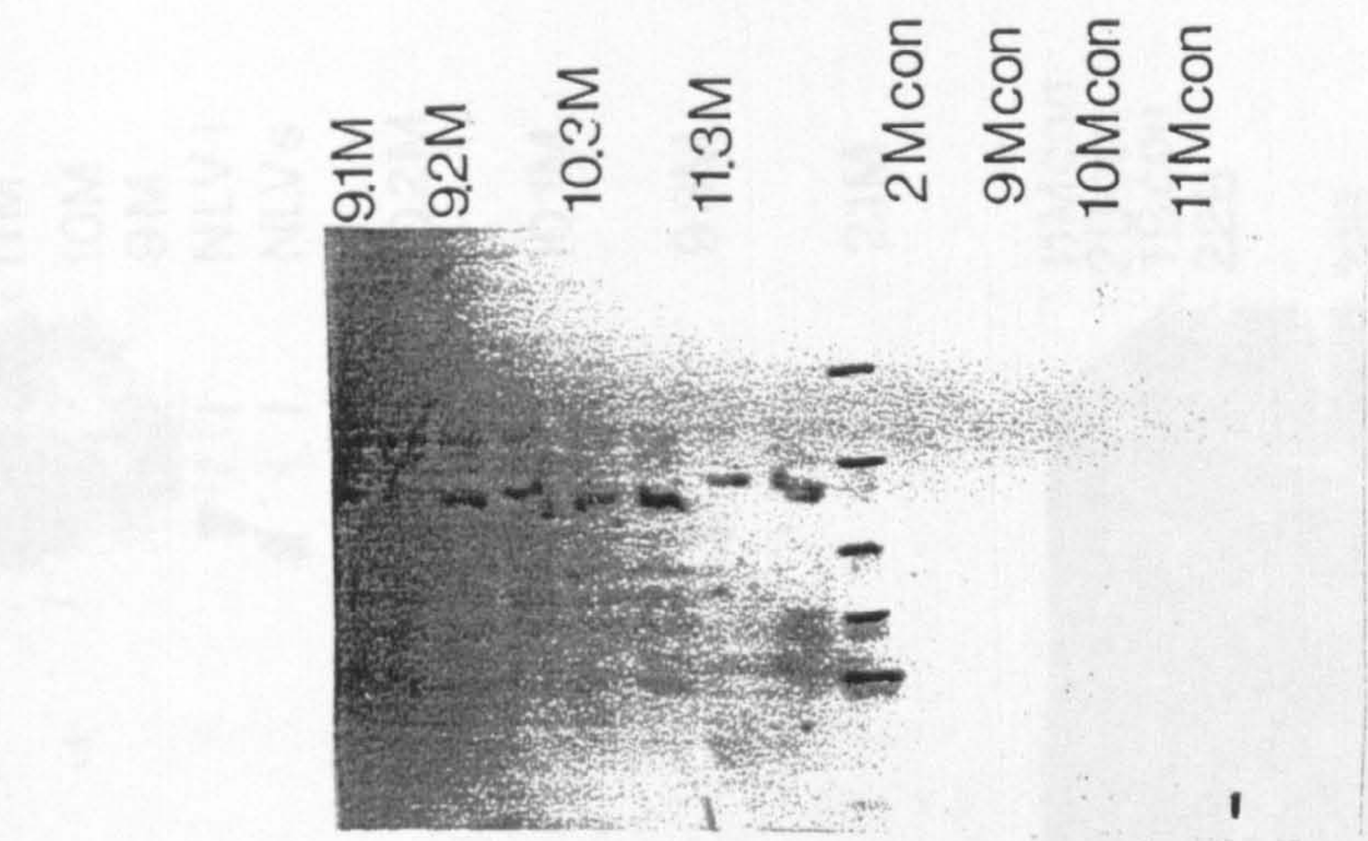
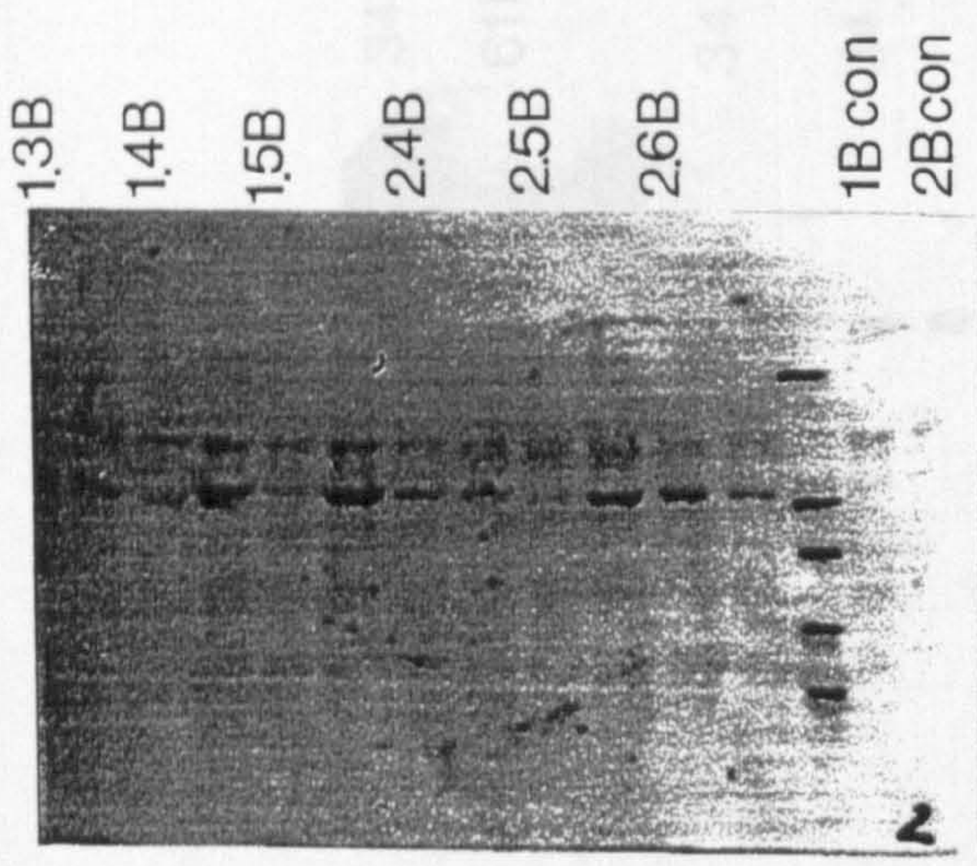


Figure 7.10 **a)**: Western blots of primary transformants of *N.benthamiana* containing either MacMV partial NIb or BaMMV complete NIb construct, 7 days post inoculation with either MacMV or NLV. **a)** Systemic and inoculated leaf protein samples probed with antisera raised to MacMV coat protein. Part b overleaf.





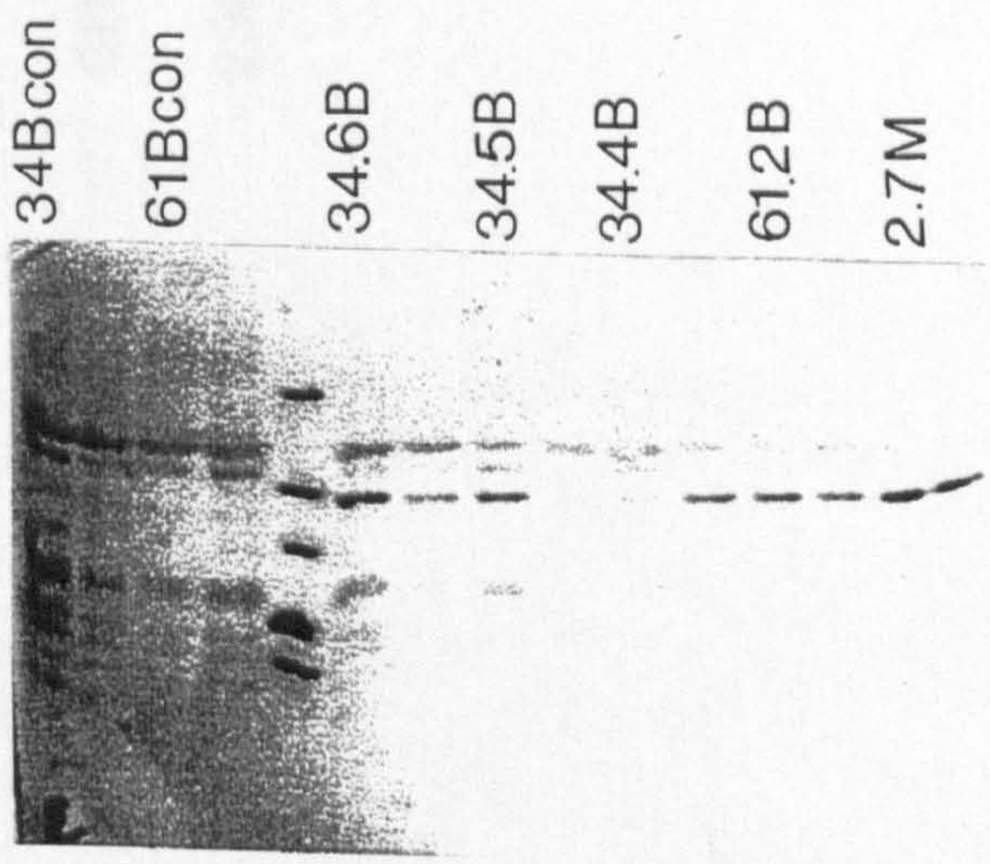


Figure 7.10 **b)**: Western blots of primary transformants of *N.benthamiana* containing either MacMV partial NIb or BaMMV complete NIb construct, 7 days post inoculation with either MacMV or NLV. **b)** Systemic and inoculated leaf protein samples probed with antisera raised to NLV coat protein .

