Caulimoviridae (Plant Pararetroviruses)

Roger Hull, *John Innes Centre, Norwich, UK*

The *Caulimoviridae* family comprises six genera whose members replicate by reverse transcription hence the fact that they are also called pararetroviruses. The family *Caulimoviridae* is the only plant virus taxon whose members have double-stranded DNA genomes.

Secondary article

Article Contents

Introduction

The family *Caulimoviridae* is the only plant virus taxon whose members have double-stranded DNA genomes. They are also termed plant pararetroviruses because their replication involves a reverse transcription step. Pararetroviruses are distinguished from the retroviruses of animals in that the encapsidated genome is DNA and that the viral replication does not involve integration of the viral genome into that of the host, all the replication being episomal (Hull and Will, 1989).

Classification

The family *Caulimoviridae* comprises six genera (**Table 1**), which are distinguished from each other primarily by their genome organizations. These genera fall into two subgroups, at present not formally recognized. One subgroup, based on the genus *Caulimovirus* and known as caulimoviruses, has isometric particles that are usually found in cytoplasmic proteinaceous inclusion bodies, and many of the members are transmitted by aphids (**Table 2**); the other subgroup comprising the genera Badnavirus and 'rice tungro bacilliform-like viruses' are known generically as badnaviruses and have bacilliform particles and no cytoplasmic inclusion bodies; many members of this group are transmitted by mealybugs (**Table 2**).

Historical Perspective

CaMV (see **Table 1** for abbreviations) was the first plant virus shown to have a double-stranded DNA genome (1968) and in 1982 was nominated by the International Committee on the Taxonomy of Viruses as the type member of a new virus group. Other viruses with circular double-stranded DNA genomes have been placed in this group, which has recently been raised to the status of a family. Sequencing of genomes has shown them to differ in the organization of the coding regions and this has led to the creation of several genera. The caulimoviruses attracted attention in the early days of plant molecular biology because of their possible use as gene vectors. However, this potential has not been realized, although some functional regions of the viral genome have found a use. In studying the replication of CaMV, it became apparent in 1983 that reverse transcription was involved in the process. This was the first example of reverse transcription outside animal retroviruses and differed in several features from the animal system, and this led to the use of the term 'pararetroviruses' for these viruses to distinguish them from retroviruses.

The DNA genome of badnaviruses was described in 1990 (Lockhart, 1990) and it then became apparent that these viruses had many features in common with caulimoviruses. These features are in the virus genomes and are associated with the reverse transcription mode of replication. As noted above, badnaviruses are distinguished from caulimoviruses by their particle shape and lack of cytoplasmic inclusion bodies; they also differ in some aspects of their genome organization.

Structure of Virus Particles

Particles of members of the caulimovirus subgroup are isometric and about 50 nm in diameter (**Figure 1a**). There are, as yet, no detailed X-ray crystal data on the particle structure but studies using cryoelectron microscopy and neutron small-angle scattering have given considerable information the structure of CaMV (Cheng *et al*., 1992). The particles are made up of three concentric layers. The outermost is composed of 420 coat protein subunits arranged in a triangulation number $T=7$ icosahedral symmetry, the first example of this class of icosahedral symmetry in a virus particle. The other two layers are made up of coat protein and genomic DNA.