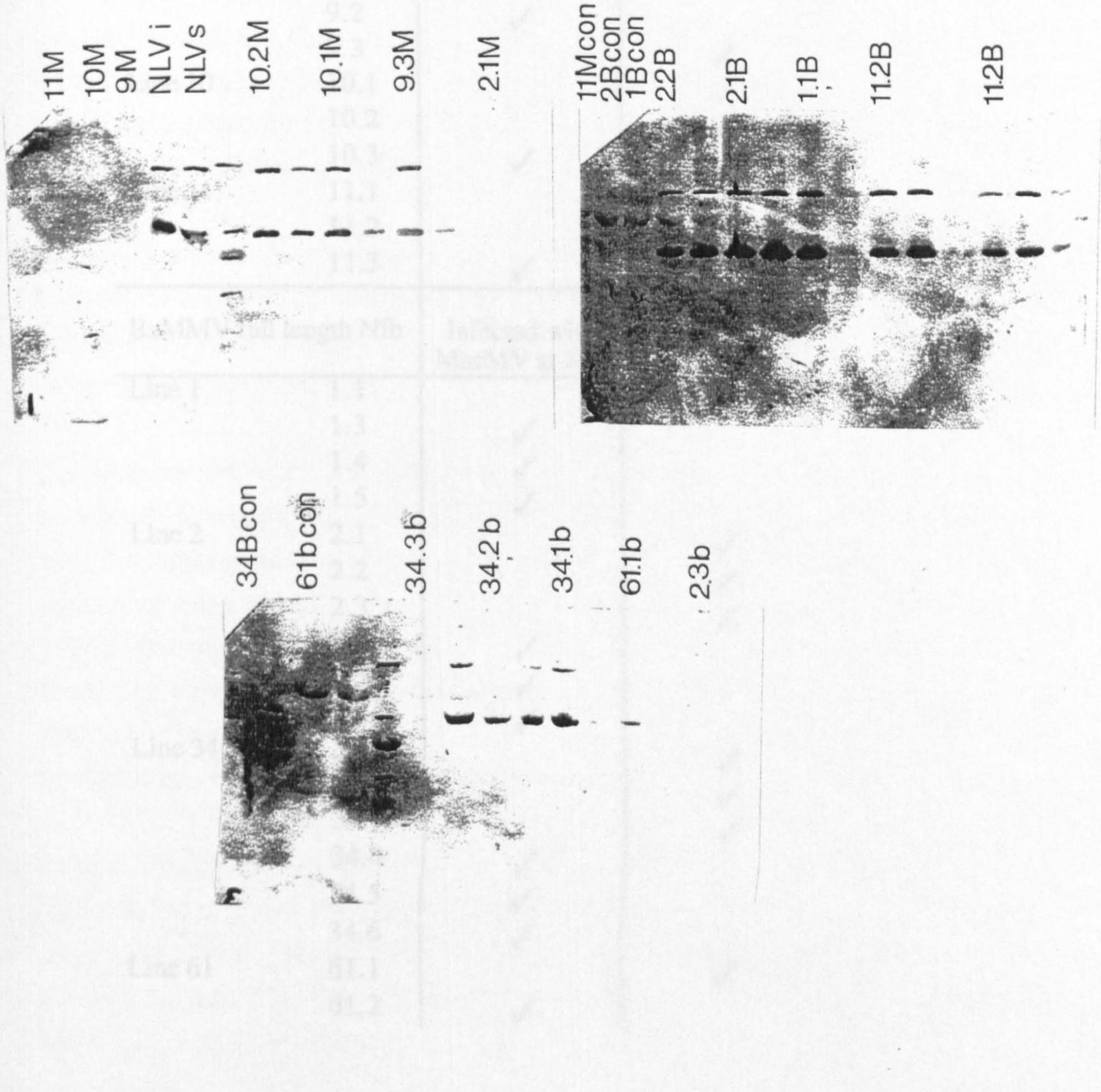


Table 7.3: Summary of results shown in figure 7.10 a, b

MacMV partial NIb construct		Infected with MacMV at 7dpi	Infected with NLV at 7dpi
Line 2	2.1		×
	2.7	✓	
Line 9	9.1	✓	
	9.2	✓	
	9.3		✓
Line 10	10.1		✓
	10.2		✓
	10.3	✓	
Line 11	11.1		✓
	11.2		✓
	11.3	✓	

BaMMV full length NIb		Infected with MacMV at 7dpi	Infected with NLV at 7dpi
Line 1	1.1		✓
	1.3	✓	
	1.4	✓	
	1.5	✓	
Line 2	2.1		✓
	2.2		✓
	2.3		×
	2.4	✓	
	2.5	✓	
	2.6	✓	
Line 34	34.1		✓
	34.2		✓
	34.3		✓
	34.4	✓	
	34.5	✓	
	34.6	✓	
Line 61	61.1		✓
	61.2	✓	

The uninoculated controls for each line remained virus free. It should be noted that all plants from every line became systemically infected by the end of each experiment.

Discussion

Three different types of tobacco plant were transformed by *Agrobacterium* containing two different viral replicase constructs. *Agrobacterium* transconjugants and either primary transformed plants or selfed progeny were successfully tested for the presence of the construct by plant PCR or Southern analysis. These transformed lines were subject to inoculation with NLV, MacMV or PVY^N in order to assess the possibility that they may demonstrate single and/or dual resistance.

No single resistant plant was found in the several screens employed. The screen of *N. clevelandii* transformed using MacMV partial NIb was very limited, using only seven selfed progeny from one plant line. The delay in virus accumulation in the systemic leaf of one plant is not significant.

The technical difficulties experienced with *N. clevelandii* transformation may have been additionally hampered by deleterious effects of the constructs used. It is interesting to note that other experimenters have reported difficulty in regenerating transformants containing deleted and full length versions of the potyvirus NIb sequence (Audy *et al.*, 1994). Detrimental effects due to the presence of the NIb gene sequence and its C-terminally truncated derivative were observed. These plants grew significantly slower than non-transformed tobacco and calli containing the C-terminally truncated NIb transgene grew so slowly that they had to be rescued on non-selective medium (Audy *et al.*, 1994). It is possible that a similar effect occurred in the presence of the MacMV deleted NIb sequence in *N. clevelandii*. However, controls using empty pROK2 were not used, and thus comparisons to transformed plants not containing the NIb sequence could not be made. These problems were not experienced with other tobacco species, although the possibility remains that the NIb sequence has some effect on *N. clevelandii* that is unique to the species.

The test species, *N. tabacum*, transformed with MacMV partial NIb construct yielded the most interesting results. At 12 dpi, only one line had all plants systemically infected, with the other five lines showing varying degrees of delayed accumulation of virus in the systemic leaf. However, since all plants were systemically infected by 19 dpi, none can be said to be completely resistant. Day 12 may have represented a window of low levels of virus in the systemic leaf beyond the limits of detection by western analysis.

The screen using transformed *N. benthamiana* lines was the largest, but it still failed to detect any resistant individuals. Although two individual plants did not contain viral coat protein in either inoculated or systemic leaf samples, it is more likely that this was due to experimental error rather than resistance. Again, all plants were systemically infected by 13 dpi, and the same plants tested in a duplicate experiment became infected on a similar time scale (results not presented).

It is perhaps not surprising that transformation with bymovirus or macluravirus replicase sequences did not yield any completely resistant lines in this case. Literature reviews show that highly resistant lines obtained by the insertion of non-structural viral sequences are rare. In the first experiment to use a defective replicase construct, 18 independently transformed lines were obtained, and 7 of those showed a resistant phenotype (Anderson *et al., 1992*). Similarly, the first potyviral NIb sequence used to confer replicase-mediated resistance generated 13 lines, testing 325 individual plants to isolate four highly resistant lines (Audy *et al.*, 1994). The experiments reported here used a comparatively small number of lines (6 for each construct) and individual plants due to the constraints of time and growth space. The unpredictable insertion of transgene constructs into the plant genome causing different positional and copy number effects could generate a whole spectrum of variation of temporal and quantitative expression in the regenerated plants. Hence, if a larger number of transgenic lines were tested, a resistant phenotype may have been observed.

Another factor in the search for resistant lines, is the level of resistance required. In this case, relatively high concentrations of virus inoculum was used as a stringent test for resistance. This experiment was to be a model system, and therefore resistance needed to be of a high level in order to increase the likelihood of effective transfer to a graminaecious host or even eventually to narcissus. In this case, the level of virus concentration may have been prohibitively high and resistance may have been observed at lower levels of inoculum. Only one concentration was used to look for absolute resistance, but as little is known about the amount of bymovirus needed to cause a systemic infection in the field, the concentration used should perhaps be re-examined.

Dual Resistance

In these experiments, dual resistance to NLV and MacMV was not achieved by the presence of a partial NIb MacMV construct, or a complete BaMMV NIb construct. Partial resistance to one virus has been achieved by the transformation of tobacco plants with the sequences of another related viruses. Audy *et al.* (1994) generated transgenic plants containing the full length PVY^{O} NIb sequence which were highly resistant to infection by PVY^{O}. Three plants from each of six lines were inoculated with the

tobacco vein necrosis strain, PVY^N, and one line showed ELISA readings that were 2 to 6 times lower that in infected control plants (Audy *et al.*, 1994) suggesting that viral replication had been suppressed. However, due to the small number of plants tested, the authors suggested that this partial resistance may be more strongly manifested in other untested lines.

PVY^O and PVY^N are over 90% homologous at the amino acid level in the NIb region and only partial dual resistance was achieved. MacMV and NLV are only 76% homologous over the part of the NIb used for the MacMV construct, and this lower level of conservation may affect the efficacy of dual resistance, but this will depend on the mechanism of resistance. If only a few important conserved motifs are involved in the mechanism, such as the GDD RNA-dependent RNA polymerase motif, then the sequence identity between NLV and MacMV around this motif is 88% (nucleotide) and they are identical at the amino acid level.

Mechanism of replicase-mediated resistance

The mechanism of replicase-mediated resistance is far from clear. Viruses have a relatively simple infection cycle and therefore there are only a few windows of opportunity for the resistance mechanism to function. Resistance mechanisms could operate at one of two levels, RNA or Protein. There is much conflicting evidence for both cases and several models have been suggested.

There is evidence that the mechanism is protein mediated. Carr *et al.* (1992) used protoplasts in a transient assay system to assess the effect on viral replication of transfected constructs, after showing that a TMV 54kDa construct conferred resistance to transgenic tobacco. When non-transformed protoplasts were transfected simultaneously with a wild-type 54kDa TMV replicase construct and TMV RNA, viral replication was reduced. This result was maintained when a construct missing the first 14 N-terminal amino acids was used. However, when a mutated replicase which was prematurely terminated and produced a product only 20% of the full length protein was used, viral replication levels were not affected. Similarly, a frameshift mutation causing a premature termination in the replicase construct of pea early browning virus (PEBV) abolished resistance in transgenic plants (MacFarlane & Davies, 1992).

It is difficult to determine whether a mutated form of a replicase is essential for resistance. Resistance has been achieved with several full length intact replicases, for example TMV (Golemboski *et al.*, 1990), PEBV (MacFarlane & Davies, 1992), PVX (Braun & Hemenway, 1992) and PVY (Audy *et al.*, 1994). However, it is not certain whether the TMV and PEBV replicase genes are actually expressed *in vivo* as both are

read-through components of other larger products that are detected during viral replication. Due to the translation strategy of the potyvirus genome, a methionine initiation and termination codon need to be artificially added, thus the replicase is not truly authentic. Although the PVX full length construct was generated from an infectious clone, thereby confirming its authenticity and protecting from the possibility of PCR errors or mutations during cloning, a deleted replicase has also been shown to confer resistance (Braun & Hemenway, 1992). Other deleted replicases have been shown to confer resistance, such as PVY (Audy *et al.*, 1994) and CMV (Anderson *et al.*, 1992).

It has been suggested that deleted or mutated forms of a viral replicase could confer resistance by a "dominant negative mutation" (Herskowitz, 1987). A defective replicase would disrupt the function of the normal product, blocking its action or by competing for it's active site. This model could also be applied to the TMV and PEBV full length replicases if they are considered to be a truncated defective version of the 183kDa replicase (Carr *et al.*, 1992).

An RNA mechanism has been suggested to explain the occurrence of variation in the level of resistance conferred by a RNA-dependent RNA polymerase construct from plum pox virus (PPV) (Mueller *et al.*, 1995) and the recovery phenotype observed when a mutated form (VDD) of the GDD motif is used to transform tobacco (Guo & Garcia, 1997). This model involves the use of homology-dependent gene silencing to account for the observation that resistance occurred in a transgene dosage-dependent manner using an untranslatable replicase construct. An "unidentified genomic feature" is proposed to influence a cytoplasmic mechanism that degrades RNA with sequence homology to the silencing transgene.

<u>Conclusions</u>

Given the uncertainty that surrounds the mechanism of replicase-mediated resistance and the normal function of the replicase genes *in vivo*, it is difficult to predict the replicase sequences which will confer resistance. The experiments using a full length BaMMV NIb gene and a partial MacMV NIb gene have not generated resistant plants. However, only small numbers of plants were tested with a single inoculum dose. Other more successful replicase-mediated resistance experiments have employed large screens to find only a few resistance lines, and if the screen were expanded it may have been more successful.

The possibility that dual resistance to MacMV and NLV, or BaMMV and BaYMV could be created in one transgenic plant cannot be ruled out, as there is little evidence for or against a resistance mechanism which operates on a strict sequence homology basis.

Future work

There are obviously many areas which could be further explored on this subject. This project only allowed for two constructs to be tested and was limited by lack of success in *N. clevelandii* transformations. Future work should include the improvement of this technique, trying different methods and growing mediums, since *N. clevelandii* is the only tobacco species to have reliable symptom expression for NLV and MacMV. The screen of transformant lines could be increased.

Different construct designs could be used, a complete MacMV NIb gene would require the further sequencing of the MacMV genome. Mutated forms of the complete NIb gene, or 3' as well as 5' deletions could be used. Different levels of virus concentration could also be incorporated into the experiment.

Results from macluraviruses could therefore influence the design of constructs for resistance against bymoviruses in barley.

Chapter 8

Discussion

The central aim of this thesis was to elucidate the taxonomic position of two plant viruses, narcissus latent virus (NLV) and Maclura mosaic virus (MacMV) whose defined characters did not wholly fit those expected of any existing plant virus taxon. The results of traditional techniques to determine coat protein size, particle shape and structure, mode of vector transmission, cytology and serology had proved unable to satisfactorily classify NLV and MacMV. Therefore, this thesis aimed to use molecular tools to gain sufficient information to enable their classification.

RT-PCR using Carla-Uni

The first tool employed was RT-PCR. A carlavirus-specific primer was used since NLV had originally been classified as a carlavirus. This primer, Carla-Uni, was designed to a unique region of homology near the 3'-terminal region of the carlavirus genome. Although the presence of cytoplasmic inclusion bodies appeared to preclude NLV from the carlavirus genus the test was still performed to confirm this supposition. Interestingly, another suspected carlavirus, that formed brush-like cytoplasmic inclusion bodies in infected host cells, cowpea mild mottle virus (CMMV) reacted positively to the RT-PCR test using the Carla-Uni primer. Subsequent sequencing (Badge *et al.*, 1996) demonstrated that CMMV contained an 11K ORF and coat protein ORF with high homology to those of the carlavirus genus. A second unusual virus, red la soda virus, with no serological relationship to known carlaviruses, but all the morphological features associated with the genus also reacted positively to the Carla-Uni RT-PCR test. Two further viruses which had been classified tentatively as carlaviruses on physio-chemical properties alone, AHLV and HMV, were successfully amplified using the Carla-Uni RT-PCR test.

Having observed that two 'suspected' carlaviruses, with features previously anomalous to the accepted trends within the genus were amplified by Carla-Uni in RT-PCR, NLV and MacMV were subject to this same test. Both failed to amplify a product. However, this negative result did not conclusively prove that these viruses were not carlaviruses.

<u>The use of RT-PCR for rapid identification of viruses</u>

Negative RT-PCR results may be obtained for many reasons. Sufficient control reactions ensured that these negative results are not due to an intrinsic failure of the

experiment. Inhibitors carried through the extraction process may affect the Reverse-Transcriptase or the *Taq* polymerase. The absence of the particular target sequence may not necessarily exclude the virus from the genus in question, since this is only a tiny proportion of the genome. The speed at which this type of test can be used dictates that a set of primers, each specific to different parts of the genome, or to different genera can be employed quickly. This is an advantage over conventional means of detecting uncharacterised viruses or identifying the presence of known viruses. PCR has a high degree of sensitivity, detecting viruses at a concentration 10^4-fold lower than detectable by ELISA (Rybicki & Hughes, 1990). PCR is a rapid technique, and does not require lengthy treatments or methods to extract virus from infected tissue, total nucleic acid can be extracted quickly and used directly. Only a small amount of starting material is required, a considerable advantage when a virus is difficult to propagate in experimental hosts or is difficult to purify. PCR primers can be made to different specifications so that a general primer could detect all the members of a single genus or family, or they can be very specific, detecting only one particular virus.

One of the main disadvantages of an RT-PCR test is the unreliable nature of a negative result. Once experimental error has been ruled out, it cannot be concluded that the virus in question does not belong to the specific group tested by the primer. It is inevitable that these types of primers cannot be truly universal. If they contain a high level of degeneracy, they will be unspecific and yet restrictive sequences are likely to vary at some point once the sample size is increased. This is not to deny the usefulness of the positive results obtained with this technique (which have the support of later confirmation by sequence analysis), but perhaps it should be used with the caveat that a negative result is potentially meaningless.

NLV and MacMV are *Macluraviruses*

Following the failure with Uni-Carla, partial 3'-terminal sequence data for NLV and MacMV was obtained by more conventional means. This demonstrated conclusively that they were not carlaviruses. Although they shared greatest overall sequence identity to the bymoviruses, phylogenetic analysis using coat protein sequences showed that they were only distantly related to this genus of the *Potyviridae*. After reviewing coat protein sequence data, genome organisation and the supporting morphological and biological characteristics, it was concluded that NLV and MacMV should be assigned to a new genus, the *Macluraviruses*.

What is a *Macluravirus*?

Since this proposal has been made, it is necessary to discuss the particular features and characteristics of these two viruses which may define the members of this new taxon.

Table 8.1 lists established features of *Macluraviruses* which could be used for comparison to other unassigned viruses.

The first part of the table lists conventional parameters for assessing a virus taxon, listing particle structure, composition and properties (table 8.1, section 1). The size range for macluravirus particles, 650-672nm, is smaller than the potyviruses (720-770nm). However, while the potyviruses conform to a very small size range, all the other genera of the *Potyviridae* have particle size ranges that differ widely from the potyviruses (from 500nm to 900nm). Coat protein sizes predicted from amino acid sequence (NLV 32.8 kDa, MacMV 34.1 kDa) have been given alongside the larger estimates obtained by SDS-PAGE and western analysis in this thesis (NLV 39.5 kDa, MacMV 40kDa). They have genomes of approximately 8 kb of single-stranded RNA (table 8.1,section 3) which is smaller than other potyvirids, but yet prove to be an underestimate once sequenced.

Sections 5-8 of table 8.1 show the more recent findings about the macluraviruses. NLV and MacMV have a poor serological relationship with potyviruses and do not react to the 'universal' potyvirus antiserum (PTY-1; Jordan & Hammond, 1991). They do cross-react to one another, and it is possible that this may be a genus feature. Various amino acid motifs in the nuclear inclusion body b and the coat protein are highlighted (table 8.1, section 6). The motifs ([T/S]GXXXTXXXN[T/S]) and GDD are thought to form the core of the RNA-dependent RNA polymerase (Kamer & Argos, 1984; Domier *et al.*, 1987) and are present in all potyviruses. The motif PYMSLT appears to be only found in bymo- and macluraviruses, and was identified in the sequence of rice necrosis mosaic virus (Chapter 6). This could be an important motif to identify new macluraviruses, but may not be essential or completely conserved within the genus. The cleavage site between the NIb and coat proteins occurs between glutamine and methionine and there is a conserved residue (proline) in the NIb five residues away from the cleavage point. This cleavage site is different from those of other members of the *Potyviridae* and may represent a new consensus specific to the macluraviruses. The motif NGTS has been found in all potyvirus coat proteins, its function is unknown, and is not unique to macluraviruses.

New macluraviruses may be isolated by RT-PCR using the 'NGDD' primer. This test should be prescribed with the caveat for negative results discussed earlier. A potential macluravirus may have already been identified, cassava brown streak-associated virus (CBSaV). CBSaV has the particle morphology of a carlavirus (650-690nm) but has only a weak serological relationship with one carlavirus, CMMV (Brunt, 1996; Lennon *et al.*, 1986). Both viruses are transmitted by whitefly (*Bemisia tabaci*) and CMMV is

Table 8.1: This table (on two pages) lists the established features of the two macluraviruses identified to date.

		narcissus latent virus (NLV)	Maclura mosaic virus (MacMV)
1	**Particles** Structure	slightly flexuous filaments 13nm x 650nm [1]	slightly flexuous filaments 13nm x 672nm (2)
	Composition	Nucleic acid: 5 % of particle weight Coat protein: 32.8 kDa (predicted from amino acid sequence) [3] 39.5 kDa (western analysis) [3]	Nucleic acid: 3-4 % of particle weight Coat protein: 34.1 kDa (predicted from amino acid sequence) [3] 40 kDa (western analysis) [3]
	Properties	single sedimenting component sedimenting coefficient: 158 S buoyant density: 1.33 g/cm^3 [1]	single sedimenting component sedimenting coefficient: 155 4 S buoyant density: 1.307 g/cm^3 [2]
2	**Stability in sap**	*Nicotiana clevelandii* Dilution end point 10^{-3} to 10^{-4} Infective after 10 min 65°C but not at 70°C or after 16 and 14 days at 2°C and 20°C respectively [1]	*Tetragonia expansa* Dilution end point 10^{-3} to 10^{-4} Thermal inactivation point: 65-67°C remains infective after 3 days 20°C [2]
3	**Nucleic Acid**	single-stranded positive-sense RNA 8 kb (estimated by gel electrophoresis) [3]	single-stranded positive-sense RNA 8 kb (estimated by gel electrophoresis) [3]
4	**Cytoplasmic inclusion bodies**	cylindrical cytoplasmic inclusion bodies [4]	cylindrical cytoplasmic inclusion bodies [3]

Table 8.1: continued

			narcissus latent virus (NLV)	Maclura mosaic virus (MacMV)
5	**Serology**	Positive reactions	cross-reacts to MacMV [3, 4]	cross-reacts to NLV [3, 4] weak cross-reaction to bean yellow mosaic potyvirus [4]
		Universal potyvirus PTY-1 antiserum	No cross-reaction [3]	No cross-reaction [3]
6	**Amino acid motifs*** [3]	Nuclear inclusion body b (NIb)	^{1}SGQPSTVVDNT (RdRp core 1) ^{40}FVCNGDDNK (RdRp core 2) ^{87}PYMSLT (bymo/mac box?)	^{212}SGQPSTVVDNT (RdRp core 1) ^{251}FVCNGDDNK (RdRp core 2) ^{298}PYMSLT (bymo/mac box?)
		NIb/coat protein cleavage site	^{211}PxxxLQ/M	^{419}PxxxLQ/M
		Coat protein motifs	^{218}DVG (mutated aphid transmission motif ?)	^{426}DAE (mutated aphid transmission motif ?)
			^{364}NGTS	^{579}NGTS
7	**PCR primers** [3]	Universal potyvirus primer CN48	No amplification	No amplification
		'NGDD' bymo- and macluravirus primer	Amplifies a 1.6kb product	Amplifies a 1.6kb product
8	**3'Untranslated Region** [3]		231 nt	256 nt

* The number at the beginning of each motif indicates the position of the first amino acid of the motif. These are not genome positions since only partial 3' terminal sequence data is available. (x) Denotes any amino acid.

References: [1] Brunt, 1976; 1977 [2] Koenig & Plese, 1981 [3] work presented in this thesis [4] Mowat *et al.*, 1991

known to induce brush-like cytoplasmic inclusion bodies. There are no reports of CBSaV host cytology. It was isolated as part of a mixed infection which is thought to cause cassava brown streak disease. CBSaV could be tested using the range of specific PCR primers employed in this thesis. It has weak associations with a carlavirus that is known to amplify with Carla-Uni, is within the correct particle length range and has poor serological relationships to the carlaviruses. CBSaV has been reported to occur in combination with cassava brown streak potyvirus, however it is unclear whether they are distinct viruses which occur in a complex or only one virus is present. It would therefore be sensible to test any samples with both the universal potyvirus primer (CN48) and the NGDD primer specific to the macluraviruses.

When looking for further potential macluraviruses to investigate, particle size and lack of serological relationship appear to be the most prominent characteristics which differentiate such viruses from the carlavirus or potexvirus genera. However, the presence of cylindrical cytoplasmic inclusion bodies (CCIs), particularly if the macluraviruses can be shown to use a polyprotein translation mechanism, may be of equal importance. In cases where these potential macluraviruses may occur naturally as part of a mixed infection, careful attention should be paid to the correct assignment of CCI induction to a particular virus. This type of confusion has already occurred between NYSV and NLV, where in a mixed infection, CCIs were attributed to NYSV rather than to NLV because it was believed that NLV was a carlavirus and therefore would not be expected to produce CCIs (Chapter 5). It is possible that similar confusion has arisen over the identification of CBSaV as part of a complex formed with a potyvirus, the CCIs observed could have been incorrectly attributed.

There are still many features of the macluraviruses that have not been investigated, or conclusively established: vector transmission, definite monopartite nature of the viruses, complete genome sequence and organisation and confirmation of polyprotein protein processing. These must be borne in mind when carefully assessing potential macluraviruses. No one feature discussed here is essential to group new viruses with the macluraviruses and no single feature should exclude it.