Table 1 Classification of *Caulimoviridae*

Genus	Species	
Caulimovirus	Type	*Cauliflower mosaic virus* (CaMV)
	Other	*Blueberry red ringspot virus* (BRRV)
		Carnation etched ring virus (CERV)
		Dahlia mosaic virus (DMV)
		Figwort mosaic virus (FMV)
		Horseradish latent virus (HRLV)
		Mirabilis mosaic virus (MiMV)
		Strawberry vein-banding virus (SVBV)
		Thistle mottle virus (ThMoV)
	Tentative	Aquilegia necrotic mosaic virus (ANMV)
		Cestrum virus (CV)
		Plantago virus 4 (PlV-4)
		Sonchus mottle virus (SmoV)
'Soybean chlorotic mottle-like viruses'	Type	*Soybean chlorotic mottle virus* (SbCMV)
	Other	*Peanut chlorotic streak virus* (PCSV)
'Cassava vein mottle-like viruses'	Type	*Cassava vein mottle virus* (CVMV)
'Petunia vein-clearing-like viruses'	Type	*Petunia vein-clearing virus* (PVCV)
Badnavirus	Type	*Commelina yellow mottle virus* (CoYMV)
	Other	*Aglaonema bacilliform virus* (ABV)
		Banana streak virus (BSV)
		Cacao swollen shoot virus (CSSV)
		Canna yellow mottle virus (CaYMV)
		Citrus mosaic badnavirus (CMBV)
		Dioscorea bacilliform virus (DBV)
		Kalanchoe top-spotting virus (KTSV)
		Pineapple bacilliform virus (PBV)
		Piper yellow mottle virus (PYMoV)
		Schefflera ringspot virus (SRV)
		Sugarcane bacilliform virus (SCBV)
	Tentative	Aucuba bacilliform virus (*Aucuba japonica*) (AuBV)
		Mimosa bacilliform virus (*Albizzia julibrissin*) (MBV)
		Taro bacilliform virus (*Colocasia esculenta*) (TaBV)
		Yucca bacilliform virus (*Yucca elephantipes*) (YBV)
'Rice tungro bacilliform-like viruses'	Type	*Rice tungro bacilliform virus* (RTBV)

Badnaviruses have bacilliform particles of about 30 nm in diameter and usually about 130 nm in length (**Figure 1b**); some isolates of some viruses can form much longer particles. The particles of RTBV have been shown to be tubular structures based on a $T = 3$ icosahedron cut across its threefold axis (Hull, 1996).

Structure of Viral Genomes

The genomes of all the plant pararetroviruses are circular double-stranded DNA of between 7.2 and 8.2 kb (**Table 2**). The virion DNA is characterized by discontinuities at specific sites, in which the DNA strands usually overlap. One strand always has one discontinuity, which is taken as the zero position on the map. The other strand has between one and three discontinuities depending on virus genus or strain. These discontinuities are the positions of priming for (−)- and (+)-strand synthesis, respectively, during the reverse transcription phase of the viral replication (see below).

All the coding capacity of each of the plant pararetroviruses is on one strand, that with the single discontinuity. The genera of these viruses are distinguished by the organization of open reading frames (ORFs) (**Figure 2**; **Table 3**). Some of the products of these ORFs act as functional proteins without further processing; other ORF products are processed to give functional proteins. For instance, the product of ORF III of badnaviruses is processed to give at least the virus coat protein, an aspartate proteinase (which is considered to perform at

Table 2 Properties of *Caulimoviridae*

Virus	Vector	Particle shape	Cytoplasmic inclusion bodies	Size of DNA[a] (kb)	No. of discontinuities
Caulimovirus					
CaMV	A	S	+	8.016–8.033	2–3
BRRV	ND	S	+	ND	ND
CERV	A	S	+	7.932	3
DMV	A	S	+	8.0	3
FMV	A	S	+	7.743	4
HRLV	A	S	ND	8.0	3
MiMV	A	S	+	7.8	4
SVBV	A	S	+	7.876	ND
ThMoV	ND	S	+	7.8	3
ANMV	ND	S	ND	ND	ND
CV	ND	S	ND	ND	ND
PIV-4	ND	S	+	ND	ND
SmoV	A	S	ND	ND	ND
'Soybean chlorotic mottle-like viruses'					
SbCMV	A	S	+	8.175	3
PCSV	ND	S	+	8.174	ND
'Cassava vein mottle-like viruses'					
CVMV	ND	S	+	8.158	ND
'Petunia vein clearing-like viruses'					
PVCV	ND	S	+	7.205	2
Badnavirus					
CoYMV	M	B	–	7.489	2
ABV	ND	B	–	ND	ND
BSV	M	B	–	7.387	2
CSSV	M	B	–	7.161	2
CaYMV	ND	B	ND	ND	ND
CMBV	ND	B	ND	ND	ND
DBV	ND	B	–	ND	2
KTSV	ND	B	ND	7.35	ND
PBV	ND	B	ND	ND	ND
PYMoV	M	B	–	ND	ND
SRV	ND	B	ND	ND	ND
SCBV	M	B	–	7.568	2
AuBV	ND	B	ND	ND	ND
MBV	ND	B	ND	ND	ND
TaBV	ND	B	ND	ND	ND
YBV	ND	B	ND	ND	ND
'Rice tungro bacilliform-like viruses'					
RTBV	L*	B	–	8.000–8.002	2

Virus abbreviations given in Table 1.