Evolution and the use of higher taxa

The differences between the macluraviruses and the other genera of the *Potyviridae* need to be carefully established as they may provide important data to suggest an evolutionary history for this large family of viruses. The subject of virus evolution is closely linked to their taxonomy; characters defining divisions between viruses may also represent their divergence from other groups. As discussed in Chapter 1, virologists have long avoided the implication of evolutionary relationships as part of

virus taxonomy. The use of coat protein sequence data to construct taxonomic relationships at a species and strain level has been called the new 'molecular virus systematics' (Rybicki, 1990; Rybicki & Shukla, 1992). This system is based on the use of several computer programmes designed to evaluate genetic distance and from these calculations construct phylogenetic trees.

Higher-than-Family Taxa

There appears to be a general agreement that coat protein sequence data, whether it is the complete sequence or the partial data, can be used to construct reliable taxonomic relationships between viruses from strains to family level (Barnett, 1991;1992; Rybicki & Shukla, 1992; Koonin & Dolja, 1993). At any level of taxa higher than family, this agreement breaks down. Those proponents of three supergroups (classes) and several orders of RNA positive-sense viruses (Goldbach & Wellink, 1988; Habili & Symons, 1989; Koonin & Dolja, 1993; Ward, 1993) have all based their divisions on the use of sequence data from non-structural proteins and an observation of the nature of 'conserved arrays' of some genes. The majority of these divisions are based on the homologies between viral RNA-dependent RNA polymerase genes which are found in all positive-sense RNA viruses, but they also incorporate analyses using helicase, methyl-transferase and proteinase sequences. Two objections to this system have been raised. Firstly, that viruses are polyphyletic by nature and cannot have an artificial hierarchical structure imposed them (Rybicki, 1990) and secondly, that the cassette mechanism of gene shuffling observed in virus genomes precludes the use of higher taxa (Rybicki & Shukla, 1992). The latter argument stems from the observation that dendograms constructed from coat protein sequences do not correlate with dendograms constructed from the helicase or RNA-dependent RNA polymerase sequence. Those who apply higher taxa based on non-structural sequence data view this as evidence of gene shuffling, and not fatal to their argument (Goldbach & Wellink, 1988; Ward, 1993; Koonin & Dojla, 1993). Just as different characters may be used to create each division in a hierarchical scheme, so non-structural sequence data can be used to infer relationships above the level of family and coat protein sequences can be used to delineate families, genera, species and strains. In general, an order would include "viruses with a common theme in gene arrangement of the core replicative genes, but a significant variability of the shell [structural genes] organisation" (Koonin & Dolja, 1993).

A test of a taxonomic system is to try to fit new viruses into the scheme. Complete sequences of the RNA-dependent RNA polymerase are not available for either NLV or MacMV. However, preliminary analysis of partial NIb sequences, which contain the GDD motif essential to viral replication, suggest that these viruses are more closely

related to the bymoviruses (57-60%) than they are to other genera within the family (21-31%) (Chapter 4, table 4.1b). NLV and MacMV are distinct from the bymoviruses, since they are more similar to one another (76.3%) in this region than they are to the bymoviruses (57-60%). Following the reasoning behind the delineation of higher taxa, this preliminary data supports the allocation of NLV and MacMV to a separate genus.

The macluraviruses would be placed in the same higher taxa groupings as the potyviruses, since they belong to the potyviridae family. According to Koonin & Dolja (1993) this would place them in the class *Picornavirata* and the order *Potyvirales*. The *Potyvirales* is comprised of two families, the *Potyviridae* and the *Hypoviridae*. The *Hypoviridae* are double-stranded RNA viruses which infect fungi (Murphy *et al.*, 1995). No true virions are associated with the members of this family whose type member is Cryphonectria hypovirus 1-EP713 (CHV1-EP713) which infects chestnut blight fungus, *Cryphonectria parasitica*. The 10-13 kb dsRNA encodes a polyprotein which is autocatalytically processed. Proteinase, polymerase and helicase domains have been demonstrated to share sequence identity to the *Potyviridae* (Koonin *et al.*, 1991). However, this association is not made by Ward (1993) where the *Potyviridae* is the only family placed in order 2 of class 1 (otherwise equivalent to the *Picornavirata*) and the *Hypoviridae* are not included in the classification scheme.

Evolution of positive-sense RNA viruses

Evolutionary relationships are necessarily inferred by the creation of higher taxa, suggesting a single common ancestor (Koonin & Dolja, 1993). Until relatively recently the extent of RNA-RNA recombination in viral populations was not completely realised or appreciated. However, it has now been accepted that such events are widespread and that homologous recombination may act as a repair mechanism to counter the error prone RNA-dependent RNA polymerase (for review see Simon & Bujarski, 1994). Sequence data provides further evidence for the many unusual ways that viruses can diverge and evolve using RNA recombination. Horizontal transfer of genes between viruses (intra- and inter-specific) and from non-homologous host genomes has probably occurred through RNA recombination. For example, it has been suggested that segmented genomes could have arisen either by the split of one genome or the capture of a heterologous RNA segment. These two points are of particular interest when considering the position of the macluraviruses in the evolutionary history of the potyviruses. The bymoviruses have segmented genomes and it is possible that they have acquired a rod-like coat protein domain from the furoviruses (Koonin & Dolja, 1993; Dessens *et al.*, 1995; Dessens & Meyer, 1996). Furoviruses are transmitted by a different soil-borne fungus (*Spongospora*) to the bymoviruses (*Polymyxa*), and it appears that the area of homology between the two is limited to the domain required for

fungal transmission and does not include that specifying the different coat protein shape. The bymoviruses have also lost N-terminal domain of HC-Pro, which is associated with aphid vector specificity (Thornbury *et al.*, 1990; Atreya *et al.*, 1992), whilst retaining the proteinase domain.

There is partial sequence data to suggest that MacMV has a homologous genome organisation at its 5'-terminus (Chapter 4). The MacMV sequence homologous to the bymovirus protein P1 on RNA2 is also homologous to the proteinase of HC-Pro, however, MacMV does not encode a protein homologous to the N-terminal domain of HC-Pro. Unfortunately the sequence data do not extend far enough to compare with the suggested fungal transmission region of the bymovirus protein P2 and the equivalent region in the furoviruses. These data could provide an evolutionary link between the bymoviruses and the potyviruses and suggest whether the fungal transmission region was transferred from the furoviruses. It has yet to be conclusively established that the macluraviruses are transmitted by aphids or by some other vector but it is tantalising to suggest that the macluraviruses can be occasionally transmitted by both aphid and fungal vectors, representing a missing link between the potyviruses and the bymoviruses.

Macluraviruses and vector transmission

It is not hard to imagine a scenario where macluraviruses could loose an ability to be transmitted by an aphid vector. Mutated forms of the DAG motif present in the sequenced isolates of MacMV and NLV suggest that these isolates may once have been aphid transmitted, although it is unlikely that they could be efficiently transmitted now. NLV is known to occur in the field at times which are not necessarily prone to heavy aphid infestations (Brunt, 1971). There are no reports on the infestation of *Maclura pomifera* by aphids and it is possible that this vector may have been selected for testing due to the relationship of MacMV to the potyviruses. The *Maclura* tree and narcissus bulbs can both be said to be long-lived hosts, since bulbs are often vegetatively propagated for many years. Thus the selection pressure for vector transmission may have been exceptionally low for both these viruses. Being propagated by other means (seed transmission or bulb division) a sub-population deleted for the N-terminal of HC-Pro necessary for aphid transmission may have a selective advantage.

It has been observed that potyviruses encoding a non-functional HC-Pro can be complemented by the presence of functional HC-Pro from another potyvirus (Thornbury *et al.*, 1985). NLV is often found in complex with other potyviruses and it is possible that it is transmitted at a low level by aphids in the presence of functional helper component protein from other narcissus potyviruses.

It has been suggested that the bymoviruses are the most ancient genera of the potyvirus family and that the acquisition of a monopartite genome and the aphid transmission domain of HC-Pro lead to the precursor of the modern potyvirus (Shukla *et al.*, 1994). If this is so, the first macluraviruses that evolved may represent a link between these two states, a monopartite virus, without the complete HC-Pro. It possible that the macluraviruses can be transmitted by aphids on an inefficient and infrequent basis. Perhaps the highly transmissible potyviruses that can be observed today evolved from those ancient macluraviruses by acquiring the aphid transmission domain of HC-Pro and undergoing further modifications to the coat protein. Further sequence data from the 5'-terminus of MacMV and confirmation of its monopartite structure by northern analysis may provide further information to investigate these possibilities.

It will be informative to compare the genome structure of MacMV with that of the bymoviruses and the potyviruses. The bymoviruses have a considerably larger genome (approx. 12kb) than the potyviruses when the bipartite segments are added together. It is usual to observe streamlining of a genome over time in an environment where replication time is a selective factor, and it is possible that the potyviruses represent such a case. The macluraviruses have been sized at 8 kb by gel electrophoresis, but this is often an under estimate when genomic sequence is established. If the macluraviruses represented the evolutionary link between bymo- and potyviruses, then it may be expected that they would have an intermediate genome size.

An alternative theory is that the macluraviruses are the most ancient genera of the *Potyviridae*. They could represent a virus that was transmitted only by vegetative propagation or seed transmission. The segmentation of the genome could result in a bymovirus-like genome, which then acquired a fungal transmission domain. Such evolutionary theories are difficult to prove with viruses since there is not yet an accurate way of measuring the molecular clock of RNA genomes. More sequence data may lend weight to one theory over another, and the vector transmission studies of the sequenced isolates of NLV and MacMV would aid the discussion further.

The acquisition of a new vector by the potyviruses appears to have led to massive adaptive radiation. Potyviruses are the largest and most diverse of all the plant genera, whereas the bymoviruses have a very restricted host range. NLV appears to be confined to horticultural bulbous hosts and the natural host range of MacMV is unknown. Infections may be isolated in a single bulb for many years and therefore considerable divergence may occur between isolates. This may make it difficult to define new isolates of viruses from these types of host.

Delineations between species and strain

At the other end of the taxonomic scale, the delineation between species and strain presents its own problems. There have been arguments over the nature of the continuum between one species and the next and in evolutionary terms. It is perhaps easier for such a concept to exist in virology than in, say, zoology because there are no restrictions of fertility for asexual viruses. However, a practical need to describe disease has led to the use of coat protein sequence data to establish that strains share 90-99% sequence identity, whereas species share 38-71% (Shukla & Ward, 1988). A sub-species taxon has been suggested for the grey area between 72% and 89%. However, it is difficult to apply these assumptions on the potyviruses infecting Narcissus that were discussed in Chapter 5.

Two 'new' narcissus-infecting potyviruses were identified by RT-PCR. It was suspected, but could not be conclusively proved that they were narcissus yellow stripe virus (NYSV) and narcissus late season yellows virus (NLSYV). Both were related to one another and to several tulip-infecting viruses. This surprising result raises questions as to the nature of this relationship. Tulip top breaking virus (TTBV) is so closely related to turnip mosaic virus (TuMV) that is considered a strain, should these new narcissus potyviruses be considered strains of TuMV or of the tulip-infecting viruses? The percentage amino acid sequence identities observed between the partial coat protein sequences of narcissus potyviruses and TuMV, TTBV and tulip breaking virus (TBV) fell within the 'grey area' of sub-species (81.5-86.9% identity; figure 5.8b). Due to the isolating nature of their hosts, it may be reasonable to assume that there would be a higher degree of divergence between virus isolates than is normally observed between viral strains. This assumption would require the downward revision of the cut-off point between strain and sub-species for bulb-infecting viruses.

When discussing taxonomy at the level of strains and sub-species, it is important to bear in mind the practical use of these delineations. The connection between diagnosis and theoretical discussions is often a tenuous one, but the important point here is that those in the field be made aware that these viruses with different aetiologies may be sufficiently similar to cross-react to a single antiserum test. It is possible, for example, that NYSV and NLSYV will cross-react (it is impossible to say from the data presented in Chapter 5 whether either antisera was raised from a single virus infection). These two diseases in particular have a history of confusion.

Future work

Although most aspects of the immediate future work which could be carried out from work in this thesis have been detailed at the end of each chapter, there are some over arching aims which could be investigated.

The first is to fully explore the use of genus or family specific RT-PCR primers on new and unassigned viruses. The isolation of a third macluravirus would help to establish more clearly trends within this new genus and hence its place within the *Potyviridae*. If sequence data are to be used throughout virus taxonomy, then it will be necessary to collate as many sequences as possible to make such analyses more robust.

The second is to completely sequence MacMV. The high degree of identity shown between NLV and MacMV makes it unnecessary to completely sequence both viruses, and more data is already available for MacMV. The impact that MacMV's genome organisation could have on our understanding of the relationship between the potyviruses and the bymoviruses is potentially enormous.

Thirdly, the future work described in Chapter 5 which is necessary to fully characterise the two narcissus potyviruses isolated and sequenced, will enable the investigation of narcissus yellow stripe disease. The long history of confusion over the complex virus infections of narcissus could be decisively concluded. A survey using NLV derived nucleic acid probes rather than serological techniques, could lead to the isolation of an aphid transmitted strain.

Finally the work on transgenic resistance to the bymoviruses is of importance to commercial barley growers and may continue elsewhere. As genetic engineering seems to be one of only a limited number of options for resistance to BaYMV and BaMMV other techniques and other constructs could be pursued. Although the work to engineer resistance to MacMV or BaMMV was unsuccessful in this thesis, there are still many options open to exploration. Transgenic plants may also provide a means to study the action of the macluraviruses or bymoviruses *in vivo*.

Conclusions

The use of RT-PCR as a rapid and specific test for unassigned viruses has confirmed the classification of ALHV, HMV, CMMV, RLSV and RNMV. Investigation of the confused history of narcissus yellow stripe disease has been clarified by the amplification of two narcissus-infecting potyvirus (possibly NYSV and NLSYV). Transgenic tobacco plants were generated, but no resistant lines were isolated. The main aim of the thesis has been achieved, demonstrating that NLV and MacMV are not carlaviruses and should be assigned to a new genus, the *Macluraviruses* within the

family *Potyviridae*. It is hoped that this proposal will be submitted to the ICTV for official recognition.

References Cited

ADAMS, A.N. & BARBARA, D.J. (1982). The use of F(ab') 2-based ELISA to detect serological relationships among carlaviruses. *Annals of Applied Biology* **101**, 495-500.

ADAMS, M.J. (1989). Deployment of resistance to the barley yellow mosaic virus in the UK. *Proceedings of the IVth International Plant Virus Epidemiology Workshop, Resistance to viruses and vectors - temperate and tropical plants, Montpellier*, pp 77-78.

ALLISON, R.F., JOHNSTON, R. E.. & DOUGHERTY, W.G. (1986). The nucleotide sequence of the coding region of tobacco etch virus genomic RNA: evidence for the synthesis of a single polyprotein. *Virology* **154**, 9-20.

ANDERSON, J.M., PALUKAITIS, P. & ZAITLIN, M. (1992). A defective replicase gene induces resistance to cucumber mosaic virus in transgenic tobacco plants. *Proceedings of the National Academy of Science, USA* **89**, 8759-8763.

ATREYA, C.D., RACCAH, B. & PIRONE, T.P. (1990). A point mutation in the coat protein abolishes aphid transmission of a potyvirus. *Virology* **178**, 161-165.

ATREYA, C.D., ATREYA, P.L., THORNBURY, D.W & PIRONE, T.P. (1992). Site-directed mutations in the potyvirus HC-Pro gene affect helper component activity, virus accumulation and symptom expression in infected tobacco plants. *Virology* **191**, 106-111.

ATREYA, P.L., LOPEZ-MOYA, J.J., CHU, M., ATREYA, C.D. & PIRONE, T.P. (1995). Mutational analysis of the coat protein N-terminal amino acids involved in potyvirus transmission by aphids. *Journal of General Virology* **76**, 265-270.

AUDY, P., PALUKAITIS, P., SLACK, S.A. & ZAITLIN, M. (1994). Replicase-mediated resistance to potato virus Y in transgenic tobacco plants. *Molecular Plant-Microbe Interactions*, **7**, 15-22.

BADGE, J., BRUNT, A., CARSON, R., DAGLESS, E., KARAMAGIOLI, M., PHILLIPS, S., SEAL., S., TURNER, R., & FOSTER, G.D. (1996). A carlavirus-specific PCR primer and partial nucleotide sequence provides further evidence for the recognition of cowpea mild mottle virus as a whitefly-transmitted carlavirus. *European Journal of Plant Pathology* **102**, 305-310

BADGE, J.L., ROBINSON, D.J., BRUNT, A.A. & FOSTER, G.D. (1997a). 3'-terminal sequences of the RNA genomes of narcissus latent and Maclura mosaic viruses suggest that they represent a new genus of the *Potyviridae*. *Journal of General Virology* **78**, 253-257.

BADGE, J.L., KASHIWAZAKI, S., LOCK, S. & FOSTER, G.D. (1997b) A bymovirus-specific PCR primer and partial nucleotide sequence provides further evidence for the recognition of rice necrosis mosaic virus as a bymovirus. *European Journal of Plant Pathology* (in press).

BARNETT, O.W. (1991).Potyviridae, a proposed family of plant viruses. *Archives of Virology* **118**, 139-141.

BARNETT, O.W. (1992). A summary of potyvirus taxonomy and definitions. In: Barnett, O.W. (ed.), Potyvirus Taxonomy. Springer, Wein and New York. *Archives of Virology* **supplement 5**, pp 435-444.

BAULCOME, D.C., SAUNDERS, G.R., BEVAN, M.W., MAYO, M.O. & HARRSION, B.D. (1986).Expression of biologically active viral satellite RNA from the nuclear genome of transformed plants. *Nature* **321**, 446-449.

BENDIEK, J., DAVIDSON, A.D., SCHULZE, S.C., SCHELL, J. & STEINBISS, H.-H. (1993). Identification and classification of a resistance breaking strain of barley yellow mosaic virus. *Annals of Applied Biology* **122**, 481-491.

BERG, J. (1986). Potential metal binding domains in nucleic acid binding proteins. *Science* **232**, 485-486.

BRATTEY, C., JEFFRIES, C., GOODFELLOW, S., ROSE, G., BURNS, R. & GEORGE, E. (1996). Some characteristics of a virus infecting *Solanum tuberosum* cv. Red La Soda imported from the USA. Scottish Agricultural Science Agency, East Craigs, Edinburgh.

BRAUN, C.J. & HEMENWAY, C.L. (1992). Expression of amino-acid terminal portions or full length viral replicase genes in transgenic plants confers resistance to potato virus X infection. *The Plant Cell* **4**, 735-744.

BRUNT, A.A. (1966). Narcissus mosaic virus. *Annals of Applied Biology* **58**, 13-23.

BRUNT, A.A. (1971). Narcissus yellow stripe virus. *C.M.I.A.A.B. Descriptions of Plant Viruses* **No.4**, 4pp.

BRUNT, A.A. (1976). Narcissus latent virus. *C.M.I.A.A.B. Descriptions of Plant Viruses* **No.170**, 4pp.

BRUNT, A.A. (1977). Some hosts and properties of narcissus latent virus, a carlavirus commonly infecting narcissus and bulbous iris. *Annals of Applied Biology* **87**, 355-364.

BRUNT, A.A. (1980). A review of problems and progress in research on viruses and virus diseases of narcissus in Britain. *Acta Horticulturae* **110**, 23-30.

BRUNT, A.A. (1992). The general properties of potyviruses. In: Barnett, O.W. (ed)., Potyvirus Taxonomy. Springer, Wein and New York. *Archives of Virology*, **supplement 5**, pp 3-16.

BRUNT, A.A. (1995). Narcissus. In: Loebenstein, G., Lawson, R.H. & Brunt, A.A. (eds.) (1995). Virus and virus-like diseases of bulb and flower crops. John Wiley & Sons, Chichester, UK.pp 322-334.

BRUNT, A.A. (1996). Cassava brown streak associated (?) carlavirus. In: Brunt, A., Crabtree, K., Dallwitz, M., Gibbs, A. & Watson, L.(eds.) (1996). Viruses of Plants. Descriptions and lists from the VIDE Database. CAB International, Wallingford, Oxford, UK. pp. 331-332.

BRUNT, A.A. & ATKEY, P.T. (1967). Rapid detection of narcissus yellow stripe and two other filamentous viruses in crude negatively stained narcissus sap. *Report of the Glasshouse Research Institute* **1966**, 155-159.

BRUNT, A.A. & KENTEN, R.H. (1973). Cowpea mild mottle, a newly recognised virus infecting cowpeas (*Vigna unguiculata*) in Ghana. *Annals of Applied Biology* **74**, 67-74.

BRUNT, A.A., ATKEY, P.T. & WOODS R.D. (1983). Intracellular occurrence of cowpea mild mottle virus in two unrelated plant species. *Intervirology* **20**, 137-142.

BRUNT, A.A., ATKEY, P.T., FROST, C. & PHILLIPS, R.S. (1994). Cytoplasmic inclusions induced by narcissus latent virus. *Acta Horticulturae* **377**, 275-280.

CAMPBELL, R.N. (1996). Fungal transmission of plant viruses. *Annual Review of Phytopathology* **34**, 87-108.

CARR, J.P., MARSH, L.E., LOMONOSSOFF, G.P., SEKIYA, M.E. & ZAITLIN, M. (1992). Resistance to tobacco mosaic virus induced by the 54kDa gene sequence requires expression of the 54kDa protein. *Molecular Plant-Microbe Interactions* **5**, 397-404.

CARRINGTON, J.C. & HERNDON, K.L.(1992). Characterisation of the potyviral HC-Pro autocatalytic cleavage site. *Virology* **187**, 308-315.

CARRINGTON, J.C., CARY, S.M., PARKS, D.T. & DOUGHERTY, W.G. (1989). A second proteinase encoded by a plant potyvirus genome. *The EMBO Journal* **8**, 365-370.

CAVILEER, T.D., HALPERN, B.T., LAWRENCE, D.M., PODLECKIS, E.V., MARTIN, R.R. & HILLMAN, B.I. (1994). Nucleotide sequence of the carlavirus associated with blueberry scorch and similar diseases. *Journal of General Virology* **75**, 711-720.

CLEVELAND, D.W., FISCHER, S.G., KIRSCHNER, N.W. & LAEMMLI, U.K. (1977). Peptide mapping by limited proteolysis in sodium dodecyl suplhate and analysis by gel electrophoresis. *Journal of Biological Chemistry* **252**, 1102.

COLINET, D., KUMMERT, J. & LEPOIVRE, P. (1996). Molecular evidence that white-fly transmitted sweet potato mild mottle virus belongs to a distinct genus of the *Potyviridae. Archives of Virology* **141**, 125-135.

CREMER, M.C. & VAN DER VEKEN, J.A. (1964). Localisation of viral antigen in narcissus leaves infected with yellow stripe virus, determined by means of a fluorescein conjugated antiserum. *Netherlands Journal of Plant Pathology* **72**, 105-113.

DAVIDSON, A.D., PROLS, M., SCHELL, J. & STEINBISS, H.-H. (1991). The nucleotide sequence of RNA 2 of barley yellow mosaic virus. *Journal of General Virology* **72**, 989-993.

DEKKER, E.L., DERKS, A.F.L.M., ASJES, C.J., LEMMERS, M.E.C., BOL, J.F. & LANGEVELD, S.A. (1993). Characterisation of potyviruses from tulip and lily which cause flower-breaking. *Journal of General Virology* **74**, 881-887.