A, aphid; M, mealybug; L*, leafhopper but requiring presence of *Rice tungro spherical virus*; ND, not determined; S, spherical; B, bacilliform; +, present; –, absent.

[a] Where given to three decimal points indicates that the genome has been sequenced; range of sizes indicates that several isolates have been sequenced.

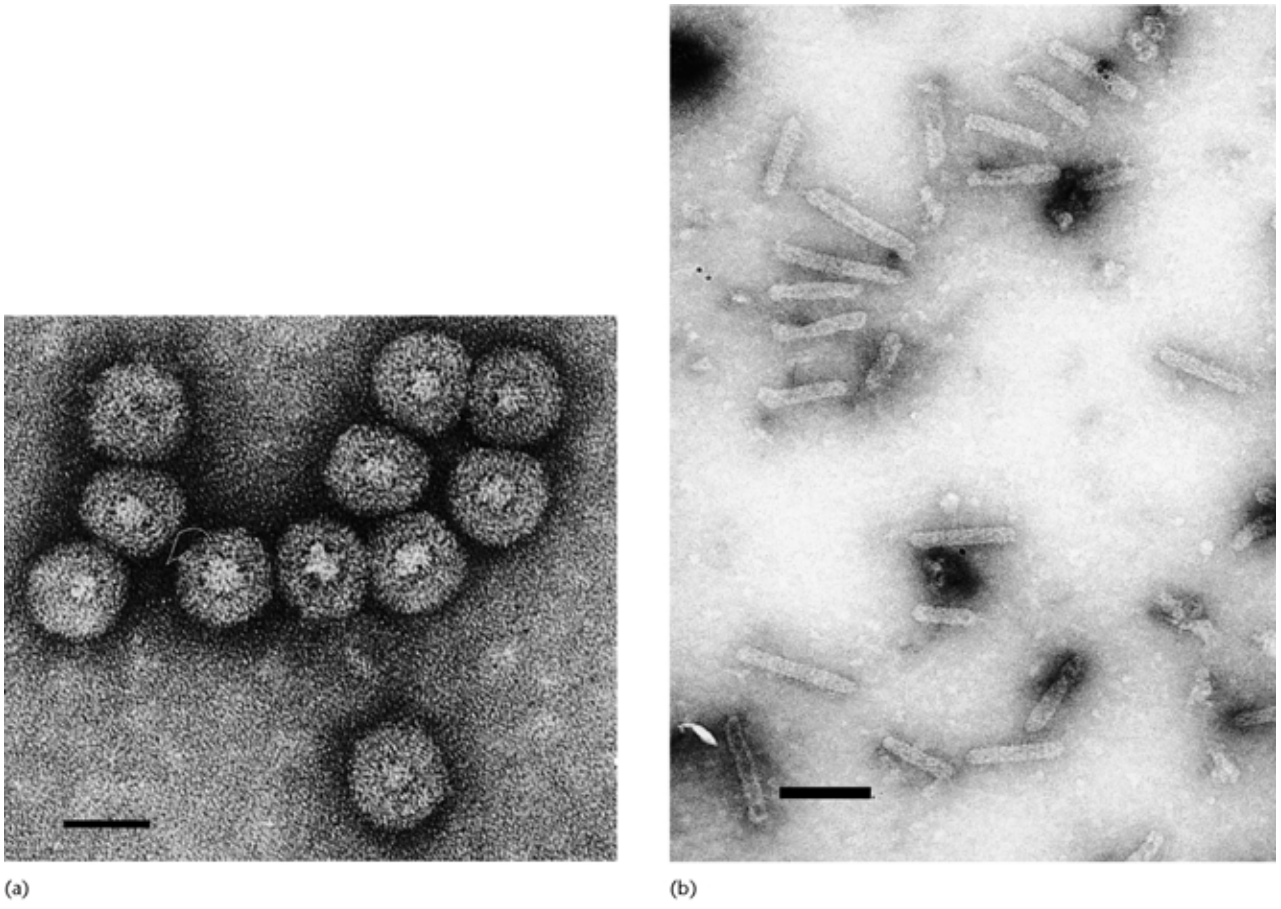

Figure 1 Electron micrographs of particles of (a) CaMV (bar, 50 nm) and (b) RTBV (bar, 100 nm) negatively stained with 2% uranyl acetate.

least some of the processing) and the reverse transcriptase and ribonuclease H enzymes involved in the virus replication. **Table 4** gives an indication of which ORF products are processed.