DERKS, A. F.L.M., HOLLINGER, TH. C. & VINK-VAN DEN ABEELE, J.L. (1985). Identification and symptom expression of four elongated viruses infecting bulbous iris. *Acta Horticulturae* **164**, 309-317

DESSENS, J.T. & MEYER, M. (1995). Characterization of fungally and mechanically transmitted isolates of barley mild mosaic virus : two strains in competition. *Virology* **212**, 383-391.

DESSENS, J.T. & MEYER, M. (1996). Identification of structural similarities between putative transmission proteins of *Polymyxa* and *Spongospora* transmitted bymoviruses and furoviruses. *Virus Genes* **12**, 95-99.

DESSENS, J.T., NGUYEN, M. & MEYER, M. (1995). Primary structure and sequence analysis of RNA2 of a mechanically transmitted barley mild mosaic virus isolate: an evolutionary relationship between bymo- and furoviruses. *Archives of Virology* **140**, 325-333.

DINANT, S., LOT, H., ALBOUY, J., KUZIAK, C., MEYER, M. & ASTIER-MANIFACIER, S. (1991). Nucleotide sequence of the 3' terminal region of lettuce mosaic potyvirus RNA shows a Gln/Val dipeptide at the cleavage site between the polymerase and the coat protein. *Archives of Virology* **116**, 235-252.

DOD, C.W. (1894). Variegation in narcissus. *Journal of the Royal Society* **17**, 22.

DOMIER, L.L., FRANKLIN, K.M., SHAHABUDDIN, M., HELLMAN, G.M., OVERMEYER, J.M., HIREMATH, S.T., SIAW, M.F.E., LOMONOSSOFF, G.P., SHAW, J.G. & RHOADS, R.E. (1986). The nucleotide sequence of tobacco vein mottling virus RNA. *Nucleic Acids Research* **14**, 5417-5430.

DOMIER, L.L., SHAW, J.G. & RHOADS, R.E. (1987). Potyvirus proteins share amino acid sequence homology with Picorna-, Como- and Caulimoviral proteins. *Virology* **158**, 20-27.

DOUGHERTY, W.G. & CARRINGTON, J.C. (1988). Expression and function of potyviral gene products. *Annual Review of Phytopathology* **26**, 123-143

DOUGHERTY, W.G. & PARKS, T.D. (1989). Molecular genetic and biochemical evidence for the involvement of the heptapeptide cleavage sequence in determining the reaction profile at two tobacco etch virus cleavage sites in cell-free assays. *Virology* **172**, 145-155.

DOUGHERTY, W.G., CARY, S.M. & PARKS, T.D. (1989). Molecular genetic analysis of a plant virus polyprotein cleavage site: a model. *Virology* **171**, 356-364.

FOSTER, G.D. (1992). The structure and expression of the genome of carlaviruses. *Research in Virology* **143**, 103-112.

FOSTER, G.D. & MILLS, P.R. (1990a). Evidence for the role of subgenomic RNAs in the production of potato virus S coat protein during *in vitro* translation. *Journal of General Virology* **71**, 1247-1249.

FOSTER, G.D. & MILLS, P.R. (1990b).Evidence for the role of subgenomic RNA species in the production of *Helenium* virus S coat protein during *in vitro* translation. *Virus Research* **17** 61-70.

FOSTER, G.D. & MILLS, P.R. (1990c). Investigation of the 5' terminal structures of genomic and subgenomic RNAs of potato virus S. *Virus Genes*, **4**, 359-366.

FOSTER, G.D. & MILLS, P.R. (1991a). Cell-free translation of American hop latent virus. *Virus Genes* **5**, 327-334.

FOSTER, G.D. & MILLS, P.R. (1991b). Evidence for subgenomic RNAs in leaves infected with an Andean strain of potato virus S. *Acta Virologica* **35**, 260-267.

FOSTER, G.D. & MILLS, P.R.(1991c). Occurrence of chloroplast ribosome binding recognition sites within conserved elements of the RNA genomes of carlaviruses. *FEBS Letters* **280**, 341-343.

FOSTER, G.D. & MILLS, P.R. (1991d). Translation of potato virus S RNA *in vitro* : evidence of protein processing. *Virus Genes* **6**, 45-52.

FOSTER, G.D. & MILLS, P.R. (1991e). Analysis of the coat proteins of the ordinary and Andean strains of potato virus S and antigenic comparisons of carlaviruses. *Acta Virologica* **35**, 184-190.

FOSTER, G.D. & MILLS, P.R. (1992). The 3'-nucleotide sequence of an ordinary strain of potato virus S. *Virus Genes,* **6**, 213-220.

FOSTER, G.D., MILLAR, A.W., MEEHAN, B.M. & MILLS, P.R. (1990a).Nucleotide sequence of the 3'-terminal region of *Helenium* virus S. *Journal of General Virology* **71**, 1877-1880.

FOSTER, G.D., MEEHAN, B.M. & MILLS, P.R. (1990b). A comparison of the nucleotide sequence homologies between isolates of the Andean and ordinary strains of potato virus S and their relationship to other carlaviruses. *Virus Genes,* **4**, 257-260.

FOULDS, I.J., LEA, V.J., SIDEBOTTOM, C., JAMES, C.M., BOULTON, R.E., BREARS, T., SLABAS, A.R., JACK, P.L. & STRATFORD, R. (1993). Cloning and sequence analysis of the coat protein gene of barley mild mosaic virus. *Virus Research* **27**, 79-89.

FRANCKI, R.I.B., FAUQUET, C.M., KNUSDEN, D.L. & BROWN, F. (1991). Classification and nomenclature of viruses. Fifth Report of the International Committee on Taxonomy of Viruses. *Archives of Virology*, **supplement 2**, 341-344

GIBBS, A. & MACKENZIE, A. (1997). A primer pair for amplifying part of the genome of all potyvirids by RT-PCR. *Journal of Virological Methods* **63**, 9-16.

GOLDBACH, R. & WELLINK, J. (1988). Evolution of plus-strand RNA viruses. *Intervirology* **29**, 260-267.

GOLEMBOSKI, D.B., LOMONOSSOFF, G.P. & ZAITLIN, M. (1990). Plants transformed with a tobacco mosaic virus nonstructural gene sequence are resistant to the virus. *Proceedings of the National Academy of Science, USA* **87**, 6311-6315.

GOTZ, R. & MAISS, E. (1995). The complete nucleotide sequence and genome organisation of the mite-transmitted brome streak mosaic virus in comparison with those of potyviruses. *Journal of General Virology* **76**, 2035-2042.

GROS, E. & WITKOP, B. (1961). Selective cleavage of the methionyl peptide in ribonuclease cyanogen bromide. *Journal of American Chemical Society* **83**, 1510-1511.

GUO, H.S. & GARCIA, J.A. (1997). Delayed resistance to plum pox potyvirus mediated by a mutated RNA replicase gene: involvement of a gene-silencing mechanism. *Molecular Plant-Microbe Interactions* **10**, 160-170.

HAASIS, F.A. (1939). Studies on narcissus mosaic. *Memoirs of Cornell University Agricultural Experiment Station* **224**, 3-20.

HABILI, N. & SYMONS, R.H. (1989). Evolutionary relationship between luteoviruses and other RNA plant viruses based on sequence motifs in their putative RNA polymerases and nucleic acid helicases. *Nucleic Acids Research* **17**, 9543-9555.

HARIRI, D.J., FOUCHARD, M. & LAPIERRE, H. (1990). Resistance to barley yellow mosaic virus and to barley mild mosaic virus in barley. *Proceedings of the First symposium of the International working Group on Plant Viruses with Fungal Vectors, Braunschweig*, pp 109-112.

HARRISON, B.D., FINCH, J.T., GIBBS, A.J., HOLLINGS, M., SHEPHERD, R.J., VALENTA, V. & WETTER, C. (1971). Sixteen groups of plant viruses. *Virology* **45**, 356-363.

HAWKER, L. (1943).Experiments in the spread of "narcissus stripe" in the field. *Annals of Applied Biology*, **30**, 184-185.

HAYLOR, M.T.M., BRUNT, A.A. & COUTTS, R.H.A. (1990). Conservation of the 3' terminal nucleotide sequence in five carlaviruses. *Nucleic Acids Research*, **18**, 6127.

HENDERSON, J., GIBBS, M.J., EDWARDS, M.L., CLARKE, V.A., GARDNER, K.A. & COOPER, J.I. (1992). Partial nucleotide sequence of poplar mosaic virus RNA confirms its classification as a carlavirus. *Journal of General Virology* **73**, 1887-1890.

HERBERT, T.T. & PANIZO, C.H. (1975). Oat mosaic virus.*CMI/ Descriptions of Plant Viruses* **No. 145**, 4pp.

HERSKOWITZ, I. (1987). Functional inactivation of a gene by dominant negative mutations. *Nature* **329**, 219-222.

HULL, R. & DAVIES, J.W. (1992). Approaches to nonconventional control of plant virus disease. *Critical Reviews in Plant Sciences* **11**, 17-33.

HUTH, W. & ADAMS, M. J. (1990). Barley yellow mosaic virus (BaYMV) and BaYMV-M: two different viruses. *Intervirology* **31**, 38-42.

INOKUCHI, Y. & HARISHIMA, A. (1987). Interference with viral infection by defective RNA polymerase. *Journal of Virology* **61**, 3946-3949.

INOKUCHI, Y. & HARISHIMA, A. (1990). Interference with viral infection by RNA replicase deleted at the carboxy-terminal region. *Journal of Biochemistry* 108, 53-58.

INOUYE, T. & FUJII, S. (1977). Rice necrosis mosaic virus. *CMI/AAB Descriptions of Plant Viruses* **No. 172**, 4pp.

INOUYE, T. & SAITO, Y. (1975). Barley yellow mosaic virus. *CMI/AAB Descriptions of Plant Viruses* **No. 143**, 4pp.

JACOBI, V., PEERENBOOM, E., SCHENK, P.M., ANTONIW, J.F., STEINBISS, H.H. & ADAMS, M.J. (1995). Cloning and sequence analysis of RNA-2 of a mechanically transmitted UK isolate of barley mild mosaic bymovirus (BaMMV). *Virus Research* **37**, 99-111.

JEYANANDARAJAH, P. & BRUNT, A.A. (1993). The natural occurrence, transmission, properties and possible affinities of cowpea mild mottle virus. *Journal of Phytopathology* **137**, 148-156.

JORDAN, R., & HAMMOND, J. (1991). Comparison and differentiation of potyvirus isolates and identification of strain-, virus-, and subgroup-specific and potyvirus-common epitopes using monoclonal antibodies. *Journal of General Virology* **72**, 25-36.

KAMER, G. & ARGOS, P. (1984). Primary structural comparison of RNA-dependent polymerases from plant, animal and bacterial viruses. *Nucleic Acids Research* **18**, 7269-7282

KAMPHIUS, I.G., DRENTH, J. & BAKER, E.N. (1985). Thiol proteases: comparative studies based on high resolution structures of papain and actinidin and on amino acid sequence information for cathepsins B & H and stem bromelain. *Journal of Molecular Biology* **182**, 317-329.

KASHIWAZAKI, S. (1996). The complete nucleotide sequence and genome organisation of barley mild mosaic virus (Na1 strain). *Archives of Virology* **141**, 2077-2089.

KASHIWAZAKI, S. & HIBINO, H. (1996) Genomic reassortment of barley mild mosaic virus : evidence for involvement of RNA 1 in pathogenicity. *Journal of General Virology* **77**, 581-585.

KASHIWAZAKI, S., HAYANO, Y., MINOBE, Y., OMURA, T., HIBINO, H. & TSUCHIZAKI, T. (1989). Nucleotide sequence of the capsid protein gene of barley yellow mosaic virus. *Journal of General Virology* **70**, 3015-3023.

KASHIWAZAKI, S., MINOBE, Y., OMURA, T. & HIBINO, H. (1990a). Nucleotide sequence of barley yellow mosaic virus RNA 1: a close evolutionary relationship with potyviruses. *Journal of General Virology* **71**, 2781-2790.

KASHIWAZAKI, S., NOMURA, K., WATANABE, K., TOSHIMA, I., IIDA, Y., USUGI, T., OGAWA, K., HIBINO, H. & TSUCHIZAKI, T. (1990b). Barley yellow mosaic virus and barley mild mosaic virus: strains and host resistance. *Proceedings of the First symposium of the International Working Group on Plant viruses with Fungal Vectors, Braunschweig*, pp 105-108.

KASHIWAZAKI, S., MINOBE, Y. & HIBINO, H. (1991) Nucleotide sequence of barley yellow mosaic virus RNA 2. *Journal of General Virology* **72**, 995-999.

KASHIWAZAKI, S., NOMURA, K., KURODA, H., ITO, H. & HIBINO, H. (1992). Sequence analysis of the 3' terminal halves of RNA-1 of two strains of barley mild mosaic virus. *Journal of General Virology* **73**, 2173-2181.

KLUG, A. & RHODES, D. (1987). 'Zinc-Fingers' : a novel protein motif for nucleic acid recognition. *Trends In Biochemical Sciences* **12**, 464-469.

KOENIG, R. (1982). The Carlavirus Group. *CMI/AAB Descriptions of Plant viruses* No. 200, 4pp.

KOENIG, R. & PLESE, N. (1981). Maclura mosaic virus. *CMI/AAB Descriptions of Plant viruses* No. 239, 3pp.

KOONIN, E.V. & DOLJA, V.V. (1993). Evolution and taxonomy of positive strand RNA viruses: implications of comparative analysis of amino acid sequences. *CRC Critical Reviews in Biochemistry and Molecular Biology* 28, 375-430.

KOONIN, E.V., CHOI, G.H., NUSS, D.L., SHAPIRA, R. & CARRINGTON, J.C. (1991). Evidence for common ancestry of a chestnut blight hypovirulence-associated double-stranded RNA and a group of positive strand RNA viruses. *Proceedings of the National Academy of Science, USA* 88, 10647-10651.

LAEMMLI, U.K. (1970). Cleavage of the structural proteins during the assembly of the head of bacteriophage T4. *Nature* 227, 680-685.

LAIN, S. REICHMANN, J.L. & GARCIA, J.A. (1990). RNA helicase: a novel catalytic activity associated with protein encoded by a positive-strand RNA virus. *Nucleic Acids Research* 18, 7003-7006.

LANGEVELD, S.A., DORE, J.-M., MEMLINK, J., DERKS, A.F.L.M., VAN DER VLUGT, C.I.M., ASJES, C.J. & BOL, J.F. (1991). Identification of potyviruses using the polymerase chain reaction with degenerate primers. *Journal of General Virology* 72, 1531-1541.

LAWRENCE, D.M., ROZANOV, M.N. & HILLMAN, B.I. (1995). Autocatalytic processing of the 223-kDa protein of blueberry scorch carlavirus by a papain-like proteinase. *Virology* 207, 127-135.

LENNON, A.M., AITON, M.M. & HARRISON, B.D. (1986). Cassava viruses from Africa. *Report of the Scottish Crop Research Institute 1985*, p168.

LEVAY, K. & ZAVRIEV, S. (1991). Nucleotide sequence and gene organisation of the 3'-terminal region of chrysanthemum virus B genomic RNA. *Journal of General Virology* 72, 2333-2337.

LI, X.H. & CARRINGTON, J.C. (1995). Complementation of tobacco etch potyvirus mutants by active RNA polymerase expression in transgenic cells. *Proceedings of the National Academy of Science , USA* **92**, 457-461.

LOMONOSSOFF, G.P. (1995). Pathogen-derived resistance to plant viruses. *Annual Review of Phytopathology* **33**, 323-343.

LOMONOSSOFF, G.P. & DAVIES, J.W. (1992). Replicase-mediated resistance to plant viruses. *Scientific Progress Oxford* **76**, 537-551.

LONGSTAFF, M., BRIGNETI, G., BOCCARD, F., CHAPMAN, S. & BAULCOMBE, D. (1993). Extreme resistance to potato virus X infection in plants expressing a modified component of the putative viral replicase. *The EMBO Journal* **12**, 379-386.

LWOFF, A., HORNE, R. & TOURNIER, P. (1962). A system of viruses. *Cold Spring Harbour Series on Quantitative Biology* **27**, 51-55.

MACFARLANE, S. A. & DAVIES, J.W. (1992). Plants transformed with a region of the 201-kilodalton replicase gene of pea early browning virus are resistant to infection. *Proceedings of the National Academy of Science, USA* **89**, 5829-5833.

MACKENZIE, D.J., TREMAINE, J.H. & STACE-SMITH, R. (1989). Organisation and interviral homologies of the 3'-terminal portion of potato virus S. *Journal of General Virology* **70**, 1053-1063.

MAISS, E., TIMPE, U., BRISSKE, A., JELKMANN, W., CASPER, R., HIMMLER, G., MATTANOVICH, D. & KATINGER, H.W.D (1989). The complete nucleotide sequence of plum pox virus RNA. *Journal of General Virology* **70**, 513-524.

MATTHEWS, R.E.F. (1985). Viral Taxonomy for the nonvirologist. *Annual Review of Microbiology* **39** 451-474.

McWHORTER, F.P. (1932). Narcissus mosaic symptoms. *Phytopathology*, **22**, 998.

MEEHAN, B.M. & MILLS, P.R. (1991). Nucleotide sequence of the 3'-terminal region of carnation latent virus. *Intervirology*, **32**, 262-267.

MEMELINK, J., VAN DER VLUGT, C.I.M, LINTHORST, H.J.M., DERKS, A.F.L.M, ASJES, C.J. & BOL, J.F. (1990). Homologies between the genomes of a carlavirus (lily symptomless virus) and a potexvirus (lily virus X) from lily plants. *Journal of General Virology* **71**, 917-924.

MEYER, M. & DESSENS, J.T. (1996). The complete nucleotide sequence of barley mild mosaic virus RNA1 and its relationship with other members of the *Potyviridae*. *Virology* **219**, 268-273.

MOWAT, W.P., DUNCAN, G.H. & DAWSON, S. (1988 a). An appraisal of the identities of potyviruses infecting narcissus. *Acta Horticulturae* **234**, 207-208.

MOWAT, W.P., DUNCAN, G.H. & DAWSON, S. (1988 b). Narcissus late season yellows potyvirus: symptoms, properties and serological detection. *Annals of Applied Biology* **113**, 531-544.

MOWAT, W.P., DAWSON, S. & DUNCAN, G.H. (1989). Production of antiserum to a non-structural potyviral protein and its use to detect narcissus yellow stripe and other potyviruses. *Journal of Virological Methods* **25**, 199-210.

MOWAT, W.P., DAWSON, S., DUNCAN, G.H. & ROBINSON, D.J. (1991). Narcissus latent, a virus with filamentous particles and a novel combination of properties. *Annals of Applied Biology* **119**, 31-46.

MUELLER, E., GILBERT, J., DAVENPORT, G., BRIGNETI, G. & BAULCOMBE, D. (1995). Homology-dependent resistance: transgenic virus resistance in plants related to homology-dependent gene silencing. *The Plant Journal* **7**, 1001-1013.

MURPHY, F.A., FAUQUET, C.M., BISHOP, D.H.L., GHABRIAL, S.A., JARVIS, A.W., MARTELLI, G.P., MAYO, M.A. & SUMMERS, M.D. (eds.) (1995) Virus Taxonomy, Classification and Nomenclature of Viruses. Sixth report of the International Committee on Taxonomy of Viruses, *Archives of Virology* **supplement 10**.

NIBLETT, C.L., ZAGULA, K.R., CALVERT, L.A., KENDALL, T.L., STARK, D.M., SMITH, C.E., BEACHY, R.N. & LOMMEL, S.A. (1991). cDNA cloning and nucleotide sequence of wheat streak capsid protein gene. *Journal of General Virology* **72**, 499-504.

OH, C.-S. & CARRINGTON, J.C. (1989). Identification of essential residues in potyvirus proteinase HC-Pro by site-directed mutagenesis. *Virology* **173**, 692-699.

PAPPU, S.S., BRAND, R., PAPPU, H.R., RYBICKI, E.P., GOUGH, K.H., FRENKEL, M.J. & NIBLETT, C.L. (1993). A polymerase chain reaction method adapted for selective amplification and cloning of 3' terminal sequences of potyviral genomes : application to mosaic virus. *Journal of Virological Methods* **41**, 9-20.

PEERENBOOM, E., PROLS, M., SCHELL, J., STEINBISS, H.H. & DAVIDSON, A.D. (1992). The complete nucleotide sequence of RNA 1 of a German isolate of barley yellow mosaic virus and its comparison with a Japanese isolate. *Journal of General Virology* **73**, 1303-1308.

PEERENBOOM, E., JACOBI, V., ANTONIW, J.F., SCHLICHTER, U., CARTWRIGHT, E.J., STEINBISS, H.H. & ADAMS, M.J. (1996). The complete nucleotide sequence of RNA-2 of a fungally-transmitted isolate of barley mild mosaic virus and identification of amino acid combinations possibly involved in fungus transmission. *Virus Research* **40**, 149-159.

PLESE, N. & MILICIC, D. (1973). Two viruses isolated from *Maclura pomifera*. *Phytopathologische Zeitschrift* **77**, 178-183.

PLESE, N. & WRISCHER, M. (1978). Light and electron microscopy of cells infected with maclura mosaic virus. *Acta Botanica Croatica* **37**, 47-51.

PLESE, N., KOENIG, R., LESEMANN, D. E. & BOZARTH, R.F. (1979). Maclura mosaic virus - an elongated plant virus of uncertain classification. *Phytopathology* **69**, number 5, 471-475.

POCH, O., SAUVAGET, I., DELARUE, M. & TORDO, N.(1989). Identification of four conserved motifs among the RNA-dependent polymerase encoding elements. *The EMBO Journal*, **8**, 3867-3874.

PROLS, M., DAVIDSON, A., SCHELL, J. & STEINBISS, H.H. (1990). *In vitro* translation studies with cDNA clones corresponding to the RNA's of barley yellow mosaic virus and barley mild mosaic virus. *Journal of Phytopathology* **130**, 249-259.

REES, A.R. (1966). Effects of infection with narcissus mosaic, narcissus yellow stripe and tobacco rattle viruses on the growth and flowering of narcissus cv. 'Minster Talma'. *Annals of Applied Biology* **58**, 25-30.

REICHMMANN, J.L., LAIN, S. & GARCIA, J.A. (1992). Highlights and prospects of potyvirus molecular biology. *Journal of General Virology* **73**, 1-16.

RIDLEY, M. (1993). Evolution. Blackwell Scientific Publications, Inc.

ROBAGLIA, C., DURAND -TARDIF, M., TRONCHET, M., BOUDAZIN, G., ASTIER-MANIFACIER, S. & CASSE-DELBAT, F. (1989). Nucleotide sequence of potato virus Y (N strain) genomic RNA. *Journal of General Virology* **70**, 935-947.

RUPASOV, V.V., MOROZOV, S.YU., KANYUKA, K.V. & ZAVRIEV, S.K. (1989). Partial nucleotide sequence of potato virus M RNA shows similarities to potexviruses in gene arrangement and the encoded amino acid sequences. *Journal of General Virology* **70**, 1861-1869.

RYBICKI, E. (1990). The classification of organisms at the edge of life or Problems with virus systematics. *South African Journal of Science* **86**, 182-186.

RYBICKI, E.P. & HUGHES, F.L. (1990). Detection and typing of maize streak virus and other distantly related geminiviruses of grasses by polymerase chain reaction amplification of a conserved viral sequence. *Journal of General Virology* **71**, 2519-2526.

RYBICKI, E.P. & SHUKLA, D.D. (1992). Coat protein phylogeny and systematics of potyviruses. *Archives of Virology* **supplement 5**, 139-170.