Genome Expression

Genome expression strategies have been studied for the caulimoviruses and for RTBV. Two major RNA species are transcribed from the CaMV genome (**Figure 2**), the 35S RNA which is more than genome length, with a terminal repeat of about 180 nucleotides, and the 19S, which is 3′-coterminal with the 35S RNA. These transcripts are expressed from promoters in the virus genome. The 35S promoter has been extensively studied and is widely used in constructs for the genetic modification of plants. The 19S RNA is the monocistronic messenger RNA (mRNA) for the ORF VI product, the inclusion body protein, whereas most, if not all, the other ORFs are expressed from the 35S RNA. The inclusion body protein (also known as the transactivator protein, TAV) has a transactivation function which facilitates the expression of proteins from the polycistronic 35S RNA (see Hohn and Fütterer, 1997). There is some evidence for splicing of the 35S RNA but the relevance of this to the expression of ORFs I–V is unknown.

ORFs I–III of RTBV are expressed from the 35S RNA in a cascade manner, details of which are described in Hull (1996). ORF I lacks a conventional AUG start codon and its translation initiates at an AUU codon which is inefficient, only recruiting about 10% of the ribosomes. The AUG start codon of ORF II is also inefficient, whereas that of ORF III is efficient. Thus, three ORFs are translated from this polycistronic mRNA. The mRNA for ORF IV is formed by splicing the 35S RNA between a position close to its 5′-end and the start of ORF IV, giving an intron of 6.3 kb (Fütterer *et al*., 1994).

The leader sequences of the pararetrovirus 35S RNAs contain several small ORFs which could affect translation of the major downstream ORFs. There is considerable evidence that, owing to RNA folding and the possible involvement of a viral (the TAV in the case of CaMV) or even host gene product, ribosomes 'shunt' across from

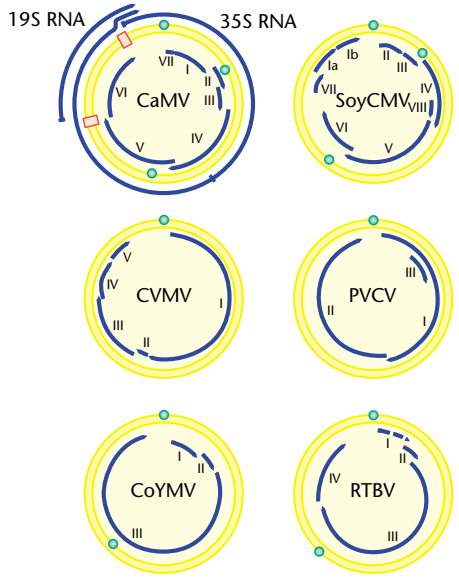

Figure 2 Genome organizations of members of the *Caulimoviridae*; abbreviations of virus names in **Table 1**. The yellow double circle represents the double-stranded DNA genome with the green circles indicating the positions of the single-strand discontinuities. The red blocks on the CaMV map show the positions of the two promoters. The inner blue arcs represent the ORFs; RTBV ORF I is dashed to indicate that it does not have a conventional start codon. The positions of the RNA transcripts of CaMV are indicated by the outer blue circle and arc with the 5′-end shown by filled box and 3′-end by arrowhead. (Data for CaMV, Franck *et al.* (1980); SoyCMV, Hasegawa *et al.* (1989); CVMV, Calvert *et al.* (1995); PVCV, Richert-Pöggeler and Shepherd (1997); CoYMV, Medberry *et al.* (1990); RTBV, Hay *et al.* (1991), Qu *et al.* (1991).)