SALM, S.N., REY, M.E.C. & RYBICKI, E.P. (1996a). Phylogenetic justification for splitting the *Rymovirus* genus of the taxonomic family *Potyviridae*. *Archives of Virology* **141**, 2237-2242.

SALM, S.N., REY, M.E.C., ROBERTSON, N.L., FRENCH, R., RABENSTEIN, F. & SCHUBERT, J. (1996b). Molecular cloning and nucleotide sequencing of the

partial genomes of Agropyron and Hordeum mosaic viruses, two members of the *Rymovirus* genus in the taxonomic family *Potyviridae*. *Archives of Virology* **141**, 2115-2127.

SALM, S.N., REY, M.E. & FRENCH, R. (1996c). Molecular cloning and nucleotide sequencing of the 3'-terminal region of a South African isolate of ryegrass mosaic virus RNA and *in vitro* expression of the coat protein gene. *Archives of Virology* **141**, 185-195.

SANFORD, J.C. & JOHNSON, S.A. (1985). The concept of parasite-derived resistance: deriving resistance genes from the parasites own genome. *Journal of Theoretical Biology* **113**, 395-405.

SCHENK, P., SOHN, A., ADAMS, M.J., ANTONIW, J.F., HAMACHER, J. & STEINBISS, H.H. (1995). Movement of bymoviruses and functions of RNA2-encoded proteins of barley yellow mosaic virus. *Agronomie* **15**, 387-391.

SCHLICHTER, U., SOHN, A., PEERENBOOM, E., SCHELL, J. & STEINBISS, H.H. (1993). Molecular analysis of the capsid protein of a German isolate of barley mild mosaic virus. *Plant Cell Reports* **12**, 237-240.

SCHUBERT, J., & RABENSTEIN, F. (1995). Sequence of the 3'-terminal region of the RNA of a mite transmitted potyvirus from *Hordeum murinum* L. *European Journal of Plant Pathology* **101**, 123-132.

SCHUBERT, J., RABENSTEIN, F. & PROLL, E. (1995). Sequence of the 3'-part of the RNA of ryegrass mosaic virus, a potyvirus. *Agronomie* **15**, 447-452.

SHUKLA, D.D. & WARD, C.W. (1988). Amino acid sequence homology of coat protein as a basis for identification and classification of the potyvirus group. *Journal of General Virology* **69**, 2703-2710.

SHUKLA, D.D., WARD, C.W., & BRUNT, A.A., (1994). The Potyviridae. CAB International, Wallingford.

SIMON, A.E. & BUJARSKI, J.J. (1994). RNA-RNA recombination and evolution in virus-infected plants. *Annual Review of Phytopathology* **32**, 337-362.

SOHN, A., SCHENK, P., SIGNORET, P.A., SCHMITZ, G., SCHELL, J. & STEINBISS, H.-H. (1994). Sequence analysis of the 3' terminal half of RNA 1 of wheat spindle streak mosaic virus. *Archives of Virology* **135**, 279-292.

THORNBURY, D.W, HELLMAN, G.M., RHOADS, R.E. & PIRONE, T.P. (1985). Purification and characterisation of potyvirus helper component. *Virology* **144**, 260-267.

THORNBURY, D.W, PATTERSON, C.A., DESSENS, J.T. & PIRONE, T.P. (1990). Comparative sequence of the helper component (HC) region of potato virus Y and a HC-defective strain, potato virus C. *Virology* **178**, 573-578.

TIMPE, U. & KUHNE, T. (1994). The complete nucleotide sequence of RNA 2 of barley mild mosaic virus (BaMMV). *European Journal of Plant Pathology* **100**, 233-241.

TREMBLAY, M-F., NICOLAS, O., SINHA, R.C., LAZURE, C. & LALIBERTE, J-F. (1990). Sequence of the 3'-terminal region of turnip mosaic virus RNA and the capsid protein gene. *Journal of General Virology* **71**, 2769-2772.

USUGI, T. & SAITO,Y. (1976) Wheat yellow moasic virus. *Annals of Phytopathology Society of Japan* **42**, 12-20.

USUGI, T., KASHIWAZAKI, S., OMURA, T. & TSUCHIZAKI, T. (1989). Some properties of nucleic acids and coat proteins of soil-borne filamentous viruses. *Annals of the Phytopathological Society of Japan,* **55**, 26-31.

VAN DUK, P. (1993). Carlavirus isolates from cultivated Allium species represent three viruses. *Netherlands Journal of Plant Pathology* **99**, 233-257.

VAN REGENMORTEL, M.H.V. (1990). Virus species, a much overlooked but essential concept in virus classification. *Intervirology* **31** 241-254.

VAN SLOGTEREN, E. & OUBOTER, M.P. (1946). Investigations on virus diseases of narcissus. *Royal Horticultural Society*, The Daffodil and Tulip Yearbook. **12**, 3-20.

VERCHOT, J., KOONIN, E.V. & CARRINGTON, J.C. (1991). The 35-kDa protein from the N-terminus of the potyviral polyprotein functions as a third virus-encoded proteinase. *Virology* **185**, 527-535.

WALKEY, D.G.A. (1990). Virus diseases. In : Rabinowitch, H.D. & Brewster J.L. (eds.), Onions and allied crops, Volume II. CRC Press inc., Boca Raton, Florida. pp 191-212.

WARD, C.W. (1993). Progress towards a higher taxonomy of viruses. *Research in Virology* **144** 419-453.

WARD, C., WEILLER, G., SHUKLA, D.D. & GIBBS, A.J. (1995). Molecular systematics of the *Potyviridae*, the largest plant virus family. In: Gibbs, A.J., Calisher, C.H. & Garcia-Arenal, F. (eds.) Molecular Basis of Viral Evolution. Cambridge University press, Cambridge, UK. pp. 477-500.

WETTER, C.& MILNE, R.G. (1981). Handbook of plant virus infections comparative diagnosis (E. Kurstak). Elsevier/North-Holland Biomedical Press, Amsterdam.

ZAVRIEV, S.K., KANYUKA, K.V. & LEVAY, K.E. (1991). The genome organisation of potato virus M RNA. *Journal of General Virology* **72**, 9-14.

ZUIDEMA, D., LINTHORST, H.J.M., HUISMAN, M.J., ASJES, C.J. & BOL, J.F. (1989). Nucleotide sequence of narcissus mosaic virus RNA. *Journal of General Virology* **70**, 267-276.

Abbreviations

BSA	bovine serum albumin
DEPC	di-ethyl pyrocarbonate
dGTP etc..	2' deoxyriboguanosine tri-phosphate
EDTA	ethylenediaminetetra-acetate
HEPES	N-2-hydroxyethylpiperazine-N'-2-ethanesulfonic acid
IPTG	*iso* propyl β-D-thiogalactopyranoside
LB	Luria and Bertani (media: See Appendix I))
$MgCl_2$	magnesium chloride
NA	nutrient agar (See Appendix I)
NaCl	sodium chloride
NaOH	sodium hydoxide
NB	nutrient broth (See Appendix I)
MOPS	3 [*N*-morpholino]propanesulfonic acid (see Appendix I)
PCR	polymerase chain reation
SDS	sodium dodecyl-sulphate
SDW	sterile distilled water
SSC	salt, sodium, citrate
TAE	tris, acetate, EDTA
TBE	tris, borate, EDTA
TEMED	N,N,N',N'-tetramethylenediamine
Tris	tris[hydroxymethyl]aminomethane
X-GAL	5-bromo-4-cholor-3-indoyl β-D-galactopyranoside

Appendix I

Solutions

Media

<u>LB broth and agar</u>

Luria & Bertani Media:

Per Litre:	Mix
Bacto-tryptone	10g
Bacto-Yeast Extract	5g
NaCl	10g

Add the above solids to 950ml distilled water and dissolve by shaking. Adjust pH to 7.0 (with approx. 0.2ml 5N NaOH) and bring final volume to 1000ml with water Sterilise by autoclaving for 20 min at 15lb/sq in on the liquid cycle. For plates, add 15g agar.

Nutrient Broth and agar are made according to the manufacturers instructions.

<u>Blue white colour selection</u>

IPTG 100mM: 0.023g in 10ml SDW

X-GAL 20mg/ml : add 5ml of N,N-dimethylformamide to a 100mg bottle of X-GAL

Aliquot and store at -20°C, protect X-GAL from light in a dark bottle or tubes wrapped in foil.

To a 400ml bottle of LA, add 800µl of X-GAL stock and 480µl of IPTG stock.

<u>Plant transformation media</u>

MSO: for 1l, mix 1x MSO salts and 3% (w/v) sucrose.

MSD 4x2: for 1l, mix MSO as above and add :

Naphthalene acetic acid (NAA) 0.1mg/l

6-Benzylaminopurine (6-BAP) 1.0 mg/l (stock solution dissolve in HCl)

Hybridisation solutions

<u>Denhardts solution</u>

100X stock can be kept in aliquots at -20°C.

2% Ficoll

2% polyvinyl pyrrolidone

2% BSA

Oligo Labelling Buffer
Use a stock solution of BSA 10mg/ml

Mix A, B and C in the ratio 2:5:3

A:

2M Tris HCl pH8	625µl
5M $MgCl_2$	25µl
SDW	350µl
2-βmercaptoethanol	18µl
100mM dATP	5µl
100mM dGTP	5µl
100mM dTTP	5µl

All dNTPS in stock solution of 3mM Tris, 0.2mM EDTA pH 7.0

B: 2M HEPES pH6.6

C: hexadeoxyribonucleotides (Pharmacia)
resuspend at 90_{OD} units/ml in 3mM Tris HCl, 0.2mM EDTA pH7.0

SSC
20X stock solution:

1.5M tri-sodium citrate

3M NaCl

RNA solutions

MOPS buffer
A 10X stock solution is prepared.

	Final conc.	Mix
MOPS	200mM	10.47g
sodium acetate	50mM	1.03g
EDTA	10mM	0.93g

Make up to 250ml with distilled water and adjust pH to 7.0 with glacial acetic acid.
Store at 4°C for a maximum of 1 month.

Borate Buffer
Used for viral preparations:
0.5M borate buffer pH 7.8 (0.5 M Boric Acid, 0.125M Tetraborate)
0.2% ascorbic acid
0.2% Na_2SO_3

Appendix II

Suppliers

ABI sequence Navigator
Perkin Elmer
Applied Biosystems division
850 Lincoln Centre RR
Foster City CA 94404-1128
USA

Adgen Diagnostic systems
Watson Peat Building
Ayr
Scotland KA6 5HW

Adgia Inc.
30380 Country Road 6
Elkhart IN 46514
USA

Advanced Biotechnologies Ltd.
Unit 7 Mole Business Park
Randalls Road
Leatherhead
Surrey KT22 7BA

Amersham International Plc
Amersham Place
Little Chalfont
Buckinghamshire
HP7 9NA

Bio 101 Inc.
P.O. Box 2284
La Jolla CA 92038-2284
USA

BioRad Laboratories Ltd.
BioRad House
Maylands Avenue
Hemel Hempstead
Hertfordshire
HP2 4BA

Gene Jockey I & II
Distributed by:
Biosoft
49, Bateman Street
Cambridge
CB2 1LR

GibCo BRL
Life Technologies Ltd.
3 Fountain Drive
Inchinnnan Business Park
Paisley
PA4 9RF

Invitrogen Corporation
3985 B Sorrento Valley Blvd.
San Diego CA 92121
USA

Perkin Elmer Corporation
761 Main Avenue
Norwalk CT 06859-0012
USA

Pharmacia Biotech
Head Office
Uppsala, Sweden

Promega Corporation
2800 Woods Hollow Road
Madison WI 53711-5399
USA

Sigma
Chemical Company
P.O.Box 14508
St Louis MO 63178
USA

Stratagene Cloning Systems
11011 North Torrey Pines Road
La Jolla CA 92037
USA

USB and 3M
Distributed by:
Fisher Scientific UK
Bishop Meadow Road
Loughborough Leicestershire
LE11 5RG

PAGES EXCLUDED UNDER INSTRUCTION FROM UNIVERSITY