Table 3 Properties of gene products of *Caulimoviridae*: sizes of open reading frames (kDa)

	Open reading frame					
Virus genus	I	II	III	IV	V	VI
Caulimovirus	37	18	15	57	79	61
'SoyCMV-like'	35/13[a]	18	23	52	86	53
'CVMV-like'	186	9	77	24	26	–
'PVCV-like'	126	125	13	–	–	–
Badnavirus	23	15	216	–	–	–
'RBTV-like'	24	12	194	46	–	–

[a] ORF I of SoyCMV is expressed from two reading frames (Figure 1).

close to the 5′-end of the RNA to the start of ORF I, omitting most of the leader sequence (see Fütterer *et al.*, 1993).

Functions have been ascribed to several of the gene products of some of the members of the *Caulimoviridae* (**Table 4**; see Hohn and Fütterer, 1997). As well as the viral coat proteins and the enzymes involved in virus replication, reverse transcriptase, ribonuclease H and aspartate proteinase, the gene products which facilitate virus movement both within and between plants have been identified for CaMV and some other, if not all, members of the genus *Caulimovirus*. The gene I product of CaMV and the N-terminal portion of ORF III of CoYMV make tubular structures which pass through the cell wall, enabling virus particles to move from cell to cell. These are termed movement proteins and they are likely to be encoded by most, if not all, members of this virus family. The gene II product of the genus *Caulimovirus* is involved in transmission of the virus particles by aphids and is described below. Caulimovirus gene III product is considered to assist with virus particle assembly and the gene VI product has several functions, including forming the matrix of the inclusion bodies in which replication occurs and in transactivating the expression of proteins from the 35S RNA. No function has formally been ascribed to the products of badnavirus ORFs I or II, to the RTBV ORF IV gene product or to the putative ORF VII of CaMV.

Genome Replication

The replication cycle of CaMV has been studied in much detail and the basic findings can be applied to most, if not all, members of the *Caulimoviridae*. The replication has two phases: transcription of the virion DNA in the nucleus to give the various transcripts; and reverse transcription of the major transcript in the cytoplasm.

The capsid is removed from the virus particles by an unknown mechanism and the virus genome enters the nucleus, where the discontinuities are sealed to give a supercoiled molecule. This molecule associates with histones to form a minichromosome, which is the template for transcription by the host DNA-dependent RNA polymerase II, to give the various transcripts described above. The transcripts pass to the cytoplasm where they act both as templates for translation and, in the case of the 35S RNA, template for reverse transcription. There is evidence that the reverse transcription phase of CaMV replication takes place in virus-like particles in the cytoplasmic inclusion bodies formed from the gene VI product (Thomas *et al.*, 1985). It is likely that there are similar virus particle-like replication complexes for all these viruses, though those of members of the badnavirus subgroup are not in proteinaceous inclusion bodies. Reverse transcription is a complex process involving priming of (−)-strand DNA synthesis by the 3′-end of $\mathrm{tRNA}_{\mathrm{init}}^{\mathrm{met}}$, priming of (+)-strand DNA synthesis usually by polypurine-rich regions of the 35S RNA and various strand switches to give the circular DNA molecule; the discontinuities are at the sites of (−)- and (+)-strand priming. Detailed descriptions of the replication mechanism of these viruses can be found in Mason *et al.* (1987) and Hull (1996).

Table 4 Properties of gene products of *Caulimoviridae*: gene functions[a]

Virus genus	CP	RT	RH	AP	CTC	HC
Caulimovirus	IV	V	V	V	I	II
'SoyCMV-like'	IV	V	V	V	Ia	ND
'CVMV-like'	I	III	III	III	I	ND
'PVCV-like'	II	II	II	II	I	ND
Badnavirus	III	III	III	III	III	ND
'RTBV-like'	III	III	III	III	ND	ND

CP, coat protein; RT, reverse transcriptase; RH, ribonuclease H; AP, aspartate proteinase; CTC, cell-to-cell spread (movement) protein; HC, insect vector helper protein.
[a]Open reading frame encoding gene function (**Figure 1**).

Pathogenesis

Viruses in the *Caulimoviridae* mostly have narrow host ranges, with those of the caulimovirus subgroup being restricted to dicotyledonous plants; members of the badnavirus subgroup infect either dicotyledonous or monocotyledonous plants. The symptoms induced by members of the caulimovirus subgroup are mainly mottles and mosaics in leaves of infected plants. Members of the badnavirus subgroup induce a variety of symptoms, ranging from chlorotic leaf streaks (mainly in monocotyledonous plants) to mottles in leaves of dicotyledonous plants and growth deformation.

Most of the viruses of the *Caulimoviridae* infect most cell types, though some of the members of the badnavirus subgroup, e.g. RTBV, are restricted to the vascular tissues. A major feature distinguishing the two subgroups is that cells infected with caulimoviruses contain characteristic cytoplasmic proteinaceous inclusion bodies which are not found in cells infected with badnaviruses. The caulimovirus inclusion bodies are not surrounded by membranes and are mainly composed of the product of ORF VI; some consist of the insect transmission factor, the product of ORF II. The inclusion bodies contain most of the virus particles and, as noted above, those made up of the gene VI product are thought to be the sites of the reverse transcription phase of virus replication. As the gene II product interacts with cellular tubulins, it is considered that the aggregation of this protein into inclusion bodies prevents the protein from damaging the cell (Pirone and Blanc, 1996).

Transmission and Epidemiology

Many members of the caulimovirus subgroup are naturally transmitted in the semipersistent manner by aphids (**Table 2**), with often little specificity for aphid species. For instance, CaMV can be transmitted by at least 27 aphid species and DaMV by 13 species. Transmission requires the presence of a virus-coded aphid transmission factor (see above) and there have been extensive studies on that of CaMV (Pirone and Blanc, 1996). The CaMV aphid transmission factor is a protein of 18 kDa whose predicted secondary structure has an N-terminal domain, which is predominantly β sheet, and a C-terminal domain of mainly a helix; the domains are separated by a region of random structure. This has led to the hypothesis that the helper factor is bifunctional, with one domain interacting with the virus particle and the other with the vector. *In vitro* studies have shown that the C-terminal domain interacts with virus particles.

The epidemiology of members of the caulimovirus subgroup with annual hosts, e.g. CaMV, is similar to that of other semipersistent aphid-transmitted viruses spreading into the crop from overwintering sources during the host-finding flights of aphids, giving primary infection foci. Secondary spread from these foci is usually by aphids breeding on the host and moving away from the initial plants when they get overcrowded. Viruses of this group with perennial hosts, e.g. CERV, DaMV, SVBV, are mainly spread by vegetative propagation, although initial infections will be by aphids.

Most members of the badnavirus subgroup are transmitted in the semipersistent manner by mealybugs (**Table 2**). Little is known about the interactions between the virus and the vector. RTBV is transmitted by leafhoppers in the semipersistent manner, the main vector being the rice green leafhopper, Nephotettix virescens. Transmission of RTBV requires the presence or association with RTSV, a plant picornavirus with an RNA genome, the two viruses making up the complex that induces rice tungro virus disease. It is thought that RTSV encodes a transmission factor which facilitates the transmission of RTBV. No vector is known for RTBV alone.

Badnaviruses cause considerable losses to tropical crops. For instance, in attempts to control CSSV, more than 180 million cacao trees have been eradicated in Ghana; rice tungro disease (RTBV + RTSV) causes annual losses in excess of \$US 1.5×10^9 in South and Southeast Asia. The main mode of transmission of many of the badnaviruses of perennial tropical crops, e.g. BSV, SCBV, is by vegetative propagation of that crop. Mealy-

bugs are the important vector in other viruses, e.g. CSSV. The spread of RTBV is associated with that of RTSV, and is by leafhoppers moving from infected crops or wild hosts into healthy crops.

Control

The main approach to control of insect-transmitted members of the *Caulimoviridae* is by the application of insecticides. Because the vector can transmit virus even after a short feed on a host, insecticides are not a particularly effective approach to preventing primary infections. Approaches to minimize primary infections include the use of mineral oil sprays and reflective backgrounds, which deter host-finding insects from landing.

Control of infection of perennial crops is often attempted by eradication of symptomatic plants. As noted above, there has been an eradication scheme in Ghana for the control of CSSV. This approach has not been proven to be very effective as it is labour intensive and requires compensation schemes to cover the losses to the grower.

References

Calvert LA, Ospina MD and Shepherd RJ (1995) Characterization of cassava vein mosaic virus: a distinct plant pararetrovirus. *Journal of General Virology* **76**: 1271–1276.

Cheng RH, Olson NH and Baker TS (1992) Cauliflower mosaic virus; 1 420 subunit ($T = 7$), multilayer structure. *Virology* **186**: 655–668.

Franck A, Guilley H, Jonard G, Richards K and Hirth L (1980) Nucleotide sequence of cauliflower mosaic virus DNA. *Cell* **21**: 285–294.

Fütterer J, Kiss-Lásló Z and Hohn T (1993) Non-linear ribosome migration on cauliflower mosaic virus 35S RNA. *Cell* **73**: 789–802.

Fütterer J, Potrykus I, Valles-Brau MP *et al*. (1994) Splicing in a plant pararetrovirus. *Virology* **198**: 663–670.

Hasegawa A, Verver J, Shimada A (1989) The complete sequence of soybean chlorotic mottle virus DNA and identification of a novel promoter. *Nucleic Acids Research* **17**: 9993–10013.

Hay JM, Jones MC, Blakebrough ML *et al*. (1991) An analysis of the sequence of an infectious clone of rice tungro bacilliform virus, a plant pararetrovirus. *Nucleic Acids Research* **19**: 2615–2621.

Hohn T and Fütterer J (1997) The proteins and functions of plant pararetroviruses: knowns and unknowns. *Critical Reviews in Plant Science* **16**: 133–161.

Hull R (1996) Molecular biology of rice tungro viruses. *Annual Review of Phytopathology* **34**: 275–297.

Hull R and Will H (1989) Molecular biology of viral and nonviral retroelements. *Trends in Genetics* **5**: 357–359.

Lockhart BEL (1990) Evidence for a double-stranded circular DNA genome in a second group of plant viruses. *Phytopathology* **80**: 127–131.

Mason WS, Taylor JM and Hull R (1987) Retroid virus genome replication. *Advances in Virus Research* **32**: 35–95.

Medberry SL, Lockhart BEL and Olszewski NE (1990) Properties of *Commelina* yellow mottle virus's complete DNA sequence, genomic discontinuities and transcript suggests that it is a pararetrovirus. *Nucleic Acids Research* **18**: 5505–5513.

Pirone TP and Blanc S (1996) Helper-dependent vector transmission of plant viruses. *Annual Review of Phytopathology* **34**: 227–247.

Qu R, Bhattacharyya M, Laco GS *et al*. (1991) Characterization of the genome of rice tungro bacilliform virus: comparison with *Commelina* yellow mottle virus and caulimoviruses. *Virology* **185**: 354–364.

Richert-Pöggeler KR and Shepherd RJ (1997) Petunia vein-clearing virus: a plant pararetrovirus with the core sequence for an integrase function. *Virology* **236**: 137–146.

Thomas CM, Hull R, Bryant JA and Maule AJ (1985) Isolation of a fraction from cauliflower mosaic virus-infected protoplasts which is active in the synthesis of (+) and (−) strand viral DNA and reverse transcription of primed RNA templates. *Nucleic Acids Research* **13**: 4557–4576.

Further Reading

Hohn T and Fütterer J (1997) The proteins and functions of plant pararetroviruses: knowns and unknowns. *Critical Reviews in Plant Science* **16**: 133–161.

Hull R (1995) Genus *Caulimovirus*. In: Murphy FA, Fauquet CM, Bishop DHL *et al*. (eds) *Virus Taxonomy*. Sixth Report of the International Committee on Taxonomy of Viruses, pp. 189–192. Vienna: Springer-Verlag.

Lockhart BEL, Olszewski NE and Hull R (1995) Genus *Badnavirus*. In: Murphy FA, Fauquet CM, Bishop DHL *et al*. (eds) *Virus Taxonomy*. Sixth Report of the International Committee on Taxonomy of Viruses, pp. 185–188. Vienna: Springer-Verlag.

Mason WS, Taylor JM and Hull R (1987) Retroid virus genome replication. *Advances in Virus Research* **32**: 35–95.

Rothnie HM, Chapdelaine Y and Hohn T (1994) Pararetroviruses and retroviruses: a comparative review of viral gene structure and gene expression strategies. *Advances in Virus Research* **44**: 1–67.